

芍药 FPPS 基因的克隆及生物信息学分析

徐 杰¹, 高水平^{2*}, 史国安¹, 范丙友^{1*}

1. 河南科技大学农学院, 河南 洛阳 471003

2. 河南科技大学林学院, 河南 洛阳 471003

摘 要: 目的 克隆芍药萜类物质生物合成关键酶法呢基焦磷酸合酶 (farnesyl pyrophosphate synthase, FPPS) cDNA 全长序列并对其进行生物信息学分析。方法 根据芍药转录组测序数据, 设计一对特异性 PCR 引物, 应用 RT-PCR 技术扩增出芍药 FPPS 基因目标条带并克隆至 pMD18-T 载体上, 菌落 PCR 和质粒 PCR 鉴定出阳性重组子后进行序列测定及序列同源性搜索、分子系统进化树构建、结构域搜索及 3D 结构预测等生物信息学分析。结果 测序结果表明芍药 FPPS 基因全长 1 315 bp, 含 60 bp 的 5'-UTR、1 050 bp 的 CDS 和 205 bp 的 3'-UTR, 共编码 349 个氨基酸, GenBank 登录号为 KP708571; Blastn 和 Blastp 分析结果显示芍药 FPPS (KP708571) 及其编码的蛋白 (AKJ26301) 在核苷酸水平和氨基酸水平上与多种植物的 FPPS 基因和 FPPS 蛋白具同源性; 分子进化树分析结果表明芍药 FPPS 蛋白与其他物种 FPPS 蛋白的亲缘关系相对较远; 预测芍药 FPPS 蛋白相对分子质量为 40 200, 等电点为 5.33, 是一个定位于胞液、不含跨膜结构域、不含信号肽分子的亲水性、稳定蛋白; 发现芍药 FPPS 含有底物结合位点、底物-Mg²⁺结合位点、催化位点、富含天冬氨酸位点 1 及富含天冬氨酸位点 2 共 7 个保守结构域; 3D 结构中 α -螺旋所占比例高且 α -螺旋之间由 β -转角 (loop) 相连接; 中央腔 (central cavity) 由大约 10 个核心 α -螺旋排列而成。结论 首次从芍药中克隆了 FPPS 基因并对其进行了初步的生物信息学分析。

关键词: 芍药; 法呢基焦磷酸合酶; 基因克隆; 生物信息学分析; 萜类物质

中图分类号: R282.12 文献标志码: A 文章编号: 0253-2670(2016)04-0655-06

DOI: 10.7501/j.issn.0253-2670.2016.04.021

Cloning and bioinformatic analysis of FPPS gene from *Paeonia lactiflora*

XU Jie¹, GAO Shui-ping², SHI Guo-an¹, FAN Bing-you¹

1. College of Agriculture, Henan University of Science and Technology, Luoyang 471003, China

2. College of Forestry, Henan University of Science and Technology, Luoyang 471003, China

Abstract: Objective To clone the full-length cDNA sequence which encodes one of the key enzymes of terpene biosynthesis, farnesyl pyrophosphate synthase (FPPS) and analyze the bioinformation. **Methods** Based on the transcriptome data of *Paeonia lactiflora* (Pl), a pair of specific PCR primers were designed. The objective PCR band of PIFPPS was successfully amplified using RT-PCR technique and then it was cloned to pMD18-T vector. After clone and plasmid PCR characterization, the recombinants were sequenced and then a series of bioinformatic analysis was carried out including sequence homology search, construction of molecular phylogenetic tree, domain search, and 3D structure prediction, etc. **Results** Sequencing results showed that the full-length PIFPPS was 1 315 bp, which contained 60 bp 5'-UTR, 1 050 bp CDS, and 205 bp 3'-UTR. It encoded 349 amino acids and the accession number of GenBank was KP708571. Blastn and Blastp analysis revealed that PIFPPS (KP708571) and its encoding protein (AKJ26301) had high homology with FPPS genes and proteins from several plants at the nucleotide and amino acid levels. Phylogenetic tree analysis indicated that the genetic relationship of PIFPPS with FPPS from other plants was relatively far. It predicted that the molecular weight and isoelectric point of PIFPPS were 40 200 and 5.33, which was a hydrophilic and stable protein, located in cytoplasm without transmembrane domain and signal peptide. It contained seven conservative domains like substrate binding pocket, substrate-Mg²⁺ binding site, catalytic site, aspartate-rich region 1, and aspartate-rich region 2. The proportion of α -helix in its 3D structure was high and the α -helix was linked by β -loop. The central cavity was composed of 10 core α -helices. **Conclusion** The full-length cDNA sequence of FPPS is cloned firstly from *P. lactiflora* and the bioinformatic analysis is then carried out.

Key words: *Paeonia lactiflora* Pall.; farnesyl pyrophosphate synthase; gene cloning; bioinformatic analysis; terpene

收稿日期: 2015-08-23

基金项目: 国家自然科学基金资助项目 (U1204323); 河南省科技厅国际合作项目 (134300510052)

作者简介: 徐 杰 (1987—), 男, 河南沈丘人, 硕士生, 从事植物分子生物学研究。E-mail: comdisc@163.com

*通信作者 高水平 (1976—), 女, 河南洛阳人, 硕士, 讲师, 从事园林植物与观赏园艺研究。E-mail: gaoshuiping1949@163.com

范丙友 (1974—), 男, 河南伊川人, 博士, 教授, 从事植物分子生物学研究。E-mail: fanbingyou2005@163.com

芍药 *Paeonia lactiflora* Pall. (Pl) 为芍药科芍药属芍药组多年生宿根草本植物, 兼具观赏和药用价值^[1]。作为我国传统的中药材, 芍药根可加工为赤芍 *Paeoniae Radix Rubra* 和白芍 *Paeoniae Radix Alba*, 二者均具有免疫调节、镇痛、镇静等作用^[2]。赤芍和白芍的主要药效成分有芍药苷 (paeoniflorin)、氧化芍药苷 (oxypaeoniflorin)、苯甲酰芍药苷 (benzoylpaeoniaflorin)、白芍苷 (albiflorin) 等双环单萜类化合物、三萜类化合物、黄酮类化合物、挥发油等物质^[3-5]。植物体内萜类物质的生物合成有 2 条代谢途径^[6], 第 1 条为位于细胞质中的甲羟戊酸途径 (mevalonate pathway, MVA)^[7]; 第 2 条为位于质体中的脱氧木酮糖-5-磷酸途径 (1-deoxy-D-xylulose-5-phosphate pathway, DXP) 或甲基赤藓醇-4-磷酸途径 (methylerythritol-4-phosphate pathway, MEP)^[8]。萜类物质的代谢过程可分为 3 个阶段: 中间体异戊烯基焦磷酸 (isopentenylpyro-phosphate, IPP) 及其双键异构体二甲基烯丙基焦磷酸 (dimethylallyl pyrophosphate, DMAPP) 的生成、直接前体物质的生成、萜类的生成及修饰^[9]。法呢基焦磷酸合酶 (farnesyl pyrophosphate synthase, FPPS) 是植物体内萜类物质生物合成的关键酶^[10], 高等植物体内 IPP 和 DMAPP 形成后, 2 分子的 IPP 和 1 分子的 DMAPP 在 FPPS 的催化下头尾相连缩合形成 C₁₅ 骨架的法呢基焦磷酸 (farnesyl diphosphate, FPP)。目前, 已经从杯萼海桑^[11]、三七^[12]和罗汉果^[13]等多种植物中分离并克隆了 FPPS 的 cDNA 序列。本研究基于芍药转录组测序数据 (DRX027794), 发现了一条与 FPPS 基因高度同源的序列。本研究, 应用 Primer Premier 5.0 软件设计了一对特异性 PCR 扩增引物, 基于 RT-PCR 技术成功扩增出 PCR 产物, 用常规方法对 PCR 产物进行了克隆测序并对其进行了生物信息学分析。本研究为下一步应用实时定量 PCR 技术分析芍药 FPPS 基因的时空表达特性并鉴定其生物学功能奠定了基础。

1 材料与方法

1.1 材料与试剂

1.1.1 材料 “桃花飞雪”芍药花瓣采自河南科技大学芍药园, 由范丙友教授鉴定, 带回实验室后液氮速冻并立即置于 -80 °C 冰箱备用。大肠杆菌菌株 DH5 α 保存于洛阳市牡丹生物学重点实验室。

1.1.2 试剂 Recombinant DNase I、PrimerScriptTM 1st strand cDNA synthesis kit、Takara MiniBEST Agarose Gel DNA Extraction Kit Ver. 4.0、克隆载体

pMD18-T、Premix TaqTM、DNA Ligation Kit Ver. 2.0、DL 2000 DNA Marker 均购自于宝生物工程 (大连) 有限公司, 其余试剂均为进口或国产分析纯。

1.2 方法

1.2.1 引物设计 基于本课题组完成的芍药转录组测序数据 (DRX027794), 利用 Primer Premier 5.0 软件设计一对特异性 PCR 扩增引物 5'-AGAACAGG-CATTTTTTCCAC-3' 和 5'-TGATTCATTCCTC-CTCCAACC-3', 引物序列由北京华大基因研究中心合成。

1.2.2 芍药总 RNA 的提取、纯化及 cDNA 的合成 采用 CTAB-LiCl 法^[14]提取“桃花飞雪”芍药花瓣总 RNA, 用 Recombinant DNase I 除去芍药总 RNA 样品中的 DNA 污染, 1.4% 琼脂糖凝胶电泳检测总 RNA 纯化效果。参照 PrimerScriptTM 1st strand cDNA synthesis kit 说明书进行 cDNA 的合成。

1.2.3 芍药 FPPS 基因的 RT-PCR 扩增 RT-PCR 反应体系总体积为 50 μ L, 含 Premix Taq 25 μ L, cDNA 2 μ L, 上游引物和下游引物各 2 μ L, ddH₂O 19 μ L。PCR 扩增程序 98 °C、10 s, 55 °C、30 s, 72 °C、60 s, 共 30 个循环; 1% 琼脂糖凝胶电泳检测 PCR 扩增产物。

1.2.4 PCR 产物的克隆及测序 用 DNA 凝胶回收试剂盒对 PCR 产物进行纯化回收并与 pMD18-T 载体连接, 随后将连接产物转化大肠杆菌 DH5 α 感受态细胞, 对长出的单菌落进行菌落 PCR 检测, 阳性菌 37 °C 过夜培养, 提取质粒并进行质粒 PCR 鉴定, 鉴定出的阳性重组子送至北京华大基因研究中心进行测序。

1.2.5 序列分析 应用在线的 Blastn 和 Blastp 程序 (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) 在核苷酸水平上和氨基酸水平上对芍药 FPPS 基因及其编码的蛋白序列进行同源性分析; 应用 DNAMAN 7.0 软件对芍药 FPPS 蛋白序列及其同源序列进行多序列比对分析。应用 MEGA 5.0 软件, 采用 NJ 法构建分子系统进化树。

1.2.6 芍药 FPPS 蛋白的物理特性预测 利用 ProtParam 软件 (<http://cn.expasy.org/tools/protparam.html>) 分析相对分子质量、等电点等^[15]; 利用 DAS-TMfilter (<http://mendel.imp.ac.at/sat/DAS/DAS.html>) 分析跨膜结构区域^[16]; 应用在线分析工具 SubLoc (<http://www.bioinfo.tsinghua.edu.cn/SubLoc/>) 分析亚细胞定位^[17]; 利用 SignalP 软件 (<http://www.cbs.dtu.dk/services/SignalP/>) 分析信号肽^[18]。

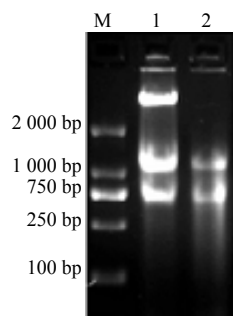
1.2.7 芍药 FPPS 蛋白的结构域分析 基于 SUPERFAMILY Sequence Search 在线工具 (<http://supfam.org/SUPERFAMILY/hmm.html>) 分析芍药 FPPS 蛋白序列; 应用 CDD 数据库 (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) 搜索芍药 FPPS 蛋白的结构域^[19]。

1.2.8 芍药 FPPS 蛋白的 3D 结构预测 采用 Swiss-model (<http://swissmodel.expasy.org/>) 同源建模方法进行 3D 结构预测^[20]。

2 结果与分析

2.1 芍药总 RNA 的提取及纯化

“桃花飞雪”芍药花瓣总 RNA 的 28 S 和 18 S 条带清晰, 表明提取的总 RNA 完整性较好, 经 DNase I 处理后总 RNA 样品中不含 DNA 污染 (图 1), 可以用于下一步 cDNA 的合成。



1-未纯化总 RNA 2-纯化后总 RNA M-Marker, 下同
1-unpurified total RNA 2-purified total RNA M-Marker, same as below

图 1 芍药总 RNA 琼脂糖凝胶电泳图

Fig. 1 Agrose gel electrophoresis of total RNA extracted from *P. lactiflora*

2.2 芍药 FPPS 基因的 RT-PCR 扩增

用反转录得到的芍药 cDNA 为模板, 基于 RT-PCR 技术用 Ex Taq DNA 聚合酶成功扩增出了一条大小约为 1 300 bp 的 PCR 条带 (图 2), 所得片段的大小与预期理论值一致。

2.3 芍药 FPPS 基因的克隆及测序

对转化后长出的单菌落进行菌落 PCR 检测, 进一步用质粒 PCR 鉴定阳性重组子 (图 3), 质粒 PCR 产物大小与预测片段大小一致, 表明成功克隆了芍药 FPPS 基因。测序结果表明芍药 FPPS 基因 cDNA 全长 1 315 bp, 包含 60 bp 的 5'-UTR、1 050 bp 的 CDS (coding sequence, 编码区) 和 205 bp 的 3'-UTR; CDS 区涵盖起始密码子 ATG 及终止密码子 TAG, 共编码 349 个氨基酸 (图 4), GenBank 登录号为 KP708571。

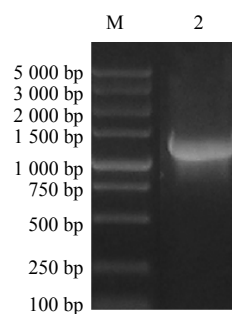


图 2 芍药 FPPS 基因的 RT-PCR 扩增

Fig. 2 RT-PCR amplification of *PIFPPS*

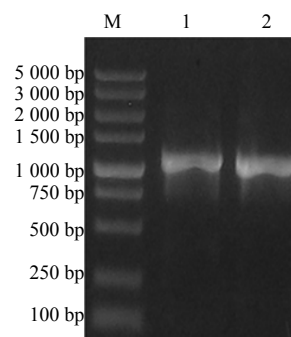


图 3 质粒 PCR 鉴定阳性重组子

Fig. 3 Characterization of recombinants by plasmid PCR

2.4 序列分析

Blastn 在线分析结果表明芍药 FPPS 基因在核苷酸水平上与多种植物的 FPPS 基因具高度同源性, 其中与毛果杨 *Populus trichocarpa* Torr. & Gray (EF147135)、橡胶 *Hevea brasiliensis* (Willd. ex A. Juss.) Muell. Arg. (AY135188)、白羽扇豆 *Lupinus albus* L. (U15777)、西洋梨 *Pyrus communis* L. (KF855953)、黄芪 *Astragalus membranaceus* (Fisch.) Bunge (KP233831)、甘草 *Glycyrrhiza uralensis* Fisch. (GQ214505)、茉莉 *Jasminum sambac* (L.) Aiton (KM395813) FSPS 基因的同源性达 81%~83%; Blastp 在线分析结果表明芍药 FPPS 蛋白 (AKJ26301) 在氨基酸水平上与多种植物的 FPPS 蛋白具高度同源性, 其中与毛果杨 (XP_002308751)、茉莉 (AIY24422)、橡胶 (AAM98379、BAF98301、ABR09548)、假马齿苋 *Bacopa monnieri* (L.) Wettst (ADV03080)、中粒咖啡 *Coffea canephora* Pierre ex Froehn (CDP04546)、树棉 *Gossypium arboreum* L. (KHG28027)、雷蒙德氏棉 *Gossypium raimondii* L. (XP_012445510) 等植物的 FPPS 蛋白同源性达 84%~89%; 由此表明分离的序列即为 FPPS 基因。

芍药 FPPS 蛋白 (AKJ26301) 与毛果杨 (XP_002308751)、茉莉 (AIY24422)、橡胶 (AAM98379)、假马齿苋 (ADV03080) FPPS 的蛋白多序列比对分析结果见图 5, 序列相似性达 90.54%, 表明 FPPS 蛋白的氨基酸序列在不同物种之间较为保守。为了探索芍药 FPPS 蛋白与其他近缘植

物 FPPS 的进化关系, 选取来自橡胶、毛果杨、树棉、茉莉、假马齿苋、辣薄荷 *Mentha × piperita* Linn. 等物种的 9 个 FPPS 序列, 基于完整的氨基酸序列构建 NJ 进化树 (图 6)。结果表明 9 个 FPPS 聚合成 2 个分枝, 茉莉、假马齿苋、辣薄荷 FPPS 蛋白形成 1 个分枝, 它们与芍药 FPPS 亲缘关系较远; 来

```

agaacaggcattttttccacgaactctttcgtttctctcatatataaaaaatacaaacATGGCTGACGTGAATGGAACCAGGATGGAT 90
CTGAGGTCAAAGTTTTTGAATGTATACTCCGTTCTGAAATCGGAGCTCCTTGATGATCCTGCTTTGAATTCACCAATGAATCTCGTCAA 10
L R S K F L N V Y S V L K S E L L D D P A F E F T N E S R Q 40
TGGGTCGAGCGGATGCTGGACTATAACGTTCTGGAGGTAAGTTGAACAGGAGGACTCTCTGTTATTGACAGCTACCAAATCCTAAAGGAT 270
W V E R M L D Y N V P G G K L N R G L S V I D S Y Q I L K D 70
GGGAAGGAACCTTACTGAAGAAGAGATATTTCTAACAAAGTCTTGGTTGGTGTATTGAATGGCTCCAAGCATATTTCTTGTCTTGAT 360
G K E L T E E I F L T S A L G W C I E W L Q A Y F L V L D 100
GATATTATGGACGGTCTCACACACGACGCGGTACGCTTGTGGTTCAGATTGCTAAGGTTGGTCTTATCGCAGCTAATGATGGAATT 450
D I M D G S T H T R R G G Q P C W G L I A A N D G I 130
ATACTTCGCAACCATCTAGGATTCTGAAAAACCATCTTAGAGAAAGGAGTTATTACGTGGATCTTCTGGATTGTTAATGAGGTG 540
I L R N H I P R I L K N H F R E R S Y Y V D L L D L F N E V 160
GAATTCCAAACAGCTTCAGGACAGATGATGATTGATTACTACACTTGAAGGAGAAAAAGATCTGCCAAATACTCATTGTCACTTAC 630
E F Q T A S G Q M I D L I T T L E G E K D L S K Y S L S L H 190
CGTCGCATTGTTCAATTCAAGACTGCTTACTACTCATTTTACCTTCGGTTGCATCGCATGTTGTTATGTCGGGTGAGAATCTGGACAAC 720
R R I V Q F K T A Y Y S F L P V A C A L V M S G E N D N 220
CACGTTGCTGTGAAGGACATCTTGTGAGATGGGACTTATTTCAAGTACAGGATGATTATCTGGATTGCTACGGTGAGCCTGAAAAG 810
H V A V K D I L V E M G T Y F T Q Q D D Y L D C Y F P E K 250
ATTGGAAGATTGGAACAGACATTGAAGATTCAAGTGCTCGTGGATGGTGAAAGCATTGGAACCTGTAATGAGGAACAGAGAAGAAA 900
I G K I G T D I E D F K C S W M V V K A L E L C N E E Q K K 280
ATATTACATGAACACTATGGGAAAGCAGACCCAGCTGATGTTGCAAAAGTGAAGGCCCTTTATAATGAACCTTGATCTTCAGGTTGATTT 990
I L H E H Y G K A D P A D V A K V K A L Y N E L D L Q G V F 310
GCCGAGTATGAGAACAAGAGCTATGAGAACTCCATTGAAGCACATCCAAAGCAAGCAGTGCAAGCAGTGTGAAGTCCTTC 1080
A E Y E N K S Y E K L I T S I E A H P S K A V Q A V L K S F 340
TTGGGTAAGATATACAAGAGGCAGAAATAGaggagattcaagaacgttgcggagaaaataggttgcgtcaggaggtatctcatgattct 1170
L G K I Y K R Q K * 349
cccgcctttagtttattataaacgatttcattgttttttaatacagggatgctcatttgtagttgttacttagaggt 1260
atcatcttttttttttttttcgcttaattgttgaggagaagaatgaatca 1315

```

ATG 为起始密码子, TAG 为终止密码子
ATG is start codon, TAG is stop codon

图 4 芍药 FPPS 基因核苷酸及编码的氨基酸序列

Fig. 4 Nucleotide sequences of FPPS gene and its encoded amino acids in *P. lactiflora*

芍药 AKJ26301	MADVNGTRMDLRSKFLNVSVLKSELNDPAFEFTNESRCWVERMLDYNVPGGKLNRLGS	60
毛果杨 XP_002308751	MADLKSFLNVSVLKSELNDPAFEWSPDSRDWVRMLDYNVPGGKLNRLGS	53
茉莉 AIY24422	MANONGANSDLRTKFLNVSVLKSELNDPAFEWTDNSRCWVERMLDYNVPGGKLNRLGS	60
橡胶 AAM98379	MADLKSFLNVSVLKSELNDPAFEWTFDSRCWVERMLDYNVPGGKLNRLGS	53
假马齿苋 ADV03080	MANHNGQTTDLREAFGLNVSVLKSELNDPAFEWTDASRCWVERMLDYNVPGGKLNRLGS	60
芍药 AKJ26301	VIDSYQILKDKBELTEBEIFLTSALGWCIEWLQAYFLVLDDIMDGSHTRRGQPCWFRLEK	120
毛果杨 XP_002308751	VIDSYKYLKEGKELTEBEIFLTSALGWCIEWLQAYFLVLDDIMDSSHTRRGQPCWFRLEK	113
茉莉 AIY24422	VIDSYKLLKEGKELTEBEIFLTSALGWCIEWLQAYFLVLDDIMDSSHTRRGQPCWFRLEK	120
橡胶 AAM98379	VIDSYKLLKEGKELTEBEIFLTSALGWCIEWLQAYFLVLDDIMDSSHTRRGQPCWFRLEK	113
假马齿苋 ADV03080	VIDSYKLLKEGKELTEBEIFLTSALGWCIEWLQAYFLVLDDIMDSSHTRRGQPCWFRLEK	120
芍药 AKJ26301	VGLIAANDGILRNHPIRLKHNHFRERSYYVDLLDLFNEVEFQTASGQMIDLITTEGEK	180
毛果杨 XP_002308751	VGLIAANDGILRNHPIRLKHNHFRDKAYVDLLDLFNEVEFQTASGQMIDLITTEGEK	173
茉莉 AIY24422	VGLIAANDGILRNHPIRLKHNHFRDKPYVDLLDLFNEVEFQTASGQMIDLITTEGEK	180
橡胶 AAM98379	VGLIAANDGILRNHPIRLKHNHFRDKAYVDLLDLFNEVEFQTASGQMIDLITTEGEK	173
假马齿苋 ADV03080	VGLIAVNDGILRNHPIRLKHNHFRKPYVDLLDLFNEVEFQTASGQMIDLITTEGEK	180
芍药 AKJ26301	DLSKYSLSLHRRIVQFKTAYYSFYLPVACALVMSGENLDNEIVKVDLLVEMGTYFQVQDD	240
毛果杨 XP_002308751	DLSKYTLSLHRRIVQKTAAYYSFYLPVACALLMAGENLDNEIVKVDLLVEMGTYFQVQDD	233
茉莉 AIY24422	DLSKYSLSLHRRIVQKTAAYYSFYLPVACALLMAGENLDNEIVKVDLLVEMGTYFQVQDD	240
橡胶 AAM98379	DLSKYTLSLHRRIVQKTAAYYSFYLPVACALLIAGENLDNEIVKVDLLVEMGTYFQVQDD	233
假马齿苋 ADV03080	DLSKYSLSLHRRIVQKTAAYYSFYLPVACALLMAGENLDNEIVKVDLLVEMGTYFQVQDD	240
芍药 AKJ26301	YLDCEGPEKIGKIGTDIEDFKCSWLVVKALEHCNEEQKKLLFHHYKGDPAVAVKIKAL	300
毛果杨 XP_002308751	YLDCEGAPETIGKIGTDIEDFKCSWLVVKALEHCNEEQKKLLFHHYKGDPAVAVKIKAL	293
茉莉 AIY24422	YLDCEGPEKIGKIGTDIEDFKCSWLVVKALEHCNEEQKKLLFHHYKGDPAVAVKIKAV	300
橡胶 AAM98379	YLDCEGPEKIGKIGTDIEDFKCSWLVVKALEHCNEEQKKLLFHHYKGDPAVAVKIKAV	293
假马齿苋 ADV03080	YLDCEGPEKIGKIGTDIEDFKCSWLVVKALEHCNEEQKKLLFHHYKGDPAVAVKIKAL	300
芍药 AKJ26301	YNELDLQGVFAEYENKSYEKLITSIEAHPSKQVQAVLKSFLGKIYKRQ	348
毛果杨 XP_002308751	YNELNLQGVFAEYENKSYEKLITSIIEAHPSKQVQAVLKSFLAKIYKRQ	341
茉莉 AIY24422	YNELDLQGVFAEYENKSYEKLITSIEAHPSKQVQAVLKSFLGKIYKRQ	348
橡胶 AAM98379	YNELDLQGVFAEYENKSYEKLITSIIEAHPSKQVQAVLKSFLAKIYKRQ	347
假马齿苋 ADV03080	YNELNLQGVFAEYENKSYEKLITSIEAHPSKQVQAVLKSFLGKIYKRQ	348

图 5 芍药 FPPS 与近缘物种 FPPS 蛋白的多序列比对

Fig. 5 Multiple sequence alignment of FPPS among *P. lactiflora* and other closely related plants

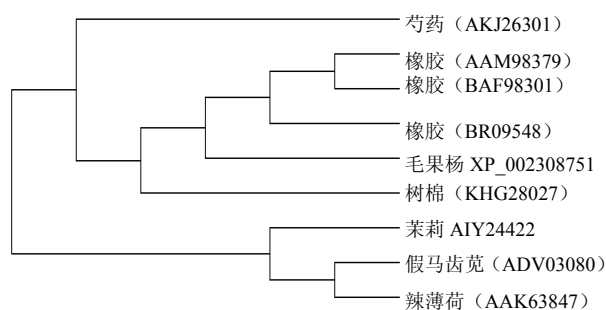


图 6 FPPS 的分子系统进化树

Fig. 6 Phylogenetic tree of FPPS protein

自橡胶的 3 个 FPPS 的亲缘关系较近，它们与毛果杨、树棉 FPPS 聚合成 1 个亚分，芍药 FPPS 独立地形成 1 个亚分，这与芍药属独立成科^[21]的形态学分类地位较为一致。

2.5 芍药 FPPS 蛋白物理特性预测

预测芍药肌动蛋白的相对分子质量为 40 200，等电点为 5.33，分子式为 $C_{1816}H_{2825}N_{469}O_{533}S_{14}$ ，平均亲水性 (grand average of hydropathicity, GRAVY) 为 -0.269，表明为亲水蛋白；预测芍药 FPPS 蛋白是一个亚细胞定位于胞液^[22]、不含跨膜结构域、不含信号肽分子的亲水性、稳定蛋白。

2.6 芍药 FPPS 蛋白的结构域分析

SUPERFAMILY Sequence Search 在线分析结果表明芍药 FPPS 蛋白属于异戊烯基二磷酸合酶 (isoprenyl diphosphate synthases) 家族，隶属于萜类合酶 (terpenoid synthases) 超级家族。CDD 数据库搜索结果表明，芍药 FPPS 蛋白含底物结合位点 L₉₂Y₉₅F₉₆L₉₇L₉₉D₁₀₀D₁₀₁M₁₀₃D₁₀₄R₁₀₉R₁₁₀Q₁₆₈D₁₇₁K₁₉₇T₁₉₈Y₂₀₁D₂₃₉D₂₄₀D₂₄₃K₂₅₃D₂₅₇K₂₆₂、底物-Mg²⁺结合位点 D₁₀₀D₁₀₁D₁₀₄R₁₀₉R₁₁₀D₁₇₁K₁₉₇D₂₃₉D₂₄₀D₂₄₃K₂₅₃D₂₅₇K₂₆₂、活性位点盖残基 S₁₀₆H₁₀₇T₁₀₈R₁₀₉R₁₁₀G₁₁₁Q₁₁₂P₁₁₃C₁₁₄W₁₁₅F₁₁₆R₁₁₇L₁₁₈P₁₁₉K₁₂₀G₂₄₆E₂₄₇K₂₅₃I₂₅₄G₂₅₅T₂₅₆D₂₅₇D₂₆₀F₂₆₁K₂₆₂C₂₆₃S₂₆₄、链长度决定区 Y₉₅F₉₆L₉₇V₉₈L₉₉D₁₀₀D₁₀₁I₁₀₂M₁₀₃D₁₀₄、催化位点 D₁₀₀D₁₀₁D₁₀₄R₁₀₉R₁₁₀K₁₉₇D₂₃₉D₂₄₀、富含天冬氨酸区 1 D₁₀₀D₁₀₁D₁₀₄R₁₀₉R₁₁₀D₁₇₁K₁₉₇、富含天冬氨酸区 2 D₂₃₉D₂₄₀D₂₄₃K₂₅₃D₂₅₇K₂₆₂ 7 个保守结构域。

2.7 芍药 FPPS 蛋白结构预测

芍药 FPPS 蛋白与 PDB (Protein Data Bank, 蛋白质结构数据库) 中绢蒿 *Artemisia spiciformis* Osterh 的重组 FPPS 嵌合体蛋白 (PDB 序列号 4kk2) 高度同源^[23]。模拟出的芍药 FPPS 蛋白的 3D 结构见图 7，表明芍药 FPPS 以二聚体形式聚合， α -螺旋所占比例高， α -螺旋之间由 β -转角 (loop) 相连接，

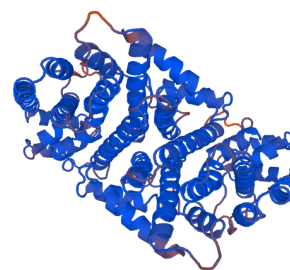


图 7 芍药 FPPS 蛋白的 3D 结构预测

Fig. 7 3D structure prediction of PIFPPS

中央腔 (central cavity) 由大约 10 个核心 α -螺旋排列而成。

3 讨论

本研究基于芍药转录组测序数据 (DRX027794)，发现了一条与 FPPS 基因高度同源的序列，应用 Primer Premier 5.0 软件设计了一对特异性 PCR 引物，基于 RT-PCR 技术克隆了芍药 FPPS 基因。测序结果表明芍药 FPPS 基因全长 1 315 bp，包含 60 bp 的 5'-UTR、1 050 bp 的 CDS 和 205 bp 的 3'-UTR，共编码 349 个氨基酸，GenBank 登录号为 KP708571；序列同源性分析表明芍药 FPPS (KP708571) 及其编码的蛋白 (AKJ26301) 在核苷酸水平和氨基酸水平上与多种植物的 FPPS 基因和 FPPS 蛋白具高同源性；分子进化树构建结果表明芍药 FPPS 蛋白与其他物种 FPPS 蛋白的亲缘关系较远；预测芍药 FPPS 蛋白的相对分子质量为 40 200，等电点为 5.33，是一个亚细胞定位于胞液、不含跨膜结构域、不含信号肽分子的亲水性、稳定蛋白；结构域分析表明芍药 FPPS 含有底物结合位点、底物-Mg²⁺结合位点、催化位点、富含天冬氨酸位点 1 及富含天冬氨酸位点 2 共 7 个保守的结构域；其 3D 结构中 α -螺旋所占比例高且由 β -转角相连接；中央腔由大约 10 个核心 α -螺旋排列而成。本研究结果为下一步基于实时定量 PCR 技术分析芍药 FPPS 基因时空表达特性及其调控芍药萜类物质生物合成机制奠定了一定基础。然而，克隆所得到的基因能否通过转基因技术提高其萜类化合物的量还有待于进一步研究。

参考文献

- [1] 李嘉珏. 中国牡丹与芍药 [M]. 北京: 中国林业出版社, 1999.
- [2] 中国药典 [S]. 一部. 2010.
- [3] 张晓燕. 芍药的化学成分研究 [D]. 沈阳: 沈阳药科大学, 2001.

- [4] 康晓飞. 芍药根内主要活性物质与营养成分的研究 [D]. 泰安: 山东农业大学, 2011.
- [5] Kamiya K, Yoshioka K, Saiki Y, *et al.* Triterpenoids and flavonoids from *Paeonia lactiflora* [J]. *Phytochemistry*, 1997, 44(1): 141-144.
- [6] 占爱瑶, 由香玲, 詹亚光. 植物萜类化合物的生物合成及应用 [J]. *生物技术通报*, 2010, 21(1): 131-135.
- [7] Chappell J, Wolf F, Proulx J, *et al.* Is the reaction catalysed by 3-hydroxy-3-methylglutaryl coenzyme A reductase a rate-limiting step for isoprenoid biosynthesis in plants? [J]. *Plant Physiol*, 1995, 109(4): 1337-1343.
- [8] Rohmer M, Knani M, Simonin P, *et al.* Isoprenoid biosynthesis in bacteria: a novel pathway for the early steps leading to isopentenyl diphosphate [J]. *Biochem J*, 1993, 295(2): 517-524.
- [9] Yuan Y, Yu J, Jiang C, *et al.* Functional diversity of genes for the biosynthesis of paeoniflorin and its derivatives in *Paeonia* [J]. *Int J Mol Sci*, 2013, 14: 18502-18519.
- [10] 邢朝斌, 龙月红, 何 闪, 等. 刺五加法尼基焦磷酸合酶基因的克隆、生物信息学及表达分析 [J]. *中国中药杂志*, 2012, 37(12): 1725-1730.
- [11] 王 峰, 吴秋红, 高 辉, 等. 红树林植物杯萼海桑法尼基焦磷酸合酶的基因克隆及序列分析 [J]. *中草药*, 2013, 44(16): 2294-2299.
- [12] 陈 莉, 蓝秀万, 朱 华, 等. 三七法尼基焦磷酸合酶的基因克隆及序列分析 [J]. *中草药*, 2006, 37(7): 1080-1083.
- [13] 蒙姣荣, 陈本勇, 黎起秦, 等. 罗汉果法尼基焦磷酸合酶基因的克隆及其序列分析 [J]. *中草药*, 2011, 42(12): 2512-2517.
- [14] Gasic K, Hernandez A, Korban S S. RNA extraction from different apple tissues rich in polyphenols and polysaccharides for cDNA library construction [J]. *Plant Mol Biol Rep*, 2004, 22: 437a-437g.
- [15] Walker J M. *The Proteomics Protocols Handbook* [M]. Totawa: Humana Press Inc, 2005.
- [16] Cserző M, Eisenhaber F, Eisenhaber B, *et al.* On filtering false positive transmembrane protein predictions [J]. *Protein Eng*, 2002, 15(9): 745-752.
- [17] Hua S J, Sun Z R. Support vector machine approach for protein subcellular localization prediction [J]. *Bioinformatics*, 2001, 17(8): 721-728.
- [18] Petersen T N, Brunak S, von Heijne G, *et al.* SignalP 4. 0: discriminating signal peptides from transmembrane regions [J]. *Nat Methods*, 2011, 8: 785-786.
- [19] Marchler-Bauer A, Derbyshire M K, Gonzales N R, *et al.* CDD: NCBI's conserved domain database [J]. *Nucl Acids Res*, 2015, 43: D222-D226.
- [20] Biasini M, Bienert S, Waterhouse A, *et al.* SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information [J]. *Nucl Acids Res*, 2014, 42: W252-W258.
- [21] 何春年. 芍药属药用植物亲缘学研究 [D]. 北京: 中国协和医科大学, 2010.
- [22] Ortiz-Gomez A, Jimenez C, Estevez A M, *et al.* Farnesyl diphosphate synthase is a cytosolic enzyme in *Lieshmania major* Promastigotes and its over expression confers resistance to risedronate [J]. *Eukaryotic Cell*, 2006, 5(7): 1057-1064.
- [23] Tarshis L C, Yan M, Poulter C D, *et al.* Crystal structure of recombinant farnesyl diphosphate synthase at 2.6-Å resolution. [J]. *Biochemistry*, 1994, 33(36): 10871-10877.