

• 药材与资源 •

地黄 RgGGPPS2 基因克隆、生物信息学分析及表达分析

赵 乐^{1,2}, 马利刚^{1,2}, 俎梦航¹, 冯卫生^{1,2}, 郑晓珂^{1,2}, 朱昀昊^{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: **目的** 为了研究地黄萜类化合物生物合成途径的功能基因, 从地黄中克隆了一条新的地黄 GGPPS 基因 (RgGGPPS2), 进行生物信息学分析, 原核表达、纯化以及基因表达模式分析。**方法** 根据地黄转录组数据中 RgGGPPS2 基因的序列信息, 设计特异性引物, 克隆了 RgGGPPS2 基因的开放阅读框 (ORF) 序列, 构建 pET32a-RgGGPPS2 原核表达载体, 用 IPTG 在大肠杆菌 BL21 (DE3) 中诱导表达 RgGGPPS2 重组蛋白, 荧光定量 PCR 检测 RgGGPPS2 基因的组织特异性表达。**结果** RgGGPPS2 基因的 ORF 为 867 bp, 编码 289 个氨基酸。生物信息学分析结果发现 RgGGPPS2 蛋白含有 2 个富含天冬氨酸的基序 FARM (first aspartate-rich motif) 和 SARM (second aspartate-rich motif), RgGGPPS2 与芝麻、长春花等双子叶植物中 GGPPS 蛋白的同源性较高。通过构建 pET-32a-RgGGPPS2 原核表达载体在大肠杆菌中成功表达 RgGGPPS2 重组蛋白, 利用 Ni²⁺亲和色谱得到了纯化的 RgGGPPS2 重组蛋白。荧光定量 PCR 结果显示 RgGGPPS2 基因在根中表达量最高, 其次是叶, 茎中最低。**结论** 克隆了 RgGGPPS2 基因, 获得了纯化的 RgGGPPS2 重组蛋白, 为研究 RgGGPPS2 基因在地黄萜类化合物生物合成途径中的功能奠定了基础。

关键词: 地黄; RgGGPPS2; 基因克隆; 生物信息学分析; 表达分析**中图分类号:** R282.12**文献标志码:** A**文章编号:** 0253-2670(2017)11-2269-10**DOI:** 10.7501/j.issn.0253-2670.2017.11.020Cloning, bioinformatic analysis and expression analysis of RgGGPPS2 gene from *Rehmannia glutinosa*ZHAO Le^{1,2}, MA Li-gang^{1,2}, ZU Meng-hang¹, FENG Wei-sheng^{1,2}, ZHENG Xiao-ke^{1,2}, ZHU Yun-hao^{1,2}

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

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

Abstract: **Objective** In order to investigate the functional genes involved in terpenoids biosynthesis pathway of *Rehmannia glutinosa*, a new geranylgeranyl pyrophosphate synthase gene (RgGGPPS2) was isolated from *R. glutinosa*. Meanwhile, the bioinformatic analysis, prokaryotic expression, purification, and expression patterns of RgGGPPS2 gene were carried out. **Methods** Based on the transcriptome data of *R. glutinosa*, specific primers of RgGGPPS2 gene were designed, and an open reading frame (ORF) of RgGGPPS2 gene was isolated from *R. glutinosa*. By constructing the prokaryotic expression vector pET32a-RgGGPPS2, the recombinant RgGGPPS2 protein was expressed in *Escherichia coli* BL21 (DE3) cells under IPTG induction. The expression pattern of RgGGPPS2 in different tissues was detected by real-time PCR. **Results** RgGGPPS2 had an ORF of 867 bp, which encoded a protein of 289 amino acid residues. Bioinformatic analysis indicated that RgGGPPS2 protein contains the two conserved motifs FARM (first aspartate-rich motif, DDxxxxD) and SARM (second aspartate-rich motif, DDxxD). Phylogenetic analysis revealed that RgGGPPS2 protein showed the highest homology with GGPPS protein from *Sesamum indicum*, *Catharanthus roseus* and other dicots. Through the construction of prokaryotic expression vector pET-32a-RgGGPPS2, the recombinant RgGGPPS2 protein was successfully expressed in *E. coli* BL21 (DE3) cells and the recombinant RgGGPPS2 protein was purified by Ni²⁺ affinity chromatography. Real-time PCR analysis indicated that RgGGPPS2 was expressed in high transcript level in roots, lower level in leaves and the lowest level in stems.

收稿日期: 2016-11-14

基金项目: 国家“十二五”科技支撑计划项目 (2011BAI06B02); 国家自然科学基金项目 (81603232); 河南省科技攻关计划项目 (162102310468); 河南中医学院博士科研基金 (BSJJ2011-07, BSJJ2011-18)

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

*通信作者 朱昀昊, 男, 博士, 讲师, 主要从事药用植物资源研究。E-mail: guxinhan123@163.com

Conclusion The RgGGPPS2 gene was isolated from *R. glutinosa* and the recombinant RgGGPPS2 protein was obtained. The results of this study provide a foundation for functional characterization of RgGGPPS2 gene involved in terpenoids biosynthesis pathway of *R. glutinosa*.

Key words: *Rehmannia glutinosa* Libosch; RgGGPPS2; gene cloning; bioinformatic analysis; expression analysis

地黄 *Rehmannia glutinosa* Libosch 是玄参科多年生草本植物, 以其块根入药, 是临床常用的大宗药材, 具有滋阴生津功效, 是“四大怀药”之一^[1]。通过研究地黄化学成分, 发现地黄含有环烯醚萜、苯乙醇、三萜、黄酮、酚酸、木脂素及多糖等多种化合物, 对心脑血管系统、神经系统、免疫系统均有显著的药理作用^[2], 其中萜类化合物是地黄药效物质基础之一, 具有神经保护、抗炎、抗氧化、抗肿瘤、降血糖等多种生物学活性^[3], 利用分子生物学技术研究地黄中萜类化合物的生物合成途径, 克隆和分析关键基因, 为今后调控萜类化合物的生物合成奠定基础。

在植物中萜类化合物是通过位于细胞质的甲羟戊酸 (mevalonate, MVA) 途径和位于质体的甲基赤藓醇磷酸 (methylerythritol phosphate, MEP) 途径衍生而来^[4], 异戊烯基焦磷酸 (isopentenyl pyrophosphate, IPP) 是沟通 2 条代谢途径的中间体, 二甲基烯丙基焦磷酸 (dimethylallyl pyrophosphate, DMAPP) 是 IPP 的双键异构体, IPP 和 DMAPP 是萜类化合物的前体, 经过催化可以生成各种复杂的萜类化合物^[5]。牻牛儿苗基牻牛儿苗基焦磷酸合成酶 (geranylgeranyl pyrophosphate synthase, GGPPS) 能够催化 1 分子 IPP 和 3 分子 DMAPP 头尾相连生成 GGPP, 或者催化 1 分子法尼基焦磷酸 (farnesyl diphosphate, FPP) 和 1 分子 IPP 反应生成 GGPP^[6], GGPP 能够为植物中叶绿素、类胡萝卜素、脱落酸、赤霉素、醌、二萜等物质的合成提供前体物^[7]。在拟南芥中 GGPP 的量能够影响赤霉素 A1、类胡萝卜素等次生代谢产物的合成^[8], 所以催化合成 GGPP 的 GGPPS 是植物次生代谢过程中的关键酶之一, GGPPS 在植物萜类化合物生物合成过程中也起着重要的作用, 在红豆杉的悬浮细胞中加入茉莉酸甲酯 (methyl jasmonate, MeJA) 能够迅速诱导 TbGGPPS 蛋白活性增高, 最终提高悬浮细胞中紫杉醇的量^[9], 目前已在多种植物中对 GGPPS 基因进行了克隆和功能分析, 如拟南芥^[10]、烟草^[11]、丹参^[12]、独行菜^[13]等。

在高等植物中, GGPPS 通常由旁系同源基因编码, 形成 GGPPS 基因家族, 在不同的植物基因组

中拥有 2~12 个基因家族成员, 在拟南芥基因组中有 12 个 GGPPS 基因家族成员^[10], 拟南芥 GGPPS (AtGGPPS) 家族成员分布在不同的器官具有组织特异性, 而且位于不同的亚细胞组分, 如线粒体、叶绿体、细胞质、内质网等, 在这些细胞组分 AtGGPPS 蛋白发挥特定的生物学功能^[14]。拟南芥 AtGGPPS1 蛋白主要位于拟南芥的地上部分, 根中量极少, AtGGPPS3 主要位于拟南芥的根部, 在细胞中 AtGGPPS1 和 AtGGPPS3 都定位于脂质体, 可能参与赤霉素、类胡萝卜素、脱落酸、叶绿素等物质的合成; AtGGPPS2 主要位于花, 少量位于根部, AtGGPPS4 主要位于根的维管组织以及花的柱头, 在细胞中 AtGGPPS2 和 AtGGPPS4 都定位于细胞质, 可能参与细胞质中萜类、多萜醇的合成以及蛋白质的异戊二烯化^[7]; AtGGPPS6 主要位于线粒体, 参与辅酶 Q 的合成^[14]。

前期通过对地黄 GGPPS (RgGGPPS) 基因进行研究, 克隆了地黄 GGPPS 基因家族的 RgGGPPS1 基因, 进行序列分析和组织特异性分析, 发现 RgGGPPS1 可能定位在叶绿体中, 在根中表达量最高^[15]。在本实验室获得的地黄转录组数据中, 发现了另一条注释为 GGPPS 的基因序列, 命名为 RgGGPPS2, 本研究克隆了 RgGGPPS2 基因的开放阅读框 (open reading frame, ORF) 序列, 进行生物信息学分析、原核表达以及表达模式分析, 为今后研究地黄中萜类化合物的生物合成途径奠定基础。

1 材料与方法

1.1 材料

地黄野生种与栽培种“北京 3 号”种植于河南中医药大学河南中药植物园, 经笔者鉴定为地黄 *Rehmannia glutinosa* Libosch。在地黄成熟期 (10 月 20 日), 收集野生种和栽培种的根、茎和叶片, 样品经液氮速冻后, 保存于 -80 °C 冰箱中, 备用。

植物 RNA 提取试剂盒、RNase-Free DNase I、Ni-Agarose His 标签蛋白纯化试剂盒购自康为世纪公司; 质粒小提试剂盒、DNA 产物纯化试剂盒、琼脂糖凝胶 DNA 回收试剂盒购自天根公司; 反转录试剂盒 TransScript II Reverse Transcriptase 和大肠杆菌感受态细胞 DH5 α 和 BL21 (DE3) 菌株购自全式

金公司; Ex Taq 酶、限制性核酸内切酶 (*EcoR* I、*Xho* I)、T₄DNA 连接酶、pMD19-T vector 购自 Takara 公司; QuantiNova SYBR Green PCR Kit 购自 QIAGEN 公司; 引物合成、样品测序由三博远志公司完成; pET-32a 载体为本实验室保存。

1.2 RNA 的提取与反转录

使用植物 RNA 提取试剂盒提取地黄野生种和栽培种的根、茎和叶片总 RNA, 用 1% 琼脂糖凝胶电泳检测总 RNA 的完整性, 用 NanoDrop 2000 测定总 RNA 的浓度和纯度。按照反转录试剂盒说明书, 以提取的地黄各组织的总 RNA 为模板, Oligo(dT)₂₀ 为引物, 反转录合成得到地黄各组织的 cDNA。

1.3 RgGGPPS2 基因克隆

根据地黄转录组数据中 RgGGPPS2 基因的序列信息, 用 Oligo7 软件设计 RgGGPPS2 基因 ORF 序列的特异性引物 (RgGGPPS2-Exp-F 和 RgGGPPS2-Exp-R), 以反转录得到的地黄野生种根 cDNA 为模板, 扩增 RgGGPPS2 基因的 ORF 序列, PCR 反应条件为 95 °C、2 min; 95 °C、10 s, 59.4 °C、45 s, 72 °C、1 min, 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, 切胶回收目的条带, 将其连接到 pMD19-T 载体上, 然后转化大肠杆菌 Trans 5α 感受态细胞, 均匀涂布在 LB 平板 (含有氨苄青霉素) 上, 经菌落 PCR 鉴定后挑选阳性克隆送三博公司测序。

1.4 RgGGPPS2 基因的生物信息学分析

使用 NCBI ORF Finder 查找地黄转录组数据中 RgGGPPS2 转录本的开放阅读框, 并通过 NCBI blastp 与非冗余蛋白数据库 (Non-redundant protein sequences, nr) 进行序列比对分析。用 ExPASy 提供的 ProtParam tool 预测 RgGGPPS2 基因编码蛋白的理化性质; 使用 InterPro Scan 预测 RgGGPPS2 蛋白的保守结构域; 采用 SWISS-MODEL 进行蛋白质的三维同源建模; 使用 SignalP 4.1 进行信号肽的预测; 利用 Target P 1.1 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 MEGA5 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-RgGGPPS2 原核表达载体的构建与诱导表达

用 *EcoR* I 和 *Xho* I 对测序正确的 pMD19-T-RgGGPPS2 质粒进行双酶切, 切胶回收

目的基因片段, 同样用 *EcoR* I 和 *Xho* I 对原核表达载体 pET-32a 进行双酶切, 回收载体片段, 然后用 T₄ DNA Ligase 将目的基因片段和载体片段在 16 °C 连接过夜, 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), 挑取单克隆进行菌液 PCR 和双酶切鉴定, 菌液 PCR 和酶切鉴定正确的单克隆送北京三博公司测序。

1.6 RgGGPPS2 重组蛋白的诱导表达与纯化

将测序正确的单克隆 (含 pET-32a-RgGGPPS2 质粒) 接种到 LB 液体培养基中 (含氨苄青霉素), 37 °C、220 r/min 培养过夜后, 第 2 天按照 1:100 比例接种到 LB 液体培养基中 (含氨苄青霉素), 在 37 °C、220 r/min 条件下培养至对数生长期 (*A*₆₀₀ 值至 0.6), 加入 IPTG 使其终浓度为 0.4 mmol/L, 然后在 28 °C、150 r/min 条件下继续培养 8 h, 诱导地黄 RgGGPPS2 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 RgGGPPS2 重组蛋白的诱导表达条件, 大量培养含 pET-32a-RgGGPPS2 质粒的大肠杆菌 BL21 (DE3) 菌株, 诱导完成后在 4 °C、8 000 ×g 条件下离心 5 min 收集菌体, 用 PBS 缓冲液洗涤菌体, 然后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、10 000 ×g 条件下离心 15 min, 分别收集上清和沉淀, 将沉淀重悬于结合缓冲液 (20 mmol/L Tris-HCl, 5 mmol/L 咪唑, 0.5 mol/L NaCl, 8 mol/L 尿素) 中, 充分混匀使包涵体充分溶解, 在 10 000 ×g 条件下离心 20 min, 收集上清, 经 0.45 μm 滤器滤过后, 利用 Ni²⁺亲和色谱, 使用 Ni-Agarose His 标签蛋白纯化试剂盒纯化 RgGGPPS2 重组蛋白, 用 SDS-PAGE 检测不同咪唑浓度 (50、100、200、300、500 mmol/L) 洗脱目的蛋白, 确定最佳咪唑洗脱浓度, 收集单一目的蛋白条带的样品经过透析和超滤, 得到纯化的 RgGGPPS2 重组蛋白。

1.7 RgGGPPS2 基因荧光定量 PCR (Real-time PCR) 分析

为了检测 RgGGPPS2 基因在地黄不同组织部位的表达情况, 利用 Real-time PCR 检测 RgGGPPS2 基因在地黄野生种和栽培种“北京 3 号”的根、茎、叶中的表达量, 地黄各组织的 cDNA 稀释 10 倍作为 Real-time PCR 反应模板。以 RgTIP41 基因^[16]为

内参 (GenBank 登录号为 KT306007), Real-time PCR 的反应体系为: $2\times$ QuantiNova SYBR Green PCR Master Mix 10 μ L、正反向引物各 0.4 μ L、cDNA 模板 2 μ L、加 ddH₂O 补至 20 μ L, 反应程序为: 95 $^{\circ}$ C、3 min, 95 $^{\circ}$ C、10 s, 56 $^{\circ}$ C、20 s, 72 $^{\circ}$ C、30 s, 35 个循环, 每个样品重复 3 次, 根

据溶解曲线判断 Real-time PCR 产物的特异性, 相对定量分析采用 $2^{-\Delta\Delta C_t}$ 方法^[17], 所得实验数据用统计分析软件 SPSS 13.0 进行单因素方差分析, 以地黄野生种根的 C_t 值的平均数设为 1, 进行 RgGGPPS2 基因的组织特异性分析, Real-time PCR 引物序列见表 1。

表 1 基因克隆和荧光定量 PCR 引物

Table 1 Sequences of primers designed for gene clone and real-time PCR

引物用途	引物名称	引物序列 (5'→3')
基因克隆	RgGGPPS2-Exp-F	CGGAATTCATGCTCGAAAAAGCCA(下划线部分为 <i>Eco</i> R I 酶切位点)
	RgGGPPS2-Exp-R	CCGCTCGAGTTAATTATCCCTATAA(下划线部分为 <i>Xho</i> I 酶切位点)
Real-time PCR 内参基因引物	TIP41-F	ATTGGGTAGATTGCCAGGAG
	TIP41-R	CCATTGTCAGCCAATTCATC
Real-time PCR RgGGPPS2 引物	RgGGPPS2-rF	TGTTTCAGGTGGTGGATGAT
	RgGGPPS2-rR	TCCCAAGAAGCTTCGGATAC

2 结果与分析

2.1 RgGGPPS2 基因的克隆

通过分析本实验室前期获得的地黄转录组数据, 除了 RgGGPPS1 以外, 还有一条注释为 GGPPS 的 unigene c45479 大小为 1 002 bp, 使用 NCBI ORF Finder 检测发现其包含一个完整的开放阅读框, 将序列提交到 NCBI blastp 进行序列比对, 结果显示其与芝麻 *Sesamum indicum* L.、茉莉 *Jasminum sambac* (L.) Ait 等植物的 GGPPS 蛋白的氨基酸序列一致性较高 (>90%)。根据该基因的 ORF 序列信息, 设计一对特异性引物 (RgGGPPS2-Exp-F 和 RgGGPPS2-Exp-R, 表 1), 以反转录得到的地黄野生种根 cDNA 为模板, 进行 PCR 扩增, PCR 产物约为 900 bp, 电泳结果如图 1 所示, 测序后得到正确序列, 经 NCBI ORF Finder 检测发现该序列包含一个完整的 ORF, 大小为 867 bp, 编码 289 个氨基酸, 命名为 RgGGPPS2, 该基因的序列信息已提交到 NCBI GenBank, 登录号为 KU258807。

2.2 RgGGPPS2 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 提供的 ProtParam tool 预测 RgGGPPS2 蛋白的理化性质, RgGGPPS2 蛋白的相对分子质量为 31 140, 等电点是 5.18, 分子式为 C₁₂₆₂H₂₂₂₈N₃₇₈O₄₂₂S₁₅, 带负电荷的残基 (Asp+Glu): 44, 带正电荷的残基 (Arg+Lys): 32, 不稳定系数是 38.89, 说明该蛋白属于稳定蛋白。

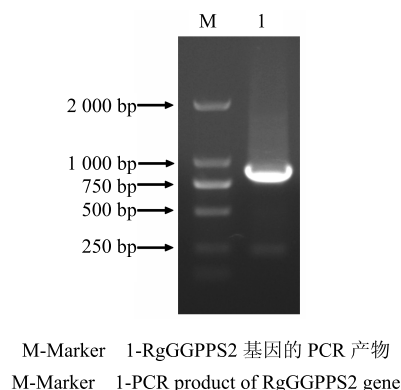


图 1 PCR 扩增 RgGGPPS2 基因

Fig. 1 PCR amplification of RgGGPPS2 gene

RgGGPPS2 基因 ORF 的核苷酸序列及对应氨基酸序列如图 2 所示。

2.2.2 保守结构域预测 用 InterProScan 预测 RgGGPPS2 蛋白的保守结构域: 多聚异戊烯基合成相关酶结构域 (IPR017446) 位于 1~288 位氨基酸序列、多聚异戊烯基合成酶结构域 (IPR000092) 位于 24~252 位氨基酸序列、异戊二烯合成酶结构域 (isoprenoid synthase domain, IPR008949) 位于 3~266 氨基酸序列, 说明 RgGGPPS2 蛋白属于异戊二烯合成酶家族的成员。

2.2.3 信号肽、亚细胞定位和跨膜区预测 根据 SignalP 预测, RgGGPPS2 蛋白不含信号肽, TargetP 预测显示 RgGGPPS2 蛋白可能位于细胞质中,

TMHMM 预测结果显示 RgGGPPS2 蛋白不含跨膜区。

2.2.4 三级结构预测 用 SWISS-MODEL 预测 RgGGPPS2 蛋白的三维结构, 以辣薄荷 *Mentha piperita* L. 的牻牛儿苗基焦磷酸合酶 (PDB ID:

3krp) 的晶体结构为模板, 序列的一致性为 85.07%, 在第 1~288 位氨基酸残基建模, 模型覆盖率为 100%, 根据预测结果 RgGGPPS2 蛋白可能以单体的形式发挥作用 (图 3)。

```

1 ATGCTCGAAAAAGCCATTCCGTCAACAAAGCCTTAGAAGACGCGGTCTCACTAAGAGATCCCTTAAAAATCCACGAA
  M L E K A N S V N K A L E D A V S L R D P L K I H E
79 TCCATGAGATACTCCCTTCTCGCAGGCGGTAAAGAGTGAGACCCATGCTGTGCATCGCCGCTCGCAACTCGTCGGC
  S M R Y S L L A G G K R V R P M L C I A A C E L V G
157 GGAGATGAATCCACCGCCATGCCGGTGCCTGCGCGGTGGAGATGATCCACACCATGTCCCTAATGCACGACGACCTC
  G D E S T A M P A A C A V E M I H T M S L M H D D L
235 CCCTGTATGGACACGACGATTGCGCAGAGGTAAGCCACGAACCAAGATCTACGGCGAGGACGTGGCTGTTTTTA
  P C M D N D D L R R G K P T N H K I Y G E D V A V L
313 GCCGGGATGCGCTTTTAGCATTTTCGTTCCAGCATATCGCCATAGCCACGAAAGGGGTTTCTCAGAAAGGATTGTT
  A G D A L L A F S F E H I A I A T K G V S E R I V
391 AGAGTGATTGGTGAATTGGCTAAGTGTATTGGTTCAGAAAGGCTTGTGGCAGGGCAAGTGGTGATATTGCTCAGAA
  R V I G E L A K C I G S E G L V A G Q V V D I C S E
469 GGGATGACTGAGGTTGGGTTAGACCATTTGGAGTTTATTCATCTGCATAAACTGCAGCTTTATTGGAGGGATCAGTG
  G M T E V G L D H L E F I H L H K T A A L L E G S V
547 GTTTTGGTGCAATTTTAGGTGGTGTCTAATGATGAAGAAGTTGAAAACTTAGAAAAATTGCAAGAAATATTGGGCTT
  V L G A I L G G A N D E E V E K L R K F A R N I G L
625 TTGTTTCAGGTGGTGGATGATATCTTGATGTCACAAATCATCTCAAGAATTAGGGAAAAGTCTGGTAAAGATTG
  L F Q V V D D I L D V T K S S Q E L G K T A G K D L
703 GTGGCTGATAAACTACGTATCCGAAGCTTCTTGGGATCGAAAAATCGAGGGATTTTGTGAGAAATGAATAGGGAA
  V A D K T T Y P K L L G I E K S R D F A E K L N R E
781 GCTCAAGAACAGCTTGTGGATTGATTCTGTTAGAGTGTCTCTTGTATTGCTTTAGCTAACTATATTGCTTATAGG
  A Q E Q L V G F D S V R A A P L I A L A N Y I A Y R
859 GATAATTAA
      D N *

```

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

The marker "*" represents termination codon, the left number indicates nucleotide position, the right number indicates amino acid position

图 2 RgGGPPS2 基因的 ORF 序列及对应的氨基酸序列

Fig. 2 Nucleotide sequence and amino acid sequence of RgGGPPS2 gene ORF

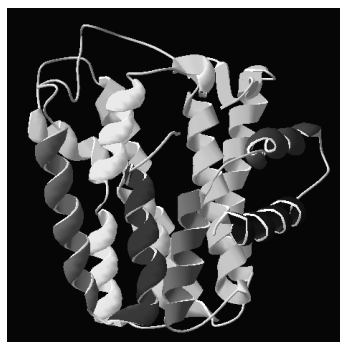


图 3 RgGGPPS2 蛋白质三级结构预测

Fig. 3 Prediction of three-dimensional structure of RgGGPPS2 protein

2.2.5 序列比对分析 将 RgGGPPS2 蛋白的氨基酸序列通过 NCBI blastp 比对 nr 非冗余蛋白数据库, 发现 RgGGPPS2 蛋白与芝麻 (SiGGPPS, XP_011097207)、长春花 *Catharanthus roseus* (Linn.) G. Don (CrGGPPS, AEI53622) 等植物的 GGPPS 蛋白氨基酸序列的一致性较高, 分别为 91%、89%, 与拟南芥的 AtGGPPS1 (NP_195399)、AtGGPPS2 (AAO63392)、AtGGPPS4 (NP_179454) 和 AtGGPPS12 (NP_195558) 等 GGPPS 蛋白氨基酸序列的一致性分别为 62%、52%、50%、35%, 与

RgGGPPS1 蛋白的一致性仅为 35%。将这 8 个 GGPPS 蛋白的氨基酸序列进行比对, 发现 GGPPS 蛋白 N 端变异较大, 中部比较保守, 都含有异戊二烯合成酶家族 2 个富含天冬氨酸的保守序列 FARM (first aspartate-rich motif, DDxxxxD) 和 SARM (second aspartate-rich motif, DDxxD), 这 2 个保守序列是和底物以及产物 GGPP 结合的部位, 对 GGPPS 催化活性的维持是必须的^[18]。在 RgGGPPS2 蛋白中这 2 个保守序列分别是 DDLPCMD 和 DDILD, 但是在 RgGGPPS1 和 AtGGPPS12 中 SARM 序列的最后一个氨基酸发生了突变, 序列变为 DDILE, 而 SARM 序列的突变则会引起 GGPPS 催化效率和最大反应速率的显著下降^[18]。此外只有在 RgGGPPS1 和 AtGGPPS12 中存在 2 个 CxxxC 基序, RgGGPPS2 和其他的 GGPPS 蛋白中只存在一个 CxxxC 基序 (图 4), 这 2 个 CxxxC 基序参与牻牛儿苗基焦磷酸合酶 (GPPS) 的 2 个亚基的相互作用^[19], 在这 8 个 GGPPS 蛋白中 RgGGPPS1 和 AtGGPPS12 的序列一致性最高为 65%, 与 RgGGPPS2 的一致性仅为 35%, 根据序列比对分析结果 RgGGPPS1 可能和 RgGGPPS2 以及其他植物的 GGPPS 蛋白的功能不同, 而与 AtGGPPS12 功能类似。

RgGGPPS1	MIFS.TVITSSPHVCLPKATRP.....IFRRT...IRCSATISIS.....	35
RgGGPPS2		0
SiGGPPS	MSALVNPIITWSRTSSVKDVNGSRFR.....STPFILPHPLHSNLFVSFCFSDPTRKPV.SFSLTV	61
CrGGPPS	.MSFVNSITTWVPAQSIYCLNGRSS.....SMRNLCHPLKNQLPISFFLSGTIRKPIFSCSRLSI	61
AtGGPPS1	.MASVT.LGSWIVVHHHHHHHPSSIL.....TKSRSRSCPTLTLPISFRS...KRTVSSSSSIYS	56
AtGGPPS2	MEPQILFLYLSLFLSLNFFFTNLKP.....RLVRLFPQSLESRVKTALLSRKEVAFAFLDSSIPVEDE	62
AtGGPPS4	MEAQNIFLYLLVFLSLHFVFTTLKGRLSPANTRLRLRLHIPIKSPVAAAIFARKDTREFLDSSIKLVN	70
AtGGPPS12	MLFSGSAIPLSSFCSLPEKPHT.....LPMKLSPAAIRSSSSSAPGS.....	42
Consensus		
RgGGPPS1	PREFDKTWTDLIKEIDQKLDPAIPVQYENQIYEMRYSVLAKG	79
RgGGPPS2	CREATED.....MLEKANSVNKALEDAVSLRDPKTHESMRYSLLAG.	42
SiGGPPS	SAVLVKQKEESK.....IEKAPTDFKNNMLGKASSVNKALEBAILLRPKTHESMRYSLLAG.	122
CrGGPPS	SAITTKETQEESESSEKKEVAFSSSTSSFDKAMMIGKANSVNKALEDAVLRPKTHESMRYSLLAG.	130
AtGGPPS1	SSVVTKEENLRQS.....EPSSFDFMSYIITKAELVNKLDSAVPLRPKTHESMRYSLLAG.	114
AtGGPPS2	EGEEREDEEGGI.....VSNANFTFEFDPYMSKAESVNKALEBAIPVGPETHESMRYSLLAG.	123
AtGGPPS4	E.....EDDFG.....FSDFKPYMISKAETINRLDEAIPLEPETHESMRYSLLAG.	119
AtGGPPS12	LNFDLRTWTTLITEINQKLDPAIPVKHAGIYEMRYSVLAKG	86
Consensus		
RgGGPPS1	AKRAPVNVQVAAACLFEGGNRLAEPACAEEMVHAASLIHDDLPCMDIPSRRGKESNHTIYGVDMAILA	149
RgGGPPS2	GKRVREMICTAACELVGGDESTAMPAAACAVEMIHMSILHDDLPCMDIDDIRRGKPTNHKIVGEDVAVLA	112
SiGGPPS	GKRVREMICTAACELVGGDESTAMPAAACAVEMIHMSILHDDLPCMDIDDIRRGKPTNHKIVGEDVAVLA	192
CrGGPPS	GKRVREMICTAACELVGGDESTAMPAAACAVEMIHMSILHDDLPCMDIDDIRRGKPTNHKIVGEDVAVLA	200
AtGGPPS1	GKRVREMICTAACELVGGDESTAMPAAACAVEMIHMSILHDDLPCMDIDDIRRGKPTNHKIVGEDVAVLA	186
AtGGPPS2	GKRVREMICTAACELVGGDESTAMPAAACAVEMIHMSILHDDLPCMDIDDIRRGKPTNHKIVGEDVAVLA	193
AtGGPPS4	GKRVREMICTAACELVGGDESTAMPAAACAVEMIHMSILHDDLPCMDIDDIRRGKPTNHKIVGEDVAVLA	189
AtGGPPS12	AKRAPVNVQVAAACLFEGGNRLAEPACAEEMVHAASLIHDDLPCMDIPSRRGKESNHTIYGVDMAILA	156
Consensus	kr p c a c l gg a aca em h sl d d l p c m d d r r g k p h g a l	
RgGGPPS1	GDALFPLGFHIVSHPTDLVPETRLRVLAEIARAVGSTMAAGQFVLEGGP.....NAVDFVQE	211
RgGGPPS2	GDALLAFSFHIAIATKG...VSSERIVRVICELAKCIGSEGLVAGQVVDICSEG...MTEVGLDHLFIHL	178
SiGGPPS	GDALLAFSFHIVATATKG...VSSARIVRVICELAKCIGSEGLVAGQVVDICSEG...LSEVGLDHLFIHL	258
CrGGPPS	GDALLAFSFHIAIATKG...VSSERIVRVICELAKCIGSEGLVAGQVVDICSEG...IADVGLHLEFIHL	266
AtGGPPS1	GDALLSFAPFHLASATSSDVSPVRVRAVGEIAKATCTEGLVAGQVVDISSEGLDLNDVGLHLEFIHL	254
AtGGPPS2	GCALLSLAFBHTTAEIS...SERMVAVREIARSICTRGLVAGQVVDISSEGLDLNDVGLHLEFIHL	259
AtGGPPS4	GCALLSLAFBHTTAEIS...SKMVRVAVGEIAKATCTEGLVAGQVVDISSEGLDLNDVGLHLEFIHL	255
AtGGPPS12	GDALFPLGFHIVSHPTDLVPETRLRVLAEIARAVGSTMAAGQFVLEGGP.....NAVDFVQE	218
Consensus	g al f h p p d l v p r a t i l r l i t e a r t v g s t m a a g q v v l e g g p f p s f v q e	
RgGGPPS1	KKFGEMGCSAVCGGLLAGASDDIQRIRNYGRAVGMVQVDDILEAKTETEENKKKGK.....SY	273
RgGGPPS2	HKTAALLTGSVVIGAILGGANDEEVEKLRKFARNIGLQVQVDDILDVTKSSQELGKTAGKDLVADKTTY	248
SiGGPPS	HKTAALLTGSVVIGAILGGANDEEVEKLRKFARNIGLQVQVDDILDVTKSSQELGKTAGKDLVADKTTY	328
CrGGPPS	HKTAALLTGSVVIGAILGGANDEEVEKLRKFARNIGLQVQVDDILDVTKSSQELGKTAGKDLVADKTTY	336
AtGGPPS1	HKTAALLTGSVVIGAILGGANDEEVEKLRKFARNIGLQVQVDDILDVTKSSQELGKTAGKDLVADKTTY	324
AtGGPPS2	HKTAALLTGSVVIGAILGGANDEEVEKLRKFARNIGLQVQVDDILDVTKSSQELGKTAGKDLVADKTTY	329
AtGGPPS4	HKTAALLTGSVVIGAILGGANDEEVEKLRKFARNIGLQVQVDDILDVTKSSQELGKTAGKDLVADKTTY	325
AtGGPPS12	KKFGEMGCSAVCGGLLAGASDDIQRIRNYGRAVGMVQVDDILEAKTETEENKKKGK.....SY	273
Consensus	k l l v t g g g l l g a s d d i q r i r n y g r a v g m v q v d d i l e a k t e t e e n k k k g k s y	
RgGGPPS1	VSLYGVDK.AMEVAEDIRSOAKKELDALE.KYG..EKVLPYISFVDYAADRGFSFAAP	327
RgGGPPS2	PKLLGLEK.SRDFAEKINRBAQEOLVGFD....SVRAAPLIALANYIAYRDN....	295
SiGGPPS	PKLLGLEK.SRDLAERINTEBAQEOLVGFD....PIKAAPLIALANYIAYRDN....	375
CrGGPPS	PKLLGLEK.SRDFAEKINRBAQEOLVGFD....PEKAAPLIALANYIAYRDN....	383
AtGGPPS1	PKIMGLEK.SRDFAEKINRBAQEOLVGFD....SDKVAPLIALANYIAYRDN....	371
AtGGPPS2	PKLLGLEK.SKEFVKRTKDKARHQLQGES....SEKVAPIVATTTFLANRNK....	376
AtGGPPS4	PKVIGVEK.SKEFVKRTKDKARHQLQGES....SDKVAPLIALANYIAYRDN....	372
AtGGPPS12	CAEKGMMEAEELKEKAKKELQVFDNKYGGGDTLVPYTFVVDYAAH...HFLLP	325
Consensus	g k l a l p l a r	

2 个富含天冬氨酸的保守序列用黑色框标出, 2 个 CxxxC 基序用虚线标出; 一致性为 100% 的氨基酸序列用黑色阴影表示, 一致性为 75% 以上的氨基酸序列用灰色阴影表示

The two Asp-rich conserved sequences were boxed in black, the two CxxxC motifs were marked with dashed lines; Identity for 100% of amino acid sequence highlighted in black, identity for more than 75% of amino acid sequence highlighted in gray

图 4 RgGGPPS 蛋白与其他植物 GGPPS 蛋白的多序列比对

Fig. 4 Multiple sequence alignment of RgGGPPS and GGPPS proteins from other plant species

2.2.6 系统进化分析 GGPPS 蛋白在植物和微生物中都普遍存在, 从 NCBI blastp 的比对结果中, 选取不同物种来源的 GGPPS 蛋白序列和 RgGGPPS1、RgGGPPS2 一起, 包括药用植物丹参、三七、银杏等, 模式植物拟南芥、水稻等, 苔藓植物和蕨类植物中完成基因组测序的小立碗藓和江南卷柏, 以及绿藻和蓝藻等构建系统进化树 (图 5)。从结果中可以看出, 不同物种来源的 GGPPS 蛋白

在进化上归属不同的分支, 分别聚为被子植物、蕨类植物、苔藓植物、裸子植物、绿藻、蓝藻等几大类, RgGGPPS2 与芝麻、长春花等植物亲缘关系较近, 与丹参、三七、拟南芥 (AtGGPPS1、AtGGPPS2、AtGGPPS4) 等植物聚为一类, 都属于被子植物的双子叶植物; 水稻、玉米、高粱等禾本科植物聚为一类, 都属于被子植物的单子叶植物; 小立碗藓和江南卷柏在进化上较为原始, 分别聚为 2 类, 属于

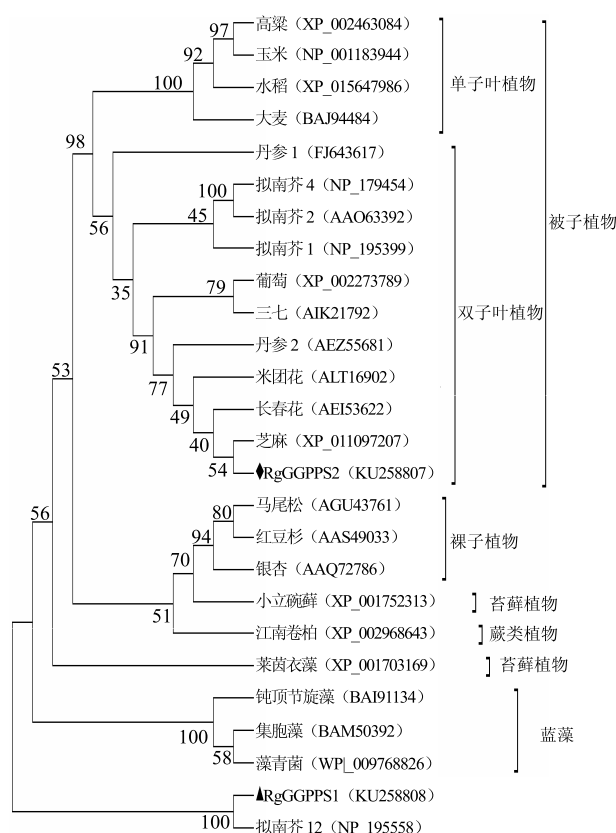


图 5 RgGGPPS2 蛋白与其他物种 GGPPS 蛋白的系统进化树分析

Fig. 5 Phylogenetic tree analysis of RgGGPPS2 protein and GGPPS proteins from other species

苔藓植物和蕨类植物；银杏、马尾松和红豆杉等植物聚为一类，属于裸子植物；水生藻类聚为 2 类，分别是绿藻和蓝藻；而 RgGGPPS1 和 AtGGPPS12 单独聚为一类（图 5），未与 RgGGPPS2 和其他植物的 GGPPS 成员聚为一类，说明 RgGGPPS1 与 RgGGPPS2 以及其他植物的 GGPPS 同源性较低，而与 AtGGPPS12 同源性较高，根据以前的研究，在高等植物中，GGPPS 通常由旁系同源基因编码^[10]，RgGGPPS1 和 RgGGPPS2 也属于旁系同源基因，RgGGPPS1 可能与 RgGGPPS2 功能不同，而与 AtGGPPS12 功能类似。

2.3 RgGGPPS2 原核表达载体的构建

用 *EcoR* I 和 *Xho* I 对测序正确的 pMD19-T-RgGGPPS2 质粒和原核表达载体 pET-32a 进行双酶切，切胶回收基因片段和载体片段，将基因片段和载体片段用 T₄ DNA Ligase 过夜连接，构建原核表达载体 pET-32a-RgGGPPS2，将连接产物

转化大肠杆菌 BL21 (DE3) 菌株，挑单克隆提取质粒，用 *EcoR* I 和 *Xho* I 进行双酶切鉴定，结果见图 6，从结果中可以看出有 900 bp 左右的目的基因片段，说明 RgGGPPS2 基因已插入到载体 pET-32a 中。将酶切正确的单克隆送测序公司测序，测序结果表明插入载体中的基因序列与目的基因序列完全一致，未发生碱基突变、缺失或插入，说明 RgGGPPS2 基因的原核表达载体构建成功。

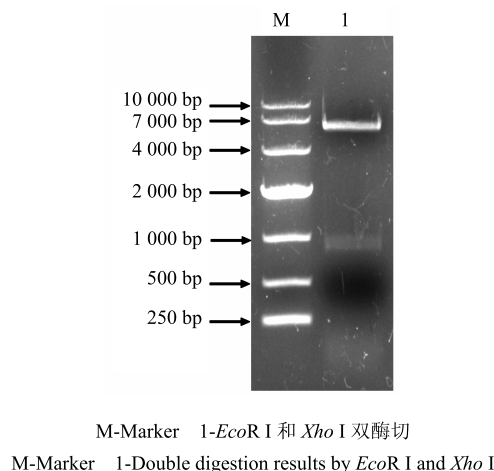
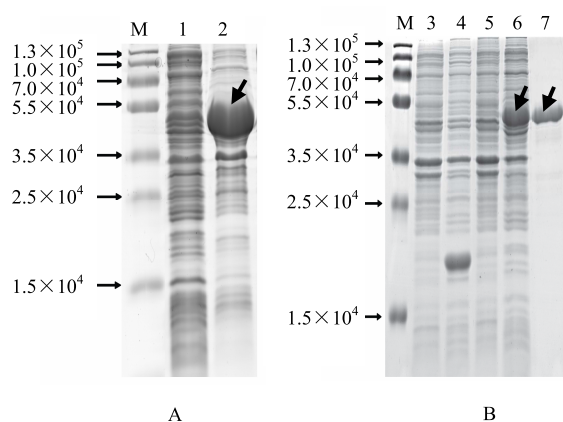


图 6 原核表达载体 pET-32a-RgGGPPS2 的双酶切鉴定

2.4 RgGGPPS2 重组蛋白的原核表达与纯化

将测序正确的大肠杆菌 BL21 (DE3) 单克隆（含 pET-32a-RgGGPPS2）振荡培养，当大肠杆菌生长至 A_{600} 达到 0.6 时，在 28 °C、0.4 mmol/L IPTG、150 r/min 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-RgGGPPS2 质粒的菌株经 IPTG 诱导后，在约 5.1×10^4 处出现目的蛋白条带，而未经 IPTG 诱导的样品在 5.1×10^4 处没有目的蛋白条带（图 7-A）。RgGGPPS2 基因编码的蛋白预测相对分子质量为 31 140，而当 pET-32a-RgGGPPS2 质粒在大肠杆菌中表达时，表达的 RgGGPPS2 重组蛋白的 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 51 140，所以在 SDS-PAGE 结果中约 5.1×10^4 处出现目的蛋白条带，与预测的 RgGGPPS2 重组蛋白分子量相符。超声破碎菌体后取上清和沉淀进行蛋白可溶性检测，结果显示 RgGGPPS2 重组蛋白主要以包涵体形式存在于沉淀中（图 7-A）。含有空载体 pET-32a 的菌株经 IPTG 诱导后，未在 5.1×10^4 处出现条带，而



A-RgGGPPS2 重组蛋白的可溶性分析 B-RgGGPPS2 重组蛋白的原核表达与纯化 M-Marker 1-IPTG 诱导含 pET-32a-RgGGPPS2 质粒 *E. coli* 的可溶性蛋白 2-IPTG 诱导含 pET-32a-RgGGPPS2 质粒 *E. coli* 的沉淀 3-未诱导的含 pET-32a 空载体的 *E. coli* 菌株 4-IPTG 诱导的含 pET-32a 空载体的 *E. coli* 菌株 5-未诱导的含 pET-32a-RgGGPPS2 质粒的 *E. coli* 菌株 6-IPTG 诱导的含 pET-32a-RgGGPPS2 质粒的 *E. coli* 菌株 7-纯化的 RgGGPPS2 重组蛋白; 箭头显示为 RgGGPPS2 重组蛋白

A-solubility analysis of recombinant RgGGPPS2 protein B-prokaryotic expression and purification of recombinant RgGGPPS2 protein M-Marker 1-soluble protein from IPTG induced *E. coli* containing pET-32a-RgGGPPS2 2-insoluble protein from IPTG induced *E. coli* containing pET-32a-RgGGPPS2 3-non-induced *E. coli* containing pET-32a used as control 4-*E. coli* containing pET-32a under IPTG inducement 5-non-induced *E. coli* containing pET-32a-RgGGPPS2 6-*E. coli* containing pET-32a-RgGGPPS2 under IPTG inducement 7-purified recombinant RgGGPPS2 protein; The arrows showed the recombinant RgGGPPS2 proteins

图 7 SDS-PAGE 分析 RgGGPPS2 重组蛋白

Fig. 7 SDS-PAGE analysis of recombinant RgGGPPS2 protein

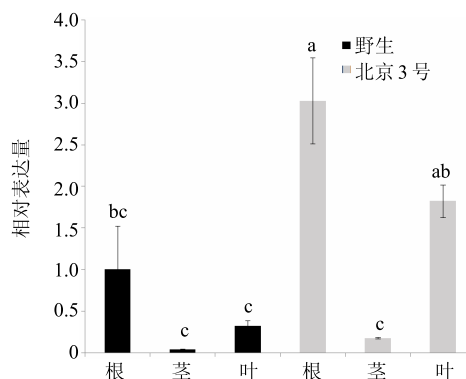
在 2.0×10^4 处出现条带, 因为 pET-32a 载体的标签序列 (Trx-Tag、His-Tag 和 S-Tag) 在 IPTG 诱导时也会表达融合蛋白, 大小约为 2.0×10^4 (图 7-B), 结果表明, RgGGPPS2 重组蛋白成功在大肠杆菌 BL21 (DE3) 菌株中表达。

根据 RgGGPPS2 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, 将沉淀重悬于含 8 mol/L 尿素的结合缓冲液, 使其充分溶解, 根据 Ni^{2+} 亲和色谱, 按照 Ni-Agarose His 标签蛋白纯化试剂盒说明书纯化 RgGGPPS2 重组蛋白, 最终得到纯化的 RgGGPPS2 重组蛋白 (图 7-B)。

2.5 RgGGPPS2 基因表达模式分析

Real-time PCR 结果显示, RgGGPPS2 基因在地黄野生种和栽培种“北京 3 号”的各组织中均有表达, 而且都是根中表达量最高, 其次是叶, 茎中最

低, 说明 RgGGPPS2 基因在地黄中的表达具有组织特异性, 以在根中表达为主 (图 8), 在栽培种“北京 3 号”的根、茎、叶中, RgGGPPS2 基因的表达量均高于野生种 (图 8)。



不同字母表示野生种和栽培种之间存在显著差异 ($P < 0.05$)

Different letters indicate significant difference among wild species and cultivar ($P < 0.05$)

图 8 RgGGPPS2 基因在地黄根、茎和叶的相对表达量

Fig. 8 Relative expression level of RgGGPPS2 gene in root, stem and leaf of *R. glutinosa*

3 讨论

本实验室在前期的研究中克隆了 RgGGPPS1 基因^[15], 通过序列分析和组织特异性分析, 发现 RgGGPPS1 可能定位在叶绿体中, 在根中表达量最高, 推测 RgGGPPS1 功能与拟南芥 AtGGPPS1 和 AtGGPPS3 类似可能参与赤霉素、叶绿素等次生代谢产物的合成。本研究克隆了 RgGGPPS2 基因, 其编码蛋白的氨基酸序列与 RgGGPPS1 的一致性只有 35%, 虽然 RgGGPPS2 和 RgGGPPS1 一样都含有异戊二烯合成酶家族 2 个富含天冬氨酸的保守序列 FARM 和 SARM, 但是在 RgGGPPS1 中 SARM 的最后一个氨基酸发生了突变, SARM 序列变为 DDILE, SARM 序列突变会引起 GGPPS 催化效率的下降, 而且 RgGGPPS1 含有 2 个 CxxxC 基序, RgGGPPS2 和其他植物的 GGPPS 蛋白只有一个 CxxxC 基序, 说明 RgGGPPS1 可能与 RgGGPPS2 发挥不同的功能。根据序列比对分析 (图 4) 和系统进化分析 (图 5) 结果, RgGGPPS1 和 RgGGPPS2 序列一致性较低且没有聚为一类, 而是和 AtGGPPS12 序列一致性最高且单独聚为一类, 说明 RgGGPPS1 和 RgGGPPS2 属于旁系同源基因, RgGGPPS1 发挥与 RgGGPPS2 不同的功能, 可能与 AtGGPPS12 功能类似。前期的研究中发现

AtGGPPS12 没有 GGPP 合成酶的活性,可能具有牻牛儿苗 GPPS 小亚基的功能,因为 AtGGPPS12 的 SARM 序列发生突变而且含有两个 CxxxC 基序,而拟南芥的其他 GGPPS 家族成员只有 1 个 CxxxC 基序,2 个 CxxxC 基序对于 GPPS 2 个亚基之间的相互作用极为重要^[19],序列比对和系统进化分析结果显示 RgGGPPS1 与 AtGGPPS12 功能相似,可能 RgGGPPS1 没有 GGPP 合成酶活性而是具有 GPPS 小亚基的功能。

根据生物信息学和表达模式分析,发现 RgGGPPS2 可能定位于细胞质中,在根中表达量最高,其次是叶、茎中最低,位于细胞质中 RgGGPPS2 能够利用来源于甲羟戊酸 (mevalonate, MVA) 途径的 IPP 合成 GGPP,为萜类化合物的合成提供前体。RgGGPPS2 的氨基酸序列含有 2 个完整的异戊二烯合成酶家族富含天冬氨酸的保守序列 FARM 和 SARM,以及一个 CxxxC 基序,这与拟南芥的其他 GGPPS 家族成员的氨基酸序列组成相似,而且系统进化树分析显示 RgGGPPS2 和 AtGGPPS2、AtGGPPS4 聚为一类 (图 5),推测其功能可能与拟南芥的 AtGGPPS2 和 AtGGPPS4 类似,参与细胞质中多聚类异戊二烯和寡聚异戊烯醇 (如多萜醇、二萜等) 的合成,以及蛋白质的异戊二烯化^[20],可以为地黄萜类化合物的合成提供前体物 GGPP。Real-time PCR 结果显示 RgGGPPS2 在地黄在根中表达量最高、其次是叶、茎中最低,而且在栽培种“北京 3 号”的根、茎、叶中 RgGGPPS2 的表达量均高于野生种,这种表达模式可能和地黄以块根作为入药部位有关,膨大的块根中积累较多的萜类成分而在叶片中积累较少^[21],地黄块根的膨大程度直接决定地黄的品质和产量,栽培地黄的根在成熟期发育为膨大块根,积累较多的萜类成分,品质和产量较高,而野生地黄的根在成熟期则未能发育为膨大的块根^[22],所以在栽培地黄“北京 3 号”的根、茎、叶中 RgGGPPS2 的表达量均高于野生地黄,这些结果说明 RgGGPPS2 可能在萜类化合物的合成过程中起关键作用。

本研究通过构建原核表达载体 pET-32a-RgGGPPS2,成功在大肠杆菌 BL21 (DE3) 菌株中表达了 RgGGPPS2 重组蛋白,通过对重组蛋白进行 SDS-PAGE 分析,发现重组蛋白 RgGGPPS2 主要以不溶的包涵体形式存在。当真核生物基因在原核表达系统大量表达时,有时候无法形成有活性

的重组蛋白,往往以包涵体形式存在,导致包涵体形成的原因较多,如,由于目的蛋白表达量过高,蛋白合成速度太快,使重组蛋白没有足够的时间折叠,二硫键无法正确配对;缺乏折叠过程中需要的酶和辅助因子 (如分子伴侣),使目的蛋白无法正确折叠;由于目的蛋白的序列中含有分泌序列或跨膜序列,也导致重组蛋白以不溶的包涵体形式存在^[23]。RgGGPPS2 蛋白可能由于表达量过高或者缺乏辅助因子,最终导致 RgGGPPS2 重组蛋白在大肠杆菌中以包涵体的形式存在。下一步研究可以通过包涵体溶解与复性的方式,或者利用酵母真核表达系统表达 RgGGPPS2 蛋白,获得有活性的 RgGGPPS2 蛋白,在体外研究其酶学特征及催化活性,也可以将纯化的 RgGGPPS2 重组蛋白制备抗体,在蛋白水平研究 RgGGPPS2 的表达模式,为研究 RgGGPPS2 基因在地黄萜类化合物生物合成途径中的功能奠定了基础。

参考文献

- [1] 夏至,王璐静,黄勇,等.地黄属植物 DNA 条形码鉴定及地黄栽培起源研究 [J]. 中草药, 2016, 47(4): 648-654.
- [2] 李红伟,孟祥乐.地黄化学成分及其药理作用研究进展 [J]. 药物评价研究, 2015, 38(2): 218-228.
- [3] 董焱,陈长勋.梓醇药理作用的研究进展 [J]. 中成药, 2013, 35(5): 1047-1051.
- [4] Eisenreich W, Bacher A, Arigoni D, et al. Biosynthesis of isoprenoids via the non-mevalonate pathway [J]. *Cell Mol Life Sci*, 2004, 61(12): 1401-1426.
- [5] Capell T, Christou P. Progress in plant metabolic engineering [J]. *Curr Opin Biotechnol*, 2004, 15(2): 148-154.
- [6] Zhang M, Su P, Zhou Y J, et al. Identification of geranylgeranyl diphosphate synthase genes from *Tripterygium wilfordii* [J]. *Plant Cell Rep*, 2015, 34(12): 2179-2188.
- [7] Beck G, D Coman, E Herren, et al. Characterization of the GGPP synthase gene family in *Arabidopsis thaliana* [J]. *Plant Mol Biol*, 2013, 82(4): 393-416.
- [8] Aharoni A, Giri A P, Deuerlein S, et al. Terpenoid metabolism in wild-type and transgenic *Arabidopsis* plants [J]. *Plant Cell*, 2003, 15(12): 2866-2884.
- [9] Laskaris G, Bounkhay M, Theodoridis G, et al. Induction of geranylgeranyl diphosphate synthase activity and taxane accumulation in *Taxus baccata* cell cultures after elicitation by methyl jasmonate [J]. *Plant Sci*, 1999, 147(1): 1-8.
- [10] Coman D, Altenhoff A, Zoller S, et al. Distinct

- evolutionary strategies in the GGPPS family from plants [J]. *Front Plant Sci*, 2014, 5: 1-12.
- [11] 李 锋, 李 明, 金立锋, 等. 烟草牻牛儿苗基牻牛儿苗焦磷酸合成酶基因的克隆及分析 [J]. 烟草科技, 2012, (5): 60-64.
- [12] 化文平, 宋双红, 智 媛, 等. 丹参 SmGGPPS3 基因的克隆及表达分析 [J]. 植物科学学报, 2014, 32(1): 50-57.
- [13] 马利刚, 赵 乐, 李英超, 等. 北葶苈子 GGPS 基因的克隆、序列分析与原核表达 [J]. 世界科学技术—中医药现代化, 2015, 17(3): 485-491.
- [14] Okada K, Saito T, Nakagawa T, *et al.* Five geranylgeranyl diphosphate synthases expressed in different organs are localized into three subcellular compartments in *Arabidopsis* [J]. *Plant Physiol*, 2000, 122(4): 1045-1056.
- [15] 赵 乐, 史晶晶, 马利刚, 等. 地黄 GGPPS1 基因克隆及表达分析 [J]. 西北植物学报, 2016, 36(5): 888-895.
- [16] 侯维海, 孙 鹏, 陈全家, 等. 地黄实时定量 PCR 内参基因的筛选 [J]. 中国农学通报, 2011, 27(17): 76-82.
- [17] Livak K J, Schmittgen T D. Analysis of relative gene expression data using Real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method [J]. *Methods*, 2001, 25(4): 402-408.
- [18] Joly A, Edwards P A. Effect of site-directed mutagenesis of conserved aspartate and arginine residues upon farnesyl diphosphate synthase activity [J]. *J Biol Chem*, 1993, 268(36): 26983-26989.
- [19] Wang G R, Dixon A. Heterodimeric geranyl (geranyl) diphosphate synthase from hop (*Humulus lupulus*) and the evolution of monoterpene biosynthesis [J]. *Proc Natl Acad Sci USA*, 2009, 106(24): 9914-9919.
- [20] Vranová E, Hirschhoffmann M, Grussem W. AtIPD: a curated database of *Arabidopsis* isoprenoid pathway models and genes for isoprenoid network analysis [J]. *Plant Physiol*, 2011, 156(4): 1655-1660.
- [21] Zhang Y L, Feng W S, Zheng X K, *et al.* Three new ursane-type triterpenes from the leaves of *Rehmannia glutinosa* [J]. *Fitoterapia*, 2013, 89(1): 15-19.
- [22] 王鹏飞, 李鑫宇, 李明杰, 等. 地黄块根膨大发生和驱动的组织观察及激素相关基因的调控分析 [J]. 中国中药杂志, 2014, 39(17): 3245-3253.
- [23] Singh S M, Panda A K. Solubilization and refolding of bacterial inclusion body proteins [J]. *J Biosci Bioeng*, 2005, 99(4): 303-310.