

· 药材与资源 ·

滇龙胆 *GrGPPS* 基因的克隆及其序列分析与原核表达王彩云¹, 李富生¹, 李涛², 李彩霞², 张晓东^{2*}, 王元忠^{3*}

1. 云南农业大学农学与生物技术学院, 云南 昆明 650201

2. 玉溪师范学院资源与环境学院, 云南 玉溪 653100

3. 云南省农业科学院药用植物研究所, 云南 昆明 650223

摘要: 目的 从滇龙胆 *Gentiana rigescens* 幼叶中克隆单萜化合物合成的关键酶牻牛儿基焦磷酸合成酶基因 *GrGPPS*, 进行序列特征分析和原核表达。方法 根据三年生滇龙胆转录组 *GrGPPS* 基因序列, 设计特异性引物, 通过 RT-PCR 扩增得到 *GrGPPS* cDNA 序列, 并进行 TA 克隆、测序及序列分析; 构建原核表达载体 pGEX-4T-1-*GrGPPS*, 转入 *Escherichia coli* Rosetta (DE3) 中, 在 37 °C、1.0 mmol/L IPTG 诱导下进行表达。结果 *GrGPPS* cDNA 全长 1 107 bp, 编码 369 个氨基酸; 序列分析表明, *GrGPPS* 基因是异戊烯基合成酶家族的成员; 氨基酸序列系统发育分析表明, *GrGPPS* 与金鱼草 *AmGPPS* 亲缘关系最近; 构建 pGEX-4T-1-*GrGPPS* 重组质粒, 获得稳定的 pGEX-4T-1-*GrGPPS* 原核表达体系。SDS-PAGE 结果表明所表达蛋白与预期蛋白大小一致。结论 克隆了 *GrGPPS* 基因, 建立 pGEX-4T-1-*GrGPPS* 稳定的原核表达体系, 为进一步纯化和鉴定 GPPS 蛋白并研究其结构和功能奠定基础。

关键词: 滇龙胆; 牻牛儿基焦磷酸合成酶基因; 基因克隆; 序列分析; 原核表达

中图分类号: R282.12 文献标志码: A 文章编号: 0253-2670(2014)14-2060-09

DOI: 10.7501/j.issn.0253-2670.2014.14.020

Cloning, sequence analysis, and prokaryotic expression of *GrGPPS* gene in *Gentiana rigescens*WANG Cai-yun¹, LI Fu-sheng¹, LI Tao², LI Cai-xia², ZHANG Xiao-dong², WANG Yuan-zhong³

1. College of Agriculture and Biological Technology, Yunnan Agriculture University, Kunming 650201, China

2. College of Resources and Environment, Yuxi Normal University, Yuxi 653100, China

3. Institute of Medicinal Plants, Yunnan Academy of Agricultural Sciences, Kunming 650223, China

Abstract: Objective To obtain the indispensable key enzyme involved in the monoterpene biosynthesis, the geranyl diphosphate synthase gene *GrGPPS* was cloned from *Gentiana rigescens*, and its sequence analysis and prokaryotic expression were performed. **Methods** According to the *GrGPPS* gene sequence of transcriptome of triennial *G. rigescens*, a pair of specific primers was designed, and the full length of cDNA sequences was obtained by RT-PCR. Then TA cloning, sequencing and sequence analysing were performed. Prokaryotic expression vector pGEX-T-1-*GrGPPS* was constructed and transformed into *Escherichia coli* Rosetta for expression under 37 °C and induced by 1 mmol/L IPTG. **Results** The *GrGPPS* cDNA had a length of 1 107 bp coding for 369 amino acids. Sequence analysis showed that *GrGPPS* was the member of “short-chain prenyltransferases” super family. Results of phylogenic analysis showed that *GrGPPS* was at the same evolutionary branch with *AmGPPS*. The SDS-PAGE results displayed that the expressed proteins were consistent with the anticipated size. **Conclusion** The *GrGPPS* gene is cloned from *G. rigescens*, and the stable prokaryotic expression system of pGEX-T-1-*GrGPPS* is constructed. This work will provide a foundation for further purification, structure, and functional research of *GrGPPS* protein.

Key words: *Gentiana rigescens* Franch.; *GrGPPS*; gene cloning; sequence analysis; prokaryotic expression

龙胆苦苷 (gentiopicoside) 广泛存在于龙胆科植物中, 是滇龙胆 *Gentiana rigescens* Franch. 等药用植物中一类重要的次生代谢产物^[1], 也是龙胆泻肝丸等多种中药产品的主要药效成分^[2]。现代药理学

收稿日期: 2014-03-16

基金项目: 国家自然科学基金项目 (81260608); 云南省教育厅科学研究基金重点项目 (2013Z075); 科技部“十二五”国家科技支撑计划项目 (2011BAI13B02-04)

作者简介: 王彩云 (1989—), 女, 硕士在读, 研究方向为药用植物资源评价与利用。Tel: 15108713725 E-mail: wangcaiyun0716@126.com

*通信作者 张晓东, 男, 博士, 讲师, 主要从事植物分子生物学方面的研究。Tel: 15925236961 E-mail: zxd95@126.com

王元忠, 男, 助理研究员, 主要从事药用植物资源评价与利用的研究。Tel: 13888829994 E-mail: boletus@126.com

研究证实龙胆苦苷具有保肝、利胆、镇痛、抗炎、抗菌、健胃、抗肿瘤以及诱发神经轴突生长等作用^[3-6]。龙胆苦苷属于裂环烯醚萜类化合物,同所有的单萜化合物一样起源于相同的底物,即牻牛儿基焦磷酸(GPP, C10),而该化合物的C10骨架由牻牛儿基焦磷酸合成酶(geranyl diphosphate synthase, GPPS)催化异戊烯基焦磷酸(isopentenyl diphosphate, IPP)及其异构体二甲基丙烯基焦磷酸(dimethylallyl diphosphate, DMAPP)头尾缩合而成^[7-8]。GPP能经过多步酶学催化过程随后转换为香叶醇、香叶醇10-羟化酶、马钱苷酸、裂环番木鳖酸、獐牙菜苷、獐牙菜苦苷,最后形成龙胆苦苷^[9-10]。GPPS是短链异戊烯基合成酶家族成员,催化1分子的IPP与DMAPP合成GPP,是异戊二烯途径中一个很重要的酶,在吸引传粉者和次生代谢产物的防御中起重要作用^[11]。

Burke等^[12]第一次从薄荷属植物辣薄荷 *Mentha piperita* L. 中分离出2个亚基组成的异型二聚体GPPS蛋白,并从cDNA文库中筛选得到了GPPS基因全长cDNA序列。随后GPPS基因相继从拟南芥 *Arabidopsis thaliana* L.^[13]、啤酒花 *Humulus lupulus* L.^[14]、长春花 *Catharanthus roseus* (Linn.) G. Don^[15]、金鱼草 *Antirrhinum majus* L.^[16]等植物中克隆得到,并发现GPPS分布在叶绿体以及分泌腺细胞中,且以同型二聚体、异质二聚体两种形式存在^[15]。据报道,同聚肽GPPS在2种裸子植物冷杉 *Abies grandis* Lindl.、挪威云杉 *Picea abies* L. Karst^[17-19]和4种被子植物拟南芥、番茄 *Lycopersicon esculentum* Miller、夏栎 *Quercus robur* L.、蝴蝶兰 *Phalaenopsis bellina* L. 中存在^[8,13,20-21];杂聚肽GPPS只在被子植物辣薄荷、金鱼草、仙女扇 *Clarkia breweri*、啤酒花中存在^[12,14,16]。在结构上,杂聚肽GPPS包含非催化活性小亚基(SSU)和一个大亚基(LSU),这些亚基在GPPS单独存在时无活性和功能,2种亚基的相互作用导致杂聚肽GPPS的活性^[15]。

滇龙胆为龙胆科龙胆属多年生宿根草本植物,是我国特有物种,主要分布于云南、贵州、四川、广西等省,其中云南是主要产区。野外主要生长在海拔1100~3000m杂木林下、荒坡地、山谷灌木丛旁^[22]。滇龙胆在云南有悠久的药用历史^[23],被《中国药典》2010年版收录,为传统中药材龙胆的植物来源之一^[24]。目前,国内外对滇龙胆的研究主要集中在栽培、农艺性状、化学成分研究、矿质元素测

定等方面^[25-27],而对滇龙胆GPPS基因的扩增和功能分析至今未见报道。本研究根据三年生滇龙胆转录组牻牛儿基焦磷酸合成酶(*GrGPPS*)基因序列,设计特异性引物,从滇龙胆叶中成功提取到RNA,反转录为cDNA,进而扩增到*GrGPPS*基因,并进行酶切鉴定、测序以及序列分析,然后进行原核表达。结果表明*GrGPPS*转化菌在37℃、终浓度为1mmol/L IPTG诱导下成功表达出目的蛋白。本研究为滇龙胆及龙胆科植物中裂环烯醚萜类化合物生物合成途径的研究奠定基础,也为其他生物中GPPS蛋白的研究提供参考。

1 材料和试剂

1.1 材料

滇龙胆 *Gentiana rigescens* Franch. 栽培于云南省农业科学院药用植物研究所种质资源圃(25°08'04.50"N, 102°46'15.05"E)。栽培地海拔1942m,年平均气温14.7℃,年平均降水量980~1050mm,极端最高温度30.4℃,极端最低温度-0.2℃。试验材料为滇龙胆长势良好的幼叶。大肠杆菌 *Escherichia coli* Trans 5α和Rosetta(DE3)菌种购于宝生物工程(大连)有限公司。

1.2 试剂

RNAiso for Polysaccharide-rich Plant Tissue Reagent、反转录试剂盒、限制性内切酶及异丙基硫代半乳糖苷(isopropyl-1-thio-β-D-galactopyranoside, IPTG)、X-gal等均为宝生物工程(大连)有限公司产品;高纯质粒小量制备试剂盒、多功能DNA纯化回收试剂盒均购自北京百泰克生物技术有限公司;质粒pGEX-4T-1由昆明理工大学生物工程技术研究中心植物组保存;引物由上海捷瑞生物工程技术服务有限公司合成,测序由上海生工生物工程技术服务有限公司完成。

2 方法

2.1 叶片总RNA提取及*GrGPPS*全长cDNA的克隆

按照RNAiso for Polysaccharide-rich Plant Tissue试剂盒说明书,提取滇龙胆幼叶的总RNA;按照Reverse Transcriptase M-MLV(RNase H-)试剂盒说明书合成cDNA,-20℃保存备用。根据三年生滇龙胆转录组*GrGPPS*基因序列和原核表达载体pGEX-4T-1多克隆酶切位点,设计一对特异引物*GrGPPSSall*-F: 5'-GTCGACATGGCTTTGATATAT-TCTACTCCATC-3';*GrGPPSXho*-R: 5'-CTCGAG-

TCAATTATCCCTATAAGCAATGTAATT-3'。以 cDNA 为模板进行 PCR 扩增, 反应条件: 94 °C、3 min; 94 °C、30 s, 53 °C、30 s, 72 °C、67 s, 30 个循环; 72 °C 延伸 7 min。

2.2 *GrGPPS* 基因的测序与分析

GrGPPS PCR 产物经 1.0% 的琼脂糖凝胶电泳检测, 用胶回收试剂盒对目的片段进行纯化, 纯化后将其连接到 pMD19-T 载体上, 转化 *E. coli* Trans 5 α 感受态细胞, 涂布于添加氨苄青霉素 (100 mg/L)、IPTG、X-gal 的 LB 平板上, 37 °C 培养 12~16 h 后随机挑取阳性克隆摇菌后提取质粒, 经 PCR 检测和酶切鉴定正确后进行测序, 获得重组载体 pMD19T-*GrGPPS*。

2.3 原核表达载体的构建

对 pMD19T-*GrGPPS* 重组质粒和 pGEX-4T-1 载体分别进行 *Sal*I 和 *Xho*I 双酶切, 回收目的基因和载体片段, 按摩尔比 1:4 混合后经 Ligation Solution I 连接后转化 *E. coli* Trans 5 α 感受态细胞, 涂布于添加氨苄青霉素 (100 mg/L) 的 LB 平板, 次日随机挑取阳性克隆摇菌后提取质粒, 经 PCR 检测和酶切鉴定正确后进行测序, 获得融合表达载体 pGEX-4T-1-*GrGPPS*。

2.4 *GrGPPS* 的生物信息学分析

利用 NCBI 网站上的 BLAST 程序进行序列比对, 应用 DNAMAN 软件推测和比对氨基酸序列, 并预测蛋白质相对分子质量和等电点等; 应用 Mega4.0 软件构建系统进化树; 利用数据库资源 http://molbiol.edu.ru/eng/scripts/01_11.html 进行稀有密码子分析。

2.5 重组质粒 pGEX-4T-1-*GrGPPS* 在大肠杆菌中的表达

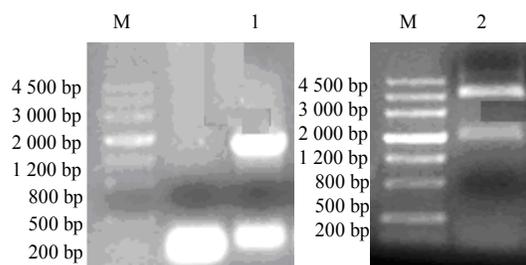
利用热激法将重组质粒 pGEX-4T-1-*GrGPPS* 转化大肠杆菌 Rosetta (DE3) 感受态细胞, 挑取单菌落接种于 4 mL LB 液体培养基 (含 100 mg/L 氨苄青霉素) 中, 37 °C 200 r/min 培养过夜。次日以 1:100 的比例转接到新的不含氨苄青霉素的 LB 液体培养基中, 37 °C 200 r/min 培养至吸光度 (A_{600}) 值 0.6~0.8, 在 37 °C 条件下, 加入 IPTG (终浓度为 1 mmol/L) 进行诱导表达, 同时以加入 IPTG (终浓度 1 mmol/L) 的 pGEX-4T-1 质粒转化菌为对照。诱导 0、1、2、4、6 和 8 h 后分别收集菌液 2 mL。12 000 r/min 常温离心 1 min 后, 弃上清, 加入 20 μ L 的 5 \times SDS-PAGE 上样缓冲液 [250 mmol/L Tris-HCl (pH 6.8)、10% SDS、0.5% 溴酚

蓝、50%甘油、5% β -巯基乙醇], 100 μ L Milli-Q water, 震荡混匀, 100 °C 煮沸 15 min。室温 12 000 r/min 离心 5 min, 取 20 μ L 样品上样, 进行 SDS-PAGE (4%浓缩胶和 12%分离胶) 电泳检测。

3 结果与分析

3.1 滇龙胆 *GrGPPS* cDNA 的克隆

以滇龙胆 cDNA 为模板扩增出的 PCR 产物经琼脂糖凝胶电泳检测, 结果表明扩增片段在 1 100 bp 左右, 是预期的目的片段 (图 1)。重组质粒 pMD19T-*GrGPPS* 经酶切验证连接正确, 有明显的、单一的目的条带, 且目的条带与预期条带的大小吻合 (图 1)。测序结果表明 *GrGPPS* (登录号为 KF922375) 全长 1 107 bp。



M-Marker 1-*GrGPPS* 扩增产物 2-重组质粒 pMD19T-*GrGPPS* *Sal*I 和 *Xho*I 双酶切产物
M-Marker 1-PCR product of *GrGPPS* 2-product of recombinant plasmid pMD19T-*GrGPPS* digested by *Sal*I and *Xho*I

图 1 RT-PCR 扩增 *GrGPPS* 基因及双酶切验证

Fig. 1 Fragment of *GrGPPS* amplified by RT-PCR and restriction verification of pMD19-*GrGPPS*

3.2 *GrGPPS* 序列分析

3.2.1 氨基酸序列同源性分析 利用 DNAMAN 以及 Genetyx 软件对 *GrGPPS* 的序列进行分析, 结果显示 *GrGPPS* 基因的开放阅读框 (ORF) 长 1 107 bp, 编码 369 个氨基酸 (图 2)。

利用 GeneBank 数据库中的 Blastp 程序对 *GrGPPS* 的氨基酸进行同源性分析, 结果表明滇龙胆 *GrGPPS* 与金鱼草 *AmGPPS* 蛋白的相似性最高, 为 80%; 与辣薄荷 *MpGPPS* 和加拿大薄荷 *McGPPS* 蛋白序列的相似性最低, 为 68% (表 1)。

利用 Mega4.0 将 *GrGPPS* 氨基酸序列与从 NCBI 中挑选的同源性较高的部分已知序列进行系统进化分析, 结果表明滇龙胆 *GrGPPS* 与长春花 *CrGPPS* 在同一进化枝上 (图 3), 表明其亲缘关系较近; 与芒果 *MiGPPS*、啤酒花 *HIGPPS* 的遗传距离较远 (图 3), 表明滇龙胆 *GrGPPS* 与芒果 *MiHPPS*、啤酒花 *HIGPPS* 的亲缘关系较远。

```

1 ATGGCTTTGATATATTTCTACTCCATCATGGGTACAAGCACATACTATTTCCATATACCATGGTAATGGATCCAGTTTT
1 M A L I Y S T P S W V Q A H T I S I Y H G N G S S F
79 TTCCCTGTATCTTTCCAAGAACAAGCTCCCGTCTTTCTATCAAACCCATGTAAAAACCCAACCTTGGCCGTTCC
27 F P C Y L S K N K A P V F L S N P C K K P N L G R S
157 CCATTGTCAATCTGTGCGATTCTCACAAAAGAAGAGAGCAAAATCAAGAAAGCCCATGACTTTTCATTCAATTTCAA
53 P L S I C A I L T K E E S K I K K A H D F S F N F K
235 GACTACATGCTTGAAGGGCGGATTTCAGTGAACAAAGCTTTGGAACAAGCGGTTTCGATTTCGTGAACCATAAAAATC
79 D Y M L E K A D S V N K A L E Q A V S I R E P L K I
313 CACGAGTCAATGAGGTACTCACTTCTTGTCTGGGGTAAAAGGGTCCGTCCTCAATGCTCTGTATTGCTGCTGTGAGCTA
105 H E S M R Y S L L A G G K R V R P M L C I A A C E L
391 TTTGGCGGGATGAATCTGTTGCCATTCCTTGTGCTGTGAGATGATTCACACCATGCCTTAATGCATGAT
131 F G G D E S V A M P S A C A V E M I H T M S L M H D
469 GATCTTCCTTGTATGGATAATGATGATTTAAGAAGAGGGAAAGCCCACTAATCATAAGGTTTATGGTGAAGATGTGGCT
157 D L P C M D N D D L R R G K P T N H K V Y G E D V A
547 GTTTTAGCTGGTGATGCACCTTCTTGCATTTGACATTTGAGCATATTGCAACTTCTACAAAAGGGTAACATCAGAAAAGA
183 V L A G D A L L A F A F E H I A T S T K G V T S E R
625 ATTTGTAGAGTAATTTGGAGAATTTAGCAAAATGTATTTGGCTCAGAAAGGATTTAGTAGCAGGGCAGATTGTTGATGTGTGT
209 I V R V I G E L A K C I G S E G L V A G Q I V D V C
703 TCTGAAGGGATATCTGATGTGGGTTTACAGCATTTAGAATTCATTCACATTCACAAGACTGCTGCTTTATTAGAAGGT
235 S E G I S D V G L Q H L E F I H I H K T A A L L E G
781 TCTGTGGCTATGGGGCAATTTCTGGGAGGTGCAGATGATGAAGAAGTTTCTAAATTTGAGAAAGTTTGAAGAGGCATT
261 S V A M G A I L G G A D D E E V S K L R K F G I
859 GGCTTGTGTTTCAAGTAGTGGATGATATCTTGGATGTTACAAAATCTTCGAAAGAATTGGGAAAACCTGCTGCAAAA
287 G L L F Q V V D D I L D V T K S S K E L G K T A A K
937 GATTTGGTAGCAGATAAGGTGACTTATCCGAAACTGATCGTATCGATAAGTCTAGAGAATTTGCTGAGAAAGTTGAAAT
313 D L V A D K V T Y P K L I G I D K S R E F A E K L N
1 015 AGAGAAGCTCAAGATCAACTTGTCTGGTTTTGATTCTGAAAAGGCTGCTCCTTTGATTGCTCTTGTCTAATTACATTGCT
339 R E A Q D Q L A G F D S E K A A P L I A L A N Y I A
1 093 TATAGGGATAATTGA
365 Y R D N *
    
```

图 2 GrGPPS 的 cDNA 序列及其预测的编码蛋白质的氨基酸序列

Fig. 2 cDNA sequence of GrGPPS and predicted amino acid sequence coding protein

表 1 GrGPPS 氨基酸序列与 NCBI 中已知植物 GPPS 序列比对结果

Table 1 BLASTn result of GrGPPS amino acids with other GPPS sequences in plant of NCBI

物种	相似性 / %	基因	登录号
长春花 <i>Catharanthus roseus</i>	77	<i>CrGPPS</i>	AGL91645.1
金鱼草 <i>Antirrhinum majus</i>	80	<i>AmGPPS</i>	AAS82860.1
胡黄连 <i>Picrorhiza kurrooa</i>	70	<i>PkGPPS</i>	AAW66658.1
辣薄荷 <i>Mentha piperita</i>	68	<i>MpGPPS</i>	AAF08793.1
啤酒花 <i>Humulus lupulus</i>	69	<i>HlGPPS</i>	ACQ90682.1
加拿大薄荷 <i>Mentha canadensis</i>	68	<i>McGPPS</i>	ABR15420.1
丹参 <i>Salvia miltiorrhiza</i>	77	<i>SmGPPS</i>	AEZ55681.1
芒果 <i>Mangifera indica</i>	75	<i>MiGPPS</i>	AFJ52722.1

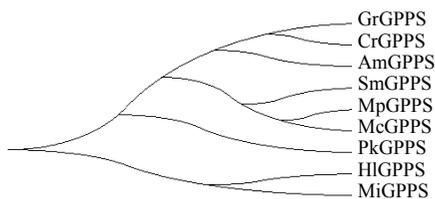


图 3 GrGPPS 与其他 GPPS 蛋白的系统发育分析

Fig. 3 Phylogenetic relationship of GrGPPS and some other GPPS proteins

利用 DNAMAN 将 GrGPPS 氨基酸序列与从 NCBI 中挑选的同源性较高的部分已知序列进行多序列比对分析, 结果表明 GrGPPS 蛋白与已知蛋白

序列在 N 端高度保守, 且含有异戊烯基结构域 (图 4), 进一步表明其为类异戊烯基超家族成员。

3.2.2 理化性质分析 通过使用 ExPASy ProtParam tool 进行分析, GrGPPS 蛋白的相对分子质量为 40 050, 理论等电点为 6.14, 分子式为 $C_{1775}H_{2848}N_{480}O_{533}S_{19}$ 。半衰期 (mammalian reticulocytes, *in vitro*) 为 30 h。不稳定指数 (instabilityindex) 为 34.85, 脂肪族指数 (aliphatic index) 为 95.22, 总平均疏水性 (grand average of hydropathicity, GRAVY) 为 -0.035。蛋白质的疏水性通常依据蛋白的 GRAVY 值来预测, GRAVY 值在 2~-2, 若为正值, 则此蛋白为疏水性蛋白, 反

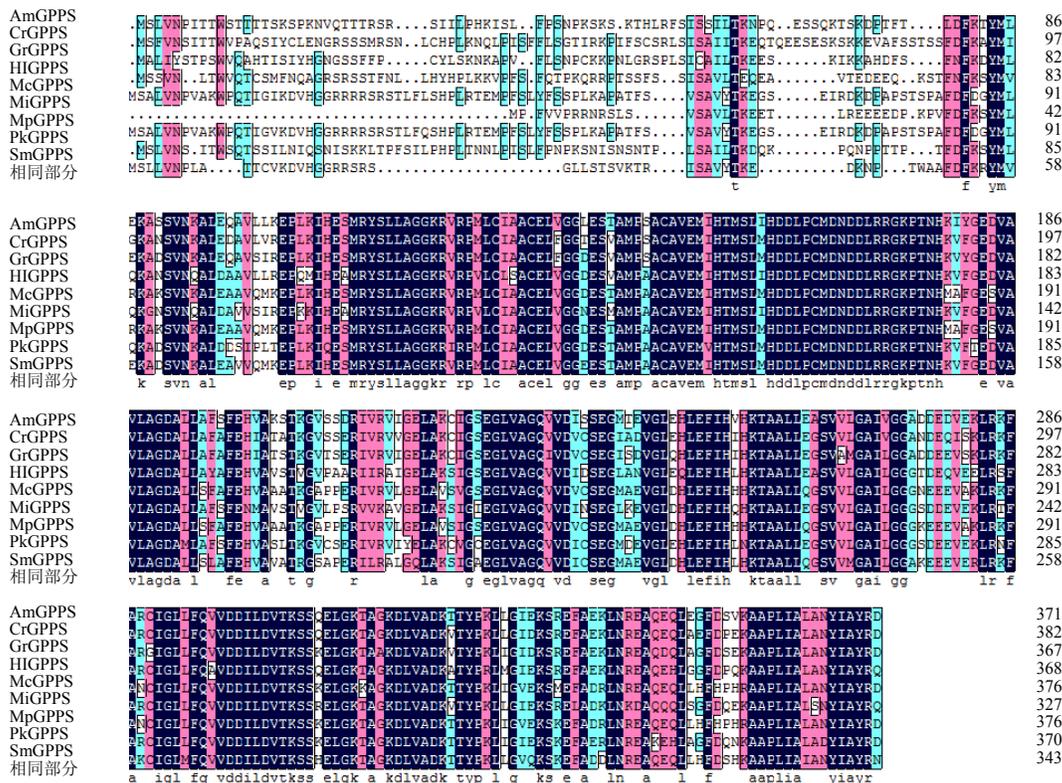


图 4 GrGPPS 与其他植物中 GPPS 氨基酸序列的多序列比对结果

Fig. 4 Multiple sequence alignment of GrGPPS amino acid sequence with GPPS in other plants

之则为亲水性蛋白。从 ProtParam 软件分析结果来看, GRAVY 值为-0.708, 说明 GrGPPS 为亲水性蛋白。GrGPPS 蛋白含有 20 种氨基酸, 其中丙氨酸 (Ala) 的量最高, 为 10.6%; 其次是亮氨酸 (Leu), 为 10.1%; 色氨酸 (Trp) 量最低, 为 0.3%。

对 GrGPPS 进行稀有密码子分析 (http://molbiol.edu.ru/eng/scripts/01_11.html), 结果表明 *GrGPPS* 基因中稀有密码子占 0.8%, 无二联或三联稀有密码子连续出现的情况, 因此可以选用 BL21 或 Rosetta (DE3) 进行原核表达。

3.2.3 蛋白质二级结构和三维建模 利用 SSpro 方法 (<http://scratch.proteomics.ics.uci.edu/>) 对 GrGPPS 进行二级结构分析。结果表明该蛋白二级结构中 α -螺旋 (H) 占 64.4%, β -折叠 (E) 占 2.45%, 无规则卷曲 (C) 占 33.15%。

利用 CPHmodels 3.2 Server 在线软件 (<http://www.cbs.dtu.dk/services/CPHmodels/>) 预测 GrGPPS 蛋白的三维结构 (图 5), 从图中可以看到 GrGPPS 的三维结构主要由螺旋结构和无规则卷曲组成, 与二级结构预测结果一致; 另外, 其三维结构形成“穴型”功能区域, 暗示其在催化反应中起关键作用。

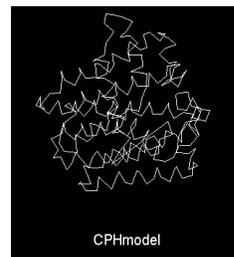


图 5 GrGPPS 的三维结构预测

Fig. 5 Predicted three dimensional structure of GrGPPS protein

3.2.4 结构域预测 采用 InterProScan 在线工具预测 GrGPPS 蛋白的保守结构域, GrGPPS 具有 7 个保守结构域 (图 6), 主要是聚丙烯合成酶 (polyprenyl synthetase) (IPR000092), 分别位于 105~354 位 (PF00348)、286~298 位 (PS00444) 和 153~169 位 (PS00723); 萜类合成酶 (terpenoid synthase) (IPR008949), 分别位于 75~366 位 (1.10.600.10) 和 77~363 位 (SSF48576); 聚丙烯相关合成酶 (polyprenyl synthetase-related) (IPR017446), 位于 61~365 位 (PTHR12001)。另外, 还有牻牛儿基焦磷酸合成酶结构域, 在 IPR 中没有明确分类, 位于

61~365 位 (PTHR12001:SF8)。

利用推测的蛋白序列检索 NCBI 的 CDD (Conserved Domain Database) 数据库, 发现 GrGPPS

蛋白属于类异戊二烯超家族成员 (图 7), 其功能可能与单萜类物质的合成有关。

3.2.5 信号肽、跨膜区、亚细胞定位及功能预测分析

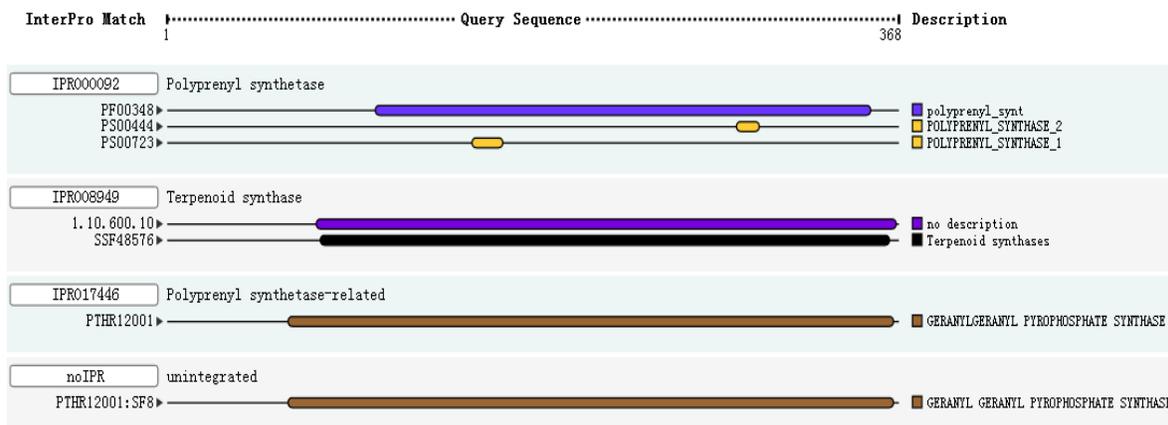


图 6 InterProScan 预测的 GrGPPS 蛋白保守结构域

Fig. 6 Conserved domains of GrGPPS protein predicted by InterProScan

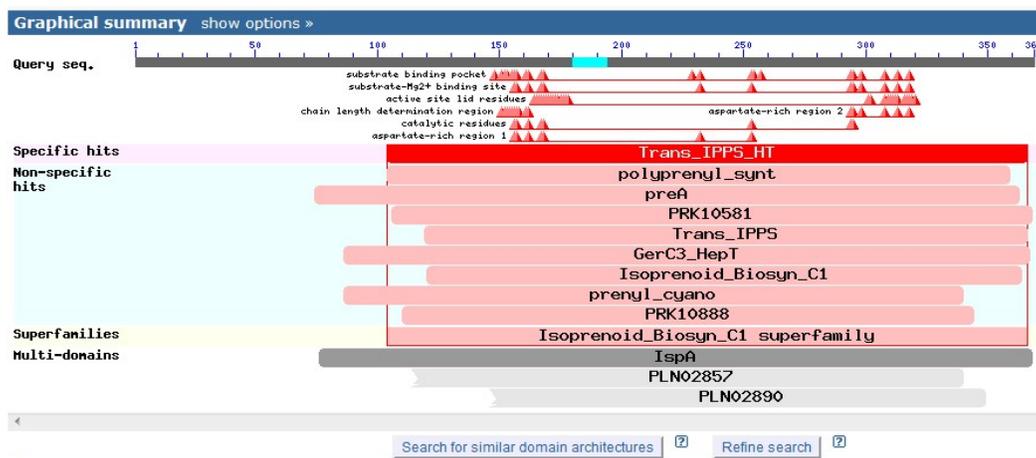


图 7 GrGPPS 蛋白保守域分析

Fig. 7 Conserved domain prediction of GrGPPS protein

利用 ExPASy SignalP 4.0 Server (<http://www.cbs.dtu.dk/services/SignalP/>) 分析 GrGPPS 蛋白, 并没有发现信号肽, 表明该蛋白为非分泌型蛋白。

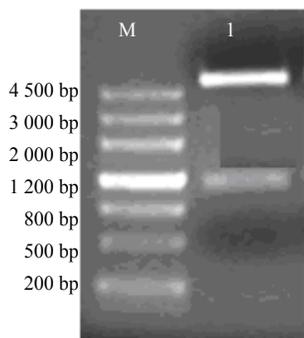
利用 ExPASy 中的 TMHMM 工具 (<http://www.cbs.dtu.dk/services/TMHMM-2.0/>) 预测 GrGPPS 蛋白的跨膜螺旋区, 结果表明 GrGPPS 蛋白非膜蛋白, 无跨膜螺旋区。

利用在线工具 WOLF PSORT (<http://www.wolfpsort.org/>) 预测 GrGPPS 的亚细胞定位情况, 结果显示该蛋白可能定位于叶绿体, 定位系数为 14.0 (chlo: 14.0)。

使用 ProtFun 软件基于已知的具有相似功能蛋白的搜索和比对, 对 GrGPPS 蛋白进行功能预测,

结果显示 GrGPPS 蛋白嘌呤碱和嘧啶碱、翻译、能量代谢的功能的可能性为 0.521、0.243、0.293, 其可能性的数值远高于其他功能, 这些数据都为后续的研究提供了一定的依据。

3.2.6 GrGPPS 原核表达载体构建 使用 *Sal*I 和 *Xho*I 双酶切重组质粒 pGEX-4T-1-GrGPPS, 可切出约 1 107 bp 的片段 (图 8), 表明扩增的 GrGPPS 片段已插入载体 pGEX-4T-1 中。在此基础上, 对重组表达质粒 pGEX-4T-1-GrGPPS 进行测序。结果表明, 连入表达载体中的基因片段与目的序列一致, 并且酶切位点连接处序列也都正确, 未出现碱基突变及移码现象。表明已获得了正确的 GrGPPS 的原核表达重组质粒。

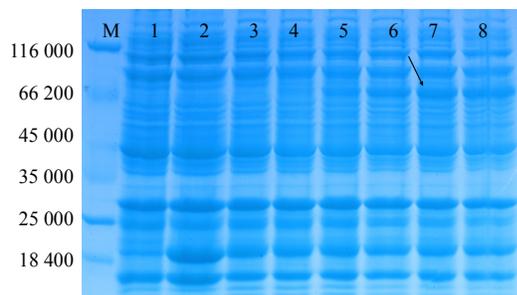


M-Marker 1-pGEX-4T-1-GrGPPS 双酶切结果
M-Marker 1-digestion result of recombinant plasmid pGEX-4T-1-GrGPPS by *Sall* and *XhoI*

图 8 pGEX-4T-1-GrGPPS 酶切检测

Fig. 8 Digestive detection of pGEX-4T-1-GrGPPS plasmid

3.2.7 GrGPPS 蛋白的表达 将重组质粒 pGEX-4T-1-GrGPPS 转化大肠杆菌 Rosetta 后进行诱导表达。在终浓度为 1 mmol/L IPTG、37 °C 下分别诱导表达 0、1、2、4、6 和 8 h 后，提取大肠杆菌总蛋白进行 SDS-PAGE 分析。结果表明，插入有外源片段的重组质粒 pGEX-4T-1-GrGPPS 经 IPTG 诱导后，在预期的蛋白相对分子质量 66 050（包括载体上 GST 蛋白）左右有 1 条蛋白条带，而未诱导时的转化重组质粒和 pGEX-4T-1 对照质粒均未出现这条蛋白带，表明重组质粒 pGEX-4T-1-GrGPPS 在大肠杆菌 Rosetta (DE3) 中诱导表达了 GrGPPS 蛋白。当温度为 37 °C、诱导时间为 4 h 时，蛋白表达量已接近最大，8 h 时最大（图 9）。



M-Marker 1~2-对照：37 °C，IPTG 终浓度为 1 mmol/L 时 pGEX-4T-1 空载体转化子诱导 0 和 8 h 的总蛋白；3~8-相同条件下，融合表达工程菌诱导 0、1、2、4、6、8 h 的总蛋白
M-Marker; 1—2-CK is the expressed product of pGEX-4T-1 with 1 mmol/L of IPTG induction for 0 and 8 h at 37 °C; 3—8-The expressed product of pGEX-4T-1-GrGPPS with 1 mmol/L of IPTG induction for 0, 1, 2, 4, 6 and 8 h separately at 37 °C.

图 9 37 °C 不同时间诱导对重组蛋白表达量的影响

Fig. 9 Effect of different time on expression of recombinant protein at 37 °C

4 讨论

到目前为止，GPPS 的活性仅在少数植物中被研究，且 GPPS 蛋白主要存在于产生大量单萜的组织中，纯化非常困难^[28]，对 GPPS 的研究主要集中在其对单萜和倍半萜的影响，如 Martin 等^[29]对葡萄 *Vitis vinifera* 中 *VvGPPS* 的转录进行分析，表明 *VvGPPS* 基因在单萜生物合成中起重要作用；Chang 等^[30]研究发现辣薄荷中的 *MpGPPS* 包含 2 个具有催化功能的大亚基和 2 个非催化的具有调控功能的小亚基，并且证实与薄荷醇的生物合成有关；Rai 等^[15]从长春花中克隆到 *CrGPPS* 基因，发现 *CrGPPS* 在长春花中以杂聚肽和同聚肽 2 种形式存在，且发现只有含有 *GrGPPS.SSU* 的杂聚肽 GPPS 能调控 MIA 生物合成中的 GPP 流。然而 Van Schie 等^[20]研究发现由于病毒诱导基因沉默引起的番茄中 *LeGPPS* 表达量降低导致植株的严重矮小，进一步研究发现这些矮小植株中赤霉素的量降低，而类胡萝卜素和叶绿素的含量未发生变化，表明 GPPS 不仅在萜类的生物合成中起作用，而且与赤霉素的生物合成有关。

尽管 GPPS 在单萜生物合成中起着重要作用，但 GPPS 在植物中的功能以及在 GPP 代谢流中的调控机制研究的还很少。克隆基因的方法有 cDNA 差异显示法^[31]、设计简并引物同源克隆法^[32-33]以及分析 EST 数据库，根据基因预测设计特异引物，直接从 cDNA 文库或基因组文库中扩增基因为策略的一些方法，例如 RACE、Tail-PCR 等。本研究根据滇龙胆转录组 *GrGPPS* 基因序列，设计一对特异性引物，应用 RT-PCR 技术进行扩增，得到完整的 cDNA 序列，并命名为 *GrGPPS*。经序列分析，发现 *GrGPPS* 全长 1 107 bp，编码 369 个氨基酸。BLASTp 分析结果显示，GrGPPS 的同源蛋白较少，但与其他物种 GPPS 蛋白的相似性较高。以上这些结果表明 *GrGPPS* 基因属于类异戊二烯家族成员。原核表达载体 pGEX-4T-1 带有谷胱甘肽转移酶 GST 标签，由于 GST 标签蛋白较大，相对分子质量约 2 600，在实验中常需要用凝血酶等将标签蛋白酶切后再进行目的蛋白的后续研究。

本研究成功构建了滇龙胆 *GrGPPS* 基因原核表达载体 pGEX-4T-1-GrGPPS，转入 *E. coli* Rosetta (DE3)，并成功诱导表达出 GrGPPS 蛋白。SDS-PAGE 结果表明龙胆叶片中的 *GrGPPS* 基因编码蛋白为单条带（图 9），因此 GrGPPS 蛋白可能为单亚基酶或二亚基同聚体，这需要进一步的实验验证。

原核生物具有繁殖快、周期长等特点,为下一步纯化 GrGPPS 蛋白以及进行生物功能分析奠定基础。以体外实验模拟体内环境来生产生物体内化学成分是目前的科研热点,采用生物技术来缓解资源短缺等问题也是中医药现代化的目的之一。该研究有助于深入了解龙胆苦苣生物合成途径的分子基础,也为单萜类化合物的生物工程提供候选基因。

参考文献

- [1] Wang Y M, Xu M, Wang D, *et al.* Review on “Long-Dan”, one of the traditional Chinese medicinal herbs recorded in *Chinese Pharmacopoeia* [J]. *Nat Prod Bioprospect*, 2012, 2: 1-10.
- [2] 张勇, 薛昆鹏, 何美, 等. 固相萃取/超高效液相色谱法测定龙胆泻肝丸中栀子苷、龙胆苦苣与黄芩苷 [J]. *分析测试学报*, 2013, 32(1): 122-126.
- [3] Wang C G, Zhang T, Cui X M, *et al.* Hepatoprotective effects of a Chinese herbal formula, longyin decoction, on carbon-tetrachloride-induced liver injury in chickens [J]. *Evid-Based Compl Alt*, 2013, 2013: 1-9.
- [4] Wang Y M, Xu M, Wang D, *et al.* Anti-inflammatory compounds of “Qin-Jiao”, the roots of *Gentiana dahurica* (Gentianaceae) [J]. *J Ethnopharmacol*, 2013, 147(2): 341-348.
- [5] Jia L Y, Guo H Y, Jia B Z, *et al.* Anti-tumour activities and a high-performance liquid chromatography mass spectrometric method for analysis of the constituents of *Lomatogonium carinthiacum* [J]. *Nat Prod Res*, 2011, 25(2): 100-107.
- [6] Chiba K, Yamazaki M, Kikuchi M, *et al.* New physiological function of secoiridoids: neuritogenic activity in PC12h cells [J]. *J Nat Med*, 2011, 65(1): 186-190.
- [7] Ogura K, Koyama T. Enzymatic aspects of isoprenoid chain elongation [J]. *Chem Rev*, 1998, 98(4): 1263-1276.
- [8] Schmidt A, Gershenzon J. Cloning and characterization of two different types of geranyl diphosphate synthases from Norway spruce (*Picea abies*) [J]. *Phytochemistry*, 2008, 69(1): 49-57.
- [9] Coscia C J, Guarnaccia R. Natural occurrence and biosynthesis of a cyclopentanoid monoterpene carboxylic acid [J]. *Chem Commun*, 1968, 3: 138-140.
- [10] Coscia C J, Botta L, Guarnaccia R. On the mechanism of iridoid and secoiridoid monoterpene biosynthesis [J]. *Arch Biochem Biophys*, 1970, 136(2): 498-506.
- [11] Gershenzon J, Dudareva N. The function of terpene natural products in the natural world [J]. *Nat Chem Biol*, 2007, 3(7): 408-414.
- [12] Burke C C, Wildung M R, Croteau R. Geranyl diphosphate synthase: cloning, expression, and characterization of this prenyltransferase as a heterodimer [J]. *P Natl Acad Sci USA*, 1999, 96(23): 13062-13067.
- [13] Bouvier F, Suire C, d'Harlingue A, *et al.* Molecular cloning of geranyl diphosphate synthase and compartmentation of monoterpene synthesis in plant cells [J]. *Plant J*, 2000, 24(2): 241-252.
- [14] Wang G, Dixon R A. Heterodimeric geranyl (geranyl) diphosphate synthase from hop (*Humulus lupulus*) and the evolution of monoterpene biosynthesis [J]. *P Natl Acad Sci USA*, 2009, 106(24): 9914-9919.
- [15] Rai A, Smita S S, Singh A K, *et al.* Heteromeric and homomeric geranyl diphosphate synthases from *Catharanthus Roseus* and their role in monoterpene indole alkaloid biosynthesis [J]. *Mol Plant*, 2013, 6(5): 1531-1549.
- [16] Tholl D, Kish C M, Orlova I, *et al.* Formation of monoterpenes in *Antirrhinum majus* and *Clarkia breweri* flowers involves heterodimeric geranyl diphosphate synthases [J]. *Plant Cell*, 2004, 16(4): 977-992.
- [17] Burke C, Croteau R. Geranyl diphosphate synthase from *Abies grandis*: cDNA isolation, functional expression, and characterization [J]. *Arch Biochem Biophys*, 2002, 405(1): 130-136.
- [18] Burke C, Croteau R. Interaction with the small subunit of geranyl diphosphate synthase modifies the chain length specificity of geranylgeranyl diphosphate synthase to produce geranyl diphosphate [J]. *J Biol Chem*, 2002, 277(5): 3141-3419.
- [19] Schmidt A, Wächtler B, Temp U, *et al.* A bifunctional geranyl and geranylgeranyl diphosphate synthase is involved in terpene oleoresin formation in *Picea abies* [J]. *Plant Physiol*, 2010, 152(2): 639-655.
- [20] Van Schie C C, Ament K, Schmidt A, *et al.* Geranyl diphosphate synthase is required for biosynthesis of gibberellins [J]. *Plant J*, 2007, 52(4): 752-762.
- [21] Hsiao Y Y, Jeng M F, Tsai W C, *et al.* A novel homodimeric geranyl diphosphate synthase from the orchid *Phalaenopsis bellina* lacking a DD (X) 2-4D motif [J]. *Plant J*, 2008, 55(5): 719-733.
- [22] 中国科学院中国植物志编辑委员会. 中国植物志 (第62卷) [M]. 北京: 科学出版社, 1988.
- [23] 兰茂. 滇南本草 (第2卷) [M]. 昆明: 云南人民出版社, 1977.
- [24] 中国药典 [S]. 一部. 2010.
- [25] 杨雁, 邵爱娟, 金航, 等. 云贵高原滇龙胆不同居群形态特征变异研究 [J]. *中草药*, 2012, 43(8):

- 1604-1610.
- [26] 杨美权, 张金渝, 沈涛, 等. 不同栽培模式对滇龙胆中龙胆苦苷含量的影响 [J]. 江苏农业科学, 2011(1): 287-289.
- [27] 沈涛, 杨美权, 赵振玲, 等. 滇龙胆中萜类物质积累的动态变化 [J]. 植物学报, 2011, 46(6): 652-657.
- [28] Tholl D, Croteau R, Gershenzon J. Partial purification and characterization of the short-chain prenyltransferases, geranyl diphosphate synthase and farnesyl diphosphate synthase, from *Abies grandis* (Grand Fir) [J]. *Arch Biochem Biophys*, 2001, 386(2): 233-242.
- [29] Martin D M, Chiang A, Lund S T, *et al.* Biosynthesis of wine aroma: transcript profiles of hydroxymethylbutenyl diphosphate reductase, geranyl diphosphate synthase, and linalool/nerolidol synthase parallel monoterpenol glycoside accumulation in Gewürztraminer grapes [J]. *Planta*, 2012, 236(3): 919-929.
- [30] Chang T H, Hsieh F L, Ko T P, *et al.* Structure of a heterotetrameric geranyl pyrophosphate synthase from mint (*Mentha piperita*) reveals intersubunit regulation [J]. *Plant Cell*, 2010, 22(2): 454-467.
- [31] Sperotto R A, Boff T, Duarte G L, *et al.* Increased senescence-associated gene expression and lipid peroxidation induced by iron deficiency in rice roots [J]. *Plant Cell Rep*, 2008, 27(1): 183-195.
- [32] Sun C, Palmqvist S, Olsson H, *et al.* A novel WRKY transcription factor, SUSIBA2, participates in sugar signaling in barley by binding to the sugar-responsive elements of the *iso1* promoter [J]. *Plant Cell*, 2003, 15(9): 2076-2092.
- [33] Guo Z J, Kan Y C, Chen X J, *et al.* Characterization of a rice *WRKY* gene whose expression is induced upon pathogen attack and mechanical wounding [J]. *Acta Bot Sin*, 2004, 46(8): 955-964.