

刺五加鲨烯合酶基因 2 cDNA 的克隆与蛋白结构分析

龙月红, 邢朝斌, 梁能松, 何 闪, 李宝财, 朱金丽

(河北联合大学 生命科学院, 河北 唐山 063000)

摘 要:采用 RT-PCR 和 RACE 法克隆了刺五加鲨烯合酶基因 2(squalene synthase gene 2, SS2) cDNA 的全长序列(GenBank 登录号: JN714465)。该序列全长 1 463 bp, 其中 5' 端非翻译区长 10 bp, 3' 端非翻译区长 208 bp, 开放阅读框长 1 245 bp, 编码 414 个氨基酸。该蛋白具有 1 个 Trans_IPPS_HH 保守区域、2 个鲨烯合酶和八氢番茄红素合成酶特异性识别区域及 2 个跨膜区。氨基酸序列比对结果表明, 该蛋白与五加科其它植物 SS 的一致性均大于 90%, 与刺五加 SS1 的一致性为 91.79%。

关键词:鲨烯合酶基因; 刺五加; RT-PCR; RACE

中图分类号:S 759.82 **文献标识码:**A **文章编号:**1001-0009(2012)02-0117-04

鲨烯合酶定位于内质网膜, 催化两分子的法呢酰基二磷酸(Farnesyl diphosphate, FPP)缩合成为一分子鲨烯的关键酶^[1-2], 而鲨烯是生物合成三萜、甾醇类化合物的第一个前体物质^[3]。在烟草细胞培养物中加入真菌诱导子, SS 的酶活性迅速降低, 甾醇的生物合成和积累受到显著抑制^[4], 将人参的 SS 基因转入刺五加后, SS 的过量表达能促进甾醇和三萜类化合物的大量合成^[5]。目前已有酵母、鼠、人、拟南芥和烟草等 14 种物种的 SS 基因被克隆和测序^[6], 不同物种中含有的 SS 基因拷贝数不同^[7], 不同植物 SS 基因表达的组织特性也存在差异^[7-8]。

在拟南芥中发现了 6 个 SS 基因的异构型(Isoform), 其中异构型 1~3 编码有功能的鲨烯合酶, 4~6 无催化活性^[9]; 异构型 SQS1 在所有组织均有表达, SQS2 主要在叶脉中表达^[10]。人参的 SS 基因存在 3 个异构型, 且均具有催化鲨烯合成的作用, 茉莉酸甲酯处理后 SS 基因的 3 个异构型均不同程度的提高了表达量, 但其转录模式不同, 异构型 PgSS2、PgSS3 的转录物积累位点无变化, PgSS1 却由全部器官高表达变为叶柄维管束专一性高表达^[1]。北柴胡中也克隆到了 2 个氨

基酸同源性为 96% 的 SS 基因^[11]。

刺五加(*Eleutherococcus senticosus* Harms)是我国传统的珍贵药用植物, 其具有的丰富皂苷类化合物为主要活性成分之一^[12]。在克隆到刺五加 SS1 的基础上, 利用 RACE 法克隆了刺五加 SS2 基因的 cDNA 全长序列, 并进一步分析了 SS2 基因编码蛋白的结构。旨在为研究刺五加中三萜类化合物的生物合成奠定基础, 同时也为对其进行表达调控提供了理论依据。

1 材料与方法

1.1 试验材料

1.1.1 植物材料 供试刺五加采自吉林省伊通满族自治县, 经河北联合大学生命科学学院邢朝斌副教授鉴定为五加科植物刺五加(*Eleutherococcus senticosus* Harms)。以清水冲洗杂质, 滤纸吸干水分后的叶片为刺五加总 RNA 提取材料。

1.1.2 试剂 RevertAid™ First strand cDNA synthesis Kit 购自 Fermentas 公司; 3'-Full RACE core set Ver. 2.0 购自宝生物工程(大连)有限公司; IPTG、X-gal、LA Taq DNA 聚合酶、dNTPs、质粒小提试剂盒购自北京拜尔迪生物技术有限公司; 植物 RNA 提取试剂盒、琼脂糖凝胶 DNA 回收试剂盒、PGM-T 克隆试剂盒、TOP-10 感受态细胞购自天根生化科技(北京)有限公司, 其它试剂均为国产分析纯。

1.1.3 引物 根据 GenBank 中登陆的人参、三七、龙牙楸木和西洋参的 SS 基因序列, 通过序列比对找到保守区, 保守区扩增引物为: ES1: 5'-TAGAGAGAAAATGG-GAAG-3' 和 SS2X1: 5'-CTATTGCCTCCTGGAAACCG-3'。3' RACE outer primer: 5'-TACCGTCGTTCCAC-

第一作者简介:龙月红(1974-), 女, 助理实验师, 现主要从事药用植物的研究工作。E-mail: xzbheuu@126.com。

责任作者:邢朝斌(1975-), 男, 硕士, 副教授, 现主要从事分子生物学及药用植物细胞工程研究工作。E-mail: xingzhaobin@yahoo.com.cn。

基金项目:国家自然科学基金资助项目(30701086); 河北省自然科学基金资助项目(C2009001252)。

收稿日期:2011-10-27

TAGTGATTT -3' 和 3' RACE inner primer: 5'-CGCG-GATCCTCCACTAGTGATTTCATATAGG-3' 为 3'-Full RACE core set Ver. 2.0 中自带引物。引物 ESS2X: 5'-AAATGGGAAGTTTGGGGGCA-3' 和引物 SS2X2: 5'-TTGGAAGCGGTTTCCAGGAG-3' 为根据 ES1 和 SS2X1 引物克隆基因片段, 利用 DNA star 软件设计的用于 3'RACE 的特异性引物。引物 SS2X3: 5'-TCACTG-GTTGTTTCGGTAGG-3' 是根据 3'RACE 扩增产物序列, 利用 DNA star 软件设计的扩增刺五加 SS2 基因 cDNA 全长的特异性引物。所有引物均由生工生物工程(上海)有限公司合成, 经 PAGE 纯化。

1.2 试验方法

1.2.1 总 RNA 的提取 称取 0.1 g 刺五加叶片, 按照植物 RNA 提取试剂盒的说明书提取刺五加总 RNA。用琼脂糖凝胶电泳检测 RNA 的完整性。

1.2.2 基因保守片段的克隆 取总 RNA 3 μ L, 以 Oligo (dT)₁₈ 为引物, 根据 RevertAid™ First strand cDNA synthesis Kit 的说明进行逆转录反应。以逆转录获得的 cDNA 为模板 PCR 扩增刺五加 SS2 基因保守片段。反应体系 25 μ L, 其中引物 ES1 和 SS2X1 各 1 μ L, 10 \times LA Taq Buffer (含 15 mmol/L MgCl₂) 2.5 μ L, 2.5 mmol/L dNTP 2 μ L, 模板 cDNA 1 μ L, LA Taq 酶 1 μ L, 补 dd H₂O 至 25 μ L。反应条件为: 预变性 94℃、5 min; 变性 94℃、1 min; 退火 50℃、30 s; 延伸 72℃、30 s。35 个循环后 72℃ 延伸 10 min。PCR 产物经 1.2% 琼脂糖凝胶电泳后, 切胶回收纯化, 采用 T-A 克隆法将其克隆至 PGM-T vector 载体。转化大肠杆菌 TOP10 后提取重组质粒并测序, 测序工作由 Takara 公司完成。根据测序结果设计 3'RACE 的特异性引物, 扩增刺五加 SS2 的 3' 端序列。

1.2.3 3'RACE 方法 取 3 μ L 总 RNA, 以 3'RACE Adaptor 为引物, 参照 3'-Full RACE core set Ver. 2.0 逆转录部分的说明进行逆转录反应, 以 3'RACE outer primer 和特异性外侧引物 ESS2X 进行套式 PCR 的 outer PCR 反应, 反应体系按照说明书进行。反应条件为 94℃、3 min; 94℃、30 s; 55℃、30 s; 72℃、2 min, 20 个循环后 72℃ 延伸 10 min。取 outer PCR 产物 1 μ L, 以 3'RACE inner primer 和特异性内侧引物 SS2X2 进行套式 PCR 的 inner PCR 反应, 反应体系按说明书进行。反应条件为 94℃、3 min; 94℃、30 s; 57℃、30 s; 72℃、1 min 30 s, 30 个循环后 72℃ 延伸 10 min。套式 PCR 产物经 1.2% 的琼脂糖凝胶后参照 1.2.2 中的方法回收、克隆测序。

1.2.4 基因全长的克隆 以扩增基因保守区的 ES1 和

根据 3'RACE 测序结果设计的 SS2X3 为引物, 以 1.2.2 中逆转录的 cDNA 为模板, PCR 扩增刺五加 SS2 cDNA 的全长序列。PCR 反应体系为 25 μ L, 其中引物 ES1 和 SS2X3 各 1 μ L, 10 \times LA Taq Buffer (含 15 mmol/L MgCl₂) 2.5 μ L, 2.5 mmol/L dNTP 2 μ L, 模板 cDNA 1.5 μ L, LA Taq 酶 1 μ L, 补 dd H₂O 至 25 μ L。反应条件为 94℃、5 min; 94℃、1 min; 51℃、30 s; 72℃、1 min 20 s, 35 个循环后 72℃ 延伸 10 min。PCR 产物经 1.2% 的琼脂糖凝胶后参照 1.2.2 中的方法回收、克隆测序。

1.2.5 生物信息学分析 通过 DNAMAN 6.0 将测序获得的刺五加 SS2 cDNA 全长序列翻译成氨基酸序列; 在线工具 <http://web.expasy.org/protparam/> 预测蛋白质的基本理化性质; 利用 <http://www.cbs.dtu.dk/services/TMHMM/> 进行蛋白质序列的跨膜区分分析; http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html 进行蛋白质二级结构分析; <http://prosite.expasy.org/> 进行蛋白质功能结构域分析。使用 MEGA 5.05 软件中的 Neighbor-Joining (邻位相连法, NJ) 法构建系统发育树。

2 结果与分析

2.1 刺五加 SS2 基因保守片段的克隆

根据已知五加科植物 SS 的保守序列设计引物, RT-PCR 扩增得到约为 430 bp 的条带(图 1), PCR 产物连接 PGM-T 载体转化大肠杆菌 TOP10, PCR 鉴定后, 选择阳性质粒测序得到 431 bp 片段, 经 NCBI 的 Blast 比对, 确定 ES1 和 SS2X1 的扩增产物为刺五加 SS2 基因 cDNA 的部分片段, 且该片段中包含起始密码子 ATG。

2.2 3'RACE 结果

利用引物 SS2X2 和 3'RACE inner primer, 通过 3'RACE 扩增得到约 1 000 bp 的条带(图 1)。将其转化入 T 质粒载体, PCR 鉴定为阳性的质粒测序得到 1 061 bp 片段, 该片段中包含终止密码子 TGA。

2.3 刺五加 SS2 基因 cDNA 全长的克隆与分析

根据刺五加 SS2 基因保守片段和 3'RACE 的测序结果, 把 5' 和 3' 末端进行拼接, 得到刺五加 SS2 基因 cDNA 全序列。结果显示, 刺五加 SS2 cDNA 全长共 1 463 bp, 其中, 5' 端非翻译区 (5' UTR) 长 10 bp, 3' 端非翻译区 (3' UTR) 长 208 bp, 开放阅读框 (ORF) 长 1 245 bp, 编码 414 个氨基酸。预测的蛋白质分子质量为 47.038×10^3 , 理论等电点 (pI) 为 6.67。该文所克隆的基因序列已提交 GenBank 数据库 (GenBank 登录号: JN714465)。根据拼接的 SS2 全序列, 设计的扩增 ORF 全长引物 ES1 和 SS2X3 的 PCR 产物与预期大小相当

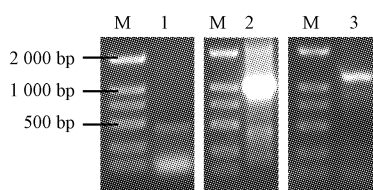


图1 RT-PCR与RACE产物琼脂糖电泳

注:M;DL2000;1;保守区 PCR 产物;2;3'RACE 扩增产物;3;SS2 全长扩增产物。

Fig.1 Agarose electrophoresis result of RT-PCR and RACE product

Note:M; DL2000;1: PCR product of conserved domine;2: 3'RACE amplifying product;3: full-length product of SS2.

(图1),测序结果显示与拼接序列相同。

2.4 刺五加 SS2 基因编码蛋白序列分析

通过 TMHMM 软件分析得知,刺五加 SS2 蛋白具有 2 个跨膜区域,分别位于第 281~303 位和第 385~407 位,刺五加 SS2 蛋白二级结构中含有 286 个 α 螺旋(alpha helix),占 69.08%;23 个延伸链(extended strand),占 5.56%;18 个 β 折叠(beta turn),占 4.35%;87 个无规则蜷曲(random coil),占 21.01%。功能结构域的预测表明,该基因编码的氨基酸序列含有 Trans_IPPS_HH 的保守区域,且该蛋白的 168~183 和 201~229 位氨基酸为鲨烯合酶和八氢番茄红素合成酶特异性识别区域。通过 NCBI 的 BLAST 比对发现,刺五加 SS2 蛋白与已知的蛋白质结构功能数据库中其它物种的鲨烯合酶具有相似的结构功能域。

2.5 刺五加与其它物种 SS 蛋白序列同源性分析

根据核苷酸序列推导出的氨基酸与龙牙楸木 *Aralia elata* (GU354313.1)、人参 *Panax ginseng* (GQ468527.2)、三七 *P. notoginseng* (DQ186630.1)、西洋参 *P. quinquefolius* (AM182456.1)、积雪草 *Centella asiatica* (AY787628.1)、甘草 *Glycyrrhiza uralensis* (HM012844.1)、柴胡 *Bupleurum chinense* (GQ889268.1)、黄芪 *Astragalus membranaceus* (HQ829974.1)、大豆 *Glycine max* (AB007503.1)、烟草 *Nicotiana tabacum* (U60057.1)、青蒿 *Artemisia annua* (AY445506.1)、番茄 *Solanum lycopersicum* (GU075687.1)、辣椒 *Capsicum annuum* (AF124842.1)、棉花 *Gossypium hirsutum* (EF688567.1)、玉米 *Zea mays* (NM_001111369.1)、拟南芥 *Arabidopsis thaliana* (NM_119630.3)、褐家鼠 *Rattus norvegicus* (AAA42179.1)、人 *Homo sapiens* (NP_004453.3)、酵母 *Schizosaccharomyces pombe* (NP_595363.1) 的 SS 及刺五加 SS1 (HQ456918.1) 的氨基酸序列进行对比分析,建立系统进化树(图2)。刺五加的 SS2 与同为五加科的三七和龙牙楸木的一致性高达 97% 和 96%,三者首先聚为同一分枝,次之是伞形目的

柴胡,再次之是伞形目的积雪草、西洋参、人参和刺五加的 SS1,与其它植物、动物和人的进化距离依次变远。

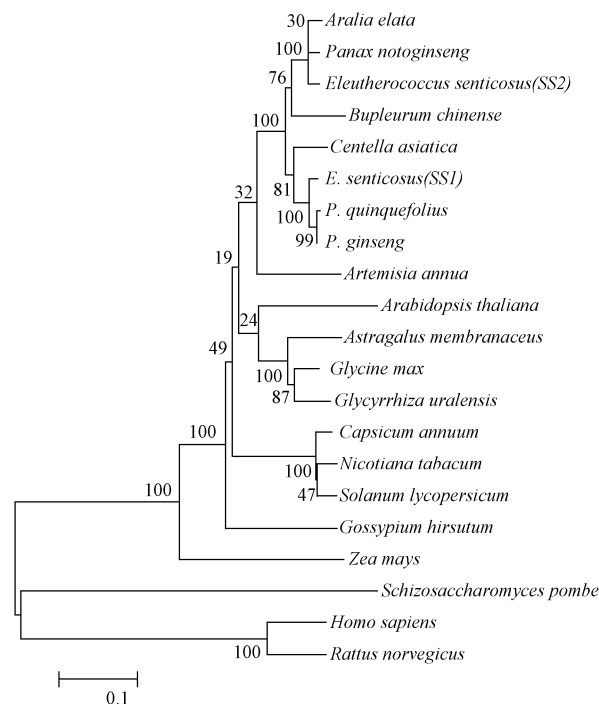


图2 鲨烯合酶氨基酸序列系统发育树

Fig.2 Phylogenetic tree of squalene synthase based on amino acid sequences

3 结论与讨论

该试验所克隆的刺五加 SS2 基因与龙牙楸木、三七、人参、西洋参等五加科植物的 SS 基因序列一致性在 90% 以上,均属于鲨烯合成酶家族。SS 在植物甾醇和三萜皂苷的生物合成中具有很强的调控功能,一般认为它的转录物存在于各种组织中,并且在枝叶类和根部中含量最高^[13]。此次克隆的刺五加 SS2 及先期克隆的 SS1 的取材部位均为叶片,同时以根的总 RNA 逆转录产物为模板的 RT-PCR 扩增中也获得了较好的结果,说明刺五加 SS2 在叶和根中均有表达。

通过对 SS 蛋白的纯化和序列测定,以及对 SS 基因的克隆和序列测定,发现 SS 的氨基酸序列中存在 6 个保守的区域,其中 III、IV 和 V 3 个区域高度保守,与活性位点相关,在动物、植物和真菌中几乎完全一致^[14],它们在氨基酸序列上的位置分别为 168~183 的 YCHYVAGLVGL-GLSKL、202~223 的 GLFLQKTNIIRDYLEINEIPK 和 280~298 的 PAIFRFCAIPQIMAIGTLA,这些结构域中的氨基酸是该酶催化反应所必须的^[1]。而其 C 端区域为疏水序列具有广泛的多样性,用于锚定在内质网上^[15]。这些区域也同样存在于刺五加的鲨烯合酶中。

早期曾认为 SS 是以单一形式存在的,因为人和酵

母的 SS 基因都以单拷贝的形式存在,但是从拟南芥和欧亚甘草(licorice)等却分离出 2 种形式的 SS 基因,目前还难以判断有多少物种含有 2 种或更多种形式的鲨烯合酶基因^[6]。该试验中得到的刺五加 SS2 基因的氨基酸序列与刺五加 SS1 存在 34 个氨基酸的差别,一致性仅为 91.79%,而与人参 SS1 仅存在 17 个氨基酸的差别,一致性高达 95.89%。因此初步推断,所克隆的刺五加 SS2 基因应在和人参中全部器官均高表达的 SS1 为同类的可能性较大。跨膜区域分析的结果表明,刺五加 SS2 和 SS1 具有相同的跨膜结构,这说明二者在内质网膜上的镶嵌模式相似。

GenBank 序列查询表明,目前已登录的刺五加属植物 SS 序列只有先期克隆的刺五加 SS1 cDNA,该试验首次克隆了刺五加鲨烯合酶基因 SS2 的 cDNA 全长序列,为后续刺五加 SS 基因表达与刺五加皂苷生物合成的相互关系与调控措施研究奠定了基础。

参考文献

- [1] Kim T D, Han J Y, Huh G H, et al. Expression and functional characterization of three squalene synthase genes associated with saponin biosynthesis in *Panax ginseng*[J]. Plant Cell Physiol, 2011, 52(1): 125-137.
- [2] Kim Y S, Cho J H, Park S, et al. Gene regulation patterns in triterpene biosynthetic pathway driven by overexpression of squalene synthase and methyl jasmonate elicitation in *Bupleurum falcatum*[J]. Planta, 2011, 233(2): 343-355.
- [3] Haralampidis K, Trojanowska M, Osbourn A. Biosynthesis of triterpenoid saponins in plants[J]. Adv Biochem Eng Biotechnol, 2002, 75: 31-49.
- [4] Vogeli U, Chappell J. Induction of sesquiterpene cyclase and suppression of squalene synthetase activities in plant cell cultures treated with fungal elicitor[J]. Plant Physiol, 1988, 88(4): 1291-1296.
- [5] Seo J W, Jeong J H, Shin C G, et al. Overexpression of squalene synthase in *Eleutherococcus senticosus* increases phytosterol and triterpene accumulation[J]. Phytochemistry, 2005, 66(8): 869-877.
- [6] 赵明文, 钟家禹, 王南, 等. 鲨烯合酶的研究进展[J]. 微生物学报, 2003, 43(5): 676-680.
- [7] Uchida H, amashita H, Kajikawa M, et al. Cloning and characterization of a squalene synthase gene from a petroleum plant, *Euphorbia tirucalli* L. [J]. Planta, 2009, 229(6): 1243-1252.
- [8] 吴耀生, 朱华, 李坤, 等. 三七鲨烯合酶基因在三七根、茎、芦头中的转录表达与三萜皂苷合成[J]. 中国生物化学与分子生物学报, 2007, 23(12): 1000-1005.
- [9] Rasbery J M, Shan H, Leclair R J, et al. *Arabidopsis thaliana* squalene epoxidase 1 is essential for root and seed development[J]. J. Biol. Chem, 2007, 282(23): 17002-17013.
- [10] Busquets A, Keim V, Closa M, et al. *Arabidopsis thaliana* contains a single gene encoding squalene synthase[J]. Plant Mol. Biol., 2008, 67(1-2): 25-36.
- [11] 隋春, 魏建和, 战晴晴, 等. 北柴胡鲨烯合酶基因及其编码区 cDNA 克隆与序列分析[J]. 园艺学报, 2010, 37(2): 283-290.
- [12] 张宝香, 姜英, 张雅凤. 刺五加叶研究新进展[J]. 特产研究, 2009, 31(4): 69-70, 77.
- [13] 张凤侠, 梁新华, 王俊. 植物三萜皂苷生物合成及关键酶鲨烯合酶的研究进展[J]. 农业科学研究, 2009, 30(3): 64-68.
- [14] Robinson G W, Tsay Y H, Kienzle B K, et al. Conservation between human and fungal squalene synthetases: similarities in structure, function, and regulation[J]. Mol. Cell Biol., 1993, 13(5): 2706-2717.
- [15] Kribbi R, Arro M, Del A A, et al. Cloning and characterization of the *Arabidopsis thaliana* SQS1 gene encoding squalene synthase-involvement of the C-terminal region of the enzyme in the channeling of squalene through the sterol pathway[J]. Eur. J. Biochem, 1997, 249(1): 61-69.

Cloning of the Squalene Synthase Gene 2 cDNA from *Eleutherococcus senticosus* and Analysis of Its Putative Protein Structure

LONG Yue-hong, XING Zhao-bin, LIANG Neng-song, HE Shan, LI Bao-cai, ZHU Jin-li
(College of Life Science, Hebei United University, Tangshan, Hebei 063000)

Abstract: Full-length of *E. senticosus* squalene synthase gene 2 cDNA (GenBank number: JN714465) was cloned by RT-PCR and RACE methods. Sequence analysis revealed a 1 463 bp cDNA containing 10 bp 5'UTR, 208 bp 3'UTR and 1 245 bp ORF encoding 414 amino acids. The protein had one Trans_IPPS_HH conserved domain, two specificity target region of phytoene synthase and two transmembrane helixes. Amino acids comparison showed that the consistence of the protein and SS in other *Araliaceae* plants was more than 90%. The consistence with *E. senticosus* SS1 was 91.79%.

Key words: squalene synthase gene; *Eleutherococcus senticosus*; RT-PCR; RACE