

• 药材与资源 •

红花栝楼鲨烯合酶基因的克隆及其序列分析

陶晨陈¹, 马成通¹, 吴耀生^{1,2*}, 周青鸟¹, 苏荷玲², 晁耐霞², 罗育¹

1. 广西医科大学 生物化学与分子生物学教研室, 广西 南宁 530021

2. 广西高校生物分子医学研究重点实验室, 广西 南宁 530021

摘要: 目的 克隆红花栝楼鲨烯合酶(squalene synthase, SS)的基因并进行序列分析, 为进一步研究葫芦科植物三萜合成通路关键酶 SS 的正选择位点与功能的关联性分析奠定基础。方法 根据绞股蓝与罗汉果 SS 基因 cDNA 比对分析的结果, 设计 SS 的 5'端简并引物, 采用 3'RACE 扩增红花栝楼 SS 全长 cDNA 基因。结果 获得红花栝楼 SS cDNA 全长共 1 466 个核苷酸, 其中包括一个含有 1 254 个核苷酸的开放读码框, 编码 417 个氨基酸残基。通过 NCBI 的 Blast 比对, 发现红花栝楼 SS 基因编码的氨基酸序列与已知植物 SS 基因编码的氨基酸序列同源率为 78%~94%, 核苷酸的同源率为 74%~93%。结论 成功克隆了红花栝楼 SS 的全长 cDNA, 为进一步研究红花栝楼 SS 基因结构、基因表达、基因突变提供了基础, 并为葫芦科植物三萜合成通路关键酶 SS 的正选择位点与功能的关联性分析提供了数据支持。

关键词: 红花栝楼; 鲨烯合酶; RACE; 基因克隆; 序列分析

中图分类号: R282.12 文献标志码: A 文章编号: 0253-2670(2015)07-1034-08

DOI: 10.7501/j.issn.0253-2670.2015.07.018

Cloning and sequence analysis of squalene synthase gene from *Trichosanthes rubriflos*TAO Chen-chen¹, MA Cheng-tong¹, WU Yao-sheng^{1,2}, ZHOU Qing-niao¹, SU He-ling², CHAO Nai-xia², LUO Yu¹

1. Department of Biochemistry and Molecular Biology, Nanning 530021, China

2. Key Laboratory of Biological Molecular Medicine Research of Guangxi High Education. Guangxi Medical University, Nanning 530021, China

Abstract: Objective To clone the full length cDNA encoding squalene synthase (SS) from *Trichosanthes rubriflos* and to carry on its sequence analysis, so as to lay the foundation for the further study on the positively selected sites and function correlation analysis of SS which is the key enzyme for triterpene synthesis pathway. **Methods** According to the cDNA comparison on SS gene from *Gynostemma pentaphyllum* and *Siraitia grosvenorii*, 5'-upstream degenerate primers of the cDNA of SS gene from *T. rubriflos* were designed and the full length cDNA of SS gene from *T. rubriflos* was amplified by 3'RACE kit. **Results** The full length cDNA of SS gene from *T. rubriflos* composed of 1 466 nucleotides was obtained. The open reading frame (ORF) of SS gene from *T. rubriflos* was 1 254 bp in length, corresponding to a predicted polypeptide of 417 amino acid residues. The results of homologous alignment analysis in GenBank demonstrated that the cDNA sequence of SS gene from *T. rubriflos* had 78%—94% similarity on the nucleotide sequence compared with SS from known plant and 74%—93% similarity on the deduced amino acid sequence compared with SS gene from other plants. **Conclusion** The full length cDNA of SS gene from *T. rubriflos* has been cloned, which not only lays a foundation for the further study on the gene expression, gene structure, and gene mutation, but also provides the important data base for the association study between the positively selected sites and function correlation analysis of which is the key enzyme for triterpene synthesis pathway.

Key words: *Trichosanthes rubriflos* Thoms. ex Cayla; squalene synthase; rapid amplification of cDNA ends; gene cloning; sequence analysis

三萜皂苷是一类重要的植物次生代谢产物, 广泛分布于植物组织中, 具有抗菌和抗虫害的作用, 可以用作药物, 具有重要的商业价值^[1]。例如, 人参皂苷是人参的主要活性成分之一, 其各单体成分

普遍具有调节免疫系统、降低血清中胆固醇的量、防治心血管疾病、抗炎、解毒、抗癌等生理功效^[2]。研究表明, 三萜皂苷合成途径中的一些关键酶对组织中三萜皂苷的量和组成具有重要的调节作用^[3]。

收稿日期: 2014-11-30

基金项目: 国家自然科学基金资助项目(31260069); 广西高等学校科研项目(201203YB038)

作者简介: 陶晨陈(1985—), 男, 安徽合肥人, 在读硕士研究生, 研究方向为生物化学与分子生物学。

Tel: 18376645670 E-mail: 332931654@qq.com

*通信作者 吴耀生, 女, 博士, 教授, 硕士生导师, 主要从事植物分子生物学和基因工程方面的研究。

Tel: 15177171863 E-mail: wuyaosheng2012@gxmu.edu.cn

鲨烯合酶 (squalene synthase, EC 2.5.1.21, 简称 SS) 是三萜类化合物生物合成途径中的一个重要调控酶^[4]。SS 能催化 2 分子的法呢基焦磷酸 (farnesyl pyrophosphate, FPP) 缩合生成 1 分子鲨烯, 是三萜、甾醇、胆固醇等萜烯类重要物质生物合成的共同前体^[5]。因此, SS 的研究对了解三萜皂苷合成途径具有重要的参考价值。目前, 三萜生物合成通路关键酶 SS 的基因克隆、进化分析已有相关报道, 如陆生植物 SS 适应性进化正选择位点分析^[6], 绞股蓝鲨烯环氧酶基因的克隆与序列分析^[7], 以及人参、金铁锁、青蒿、丹参^[8-13]等。葫芦科植物三萜生物合成通路关键酶 SS 基因序列在 GenBank 中已收录的目前只有罗汉果 (HQ128565) 和绞股蓝 (FJ906799)。

红花栝楼 *Trichosanthes rubriflos* Thoms. ex Cayla 为葫芦科 (Cucurbitaceae) 栝楼属 *Trichosanthes* L. 植物, 具有一定的药用价值, 国内主要分布在广东、广西、贵州等省区, 国外主要分布于印度东北部、缅甸、泰国等地区。根据《中国药典》2010 年版所收载, 中药瓜蒌的植物基原是葫芦科植物栝楼 *Trichosanthes kirilowii* Maxim. 及双边栝楼 *Trichosanthes rosthornii* Harms 的干燥成熟果实, 而红花栝楼等其他栝楼属植物的果实则是混杂品。但也有些本草记载, 红花栝楼的根部也可药用, 其味甘、微苦, 性寒, 有小毒; 主要功能为清肺化痰、解毒散结; 主肺热咳嗽、胸闷胸痛、便秘、疟疾、疮疖肿毒^[14]。有研究表明, 在栝楼属植物中可分离得到葫芦素^[15], 但有关红花栝楼的研究仍未见报道, 亦未知是否含有葫芦素。SS 是三萜合成通路关键酶, 因此红花栝楼 SS 基因序列的克隆分析对了解其三萜生物合成通路具有重要意义。

cDNA 末端快速扩增技术 (rapid amplification of cDNA ends, RACE) 是一种始于 mRNA 的 3' 或 5' 端, 最终获得目的基因全长 cDNA 的分子生物学技术^[16]。本课题组已运用 RACE 技术成功克隆出了绞股蓝 SS、SE、FPS 等基因^[7,17], 现继续运用该技术对红花栝楼三萜合成途径关键酶 SS 基因进行克隆及分析。

1 材料与方法

1.1 材料与试剂

样品于 2013 年 9 月采集自广西南宁市广西中医药大学药圃。经广西中医药大学朱意麟讲师鉴定为红花栝楼 *Trichosanthes rubriflos* Thoms. ex Cayla, 存放于广西医科大学分子医学实验室 -80 °C 超低温冰箱。大肠杆菌 DH5 α 感受态细胞、LATAq DNA

聚合酶限制性内切酶、3'-Full RACE Core Set Ver 2.0、柱式胶回收试剂盒及质粒小量提取试剂盒等为 TaKaRa 公司产品; DNA 寡核苷酸引物由上海生物工程公司合成; PEASY-T1 载体为全式金公司产品; RCR 仪为 Biometra An Analytik Jena Company 公司产品; 其他试剂均为国产分析纯产品; 本研究利用的不同植物 SS 的相关信息见表 1。

1.2 红花栝楼总 RNA 提取

参照蒋军富等^[7]的异硫氰酸胍法, 提取红花栝楼叶片总 RNA。

1.3 3'RACE PCR 克隆

以绞股蓝和罗汉果的 SS 基因为基础, 经比对分析找同源区, 设计出扩增红花栝楼 SS cDNA 全长所需的 5' 端简并引物 ssjbf2 和 ssjbf3 (表 2), 由上海生物工程有限公司合成。以红花栝楼总 RNA 为模板, 加入 3'RACE Adaptor 等试剂, 按试剂盒说明书操作反转录合成第一链 cDNA。反应体系包括红花栝楼总 RNA 1 μ L、3'RACE Adaptor (5 μ mol/L) 1 μ L、5 \times Prime Script Buffer 2 μ L、dNTP Mixture (10 mmol/L each) 1 μ L、RNase Inhibitor (40 U/ μ L) 0.25 μ L、PrimeScript RTase (200 U/ μ L) 0.25 μ L, 用 RNase Free dH₂O 补至总体积 10 μ L。42 °C 反转录 60 min, 70 °C 作用 15 min, 反应产物用于巢式 PCR。以反转录得到的第一链 cDNA 为模板, 分别以 ssjbf2、ssjbf3 与 3'RACE 试剂盒中的 3'RACE Outer Primer、3'RACE Inner Primer 为引物, 按照 3'-Full RACE Core Set Ver 2.0 说明进行 3'RACE 外侧、内侧 2 轮巢式 PCR 扩增反应。第 1 轮 PCR 反应条件为 94 °C 预变性 3 min, 然后以 94 °C 变性 30 s、55 °C 退火 30 s、72 °C 延伸 100 s 运行 20 个循环, 72 °C 延伸 10 min。以第 1 轮 PCR 产物为模板进行第 2 轮巢式 PCR。第 2 轮 PCR 反应条件为 94 °C 预变性 3 min, 然后以 94 °C 变性 30 s、55 °C 退火 30 s、72 °C 延伸 100 s 运行 30 个循环, 72 °C 延伸 10 min。用琼脂糖凝胶电泳判断 PCR 扩增结果。Inner PCR 产物纯化后, 与 PEASY-T1 连接, 连接产物转化大肠杆菌 DH5 α , 在 X-gal/IPTG/Amp LB 琼脂平板上挑取白色菌落摇菌扩增培养, 提取质粒并进行 PCR 鉴定, 将阳性克隆送华大基因公司测序。

1.4 红花栝楼 SS 基因 cDNA 全长序列拼接与生物信息学分析

利用 Vector NTI Suite 6.0 软件对 3'RACE 所得片段的测序结果进行分析与拼接。利用 DNAMAN

表 1 不同植物 SS 的相关信息

Table 1 Related information of SS gene from different plants

| 序号 | 种名 | 拉丁名 | 科名 | 编码氨基酸 残基数 | 提交 国家 | GenBank 核酸 序列登录号 | GenBank 氨基酸 序列登录号 |
|----|-------|-----------------------------------|------|--------------|----------|---------------------|----------------------|
| 1 | 楸木 | <i>Aralia elata</i> | 五加科 | 414 | 中国 | GU354313 | ADC32 654.1 |
| 2 | 百脉根 | <i>Lotus japonicus</i> | 豆科 | 413 | 日本 | AB102688 | BAC56 854.1 |
| 3 | 北柴胡 | <i>Bupleurum chinense</i> | 伞形科 | 415 | 中国 | GQ889267 | ACX42 424.1 |
| 4 | 本氏烟 | <i>Nicotiana benthamiana</i> | 茄科 | 411 | 美国 | U46000 | AAA87 048.1 |
| 5 | 布朗葡萄藻 | <i>Botryococcus braunii</i> | 葡萄藻科 | 461 | 美国 | AF205791 | AF205791_1 |
| 6 | 大豆 | <i>Glycine max</i> | 豆科 | 413 | 日本 | AB007503 | BAA22 559.1 |
| 7 | 丹参 | <i>Salvia miltiorrhiza</i> | 唇形科 | 413 | 中国 | FJ768961 | ACR57 219.1 |
| 8 | 东北红豆杉 | <i>Taxus cuspidata</i> | 红豆杉科 | 409 | 中国 | DQ836053 | DQ836 053.1 |
| 9 | 番茄 | <i>Solanum lycopersicum</i> | 茄科 | 411 | 美国 | NM_001247787 | NP_001234 716.1 |
| 10 | 甘草 | <i>Glycyrrhiza uralensis</i> | 豆科 | 412 | 中国 | AM182330 | CAJ77 653.1 |
| 11 | 高粱 | <i>Sorghum bicolor</i> | 禾本科 | 403 | 美国 | XM_002466214 | XP_002466 259.1 |
| 12 | 黄甘草 | <i>Glycyrrhiza eurycarpa</i> | 豆科 | 413 | 中国 | AM182331 | CAJ77 654.1 |
| 13 | 黄花蒿 | <i>Artemisia annua</i> | 菊科 | 418 | 中国 | AY445506 | AAR20 329.1 |
| 14 | 积雪草 | <i>Centella asiatica</i> | 伞形科 | 415 | 韩国 | AY787628 | AAV58 897.1 |
| 15 | 假马齿苋 | <i>Bacopa monnieri</i> | 玄参科 | 414 | 印度 | GU734711 | ADX01 171.1 |
| 16 | 江南卷柏 | <i>Selaginella moellendorffii</i> | 卷柏科 | 392 | 美国 | XM_002971905 | XP_002971 951.1 |
| 17 | 绞股蓝 | <i>Gynostemma pentaphyllum</i> | 葫芦科 | 417 | 中国 | FJ906799 | ACQ90 302.1 |
| 18 | 金铁锁 | <i>Psammosilene tunicoides</i> | 石竹科 | 414 | 中国 | EF585250 | ABQ96 265.1 |
| 19 | 辣椒 | <i>Capsicum annuum</i> | 茄科 | 411 | 韩国 | AF124842 | AAD20 626.1 |
| 20 | 陆地棉 | <i>Gossypium hirsutum</i> | 锦葵科 | 412 | 中国 | EF688567 | ABX10 442.1 |
| 21 | 绿玉树 | <i>Euphorbia tirucalli</i> | 大戟科 | 411 | 日本 | AB433916 | BAH23 428.1 |
| 22 | 罗汉果 | <i>Siraitia grosvenorii</i> | 葫芦科 | 417 | 中国 | HQ128565 | AEM42 980.1 |
| 23 | 马铃薯 | <i>Solanum tuberosum</i> | 茄科 | 411 | 日本 | AB022599 | BAA82 093.1 |
| 24 | 毛果杨 | <i>Populus trichocarpa</i> | 杨柳科 | 413 | 美国 | XM_002313729 | XP_002313 765.1 |
| 25 | 膜荚黄芪 | <i>Astragalus membranaceus</i> | 蝶形花科 | 413 | 中国 | HQ829974 | ADW27 427.1 |
| 26 | 木榄 | <i>Bruguiera gymnorhiza</i> | 红树科 | 416 | 中国 | GU478981 | ADD54 645.1 |
| 27 | 人参 | <i>Panax ginseng</i> | 五加科 | 415 | 韩国 | AB115496 | BAD08 242.1 |
| 28 | 三七 | <i>Panax notoginseng</i> | 五加科 | 415 | 中国 | DQ186630 | ABA29 019.1 |
| 29 | 蛇足石杉 | <i>Huperzia serrata</i> | 石杉科 | 420 | 中国 | JQ004938 | AEX58 673.1 |
| 30 | 柿 | <i>Diospyros kaki</i> | 柿科 | 415 | 中国 | FJ687954 | ACN69 082.1 |
| 31 | 鼠耳芥 | <i>Arabidopsis thaliana</i> | 十字花科 | 410 | 日本 | D29017 | BAA06 103.1 |
| 32 | 水稻 | <i>Oryza sativa</i> | 禾本科 | 403 | 日本 | AB007501 | BAA22 557.1 |
| 33 | 睡茄 | <i>Withania somnifera</i> | 茄科 | 411 | 印度 | GU474427 | ADC95 435.1 |
| 34 | 西洋参 | <i>Panax quinquefolius</i> | 五加科 | 415 | 中国 | GU997681 | AED99 863.1 |
| 35 | 烟草 | <i>Nicotiana tabacum</i> | 茄科 | 411 | 美国 | U60057 | AAB08 578.1 |
| 36 | 洋甘草 | <i>Glycyrrhiza glabra</i> | 豆科 | 413 | 日本 | D86409 | BAA13 083.1 |
| 37 | 银柴胡 | <i>Bupleurum falcatum</i> | 石竹科 | 415 | 韩国 | AY964186 | AAV46 017.1 |
| 38 | 玉蜀黍 | <i>Zea mays</i> | 禾本科 | 401 | 日本 | AB007502 | BAA22 558.1 |
| 39 | 远志 | <i>Polygala tenuifolia</i> | 远志科 | 413 | 韩国 | DQ672339 | ABG66 304.1 |
| 40 | 蒺藜苜蓿 | <i>Medicago truncatula</i> | 豆科 | 423 | 美国 | XM_003606992 | XP_003607 040.1 |

表 2 SS 的 3'RACE 引物
Table 2 Primers of 3'RACE for SS

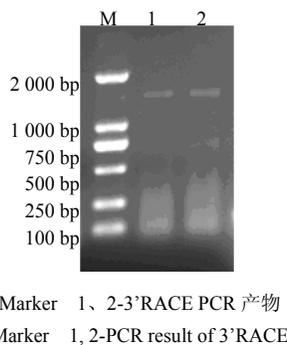
| 引物 | 序列 | 碱基数 |
|---------------------|----------------------------------|-----|
| ssjbf2 | ATATWKAGAGVSAGAAATGGGCA | 23 |
| 3'RACE Outer Primer | TACCGTCGTTCCACTAGTGATT | 23 |
| ssjbf3 | ATATGGGVAGYTTKGGRCRAT | 22 |
| 3'RACE Inner Primer | CGCGGATCCTCCACTAGTGATTTCACTATAGG | 32 |

软件分析红花栝楼 SS 基因 cDNA 全长序列性质, 序列测定结果采用 NCBI Blast 进行同源性比对搜索。用 Vector NTI Suite 6.0 软件对红花栝楼和已报道的植物 SS 氨基酸序列进行相似性比对, 构建 SS 基因系统进化树, 并进行相关生物信息学分析^[18-20]。

2 结果与分析

2.1 红花栝楼 SS 基因 3'RACE

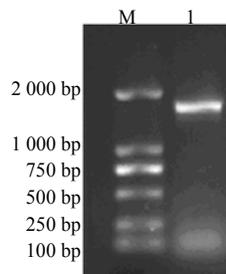
通过 3'RACE 得到 1 700 bp 左右的特异条带, 见图 1。通过菌落 PCR 得到 1 800 bp 左右的特异条带, 见图 2。



M-Marker 1、2-3'RACE PCR 产物
M-Marker 1、2-PCR result of 3'RACE

图 1 3'RACE 产物电泳图

Fig. 1 Electrophoretogram of 3'RACE product



M-Marker 1-SS PCR 产物
M-Marker 1-PCR product of SS

图 2 菌落 PCR 电泳图

Fig. 2 Electrophoretogram of colony PCR

2.2 红花栝楼 SS cDNA 基因的核苷酸序列分析

根据 3'RACE 扩增测序结果, 由 Vector NTI Suite 6.0 软件拼接及人工校对后得到红花栝楼 SS 基因全长序列。该基因 cDNA 全长 1 466 bp, 翻译

起始位点为 3 碱基处, 翻译终止点位于 1 255 碱基处, 开放读码框由 1 254 个碱基组成。

红花栝楼 SS 基因 cDNA 核苷酸及推衍的氨基酸序列见图 3。

2.3 红花栝楼 SS 基因推导的氨基酸序列分析

由 SWISS-MODEL Workspace 在线分析软件建立红花栝楼 SS 基因编码蛋白三维结构模型如图 4 所示。根据 Protscale 中的 Kyte & Doolittle 算法对蛋白质疏水性与吸水性预测可知 (图 5), MIN: -2.300, MAX: 3.344, 正值越大表示越疏水, 负值越大表示越亲水, 介于 0.5~-0.5 的主要为两性氨基酸, 其中 Arg 有最低的亲水性分值-4.5, 即亲水性最强, 而 Ile 疏水性最强, 其分值为 4.5。根据整体肽链可知, 其中在 386~408 区域的疏水性最强, 其次 335~337、283~340、175~181、69~76、46~49 区域具有较强的疏水性; 214~254 区域的亲水性最强, 其次 17~21、25~37、104~115、137~144、161~162、164~166、350~361、367~370、373~383 区域具有较强的亲水性。红花栝楼 SS 多肽链的亲水区域大于疏水区域, 预测该蛋白属于亲水性蛋白。用 Topcons 预测红花栝楼 SS 跨膜区域, 此蛋白含有 3 个跨膜区, 分别位于 58~78 位、281~301 位、386~406 位。

通过与葫芦科植物罗汉果和绞股蓝的 SS 氨基酸序列比对分析, 得到红花栝楼 SS 基因的开放阅读框编码由 417 个氨基酸残基构成。通过 <http://www.expasy.org/> 网站在线分析, 得到红花栝楼 SS 的氨基酸序列等电点为 7.90, 相对分子质量为 47 594.33, 二级结构中含 α 螺旋 (alpha helix) 68.82%, β 转角 (beta turn) 2.88%, 扩展链 (extended strand) 4.56%, 无规卷曲 (random coil) 23.74%。

利用 PSORT Prediction 对红花栝楼 SS 蛋白的亚细胞定位分析得到 SS 位于膜结构上。SS 定位于质膜上的可能性为 0.73, 明显大于微粒体、线粒体内膜、内质网膜上的可能性, 所以 SS 定位于质膜上

```

1  at ATGGGGAGTTTGGGGCGATTGAGACACCCAGATGACTTTTACCCGCTTTGAAACT
1  M G S F G A I L R H P D D F Y P L L K L
62  GAAAATGGCTGCTAGACATGCGGAGAAGCAGATCCCGTCCGAGCCTCATTGGGGATTTG
21  K M A A R H A E K Q I P S E P H W G F C
122  CTATTCCATGCTGCATAAGGTCTCACGAAGTTTGGCTCTCGTTATTTCAGCAGCTTACGCC
41  Y S M L H K V S R S F A L V I Q Q L T P
182  CGACCTTCGAAATGCCGTATGCATATTTTATCTTGTCTGAGAGCCCTCGACACTGTTGA
61  D L R N A V C I F Y L V L R A L D T V E
242  GGATGATACAAGCATACAACAGATATCAAAGTGCCCATTTTGAAGCTTTTCACTGTCA
81  D D T S I Q T D I K V P I L K A F H C H
302  CATATAACCGTGATTGGCATTTTTCATGTGCCACAAAGGACTATAAAGTCTTATGGA
101  I Y N R D W H F S C G T K D Y K V L M D
362  TGAGTTTCATCATGTTTCACTGCGTTTCTAGAAGTTGGGAAAGGGTACCAGGAAGCGAT
121  E F H H V S T A F L E L G K G Y Q E A I
422  TGAGGATATCACGAAACGAATGGGTGCAGGAATGGCTAAATTCATTGCAAAGAGGTAGA
141  E D I T K R M G A G M A K F I C K E V E
482  GACAGTTGACGATTATGATGAATATTGCCACTATGTAGCTGGACTTGTCCGACTAGVTT
161  T V D D Y D E Y C H Y V A G L V G L G L
542  GTCTAAGTTGTTCCACGCTTTCGAAGTTAGAGGATTGGCACCTGATTCTCTTCCAATTC
181  S K L F H A S K L E D L A P D S L S N S
602  TATGGGATTGTTTCTTCAGAAAACAAACATTATTTCGAGATTACTTGGAAAGACATCAATGA
201  M G L F L Q K T N I I R D Y L E D I N E
662  GATACCAAAGTCTCGGATGTTTTGGCCTCGTGAGATTGGAGCAAATATGCTGATAAAT
221  I P K S R M F W P R E I W S K Y A D K L
722  AGAGGATTTAAATATGAGGAGAACTCGGTCAAGGCCGTGCAATGCCTCAATGATTCCGT
241  E D F K Y E E N S V K A V Q C L N D S V
782  CACTAATGCTTTGAACCATGTGGAGGATTGTCTGAAATACATGCTCAACTAAGAGATCT
261  T N A L N H V E D C L K Y M S N L R D L
842  TTCCATATTTGATTTTGTGCAATCCCTCAGATTATGGCGATTGGAACCCCTAGCATTATG
281  S I F R F C A I P Q I M A I G T L A L C
902  CTACAACAACGTCGAAGTCTTCAGGGGGGTGGTCAAATGCGTCCGGGTCTTACTGCAA
301  Y N N V E V F R G V V K M R R G L T A K
962  GATCATTGATCGGACAAAACGATTGCGAGCGTCTATGGATCTTTCTTTGATTTTCTGT
321  I I D R T K T I A D V Y G S F F D F S V
1 022  TATGCTGAAGGCTAAGGTCAACAGCAATGATCCTAATGCTGCTAAAACCCCTCAGCAGAAT
341  M L K A K V N S N D P N A A K T L S R I
1 082  CGAGGCAATACAGAAAACATGCAAGGAGTCCGGACTCTTGAACAAGAGGAAATGTATGT
361  E A I Q K T C K E S G L L N K R K L Y V
1 142  GGTAAAGAAGTGAGCCAATGTACAATCCAGCTGCGATTGTCATACTTTTCCAGCCTAATATG
381  V R S E P M Y N P A A I V I L F S L I C
1 202  CATCATTCTCGCATACTCTCTGCCAAACAGTTACCAGCTAACCAACCCATA TGAgcatc
401  I I L A Y L S A K Q L P A N Q P I * - -
1 262  gaactccggatacaaagttcagcagagcacaccgtaaacttggagcggtatgattagttt
1 322  ctggctcctttatctctcctcagagctttgtttaaatattacaatggctgtaagttactgt
1 382  ttgtttagtacagaccaggaaacaagtgagatctctgtagctctctaatgggtgtagtgtaa
1 442  tggttgacatggaaaaaaaaaaaa

```

斜体 ATG 为起始密码子，斜体加下划线 TGA 为终止密码子；小写字母表示 5'和 3'端非翻译区，大写字母表示编码区；上行为核苷酸序列，下行为氨基酸序列

ATG is initiation codon in bold and italics, TGA is termination codon in bold and indicated by underlined; Lowercase letters represent 5'and 3'untranslation regions, Capital letters represent encoding regions; Up rows indicate nucleotide sequence; Down rows indicate amino acid squence

图 3 红花栝楼 SS cDNA 序列及推导的氨基酸序列

Fig. 3 cDNA sequence and deduced amino acid sequence for SS gene from *T. rubriflos*

的可能性最大。

2.4 陆生植物 SS 基因序列比较及进化树分析

经 NCBI Blast 比对，发现本实验克隆得到的红花栝楼 SS 基因的氨基酸序列与已知植物 SS 的同源性为 78%~94%，核酸序列同源性为 74%~93%。其中，氨基酸序列与罗汉果和绞股蓝相似性分别为 94%和 88%，核苷酸序列与罗汉果和绞股蓝相似性分别为 93%和 88%。保守序列 (conservedsequence) 对推测进化结果具有特殊意义，在 NCBI 的蛋白保

守结构域数据库 (conserved domain database, CDD) 中对红花栝楼 SS 进行蛋白保守区预测，结果表明与该基因匹配的蛋白为 SS，保守序列在 N 端和中间 (图 6)。用 Vector NTI Suite 6.0 软件将得到的红花栝楼 SS 与不同植物的 SS 的氨基酸序列进行多重比对，构建得到进化树 (图 7)。

3 讨论

药用植物的生物合成途径关键酶基因的克隆、表达和调控是近年来的研究热点。植物体内三萜皂

昔主要通过甲羟戊酸 (mevalonic acid, MVA) 途径合成^[5]。MVA 途径中, SS 处于 FPP 到其他产物的分支点上, FPP 除可以被 SS 催化产生鲨烯 (SQ) 外, 还可以在其他酶的催化下产生赤霉素、类胡萝



图 4 红花栝楼 SS 基因编码蛋白三维结构模型

Fig. 4 3D structure of SS-encoding protein from *T. rubriflos*

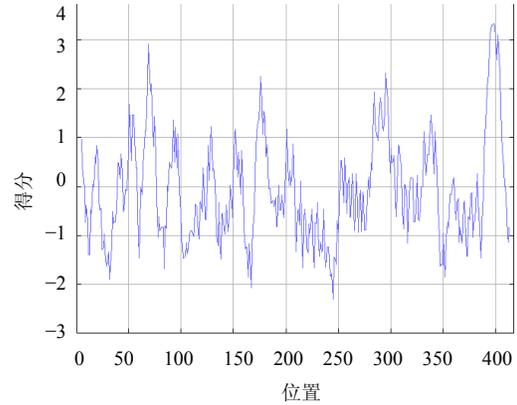


图 5 红花栝楼 SS 的疏水性与亲水性预测

Fig. 5 Hydrophobicity and hydrophilicity prediction of SS gene from *T. rubriflos*

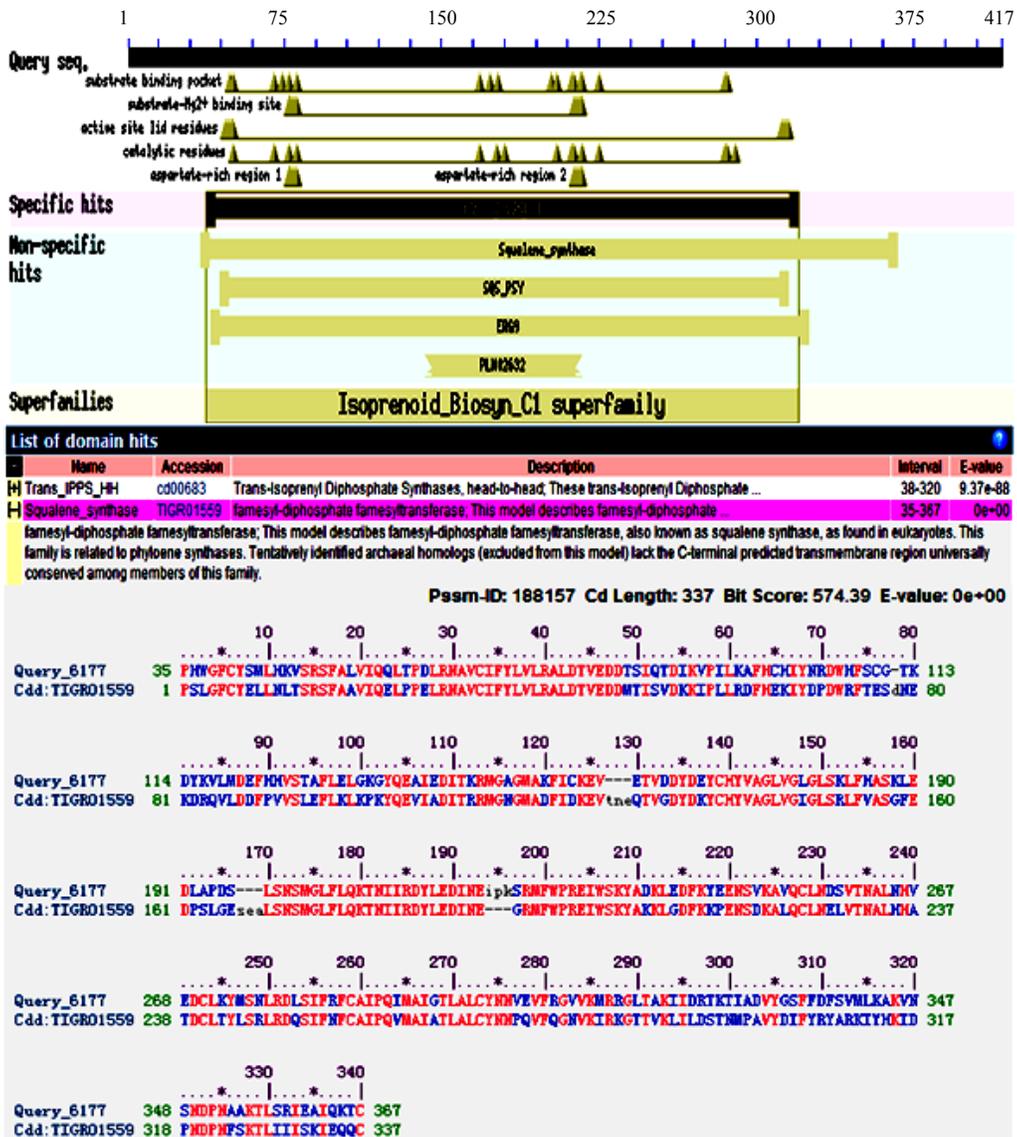


图 6 红花栝楼 SS 保守结构区域的预测

Fig. 6 Prediction of conserved domain of SS protein from *T. rubriflos*

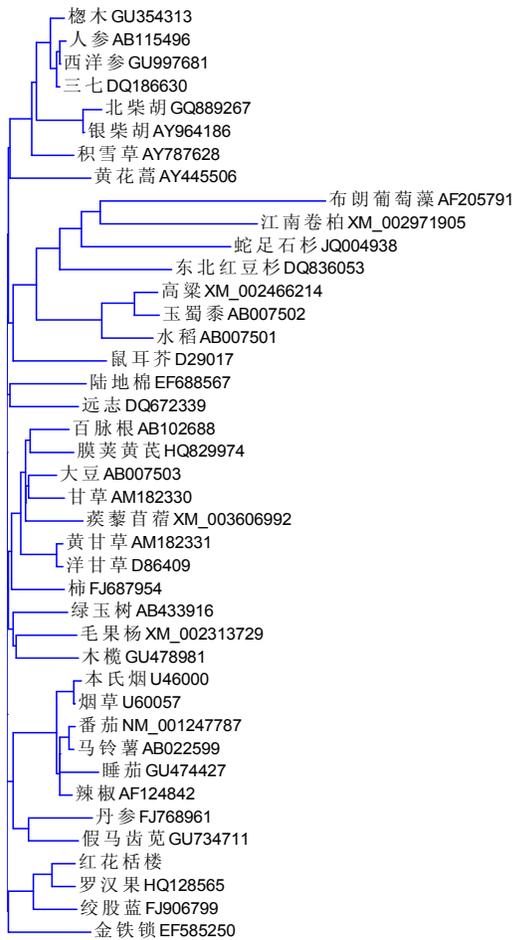


图 7 不同植物 SS 氨基酸序列进化树

Fig. 7 Phylogenetic tree based on amino acid sequence of SS from different plants

毒素等。SQ 是所有三萜皂苷、甾醇、胆固醇等萜烯类重要物质的共同前体^[21]，因此对 SS 的研究具有重要意义。

通过 GenBank 序列查询表明，目前尚无关于红花栝楼 SS 基因序列的报道。本研究首次克隆出了红花栝楼 SS cDNA 基因并对其进行了相关的生物信息学分析，为进一步研究红花栝楼 SS 基因结构、基因表达、基因突变提供了基础，并为三萜合成通路关键酶 SS 的正选择位点与功能的关联性分析提供了数据支持。

根据已知的 SS 基因设计出红花栝楼 SS 基因 5' 端的简并引物，运用 3'RACE 试剂盒直接扩增出 SS 基因的全长，省去了 5'RACE 的步骤，方法简单可行，节省经费开支，为后续研究提供了基础。本研究克隆得到的红花栝楼 SS 基因与已知的罗汉果 SS 基因的编码区比较，一致性达到 93%，具有较高的同源性。通过 NCBI 网站和一些生物信息学软件对

红花栝楼 SS 基因的碱基分布、氨基酸组成进行分析，结果发现，该基因中 G+C 的碱基量约为 41.54%，低于 50%，表明发生错配的概率较低，核苷酸处于较稳定状态。通过 Clustal X 等生物信息学软件将红花栝楼 SS 氨基酸与其他物种氨基酸的序列进行多重比对和进化分析可知，序列之间的同源性较高。例如，红花栝楼氨基酸序列：76~82 (LDTVEDD)、209~214 (NIIRDY)、216~220 (EDINE) 等区段同源性达 100%，说明其在进化过程中具有高度保守性，对其他物种 SS 基因的克隆与分离提供重要的数据支持。同时这些保守区基因信息为其他物种中该基因的克隆提供了十分有价值的序列信息，为加快基因克隆及 SS 基因的分子调控研究奠定基础。从植物 SS 氨基酸序列的进化树看到，亲缘关系近的科属，其 SS 的氨基酸序列一致性更高。其中，五加科与伞形科 SS 的进化关系更密切。葫芦科植物 SS 的序列尚有待增加。较大差异的基因区段主要集中在 C 端。今后可通过实验研究进一步验证通过 NCBI 的蛋白保守结构域数据库 (conserved domain database, CDD) 中对红花栝楼 SS 进行蛋白保守区预测的正确性。这些差异性较大的区段可能与基因的进化与变异相关，这些仍待确定。

参考文献

- [1] Henry M, Rahier A, Taton M. Effect of gypsogenin 3-O-glucuronide pretreatment of *Gypsophila paniculata* and *Saponaria officinalis* cell suspension cultures on the activities of microsomal 2,3-oxidosqualene cycloartenol and amyrrin cyclases [J]. *Phytochemistry*, 1992, 31(11): 3855-3859.
- [2] 张丹, 刘耀平, 鱼红闪, 等. 人参皂苷 8-葡萄糖苷酶的分离纯化及其酶学特性 [J]. *应用与环境生物学报*, 2003, 9(3): 259-262.
- [3] Haralampidis K, Trojanowska M, Osbourn A E. Biosynthesis of triterpenoid saponins in plants [J]. *Adv Biochem Eng Biotechnol*, 2002(75): 31-49.
- [4] 卢虹玉, 刘敬梅, 阳文龙, 等. 甘草鲨烯合酶基因的表达及植物表达载体的构建 [J]. *药物生物技术*, 2007, 14(4): 255-258.
- [5] Jennings S M, Tsay Y H, Fisch T M, et al. Molecular cloning and characterization of the yeast gene for squalene synthetase [J]. *Proc Natl Acad Sci USA*, 1991, 88(14): 6038-6042.
- [6] 刘镛, 吴耀生, 胡艳玲, 等. 陆生植物鲨烯合酶适应性进化正选择位点分析 [J]. *中国生物化学与分子生物*

- 学报, 2013, 29(1): 91-97.
- [7] 蒋军富, 李雄英, 吴耀生, 等. 绞股蓝鲨烯环氧酶基因的克隆与序列分析 [J]. 西北植物学报, 2010, 30(8): 1520-1526.
- [8] 毛乐心, 赵昶灵, 支伟特, 等. 人参属植物鲨烯合酶编码基因及其氨基酸序列的生物信息学分析 [J]. 中国农学通报, 2012, 28(31): 220-226.
- [9] 戴住波, 钱子刚, 胡运乾, 等. 金铁锁鲨烯合酶 cDNA 的克隆和功能鉴定 [J]. 药学学报, 2008, 43(12): 1245-1250.
- [10] 常 晶, 郭春华, 尹永志, 等. 青蒿鲨烯合酶 cDNA 的克隆与序列分析 [J]. 安徽农业科学, 2010, 38(10): 4996-4998.
- [11] 张 毅, 刘 彦, 王 红. 转青蒿反义鲨烯合酶基因对烟草鲨烯合酶基因表达的影响 [J]. 农业生物技术学报, 2005, 13(4): 416-422.
- [12] 马艺沔, 袁丽钗, 张林甦, 等. 2 个丹参鲨烯合酶基因的克隆和鉴定 [J]. 中草药, 2014, 45(9): 1307-1312.
- [13] 卢虹玉, 刘敬梅, 阳文龙. 甘草鲨烯合成酶基因的分离及植物表达载体的构建 [J]. 药物生物技术, 2007, 14(4): 255-258.
- [14] 中国科学院中国植物志编辑委员会. 中国植物志 [M]. 北京: 科学出版社, 1986.
- [15] Takahashi N, Yoshida Y, Sugiura T, *et al.* Cucurbitacin D isolated from *Trichosanthes kirilowii* induces apoptosis in human hepatocellular carcinoma cells *in vitro* [J]. *Int Immunopharmacol*, 2009, 9(4): 508-513.
- [16] Scotto-Lavino E, Du G, Frohman M A. Amplification of 5'end cDNA with 'new RACE' [J]. *Nat Protoc*, 2006, 1(6): 3056-61.
- [17] 蒋 东, 唐银琳, 陶晨陈, 等. 绞股蓝法呢基焦磷酸合酶基因的克隆及其序列分析 [J]. 生物技术通讯, 2014, 25(2): 198-202.
- [18] 张 萍, 刘迪秋, 葛 锋, 等. 三七 3-羟基-3-甲基戊二酸单酰辅酶 A 还原酶基因的克隆和生物信息学分析 [J]. 中草药, 2014, 45(18): 2684-2690.
- [19] 陈 莉, 蓝秀万, 李 坤, 等. 三七法呢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.
- [20] Mount D W. 序列与基因组分析 [M]. 北京: 高等教育出版社, 2003.
- [21] 张风侠, 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.