

## • 药材与资源 •

## 珊瑚菜 GIPS1、GIPS2 基因克隆及表达分析

宋洁洁<sup>1</sup>, 罗红梅<sup>2</sup>, 马小晶<sup>1</sup>, 高婷<sup>1,3\*</sup>

1. 青岛农业大学生命科学院, 山东省高校植物重点实验室, 山东 青岛 266109

2. 中国医学科学院 北京协和医学院药用植物研究所, 中草药物质基础与资源利用教育部重点实验室, 北京 100193

3. 上海市资源植物功能基因组学重点实验室(上海辰山植物园), 上海 201602

**摘要:** 目的 克隆珊瑚菜 *Glehnia littoralis* 补骨脂素合成酶 (psoralen synthase, PS) 基因全长 cDNA 序列, 并进行生物信息学和表达量分析。方法 在前期珊瑚菜转录组测序的基础上, 筛选出表达量较高的 2 条 PS 基因 GIPS1 和 GIPS2 序列。利用 cDNA 末端快速扩增 (rapid amplification of cDNA ends, RACE) 方法对 2 个基因 cDNA 序列 3'端进行克隆, DNAMAN 拼接后得到基因全长 cDNA 序列, 并对其编码蛋白进行生物信息学分析。利用实时荧光定量 PCR (RT-qPCR) 方法检测 GIPS1 和 GIPS2 基因的组织特异性表达。结果 GIPS1 和 GIPS2 基因 cDNA 序列全长分别为 1 885 bp 和 1 971 bp, 编码 495 和 502 个氨基酸残基, 蛋白相对分子质量分别为 55 740.7 和 56 363.9, 等电点为 8.28 和 6.62, 均属于细胞色素 P450 超家族, 含有 1 个跨膜区, 为亲水性蛋白。系统进化分析表明, GIPS1 和 GIPS2 与伞形科欧防风 *Pastinaca sativa*、旱芹 *Apium graveolens*、大阿米芹 *Ammi majus* PS 蛋白亲缘关系较近。RT-qPCR 结果显示 GIPS1 基因在根中表达量较高, 在叶中表达量较低, 而 GIPS2 基因在花中表达量较高, 在根中表达量较低。结论 首次获得珊瑚菜 GIPS1 和 GIPS2 基因的全长 cDNA 序列并进行了表达分析, 为进一步开展珊瑚菜补骨脂素合成酶基因的功能和遗传调控机制研究奠定基础。

**关键词:** 珊瑚菜; 补骨脂素合成酶; RACE; 生物信息学分析; 组织特异性表达

中图分类号: R282.12 文献标志码: A 文章编号: 0253-2670(2018)05-1139-07

DOI: 10.7501/j.issn.0253-2670.2018.05.023

## Cloning and expression analysis of GIPS1, GIPS2 gene in *Glehnia littoralis*

SONG Jie-jie<sup>1</sup>, LUO Hong-mei<sup>2</sup>, MA Xiao-jing<sup>1</sup>, GAO Ting<sup>1,3</sup>

1. Key Laboratory of Plant Biotechnology in Universities of Shandong Province, College of Life Sciences, Qingdao Agricultural University, Qingdao 266109, China

2. Key Laboratory of Bioactive Substances and Resources Utilization of Chinese Herbal Medicine, Ministry of Education, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100193, China

3. Shanghai Key Laboratory of Plant Functional Genomics and Resources (Shanghai Chenshan Botanical Garden), Shanghai 201602, China

**Abstract: Objective** To clone the full-length cDNA sequences of psoralen synthase (PS) genes from *Glehnia littoralis* so as to perform the bioinformatic and expression pattern analysis. **Methods** Based on our previous transcriptome sequencing data of *G. littoralis*, the gene sequences GIPS1 and GIPS2 with high expression level were screened. The 3'cDNA ends of GIPS1 and GIPS2 genes were cloned by the RACE (rapid amplification of cDNA ends) method and the full-length cDNA of genes were assembled by using DNAMAN software. And then encoded GIPS proteins were analyzed by the bioinformatics tools. The issue specific expression of GIPS1 and GIPS2 genes were detected using qPCR. **Results** The full-length cDNA of GIPS1 gene was 1 885 bp, which encoding a protein of 495 amino acids with a predicted molecular weight of 55 740.7 and isoelectric point of 8.28; The full-length cDNA of GIPS2 gene was 1 971 bp, which encoding a protein of 502 amino acids with a predicted molecular weight of 56 363.9 and isoelectric point of 6.62. GIPS1 and GIPS2 proteins belong to the cytochrome P450 superfamily, which share one transmembrane zone acting as hydrophilic protein. Phylogenetic analysis showed GIPS1 and GIPS2 were genetically closely related to the PS of *Pastinaca sativa*,

收稿日期: 2017-08-09

基金项目: 国家自然科学基金青年项目(31500230); 上海市资源植物功能基因组学重点实验室开放课题(PFGR20170); 山东省青岛市青年教师成长计划经费资助项目

作者简介: 宋洁洁(1993—), 女, 在读硕士, 主要从事植物资源学和分子生物学研究。E-mail: sjj\_green@163.com

\*通信作者: 高婷(1982—), 女, 博士, 副教授, 主要从事植物资源学和分子生物学研究。E-mail: gt\_kelly@163.com

*Apium graveolens*, *Ammi majus*. Higher expression of GIPS1 gene was observed in roots than leaves. However, GIPS2 gene was expressed at a relatively higher level in flowers than in roots. **Conclusion** The full-length cDNA of GIPS1 and GIPS2 genes were obtained and the expression patterns were explored in *G. littoralis* for the first time, which provided a foundation for further studies on gene function and genetic regulatory mechanism of GIPS.

**Key words:** *Glehnia littoralis* Fr. Schmidt ex Miq.; psoralen synthase; RACE; bioinformatics analysis; tissue specific expression

珊瑚菜 *Glehnia littoralis* Fr. Schmidt ex Miq. 为伞形科珊瑚菜属多年生药用草本植物，渐危种，国家二级重点保护植物，原野生于沿海地区，现多为家种栽培。北沙参 *Glehniae Radix* 为珊瑚菜的干燥根<sup>[1]</sup>，是我国传统中药材，也是临床常用的滋阴药，其味甘、微苦，性微寒，具有养阴清肺、益胃生津的功效，用于肺热燥咳、劳嗽痰血、胃阴不足等<sup>[2]</sup>。呋喃香豆素是珊瑚菜主要的活性成分，补骨脂素是一种重要的呋喃香豆素，具有较强的光敏化和镇静、解痉、止血等作用<sup>[3-4]</sup>，是治疗白癜风、斑秃、牛皮癣以及瘤样皮肤病的有效药物。高浓度补骨脂素对不同类型白血病细胞株（HL-60、Raji、K562）有显著杀伤作用。近来研究表明，补骨脂素对多种动物及人的肿瘤生长有明显抑制作用<sup>[5-6]</sup>。补骨脂素是衡量珊瑚菜内在质量优劣的重要标准，增加珊瑚菜中补骨脂素的含量，可有效提升珊瑚菜的药用价值，具有重要的理论和实践意义。

补骨脂素来源于植物苯丙氨酸代谢途径，苯丙氨酸在苯丙氨酸解氨酶作用下生成肉桂酸，然后经过羟化酶、连接酶和内酯化作用生成伞形花内酯，在印度枸橘素合成酶催化下生成异紫花前胡内酯，最终异紫花前胡内酯在补骨脂素合成酶作用下生成补骨脂素<sup>[7-10]</sup>。补骨脂素合成酶（psoralen synthase, PS）是其生物合成途径的关键酶，但目前补骨脂素合成酶的相关研究很少，只在大阿米芹 *Ammi majus* L.、欧防风 *Pastinaca sativa* L. 和旱芹 *Apium graveolens* L. 中获得了该基因的克隆及部分功能分析<sup>[7, 11]</sup>。

本研究在前期已完成的珊瑚菜转录组高通量测序数据（GenBank NO. PRJNA387325）的基础上，对筛选获得的表达量较高的 2 条 GIPS 基因 GIPS1 和 GIPS2 进行分析，利用 cDNA 末端快速扩增（RACE）技术获得 2 条基因的全长 cDNA 序列，并初步进行生物信息学和表达模式分析，为揭示二者在珊瑚菜补骨脂素合成途径中的功能奠定基础。

## 1 材料与试剂

### 1.1 材料

在山东省莱阳市高格庄镇采集新鲜的 4 年生珊

瑚菜，由青岛农业大学辛华教授鉴定为珊瑚菜 *Glehnia littoralis* Fr. Schmidt ex Miq.。清洁干净后将根、叶和花放入液氮速冻，-80 ℃保存。

### 1.2 试剂

SMARTer® RACE5'3'Kit 试剂盒、TaKaRa Ex Taq 酶、SYBR® Premix Ex Taq™ II (Tli RNaseH Plus) 定量试剂盒、凝胶回收试剂盒、DL5000 DNA Marker、质粒载体 pMD18-T 购自宝生物工程(大连)有限公司，大肠杆菌 *DH5α* 感受态购自天根生化科技(北京)有限公司。

## 2 方法

### 2.1 总 RNA 的提取

取-80 ℃保存的珊瑚菜的根、叶和花，用 Trizol 法提取总 RNA。经核酸测定仪和琼脂糖凝胶电泳检测浓度和纯度后，-80 ℃保存备用。

### 2.2 RACE 基因克隆

根据转录组的高通量测序获得 GIPS1 和 GIPS2 基因的部分序列(已有 5'末端)，通过 Primer Premier 5.0 软件设计 3' RACE 特异性引物 GIPS1-3'RACE 和 GIPS2-3'RACE (表 1)，由生工生物工程(上海)股份有限公司合成。参照 SMARTer® RACE 5'3' Kit User Manual 方法进行反转录，生成第 1 链 cDNA。反应体系参照 SMARTer® RACE 5'3' Kit User Manual，利用 Touchdown PCR 程序扩增 3'末端序列：第 1 步，94 ℃，30 s；72 ℃，1 min；5 个循环；第 2 步，94 ℃，30 s；70 ℃，30 s；72 ℃，1 min；5 个循环；第 3 步，94 ℃，30 s；68 ℃，30 s；72 ℃，1 min；25 个循环。

PCR 产物经 1% 琼脂糖凝胶电泳检测后，用琼脂糖凝胶 DNA 回收试剂盒回收目的条带。回收产物与 pMD18-T 载体连接，连接产物转化 *DH5α* 感受态细胞，在含有氨苄的 LB 固体培养基平板上进行培养，经菌落 PCR 检验后，选取 3 个阳性克隆送至生工生物工程(上海)股份有限公司测序。

### 2.3 基因 ORF 的克隆

使用 DNAMAN 软件将转录组中已知序列片段和测序得到的序列片段进行拼接，得到 GIPS1

**表1 GIPS1 和 GIPS2 基因扩增及 RT-qPCR 所用的引物序列**  
**Table 1 Primers used for amplification and qPCR of GIPS1 and GIPS2 genes**

引物	序列(5'→3')
GIPS1-3'RACE	TGACTGGTCCACTGGCATTACTCTTC
GIPS2-3'RACE	AAGAGTTGGACATGGATGCAGAGACGGG
GIPS1-F	ATGAAGATGCAGGAACAAATCCCTG
GIPS1-R	TCAAAACATGTGGTGTGGCAACCACC
GIPS2-F	ATGATGGACCAACAATCCTTGTTC
GIPS2-R	TCAGACACGTGATGTTGCAATCAC
GIPS1-qF	TCAAATTGGTCCTGACCCGC
GIPS1-qR	GGGTTTCACAACCTCCCGA
GIPS2-qF	GTTAACCAAGACCCGCACAA
GIPS2-qR	GCACCAGCTGAAGAGACAACA
GlActin-F	TGGTATTGTGCTGGATTCTGGT
GlActin-R	TGAGATCACGACCAGCAAGG

和 GIPS2 基因的全长 cDNA 序列。根据基因编码区序列设计特异引物 GIPS1-F/GIPS1-R 和 GIPS2-F/GIPS2-R (表 1)，以珊瑚菜叶 cDNA 为模板，参照 TaKaRa Ex Taq 酶说明书配制体系，进行 PCR 程序：94 °C 预变性 5 min；94 °C、30 s，55 °C、30 s，72 °C、2 min，28 个循环；72 °C、10 min，4 °C 终止反应。将目的条带切胶回收后克隆至 pMD18-T 载体并进行测序，分析测序结果与所拼接序列的一致性。

#### 2.4 生物信息学分析

使用 NCBI 的 ORF Finder 和 Conserved Domains 在线网站分别进行开放阅读框 (ORF) 和保守结构域分析；使用 ExPASy 的 Protscale 及 TMHMM Server 在线网站分别分析蛋白亲疏水性并预测跨膜区；使用 SOMPA 在线网站进行蛋白二级结构分析；使用 SWISS-MODEL 在线网站构建三维模型并用 PyMOL 软件对其进行分析；最后使用 MEGA5.0 软件构建 NJ 系统进化树。

#### 2.5 基因的组织特异性表达

利用 Primer Premier 5.0 软件设计 GlIPS1 和 GlIPS2 基因的 RT-qPCR 引物 GIPS1-qF/GIPS1-qR 和 GIPS2-qF/GIPS2-qR (表 1)，扩增片段长度分别为 137 bp 和 110 bp。以珊瑚菜 Actin 作为内参基因 (GlActin-F/GlActin-R)。按照试剂盒 SYBR® Premix Ex Taq™ II (Tli RNaseH Plus) 说明书构建体系，以珊瑚菜根、叶和花总 RNA 反转录出的 cDNA 为模板，每个反应做 3 个生物学重复，采用 3 步法程

序进行 RT-qPCR：第 1 步，95 °C，10 min；第 2 步，95 °C，30 s；55 °C，30 s；72 °C，20 s；40 个循环；第 3 步，95 °C，1 min；55 °C，30 s；95 °C，30 s。

#### 3 结果与分析

##### 3.1 GIPS1 和 GIPS2 基因克隆及序列分析

利用已知 GIPS 序列设计引物 3'GSP，通过 RACE 技术，克隆出 GIPS1 和 GIPS2 基因 3'末端序列分别为 261 bp 和 410 bp (图 1)。将测序获得的序列与转录组已知序列进行拼接，获得 GIPS1 和 GIPS2 基因全长 cDNA 序列为 1 885 bp 和 1 971 bp。

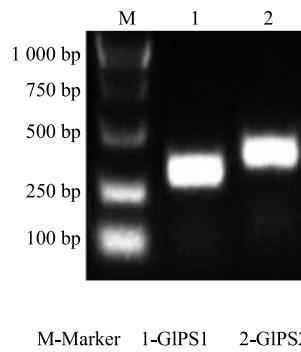


图 1 GIPS1 和 GIPS2 3'RACE PCR 扩增结果

Fig. 1 PCR amplification of 3'RACE of GIPS1 and GIPS2

GIPS1 基因 cDNA 序全长 1 885 bp，编码 495 个氨基酸，其中 ORF1 488 bp，5'非编码区 232 bp，3'非编码区 165 bp，包括 28 bp poly A 尾 (图 2)。GIPS2 基因 cDNA 序全长 1 971 bp，编码 502 个氨基酸，其中开放阅读框 1 509 bp，5'非编码区 160 bp，3'非编码区 302 bp，包括 29 bp poly A 尾 (图 2)。根据 GIPS1 和 GIPS2 基因编码区序列分别设计特异引物，以珊瑚菜叶 cDNA 为模板扩增出 1 488 bp 和 1 509 bp 的条带 (图 3)，2 条条带测序结果与所拼接得到的序列一致，命名基因为 GIPS1 和 GIPS2 (GenBank 登录号为 MF632070 和 MF632071)。

#### 3.2 生物信息学分析

**3.2.1 同源性分析及结构域预测** 将 GIPS1 和 GIPS2 蛋白序列与 NCBI 上植物的同源序列进行比对。结果表明，GIPS1 与大阿米芹 *Ammi majus* L. (Q6QNI4)、欧防风 *Pastinaca sativa* L. (C0SJS2)、旱芹 *Apium graveolens* L. (C0SJS4) 的蛋白序列相似性分别为 85%、86%、84%；GIPS2 与毒胡萝卜素属的 *Thapsia Garganica* L. (AKJ23348)、独活属的 *Heracleum mantegazzianum* Somm. et Lev. (AKJ23344)、大

A

1 GTGGCGGTTGGTGGCAAGCCGAGACGTACGATTCAGATTCTATATATAAACTCTGTAACGATCAGGGTTGG  
2 GCGGTTTTGTCAGTTGATCAGCCGACAGTCGAGGTTCTGGCTTGATCAGGGCGGACGAA  
3 1 AACCATAAGTTGGCGCACCTACTCTGGTATATATATATGTTAGCTTCTTGGACGAA  
26 ATAGCA  
33 ATAGAATAGCAGGAAACAAAATCCCTGTACCTATATTTTCATIGTCTTCCTGCACAAATTTCCTACAAA  
1 M K M Q E Q N P L Y L F F S L S F V T I F L Y K  
308 TGTTAAACATCATAAAAACACCATCAAAAAATCTGCACCTTCCGGCACAATTCTACATAAGGAAACCTT  
1 W L C H T K K T P S K N L P S P Q P I F I G N L  
383 CATACAAAGTTGGCTGCGGCCAGGGTCTTACTCTGGAGACTGCTTAAAGATGTGCTAACATGCTACTAAAG  
1 M H Q I G P D P Q V S S L R D L A K K Y G S L M  
485 TTGGAGGTGACCTGTTCTGGCTTCTTCAGCTGATGCGCTCGGGAGGTTGTGAAACCCATGATCTGGT  
76 F G S V V V V S S A D A A C C T C T C T C H T D L V  
333 TTGCAGACGGACCTTAACTAAGITGTCGCAATAGATCTTACAGGAAGGGACATGGTGTGCTCTGTT  
101 F F A D R P P Y S S V V A N R I F Y N G R D M V F A Y  
608 ACTGAGATTAGGAGACAGAAGAACATGTTGCTACGCTCTAAAGTAAACAGGGGTTCTTCGTTCTAC  
126 T E Y W Q R Q V K S T C V T Q L L S N K R V L S F H  
683 AAATGTCAGAGAGAAAGAAGATGTCCTTAACTCCTAAAGGAAATTCGTCIAAAAGTAGTAATTAACTG  
151 N V R E E E V A L L I Q K I E N S R S K V V N L S  
758 GAACATTAACTGCAATCAGGCCAATGAGTCGAGGGCTGCACTAGAAGATGTCAGTAAAGGGTACATAC  
176 E H L I E L T G N V C A C A G G S H E V D S Y  
833 AAGTGTGACTTTGGAAATAATGGATGTTAGGCTTACCGCAGCATAGGAGTTTTTCCAACTGGT  
201 K L L L L E I M D L M G Y S R S E I F D F M  
908 TGGTTGTGATGGCTACTGGATTAAGGGGAAAGGTGAGAAAGCAGCCAAAGGGCTGATCTTCTGGAAAGGT  
226 W V D W T L G K G V E K A A K G V D F T L E G  
983 GTCTTACAGAGATCATCTGCTACATGATAGTGTCTGCACCAAAGAACATTGTTATCTTGTGAGTCA  
251 V L R D H P A P S N G S A T K C D F V S I L L E I Q  
506 GAAGCGGGATGTCGCTCTTATGATAAGAAGATGTCATACATTCTATCTGGGATATGTTGGGTGCTGGAACT  
276 E A D A G F S I D K E C I K S L I W D M L G A G T  
133 GAACAAATCATGACAGCTTGGAGTGGACATAGCAGGCCCTAAAGAACCTGAGTCATGCTCAAAATGCG  
301 E T I S T A L E W T I A A L K S P D A M L K Q L  
208 ATAGGATGTCAGAGAAATAGCGAACAGGAACATTCAGATACAGGGTGTAGCTGATCAGTACATCTGGAA  
983 GCAGTAAAGGAAGAGATGTCGATCTTACAGGCCACTTCTAGTCTCTAGAGAAACAGGCAGACCTGG  
351 A V A M K E S C M R L Y F T A P L L V P R E T R Q D V  
358 AAAATCATGGGGTGTGACATAAAAGGAGGAAACAGGACTGATAAAATGCTGGCAATTCGAGAACAGCTTC  
376 K F M G M O D Y D I K A G T Q V L I N A W A I J R D P S  
433 TTGGGGACACCAAGGAGGATCTGGCGCTGAGGAGTTCTGAGATGTTAGCTGATCTAACAAAGGCTTCTACAT  
401 L W D N M P E F R P E R F L L N S R I D Y K G F H Y  
505 GAGTATCTTCCCCTTGGAGCGCTGGAGGGGGTCTCCCTGGATCAGGTGTCATAGCTGTTAATGAGCTGCA  
426 E Y L P F G A G R R G C P G I Q F A I A V N E L A  
583 GTGCAATGATGTCGCAAGATGTTGAGCTGGCCACTTCTAGTCTCTAGAGAAACAGGCAGACCTGG  
451 V A N V V H K F A T T F E L L R G R D L E D L D M T G  
658 TCCACTGGATCTACTCTCTACATAAAATCTCTCTTGTGTTGGCTGACACCATGTTG1720  
476 S T G I T L H K K S P L V V L A T P T H \*  
721 TCTCTTATGTCATTTGAGTGTCTTGTGATGTCGCAACTAAAGACTGCTTGGCTACCTCACAACACTGT  
796 TGGTACTTACCAACTCTAGCGCTTAAATAAAATGTTAATAAGGGCACCTCTAGCCCCAAAAAA  
871 AAAAaaaaaaaaaaa 188

B

1 CATTATATGGGTACAGGTTCTAGCTGTAAGCTCTGATGAGTGTACGATTCTATGACTCTTAAAGATGTTGTT  
2 TTCAACCGAACAAAATTTATGTCAGGAGACACATAGTTGACGGGTTAGATTATGTCGAACTTAAAGTGT  
3 TUAACCMG  
4 ATGATGACCAACAACTCTGTTTCTATCTTTAAGTCTTCTGTTGCTCTGTTGTTGCTACATGGTG  
5 M D Q Q S L S L S L F V L F Y I W  
6 151 161 171 181 191 201 211 221 231 241 251 261 271 281 291 301 311 321 331 341 351 361 371 381 391 401 411 421 431 441 451 461 471 481 491 501 511 521 531 541 551 561 571 581 591 601 611 621 631 641 651 661 671 681 691 701 711 721 731 741 751 761 771 781 791 801 811 821 831 841 851 861 871 881 891 901 911 921 931 941 951 961 971 981 991 1001 1011 1021 1031 1041 1051 1061 1071 1081 1091 1101 1111 1121 1131 1141 1151 1161 1171 1181 1191 1201 1211 1221 1231 1241 1251 1261 1271 1281 1291 1301 1311 1321 1331 1341 1351 1361 1371 1381 1391 1401 1411 1421 1431 1441 1451 1461 1471 1481 1491 1501 1511 1521 1531 1541 1551 1561 1571 1581 1591 1592 1601 1611 1621 1631 1641 1651 1661 1671 1681 1691 1701 1711 1721 1731 1741 1751 1761 1771 1781 1791 1801 1811 1821 1831 1841 1851 1861 1871 1881 1891 1901 1911 1921 1931 1941 1951 1961 1971 1981 1991 2001 2011 2021 2031 2041 2051 2061 2071 2081 2091 2101 2111 2121 2131 2141 2151 2161 2171 2181 2191 2201 2211 2221 2231 2241 2251 2261 2271 2281 2291 2301 2311 2321 2331 2341 2351 2361 2371 2381 2391 2392 2401 2411 2421 2431 2441 2451 2461 2471 2481 2491 2492 2501 2511 2521 2531 2541 2551 2561 2571 2581 2591 2592 2601 2611 2621 2631 2641 2651 2661 2671 2681 2691 2692 2701 2711 2721 2731 2741 2751 2761 2771 2781 2791 2792 2801 2811 2821 2831 2841 2851 2861 2871 2881 2891 2892 2901 2911 2921 2931 2941 2951 2961 2971 2981 2991 2992 3001 3011 3021 3031 3041 3051 3061 3071 3081 3091 3092 3101 3111 3121 3131 3141 3151 3161 3171 3181 3191 3192 3201 3211 3221 3231 3241 3251 3261 3271 3281 3291 3292 3301 3311 3321 3331 3341 3351 3361 3371 3381 3391 3392 3401 3411 3421 3431 3441 3451 3461 3471 3481 3491 3492 3501 3511 3521 3531 3541 3551 3561 3571 3581 3591 3592 3601 3611 3621 3631 3641 3651 3661 3671 3681 3691 3692 3701 3711 3721 3731 3741 3751 3761 3771 3781 3791 3792 3801 3811 3821 3831 3841 3851 3861 3871 3881 3891 3892 3901 3911 3921 3931 3941 3951 3961 3971 3981 3991 3992 4001 4011 4021 4031 4041 4051 4061 4071 4081 4091 4092 4101 4111 4121 4131 4141 4151 4161 4171 4181 4191 4192 4201 4211 4221 4231 4241 4251 4261 4271 4281 4291 4292 4301 4311 4321 4331 4341 4351 4361 4371 4381 4391 4392 4401 4411 4421 4431 4441 4451 4461 4471 4481 4491 4492 4501 4511 4521 4531 4541 4551 4561 4571 4581 4591 4592 4601 4611 4621 4631 4641 4651 4661 4671 4681 4691 4692 4701 4711 4721 4731 4741 4751 4761 4771 4781 4791 4792 4801 4811 4821 4831 4841 4851 4861 4871 4881 4891 4892 4901 4911 4921 4931 4941 4951 4961 4971 4981 4991 4992 5001 5011 5021 5031 5041 5051 5061 5071 5081 5091 5092 5101 5111 5121 5131 5141 5151 5161 5171 5181 5191 5192 5201 5211 5221 5231 5241 5251 5261 5271 5281 5291 5292 5301 5311 5321 5331 5341 5351 5361 5371 5381 5391 5392 5401 5411 5421 5431 5441 5451 5461 5471 5481 5491 5492 5501 5511 5521 5531 5541 5551 5561 5571 5581 5591 5592 5601 5611 5621 5631 5641 5651 5661 5671 5681 5691 5692 5701 5711 5721 5731 5741 5751 5761 5771 5781 5791 5792 5801 5811 5821 5831 5841 5851 5861 5871 5881 5891 5892 5901 5911 5921 5931 5941 5951 5961 5971 5981 5991 5992 6001 6011 6021 6031 6041 6051 6061 6071 6081 6091 6092 6101 6111 6121 6131 6141 6151 6161 6171 6181 6191 6192 6201 6211 6221 6231 6241 6251 6261 6271 6281 6291 6292 6301 6311 6321 6331 6341 6351 6361 6371 6381 6391 6392 6401 6411 6421 6431 6441 6451 6461 6471 6481 6491 6492 6501 6511 6521 6531 6541 6551 6561 6571 6581 6591 6592 6601 6611 6621 6631 6641 6651 6661 6671 6681 6691 6692 6701 6711 6721 6731 6741 6751 6761 6771 6781 6791 6792 6801 6811 6821 6831 6841 6851 6861 6871 6881 6891 6892 6901 6911 6921 6931 6941 6951 6961 6971 6981 6991 6992 7001 7011 7021 7031 7041 7051 7061 7071 7081 7091 7092 7101 7111 7121 7131 7141 7151 7161 7171 7181 7191 7192 7201 7211 7221 7231 7241 7251 7261 7271 7281 7291 7292 7301 7311 7321 7331 7341 7351 7361 7371 7381 7391 7392 7401 7411 7421 7431 7441 7451 7461 7471 7481 7491 7492 7501 7511 7521 7531 7541 7551 7561 7571 7581 7591 7592 7601 7611 7621 7631 7641 7651 7661 7671 7681 7691 7692 7701 7711 7721 7731 7741 7751 7761 7771 7781 7791 7792 7801 7811 7821 7831 7841 7851 7861 7871 7881 7891 7892 7901 7911 7921 7931 7941 7951 7961 7971 7981 7991 7992 8001 8011 8021 8031 8041 8051 8061 8071 8081 8091 8092 8101 8111 8121 8131 8141 8151 8161 8171 8181 8191 8192 8201 8211 8221 8231 8241 8251 8261 8271 8281 8291 8292 8301 8311 8321 8331 8341 8351 8361 8371 8381 8391 8392 8401 8411 8421 8431 8441 8451 8461 8471 8481 8491 8492 8501 8511 8521 8531 8541 8551 8561 8571 8581 8591 8592 8601 8611 8621 8631 8641 8651 8661 8671 8681 8691 8692 8701 8711 8721 8731 8741 8751 8761 8771 8781 8791 8792 8801 8811 8821 8831 8841 8851 8861 8871 8881 8891 8892 8901 8911 8921 8931 8941 8951 8961 8971 8981 8991 8992 9001 9011 9021 9031 9041 9051 9061 9071 9081 9091 9092 9101 9111 9121 9131 9141 9151 9161 9171 9181 9191 9192 9201 9211 9221 9231 9241 9251 9261 9271 9281 9291 9292 9301 9311 9321 9331 9341 9351 9361 9371 9381 9391 9392 9401 9411 9421 9431 9441 9451 9461 9471 9481 9491 9492 9501 9511 9521 9531 9541 9551 9561 9571 9581 9591 9592 9601 9611 9621 9631 9641 9651 9661 9671 9681 9691 9692 9701 9711 9721 9731 9741 9751 9761 9771 9781 9791 9792 9801 9811 9821 9831 9841 9851 9861 9871 9881 9891 9892 9901 9911 9921 9931 9941 9951 9961 9971 9981 9991 9992 10001

图 2 GIPS1 (A) 和 GIPS2 (B) 基因全长 cDNA 序列

**Fig. 2** Full-length cDNA sequences of GlPS1 (A) and GlPS2 (B) genes

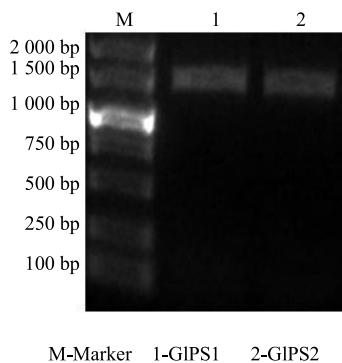


图 3 GIPS1 和 GIPS2 基因 ORF 的 PCR 扩增

**Fig. 3** PCR amplification products of ORF of GIPS1 and GIPS2 genes

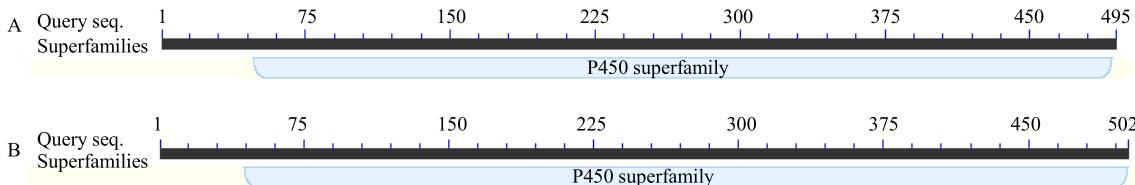


图 4 GIPS1 (A) 和 GIPS2 (B) 蛋白的保守结构域预测

**Fig. 4** Prediction of conserved domains of GIPS1 (A) and GIPS2 (B) proteins

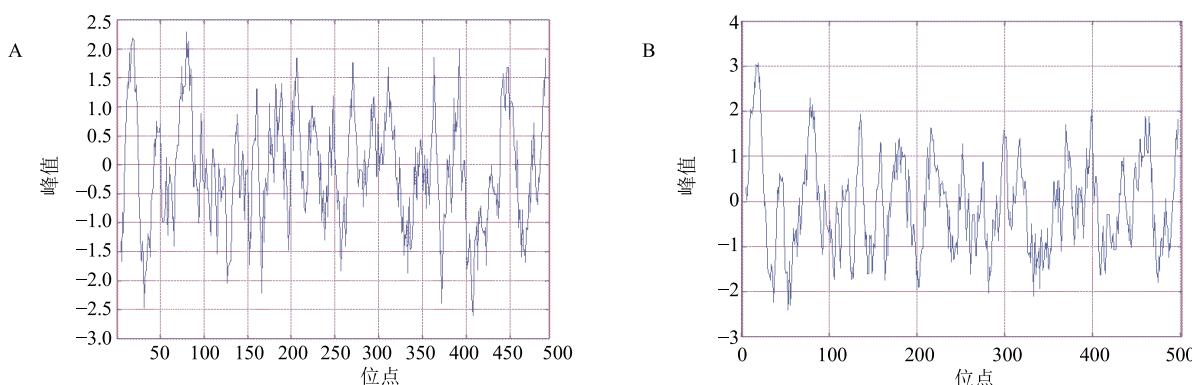


图 5 GIPS1 (A) 和 GIPS2 (B) 蛋白亲疏水性预测  
Fig. 5 Hydrophobicity analysis on GIPS1 (A) and GIPS2 (B) proteins

白在第 53 个氨基酸位点, 最低峰值是 -2.411, 在第 19 个氨基酸位点, 最高峰值是 3.067, 总平均亲水性 -0.089。2 个蛋白均具有一个明显的跨膜区, 为一次跨膜亲水性蛋白。

**3.2.3 蛋白二维、三维结构预测** 利用在线软件 SOPMA 预测 GIPS1 和 GIPS2 的二级结构(图 6)。GIPS1 二级结构中  $\alpha$ -螺旋(蓝色)占 43.03%、延伸链(红色)占 20.20%、 $\beta$ -折叠(绿色)占 6.87%、无规卷曲(紫色)占 29.90%。GIPS2 二级结构中  $\alpha$ -螺旋(蓝色)占 46.41%、延伸链(红色)占 17.73%、 $\beta$ -折叠(绿色)占 8.17%、无规卷曲(紫

色)占 27.69%。GIPS1 和 GIPS2 中主要结构均为  $\alpha$ -螺旋和无规卷曲,  $\alpha$ -螺旋是最稳定的蛋白质空间结构, 二者中  $\alpha$ -螺旋的比例均在 40%以上,  $\alpha$ -螺旋是最稳定的蛋白质空间结构, 因此推测两者蛋白相对稳定。

通过在线软件 SWISS-MODEL 分别对 GIPS1 和 GIPS2 蛋白进行三维模型构建(图 7)。GIPS1、GIPS2 均以 5tfu.4.A 蛋白作为模体, 相似性分别为 26.28% 和 25.88%, 三维结构均包括 8 个  $\beta$ -折叠, GIPS1 具有 18 个  $\alpha$ -螺旋, GIPS2 包括 19 个  $\alpha$ -螺旋。由图 7 可以看出 2 个蛋白三维结构较为相似但又有所区别。

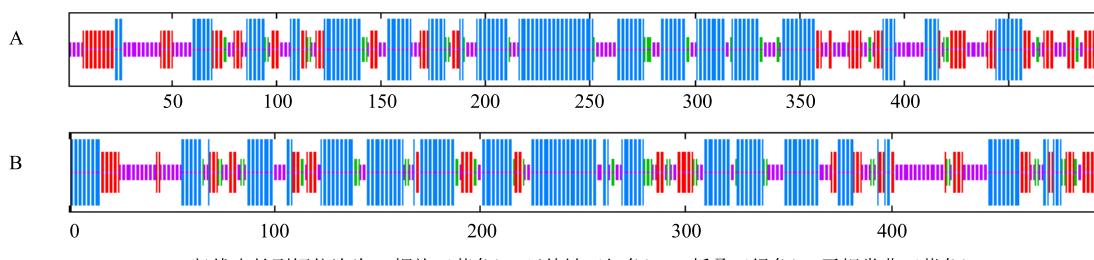


图 6 GIPS1 (A) 和 GIPS2 (B) 蛋白的二级结构预测  
Fig. 6 Secondary structure prediction of GIPS1 (A) and GIPS2 (B) proteins

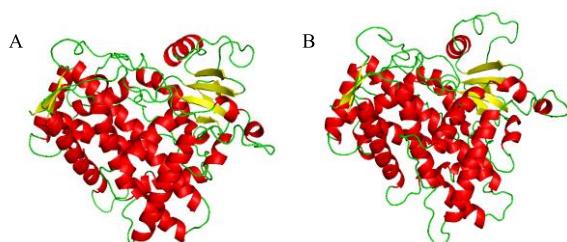


图 7 GIPS1 (A) 和 GIPS2 (B) 蛋白三维结构预测  
Fig. 7 Tertiary structure prediction of GIPS1 (A) and GIPS2 (B) proteins

**3.2.4 系统进化分析** 从 NCBI 数据库选取已发表植物的 PS 蛋白序列, 利用 MEGA5.0 软件, 利用 NJ 法构建系统进化树(图 8)。结果表明, GIPS1 和 GIPS2 与伞形科植物欧防风 *Pastinaca sativa* L.、旱芹 *Apium graveolens* L.、大阿米芹 *Ammi majus* L. 的 PS 蛋白亲缘关系较近, 而二者与川桑 *Morus notabilis* Schneid.、紫花风铃木 *Handroanthus impetiginosus* (Mart. ex DC.) Mattos、浆果状辣椒 *Capsicum baccatum* L. 的 PS 蛋白亲缘关系较远。

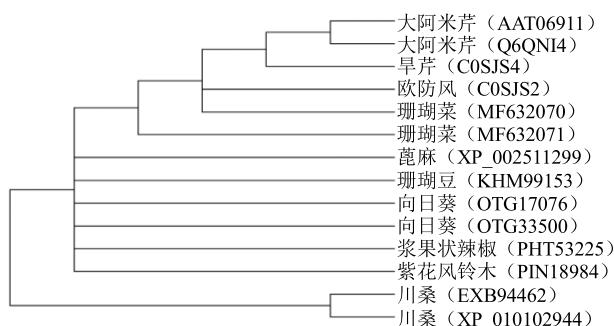


图 8 PS 蛋白的 NJ 系统进化树

Fig. 8 Neighbor-joining phylogenetic tree of PS proteins

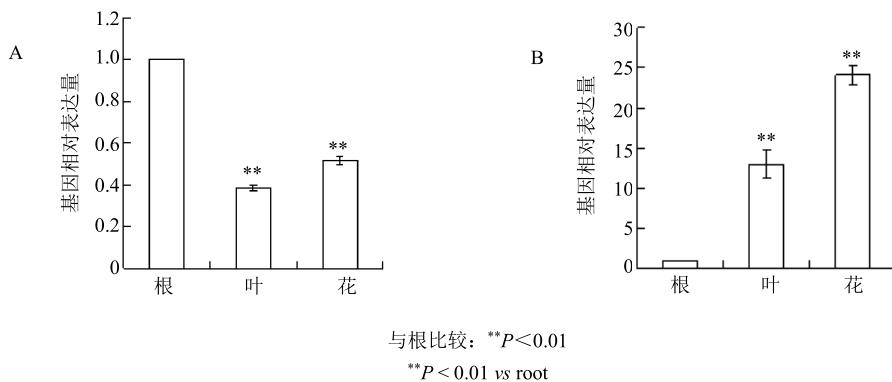


图 9 GIPS1 (A) 和 GIPS2 (B) 基因的组织特异性表达

Fig. 9 Tissue specific expression of GIPS1 (A) and GIPS2 (B) genes

分别命名为 GIPS1 和 GIPS2。同源分析结果显示这 2 个基因与其他植物 PS 基因的同源性最高分别为 90% 和 94%。补骨脂素合成酶属于细胞色素 P450 超家族，GIPS1 和 GIPS2 蛋白具有 P450 超家族保守结构域，因此推测 2 个基因都属于 PS 基因家族的成员。

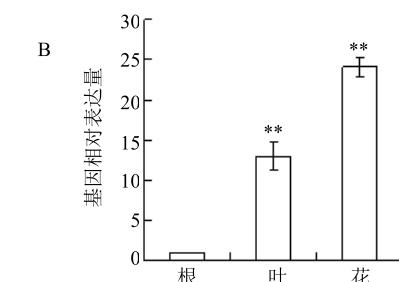
细胞色素 P450 为一类亚铁血红素-硫醇盐蛋白（即 b 型血红蛋白）的超家族，它参与内源性物质和包括药物、环境化合物在内的外源性物质的代谢，催化许多内源性底物和外源性化学物质的单加氧化反应<sup>[12]</sup>，在动物、植物、微生物中均有存在<sup>[13]</sup>。在苯丙烷代谢途径中，CYP71、CYP73、CYP75、CYP81、CYP84、CYP93、CYP98 等家族在黄酮、类黄酮、木脂素中间体、呋喃香豆素合成中发挥重要作用，主要参与脂肪族、芳香族碳羟基化反应以及芳基转移反应、亚甲基二氧桥生成等<sup>[14]</sup>。补骨脂素合成酶属于 CYP71 家族，基因家族各成员表达可能具有组织器官特异性，从而担负不同的功能。本实验中 GIPS1 和 GIPS2 基因表达结果显示，GIPS1 基因在根中表达量较高，在叶中表达量较低；而

### 3.3 GIPS1 和 GIPS2 基因的组织特异性表达

利用 qPCR 技术和  $2^{-\Delta\Delta Ct}$  计算方法分析 GIPS1 和 GIPS2 基因在珊瑚菜根、叶、花中的相对表达量。结果表明，GIPS1 和 GIPS2 基因的转录本在根、叶、花均能检测到，但在不同组织中的表达量存在极显著差异（*t* 体验）。GIPS1 基因在根中表达量最高，其次为花，在叶中表达量最低；而 GIPS2 基因在花中表达量最高，其次为叶，在根中表达量最低（图 9）。

### 4 讨论

本研究通过 RACE 技术克隆得到珊瑚菜 2 条 PS 基因的 cDNA 序列全长，对其编码序列进行验证后，



GIPS2 基因在花中表达量较高，在根中表达量较低，2 条基因在根中的表达情况相反。刘玉函等<sup>[15]</sup>研究发现，补骨脂素在珊瑚菜的不同组织中含量存在显著差异，且在根中补骨脂素含量显著高于叶中。由此推测二者的基因功能有所区别，GIPS1 可能在补骨脂素合成中起主要作用，有待进一步实验验证。

本研究克隆获得了珊瑚菜 GIPS1 和 GIPS2 基因的全长 cDNA 序列并分析了其表达模式，为阐明其在珊瑚菜补骨脂素生物合成途径中的功能奠定了研究基础，同时也对未来通过生物技术手段调控珊瑚菜补骨脂素含量提供可能。本研究为今后利用基因工程方法提高 GIPS1 和 GIPS2 基因在珊瑚菜中各器官的表达量，最终获得高产植株做好准备。预期可以通过改良单个或多个关键酶促步骤，实现人工调控珊瑚菜补骨脂素在植物细胞中的产量，从分子水平提高补骨脂素的生产率是北沙参产业发展的新途径。

### 参考文献

- [1] 王爱兰, 王贵琳, 李维卫. 濒危物种珊瑚菜遗传多样性 的 ISSR 分析 [J]. 西北植物学报, 2015, 35(8): 1541-1546.

- [2] 中国药典 [S]. 一部. 2015.
- [3] 朱兆荣, 李玉强, 刘娟, 等. 补骨脂总黄酮抗氧化作用研究 [J]. 中国兽药杂志, 2005, 39(2): 18-20.
- [4] Herrmann G, Brenneisen P, Wlaschek M, et al. Psoralen photoactivation promotes morphological and functional changes in fibroblasts *in vitro* reminiscent of cellular senescence [J]. *J Cell Sci*, 1998, 111(6): 759-767.
- [5] Fisher G J, Wang Z Q, Datfa S C, et al. Pathophysiology of premature skin aging induced by ultraviolet light [J]. *New England J Med*, 1997, 337(20): 1419-1429.
- [6] 李祝, 宋宝安. 补骨脂素抗肿瘤机理及其应用研究进展 [J]. 山地农业生物学报, 2007, 26(3): 255-260.
- [7] Larbat R, Hehn A, Hans J, et al. Isolation and functional characterization of CYP71AJ4 encoding for the first P450 monooxygenase of angular furanocoumarin biosynthesis [J]. *J Biol Chem*, 2009, 284(8): 4776-4785.
- [8] Bourgaud F, Hehn A, Larbat R, et al. Biosynthesis of coumarins in plants: A major pathway still to be unravelled for cytochrome P450 enzymes [J]. *Phytochem Rev*, 2006, 5(2): 293-308.
- [9] Karamat F, Olry A, Munakata R, et al. A coumarin-specific prenyltransferase catalyzes the crucial biosynthetic reaction for furanocoumarin formation in parsley [J]. *Plant J Cell Mol Biol*, 2014, 77(4): 627-638.
- [10] Munakata R, Olry A, Karamat F, et al. Molecular evolution of parsnip (*Pastinaca sativa*) membrane-bound prenyltransferases for linear and/or angular furanocoumarin biosynthesis [J]. *New Phytol*, 2016, 211(1): 332-344.
- [11] Larbat R, Kellner S, Specker S, et al. Molecular cloning and functional characterization of psoralen synthase, the first committed monooxygenase of furanocoumarin biosynthesis [J]. *J Biol Chem*, 2007, 282(1): 542-554.
- [12] 李泽莉. *Bacillus sp.C3 细胞色素 P450: CYP102A16 基因的克隆、优化表达及其酶性质研究* [D]. 杭州: 浙江大学, 2012.
- [13] 许华夏, 李培军, 刘宛, 等. 生物细胞色素 P450 的研究进展 [J]. 农业环境科学学报, 2002, 21(2): 188-191.
- [14] 解敏敏, 龚达平, 李凤霞, 等. 烟草细胞色素 P450 的基因组学分析 [J]. 遗传, 2013, 35(3): 379-387.
- [15] 刘玉函, 刘汉柱, 辛华. 珊瑚菜不同部位香豆素含量的研究 [J]. 植物科学学报, 2010, 28(1): 114-117.