

## 20种中药基原植物鲨烯合酶基因的生物信息学分析

王琴波, 王雨晴, 杨 谨, 陈观水\*

福建农林大学生命科学学院, 福建 福州 350002

**摘要:**目的 鲨烯合酶(squalene synthase, SQS)是三萜化合物合成途径中的关键限速酶,对其蛋白特性进行分析与预测。方法 利用生物信息学相关分析工具对20种《中国药典》2015年版收录中药基原植物SQS的理化性质、结构特征及功能特点进行分析。结果 20种中药基原植物有关各条SQS氨基酸序列,其在氨基酸长度、氨基酸组成、相对分子质量及平均亲水系数方面基本一致;所有的SQS均未发现信号肽;二级结构具有明显的一致性,并且 $\alpha$ -螺旋和无规则卷曲是主要的结构元件;所有SQS蛋白均含有4个高度保守的结构域和在氨基酸的C端有1~2个跨膜结构;系统进化树结果与植物系统分类学结果基本一致。结论 这20种中药基原植物的SQS的特性差别不大,进化距离也比较近,显示SQS具有较高的保守性和稳定性。分析结果为SQS的功能和作用机制研究提供依据。

**关键词:** 基原植物; 三萜类; 鲨烯合酶; 生物信息学; 结构域

中图分类号: R282.12 文献标志码: A 文章编号: 0253-2670(2021)03-0838-07

DOI: 10.7501/j.issn.0253-2670.2021.03.027

## Bioinformatics analysis of squalene synthase gene from 20 Chinese herbal plants

WANG Qin-bo, WANG Yu-qing, YANG Jin, CHEN Guan-shui

College of Life Science, Fujian Agriculture and Forestry University, Fuzhou 350002, China

**Abstract: Objective** Squalene synthase (SQS) is a key rate limiting enzyme in the triterpenoid biosynthesis pathway. The present study aimed to investigate the protein characteristics and phylogenetics of SQS. **Methods** The physicochemical properties, structural characteristics and functional characteristics of squalene synthases from 20 Chinese herbal plants included in the 2015 edition of Chinese Pharmacopoeia were analyzed by bioinformatics tools. **Results** The SQS amino acid sequences from 20 Chinese herbal plants showed strong consistency in amino acid numbers, relative molecular weight, amino acid compositions and grand average of hydropathicity. Signal peptides were not detected in the all SQS proteins. The secondary structure prediction results showed that the amino acid sequence of all squalene synthases had Alpha helix and random coil as the main components. The results revealed SQS proteins existed four highly conserved domains and 1—2 transmembrane regions in the C-terminal. The result of phylogenetic analysis was consistent with that of plant classical taxonomy. **Conclusion** The protein characteristics of SQSs have less differences among the 20 sequences, and the evolutionary differences are also close, which hint that SQSs have high conservation in evolution. The results provide the basis for the study of SQS function and mechanism.

**Key words:** Chinese herbal plants; triterpenoid; squalene synthase; bioinformatics; domain

三萜类化合物是一类具有抗肿瘤、抗病毒、降血糖、调血脂等多种生物活性的植物次生代谢产物,在太子参、人参、甘草等中药基原植物中广泛存在<sup>[1-3]</sup>。三萜类化合物的结构与生物合成过程比较复杂,关于合成途径中的关键酶的研究是热点之一<sup>[4]</sup>。研究表明,三萜类化合物主要有两条合成

途径生成,即甲羟戊酸(mevalonate pathway, MVA)和甲基-赤藓醇-4-磷酸途径(2-methyl-D-erythritol-4-phosphate, MEP)途径,其中MVA途径是植物三萜类化合物合成的主要途径<sup>[5-8]</sup>。在MVA途径中,鲨烯合酶(squalene synthase, SQS)位于三萜合成通路的分支点上,决定着法尼基焦磷酸的流向,是

收稿日期: 2020-07-09

基金项目: 福建省教育厅中青年骨干教师科研项目(JT180138); 福建农林大学科技创新基金资助项目(KFA17580A)

作者简介: 王琴波(1994—),男,学士,助理研究员,主要从事中草药生物技术研究。E-mail:1791012905@qq.com

\*通信作者: 陈观水(1979—),男,博士,副教授,主要从事中草药生物技术研究。Tel: (0591)83789494 E-mail: gshchenfafu@163.com

三萜类化合物生物合成途径中一个关键酶,其含量及活性决定了植物该物质的产量,也是三萜皂苷合成代谢中的一个重要的限速酶<sup>[9-14]</sup>。目前,已从各种各样的生物,如动物<sup>[15]</sup>、人类<sup>[16]</sup>、真菌<sup>[17]</sup>、植物<sup>[9]</sup>中成功分离了鲨烯合酶编码基因序列近 1000 条。通过对 SQS 的结构及理化特性分析表明,不同来源的 SQS 在功能位点、结构特征存在一定的差异。因此,鲨烯合酶作为三萜类化合物代谢途径中关键调控酶备受关注,具有重要的研究意义。

本研究运用生物信息学相关软件及其分析方法,对太子参等《中国药典》2015 年版<sup>[18]</sup>收录的 20 种中药基原植物中 SQS 蛋白氨基酸序列的理化特性、跨膜结构域、亚细胞定位、亲水性进行比较与分析,并构建 SQS 蛋白家族的系统发育树。为进一步研究植物 SQS 的功能与结构特征和调控植物三萜类化合物的生物合成提供依据。

## 1 数据来源

设定 SQS 为关键词进行搜索,从美国国立生物技术信息中心(National Center of Biotechnology Information, NCBI)下载并筛选出了太子参 *Pseudostellaria heterophylla* (Miq.) Pax ex Pax et Hoffm(登录号 AQV11962.1)、人参 *Panax ginseng* C. A. Meyer (登录号 ACV88718.1)、三七 *Panax notoginseng* (Burkill) F. H. Chen ex C. H. (登录号 AIK21786.1)、西洋参 *Panax quinquefolium* L. (登录号(AED99863.1)、竹节参 *Panax japonicas* (T. Nees) C. A. Mey. (登录号 ALB38664.1)、刺五加 *Acanthopanax senticosus* (Rupr. Maxim.) Harms (登录号 AER23670.1)、柴胡 *Bupleurum chinense* D. C. (登录号(ACX42425.1)、甘草 *Glycyrrhiza uralensis* Fisch. (登录号 ACS66749.1)、膜荚黄芪 *Astragalus membranaceus* (Fisch.) Bunge (登录号(ALA23415.1)、千金子 *Euphorbia lathyris* (L.) Nees (登录号(AFZ93644.1)、栀子 *Gardenia jasminoides* Ellis. (登录号 AYC62335.1)、瓜蒌 *Trichosanthes kirilowii* Maxim. (登录号(ARE29883.1)、金铁锁 *Psammosilene tunicoides* W. C. Wu & C. Y. Wu (登录号(ABQ96265.1)、远志 *Polygala tenuifolia* Willd. (登录号 ABG66304.1)、罗汉果 *Siraitia grosvenorii* (Swingle) C. Jeffery (登录号(ANM71227.1)、丹参 *Salvia miltiorrhiza* Bunge (登录号 ACR57219.1)、黄花蒿 *Artemisia annua* L. (登录号 AAR20329.1)、

京大戟 *Euphorbia pekinensis* Rupr. (ER) (登录号 AFT92039.1)、光果甘草 *Glycyrrhiza glabra* L. (登录号(AMR98504.1)、积雪草 *Centella asiatica* (L.) Urb. (登录号 AAV58897.1) 20 种《中国药典》2015 年版中收录的中药基原植物完整的 SQS 氨基酸序列。

## 2 方法

SQS 氨基酸序列利用 NCBI 网站进行在线分析;氨基酸序列的组成、相对分子量、等电点、不稳定系数等理化性质利用 Protparam (<http://web.expasy.org/protparam/>) 在线进行分析;跨膜结构域用 TMHMM Serverv.2.0 (<http://www.cbs.dtu.dk/services/TMHMM-2.0/>) 进行预测;亚细胞定位利用 PSORTPrediction (<https://wolfsort.hgc.jp/>) 进行分析;信号肽利用 SignalP 软件 5.0 版(<http://www.cbs.dtu.dk/services/SignalP/>) 进行预测;采用 SOPMA ([https://npsa-prabi.ibcp.fr/cgi-bin/npsa\\_automat.pl/page=npsa\\_sopma.html](https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl/page=npsa_sopma.html)) 进行蛋白二结构的预测分析;采 CDD 在线工具分析 (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) 进行功能域预测;蛋白质磷酸化位点利用 NetPhos3.1 Server (<http://www.cbs.dtu.dk/services/NetPhos/>) 进行预测;利用 DNAMAN 9.0 进行 SQS 蛋白氨基酸序列的多重比对分析;采用 MEGA 10.0 软件构建 SQS 蛋白的系统进化树;利用 MEME 在线分析工具 (<http://meme-suite.org/tools/meme>) 进行 SQS 蛋白的保守基序进行分析。

## 3 结果与分析

### 3.1 不同中药基原植物鲨烯合酶氨基酸序列理化性质分析

使用在线软件 Protparam 对 20 条 SQS 的氨基酸序列进行理化性质预测分析(表 1)。预测分析表明,20 条氨基酸序列的等电点在 6.19~8.53,呈现弱酸性或弱碱性,正负电荷比例约为 1:1;同时对氨基酸序列的组成成分也进行了分析,发现所有 SQS 蛋白中亮氨酸(L)含量最高,最高可达 11.8%。相对分子量之间差别不大,在 46 960~47 890;总平均亲水系数均为负值,亲水性较好,表明 SQS 均属于亲水性蛋白;预测的蛋白的不稳定指数在 35.77~44.38,表明 20 种蛋白中既有稳定性蛋白也有不稳定性蛋白。

表 1 20 种中药基原植物 SQS 蛋白理化性质的预测分析

Table 1 Prediction and analysis of physical and chemical properties of SQS protein in 20 Chinese herbal plants

中药	氨基酸长度	理论等电点	相对分子质量	亮氨酸/%	带负/正电荷氨基酸	平均亲水系数	不稳定指数
人参	415	6.19	47 130	11.1	12.0/10.8	-0.046	39.59
三七	415	6.50	47 070	10.8	12.3/11.6	-0.076	42.03
西洋参	415	6.24	47 130	11.1	12.5/11.3	-0.062	41.34
竹节参	415	6.36	47 070	10.8	12.3/11.3	-0.076	42.11
刺五加	414	6.67	47 040	10.9	11.8/11.4	-0.057	39.78
柴胡	414	6.36	47 200	10.9	11.8/10.9	-0.071	39.76
甘草	412	8.21	47 080	11.4	11.7/12.4	-0.087	36.15
膜荚黄芪	413	6.70	47 230	11.4	11.6/11.4	-0.036	42.51
千金子	410	7.54	46 960	10.7	12.4/12.7	-0.100	41.79
栀子	414	6.28	47 130	10.6	13.0/12.1	-0.095	40.55
瓜蒌	417	8.14	47 580	11.0	11.5/12.2	-0.034	41.74
太子参	414	7.18	47 390	10.1	12.1/12.1	-0.070	43.07
金铁锁	414	7.58	47 590	10.1	12.1/12.3	-0.065	44.38
远志	413	6.94	47 250	11.1	11.8/11.6	-0.108	35.77
罗汉果	417	7.15	47 560	11.8	11.8/11.8	-0.043	41.94
丹参	413	6.60	47 400	9.9	12.8/12.3	-0.082	42.38
黄花蒿	418	8.53	47 890	10.5	11.5/12.7	-0.079	43.80
京大戟	411	8.16	47 160	9.7	12.2/12.9	-0.123	38.82
光果甘草	413	8.18	47 250	11.1	11.6/12.3	-0.067	43.20
积雪草	415	7.16	47 330	11.3	11.8/11.8	-0.101	41.73

3.2 亚细胞定位、信号肽及跨膜结构域预测与分析

利用在线软件 TMHMM 对太子参等 20 个 SQS 蛋白进行跨膜预测 (表 2), 显示除来源于栀子和丹参的 SQS 蛋白只有 1 个跨膜, 其余 SQS 蛋白均含有 2 个跨膜结构。亚细胞定位预测表明 (表 2), 20 个 SQS 蛋白均具有多个亚细胞定位, 其主要分布于细胞质和质膜中, 并在细胞核、线粒体、高尔基体、内质网、液泡中也有少量分布,

说明这些蛋白质主要在细胞质和质膜中行使其功能。利用在线工具 Signal P 对 20 个 SQS 蛋白进行信号肽预测 (表 2), 结果显示, 所有的 SQS 蛋白均不存在信号肽。

3.3 SQS 蛋白二级结构预测

利用 SOPMA 对 20 种中药 SQS 蛋白的二级结构进行预测 (表 3)。结果表明, 所有 SQS 蛋白均由  $\alpha$ -螺旋、延伸连和无规则卷曲 3 种构件组成, 其

表 2 不同中药基原植物 SQS 蛋白跨膜、亚细胞定位及信号肽预测分析

Table 2 SQS protein sequences prediction analysis of transmembrane, subcellular localization and signal peptide from different Chinese herbal plants

中药	跨膜螺旋数量	亚细胞定位	信号肽
人参	2	细胞质、细胞核、高尔基体、叶绿体、线粒体、质膜、内质网	无
三七	2	质膜、内质网、细胞质、液泡、细胞核	无
西洋参	2	质膜、内质网、细胞核、细胞质、液泡	无
竹节参	2	质膜、内质网、细胞质、液泡、细胞核	无
刺五加	2	内质网、质膜、细胞核、细胞质、液泡	无
柴胡	2	细胞质、高尔基体、细胞核、内质网、叶绿体、质膜	无
甘草	2	细胞质、内质网、高尔基体、叶绿体、细胞核、线粒体、质膜	无
膜荚黄芪	2	质膜、内质网、细胞质、液泡	无
千金子	2	叶绿体、细胞质、内质网、高尔基体、细胞核、质膜	无
栀子	1	细胞质、细胞核、高尔基体、叶绿体、质膜、内质网	无
瓜蒌	2	细胞质、叶绿体、内质网、高尔基体、细胞核、质膜、Pero	无
太子参	2	细胞质、细胞核、叶绿体、液泡、内质网、高尔基体	无
金铁锁	2	细胞质、内质网、高尔基体、叶绿体、细胞核、质膜	无
远志	2	细胞质、内质网、高尔基体、叶绿体、细胞核、质膜	无
罗汉果	2	细胞质、细胞核、叶绿体、高尔基体、质膜、内质网、Pero	无
丹参	1	细胞质、细胞核、内质网、叶绿体、高尔基体、细胞核、线粒体	无
黄花蒿	2	细胞质、高尔基体、细胞核、线粒体、叶绿体、内质网	无
京大戟	2	细胞质、叶绿体、内质网、质膜、细胞核	无
光果甘草	2	细胞质、细胞核、质膜、内质网、高尔基体、Pero	无
积雪草	2	细胞质、高尔基体、叶绿体、细胞核、质膜、内质网	无

表3 不同中药基原植物 SQS 蛋白二级结构元件比例

Table 3 Secondary structure element ratio of SQS in different Chinese herbal plants

中药	α-螺旋	延伸链	无规则卷曲
太子参	37.44	24.15	38.41
人参	37.83	20.72	41.45
三七	39.90	20.72	43.37
西洋参	37.83	20.72	41.45
竹节参	39.04	18.80	42.17
刺五加	42.51	18.84	38.65
柴胡	41.55	17.15	41.30
甘草	46.84	17.23	35.92
膜荚黄芪	44.07	18.64	37.29
千金子	42.68	20.00	37.32
桅子	39.61	21.01	39.37
瓜蒌	41.01	16.31	42.69
金铁锁	38.65	22.22	39.13
远志	47.94	17.43	34.62
罗汉果	42.69	16.79	40.53
丹参	45.76	19.61	34.62
黄花蒿	44.74	18.18	37.08
京大戟	38.93	19.22	41.85
光果甘草	41.89	20.34	37.77
积雪草	38.55	19.28	42.17

中以 α-螺旋和无规则卷曲为主。以太子参 SQS 为例，其 α-螺旋、延伸连和无规则卷曲 3 种构件的比例分别为 38.65%、22.22% 和 39.13%。因此推测，α-螺旋和无规则卷曲是植物 SQS 蛋白中主要存在的结构元件，并且分散在整个多肽链中。

### 3.4 SQS 蛋白的保守结构域分析

利用 CDD 在线工具分析太子参等 20 种 SQS 氨基酸序列的功能结构域。结果表明，20 种 SQS 氨基酸的保守结构域中均含有底物结合区、底物-镁离子结合位点、活性位点盖残基、催化残基和 2 个天冬氨酸富集区，具有典型的多聚异戊二烯基合成酶活性结构域和鲨烯/八氢番茄红素合成酶活性结构域，属于 Isoprenoid-Biosyn-C1 超家族，为类异戊二烯生物合成酶。太子参 SQS 的保守结构域见图 1。

### 3.5 SQS 蛋白的磷酸化位点

利用 NetPhos 3.1 Server 对 20 种中药基原植物

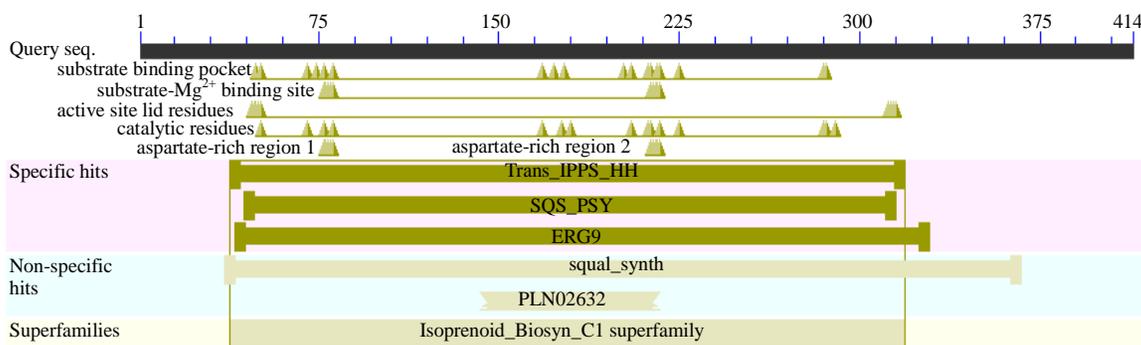


图1 太子参 SQS 的保守结构域分析

Fig. 1 Conserved domain analysis of PhSQS

SQS 蛋白磷酸化位点进行预测，位点数在 29~38 个，其中丝氨酸磷酸化位点数在 12~20 个，苏氨酸位点有 5~15 个，酪氨酸磷酸化位点有 5~8 个。以太子参 SQS 为例，共有 32 个磷酸化位点，其中丝氨酸位点有 13 个，苏氨酸位点有 11 个，酪氨酸磷酸化位点有 8 个，其他中草药植物的 SQS 蛋白的氨基酸磷酸化位点差别。见表 4。

### 3.6 SQS 蛋白多序列比对与系统进化树分析

20 种中药基原植物的 SQS 氨基酸序列的多重分析表明 (图 2)，所有 20 种 SQS 均含有 4 个高保守 (I~IV) 且典型的 17~23 个氨基酸长度的结构域和 1 个高度差异且几乎以疏水氨基酸残基为主的高度疏水区结构域 (V)。研究表明，这些结构域与 SQS 的结合、调节及催化活性功能密切相关。用 MEGA 10.0 软件对包括太子参在内的 20 种有代表性的 SQS 蛋白构建系统进化树 (图 3)。

表4 20 种中药基原植物 SQS 蛋白磷酸化作用位点预测结果  
Table 4 Prediction of phosphorylation sites of SQS protein from 20 Chinese herbal plants

中药	氨基酸磷酸化位点数			位点总数
	丝氨酸	苏氨酸	酪氨酸	
人参	17	11	8	36
三七	20	11	7	38
西洋参	18	10	7	35
竹节参	18	10	6	34
刺五加	19	9	6	34
柴胡	20	9	9	38
甘草	16	12	7	35
膜荚黄芪	18	11	6	35
千金子	13	12	5	30
桅子	17	11	7	35
瓜蒌	16	14	7	37
太子参	13	11	8	32
金铁锁	13	12	6	31
远志	14	8	7	29
罗汉果	12	14	6	32
丹参	15	15	6	36
黄花蒿	19	11	7	37
京大戟	14	10	7	31
光果甘草	16	10	6	31
积雪草	18	10	8	36

太子参 (AQV11962.1)	ISLIGALLKRPDEIYPLIKLRFARFAERKTEPEEHWGQVNSMLKVSRSSEAVI100LDTCLNNAVGFYLVLRALDVTDEDDTSIPTEEMKVPILHREKHFHYDRD	105
人参 (ACV88718.1)	IGSLIGALLKRPDDFYPLIKLRFARFAERKTEPEEHWGQVNSMLKVSRSSEAVI100LGPCLDADAVGFYLVLRALDVTDEDDTSISSTEMKVPILHREKHFHYDND	105
三七 (AIK21786.1)	IGSLIGALLKRPDDFYPLIKLRFARFAERKTEPEEHWGQVNSMLKVSRSSEAVI100LGPCLDADAVGFYLVLRALDVTDEDDTSISSTEMKVPILHREKHFHYDND	105
西洋参 (AED99863.1)	IGSLIGALLKRPDDFYPLIKLRFARFAERKTEPEEHWGQVNSMLKVSRSSEAVI100LGPCLDADAVGFYLVLRALDVTDEDDTSIPTEEMKVPILHREKHFHYDKD	105
竹节参 (ALB38664.1)	IGSLIGALLKRPDDFYPLIKLRFARFAERKTEPEEHWGQVNSMLKVSRSSEAVI100LGPCLDADAVGFYLVLRALDVTDEDDTSIPTEEMKVPILHREKHFHYDKD	105
刺五加 (AER23670.1)	IGSLIGALLKRPDDFYPLIKLRFARFAERKTEPEEHWGQVNSMLKVSRSSEAVI100LGPCLDADAVGFYLVLRALDVTDEDDTSIPTEEMKVPILHREKHFHYDKD	105
柴胡 (ACX42425.1)	IGSLIGALLKRPDDFYPLIKLRFARFAERKTEPEEHWGQVNSMLKVSRSSEAVI100LGPCLDADAVGFYLVLRALDVTDEDDTSISSESEMKVPILHREKHFHYDND	105
甘草 (ACS66749.1)	IGSLIGALLKRPDDFYPLIKLRFARFAERKTEPEEHWGQVNSMLKVSRSSEAVI100LGPCLDADAVGFYLVLRALDVTDEDDTSISSTEMKVPILHREKHFHYDRD	105
膜荚黄芪 (ALA23415.1)	IGSLIGALLKRPDDFYPLIKLRFARFAERKTEPEEHWGQVNSMLKVSRSSEAVI100LGPCLDADAVGFYLVLRALDVTDEDDTSISSTEMKVPILHREKHFHYDRD	105
千金子 (AFZ93644.1)	IGSLIGALLKRPDDFYPLIKLRFARFAERKTEPEEHWGQVNSMLKVSRSSEAVI100LGPCLDADAVGFYLVLRALDVTDEDDTSISSTEMKVPILHREKHFHYDRD	105
栀子 (AYC62335.1)	IGSLIGALLKRPDDFYPLIKLRFARFAERKTEPEEHWGQVNSMLKVSRSSEAVI100LGPCLDADAVGFYLVLRALDVTDEDDTSISSTEMKVPILHREKHFHYDRD	105
瓜蒌 (ARE29883.1)	IGSLIGALLKRPDDFYPLIKLRFARFAERKTEPEEHWGQVNSMLKVSRSSEAVI100LGPCLDADAVGFYLVLRALDVTDEDDTSISSTEMKVPILHREKHFHYDRD	105
金铁锁 (ABQ96265.1)	IGSLIGALLKRPDDFYPLIKLRFARFAERKTEPEEHWGQVNSMLKVSRSSEAVI100LGPCLDADAVGFYLVLRALDVTDEDDTSISSTEMKVPILHREKHFHYDRD	105
远志 (ABG66304.1)	IGSLIGALLKRPDDFYPLIKLRFARFAERKTEPEEHWGQVNSMLKVSRSSEAVI100LGPCLDADAVGFYLVLRALDVTDEDDTSISSTEMKVPILHREKHFHYDRD	105
罗汉果 (ANM71227.1)	IGSLIGALLKRPDDFYPLIKLRFARFAERKTEPEEHWGQVNSMLKVSRSSEAVI100LGPCLDADAVGFYLVLRALDVTDEDDTSISSTEMKVPILHREKHFHYDRD	105
丹参 (ACR57219.1)	IGSLIGALLKRPDDFYPLIKLRFARFAERKTEPEEHWGQVNSMLKVSRSSEAVI100LGPCLDADAVGFYLVLRALDVTDEDDTSISSTEMKVPILHREKHFHYDRD	105
黄花蒿 (AAR20329.1)	IGSLIGALLKRPDDFYPLIKLRFARFAERKTEPEEHWGQVNSMLKVSRSSEAVI100LGPCLDADAVGFYLVLRALDVTDEDDTSISSTEMKVPILHREKHFHYDRD	105
京大戟 (AFT92039.1)	IGSLIGALLKRPDDFYPLIKLRFARFAERKTEPEEHWGQVNSMLKVSRSSEAVI100LGPCLDADAVGFYLVLRALDVTDEDDTSISSTEMKVPILHREKHFHYDRD	105
光果甘草 (AMR98504.1)	IGSLIGALLKRPDDFYPLIKLRFARFAERKTEPEEHWGQVNSMLKVSRSSEAVI100LGPCLDADAVGFYLVLRALDVTDEDDTSISSTEMKVPILHREKHFHYDRD	105
积雪草 (AAV58897.1)	IGSLIGALLKRPDDFYPLIKLRFARFAERKTEPEEHWGQVNSMLKVSRSSEAVI100LGPCLDADAVGFYLVLRALDVTDEDDTSISSTEMKVPILHREKHFHYDRD	105
太子参 (AQV11962.1)	WHFSCTGRKVMIDQEHVSAFIEIDRQYQAEIETIKKMGQGMKFIQGEVETIDYDEYCHYVAGLVGIGSKLHNSGLDIASSISNSMGLLQKNTNI	210
人参 (ACV88718.1)	WHFSCTGRKVMIDQEHVSAFIEIDRQYQAEIETIKKMGQGMKFIQGEVETIDYDEYCHYVAGLVGIGSKLHNSGLDIASSISNSMGLLQKNTNI	210
三七 (AIK21786.1)	WHFSCTGRKVMIDQEHVSAFIEIDRQYQAEIETIKKMGQGMKFIQGEVETIDYDEYCHYVAGLVGIGSKLHNSGLDIASSISNSMGLLQKNTNI	210
西洋参 (AED99863.1)	WHFSCTGRKVMIDQEHVSAFIEIDRQYQAEIETIKKMGQGMKFIQGEVETIDYDEYCHYVAGLVGIGSKLHNSGLDIASSISNSMGLLQKNTNI	210
竹节参 (ALB38664.1)	WHFSCTGRKVMIDQEHVSAFIEIDRQYQAEIETIKKMGQGMKFIQGEVETIDYDEYCHYVAGLVGIGSKLHNSGLDIASSISNSMGLLQKNTNI	210
刺五加 (AER23670.1)	WHFSCTGRKVMIDQEHVSAFIEIDRQYQAEIETIKKMGQGMKFIQGEVETIDYDEYCHYVAGLVGIGSKLHNSGLDIASSISNSMGLLQKNTNI	210
柴胡 (ACX42425.1)	WHFSCTGRKVMIDQEHVSAFIEIDRQYQAEIETIKKMGQGMKFIQGEVETIDYDEYCHYVAGLVGIGSKLHNSGLDIASSISNSMGLLQKNTNI	210
甘草 (ACS66749.1)	WHFSCTGRKVMIDQEHVSAFIEIDRQYQAEIETIKKMGQGMKFIQGEVETIDYDEYCHYVAGLVGIGSKLHNSGLDIASSISNSMGLLQKNTNI	210
膜荚黄芪 (ALA23415.1)	WHFSCTGRKVMIDQEHVSAFIEIDRQYQAEIETIKKMGQGMKFIQGEVETIDYDEYCHYVAGLVGIGSKLHNSGLDIASSISNSMGLLQKNTNI	210
千金子 (AFZ93644.1)	WHFSCTGRKVMIDQEHVSAFIEIDRQYQAEIETIKKMGQGMKFIQGEVETIDYDEYCHYVAGLVGIGSKLHNSGLDIASSISNSMGLLQKNTNI	210
栀子 (AYC62335.1)	WHFSCTGRKVMIDQEHVSAFIEIDRQYQAEIETIKKMGQGMKFIQGEVETIDYDEYCHYVAGLVGIGSKLHNSGLDIASSISNSMGLLQKNTNI	210
瓜蒌 (ARE29883.1)	WHFSCTGRKVMIDQEHVSAFIEIDRQYQAEIETIKKMGQGMKFIQGEVETIDYDEYCHYVAGLVGIGSKLHNSGLDIASSISNSMGLLQKNTNI	210
金铁锁 (ABQ96265.1)	WHFSCTGRKVMIDQEHVSAFIEIDRQYQAEIETIKKMGQGMKFIQGEVETIDYDEYCHYVAGLVGIGSKLHNSGLDIASSISNSMGLLQKNTNI	210
远志 (ABG66304.1)	WHFSCTGRKVMIDQEHVSAFIEIDRQYQAEIETIKKMGQGMKFIQGEVETIDYDEYCHYVAGLVGIGSKLHNSGLDIASSISNSMGLLQKNTNI	210
罗汉果 (ANM71227.1)	WHFSCTGRKVMIDQEHVSAFIEIDRQYQAEIETIKKMGQGMKFIQGEVETIDYDEYCHYVAGLVGIGSKLHNSGLDIASSISNSMGLLQKNTNI	210
丹参 (ACR57219.1)	WHFSCTGRKVMIDQEHVSAFIEIDRQYQAEIETIKKMGQGMKFIQGEVETIDYDEYCHYVAGLVGIGSKLHNSGLDIASSISNSMGLLQKNTNI	210
黄花蒿 (AAR20329.1)	WHFSCTGRKVMIDQEHVSAFIEIDRQYQAEIETIKKMGQGMKFIQGEVETIDYDEYCHYVAGLVGIGSKLHNSGLDIASSISNSMGLLQKNTNI	210
京大戟 (AFT92039.1)	WHFSCTGRKVMIDQEHVSAFIEIDRQYQAEIETIKKMGQGMKFIQGEVETIDYDEYCHYVAGLVGIGSKLHNSGLDIASSISNSMGLLQKNTNI	210
光果甘草 (AMR98504.1)	WHFSCTGRKVMIDQEHVSAFIEIDRQYQAEIETIKKMGQGMKFIQGEVETIDYDEYCHYVAGLVGIGSKLHNSGLDIASSISNSMGLLQKNTNI	210
积雪草 (AAV58897.1)	WHFSCTGRKVMIDQEHVSAFIEIDRQYQAEIETIKKMGQGMKFIQGEVETIDYDEYCHYVAGLVGIGSKLHNSGLDIASSISNSMGLLQKNTNI	210
太子参 (AQV11962.1)	TROYLEDINEIPKSRMFWFRITASKYDKLDDIYEENSSEKAVCLINDVINGLHVEVETIIVMSAIEHSHFRFCALPQIMSIGTLAIGHNNICVFRGVWR	315
人参 (ACV88718.1)	TROYLEDINEIPKSRMFWFRITASKYDKLDDIYEENSSEKAVCLINDVINGLHVEVETIIVMSAIEHSHFRFCALPQIMSIGTLAIGHNNICVFRGVWR	315
三七 (AIK21786.1)	TROYLEDINEIPKSRMFWFRITASKYDKLDDIYEENSSEKAVCLINDVINGLHVEVETIIVMSAIEHSHFRFCALPQIMSIGTLAIGHNNICVFRGVWR	315
西洋参 (AED99863.1)	TROYLEDINEIPKSRMFWFRITASKYDKLDDIYEENSSEKAVCLINDVINGLHVEVETIIVMSAIEHSHFRFCALPQIMSIGTLAIGHNNICVFRGVWR	315
竹节参 (ALB38664.1)	TROYLEDINEIPKSRMFWFRITASKYDKLDDIYEENSSEKAVCLINDVINGLHVEVETIIVMSAIEHSHFRFCALPQIMSIGTLAIGHNNICVFRGVWR	315
刺五加 (AER23670.1)	TROYLEDINEIPKSRMFWFRITASKYDKLDDIYEENSSEKAVCLINDVINGLHVEVETIIVMSAIEHSHFRFCALPQIMSIGTLAIGHNNICVFRGVWR	315
柴胡 (ACX42425.1)	TROYLEDINEIPKSRMFWFRITASKYDKLDDIYEENSSEKAVCLINDVINGLHVEVETIIVMSAIEHSHFRFCALPQIMSIGTLAIGHNNICVFRGVWR	315
甘草 (ACS66749.1)	TROYLEDINEIPKSRMFWFRITASKYDKLDDIYEENSSEKAVCLINDVINGLHVEVETIIVMSAIEHSHFRFCALPQIMSIGTLAIGHNNICVFRGVWR	315
膜荚黄芪 (ALA23415.1)	TROYLEDINEIPKSRMFWFRITASKYDKLDDIYEENSSEKAVCLINDVINGLHVEVETIIVMSAIEHSHFRFCALPQIMSIGTLAIGHNNICVFRGVWR	315
千金子 (AFZ93644.1)	TROYLEDINEIPKSRMFWFRITASKYDKLDDIYEENSSEKAVCLINDVINGLHVEVETIIVMSAIEHSHFRFCALPQIMSIGTLAIGHNNICVFRGVWR	315
栀子 (AYC62335.1)	TROYLEDINEIPKSRMFWFRITASKYDKLDDIYEENSSEKAVCLINDVINGLHVEVETIIVMSAIEHSHFRFCALPQIMSIGTLAIGHNNICVFRGVWR	315
瓜蒌 (ARE29883.1)	TROYLEDINEIPKSRMFWFRITASKYDKLDDIYEENSSEKAVCLINDVINGLHVEVETIIVMSAIEHSHFRFCALPQIMSIGTLAIGHNNICVFRGVWR	315
金铁锁 (ABQ96265.1)	TROYLEDINEIPKSRMFWFRITASKYDKLDDIYEENSSEKAVCLINDVINGLHVEVETIIVMSAIEHSHFRFCALPQIMSIGTLAIGHNNICVFRGVWR	315
远志 (ABG66304.1)	TROYLEDINEIPKSRMFWFRITASKYDKLDDIYEENSSEKAVCLINDVINGLHVEVETIIVMSAIEHSHFRFCALPQIMSIGTLAIGHNNICVFRGVWR	315
罗汉果 (ANM71227.1)	TROYLEDINEIPKSRMFWFRITASKYDKLDDIYEENSSEKAVCLINDVINGLHVEVETIIVMSAIEHSHFRFCALPQIMSIGTLAIGHNNICVFRGVWR	315
丹参 (ACR57219.1)	TROYLEDINEIPKSRMFWFRITASKYDKLDDIYEENSSEKAVCLINDVINGLHVEVETIIVMSAIEHSHFRFCALPQIMSIGTLAIGHNNICVFRGVWR	315
黄花蒿 (AAR20329.1)	TROYLEDINEIPKSRMFWFRITASKYDKLDDIYEENSSEKAVCLINDVINGLHVEVETIIVMSAIEHSHFRFCALPQIMSIGTLAIGHNNICVFRGVWR	315
京大戟 (AFT92039.1)	TROYLEDINEIPKSRMFWFRITASKYDKLDDIYEENSSEKAVCLINDVINGLHVEVETIIVMSAIEHSHFRFCALPQIMSIGTLAIGHNNICVFRGVWR	315
光果甘草 (AMR98504.1)	TROYLEDINEIPKSRMFWFRITASKYDKLDDIYEENSSEKAVCLINDVINGLHVEVETIIVMSAIEHSHFRFCALPQIMSIGTLAIGHNNICVFRGVWR	315
积雪草 (AAV58897.1)	TROYLEDINEIPKSRMFWFRITASKYDKLDDIYEENSSEKAVCLINDVINGLHVEVETIIVMSAIEHSHFRFCALPQIMSIGTLAIGHNNICVFRGVWR	315
太子参 (AQV11962.1)	SLTAVIVDQNSHADVSGAFIDACMLKRVKTDIETKMLSEHEATQKRSSTLNKRKLLIVKKGKGLTPAVVVFVIFILAIIMRSIISNNQPNY...	414
人参 (ACV88718.1)	SLTAVIVDQNSHADVSGAFIDACMLKRVKTDIETKMLSEHEATQKRSSTLNKRKLLIVKKGKGLTPAVVVFVIFILAIIMRSIISNNQPNY...	414
三七 (AIK21786.1)	SLTAVIVDQNSHADVSGAFIDACMLKRVKTDIETKMLSEHEATQKRSSTLNKRKLLIVKKGKGLTPAVVVFVIFILAIIMRSIISNNQPNY...	414
西洋参 (AED99863.1)	SLTAVIVDQNSHADVSGAFIDACMLKRVKTDIETKMLSEHEATQKRSSTLNKRKLLIVKKGKGLTPAVVVFVIFILAIIMRSIISNNQPNY...	414
竹节参 (ALB38664.1)	SLTAVIVDQNSHADVSGAFIDACMLKRVKTDIETKMLSEHEATQKRSSTLNKRKLLIVKKGKGLTPAVVVFVIFILAIIMRSIISNNQPNY...	414
刺五加 (AER23670.1)	SLTAVIVDQNSHADVSGAFIDACMLKRVKTDIETKMLSEHEATQKRSSTLNKRKLLIVKKGKGLTPAVVVFVIFILAIIMRSIISNNQPNY...	414
柴胡 (ACX42425.1)	SLTAVIVDQNSHADVSGAFIDACMLKRVKTDIETKMLSEHEATQKRSSTLNKRKLLIVKKGKGLTPAVVVFVIFILAIIMRSIISNNQPNY...	414
甘草 (ACS66749.1)	SLTAVIVDQNSHADVSGAFIDACMLKRVKTDIETKMLSEHEATQKRSSTLNKRKLLIVKKGKGLTPAVVVFVIFILAIIMRSIISNNQPNY...	414
膜荚黄芪 (ALA23415.1)	SLTAVIVDQNSHADVSGAFIDACMLKRVKTDIETKMLSEHEATQKRSSTLNKRKLLIVKKGKGLTPAVVVFVIFILAIIMRSIISNNQPNY...	414
千金子 (AFZ93644.1)	SLTAVIVDQNSHADVSGAFIDACMLKRVKTDIETKMLSEHEATQKRSSTLNKRKLLIVKKGKGLTPAVVVFVIFILAIIMRSIISNNQPNY...	414
栀子 (AYC62335.1)	SLTAVIVDQNSHADVSGAFIDACMLKRVKTDIETKMLSEHEATQKRSSTLNKRKLLIVKKGKGLTPAVVVFVIFILAIIMRSIISNNQPNY...	414
瓜蒌 (ARE29883.1)	SLTAVIVDQNSHADVSGAFIDACMLKRVKTDIETKMLSEHEATQKRSSTLNKRKLLIVKKGKGLTPAVVVFVIFILAIIMRSIISNNQPNY...	414
金铁锁 (ABQ96265.1)	SLTAVIVDQNSHADVSGAFIDACMLKRVKTDIETKMLSEHEATQKRSSTLNKRKLLIVKKGKGLTPAVVVFVIFILAIIMRSIISNNQPNY...	414
远志 (ABG66304.1)	SLTAVIVDQNSHADVSGAFIDACMLKRVKTDIETKMLSEHEATQKRSSTLNKRKLLIVKKGKGLTPAVVVFVIFILAIIMRSIISNNQPNY...	414
罗汉果 (ANM71227.1)	SLTAVIVDQNSHADVSGAFIDACMLKRVKTDIETKMLSEHEATQKRSSTLNKRKLLIVKKGKGLTPAVVVFVIFILAIIMRSIISNNQPNY...	414
丹参 (ACR57219.1)	SLTAVIVDQNSHADVSGAFIDACMLKRVKTDIETKMLSEHEATQKRSSTLNKRKLLIVKKGKGLTPAVVVFVIFILAIIMRSIISNNQPNY...	414
黄花蒿 (AAR20329.1)	SLTAVIVDQNSHADVSGAFIDACMLKRVKTDIETKMLSEHEATQKRSSTLNKRKLLIVKKGKGLTPAVVVFVIFILAIIMRSIISNNQPNY...	414
京大戟 (AFT92039.1)	SLTAVIVDQNSHADVSGAFIDACMLKRVKTDIETKMLSEHEATQKRSSTLNKRKLLIVKKGKGLTPAVVVFVIFILAIIMRSIISNNQPNY...	414
光果甘草 (AMR98504.1)	SLTAVIVDQNSHADVSGAFIDACMLKRVKTDIETKMLSEHEATQKRSSTLNKRKLLIVKKGKGLTPAVVVFVIFILAIIMRSIISNNQPNY...	414
积雪草 (AAV58897.1)	SLTAVIVDQNSHADVSGAFIDACMLKRVKTDIETKMLSEHEATQKRSSTLNKRKLLIVKKGKGLTPAVVVFVIFILAIIMRSIISNNQPNY...	414

图 2 20 种中药基原植物的 SQS 氨基酸序列比对分析

Fig. 2 Amino acid sequence alignment of SQS protein from 20 Chinese herbal plants

分析结果显示,在进化遗传学上亲缘越近的物种,在 SQS 的分子系统进化树上基本上距离越近。基于氨基酸序列重建的系统进化树,其结果对判断不同植物之间的亲缘关系具有一定的借鉴意义和可行性参考。

通过 MEME 软件的搜索,在 20 个 SQS 蛋白氨基酸序列中发现了 15 个保守基序 (motif)

结构 (图 3),长度在 6~50 个氨基酸,所有 SQS 蛋白含有的基序除在碳端有稍有差别外,其他大体一致,且存在有位于 SQS 保守结构域内的保守基序。同时还发现,在系统发育树的每个分支上的成员具有相同或类似的基序类型和排列顺序,显示 SQS 蛋白之间在这 20 种中药基原植物中的相对保守性。

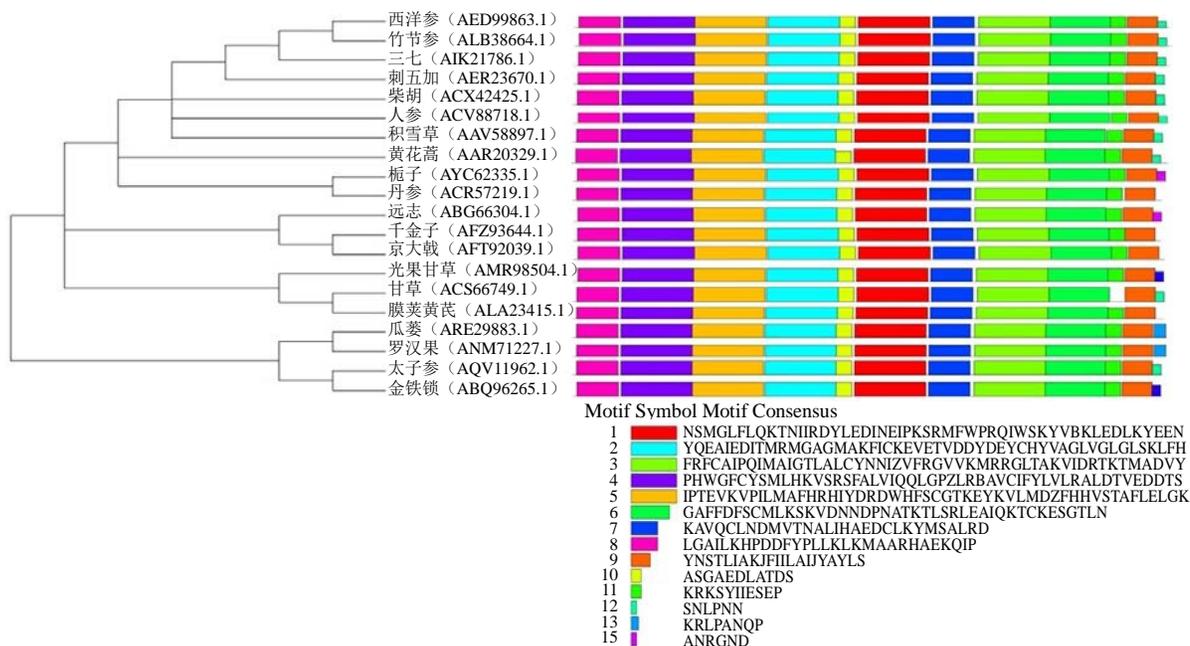


图3 20种中药基原植物SQS蛋白系统发育树及保守基序

Fig. 3 Phylogenetic tree and conserved motif of SQS proteins from 20 Chinese herbal plants

#### 4 讨论

三萜及其苷类广泛存在于自然界，在菌类、蕨类、单子叶、双子叶植物、动物及海洋生物中均有分布，尤以双子叶植物中分布最多<sup>[3]</sup>。三萜类化合物是人参、甘草、柴胡等中草药植物中主要活性物质，常具有抗肿瘤、溶血、抗菌、增加机体免疫力的作用<sup>[1-3]</sup>。通过对三萜类化合物生物合成的研究表明，三萜类化合物是由鲨烯经过不同的途径环合而成，而鲨烯是三萜类化合物合成途径中关键前体物质，它由法尼基焦磷酸 (farnesyl diphosphate, FPP) 经过鲨烯合酶作用下两分子鲨烯通过尾-尾缩合生成<sup>[5-11]</sup>。众多研究表明，鲨烯合酶是三萜类化合物的底物合成起始点的催化合成酶，是三萜类化合物合成途径中一个关键限速酶，对三萜类化合物合成途径的下游物质合成和积累具有重要决定性意义<sup>[9-10]</sup>。

利用生物信息学分析工具对源自《中国药典》2015年版且三萜类化合物为其主要活性物质的中药基原植物的20条SQS蛋白氨基酸序列进行分析。SQS蛋白质一级结构分析，表明这20种中药基原植物的SQS蛋白呈现弱酸性或弱碱性，部分不稳定，部分稳定，均属于亲水性蛋白，不具有信号肽，可推测SQS不是分泌蛋白，这与其属于细胞质中的MVA途径相一致；均存在1~2个跨膜结构域；亚细胞定位预测分析表明其最大可能定位在质膜或细

胞质上，这与前人研究报道鲨烯合酶属于膜结合蛋白相一致<sup>[12]</sup>。二级结构预测结果表明，所有的SQS均以 $\alpha$ -螺旋和任意卷曲为主要成份。多序列比对显示这20种中药基原植物的SQS氨基酸序列彼此间存在较高的相似性，含有4个高保守且典型的17~23个氨基酸长度的结构域(I~IV)包括底物结合区、底物-镁离子结合位点、活性位点盖残基、催化残基和2个天冬氨酸富集区，研究表明，这些高保守区域对于SQS发挥催化作用具有关键作用<sup>[14,19]</sup>。另外，这20条序列中均在氨基酸序列的C端序列差异性极大且属于几乎都是疏水氨基酸残基，研究表明，这些残基有助于鲨烯合酶锚定于细胞器膜上<sup>[20]</sup>。系统发育树的结果显示，20种中药基原植物的SQS蛋白氨基酸序列的聚类结果基本与植物系统分类学结果相一致。所有的信息表明，这20种中药基原植物的SQS蛋白结构和序列上的高度保守也许揭示它们在功能上的相似性。本研究通过对SQS的序列结构的预测和分析为深入研究鲨烯合酶结构与功能的关系、作用机制和代谢过程提供了参考，同时也有助于其他生物的SQS基因的克隆。

#### 参考文献

- [1] 孙丽超, 李淑英, 王凤忠, 等. 萜类化合物的合成生物学研究进展 [J]. 生物技术通报, 2017(1): 64-75.
- [2] 时敏, 王瑶, 周伟, 等. 药用植物萜类化合物的生物合成与代谢调控研究进展 [J]. 中国科学: 生命科学,

- 2018, 48(4): 352-364.
- [3] Biswas T, Dwivedi U N. Plant triterpenoid saponins: Biosynthesis, *in vitro* production, and pharmacological relevance [J]. *Protoplasma*, 2019, 256(6): 1463-1486.
- [4] 陈观水, 林思妮, 柯兰兰, 等. 太子参鲨烯环氧酶基因的克隆及其表达分析 [J]. *中草药*, 2017, 48(13): 2733-2739.
- [5] Augustin J M, Kuzina V, Andersen S B, *et al.* Molecular activities, biosynthesis and evolution of triterpenoid saponins [J]. *Phytochemistry*, 2011, 72(6): 435-457.
- [6] Moses T, Pollier J, Faizal A, *et al.* Unraveling the triterpenoid saponin biosynthesis of the African shrub *Maesa lanceolata* [J]. *Mol Plant*, 2015, 8(1): 122-135.
- [7] Phillips D R, Rasbery J M, Bartel B, *et al.* Biosynthetic diversity in plant triterpene cyclization [J]. *Curr Opin Plant Biol*, 2006, 9(3): 305-314.
- [8] Zhao Y J, Cheng Q Q, Su P, *et al.* Research progress relating to the role of cytochrome P450 in the biosynthesis of terpenoids in medicinal plants [J]. *Appl Microbiol Biotechnol*, 2014, 98(6): 2371-2383.
- [9] Kang J M, Zhang Q Y, Jiang X, *et al.* Molecular cloning and functional identification of a squalene synthase encoding gene from alfalfa (*Medicago sativa* L.) [J]. *Int J Mol Sci*, 2019, 20(18): 4499.
- [10] Fu J, Liu G, Yang M, *et al.* Isolation and functional analysis of squalene synthase gene in tea plant *Camellia sinensis* [J]. *Plant Physiol Biochem*, 2019, 142: 53-58.
- [11] Wang J R, Lin J F, Guo L Q, *et al.* Cloning and characterization of squalene synthase gene from *Poria cocos* and its up-regulation by methyl jasmonate [J]. *World J Microbiol Biotechnol*, 2014, 30(2): 613-620.
- [12] Rong Q, Jiang D, Chen Y, *et al.* Molecular cloning and functional analysis of squalene synthase 2(SQS2) in *Salvia miltiorrhiza bunge* [J]. *Front Plant Sci*, 2016, 7: 1274.
- [13] Liu M, Li L N, Pan Y T, *et al.* cDNA isolation and functional characterization of squalene synthase gene from *Ornithogalum caudatum* [J]. *Protein Expr Purif*, 2017, 130: 63-72.
- [14] Jiang D, Rong Q, Chen Y, *et al.* Molecular cloning and functional analysis of squalene synthase (SS) in *Panax notoginseng* [J]. *Int J Biol Macromol*, 2017, 95: 658-666.
- [15] Inoue T, Osumi T, Hata S. Molecular cloning and functional expression of a cDNA for mouse squalene synthase [J]. *Biochim Biophys Acta*, 1995, 1260(1): 49-54.
- [16] Summers C, Karst F, Charles A D. Cloning, expression and characterisation of the cDNA encoding human hepatic squalene synthase, and its relationship to phytoene synthase [J]. *Gene*, 1993, 136: 185-192.
- [17] Zheng F L, Liu N, Che Y C, *et al.* Cloning, expression and characterization of squalene synthase from *Inonotus obliquus* [J]. *Genes Genom*, 2013, 35(5): 631-639.
- [18] 中国药典 [S]. 一部. 2015: 283.
- [19] Ohtake K, Saito N, Shibuya S, *et al.* Biochemical characterization of the water-soluble squalene synthase from *Methylococcus capsulatus* and the functional analyses of its two DXXD(E)D motifs and the highly conserved aromatic amino acid residues [J]. *FEBS J*, 2014, 281(24): 5479-5497.
- [20] Zha L, Liu S, Su P, *et al.* Cloning, prokaryotic expression and functional analysis of squalene synthase (SQS) in *Magnolia officinalis* [J]. *Protein Expr Purif*, 2016, 120: 28-34.

[责任编辑 时圣明]