

· 药材与资源 ·

宁夏枸杞查耳酮合成酶基因 cDNA 片段的克隆及生物信息学分析

尚洁¹, 陈阿丽², 潘伟志³

1. 北方民族大学生物科学与工程学院, 宁夏 银川 750021

2. 宁夏疾病预防控制中心, 宁夏 银川 750004

3. 华中科技大学生命科学与技术学院, 湖北 武汉 430074

摘要: 目的 克隆宁夏枸杞类黄酮/异黄酮生物合成途径中的关键酶查耳酮合成酶 (CHS, EC2.3.1.74)。方法 以宁夏枸杞为材料, 利用同源克隆和 RT-PCR 技术得到 1 条 CHS 基因, 并对其进行生物信息学分析。结果 该家族的 cDNA 片段长为 1 148 bp, 编码 382 个氨基酸, 蛋白相对分子质量为 4.168×10^4 , 等电点为 6.22。氨基酸序列和结构分析显示 CHS 基因家族有一个活性位点, 即查耳酮和二苯乙烯合成酶活性位点。宁夏枸杞 CHS (LyCHS) 蛋白很可能定位在细胞质。对二级结构和三级结构进行预测, 发现 LyCHS 蛋白中无规则卷曲 126 个, α 螺旋和 β 折叠分别为 166 和 25 个, 延伸链 65 个。系统进化分析表明, LyCHS 基因与马铃薯野生种、马铃薯、番茄具有很高的同源性。该基因已在 GenBank 上注册, 基因序列登录号为 JQ964237。结论 首次克隆了 LyCHS 基因片段, 为研究其表达特性以及功能提供了基础。

关键词: 枸杞; CHS 基因; 克隆; 生物信息学分析; LyCHS 蛋白

中图分类号: R282.12 文献标志码: A 文章编号: 0253-2670(2013)20-2898-06

DOI: 10.7501/j.issn.0253-2670.2013.20.020

Cloning and bioinformation analysis of chalcone synthase gene from *Lycium barbarum* in Ningxia

SHANG Jie¹, CHEN A-li², PAN Wei-zhi³

1. College of Biological Science and Engineering, Beifang University of Nationalities, Yinchuan 750021, China

2. Ningxia Centers for Diseases Prevention Control, Yinchuan 750004, China

3. College of Life Science & Technology, Huazhong University of Science & Technology, Wuhan 430074, China

Abstract: Objective To obtain cDNA of chalcone synthase (CHS, EC 2.3.1.74) involved in the flavonoid/isoflavonoid biosynthesis pathway. **Methods** The partial sequence of CHS in *Lycium barbarum* (LyCHS) was successfully cloned by RT-PCR and using a sequence homology strategy, and the bioinformation analysis was carried out. **Results** The cDNA fragment of CHS gene was 1 148 bp in length, which encoded a protein of 382 amino acids with the predicted relative molecular weight of 4.168×10^4 . The estimated isoelectric point (pI) of the putative protein is 6.22. According to the amino acid sequence and structural analysis, it showed that this protein contained one conserved active site, namely chalcone and stilbene synthases active site. Subcellular localizations of LyCHS proteins were likely in the cytoplasm. The secondary and tertiary structures of LyCHS were abundant in α -helixes (166) and random coils (126), while were less in β -turns (25) and extended strains (65). Phylogenetic analysis showed that the genes of LyCHS were closely related to CHS in *Solanum pinnatisectum*, *S. tuberosum*, and *Lycopersicon esculentum*. The sequences had been registered in GenBank with the accession numbers JQ964237. **Conclusion** The present results provide the foundation for the next study of CHS gene about its expression and function in *L. barbarum*.

Key words: *Lycium barbarum* L.; CHS gene; cloning; bioinformation analysis; LyCHS protein

宁夏枸杞 *Lycium barbarum* L. 是一种多年生灌木, 主要生长在中国大多数地区、欧洲各国和地中海地区。在中国, 宁夏枸杞是一味传统常用中药和滋补品。研究表明, 类黄酮是枸杞叶片中一种重要

的活性成分^[1-3], 其与人类的健康有着密切的关系, 具有消除氧自由基、抗炎、抗过敏、抗癌、抗菌、保护心脑血管系统等多种药理作用^[4-11]。

查耳酮合成酶 (CHS, EC2.3.1.74) 催化苯丙烷

收稿日期: 2013-04-28

基金项目: 北方民族大学校级科研项目 (2010Y046)

作者简介: 尚洁 (1979—), 女, 硕士, 讲师, 研究方向为植物分子生物学。Tel: (0951)2067893 E-mail: shangjie@126.com

代谢途径中类黄酮合成支路的初始反应,是关键酶和限速酶,类黄酮是高等植物重要的次生代谢物,在花色形成、抗紫外线照射、生殖性能、抗真菌感染、固氮细菌的募集等过程中发挥不同的功能^[12]。在类黄酮代谢途径中,CHS被认为是最重要的酶^[13]。CHS催化1分子4-对香豆酰-CoA和3分子丙二酰-CoA形成4分子二氧化碳、3分子辅酶A和1分子柚皮素查耳酮。CHS是一种同型二聚体蛋白质,每个单体相对分子质量 $4.2 \times 10^4 \sim 4.5 \times 10^4$ ^[14]。CHS主要定位于细胞质^[15],与内质网膜相连。而另一项研究显示,CHS和查耳酮异构酶(CHI)共同定位于细胞核^[16]。

目前已经对多种植物蝴蝶兰 *Phalaenopsis hybrid* cv. Formosa rose^[17]、马铃薯 *Solanum tuberosum* L.^[18]、核桃 *Juglans regia* L.^[19]、羽扇豆 *Lupinus luteus* L.^[20]、银杏 *Ginkgo biloba* L.^[21]、苔藓 *Physcomitrella patens* (Hedw.) Bruch et Schimp.^[22]等CHS基因进行了cDNA克隆和序列分析、表达。大多数CHS属于多基因家族,但是不同物种的CHS拷贝数差异较大,在拟南芥中只有1个^[23]。CHS的编码区十分保守,约为 1.2×10^3 个碱基,编码近400个氨基酸。苜蓿等^[24]植物的CHS蛋白晶体结构已经解析,蛋白的保守域和活性氨基酸位点的研究已较为深入。然而关于宁夏枸杞CHS基因尚未报道。本研究对CHS基因进行克隆、序列分析,以期利用基因工程技术提高宁夏枸杞对不良环境的抵御能力以及为调控宁夏枸杞类黄酮的积累奠定基础。

1 材料与仪器

1.1 材料

供试材料为宁夏枸杞 *Lycium barbarum* L.,于2011年采自宁夏回族自治区中宁县,后种植于北方民族大学生物化学实验室,采集新鲜幼苗,立即冰冻于液氮中,置于 $-80\text{ }^\circ\text{C}$ 超低温冰柜中,待用。

1.2 主要试剂与菌种

DNA Polymerase、RNAiso for Polysaccharide-rich Plant Tissue、RevertAid First Strand cDNA Synthesis Kit 反转录试剂盒、DL2000 DNA marker 购自TakaRa公司;DNA胶回收试剂盒购自上海生工生物公司;质粒DNA小量试剂盒购自Axygen公司;pEASYTM-Blunt Zero Cloning Kit、Trans1-T1感受态细胞购自Transgen公司;一些常规试剂均为国产分析纯;引物由上海捷瑞生物工程有限公司合成;测序由华中科技大学完成。

2 方法

2.1 总RNA的提取与cDNA链的合成

利用TakaRa公司的RNAiso for Polysaccharide-rich Plant Tissue试剂提取总RNA,操作步骤参考说明书。按照反转录试剂盒PrimeScript RT reagent Kit (TaKaRa)说明书合成第一条链cDNA。

2.2 简并引物设计

CHS的简并引物是根据GenBank上收录的茄科植物 *Solanum lycopersicum* L.、*Petunia x hybrid* Vilm.、*Solanum tuberosum* L.、*Nicotiana tabacum* L.、*Nicotiana alata* Link et Otto、*Nicotiana tomentosiformis* L.、*Solanum pinnatisectum* Dunal、*Nicotiana benthamiana* Domin的氨基酸序列,使用clustalw2进行序列比对,根据比对结果为上下游引物确定两个合适的保守区域,应用Primer Premier 5.0设计简并引物,结合密码子的偏向性,对简并引物进行修饰。

PCR反应条件: $94\text{ }^\circ\text{C}$ 预变性180 s; $94\text{ }^\circ\text{C}$ 变性30 s, $47.5\text{ }^\circ\text{C}$ 退火30 s, $72\text{ }^\circ\text{C}$ 延伸80 s; 共计35循环; $72\text{ }^\circ\text{C}$ 后延伸7 min。

2.3 宁夏枸杞CHS基因序列的生物信息学分析

通过Proteomics Server (<http://web.expasy.org/protparam/>)预测氨基酸序列的相对分子量、等电点、稳定性指数等;利用ProtScale (<http://web.expasy.org/protscale/>)进行疏水性分析;利用TMHMM (<http://www.cbs.dtu.dk/services/TMHMM/>)进行跨膜结构域分析;利用Eslpred^[25] (<http://www.imtech.res.in/raghava/eslpred/>)预测亚细胞定位;利用SignalP4.1^[26] (<http://www.cbs.dtu.dk/services/SignalP/>)预测蛋白质的信号序列;利用FoldIndex^[27] (<http://bioportal.weizmann.ac.il/fldbin/findex>)分析氨基酸序列的无序化特性。用SUMOsp 2.0^[28] (<http://sumosp.biocuckoo.org/down.php>)进行泛素化预测;用MotifScan (http://myhits.isb-sib.ch/cgi-bin/motif_scan)和InterPro^[29] (<http://www.ebi.ac.uk/interpro/>)预测编码蛋白序列上潜在的结构域和功能位点;用NetPhos^[30] (<http://www.cbs.dtu.dk/services/NetPhos/>)预测蛋白磷酸化位点;使用SOPMA^[31] (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html)对蛋白质二级结构进行预测;用SWISS-MODEL^[32-34] (<http://swissmodel.expasy.org/>)预测蛋白三级结构;利用MEGA 5.10软件,采用Neighbour-joining (NJ)法构建系统发育树。

3 结果与分析

3.1 宁夏枸杞 CHS 基因大片段克隆

根据亲缘关系最近的几种茄科植物的 CHS 序列比对, 并使用 Primer Premier 5.0 软件设计得到一对简并引物, 每个引物的简并度都比较小, CHS-F 的简并度仅为 12, CHS-R 为 4, 见表 1。应用宁夏枸杞总 RNA 反转录获得的 cDNA 为模板, 采用以上引物进行简并

PCR, PCR 产物经 1.0% 琼脂糖凝胶电泳进行检测, 见图 1。PCR 扩增出一条大小约为 1 200 bp 特异的条带, 没有出现非特异性条带, 表明此引物扩增特异性强。目标条带经回收、克隆、测序后获得 1 条长 1 148 bp 的 DNA 序列。编码 382 个氨基酸, 相对分子质量约为 4.168×10^4 , 命名为 LyCHS, 提交至 GenBank, 基因序列登录号为 JQ964237。

表 1 CHS 基因克隆引物

Table 1 Primers for CHS gene cloning

引物名称	引物序列 (5'-3')	简并度
CHS-F	CGA AAATGGTSACMGTBGAGG	12
CHS-R	ACAACAGTCTCAACWGTWAGCCC	4

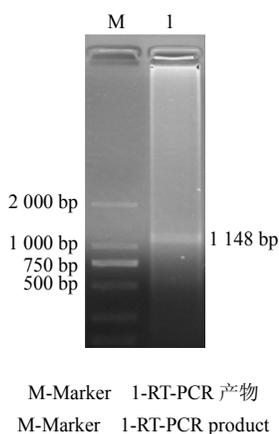


图 1 RT-PCR 产物的电泳结果

Fig. 1 Electrophoresis of RT-PCR product

3.2 LyCHS 蛋白的理化性质

通过 ExPASy Proteomics Server 预测氨基酸序列的相对分子质量、等电点、稳定性指数、脂溶性指数、两亲性指数。LyCHS 蛋白质片段共 382 个氨基酸, 碱性氨基酸残基 (Arg+Lys) 总数为 43, 酸性氨基酸残基 (Asp+Glu) 总数为 46, 相对分子质量为 41.68, 理论等电点为 6.22。在溶液中, 不稳定性指数为 38.91, 脂溶性指数为 90.89, 总平均亲水性指数为 -0.082。

3.3 亲疏水性、亚细胞定位预测

利用位于 ExPASy 的 ProtScale 进行疏水性分析, Window size 设置为 9。从图 2 得知 LyCHS 在第 342 位氨基酸处有最大值 2.422, 大约在 340~350 位氨基酸之间含有 1 个高疏水性区域; 而在第 119 位氨基酸处有最小值 -2.489, 大约在 40~60 位和 100~130 位氨基酸之间有 2 个高亲水性区域。

利用 TMHMM 进行分析, 宁夏枸杞 CHS 蛋白都没有跨膜结构域, 该蛋白位于细胞膜外。Espred

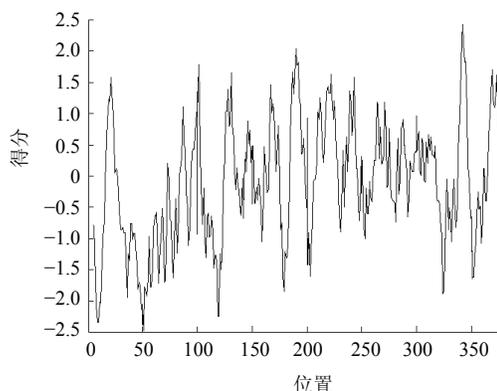


图 2 蛋白疏水区分析

Fig. 2 ProtScale output of hydrophobic area

预测宁夏枸杞 CHS 蛋白定位于细胞质的可能性最大为 53%。另外, 经 Signalp 4.1 预测 LyCHS 蛋白质序列中无明显的信号序列。

3.4 蛋白质无序化特性分析

利用 FoldIndex 分析来源于宁夏枸杞 LyCHS 基因氨基酸序列的无序化特性, 结果表明蛋白含有无序区 2 处 MVTVEEYRKAQRAEGPATVMAIGTATPSNCVDQ (1~33)、FRITDSEHKTELKEKFKRMCDKSMIKKRYMHLT (41~73), 无序残基最长 33 个氨基酸残基, 无序残基总共 66 个氨基酸残基, 见图 3。

3.5 CHS 的结构域和功能位点预测

用 MotifScan 和 InterPro (图 4) 预测 CHS 编码蛋白序列上潜在的结构域和功能位点。结果表明, 该蛋白可能存在的翻译后修饰位点包括 1 个 N-糖基化位点、7 个 N-豆蔻酰化位点、3 个蛋白激酶 C 磷酸化位点、1 个 CAMP 磷酸化位点、8 个酪蛋白激酶 II 位点。在 310~331 处可能存在 1 个亮氨酸拉链。在 156~172 处有一段高度保守的结构功能域

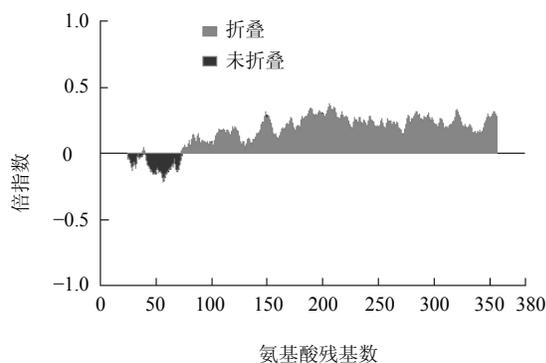


图3 无序化特性分析

Fig. 3 Disordered segment prediction

(RLMMYQQGCFA GGTVLR), 这是查耳酮/二苯乙烯合成酶的特征序列, 符合 R-[LIVMFYS]-x-[LIVM]-x-[QHG]-x-G-C-[FYNA]-[GAPV]-G-[GAC]-[STAVK]-x-[LIVMF]-[RAL]查耳酮/二苯乙烯合成酶标签序列。在第1~382位含有典型的PKSIII型聚酮合酶、类硫解酶结构域。在2~228和238~382处分别是查耳酮和二苯乙烯合成酶N端和C端结构域。除此之外, 预测结果还发现蛋白质序列中有很多与ACP合成酶III相似的结构域, 这是因为CHS与脂肪酸合成酶基因具有很高的同源性, 二者可能是由同一个祖先基因进化而来^[35]。

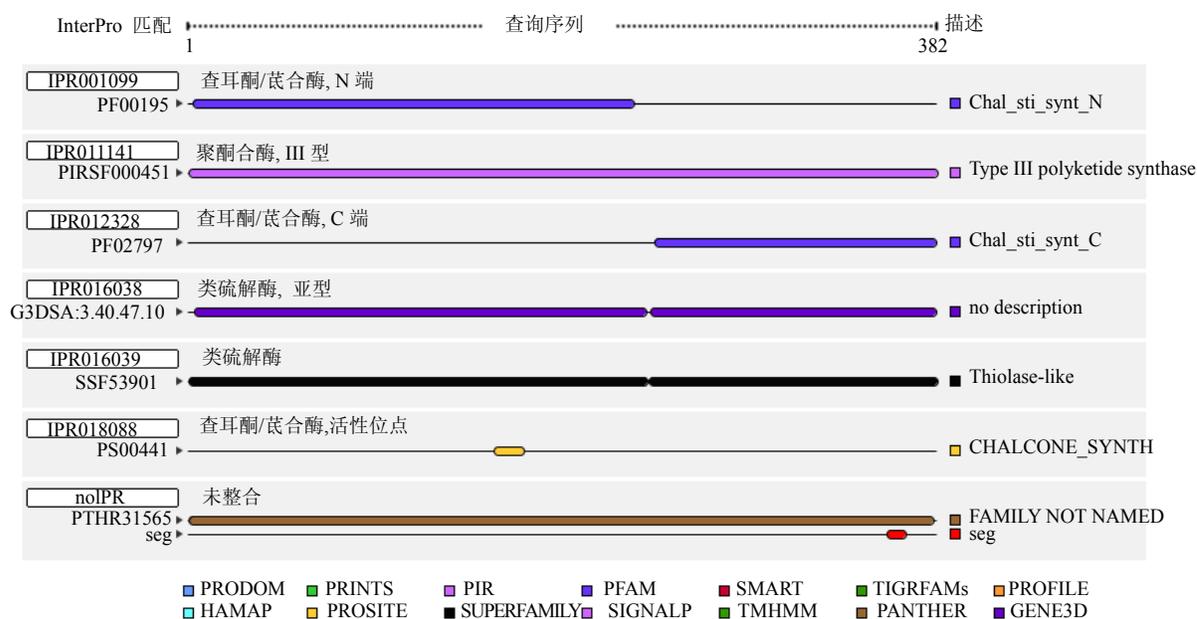


图4 LyCHS 含有的结构域

Fig. 4 Domains of LyCHS

用磷酸化位点预测程序 NetPhos 2.0, 发现 LyCHS 有 10 个潜在的 Ser 磷酸化位点、7 个 Thr 磷酸化位点、2 个 Tyr 磷酸化位点。使用 SUMO sp 2.0 软件对 LyCHS 蛋白序列进行泛素化修饰分析, 在蛋白序列中可能有两个泛素化位点 K49 (SEHKTEL, TypeII: Non-consensus) 和 K320 (LALKPEK, TypeI: Ψ-K-X-E)。

3.6 LyCHS 片段的二级结构、三级结构预测

经 SOPMA 对 LyCHS 蛋白片段二级结构进行预测, 发现蛋白中以 α 螺旋和无规则卷曲为主, β 折叠和延伸链较少。α 螺旋共 166 个, 占总二级结构的 43.46%; 延伸链共 65 个, 占总二级结构的 17.02%; β 折叠共 25 个, 占总二级结构的 6.54%;

无规则卷曲 126 个, 占总二级结构的 32.98%。

利用 SWISS-MODEL Workspace 在线分析工具对 LyCHS 蛋白三级结构进行预测。三级结构预测采用同源建模方法, 搜索蛋白质结构数据库 (PDB) 得到同源序列 1bi5A (0.156 nm) 作为宁夏枸杞 LyCHS 蛋白三级结构预测模板, 依据模板调整待预测蛋白序列主链原子位置生成预测蛋白三级结构模型 (图 5), 最后依据能量最小化原理使侧链基团处于能量最小的位置。预测信息显示, 预测模型与同源模板序列同源性为 84.03%。由于 CHS 是一种同型二聚体蛋白质, 因此预测结果是其亚基结构。

3.7 LyCHS 编码基因的系统进化分析

在 NCBI 上共选取茄科的 14 个物种 CHS 基因

序列, 以银杏 *Ginkgo biloba* CHS 基因作为外群 (outgroup), 利用 MEGA 5.10 软件, 采用 NJ 法构建系统发育树, 结果见图 6。结果显示 LyCHS 基因在进化上与马铃薯野生种 *Solanum pinnatisectum* Dunal 的 CHS、马铃薯 *S. tuberosum* 的 CHS 1a 和 1b、番茄 *Lycopersicon esculentum* 中的 CHS1 亲缘关系较近。

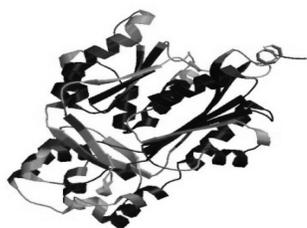


图 5 预测的宁夏枸杞 LyCHS 蛋白三级结构
Fig. 5 Predicted tertiary structure of LyCHS protein in *L. barbarum*

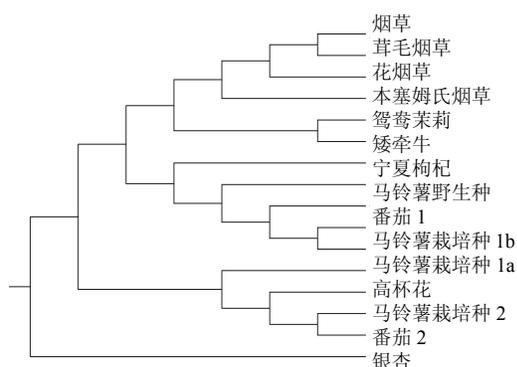


图 6 不同种植物 CHS 基因的系统进化树

Fig. 6 Phylogenetic tree of CHS genes from different plant species

4 讨论

查耳酮合成酶是类黄酮生物合成的关键酶和限速酶, 是一种同型二聚体蛋白质。它通过逐步缩合 3 分子丙二酰 CoA 和 1 分子的 4-对香豆酰-CoA 形成类黄酮的骨架柚皮素查耳酮。柚皮素查耳酮进一步派生出各种类黄酮物质^[36]。类黄酮在植物花色、生殖性能、防御反应中都起到重要作用。

本研究利用 RT-PCR 技术, 首次从宁夏枸杞中克隆到 CHS 基因家族 cDNA 片段, 命名 LyCHS。该序列长 1 148 bp, 编码 382 个氨基酸, 已在 GenBank 上注册, 基因序列登录号为 JQ964237。该基因编码的氨基酸序列分析显示, 在 156~172 处有一段高度保守的结构功能域 (RLMMYQQGCFAG-GTVLR), 这是查耳酮/二苯乙烯合成酶的特征序

列。二级结构预测中发现, 宁夏枸杞的 CHS 蛋白中含有大量的 α 螺旋和无规则卷曲, β 折叠和延伸链较少, 无明显跨膜结构, 蛋白含有无序化区域, 很可能定位在细胞质。在系统发生树中发现, 宁夏枸杞的 CHS 基因与马铃薯野生种、马铃薯、番茄的部分 CHS 基因亲缘关系较近。CHS 基因的克隆及序列分析将有助于深入了解 CHS 在宁夏枸杞类黄酮生物合成代谢中的作用, 更好调控类黄酮在枸杞中的表达, 包括表达的时间及部位。

参考文献

- [1] Huang Y, Tan A, Shen Y, *et al.* Scavenging effect of total flavonoids of *Lycium barbarum* L. on active oxygen radicals and inhibitory effects on heat output from L1210 cells [J]. *J Hyg Res*, 1998, 27(2): 109-111.
- [2] Huang Y, Lu J, Shen Y. The protective effects of total flavonoids from *Lycium barbarum* L. on lipid peroxidation of liver mitochondria and red blood cell in rats [J]. *J Hyg Res*, 1999, 28(2): 115-116.
- [3] Dong J Z, Lu D Y, Wang Y. Analysis of flavonoids from leaves of cultivated *Lycium barbarum* L. [J]. *Plant Foods Hum Nutr*, 2009, 64(3): 199-204.
- [4] Yamamoto Y, Gaynor R B. Therapeutic potential of inhibition of the NF- κ B pathway in the treatment of inflammation and cancer [J]. *J Clin Invest*, 2001, 107(2): 135-142.
- [5] Cushnie T P, Lamb A J. Recent advances in understanding the antibacterial properties of flavonoids [J]. *Int J Antimicrob Agents*, 2011, 38(2): 99-107.
- [6] Schuier M, Sies H, Illek B, *et al.* Cocoa-related flavonoids inhibit CFTR-mediated chloride transport across T84 human colon epithelia [J]. *J Nutr*, 2005, 135(10): 2320-2325.
- [7] Ruela de Sousa R R, Queiroz K C S, Souza A C S, *et al.* Phosphoprotein levels, MAPK activities and NF- κ B expression are affected by fisetin [J]. *J Enzyme Inhib Med Chem*, 2007, 22(4): 439-444.
- [8] Cushnie T P, Lamb A J. Antimicrobial activity of flavonoids [J]. *Int J Antimicrob Agents*, 2005, 26(5): 343-356.
- [9] 史玉俊. 千根草中的抗病毒类黄酮 [J]. *中草药*, 2000, 30(10): 88.
- [10] 徐汝明. 山金车属植物中类黄酮的细胞毒活性 [J]. *中草药*, 1996, 37(5): 315.
- [11] 张敏. 苦参中异戊二烯型类黄酮对磷脂酶 C γ 1 的抑制作用 [J]. *中草药*, 1998, 39(2): 137.
- [12] Cain C C, Saslowsky D E, Walker R A, *et al.* Expression

- of chalcone synthase and chalcone isomerase proteins in Arabidopsis seedlings [J]. *Plant Mol Biol*, 1997, 35(3): 377-381.
- [13] Crosby K C, Pietraszewska-Bogiel A, Gadella T W J, *et al.* Förster resonance energy transfer demonstrates a flavonoid metabolon in living plant cells that displays competitive interactions between enzymes [J]. *FEBS Lett*, 585(14): 2193-2198.
- [14] Austin M B, Noel J P. The chalcone synthase superfamily of type III polyketide synthases [J]. *Nat Prod Rep*, 2003, 20(1): 79-110.
- [15] Hrazdina G, Jensen R A. Spatial organization of enzymes in plant metabolic pathways [J]. *Annu Rev Plant Biol*, 1992, 43(1): 241-267.
- [16] Saslowsky D, Winkel-Shirley B. Localization of flavonoid enzymes in Arabidopsis roots [J]. *Plant J*, 2001, 27(1): 37-48.
- [17] Han Y Y, Ming F, Wang J W, *et al.* Cloning and characterization of a novel chalcone synthase gene from *Phalaenopsis hybrida* orchid flowers [J]. *Russ J Plant Physiol*, 2006, 53(2): 223-230.
- [18] Jeon J H, Kim H S, Choi K H, *et al.* Cloning and characterization of one member of the chalcone synthase gene family from *Solanum tuberosum* L. [J]. *Biosci Biotechnol Biochem*, 1996, 60(11): 1907-1910.
- [19] Claudot A C, Ernst D, Sandermann Jr H, *et al.* Cloning and characterization of two members of the chalcone synthase gene family from walnut [J]. *Plant Physiol Biochem*, 1999, 37(10): 721-730.
- [20] Narozna D, Pas J, Schneider J, *et al.* Two sequences encoding chalcone synthase in yellow lupin (*Lupinus luteus* L.) may have evolved by gene duplication [J]. *Cell Mol Biol Lett*, 2004, 9(1): 95-105.
- [21] Pang Y, Shen G, Wu W, *et al.* Characterization and expression of chalcone synthase gene from *Ginkgo biloba* [J]. *Plant Sci*, 2005, 168(6): 1525-1531.
- [22] Jiang C, Schommer C K, Kim S Y, *et al.* Cloning and characterization of chalcone synthase from the moss, *Physcomitrella patens* [J]. *Phytochemistry*, 2006, 67(23): 2531-2540.
- [23] Feinbaum R, Ausubel F. Transcriptional regulation of the *Arabidopsis thaliana* chalcone synthase gene [J]. *Mol Cell Biol*, 1988, 8(5): 1985-1992.
- [24] Ferrer J L, Jez J M, Bowman M E, *et al.* Structure of chalcone synthase and the molecular basis of plant polyketide biosynthesis [J]. *Nat Struct Biol*, 1999, 6: 775-784.
- [25] Bhasin M, Raghava G P. ESLpred: SVM-based method for subcellular localization of eukaryotic proteins using dipeptide composition and PSI-BLAST [J]. *Nucl Acids Res*, 2004, 32(Suppl 2): W414-W419.
- [26] Petersen T N, Brunak S, von Heijne G, *et al.* SignalP 4.0: discriminating signal peptides from transmembrane regions [J]. *Nat Method*, 2011, 8(10): 785-786.
- [27] Prilusky J, Felder C E, Zeev-Ben-Mordehai T, *et al.* FoldIndex[®]: a simple tool to predict whether a given protein sequence is intrinsically unfolded [J]. *Bioinformatics*, 2005, 21(16): 3435-3438.
- [28] Ren J, Gao X, Jin C, *et al.* Systematic study of protein sumoylation: Development of a site-specific predictor of SUMOsp 2.0 [J]. *Proteomics*, 2009, 9(12): 3409-3412.
- [29] Hunter S, Jones P, Mitchell A, *et al.* InterPro in 2011: new developments in the family and domain prediction database [J]. *Nucl Acids Res*, 2012, 40(1): D306-D312.
- [30] Blom N, Gammeltoft S, Brunak S, *et al.* Sequence and structure-based prediction of eukaryotic protein phosphorylation sites [J]. *J Mol Biol*, 1999, 294(5): 1351-1362.
- [31] Geourjon C, Deleage G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments [J]. *Comput Appl Biosci*, 1995, 11(6): 681-684.
- [32] Schwede T, Kopp J, Guex N, *et al.* SWISS-MODEL: an automated protein homology-modeling server [J]. *Nucl Acids Res*, 2003, 31(13): 3381-3385.
- [33] Arnold K, Bordoli L, Kopp J, *et al.* The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling [J]. *Bioinformatics*, 2006, 22(2): 195-201.
- [34] Guex N, Peitsch M C. SWISS-MODEL and the Swiss-Pdb viewer: An environment for comparative protein modeling [J]. *Electrophoresis*, 1997, 18(15): 2714-2723.
- [35] Abe I, Morita H M. Structure and function of the chalcone synthase superfamily of plant type III polyketide synthases [J]. *Nat Prod Rep*, 27(6): 809-838.
- [36] Koes R E, Quattrocchio F, Mol J N M. The flavonoid biosynthetic pathway in plants: function and evolution [J]. *Bioessays*, 1994, 16(2): 123-132.