

甘草中新 NADPH 细胞色素 P450 还原酶基因的克隆与相关生物信息学分析

孙宇峰¹, 厉 妲², 黄 颖¹, 杨晓涵¹, 刘春生^{1*}

1. 北京中医药大学中药学院, 北京 102488

2. 北京城市学院 生物医药学部, 北京 100094

摘要:目的 从甘草 *Glycyrrhiza uralensis* 中克隆 1 个新的 NADPH 细胞色素 P450 还原酶基因(GuCPR, 登录号: MH401048) 并进行生物信息学分析。方法 提取甘草根部总 RNA 后逆转录为 cDNA, 通过转录组数据库筛选得到 GuCPR 基因全长, 利用 NCBI ORF finder 得到其开放阅读框并翻译得氨基酸序列, 设计引物进行 PCR 扩增, 构建克隆重组质粒。利用生物信息分析预测蛋白性质、结构等及模拟蛋白三级结构, 并构建同源系统进化树。结果 克隆得到 GuCPR 基因 cDNA 全长 2 118 bp, 编码 705 个氨基酸残基, 相对分子质量 78 450, 等电点 (pI) 5.19。GuCPR 蛋白为非分泌蛋白, 不具有信号识别功能。跨膜预测结果显示蛋白的第 44~64 位氨基酸为跨膜区, 同时, 亚细胞定位预测结果显示蛋白位于内质网上。结论 克隆得到一个新的 GuCPR, 并进行了生物信息学分析, 为进一步研究提供了理论基础。

关键词: 甘草; NADPH 细胞色素 P450 还原酶; 基因克隆; 生物信息学分析; PCR

中图分类号: R282.12 文献标志码: A 文章编号: 0253-2670(2019)07-1676-06

DOI: 10.7501/j.issn.0253-2670.2019.07.027

Cloning and related bioinformatics analysis of new cytochrome NADPH P450 reductase gene in *Glycyrrhiza uralensis*

SUN Yu-feng¹, LI Da², HUANG Ying¹, YANG Xiao-han¹, LIU Chun-sheng¹

1. College of Chinese Pharmacy, Beijing University of Chinese Medicine, Beijing 102488, China

2. School of Biology and Medicine, Beijing City University, Beijing 100094, China

Abstract: Objective To clone a new NADPH cytochrome P450 reductase gene (GuCPR) (Login number: MH401048) from *Glycyrrhiza uralensis* and do some bioinformatics analysis. **Methods** Total RNA was extracted from the roots of *G. uralensis* and then transcription reversed into cDNA. The GuCPR gene was obtained by screening the transcriptome database. The NCBI ORF finder was used to obtain its open reading frame and translated amino acid sequence, design primers for PCR amplification, construction of recombinant cloned plasmid. Bioinformatics analysis was used to predict the protein properties, structure and model the tertiary structure of the protein, and homologous phylogenetic tree construction was performed. **Results** cDNA of GuCPR gene was cloned to a total length of 2 118 bp and encoded 705 amino acid residues with a relative molecular mass of 78 450. Electricity (pI) 5.19. GuCPR protein was a non-secretory protein and did not have a signal recognition function. The results of transmembrane prediction showed that the amino acids 44—64 of the protein were transmembrane regions. Meanwhile, the subcellular localization prediction results showed that the protein was located on the endoplasmic reticulum. **Conclusion** A new GuCPR was cloned and its bioinformatics analysis laid the foundation for further research.

Key words: *Glycyrrhiza uralensis* Fisch.; NADPH-cytochrome P450 reductase; gene cloning; bioinformatics analysis; PCR

甘草是豆科植物甘草 *Glycyrrhiza uralensis* Fisch.、胀果甘草 *G. inflata* Bat. 或光果甘草 *G. glabra* L. 的干燥根和根茎, 具有补脾益气、清热

解毒等功效, 被广泛应用于临床配方, 素有“国老”之称。甘草酸是甘草中的主要活性成分, 具有广泛的药理活性^[1]。

细胞色素 P450 酶 (CYP450) 系是一个非常重

要的氧化酶系, 在内源化合物与外源化合物的代谢

收稿日期: 2018-09-09

基金项目: 国家自然科学基金资助项目 (81773838)

作者简介: 孙宇峰 (1994—), 女, 硕士研究生, 研究方向为药用植物与分子生药学。Tel: 18810620122 E-mail: sunyufeng0604@163.com

*通信作者 刘春生, 教授, 博士生导师, 研究方向为药用植物与分子生药学。Tel: (010)84738624 E-mail: max_liucs@263.net

作用,包括动物、植物、细菌等有氧机体。CYPP450 酶系参与三萜类化合物次生代谢途径已有大量报道,如大豆中 CYP93E1 参与大豆皂醇 B 和齐墩果-12-烯-3 β ,24-二醇合成^[2],燕麦中 CYP51H10 参与燕麦根皂苷合成^[3]。甘草酸属于三萜皂苷类化合物,其生物合成途径中大多数关键酶已被表征,包括 β -香树脂醇合成酶^[4],2 个细胞色素 P450 酶(CYP88D6, CYP72A154)^[5-6],催化甘草次酸生成甘草酸最后反应的糖基转移酶已由本课题组成功表征^[7-8]。NADPH-CYPP450 还原酶(EC 1.6.2.4, NADPH-cytochrome P450 reductase, CPR)是细胞色素酶系的重要组成部分。CPR 属于膜结合蛋白,由 CPR 通过 2 个辅基 FMN 与 FAD 将电子传递给 P450s 后, CYPP450 才能与底物结合,发生下一步的氧化反应,参与内外源物质的代谢^[9]。在真核生物中, CPR 是所有微粒体的唯一电子供体,是 P450 单加氧酶反应过程中的限速因子^[10-12]。虽然此前已报道了一种甘草 CPR (GuCPR) 基因的克隆及其在酵母中的功能表达^[13],但是挖掘 CPR 家族新基因依然重要,以实现甘草酸的多通路合成。

本研究整合转录组学和生物信息学,通过 PCR 进行 cDNA 末端快速克隆(rapid-amplification of cDNA ends, RACE)技术和克隆重组质粒构建技术,成功地从甘草根中克隆了一条尚未报道的甘草细胞色素 P450 还原酶基因(Cytochrome P450 reductase, GuCPR)(登录号:MH401048),并结合相关的生物信息学分析,对 GuCPR 蛋白进行了预测和理化性质分析,为下一步的原核表达与功能验证提供参考,亦为甘草植物的分子育种、定向培育及活性成分的合成生物学研究奠定基础。

1 材料与试剂

1.1 材料

甘草种子由本课题组保存,经北京中医药大学中药鉴定系主任刘春生鉴定为豆根植物甘草 *Glycyrrhiza uralensis* Fisch. 种子。

1.2 试剂

HS DNA Polymerase 购于 TaKaRa 公司; FastQuant cDNA 第一链合成试剂盒购于天根生化科技(北京)有限公司;pEASY-Blunt Zero Cloning Kit、Trans1-T1 E. coli 感受态细胞、EasyPure Plasmid MiniPrep Kit 购于北京全式金生物技术有限公司; EASYspin 通用植物 RNA 快速提取试剂盒、琼脂糖凝胶纯化回收试剂盒、DNA Marker (BM-2000、

BM-5000) 购于北京博迈德基因技术有限公司; LB 培养基所需试剂(Tryptone、Yeast Extract、NaCl、琼脂粉)、浓硫酸购于北京拜尔迪生物技术有限公司;引物合成及测序均由生工生物工程公司完成。

2 方法

2.1 GuCPR 基因的克隆

取少量甘草种子,浓硫酸浸泡 2 h 后用清水反复冲洗数十次,清水浸泡过夜后,取出种子轻轻摆放于基质土壤上,上层轻轻覆土,2~3 d 浇水 1 次,3 周后取小苗根部提取总 RNA,并逆转录成 cDNA。以 cDNA 为模板,使用 HS DNA Polymerase 进行 PCR 扩增。本课题组前期对甘草进行的转录组深度测序数据,获得多条细胞色素 P450 还原酶家族基因,通过数据分析筛选得到的基因编码区全长序列,运用斯坦福大学在线引物设计系统(<https://www.yeastgenome.org/primer3>)设计特异性引物:g1F 5'-ATGCAGGATTCAAACCTCCATGA-AG-3'; g1R 5'-TCACCATACATCACGCAAATACCTGCC-3'。扩增程序如下:98 °C 预变性 3 min; 98 °C 变性 30 s; 55 °C 退火 5 s; 72 °C 延伸 2 min; 35 个循环; 72 °C 延伸 8 min。用 1% 琼脂糖凝胶电泳检测扩增结果,切下目的片段进行回收纯化。PCR 产物经回收纯化后,连接至 pEASY-Blunt Zero Cloning Kit 构建 GuCPR-pEASY 克隆重组质粒,并转化 Trans1-T1 E. coli 感受态细胞(北京全式金有限公司)。挑取单克隆进行菌落 PCR 鉴定,阳性菌株进行测序。

2.2 GuCPR 基因的生物信息学分析

采用 NCBI 网站 ORF finder 页面找到转录组中基因序列的开放阅读框(ORF)并翻译为氨基酸序列;双向测序结果采用 Conting Express 软件进行拼接,得到 GuCPR 基因全长序列;使用 Protparam 工具(<http://www.expasy.org/tools/protparam.html>)预测基因编码蛋白的理化性质;使用 TMHMM(http://www.ch.embnet.org/software/TMPRED_form.html)与 TMpred 工具(<http://www.cbs.dtu.dk/services/TMHMM/>)对蛋白进行跨膜预测和跨膜区位置计算及分值;使用 ProtScale 工具(<http://ca.expasy.org/tools/protscale.html>)进行蛋白亲疏水性分析。使用 SignalP 4.1Server(<http://www.cbs.dtu.dk/services/SignalP/>)进行蛋白信号肽预测。使用 NCBI(<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>)数据库分析蛋白的结构域。使用 ProtCompv.9.0(<http://>

linux1.softberry.com/berry.phtml?topic=protcompan&group=programs&subgroup=proloc) 进行蛋白亚细胞定位预测。使用 SWISS-MODEL (<http://swissmodel.expasy.org/repository/>) 进行蛋白质三级模拟。采用 MEGA 6.0 软件进行蛋白序列比对分析。

3 结果与分析

3.1 GuCPR 基因的克隆

RNA 提取结果如图 1 所示, 18 S 和 28 S 2 条带清晰明亮, 说明提取的 RNA 质量较好, 完整性较高。逆转录成 cDNA 后, 利用特异性引物进行 PCR 扩增, 结果如图 1 所示, 使用 Conting Express 拼接双向测序结果, 得到 2 118 bp 的基因全长 mRNA 序列, 编码 705 个氨基酸残基。将获得的 ORF 核苷酸序列和氨基酸序列在 NCBI 数据库进行 BLAST 比对, 结果显示, 氨基酸序列与豆科植物鹰嘴豆 *Cicer arietinum* L. 的细胞色素 P450 还原酶相似性最高, 达到 90%。

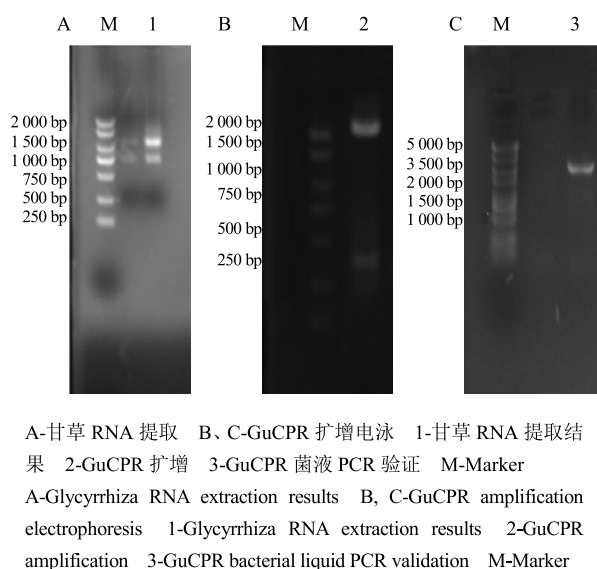


图 1 GuCPR 基因克隆电泳结果

Fig. 1 GuCPR gene cloning results

3.2 GuCPR 蛋白理化性质的预测

采用 protparam 在线分析工具 (<http://www.expasy.org/tools/protparam.html>) 对 GuCPR 蛋白的理化性质进行预测, 蛋白的相对分子质量 78 450, 编码 705 个氨基酸, 等电点 (pI) 为 5.19, 非极性氨基酸残基 (Asp+Glu) 102 个, 极性氨基酸残基 (Arg+Lys) 76 个, 脂溶系数 (Aliphatic index, AI) 为 82.87, 平均疏水性 (grand average of hydropathy, GRAVY) 为 -0.295 (图 2), 不稳定系数 (II) 为 45.05,

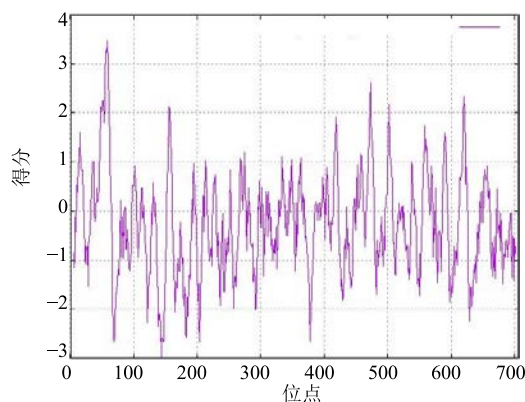


图 2 GuCPR 氨基酸亲水性/疏水性预测

Fig. 2 Predicted hydrophobicity/hydrophilicity of amino acid sequence of GuCPR

是一个不稳定的亲水性蛋白。

3.3 GuCPR 蛋白信号肽预测

在起始密码子后, 有一段编码疏水性氨基酸序列的 RNA 区域, 负责把蛋白质引导到细胞中含不同膜结构的亚细胞器内, 即为信号肽序列。本研究采用在线分析工具 SignalP 4.1 server 预测 GuCPR 的 N 端信号肽, 默认参数下, 结果显示 GuCPR 蛋白的 D 值为 0.132 (图 3), 小于设定阈值 0.500。因此, 预测 GuCPR 可能不含有信号肽, 并非为分泌蛋白且不具有信号识别功能。

3.4 GuCPR 蛋白结构域分析

使用 NCBI (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) 数据库分析 GuCPR 蛋白的结构域。结果显示 GuCPR 蛋白在 287~691 氨基酸序列存在 NADPH-cytochrome P450 reductase 保守结构域 (图 4), 属于细胞色素 P450 还原酶超家族。

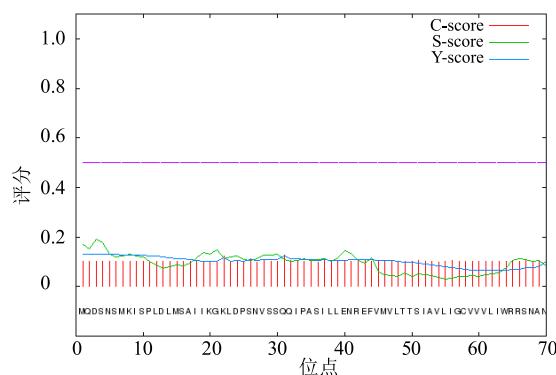


图 3 GuCPR 蛋白的信号肽预测

Fig. 3 Signal peptide prediction of GuCPR protein

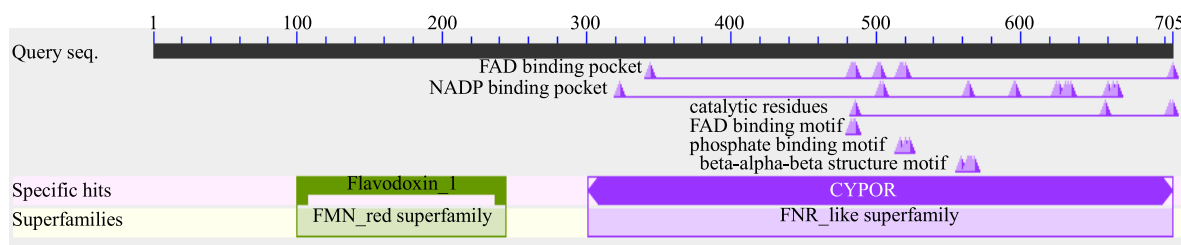


图 4 GuCPR 蛋白的结构域分析

Fig. 4 Domain analysis of GuCPR protein

3.5 GuCPR 蛋白跨膜区预测

采用 TMHMM Server v.2.0 进行 GuCPR 的跨膜区预测, 结果表明 GuCPR 有一段跨膜螺旋区, 即蛋白的第 44~64 位氨基酸。蛋白第 1~43 位氨基酸在膜内侧, 第 65~705 位氨基酸在膜外侧(图 5)。

3.6 GuCPR 蛋白亚细胞定位预测

采用 ProtCompv. 9.0 进行 GuCPR 的亚细胞定位预测, 结果表明 GuCPR 应定位于内质网上, 综合打分结果最高, 为 9.19 (图 6)。

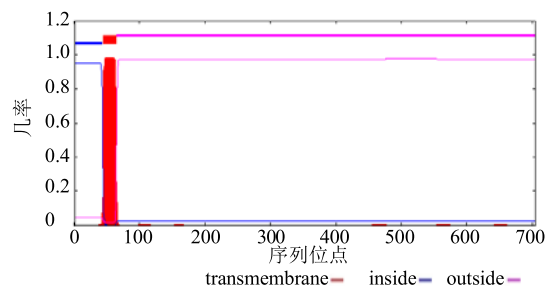


图 5 GuCPR 跨膜区预测

Fig. 5 Transmembrane helices prediction of GuCPR

Integral Prediction of protein location: Membrane bound Endoplasmic reticulum with score 9.2					
Location weights:	LocDB /	PotLocDB /	Neural Nets /	Pentamers /	Integral
Nuclear	0.0 /	0.0 /	0.14 /	0.11 /	0.01
Plasma membrane	0.0 /	0.0 /	0.00 /	0.00 /	0.80
Extracellular	0.0 /	0.0 /	0.00 /	0.00 /	0.00
Cytoplasmic	0.0 /	0.0 /	0.01 /	0.00 /	0.00
Mitochondrial	0.0 /	0.0 /	0.00 /	0.05 /	0.00
Endoplasm. retic.	10.0 /	3.0 /	0.11 /	2.92 /	9.19
Peroxisomal	0.0 /	0.0 /	0.05 /	0.00 /	0.00
Lysosomal	0.0 /	0.0 /	0.34 /	0.00 /	0.00
Golgi	0.0 /	0.0 /	0.00 /	0.25 /	0.00
Vacuolar	0.0 /	0.0 /	2.35 /	0.00 /	0.00

图 6 GuCPR 亚细胞定位预测

Fig. 6 Sub-cellular location prediction of GuCPR

3.7 GuCPR 蛋白三级结构预测

SWISS-MODEL (<http://swissmodel.expasy.org/repository/>) 是一个蛋白质 3D 结构数据库, 库中收录的蛋白质结构都是使用 SWISS-MODEL 同源建模方法 (homology-modelling) 得来的。利用 SWISS-MODEL 在线软件构建 GuCPR 蛋白的三维结构, 使用 2vg8.1.A 作为模板序列, 结果见图 7。

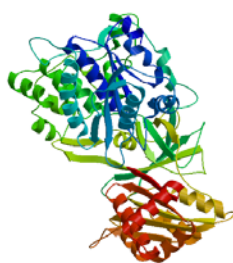


图 7 模拟 GuCPR 蛋白三级结构

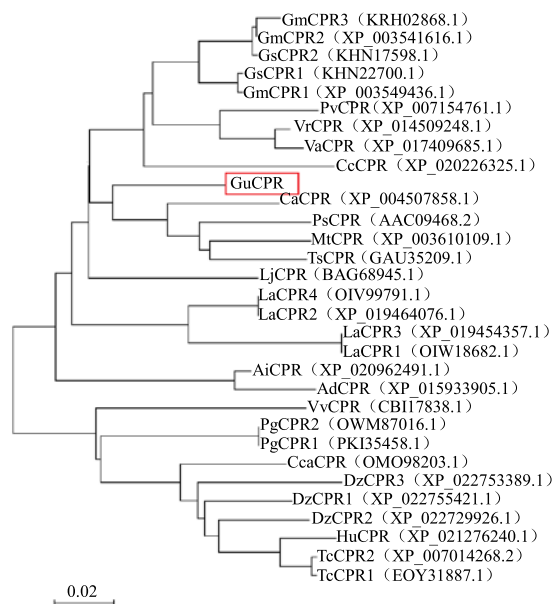
Fig. 7 Predicted three-dimensional structure of GuCPR

3.8 GuCPR 系统进化分析

通过 NCBI 的 blast 比对 GuCPR 蛋白序列, 按照相似度从高到低排序, 从 Genbank 中下载现有报道的 CPR 蛋白质序列, 采用 Clustal W 进行多重序列比对分析, 然后采用 MEGA 6.0 的 Neighbor-Joining (NJ) 算法构建系统发育树 (Bootstrap 1 000 次分析重复, 图 8)。结果显示 GuCPR 与鹰嘴豆 *Cicer arietinum* L. (CaCPR, XP_004507858.1)、豌豆 *Pisum sativum* L. (PsCPR, AAC09468.2)、地三叶 *Trifolium subterraneum* L. (TsCPR, GAU35209.1)、蒺藜 *Tribulus terrestris* L. (MtCPR, XP_003610109.1) 可聚为一支, 显示较近亲缘关系。

4 讨论

甘草是在中医药中使用最频繁的药材, 其主要成分有甘草酸、甘草次酸等三萜类化合物 20 余种,



Bootstrap 分析重复 1 000 次: 方框表示甘草 GuCPR 的系统进化位置

Bootstrap analysis repeated 1 000 times: box indicates phylogenetic position of licorice GuCPR

图 8 基于 NJ 算法的 GuCPR 系统发育树

Fig. 8 GuCPR phylogenetic tree based on NJ algorithm

甘草苷、异甘草苷、甘草素、异甘草素等黄酮类化合物 300 余种以及各种多糖类化合物, 具有抗炎、抗病毒等活性^[14-15], 目前甘草酸完全依赖于从甘草原植物中提取, 过度采挖造成甘草资源严重匮乏, 而甘草酸的合成生物学研究能大大降低甘草资源的消耗, 为甘草酸产业的发展提供充足的原料, 是甘草生态资源可持续利用的重要保障。本研究成功克隆了甘草酸生物合成途径中 CPR 基因, 并结合相关的生物信息学分析, 预测了 GuCPR 蛋白的理化性质, 结果显示 GuCPR 基因 cDNA 全长 2 118 bp, 编码 705 个氨基酸残基, 相对分子质量 78 450, pI 为 5.19。GuCPR 蛋白为分泌蛋白, 不具有信号识别功能。跨膜预测结果显示蛋白的第 44~64 位氨基酸为跨膜区, 同时, 亚细胞定位预测结果显示蛋白位于内质网上, 与一般报道的 CPR 蛋白的定位结果一致^[16]。

研究过程中, 使用浓硫酸对甘草种子进行预处理, 可增加种皮和胚乳的透气性, 以促进萌发, 缩短种子的催芽时间。本研究对 PCR 条件做了优化考察, 结果显示退火温度与退火时间对扩增结果有显著影响, 尝试分别使用 5 S 与 30 S, 55、58、62、65 °C 梯度退火, 结果显示退火温度 55 °C, 退火时

间 5 s 条件下扩增效果最好, 特异性最高。本研究使用零背景克隆载体代替传统 pMD19-T 克隆载体, 有效地缩短了实验时间并提高了阳性克隆菌斑数目, 提高了实验效率。GuCPR 为甘草酸生物合成途径中的关键基因, 本研究为进一步表征 GuCPR 的功能, 实现甘草酸多通路生物合成奠定前期基础。

综上所述, 本研究从甘草中克隆出 GuCPR 新基因 (登录号 MH401048), 并结合相关的生物信息学分析方法对 GuCPR 蛋白进行预测和性质分析, 为进一步研究该酶的活性与功能提供理论基础及参考。同时, GuCPR 在甘草酸的生物合成途径中发挥重要作用, 也为解析甘草酸的生物合成途径奠定基础。

参考文献

- [1] 中国药典 [S]. 一部. 2015.
- [2] Shibuya M, Hoshino M, Katsube Y, *et al.* Identification of β -amyrin and sophoradiol 24-hydroxylase by expressed sequence tag mining and functional expression assay [J]. *FEBS J*, 2006, 273(5): 948-959.
- [3] Qi X, Bakht S, Qin B, *et al.* A different function for a member of an ancient and highly conserved cytochrome P450 family: From essential sterols to plant defense [J]. *PNAS*, 2006, 103(49): 18848-18853.
- [4] 沈湛云, 刘春生, 王学勇. 甘草 β -香树脂醇合成酶的编码区克隆与序列分析 [J]. *中国中药杂志*, 2009, 34(19): 2438-2440.
- [5] Seki H, Ohyama K, Sawai S, *et al.* Licorice β -amyrin 11-oxidase, a cytochrome P450 with a key role in the biosynthesis of the triterpene sweetener glycyrrhizin [J]. *PNAS*, 2008, 105(37): 14204-14209.
- [6] Seki H, Sawai S, Ohyama K, *et al.* Triterpene functional genomics in licorice for identification of CYP72A154 involved in the biosynthesis of glycyrrhizin [J]. *Plant Cell*, 2011, 23(11): 4112-4123.
- [7] Xu G, Cai W, Gao W, *et al.* A novel glucuronosyltransferase has an unprecedented ability to catalyse continuous two-step glucuronosylation of glycyrrhetic acid to yield glycyrrhizin [J]. *New Phytol*, 2016, 212(1): 123-135.
- [8] 徐国杰. 甘草酸生物合成相关糖基转移酶的研究 [D]. 北京: 北京中医药大学, 2017.
- [9] Pandey A V, Fluck C E. NADPH P450 oxidoreductase: Structure, function, and pathology of diseases [J]. *Pharm Therap*, 2013, 138(2): 229-254.
- [10] Jensen K, Moller B L. Plant NADPH-cytochrome P450

- oxidoreductases [J]. *Phytochemistry*, 2010, 71(2/3): 132-141.
- [11] Riddick D S, Ding X X, Wolf C R, *et al.* NADPH-cytochrome P450 oxidoreductase: Roles in physiology, pharmacology, and toxicology [J]. *Drug Metab Dispos*, 2013, 41(1): 12-23.
- [12] 汪 丹. 棉铃虫 NADPH-细胞色素 P450 还原酶基因的克隆、时空表达及 RNA 干扰 [D]. 南京: 南京农业大学, 2011.
- [13] Zhu M, Wang C, Sun W, *et al.* Boosting 11-oxo-beta-amyrin and glycyrrhetic acid synthesis in *Saccharomyces cerevisiae* via pairing novel oxidation and reduction system from *leGume* plants [J]. *Metab Eng*, 2011, 45: 43-50
- [14] Fiore C, Eisenhut M, Krauss R, *et al.* Antiviral effects of *Glycyrrhiza* species [J]. *Phytother Res*, 2008, 22(2): 141-148.
- [15] Hayashi H, Sudo H. Economic importance of licorice [J]. *Plant Biotechnol*, 2009, 26(1): 101-104.
- [16] 何 海, 郭继云, 舒少华, 等. 茯苓细胞色素 P450 还原酶基因的克隆与生物信息学分析 [J]. 中草药, 2016, 47(16): 2909-2915.