

文章编号: 1001-1498(2009)02-0177-05

油桐种仁 cDNA 文库的构建及其油体蛋白 *oleosin* 基因的生物信息学分析

周冠^{1,2}, 汪阳东², 陈益存², 李鹏^{1,2}, 张姗姗^{1,2}, 张小平^{1*}

(1. 安徽师范大学生命科学学院, 安徽 芜湖 241000; 2. 中国林业科学研究院亚热带林业研究所, 浙江 富阳 311400)

摘要:以木本油料植物油桐为研究对象,通过预实验选择种子成熟过程中脂肪酸转变关键时期,构建油桐种仁 cDNA 文库,并进行部分测序。初级文库的滴度为 1×10^6 pfu · mL⁻¹,重组率为 99.7%,插入片段大小为 0.5 ~ 2.5 kb,平均约 1 kb。通过测序获得的功能基因类型主要包括抗性基因、油体蛋白基因、贮藏蛋白基因及种子发育相关基因等,还有大量的未知功能基因和其他基因存在,各类基因的表达丰度不一。着重对首次分离到的油桐油体蛋白 *oleosin* 基因序列进行了同源性和蛋白结构预测分析。

关键词:油桐;种仁;cDNA文库;*oleosin*;生物信息学

中图分类号: S794.3

文献标识码: A

Construction of Kernel cDNA Library and Bioinformatic Analysis on *oleosins* Gene of *Vernicia fordii*

ZHOU Guan^{1,2}, WANG Yang-dong², CHEN Yi-cun², LI Peng^{1,2}, ZHANG Shan-shan^{1,2}, ZHANG Xiao-ping¹

(1. College of Life Science, Anhui Normal University, Wuhu 241000, Anhui, China;

2. Research Institute of Subtropical Forestry, CAF, Fuyang 311400, Zhejiang, China)

Abstract: Seed kernel of *Vernicia fordii* was used as materials to construct a cDNA library at the key stage of fatty acids synthesis. A primary library of 1×10^6 pfu · mL⁻¹ was obtained. The recombinant efficiency was 99.7%, and the sizes of inserted fractions ranged from 0.5 to 2.5 kb, with an average value of 1 kb. These sequences mainly involved oil synthesis related genes, seed development genes with respective abundance. And there are also many other genes to be unknown. Furthermore, bioinformatic analysis on oil body protein *oleosins* gene from *V. fordii* was done, which was firstly obtained from the cDNA library.

Key words: *Vernicia fordii*; seed kernel; cDNA library; *oleosin*; bioinformatics

油桐 (*Vernicia fordii* (Hemsl.) Airy-Shaw), 又叫三年桐, 隶属大戟科 (Euphorbiaceae) 油桐属 (*Vernicia* Lour.), 是我国南方重要的木本油料经济树种。油桐干种仁含油率达 60% ~ 70%, 桐油含有的脂肪酸主要包括软脂酸、硬脂酸、油酸、亚油酸、亚麻酸、桐酸, 其中桐酸约占总脂肪酸含量的 80%, 是决定桐油性质的主要物质^[1]。桐油是最好的干性油之

一, 是环保型新型化工产品原料, 具有绝缘、耐酸碱、防腐防锈等优良性能, 可用来研制新型天然涂料、合成树脂、粘合剂、药品等, 广泛应用于化工制造业。近年研究还发现, 桐油具有抗肿瘤作用^[2]。我国是世界上最大的桐油生产国, 年产量达 10 万 t 以上, 占世界桐油产量的 80%。

油桐还可用来开发生物柴油原料, 与其他许多

收稿日期: 2008-08-25

基金项目: 中国林科院中央级公益性科研院所基本科研业务费专项重点项目 (060702) 资助

作者简介: 周冠 (1984—), 男, 安徽濉溪人, 在读硕士生, 主要从事植物生物技术研究。

* 通讯作者

生物柴油原料植物类似,桐油中多元脂肪酸含量高,制造生物柴油稳定性较低,残留较多,需要进一步进行遗传改良,这有赖于阐明油桐脂肪酸合成及油脂形成、转化、贮藏分子机制,从而有效调控油脂含量和品质,而目前国内外关于这方面的研究还比较欠缺。2006年,Shockey等^[3]报道了油桐两种类型的二酰甘油酰基转移酶(diacylglycerol acyltransferase, DGAT1和DGAT2)在三酰甘油合成中具有至关重要的调控作用。汪阳东等^[4-6]分离得到桐酸(18:3^{9cis,11trans,13trans})合成的关键酶基因FADX,并初步分析了其功能。为了获取更好的与脂肪酸、油脂相关的功能基因,阐述合成调控机理,从而通过对关键基因的调控进行定向遗传改良,本研究通过构建优质油桐种仁cDNA文库,为筛选获得油桐脂肪酸油脂合成调控相关基因及遗传育种、分子改良奠定基础。

1 材料与方法

1.1 材料

油桐(小米桐品种群,浙江丛生球桐)种子于8月下旬采自浙江省富阳市中国林科院亚热带林业研究所后山,去种皮后于液氮中速冻,-70℃保存备用。

1.2 试剂

Trizol购自Invitrogen公司,Absolutely mRNATM Purification Kit,cDNA Synthesis Kit,ZAP-cDNA Synthesis Kit和ZAP-cDNA Gigapack均购自Stratagene公司,CHROMA SPN-400 Column购自Clontech公司,其余试剂均为进口分析纯。

1.3 方法

1.3.1 总RNA及mRNA纯化 用Trizol提取油桐种仁总RNA。将样品在液氮中迅速研磨充分,取约0.1g样品加入盛有预冷的1mL Trizol的1.5mL离心管中,室温放置5min;4℃,12000r·min⁻¹离心5~10min;取上清,每1mL Trizol加300μL氯仿,剧烈振荡15s,室温静置3min;4℃,12000r·min⁻¹离心10~15min;取水相,加等体积氯仿重复抽提2次;取上清加等体积异丙醇混匀,室温放置10~20min;4℃,12000r·min⁻¹离心10min;弃上清,加1mL 75%乙醇洗涤沉淀,4℃,7000r·min⁻¹离心5min;弃上清,晾干后加30~100μL DEPC水溶解。用Absolutely mRNATM Purification Kit(磁珠法)进行mRNA分离纯化。

1.3.2 cDNA合成及文库构建 参照Stratagene公司的cDNA Synthesis Kit提供的方法:以5μg mRNA为模板合成cDNA,对cDNA末端进行平滑化,连接EcoRI Adapters,用XhoI酶切后的产物按照Clontech公司建议的方法过柱,收集过柱液约40μL,取3μL电泳检测,然后收集大于500bp的cDNA片段。将目的片段接入Uni-ZAP XR vector载体后,用Gigapack III Gold Packaging Extract体外包装,得到cDNA初级文库,同步用testRNA做对照实验。

1.3.3 cDNA文库的质量分析 参照Stratagene公司的cDNA Synthesis Kit提供的方法,对初级文库和扩增文库的滴度及重组率进行测定。设计一对引物:T7(5'-GTAAACGACTCACTATAGG-3')和SK(5'-AATTAACCTCACTAAAG-3')对cDNA克隆进行PCR扩增,检测插入片段大小,挑取单个噬菌斑环化后提取质粒DNA,EcoRI和XhoI双酶切,用1%琼脂糖凝胶电泳检测PCR结果;随机挑选阳性克隆进行测序。

1.3.4 序列的生物信息学分析 将得到的基因序列在NCBI上进行BLAST;用GenBank等数据库中的在线软件对一些重要功能基因的蛋白质结构进行预测和分析。

2 结果与分析

2.1 总RNA的提取及mRNA的纯化

经过对不同时期的油桐种仁脂肪酸含量的测定,选定油桐脂肪酸转变关键时期(8月下旬)的油桐种仁提取总RNA,但按常规方法提取的总RNA得率较低、质量较差,故在用Trizol提取时增加了氯仿抽提次数,得到的总RNA利用1%琼脂糖凝胶电泳可检测到28S RNA、18S RNA和5S RNA 3条带,且28S RNA与18S RNA的亮度比例约为2:1,5S RNA条带较弱(图1),说明总RNA没有降解,较完整,完全符合构建高质量cDNA文库的要求。用分光光度计检测OD₂₆₀/OD₂₈₀比值在1.9~2.0之间。在用磁珠法分离纯化mRNA时,将总RNA浓缩后再过磁珠得到的mRNA质量较好。

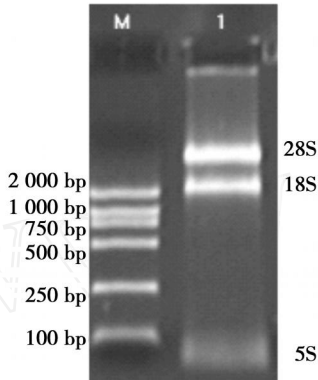
2.2 ds cDNA合成及分级分离

mRNA反转录合成的双链cDNA在凝胶电泳上检测,得知其分子大小范围在0.3~4.0kb之间(图2),基因丰度满足试验研究需要;为确定cDNA的大小,将分级分离后的cDNA片段通过琼脂糖凝胶电泳,收集片段长度在500bp以上的过柱液并混合到

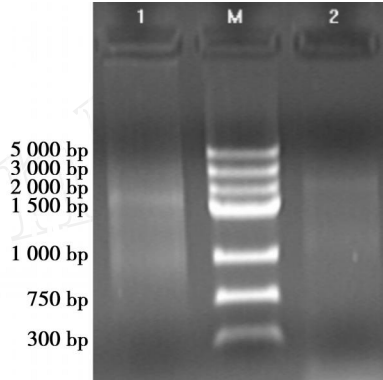
同一管中。

由于按照 Stratagene公司的 ZAP-cDNA Gigapack III Gold Cloning Kit使用说明,在分级分离时,需要进行装柱,时间比较长,且容易产生气泡,故改选用

Clnitech公司的 Creator™ SMART™ cDNA Library Construction Kit中分级分离部分的 CHROMA SPN-400 Column,操作简单且时间短,且能得到较好的分级分离效果。



1:油桐种仁总 RNA;M: DNA Marker
图 1 油桐种仁总 RNA

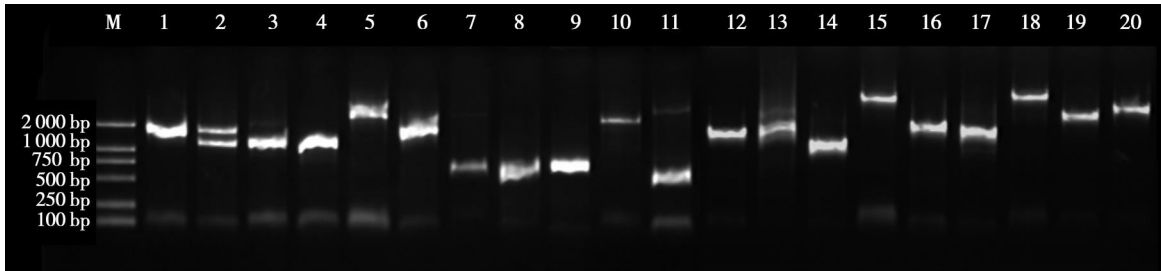


1:单链 cDNA; 2:双链 cDNA;M: DNA Marker
图 2 单链、双链 cDNA的大小范围

2.3 cDNA文库质量的检测

按照 Stratagene公司的 ZAP-cDNA Gigapack III Gold Cloning Kit检测,所建立的油桐种仁 cDNA 初级文库的滴度为 1×10^6 pfu · mL⁻¹,重组率为 99.7%,扩增后文库的滴度为 1.2×10^9 pfu · mL⁻¹,随机挑选

20个阳性克隆进行 PCR 检测,琼脂糖凝胶电泳检测结果表明:插入片段大小在 0.5~2.5 kb之间,平均约 1 kb(图 3),对环化后质粒 DNA 进行 *EcoR* I和 *Xho* I 双酶切,检测结果与 PCR 一致。文库的滴度、重组率及完整性均符合构建高质量文库标准。



1~20: cDNA文库随机插入片段 PCR 检测
图 3 cDNA文库随机插入片段 PCR 检测

2.4 cDNA文库测序与分析

将文库进行大量测序,共测得 3 000 条序列,并向 NCBI提交了 10 个基因序列(序列号为: FJ362588~FJ362597),对文库初步分析显示,与拟南芥 (*Arabidopsis thaliana* L.)、毛果杨 (*Populus trichocarpa* Torr)、麻疯树 (*Jatropha curcas* L.)、蓖麻 (*Ricinus communis* L.)等一些物种的同源性较高。挑选其中 93 条原始序列做了详细的分析,去除载体后,用 Phrap 软件对序列进行 cluster 分析得到重叠群(contig),然后将 contig 和 siglet 序列在 NCBI 上进行 BLASTX 分析,初步预测功能基因的类型及分布的结果见表 1,包含有抗性基因、油体蛋白基因、贮藏蛋白基因等,另外还有大量未知功能基因及其他

基因在种仁中表达^[7-8]。

表 1 油桐种仁中部分表达基因分类

基因类别	条数
油体蛋白基因	3
氨基酸合成基因	2
贮藏蛋白基因	12
抗性基因	6
基因表达调控基因	5
种子成熟和胚胎发育相关基因	7
蛋白质降解基因	5
信号转导基因	4
未知功能基因	11
其他基因	56

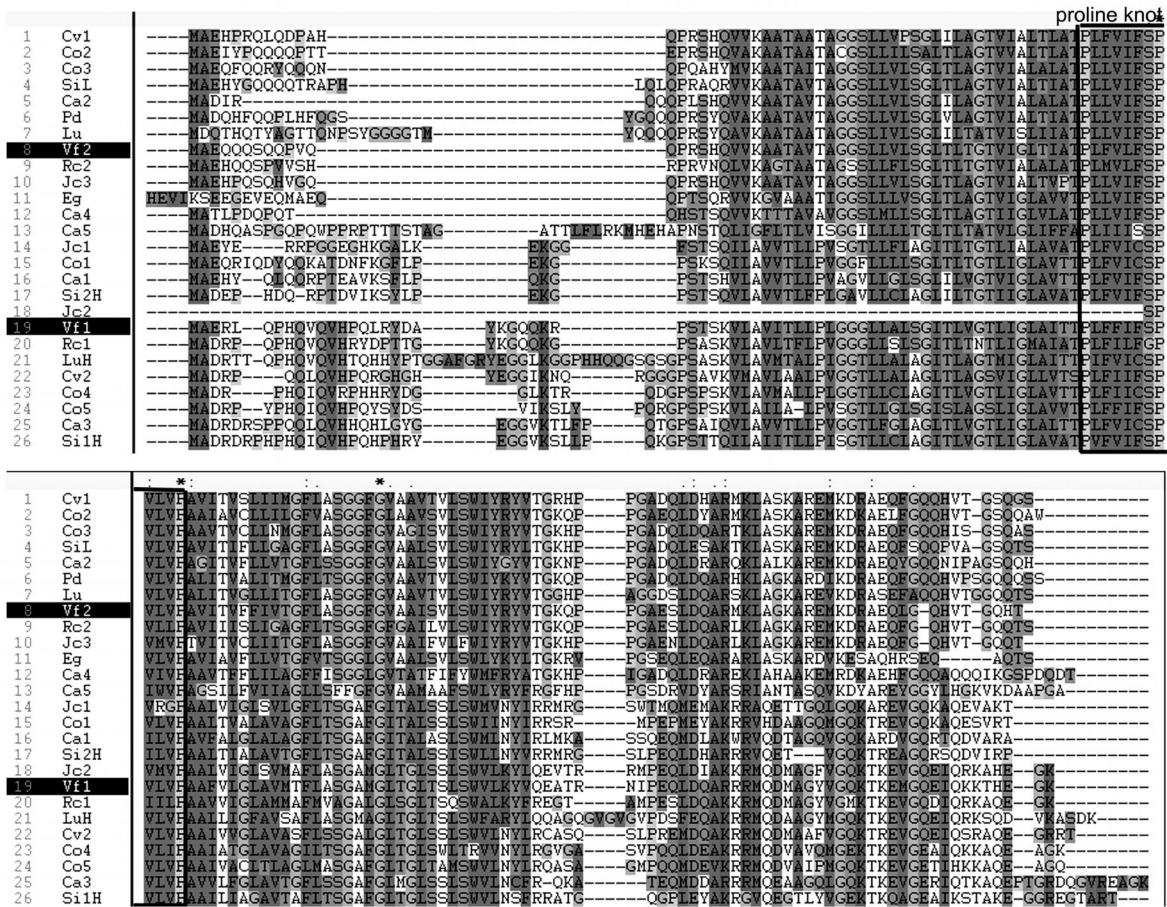
2.5 油体蛋白生物信息学分析

oleosin 是一类特殊贮藏蛋白,主要在植物种子特异表达,覆盖于油体表面,在油体发生到分解消失过程中起着重要的生物学作用,在植物基因工程研究中有重要的应用价值。目前,oleosin 基因已在很多作物中分离获得,包括油菜(Brassica campestris L.)、玉米(Zea mays L.)、大豆(Glycine max (L.) Merrill)、拟南芥、向日葵(Helianthus annuus L.)、棉花(Gossypium hirsutum L.)、芝麻(Sesamum indicum L.)、水稻(Oryza Sativa L.)、小麦(Triticum aestivum L.)和木本油料植物油茶(Camellia oleifera Abel)、麻疯树[9-10]。

作者测序得到油桐 2 个编码油体蛋白 oleosin 的转录本序列,根据其编码氨基酸的大小可分为低分子量 oleosin 和高分子量 oleosin。与其他油料植物 oleosin 蛋白的同源性分析,结果显示与蓖麻、麻风树的序

列同源性最高(图 4)。

生物信息学关于蛋白一级结构和二级结构的预测表明,油桐低分子量 oleosin (L-OLE)序列编码 154 个氨基酸,相对分子量是 16.188 kD,理论等电点是 9.93;二级结构预测表明,该蛋白含有 7 个螺旋,5 个折叠。应用 TMHMM 预测模型预测跨膜螺旋及拓扑结构,该序列存在 3 个跨膜螺旋,分别位于 34~56 aa, 58~80 aa, 85~107 aa,与在线工具(http://www.expasy.org/cgi-bin/protscale.pl)上预测的疏水区相一致。油桐高分子量 oleosin (H-OLE)序列编码 137 个氨基酸,相对分子量是 14.443 kD,理论等电点是 9.87;二级结构预测表明,该蛋白含有 3 个螺旋,5 个折叠。应用 TMHMM 预测模型预测跨膜螺旋及拓扑结构,该序列存在 2 个跨膜螺旋,分别位于 32~54 aa, 69~91 aa,此结果和预测的



Jc: 麻疯树 (Jc1, Jatropha curcas OLE1; Jc2, Jatropha curcas OLE2; Jc3, Jatropha curcas OLE3); Vf: 油桐 (Vf1, Vemicial fordii H-OLE; Vf2, Vemicial fordii L-OLE); Rc: 蓖麻 (Rc1, Ricinus communis OLE1; Rc2, Ricinus communis OLE2); Cv: 榛木 (Cv1, Corylus avellana OLE; Cv2, Corylus avellana OLE); Ca: 咖啡 (Ca1, Coffea Arabica OLE1; Ca2, Coffea canephora OLE2; Ca3, Coffea canephora OLE3; Ca4, Coffea canephora OLE4; Ca5, Coffea canephora OLE5); Si: 芝麻 (Si1H, Sesamum indicum; Si2H, Sesamum indicum OLE); SL, Sesamum indicum OLE); Co: 油茶 (Co1, Camellia oleifera Ole ; Co2, Camellia oleifera Ole ; Co3, Camellia oleifera Ole ; Co4, Camellia oleifera Ole ; Co5, Camellia oleifera Ole); Lu: 亚麻 (Lu, Linum usitatissimum L-OLE; LuH, Linum usitatissimum H-OLE); Eg: 油棕 (Eg, Elaeis guineensis OLE); Pd: 杏仁 (Pd, Prunus Dulcis OLE)

图 4 油料植物中 oleosin 蛋白同源性分析 (其中油桐 oleosin 序列与蓖麻、麻疯树的序列同源性最高)

含有两性 N 末端区、中心疏水区 (32 ~ 89 aa) 和两性 C 末端区一致。

利用 SMART 对 *oleosin* 功能结构域的分析显示, *oleosin* 蛋白中间的疏水区域很保守, 由 2 个反向平行的 折叠和 1 个脯氨酸结组成, 脯氨酸结 “proline knot” 为 -PX₅SPX₃P- (图 4), 该结构域对 *oleosin* 靶向结合油体起着重要作用^[11]。*oleosin* 通过决定油体大小而在油脂调控中发挥重要作用^[12]。

3 讨论

植物基因组的研究已经由以全基因组测序为目标的基因组学转向功能基因组学研究。近年来, 通过 cDNA 文库筛选获得功能基因成为功能基因组学研究的重要方法之一。cDNA 文库代表了 mRNA 的反转录复本^[13-15], 其质量高低是功能基因筛选成功与否的关键。由于油桐种子含大量的多糖、多酚及一些成分复杂的次生代谢物质, 按常规方法提取 RNA 得率低、质量差, 本试验进行了改进, 增加氯仿抽提次数, 得到较完整的总 RNA。

本试验构建的文库插入片段大小在 0.5 ~ 2.5 kb 之间, 平均约为 1 kb。文库的滴度、重组率及完整性都符合鉴定文库质量的重要标准^[16]。部分测序结果表明, 油桐种子发育过程中相关的功能基因主要有油体蛋白基因、抗性基因、贮藏蛋白基因、种子发育相关基因等, 另外还有大量的未确定功能的基因, 其中, 作者通过文库首次分离到油桐油体蛋白 *oleosin* 两条异构体序列, 并进行了生物信息学分析, 预测 *oleosin* 对油体结构的稳定及决定油脂含量高低都起着重要作用。在此基础上, 作者正在进一步利用试验分析其功能, *oleosin* 的功能分析对于调控桐油合成具有关键性作用。

油桐作为原产于我国的重要木本经济油料植物, 具有生长快、适生能力强的特征。和油料作物如油菜、大豆等相比较, 国内外对木本油料植物的开发利用还处于研究的起步阶段。油桐种仁高质量 cDNA 文库的建立, 为后期通过大规模测序、芯片等技术分离获得大量与油脂合成、调控和贮藏相关的基因奠定了基础, 从而为深入阐述木本油料植物油脂合成调控分子机制提供了平台。同时, 油桐种仁 cDNA 文库的建立有助于了解木本油料植物种子发

育过程。

参考文献:

- [1] 方嘉兴, 何方. 中国油桐 [M]. 北京: 中国林业出版社, 1998: 333 - 334
- [2] Miki I, Teruo M. Newly recognized cytotoxic effect of conjugated trienoic fatty acids on cultured human tumor cells [J]. *Cancer Letters*, 2000, 148: 173 - 179
- [3] Shockey J M, Gidda S K, Chapital D C, et al. Tung tree DGAT1 and DGAT2 have nonredundant functions in triacylglycerol biosynthesis and are localized to different subdomains of the endoplasmic reticulum [J]. *Plant Cell*, 2006, 18: 2294 - 2313
- [4] 汪阳东, 李春秀, 齐力旺, 等. 中间锦鸡儿 *fad2* 基因片段克隆、反义表达载体构建及遗传转化初步研究 [J]. *林业科学研究*, 2007, 20(1): 6 - 9
- [5] 李元, 汪阳东, 李鹏, 等. 油桐种子 *FAD3* 基因的克隆和序列分析 [J]. *安徽农业科学*, 2008, 36(11): 4753 - 4755
- [6] 李鹏, 汪阳东, 陈益存, 等. 油桐 ISSR-PCR 最佳反应体系的建立 [J]. *林业科学研究*, 2008, 21(2): 194 - 199
- [7] White J, Todd J, Newman T, et al. A new set of Arabidopsis expressed sequence tags from developing seeds [J]. *Plant Physiology*, 2000, 124: 1582 - 1594
- [8] 谭晓风, 胡芳名, 谢禄山, 等. 油茶种子 EST 文库构建及主要表达基因的分析 [J]. *林业科学*, 2006, 42(1): 43 - 48
- [9] 睦顺照, 祝钦泷, 李名扬, 等. 植物蛋白 *oleosin* 及其在植物基因工程中的应用 [J]. *中国生物工程杂志*, 2003, 23(6): 17 - 20
- [10] Kim Hu, Hsien K, Katayake C, et al. A novel group of oleosins is present inside the pollen of *Arabidopsis* [J]. *J Biol Chem*, 2002, 277(25): 22677 - 22684
- [11] van Rooijen G J H, Moloney M M. Structural requirements of oleosin domains for subcellular targeting to the oilbody [J]. *Plant Physiology*, 1995, 109: 1353 - 1361
- [12] Sibto R M P, Findlay K, Lopez - Villalobos A, et al. The accumulation of oleosins determines the size of seed oilbodies in *Arabidopsis* [J]. *Plant Cell*, 2006, 18: 1961 - 1974
- [13] Tian B, Lin Z B, Ding Y, et al. Cloning and characterization of a cDNA encoding Ran binding protein from wheat [J]. *DNA Seq*, 2006, 17(2): 136 - 142
- [14] Li F X, Jin Z P, Qu W Q, et al. Cloning of a cDNA encoding the *Saussurea medusa* chalcone isomerase and its expression in transgenic tobacco [J]. *Plant Physiol Biochem*, 2006, 44(7 - 9): 455 - 461
- [15] Tan W, Chen Y, Zhang L, et al. Construction and characterization of a cDNA library from liver tissue of Chinese Banna minipig inbred line [J]. *Transplant Proc*, 2006, 38(7): 2264 - 2266
- [16] 萨姆布鲁克 J, 拉塞尔 D W. 分子克隆实验指南 [M]. 黄培堂等, 译. 3 版. 北京: 科学出版社, 2002