

# 野生大豆种子 cDNA 文库构建及其球蛋白基因克隆

王清连, 石明旺

(河南科技学院, 河南 新乡 453003)

**摘要:** 以优质野生大豆(*Glycine soja*)的近成熟种子为材料成功构建了野生大豆 cDNA 文库, 库容量为  $9.594 \times 10^5$  克隆, 文库的重组率约为 93.7%。PCR 检测重组克隆的插入片段平均大于 1 000bp, 表明构建的近成熟种子 cDNA 文库质量较高。从文库中随机挑选 252 个克隆进行 3'端测序, 得到 217 条表达序列标签(EST), 测序结果表明, 片断大于 500bp; 野生大豆与栽培大豆(*Glycine max*)球蛋白的 cDNA 序列存在 99%相似性。

**关键词:** 野生大豆; 种子; cDNA 文库; 大豆球蛋白

中图分类号: S565.1 文献标识码: A 文章编号: 1004-3268(2006)01-0029-04

## Construction of cDNA Library and Cloning of Globulin Gene from Wild Soybean(*Glycine soja*) Seeds

WANG Qing-lian, SHI Ming-wang

(Henan Institute of Science and Technology, Xinxiang 453003, China)

**Abstract:** The cDNA library containing  $9.594 \times 10^5$  clones was established. Using the ripe seeds of high-quality wild soybeans, the percentage of recombination was as high as 93.7%, and PCR results showed that the average size of inserts was larger than 1 000bp, 217 expressed sequence tags were gained from 252 clones, randomly from the library selected by sequencing at the 3' end. Sequencing length of fragments was with over 500bp. 99% globulin cDNA sequence was similar between *Glycine soja* and *Glycine max*.

**Key words:** *Glycine soja*; Seed; cDNA library; Glycinin

我国野生大豆(*Glycine soja*)资源十分丰富, 野生大豆具多花荚、多分枝, 高油、高蛋白等特点, 有潜在优良基因和利用价值<sup>[1-3]</sup>, 有助于拓宽大豆育种遗传基础, 提高大豆育成品种水平<sup>[4]</sup>。本研究选择的野生大豆是经测定属于高油、高蛋白的“双高”材料类型, 作为一种新的种质材料, 其成熟种子不仅含有较高的油脂, 而且同时含有较高的储藏蛋白质, 从其种子中分离、鉴定与高油、高蛋白的生物合成相关的酶基因及研究该材料“双高”形成机理、利用其优

良基因具有重要意义。

cDNA 文库不仅在基因分离和克隆及筛选新基因中具有重要作用<sup>[5,6]</sup>, 而且 cDNA 文库的构建已成为研究特定组织基因表达情况和筛选目的基因的有利工具<sup>[7]</sup>。为此, 我们构建了野生大豆近成熟的发育种子 cDNA 文库, 并进行了大量测序, 获得了一些有价值的基因片断和新基因, 对实现大豆产量的突破和品质的改良, 为“双高”育种及转基因研究提供了材料和研究基础。

收稿日期: 2005-09-23

基金项目: 河南省杰出人才创新基金项目(0321001400)

作者简介: 王清连(1956-), 男, 河南浚县人, 教授, 主要从事植物生理生化研究与教学工作。

## 1 材料与方法

### 1.1 材料

野生大豆材料种植在河南科技学院大豆育种圃中,从 2004 年 9 月初始,每 3d 进行观察和取样。取样时将叶片、豆夹一起取下,将种子放入  $-80^{\circ}\text{C}$  冰箱或液氮中以备用,记录日期,样品种类。用种子含油转化和蛋白转化高峰期接近成熟的健康种子提取 mRNA,取 20g 种子进行建库。

### 1.2 药品

cDNA 双链的合成采用 Promega 公司的 cDNA 合成试剂盒和回收试剂盒,为 QIAEX Extraction Kit,其余药品按实验要求进行配制。

### 1.3 方法

1.3.1 RNA 的提取等 采用裂解法<sup>[8]</sup>提取总 RNA,采用过柱法分离纯化 mRNA,并按 Promega 公司提供的试剂盒说明书进行纯化。总 RNA、mRNA 的浓度测定:分别取  $1\mu\text{l}$  RNA 和 mRNA 样品,稀释至 100 倍,紫外测定  $\text{OD}_{260}$ 、 $\text{OD}_{280}$  确定浓度,其纯度由 260 nm 吸收值决定,并计算 RNA 和 mRNA 的量。

总 RNA、mRNA 的电泳检测:各取  $5\mu\text{l}$  总 RNA 和 mRNA 上样,稳压 85V,1% 的  $0.5\times\text{TBE}$  琼脂糖凝胶电泳 30~40 min,在自动凝胶成像系统上观察电泳结果,照相并保存。

1.3.2 cDNA 文库的构建 以 mRNA 为模板,以带有 Xho I 酶切位点的 Oligod(T)<sub>18</sub> 为引物,在 SuperScript<sup>TM</sup> II RNAase H-Reverse Transcriptase 作用下转录出 cDNA 第一链,之后参照 Promega 试剂说明书在 DNA Polymerase I 作用下合成 cDNA 第二链<sup>[9]</sup>。双链 cDNA 经加接头、末端补平、末端磷酸化及 Xho I 酶切后进行回收,具体方法参照 QIAEX Extraction Kit 回收试剂盒说明书进行。cDNA 与质粒载体的连接,依 cDNA 的电泳定量结果,每个样品设置 3~4 个比例的连接,  $14^{\circ}\text{C}$  室温反应 3 h 或者  $4^{\circ}\text{C}$  过夜。向 *E. coli* 中导入已连接处理过的 cDNA,参照《分子克隆》,利用  $\text{CaCl}_2$  法<sup>[10]</sup>,制备感受态细胞 DH10B,留适当体积的转化细胞做文库质量测定,其余转化细胞加入 60% LB 与 40% 甘油混合液 1 ml,混匀后在  $-70^{\circ}\text{C}$  储存。培养 12 h 后得到 cDNA 克隆。分别挑取单克隆到含 1 ml LB 液体培养基的离心管中,培养后抽提质粒 DNA,在  $-20^{\circ}\text{C}$  保存。

1.3.3 cDNA 文库质量测定 用稀释缓冲液以

1:5, 1:10, 1:15 的比例稀释,上述转化细胞在含有 IPTG、X-gal 和含氨苄青霉素 ( $100\mu\text{g/ml}$ ) 的 LB 培养基上,  $37^{\circ}\text{C}$  过夜,培养筛选。用 Quantity one 4.4 拍照后计算每平皿中的克隆数,蓝白斑数的统计将用于重组子百分比和库容量的计算等<sup>[3]</sup>。

1.3.4 插入片段的大小测定与克隆测序 在含氨苄青霉素 ( $100\mu\text{g/ml}$ ) 的 LB 平板上,挑选分隔良好的单菌落接种在新的含氨苄青霉素 ( $100\mu\text{g/ml}$ ) 的 LB 培养基的平皿中,过夜培养后,将每个克隆编号,然后用无菌牙签从每平皿中逐一挑取少许菌落,即 cDNA 克隆,置入含氨苄青霉素 ( $50\mu\text{g/ml}$ ) 的 LB 液体的 1.5 ml 离心管内,在良好通气的摇床上  $37^{\circ}\text{C}$  培养过夜,利用碱裂解法制备质粒 DNA,剩余的克隆在  $4^{\circ}\text{C}$  下保存。在华大基因公司进行 DNA 测序,检测 cDNA 文库重组克隆插入片段的大小,之后,从文库中随机挑选一定数量的克隆,由华大基因公司进行测序。

## 2 结果与分析

### 2.1 总 RNA 的提取和 mRNA 的分离

高质量的总 RNA 和 mRNA 是构建高质量的 cDNA 文库的前提条件。采用变性裂解液提取大豆种子的总 RNA,取  $1\mu\text{l}$  电泳检测,通过 1% 的琼脂糖变性凝胶电泳检测,具有 18S 和 28S 条带(图 1),没有降解,说明实验得到了高质量的 RNA 分子,OD 值为:  $A_{260} = 1.396$ ,  $A_{280} = 0.750$ ;  $A_{260}/A_{280} = 1.86$ ; 总量为  $2\ 200\mu\text{g}$ 。利用 Poly(A) Quik mRNA Isolation Kit 分离纯化了野生大豆的 mRNA,并取  $5.5\mu\text{l}$  电泳检测。

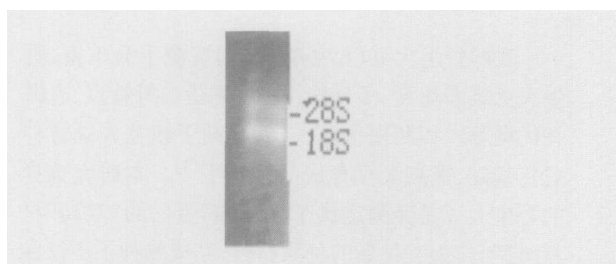


图 1 总 RNA、mRNA 电泳检测结果

### 2.2 双链 cDNA 合成

以 mRNA 为模板和 Oligod(T) 为引物,反转录酶催化合成第一链 cDNA,取  $2\mu\text{l}$  电泳检测符合要求后,再由 RNaseH DNA 聚合酶 I 共同作用合成第二链 cDNA。形成的双链,在 T4 DNA polymerase 的作用下将双链 cDNA 的末端补平,补平后的 cDNA 需要经过酚/氯仿/异戊醇纯化,去除蛋白等杂质。

EcoR I Adaptor 在 T4 DNA Ligase 的作用下连接过夜, 加接头的 cDNA 双链先在 T4 PNK 的作用下进行末端磷酸化, 后经内切酶 Xho I 酶切消化后产生 Xho I 粘性末端。形成的双链 cDNA 经紫外测定结果为  $A_{260}/A_{280}=1.75$ ,  $A_{260}/A_{230}=1.94$ , 表明 cDNA 样品纯度较高。采用 QIAEXII Gel Extraction Kit 回收, 通过 DNA 定量 Marker 电泳定量, 1.0 ~ 4.0kb 的片段浓度大约为 5 ng/ $\mu$ l。

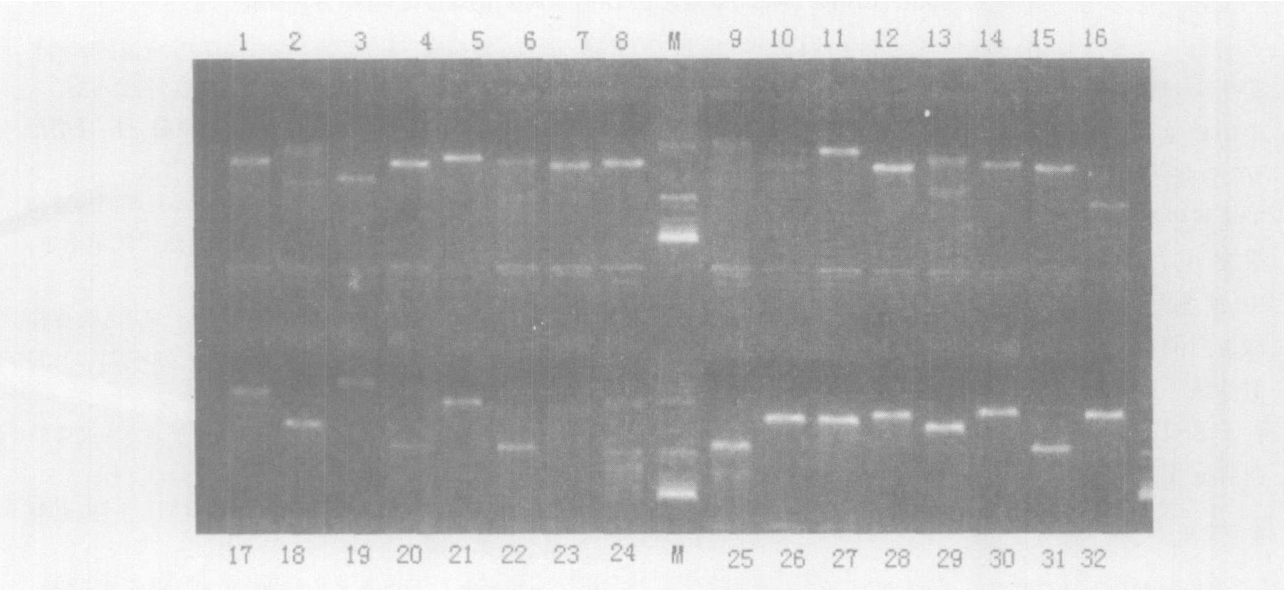
2.3 cDNA 文库质量检测

2.3.1 库容检测 从插入片段:载体=1:1 和 1:3 及 3:1 等比例中, 筛选出最合适的比例为插入片段:载体=3:1, 将插入片段和 pBluescript II SK(+) XR 在 T4 DNA Ligase 的作用下连接过夜。连接体

系 15  $\mu$ l。第 2 天取 1  $\mu$ l 连接产物做 PCR, 检测连接效率。分别以 T3 和 T7 作为引物进行 PCR 检测插入片断。

所有的连接产物在做转化前通过 Millipore 的膜纯化去除盐离子; 取 1  $\mu$ l 连接产物电转化感受态细胞 DH10B, 37  $^{\circ}$ C 过夜培养后观察生长情况。转化结果: 5  $\mu$ l 涂板共长了 123 个单菌落, 转化后共得 1 300  $\mu$ l 菌液, 经计算库容为  $9.594 \times 10^5$  克隆。

2.3.2 菌落 PCR 检测 以 T3 和 T7 为引物对 cDNA 文库进行 PCR 检测, 随机挑选 32 个菌落, 经 PCR 法特异扩增, 扩增产物经 1% 琼脂糖凝胶电泳检测, 结果如图 2。从图 2 可以看出, 30 个重组克隆 cDNA 都带有插入片段, 片段长度大于 500bp, 说明



M: 1kb ladder; 1 ~ 32 为样品  
图 2 菌落 PCR 电泳检测结果

获得的文库质量较高。

2.3.3 蓝白斑筛选与球蛋白 cDNA 测序分析 经蓝白斑筛选测定表明, 文库的重组率约为 93.7%。从文库中随机挑选 252 个克隆进行 3'端测序, 得到 217 条表达序列标签 (EST), 测序结果表明, 片断平均长度大于 500bp。测序检测结果, 获得了近成熟种子 200 多个克隆, 统计结果表明, 3'端测序平均长度大于 500bp cDNA 序列, 将数据与 GenBank 中数据进行比对, 野生大豆球蛋白的 cDNA 序列 (长度 589bp) 与栽培大豆 (*Glycine max*) 球蛋白的 A2B1a 亚基 cDNA 序列之间存在 99% 一致性 (Identities), 即相似性, 比对参数分别为: Score=1 148 bits (579), Expect=0.0, Identities=582/583 (99%), Gaps=0/583 (0%), Strand=Plus/Plus, 因篇幅关系碱基比

对结果列出一部分 (图 3), 图中 119 表示野生大豆所测序列的 cDNA 编号, SOYGA2B1A 和 SOYGLYBSU 代表栽培大豆的球蛋白 cDNA 的 GenBank 中的登录号, Consensus 表示了比对 cDNA 序列的保守序列, 图 3 中用不同颜色表示碱基差异, 相同的颜色表示相同的碱基, 如图 3 中的浅色部分表示它们的相同碱基, 即比较保守的序列, 深色等表示有差异的碱基。

3 结论与讨论

以双高型优质野生大豆的近成熟种子为材料, 采用裂解法并经改进提取总 RNA 之后, 经过柱分离并纯化得到 mRNA 等实验, 成功地构建了双高型优质野生大豆的近成熟种子 cDNA 文库。经检测,

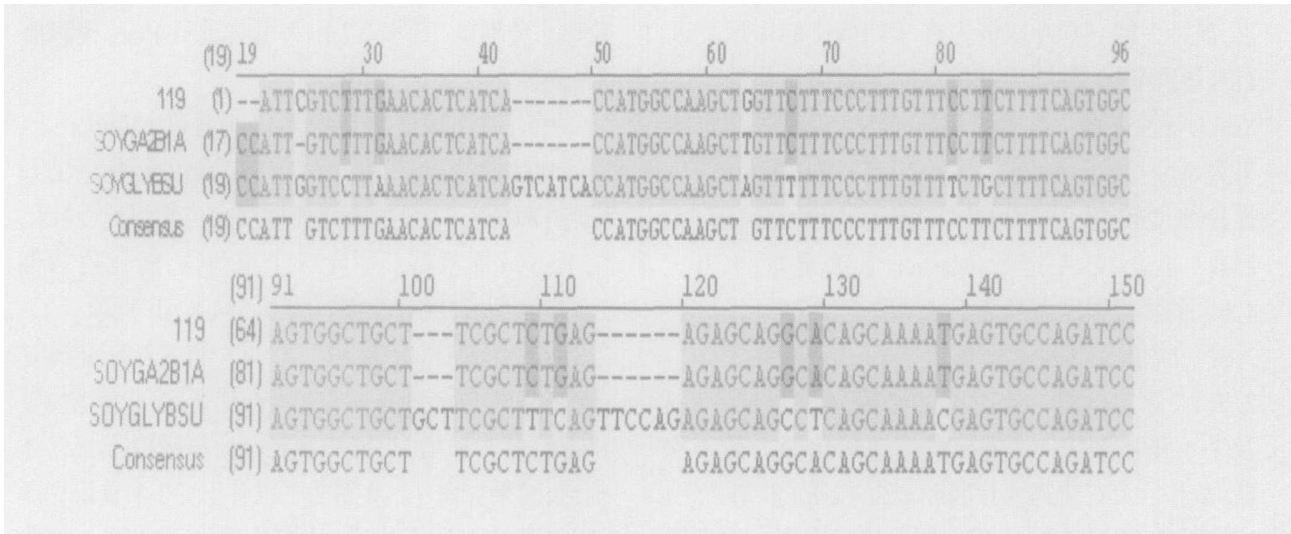


图 3 野生大豆 (*Glycine soja*)与栽培大豆(*Glycine max*)的球蛋白 cDNA 序列比较

库容量达  $9.594 \times 10^5$  克隆, 库容量相对较小, 这与野生大豆植物组织 RNA 的提取较难有关。因为无论质粒文库还是噬菌体文库, 一般可使库容量达  $10^6$  克隆<sup>[11, 12]</sup>。从文库中随机挑选 252 个克隆进行 3'端测序, 得到 217 条表达序列标签 (EST), 测序结果表明, 片断大于 500bp, 表示文库质量较高。其中, 所测野生大豆球蛋白的 mRNA 序列与栽培大豆球蛋白的 A2B1a 亚基 cDNA 序列之间存在 99%相似性或一致性。上述工作, 为今后构建野生大豆的种子 ESTs 数据库, 进一步克隆野生大豆高油高蛋白合成相关基因等奠定了良好物质和技术基础。

参考文献:

[ 1 ] Wang D, Graef G L, Procopiuk A M, et al. Identification of putative QTL that underlie yield in interspecific soybean backcross populations[ J ]. Theor Appl Genet, 2004, 108(3): 458—467.  
[ 2 ] 程玉忠, 米景九. 野生大豆基因文库的构建[ J ]. 遗传学报, 1990, 17(6): 455—460.  
[ 3 ] Fukuda T, Maruyama N, Kanazawa A, et. al. Molecular analysis and physicochemical properties of electrophoretic variants of wild soybean (*Glycine soja*) storage proteins[ J ]. J Agric Food Chem, 2005, 53(9): 3658—3665.

[ 4 ] 林红, 姚振纯, 齐宁, 等. 大豆优异种质资源的利用与创新[ J ]. 植物遗传资源科学, 2001, 2(3): 32—35.  
[ 5 ] 李海红, 王秦秦. cDNA 文库的构建策略[ J ]. 昆明医学院学报, 2003, 24(4): 22—25, 34.  
[ 6 ] 何东苟, 余新炳, 吴忠道. 全长 cDNA 文库的构建和新基因全长 cDNA 克隆的策略[ J ]. 热带医学杂志, 2003, 3(4): 473—476.  
[ 7 ] 石明旺, 谭晓风, 王义强, 等. 油茶种子 cDNA 文库构建中的问题与对策[ J ]. 经济林研究, 2004, 22(2): 53—55.  
[ 8 ] 胡芳名, 谭晓风, 石明旺, 等. 油茶种子 cDNA 文库的构建[ J ]. 中南林学院学报, 2004, 24(5): 1—4.  
[ 9 ] 李德堡, 周雪平. 基因工程操作技术[ M ]. 上海: 上海科学技术出版社, 1996.  
[ 10 ] Sambrook J, Fritsch E F. Molecular Cloning: A laboratory manual (2<sup>nd</sup> ed. )[ M ]. New York: Cold Spring Harbor Laboratory Press 1989.  
[ 11 ] 吴耀荣, 赵双宜, 夏光敏. 小麦幼叶 cDNA 文库的构建[ J ]. 山东大学学报(理学版), 2003, 38(2): 101—104.  
[ 12 ] 赵志辉, 李宁. EST 序列测定时 cDNA 文库的构建和参数评估[ J ]. 农业生物技术学报, 2003, 11(4): 422—425.