Theor Appl Genet (1999) 98: 638–646

Y.-W. Nam · R. V. Penmetsa · G. Endre
P. Uribe · D. Kim · D. R. Cook

Construction of a bacterial artificial chromosome library of *Medicago truncatula* and identification of clones containing ethylene-response genes

Received: 27 July 1998 / Accepted: 5 August 1998

Abstract To facilitate genome analysis and map-based cloning of symbiotic genes in the model legume *Medicago truncatula*, a bacterial artificial chromosome (BAC) library was constructed. The library consists of 30 720 clones with an average insert size of approximately 100 kb, representing approximately five haploid-genome equivalents. The frequency of BAC clones carrying inserts of chloroplast DNA was estimated to be 1.4%. Screening of the library with single- or low-copy genes as hybridization probes resulted in the detection of 1–12 clones per gene. Hybridization of the library with repeated sequences such as rDNA genes and transposon-like elements of *M. truncatula* revealed the presence of 60 and 374 BAC clones containing the two sequences, respectively. The BAC library was pooled for screening by polymerase chain reaction (PCR)-amplification. To demonstrate the utility of this system, we used primers designed from a conserved region of the *ein3*-like loci of *Arabidopsis thaliana* and isolated six unique BAC clones from the library. DNA gel-blot and sequence analyses showed that these *ein3*-like clones could be grouped into three classes, an observation consistent with the presence of multiple *ein3*-like loci in *M. truncatula*. These results indicate that the BAC library represents a central resource for the map-based cloning and physical mapping in *M. truncatula* and other legumes.

Communicated by G. Wenzel

Y.-W. Nam · R.V. Penmetsa · P. Uribe
D. Kim · D. R. Cook (✉)
Department of Plant Pathology and Microbiology,
and Crop Biotechnology Center, Texas A&M University,
College Station, TX 77843, USA
Fax: +1 409 862 4790
E-mail: drc1653@acs.tamu.edu

G. Endre
Institute of Genetics, Biological Research Center of the Hungarian Academy of Sciences, P.O. Box 521, Szeged, H-6701, Hungary

Key words *Medicago truncatula* · BAC library · PCR-amplification screening · Nodulation · Ethylene-response genes

Introduction

Medicago truncatula Gaertn. ("barrel medic") is an annual, self-fertile, plant species of Mediterranean origin that is closely related to the agronomically important forage legume alfalfa (*Medicago sativa* L.). Recently, *M. truncatula* has attracted interest from researchers studying symbiotic nitrogen fixation (Cook et al. 1997). Traditionally, investigation of the *Rhizobium*-legume symbiosis has been conducted with major legume crops such as pea, soybean and alfalfa. However, features such as large genome size, abundant repetitive DNA, complex ploidy, and/or difficulties with the regeneration of transgenic plants have limited the progress of molecular genetic analysis in these species. Several attributes of *M. truncatula* make it suitable as a model plant for legume biology and symbiosis research, including its small diploid genome, short life cycle, and its capacity for rapid transformation and regeneration (Barker et al. 1990; Blondon et al. 1994; Trieu and Harrison 1996).

Several symbiotic mutants have been identified from EMS (ethylmethanesulfonate) and γ ray-mutagenized populations of *M. truncatula* (Sagan et al. 1995; Penmetsa and Cook, manuscript in preparation). One particular mutant exhibited failure in selective nodulation arrest and developed abnormally numerous nodules. This 'sickle' mutant also displayed an ethylene-insensitive phenotype which was determined by examining its growth response upon treatment of germinating seedlings with ethylene (Penmetsa and Cook 1997). Thus, in addition to its role in regulating plant growth, development, and response to plant pathogens (O'Donnell et al. 1996; Ecker 1995), ethylene is implicated as a key

regulator of rhizobial infection, and thereby symbiotic nodule development in, legumes (Peters and Crist-Estes 1989; Lee and LaRue 1992; Heidstra et al. 1997; Penmetsa and Cook 1997). Therefore, the gene product responsible for the 'sickle' (*skl*) phenotype is likely to play an important role in the signal transduction pathway that mediates ethylene perception and nodulation. Our present goals include cloning the wild-type allele for *skl* as well as other symbiotic loci defined by their early nodulation phenotypes.

The recent technological advances in the map-based cloning of plant genes (Arondel et al. 1992; Martin et al. 1993; Song et al. 1995) encourages the molecular genetic analysis of plant nodulation mutants. In addition to the need for tightly linked molecular markers, the availability of large insert DNA libraries, such as YAC (yeast artificial chromosome) and BAC (bacterial artificial chromosome) is an essential element in map-based cloning strategies (Tanksley et al. 1995). The BAC system has recently emerged as a preferred method for cloning large insert DNA because of its relative absence of chimerism and ease of manipulation (Green et al. 1991; Shizuya et al. 1992; Woo et al. 1994). Several BAC libraries have been constructed from higher-plant DNAs, including libraries from *Arabidopsis thaliana* (Choi et al. 1995; Wang et al. 1996), rice (Wang et al. 1995; Zhang et al. 1996; Nakamura et al. 1997), sorghum (Woo et al. 1994), soybean (Danesh et al. 1998), and lettuce (Frijters et al. 1997).

We report here the construction and characterization of a BAC library of *M. truncatula* containing over 30 000 clones. For efficient PCR-amplification screening of the library, an ordered array of BAC DNA pools was also constructed. We demonstrate the usefulness of this system by isolating *M. truncatula* genes that are homologous to the corresponding *Arabidopsis* genes of the ethylene-response pathway.

Materials and methods

Plant material and isolation of high-molecular-weight DNA

The *M. truncatula* genotype A17 (cultivar "Jemalong") was used as the source of high-molecular-weight (HMW) DNA. Plants were grown in flats in a growth chamber for 4–6 weeks. Harvested leaf tissue was ground to a powder using a cold mortar and pestle and liquid nitrogen. Intact nuclei were isolated from 120 g of tissue and embedded in 5 ml of low-melting-point (LMP) agarose microbeads as described by Zhang et al. (1995). Nuclei prepared from 120 g of tissue were embedded in 5 ml of LMP agarose microbeads. HMW DNA was prepared by the lysis of nuclei and washed as described by Zhang et al. (1995).

Partial digestion and size-fractionation of HMW DNA

HMW DNA embedded in agarose microbeads was equilibrated with 1 × *Hin*dIII reaction buffer (50 mM Tris-HCl, pH 8.0, 50 mM NaCl, 10 mM $MgCl_2$) and 2 mM spermidine for 30 min on ice prior to the addition of *Hin*dIII (Gibco-BRL, Gaithersburg, M.D.). For partial digestion, the *Hin*dIII enzyme was diluted by serial transfer to obtain concentrations of 4, 2, 1 and 0.5 units of *Hin*dIII per 100 µl of microbeads. After enzyme addition, the mixture was again incubated for 30 min on ice. Digestion was carried out for 15 min at 37°C and stopped by the addition of 50 mM of EDTA. Microbeads containing the partially digested DNA were loaded on a 1.0% LMP agarose gel, and separated by pulsed-field gel electrophoresis (PFGE) in 1 × TAE (40 mM Tris-acetate, 1 mM EDTA, pH 8.0) using a CHEF Mapper (Bio-Rad, Hercules, Calif.) at 11°C for 18 h (6 V/cm, 120° angle, 90 s pulse time). After electrophoresis, gel slices containing DNA ranging in size from 200 kb to 500 kb were excised and used for ligation.

Preparation of BAC vector

The BAC cloning vector pBeloBACII was a gift from Drs. H. Shizuya and M. Simon (California Institute of Technology). A culture of *Escherichia coli* strain DH10B cells carrying the vector was grown in 4 l of LB containing 12.5 µg/ml of chloramphenicol overnight. The plasmid was isolated by the alkaline-lysis method (Sambrook et al. 1989) and further purified by CsCl density gradient centrifugation in an NVT 90 rotor (Beckman, Palo Alto, Calif.) at 70000 rpm for 12 h. After centrifugation, supercoiled DNA was collected and washed repeatedly using a Centricon 30 microconcentrator (Amicon, Beverly, Mass.). The BAC vector was digested with *Hin*dIII to completion and extracted with phenol/chloroform. Ends of the linearized vector were de-phosphorylated with shrimp alkaline phosphatase (US Biochemical, Cleveland, Ohio) according to the manufacturer's specification. After de-phosphorylation, the phosphatase was inactivated at 65°C for 20 min, and approximately 50-ng aliquots were stored at −80°C. The extent of de-phosphorylation was examined by a self-ligation and transformation assay.

Library construction

The LMP agarose-gel slice containing size-selected *M. truncatula* genomic DNA fragments was melted at 68°C for 15 min and treated with GELase (Epicentre Technologies, Madison, Wis.) at 45°C for 1 h. For ligation, the insert (approximately 50 ng) and the vector were mixed at a molar ratio of 1:25 and incubated at 60°C for 10 min prior to the addition of the reaction buffer. Then 50 µl of ligation mixture was incubated with 2 units of T4 DNA ligase (Gibco-BRL) at 16°C for 16 h, and 1 µl was transformed into electrocompetent *E. coli* DH10B cells (Gibco-BRL). The transformation mixture was transferred to 1 ml of SOC (2% Bacto tryptone, 0.5% Bacto yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 20 mM glucose, pH 7.0) and incubated with shaking at 37°C for 1 h. Transformed cells were selected by growth on LB plates containing chloramphenicol (12.5 µg/ml). Insert-containing BAC clones appeared white on media supplemented with 5-bromo-4-chloro-3-indolyl-β-galactoside (X-gal) (50 µg/ml) and isopropylthio-β-D-galactoside (IPTG) (25 µg/ml). A total of 30 720 white colonies were transferred to 80 384-well microtiter plates (Nunc, Naperville, Ill.) containing LB freezing buffer (Woo et al. 1994). Following overnight growth at 37°C, the plates were stored at −80°C.

Screening of the library by filter hybridization

BAC replica filters were prepared using a high-density replicating tool installed in a Biomek 2000 robot (Beckman). Hybond N(+) nylon filters (8 × 12 cm^2) (Amersham, Arlington Heights, Ill.) were

placed on disposable 384-well plate lids (Corning Glass Works, Corning, N.Y.) containing LB agar and 12.5 μg/ml chloramphenicol. Each clone was inoculated twice as symmetrical points in a 3 × 3 grid pattern. A library of 30 720 clones was blotted onto 20 filters, each of which contained 1536 clones. Bacterial colonies grown on the filter surface were lysed and the DNA was fixed by treatment with NaOH (Woo et al. 1994). The filters were washed as described by Woo et al. (1994) and baked at 80°C for 2 h prior to storage at 4°C. BAC filters were pre-hybridized at 65°C for 3 h in a solution of 0.25 M sodium phosphate, pH 7.2, 7% SDS and 1 mM EDTA (1% bovine albumin and 10 μg/ml of sheared and denatured herring sperm DNA were added in some experiments). DNA probes were labeled with α-^{32}P-dCTP (3000 Ci/mmole, Amersham, USA) using a Ready-to-Go DNA Labeling Kit (Pharmacia, Piscataway, N.J.) and eluted through a Nick spin column (Pharmacia). Heat-denatured probe was mixed with fresh buffer, and hybridization was carried out at 65°C for at least 15 h. For low-stringency hybridization, incubation was at 50°C. After hybridization, filters were washed successively at 65°C (50°C for low-stringency washing) in solutions containing 0.1% SDS and decreasing concentrations of SSPE (2 × SSPE, 1 × SSPE, and 0.5 × SSPE) for 30 min each.

DNA probes

M. truncatula actin 2, histone H3, ENOD 40 and leghemoglobin-1 clones were obtained by subcloning PCR-amplified fragments of A17 genomic DNA into pBluescript vector (Stratagene, La Jolla, Calif.). The primers were designed from conserved regions of related sequences available at the NCBI GenBank databases, and the identities of individual clones were verified by DNA sequence analysis. The ribosomal DNA probe was a 3-kb fragment amplified by PCR from A17 genomic DNA with two primers specific to the 18S and 25S rDNA genes (a gift from Dr. T. Huguet, CNRS-INRA, Castanet-Tolosan, France). The transposon-like (TL) probe was a mixture of five clones that share a common 171-bp inverted repeat, similar to that described by Peng et al. (1996, and unpublished data). Barley chloroplast *rbcL* and *psbA* cDNA clones were gifts from Dr. J. E. Mullet (Texas A&M University). The *A. thaliana ein3* cDNA clone was a gift from Dr. J. R. Ecker (University of Pennsylvania). The nucleotide sequences of the EIL-B and EIL-D degenerate primers used to amplify the *M. truncatula ein3* homologues are: 5′-GCACARGAYGGIATHYTIAARTAYATG-3′ (EIL-B) and 5′-ATGACNSCIRTIARIA CNCCIAYYTTCCA-3′ (EIL-D), respectively. C_0t-1 DNA was prepared from A17 genomic DNA according to Zwick et al. (1997).

BAC clone analysis

Individual BAC clones were inoculated in 5 ml of LB containing 12.5 μg/ml of chloramphenicol and grown at 37°C overnight. BAC DNA was prepared by the alkaline-lysis method and RNA was removed partially by precipitation in cold 2.5 M ammonium acetate. The final DNA pellet was re-suspended in 50 μl of TE (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) and 5 μl was digested with *Not*I (Gibco-BRL). The digested DNA was loaded in a 0.9% agarose gel and separated by PFGE in 0.5 × TBE (45 mM Tris-borate, 1 mM EDTA, pH 8.0) at 11°C for 18 h. The CHEF Mapper was set at 6 V/cm, 120° angle, 5 s initial and 15 s final pulse times with a linear ramp. The insert size of BAC clones was determined by comparison with λ-concatemer size standards (New England Biolabs, Beverly, Mass.) after staining the gel with ethidium bromide. The gel containing BAC DNA was nicked by irradiation with UV light (60 mJ) in a GS Gene Linker UV chamber (Bio-Rad) and blotted onto a Hybond N(+) nylon filter with 0.4 N NAOH, 1.5 M NaCl. Filters were hybridized as described.

DNA sequencing

To sequence *ein3*-like clones, a 450-bp fragment from PCR-amplification was purified from an agarose gel using the GENECLEAN kit (Bio 101, Vista, Calif.) and used in dye terminator cycle sequencing with Thermo Sequenase (Amersham). Automated sequencing was carried out with a ABI Prism 377 DNA Sequencer (Perkin-Elmer Applied Biosystems, Foster City, Calif.) and data analysis was facilitated by Sequencher 3.0 software (Gene Codes Co., Ann Arbor, Mich.). Alignment of the predicted amino-acid sequences was performed with the Lasergene Biocomputing Software (DNASTAR, Madison, Wis.).

Pooling of the library for PCR screening

To prepare BAC DNA pools, the 30 720 clones were cultured overnight in 80 384-well microtiter plates containing LB freezing buffer (Woo et al. 1994) supplemented with 12.5 μg/ml of chloramphenicol. The 80 plates were organized into a composite 8 × 10 grid containing 192 columns and 160 rows. The 30 720 clones were condensed into 352 samples by reducing each of the 192 columns and 160 rows into pools consisting of a 25-μl aliquot from each clone. DNA was isolated from the pooled samples by the alkaline-lysis method (see above), and dissolved in 100 μl of TE. These 352 primary BAC DNA pools were further reduced to 88 superpools by combining each set of four adjacent column or row pools. DNA in each superpool was diluted 1/50 in TE and used as a template for PCR screening of the library. To facilitate pipetting, the DNA superpools were transferred into 0.2-ml PCR tube strips in a 8 × 12 microtiter format and stored at −20°C. For PCR, 5 μl of diluted BAC DNA templates were transferred with an 8-channel pipettor into a 96-well reaction plate (Perkin-Elmer Applied Biosystems) containing 5 μl of reaction mixture (2.5 mM $MgCl_2$, 50 μM dNTPs, 0.25 pmole primers and 0.1 unit of *Taq* polymerase). Typical amplification conditions were 94°C, 30 s; 55°C, 20 s; 72°C, 1 min for 35 cycles. After PCR, the products were analyzed in a 1.4% agarose gel in a wide-format electrophoresis system (Owl Scientific, Woburn, Mass.). Using this strategy, three rounds of PCR analysis were required to identify individual BAC clones. The first round of 88 reactions identified PCR-positive rows and columns in the superpooled library; the second round of PCR identified the specific column and row intersects; and the third round of PCR identified positive BAC clones in the non-pooled library.

Results

Construction of a BAC library

The *M. truncatula* BAC library was constructed from HMW DNA isolated from nuclei embedded in agarose microbeads according to Zhang et al. (1995). DNA was released from nuclei by lysing nuclear envelopes before partial digestion with *Hin*dIII to ensure the exposure of all the DNA to enzyme. Based on the analysis of ethidium bromide-stained pulsed-field gels, one unit of *Hin*dIII per 100 μl of microbeads produced the largest amount of DNA in the 200–500 kb size range (data not shown). The library was constructed with one size selection (Zhang et al. 1996) of partially digested DNA, and four separate ligation reactions gave rise to the library consisting of 30 720 clones.

BAC insert sizes were determined by digestion of randomly selected BAC clones with the *Not*I enzyme followed by PFGE. Out of 160 BAC clones tested, 134 contained *M. truncatula* insert-DNA and approximately 16% of the clones contained no inserts. The insert sizes ranged from 7 to 215 kb with an average of 101.8 kb (Fig. 1 A). The size distribution of these clones (Fig. 2) appears to represent a normal distribution, with the majority of clones in the 80–120 kb size range. Based on an estimated size of 5×10^8 bp per haploid genome for *M. truncatula* (Blondon et al. 1994), the library provides approximately five haploid-genome equivalents.

Characterization of the library

To examine the extent of repetitive DNA sequences in the library, the BAC DNA blot prepared from the pulsed-field gel (Fig. 1 A) was hybridized with *M. truncatula* C_0t-1 DNA. Of 11 clones, six hybridized strongly with the C_0t-1 DNA, whereas the other five showed little or no hybridization signal (Fig. 1 B). Based on this analysis, we estimate that 55% of the clones contained repetitive sequences and 45% contained low- or single-copy sequences.

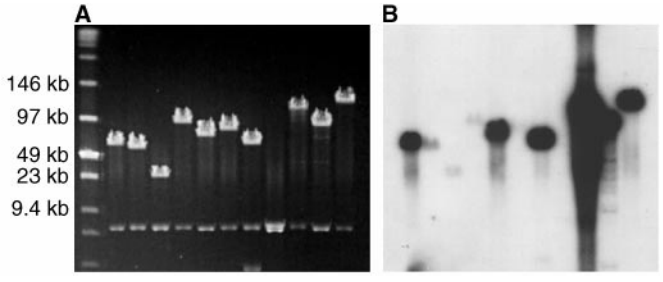

Fig. 1A, B An analysis of *M. truncatula* BAC clones by PFGE. **A** *Lane 1* Lambda concatemer size standards. *Lanes 2–12* fragments of *Not*I-digested DNA isolated from randomly selected BAC clones. PFGE was performed in a 0.9% agarose-0.5 × TBE gel by a CHEF Mapper. **B** Southern hybridization of the gel in **A** probed with radioactively labeled *M. truncatula* C_0t-1 DNA

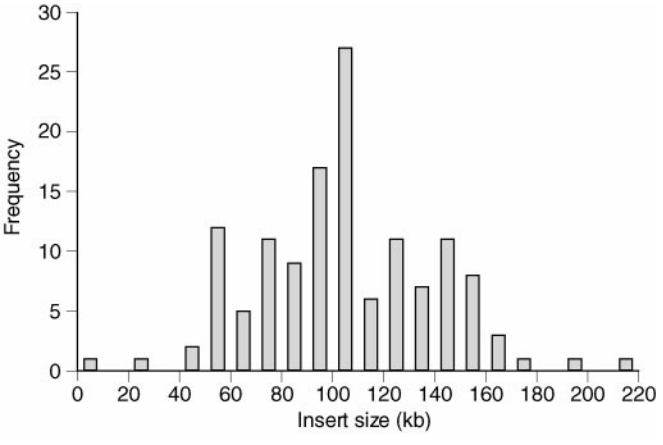

Fig. 2 Distribution of insert sizes of randomly selected BAC clones. DNA was isolated from 134 random clones and analyzed as described in Fig. 1. The insert size of each clone is plotted versus its frequency

Further characterization of the library was carried out using high-density BAC DNA filter sets. The entire library was replicated onto 20 high-density nylon membranes by means of a Beckman Biomek 2000 robot. Each $8 \times 12\ cm^2$ membrane contained colonies from four 384-well plates (for a total of 1536 clones) spotted in duplicate. To determine the percentage of BAC clones containing chloroplast DNA sequences, the high-density filters were probed with genes from the barley chloroplast genome; 284 clones hybridized with a ^{32}P-labeled mixture (1:1) of chloroplast *rbcL* and *psbA* cDNAs. Based on the size of the circular chloroplast chromosome, and the relative position of *rbcL* and *psbA*, two-thirds of the BAC clones containing chloroplast DNA inserts of 60 kb or larger should hybridize to one of these probes (Woo et al. 1994). Thus, we estimate that 1.4% of the library clones carry chloroplast sequences.

The library was screened with various DNA probes including both single- and low-copy genes, and high-copy or repetitive sequence elements. Positive BAC clones identified by hybridization to high-density filters were analyzed by *Not*I digestion followed by PFGE. The gel was blotted and the filter was hybridized with the same DNA probe again to verify the initial screening result. As shown in Table 1, the number of positive clones determined in this manner varied from 1 to 12 per single- or low-copy gene probe. At least one BAC clone was isolated for each gene probe tested. As expected, clones detected with probes belonging to a small gene family, such as histone H3, were more abundant than clones hybridizing to single-copy gene probes, such as *rip1*. By contrast, 60 clones hybridized with a PCR fragment amplified from the rDNA gene cluster (data not shown). Furthermore, probing the library with a repetitive sequence element, such as the *rip1*-associated transposon-like (TL) element of *M. truncatula* (Peng et al. 1996; Peng, 1997), resulted in hybridization to 374 clones comprising nearly 1.2% of the library (data not shown). Apparently the TL element, of which one copy is located in the promoter region of the *Rhizobium*-induced peroxidase gene *rip1*, is widespread in the *M. truncatula* genome and the BAC library is useful in demonstrating it.

Isolation of *M. truncatula* homologues of *Arabidopsis* ethylene-response genes

In an effort to elucidate molecular mechanisms of ethylene-mediated rhizobial infection arrest, we initiated

Table 1 Identification of BAC clones corresponding to gene-specific and RFLP probes

Probe	BAC clones hybridized (size in kb)[a]	Source or reference
actin2[b]	1N1 (50), 61M24 (70), 69H24 (30), 69K17 (70), 80L24 (75)	This study
aldolase[c]	8D23 (120), 37P16 (60), 40B11 (65), 50G17 (25)	Vance et al. (1994)
ein3[d]	2L23 (80), 3K16 (50), 40E9 (50), 42H9 (105), 64B21 (70), 76B24 (25)[e]	Chao et al. (1997)
ENOD8[b]	19N23 (70), 31E11 (70), 51I19 (70), 58P7 (45), 64N13 (125)	R. Dickstein, unpublished
ENOD40[b]	23A6 (25), 39I20 (7)	Crespi et al. (1994)
histone H3[b]	1F24 (45), 9J21 (30), 17E9 (75), 20D14 (65), 35L16 (60), 38P5 (45), 44C17 (30), 47M3 (20), 58B8 (45), 59O14 (50), 62G18 (85), 69G22 (65)	This study
leghemoglobin-154[b]	1K12 (115), 35C18 (70), 35O12 (85), 45J13 (80)	Barker et al. (1988)
rip1 exon3[b]	31P6 (80), 44L23 (60), 53P16 (80), 57K2 (65)	Cook et al. (1995)
U224[f]	2I4 (80), 14B3 (15), 17N5 (125), 19G24 (65), 32P13 (45), 35C8 (60), 58K12 (65), 61I3 (30), 80D1 (30)	Kiss et al. (1998)
U492[f]	60O16 (60)	Kiss et al. (1998)

[a] BAC clones were identified by hybridization of high-density filters and verified by repeated hybridization of PFGE gel blots
[b] *M. truncatula* cDNA or genomic DNA clones
[c] *M. sativa* cDNA clone
[d] *A. thaliana* cDNA clone
[e] Identified by PCR-amplification screening of BAC DNA pools
[f] RFLP clones of diploid *M. sativa*

the identification of *M. truncatula* homologues of cloned ethylene-response genes of *Arabidopsis* such as *etr1* (Chang et al. 1993), *ein3* (Chao et al. 1997) and *hls1* (Lehman et al. 1996). Genomic Southern blots of *M. truncatula* established that the *Arabidopsis ein3* cDNA clone cross-hybridized effectively with *M. truncatula* DNA under a low-stringency condition. Screening of the BAC library with radiolabeled *Arabidopsis ein3* cDNA identified five positive clones (Table 1). To determine whether these clones represent one or more *M. truncatula ein3*-like loci, DNA isolated from each clone was subjected to restriction-enzyme fingerprinting with *Eco*RI and *Hin*dIII (Fig. 3). The patterns of enzyme digestion differentiated two classes of BAC clones: one shared by clones 3K16 and 40E9, and the other shared by clones 2L23, 42H9 and 64B21. These two *M. truncatula ein3*-like loci were designated Mt-*eil1* and Mt-*eil2*, respectively.

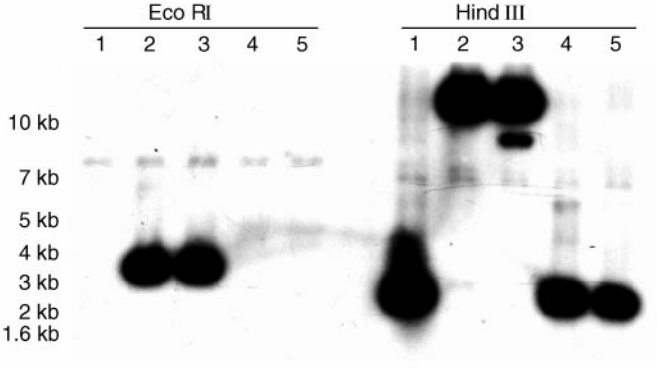

Fig. 3 Characterization of *M. truncatula* BAC clones hybridizing to *Arabidopsis ein3* cDNA. Positive BAC DNAs isolated from the library based on the hybridization of high-density filters were digested with the designated restriction enzymes. The fragments were resolved in a 0.6% agarose gel and Southern blotted. Probing of this blot with an *ein3* probe reveals two distinct patterns of hybridization (*lanes 2 and 3* vs. *lanes 1, 4 and 5*), indicating two classes of *M. truncatula ein3*-like gene. *Lanes 1, 2, 3, 4 and 5* represent BAC clones 2L23, 3K16, 40E9, 42H9 and 64B21, respectively

Preparation of BAC DNA pools and PCR-amplification screening of the library

To facilitate efficient PCR-based screening of the library, we developed a BAC DNA pool system. The entire 30 720 BAC clones were pooled two-dimensionally into 192-column and 160-row pools by arranging 80 library plates in a 8×10 configuration. Every pool was subject to BAC DNA isolation. The resulting 352 primary DNA pools were pooled again one-dimensionally to generate a final set of 88 (48 column and 40 row) DNA superpools.

The BAC DNA pools were screened by PCR for *ein3* and its closely related genes *eil1*, *eil2* and *eil3* (Chao et al. 1997). A degenerate primer pair was designed against conserved regions of the *Arabidopsis Ein3* family of proteins (Chao et al. 1997), corresponding to amino-acid residues 95–106 (AQDGILKYMLKM; EIL-B) and 247–256 (WKVGVLTAVI; EIL-D), respectively. PCR-amplification of BAC DNA superpools identified eight and ten putative positives from the column and row pools, respectively (Fig. 4). As expected, the size of the amplified fragments was approximately 450 bp. Using the same primer pair, a second screening was carried out with the four individual column or row primary DNA pools that comprised each positive superpool. This yielded eight and seven positives from the column and row pools, respectively (data not shown). The result from the second screening also provided information about the column and row coordinates that would determine the final positive BAC

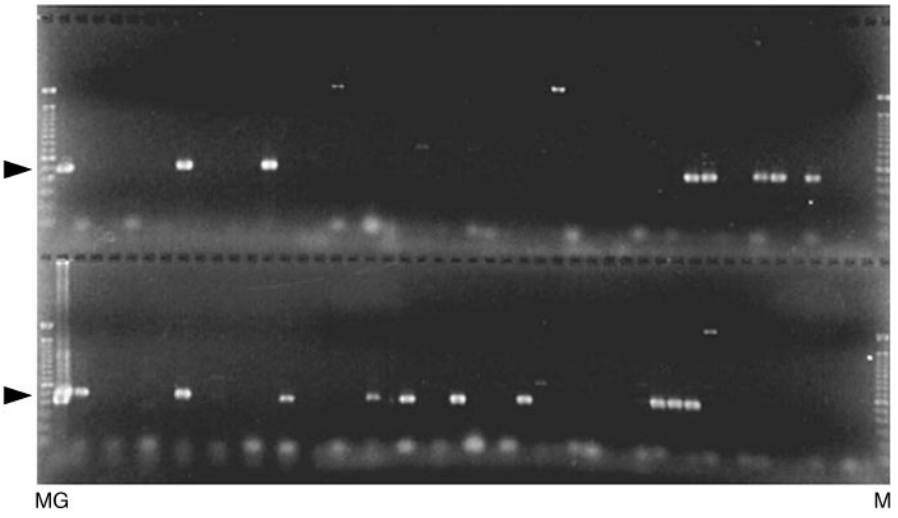

Fig. 4 PCR products of BAC DNA superpools after amplification for *ein3*-like clones. The agarose gel contains fragments amplified with EIL-B and EIL-D primers from 48-column (upper half of the gel) and 40-row (lower half of the gel) pools, respectively: '*M*' 100-bp DNA size standards; '*G*' PCR fragment amplified from A17 genomic DNA. *Arrow heads* indicate the expected size (about 450 bp) of the PCR product

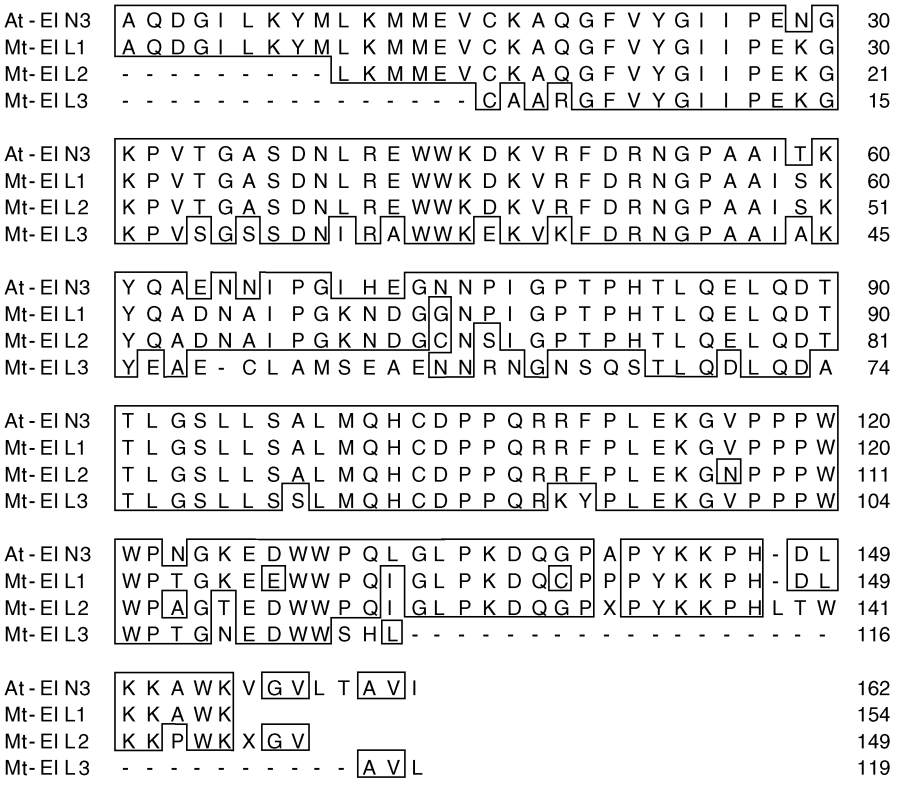

Fig. 5 Alignment of predicted partial amino-acid sequences of three *M. truncatula* BAC clones with *Arabidopsis ein3* (At-*EIN3*). The sequences were derived from nucleotide sequencing of three distinct classes of *M. truncatula ein3*-carrying BAC clones identified by low-stringency filter hybridization (Mt-*EIL1* and Mt-*EIL2*) or by PCR-amplification screening of BAC DNA pools (Mt-*EIL3*). Only highly conserved regions of *ein3* sequence are shown for comparison

clones. Individual BAC clones corresponding to every possible combination of column and row coordinates were then tested for carrying an *ein3*-like sequence. These experiments resulted in the verification of a total of six positive *ein3*-like clones (data not shown). Of these, five corresponded to the BAC clones previously identified by filter hybridization, but the sixth one (76B24) was new (Table 1). Southern analysis of restriction-digested fragments of clone 76B24 revealed a completely different pattern of hybridization (data not shown), defining a novel locus. We therefore designated this clone Mt-*eil3*.

To verify that the three putative *M. truncatula ein3* homologues are indeed related to *Arabidopsis ein3*, we sequenced PCR products amplified from individual BAC clones using the degenerate primers EIL-B and EIL-D. The deduced amino-acid sequences of the three *M. truncatula* loci exhibited strong homology ($>70\%$ nucleotide identity; $>90\%$ amino acid identity) with the *Arabidopsis ein3* sequence (Fig. 5).

Discussion

The *M. truncatula* BAC library

We have constructed a five-genome-equivalent BAC library of *M. truncatula* using partial *Hin*dIII digests. The library consists of 30 720 clones with an average insert size of approximately 100 kb. Based on an estimation that 16% of the library clones lack inserts and 1.4% contain chloroplast DNA, the library comprises a total of 2500 Mb of nuclear DNA. This accounts for a 99% probability of containing any sequence in the *M. truncatula* genome at least once (Meyerowitz 1994). Screening of the *M. truncatula* BAC library with single- or low-copy number gene-specific or RFLP probes revealed the presence of typically four or five clones (Table 1), consistent with the calculated genome coverage of the library representing five haploid-genome equivalents.

The analysis of BAC clones by *Not*I digestion followed by PFGE showed that the majority of *M. truncatula* DNA inserts were typically present as single *Not*I fragment inserts (Fig. 1). Thus, the *M. truncatula* genome apparently contains few *Not*I sites, a feature commonly observed with the genomes of other dicot species (Choi et al. 1995; Frijters et al. 1997; Danesh et al. 1998) and contrary to the results obtained with monocot species (Woo et al. 1994; Wang et al. 1995; Zhang et al. 1996).

The sizes of inserts recovered from positive BAC clones identified by library screening (Table 1) were often considerably smaller than those of the clones sampled randomly from the library (Fig. 2). This result could be due to the presence of disproportionately high numbers of *Hin*dIII sites in genomic regions containing the target genes. For instance, the single *ENOD40* locus of *M. truncatula* contains four *Hin*dIII sites in the coding and immediate flanking regions, potentially explaining the small sizes of BAC inserts containing the *ENOD40* gene. This mixture of small and large insert clones in the library may facilitate the characterization of genomic targets by reducing the demand for extensive subcloning (Wang et al. 1996). Libraries with moderately sized inserts may also be sufficient for map-based cloning, as methods for the saturation of a genomic target region with DNA markers are now relatively routine (Michelmore et al. 1991; Vos et al. 1995). Nevertheless, construction of a second BAC library using partial digests with an alternative restriction enzyme such as *Eco*RI may provide a useful complement to the present *Hin*dIII library, for example by increasing genome coverage and by increasing the total number of large insert clones.

An efficient PCR-based screening of the BAC library

Because screening of a BAC library based on conventional filter hybridization with radiolabeled probes is laborious and time-consuming, we developed an alternative PCR-based screening strategy. We chose to pool the library two-dimensionally into columns and rows (Kwiatkowski et al. 1990; Leister et al. 1997) rather than combining all clones from a single plate into a subpool (Green and Olson 1990; Danesh et al. 1998). A second one-dimensional pool was used to condense the library into 88 superpools, where the final clone complexity was 1/640 and 1/768 for column and row pools, respectively. To avoid errors that might be introduced by unequal growth of clones in a mixed culture, we pooled clones from a fully grown replicate of the library. The initial 352 column and row mixtures were subject to DNA isolation and the resulting BAC DNA pools were used to generate the superpools. For routine analysis, all 88 PCR reactions from the DNA superpools can be analyzed in a single 96-well PCR format and the resulting amplification products can be resolved in a single 100-well agarose gel.

The BAC DNA pools were tested for PCR amplification with a number of primer sets designed from specific sequences of the DNAs that were used as probes for filter hybridization (Table 1). The results from PCR screening with these primers were consistent with the results from filter hybridization. An example is illustrated by the identification of the ethylene-response genes of the *ein3* family (Fig. 4). There were slight differences in the number of positive clones identified at each step of PCR amplification: eight-column and ten-row positives were identified from the first screening, whereas eight-column and seven-row positives were detected from the second screening and only six clones were verified finally. These discrepancies could be due to the occurrence of nonspecific amplification caused by the use of degenerate primers. However, only a single major class of amplified fragments (about 450 bp) was evident from the amplification (Fig. 4) – (a minor 1.5-kb fragment detected in the first screening (Fig. 4) was not traceable through further analyses). This suggests that the amplified N-terminal region of the *EIN3*-like sequence is extremely conserved among species. In addition to the five BAC clones identified by filter hybridization, PCR amplification led to the discovery of a novel clone related to *ein3*, defining a third locus of *ein3*-like genes in *M. truncatula*. This observation is in agreement with the finding that four distinct *ein3*-like loci are present in the *Arabidopsis* genome (Chao et al. 1997).

Our PCR-based approach of BAC library screening is efficient and reliable. It considerably shortens the time required for library screening (within 2 days) compared with the time for conventional filter-hybridization screening and verification by Southern blotting (more than a week). Coupled with the capacity of direct automated sequencing of BAC clone ends and the generation of sequence-specific PCR primers (Kim and Cook, manuscript in preparation), this should allow us to rapidly construct BAC physical

contigs of overlapping clones spanning a specific genomic region.

Usefulness of the library for the genome analysis of legumes

We expect that the BAC library will be instrumental for map-based cloning of plant genes associated with legume-microbe interactions. For example, we have recently identified a hyper-nodulating mutant of *M. truncatula* that is insensitive to ethylene (Penmetsa and Cook 1997). Cloning *M. truncatula* homologues of ethylene-response genes, such as the *ein3* gene family described herein, should provide a simple means to develop gene-specific markers against this conserved pathway, and thereby to explore possible linkage between ethylene-response loci and *M. truncatula* mutant alleles. We also anticipate that the BAC library will serve as a resource for the comparative genome analysis of legumes by facilitating the physical mapping of selected genomic regions. For instance, five of the ten genes used as probes to characterize the BAC library (Table 1), including *ENOD40*, leghemoglobin-1, *rip1*, and the two RFLP probes (U492 and U224), are located in the same linkage group of *M. truncatula* (Kim and Cook, personal communication) and *M. sativa* (Endre et al. 1998; Kiss et al. 1998). The corresponding chromosomal region in pea (*Pisum sativum* L.) has also been found to be enriched with many symbiosis-related-genes (Peter Kalo and Noel Ellis, personal communication). We are currently using the BAC library to construct a physical map of selected regions within this conserved legume linkage group.

Acknowledgments A copy of the *M. truncatula* BAC library has been deposited in the Clemson University Genomics Institute. We thank Drs. Hong-Bin Zhang (Texas A&M University) and Rod Wing (Clemson University) for providing valuable advice in library construction. We also thank Drs. Joe Ecker, John Mullet, Carrol Vance (University of Minnesota) and Rebecca Dickstein (Drexel University) for the probes. This work was supported by grants from the Samuel Roberts Noble Foundation and from the National Science Foundation Integrative Plant Biology Program, award number IBN-9507535. R. V. Penmetsa is a Tom Slick Senior Graduate Fellow. We declare that the experiments described in this paper comply with the current laws of USA in which they were performed.

References

Arondel V, Lemieux B, Hwang I, Gibson S, Goodman HM, Somerville CR (1992) Map-based cloning of a gene controlling omega-3 fatty acid desaturation in *Arabidopsis*. Science 258:1353–1354

Barker DG, Gallusci P, Lullien V, Khan H, Ghérardi M, Huguet T (1988) Identification of two groups of leghemoglobin genes in alfalfa (*Medicago sativa*) and a study of their expression during root nodule development. Plant Mol Biol 11:761–772

Barker DG, Bianchi S, Blondon F, Dattee Y, Duc G, Essad S, Flament P, Gallusci P, Genier G, Guy P, Muel X, Tourneur J, Denarie J, Huguet T (1990) *Medicago truncatula*, a model plant for studying the molecular genetics of the *Rhizobium*-legume symbiosis. Plant Mol Biol Rep 8:40–49

Blondon F, Marie D, Brown S, Kondorosi A (1994) Genome size and base composition in *Medicago sativa* and *M. truncatula* species. Genome 37:264–270

Chang C, Kwok SF, Bleecker AB, Meyerowitz EM (1993) *Arabidopsis* ethylene-response gene *ETR1*: similarity of product to two-component regulators. Science 262:539–544

Chao Q, Rothenberg M, Solano R, Roman G, Terzaghi W, Ecker JR (1997) Activation of the ethylene-gas response pathway in *Arabidopsis* by the nuclear protein ETHYLENE-INSENSITIVE3 and related proteins. Cell 89:1133–1144

Choi S, Creelman RA, Mullet JE, Wing R A (1995) Construction and characterization of a bacterial artificial chromosome library of *Arabidopsis thaliana*. Plant Mol Biol Rep 13:124–128

Cook D, Dreyer D, Bonnet D, Howell M, Nony E, VandenBosch K (1995) Transient induction of a peroxidase gene in *Medicago truncatula* precedes infection by *Rhizobium meliloti*. Plant Cell 7:43–55

Cook DR, VandenBosch KA, de Bruijn F, Huguet T (1997) Model legumes get the nod. Plant Cell 9:275–281

Crespi MD, Jurkevitch E, Poiret M, d'Aubenton-Carafa Y, Petrovics G, Kondorosi E, Kondorosi A (1994) *enod*40, a gene expressed during nodule organogenesis, codes for a non-translatable RNA involved in plant growth. EMBO J 13:5099–5112

Danesh D, Penuela S, Mudge J, Denny RL, Nordstrom H, Martinez JP, Young ND (1998) A bacterial artificial chromosome library for soybean and identification of clones near a major cyst nematode resistance gene. Theor Appl Genet 96:196–202

Dickstein R, Prusty R, Peng T, Ngo W, Smith ME (1993) ENOD8, a novel nodule-specific gene, is expressed in empty alfalfa nodules. Mol Plant-Microbe Interact 6:715–721

Ecker JR (1995) The ethylene signal transduction pathway in plants. Science 268:667–675

Endre G, Kalo P, Kereszt A, Kiss P, Kiss GB (1998) Genetic analysis of *Medicago sativa* nodulation genes. In: Elmerich C, Kondorosi A, Newton WE (eds) Biological nitrogen fixation for the 21st Century. Kluwer Academic Publishers, Dordrecht, pp 315–316

Frijters ACJ, Zhang Z, Van-Damme M, Wang GL, Ronald PC, Michelmore RW (1997) Construction of a bacterial artificial chromosome library containing large *Eco*RI and *Hin*dIII genomic fragments of lettuce. Theor Appl Genet 94:390–399

Green ED, Olson MV (1990) Systematic screening of yeast artificial chromosome libraries by use of the polymerase chain reaction. Proc Natl Acad Sci USA 87:1213–1217

Green ED, Riethman HC, Dutchik JE, Olson MV (1991) Detection and characterization of chimaeric yeast artificial chromosome clones. Genomics 11:658–669

Heidstra R, Yang WC, Yalcin Y, Peck S, Emons A, van Kammen A, Bisseling T (1997) Ethylene provides positional information on cortical cell division but is not involved in Nod factor-induced tip growth in *Rhizobium*-legume interaction. Development 124:1781–1787

Kiss GB, Kalo P, Kiss P, Felfoldi K, Kereszt A, Endre G (1998) Construction of an improved genetic map of diploid alfalfa (*Medicago sativa*) using a novel linkage analysis for chromosomal regions exhibiting extreme distorted segregation. In: Elmerich C, Kondorosi A, Newton WE (eds) Biological nitrogen fixation for the 21st Century. Kluwer Academic Publishers, Dordrecht, p 313

Kwiatkowski TJ, Zoghbi HY, Ledbetter SA, Ellison KA, Chinault AC (1990) Rapid identification of yeast artificial chromosome clones by matrix pooling and crude lysate PCR. Nucleic Acids Res 18:7191–7192

Lee KH, LaRue TA (1992) Exogenous ethylene inhibits nodulation of *Pisum sativum* L. cv Sparkle. Plant Physiol 100:1759–1763

Lehman A, Black R, Ecker JR (1996) *HOOKLESS1*, an ethylene-response gene, is required for differential cell elongation in the *Arabidopsis* hypocotyl. Cell 85:183–194

Leister D, Berger A, Thelen H, Lehmann W, Salamini F, Gebhardt C (1997) Construction of a potato YAC library and identification of clones linked to the disease resistance loci *R1* and *Gro1*. Theor Appl Genet 95 : 954–960

Martin GB, Brommonschenkel SH, Chunwongse J, Frary A, Ganal MW, Spivey R, Wu T, Earle ED, Tanksley SD (1993) Map-based cloning of a protein kinase gene conferrin g disease-resistance in tomato. Science 262 : 1432–1436

Meyerowitz EM (1994) Structure and organization of the *Arabidopsis thaliana* nuclear genome. In: Meyerowitz EM, Somerville CR (eds) *Arabidopsis*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, pp 21–36

Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci USA 88 : 9828–9832

Nakamura S, Asakawa S, Ohmido N, Fukui K, Shimizu N, Kawasaki S (1997) Construction of an 800-kb contig in the near-centromeric region of the rice blast resistance gene *Pi-ta*^{2} using a highly representative rice BAC library. Mol Gen Genet 254 : 611–620

O'Donnell PJ, Calvert C, Atzorn R, Wasternack C, Leyser HMO, Bowles DJ (1996) Ethylene as a signal mediating the wound response of tomato plants. Science 274 : 1914–1917

Peng H-M (1997) Characterization and regulation of the early nodulin gene *rip1* in *Medicago truncatula*. PhD thesis, Texas A&M University, College Station, Texas, USA

Peng H-M, Dreyer DA, VandenBosch KA, Cook DR (1996) Gene structure and differential regulation of the *Rhizobium*-induced peroxidase gene *rip1*. Plant Physiol 112 : 1437–1446

Penmetsa RV, Cook DR (1997) A legume ethylene-insensitive mutant hyper-infected by its rhizobial symbiont. Science 275 : 527–530

Peters NK, Crist-Estes DK (1989) Nodule formation is stimulated by the ethylene inhibitor aminoethoxyvinylglycine. Plant Physiol 91 : 690–693

Pillai SV, Funke RP, Gresshoff PM (1996) Yeast and bacterial artificial chromosome (YAC and BAC) clones of the model legume *Lotus japonicus*. Symbiosis 21 : 149–164

Sagan M, Morandi D, Tarenghi E, Duc G (1995) Selection of nodulation and mycorrhizal mutants in the model plant *Medicago truncatula* (Gaertn.) after γ-ray mutagenesis. Plant Sci 111 : 63–71

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual (2nd edn). Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Shizuya H, Birren B, Kim U-J, Mancino V, Slepak T, Tachiiri Y, Simon M (1992) Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. Proc Natl Acad Sci USA 89 : 8794–8797

Song WY, Wang G-L, Chen LL, Kim HS, Pi L-Y, Gardner J, Wang B, Holsten T, Zhai W-X, Zhu L-H, Fauquet C, Ronald P (1995) A receptor kinase-like protein encoded by the rice disease resistance gene *Xa21*. Science 270 : 1804–1806

Tanksley SD, Ganal MW, Martin GB (1995) Chromosome landing: a paradigm for map-based gene clonin g in plants with large genomes. Trends Genet 11 : 63–68

Trieu AT, Harrison MJ (1996) Rapid transformation of *Medicago truncatula*: regeneration via shoot organogenesis. Plant Cell Rep 16 : 6–11

Vance CP, Gregerson RG, Robinson DL, Miller SS, Gantt JS (1994) Primary assimilation of nitrogen in alfalfa nodules: molecular features of the enzymes involved. Plant Sci 101 : 51–64

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23 : 4407–4414

Wang G-L, Holsten TE, Song W-Y, Wang H-P, Ronald PC (1995) Construction of a rice bacterial artificial chromosome library and identification of clones linked to the *Xa-21* disease resistance locus. Plant J 7 : 525–533

Wang G-L, Warren R, Innes R, Osborne B, Baker B, Ronald PC (1996) Construction of an *Arabidopsis* BAC library and isolation of clones hybridizing with disease-resistance, gene-like sequences. Plant Mol Biol Rep 14 : 107–114

Woo S-S, Jiang J, Gill BS, Paterson AH, Wing RA (1994) Construction and characterization of a bacterial artificial chromosome library of *Sorghum bicolor*. Nucleic Acids Res 22 : 4922–4931

Zhang H-B, Zhao X, Ding X, Paterson AH, Wing RA (1995) Preparation of megabase-size DNA from plant nuclei. Plant J 7 : 175–184

Zhang H-B, Choi S, Woo S-S, Li Z, Wing RA (1996) Construction and characterization of two rice bacterial artificial chromosome libraries from the parents of a permanent recombinant inbred mapping population. Mol Breed 2 : 11–24

Zwick MS, Hanson RE, McKnight TD, Islam-Faridi MN, Stelly DM, Wing RA, Price HJ (1997) A rapid procedure for the isolation of C_0t-1 DNA from plants. Genome 40 : 138–142