Theor Appl Genet (1999) 99: 305–313 © Springer-Verlag 1999

O. Moullet · H.-B. Zhang · E. S. Lagudah

Construction and characterisation of a large DNA insert library from the D genome of wheat

Received: 28 August 1998 / Accepted: 28 November 1998

Abstract A large DNA fragment library consisting of 144 000 clones with an average insert size of 119 kb was constructed from nuclear DNA isolated from root and leaf tissue from *Triticum tauschii* (syn. *Aegilops tauschii*), the D-genome progenitor of wheat. The library was made in a binary vector that had previously been shown to stably maintain large inserts of foreign DNA in *Escherichia coli*. The use of root nuclei reduced considerably the proportion of the library containing clones derived from chloroplast DNA. Several experimental parameters were investigated and optimised, leading to a high cloning efficiency. Only three ligations were needed to construct the library which was estimated to be equivalent to 3.7 haploid genomes. The accuracy of this estimation was demonstrated by screening this library with three well-defined probes. One probe containing a glutenin gene sequence identified 5 clones covering at least 230 kb of the *Glu-D1* locus and contained the two tightly linked high-molecular-weight glutenin genes *Glu-D1x* and *-D1y*. Each of the other two single-copy probes derived from the *Cre3* cereal cyst nematode resistance gene locus hybridised with 4 clones containing gene sequences encoding nucleotide binding sites and a leucine-rich region. This is the first representative large-insert DNA library for wheat, and the results indicated that large molecules of wheat DNA can be efficiently cloned, stably maintained and manipulated in a bacterial system.

Communicated by G. E. Hart

O. Moullet · E. S. Lagudah (✉)
Commonwealth Scientific and Industrial Research Organisation, Plant Industry, G.P.O. Box 1600, Canberra ACT 2601, Australia
Fax: +61 2 624 65000
E-mail: e.lagudah@pi.csiro.au

H.-B. Zhang
Department of Soil and Crop Sciences and Crop Biotechnology Center, Texas A&M University, College Station, TX 77843-2123, USA

Key words Binary cosmid · BAC · Wheat D genome · Glutenin · Disease resistance

Introduction

Wheat has been extensively studied because of its economic importance as an agricultural crop species. It also has a highly developed genetic linkage map which is colinear with the diploid "A-" and "D-"genome progenitor species. Wheat has a very large genome with a complex organisation consisting of unique or low-copy sequences surrounded by regions of highly repetitive DNA (Verdel and Delseny 1987). DNA libraries currently available are inadequate for molecular-genetic analysis when dealing with physical mapping of the wheat genome. Recent developments in large DNA fragment manipulation and cloning technologies, such as yeast artificial chromosome (YAC) (Burke et al. 1987) and bacterial artificial chromosome (BAC) (Shizuya et al. 1992) made it possible to overcome these difficulties. Apart from maize and barley YAC libraries and a lettuce BAC library, complete large-insert libraries has been restricted to species with relatively small genomes.

Based on either the conservation of linkage relationships or colinearity observed between genomes of grasses (Kurata et al. 1994; Moore et al. 1995), large-insert libraries from small-genome species like rice or sorghum have been presented as model systems for the cloning of genes of interest in other cereal crops with larger genomes. However, the success of this approach is limited because colinearity tends to break down in some regions of the grass genome due to inversions, translocations and segmental duplications of DNA. Furthermore, cross-species comparisons at a molecular level demonstrated that the nature and the amount of intergenic DNA varies considerably between species (Chen et al. 1997; Moore et al. 1995). Kurata et al.

(1994) also showed that only 40% of a set of wheat single-copy markers could be mapped in rice, indicating that large regions of the wheat genome are absent from rice. Furthermore, of 26 rice YAC clones tested, Dunford et al. (1995) showed that only one YAC insert contained termini where both hybridised to single-copy sequences in the wheat genome. Among the grasses the conservation of linkage relationships closely matching the hexaploid wheat genome are found in their diploid progenitor species.

Tao and Zhang (1998) recently reported the successful cloning of large DNA fragments (over 300 kb) in conventional binary cosmids that were stably maintained in an appropriate *E. coli* culture. Improvements in *E. coli* transformation technology was cited by the authors as a key factor for the successful cloning of the large inserts. The size range of the cloned inserts in the binary cosmids was equivalent to those reported in BAC libraries. In this paper, we report the construction and characterisation of a diploid wheat D-genome large-insert library made in a binary cosmid vector that is suitable for *Agrobacterium*-mediated plant transformation. By using root tissue in comparison with leaf tissue a 30- to 90-fold reduction in the amount of chloroplast (cp) DNA clones was achieved. The usefulness of the present library was demonstrated by the identification of clones carrying a pair of closely linked high-molecular-weight (HMW) glutenin genes and another set of clones containing a family of nucleotide binding site and leucine-rich region (NBS-LRR) sequences tightly linked to the *Cre3* cereal cyst nematode (CCN) resistance locus.

Materials and methods

Plant material

Triticum tauschii accession Aus18913 was the source of genomic DNA for the library construction. This wheat relative carries the *Cre3* (syn. *CcnD1*) gene conferring a high level of resistance to the Australian pathotype of CCN, *Heterodera avenae* (Eastwood et al. 1991), the HMW glutenin subunits Dx2 and DyT2 and an unusual HMW gliadin designated T1 (Lagudah and Halloran 1988) as well as resistance to stem, leaf and stripe rust. Leaf and root samples were collected from seedlings grown at 25°C for 3–5 weeks in hydroponic culture supplemented with nutrients. Samples were frozen in liquid nitrogen and stored at −80°C.

Preparation of HMW DNA

Nuclei were extracted from leaf tissue and embedded in low-melting-point (LMP) agarose (BRL, USA) microbeads as described by Zhang et al. (1995). Root nuclei were isolated using the following procedure. Approximately 5 g of root tissue was ground into powder using a cold mortar and pestle in liquid nitrogen. The powder was suspended in 100 ml of cold homogenisation buffer (HB) (10 m*M* Trizma base, pH 9.4, 80 m*M* KCl, 0.5 *M* sucrose, 10 m*M* Na_2EDTA, 1 m*M* spermidine, 1 m*M* spermine, 0.5% Triton-X 100, 0.15% β-mercaptoethanol). A subsequent homogenisation of this mixture in a glass–glass hand homogenizer was necessary to improve the yield of the nuclei extraction. After filtration through four layers of cheesecloth and two layers of miracloth, the mixture was centrifuged at 4000 *g* for 10 min at 4°C. The pellet was resuspended in 40 ml of HB without spermine and spermidine (HB^-), and aliquots of 10 ml were layered onto cushions made of 5 ml of 80% Percoll in $HB^-$and centrifuged at 4000 *g* for 30 min at 4°C in a swinging bucket centrifuge. The pelleted nuclei, washed once in $HB^-$and twice in HB, was resuspended in 800 µl of HB without Triton-X 100 and without β-mercaptoethanol. A 200-µl aliquot of 2.5% LMP agarose (NuSieve, FMC, Australia) in HB without Triton-X 100 and without β-mercaptoethanol was added to the suspension of nuclei at 50°C, and the mixture was solidified in 25-µl plugs onto an ice-cold surface. Leaf and root DNA was released by digestion with ESP (0.5 *M* EDTA, 1% sodium lauryl sarcosine, 0.25 mg ml^{-1}proteinase K) followed by extensive washes (Zhang et al. 1995). HMW DNA was stored in TE at 4°C without degradation for up to 8 months.

Partial digestion of HMW and size fractionation by pulse field gel electrophoresis (PFGE)

Microbeads (1.5 ml) or plugs (1.5 ml) containing approximately 15 µg of DNA were first equilibrated in 15 ml of 1 × *Hin*dIII reaction buffer (BRL, USA) supplemented with 4 m*M* spermidine and 1 m*M* DTT for 1 h on ice with one change of buffer after 30 min. The supernatant was replaced by 1.5 ml of fresh solution of 1 × *Hin*dIII reaction buffer containing 60 units (U) of *Hin*dIII (BRL, USA) for the microbeads or 240 U of *Hin*dIII for the plugs. The reactions were incubated on ice for an additional 30 min, then transferred into a 37°C water bath and incubated 10 min for the microbeads or 30 min for the plugs. The digestion was stopped with 30 µl of 0.5 *M* EDTA pH 8.0, and the samples were incubated on ice for 30 min. Partially digested DNA was fractionated on a 1% LMP agarose gel in 0.5 × TBE (45 m*M* TRIS-borate, pH 8.0, 1 m*M* EDTA) by PFGE at 6 V cm^{-1}, 95-s switch time for 19 h at 11°C (CHEF DRIII system, BioRad, USA). DNA fractions ranging from 125 kb to 200 kb and 200 kb to 400 kb were excised from the gel and directly used for the second size fractionation in a 0.8% LMP agarose using conditions that focused DNA larger than 150 kb into a thin band (4 V cm^{-1}, 5-s switch time, 9 h, 11°C, 0.5 × TBE). The compressed DNA band was cut and used for the ligation reaction or subjected to a third round of size selection carried out under the same conditions as the second one. The agarose slices were equilibrated for 1 h on ice in 10 m*M* TRIS-HCl, pH 7.5, 1 m*M* EDTA and 40 m*M* NaCl with one change of buffer after 30 min, melted at 68°C for 5 min and digested with 1 U of gelase (Epicentre, USA) per 100 mg LMP agarose for 30 min at 45°C. The enzyme was heat-inactivated at 70°C for 5 min and the DNA solution used directly for the ligation reaction.

Isolation and preparation of vector DNA and ligation reaction

The binary vector pCLD04541 DNA (Jones et al. 1992; Bent et al. 1994; Tao and Zhang 1998) was isolated from a 4-l overnight culture by the alkaline lysis method followed by PEG precipitation and a cesium chloride/ethidium bromide equilibrium centrifugation at 35 000 rpm. for 72 h at 20°C (SW 50.1, Beckman, USA) as described by Sambrook et al. (1989). The cosmid vector was digested to completion with *Hin*dIII (30 U per 10 µg of vector) for 1 h 30 min at 37°C and assayed by gel electrophoresis. After a chloroform/isoamyl alcohol (IAA) extraction and an ethanol precipitation, the vector was dephosphorylated with 0.25 U of calf intestinal alkaline phosphatase (CIAP) (BRL, USA) per picomole of DNA for 30 min at 37°C. CIAP was heat-inactivated at 75°C for 10 min in the presence of 100 m*M* NaCl, 1 m*M* EDTA and 0.25% SDS. The DNA, extracted with chloroform/IAA and precipitated with ethanol, was

resuspended in TE pH 8.0 at a concentration of 20 ng μl^{-1} and directly used in the ligation reaction. The ligation was carried out in a 100- to 150-µl volume in which about 200–400 ng of *Hin*dIII partially digested wheat DNA was ligated to 60–120 ng of vector (molar ratio of about 3 : 1 with vector excess) with 6 U of T4 DNA ligase (UPS, USA) and incubated overnight at 16°C.

Bacterial transformation

The recombinant cosmids were introduced into ElectroMAX DH10B cells (BRL, USA) by electroporation using a Cell Porator and Voltage Booster system (BRL, USA). A 1.5-µl aliquot of ligation mixture was added to 18 µl of *E. coli* cells for a single electroporation. The Cell Porator settings were 350 V, 330 µF capacitance, low ohms impedance and fast charge rate, and the Voltage Booster setting was 4 kohms resistance. After electroporation the cells were transferred to 1 ml SOC solution (2% bacto tryptone, 0.5% bacto yeast extract, 10 m*M* NaCl, 2.5 m*M* KCl, 10 m*M* $MgCl_2$, 10 m*M* $MgSO_4$, 20 m*M* glucose, pH 7.0) and incubated at 37°C with shaking at 200 rpm. for 1 h. The cells were spread on LB plates containing tetracycline (15 µg ml^{-1}), X-gal (30 µg ml^{-1}) and isopropylthio-β-D-galactoside (20 µg ml^{-1}), grown at 37°C for 20 h and stored at 4°C. White colonies containing wheat DNA inserts were transferred to 384-well microtiter plates (Medos, Australia) containing 75 µl of LB-freezing buffer [36 m*M* K_2HPO_4, 13.2 m*M* KH_2PO_4, 1.7 m*M* citrate, 0.4 m*M* $MgSO_4$, 6.8 m*M* $(NH_4)_2SO_4$, 4.4% glycerol, 15 µg ml^{-1} tetracycline, LB] per well, incubated at 37°C for 24 h and stored at − 80°C.

Isolation of recombinant DNA and insert analysis

Clones were streaked out onto LB plates containing 15 µg ml^{-1} tetracycline. Single colonies were picked and grown in 3 ml LB with antibiotic (15 µg ml^{-1} tetracycline) at 37°C for 20 h. The alkaline lysis method (Sambrook et al. 1989) was used for preparing the plasmid DNA, and samples were digested with *Not*I (NEB, USA) for 4 h and then loaded on a 0.8% agarose gel. The gel was subjected to PFGE for 48 h at 370 V in 0.5 × TBE at 16°C using a 4- to 90-s discontinuous ramped switch time (GeneLine II, Beckman, Australia). The insert sizes were estimated by comparison with a lambda mid-range size standard (NEB, Australia) ran in the same gel.

Filter preparation and screening

High-density filters were made by replicating nine 384-well microtiter plates onto single Hybond N + filters (8 × 12 cm) (Amersham, Australia) that had been prewet on LB/agar containing 15 µg ml^{-1} tetracycline using a BIOMEK 2000 robot (Beckman, Australia). Filters were incubated overnight on the same medium. The fixation of plasmid DNA to the filter was performed according to the filter manufacturer's recommendations. DNA hybridisation and washing conditions were the same as described by Lagudah et al. (1991). Probes were labeled using a random priming extension kit (Amersham, Australia) following the manufacturer's instructions.

DNA probes

The chloroplast genome-encoded RNA polymerase B gene from *T. tauschii* was obtained by polymerase chain reaction (PCR) amplification of a 570-bp fragment from leaf DNA using two primers (RPOB5′-CTTCCGAATTATATGTATCCGCG and RPOB3′-CGATTCATATTTCGTCGACCAAC) derived from the maize sequence of the RNA polymerase B gene. The HMW glutenin probe was a subclone containing the 5′ end of the *Glu-B1x17* gene from hexaploid wheat (Reddy and Appels 1993). The G4sp and G12sp probes, specific markers for the CCN resistance gene at the *Cre3* locus, were isolated from lambda clones (Lagudah et al. 1997; unpublished results).

Stability tests of recombinant clones

Two clones containing 120-kb and 130-kb *T. tauschii* DNA inserts, respectively, were grown in 500 ml LB containing 15 µg ml^{-1} tetracycline at 37°C for 24 h. A 1- µl aliquot of these cultures was used to inoculate fresh 500 ml LB with antibiotic for an additional 24 h of growth. This process was repeated for 5 consecutive days. Each passage was considered to represent about 20 generations of *E. coli* (Kim et al. 1992). The *Not*I restriction patterns of the plasmids DNA were checked using PFGE.

Southern blot analysis

Gels stained in ethidium bromide were UV-irradiated for 10 min, denatured for 30 min in 1.5 *M* NaCl and 0.5 *M* NaOH, neutralised for 40 min in 1.5 *M* NaCl, 0.5 *M* TRIS-HCl, pH 7.2, and 1 m*M* EDTA and washed with 2 × SSC (0.3*M* NaCl, 30 m*M* triNa citrate). The transfer onto Hybond N + (Amersham, Australia) membranes was performed with 20 × SSC for 20–36 h. After blotting, the membrane was treated with 0.4 *M* NaOH.

Results

Construction of the library

Several parameters, as described below, were optimised after failure to generate a good library using previously described protocols for large DNA fragment libraries. A relatively high amount of sheared DNA (between 50–800 kb) occurred when HMW DNA was isolated from *T. tauschii* leaves compared with other plant species under the same extraction conditions (Zhang et al. 1995). The same proportion of sheared DNA as well as low yields were obtained when this extraction method was applied to isolate root HMW DNA from *T. tauschii*. In this study, an additional homogenisation step and the purification of intact root nuclei on a Percoll cushion significantly increased the amount of HMW DNA extracted (approximately 10 times) without increasing the proportion of sheared DNA. Purification of the HMW DNA on pulse field gel electrophoresis (PFGE) (Gill et al. 1996) or removal of the small DNA fragments from the plugs on tube electrophoresis run for 24 h at 30 V in TAE resulted in a better quality HMW DNA. However, this purified DNA failed to generate a high number of clones.

The concentration of restriction enzyme (2–4 U) producing the maximum amount of partially restricted DNA in the 200- to 500-kb range generally used to construct large insert libraries in BAC vectors yielded only a few recombinant cosmid clones from *T. tauschii*. The optimal *Hin*dIII concentration (4–8 U) used in this

Table 1 Characteristics of the wheat D-genome large-insert library

Ligation reaction	Source of HMW DNA	Number of size selection steps	Size selection between (kb)	Transformation efficiency[a]	Number of clones picked	Average insert size[b] (kb)	Percentage of clones with cpDNA[c]
L-1	Leaf	3	125–200	> 800	72 960	128	0.3
R-1	Root	3	125–200	> 600	40 320	122	0.01
L-2	Leaf	2	200–300	> 5000	30 720	92	0.9

[a] Number of white colonies per electroporation
[b] Average insert size based on 57, 41 and 59 randomly selected clones for ligation reaction L-1, R-1 and L-2, respectively
[c] Percentage of clones hybridising with RNA polymerase B subunit from the chloroplast (cp) genome

study generated DNA fragments all below 500 kb with the maximum amount of DNA in the 50- to 200-kb range.

The large-insert library from *T. tauschii* consisting of 144 000 clones with an average insert size of 119 kb was obtained from three separate ligation experiments, L-1, R-1 and L-2 (Table 1). In experiment L-2, the DNA was subjected to two rounds of size selection between 200 and 300 kb on a pulse field gel prior to the ligation reaction. This resulted in a high total number of clones (over 500 000). The average insert size distribution from 59 of these clones randomly selected was determined by digestion with *Not*I and separation on PFGE. The smallest and largest clones were 45 and 145 kb, respectively, and 73% of the inserts were distributed between 70 and 110 kb with an average size of 92 kb (Fig. 1). The average insert size was successfully increased through a third round of size selection between 125 and 200 kb in experiments L-1 and R-1 (Table 1). A total number of over 80 000 and over 60 000 clones were obtained in the L-1 and R-1 experiments, respectively. Analysis of 57 randomly selected clones from the L-1 experiments showed that 82% of the clones contained an insert that ranged between 110 and 160 kb. The largest insert was estimated at 161 kb and the smallest at 73 kb. In experiment R-1, analysis of 41 random clones revealed that more than half of the tested clones (56%) contained inserts between 100 and 130 kb; the largest insert was 186 kb and the smallest 68 kb (Fig. 1). The estimated average insert size was 128 and 122 kb, respectively for L-1 and R-1. Parallel experiments demonstrated that three rounds of size selections of DNA between 200 and 300 kb prior to ligation failed to increase the average insert size and reduced the cloning efficiency (data not shown). Out of 157 random clones tested, all contained an insert.

Based on an estimated genome size of 4.7×10^9 kb for the D genome of wheat (Lee et al. 1997), the *T. tauschii* large-insert library reported in this study is equivalent to 3.7 haploid genomes. Theoretically, the probability of identifying a particular clone from this library corresponding to any *T. tauschii* DNA sequence is greater than 97%.

The relative prevalence of *Not*I restriction endonuclease sites in monocot genomes (Wang et al. 1995; Woo et al. 1994) was evident from our *T. tauschii* library. The majority of the clones (82%) contained at least one internal *Not*I restriction site, and over 48% of them are cleaved two to five times with this enzyme (data not shown).

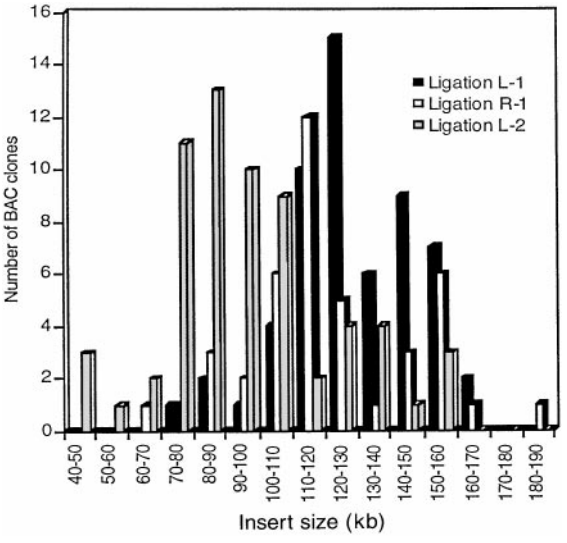

Fig. 1 Insert size distribution in the large-insert library. Fifty-nine random clones from ligation reaction L-1, 41 random clones from ligation reaction R-1 and 59 random clones from ligation reaction L-2 were isolated and sized using PFGE

Frequency of cpDNA clones in the library

Approximately one third of the library was constructed with root DNA, which was expected to reduce the proportion of clones derived from cpDNA. In order to determine the percentage of clones containing chloroplast sequences, 15 360 colonies from each of the three ligation experiments were probed with the RNA polymerase B subunit gene from the *T. tauschii* chloroplast genome (Table 1). On the basis of the RNA polymerase B DNA hybridisation, the clones that originated from leaf DNA were inferred to contain chloroplast-encoded sequences at a frequency of 0.3% (51 clones) of L-1

clones and 0.9% (138 clones) of L-2 clones. By contrast, construction of the large-insert library with root tissue reduced considerably the proportion of clones carrying chloroplast sequences. Only 2 of the R-1 clones tested were identified under the same DNA hybridisation conditions (0.01%). Furthermore, comparison of L-1 and L-2 indicated that an increase in size selection efficiency which increased the average insert size, reduced the fraction of cpDNA clones. Presumably this is due to the relatively small size of the wheat chloroplast genome (Prombona and Subramanian 1991).

Stability of cloned wheat DNA in pCLD04541 binary vector

To investigate the stability of cloned large wheat DNA fragments maintained in *E. coli*, we grew 2 colonies (120 and 130 kb, respectively) for 100 generations, and the *Not*I restriction patterns were compared after 20, 40, 60, 80 and 100 generations. As shown in Fig. 2, the restriction pattern was identical throughout this experiment for both clones. This was confirmed by comparing the *Hin*dIII restriction pattern for each clone across the five generation levels (data not shown). These results indicate that cloned wheat DNA can be stably maintained in the pCLD04541 vector in *E. coli*. Furthermore, repeated culturing of 40 other clones produced a consistent restriction fragment profile upon analysis with six base-pair cutter restriction endonucleases. Tao and Zhang (1998) also reported the stability of 20 cloned large inserts in pCLD04541 derived from sorghum DNA after 5 days of continuous growth in the same *E. coli* strain as used in the present study.

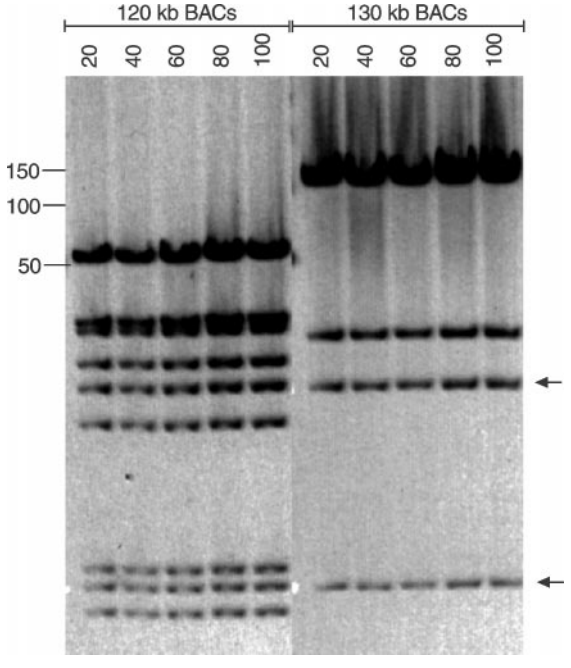

Fig. 2 Comparison of *Not*I restriction enzyme patterns of 2 clones (120 and 130 kb, respectively), after 20, 40, 60, 80 and 100 generations of culturing. *Arrows* denote the vector bands. Sizes are given in kilobases (kb)

Identification of clones containing the HMW glutenin genes

Previous studies on the wheat D-genome donor line, Aus18913, used in constructing the present library showed that it contained the two tightly linked genes, *Glu-D1x* and *-D1y*, encoding polypeptides designated as Dx2 and DyT2, respectively (Lagudah and Halloran 1988). DNA hybridisation analysis using a HMW glutenin probe (Reddy and Appels 1993) that detects *Glu-D1x* and *-D1y* gene sequences served as markers in characterising and confirming the level of gene sequence representation in the *T. tauschii* library. Five distinct clones, referred to as BHMWGlu-1 to -5, were identified as containing wheat sequences of 116–127 kb. Recombinant cosmid DNA isolated from these clones was digested with *Eco*RI enzyme and the restriction fragments separated on PFGE (Fig. 3a). Previous studies in hexaploid wheat demonstrated that the size of the *Eco*RI fragments containing the HMW glutenin genes

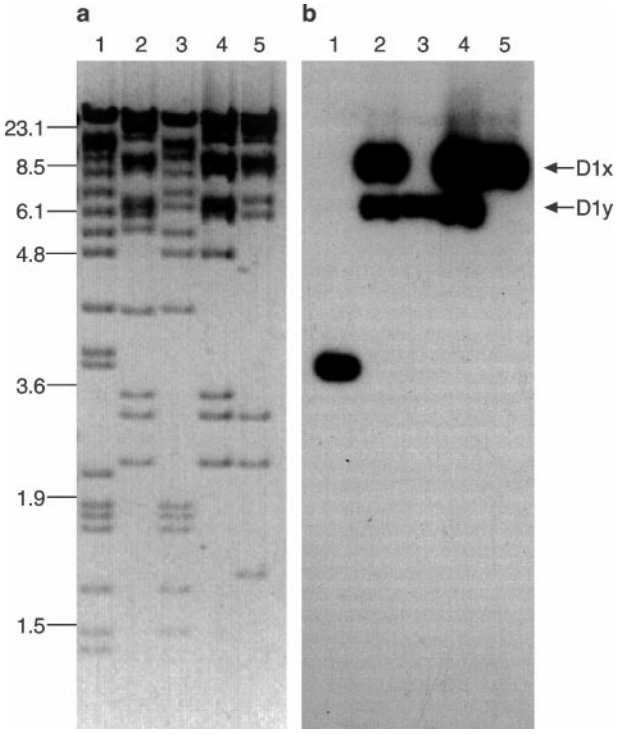

Fig. 3a, b Wheat clones containing the HMW glutenin gene sequences. **a** Ethidium bromide-stained agarose gel showing an *Eco*RI restriction digest of 5 clones (BHMWGlu-1 to -5) identified by screening the library with a probe hybridising with the HMW glutenin genes. **b** Autoradiogram showing Southern blot hybridisation of the gel in **a** with a probe hybridising with the HMW glutenin genes *Glu-D1x* and *D1y*. Sizes are given in kilobases

provide restriction fragment length polymorphisms (RFLPs) diagnostic of the *Glu-1x* and *Glu-1y* loci for their corresponding A, B or D genes (Anderson et al. 1988). The sizes of the *Eco*RI RFLPs from the clones BHMWGlu-2 to -5 were identical to the predicted sizes of the *Glu-D1x* and *-D1y* loci. Both the *Glu-D1x* and *-D1y* genes were present in BHMWGlu-2 and 4, whereas BHMWGlu-3 contained only the *D1y* gene and BHMWGlu-5 possessed only the *D1x* gene (Fig. 3b). A *Hin*dIII digest of these clones confirmed these results and showed that BHMWGlu-1 contained the 5′ end of the *D1y* gene (data not shown).

As the insert size of the smallest clone containing both the *Glu-D1x* and *-D1y* genes (BHMWGlu-4) is 117 kb, we can conclude that the distance separating these two genes is below 117 kb. Based on screening 3.7 genome equivalents and a physical distance separating the two genes deduced to be below the average insert size of the library, at least 4 clones were expected to be identified. The detection of 5 clones with the HMW glutenin probe suggests that the region of chromosome 1D carrying the *Glu-D1x* and *-D1y* loci is correctly represented in the library.

Screening the library with markers cosegregating with the cyst nematode resistance locus, *Cre3*

Another gene region of interest present in the donor line used in the large insert library construction is the *Cre3* locus, which confers resistance to CCN in *T. tauschii* and hexaploid wheat (Eastwood et al. 1991). The library was screened by DNA hybridisation with two *Cre*3 probes, designated G4sp and G12sp, that hybridise to markers that cosegregate with *Cre3* (Lagudah et al. 1997; unpublished results). These markers are found as a single copy in *Cre3*-derived lines. For each of these markers, 4 corresponding clones were identified and referred to as BG4-1 to -4 and BG12-1 to -4. The DNA inserts of these 8 clones ranged from 60 to 140 kb (Fig. 4a). The identification of 4 clones with each probe revealed that the distal part of chromosome 2DL, the location of the *Cre3* locus, is represented in the library as predicted by the screening of 3.7 genome equivalents.

In a previous study, we isolated a gene family predicted to encode a NBS-LRR protein at the *Cre3* locus (Lagudah et al. 1997). A member of this family, CD2 root cDNA, was used as a probe on the *Not*I digest of the DNA extracted from the BG4 and BG12 clones (Fig. 4). Six clones (BG4-1, BG4-3 and BG12-1 to -4) showed one hybridising fragment and the other clones contained two (BG4-2) or three (BG4-4) hybridising sequences.

The BG clones were also digested with *Pst*I which was previously used to map the CD2 RFLPs cosegregating with *Cre3* CCN resistance in a segregating population. The size of the fragments from the clones detected by the CD2 probe was identical to the size of

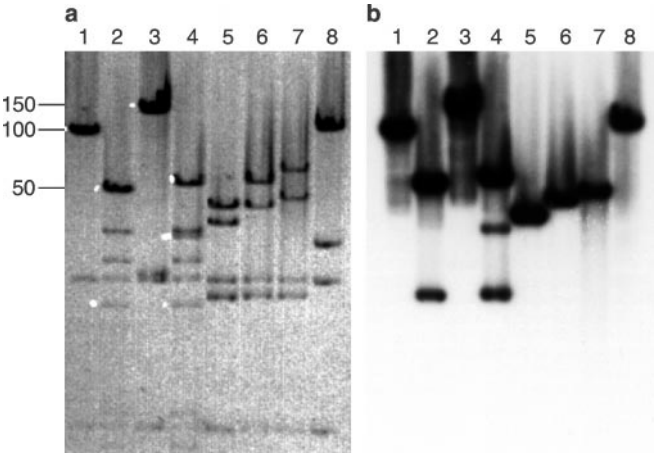

Fig. 4a, b Wheat clones originating from the *Cre3* cyst nematode resistance locus. **a** Ethidium bromide-stained agarose gel showing a *Not*I digest of the clones BG4-1 to -4 (*lanes 1–4*) and BG12-1 to -4 (*lanes 5–8*) identified by screening the library with the G4sp and G12sp probes cosegregating with the *Cre3* locus. **b** Autoradiogram showing Southern blot hybridisation of the gel in **a** with the CD2 probe containing the *Cre*-derived NBS-LRR sequences. Sizes are given in kilobases

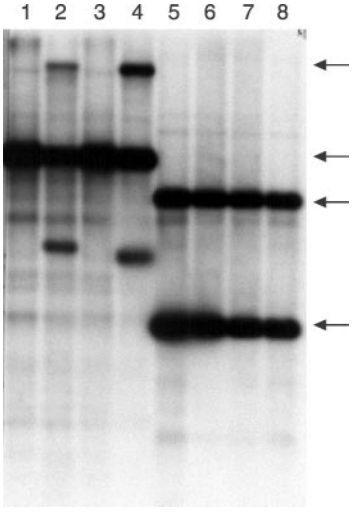

Fig. 5 *Pst*I restriction digest patterns of the BG4-1 to -4 (*lanes 1–4*) and BG12-1 to -4 (*lanes 5–8*) clones probed with CD2 containing the *Cre*-derived NBS-LRR sequences. *Arrows* denote the positions of the RFLPs previously mapped at the *Cre3* locus

*Pst*I RFLPs that cosegregate with *Cre3* in genomic DNA of the donor line (Fig. 5). This analysis confirmed that the BG clones were derived from sequences located at the *Cre3* locus. The markers to which G4sp and G12sp hybridise were never found in a single clone. We are currently investigating the BG clones in more detail to construct a complete contig of the *Cre3* region.

Discussion

Construction and characterisation of a *Triticum tauschii* large-insert library

A wheat D-genome library of 144 000 clones with an average insert size of 119 kb was constructed

representing 3.7 genome equivalents. Screening of the library with 3 clones from regions of well-defined gene sequences suggests that our numerical estimates for clone representation is correct.

The correlation between a single and a double round of size selections of genomic DNA prior to ligation with the average insert size of BAC clones has already been demonstrated (Woo et al. 1994; Zhang et al. 1996). Two size-selection steps resulted in an increase in the average insert size of 30 kb for a rice BAC library and 68 kb for a sorghum BAC library. In the present study, two size selections resulted in a relatively small average insert size (92 kb), presumably due to the poor quality of the HMW DNA extracted from *T. tauschii*. An increase of over 30 kb in the average insert size was obtained with a third round of size selection. This was attributed to an efficient reduction in the amount of small DNA fragments entrapped in the large DNA molecules on PFGE. Two size selections were insufficient to eliminate the entrapped DNA fragments. Similar observations have been made in the construction of plant BAC libraries where the average insert sizes were approximately 100–200 kb below those of the smaller selected DNA fragments prior to ligation (Danesh et al. 1998; Frijters et al. 1997; Wang et al. 1995; Woo et al. 1994; Zhang et al. 1996). When three size selections were performed, the average insert size (122 and 128 kb) was close to the size of the DNA selected for ligation on PFGE (125–200 kb).

The quantity and quality of input genomic DNA has been shown to be two of the most critical factors for the construction of a large-insert library. When the HMW DNA extraction results in significant shearing, a lower transformation efficiency has to be compensated for by increasing the amount of partially digested DNA loaded on PFGE. This leads to excessive trapping of smaller DNA molecules which co-migrate with larger DNA fragments and decreases significantly the average insert size (Frijters et al. 1997; data not shown). A third round of size selection reduced the effect of the quality of input DNA on the cloning efficiency by allowing more flexibility in the quantity of DNA used on PFGE without decreasing the average insert size.

Increasing the number of size-selection steps reduced the cloning efficiency as demonstrated in previous studies. With three rounds of size selection, we generated 600–800 recombinants per electroporation. This is higher than the cloning efficiency obtained for most of the BAC libraries constructed. The same transformation efficiency was observed when the pBeloBAC II vector was used instead of the pCLD04541 vector (data not shown). As Danesh et al. (1998) reported, an efficient dephosphorylation of the vector was essential for a high cloning efficiency. We found that besides the dephosphorylation, the purification of the vector by a CsCl gradient improved significantly the number of transformants, probably because it eliminated contaminations with linearised plasmid which may be ligated at one end of the plant DNA thereby preventing the formation of a circular plasmid.

Reduction of chloroplast-encoded sequences in the clones by the use of root nuclei

A high frequency of cpDNA clones in large-insert libraries increases the number of clones required to be generated and screened in order to optimise the chances of isolating a particular nuclear DNA clone. Consequently, a minimum proportion of clones originating from organellar DNA is preferable in a library. We have demonstrated that the frequency of clones containing cpDNA sequences can be considerably reduced by the use of HMW DNA isolated from root tissue.

Cheung and Gale (1990) extracted high-quality HMW DNA using leaf protoplasts from cereal crop species. Such high-quality HMW DNA is not easily obtained from intact wheat nuclei (Cheung and Gale 1990; Gill et al. 1996). The quality of the HMW DNA is a critical factor in the construction of large-insert libraries. However, the majority of the large-insert libraries made with HMW DNA isolated from leaf protoplasts have a high percentage (5–26%) of chloroplast clones. The use of protoplasts isolated from root tissue may be a means of obtaining high-quality HMW DNA for the construction of large-insert libraries without the inconvenience of a high frequency of clones derived from the chloroplast genome.

Usefulness of the large-insert library for molecular analysis in the D genome of wheat

Apart from insert size, the stability of the clones, genome representation and contamination by organellar DNA clones, the quality of a large-insert library may be best assessed by its potential to provide specific clones corresponding to chromosomal regions of interest. So far, 12 gene sequences, including storage protein genes, a range of starch synthesis and modifying genes, and different classes of putative resistance gene sequences used in screening the present library have led to the isolation of all of the targeted large insert clones (K. Turnbull, S. Rahman and W. Spielmeyer, personal communication). This library represents the first representative bacterial large-insert library for a Triticeae species and is thus suitable for chromosome walking.

The HMW glutenins, a multigene family encoding storage proteins, are one of the most thoroughly studied gene families in relation to wheat quality. They are located in the long arm of the group-1 homoeologous chromosomes at the *Glu-A1*, *Glu-B1* and *Glu-D1* loci of the hexaploid wheat. Each locus contains two HMW glutenin genes (Anderson et al. 1988) which have rarely been separated by recombination. Analysis of the flanking regions surrounding the HMW glutenin genes have been undertaken in order to understand the relative levels of expression of these genes and the degree of accumulation of individual HMW glutenins. PCR amplification methods were successfully applied to analyse approximately 1.5 kb of promoter sequences (Anderson et al. 1998). No conclusive evidence for levels of gene expression was found based on the limited flanking sequences. However, owing to the lack of a wheat large-insert library, information on the organisation and chromosomal context of the *Glu-1* locus is still limited. Our data suggest that we have cloned a region of some 230 kb from the *Glu-D1* locus which can be further extended by chromosome walking.

Since the first application of the large-insert cloning system to plant genomes (Guzman and Ecker 1988), several agronomically important genes, such as disease resistance genes, have been isolated and characterised by positional cloning with the aid of large-insert libraries in tomato, sugar beet and rice. One of our objectives in making the present large-insert library was to construct a contig spanning *Cre3* (Eastwood et al. 1991; Lagudah et al. 1997). The identification of two contigs

originating from the *Cre3* locus will greatly facilitate the identification of functional and disrupted homologues of the gene family for cyst nematode resistance.

Clones containing less than 10 kb of single-copy DNA are often not effective in detecting signals from fluorescence in situ hybridisation (FISH) in plants. DNA libraries with large inserts are thus required. Recent studies have demonstrated that BAC clones are useful tools for the physical mapping of single or low-copy sequences in barley (Lapitan et al. 1997) and rice (Jiang et al. 1995) chromosomes by FISH. The availability of a *T. tauschii* large-insert library will enable physical mapping of single- or low-copy sequences by FISH in order to correlate the genetic and physical maps of the wheat D genome.

In plant systems, the verification that a potential clone isolated by positional cloning carries the gene of interest requires complementation by the tedious step of plant transformation. Stable transformation of tobacco plants with 150 kb of human DNA was achieved (Hamilton et al. 1996) using a binary BAC vector (BIBAC) (Hamilton 1997). The construction of the present diploid wheat large-insert library in a binary vector designed to introduce foreign DNA into plant genomes by *Agrobacterium*-mediated transformation makes this library a valuable resource. The ability to test large-insert clones in transformation experiments as a first step toward delineating candidate genes is of crucial importance, particularly in plants with large genomes such as wheat. The recent success of the *Agrobacterium* transformation technology in monocot plant species including rice, maize, barley and wheat, will facilitate the process of functional gene verification from the binary large-insert library.

Acknowledgements We are grateful to Peggy Kooij, Sutha Chandramohan, Libby Viccars and Robin Chapple for their excellent support, Sadiq Rahman for providing the glutenin clones and Bill Taylor for the maize RNA polymerase B oligonucleotides. This work was partly funded by the Grains Research and Development Corporation, Australia.

References

Anderson OD, Halford NG, Forde J, Yip RE, Shewry PR Greene FC (1988) Structure and analysis of the high molecular weight glutenin subunits from *Triticum aestivum* L. cv Cheyenne. In: Miller TE, Koebner RMD (eds) Proc 7th Int Wheat Genet Symp. Bath Press, Bath, UK, pp 699–704

Anderson OD, Abraham-Pierce FA, Tam A (1998) Conservation in wheat high-molecular-weight glutenin gene promoter sequences: comparisons among loci and among alleles of *GLU-B1-1* locus. Theor Appl Genet 96: 568–576

Bent A, Kunkel BN, Dahlbeck D, Brown KL, Schmidt R, Giraudat J, Leung J, Staskawicz BJ (1994) *RPS2* of *Arabidopsis thaliana*: a leucine-rich repeat class of plant disease resistance genes. Science 265: 1856–1860

Burke DT, Carle GF, Olson M (1987) Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. Science 236: 806–812

Chen M, SanMiguel P, DeOliviera AC, Woo S-S, Zhang H, Wing RA, Bennetzen JL (1997) Microcolinearity in *sh2*-homologous regions of the maize, rice and sorghum genomes. Proc Natl Acad Sci USA 94: 3431–3435

Chueng WY, Gale MD (1990) The isolation of high molecular weight DNA from wheat, barley and rye for analysis by pulse-field gel electrophoresis. Plant Mol Biol 14: 881–888

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

Dunford RP, Kurata N, Laurie DA, Money TA, Minobe Y, Moore G (1995) Conservation of fine-scale DNA marker order in the genomes of rice and Triticeae. Nucleic Acids Res 23: 2724–2728

Eastwood RF, Lagudah ES, Appels R, Hannah M, Kollmorgen JF (1991) *Triticum tauschii*: a novel source of resistance to cereal cyst nematode (*Heterodera avenae*). Aust J Agric Res 42: 69–77

Frijters ACJ, Zhang Z, van Damme M, Wang G-L, 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

Gill KS, Nasuda S, Gill BS (1996) Isolation, cloning and gel blot analysis of high molecular weight wheat DNA. BioTechniques 21: 572–576

Guzman P, Ecker J (1988) Development of large DNA methods for plants: molecular cloning of large segments of *Arabidopsis* and carrot DNA into yeast. Nucleic Acids Res16: 11091–11105

Hamilton CM (1997) A binary-BAC system for plant transformation with high-molecular-weight DNA. Gene 200: 107–119

Hamilton CM, Frary A, Lewis C, Tanksley D (1996) Stable transfer of intact high molecular weight DNA into plant chromosomes. Proc Natl Acad Sci USA 93: 9975–9979

Jiang J, Gill BS, Wang G, Ronald PC, Ward DC (1995) Metaphase and interphase fluorescence in situ hybridization mapping of the rice genome with bacterial artificial chromosomes. Proc Natl Acad Sci USA 92: 4487–4491

Jones JDG, Shlumukov L, Carland F, English J, Scofield SR, Bishop GJ, Harrison K (1992) Effective vectors for transformation, expression of heterologous genes, and assaying transposon excision in transgenic plants. Transgenic Res 1: 285–297

Kim U-J, Shizuya H, de Jong PJ, Birren B, Simon MI (1992) Stable propagation of cosmid sized human DNA inserts in an F factor based vector. Nucleic Acids Res 20: 1083–1085

Kurata N, Moore G, Nagamura Y, Foote T, Yano M, Minobe Y, Gale M (1994) Conservation of genome structure between rice and wheat. Bio/Technology 12: 276–278

Lagudah ES, Halloran GM (1988) Phylogenetic relationships of *Triticum tauschii* the D genome donor to hexaploid wheat. 1. Variation in HMW subunits of glutenin and gliadins. Theor Appl Genet 75: 592–598

Lagudah ES, Appels R, McNeil D, Brown AHD (1991) The molecular-genetic analysis of *Triticum tauschii*, the D genome donor to hexaploid wheat. Genome 34: 375–386

Lagudah ES, Moullet O, Appels R (1997) Map-based cloning of a gene sequence encoding a nucleotide-binding domain and a leucine-rich region at the *Cre3* nematode resistance locus of wheat. Genome 40: 659–665

Lapitan NLV, Brown SE, Kennard W, Stephens JL, Knudson DL (1997) FISH physical mapping with barley BAC clones. Plant J 11: 149–156

Lee J-H, Yen Y, Arumuganathan K, Baenziger PS (1997) DNA content of wheat monosomics at interphase estimated by flow cytometry. Theor Appl Genet 95: 1300–1304

Moore G, Devos KM, Wang Z, Gale MD (1995) Grasses, line up and form a circle. Curr Biol 5: 737–739

Prombona A, Subramanian AR (1991) Construction of a chloroplast DNA library from rye (*Secale cereale*), a cereal plant of temperature-inducible chloroplast ribosome deficiency. Biochem Int 24: 559–566

Reddy P, Appels R (1993) Analysis of a genomic DNA segment carrying the wheat high-molecular-weight (HMW) glutenin Bx17 subunit and its use as an RFLP marker. Theor Appl Genet 85:616–624

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, 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 of *Escherichia coli* using an F-factor-based vector. Proc Natl Acad Sci USA 89: 8794–8797

Tao Q, Zhang H-B (1998) Cloning and stable maintenance of DNA fragments over 300 kb in *Escherichia coli* with conventional plasmid-based vectors. Nucleic Acids Res 26: 4901–4909

Verdel F, Delseny M (1987) Repetitivity and variability of higher plant genomes. Plant Physiol Biochem 25:191–210

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

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