Research Article

BAC Libraries from Wheat Chromosome 7D: Efficient Tool for Positional Cloning of Aphid Resistance Genes

Hana Šimková,1 Jan Šafář,1 Marie Kubaláková,1 Pavla Suchánková,1 Jarmila Čihalíková,1 Heda Robert-Quatre,1 Perumal Azhaguvel,2 Yiqun Weng,2,3 Junhua Peng,4,5 Nora L.V. Lapitan,4 Yaqin Ma,6 Frank M. You,6 Ming-Cheng Luo,6 Jan Bartoš,1 and Jaroslav Doležel1

1 Laboratory of Molecular Cytogenetics and Cytometry, Institute of Experimental Botany, Sokolovská 6, CZ 772 00 Olomouc, Czech Republic
2 Texas AgriLife Research, 6500 Amarillo Boulevard W., Amarillo, TX 79106, USA
3 Vegetable Crops Research Unit, USDA-ARS, Department of Horticulture, University of Wisconsin, Madison, WI 53706, USA
4 Department of Soil and Crop Sciences, Colorado State University, Fort Collins, CO 80523, USA
5 Key Laboratory of Plant Germplasm Enhancement and Specialty Agriculture, and Wuhan Botanical Garden, Chinese Academy of Science, Wuhan, Hubei 430074, China
6 Department of Plant Sciences, University of California, Davis, CA 95616, USA

Correspondence should be addressed to Hana Šimková, simkovah@ueb.cas.cz

Received 30 July 2010; Accepted 26 October 2010

Academic Editor: Angela Sessitsch

Copyright © 2011 Hana Šimková et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Positional cloning in bread wheat is a tedious task due to its huge genome size and hexaploid character. BAC libraries represent an essential tool for positional cloning. However, wheat BAC libraries comprise more than million clones, which makes their screening very laborious. Here, we present a targeted approach based on chromosome-specific BAC libraries. Such libraries were constructed from flow-sorted arms of wheat chromosome 7D. A library from the short arm (7DS) consisting of 49,152 clones with 113 kb insert size represented 12.1 arm equivalents whereas a library from the long arm (7DL) comprised 50,304 clones of 116 kb providing 14.9x arm coverage. The 7DS library was PCR screened with markers linked to Russian wheat aphid resistance gene DnCl2401, the 7DL library was screened by hybridization with a probe linked to greenbug resistance gene Gb3. The small number of clones combined with high coverage made the screening highly efficient and cost effective.

1. Introduction

Bread wheat (Triticum aestivum L.) is one of the most important crop species providing the staple food for 40% of the world’s population. As with other crops, the yields of wheat are annually significantly reduced due to attack of a large variety of pests and diseases. Supplementation of conventional wheat breeding for pathogen resistance with marker-assisted breeding and direct gene transfer by molecular methods promises to enhance the efficiency of plant breeding. Prerequisites of this approach are the saturation of genetic maps in the region of interest and isolation of the resistance gene. Positional or map-based cloning is an experimental approach to isolate unknown genes based on their position in the genome. This approach involves construction of a high-density genetic map covering the target locus and a physical map spanning the region of interest. Physical maps are produced by ordering DNA clones from large-insert (usually BAC) libraries into contigs on the basis of clone fingerprint pattern.

Whereas some wheat genes can be mined from smaller genomes of related or model species, many agronomically important genes, including those for pathogen resistance, yield, or grain quality factors, can only be extracted from
the hexaploid wheat genome. However, only a few bread wheat genes, including those conferring pathogen resistances, grain protein content, vernalization requirement, and domestication traits [1–11], have been successfully cloned. This is mainly due to specific features of the wheat genome. *T. aestivum* is an allohexaploid species (2n = 6x = 42, AABBDD genome), which originated from a spontaneous hybridization of three diploid wheat ancestors, donors of the A, B, and D genomes, respectively. They contributed to the enormous size of the bread wheat genome (1C∼17 Gbp), which is composed of ∼1% of genes [12, 13] interspersed by huge amounts of repetitive elements.

A variety of genomic resources has been developed to enable positional cloning in bread wheat including a number of BAC libraries [14–18]. They comprise ∼400,000 to 1,200,000 clones, representing 3.1–9.3 genome equivalents. While fingerprinting so many clones is technically feasible using the SNaPshot-based HiCF procedure [19] and recent results indicate that existing computational techniques may allow for reliable assignment of fingerprinted BAC clones to particular homoeologous chromosomes [20], handling and screening libraries composed of more than million clones remain expensive and tedious tasks. These obstacles can be overcome by working with a library derived from a smaller part of the wheat genome.

One option is to use diploid genome-donor species or their relatives as surrogates, taking advantage of smaller genome size and absence of polyploidy. Several BAC libraries of diploid wheat progenitors have been constructed including those of *T. monococcum* [21] and *Ae. speltoides* [22], close relatives of the A- and B-genome ancestors, respectively. Both BAC and BiBAC libraries were produced from *Ae. tauschii*—the wheat D-genome donor [22–25]. Some of the diploid wheat libraries and the arising *Ae. tauschii* physical map (see [25], http://www.wheat.ucdavis.edu) were employed in cloning agronomically important genes [1, 2, 4–7, 11]. However, clone numbers in these libraries are still too large to be screened easily. Moreover, wheat genomes have undergone revolutionary changes following the polyploidization events including losses of DNA. This partial diploidization and other genomic changes [26–29] suggest that physical maps and genomic sequences of wheat diploid ancestors, although useful resources for wheat genomics, cannot fully substitute for the genomic sequence of hexaploid wheat itself.

Recently, we proposed an alternative approach utilizing flow cytometry to dissect the hexaploid wheat genome into small fractions—chromosomes or chromosome arms, which represent only a few percent of the hexaploid wheat genome [30]. The chromosome-based strategy was made possible by developing procedures for purification of particular wheat chromosomes by flow sorting [31] and a protocol for preparation of intact DNA from sorted chromosomes suitable for cloning [32]. This allowed us to create the first ever BAC library derived from a single chromosome of a higher eukaryote [33]. Our ability to purify single chromosome arms from hexaploid wheat [34] enabled us to construct BAC libraries from 13 of the 21 wheat chromosomes thus far (see [35–37], http://olomouc.uab.cas.cz/dna-libraries/cereals), as well as from the short arm of rye 1R chromosome, which is a component of a number of wheat varieties [38].

Here, we report on the construction of BAC libraries from both arms of wheat chromosome 7D, which harbors numerous genes underlying agronomically important traits, mainly resistance genes. The long arm of 7D is known to carry several greenbug (*Schizaphis graminum*) resistance genes [39, 40] as well as QTLs influencing Russian wheat aphid (RWA, *Diuraphis noxia*) resistance [41]. The short arm of 7D comprises several major genes and QTLs for RWA resistance [41–44], the recently dissected *Lr34* locus underlying resistance to leaf rust, yellow rust, stem rust and powdery mildew [11], genes *Stb4* and *Stb5* underlying resistance to *Septoria tritici* blotch [45, 46] as well as several yield-related QTLs [47, 48].

In order to demonstrate how the chromosome-arm-specific BAC libraries can facilitate positional gene cloning in a supersized plant genome, we report on a hybridization-based screening of the 7DL library with markers for the greenbug resistance gene *Gb3*, and a PCR-based screening of the 7DS library with markers for Russian wheat aphid resistance gene *DnCI2401*.

### 2. Materials and Methods

#### 2.1. Plant Material

A double ditelosomic line of wheat *Triticum aestivum* L. cv. Chinese Spring carrying both arms of chromosome 7D as telosomes (2n = 40 + 27DS + 27DL) was used to sort the arms. The seeds were provided by Dr. Bikram Gill (Kansas State University, Manhattan, USA). Four thousand two hundred seeds subdivided into 184 batches of 20–25 were germinated in the dark at 25°C on moistened filter paper for 55–60 h to reach root length of 2–3 cm.

#### 2.2. BAC Libraries

##### 2.2.1. Preparation of Chromosome Suspensions and Sorting of Chromosomes

Cell-cycle synchronization, accumulation of metaphases in root tips and preparation of chromosome suspensions were performed as described [31]. Briefly, chromosome suspensions were prepared by mechanical homogenization of 20–25 root-tip meristems enriched for metaphase cells in 1 ml ice-cold isolation buffer (IB, [32]). Chromosomes in suspension were stained with 2 µl of DAPI (4',6-diamidino-2-phenylindole) and analyzed using a FACS Vantage SE flow cytometer (Becton Dickinson, San Jose, USA). Both arms were sorted separately from the same sample in batches of 200,000 into 320 µl of 1.5 x IB. The purity in sorted fractions was checked regularly by FISH with probes for telomeric and GAA repeats as described [36].

##### 2.2.2. BAC Library Construction

Preparation of high molecular weight DNA (HMW DNA) and library construction were performed as described [32, 33] with some modifications. Briefly, each batch of flow-sorted arms was spun down and the pelleted chromosomes were mixed with low-melting point agarose to form 20-µl miniplugs, which were incubated in lysis buffer containing proteinase K to purify...
the chromosomal DNA. The isolated HMW DNA was partially digested with HindIII (New England Biolabs, Beverly, Mass., USA) and subjected to two rounds of size selection. At the first round, the partially digested DNA was size-separated in 1% SeaKem Gold Agarose gel (Lonza, Rockland, Ill., USA) in 0.25x TBE under the following conditions of pulsed-field gel electrophoresis (PFGE): voltage 6 V/cm, switch time 1–50 s, run time 17 hours. Size fraction of 100–210 kb was excised from the gel and split into two parts. Fraction B comprised fragments of 100–150 kb whereas fraction M represented a fraction of 150–210 kb. Both fractions were subjected to a second round of size selection in 0.9% SeaKem Gold Agarose gel in 0.25x TBE under the following conditions: voltage 6 V/cm, switch time 3 s, run time 17 h. A gel zone corresponding to 100–150 kb was excised from the lane containing the B fraction and was subdivided into two slices comprising a fraction of 100–120 kb (B1) and 120–150 kb (B2), respectively, whereas only one gel slice (~150–200 kb) was excised from the M lane. The DNA of particular fractions was electroeluted from the gel and ligated into HindIII-digested dephosphorylated pBludigoBAC-5 vector (Epigenetix, Madison, Wisc., USA) in 1:4 molar ratio (DNA:vector). The recombinant vector was used to transform Escherichia coli ElectroMAX DH10B competent cells (Invitrogen, Carlsbad, Calif., USA). The library was ordered by Qbot (Genetix, New Milton, UK) into 384-well plates filled with 75 µl freezing medium consisting of 2YT [49], 6.6% glycerol and 12.5 µg/ml chloramphenicol. The clones were stored at −80°C.

2.2.3. Insert-Size Analysis. A total of 121 BAC clones from the 7DL (63 from the B1, 32 from the B2, and 27 from the M fraction) and 184 clones from the 7DS library (56 from the B1, 76 from the B2 and 52 from the M fraction) were analyzed to estimate average insert size and percentage of empty BAC clones. The BAC DNA was isolated after overnight incubation of particular BAC clones in 1.5 ml 2YT supplemented with 12.5 µg/ml chloramphenicol using a standard alkaline lysis method. The DNA was digested with 0.4 U NorI in 20-µl reaction volumes. Resulting DNA fragments were separated in a 1% agarose gel in 0.25x TBE buffer by PFGE. The size of the fragments was estimated by comparing with two size markers: Lambda Ladder PFG Marker and MidRange Marker I (New England Biolabs).

2.3. Hybridization Screening

2.3.1. Preparation of Colony Filters. High-density colony filters were prepared from the 7DL-specific BAC library by spotting BAC clones in duplicate in a 4 × 4 gridding pattern onto Hybond N+ nylon membranes (GE Healthcare, Piscataway, N.J., USA) using the GeneTAC G3 robot (Genomic Solutions, Ann Arbor, Mich., USA). The entire library was placed on three 22 × 22.5 cm filters. The filters were treated as described [22] with some modifications.

2.3.2. Preparation of Hybridization Probes and Hybridization. An EST-derived STS marker (STS-Aug-08-28) was mapped from genomic DNA of a greenbug-susceptible Ae. tauschii accession AL8/78. The PCR product was separated by agarose gel electrophoresis and a fragment of desired size was eluted from the gel using Qiaquick Gel Extraction kit (Qiagen, Germantown, Md., USA). The fragment was then 32P-dATP radio-labeled by NEBlot kit (New England Biolabs) using manufacturer’s protocol. The labeled probe was purified in a Sephadex G50 column (GE Healthcare) and denatured at 100°C for 10 min. For prehybridization, overnight incubation of colony filters in hybridization solution (2x SSPE, 0.5%SDS, 5x Denhardt’s reagent [48], 40 µg/ml herring sperm DNA) was done in rotary glass tubes at 65°C. The labeled probe was mixed with 5 ml of hybridization solution and colony filters were incubated at 65°C overnight. To remove the unbound probe, we washed the filters twice in washing solution containing 2x SSPE and 0.5% SDS and rinsed with 1x SSC. The washed filters were exposed to X-ray film for one to three days based on the signal intensity to identify positive clones. To complete the assembly of the contig spanning the Gb3 region, we conducted three rounds of 7DL BAC library screening. The STS-Aug-08-28 marker was used for the first round of the screening whereas probes derived from protruding ends of BAC no. 22 and BAC no. 25, respectively, were used for the second and third round of screening (Figure 3).

2.3.3. Building a Local Contig Spanning the Gb3 Region. The positive BAC clones resulting from each BAC library screening were grown overnight in LB media supplemented with 12.5 µg/ml chloramphenicol. BAC DNA was extracted by standard alkaline lysis procedure [49] and subjected to HindIII digestion to create a fingerprint for each positive clone. Overlaps between the clones were detected manually by identifying shared fragments in the fingerprints. Five micrograms DNA of all positive clones in each screen were extracted with QIAGEN plasmid kit (Qiagen) and used for direct cycle sequencing of BAC ends using T7 (5′ TAATACGACTCACTATAGGG 3′) and M13R (5′ CAG- GAAACAGCTATGACC 3′) primers. Moreover, three BAC clones (BAC 22, BAC 25 and BAC 72) were sequenced completely by Sanger technology at the Genome Center, Washington University, St. Louis, Mo., USA. The alignment of BAC end sequences with the fully sequenced BAC clones was done with Seqman Module in the software DNASTar (Lasergene Corp., Madison, Wisc., USA). The insert size of the positive BAC clones was determined by aligning BAC end sequences with the reference sequences of the completely sequenced BAC clones. BAC end sequences were used to develop sequence tagged site (STS) or cleaved amplified polymorphic sequence (CAPS) markers, which were used for further screening of the library to identify and/or confirm overlapping BACs to span the region of interest.

2.4. PCR Screening

2.4.1. Clone Pooling. To enable PCR screening of the 7DS library, we generated three types of BAC pools: plate pools, superpools and three-dimensional (3D) pools. First,
128 plate pools were prepared by pooling clones from each of 584-well plates comprising the library. Rectangular dishes (Nunc, Rochester, NY, USA) containing agarose (1.6%) 2YT medium with 12.5 µg/ml chloramphenicol were inoculated with all 384 clones of single plates using a GeneTAC G3 robot. Each clone was spotted 16x. After 16-h incubation at 37°C, the clones were suspended in 5–10 ml TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0), bacteria were pelleted by centrifugation and subsequently resuspended in 600 µl GET buffer (50 mM glucose, 25 mM Tris, 10 mM EDTA, pH 8.0). The suspension was subdivided into six aliquots, which were processed separately. DNA was isolated using standard alkaline lysis protocol [49] supplemented with RNAse treatment and precipitation of contaminating compounds by ammonium acetate. The DNA of each aliquot was dissolved in 20 µl sterile deionized water; the aliquots were combined and stored at −20°C. For PCR screening these stocks were diluted to 10 ng/µl DNA.

To prepare the 3D pools, we subdivided the library into 16 stacks, each comprising 8 plates. Plate, row and column pools were generated for each of the 16 stacks. We created 48 pools (8 plate, 16 row and 24 column pools) for each stack, thus totally 768 pools represented the entire library. To prepare the pools, we spotted bacterial clones in duplicate on plates with agarose medium as described above. Bacteria grown for 48 h were suspended in 5 ml TE buffer and boiled for 30 min to release DNA. Remnants of bacteria were pelleted at 3000 g for 60 min and 1.4 ml of the supernatants were deposited at −20°C as the particular pools. For PCR screening these stocks were diluted 50x with sterile deionized water. Superpools were created for each of the 8-plate stacks by combining all clones from the respective 8 plates. Spotting, growth of bacteria and DNA isolation were done as for the 3D pools.

2.4.2. PCR Screening Procedure. The 7DS-specific library was screened with microsatellite markers Xefa68, Xgwm473 and Xbarc214 closely linked to the Dn gene. Primer sequences were retrieved from the GrainGenes database (http://wheat.pw.usda.gov/GG2/index.shtml). The PCR reaction mix (10 µl volume) consisted of 2 µl template DNA, 1× Buffer for DyNAzyme DNA Polymerase (Finnzymes, Espoo, Finland), 0.01% Cresol Red, 1.5% saccharose, 0.2 mM dNTPs, 1 µM primers and 0.4 U DyNAzyme II DNA Polymerase (Finnzymes). The PCR reaction was performed under the following conditions: denaturation for 5 min at 95°C followed by 35 cycles of denaturation at 95°C for 30 s, annealing at the appropriate temperature for 30 s, extension at 72°C for 30 s, and a final extension at 72°C for 10 min. The presence of PCR products was detected by electrophoresis in 1.2% agarose gel run in 0.5x TBE buffer. The precise size of individual PCR products was determined using ABI 3730xl sequencer (Applied Biosystems, Foster City, Calif., USA). The fragment sizes were estimated relative to the GeneScan-500 LIZ Size Standard (Applied Biosystems) by GeneMarker V1.75 (SoftGenetics, LLC, USA).

2.4.3. Fingerprinting of BAC Clones and Contig Assembly. All clones of the 7DS-specific library were fingerprinted using the SNaPshot-based high-information-content fingerprinting technology [19]. The fingerprints were automatically edited with the computer program package FPMiner (www.bioinforsoft.com) and GenoProfiler [50] and assembled using FPC V9.3 [51] at an initial cutoff of 1 × 10−45, followed by various steps of DQing and end merging.

3. Results and Discussion

3.1. Flow-Sorting and Purity of the Sorted Fractions. The short and the long arms of the 7D chromosome were sorted simultaneously from the 7D double-ditelosomic (DDt) 7D line. The flow karyotype obtained from this line consisted of two peaks representing 7DL and 7DS telocentric chromosomes, respectively, three composite peaks comprising groups of various wheat chromosomes and the rightmost peak corresponding to the largest wheat chromosome 3B (Figure 1). The leftmost peak, representing a chromosome with the smallest relative DNA content, actually comprised telosome 7DL and not 7DS, as the arm designated 7DS based on homology with the wheat chromosome arms 7AS and 7BS is physically longer and hence has greater relative DNA content [52]. The purity of sorted telosomes as estimated by FISH varied between 84–92% (average 88.8%) for the 7DL and between 79–86% (average 84.1%) for the 7DS telosome. The 7D chromosome is metacentric and the 7DL and 7DS arms are of similar size (2.04% of the wheat genome alias 346 Mbp for 7DL and 2.25% of the wheat genome alias 381 Mbp for 7DS, [37, 53]). Consequently the greatest number of contaminating particles came from the opposite arm. The 7DS constituted 1.3% of the sorted 7DL whereas 7DL represented 1.1% of the sorted 7DS fraction. The remaining contamination was composed of a mixture of chromosomes without a prevalence of a particular type.

3.2. BAC Libraries. In total, 31 agarose miniplugs comprising 5,900,000 7DL telosomes corresponding to ~4.18 µg DNA were prepared. In parallel, 6,000,000 7DS telosomes (~4.67 µg) were embedded in 30 agarose plugs. BAC libraries were created from both telosomes using the HindIII cloning site. The library derived from the long arm of the chromosome 7D called TaaCsp7DLhA comprised 50,304 clones ordered in 131 384-well plates. The average insert size of the whole library reached 116 kb and, excluding 0.5% empty clones and considering the 11% contamination by other chromosomes and the size of 346 Mbp [37], the library provided 14.9x coverage of 7DL. The library derived from the short arm of the chromosome 7D named TaaCsp7DShA consisted of 49,152 clones ordered in 128 plates. Having 113 kb mean insert size and 1.4% empty clones, the library constituted 12.1 arm equivalents if the 16% contamination with other chromosomes and 7DS molecular size of 381 Mbp [37] were considered. Both libraries comprised three size fractions. The representation and mean insert size of the particular fractions are shown in Table 1.

The overall distribution of insert sizes, which were estimated for 121 clones from the 7DL library and 184 clones
Table 1: Characteristics of individual size fractions of the 7DL- and 7DS-specific BAC libraries.

<table>
<thead>
<tr>
<th>Library</th>
<th>7DL</th>
<th>7DS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size fraction</td>
<td>B1</td>
<td>B2</td>
</tr>
<tr>
<td>No. of clones</td>
<td>21,504</td>
<td>23,424</td>
</tr>
<tr>
<td>Portion of the library</td>
<td>43%</td>
<td>46%</td>
</tr>
<tr>
<td>Mean insert size (kb)</td>
<td>101</td>
<td>123</td>
</tr>
</tbody>
</table>

Figure 1: Histogram of relative fluorescence intensity (flow karyotype) obtained after the analysis of DAPI-stained suspension of mitotic metaphase chromosomes prepared from double ditelosomic line 7D. The inset shows examples of flow-sorted 7DL and 7DS telosomes after FISH with probes for the telomeric repeat (red color) and the GAA repeat (turquoise color), which were used to identify the telosomes in the sorted fractions and estimate their purities.

Figure 2: Insert size distribution in 7DL- and 7DS-specific BAC libraries.

from the 7DS library, ranged from 10 to 200 kb and differed slightly between the libraries as shown in Figure 2.

Whereas most clones (39%) were found in the fraction of 125–149 kb in the 7DL library, the most numerous fraction (42% of clones) of the 7DS library is in the size range of 100–124 kb. This was due to greater representation of B2 compared with B1 fraction in the 7DL library. On the other hand, the 7DS library contained a greater percentage of large clones more than 150 kb (9% versus 7% for 7DL) although the average insert size of the M fraction was significantly greater in the 7DL library (147 kb for 7DL-M versus 130 kb for 7DS-M). This seeming discrepancy was due to a greater proportion of the M fraction in the 7DS library (11% for 7DL-M versus 18% for 7DS-M).

The portion of BAC clones smaller than 50 kb, which are usually noninformative in HICF analysis applied in construction of physical maps, was negligible in both libraries (below 2%). On the other hand, presence of short clones less than 75 kb in all fractions of both libraries suggested that the ligation ratio between the wheat DNA and the vector was set up correctly preventing ligation of more than a single DNA fragment into one vector due to excess wheat DNA. An improper ratio would have led to the creation of chimeric clones, which significantly compromise the quality of libraries. Absence of chimeric clones in the 7D libraries was further supported by the fact that no inserts exceeding sizes excised from the gel at the size selections were found for the B1, B2 and M fractions, respectively.

The mean insert size was estimated to be 116 kb for the 7DL and 113 kb for the 7DS library; both were significantly larger than for early wheat chromosome-specific BAC libraries, which reached 82 kb in the 1BS-specific library [36], 85 kb in a composite library constructed from chromosomes 1D, 4D and 6D [35] and 103 kb in the 3B-specific library [33]. The increase in insert size was achieved by including a second size-selection step in the BAC library construction procedure. Available genomic BAC libraries created from hexaploid wheat, which can be used as a standard for evaluating the quality of the 7D-specific libraries, varied significantly in their average insert size, which ranged from 75 kb for a library constructed from wheat cv. Norstar [17] to 157 kb for the “Chinese Spring” library constructed by Shen et al. [18]. In this regard, the 7D-specific libraries reached average values. However, in a comparison of the coverage of the libraries, even the most representative of the wheat genomic libraries constructed by Allouis et al. [14] providing 9.3x wheat genome coverage does not reach the coverage of our chromosome-specific libraries (14.9 and 12.1x for the 7DL and 7DS libraries, resp.).

3.3. Hybridization Screening of the 7DL Library. EST-derived marker STS-Aug-08-28 (Azhaguvel et al. in preparation) tightly linked with the Gb3 gene was used as a hybridization probe to screen the 7DL-specific BAC library (Figure 3, Screen I) providing 14 positive clones. All 14 were BAC-end sequenced and fingerprinted. Marker STS-BAC 22-R developed from the protruding end of BAC 22 clone was used as a probe for a second round of library screening and detected 22 positive BAC clones including two selected in the previous screen (Figure 3, Screen II). A third round of library screening was performed using marker STS-BAC
no. 25-T7 derived from the protruding end of BAC no. 25 (Figure 3, Screen III). This marker identified 23 positive BAC clones (including five BAC clones detected in the Screen II). Three BAC clones (BAC 22, BAC 25 and BAC 72) were fully sequenced and assembled. The sizes of BAC 22, BAC 25 and BAC 72 were 200.56 kb, 110.95 kb and 117.09 kb, respectively. Markers that were continuously derived during the contig assembly from available BAC sequences cosegregated with the resistance gene. Only a marker derived from the end of the BAC 72 showed recombination with the gene and flanked the Gb3 region from the opposite side. Thus three rounds of screening the 7DL library were sufficient to build a contig spanning the region between markers flanking the Gb3 gene. Positional cloning of Gb3 and a detailed annotation of this contiguous sequence will be described elsewhere (Azhaguvel et al., in preparation).

The number of positive clones selected by hybridization with single-copy probes and confirmed by contig assembly and sequencing outputs indicates library coverage greater than estimated based on insert size analysis. In fact, the number of positive clones confirmed by fingerprinting was 20 or more in all screens (21.7 at average) but in the screen I, BAC-end sequences were available for 14 of 20 clones, so only these clones were included in the final contig assembly. As all these screens were conducted in one region of the 7DL arm, we do not conclude that this number (21.7) shows evidence of overall underestimation of coverage based on insert size analysis but rather indicated overrepresentation of this region in the library. Differences in local coverage were observed also in the 7DS library for particular markers (Table 2 and Figure 4).

The above mentioned numbers of positive BAC clones were obtained after screening only three high-density filters that comprised the whole 7DL library. For comparison, BAC libraries of diploid wheat relatives T. urartu, Ae. speltoides and Ae. tauschii [22] representing 3.7, 5.4 and 4.1 equivalents of the respective genomes occupied 9, 13 and 10 high-density filters, respectively. Consequently, one filter comprising 18,432 clones of a diploid library represents only 0.41x genome coverage. An even greater workload would be required in hybridization screening of the polyploid wheat libraries. The BAC library of tetraploid T. Turgidum [54] with 5.8x coverage occupies 28 high-density filters; thus one filter carries just 0.2x genome equivalents. Hexaploid wheat BAC library with 3.1x coverage constructed from cv. Glenlea [15] was spotted on 24 filters with a greater density of 27,360 clones per filter. Despite of that, one filter represented only 0.13x genome equivalents. On the other hand, the 7DL library with coverage of 14.9x was placed on only three filters and thus the 18,432 clones spotted on one filter covered the 7DL arm 5.46x. Such coverage would be sufficient for positional cloning providing 99.5% probability of recovering any sequence present on the arm [55] and poses a strong argument for the chromosome-based genomics in wheat [56].

3.4. PCR Screening of the 7DS Library. Microsatellite markers Xcfd68, Xbarc214 and Xgwmn473 were found to be tightly linked to the Dn gene (Fazelnajafabadi and Lapitan, unpublished). They delimit an interval of 2.7 cM encompassing the gene. These markers were applied to screen the 7DS library using the set of BAC pools. As the first step, PCR was run on the superpools representing stacks of 8 plates, followed by screening on the plate pools and finally 3D pools. Although some of the superpools did not provide a PCR product, screening was performed on all plate pools to verify the reliability of the information obtained from superpools. The results from superpools and plate pools, respectively, were in full accordance for all markers tested. This implies the superpools can be used to preselect stacks of plates to be subjected to further screening and thus reduce the number of PCR reactions needed for library screening. However, considering the small number of 3D pools and the high coverage of the library due to which 37.5–75% of the superpools (depending on the locus) were found to be positive, this step is not essential. Similarly, PCR was run on all plate pools and 3D pools, respectively, to compare the results. In 6 of 47 cases (12.8%) the information from the plate pools helped in identifying the positive plates in the 3D pool set. These were cases when the PCR product obtained from 3D pools, which were prepared by a simplified procedure, was too weak for an unambiguous identification of the positive plate. Thus the set of plate pools prepared by a more advanced DNA-isolation procedure proved to be useful in resolving the position of the positive clones. However, as demonstrated below, its role can be substituted by availability of data from the BAC contig assembly.

Among the positive BAC clones selected by Xbarc214 and Xgwmn473, several BAC clones were found that provided products of an unexpected size. By comparing these products with products obtained by PCR on DNA of flow-sorted 7DS arms we verified that these products did not relate to the 7DS arm. These BAC clones remained as singletons in the contig assembly, which implies that they come from contamination of the sorted fraction by other chromosomes. Such clones were not considered positives. However, screening the 7DS library with the Xgwmn473 marker provided an even more complex spectrum of PCR products. Besides clones bearing a double-band of 220 and 226 bp, respectively, which has been characteristic of the DnCI2401-linked marker, and a few clones of unexpected size, we also found numerous clones that generated a PCR product of 200 bp. This product was also visible as a weaker band when amplifying DNA of flow-sorted 7DS. This product may correspond to another Xgwmn473 locus present on 7DS. This presumption was supported by finding a BAC contig ctg135 comprising clones with the 200 bp PCR product (Figure 4(d)). Sequencing the locus both from positive clones of ctg285 and ctg135 confirmed that the two loci differed. Comparing sequences of the 226 and the 200 bp band revealed a 26 bp deletion and 92% homology in the remaining sequence. The results suggest we identified a paralogous Xgwmn473 locus situated on 7DS. This locus is probably not polymorphic in available mapping populations as there is no evidence about it in databases as GrainGenes or Gramene.

The results of the screening the 7DS library with Xcfd68, Xbarc214 and Xgwmn473 are shown in Figure 4 and Table 2. Most of the positive clones were identified unambiguously based on PCR screening of plate and 3D pools. The Xcfd68
marker selected nine positive clones in the screening and two additional BAC clones were revealed after integrating data from the BAC contig assembly as row or column information for these clones was missing (Figure 4(a)). Similarly, for Xbarc214 15 positive BAC clones were identified by library screening whereas BAC addresses of another two clones could be completed only after considering data from the BAC contig assembly (Figure 4(b)). In case of the Xgwm473 marker, nine and 16 positive clones, respectively, were identified in ctg285 and 135, respectively (Figures 4(c) and 4(d)). All BAC addresses were complete; however, three of ten BAC clones providing the 220 and 226 bp products were absent from the contig assembly, thus their position in ctg285 could not be confirmed. They might have been missing in the library replica used for fingerprinting or excluded by the GenoProfiler software due to a low quality of fingerprint or cross-contamination. Our results imply that the proposed system of BAC pools is sufficiently powerful to reveal positions of positive BAC clones. Having contig assembly data in hand makes the deconvolution of BAC addresses easier and enables further reducing the number of clones to be screened. The local coverage differed among the tested loci and ranged between 10 and 17x, averaging 13.5x. This is in agreement with the coverage estimate based on insert size (12.1x).

### Table 2: Results of 7DL library screening with microsatellite markers Xcfd68, Xbarc214, and Xgwm473.

<table>
<thead>
<tr>
<th>Marker</th>
<th>Xcfd68</th>
<th>Xbarc214</th>
<th>Xgwm473</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product size</td>
<td>207 bp</td>
<td>217 bp</td>
<td>220 and 226 bp</td>
</tr>
<tr>
<td>Contig</td>
<td>ctg244</td>
<td>ctg148</td>
<td>ctg285</td>
</tr>
<tr>
<td>No. of hit BAC clones</td>
<td>9</td>
<td>15</td>
<td>7 + 3³</td>
</tr>
<tr>
<td>Total no. of positive clones</td>
<td>11</td>
<td>17</td>
<td>7 + 3³</td>
</tr>
</tbody>
</table>

1. Positive clones identified by PCR screening only.
2. Hit BAC clones plus clones identified with the use of contig assembly data.
3. Three of ten BAC clones providing the 220 and 226 bp products were missing from the assembly, thus their location in ctg285 could not be confirmed.

Various pooling strategies have been proposed for PCR-based screening of wheat libraries for cloning of genes. Nil-malagoda et al. [15] employed the analogous pooling strategy as applied in this study involving superpools (combining clones from ten consecutive plates) and 3D (plate, row, and column) pools created for stacks of ten plates. The “Glenlea” wheat library comprising 3.1 genome equivalents was represented by 171 superpools and 8,550 3D pools.
In their case, assaying superpools played a substantial role in reducing the overall number of PCR reactions. In this two-step screening, a single positive BAC clone could be reached in 221 (171 superpools + 50 3D pools for the selected superpool) PCR reactions. However, for obtaining several clones, which is preferable in positional cloning, significantly more pools would have to be assayed. A different pooling strategy was proposed by Febrer et al. [57] who screened a part of the “Chinese Spring” wheat genomic library constructed by Allouis et al. [14] for a wheat dwarfing gene Rht. The screen of 715,776 clones from the library (∼5 genome equivalents) was based on pooling of DNA from BAC clones into 675 superpools arrayed in a three-dimensional configuration. This enabled identification of candidate 384-well plates. A second round of PCR was used to detect a specific BAC clone within the candidate plate that corresponded to the gene of interest. So assaying at least three 384-well plates was needed to identify a single copy of the target gene. For identifying all three homoeologues of the Rht gene, 17 candidate plates were screened providing two-step screening, a single positive BAC clone could be reached in 221 (171 superpools + 50 3D pools for the selected superpool) PCR reactions. However, for obtaining several clones, which is preferable in positional cloning, significantly more pools would have to be assayed. A different pooling strategy was proposed by Febrer et al. [57] who screened a part of the “Chinese Spring” wheat genomic library constructed by Allouis et al. [14] for a wheat dwarfing gene Rht. The screen of 715,776 clones from the library (∼5 genome equivalents) was based on pooling of DNA from BAC clones into 675 superpools arrayed in a three-dimensional configuration. This enabled identification of candidate 384-well plates. A second round of PCR was used to detect a specific BAC clone within the candidate plate that corresponded to the gene of interest. So assaying at least three 384-well plates was needed to identify a single copy of the target gene. For identifying all three homoeologues of the Rht gene, 17 candidate plates were screened providing four to 12 copies of a particular microsatellite locus.

Once a BAC contig assembly has been completed, screening can be further simplified by preparing BAC pools from MTP clones only. In case of the wheat 3B-chromosome-specific BAC library [33] consisting of 67,968 clones, the MTP comprised 7,440 BAC clones ordered in 20 plates. Thus the library was represented by 60 three-dimensional pools [58], which were sufficient for screening the whole library.

In their case, assaying superpools played a substantial role in reducing the overall number of PCR reactions. In this two-step screening, a single positive BAC clone could be reached in 221 (171 superpools + 50 3D pools for the selected superpool) PCR reactions. However, for obtaining several clones, which is preferable in positional cloning, significantly more pools would have to be assayed. A different pooling strategy was proposed by Febrer et al. [57] who screened a part of the “Chinese Spring” wheat genomic library constructed by Allouis et al. [14] for a wheat dwarfing gene Rht. The screen of 715,776 clones from the library (∼5 genome equivalents) was based on pooling of DNA from BAC clones into 675 superpools arrayed in a three-dimensional configuration. This enabled identification of candidate 384-well plates. A second round of PCR was used to detect a specific BAC clone within the candidate plate that corresponded to the gene of interest. So assaying at least three 384-well plates was needed to identify a single copy of the target gene. For identifying all three homoeologues of the Rht gene, 17 candidate plates were screened providing four to 12 copies of a particular microsatellite locus.

Once a BAC contig assembly has been completed, screening can be further simplified by preparing BAC pools from MTP clones only. In case of the wheat 3B-chromosome-specific BAC library [33] consisting of 67,968 clones, the MTP comprised 7,440 BAC clones ordered in 20 plates. Thus the library was represented by 60 three-dimensional pools [58], which were sufficient for screening the whole library.
In the case of a library from the short arm of chromosome 3D [59] containing 36,864 clones, the MTP comprised 3,823 BAC clones occurring in 50 three-dimensional pools [60]. Screening this reduced number of pools can identify a BAC contig containing the marker; however, as the MTP may not comprise the complete chromosomal sequence and its coverage is ~1x, the risk of losing some information is relatively high and thus screening the 3D MTP pools of the 3B library was combined with screening plate pools prepared from the whole library [61].

Microsatellites represent a frequently used marker system as they can be easily extracted from genomic (e.g., BAC-end or shotgun) sequences, are abundant and highly polymorphic. They are a typical marker of choice in efforts to place genes of interest on a genetic map. However, microsatellites proved to be of limited use in large-scale screening of a soybean genomic BAC library to anchor the physical contig map because of their multilocus character [62]. Screening by microsatellites in a hexaploid wheat library would be even more complicated due to the presence of homoeologous genomes. This view is supported by findings of Nilmalgoda et al. [15] who screened a hexaploid wheat library of 3.1x coverage both with gene-derived and microsatellite markers. Whereas gene sequences identified 2.7 positive BAC clones on average, the average number of clones hit by microsatellites was twofold (5.5 clones). Our small-scale screening detected two loci relating to one of the microsatellite markers (Xgwm473) even on one chromosome arm. Similarly, multiple loci were found for another microsatellite marker (Xgwm444) when screening the 7DS library and analyzing the selected clones using the available contig assembly (Šimková, unpublished). This indicates that results of screening BAC libraries with microsatellite markers must be interpreted with caution even in the case of chromosome-arm specific libraries.

The screening of the library with markers for the DnCI2401 gene is a first step towards positional cloning of this gene. The work is in progress to derive new markers from colinear regions of related genomes (rice, Brachypodium) that were identified based on sequencing BAC clones from contigs containing the markers.

4. Conclusions

Two BAC libraries specific for both the long and the short arms of the bread wheat chromosome 7D were constructed with parameters challenging available wheat genomic libraries. The libraries represent the first subgenomic BAC resources available for wheat homoeologous chromosome group 7 and will facilitate advancement in many areas of wheat genomics, including physical map construction and map-based cloning of genes. Due to the small number of clones, high genome coverage and absence of clones from homoeologous genomes, these libraries make screening with markers for genes of interest highly efficient and cost-effective. For example, one high-density filter prepared from the 7DL library provided coverage of 5.4x, which ensures 99.5% probability of finding any sequence located on the 7DL chromosome arm. A simple pooling strategy was elaborated based on which 6x equivalents of the 7DS arm could be screened by 384 PCR reactions. After completing physical map assembly of the 7DS arm, the number of pools to be screened can be further reduced. All these features and the results of screening these libraries with markers for aphid resistance genes indicate that chromosome-arm-specific BAC libraries are a powerful resource for cost-effective positional gene cloning in bread wheat.

Acknowledgments

The authors are grateful to Professor B. S. Gill (Kansas State University, Manhattan, USA) for seeds of the double ditelosomic 7D line of wheat Triticum aestivum L. cv. Chinese Spring. We thank our colleagues, Jana Dostálová, Bc., Radka Tušklová, Helena Tvardíková and Ing. Marie Seifertová for excellent technical assistance. The authors thank Dr. Mehdi Fazelnajafabadi for the genetic map data of DnCI2401. This work was supported by the Czech Science Foundation (Grant Award 521/07/1573), Czech Republic Ministry of Education, Youth and Sports (Grant Award LO6004), and National Basic Research Program of China (973 Program, 2009CB118300).

References


