Theor Appl Genet (2004) 109: 1337–1345
DOI 10.1007/s00122-004-1768-8

ORIGINAL PAPER

Jaroslav Janda · Jan Bartoš · Jan Šafář · Marie Kubaláková · Miroslav Valárik · Jarmila Číhalíková · Hana Šimková · Michel Caboche · Pierre Sourdille · Michel Bernard · Boulos Chalhoub · Jaroslav Doležel

Construction of a subgenomic BAC library specific for chromosomes 1D, 4D and 6D of hexaploid wheat

Received: 10 May 2004 / Accepted: 4 July 2004 / Published online: 10 September 2004

Abstract The analysis of the hexaploid wheat genome (*Triticum aestivum* L., 2*n*=6*x*=42) is hampered by its large size (16,974 Mb/1C) and presence of three homoeologous genomes (A, B and D). One of the possible strategies is a targeted approach based on subgenomic libraries of large DNA inserts. In this work, we purified by flow cytometry a total of 10^7 of three wheat D-genome chromosomes: 1D, 4D and 6D. Chromosomal DNA was partially digested with *Hin*dIII and used to prepare a specific bacterial artificial chromosome (BAC) library. The library (designated as TA-subD) consists of 87,168 clones, with an average insert size of 85 kb. Among these clones, 53% had inserts larger than 100 kb, only 29% of inserts being shorter than 75 kb. The coverage was estimated to be 3.4-fold, giving a 96.5% probability of identifying a clone corresponding to any sequence on the three chromosomes. Specificity for chromosomes 1D, 4D and 6D was confirmed after screening the library pools with single-locus microsatellite markers. The screening indicated that the library was not biased and gave an estimated coverage of sixfold. This is the second report on BAC library construction from flow-sorted plant chromosomes, which confirms that dissecting of the complex wheat genome and preparation of subgenomic BAC libraries is possible. Their availability should facilitate the analysis of wheat genome structure and evolution, development of cytogenetic maps, construction of local physical maps and map-based cloning of agronomically important genes.

Communicated by P. Langridge

J. Janda · J. Bartoš · J. Šafář · M. Kubaláková · M. Valárik · J. Číhalíková · H. Šimková · J. Doležel (✉)
Laboratory of Molecular Cytogenetics and Cytometry, Institute of Experimental Botany,
Sokolovská 6,
77200 Olomouc, Czech Republic
e-mail: dolezel@ueb.cas.cz
Fax: +420-585 205 852

M. Kubaláková · J. Číhalíková · J. Doležel
Department of Cell Biology and Genetics, Palacký University,
Šlechtitelů 11,
78371 Olomouc, Czech Republic

M. Caboche · B. Chalhoub (✉)
Laboratory of Genome organization, Unité de Recherches en Génomique Végétale (INRA-URGV),
2 rue Gaston Crémieux CP 5708,
91057 Évry Cedex, France

P. Sourdille · M. Bernard
Génétique Moléculaire des Céréales, UMR INRA-UBP,
Domaine de Crouelle, 234 Avenue du Brézet,
63039 Clermont-Ferrand Cedex 2, France

Introduction

Wheat (*Triticum aestivum* L.) is grown on more areas of land than any other crop plant and is a staple food for 40% of the world's population, providing 20% of the calories consumed (Gill et al. 2004). It is an allohexaploid species (2*n*=6*x*=42, genome formula AABBDD) that resulted from two independent hybridization events. The first lead to the allotetraploid species *T. turgidum*, in which the A genome of the diploid species *T. urartu* and the B genome of a yet unidentified diploid species were combined. *Aegilops speltoides* is the living species believed to be most closely related to the B genome of hexaploid wheat (Daud and Gustafson 1996). The second hybridization occurred only about 8,000 years ago and involved *T. turgidum* and *Ae. tauschii*, a diploid species carrying the D genome, and resulted in hexaploid wheat (Feldman et al. 1995). Due to the presence of three closely related (homoeologous) subgenomes (A, B and D), sequence redundancy is a characteristic feature of the wheat genome. This, together with a genome size of 16,974 Mb/1C (Bennett and Smith 1991)—which is fivefold the human genome, and more than 80% of repetitive DNA (Smith and Flavell 1975)—represents a great challenge for gene discovery and genome sequencing efforts.

One way to simplify the molecular genetic analyses is to create resources specific for the three subgenomes of wheat. Bacterial artificial chromosome (BAC) libraries, which represent one of the most useful genome resources

(Zhang and Wu 2001), were constructed for the D genome of *T. tauschii* (Moullet et al. 1999) and for the A^m genome of domesticated diploid wheat *T. monococcum* (Lijavetzky et al. 1999), which is related to *T. urartu* [known to be the putative donor of the A genome of wheat (Dvorak et al. 1993)]. More recently, Cenci et al. (2003) constructed a BAC library for tetraploid durum wheat (*T. turgidum* var. *durum*, genome formula AABB). These BAC resources turned out to be very useful and facilitated gene cloning (Faris et al. 2003; Ling et al. 2003; Yan et al. 2003) and discovery of organization of the wheat genome into gene-rich islands, where gene densities are similar to that of *Arabidopsis*, and gene-poor regions loaded with retro-elements (Wicker et al. 2001; Brooks et al. 2002). These observations showed that despite the initial scepticism, map-based cloning is viable in wheat.

Unfortunately, the diploid BAC libraries have limitations stemming mainly from a dynamic and rapid genome evolution in wheat species and changes induced by polyploidization (Ozkan et al. 2001). For example, a comparison of a *T. tauschii* BAC clone sequence with orthologous regions of the wheat A and B genomes revealed rapid sequence divergence (Anderson et al. 2003). Sequencing and annotation of two physical contigs covering orthologous glutenin loci on chromosome 1AS of *T. monococcum* and *T. turgidum* var. *durum* showed that sequence conservation was limited to the regions containing glutenin genes. The rest of the two contigs, representing 90% of the total sequence, was not conserved (Wicker et al. 2003). These data imply that BAC libraries created from wheat progenitors and related species cannot fully replace BAC resources made from hexaploid wheat. The recent progress in BAC cloning methodology facilitated preparation of several BAC libraries for hexaploid wheat (Allouis et al. 2003; Liu et al. 2000; Nilmalgoda et al. 2003; Ma et al. 2000). Although they represent invaluable tools, due to the large number of clones needed to achieve a desired genome coverage (0.5–1×10^6 clones), their maintenance, handling and screening represents a formidable challenge.

A strategy is needed to reduce the complexity of the huge genome of hexaploid wheat, without sacrificing the access to large genomic DNA clones. The creation of chromosome and chromosome-arms-specific BAC libraries is an attractive possibility. Representing only a few per cent of the total genome, they would open completely new horizons for wheat genome analysis and gene cloning. With this in mind, Vrána et al. (2000) developed protocols that could be used to purify wheat chromosomes in large numbers and high purity, using laser flow cytometry. Subsequently, Kubaláková et al. (2002) demonstrated that any chromosome arm of wheat could be sorted from telosomic lines, and Šimková et al. (2003) developed a protocol for preparation of high-molecular-weight (HMW) DNA suitable for cloning from sorted chromosomes. These advances facilitated the development of protocols for the cloning of individual flow-sorted chromosomes into BAC libraries and the construction of the first chromosome-specific BAC library in plants: the chromosome 3B library (Šafář et al. 2004). Consisting of only 67,968 clones, with an average insert size of 103 kb, the TA-3B library represents an easily manageable resource with over a sixfold coverage of the chromosome (Šafář et al. 2004).

In this work, we continued cloning of hexaploid wheat chromosomes and report the construction of the second BAC library ever produced from flow-sorted wheat chromosomes. The library was produced for chromosomes 1D, 4D and 6D that were sorted from cv. Chinese Spring. Creation of the subgenomic library confirms a possibility to prepare BAC libraries from DNA of sorted wheat chromosomes. The availability of the subgenomic BAC library will aid in comparative genome analysis as well as gene cloning from the three wheat D-genome chromosomes.

Materials and methods

Plant material

Seeds of hexaploid wheat *T. aestivum* L. cv. Chinese Spring ($2n=6x=42$) were germinated in the dark at 25 ±0.5°C on moistened filter paper in glass petri dishes for 3 days to achieve optimal root length (2–3 cm). In total, 4,000 seeds were germinated and used in batches of 20 for preparation of chromosome suspensions.

Chromosome analysis and sorting

Suspensions of intact mitotic chromosomes were prepared from synchronized root tip meristems as described by Vrána et al. (2000). Mitotic chromosomes were released into the modified isolation buffer (IB) of Šimková et al. (2003), stained by 2 μg/ml 4′,6-diamidino-2-phenylindole (DAPI) and analysed using the FACSVantage flow cytometer (Becton Dickinson, San José, Calif., USA). In order to sort chromosomes 1D, 4D and 6D, a sort window was set on a dot plot of fluorescence pulse area versus fluorescence pulse width. Chromosomes were sorted in aliquots of 1.0×10^5 into 160 μl 1.5× IB buffer. In order to check the purity in sorted fractions, 2,000 chromosomes were sorted at regular intervals onto a microscope slide into a 15-μl drop of PRINS buffer supplemented with 5% sucrose and air-dried. Sorted chromosomes were identified after fluorescence in situ hybridization (FISH), with a probe for the *Afa* repeat (Kubaláková et al. 2002).

BAC library construction

Preparation of HMW DNA and BAC library construction were performed as described previously (Šafář et al. 2004). Briefly, each batch of 10^5 flow-sorted chromosomes was pelleted, and the chromosomes were embedded in 12 μl of low-melting-point agarose. In total, 100 agarose miniplugs, representing 1×10^7 sorted chromo-

somes, were prepared and used. Chromosomal DNA was partially digested with *Hin*dIII and size-selected by pulsed field gel electrophoresis (PFGE). Four regions of the gel corresponding to approximately 50–100 kb (H0 fraction), 100–150 kb (H1 fraction), 150–200 kb (H2 fraction) and 200–250 kb (H3 fraction) were excised. HMW DNA was electroeluted and ligated into a dephosphorylated vector pIndigoBAC (Caltech, Pasadena, Calif., USA), prepared according to Chalhoub et al. (2004). *Escherichia coli* ElectroMAX DH10B competent cells (GIBCO–BRL, Gaithersburg, Md., USA) were transformed by electroporation. The library was ordered in 384-well plates in FM medium (Woo et al. 1994). The plates were incubated at 37°C overnight, replicated and stored at −80°C.

BAC library characterization

Insert size of BAC clones selected randomly from all four ligations was determined as described previously (Šafář et al. 2004). The library was pooled and screened using microsatellite (SSR) markers specific for 1D, 4D and 6D chromosomes as well as markers specific for the remaining 18 wheat chromosomes (Tables 1, 2). Screening was done also with mit3 and mit4 markers specific for mitochondrial DNA (B. Chalhoub et al., unpublished data) and markers specific for chloroplast genes *rbcL* and *psbA* (Ogihara et al. 2000). BAC clones from each of the 384-well plates were pooled, pelleted and resuspended in 4 ml TE (10:1 mM) buffer. Bacterial suspensions were lysed at 95°C for 30 min and pelleted at 3,000 *g* for 60 min. Supernatant was diluted 20-fold by deionized water and used for PCR. The reaction mix (10 µl) consisted of 2 µl template DNA, 1× PCR buffer, 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 1.25 µM primers and 0.5 U AmpliTaq DNA polymerase (Roche, Mannheim, Germany). PCR reaction was performed as follows: nine cycles consisting of 5 min at 94°C, 12 s at 65°C and 18 s at 72°C; 35 cycles of 12 s at 94°C, 12 s at 55°C and 18 s min at 72°C; and final extension at 72°C for 7 min. Presence of PCR products was checked by electrophoresis on 4% MetaPhor agarose gel (FMC Bioproducts, Philadelphia, Pa., USA).

FISH with 'low copy' BAC clones

A total of 1,536 BAC clones of the H0 sublibrary (four 384-well plates) were spotted onto two 8×12-cm Hybond-N^+ filters (AP Biotech, Piscataway, N.J., USA). The filters were hybridized with digoxigenin-labelled C_0t-1 fraction of genomic DNA (Zwick et al. 1997) to reveal clones with low proportions of repetitive sequences. Twenty-one 'low-copy' BAC clones (Table 3) showing no or weak hybridization signals were chosen for FISH. Preparations of digoxigenin-labelled probes from BAC inserts and FISH on flow-sorted chromosomes were performed as described by Šafář et al. (2004). When needed, dispersed repetitive DNA was blocked by C_0t-1 fraction at various ratios of probe/blocking DNA. In order to identify individual chromosomes, a biotin-labelled probe for the *Afa* repeat was included in the hybridization mix. The sites of probe hybridization were detected using anti-digoxigenin-FITC and streptavidin-Cy3, and the chromosomes were counterstained by DAPI. The preparations were evaluated using an Olympus BX60 microscope (Olympus, Tokyo, Japan) equipped with optical filter sets appropriate for DAPI, fluorescein and Cy3 fluorescence. The images of DAPI, fluorescein and Cy3 fluorescence were acquired separately with a b/w CCD camera, which was interfaced to a PC running the ISIS software (Metasystems, Altlussheim, Germany).

Table 1 Library screening with SSR markers specific for chromosomes 1D, 4D and 6D

1D specific		4D specific		6D specific	
Marker[a]	Positive pools	Marker[a]	Positive pools	Marker[a]	Positive pools
cfa2147	5	*cfa2173*	14	*cfd13*	4
cfa2158	16	*cfd106*	8	*cfd132*	3
cfd15	10	*cfd193*	4	*cfd135*	1
cfd21	3	*cfd23*	15	*cfd213*	5
cfd27	7	*cfd81*	5	*cfd287*	9
cfd28	9	*cfd84*	3	*cfd33*	16
cfd282	5	*cfd89*	7	*cfd37*	3
cfd58	8	*gwm165*	9	*cfd38*	3
cfd61	6	*gwm194*	6	*cfd42*	8
cfd63	3	*gwm609*	2	*cfd45*	5
cfd65	5	*gwm624*	0	*cfd47*	7
cfd72	9	–	–	*cfd49*	10
cfd83	6	–	–	*cfd5*	6
cfd92	2	–	–	*cfd60*	2
gwm232	2	–	–	*cfd75*	4
gwm337	7	–	–	*cfd76*	4
gwm458	5	–	–	*cfd80*	5
gwm642	5	–	–	*cfd95*	7
cfd32	0	–	–	*gwm325*	2
–	–	–	–	*gwm469*	10
Mean	5.94		6.34		5.70

[a]*gwm*- Röder et al. (1998), *cfd*- Guyomarc'h et al. (2002)

Results

Flow cytometric analysis of relative fluorescence intensity of DAPI-stained chromosomes isolated from cv. Chinese Spring resulted in a distribution (flow karyotype) on which four peaks (I, II, III and 3B) could be discriminated (Fig. 1). While the peak 3B represented solely the chromosome 3B, peaks I, II, and III represented groups of 3, 7 and 10 chromosomes, respectively (Kubaláková et al. 2002). Chromosomes 1D, 4D and 6D from peak I (Fig. 1, insert) were sorted to prepare the subgenomic BAC library. In total, 10^7 chromosomes were sorted and used to prepare 100 miniplugs (10^5 chromosomes each), representing approximately 13.4 µg of DNA. The sorting

Table 2 Library screening with SSR markers specific for potentially contaminating chromosomes

Chromosome	A genome		B genome		D genome	
	Marker[a]	Positive pools	Marker[a]	Positive pools	Marker[a]	Positive pools
1	*Xgwm99*	0	*Xgmw11*	0	–	–
	Xgwm135	1	*Xgwm413*	0	–	–
2	*Xgwm312*	0	*Xgwm120*	0	*Xgwm261*	1
	Xgmw372	0	*Xgwm257*	0	*Xgwm539*	0
3	*Xgwm5*	0	*Xgpw1180*	0	*Xgwm161*	0
	Xgwm480	0	*Xgpw8100*	0	*Xgwm664*	0
4	*Xgwm610*	0	*Xgwm149*	0	–	–
	Xgwm601	Not functional	*Xgwm251*	0	–	–
5	*Xgwm415*	0	*Xgwm234*	2	*Xgwm190*	0
	Xgwm186	0	*Xgwm408*	0	*Xgwm272*	0
6	*Xgwm169*	0	*Xgwm132*	0	–	–
	Xgwm427	0	*Xgwm219*	0	–	–
7	*Xgwm233*	0	*Xgwm46*	1	*Xgwm44*	0
	Xgwm260	0	*Xgwm400*	0	*Xgwm437*	0

[a]*gwm*- Röder et al. (1998), *gpw*- Sourdille et al. (2001)

took 18 working days, and 4,000 seeds were needed to prepare 200 samples of chromosome suspensions. Purity, in sorted fractions as determined microscopically after FISH with the *Afa* repeat, was 91%. It was found that sorted chromosomes were contaminated at an average frequency of 9% by various chromosomes and chromosome arms, without an apparent prevalence of certain chromosome types.

Four BAC sublibraries were created from four independent ligation reactions with different size classes of partially digested chromosomal DNA, namely the sublibraries H0, H1, H2 and H3. Ninety-six BAC clones were selected randomly from the four sublibraries and used to estimate the average insert size using PFGE (Fig. 2a). The most abundant sublibrary, H0, contained 42,240 clones and its average insert size was 70 kb. The H1 sublibrary

Table 3 Cytogenetic mapping of putatively 'low copy' clones selected from the subgenomic bacterial artificial chromosome (*BAC*) library specific for chromosomes 1D, 4D and 6D. Clones determined after hybridization with C_0t-1 fraction of genomic DNA

BAC clone	Insert (kb)	Genomic distribution	
		Localization on chromosomes 1D, 4D and 6D[a]	Hybridization to other wheat chromosomes
214/J1	90	6D - distal (S)	–[b]
214/02	55	1D - distal (S), centromere	–
214/C8	60	4D - interstitial (L) 6D - distal (S)	–
109/F2	70	1D - distal (S), centromere 4D - distal (L), centromere	7D
110/K6	60	4D - distal (L) 6D - centromere	4A
109/C22	45	1D - distal (S, L), centromere 4D - distal (S, L), centromere	3A, 4A, 7A, 4B, 2D, 7D
110/C2	65	Centromere	Centromeres of all chromosomes
109/019	45	Centromere	Centromeres of all chromosomes
109/I1	50	Dispersed	All
214/K1	55	Dispersed	All
215/B8	60	Dispersed	All
215/C23	70	Dispersed	All
110/K3	55	Dispersed	All
215/C21	90	Dispersed	All
214/M16	65	Dispersed	All
215/C20	100	Dispersed	All
110/F24	90	Dispersed	All
109/K24	90	Dispersed	All
109/K17	105	Dispersed	All
214/N21	90	Dispersed	All
215/M22	95	Dispersed	All

[a]Localization to discrete loci required suppression of dispersed repeats, using C_0t-1 fraction of genomic DNA. *S* short arm, *L* long arm
[b]No hybridization

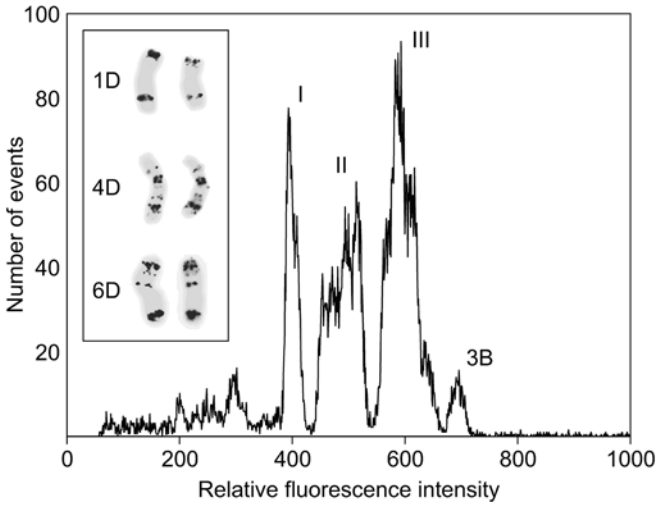

Fig. 1 Histogram of relative fluorescence intensity (flow karyotype) obtained after flow cytometric analysis of 4′,6-diamidino-2-phenylindole (DAPI)-stained suspension of chromosomes prepared from hexaploid wheat cv. Chinese Spring. The flow karyotype consists of three composite peaks (I, II and III) representing groups of chromosomes and a clearly discriminated peak representing chromosome 3B. Examples of chromosomes 1D, 4D and 6D, which were sorted from peak I, are shown in the *insert* after fluorescence in situ hybridization (FISH) with a probe for *Afa* repeat (*black signals*) that facilitated their identification

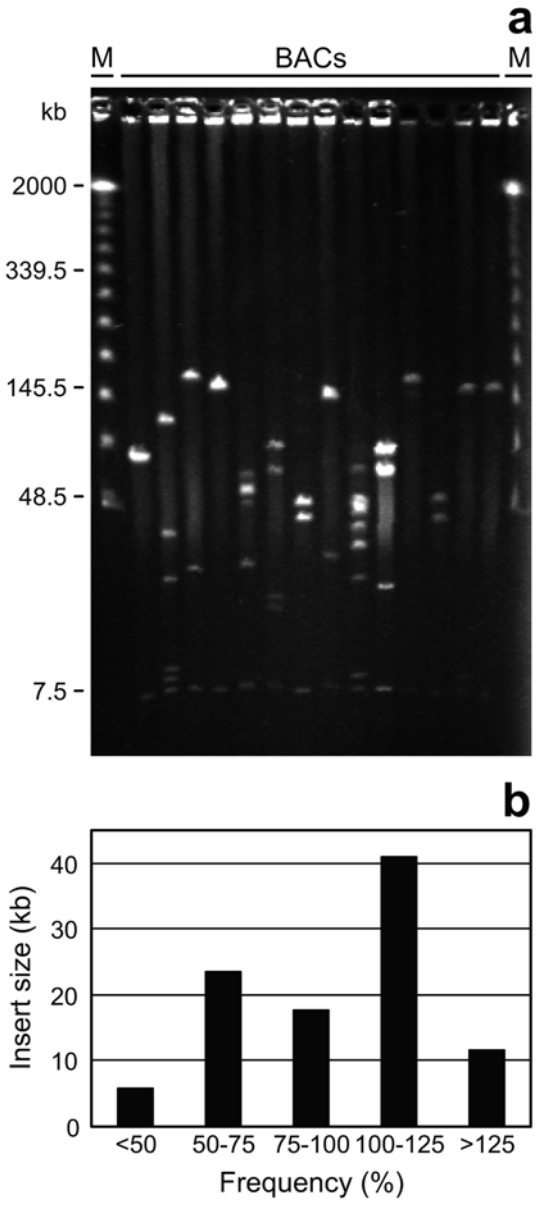

Fig. 2 Insert-size analysis of randomly chosen clones from the TA-subD bacterial artificial chromosome (BAC) library. **a** BAC DNA was digested with *Not*I enzyme to release the insert, separated by pulsed field gel electrophoresis (PFGE) and stained with ethidium bromide. *Lanes M* contain λPFGE marker. The 7.5-kb band represents the BAC vector. **b** Insert-size distribution of 96 randomly chosen BAC clones from the BAC library. Fifty-three per cent of BAC clones contain inserts larger than 100 kb; only 29% of BAC clones have inserts shorter than 75 kb

had 33,024 clones and an average insert size of 100 kb. The least abundant sublibrary with only 3,456 clones was H2, with 90 kb of average insert size, while the H3 sublibrary had 8,448 clones, and the average insert size was 92 kb. The complete BAC library (designated as TA-subD) consisted of 87,168 clones ordered in 227 384-well plates, with an average insert size of 85 kb. Of the clones, 53% showed inserts larger than 100 kb; only 29% of BAC clones have inserts shorter than 75 kb (Fig. 2b).

Considering the wheat genome size of 16,974 Mb/1C (Bennett and Smith 1991) and relative chromosome lengths given by Gill et al. (1991), molecular sizes of the three chromosomes were estimated as 606 Mb for 1D, 649 Mb for 4D and 714 Mb for 6D. Together they represent 1,969 Mb, or 11.6%, of the hexaploid wheat genome. If a contamination with other chromosomes of 9% is considered, the subgenomic BAC library represents 3.4 size equivalents of the three chromosomes. Using the formula of Clarke and Carbon (1976), a probability of recovering any DNA sequence present on 1D, 4D and 6D chromosomes in the library was estimated to be 96.5%. Conversely, a probability of finding DNA sequences from chromosomes that contaminated the sorted fraction was calculated to be 4.8%.

With the aim to confirm the chromosome specificity and verify the genome coverage of 3.4-fold, the subgenomic BAC library was screened by PCR with SSR markers. The library was pooled into 227 pools, each pool representing one 384-well plate. The screening was done using 19 SSR markers specific for chromosome 1D, 11 SSR markers specific for chromosome 4D, and 20 SSR markers specific for chromosome 6D. The number of pools that gave PCR products of expected length with each SSR marker is listed in Table 1. Based on these results, the coverage for chromosomes 1D, 4D, and 6D was estimated to be 5.95-fold, 6.34-fold and 5.7-fold, respectively, indicating equal representation of the three chromosomes in the library. The coverage of the complete BAC library estimated using the SSR markers was thus sixfold. The extent of contamination with DNA of other 18 chromosomes was estimated using 36 SSR markers. Only a scarce contamination by

DNA of chromosomes 1A, 2D, 5B and 7B was detected (Table 2).

While no positive pools were found after screening with mitochondria-specific primers, few positive pools were found after screening with markers specific for chloroplast DNA. However, simultaneous occurrence of products with both chloroplast primer pairs in one pool was observed only in 10% of the positive pools. Considering the fact that the *rbcL* and *psbA* genes are separated only by 55 kb in the wheat chloroplast genome (Ogihara et al. 2000), our observation indicates that sequences integrated in nuclear genome were detected rather than DNA coming from organelles.

One of the uses of the library will be the construction of detailed cytogenetic maps. To test this approach, 1,536 clones from the H0 sublibrary were hybridized with C_0t-1 fraction of genomic DNA to reveal clones with the lowest amount of repetitive DNA and potentially containing low- or single-copy sequences. Twenty-one of these were chosen for cytogenetic mapping using FISH. Although the clones were selected for low proportion of repeats, FISH with 13 of them resulted in complete labelling of all wheat chromosomes (Table 3). FISH localization of the eight remaining BAC clones on chromosomes 1D, 4D and 6D is shown in Fig. 3 and summarized in Table 3. Two of the BAC clones (110/C2 and 109/O19) localized in centromeres of all wheat chromosomes. It is interesting to note that these clones localized specifically to centromeres of all barley chromosomes, too (not shown). Two other BAC clones (214/J1 and 214/O2) gave specific signals on only one of the D-genome chromosomes. Four clones (214/C8, 109/F2, 110/K6 and 109/C22) localized to discrete loci on two of the three D chromosomes from which the library was created. BAC clone 110/K6 hybridized to chromosome 4A and BAC clone 109/F2 to chromosome 7D. While the BAC clone 214/C8 gave no signal on the other 19 wheat chromosomes, BAC clone 109/C22 hybridized also to six other wheat chromosomes (Table 3).

Discussion

This paper reports on the construction of a subgenomic BAC library (TA-subD) specific for chromosomes 1D, 4D and 6D of hexaploid wheat. This is the second BAC library ever produced from particular plant chromosomes, after the chromosome 3B-specific BAC library created by Šafář et al. (2004). Its construction proves that chromosomes purified using flow cytometry may be used for preparation of chromosome-specific BAC libraries in wheat and other plant species as the methods for chromosome sorting are available for other cereal and legume species (Doležel et al. 2004).

The library screening with SSR markers proved a representative coverage of the three chromosomes and their equal representation in the library. The genomic coverage of sixfold estimated after screening with SSR markers was higher compared to that estimated with average insert size (3.4-fold). A similar phenomenon was observed for the TA-3B BAC library specific for wheat chromosome 3B (Šafář et al. 2004). As only PCR products with expected length were considered, the sixfold coverage could not be due to the presence of duplicated loci. Therefore, as with the TA-3B BAC library (Šafář et al. 2004), the discrepancy could either be due to overestimation of molecular size of the three chromosomes or to underestimation of the average insert size. The subgenomic library compares well with large-insert DNA libraries created using genomic DNA of hexaploid wheat.

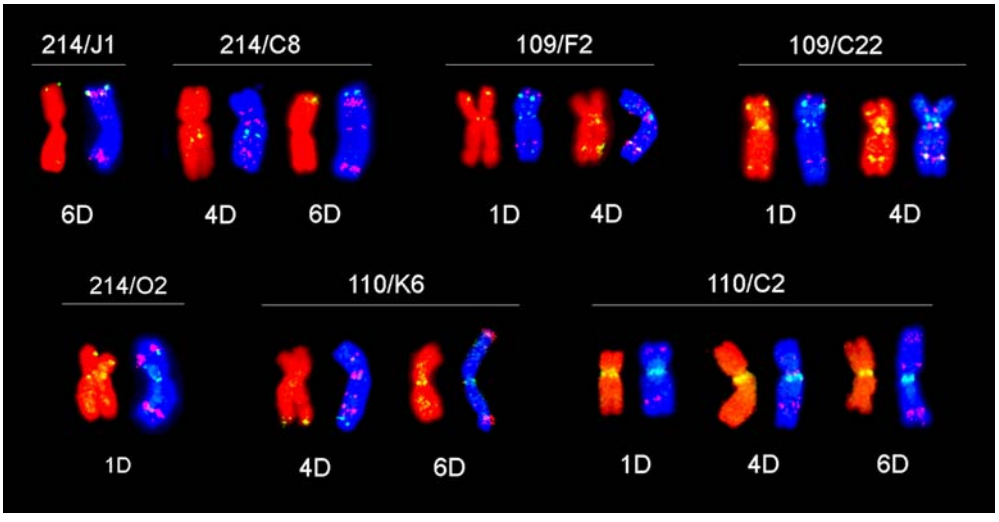

Fig. 3 Physical mapping of selected BAC clones, using FISH on wheat chromosomes. The addresses of individual BAC clones are given *above the lines*. For each BAC probe, two representative examples of FISH are given. In each pair of chromosomes on the figure, the chromosome on the *left* shows hybridization with a BAC probe detected by fluorescein (*yellow signals*); the chromosomes were counterstained by propidium iodide (*red*). The chromosomes on the *right* show simultaneous hybridization with the BAC probe detected by fluorescein (*yellow signals*) and a probe for the *Afa* repeat, detected with Cy3 (*red*), which facilitated chromosome identification; the chromosomes were counterstained by DAPI (*blue*)

Liu et al. (2000) reported on the construction of a transformation-competent artificial chromosome (TAC) library with average insert size of 54 kb and 3.07-fold genomic coverage. A BAC library created by Nilmalgoda et al. (2003) with average insert size of 79 kb, has a similar coverage of 3.1-fold. Recently, we achieved a much better coverage (9.3-fold) by constructing a BAC library from cv. Chinese Spring, consisting of 1.2×10^6 clones with average insert sizes of 130 kb (Allouis et al. 2003) as well as from cv. Renan (unpublished results). However, maintenance and handling of these two libraries with a huge number of BAC clones remains a non-trivial task.

Specificity of the TA-subD BAC library for chromosomes 1D, 4D and 6D was independently confirmed by two approaches. First, the sorted fractions were periodically checked microscopically for the presence of contaminating chromosomes. On average, 9% of chromosomes other than 1D, 4D and 6D were detected after FISH with the *Afa* repeat. This approach facilitates unambiguous identification of the D-genome chromosomes (Kubaláková et al. 2002). The second approach involved library screening with 36 SSR markers specific to chromosomes other than 1D, 4D and 6D (one marker per each chromosome arm). Although only a small number of SSR markers were used, in principle the results confirmed low contamination as determined by FISH. Only markers specific for chromosomes 1A, 2D, 5B and 7B were detected, demonstrating the high specificity of the library (Table 2).

One of the uses of the BAC library will be the construction of a detailed molecular cytogenetic map. Cytogenetic maps consisting of a series of molecular landmarks aid in construction of physical contig maps by resolving the order of local contigs and closely spaced clones, and confirming the precise physical position of centromere and markers at the end of linkage groups (Islam-Faridi et al. 2002; Kim et al. 2002). Cytogenetic maps are also very useful to study chromosome organization and karyotype evolution (Jackson et al. 2000; Ziolkowski and Sadowski 2002). Our results with a limited number of BAC clones indicate that the subgenomic BAC library will be a useful source of D-genome cytogenetic markers. Cytogenetically mapped low-copy BAC clones could serve as anchor makers and facilitate construction of physical contig maps (Stein et al. 2000; Faris et al. 2003). The reason that most of the 'low-copy BAC clones' as identified by no or weak hybridization with C_0t-1 fraction of genomic DNA were not chromosome specific is probably due the fact that these clones still contain few but highly repeated sequences which is not surprising considering the wheat genome structure (Akhunov et al. 2003; Chantret et al. 2004).

Our study indicates a possibility to isolate BAC clones specific to single-wheat chromosomes. This observation is surprising when considering the presence of three homoeologous chromosome sets in hexaploid wheat, and may be explained by a rapid and dynamic evolution of the *Triticum* genomes and divergence of homologous regions (Anderson et al. 2003; Wicker et al. 2003). Recently, Zhang et al. (2004) identified several BAC clones whose fragments hybridized only to specific chromosome pairs of wheat, thus supporting our observations. Localization of some BACs to non-homologous chromosomes could be due to interchromosomal duplications. Alternatively, these observations may reflect the structural changes accompanying the evolution of the hexaploid wheat karyotype (Gill et al. 1991). As none of the BAC clones has been sequenced, the nature of the presumably homologous sequences is not known. Identification of BAC clones containing specific repeats, such as the two clones containing centromere-specific repeats, can facilitate the analysis of molecular organization of functional domains of wheat chromosomes.

We expect that the most attractive application of the library will be gene cloning, as a number of agronomically interesting genes are known to be localized on one of the three D-genome chromosomes. These include seed storage protein genes (glutenins and gliadins) on chromosomes 1D and 6D (Anderson and Greene 1997; Sabelli and Shewry 1991; Spielmeyer et al. 2000) and genes related to disease resistance on chromosome 1D (Brooks et al. 2002). The *Rht-D1* dwarfing gene was mapped to the short arm of chromosome 4D (King et al. 1996; Cadalen et al. 1998; Sourdille et al. 1998). A major gene conferring aluminium tolerance as well as a locus controlling the K^+/Na^+ discrimination in saline environments were localized on the long arm of chromosome 4D (Rodriguez Milla and Gustafson 2001). Although the interesting genes may be cloned using the existing genomic BAC libraries (Moullet et al. 1999; Cenci et al. 2003; Allouis et al. 2003), the newly created BAC library specific for chromosomes 1D, 4D and 6D, which represents only about 11% of the whole genome, offers a targeted and hence more efficient and faster approaches for map-based cloning.

Targeting the gene-rich regions of the D genome is another attractive use of the library. This could be achieved by screening the library with appropriate probes such as ESTs to reveal gene-rich BAC clones. Alternatively, the library could be pooled and the chromosome DNA used for Cot analysis (Peterson et al. 2002; Yuan et al. 2003) to generate a shotgun library enriched in low-copy sequences for targeted sequencing of the genic fraction. Although the analysis could also be performed using genomic DNA, working with only about one tenth of the wheat genome should simplify the sequence annotation and assembly.

Rapid production of complete chromosome contigs is now possible using the subgenomic TA-subD BAC library. Their production would be greatly simplified by already existing genetic (Boyko et al. 1999, 2002) and physical (http://wheat.pw.usda.gov/PhysicalMapping/) maps for the diploid D genome of *Ae. tauschii*, which could serve as a framework and also by genetic/physical wheat map relationships that were recently published (Sourdille et al. 2004). The construction of complete BAC contigs for the three D-genome chromosomes has a potential to transform the positional cloning of their genes into a trivial bioinformatic exercise (Lagudah et al. 2001).

The TA-subD library will allow appreciation of the D-genome evolution in the hexaploid context and facilitate studies to explain the lack of polymorphism and difficulties associated with the development of genetic maps for the D genome. Is it not clear whether this is because the D genome is more conserved (if yes, why?): because it is less prone to evolution or because the recombination in the D genome is inhibited? A physical approach should allow answering these questions.

In conclusion, this study confirmed that the large genome of hexaploid wheat may be dissected into small and defined fractions corresponding to one or few chromosomes. While this paper reports on a BAC library specific for chromosomes 1D, 4D and 6D, we previously reported the preparation of the TA-3B BAC library specific for chromosome 3B (Šafář et al. 2004). These unique resources can greatly facilitate genome sequencing and gene cloning as well as comparative analyses aimed at revealing the genome changes accompanying the evolution of polyploid wheat. The work is in progress to construct BAC libraries specific for other chromosomes of hexaploid wheat.

Acknowledgements We thank our colleagues Arnaud Bellec, Michaela Libosvárová, Stephanie Pateyron, Radka Tušková and Jitka Weiserová for excellent technical assistance. This work was supported by research grants from the Grant Agency of the Czech Republic (522/03/0354 and 521/04/0607), Ministry of Agriculture of the Czech Republic (QC 1336), research grants from the French Ministry of Research and Agriculture, and by a collaborative project 'Barrande' (reg. no. 2003-035-2) between the Ministry of Education, Youth and Sports of the Czech Republic and the French Ministry of Foreign Affairs and Education. J. Janda acknowledges the receipt of a Marie Curie Ph.D. Student Fellowship at URGV-INRA.

References

Akhunov ED, Akhunova AR, Linkiewicz AM, et al (2003) Synteny perturbations between wheat homoeologous chromosomes caused by locus duplications and deletions correlate with recombination rates. Proc Natl Acad Sci USA 100:10836–10841

Allouis S, Moore G, Bellec A, Sharp R, Faivre P, Montimer K, Pateyron S, Foote T, Griffiths S, Caboche M, Chalhoub B (2003) Construction and characterisation of a hexaploid wheat (*Triticum aestivum* L.) BAC library from the reference germplasm 'Chinese Spring'. Cereal Res Comm 31:331–338

Anderson OD, Greene FC (1997) The alpha-gliadin gene family. II. DNA and protein sequence variation, subfamily structure, and origins of pseudogenes. Theor Appl Genet 95:59–65

Anderson OD, Rausch C, Moullet O, Lagudah E (2003) The wheat D-genome HMW-glutenin locus: BAC sequencing, gene distribution, and retrotransposon clusters. Funct Integr Genomics 3:56–68

Bennett MD, Smith JB (1991) Nuclear DNA amounts in angiosperms. Proc R Soc Lond B 334:309–345

Boyko EV, Gill KS, Mickelson-Young L, Nasuda S, Raupp WJ, Ziegle JN, Singh S, Hassawi DS, Fritz AK, Namuth D, Lapitan NL, Gill BS (1999) A high-density genetic linkage map of *Aegilops tauschii*, the D-genome progenitor of bread wheat. Theor Appl Genet 99:16–26

Boyko E, Kalendar R, Korzun V, Fellers J, Korol A, Schulman AH, Gill BS (2002) A high-density cytogenetic map of the *Aegilops tauschii* genome incorporating retrotransposons and defense-related genes: insights into cereal chromosome structure and function. Plant Mol Biol 48:767–790

Brooks SA, Huang L, Gill BS, Fellers JP (2002) Analysis of 106 kb of contiguous DNA sequence from D genome of wheat reveals high gene density and a complex arrangement of genes related to disease resistance. Genome 45:963–972

Cadalen T, Sourdille P, Charmet G, Tixier MH, Gay G, Boeuf C, Bernard S, Leroy P, Bernard M (1998) Molecular markers linked to genes affecting plant height in wheat using a doubled-haploid population. Theor Appl Genet 96:933–940

Cenci A, Chantret N, Kong X, Gu Y, Anderson OD, Fahima T, Distelfeld A, Dubcovsky J (2003) Construction and characterization of a half million clone BAC library of durum wheat (*Tridicum turgidum* ssp. *durum*). Theor Appl Genet 107:931–939

Chalhoub B, Belcram H, Caboche M (2004) Efficient cloning of plant genomes into bacterial artificial chromosome (BAC) libraries with larger and more uniform insert size. Plant Biotechnol J 2:181–188

Chantret N, Cenci A, Sabot F, Anderson O, Dubcovsky J (2004) Sequencing of the *Triticum monococcum* hardness locus reveals good microcolinearity with rice. Mol Genet Genomics 271:377–386

Clarke L, Carbon J (1976) A colony bank containing synthetic Col El hybrid plasmids representative of the entire *E. coli* genome. Cell 9:91–99

Daud HM, Gustafson JP (1996) Molecular evidence for *Triticum speltoides* as a B-genome progenitor of wheat (*Triticum aestivum*). Genome 39:543–548

Doležel J, Kubaláková M, Bartoš J, Macas J (2004) Flow cytogenetics and plant genome mapping. Chromosome Res 12:77–91

Dvorak J, di Terlizzi P, Zhang HB, Resta P (1993) The evolution of polyploid wheats—Identification of the A-genome donor species. Genome 36:21–31

Faris JD, Fellers JP, Brooks SA, Gill BS (2003) A bacterial artificial chromosome contig spanning the major domestication locus Q in wheat and identification of a candidate gene. Genetics 164:311–321

Feldman M, Lupton FGH, Miller TE (1995) Wheats. In: Smartt J, Simmonds NW (eds) Evolution of crops, 2nd edn. Longman Scientific, London, pp 184–192

Gill BS, Appels R, Botha-Oberholster AM, Buell RC, Bennetzen JL, Chalhoub B, Chumley F, Dvorak J, Iwanaga M, Keller B, Li W, McCombie WR, Ogihara Y, Quetier F, Sasaki T (2004) A Workshop Report on Wheat Genome Sequencing. International Genome Research on Wheat (IGROW) Consortium. Genetics (in press)

Gill BS, Friebe B, Endo TR (1991) Standard karyotype and nomenclature system for description of chromosome bands and structural aberrations in wheat (*Triticum aestivum*). Genome 34:830–839

Guyomarc'h H, Sourdille P, Charmet G, Edwards KJ, Bernard M (2002) Characterisation of polymorphic markers from *T. tauschii* and transferability to the D-genome of bread wheat. Theor Appl Genet 104:1164–1172

Islam-Faridi MN, Childs KL Klein PE, Hodnett G, Menz MA, Klein RR, Rooney WL, Mullet JE, Stelly DM, Price HJ (2002) A molecular cytogenetic map of sorghum chromosome 1: fluorescence in situ hybridization analysis with mapped bacterial artificial chromosomes. Genetics 161:345–353

Jackson SA, Cheng Z, Wang ML, Goodman HM, Jiang J (2000) Comparative fluorescence in situ hybridization mapping of a 431-kb *Arabidopsis thaliana* bacterial artificial chromosome contig reveals the role of chromosomal duplications in the expansion of the *Brassica rapa* genome. Genetics 156:833–838

Kim JS, Childs KL, Islam-Faridi MN, Menz MA, Klein RR, Klein PE, Price HJ, Mullet JE, Stelly DM (2002) Integrated karyotyping of sorghum by in situ hybridization of landed BACs. Genome 45:402–412

King IP, Cant KA, Law CN, Worland AJ, Orford SE, Reader SM, Miller TE (1996) An assessment of the potential of 4DS.4DL-4sL translocation lines as a means of eliminating tall off types in semi-dwarf wheat varieties. Euphytica 89:103–106

Kubaláková M, Vrána J, Číhalíková J, Šimková H, Doležel J (2002) Flow karyotyping and chromosome sorting in bread wheat (*Triticum aestivum* L.). Theor Appl Genet 104:1362–1372

Lagudah E, Dubcovsky J, Powell W (2001) Wheat genomics. Plant Physiol Biochem 39:1–10

Lijavetzky D, Muzzi G, Wicker T, Keller B, Wing R, Dubcovsky J (1999) Construction and characterization of a bacterial artificial chromosome (BAC) library for the A genome of wheat. Genome 42:1176–1182

Ling HQ, Zhu Y, Keller B (2003) High-resolution mapping of the leaf rust disease resistance gene *Lr1* in wheat and characterization of BAC clones from the *Lr1* locus. Theor Appl Genet 106:875–882

Liu YG, Nagaki K, Fujita M, Kawaura K, Uozumi M, Ogohara Y (2000) Development of an efficient maintenance and screening system for large-insert genomic DNA libraries of hexaploid wheat in a transformation-competent artificial chromosome (TAC) vector. Plant J 23:687–695

Ma Z, Weining S, Sharp PJ, Liu C (2000) Non-gridded library: a new approach for BAC (bacterial artificial chromosome) exploitation in hexaploid wheat (*Triticum aestivum*). Nucleic Acids Res 28:106–111

Moullet O, Zhang HB, Lagudah ES (1999) Construction and characterisation of a large DNA insert library from the D genome of wheat. Theor Appl Genet 99:305–313

Nilmalgoda SD, Cloutier S, Walichnowski AZ (2003) Construction and characterization of a bacterial artificial chromosome (BAC) library of hexaploid wheat (*Triticum aestivum* L.) and validation of genome coverage using locus-specific primers. Genome 46:870–878

Ogihara Y, Isono K, Kojima T, Endo A, Hanaoka M, Shiina T, Terachi T, Utsugi S, Murata M, Mori N, Takumi S, Ikeo K, Gojobori T, Murai R, Murai K, Matsuoka Y, Ohnishi Y, Tajiri H, Tsunewaki K (2000) Chinese spring wheat (*Triticum aestivum* L.) chloroplast genome: complete sequence and contig clones. Plant Mol Biol Rep 18:243–253

Ozkan H, Levy AA, Feldman M (2001) Allopolyploidy-induced rapid genome evolution in the wheat (*Aegilops Triticum*) group. Plant Cell 13:1735–1747

Peterson DG, Wessler SR, Paterson AH (2002) Efficient capture of unique sequences from eukaryotic genomes. Trends Genet 18:547–550

Röder MS, Korsun V, Wendehake K, Plaschke J, Tixier MH, Leroy P, Ganal MW (1998) A microsatellite map of the wheat genome. Genetics 149:2007–2023

Rodriguez Milla MA, Gustafson JP (2001) Genetic and physical characterization of chromosome 4DL in wheat. Genome 44:883–892

Sabelli PA, Shewry RB (1991) Characterization and organization of gene families at the *Gli-1* loci of bread and durum wheats by restriction fragment analysis. Theor Appl Genet 83:209–216

Šafář J, Bartoš J, Janda J, Bellec A, Kubaláková M, Valárik M, Pateyron S, Weiserová J, Tušková J, Číhalíková J, Vrána J, Šimková H, Faivre-Rampant P, Sourdille P, Caboche M, Bernard M, Doležel J, Chalhoub B (2004) Dissecting large and complex genomes: flow sorting and BAC cloning of individual chromosomes from bread wheat. Plant J 39:960–968

Šimková H, Číhalíková J, Vrána J, Lysák MA, Doležel J (2003) Preparation of HMW DNA from plant nuclei and chromosomes isolated from root tips. Biol Plant 46:369–373

Smith DB, Flavell RB (1975) Characterization of the wheat genome by renaturation kinetics. Chromosoma 50:223–242

Sourdille P, Charmet G, Trottet M, Tixier MH, Boeuf C, Nègre S, Barloy D, Bernard M (1998) Linkage between RFLP molecular markers and the dwarfing genes *Rht-B1* and *Rht-D1* in wheat. Hereditas 128:41–46

Sourdille P, Guyomarc'h H, Baron C, Gandon B, Chiquet V, Artiguenave F, Edwards K, Foisset N, Dufour P, Bernard M (2001) Improvement of the genetic maps of wheat using new microsatellite markers. In: Proceedings of 9th plant and animal genome. Final abstract guide. San Diego, p 167

Sourdille P, Singh S, Cadalen T, Brown-Guedira GL, Gay G, Qi L, Gill BS, Dufour P, Murigneux A, Bernard M (2004) Microsatellite-based deletion mapping system for the establishment of genetic map—physical map relationships in wheat. Funct Integr Genomics 4:12–25

Spielmeyer W, Moullet O, Laroche A, Lagudah ES (2000) Highly recombinogenic regions at seed storage protein loci on chromosome 1DS of *Aegilops tauschii*, the D-genome donor of wheat. Genetics 155:361–367

Stein N, Feuillet C, Wicker T, Schlagenhauf E, Keller B (2000) Subgenome chromosome walking in wheat: a 450-kb physical contig in *Triticum monococcum* L. spans the *Lr10* resistance locus in hexaploid wheat (*Triticum aestivum* L.). Proc Natl Acad Sci USA 97:13346–13444

Vrána J, Kubaláková M, Šimková H, Číhalíková J, Lysák MA, Doležel J (2000) Flow-sorting of mitotic chromosomes in common wheat (*Triticum aestivum* L.). Genetics 156:2033–2041

Wicker T, Stein N, Albar L, Feuillet C, Schlagenhauf E, Keller B (2001) Analysis of a continuous 211 kb sequence in diploid wheat (*Triticum monococcum* L.) reveals multiple mechanisms of genome evolution. Plant J 26:307–316

Wicker T, Yahiaoui N, Guyot R, Schlagenhauf E, Liu ZD, Dubcovsky J, Keller B (2003) Rapid genome divergence at orthologous low molecular weight glutenin loci of the A and A^m genomes of wheat. Plant Cell 15:1186–1197

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

Yan L, Loukoianov A, Tranquilli G, Helguera M, Fahima T, Dubcovsky J (2003) Positional cloning of the wheat vernalisation gene VRN1. Proc Natl Acad Sci USA 100:6263–6268

Yuan Y, SanMiguel PJ, Bennetzen JL (2003) High-Cot sequence analysis of the maize genome. Plant J 34:249–255

Zhang HB, Wu C (2001) BAC as tools for genome sequencing. Plant Physiol Biochem 39:195–209

Zhang P, Li W, Fellers J, Friebe B, Gill BS (2004) BAC-FISH identifies chromosome landmarks consisting of different types of transposable elements. Chromosoma 112:288–299

Ziolkowski PA, Sadowski J (2002) FISH-mapping of rDNA and *Arabidopsis* BACs on pachytene complements of selected *Brassicas*. Genome 45:189–197

Zwick MS, Hanson RE, McKnith TD, Islam-Faridi MN, Stelly DM, Wing RA, Price HJ (1997) A rapid procedure for the isolation of C(0)t-1 DNA from plants. Genome 40:138–142