Chromosoma (2004) 112:288–299
DOI 10.1007/s00412-004-0273-9

RESEARCH ARTICLE

Peng Zhang · Wanlong Li · John Fellers · Bernd Friebe · Bikram S. Gill

BAC-FISH in wheat identifies chromosome landmarks consisting of different types of transposable elements

Received: 26 August 2003 / Revised: 16 December 2003 / Accepted: 20 January 2004 / Published online: 18 February 2004

Abstract Fluorescence in situ hybridization (FISH) has been widely used in the physical mapping of genes and chromosome landmarks in plants and animals. Bacterial artificial chromosomes (BACs) contain large inserts making them amenable for FISH mapping. We used BAC-FISH to study genome organization and evolution in hexaploid wheat and its relatives. We selected 56 restriction fragment length polymorphism (RFLP) locus-specific BAC clones from libraries of *Aegilops tauschii* (the D-genome donor of hexaploid wheat) and A-genome diploid *Triticum monococcum*. Different types of repetitive sequences were identified using BAC-FISH. Two BAC clones gave FISH patterns similar to the repetitive DNA family pSc119; one BAC clone gave a FISH pattern similar to the repetitive DNA family pAs1. In addition, we identified several novel classes of repetitive sequences: one BAC clone hybridized to the centromeric regions of wheat and other cereal species, except rice; one BAC clone hybridized to all subtelomeric chromosome regions in wheat, rye, barley and oat; one BAC clone contained a localized tandem repeat and hybridized to five D-genome chromosome pairs in wheat; and four BAC clones hybridized only to a proximal region in the long arm of chromosome 4A of hexaploid wheat. These repeats are valuable markers for defined chromosome regions and can also be used for chromosome identification. Sequencing results revealed that all these repeats are transposable elements (TEs), indicating the important role of TEs, especially retrotransposons, in genome evolution of wheat.

Communicated by P.B. Moens

P. Zhang · W. Li · B. Friebe (✉) · B. S. Gill
Wheat Genetics Resource Center, Department of Plant Pathology,
4024 Throckmorton Plant Sciences Center,
Kansas State University,
Manhattan, KS 66506-5502, USA
e-mail: friebe@ksu.edu

J. Fellers
USDA-ARS, Plant Science & Entomology Unit,
Kansas State University,
Manhattan, KS 66506-5502, USA

Introduction

Fluorescence in situ hybridization (FISH) has been used for physical mapping of repetitive DNA sequences on chromosomes and genomes. Highly repetitive DNA sequences can be divided into two types based on their organization and distribution pattern in the genome. One type comprises localized tandem repeats that are arranged in tandem arrays and are present as clusters along chromosomes, usually associated with heterochromatin, such as the tandem repetitive DNA sequences present in clones pSc74 (Bedbrook et al. 1980) and pSc119.2 (Bedbrook et al. 1980; McIntyre et al. 1990) isolated from *Secale cereale* L., pAs1 isolated from *Aegilops tauschii* Coss. (Rayburn and Gill 1986b), pAesKB52 isolated from *Ae. speltoides* Tausch (Anamthawat-Jonsson and Heslop-Harrison 1993), and pHvG38 isolated from *Hordeum vulgare* L. (Pedersen et al. 1996). In the tribe Triticeae, localized tandem repeated sequences are useful as cytogenetic markers for chromosome identification (Bedbrook et al. 1980; Appels et al. 1981; Jones and Flavell 1982; Rayburn and Gill 1986b; McIntyre et al. 1990; Mukai et al. 1992; Anamthawat-Jonsson and Heslop-Harrison 1993; Pedersen and Langridge 1997). Another type of repetitive sequences comprises dispersed repeats that are interspersed with unrelated repeated or unique DNA sequences (Flavell 1986). Dispersed repeats can be found at a few chromosomal regions or distributed throughout an entire genome, such as the rye-specific dispersed repetitive DNA sequence present in clone pSc119.1 (McIntyre et al. 1990). Dispersed repetitive sequences can be used to identify parental genomes in interspecific hybrids (Hutchinson and Lonsdale 1982; Appels and Moran 1984; McIntyre et al. 1990; Anamthawat-Jonsson and Heslop-Harrison 1993; Vershinin et al. 1994) and to study plant genome evolution (Dvorák and Zhang 1992).

The use of genomic DNA cloned in large-insert vectors such as bacterial artificial chromosomes (BACs) (Shizuya et al. 1992) in combination with FISH, called BAC-FISH, has been an effective approach for physically mapping specific DNA sequences and identifying individual chromosomes in humans and in plants with small genomes such as rice (Jiang et al. 1995; Cheng et al. 2001, 2002), cotton (Hanson et al. 1995), sorghum (Gomez et al. 1997; Kim et al. 2002), *Arabidopsis* (Fransz et al. 2000; Jackson et al. 2000), and potato (Dong et al. 2000). Here we report BAC-FISH analysis of the wheat genome and identification of markers for various chromosome domains.

Materials and methods

Plant material

Eleven diploid, seven tetraploid, and four hexaploid *Aegilops* species, and two each of diploid, tetraploid and hexaploid *Triticum* species were analyzed. In addition, *Oryza sativa* L., *Zea mays* L., *Avena sativa* L., *S. cereale* , and *H. vulgare* also were included. Accession numbers, genome constitutions, and chromosome numbers are given in Table 1. The nomenclature for *Aegilops* and *Triticum* follows that of van Slageren (1994) except that *Ae. mutica* is considered as belonging to the genus *Aegilops*, not the genus *Amblyopyrum* as suggested by van Slageren. The DNA amount is expressed as a 1C value according to Bennett and Smith (1976) and Bennett and Leitch (1995). All materials are maintained at the Wheat Genetics Resource Center at Kansas State University.

Screening of BAC clones

Altogether, 56 BAC clones from three libraries were used, including 29 derived from a *T. monococcum* subsp. *monococcum* BAC library (Lijavetzky et al. 1999), 11 derived from an *Ae. tauschii* binary BAC (BIBAC) library (Moullet and Lagudah 1998; Moullet et al. 1999), and 16 from an *Ae. tauschii* BAC library developed within a National Science Foundation project (http://wheat.pw.usda.gov/PhysicalMapping/) (Table 2).

Selection of BAC clones from the above libraries was done by screening high-density BAC filters with genetically mapped restriction fragment length polymorphism (RFLP) markers. Among the 29 *T. monococcum* BAC clones, 27 were previously mapped to the short arms of group-5 chromosomes and two to chromosome arm 3AL (Li and Gill 2002; L.L. Qi and B.S. Gill, personal communication). Eleven *Ae. tauschii* BIBAC clones were mapped to chromosome arm 1DS (Spielmeyer et al. 2000). The 16 *Ae. tauschii* BAC clones were selected using pools of RFLP probes that

Table 1 Plant materials used in this study

Species	Accession	Genome	1C DNA (pg)
Secale cereale L. cv. Blanco	TA9020	2 *n* =2 *x* =14, RR	8.3
Hordeum vulgare L. cv. Betzes	TA9001	2 *n* =2 *x* =14, HH	5.5
Avena sativa L. cv. Starter	TA9012	2 *n* =6 *x* =42, AACCDD	11.7
Zea mays L.		2 *n* =2 *x* =20	2.8
Oryza sativa L. ssp. *indica*		2 *n* =2 *x* =24	0.5
Aegilops speltoides Tausch	TA1770, TA1778	2 *n* =2 *x* =14, SS	5.1
Ae. longissima (Schweinf. & Muschl. *in* Muschl.) Eig	TA1910, TA1912	2 *n* =2 *x* =14, S^lS^l	6.9
Ae. sharonensis Eig	TA1995, TA1996	2 *n* =2 *x* =14, $S^{sh}S^{sh}$	7.3
Ae. searsii Feldman & Kislev *ex* K. Hammer	TA1841, TA2355	2 *n* =2 *x* =14, S^sS^s	5.9
Ae. bicornis (Forsskal) Jaub. & Spach.	TA1942, TA1954	2 *n* =2 *x* =14, S^bS^b	7.2
Ae. tauschii Coss.	AL8/78, TA2467	2 *n* =2 *x* =14, DD	4.2
Ae. mutica Boiss	TA2753	2 *n* =2 *x* =14, TT	6.3
Ae. caudata L.	TA1908, TA1909	2 *n* =2 *x* =14, CC	4.6
Ae. comosa Sm. *in* Sibth. & Sm. var. *comosa*	TA2104, TA2731	2 *n* =2 *x* =14, MM	6.2
Ae. comosa Sm. *in* Sibth. & Sm. var. *subventricosa* Bioss.	TA1968	2 *n* =2 *x* =14, M^hM^h	6.2
Ae. uniaristata Vis.	TA2688, TA2696	2 *n* =2 *x* =14, NN	6.3
Ae. umbellulata Zhuk.	TA1831	2 *n* =2 *x* =14, UU	5.1
Ae. biuncialis Vis.	TA1959	2 *n* =4 *x* =28, $U^bU^bM^bM^b$	11.3
Ae. columnaris Zhuk.	TA2105	2 *n* =4 *x* =28, $U^cU^cM^cM^c$	10.5
Ae. crassa Boiss.	TA2217	2 *n* =6 *x* =42, $X^{cr}X^{cr}D^{cr1}D^{cr1}D^{cr2}D^{cr2}$	15.7
Ae. cylindrica Host	TA2203	2 *n* =4 *x* =28, $C^cC^cD^cD^c$	8.8
Ae. geniculata Roth	TA1798, TA1811, TA1818	2 *n* =4x=28, $U^gU^gM^gM^g$	9.2
Ae. juvenalis (Thell.) Eig	TA2116	2 *n* =6 *x* =42, $X^jX^jD^jD^jU^jU^j$	18.8
Ae. neglecta Req. *ex* Bertol.	TA1868	2 *n* =6 *x* =42, $U^nU^nM^nM^nN^nN^n$	16.4
Ae. peregrina Hackel	TA2775	2 *n* =4 *x* =28, $S^pS^pU^pU^p$	13.8
Ae. triuncialis L.	TA1732	2 *n* =4 *x* =28, $U^tU^tC^tC^t$	9.5
Ae. vavilovii (Zhuk.) Chennav.	TA2655	2 *n* =6 *x* =42, $X^{va}X^{va}D^{va}D^{va}S^{va}S^{va}$	18.3
Ae. ventricosa Tausch	TA1993	2 *n* =4 *x* =28, $D^vD^vN^vN^v$	9.8
Triticum urartu Tumanian *ex* Gandilyan	TA830, TA858	2 *n* =2 *x* =14, AA	5.7
T. monococcum L. subsp. *monococcum*	TA141, TA142	2n=2 *x* =14, A^mA^m	5.8
T. timopheevii (Zhuk.) Zhuk. subsp. *armeniacum* (Jakubz.) MacKey	TA1475, TA1557	2 *n* =4 *x* =28, A^tA^tGG	10.1
T. timopheevii (Zhuk.) Zhuk. subsp. *timopheevii*	TA103, TA140	2 *n* =4 *x* =28, A^tA^tGG	9.6
T. turgidum L. subsp. *durum* cv. Langdon & Rubezh	TA3028, TA3029	2 *n* =4 *x* =28, AABB	13.5
T. zhukovskyi Menabde & Ericz.	TA2610	2 *n* =6 *x* =42, AAA^mA^mGG	15.4
T. aestivum L. subsp. *aestivum* cv. Chinese Spring	TA3008	2 *n* =6 *x* =42, AABBDD	17.9

Table 2 List of bacterial artificial chromosome (BAC) clones used in fluorescent in situ hybridization (FISH) experiments

Triticum monococcum BACs		*Aegilops tauschii* BIBACs	*Ae. tauschii* BACs
563D4 [a]	647D3 [a]	M11 [c]	1I9 [d]
578A12 [a]	673A20 [a]	I11 [c]	2O17 [d]
652O9 [a]	678F8 [a]	A6 [c]	3H18 [d]
660M17 [a]	691C20 [a]	D9 [c]	3I19 [d]
694M20 [a]	692G13 [a]	H11 [c]	4D22 [d]
705E10 [a]	713O20 [a]	M4 [c]	4P6 [d]
363M9 [a]	535A1 [a]	O10 [c]	5D10 [d]
689N10 [a]	539G1 [a]	B1 [c]	5L16 [d]
369N3 [a]	584B8 [a]	D8 [c]	6A23 [d]
616H8 [a]	585D8 [a]	E4 [c]	6C6 [d]
676D4 [a]	605E16 [a]	F12 [c]	6D16 [d]
691K15 [a]			6E9 [d]
439J21 [a]	683A21 [b]		7D1 [d]
532J13 [a]	611L12 [b]		8C19 [d]
542H3 [a]			9I10 [d]
561N11 [a]			9M13 [d]

[a] BAC clones selected using restriction fragment length polymorphism (RFLP) probes that hybridized to short arms of group-5 chromosomes (L.L. Qi and B.S. Gill, personal communication)
[b] BAC clones selected using RFLP probes that hybridized to the long arm of chromosome 3A (Li and Gill 2002)
[c] BAC clones mapped to chromosome arm 1DS (Spielmeyer et al. 2000)
[d] BAC clones selected using pools of RFLP probes that hybridized to chromosome groups 3, 5, 6, and 7 (W. Li and B.S. Gill, unpublished data)

hybridize to group 3, 5, 6, and 7 chromosomes (W. Li and B.S. Gill, unpublished data) (Table 2).

Fluorescence in situ hybridization analysis

Seed germination, root-tip pretreatment, squash preparations, slide pretreatment and denaturation were according to Zhang et al. (2001).

Isolation of BAC and plasmid DNA was done using a Qiagen Plasmid Midi Kit and a QIAprep Spin Miniprep Kit (Qiagen, Valencia, Calif.), respectively. One microgram of BAC or plasmid DNA was labeled with biotin-14-dATP using the BioNick Labeling System (Invitrogen Life Technologies, Carlsbad, Calif.) or with rhodamine-6-dUTP using Rhodamine-Nick Translation Mix (Roche Applied Science, Indianapolis, Ind.) using nick translation according to the manufacturer's protocols. Probes were purified with QIAquick Nucleotide Removal Kit (Qiagen, Valencia, Calif.).

Fluorescence in situ hybridization was as described by Zhang et al. (2001) with a few modifications. The hybridization solution (approximately 75% hybridization stringency), containing 50% deionized formamide (Fisher Scientific, Pittsburgh, Pa.), 2×SSC, 10% dextran sulfate (Sigma, St. Louis, Mo.), 0.3 mg/ml of sheared salmon testes DNA (Sigma, St. Louis, Mo.), and about 1 μg/ml of labeled probes, was denatured by boiling for 5 min. After chilling on ice for 5 min, 30 μl of the denatured hybridization mixture was applied to each slide and hybridization was performed in a humid chamber at 37°C overnight. Post-hybridization washes were in 2×SSC at 42°C for 10 min, 50% formamide in 2×SSC at 42°C for 10 min (approximately 80% washing stringency), and 2×SSC at 42°C for 10 min. The biotin-labeled probes were detected with fluorescein-avidin DN (Vector Laboratories, Burlingame, Calif.). Chromosomes were counterstained with propidium iodide (Molecular Probes, Eugene, Ore.) in Vectashield (Vector Laboratories, Burlingame, Calif.). The rhodamine-labeled probes were counterstained with 4′,6-diamidino-2-phenylindole (Molecular Probes, Eugene, Ore.) in Vectashield. Slides were analyzed with an epifluorescence Zeiss Axioplan 2 microscope. Images were captured using a SPOT CCD (charge-coupled device) camera operated with SPOT 2.1 software (Diagnostic Instruments, Sterling Heights, Mich.) and processed with Adobe Photoshop v5.5 (Adobe Systems, San Jose, Calif.).

Shotgun subcloning

The TOPO Shotgun Subcloning Kit (Invitrogen Life Technologies, Carlsbad, Calif.) was used for subcloning. Approximately 15 μg of BAC DNA isolated with a Qiagen Plasmid Midi Kit was nebulized at 15 psi (pounds/in^2) for 10 min. Following precipitation and resuspension, the sheared DNA was electrophoresed, and DNA fragments of 500 bp to 1.5 kb were selected. DNA was gel purified using a NucleoTrap Gel Extraction Kit (Clontech, Palo Alto, Calif.), blunt-end repaired, dephosphorylated, and ligated into the pCR®4Blunt-TOPO® vector. The ligation was transformed into TOP10 *Escherichia coli* using electroporation. The transformation culture was plated out on petri dishes with LB-carbenicillin, isopropyl-1-thio-galactopyranoside, 5-bromo-4-chloro-3-indolyl-galactopyranoside and incubated at 37°C overnight. White colonies were picked and transferred into 384-well microtiter plates.

A replicator was used to make a copy of the colonies in the 384-well plate onto N^+ Hybond membrane (Amersham Biosciences, Piscataway, N.J.). Bacteria were allowed to grow overnight, lysed, denatured on membrane in 0.4 M NaOH for 5 min, neutralized in 2×SSC for 5 min, and vigorously shaken in 2×SSC for 15 min. Colony-blot hybridization was carried out with the genomic DNA of either *T. monococcum* or *Ae. tauschii* as a probe depending on the BAC clone. Colonies that gave strong hybridization signals were selected and plasmid DNA was isolated using a QIAprep Spin Miniprep Kit for FISH analysis.

Sequencing analysis

BAC subclones that generated the same FISH hybridization patterns as the original BAC clones were sequenced. Contigs were assembled using phrap version 0.990319 (http://www.phrap.org) and AssemblyLIGN™ version 1.0.9c (Oxford Molecular, Madison, Wis.) (Brooks et al. 2002). Sequences were then tested for similarities to other known sequences in the public databases using the BLASTn search.

DNA dot-blot analysis

Plant genomic DNA was isolated according to Faris et al. (2000). Plant genomic DNA and plasmid DNA concentrations were determined by spectrophotometry and gel electrophoresis. Five

microliters of each plant genomic DNA (200C/μl) were spotted onto a N^+ Hybond membrane using a multichannel pipet. A serial dilution of the plasmid DNAs was used as a standard with copy number ranging from 10^4 to 10^8. After spotting the DNA onto the membrane, the membrane was denatured in 0.4 M NaOH for 5 min, neutralized with 2×SSC for 5 min, and the DNA was cross-linked to membrane using UV light and then dried. These membranes then were used for dot-blot hybridization.

Prehybridization, hybridization, probe labeling by the random hexamer method with [α-^{32}P]dCTP, and filter washing were performed as described by Faris et al. (2000). The exposure time of membranes to X-ray film depended on signal intensity and ranged from a few hours at room temperature to 3–7 days at −80°C. The copy number of specific sequences in plant genomic DNA was estimated by comparing the hybridization intensity with the standard.

Results

All BAC clones were first hybridized to mitotic metaphase chromosomes of the hexaploid wheat cv. Chinese Spring (CS). Forty BAC clones hybridized to all chromosomes, and the hybridization signals covered the entire chromosome length (Fig. 1a). This type of hybridization pattern indicates that these BACs contain repetitive DNA sequences that are shared among all wheat chromosomes and/or that different types of repeated sequences are interspersed. Three clones, 660M17, 532J13, and 539G1, did not give a hybridization signal on any wheat chromosome. These BACs were shown not to be empty clones or chloroplast DNA contamination using RFLP analysis (data not shown); the reason for their lack of hybridization is not known at present. One BAC clone from the *Ae. tauschii* BIBAC library, M11, and two BAC clones from the *T. monococcum* library, 689N10 and 678F8, gave FISH patterns similar to those of the repetitive DNA families pAs1 and pSc119 (data not shown) (Bedbrook et al. 1980; Rayburn and Gill 1986b; McIntyre et al. 1990; Mukai et al. 1993). The 120 bp repeat present in clone pSc119.2 was classified as a Ty*3*/ *gypsy* retrotransposon-like sequence and designated as Daniela. This sequence has a very strong similarity at the amino acid level to putative retrotransposon polyproteins from *O. sativa*, *S. bicolor* (L.) Moench, *Z. mays* , and *Arabidopsis thaliana* (L.) Heynh. (Johnston and Pickering 2002). One *T. monococcum* BAC and two *Ae. tauschii* BACs contain repeats that preferentially hybridized to the A- and D-genome chromosomes, respectively; these are discussed elsewhere (Zhang 2002).

Centromere-specific repeat

BAC clone 6C6 from the *Ae. tauschii* library hybridized to the centromeric regions of wheat (Fig. 1b), rye, barley, and maize but not to rice centromeres (data not shown). The hybridization intensity was similar in wheat and rye but lower in barley and maize.

Eight shotgun subclones, ranging in size from 550 bp to 1.5 kb, were selected that gave strong hybridization signals in colony-hybridization blots using the genomic DNA of *Ae. tauschii* as a probe. Among these subclones, five hybridized to the centromeric region. These five positive subclones were sequenced and assembled into two contigs, 3 (GenBank Accession number AY249982) and 4 (GenBank Accession number AY249981; Table 3), which are 1084 bp and 1138 bp, respectively. Sequencing showed that they have 92% sequence identity to the Ty*3*/ *gypsy* centromeric retrotransposon *cereba* gag-pol polyprotein in *H. vulgare*, including a putative RNaseH catalytic domain and a putative integrase catalytic domain (Presting et al. 1998; Hudakova et al. 2001); 84% sequence identity to the gag-pol polyprotein of the CRM centromeric retrotransposon; and 84% sequence identity to a putative retroelement that is similar to the gag-pol polyprotein in *O. sativa* (NCBI database) (Table 4). This repeat has 80% homology with the dispersed centromeric-repeat family RCS1 (Dong et al. 1998).

Dot-blot analysis using subclone pAet6-J9 (Table 3) as a probe revealed that the copy number of this sequence was about 700 in rye, 400 in barley, and ranged from 700 to 900 in diploid *Triticum/Aegilops* species, from 4000 to 8000 in tetraploid species, and from 4000 to 9000 in hexaploid species.

pRCS1, a centromere-specific sequence isolated from rice, hybridizes to the centromeric regions of rice, maize, and wheat but not to barley (Dong et al. 1998). In order to test whether the centromere-specific sequence in BAC 6C6 is similar to that present in pRCS1, BAC 6C6 was digested with three different enzymes, *Eco*RI, *Bam*HI, and *Hin*dIII, and two blots were made. The different Southern hybridization patterns found using subclone pAet6-K11 of BAC 6C6 and pRCS1 (Dong et al. 1998) as probes indicate that the centromere-specific sequence in BAC 6C6 is different from that present in pRCS1 (data not shown). In addition, sequencing results also revealed differences between the centromeric repeats in BAC 6C6 and pRCS1. Although both sequences have high sequence identity to the Ty*3*/ *gypsy* centromeric retrotransposon *cereba* gag-pol polyprotein of *H. vulgare*, they match different domains in *cereba*. The centromere-specific repeat in BAC 6C6 matched the RNase H catalytic and integrase catalytic domains, whereas the repeat in pRCS1 matched the integrase catalytic domain and the long terminal repeat (LTR).

Subtelomeric repeat

Clone A6 from the *Ae. tauschii* BIBAC library hybridized strongly to all subtelomeric chromosome regions in wheat (Fig. 1c) and rye (Fig. 1d), indicating that this repeat is conserved in the subtelomeric regions of cereal genomes. As expected, this BAC gave similar chromosome FISH patterns in *T. urartu*, *T. monococcum* , *Ae. speltoides*, and *Ae. tauschii* (data not shown). This BAC also hybridized to the subtelomeric chromosome regions in barley and oat but at a much lower intensity (data not shown). The hybridization signal was stronger in oat than in barley. Dot-blot results using one of the subclones, pAet7-L3, as

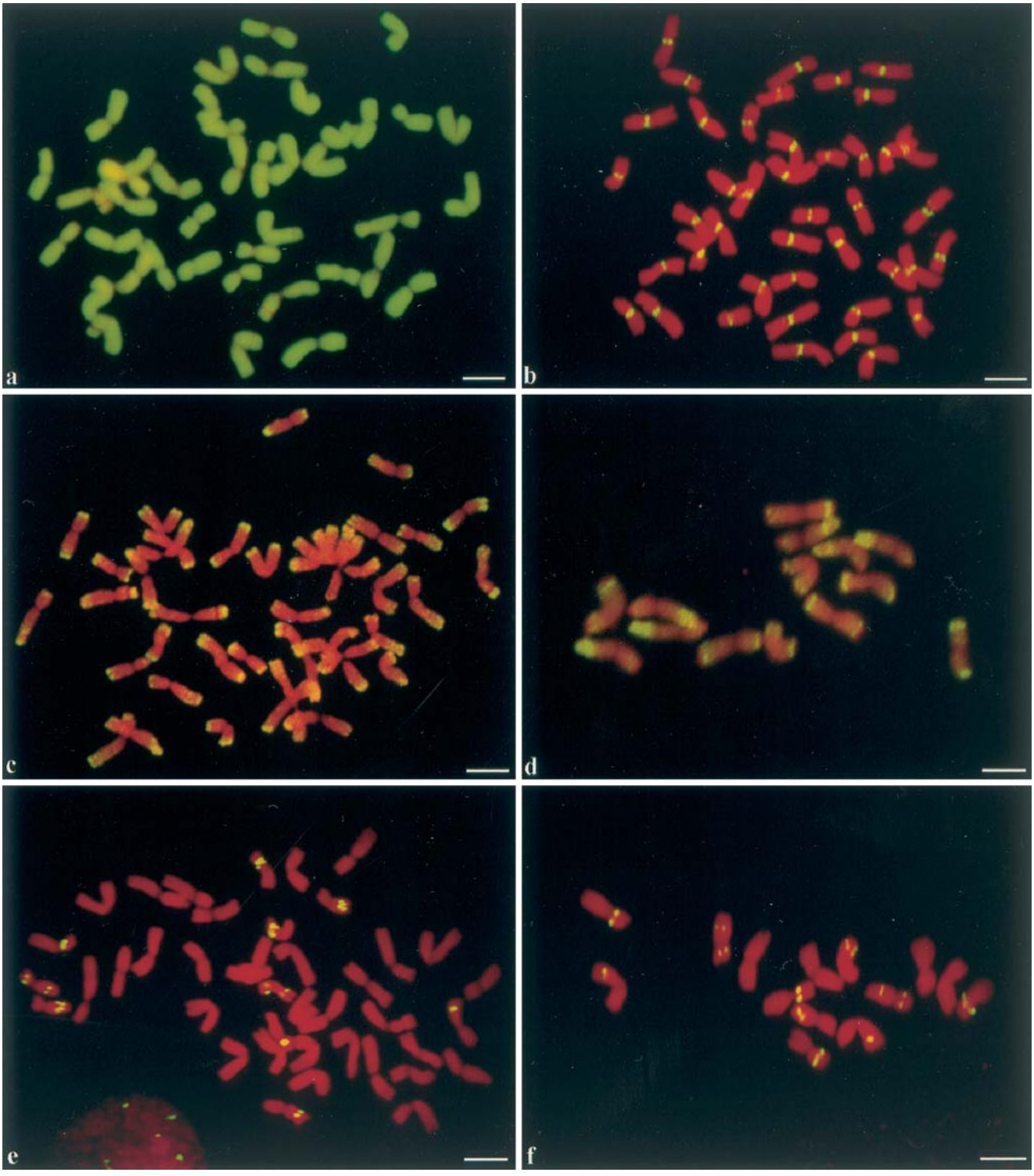

Fig. 1 Fluorescence in situ hybridization (FISH) pattern of bacterial artificial chromosome (BAC) clone 363M9 (**a**), 6C6 (**b**), A6 (**c**, **d**), and 4P6 (**e**, **f**) on mitotic metaphase chromosomes of *Triticum aestivum* cv. Chinese Spring (CS) (**a**–**c**, **e**), *Secale cereale* (**d**), and *Aegilops tauschii* (**f**). The BAC DNA was labeled with biotin-14-dATP and detected with fluorescein-avidin DN, which was visualized by yellow-green fluorescence. Chromosomes were counterstained with propidium iodide and fluoresced red. **a** BAC 363M9 hybridized to all chromosomes of CS and the hybridization signals covered the entire chromosome length. **b** FISH pattern of BAC 6C6, which hybridized strongly to the centromeric regions in *T. aestivum*. **c**, **d** FISH pattern of BAC A6, which strongly hybridized to all subtelomeric chromosome regions in wheat (**c**) and rye (**d**) with different hybridization intensities. **e**, **f** FISH pattern of BAC 4P6, which hybridized to five pairs of D-genome chromosomes in *T. aestivum* (**e**) and in *Ae. tauschii* (**f**). *Bars* represent 10 μm

Table 3 Results of contig assembly. The subclones of the individual bacterial artificial chromosome (BAC) clones that generated the same fluorescent in situ hybridization (FISH) hybridization patterns as the original BAC clones were sequenced and then assembled into contigs

BAC clones	Repeat	GenBank Accession number	Contig number	Subclones
6C6	Centromere-specific repeat	AY249982	3	6-L4
		AY249981	4	6-K11, 6-J9 (pAet6-J9)
A6	Subtelomeric repeat	AY249979	8	7-A6, 7-M5
		AY249980	10	7-P1, 7-L3 (pAet7-L3)
4P6	D-genome tandem repeat	AY249987	2	4-H18
		AY249986	9	4-P23 (pAet4-P23), 4-O2
4P6 and 9M13	D-genome tandem repeat	AY249985	14	9-P7, 4-E8, 4-J9, 4-O1, 9-K24, 9-M24, 9-K21, 9-E21, 9-P18, 9-N11
673A20	Tandem repeat specific for chromosome arm 4AL of Chinese Spring	AY249984	1	20-P7 (pTm20-P7)
		AY249983	12	20-C1, 20-D22

Table 4 DNA sequence similarity between the repetitive sequences present in bacterial artificial chromosome (BAC) clones of *Aegilops tauschii* and *Triticum monococcum* and those present in other cereals

BAC clone	Contig	Species	Repetitive sequence	Best DNA sequence identity (%)	E-value	GenBank Accession number	Reference
6C6	3	*Hordeum vulgare*	Ty*3*/ *gypsy* retrotransposon *cereba* gag-pol polyprotein	86	0	AY040832	Hudakova et al. (2001)
	4	*H. vulgare*	Ty*3*/ *gypsy* retrotransposon *cereba* gag-pol polyprotein	92	0	AY040832	Hudakova et al. (2001)
A6	8	*Aegilops tauschii*	Tandem repetitive *Afa*-family sequence	87	3×10^{-13}	AB003255	Nagaki et al. (1998)
	10	*Triticum turgidum*	CACTA transposon Caspar	89	0	AY146588	Wicker et al. (2003a, 2003b)
4P6	14	*T. monococcum*	*gypsy* -like retrotransposon Romani	87	0	AF459088	Yan et al. (2002)
673A20	1	*H. vulgare*	Nikita LTR-1	86	5×10^{-29}	AF254799	Shirasu et al. (2000)
	12	*H. vulgare*	Nikita LTR-1	87	2×10^{-32}	AF254799	Shirasu et al. (2000)

a probe showed that the copy numbers ranged from 250 to 700 in diploid *Triticum* / *Aegilops* species, from 1000 to 2000 in tetraploid species, and from 3000 to 4000 in hexaploid species.

Fourteen shotgun subclones with sizes between 750 bp and 1.7 kb were selected based on strong hybridization signals in the colony-hybridization blot using genomic DNA of *Ae. tauschii* as a probe. Among these subclones, five hybridized to subtelomeric chromosome regions as did BAC clone A6. The remaining nine clones did not give any signal in the FISH experiment. Sequencing of the positive subclones indicated that the sequence hybridizing to the subtelomeric regions has 89% sequence identity with the putative CACTA transposon Caspar (Wicker et al. 2003a, 2003b) and 93% sequence identity with the tandem repetitive *Afa* -family sequences from *Ae. tauschii*, *T. turgidum* subsp. *durum*, and *T. aestivum* subsp. *aestivum* (Nagaki et al. 1998) (Table 4). The sequences of five positive shotgun subclones were assembled into two contigs, 8 (GenBank Accession number AY249979) and 10 (GenBank Accession number AY249980; Table 3) that are 1433 bp and 1373 bp, respectively. Sequence comparison indicates that this sequence is a tandem repeat and has two copies in contig 8 and four in contig 10 (data not shown).

D-genome tandem repeat

BAC clone 4P6 from the *Ae. tauschii* library contained a localized tandem repeat and hybridized to five pairs of chromosomes in CS wheat (Fig. 1e). The hybridization sites are different from those produced by clone pAs1 (Rayburn and Gill 1986a, 1986b; Mukai et al. 1993). Clone 4P6 gave an identical FISH pattern on chromosomes of *Ae. tauschii* with hybridization sites on five chromosome pairs (Fig. 1f). These data suggest that the five labeled chromosome pairs in CS wheat belong to the D genome. This clone did not hybridize to *T. urartu*, *T. monococcum* , *T. turgidum* subsp. *durum*, or *T. timopheevii* subsp. *armeniacum*, but hybridized to three chromosome pairs in *Ae. speltoides*, one pair in *T. timopheevii* subsp. *timopheevii*, five pairs in *Ae. cylindrica*, and six pairs in *Ae. ventricosa* (data not shown).

Six subclones, ranging in size from 550 bp to 1.1 kb, were selected that gave strong hybridization signals on the colony-hybridization blot using genomic DNA of *Ae. tauschii* as a probe. All these clones gave the same hybridization patterns as BAC 4P6. In addition, seven shotgun subclones from another BAC clone, 9M13, also gave the same hybridization patterns as BAC 4P6. These positive clones were sequenced and assembled into three

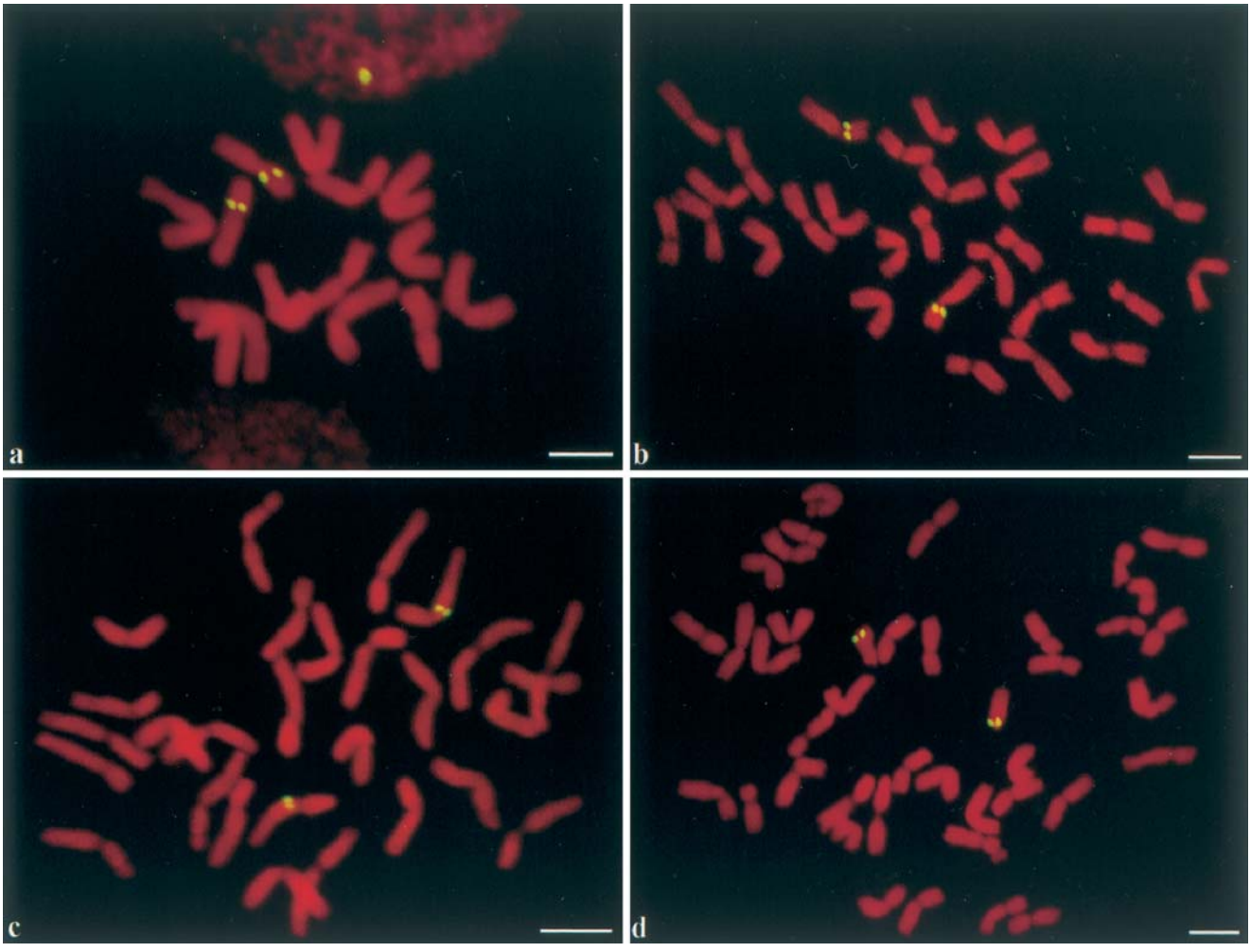

Fig. 2 Fluorescence in situ hybridization (FISH) pattern of BAC 673A20 on mitotic metaphase chromosomes of *Triticum urartu* (**a**), *T. timopheevii* (**b**), *T. turgidum* (**c**), and *T. aestivum* (**d**). The FISH procedure was the same as in Fig. 1. **a**, **b** FISH pattern of BAC 673A20 on *T. urartu* (**a**) and *T. timopheevii* (**b**); the FISH sites were in the proximal region of the short arm of chromosome 4A. **c** FISH pattern on *T. turgidum*; the FISH sites were in the proximal region of the long arm of chromosome 4A. **d** FISH pattern of BAC 673A20 on the stock ditelosomic for chromosome arm 4AL, the signals were in proximal regions of 4AL. *Bars* represent 10 μm

contigs, 2 (GenBank Accession number AY249987), 9 (GenBank Accession number AY249986), and 14 (GenBank Accession number AY249985; Table 3), which are 858 bp, 1617 bp, and 1557 bp, respectively. The sequence that is responsible for the FISH pattern has 87% sequence identity with the *gypsy* -like retrotransposon Romani from *T. monococcum* (Yan et al. 2002) (Table 4).

Dot-blot results using subclone pAet4-P23 as a probe revealed that the copy numbers ranged from 500 to 850 in some diploid *Aegilops* species including *Ae. speltoides*, *Ae. longissima*, *Ae. sharonensis*, *Ae. searsii*, *Ae. bicornis* , *Ae. caudata*, and *Ae. comosa*. About 1500 copies are present in *Ae. tauschii*, 1000 to 2500 copies are present in tetraploid *Aegilops* species, and about 7000 copies are present in *T. aestivum*. Dot-blot analysis in *T. urartu*, *T. monococcum*, *Ae. mutica*, *Ae. uniaristata*, *Ae. umbellulata*, *T. timopheevii*, *T. turgidum*, and *T. zhukovskyi* did not detect any hybridization. These species either do not contain this sequence, the amount of this sequence is too low to be detected by FISH, or homology between the sequences is too low to be detected.

Chromosome arm-specific tandem repeat

Four BAC clones (673A20, 584B8, 705E10, and 691K15, with sizes of 65.5, 87.3, 98.4, and 116.4 kb, respectively) from the *T. monococcum* library hybridized only to a proximal region in the long arm or the short arm of one pair of diploid, tetraploid, or hexaploid wheat chromosomes. BAC-FISH analysis using the ditelosomic stocks of CS mapped these BACs to a proximal region of the long arm of chromosome 4A, which corresponds to a dark C-band (Fig. 2d). All four BACs were selected using RFLP probes that hybridized to the proximal region on the short arm of group-5 chromosomes (L.L. Qi and B.S. Gill, personal communication).

These clones hybridized to a proximal region in the long arm of chromosome 4A in *T. turgidum* (Fig. 2c) and to the proximal region of the short arm of chromosome 4A in both *T. urartu* (Fig. 2a) and *T. monococcum* (data not shown), as well as in *T. timopheevii* (Fig. 2b). Two pairs of 4A chromosomes in *T. zhukovskyi* had hybridization signals in the proximal regions of their short arms (data not shown). They did not hybridize to chromosomes of either *Ae. speltoides* or *Ae. tauschii* (data not shown). Our data suggest that these BACs contain a specific repetitive sequence that is only amplified in a small region of the long arm of chromosome 4A of *T. turgidum* and *T. aestivum* and in a small region of the short arm of chromosome 4A of *T. urartu*, *T. monococcum*, *T. timopheevii*, and *T. zhukovskyi*.

BAC 673A20 was chosen for shotgun subcloning because of its size. Clone 673A20 is the smallest of the four BAC clones that gave the same hybridization pattern. Eight subclones with size ranging from 500 bp to 2.2 kb were selected based on the colony-hybridization blot using genomic DNA of *T. monococcum* as a probe. Among these clones, five did not give a hybridization signal, and three clones including pTm20-P7 gave the same hybridization pattern as that of BAC 673A20. Sequencing of the positive subclones showed that this repetitive sequence has 88% sequence identity with Nikita LTR-1 from *H. vulgare* (Shirasu et al. 2000) (Table 4). The sequences of these three positive subclones were assembled into two contigs, 1 (GenBank Accession number AY249984) and 12 (GenBank Accession number AY249983; Table 3) that are 799 bp and 2937 bp, respectively. Fourteen copies of this sequence were present in contig 12 (data not shown) indicating that this sequence is highly repeated. In addition, digestion of these four BACs with one (*Hin*dIII) and two enzymes (*Hin*dIII+*Bam*HI) and Southern hybridization analysis using the subclone pTm20-P7 as probe showed an obvious ladder pattern indicating that these four BACs contain a tandem repeat (data not shown).

Discussion

Our BAC-FISH analysis identified several novel classes of repetitive DNA elements that hybridized to specific chromosome landmarks. Similarly, BAC-FISH analysis in *Allium cepa* L. also mapped BAC clones localized to centromeric, telomeric, or several localized interstitial chromosomal regions (Suzuki et al. 2001). The BAC-FISH procedure used in the present study can effectively map repetitive DNA sequences, but was not sensitive enough to map single- or low-copy sequences in wheat as has been done in plant species with small genomes such as *Arabidopsis*, potato, cotton, rice, and sorghum. None of the BACs hybridized to the same position on a chromosome(s) as that of the corresponding RFLP fragment. In wheat, more than 80% of the genome consists of repeated DNA sequences of varying degrees of reiteration and length, a lesser proportion (12%) consists of low-copy number or unique sequences (Smith and Flavell 1975), and an even smaller proportion consists of actual genes (May and Appels 1987; Bennetzen and Freeling 1997). This high percentage of repetitive DNA accounted for the hybridization pattern generated by 40 BAC clones, which resulted in a strong labeling of all chromosomes over their entire length.

Centromere-specific repeat

We identified a sequence present in clone pAet6-J9 that has a high similarity to the Ty*3*/*gypsy* centromeric retrotransposon *cereba* gag-pol polyprotein from *H. vulgare*, including a putative RNaseH catalytic domain and a putative integrase catalytic domain (Presting et al. 1998; Hudakova et al. 2001). The presence of this centromere-specific repetitive sequence in different members of the Gramineae family together with other centromeric repeats, such as CCS1 (Aragon-Alcaide et al. 1996), pSau3A9 (Jiang et al. 1996), pRCS1 (Dong et al. 1998), and *cereba* (Presting et al. 1998), indicates that the cereal centromere may have evolved from a common progenitor prior to divergence about 60 million years ago (Kumar and Bennetzen 1999). In addition, the fact that the centromeric repeat present in clone pAet6-J9 hybridized to the centromeric regions of wheat, rye, barley, and maize but not to rice, indicates that this sequence is present with higher copy numbers in wheat, rye, barley, and maize but not in rice. Similarly, the fact that clone pRCS1 did not hybridize to barley may indicate that the centromere-specific repeat in clone pAet6-J9 is present in higher copy numbers than that in clone pRCS1 in barley.

The Ty*3*/*gypsy*-related retrotransposon DNA sequence is common to centromere-specific repeats in all cereals analyzed (Miller et al. 1998; Presting et al. 1998; Francki 2001; our study). The Ty*3*/*gypsy*-related retrotransposons in grass centromeres probably existed in the ancestral species prior to the divergence of the Triticeae species because they are highly conserved and are likely to have been amplified during centromere evolution. Centromeres are like "graveyards" for retrotransposons (Miller et al. 1998; Presting et al. 1998; Francki 2001). These retrotransposons are not intact elements and, therefore, are most likely inactive. For instance, in the centromere of *O. sativa*, the gag-pol protein sequence is probably inactive because of several stop codons in the coding region.

Three possible mechanisms were proposed for the centromere-restricted distribution pattern. First, it is possible that the retrotransposons preferentially transposed into the centromeres. Second, the retrotransposons that transposed into the cereal centromeres possibly may have been amplified by an unknown mechanism before divergence of the grass species. Third, these retrotransposons may have been eliminated from the rest of the genome at a higher rate than from centromeric regions.

Mutations and other modifications of these centromeric retrotransposon sequences may have accumulated

at a much slower rate than in retrotransposons located outside the centromeres, which resulted in the high conservation of these sequences within the centromeric regions of distantly related plant species. Another explanation is that the centromeric retrotransposon sequences might be part of the functional centromere and associated with centromere structure and/or function. Therefore, functional constraints result in high conservation. A relationship between transposable elements and centromere structure has been proposed for mammalian species (Kipling and Warburton 1997).

To date, seven centromeric repeats have been isolated from different cereal species, i.e., CCS1 from *Brachypodium sylvaticum* (Aragon-Alcaide et al. 1996), pSau3A9 from *S. bicolor* (Jiang et al. 1996), pRCS1 from *O. sativa* (Dong et al. 1998), *cereba* from *H. vulgare* (Presting et al. 1998), *Bilby* from *S. cereale* (Francki 2001), the centromere-specific repeat in clone pAet6-J9 from *Ae. tauschii* (this study), and pBs301-1 from *Ae. speltoides* (Cheng and Murata 2003). The first six repeats are Ty*3*/*gypsy*-like retrotransposons as deduced from sequence analysis. Although they are all present in centromere regions of cereals, they differ in sequence, species specificity, and copy number. The 250 bp repeats in pBs301-1 are tandemly arrayed at the centromere regions of common wheat and rye chromosomes and originated from the *cereba*-like retrotransposons in diploid wheat such as *Ae. speltoides* (Cheng and Murata 2003).

Subtelomeric repeat

We identified a repeat in BAC clone A6 that localized to all subtelomeric chromosome regions of the Triticeae species analyzed, and also oat, of the tribe Aveneae, but with much lower intensity. Barley also showed reduced signal intensity because its own subtelomeric repeat may have interfered with the hybridization of the *Ae. tauschii*-derived sequence (Belostotsky and Ananiev 1990; Schubert et al. 1998). This repeat has high sequence homology with the putative CACTA transposon *Caspar* from Triticeae (Wicker et al. 2003a, 2003b) and tandem repetitive *Afa* -family sequences from *Ae. tauschii*, *T. turgidum*, and *T. aestivum*. The repeat is present with high copy numbers in *Triticum*/*Aegilops* species, and rye, indicating that this subtelomeric tandem repeat is common within the Triticeae.

Subtelomeres are extraordinarily dynamic and variable regions near the ends of chromosomes. They are defined by their unusual structure: patchwork blocks that are duplicated near the ends of chromosomes. Subtelomeres of different species such as *Saccharomyces cerevisiae* (Pryde and Louis 1997; Pryde et al. 1997), *Drosophila melanogaster*, humans, and several plant species (Belostotsky and Ananiev 1990; The *Arabidopsis* Initiative 2000) are structurally similar in that they are composed of various repeated elements, but the sequence of the elements can vary greatly between organisms. Of all the eukaryotic organisms whose genomes have been fully sequenced, only subtelomeric chromosome regions of *Caenorhabditis elegans* lack homology to each other (Wicky et al. 1996).

The exact mechanism(s) by which subtelomeric sequences have been amplified and dispersed among many chromosome ends is unknown but likely involves several different processes. Translocation and recombination, which result in the exchange of chromosome ends; gene conversion events, resulting in the replacement of all or part of one subtelomeric region by another; or duplication by a transposition-like event are possible processes (Albini and Schwarzacher 1992; Mefford and Trask 2002).

Two models for subtelomere evolution exist (Mefford and Trask 2002). One model postulates independent block evolution, where the subtelomeric sequences of different chromosomes independently accumulate nucleotide changes and undergo no further exchange. In the second model, the subtelomere block is constantly shuffled and homogenized between different chromosome ends. The existence of shared subtelomeric repeats in all chromosomes of barley (Belostotsky and Ananiev 1990), rye (Vershinin et al. 1995; this study), and wheat (this study) supports the second model. Barley and wheat have coevolved for 12 million years (Huang et al. 2002). The presence of similar subtelomeric sequences in different Triticeae species indicates that this sequence existed in their ancestor, and the maintenance of high homology between different species implies that they are important for genome organization in the Triticeae. Triticeae species usually have a large genome size, and the amplification of this sequence may help to stabilize genome structure during the genome inflation. If subtelomeric sequences are species or genome specific then they presumably have evolved more recently (Vershinin et al. 1995). The subtelomeric repeat of barley, HvT01, hybridized to the subtelomeric regions of all 14 chromosomes in barley but not to the subtelomeric regions in wheat. On the contrary, the subtelomeric repeat present in BAC A6 hybridized to the subtelomeric regions in wheat and rye but weakly in barley, indicating that the subtelomeric repeats present in BAC A6 (such as pAet7-L3) and clone pHvT01 were amplified independently in wheat, rye and barley after barley diverged from wheat and rye.

The complex and variable nature of subtelomeres has made it difficult to assess the possible function(s) of these regions. Mutations in *S. cerevisiae* and *Plasmodium falciparum* that resulted in the deletion of subtelomeric sequences indicated that subtelomeres are not required for the viability of an organism or for proper chromosome segregation at mitosis or meiosis (Pologe and Ravetch 1988; Vergnaud 1999). However, no cases have been reported where all subtelomeric repeats were removed from an organism. Therefore, it remains possible that some subtelomeric repeats are required for viability.

Comparing the degree of sequence conservation, the telomere sequence is the most conserved, followed by centromere sequences, and subtelomeric sequences are

the least conserved among these three sequences. Subtelomeric sequences may have evolved faster than the centromere and telomere sequences.

D-genome tandem repeat

The D-genome-specific tandem repeat present in plasmid clone pAet4-P23 hybridized to five chromosome pairs in *Ae. tauschii* and also hybridized to three chromosome pairs in *Ae. speltoides*. The possible explanation for the observed hybridization pattern is that this tandem repetitive sequence was eliminated from the *Ae. speltoides* chromosomes after the polyploidization process. It is also possible that either the existent *Ae. speltoides* is not the actual B-genome donor of wheat or this repeat has undergone rapid divergence in the B-genome chromosomes. This sequence has high homology to the *gypsy*-like retrotransposon Romani from *T. monococcum* (Yan et al. 2002). This tandem repeat can be used as a cytogenetic marker for D-genome chromosomes of hexaploid wheat.

Chromosome arm-specific tandem repeat

The tandem repetitive sequence present in subclone pTm20-P7 and BACs 673A20, 584B8, 705E10, and 691K15 is 4AL chromosome arm specific. The presence of this sequence in the short arm of *T. urartu*, *T. monococcum*, and *T. timopheevii* indicates this sequence existed and was amplified at least 3 million years ago (Huang et al. 2002). The hybridization patterns of this sequence in *T. urartu*, *T. monococcum*, and *T. timopheevii* are also in agreement with the evolution of *Triticum* (Jiang and Gill 1994). The difference in the hybridization pattern compared with CS wheat is caused by a pericentric inversion that is known to be present in chromosome 4A of *T. turgidum* and *T. aestivum* and that occurred during the polyploidization process (Naranjo et al. 1987). As a result of this pericentric inversion, the original site in 4AS is relocated to 4AL in *T. turgidum* and *T. aestivum*. This sequence has a high homology with retrotransposon Nikita LTR-1 from *H. vulgare*.

A similar chromosome arm-specific repetitive sequence, WE35, was previously isolated from wheat genomic DNA and shown to hybridize to the subtelomeric regions of only one pair of wheat chromosomes (Ueng et al. 2000). WE35 has a 70–76% homology to the 5′ terminal regions of LTRs in cereal retroelements (Lucas et al. 1992; Manninen and Schulman 1993). This type of chromosome arm-specific repetitive sequence can be used as a molecular marker for genetic, phylogenetic, and evolutionary studies in the tribe Triticeae (Schmidt and Heslop-Harrison 1996).

Transposable elements provide raw material for heterochromatin

In this study, three repetitive DNA sequences (a centromere-specific repeat, a D-genome tandem repeat, and a tandem repeat specific for chromosome-arm 4AL of CS) were identified as retrotransposons. Centromeric regions of cereal species are composed of constitutive heterochromatin and centromere-specific retrotransposons contribute a large portion of the DNA in the centromeres. In addition, hybridization signals generated by either the D-genome tandem repeat or the tandem repeat specific for chromosome-arm 4AL of CS correspond to dark C-bands, which indicate the position of constitutive heterochromatin. Therefore, these retrotransposons are likely to provide raw material for the formation of constitutive heterochromatin in these regions.

Conclusions

In the present study, different classes of novel repetitive sequences were identified and mapped that can be used as cytogenetic markers for chromosome identification. The results showed that BAC-FISH is an efficient technique for identifying molecular cytogenetic markers in large-genome plants, such as wheat. Our approach allowed us rapidly to isolate repeated sequences from specific chromosome regions and study their distribution in the wheat genome.

Acknowledgements This research was supported by the Kansas Wheat Commission and a special USDA grant to the Wheat Genetics Resource Center. We also thank W. John Raupp and Duane Wilson for their excellent assistance, Drs. Li Huang, Lili Qi, and Steven Brooks for beneficial discussions, and Angie Matthews for DNA sequencing. This paper is contribution number 04-062-J from the Kansas Agricultural Experimental Station, Kansas State University, Manhattan, KS 66506-5502, USA.

References

Albini SM, Schwarzacher T (1992) In situ localization of two repetitive DNA sequences to surface-spread pachytene chromosomes of rye. Genome 35:551–559

Anamthawat-Jonsson K, Heslop-Harrison JS (1993) Isolation and characterization of genome-specific DNA sequences in Triticeae species. Mol Gen Genet 240:151–158

Appels R, Moran LB (1984) Molecular analysis of alien chromatin introduced into wheat. Stadler Genet Symp 16:529–557

Appels R, Dennis ES, Smyth DR, Peacock WJ (1981) Two repeated DNA sequences from the heterochromatic regions of rye (*Secale cereale*) chromosomes. Chromosoma 84:265–277

Aragon-Alcaide L, Miller T, Schwarzacher T, Reader S, Moore G (1996) A cereal centromeric sequence. Chromosoma 105:261–268

Bedbrook JR, Jones J, O'Dell M, Thompson R, Flavell RB (1980) A molecular description of telomeric heterochromatin in *Secale* species. Cell 19:545–560

Belostotsky DA, Ananiev EV (1990) Characterization of relic DNA from barley genome. Theor Appl Genet 80:374–380

Bennett MD, Leitch IJ (1995) Nuclear DNA amount in angiosperms. Ann Bot 76:113–176

Bennett MD, Smith JB (1976) Nuclear DNA amounts in angiosperms. Philos Trans R Soc Lond B Biol Sci 274:227–274

Bennetzen JL, Freeling M (1997) The unified grass genome: synergy in synteny. Genome Res 7:301–306

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

Cheng Z-J, Murata M (2003) A centromeric tandem repeat family originating from a part of Ty3/ *gypsy* -retroelement in wheat and its relatives. Genetics 164:665–672

Cheng Z, Presting GG, Buell CR, Wing RA, Jiang J (2001) High resolution pachytene chromosome mapping of bacterial artificial chromosomes anchored by genetic markers reveals the centromere location and the distribution of genetic recombination along chromosome 10 of rice. Genetics 157:1749–1757

Cheng Z, Buell CR, Wing RA, Jiang J (2002) Resolution of fluorescence in-situ hybridization mapping on rice mitotic prometaphase chromosomes, meiotic pachytene chromosomes and extended DNA fibers. Chromosome Res 10:379–387

Dong F, Miller T, Jackson SA, Wang GL, Ronald PC, Jiang J (1998) Rice (*Oryza sativa*) centromeric regions consist of complex DNA. Proc Natl Acad Sci U S A 95:8135–8140

Dong F, Song J, Naess SK, Helgeson JP, Gebhardt C, Jiang J (2000) Development and applications of a set of chromosome-specific cytogenetic DNA markers in potato. Theor Appl Genet 101:1001–1007

Dvorák J, Zhang HB (1992) Reconstruction of the phylogeny of the genus *Triticum* from variation in repeated nucleotide sequences. Theor Appl Genet 84:419–429

Faris JD, Haen KM, Gill BS (2000) Saturation mapping of a gene-rich recombinant hot spot region in wheat. Genetics 154:823–835

Flavell RB (1986) Repetitive DNA and chromosome evolution in plants. Philos Trans R Soc Lond B Biol Sci 312:227–242

Francki MG (2001) Identification of *Bilby*, a diverged centromeric Ty1- *copia* retrotransposon family from cereal rye (*Secale cereale* L.). Genome 44:266–274

Fransz PF, Armstrong S, de Jong JH, Parnell LD, van Drunen C, Dean C, Zabel P, Bisseling T, Jones GH (2000) Integrated cytogenetic map of chromosome arm 4S of *A. thaliana*: structural organization of heterochromatic knob and centromere region. Cell 100:367–376

Gomez MI, Islam-Faridi MN, Woo S-S, Schertz KF, Czeschin D, Zwick MS, Wing RA, Stelly DM, Price HJ (1997) FISH of a maize *sh2* -selected sorghum BAC to chromosomes of *Sorghum bicolor*. Genome 40:475–478

Hanson RE, Zwick MS, Choi S, Islam-Faridi MN, McKnight TD, Wing RA, Price HJ, Stelly DM (1995) Fluorescent in situ hybridization of a bacterial artificial chromosome. Genome 38:646–651

Huang S, Sirikhachornkit A, Su X, Faris J, Gill B, Haselkorn R, Gornicki P (2002) Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the *Triticum / Aegilops* complex and the evolutionary history of polyploid wheat. Proc Natl Acad Sci U S A 99:8133–8138

Hudakova S, Michalek W, Presting GG, ten Hoopen R, dos Santos K, Jasencakova Z, Schubert I (2001) Sequence organization of barley centromeres. Nucleic Acids Res 29:5029–5035

Hutchinson J, Lonsdale DM (1982) The chromosomal distribution of cloned highly repetitive sequences from hexaploid wheat. Heredity 48:371–376

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 1156:833–838

Jiang J, Gill BS (1994) Different species-specific chromosome translocations in *Triticum timopheevii* and *T. turgidum* support the diphyletic origin of polyploid wheats. Chromosome Res 2:59–64

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

Jiang J, Nasuda S, Dong F, Scherrer CW, Woo SS, Wing RA, Gill BS, Ward DC (1996) A conserved repetitive DNA element located in the centromeres of cereal chromosomes. Proc Natl Acad Sci U S A 93:14210–14213

Johnston PA, Pickering RA (2002) PCR detection of *Hordeum bulbosum* introgressions in an *H. vulgare* background using a retrotransposon-like sequence. Theor Appl Genet 104:720–726

Jones JDG, Flavell RB (1982) The structure, amount and chromosomal localization of defined repeated DNA sequences in species of the genus *Secale*. Chromosoma 86:613–641

Kim J-S, 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

Kipling D, Warburton PE (1997) Centromeres, CENP-B and *Tigger* too. Trends Genet 13:141–145

Kumar A, Bennetzen JL (1999) Plant retrotransposons. Annu Rev Genet 33:479–532

Li WL, Gill BS (2002) The colinearity of the *Sh2/A1* orthologous region in rice, sorghum and maize is interrupted and accompanied by genome expansion in the Triticeae. Genetics 160:1153–1162

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

Lucas H, Moore G, Murphy G, Flavell RB (1992) Inverted repeats in the long-terminal repeats of the wheat retrotransposon Wis 2–1A. Mol Biol Evol 9:716–728

Manninen I, Schulman AH (1993) *BARE-1*, a *copia* -like retroelement in barley (*Hordeum vulgare* L.). Plant Mol Biol 22:829–846

May CE, Appels R (1987) The molecular genetics of wheat: toward an understanding of 16 billion base pairs of DNA. In: Heyne EG (ed) Wheat and wheat improvement, vol 13, 2nd edn. American Society of Agronomy, Madison, Wis., pp 165–198

McIntyre CL, Pereira S, Moran LB, Appels R (1990) New *Secale cereale* (rye) DNA derivatives for the detection of rye chromosome segments in wheat. Genome 33:635–640

Mefford HC, Trask BJ (2002) The complex structure and dynamic evolution of human subtelomeres. Nat Rev Genet 3:91–102

Miller JT, Dong F, Jackson SA, Song J, Jiang J (1998) Retrotransposon-related DNA sequences in the centromeres of grass chromosomes. Genetics 150:1615–1623

Moullet O, Lagudah ES (1998) Construction of a wheat D genome bacterial artificial chromosome library. In: Slinkard AE (ed) Proceedings of the 9th International Wheat Genetics Symposium. University Extension Press, Saskatoon, Canada, pp 31–32

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

Mukai Y, Friebe B, Gill BS (1992) Comparison of C-banding patterns and in situ hybridization sites using highly repetitive and total genomic rye DNA probes of 'Imperial' rye chromosomes added to 'Chinese Spring' wheat. Jpn J Genet 67:71–83

Mukai Y, Nakahara Y, Yamamoto M (1993) Simultaneous discrimination of the three genomes in hexaploid wheat by multicolor fluorescence in situ hybridization using total genomic and highly repeated DNA probes. Genome 36:489–494

Nagaki K, Tsujimoto H, Sasakuma T (1998) Dynamics of tandem repetitive *Afa*-family sequence in Triticeae, wheat-related species. J Mol Evol 47:183–189

Naranjo T, Roca A, Goicoechea PG, Giraldez R (1987) Arm homoeology of wheat and rye chromosomes. Genome 29:873–882

Pedersen C, Langridge P (1997) Identification of the entire chromosome complement of bread wheat by two-color FISH. Genome 40:589–593

Pedersen C, Rasmussen SK, Linde-Laursen I (1996) Genome and chromosome identification in cultivated barley and related species of the Triticeae (Poaceae) by in situ hybridization with the GAA-satellite sequence. Genome 39:93–104

Pologe LG, Ravetch JV (1988) Large deletions result from breakage and healing of *P. falciparum* chromosomes. Cell 55:869–874

Presting GG, Malysheva L, Fuchs J, Schubert I (1998) A Ty3/*gypsy* retrotransposon-like sequence localizes to the centromeric regions of cereal chromosomes. Plant J 16:721–728

Pryde FE, Louis EJ (1997) *Saccharomyces cerevisiae* telomeres. A review. Biochemistry (Mosk) 62:1232–1241

Pryde FE, Gorham HC, Louis EJ (1997) Chromosome ends: all the same under their caps. Curr Opin Genet Dev 7:822–828

Rayburn AL, Gill BS (1986a) Isolation of a D-genome specific repeated DNA sequence from *Aegilops squarrosa*. Plant Mol Biol Rep 4:102–109

Rayburn AL, Gill BS (1986b) Molecular identification of the D-genome chromosomes of wheat. J Hered 77:253–255

Schmidt T, Heslop-Harrison JS (1996) High resolution mapping of repetitive DNA by in situ hybridization: molecular and chromosomal features of prominent dispersed and discretely localized DNA families from the wild beet species *Beta procumbens*. Plant Mol Biol 30:1099–1114

Schubert I, Shi F, Fuchs J, Endo TR (1998) An efficient screening for terminal deletions and translocations of barley chromosomes added to common wheat. Plant J 14:489–495

Shirasu K, Schulman AH, Lahaye T, Schulze-Lefert P (2000) A contiguous 66-kb barley DNA sequence provides evidence for reversible genome expansion. Genome Res 10:908–915

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

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

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

Suzuki G, Ura A, Saito N, Do GS, Seo BB, Yamamoto M, Mukai Y (2001) BAC FISH analysis in *Allium cepa*. Genes Genet Syst 76:251–256

The *Arabidopsis* Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408:796–815

Ueng PP, Hang A, Tsang H, Vega JM, Wang L, Burton CS, He FT, Liu B (2000) Molecular analyses of a repetitive DNA sequence in wheat (*Triticum aestivum* L.). Genome 43:556–563

van Slageren MW (1994) Wild wheats: a monograph of *Aegilops* L. and *Amblyopurum* (Jaub. & Spach) Eig (Poaceae). Wageningen Agricultural University Papers, Wageningen, The Netherlands

Vergnaud G (1999) Structure and evolution of human subtelomeric regions. J Soc Biol 193:35–40 (In French)

Vershinin A, Svitashev S, Gummesson PO, Salomon B, von Bothmer R, Bryngelsson T (1994) Characterization of a family of tandemly repeated DNA sequences in Triticeae. Theor Appl Genet 89:217–225

Vershinin AV, Schwarzacher T, Heslop-Harrison JS (1995) The large-scale genomic organization of repetitive DNA families at the telomeres of rye chromosomes. Plant Cell 7:1823–1833

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

Wicker T, Guyot R, Yahiaoui N, Keller B (2003b) CACTA transposons in Triticeae. A diverse family of high-copy repetitive elements. Plant Physiol 132:52–63

Wicky C, Villeneuve AM, Lauper N, Codourey L, Tobler H, Müller F (1996) Telomeric repeats $(TTAGGC)_n$ are sufficient for chromosome capping function in *Caenorhabditis elegans*. Proc Natl Acad Sci U S A 93:8983–8988

Yan L, Echenique V, Busso C, SanMiguel P, Ramakrishna W, Bennetzen JL, Harrington S, Dubcovsky J (2002) Cereal genes similar to Snf2 define a new subfamily that includes human and mouse genes. Mol Genet Genomics 268:488–499

Zhang P (2002) Analysis of the wheat genome by BAC-FISH. PhD thesis, Department of Plant Pathology, Kansas State University, Manhattan, USA

Zhang P, Friebe B, Lukaszewski AJ, Gill BS (2001) The centromere structure in Robertsonian wheat-rye translocation chromosomes indicates that centric breakage-fusion can occur at different positions within the primary constriction. Chromosoma 110:335–344