Theor Appl Genet (2002) 104:286–293

Q.J. Song · E.W. Fickus · P.B. Cregan

Characterization of trinucleotide SSR motifs in wheat

Received: 17 February 2001 / Accepted: 31 May 2001

Abstract Length differences among trinucleotide-based microsatellite alleles can be more easily detected and frequently produce fewer "stutter bands" as compared to dinucleotide-based microsatellite markers. Our objective was to determine which trinucleotide motif(s) would be the most-polymorphic and abundant source of trinucleotide microsatellite markers in wheat (*Triticum aestivum* L.). Four genomic libraries of cultivar 'Chinese Spring' were screened with nine trinucleotide probes. Based on the screening of 28550 clones, the occurrences of $(CTT/GAA)_n$, $(GGA/CCT)_n$, $(TAA/ATT)_n$, $(CAA/GTT)_n$, $(GGT/CCA)_n$, $(CAT/GTA)_n$, $(CGA/GCT)_n$, $(CTA/GAT)_n$, and $(CGT/GCA)_n$ repeats were estimated to be 5.4×10^4, 3.5×10^4, 3.2×10^4, 1.2×10^4, 6.3×10^3, 4.9×10^3, 4.5×10^3, 4.5×10^3 and 3.6×10^3, i.e., once every 293 kbp, 456 kbp, 500 kbp, 1.3 Mbp, 2.6 Mbp, 3.2 Mbp, 3.6 Mbp, 3.6 Mbp and 4.5 Mbp in the wheat genome, respectively. Of 236 clones selected for sequencing, 38 (93%) $(TAA/ATT)_n$, 30 (43%) $(CTT/GAA)_n$, 16 (59%) $(CAA/GTT)_n$, 3 (27%) $(CAT/GTA)_n$ and 2 (4%) $(GGA/CCT)_n$ clones contained microsatellites with eight or more perfect repeats. From these data, 29, 27 and 16 PCR primer sets were designed and tested to the $(TAA/ATT)_n$, $(CTT/GAA)_n$ and $(CAA/GTT)_n$ microsatellites, respectively. A total of 12 (41.4%) primers designed to $(TAA/ATT)_n$, four (14.8%) to $(CTT/GAA)_n$, and two (12.5%) to $(CAA/GTT)_n$ resulted in polymorphic markers. The results indicated that $(TAA/ATT)_n$ microsatellites would provide the most-abundant and the most-polymorphic source of trinucleotide microsatellite markers in wheat.

Communicated by A.L. Kahler

Q.J. Song
Building 006, Soybean Genomics and Improvement Laboratory, USDA, ARS, Beltsville Agriculture Research Center, Beltsville, MD 20705, USA
(A visiting scientist from Nanjing Agricultural University, Nanjing, 210095, P. R. China)

P.B. Cregan (✉), E.W. Fickus
Building 006, Soybean Genomics and Improvement Laboratory, USDA, ARS, Beltsville Agriculture Research Center, Beltsville, MD 20705, USA
e-mail: creganp@ba.ars.usda.gov

Keywords Wheat · Simple sequence repeat (SSR) · Microsatellites · Trinucleotide motifs · Molecular markers

Introduction

The level of molecular genetic polymorphism in hexaploid wheat is low as compared to many other species. As a result, the use of RFLP and RAPD DNA markers in adapted wheat germplasm for applications such as quantitative trait locus discovery and marker assisted selection is limited (Chao et al. 1989; Sun et al. 1998). In contrast, microsatellite DNA markers generally possess higher levels of polymorphism in most species, such as soybean (Akkaya et al. 1992; Morgante and Olivieri 1993), rice (Wu and Tanksley 1993; Panaud et al. 1996), barley (Liu et al. 1996) and tomato (Broun and Tanksley 1996), and thus are a good alternative marker system in wheat (Röder et al. 1998b). This PCR-based molecular marker is easily detected on polyacrylamide or high-resolution agarose gels and distribution of primer sequences is all that is necessary to transfer markers from laboratory to laboratory. As a consequence, there has been considerable effort devoted to developing microsatellite markers and their applications to the construction of linkage maps (Becker and Heun 1995; Akagi et al. 1996; Senior et al. 1996; Cregan et al. 1999), the study of allelic profiles of genotypes (Rongwen et al. 1995; Plaschke et al. 1995; Diwan and Cregan 1997; Fahima et al. 1998; Song et al. 1998a; Worland et al. 1998), the identification of cultivars (Bligh et al. 1999; Song et al. 1999) and the tracing of genes with agronomic importance (Ayres et al. 1997; Jin et al. 1998; Mian et al. 1999).

The efficiency of microsatellite-marker development depends on the abundance of repeats in the target species and the ease with which these repeats can be developed into informative markers. Efforts to investigate the abundance of microsatellites have been based upon database searches or screening genomic libraries. In comparison to database searches, which are largely based on gene sequences, genomic library screening offers a more com-

prehensive evaluation of the complete genome and therefore provides a more accurate estimation of SSR frequency. Based on genomic library screening, it was observed that, in plants, $(AG/TC)_n$ is more frequent than $(AC/TG)_n$ (Lagercrantz et al. 1993), whereas in humans $(AC/TG)_n$ repeats are the most common (Weber and May 1989).

In recent years, efforts have been made to develop markers based on $(AG/TC)_n$ and $(AC/TG)_n$ repeats in wheat and about 315 of these markers are now available (Korzun et al. 1997; Röder et al. 1998a; Pestsova et al. 2000a, b, c,). In contrast to $(AG/TC)_n$ repeats, there is little information on trinucleotide microsatellites, and few trinucleotide microsatellite markers have been developed in wheat. Likewise, the frequency of trinucleotide microsatellites has not been reported in wheat.

Trinucleotide repeats have been demonstrated to be highly polymorphic and stably inherited in the human genome (Edwards et al. 1991; Gastier et al. 1995; Sheffield et al. 1995). Trinucleotide-based microsatellites are generally more robust than dinucleotide-based microsatellite markers because they frequently give fewer "stutter bands" (Hearne et al. 1992; Diwan and Cregan 1997). Thus, allele sizing is less error prone using tri- rather than di-nucleotide repeats (Diwan and Cregan 1997). The ease of allele sizing using tri- vs di-nucleotide SSR markers and the ability to readily distinguish between adjacent alleles will permit the use of microsatellite markers using a range of analysis systems, including agarose gels. Because of our previous success in developing a large set of mostly trinucleotide SSR markers in soybean (Cregan et al. 1999), we decided to investigate the same approach in wheat.

Before embarking upon a program of trinucleotide-based microsatellite development in wheat, it was necessary to assess the relative frequencies and repeat lengths of these microsatellites, and to determine which motif or motifs could be most readily converted to useful markers. Therefore, our objectives were to determine the abundance of nine different trinucleotide microsatellites in the wheat genome, the repeat length distributions of each and the rates at which they could be developed into informative markers.

Materials and methods

Plant materials

Genomic DNA of 'Chinese Spring', PI 428203 (*Triticum urartu*, A genome); PI 560527 (*Triticum speltoides*, B genome); PI 220331 (*Aegilops tauschii*, D genome); 'Langdon' (*Triticum durum*, A, B genome); 'Cheyenne'; 'Canadian Red'; 'Cascade'; 'Dicklow'; 'Bunyip'; 'Catoctin'; 'Pioneer 2548'; 'Opata 85' and M6 were isolated with either a DNeasy DNA Plant Kit (QIAGEN Inc.) or a QIAGEN Genomic DNA Purification Kit (QIAGEN Inc.) after DNA extraction with the CTAB method described by Keim et al. (1988).

Wheat library construction, screening and DNA sequencing

Genomic DNA of Chinese Spring was digested with the enzymes *Sma*I, *BsrB*I, *NIa*IV, *Cac8*I and *ECI136*II, and then size-selected on a 1% agarose gel. DNA fragments in the 400- to 750-bp range were isolated from the gel using Geneclean II (Bio101 Inc., La Jolla, Calif.). Purified DNA fragments were ligated into the *SmaI* site of pBluescript (Stratagene, La Jolla, Calif.). Ligations were transformed into the XL2-blue strain of *Eschericha coli*. Libraries were plated at a density of 1500–2500 clones per Petri plate. Due to the washing temperature differences of different probes, libraries were separately screened with the five "AT-rich" trinucleotide probes, $(TAA/ATT)_{10}$, $(CAA/GTT)_{10}$, $(CTT/GAA)_{10}$, $(CAT/GTA)_{10}$, and $(CTA/GAT)_{10}$, and the four "GC-rich" trinucleotide probes, $(CGT/GCA)_{10}$, $(GGA/CCT)_{10}$, $(GGT/CCA)_{10}$ and $(CGA/GCT)_{10}$. To provide replication, two libraries were created and screened with "AT-rich" probes and two with the "GC-rich" probes. Probes were labelled with ^{32}P, hybridized to the membranes from colony lifts as described by Cregan et al. (1994). Membranes were washed in 1× SSC and 0.1% SDS for 45 min at 46°C for the "AT-rich" probes, and 50°C for the "GC-rich" probes. Two more cycles of clone selection and re-screening followed, as described by Cregan et al. (1994). After screening, putative SSR-containing clones were either cultured in the LB liquid medium, then lysed and purified by the QIAwell Plasmid Kit (QIAGEN Inc.), or heated at 100°C for 10 min before amplifying with T3 and T7 primers and purified with the QIAquick PCR Purification Kit (QIAGEN Inc.). Clones were subsequently sequenced on an ABI PRISM 377 DNA Sequencer following labelling reactions, using a dRhodamine Terminator Cycle Sequencing Kit (The Perkin-Elmer Corporation).

Primer design and evaluation

Since the informativness of microsatellites tends to increase with an increasing number of repeats (Weber and May 1989; Wu and Tanksley 1993), primers were designed only for those sequences with repeat units equal to, or greater than, eight, using the software OLIGO 5.0 (National Biosciences, Inc. Plymouth, Minn.). Primers were synthesized by BioServe Biotechnologies (Laurel, Md.). Each primer set was first examined in amplification reactions using the plasmid from which the sequence was derived and Chinese Spring genomic DNA as templates. The reaction mix contained 100 ng of template DNA, 1.5 of mM Mg^{+2}, 1.5 of mM of dNTPs, 1.5 µM of each primer, 1 µl of 10× buffer, 1.25 units of *Taq* DNA polymerase and 0.1 µl of 3000 Ci/mmol α-^{32}PdATP in a total volume of 10 µl. PCR cycling conditions were as follows: 3 min at 95°C, followed by 30 cycles of 40 s denaturation at 94°C, 40 s annealing at the optimized annealing temperature, 1 min extension at 72°C. After cycling, the reactions were incubated at 72°C for 10 min. The ^{32}P-labelled PCR products were analysed on a 6% DNA sequencing gel with 30% formamide followed by autoradiography. To determine the genome of origin and the level of polymorphism, those primer sets that produced a product of expected size in Chinese Spring were further tested using the genomic DNA of PI 428203 (*T. urartu*, A genome), PI 560527 (*T. speltoides*, B genome), PI220331 (*A. tauschii*, D genome), Langdon (*T. durum*, A, B genome), Cheyenne, Canadian Red, Cascade, Dicklow, Bunyip, Catoctin, Pioneer 2548, Opata 85 and M6. Opata 85 and M6 are the parents of the ITMI (International Triticae Mapping Initiative) mapping population.

Statistical analysis

Gene diversity was calculated for each primer set as: $1-\sum P^2_{ij}$, where P_{ij} is the frequency of the j^{th} allele for the i^{th} locus. The data were processed using the SSRFING program developed by Song et al. (1998b).

Results

Abundance of various trinucleotide motifs in the wheat genome

Replicate libraries containing approximately 143000 and 142000 clones were screened with the "AT-rich" and "GC-rich" probes, respectively. A total of 1500 clones were initially observed to contain "AT-rich" motifs, and 388 with "GC-rich" motifs. Thus, an estimated 0.7% of the clones contained a trinucleotide microsatellite. The inserts of 236 positive clones were sequenced, of which 70 contained $(CTT/GAA)_n$, 45 $(GGA/CCT)_n$, 41 $(TAA/ATT)_n$, 27 $(CAA/GTT)_n$, 14 $(GGT/CCA)_n$, 11 $(CAT/GTA)_n$, 10 $(CGA/GCT)_n$, 10 $(CTA/GAT)_n$, and eight contained $(CGT/GCA)_n$ repeats (Table 1). Four trinucleotide repeats, $(CTT/GAA)_n$, $(GGA/CCT)_n$, $(TAA/ATT)_n$ and $(CAA/GTT)_n$ predominated making up 78% of the trinucleotide microsatellite-containing clones.

Table 1 Number and percentage of clones containing perfect trinucleotide repeats of eight or more, determined via screening of genomic libraries using nine different trinuclotide repeat motifs followed by sequence analysis of selected clones

Repeat motif	Clones with eight or more repeats (no.)	Clones with eight or more repeats (%)
$(TAA/ATT)_n$	38	93
$(CAA/GTT)_n$	16	59
$(CTT/GAA)_n$	30	43
$(CAT/GTA)_n$	3	27
$(GGA/CCT)_n$	2	4

The occurrences of the remaining five repeats were much lower, each with a proportion of less than 6% of the total.

Although the frequencies of $(CTT/GAA)_n$ and $(GGA/CCT)_n$ were higher than that of $(TAA/ATT)_n$, most of these repeats were very short (2–7 repeat units). Since primers were designed only for clones with eight or more repeat units (Table 1), the proportions of clones falling into this category were only 59%, 43%, 27% and 4% for $(CTT/GAA)_n$, $(CAA/GTT)_n$, $(CAT/GTA)_n$ and $(GGA/CCT)_{n,}$ respectively, but was 93% for clones with the $(TAA/ATT)_n$ motif.

Based upon the sequence analysis, average insert size was 575 bp; the occurrences in the wheat genome of the nine microsatellites were estimated to range from 3.6×10^3 for $(CGT/GCA)_n$, the least frequent, to 5.4×10^4 for $(CTT/GAA)_n$, the most frequent microsatellite, i.e., once every 4.5 Mbp to once every 293 kbp (Table 2). The occurrences in the wheat genome of the $(TAA/ATT)_n$, $(CTT/GAA)_n$, $(CAA/GTT)_n$, $(CAT/GTA)_n$ and $(GGA/CCT)_n$ repeats with eight or more units were estimated to be 3.0×10^4, 2.3×10^4, 1.2×10^4, 2.3×10^3 and 1.5×10^3, respectively. Thus, $(TAA/ATT)_n$ is the most-abundant microsatellite containing the desired number of repeats.

Conversion rates of different motifs to polymorphic markers

Primers were selected to the SSR flanking regions of 74 of the 89 clones with eight or more perfect repeats.

Table 2 Characteristics and estimated frequency of occurrence of microsatellites in the wheat genome, as determined from analysis of Chinese Spring genomic libraries

Repeat motifs	Range of repeat length										Estimated occurrence[a] in the wheat genome (no.)
	2–4	5–7	8–11	12–15	16–19	20–23	24–27	28–31	>32	Total	
$(TAA/ATT)_n$	1	2	10	15	4	3	4	1	1	41	3.2×10^4
$(CAA/GTT)_n$	5	7	9	2	1	2	0	0	1	27	1.2×10^4
$(CTT/GAA)_n$	18	24	9	9	5	4	0	1	2	70	5.4×10^4
$(CAT/GTA)_n$	5	3	1	1	0	0	1	0	0	11	4.9×10^3
$(CTA/GAT)_n$	2	8	0	0	0	0	0	0	0	10	4.5×10^3
$(CGT/GCA)_n$	7	1	0	0	0	0	0	0	0	8	3.6×10^3
$(GGA/CCT)_n$	38	5	2	0	0	0	0	0	0	45	3.5×10^4
$(GGT/CCA)_n$	12	2	0	0	0	0	0	0	0	14	6.3×10^3
$(CGA/GCT)_n$	10	0	0	0	0	0	0	0	0	10	4.5×10^3

[a] Occurrence was estimated based upon number of positive clones on the membranes and the frequency of occurrence of the predicted microsatellite based upon the DNA seqeunce analysis of clones

Table 3 Number of primer sets that amplified single, multiple or no products with Chinese Spring genomic DNA as a template

Motif	Primers sets designed (no.)	Amplified one or two discrete products (no.)	Amplified multiple products (no.)	No products (no.)
$(TAA/ATT)_n$	29	19	8	2
$(CAA/GTT)_n$	15	11	3	1
$(CTT/GAA)_n$	27	13	9	5
$(CAT/GTA)_n$	3	2	1	0

Table 4 Primer sequences, repeat motifs, repeat lengths in Chinese Spring, and PIC values of loci that were polymorphic among a set of 14 genotypes

Markers	Primer Sequences (5′→3′)	Repeat types	Size in CS (bp)	Tm (°C)	Gene diversity (score)	Alleles (no.)
BARC001	GCG ATG CTT TTG CCT TGT TTC AG GCG GCC CCT TTG ACT CTT CAT AG	$(TAA)_8$	271	52	0.30	2
BARC002	GCG AGA TGC ACA TCA GCA CCT AAT C GCG CCA TAT TGG AAT CTT TCT TTC T	$(TTA)_{5+20}$	287	52	0.69	4
BARC004	GCG TGT TTG TGT CTG CGT TCT A CAC CAC ACA TGC CAC CTT CTT T	$(TTA)_{15}$	158	52	0.83	7
BARC005	GCGCCTGGACCGGTTTTCTATTTT GCGTTGGGAATTCCTGAACATTTT	$(TTA)_{5+8}$	278	52	0.46	3
BARC006	TTCGGTCGTTGAGGTGACCAATTATG GACAAAGGATTAGCCCAAAGTAAGAG	$(TTA)_{24}$	471	52	0.77	4
BARC008	GCG GGA ATC ATG CAT AGG AAA ACA GAA AA GCG GGG GCG AAA CAT ACA CAT AAA AAC A	$(TTA)_{15+11}$	245	50	0.72	5
BARC009	GCG GTC GGT GTC TCC AGT TTT TTT ATC A GCG ACA TGC GGA CAG TAT TTA AAT TTC	$(TTA)_6+(TTG)_6$	201	52	0.46	2
BARC011	GCG ATG CGT GTA AAG TCT GAA GAT GA GCG TCC ATG GAG CTC TGT TTT ATC TGA	$(TAA)_{9+12}$	299	50	0.75	4
BARC010	CGA CAG AGT GAT CAC CCA AAT ATA A CAT CGG TCT AAT TGT CAA TGT A	$(TAA)_{28}$	196	50	0.83	7
BARC014	GCG TTG TGG AAA CTC AGT TTT GTT GAT TTA GCG GAA AGG AAC GAA GTA CAT TTT GTA GA	$(TAA)_{18+12}$	270	50	0.80	6
BARC017	GCG CAA CAT ATT CAG CTC AAC A TCC ACA TCT CGT CCC TCA TAG TTT G	$(TAA)_{12}$	272	50	0.65	4
BARC018	CGC TTC CCA TAA CGC CGA TAG TAA CGC CCG CAT CAT GAG CAA TTC TAT CC	$(TAA)_{20}$	228	52	0.84	7
BARC003	TTCCCTGTGTCTTTCTAATTTTTTTT GCGAACTCCCGAACATTTTTAT	$(TTC)_{17}$	188	52	0.65	4
BARC007	GCG AAG TAC CAC AAA TTT GAA GGA A CGC CAT CTT ACC CTA TTT GAT AAC TA	$(TTC)_{6+3}$	267	50	0.46	2
BARC010	GCG TGC CAC TGT AAC CTT TAG AAG A GCG AGT TGG AAT TAT TTG AAT TAA ACA AG	$(TTC)_{13}$	283	52	0.32	3
BARC013	GCA GGA ACA ACC ACG CCA TCT TAC GCG TCG CAA TTT GAA GAA AAT CAT C	$(TTC)_{5+3+2}$	142	50	0.46	2
BARC015	ATG CAA AGG CCG GGG TTA TC CAC CTC TAG CCT ACG CCA ACA TT	$(CAA)_9$	102	50	0.60	3
BARC016	GCG TTG CGA GAT CTT ATG GGT TT GCG CGT TTT TCG AAT ACC TTG T	$(CAA)_9$	112	52	0.14	2

Among these, 29, 15, 27 and 3 primer sets flanked $(TAA/ATT)_n$, $(CAA/GTT)_n$, $(CTT/GAA)_n$ and $(CAT/GTA)_n$ microsatellites, respectively (Table 3). In 15 instances the microsatellite was too close to the cloning site and therefore primers could not be selected. Of the 74 primer sets tested, 45 produced one or two discrete products with the genomic DNA of Chinese Spring as a template (Table 3). About 66%, 73%, 48% and 67% of primers to $(TAA/ATT)_n$, $(CAA/GTT)_n$, $(CTT/GAA)_n$ and $(CAT/GTA)_n$ microsatellites amplified one or two discrete PCR products, respectively.

Of the 45 primer sets that amplified one or two discrete products, 18 amplified polymorphic products i.e., products that varied in size among the ten hexaploid and one tetraploid wheat genotypes. In the case of primer sets that amplified two products, polymorphism was defined as variation among the higher-molecular weight or lower-molecular-weight products or both. The number of alleles at each of these loci, and the gene diversity values based upon the 11 wheat genotypes, are given in Table 4. Among these 18 loci, 12 were based on $(TAA/ATT)_n$ repeats, while only four and two were based upon $(CTT/GAA)_n$ and $(CAA/TTG)_n$ microsatellites, respectively. The percentage of primer sets tested that resulted in a polymorphic microsatellite marker is given in Table 5. A much higher proportion (41%) of the primer sets synthesized to $(TAA/ATT)_n$ flanking regions resulted in polymorphic markers.

Markers based on trinucleotide microsatellites were highly polymorphic in the 11 wheat genotypes tested. There was an average of four alleles at the 18 loci among the 11 genotypes. The mean gene diversity values of the 18 loci was 0.60. High polymorphism of $(TAA/ATT)_n$ microsatellite markers was observed and the average

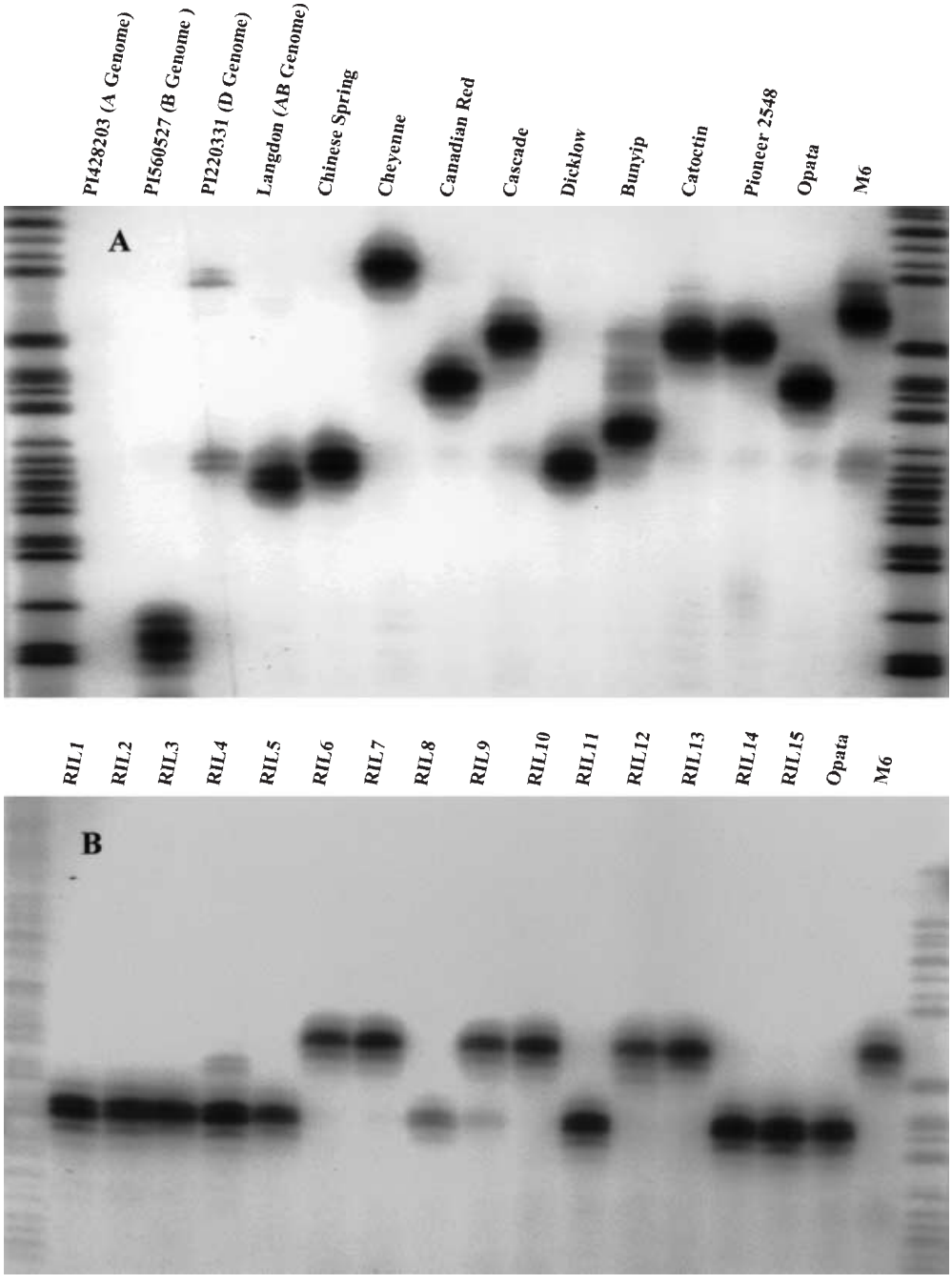

Fig. 1 **Panel A** Pattern of polymorphism detected by microsatellite BARC004 in a set of wheat accessions. **Panel B** Genotypes of 15 recombinant inbred lines (RILs) from the ITMI mapping population and ITMI parents Opata 85 and M6 at the BARC004 locus

Table 5 Percentage of primer sets producing polymorphic markers

Core motif	Primer sets synthesized and tested (no.)	Polymorphic loci developed (no.)	Primer sets producing polymorphic markers (%)
$(TAA/ATT)_n$	29	12	41
$(CAA/GTT)_n$	15	2	13
$(CTT/GAA)_n$	27	4	15
$(CAT/GTA)_n$	3	0	0

gene diversity was estimated to be 0.68 for $(TAA/ATT)_n$, compared to 0.48 and 0.37 for $(CTT/GAA)_n$ and $(CAA/TTG)_n$ microsatellite markers, respectively. Based upon the analysis of the A-, B- and D-genome donors, six of the 18 loci were assigned to the A genome, four to the B genome, three to the D genome, one to both the A and B genomes, two to both the B and D genomes, two to the A, B and D genomes. Primers that amplified prod-

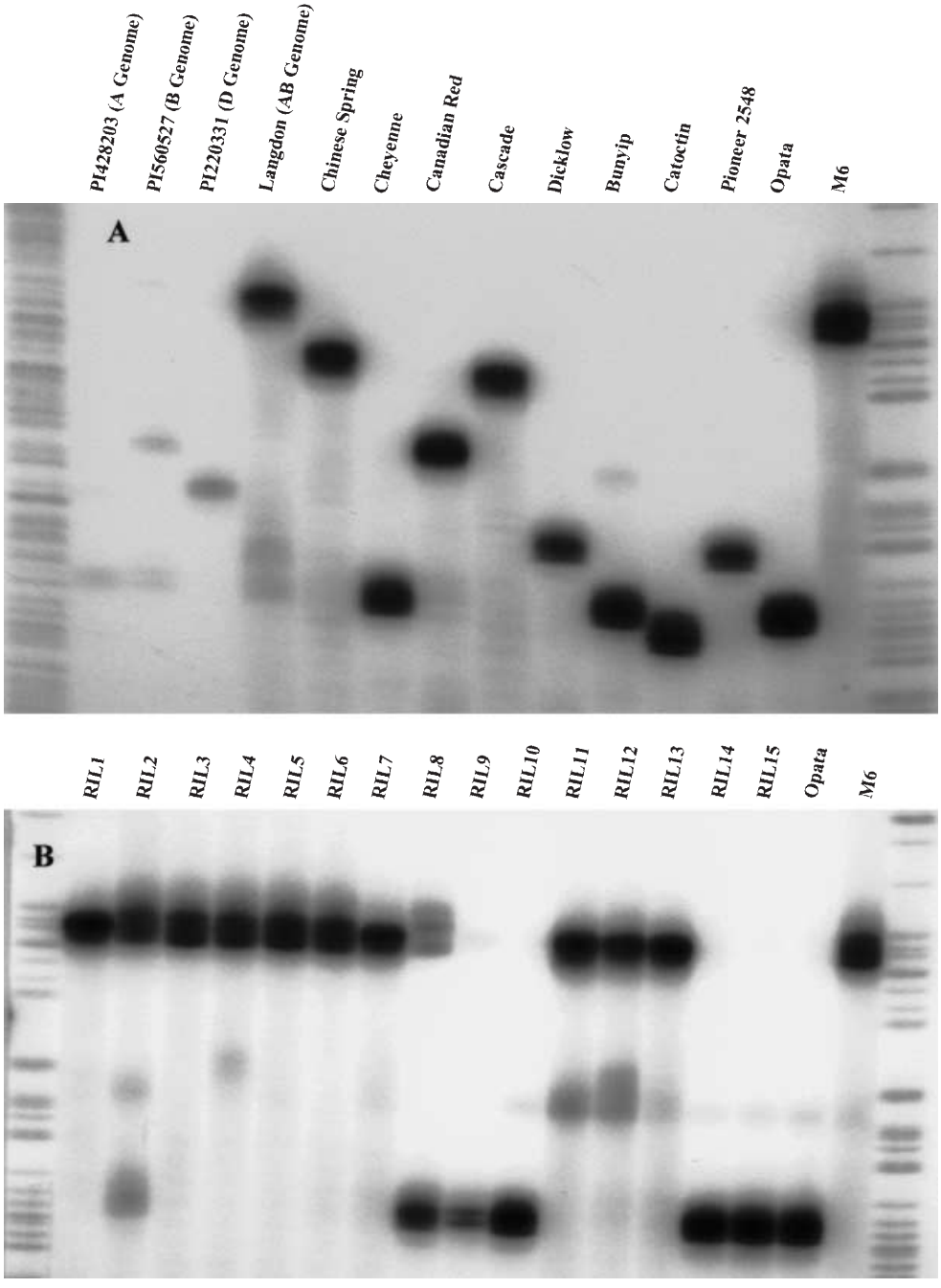

Fig. 2 **Panel A** Pattern of polymorphism detected by microsatellite BARC012 in a set of wheat accessions. **Panel B** Genotypes of 15 recombinant inbred lines (RILs) from the ITMI mapping population and ITMI parents Opata 85 and M6 at the BARC012 locus

ucts from more than one genome in the donor genomes, however, usually amplified what appeared to be one product in the hexaploid wheat. Figures 1 and 2 provide an illustration of the PCR products produced by two of the 18 loci using the 11 wheat genotypes as templates. In addition, each of these two loci were used to genotype 15 recombinant inbred lines from the ITMI mapping population derived from the cross of Opata 85 (hexaploid wheat)×M6 (synthetic hexaploid). Figures 1 and 2 demonstrate the discrete products with few stutter bands and the minimal amount of extraneous signal that were typical of most of the trinucleotide loci developed.

Discussion

The frequency of $(TAA/ATT)_n$ in the library

Our results indicated that trinucleotide repeat motifs were present in 0.7% of the clones screened via hybridization with oligo probes. If only the repeats with lengths equal to eight or more were considered, this value was estimated to be 0.26% of the clones. The occurrence of $(TAA/ATT)_n$, $(CAA/GTT)_n$, $(CTT/GAA)_n$, $(CAT/GTA)_n$ and $(GGA/CCT)_n$ was once every 235 kbp, or a total of 6.8×10^4 per haploid the genome. As a basis of comparison, we searched genomic DNA sequences of *Arabidopsis thalinana* available in

GeneBank and EMBL. In total, 436 trinucleotide repeats of eight or more repeat units were found. $(TAA/ATT)_n$, $(CAA/GTT)_n$, $(CTT/GAA)_n$, $(CAT/GTA)_n$ and $(GGA/CCT)_n$ microsatellites accounted for 96% of the total, while $(GGT/CCA)_n$, $(CGT/GCT)_n$, $(CTA/GAT)_n$ and $(CGA/GCT)_n$ accounted for only 4%. This result is quite similar to our observations in wheat where $(TAA/ATT)_n$, $(CTT/GAA)_n$ and $(CAA/GTT)_n$ accounted for about 96% of the clones with eight or more repeat units. Lagercrantz et al. (1993) reported that $(TAA/ATT)_n$ and $(CTT/GAA)_n$ were the most-common trinucleotide motifs with n≥8 in plants through a search of the EMBL database. In tomato, these two microsatellites, $(TAA/ATT)_n$ and $(CTT/GAA)_n$, were nearly as frequent as the dinucleotide repeats $(AG/TC)_n$ and $(AC/TG)_n$ as determined by genomic library screening (Broun and Tanksley 1996). Work in our laboratory demonstrated that $(TAA/ATT)_n$ was the most useful trinucleotide motif for developing SSR markers, and has resulted in the development of more than 500 SSR markers in soybean (Cregan et al. 1999). Of further note is the fact that $(TAA/ATT)_n$ repeats were also found to be the most abundant and the most frequently polymorphic among all of the trinucleotide motifs in humans (Gastier et al. 1995). Their work resulted in the development and mapping of more than 500 $(TAA/ATT)_n$ microsatellite markers (Sheffield et al. 1995).

Effectiveness of developing polymorphic markers in wheat

The conversion of microsatellite-containing sequences into useful markers is sometimes quite difficult, especially in species with large genomes. Pfeiffer et al. (1997) reported that only 20% of 36 primer pairs flanking dinucleotide repeats amplified single variable loci from spruce genomic DNA. The remaining primer pairs either gave poor amplification or produced a complex pattern with multiple bands. A similar low conversion rate from SSR-containing sequence to useful markers was observed in pine (Smith and Devey 1994; Kostia et al. 1995) and wheat (Röder et al. 1995). The low conversion rates of primer pairs to useful markers in these species were probably due to the high level of repetitive DNA sequences in their genomes. Based upon our experiment, on average, about 30% of the $(TAA/ATT)_n$, 7% of the $(CAA/GTT)_n$ and 6% of the $(CTT/GAA)_n$ containing clones identified by library screening resulted in a useful microsatellite marker. No useful markers resulted from other clones containing other repeat motifs. Obviously the $(TAA/ATT)_n$ repeat was preferable to other trinucleotide repeats. In this study, a high proportion (66%) of discrete single or double products were amplified by designing primers to the flanking sequences of $(TAA/ATT)_n$ SSRs.

Assignment of loci to A, B and D genomes

Even though some of the loci were assigned to two or three different genomes based on the evaluation of genome donors, these loci generally amplified only a single clear fragment in durum and hexaploid wheat. These loci can be readily used for mapping and other analyses. For example, Fig. 1 showed that BARC004 amplified a strong product from the B-genome donor as well as two products from the D-genome donor. However, with tetraploid or hexaploid wheats as templates, only one clear product was generally apparent. Likewise, in the case of BARC012, a weakly amplified product was apparent with each of the A-, B-, and D-genome donors; however, with the tetraploid and hexaploid wheats, a single discrete product was apparent. Thus, while five of the 18 loci were assigned to more than one genome based upon amplification of PCR products from two or three of the genome donors, these primers usually amplified a single discrete product from Chinese Spring and the other hexaploid wheats. Therefore, the use of the genome donors does not appear to be a reliable method to determine the genome from which a new SSR will amplify a product when hexaploid wheat is used as the template. This lack of reliabilty may result from the fact that PCR amplification will occur from a homoeologous locus when the specific locus to which primers were designed is not present. For example, if primers were designed to a locus in the A genome they might amplify the homoeologous locus in the B or D genome in the diploid genome donors because there is no perfectly homologous template present to compete for primer annealing. In contrast, when hexaploid wheat is used as a template the homologous locus is present as a template and it succesfully competes for primer annealing.

The use of Chinese Spring, Opata 85 and M6 to identify polymorphic primer sets

Among the 18 primer sets which amplified polymorphic products on 10 hexaploid and one tetraploid wheat genotypes, all were polymorphic in a comparison with Chinese Spring, Opata 85 and M6. Thus, these three genotypes could be effectively used as a preliminary screen for the polymorphism of new SSR loci, thereby decreasing the number of amplification reactions needed to test each primer set.

Acknowledgements The authors thank the Agricultural Technology, Utilization and transfer Project (ATUT) Project 008, and the U.S. Wheat and Barley Scab Initiative project for support of this research. Also, the excellent technical assistance of Charles Quigley, Susan Fogarty, Sarah Hyatt, Brian Boynton and P. Chris Lee is gratefully acknowledged.

References

Akagi H, Yokozeki Y, Inagaki A, Fujimura T (1996) Microsatellite DNA markers for rice chromosomes. Theor Appl Genet 93:1071–1077

Akkaya MS, Bhagwat AA, Cregan PB (1992) Length polymorphisms of simple sequence repeat DNA in soybean. Genetics 132: 1131–1139

Ayres NM, McClung AM, Larkin PD, Bligh HFJ, Jones CA, Park WD (1997) Microsatellites and a single nucleotide polymorphism differentiate apparent amylose classes in an extended pedigree of US rice germ plasm. Theor Appl Genet 94: 773–778

Becker J, Heun M (1995) Barley microsatellites: allele variation and mapping. Plant Mol Biol 27: 835–845

Bligh HFJ, Blackhall NW, Edwards KJ, McClung AM (1999) Using amplified fragment length polymorphisms and simple sequence length polymorphisms to identify cultivars of brown and white milled rice. Crop Sci 39:1715–1721

Broun P, Tanksley SD (1996) Characterization and genetic mapping of simple repeat sequences in the tomato genome. Mol Gen Genet 250:39–49

Chao S, Sharp PJ, Worland AJ, Warham EJ, Koebner RMD, Gale MD (1989) RFLP-based genetic maps of wheat homologous group-7 chromosomes. Theor Appl Genet 78:495–504

Cregan PB, Bhagwat AA, Akkaya MS, Jiang RW (1994) Microsatellite fingerprinting and mapping of soybean. Methods Mol Cell Biol 5:49–61

Cregan PB, Jarvik T, Bush AL, Shoemaker RC, Lark KG, Kahler AL, VanToai TT, Lohnes DG, Chung J, Specht JE (1999) An integrated genetic linkage map of the soybean genome. Crop Sci 39:1464–1490

Diwan N, Cregan PB (1997) Automated sizing of fluorescent-labelled simple sequence repeat (SSR) markers to assay genetic variation in soybean. Theor Appl Genet 95:723–733

Edwards A, Civitello A, Hammond HA, Caskey CT (1991) DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. Am J Hum Genet 49:746–756

Fahima T, Röder MS, Grama A, Nevo E (1998) Microsatellite DNA polymorphism divergence in *Triticum dicoccoides* accessions highly resistant to yellow rust. Theor Appl Genet 96:187–195

Gastier JM, Pulido JC, Sunden S, Brody T, Buetow KH, Murray JC, Weber JL, Hudson TJ, Sheffield VC, Duyk GM (1995) Survey of trinucleotide repeats in the human genome: assessment of their utility as genetic markers. Hum Mol Genet 4:1829–1836

Hearne CM, Ghosh S, Todd JA (1992) Microsatellites for linkage analysis of genetic traits. Trends Genet 8:288–294

Jin W, Palmer RG, Horner HT, Shoemaker RC (1998) Molecular mapping of a male-sterile gene in soybean. Crop Sci 38: 1681–1685

Keim P, Olson T, Shoemaker R (1988) A rapid protocol for isolating soybean DNA. Soybean Genet Newslett 15:150–152

Korzun V, Boerner A, Worland AJ, Law CN, Röder MS (1997) Application of microsatellite markers to distinguish intervarietal chromosome substitute lines of wheat Triticum aestivum L. Euphytica 95: 149–155

Kostia S, Varvio SL, Vakkari P, Pulkkinen P (1995) Microsatellite sequences in a conifer, *Pinus sylvestris*. Genome 38:1244–1248

Lagercrantz U, Ellegren H, Andersson L (1993) The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. Nucleic Acids Res 21:1111–1115

Liu ZW, Biyashev RM, Saghai Maroof MA (1996) Development of simple sequence repeat DNA markers and their integration into barley linkage map. Theor Appl Genet 93:869–876

Morgante M, Olivieri AM (1993) PCR-amplified microsatellites as markers in plant genetics. Plant J 3:175–182

Mian MAR, Wang T, Phillips DV, Alvernaz J, Boerma HR (1999) Molecular mapping of the *Rcs3* gene for resistance to frogeye leaf spot in soybean. Crop Sci 39:1687–1691

Pestsova E, Ganal MW, Röder MS (2000a) Isolation and mapping of microsatellite markers specific for the D genome of bread wheat. Genome 43:689–697

Pestsova E, Salina E, Börner A, Korzun V, Maystrenko OI, Röder MS. (2000b) Microsatellites confirm the authenticity of intervarietal chromosome substitution lines of wheat (*Triticum aestivum* L.). Theor Appl Genet 101: 95–99

Pestsova E, Korzun V, Goncharov NP, Hammer K, Ganal MW, Röder MS (2000c) Microsatellite analysis of *Aegilops tauschii* germplasm. Theor Appl Genet 101:100–106

Panaud O, Chen X, McCouch SR (1996) Development of microsatellite markers and characterization of simple sequence length polymorphism (SSLP) in rice (*Oryza sativa* L.). Mol Gen Genet 252:597–607

Pfeiffer A, Olivieri AM, Morgante M (1997) Identification and characterization of microsatellites in Norway spruce (*Picea abies* K.). Genome 40:411–419

Plaschke J, Ganal MW, Röder MS (1995) Detection of genetic diversity in closely related bread wheat using microsatellite markers. Theor Appl Genet 91:1001–1007

Röder MS, Plaschke J, König SU, Börner A, Sorrells ME, Tanksley SD, Ganal MW (1995) Abundance, variability and chromosome location of microsatellites in wheat. Mol Gen Genet 246:327–333

Röder MS, Korzun V, Wendehake K, Gill BS, Ganal MW (1998a) The physical mapping of microsatellite markers in wheat. Genome 41:278–283

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

Rongwen J, Akkaya MS, Lavi U, Cregan PB (1995) The use of microsatellite DNA markers for soybean genotype identification. Theor Appl Genet 90:43–48

Senior ML, Chin ECL, Lee M, Smith JSC, Stuber CW (1996) Simple sequence repeat markers developed from maize sequences found in the GenBank database: map construction. Crop Sci 36:1676–1683

Sheffield VC, Weber JL, Buetow KH, Murray JC, Even DA, Wiles K, Gastier JM, Pulido JC, Yandava C, Sunden S, Mattes G, Businga T, McClain A, Beck J, Scherpier T, Gilliam J, Zhong J, Duyk GM (1995) A collection of tri- and tetra-nucleotide repeat markers used to generate a high quality, high resolution human genome-wide linkage map. Hum Mol Genet 4: 1837–1844

Smith DN, Devey ME (1994) Occurrence and inheritance of microsatellites in *Pinus radiata*. Genome 37:977–983

Song QJ, Shim KM, Heo NK, Kim NS (1998a) Genotype fingerprinting, differentiation and association between morphological traits and SSR loci of soybean landraces. Plant Resourc 1:81–91

Song QJ, Kim NS (1998b) SSRFING: a BASIC program for plant fingerprinting with SSR information. Korea J Genet 20:29–34

Song QJ, Quigley CV, Nelson RL, Carter TE, Boerma HR, Strachan JL, Cregan PB (1999) A selected set of trinucleotide simple sequence repeat markers for soybean cultivar identification. Plant Var Seeds 12:207–220

Sun GL, Salomon B, Bothmer RV (1998) Analysis of tetraploid *Elymus* species using wheat microsatellite markers and RAPD markers. Genome 40:086–814

Weber JL, May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J Hum Genet 44:388–396

Worland AJ, Korzun V, Röder MS, Ganal MW, Law CN (1998) Genetic analysis of the dwarfing gene *Rht8* in wheat. Part II. The distribution and adaptive significance of allelic variants at the *Rht8* locus of wheat as revealed by microsatellite screening. Theor Appl Genet 96:1110–1120

Wu KS, Tanksley SD (1993) Abundance, polymorphism and genetic mapping of microsatellites in rice. Mol Gen Genet 244:225–235