BULG. J. PLANT PHYSIOL., 2002, 28(1–2), 3–10

ISOLATION OF WHEAT MICROSATELLITE DNA FRAGMENTS BY HYBRIDIZATION SELECTION

Erdogan Hakki[1], Christo Peev[2], Luchezar Karagyozov[2*], Mahinur Akkaya[1]

[1]Middle East Technical University, Department of Chemistry, 06531 Ankara, Turkey
[2]Bulgarian Academy of Sciences, Institute of Molecular Biology, 1113 Sofia, Bulgaria

Received May 7, 2002

Summary. A procedure for selection of wheat genomic fragments containing various microsatellites is described. The procedure is based on the hybridization of single-stranded DNA fragments either to filter bound microsatellite DNA or to biotinylated oligonucleotides in solution. Hybridized strands are isolated, amplified and cloned. Results presented show that enrichment for $(CA)_n/(TG)_n$ microsatellites depends on hybridization conditions and on amplification protocol. Enrichment can be significantly improved by performing several hybridization selection rounds. The sequences of two wheat microsatellite clones are presented. The method is simple, saves time and may be used for microsatellite fragment isolation from plant DNA for genetic, evolutionary or population studies.

Key words: DNA enrichment, microsatellites, wheat

INTRODUCTION

Microsatellites, known also as simple sequence repeats (SSRs) or short tandem repeats (STRs) are DNA stretches composed of simple motifs, 2 to 6 base pairs in length, tandemly repeated. Nuclear DNA of all eukaryotes contains microsatellites, scattered throughout the genome. Primers, flanking the simple sequence repeat can amplify a specific microsatellite. After amplification the microsatellite length can be estimated by electrophoresis. Results in humans showed that most microsatellite loci are polymorphic. Allele identification in SSRs is highly reproducible, which makes microsatellites convenient and highly informative markers for genetic analysis (Tautz, 1989; Weber and May, 1989; Weber, 1990).

* Corresponding author, e-mail: lkara@bio25.bas.bg

In plants, early studies demonstrated that loci with microsatellites such as $(AT)_n$, $(CA)_n$ or (AAT)n, are multi-allelic and somatically stable. This makes microsatellites good candidates for genetic markers in plants too (Akkaya et al., 1992; Rafalski and Tingey, 1993). The use of DNA markers is especially important for species with low levels of polymorphism, such as wheat. Although enough wheat microsatellites are identified to construct a microsatellite based genetic map (Roder et al., 1998), the available microsatellite sequences are still limited in number.

Wheat, one of the most important agricultural plants, has an amphiploid genome, composed of three genomes A, B and D. Hexaploid wheat (bread wheat) has an AABBDD genome and the tetraploid wheat (durum wheat) has an AABB genome. The wheat genome is extremely large (16×10^9 bp\1C, Bennett and Smith, 1976) and contains much repetitive DNA (more than 80%). On the other hand it is known that microsatellites in plants are associated with the low copy fraction of the genome (Morgante et al., 2002). This hinders the identification and isolation of microsatellite loci, when traditional library construction and screening methods are used (Roder et al., 1998).

In this paper we describe procedures for construction of a wheat genomic library enriched for microsatellites. Procedures are based on hybridization selection of genomic fragments. Enrichment depends on hybridization conditions and on the amplification protocol. Enrichment is improved by performing several cycles of hybridization and selection. The sequences of two wheat clones containing microsatellites are shown.

MATERIALS AND METHODS

DNA digestion and amplification of genomic fragments

Wheat DNA was isolated by a standard procedure (Akkaya et al., 1992) and digested with restriction endonuclease Taq I (cleavage site T/CGA). The DNA fragments (200 ng) were ligated at 8 ºC to 20 pmoles of adapter, composed of two oligonucleotides: RX24 (AGCACTCTGCAGCCTCTAGATCTC) and RX11 (CGGAGATCTAG). This adapter can be ligated to DNA fragments with 5´ CG overhangs and it contains a Xba I site. Ligated fragments (20 ng) were amplified (thermal MiniCycler, MJ Research) in 50 µl volume of 10 mM Tris-HCl, pH 8.3, 1.5 mM $MgCl_2$, 50 mM KCl, 10 µg/ml gelatin, 200 µM dNTPs, 5 µM primer RX24 (95 ºC, 30 sec., 55 ºC, 30 sec., 72 ºC, 1 min., 15 cycles). The amplified genomic fragments were subjected to enrichment.

Labeling of oligonucleotides and biotinylation

Oligonucleotides (10 pmoles, 10 µl reaction) were 5´-end labeled with 5 µCi ^{32}P [γATP] and 2 U T4 polynucleotidekinase (15 min, 37 ºC) or 3´-end biotinylated with

bio-dUTP (20 pmoles) and 10 U terminal transferase (60 min, 37°C). Oligonucleotides were ethanol precipitated and dissolved in water, conc. 1 pmole/µl.

Binding of oligonucleotides and plasmid DNA to nylon filters

Oligonucleotides are spotted on 3–4 mm pieces of nylon membrane (Hybond N+, Amersham, UK), water washed, baked and UV cross-linked as described (Karagyozov et al., 1993). Plasmid pCA24 was also used for hybrid selection. This plasmid contains a stretch of $(CA)_{25}$ dinucleotides (Fig. 2) in a 164 bp fragment of mouse DNA, cloned into the SmaI site of pBluescript SK+ (Stratagene). One pmole of pCA24 (2 µg) was spotted on a 3–4 mm piece of nylon, air dried and denatured in 0.5 M NaOH/1 M NaCl for 5 min, washed thoroughly in water and baked at 80°C for 1 hour (DNA retention on filter exceeds 90%).

Hybridization selection

The denatured amplified genomic fragments (50 ng) were hybridized overnight at 37°C to filter-bound $(GT)_{15}$ or pCA24 in a minimal volume (350 µl) of 5×SSC, 1% SDS, 50 mM sodium phosphate buffer, pH 7.0. Filters were washed (room temperature) 5 times with 2×SSC, 1% SDS, 50 mM sodium phosphate buffer. Final washes were with 0.5×SSC, 1% SDS. The filters were put into 100 µl 0.5% SDS and heated for 15 min. at 80°C. Detached strands were precipitated with carrier and amplified. Subsequent enrichment cycles were performed as indicated, using the same conditions.

The hybridization selection was also performed by another procedure (Prochazka, 1996). Fragments containing $(CA)_n$ repeats were hybridized to 3´-biotinylated $(GT)_{15}$ oligonucleotides and then separated on streptavidin coated magnetic beads. The biotinylated oligonucleotides (1 pmol) and the heat denatured amplified genomic fragments (50 ng) were hybridized overnight at 37°C in 40 µl 5×SSC, 0.1% SDS, 50 mM sodium phosphate buffer, pH 7.0. Streptavidin coated beads (Promega, 50 µl, 10 mg/ml) were washed at room temperature with 0.5 ml 5×SSC and for 30 min in 0,2 ml 5×SSC, 50 mM sodium phosphate buffer, pH 7.0, 25 µg tRNA, 1 µg single stranded M13mp18 phage DNA. The beads were then suspended in 360 µl 5×SSC, 0.1% SDS, 50 mM sodium phosphate, added to the hybridization solution and rolled in a hybridization chamber for 1 hour at 37°C. The beads were washed (room temperature, 1 min. 1 ml washes) with 5×SSC, 0,1% SDS, 2×SSC, 1×SSC and finally with 0,1×SSC (once). DNA fragments enriched for $(CA)_n$ repeats were eluted from beads in 40 µl TE buffer at 80°C for 15 min. Beads were separated and DNA amplified as indicated above.

The enrichment for microsatellites was assayed by hybridization. Samples (2 µl) of the amplified genomic fragments were dotted on nylon filter and probed with [^{32}P] 5´-end labeled $(GT)_{15}$ oligonucleotides. The hybridization intensity was assayed either by liquid scintillation counting or by autoradiography as described in Sambrook et al., 1989.

Cloning and library screening

Amplification products were digested with Xba I restriction endonuclease which cleaves the adapter and cloned into pBluescript SK+ (Stratagene). The clones were screened by hybridization with 5´-end labeled $(TG)_{15}$ (5×SSC, 5% SDS, 50 mM sodium phosphate buffer, pH 7.0, 42ºC, overnight). Insert size was estimated by PCR with M13 primers. Some of the positive clones were sequenced by the dideoxy method.

Results and Discussion

Preparation of wheat DNA fragments for amplification.

To take advantage of the benefits of the microsatellite polymorphism, the short tandem repeats must be identified and their flanking regions sequenced. Sequences are to be used to select suitable PCR primers. Screening of genomic libraries is the most common method for identification of DNA segments, containing microsatellites. In mammals, the most abundant microsatellite, the $(CA)_n$ repeat, is present in about 1% of clones with average insert size 500 bp. This makes the screening process time consuming and has provoked the introduction of a step of enrichment for genomic fragments, containing microsatellites (Karagyozov et al., 1993; Armour et al., 1994). Briefly the employed procedure is as follows. At first, short synthetic adapters are ligated to both ends of genomic DNA fragments. This converts the fragments into a population of molecules, which can be amplified (Kinzler and Vogelstein, 1989). Next, the fragments are denatured and hybridized to the desired microsatellite motif(s). The hybridized fragments are amplified and cloned.

The need for an enrichment is more pronounced with large plant genomes, such as those of *Gramineae,* where microsatellites are about 10 times less abundant than in mammals (Wu and Tanksley, 1993; Roder et al., 1995). We fragmented wheat DNA by digestion with restriction endonuclease TaqI and ligated short double stranded adapters to the fragments. In most experiments the adapter oligonucleotides (see Materials and Methods) were non-phosphorylated. In that case only one strand of the adapter is ligated, so the other strand was filled in by the Taq DNA polymerase prior to amplification (5 min., 70ºC).

Enrichment of the genomic fragments for microsatellite

To select the repeat-containing sequences we hybridized the PCR library to filter bound $(GT)_{15}$ oligonucleotides (Karagyozov et. al. 1993) or to denatured filter bound pCA24 DNA. Experiments with 5´-end labeled oligonucleotides (10 pmoles) showed that $(GT)_{15}$ binds better than $(CA)_{15}$ and that retention on filter increases upon UV irradiation. However hybridization to filter bound $(GT)_{15}$ decreased with longer exposures

to UV upon binding. This is apparently due to damage of the oligonucleotide structure by UV irradiation.

In view of these results, in later hybridization selection experiments we used plasmid pCA24 as a source of $(AC)_n/(GT)_n$ sequences (see Materials and Methods). Alternatively, biotin 3´-end labeled $(GT)_{15}$ oligonucleotides were hybridized to the modified genomic fragments and the hybridization mixture was treated with streptavidin coated beads.

After hybridization and removal of non-specifically bound material, hybridized strands were detached and amplified. At that point amplification products may be cloned. However additional rounds of hybridization selection were found useful, resulting in further enrichment. In an experiment the strands were detached from filter and amplified after each round of enrichment. Hybridization of PCR products to labeled $(CA)_{15}$ (Fig. 1) shows that hybridization signal does not increase significantly after 20–25 cycles of amplification, although the amount of amplification products still increases. Most probably this is due to reannealing of abundant PCR products during amplification. Consequently, amplification of microsatellites plateaus while non-specifically bound strands continue to amplify. Apparently, fragments after more cycles of amplification will be less useful for cloning as the proportion of positive clones will be less.

Cloning, screening the library, and characterization of clones

Genomic fragments (three rounds of hybridization selection, 15 cycles of amplification) enriched in TG/CA-repeats were cloned in *E. coli* (see Materials and Methods). A limited number of clones were screened and some positively hybridizing clones were sequenced (positive were 30% of all clones). Repeated structures of five clones

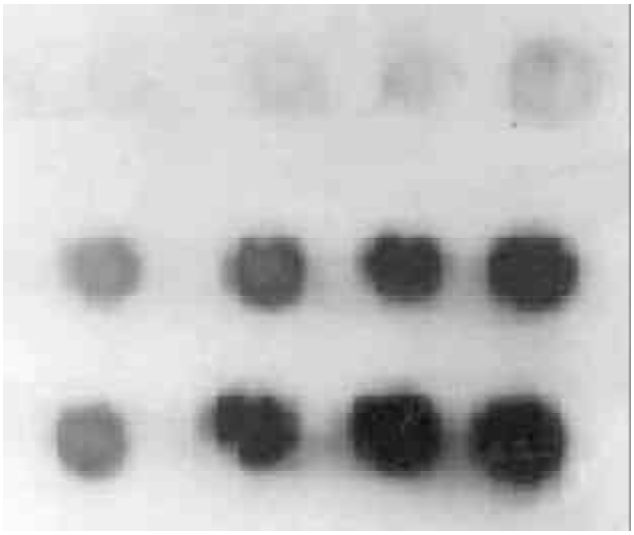

Fig. 1. Amplification of fragments detached from filter (80ºC, 15 min). Amplification products (2 ml) were spotted on nylon filter and hybridized with 5´-end labeled $(GT)_{15}$. A, B and C – first, second and third round of hybridization selection. 1, 2, 3 and 4 – PCR products after 15, 20, 25 and 30 cycles of amplification.

were classified according to Weber (1990). One clone (p14) did not have microsatellites (repeat length < 12). Two clones contained $(CT)_n/(GA)_n$ repeats: perfect $(GA)_8$ in clone p17 and imperfect $(CT)_3$ TC $(CT)_4$ GT $(CT)_{12}$ in clone p13. Three clones contained $(AC)_n/(GT)_n$ microsatellites (n>6). Clone p36 contains the compound microsatellite $(TG)_7$ $(AG)_7$ and clone p311 contains the imperfect $(CA)_8$ T $(CA)_2$ A $(CA)_5$ (Fig. 2). Clone p17 contains also the compound $(GC)_5$ $(GT)_6$.

Mouse, clone pCA24

TTTGATGTGT	GCCTTGCACT	GCAATTGGTT	CACACTTGCA	GTCTTGCTTT	50
AAGCGTGTGT	GTGTGTGTGT	GTGTGTGTGT	GTGTGTGTGT	GTGTGTGTGT	100
GTGTGATGTG	TACTTGAGTG	TAGGGGTCTA	TGGAGGCCAG	AAGAGGGCGT	150
CAGATACCCA	TTTA				

Wheat, clone p36

TCGAAGGACA	TCCACCGAAC	CTTTTTGTGT	GTGTGAAAAC	AATCCCCGAA	50
TATTACTCAA	CTTTTTTTCC	TGGAAAAGTT	CTTGACGCTA	CAATATTTCC	100
ATATATTTGA	ATTTAGCTCC	AGTTTGTGTT	TATTTGGATT	TGGACGTTTT	150
AGGTGCGCCC	TAGTTTTTTT	TAATACTGAT	TTTGTGTGTG	TGACTGAGAG	200
AGAACGTGTG	TATGTGTCCA	TATGTGTGTT	GTGGCGGAAG	GGCAATGTGG	250
GGATGGTGTG	CATGTGATAG	AAACATAGAG	AGGTGTGAAT	GTGCAAATGA	300
TCATGCATGA	GTGAGACATA	GGCAGATGCT	GATACGTTGT	GTATAAATGA	350
GAGAACGGT	CTTCTATTTT	ACTTGTGTGT	GTGTGTGAGA	GAGAGAGAGA	400
GGAGCATTTG	TAACACAACA	ACCATCTCTC	TCGA		

Wheat, clone p311

TCGATTCGTC	CGTCCCTCGC	ATGATCGCAT	GTAACACATG	CATCAACACG	50
CACATTTTGA	CACCCTCACC	CAACCCCTGC	CCACCTCAAG	GCACGCACAA	100
CCTCTTCGTG	AATACCCATT	TCTCTCCTTT	GGTTTCTTTT	CTCTCAATAT	150
CTGACACACA	CACACACACA	TCACAACACA	CACACACCCT	CCCCCGAGCA	200
CACAATTCAT	CCCCATCCAA	GGTTATAGGG	CGACCTCTCT	CTCTTGTGCT	250
TCTTTGCACC	CTAGCATGAA	CAACACCTCA	ATCTTCAAAG	AGGGCATCGA	300

Fig. 2. Clone sequences, microsatellites are underlined: A – Mouse DNA clone pCA24, SSR position 55–104; B – wheat DNA clone p36, SSR position 374–401; C – wheat DNA clone p311, SSR position 155–186.

With large plant genomes finding microsatellite loci may be speed up and improved by procedures increasing SSR frequency in screened libraries. Several methods have been proposed for that purpose, e.g. a selective second strand DNA synthesis (Ostrander et al., 1992), triplex affinity capture (Ito et al., 1992), or hybridization selection (Karagyozov et al., 1993; Kandpal et al., 1994). Here we describe enrichment procedure based on annealing of single-stranded DNA fragments to biotinylated oligonucleotides or to microsatellite plasmid DNA, bound to filter. Affinity captured strands are amplified and cloned. Degree of final enrichment depends on hybridization con-

ditions and on amplification protocol. Enrichment can be significantly improved by performing several enrichment rounds. The method is simple, saves time and may be used for microsatellite isolation from wheat or other important crops for genetic, evolutionary or population studies.

Acknowledgment: This work has been partially supported by grant CRP/BUL95-02 to L.K. from ICGEB, Trieste, Italy.

References

Armour, J. A., R. Neumann, S. Gobert, A. J. Jeffreys, 1994. Isolation of human simple repeat loci by hybridization selection. Human Molec. Genetics, 3, 599–605.

Akkaya, M. S., A. A. Bhagwat, P. B. Cregan, 1992. Length polymorphism of simple sequence repeat DNA in soybean. Genetics, 132, 1131–1139.

Bennett, M. D., J. B. Smith, 1976. Nuclear DNA amounts in angiosperms. Phil. Trans. Roy. Soc. Lond. B, 274, 227–274.

Ito, T., C. L. Smith, C. R. Cantor, 1992. Sequence-specific DNA purification by triplex affinity capture. *Proc. Natl. Acad. Sci. U S A,* 89, 495–498.

Kandpal, R. P., G. Kandpal, S. M. Weissman, 1994. Construction of libraries enriched for sequence repeats and jumping clones, and hybridization selection for region-specific markers. *Proc. Natl. Acad. Sci. U S A,* 91, 88–92.

Karagyozov, L., I. D. Kalcheva, V. Chapman, 1993. Construction of random small-insert genomic libraries highly enriched for simple sequence repeats. Nucl. Acids Res., 21, 3911–3912.

Kinzler, K. W., B. Vogelstein, 1989. Whole genomic PCR: application to the identification of sequences bound to gene regulatory proteins. Nucl. Acids Res., 10, 3645–3653.

Morgante, M., M. Hanafey, W. Powell, 2002. Microsatellites are preferentially associated with non-repetitive DNA in plant genomes. Nature Genet., 30, 194–200.

Ostrander, E. A., P. M. Jong, J. Rine, G. Duyk, 1992. Construction of small-insert genomic DNA libraries highly enriched for microsatellite repeat sequences. Proc. Natl. Acad. Sci. U S A, 1992, 89, 3419–3423.

Prochazka, M., 1996. Microsatellite hybrid capture technique for simultaneous isolation of various STR markers. Genome Res., 6, 646–649.

Rafalski, J. A., S. V. Tingey, 1993. Genetic diagnostics in plant breeding: RAPDs, microsatellites and machines. Trends Genet., 9, 275–280.

Röder, M. S., J. Plaschke, S.U. König, A. Börner, M. E. Sorrells, S.D. Tanksley, M. W. Ganal, 1995. Abundance, variability and chromosomal location of microsatellites in wheat. Mol. Gen. Genetics, 246, 327–333.

Roder, M. S., V. Korzun, K. Wendehake, J. Plaschke, M-H. Tixier, P. Leroy, M. Ganal, 1998. A microsatellite map of wheat. Genetics, 149, 2007–2023.

Sambrook, J., E. F. Fritsch, T. Maniatis, 1989. Molecular cloning: a laboratory manual (2nd edn.) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Tautz, D., 1989. Hypervariability of simple sequences as a general source for polymorphic markers. Nucl. Acids Res. ,17, 6463–6471.

Weber, J. L.,1990. Informativness of human (dC-dA)n.(dG-dT)n polymorphisms. Genomics, 7, 524–530.

Weber, J. L., P. E. May, 1989. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am. J. Hum. Genet., 44, 388–396.

Wu, K., S. D. Tanksley, 1993. Abundance, polymorphism and genetic mapping of microsatellites in rice. Mol. Gen. Genetics, 241, 225–235.