VERSITA

CELLULAR & MOLECULAR BIOLOGY LETTERS
Volume 11, (2006) pp 291 – 298
http://www.cmbl.org.pl

DOI: 10.2478/s11658-006-0023-5

Received: 25 January 2005
Revised form accepted: 18 April 2006

Short communication

IDENTIFICATION OF MICROSATELLITE MARKERS IN THE RYE GENOME

HANNA BOLIBOK[1], MONIKA RAKOCZY-TROJANOWSKA[1]*, MAŁGORZATA WYRZYKOWSKA[1], MAGDALENA RADECKA[2] and WACŁAW ORCZYK[2]

[1]Department of Plant Genetics, Breeding and Biotechnology, Warsaw Agricultural University, Nowoursynowska 166, 02-787 Warszawa, Poland
[2]Plant Breeding and Acclimatization Institute, Radzików, 05-870 Błonie, Poland

Abstract: The rye genomic library, which consists of DNA fragments in the range of 0.5-1.1 kb, was screened for the presence of tri- and tetranucleotide and compound microsatellites. Of the 1,600,000 clones analysed, 102 clones were positive and 41 were suitable for SSR primer pair design. Twenty-six primer pairs amplified specific products, and six of them were capable of detecting polymorphism among 30 rye accessions of different genetic backgrounds. Using a set of Chinese Spring-Imperial wheat-rye addition lines, it was possible to locate 3 newly identified microsatellites on chromosomes 3R, 4R and 7R.

Key words: Rye, Genomic library, SSR, Microsatellite, Wheat-rye addition lines

INTRODUCTION

Microsatellites are a class of repetitive DNA sequences present in all organisms, both eukaryotes [1, 2] and prokaryotes [3]. The molecular markers based on these sequences – SSRs, ISSRs and SAMPLs – are characterized by an extremely high level of polymorphism [4]. Thus, as numerous authors have stated, they are highly suitablefor the construction of molecular maps and other genomic research.

* Corresponding author: e-mail: Monika_Rakoczy_Trojanowska@sggw.pl

Abbreviations used: SSR – simple sequence repeat; ISSR – inter-simple sequence repeat; SAMPL – selective amplification of microsatellite polymorphic loci; EST – expressed sequence tags; IS – insert size [bp]; GmS – genome size [bp]; GS – genetic similarity

Microsatellite-based markers were applied for the molecular mapping of numerous plant species [4]. Several maps containing SSR markers are available for rye [5-9]. These currently available rye maps predominantly contain EST-SSRs, which detect the variation present only in the coding regions of the genome.

The number of genome-based SSRs, especially tri- and tetranucleotide and those consisting of compound repeats, is still relatively low for rye. Therefore, we decided to isolate them from the rye genomic library, characterize newly derived markers with regards to their usefulness for detecting polymorphism within a set of rye accessions, and, if possible, determine their chromosomal location.

MATERIALS AND METHODS

Plant material

Thirty rye inbred lines, described previously [10], were used in this study as the screening array to verify the ability of the developed SSRs to detect polymorphism. A set of Chinese Spring-Imperial wheat-rye addition lines, provided by Prof. A. Łukaszewski (University of California, Riverside, USA) was used to determine the chromosomal location of the SSR markers.

Genomic library construction

DNA was isolated from young seedlings of rye line L318 (one of the 30 inbred lines) using a standard CTAB method [11]. It was then treated with RNase A, phenol/chloroform extracted, and sodium acetate precipitated. Two DNA samples of 43 μg each were partially restricted in separate tubes with *Sau* 3AI. The reaction was performed in the manufacturer's buffer using 1 U of enzyme per 1 μg of plant DNA. Samples containing 10 μg of partially digested DNA were collected after 15, 30, 60 and 120 minutes from the first and after 5, 10 and 20 minutes from the second tube. After separation in 0.7% agarose, DNA fragments in the range 0.5-1.1 kb were excised and isolated using a QIAgen Gel Extraction Kit. To ensure the highest possible representation of the whole rye genome in the selected size range, two DNA samples were extracted independently, and after adjusting the concentration to 100 ng/μl in both isolates, they were mixed in a 2:1 vol./vol. ratio. A 40-ng sample of such DNA was used for ligation with 1 μg of ZAP Express vector predigested with *Bam* HI. Ligation, packaging, titering and amplification of the library were done according to Stratagene's instructions.

The size of the library containing any given sequence with 99% probability was estimated according to the equation $N = \ln(1 - 0.99)/\ln(1 - (IS/GmS))$, where N is the number of clones.

Library screening

The library was screened by plaque lifting onto nylon membranes (Porablot NY amp, Macherey-Nagel) and hybridization. The isolation of microsatellites from the L318 genomic library was performed during two separate experiments. In

the first experiment (exp. 1), the density of plaques was approximately 50,000 pfu/150-mm plate. A total of 1,000,000 plaques were screened with six 3' biotinylated probes: 1. $(CAAA)_6$, 2. $(TTTA)_6$, 3. $(TAAA)_6$, 4. $(TTTG)_6$, 5. $(CTTT)_6$, and 6. $(GATA)_6$. Five probes – 7. $(AAG)_{12}$, 8. $(AGC)_{12}$, 9. $(ATC)_{12}$, 10. $(TG)_5(AG)_5$, and 11. $(AC)_5(AG)_5$ – were used in the second experiment (exp. 2) to screen a total of 600,000 plaques (the density of plaques was approximately 30,000 pfu/150-mm plate).

The hybridizations and detections of signals were preformed according to the protocols provided with the AURORA-Southern blot chemiluminescent detection system for biotin-labeled probes (ICN).

To reduce the number of false positives, two to three rounds of hybridization were carried out. Positive clones identified by plaque hybridization were converted into plasmids by *in vivo* excision (Stratagene protocol).

Sequencing and sequence analysis

Sequencing was preformed on an automated sequencer using dideoxy termination fluorescence chemistry by the DNA sequencing laboratory at the Institute of Biochemistry and Biophysics, Warsaw.

To avoid the sequencing of empty clones or clones bearing an insert too short to contain both flanking sequences of a microsatellite region, the clones identified during exp. 2 were digested with *Hind* III prior to sequencing in order to determine insert length. Only those clones with 900- to 1400-bp long inserts were sequenced. The BLASTn algorithm [12] was used to compare the obtained nucleotide sequences to the sequences deposited in the public databases.

Primer design and evaluation

PCR primers for each microsatellite were designed with Primer 3.0 software (http://www.genome.wi.mit.edu/) using the default settings. The SSR markers developed in this study were sequentially named with the prefix WRM (Warsaw Rye Microsatellites), starting with 201. PCR, electrophoresis and visualization were performed as described previously [10].

Each pair of primers was initially screened for the amplification of a specific product from the DNA of line L318. Primer pairs producing fragments of predicted length were tested for their ability to detect polymorphism within the panel of 30 rye inbred lines. All the sequence information can be obtained from the authors upon request.

Chromosome assignment of SSR markers

The assignment of identified markers to specific chromosomes was accomplished through PCR amplification using template DNA from Chinese Spring, Imperial and each of 7 wheat-rye addition lines. The microsatellite loci were assigned to the chromosomes corresponding to the addition line for which the PCR product characteristic for the Imperial line was observed.

RESULTS AND DISCUSSION

In this paper, we describe the isolation, characterization, capability for polymorphism detection and chromosomal location of SSR markers isolated from the rye genomic library.

The majority of the SSRs mapped on rye linkage maps are derived from EST data bases. However, EST-derived SSR markers are generally less polymorphic than genomic SSRs, as the former are situated in more conserved regions of the genome [13]. The use of such markers results in the potential for a considerable part of possible polymorphism to be lost, as microsatellites are broadly present in the non-coding regions of genomes. On the other hand, EST-derived markers make it possible to tag and isolate genes of interest quickly. In rice, microsatellites derived from genomic libraries detected a higher level of polymorphism than those derived from ESTs contained in the GenBank database – 83.8% vs. 54.0% [14]. The other measures of genetic variability, like the number of alleles per locus, polymorphism information content, and allele size ranges, were higher in the genomic library-derived than in the EST-derived microsatellites. Similarly, in rye, EST-derived SSRs turned out to be less efficient at detecting genetic diversity than those from the genomic library – GS values of 0.3814 and 0.3221, respectively [10].

The reasons mentioned above inclined us towards the isolation of new microsatellite loci from the rye genomic library. We decided (for the first time in research on rye) to use probes detecting tri- and tetranucleotide and compound repeat motifs, as microsatellite sequences containing such core motifs are though to be the basis for producing the most polymorphic and reproducible SSR markers [15, 16].

Genomic library characterization and screening

Two independent ligation reactions of fragmented rye genomic DNA were performed with the ZAP Express vector. This resulted in two separate prime libraries of 600 µl each. The titer of the first was $2.2 \cdot 10^6$ pfu·µl^{-1} and the titer of the second was $5 \cdot 10^4$ pfu·µl^{-1}, which corresponds with 1.2 x 10^9 clones from the first and 3 x 10^7 clones from the second ligation.

The calculated number of clones representing any given rye sequence with 99% probability (with an average insert size (IS) of 800 bp and a genome size (GS) of $9.3 \cdot 10^9$ bp) is $5.3 \cdot 10^7$clones. This indicates that any library containing over $5.3 \cdot 10^7$clones represents, with 99% probability, the whole rye genome. The first library with 1.2 x 10^9 clones clearly met this requirement and was used in the subsequent work.

Of the ca. 1,600,000 clones analyzed in exp. 1 and exp. 2 with the use of each biotinylated probe, 154 were positive. The highest number of positive clones was detected with probes: 1 – $(CAAA)_6$, 10 clones; 3 – $(TAAA)_6$, 32 clones; 7 – $(AAG)_{12}$, 20 clones; and 9 – (-$(ATC)_{12}$), 25 clones.

Sequence analysis and primer design

In total, 72 clones (56 from exp. 1 and 16 from exp. 2) were subjected to sequencing. The obtained nucleotide sequences were submitted to the dbGSS division of GenBank. Thirty-one clones – 19 identified in exp. 1 (including 13 clones detected by probe 3) and 12 from exp. 2 – contained microsatellite repeats and were suitable for SSR primer design (Tab. 1), with more than one microsatellite motif present in some clones (e.g. 3.14.1). BLASTn analysis performed on those 31 clones showed the similarity of 12 clones to the known sequences deposited in data bases, with only one of them originating from *Secale cereale*. Altogether, 41 primer pairs were designed and synthesized (21 in exp. 1 and 14 in exp. 2). Of the isolated microsatellite loci, 13 contained compound core motifs, while 11 contained trinucleotide, 6 binucleotide, 5 mononucleotide, 3 tetranucleotide, 1 pentanucleotide, 1 hexanucleotide and 1 heptanucleotide core motifs.

The very low efficiency of microsatellite isolation achieved in this study may indicate a low representation of tri- and tetranucleotide or compound repeats in the rye genome; this is partially consistent with the data published by Varshney *et al.* [13], who reported tetranucleotide repeats to be represented in cereal genomes in a low proportion, between 2.6 and 6.6%, without any specific trend in relation to the abundance of a particular motif. However, the authors proved that trimeric repeats, the most frequent class of the WRMs, to be the most abundant (54% to 78%) in cereal ESTs.

The final efficiency of microsatellite isolation obtained in our study is lower than that obtained by Saal and Wricke [5], who isolated 534 positive clones out of 255,500 tested clones. Nevertheless, our obtained rate of positive clones suitable for primer design is very similar to that in the above-cited work – the authors sequenced 74 (40.7%) out of 182 positive clones, for 57 (31.3%) of them, primer pairs were designed, and 27 primer pairs resulted in specific SSR markers. Comparable results were achieved for wheat [17, 18].

SSR marker functionality and polymorphism

Of the 41 primer pairs designed, 26 produced a specific product of predicted length from the DNA of rye line L318. The other primer pairs amplified either fragments of many sizes or fragments with sizes other than expected. Six markers, namely WRM202, WRM206, WRM216 and WRM225 (from exp. 1), and WRM 229 and WRM 230 (from exp. 2) detected polymorphism in the set of 30 rye inbred lines. The remaining 20 markers were monomorphic.

The relatively low level of polymorphism detected by the newly isolated SSRs may indicate that microsatellites isolated from our library, usually composed of compound nucleotide motifs with a low number of repeats of the core motif, could be highly conserved in rye. Another possible explanation suggested by some authors may be the location of a high number of monomorphic SSRs on particular chromosomes, namely 1RS [19] or 4R [5]. Remarkably, 4A and 4D wheat chromosomes also contain a very low number of SSRs [17].

Tab. 1. Characteristics of microsatellites isolated from the genomic library of line L318, suitable for SSR primer design.

Clone No./clone designation[1)]	Symbol of primer pairs designed/chromosome assignment	Predicted product size [bp]	Core motif
1/1.189t3	WRM 201*	197	(CA)7
2/1.266(3)	WRM 202*p	266	(A)9(…)(AT)9
3/3.5	WRM 203	155	(C)6(…)(TCT)6
4/3.7	WRM 204*	192	(TTAA)4
5/3.6.2	WRM 205*	180	(TCT)4
6/3.11(3)	WRM 206*p	225	(GAG)4(…)(CTG)3
7/3.14.2	WRM 207*	156	(GAGG)3
8/3.14.1	WRM 208*	157	(CTTTC)2(CT)6
8/3.14.1	WRM 209	178	(CCT)5
9/3.22	WRM 210*	198	(GAA)4
9/3.26	WRM 211*/3R	182	(CACCCAC)2
10/4.36	WRM 212*	167	(GA)3(GTA)4
10/4.36	WRM 213*	227	(AAC)5
10/4.36	WRM 214*	176	(CA)7
2/1.266(3)	WRM 215*	249	(GT)6
11/1.189t7	WRM 216*p/4R	241	(GAAAAA)2(…)(G)6
12/2.34	WRM 217*	159	(GC)7
13/3.1.2	WRM 218*	154	(GAGG)2(CGG)2
14/3.2	WRM 219*	216	(GCG)2(C)8
15/3.6.1	WRM 220*	179	(AGA)2(CGG)3
16/3.12(3)	WRM 221*	151	(T)8
17/3.31t7	WRM 222*	165	(CG)5
18/6.27.1	WRM 223	190	(CT)6
18/6.27.1	WRM 224*	210	(GGGA)2(GAAG)3
18/6.27.1	WRM 225*p	158	(T)13
19/7.7	WRM226*	194	$(GAG)_3$
20/8.12	WRM227	157	$(CAAAAA)_2$
21/9.11	WRM228	188	$(A)_{10}$
21/9.11	WRM229*p/7R	208	$(T)_{10}$
22/11.14	WRM230*p	154	$(GAAA)_2$
23/11.9	WRM231	197	$(AGG)_3$
24/8.11	WRM232	246	$(TGA)_2\,(CT)_3$
25/7.5	WRM233	156	$(A)_8$
24/8.11	WRM234	219	$(GCC)_3$
24/8.11	WRM235	217	$(TGA)_2(CT)_3$
26/9.3	WRM236	231	$(GT)_3(GAA)_2$
27/9.6	WRM237	249	$(TTG)_3$
28/10.3	WRM238	160	$(TGG)_3$
29/11.4	WRM239	249	$(TGC)_4$
30/11.4	WRM240	234	$(GCT)_3$
31/10.6	WRM241*	153	$(GGGGC)_2$

[1)] the first number in the clone designation is the same as the probe number, *primer pairs producing a specific product of predicted length from DNA of rye line L318, p – primer pairs producing polymorphic products.

Chromosome assignment of SSR markers

Three newly developed rye SSR markers (WRM211, WRM216 and WRM229) were respectively assigned to chromosomes 3R, 4R and 7R by means of wheat-rye addition lines, as unambiguous rye-specific amplification products were obtained. The remaining primer pairs produced a strong band in wheat of the same size as in rye, or amplified no product in the Imperial, or in neither the Imperial nor Chinese Spring, thus preventing any chromosome assignment.

The final result of this study was the isolation of 26 microsatellite loci from a rye genomic library. Although the number of SSRs capable of detecting polymorphism in the materials analyzed was rather low (6), it is still an increase of ca. 14% in the number of genomic SSRs which are currently available for rye, and can be used for a variety of standard applications including linkage mapping, genetic diversity studies and identity testing. Since the chromosomal location of some markers could be determined, it can be assumed that they should also be useful in other types of research work, for instance during the characterization of clones from BAC libraries.

Nevertheless, it seems appropriate to continue efforts aimed at developing new SSR markers for rye, as the number of these excellent molecular markers is still too low to ensure a satisfactory survey of the genome of this important crop.

Acknowledgements. The authors would like to thank Prof. A. Łukaszewski (University of California, Riverside, USA) for supplying seeds from a set of wheat-rye addition lines. We are also grateful to Prof. P. Masojć (Agricultural University of Szczecin, Poland) and Dr. L. Madej (Plant Breeding and Acclimatization Institute, Radzików, Poland) for providing seeds of some inbred lines used during the polymorphism survey. This research project was financially supported by the Ministry of Scientific Research and Information Technology, project No 6P06A 007 20.

REFERENCES

1. Litt, M. and Luty, J.A. A hypervariable microsatellite revealed by *in vitro* amplification of a dinucleotide repeat within the cardiac muscle actin gene. **Am. J. Hum. Genet.** 44 (1989) 397-401.
2. Tautz, D., Trick, D. and Dover, G. Cryptic simplicity in DNA is a major source of genetic variation. **Nature** 322 (1986) 652-656.
3. Gur-Arie, R., Cohen, C.J., Eitan, Y., Shelef, L., Hallerman, E.M. and Kashi, Y. Simple sequence repeats in *Escherichia coli*: abundance, distribution, composition, and polymorphism. **Genome Res.** 10 (2000) 62-71.
4. Rakoczy-Trojanowska, M. and Bolibok, H. Characteristics and a comparison of three classes of microsatellite-based markers and their application in plants. **Cell. Mol. Biol. Lett.** 9 (2004) 221-238.
5. Saal, B. and Wricke, G. Development of simple sequence repeat markers in rye (*Secale cereale* L). **Genome** 42 (1999) 964-972.
6. Korzun, V., Malyshev, S., Voylokov, A.V. and Börner, A. A genetic map of

rye (Secale cereale L) combining RFLP, isozyme, protein, microsatellite and gene loci. **Theor. Appl. Genet.** 102 (2001) 709-717.
7. Hackauf, B. and Wehling, P. Identification of microsatellite polymorphisms in an expressed portion of the rye genome. **Plant Breed.** 121 (2002) 17-25.
8. Hackauf, B. and Wehling P. Development of microsatellite markers in rye: map construction. **Plant Breed. Seed Sci.** 48 (2003) 143-151.
9. Khlestkina, E.K., Than, M.H.M., Pestsova, E.G., Röder, M.S., Malyshev, S.V., Korzun, V. and Börner, A. Mapping of 99 new microsatellite-derived loci in rye (*Secale cereale* L.) including 39 expressed sequence tags. **Theor. Appl. Genet.** 19 (2004) 725-732.
10. Bolibok, H., Rakoczy-Trojanowska, M., Hromada, A. and Pietrzykowski R. Efficiency of different PCR-based marker systems in assessing genetic diversity among rye inbred lines. **Euphytica** 146 (2005) 109-115.
11. Murray, M.G. and Thompson, W.F. Rapid isolation of high-molecular-weight plant DNA. **Nucleic Acids Res.** 8 (1980) 4321-4325.
12. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. Basic local alignment search tool. **J. Mol. Biol.** 215 (1990) 403-410.
13. Varshney, R.K., Thiel, T., Stein, N., Langridge, P. and Graner, A. *In silico* analysis of frequency and distribution of microsatellites in ESTs of some cereals. **Cell. Mol. Biol. Lett.** 7 (2002) 537-546.
14. Cho, Y.G., Temnykh, S., Chen, X., Lipovich, L., McCouch, S.R., Ayres, N. and Cartinhour, S. Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (*Oryza sativa* L). **Theor. Appl. Genet.** 100 (2000) 713-722.
15. Schlötterer, C. Evolutionary dynamics of microsatellite DNA. **Chromosoma** 109 (2000) 365-371.
16. Sharopova, N., McMullen, M.D., Schultz, L., Schroeder, S., Sanchez-Villeda, H., Gardiner, J., Bergstrom, D., Houchins, K., Melia-Hancock, S., Musket, T., Duru, N., Polacco, M., Edwards, K., Ruff, T., Register, J.C., Brouwer, C., Thompson, R., Velasco, R., Chin, E., Lee, M., Woodman-Clikeman, W., Long, M.J., Liscum, E., Cone, K., Davis, G. and Coe Jr., E.H. Development and mapping od SSR markers for maize. **Plant Mol. Biol.** 48 (2002) 463-481.
17. Röder, M.S., Korzun, V., Wendehakem K., Plaschkem, J., Tixier, M.H., Leroy, P. and Ganal, M.W. A microsatellite map of wheat. **Genetics** 149 (1998) 2007-2023.
18. Pestsova, E., Ganal, M.W. and Röder, M.S. Isolation and mapping of microsatellite markers specific for the D genome of bread wheat. **Genome** 43 (2000) 689-697.
19. Ma, X.F., Wanous, M.K., Houchins, K., Rodriguez Milla, M.A., Goicoechea, P.G., Wang, Z., Xie, M. and Gustafson, J.P. Molecular linkage mapping in rye (*Secale cereale* L). **Theor. Appl. Genet.** 102 (2001) 517-523.