Sequencing of Simple Sequence Repeat Anchored Polymerase Chain Reaction Amplification Products of *Biomphalaria glabrata*

Roberta L Caldeira, Omar S Carvalho/[+], Regina CG Lage, Paula Cristina M Cardoso, Guilherme C Oliveira*

Laboratório de Helmintoses Intestinais and Laboratório de Parasitologia Celular e Molecular, Centro de Pesquisas René Rachou-Fiocruz, Av. Augusto de Lima 1715, 30190-002 Belo Horizonte, MG, Brasil *Programa de Pós-graduação e Pesquisa, Santa Casa de Misericórdia, Belo Horizonte, MG, Brasil

Simple sequence repeat anchored polymerase chain reaction amplification (SSR-PCR) is a genetic typing technique based on primers anchored at the 5' or 3' ends of microsatellites, at high primer annealing temperatures. This technique has already been used in studies of genetic variability of several organisms, using different primer designs. In order to conduct a detailed study of the SSR-PCR genomic targets, we cloned and sequenced 20 unique amplification products of two commonly used primers, CAA(CT)$_6$ and (CA)$_8$RY, using Biomphalaria glabrata *genomic DNA as template. The sequences obtained were novel* B. glabrata *genomic sequences. It was observed that 15 clones contained microsatellites between priming sites. Out of 40 clones, seven contained complex sequence repetitions. One of the repeats that appeared in six of the amplified fragments generated a single band in Southern analysis, indicating that the sequence was not widespread in the genome. Most of the annealing sites for the CAA(CT)$_6$ primer contained only the six repeats found within the primer sequence. In conclusion, SSR-PCR is a useful genotyping technique. However, the premise of the SSR-PCR technique, verified with the CAA(CT)$_6$ primer, could not be supported since the amplification products did not result necessarily from microsatellite loci amplification.*

Key words: *Biomphalaria glabrata* - microsatellites - simple sequence repeat - polymerase chain reaction

Microsatellites, or simple sequence repeats (SSR), are DNA sequences characterized by short (2-6 nucleotides) tandemly arranged repeats with a total length not exceeding 200 bp. They are highly polymorphic (Litty & Luty 1989, Weber & May 1989, Tautz 1989) and widely distributed in the eukaryote genome, therefore generating a unique genotypic profile that permits individual identification and relationship categorization (Beckman & Weber 1992, Goldstein et al. 1999). Consequently, assays of microsatellite loci have rapidly become established as a powerful tool for the analysis of mating systems and population structure (Queller et al. 1993). Microsatellites have been used for the study of the genetic structure of human populations (Santos et al. 1993, Santos & Tyler-Smith 1996), mammals (Ishibashi et al. 1997), birds (Gibbs et al. 1997), snails (Mavárez et al. 2000), fungi (Groppe et al. 1995), and numerous other species.

DNA amplification of microsatellites with locus non-specific primers were first employed in studies of species differentiation in fungal strains (Meyer et al. 1993). Such procedure, however, may produce a number of bands larger than expected, probably due to primer annealing at different positions within the repeats. Zietkiewicz et al. (1994) overcame this problem by anchoring CA-repeat primers at their 5' or 3' ends, reducing the number of possible annealing templates and increasing specificity. With this methodology the repeated regions and intervening sequences were targeted for amplification. The cited investigators named the technique SSR-PCR and obtained successful results with several species of eukaryotes, mainly with the (CA)$_8$RY primer, anchored at the 3'end, under high stringency conditions.

SSR-PCR has already been applied to studies on intraspecific variability of *Trypanosoma cruzi, Leishmania braziliensis* and *Schistosoma mansoni* (Oliveira et al. 1997) and to intra- and interspecific studies of snails of the genus *Biomphalaria* (Caldeira et al. 2001, Campos et al. 2002). However, a detailed characterization of the amplification targets has never been performed. Sequencing of the amplicons would be important for the exact determination of the amplified templates and the identification of microsatellite loci, without the need to generate microsatellite-enriched genomic libraries or a large number of sequences. The microsatellite loci identified in this manner could be checked for the presence of polymorphism and used for studies of population genetics, such as the dispersal processes in *B. glabrata* (Carvalho et al. 1998).

The objective of the present study was to identify SSR-PCR amplification targets by sequencing the amplified genomic DNA in order to determine the molecular templates to which the oligonucleotides anneal, using *B. glabrata* DNA.

MATERIALS AND METHODS

DNA extraction and PCR amplification - DNA was extracted from body parts, excluding the viscera, of *B. glabrata* from Pernambuco, Brazil, using the phenol-chloroform method as described by Simpson et al. (1982). SSR-PCR products were amplified using the protocol previously described by Caldeira et al. (2001).

Cloning and sequencing of PCR products - PCR products were cloned using the pCR-II TOPO kit and transformed into electrocompetent TOP10 *Escherichia coli*, according to manufacturer's instructions (Invitrogen). Twenty unique amplification products were cloned and sequenced from each PCR assay with primers CAA(CT)$_6$ and (CA)$_8$RY. Plasmid was isolated using the Quiagen plasmid miniprep kits according to the manufacturer's protocol. Sequencing was carried out using fluorescently labeled M13 reverse and forward universals primers fluorescently labelled and the Thermo Sequenase Kit (Amersham Pharmacia) in an ALF Automated Sequencer (Amersham Pharmacia). Sequences were named CT [amplified with CAA(CT)$_6$] or CA [amplified with (CA)$_8$RY], using 1 through 20 for clone number and the letters F or R for forward and reverse sequencing primer, respectively.

Sequence analysis - The sequences obtained were edited using the Bioedit software (http://jwbrown. mbio.ncsu.edu/BioEdit/bioedit.html). Blast N (www. ncbi.nlm.nih.gov/blast) was used for identity search. RepeatMasker (http://ftp.genome. washington.edu/cgi-bin/RepeatMasker) was used to detect the presence of microsatellites, and DNA tools (http://www.crc.dk/ dnatools) was used to search for complex repetitive sequences within clones.

Southern blot - A probe was constructed from clone 7CT using the PCR DIG probe synthesis kit according to manufacturer instructions (Roche Diagnostics). *B. glabrata* genomic DNA (50 ng) was digested with *Hind*III, *Pvu*II, *Xba*I and *Xho*I restriction enzymes (Life Technologies) and separated on 0.7% agarose gel. The digested DNA was transferred to a positively charged nylon membrane (Roche Diagnostics) by capillarity (Sambrook et al.

1989) and detected using the DIG Nucleic Acid Detection Kit (Roche Diagnostics) according to manufacturer's instructions.

RESULTS

PCR amplification using the CAA(CT)$_6$ and (CA)$_8$RY primers resulted in multiple bands varying in size from 250 to 800 bp (Fig. 1). We cloned 20 unique amplicons from each amplification reaction. No identities with the sequenced fragments were observed by BlastN analysis against the GenBank (accession number BH772738 to BH772817).

RepeatMasker analysis of the sequenced fragments identified 15 that contained simple sequence repeats elements (Table). We observed di-, tri-, tetra- and pentanucleotide repeats that were classified as perfect (only one repetitive motif without interruptions) or imperfect (repetitive motif intercalated with non-repetitive sequences). In all of the sequences obtained only the exact number of repeats designed in the primers was observed.

TABLE
Sequences containing intervening repeated regions

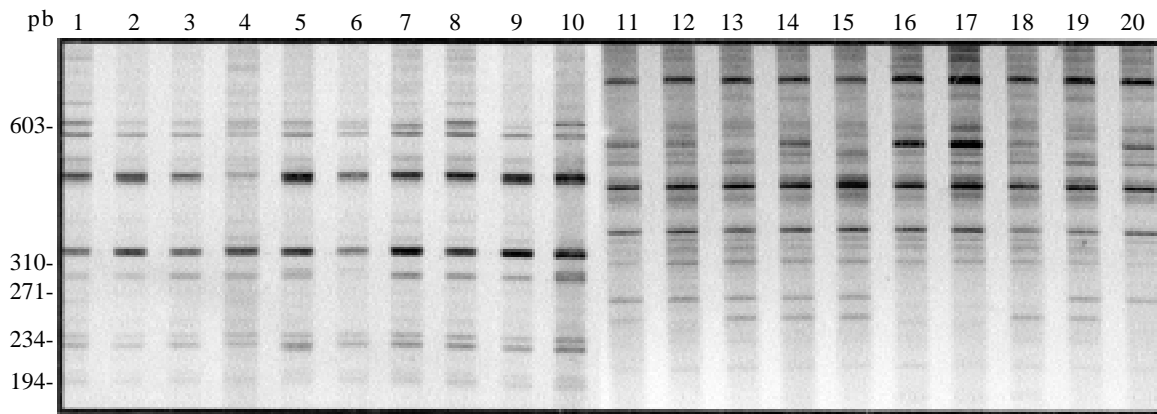| Sequence	Microsatellite-like	Classification
3-CAf	(CACG)$_4$	Perfect
3-CTf	(CGT)$_9$	Perfect
4-CAf	(GA)$_9$	Perfect
4-CAr	(CT)$_7$/ (CT)$_9$	Perfect
4-CTf	(GA)$_{12}$/ (GA)$_6$	Perfect
4-CTr	(CT)$_{14}$/ (CT)$_7$	Perfect
5-CTr	(CT)$_{10}$	Imperfect
6-CTr	(GT)$_{10}$	Imperfect
7-CTf	(CAT)$_5$	Perfect
9-CTr	(CATAG)$_5$/ (GT)$_6$	Perfect
11-CTf	(CGATA)$_4$	Perfect
12-CTf	(CA)$_6$	Perfect
15-CAf	(CA)$_{16}$	Perfect
16-CTr	(CT)$_7$	Perfect
20-CAr	(CG)$_{12}$	Perfect

Fig. 1: silver-stained 6% polyacrylamide gels showing simple sequence repeat-anchored polymerase chain reaction amplification profiles from ten *Biomphalaria glabrata* individuals. Lanes 1 to 10 show profiles obtained by utilizing the primer CAA(CT)$_6$ and lanes 11 to 20 profiles obtained by utilizing the primer (CA)$_8$RY. Molecular size markers are shown on the left of each gel.

Complex internal repetitive elements were identified in seven sequences of the CT clones after DNA tools analysis. The 7CT clone that contained the repeat (ATGTACTCC), also observed in six other amplified fragments, was used as a probe for Southern blot analysis for the verification of the distribution of the fragment in the *B. glabrata* genome. Only one band was obtained when this fragment was used as a probe (Fig. 2).
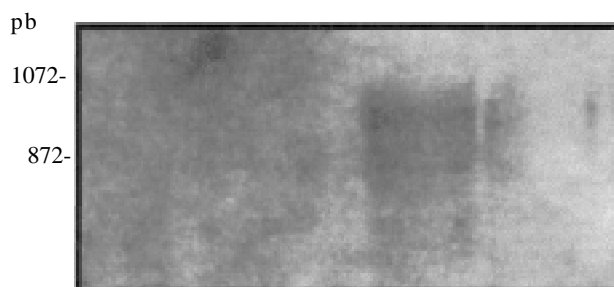

pb

1072-

872-

Fig. 2: Southern blot analysis of digestion *Biomphalaria glabrata* DNA probed with clone 7CT.

DISCUSSION

Inter-repeat region amplification by PCR results in different banding patterns that can be compared among individuals. Thus, SSR-anchored PCR methods have been applied to differentiate various animal and plant species (Zietkiewicz et al. 1994, Godwin et al. 1997). In the current study we sequenced products generated by this methodology with the commonly used primers $(CA)_8RY$ and $CAA(CT)_6$, in an attempt to characterize the amplified genomic regions. The results observed at the beginning or at end of the clones were expected for the $(CA)_8RY$ primer, since it would anneal at the 3' end of the microsatellite. However, primer $CAA(CT)_6$ was designed to anneal at the 5' end of the repeated sequence (Wu et al. 1994), anchored by the CAA nucleotides. This primer should amplify complete microsatellite regions. We observed that of the 40 sequences (20 forward and 20 reverse sequences), only 11 contained microsatellite-like repeats, and only 4 contained the CT repeat. Therefore, this primer annealed primarily to sequences other than microsatellites containing CT repeats.

It appears that the SSR-PCR technique may work in a manner similar to randomly amplified polymorphic DNA (RAPD), amplifying disperse regions in the genome and not specifically microsatellites, as expected. However, as the annealing temperature is higher, the technique shows increased specificity for the template and reproducibility (Oliveira et al. 1997). Indeed, SSR-PCR has been found to be more efficient and reliable than AP-PCR (Oliveira et al. 1997, Gomes et al. 1998). Yang et al. (2000) used SSR-PCR to differentiate three *Metagonimus* species parasitizing fish and observed that more polymorphic bands were present when anchored SSR-PCR was used compared to AP-PCR (Yu et al. 1997a) and PCR-RFLP (Yu et al. 1997b). Caldeira et al. (2001), used anchored SSR-PCR to study the genetic variability of the *B. straminea* complex. The genetic similarity coefficients obtained by these authors were similar to those obtained by Campos et al. (2002) for

B. glabrata with anchored SSR-PCR, by Vidigal et al. (1994) and Langand et al. (1999) using AP-PCR and by Mulvey and Vrijenhoek (1982) using isoenzymes.

It was very interesting to observe that some of the clones contained another microsatellite locus between the primer annealing sites. Species that do not have many sequences deposited in GenBank may benefit from the sequencing of SSR-PCR-generated amplicons for the identification of microsatellite loci (Oliveira & Johnston 2001). This would considerably lower the cost of constructing microsatellite-enriched genomic libraries or generating a large number of sequences for the identification of repetitive sequences (Gibbs et al. 1997).

In conclusion, the SSR-PCR methodology amplifies genomic fragments, but not exclusively anchored at microsatellite sequences. Therefore, the premise of this technique is not supported by this study. The results obtained by SSR-PCR should not be interpreted as the amplification of microsatellite loci, and analytical rules similar to those for RAPD or AP-PCR should be used. This methodology can also be useful for the identification of microsatellites, since there were repetitive sequences within a number of sequenced fragments.

REFERENCES

Beckman JS, Weber JL 1992. Survey of human and rat microsatellites. *Genomics 12*: 627-631.

Caldeira RL, Vidigal THDA, Simpson AJG, Carvalho OS 2001. Genetic variability in Brazilian populations of *Biomphalaria straminea* complex detected by simple repeat anchored polymerase chain reaction amplification. *Mem Inst Oswaldo Cruz 96*: 535-544.

Campos YR, Carvalho OS, Goveia CO, Romanha AJ 2002. Genetic variability of the main intermediate host of the *Schistosoma mansoni* in Brazil, *Biomphalaria glabrata* (Gastropoda: Planorbidae) assessed by SSR-PCR. *Acta Trop 83*: 19-27.

Carvalho OS, Nunes IM, Caldeira RL 1998. First report of *Biomphalaria glabrata* in the State of Rio Grande do Sul (Brazil). *Mem Inst Oswaldo Cruz 93*: 39-40.

Gibbs M, Dawson DA, McCamley C, Wardle AF, Armour JA, Burke T 1997. Chicken microsatellite markers isolated from libraries enriched for simple tandem repeats. *Anim Genet 28*: 401-417.

Godwin ID, Aitken EAB, Smith DW 1997. Application of inter simple sequence repeat (ISSR) markers to plant genetics. *Eletrophoresis 18*: 1524-1528.

Goldstein DB, Roemer GW, Smith DA, Reich DE, Bergman A, Wayne RK 1999. The use of microsatellite variation to infer population structure and demographic history in a natural model system. *Genetics 151*: 797-801.

Gomes ML, Macedo AM, Pena SDJ, Chiari E 1998. Genetic relationships between *Trypanosoma cruzi* strains isolated from chronic chagasic patients in southern Brazil as revealed by RAPD and SSR-PCR analysis. *Acta Trop 69*: 99-109.

Groppe K, Sanders I, Wiemken A, Boller T 1995. A microsatellite marker for studying the ecology and diversity of fungal endophytes (epichloe spp) in grasses. *Appl Environ Microbiol 61*: 3943-3949.

Ishibashi Y, Saitoh T, Abe S, Yoshida MC 1997. Sex-related spatial kin structure in a spring population of grey-sided voles *Clethrionomys rufocanus* as revealed by mitochondrial and microsatellite DNA analyses. *Mol Ecol 6*: 63-71.

Langand J, Theron A, Pointier JP, Delay B, Jourdane J 1999. Population structure of *Biomphalaria glabrata*, intermediate snail host of *Schistosoma mansoni* in Guadeloupe Island, using RAPD markers. *J Mol Stud 65*: 425-433.

Litt M, Luty JA 1989. A hypervariable microsatellite revealed by *in vitro* amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Am J Hum Genet 44*: 397-401.

Mavárez J, Amarista M, Pointier JP, Jarne P 2000. Microsatellite variation in the freshwater schistosome-transmitting snail *Biomphalaria glabrata*. *Mol Ecol 9*: 1009-1011.

Meyer W, Lieckfeldt E, Kuhls K, Freedman EZ, Borner T, Mitchell TG 1993. DNA and PCR fingerprinting in fungi. In SDJ Pena, R Chakraborty, JT Epplen, AJ Jeffreys (eds), *DNA Fingerprinting: State of the Science*, Birkhauser, Verlag, p. 311-320.

Mulvey M, Vrijenhoek RC 1982. Population structure in *Biomphalaria glabrata*: examination of hypothesis for the patchy distribution of susceptibility to schistosomes. *Am J Trop Med Hyg 27*: 782-786.

Oliveira GC, Johnston D 2001. Mining the schistosome DNA sequence database. *Trends Parasitol 17*: 501-503.

Oliveira RP, Macedo AM, Chiari E, Pena SDJ 1997. An alternative approach to evaluating the intraspecific genetic variability of parasites. *Parasitol Today 5*: 196-200.

Queller DC, Strassmann JE, Hughes CR 1993. Microsatellites and kinship. *Trends Ecol Evol 8*: 285-288.

Sambrook J, Fritsch EF, Maniatis T 1989. Analysis and cloning of eukaryotic genomic DNA. In *Molecular Cloning – A Laboratory Manual,* 2nd ed., Cold Spring Harbor Press, USA, p. 9.34-9.35

Santos FR, Pena SDJ, Epplen TJ 1993. Genetic population study of a Y-linked tetranucleotide repeat DNA polymorphism. *Hum Genet 90*: 655-656.

Santos FR, Tyler-Smith C. 1996. Reading the human Y chromosome, the emerging DNA markers and human genetic history. Brazil *J Genet 19*: 665-670.

Simpson AJG, Sher A, McCutchan TF 1982. The genome of *Schistosoma mansoni*: its size, bases and repetitive sequences. *Mol Biochem Parasitol 6*: 125-137.

Southern EM 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol 98*: 503-517.

Tautz D 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Res 17*: 6463-6471.

Vidigal THDA, Dias Neto E, Carvalho OS, Simpson AJG 1994. *Biomphalaria glabrata*: extensive genetic variation in Brazilian isolates by random amplified polymorphic DNA analysis. *Exp Parasitol 79*: 187-194.

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

Wu KS, Jones R, Danneberger L, Scolnik PA 1994. Detection of microsatellite polymorphisms without cloning. *Nucleic Acids Res 22*: 3257-3258.

Yang HJ, Guk SM, Han ET, Chai JY 2000. Molecular differentiation of three species of *Metagonimus* by simple sequence repeat anchored polymerase chain reaction (SSR-PCR) amplification. *J Parasitol 86*: 1170-1172.

Yu JR, Chung JS, Chal, JY 1997a. Different RAPD patterns between *Metagonimus yokogawai* and *Metagonimus* Miyata type. *Korean J Parasitol 35*: 295-298.

Yu JR, Chung JS, Huh S, Chal, JY, Lee SH 1997b. PCR-RFLP patterns of the three kinds of *Metagonimus* in Korea. *Korean J Parasitol 35*: 271-276.

Zietkiewicz E, Rafalski A, Labuda D 1994. Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. *Genomics 20*: 176-193.