CA88, a nuclear repetitive DNA sequence identified in *Schistosoma mansoni,* aids in the genotyping of nine *Schistosoma* species of medical and veterinary importance

Diana Bahia[1,2]/+, Nilton B Rodrigues[2,5], Flávio Marcos G Araújo[2], Álvaro José Romanha[2], Jerônimo C Ruiz[2], David A Johnston[3], Guilherme Oliveira[2,4]

[1]Departamento de Microbiologia, Imunologia e Parasitologia, Escola Paulista de Medicina, Universidade Federal de São Paulo, Rua Botucatu 862 6º andar, 04023-062 São Paulo, SP, Brasil [2]Instituto de Pesquisas René Rachou-Fiocruz, Belo Horizonte, MG, Brasil [3]Department of Zoology, The Natural History Museum, London, UK [4]Programa de Pós-Graduação e Pesquisa, Santa Casa de Belo Horizonte, Belo Horizonte, MG, Brasil, [5]Laboratório de Pesquisas Clínicas, Escola de Farmácia, Universidade Federal de Ouro Preto, Ouro Preto, MG, Brasil

CA88 is the first long nuclear repetitive DNA sequence identified in the blood fluke, Schistosoma mansoni. *The assembled* S. mansoni *sequence, which contains the CA88 repeat, has 8,887 nucleotides and at least three repeat units of approximately 360 bp. In addition, CA88 also possesses an internal CA microsatellite, identified as SmBr18. Both PCR and BLAST analysis have been used to analyse and confirm the CA88 sequence in other* S. mansoni *sequences in the public database. PCR-acquired nuclear repetitive DNA sequence profiles from nine* Schistosoma *species were used to classify this organism into four genotypes. Included among the nine species analysed were five sequences of both African and Asian lineages that are known to infect humans. Within these genotypes, three of them refer to recognised species groups. A panel of four microsatellite loci, including SmBr18 and three previously published loci, has been used to characterise the nine* Schistosoma *species. Each species has been identified and classified based on its CA88 DNA fingerprint profile. Furthermore, microsatellite sequences and intra-specific variation have also been observed within the nine* Schistosoma *species sequences. Taken together, these results support the use of these markers in studying the population dynamics of* Schistosoma *isolates from endemic areas and also provide new methods for investigating the relationships between different populations of parasites. In addition, these data also indicate that* Schistosoma magrebowiei *is not a sister taxon to* Schistosoma mattheei, *prompting a new designation to a basal clade.*

Key words: genotypes - microsatellite - *Schistosoma* - species - population dynamics

Micro (2-6 bp) and minisatellites (8 to more than 100 bp) are tandem repeated DNA sequences (Hamada et al. 1982, Tautz & Renz 1984). Both types of DNA repeats are abundant, occur randomly in coding and non-coding regions throughout eukaryotic genomes (MacLeod 2004, Oliveira et al. 2006) and display high levels of mutation (Jarne & Lagoda 1996, Armour et al. 1999). As a result, micro and minisatellites serve as excellent tools for high resolution molecular fingerprinting to type both individuals (Hagelberg et al. 1991) and populations (Jarne & Lagoda 1996). They can also be used to construct genetic maps and to identify loci involved in genetic diseases (Dietrich et al. 1996). Furthermore, micro and minisatellite based probes have become key elements to distinguish individual eukaryotic parasites such as *Plasmodium* (Collins et al. 2000), *Trypanosoma brucei* (MacLeod et al. 2000), *Trypanosoma cruzi* (Macedo et al. 1992) and *Theileria parva* (Oura et al. 2003).

Financial support: FAPESP (JP3 2007-50551-2, 07/54831-0), WHO/TDR (A60251), NIH-Fogarty (TW007358-01)
+ Corresponding author: dianabahia@hotmail.com
Received 17 December 2008
Accepted 10 February 2010

Given that a significant level of intra-specific variation has been detected within *Schistosoma mansoni* populations (Rodrigues et al. 2002b), there has been a growing interest in analysing how that variation is partitioned within and flows between natural populations. Polymorphic DNA sequences have been a valuable tool in determining the distribution of schistosome genotypes among intermediate hosts (Minchella et al. 1994, Dabo et al. 1997, Sire et al. 2001, Eppert et al. 2002, Liu et al. 2006). Furthermore, these DNA sequences are also useful in the examination of *S. mansoni* population structure and subdivision within definitive hosts (Rodrigues et al. 2002b).

Several microsatellite markers in *S. mansoni* have been identified and developed. These markers have proven to be highly useful in genetic and population studies of *Schistosoma* spp (Scharf et al. 1998, Durand et al. 2000, Blair et al. 2001, Prugnolle et al. 2002, Rodrigues et al. 2002a, b, 2007). It is worth mentioning that single nucleotide polymorphisms in *S. mansoni* have also been described (Simões et al. 2007). Currently, only two tandemly repeated arrays have been reported for *Schistosoma* spp.

F21 is a 62-bp tandemly repeated array in *S. mansoni* (Pena et al. 1995), while DraI is a 121-bp tandem repeat in *Schistosoma haematobium* (Hamburger et al. 2001). However, the F21 sequence appears to be unique to *S. mansoni* (DA Johnston, unpublished observations),

which limits its application to the *Schistosoma* species. Furthermore, DraI exhibits cross-hybridisation signals with DNA from *Schistosoma magrebowiei, Schistosoma bovis, Schistosoma mattheei, Schistosoma intercalatum* and *Schistosoma curassoni*. Therefore, its use is restricted to studies that do not include these species (Hamburger et al. 2001).

This is the first report of a cross-species *Schistosoma* long nuclear repetitive DNA sequence, identified as CA88. This DNA repeat is likely to be present in the genomes of all nine *Schistosoma* species examined, from both African and Asian lineages, which are of medical and veterinary importance. The CA88 profiles obtained with specific primers are used to classify the *Schistosoma* species according to four genotypes.

In addition to CA88, a panel of four microsatellite loci has been used in the characterisation of the nine *Schistosoma* species. The use of CA88 nuclear repetitive DNA sequence in conjunction with microsatellite markers resulted in the identification of inter and intra-specific differences across *Schistosoma* species. Therefore, these two elements can be used to identify and classify these species on the basis of a DNA fingerprint profile.

MATERIALS AND METHODS

Schistosoma species - Adult worms from the following nine different *Schistosoma* species, which represent the three main species groups as well as both African and Asian lineages, were used in this study: (i) *S. haematobium* (NHM-3388, from Mali), *S. magrebowiei* (NHM-2623, from Zambia), *S. mattheei* (NHM-2763, from Zambia), *S. bovis* (NHM-1289, from Kenya), *Schistosoma guineensis* (NHM-1624, from São Tome) and *S. curassoni* (NHM-C1, from Senegal), from the terminally spined African (*S. haematobium*) group, (ii) *S. mansoni* (LE, from Brazil) and *Schistosoma rodhaini* (NHM-303, from Burundi) from the laterally spined African (*S. mansoni*) group and (iii) *Schistosoma japonicum* (NHM-1314, from China) from the Asian (*S. japonicum*) group. To minimise the effects of genetic selection in laboratory hosts, the worms were collected, when possible, either directly from the field or from recent field isolates maintained for a couple of passages in the laboratory. Worms were recovered from laboratory hosts by cardiac perfusion with citrate saline and stored in liquid nitrogen. Worms were thawed into absolute ethanol for transport.

DNA extraction - DNA was extracted from 10 individual worms of each species. Worms were washed in phosphate buffered saline, placed individually into 1.5 mL microcentrifuge tubes and stored at -20°C until DNA extraction. Worms were homogenised by five cycles of pulverisation in a dry ice bath followed by thawing in a 37°C water bath. The homogenate was resuspended in 50 µL of 50 mM Tris-HCl pH 8.0, 100 mM EDTA, 100 mM NaCl and 0.5% sodium dodecyl sulfate (Grossman et al. 1990) and incubated in the presence of 20 µg/mL proteinase K, at 37°C overnight. Standard phenol/chloroform extraction and ethanol precipitation (Sambrook et al. 1989) were used to purify and recover DNA from the lysate. DNA was resuspended in 30 µL of 10 mM Tris-HCl, 1 mM EDTA, pH 8.0 (TE).

Bioinformatics tools - BLAST (Altschul et al. 1990) and FASTA (Lipman & Pearson 1985) were used to search the CA88 sequence against different databases (DB), including the non-redundant protein DB, core nucleotide and EST DBs at National Center for Biotechnology Information (NCBI), The Institute for Genomic Research and *S. mansoni* geneDB version 3.1 (http://www.genedb.org/genedb/smansoni/), hosted by the Sanger Institute. Low-complexity filters in the search algorithms were turned off to prevent the masking of low-complexity segments of sequences and repeats. Tandem repeat DNA regions were identified using Tandem Repeats Finder software (Benson 1999).

Annotation and graphical analysis of the CA88 repetitive DNA sequences were performed using a local installation of Artemis (http://www.sanger.ac.uk/Software/Artemis/) with custom PERL scripts to analyse and format the results and also provide local copies of the relevant DB.

Microsatellite analyses - Microsatellite analyses were performed by PCR using 1 ng of template DNA, 0.75 U of Taq DNA polymerase (Invitrogen), 1X PCR buffer (1.5 mM $MgCl_2$, 50 mM KCl, 10 mM Tris/HCl, pH 8.3), 10 µM of each primer and 200 µM of each dNTP. PCR amplifications were carried out in a Perkin Elmer model 9600 thermal cycler in a final reaction volume of 10 µL. The cycling conditions were: denaturation at 95°C for 3 min, followed by 35 cycles of denaturation at 95°C for 45 s, annealing at primer dependent temperature for 1 min (Table I) and extension at 72°C for 30 s. PCR primers for microsatellite amplifications were designed using the software FastPCR (Kalendar 2003). One primer of each pair was labelled at the 5' end with fluorescein to allow allele scoring using the AlleleLinks software package in an ALF automated sequencer (Amersham), as previously described (Rodrigues et al. 2002a, 2007). Amplification products (3 µL) were visualised by electrophoresis on 8% polyacrylamide gels (PAGE), followed by silver staining (Sanguinetti et al. 1994).

RESULTS

SmBr18 is a microsatellite within CA88, a long nuclear DNA repetitive sequence in S. mansoni - To further characterise CA88, we have carried out several *in silico* analyses. A 364 bp DNA sequence from *S. mansoni*, named SmBR18 (GenBank accession DQ137590.1), previously identified from an *S. mansoni* (CA) enriched library (Rodrigues et al. 2007), was used in BLAST analyses (Altschul et al. 1990) against the NCBI (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi) and SANGER (http://www.sanger.ac.uk) DBs. The following sequences, DQ137585.1, DQ137520.1, DQ137504.1, DQ137567.1, DQ137605.1, DQ137539.1, DQ137537.1, DQ137526.1, DQ137489.1, DQ137525.1, DQ137461.1 and DQ137466.1 displayed similarities with the SmBr18 sequence. These sequences were clustered using CAP3 (Huang & Madan 1999), forming a 1,980 bp consensus sequence that was designated as CA88.

Initially, CA88 was compared to the current *S. mansoni* genome assembly from GeneDB version 3.1. Three supercontigs (Smp_scaff000011, Smp_scaff000047 and

TABLE I

Characteristics of the analyzed *Schistosoma mansoni* microsatellite loci

Locus	Accession	Repeat	Primer sequences (5'-3')[a]	Approximate amplicon size	Annealing temperature (ºC)	Reference
SmBr5	L25065	(ATT)n	R: GAATTACTGTCCCTTTATCTC F: AAACTATTCATTACTGTCGGG	328 bp	58	Rodrigues et al. (2002a)
SmBr6	AF009659	(CTT)n	R: CTTAACAGACATACACGC F: GAATACAGGCTATAATCTACA	265 bp	55	Rodrigues et al. (2002a)
SmBr9	DQ137431	(TTA)n	R: ATTGGCGTCAGTAGAAGAGATT F: ATTCACCCATTGTCTTAAAACC	161 bp	60	Rodrigues et al. (2007)
SmBr18	DQ137590.1	(CA)n	R: TTTTCTGTCTACATGTTGATGAAG F: TAACCATCATTCACCAAACATTC	141 bp	60	This paper

a: the reverse primer of each pair was 5' labelled with fluorescein for automated genotyping analysis.

Smp_scaff000068) were identified as containing sequences similar to DQ137590.1 and therefore were further examined. In supercontig Smp_scaff000011, the DQ137590.1 element is located within intron 6 of the predicted gene Smp_scaff000011_0570 as a consensus repeat of 355 bp and shares 95% identity between the 3.7 copies present in the region. In supercontig Smp_scaff000068, the DQ137590.1 element is located at the end of the supercontig and was characterised as a consensus repeat of 358 bp, which shares 98% identity between the 2.1 copies present in the region. In supercontig Smp_scaff000047, only small fragments of the DQ137590.1 element were identified (Fig. 1, Supplementary data).

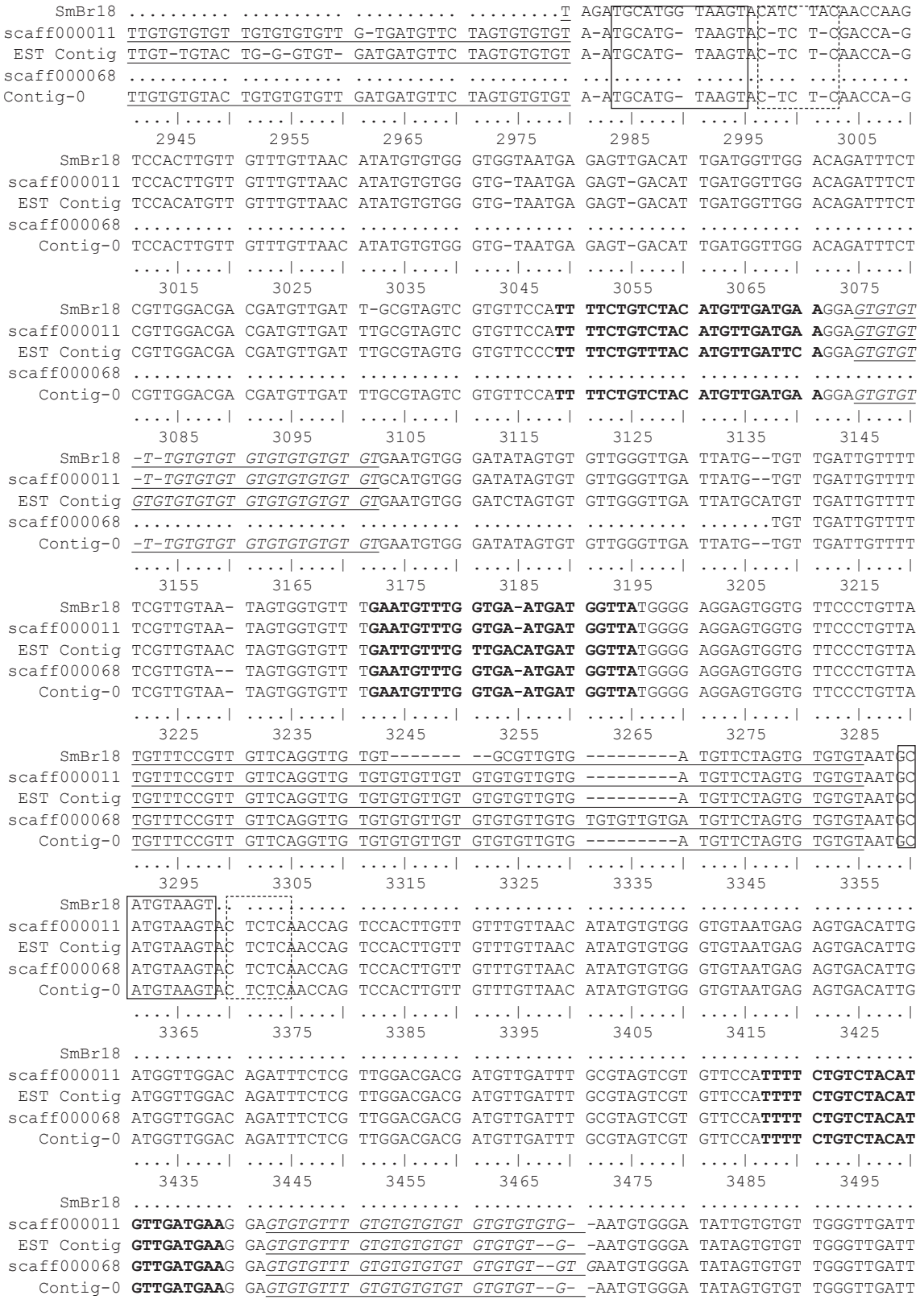
The consensus sequence assembled from database sequences matching the repeated regions of these supercontigs is 8,887 bp long and contains approximately four tandem copies of the CA88 repetitive sequence (Fig. 1, Supplementary data). The CA88 sequence contains two repeated regions consisting of a (CA) microsatellite that varies from 7-14 dinucleotide repeats and a CA rich region (Fig. 1, Supplementary data).

SmBr18 is a microsatellite repeat and is conserved among Schistosoma species - SmBr18 PCR primers were designed to amplify a fragment of approximately 145 bp that spans the (CA) microsatellite repeat of the CA88 unit. The primers were tested in PCR using DNA samples from 10 individuals of each of the nine *Schistosoma* spp. We observed band patterns in the PAGE gels consistent with the presence of more than one amplification site in each species tested. These data suggest that the SmBr18 microsatellite is present in the CA88 regions of all *Schistosoma* spp (Fig. 2A, B).

Microsatellite variability among nine Schistosoma species - In addition to SmBr18, three other microsatellite loci were also selected and used in this analysis (Table I). These markers are identified as SmBr5, SmBr6 (Rodrigues et al. 2002a) and SmBr9 (Rodrigues et al. 2007). These four microsatellite loci were designed from *S. mansoni* sequences and tested on DNA extracted from individual worms of nine different *Schistosoma* species (Table I). The microsatellites SmBr5, SmBr6 and SmBr9, which have been previously described for *S. mansoni*, exhibited a high number of alleles across the species. Furthermore, all of the species generated amplification products for the new SmBr18 locus. Notably, only the SmBr18 locus produced amplification products for *S. guineensis* and *S. curassoni* (Table II, Fig. 3, Supplementary data). The data presented in Fig. 2 demonstrate that the microsatellite locus SmBr18 generated a complex, multi-band profile for each species examined (Fig. 2A, B).

The multi-band profiles of *Schistosoma* species, which were obtained by PCR with SmBr18 primers, were identified by PAGE analysis (Fig. 2A, B) and analysed for polymorphisms based on the presence or absence of specific bands (Zalloum et al. 2005). The fragment sizes were assigned using the AlleleLinks software package in an ALF automated sequencer (Amersham). The profiles obtained are reproducible in separate assays. Four separate genotypes were identified (designated I-IV) (Fig. 2A, B, Table I). Three genotypes (I, II and IV) reflected the traditional species group to which the species are assigned (Rollinson & Southgate 1987, Littlewood & Johnston 1995, Barker & Blair 1996, Webster et al. 2006).

Genotype I contained the "African", lateral-spined egg group species *S. mansoni* and *S. rodhaini*. These two species generated identical profiles with bands of approximately 141, 158, 210 and 240 bp, in addition to bands of 400 bp or more. Genotype II comprised the "African" terminal-spined egg group species including *S. bovis, S. curassoni, S. guineensis, S. haematobium* and *S. mattheei*. The profiles consisted of bands of approximately 158, 280, 520 and 750 bp. Genotype IV, which comprises the "Asian" schistosome, *S. japonicum,* shared bands of 141 and 158 bp with genotype I, but also exhibited additional bands of approximately



☐ CA88 minisatellite begin, ⬚ CA88 minisatellite end, *SmBr18 microsatellite*, GT rich region, **forward and reverse primers**

Fig. 1: fragment of the CA88 nuclear repetitive sequence (for complete sequence see Supplementary data).

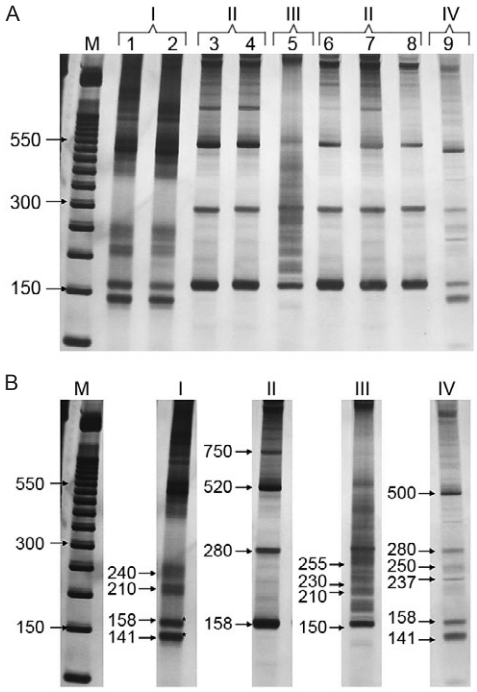


Fig. 2: polyacrylamide gels (PAGE) profiles of SmBr18 amplification products from nine *Schistosoma* species. Primers flanking SmBr18, derived from the sequence of CA88 (ACC # DQ137431), were used to amplify the DNA of individual *Schistosoma* species. The distinct profiles of bands generated by this assay were visualized by PAGE analysis. Different genotypes have been assigned to *Schistosoma* species. The band sizes were determined by using specific markers on PAGE gels. A: M: 50 bp DNA ladder. Lane 1: *Schistosoma mansoni*; 2: *Schistosoma rodhaini*; 3: *Schistosoma haematobium*; 4: *Schistosoma guineensis*; 5: *Schistosoma magrebowiei*; 6: *Schistosoma bovis*; 7: *Schistosoma curassoni*; 8: *Schistosoma mattheei*; 9: *Schistosoma japonicum*; B: assigned genotypes (I - IV) are indicated above the lanes. M: 50 bp DNA ladder. The band sizes are indicated on the left.

237, 280 and 500 bp. *S. magrebowiei*, an "African" terminal-spined egg group species, displayed a unique genotype (genotype III) (Fig. 2A, B). Genotype III, although broadly similar to that of the other members of its species group (i.e., shares a band of approximately 158 bp and lacks the smallest band seen in genotypes I and IV), showed a distinct ladder of bands ranging between 158-280 bp. The four genotypes were consistently obtained in different amplification experiments from the same individual worms and also between amplifications of different worms.

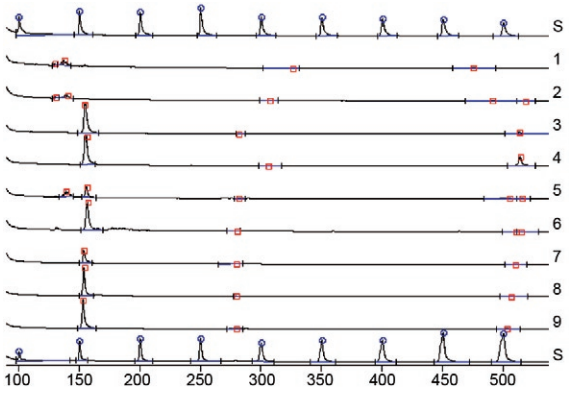


Fig. 3: automated genotype analysis of the microsatellite locus SmBr18 in nine individual worms from different *Schistosoma* species: S: 50-500 bp reference ladder. Line 1: *Schistosoma rodhaini*; 2: *Schistosoma mansoni*; 3: *Schistosoma haematobium*; 4: *Schistosoma guineensis*; 5: *Schistosoma japonicum*; 6: *Schistosoma magrebowiei*; 7: *Schistosoma curassoni*; 8: *Schistosoma bovis*; 9: *Schistosoma mattheei*. Squares indicate major amplification products of each CA88 sub-unit (compare with Fig. 2, text).

TABLE II

Microsatellite profiles for *Schistosoma* species from four different loci

Species	SmBr5 (ATT)n alleles (bp)	SmBr6 (CTT)n alleles (bp)	SmBr9 (TTA)n alleles (bp)	SmBr18 (CA)n alleles (bp)	Genotype assigned[a]
haematobium group					
Schistosoma bovis	445, 448, 457	-	146	151, 153, 155	II
Schistosoma curassoni	-	-	-	151, 153	II
Schistosoma guineensis	-	-	-	151, 153, 155, 157	II
Schistosoma haematobium	448	257	149, 152	151, 155, 157	II
Schistosoma magrebowiei	445, 448, 451	254, 257, 260	146, 149, 152	127, 129, 151, 153	III
Schistosoma mattheei	448, 451, 460, 463, 466	-	158	151, 153	II
mansoni group					
Schistosoma mansoni	451, 454, 457, 463,475,478, 481	254, 257, 260, 263, 266, 272	146, 149, 152, 158	127, 135, 137, 139	I
Schistosoma rodhaini	445	263	146, 152	127, 135, 137, 153	I
japonicum group					
Schistosoma japonicum	-	254, 260	-	151, 153, 155	IV

a: these genotypes were assigned based only on the size of the amplified SmBr18 fragments; -: no amplification achieved.

Table II demonstrates that CA88 genotypes, together with a microsatellite panel, can be used to identify and classify each species according to a DNA fingerprint profile. Furthermore, intra-specific variation has also been observed in these studies (data not shown).

DISCUSSION

CA88 is the first long nuclear repetitive DNA sequence to be reported for *S. mansoni*. It appears to occur in all nine *Schistosoma* species examined in this study, including the five species that infect humans. These nine *Schistosoma* species represent the three main species groups in both African and Asian parasite lineages. In *S. mansoni*, for which a genomic sequence is available, at least 3.7 copies of CA88 occur within the introns of predicted genes.

Amplification profiles generated by primers targeting a SmBr18 microsatellite within the CA88 long DNA repetitive sequence are used to classify the nine *Schistosoma* species into four genotypes. Three of the genotypes correspond to the traditional species groups that are delineated by egg morphology, geographic distribution and intermediate host specificity (Rollinson & Southgate 1987). Apart from the alignment of the CA88 sequence with contigs of the current *S. mansoni* genome assembly, no link has yet been established between this nuclear repetitive DNA sequence and a specific gene.

In the same fashion as DraI, a 121 bp tandem repeat from *S. haematobium* (Hamburger et al. 2001), CA88 appears to be an unusual nuclear repetitive DNA sequence that cannot be classified as a minisatellite. Presently, the 8,887 bp consensus sequence flanking CA88 contains at least three repeat units of approximately 360 bp. Furthermore, the 8,887 bp consensus sequence was produced by an alignment of sequences from the analysed DB, followed by tandem arrangement of the repeats.

The terminal-spined species group includes three species that are human parasites: *S. haematobium*, *S. intercalatum* and *S. guineensis,* the latter previously having been assigned as a sub-population of *S. intercalatum* (Webster et al. 2006). This group also comprises five species that parasitise wild and/or domestic (artiodactyls) animals: *S. bovis*, *S. curassoni*, *Schistosoma leiperi*, *S. magrebowiei* and *S. mattheei*. Therefore, this group can be considered of medical, veterinary and economic importance. Compared to the other members of the species group examined in this study, *S. magrebowiei* displays a unique genotype.

Recent multigene phylogenetic analyses, which utilised entire nuclear ssrDNA, lsrDNA and partial mitochondrial COX1 sequences, have improved our phylogenetic understanding of the species group. These studies have resulted in the relocation of *S. magrebowiei* from its original identification as the sister taxon of *S. mattheei* [following analysis of COX1 sequences (Kane et al. 2003)], to its new position as the basal clade of the species group (Webster et al. 2006). Undoubtedly, *S. magrebowiei* may have retained genomic characteristics that separate it from the more divergent members of the species group, as is suggested by its distinct genotype. In support of this hypothesis, *S. magrebowiei* is predominantly a parasite of wild antelopes (lechwe, puku and waterbuck), rather than domestic stock (Christensen et al. 1983). This organism also does not mate with other members of the species group, either in the wild or in the laboratory (which is the case for schistosomes) (Webster et al. 2006). *S. magrebowiei* also has a much larger adult body size and higher reproductive rate relative to other members of this group, as it is the largest of the *Schistosoma* species (Rollinson & Southgate 1987).

Although the markers used in these studies can discriminate species within the *Schistosoma* genus, they can also serve as tools in population genetic studies of a species. Thus, as suggested by Yin et al. (2008), it is important to pay attention to somatic mutations in schistosome microsatellites that occur during asexual reproduction in the intermediate host snail. If these mutations occur regularly and in real time (i.e., at every generation of schistosome infection in snails), then it is likely that they have the capacity to confound or at least complicate, microsatellite-mediated population genetic studies of schistosomes.

Taken together, these results suggest that *S. magrebowiei* is not a sister taxon to *S. mattheei* and therefore supports designating *S. magrebowiei* to a basal clade. This study may provide an additional tool for the analysis of schistosome phylogenetics. Future efforts should address this issue by means of phylogenetic analysis with an outgroup.

REFERENCES

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ 1990. Basic local alignment search tool. *J Mol Biol 215*: 403-410.

Armour JA, Brinkworth MH, Kamischke A 1999. Direct analysis by small-pool PCR of MS205 minisatellite mutation rates in sperm after mutagenic therapies. *Mutat Res 445*: 73-80.

Barker SC, Blair D 1996. Molecular phylogeny of *Schistosoma* species supports traditional groupings within the genus. *J Parasitol 82*: 292-298.

Benson G 1999. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res 27*: 573-580.

Blair L, Webster JP, Barker GC 2001. Isolation and characterization of polymorphic microsatellite markers in *Schistosoma mansoni* from Africa. *Mol Ecol Notes 1*: 93-95.

Christensen NO, Mutani A, Frandsen F 1983. A review of the biology and transmission ecology of African bovine species of the genus *Schistosoma*. *Z Parasitenkd 69*: 551-570.

Collins HE, Li H, Inda SE, Anderson J, Laiho K, Tuomilehto J, Seldin MF 2000. A simple and accurate method for determination of microsatellite total allele content differences between DNA pools. *Hum Genet 106*: 218-226.

Dabo A, Durand P, Morand S, Diakite M, Langand J, Imbert-Establet D, Doumbo O, Jourdane J 1997. Distribution and genetic diversity of *Schistosoma haematobium* within its bulinid intermediate hosts in Mali. *Acta Trop 66*: 15-26.

Dietrich WF, Miller J, Steen R, Merchant MA, Damron-Boles D, Husain Z, Dredge R, Daly MJ, Ingalls KA, O'Connor TJ 1996. A comprehensive genetic map of the mouse genome. *Nature 380*: 149-152.

Durand P, Sire C, Théron A 2000. Isolation of microsatellite markers in the digenetic trematode *Schistosoma mansoni* from Guadeloupe island. *Mol Ecol 9*: 997-998.

Eppert A, Lewis FA, Grzywacz C, Coura-Filho P, Caldas I, Minchella DJ 2002. Distribution of schistosome infections in molluscan hosts at different levels of parasite prevalence. *J Parasitol 88*: 232-236.

Grossman Z, Ram D, Markovics A, Tarrab-Hazdai R, Lantner F, Ziv E, Schechter I 1990. *Schistosoma mansoni*: stage-specific expression of muscle-specific genes. *Exp Parasitol 70*: 62-71.

Hagelberg E, Gray IC, Jeffreys AJ 1991. Identification of the skeletal remains of a murder victim by DNA analysis. *Nature 352*: 427-429.

Hamada H, Petrino MG, Kakunaga T 1982. A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes. *Proc Natl Acad Sci USA 79*: 6465-6469.

Hamburger J, He-Na, Abbasi I, Ramzy RM, Jourdane J, Ruppel A 2001. Polymerase chain reaction assay based on a highly repeated sequence of *Schistosoma haematobium*: a potential tool for monitoring schistosome-infested water. *Am J Trop Med Hyg 65*: 907-911.

Huang X, Madan A 1999. CAP3: A DNA sequence assembly program. *Genome Res 9*: 868-877.

Jarne P, Lagoda PJ 1996. Microsatellites, from molecules to populations and back. *Trends Ecol Evol 11*: 424-429.

Kalendar R 2003. Fast PCR: a program for fast design PCR primer, DNA and protein manipulation (version 2.5.59). Institute of Biotechnology, University of Helsinki.

Kane RA, Southgate VR, Rollinson D, Littlewood DT, Lockyer AE, Pagès JR, Tchuem Tchuentè LA, Jourdane J 2003. A phylogeny based on three mitochondrial genes supports the division of *Schistosoma intercalatum* into two separate species. *Parasitology 127*: 131-137.

Lipman DJ, Pearson WR 1985. Rapid and sensitive protein similarity searches. *Science 227*: 1435-1441.

Littlewood DT, Johnston DA 1995. Molecular phylogenetics of the four *Schistosoma* species groups determined with partial 28S ribosomal RNA gene sequences. *Parasitology 111*: 167-175.

Liu F, Lu J, Hu W, Wang SY, Cui SJ, Chi M, Yan Q, Wang XR, Song HD, Xu XN, Wang JJ, Zhang XL, Zhang X, Wang ZQ, Xue CL, Brindley PJ, McManus DP, Yang PY, Feng Z, Chen Z, Han ZG 2006. New perspectives on host-parasite interplay by comparative transcriptomic and proteomic analyses of *Schistosoma japonicum*. *PLoS Pathog 2*: e29.

Macedo AM, Martins MS, Chiari E, Pena SD 1992. DNA fingerprinting of *Trypanosoma cruzi*: a new tool for characterization of strains and clones. *Mol Biochem Parasitol 55*: 147-153.

MacLeod A 2004. Minisatellites and MVR-PCR for the individual identification of parasite isolates. *Methods Mol Biol 270*: 187-202.

MacLeod A, Tweedie A, Welburn SC, Maudlin I, Turner CM, Tait A 2000. Minisatellite marker analysis of *Trypanosoma brucei*: reconciliation of clonal, panmictic and epidemic population genetic structures. *Proc Natl Acad Sci USA 97*: 13442-13447.

Minchella DJ, Lewis FA, Sollenberger KM, Williams JA 1994. Genetic diversity of *Schistosoma mansoni*: quantifying strain heterogeneity using a polymorphic DNA element. *Mol Biochem Parasitol 68*: 307-313.

Oliveira EJ, Pádua JG, Zucchi MI, Vencovsky R, Vieira MLC 2006. Origin, evolution and genome distribution of microsatellites. *Genet Mol Biol 29*: 294-307.

Oura CA, Odongo DO, Lubega GW, Spooner PR, Tait A, Bishop RP 2003. A panel of microsatellite and minisatellite markers for the characterisation of field isolates of *Theileria parva*. *Int J Parasitol 33*: 1641-1653.

Pena HB, de Souza CP, Simpson AJ, Pena SD 1995. Intracellular promiscuity in *Schistosoma mansoni*: nuclear transcribed DNA sequences are part of a mitochondrial minisatellite region. *Proc Natl Acad Sci USA 92*: 915-919.

Prugnolle F, De Meeûs T, Durand P, Sire C, Théron A 2002. Sex-specific genetic structure in *Schistosoma mansoni*: evolutionary and epidemiological implications. *Mol Ecol 11*: 1231-1238.

Rodrigues NB, Silva MR, Pucci MM, Minchella DJ, Sorensen R, LoVerde PT, Romanha AJ, Oliveira G 2007. Microsatellite-enriched genomic libraries as a source of polymorphic loci for *Schistosoma mansoni*. *Mol Ecol Notes 7*: 263-265.

Rodrigues NB, Coura Filho P, de Souza CP, Jannoti Passos LK, Dias-Neto E, Romanha AJ 2002a. Populational structure of *Schistosoma mansoni* assessed by DNA microsatellites. *Int J Parasitol 32*: 843-851.

Rodrigues NB, LoVerde PT, Romanha AJ, Oliveira G 2002b. Characterization of new *Schistosoma mansoni* microsatellite loci in sequences obtained from public DNA databases and microsatellite enriched genomic libraries. *Mem Inst Oswaldo Cruz 97* (Suppl. I): 71-75.

Rollinson D, Southgate VR 1987. The genus *Schistosoma*: a taxonomic appraisal. In D Rollinson, AJ Simpson (eds.), *The biology of schistosomes: from genes to latrines*, Academic Press, London, p. 1-49.

Sambrook J, Fritsch EF, Maniatis T 1989. *Molecular cloning*: *a laboratory manual*, 2nd ed., Cold Spring Harbor, New York, 545 pp.

Sanguinetti CJ, Dias-Neto E, Simpson AJ 1994. Rapid silver staining and recovery of PCR products separated on polyacrylamide gels. *Biotechniques 17*: 914-921.

Scharf JM, Endrizzi MG, Wetter A, Huang S, Thompson TG, Zerres K, Dietrich WF, Wirth B, Kunkel LM 1998. Identification of a candidate modifying gene for spinal muscular atrophy by comparative genomics. *Nat Genet 20*: 83-86.

Simões M, Bahia D, Zerlotini A, Torres K, Artiguenave F, Neshich G, Kuser P, Oliveira G 2007. Single nucleotide polymorphisms identification in expressed genes of *Schistosoma mansoni*. *Mol Biochem Parasitol 154*: 134-140.

Sire C, Durand P, Pointier JP, Théron A 2001. Genetic diversity of *Schistosoma mansoni* within and among individual hosts (*Rattus rattus*): infrapopulation differentiation at microspatial scale. *Int J Parasitol 31*: 1609-1616.

Tautz D, Renz M 1984. Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Res 12*: 4127-4138.

Webster BL, Southgate VR, Littlewood DT 2006. A revision of the interrelationships of *Schistosoma* including the recently described *Schistosoma guineensis*. *Int J Parasitol 36*: 947-955.

Yin M, Hu W, Mo X, Wang S, Brindley PJ, McManus DP, Davis GM, Feng Z, Blair D 2008. Multiple near-identical genotypes of *Schistosoma japonicum* can occur in snails and have implications for population-genetic analyses. *Int J Parasitol 38*: 1681-1691.

Zalloum L, Gomes ML, Kinoshita AT, Toledo MJ, Prioli AJ, de Araújo SM 2005. *Trypanosoma cruzi*: two genetic groups in Paraná state, Southern Brazil. *Exp Parasitol 111*: 55-58.