Intra- and inter-specific variation in nuclear ribosomal internal transcribed spacer 1 of the *Schistosoma japonicum* species complex

L. VAN HERWERDEN1*, D. BLAIR1 and T. AGATSUMA2

1 Zoology Department, James Cook University, Townsville, Queensland 4811, Australia
2 Department of Bioresource Chemistry, Obihiro University of Agriculture and Veterinary Medicine, Inada, Obihiro, Hokkaido 080, Japan

(Received 15 August 1997; revised 3 November 1997; accepted 12 November 1997)

**SUMMARY**

The first internal transcribed spacer (ITS1) of the nuclear ribosomal DNA repeat was sequenced for members of the *Schistosoma japonicum* species complex (*S. mansoni*, *S. mekongi* and 2 geographical isolates of *S. japonicum*). The ITS1 is composed of 3 distinct regions: the 5′ end (23 nucleotides); a tract of approximately 90–140 nucleotides, which occurs up to 7 times in tandem, the number varying even within an individual in all species investigated in this study; the 3′ region (378 nucleotides), which lacks repeats. There is size and sequence variation among copies of the ITS1 repeat within a single individual. The relative abundances of size variants of ITS1 in *S. japonicum* have been ascertained by hybridizing genomic digests with an ITS1 probe. Multiple repeats and intra-individual variation in numbers and abundance of these is a feature of the Asian schistosomes, but not generally of African schistosomes. Possible reasons for this difference in ITS1 between African and Asian schistosomes are discussed. The ITS1 repeat sequences described for African schistosomes are different to, and cannot be aligned with, those from the Asian species described here, whereas the remainder of the ITS1 can be aligned quite easily.

Key words: ribosomal DNA, repeats, ITS1, *Schistosoma japonicum*, size variation.

**INTRODUCTION**

Schistosomes are digenetic trematodes that live as adults and lay their eggs in the vasculature of mammals, birds and crocodiles. Nineteen species of the genus *Schistosoma* have been recognized to date, of which 7 infect man: *S. mansoni* (*S. mansoni* group); *S. haematobium*, *S. intercalatum* and *S. mattheei* (*S. haematobium* group); *S. japonicum*, *S. mekongi* and *S. malayensis* (*S. japonicum* group) (McManus & Hope, 1993). The most important species infecting man are *S. mansoni*, *S. haematobium* and *S. japonicum*. Studies on phylogeny and/or intraspecific variation in *Schistosoma* species have recently been done using nuclear rDNA and mitochondrial DNA sequences (Després et al. 1992; Bowles, Blair & McManus, 1995a, Littlewood & Johnston, 1995; Barker & Blair, 1996; Blair et al. 1997). Ribosomal DNA sequences are present as tandemly repeated clusters of highly conserved genes for 18S, 5.8S and 28S rRNA separated by variable spacer sequences (Long & Dawid, 1980). There are about 100 copies of rDNA per haploid schistosome genome, each approximately 10 kb in size, of which about 4 kb per copy is spacer (Simpson et al. 1984; Walker, Rollinson & Simpson, 1986).

In Asia, the most important schistosome is *S. japonicum*, with the Chinese and Philippine strains being regarded by some workers as very distinct. There has been some controversy concerning differences between these strains. No differences were found in randomly amplified polymorphic DNA (RAPDs) or ribosomal ITS2 sequences and only 0.5% difference was found between these strains in mitochondrial CO1 DNA sequence (Bowles et al. 1995a). However, TPI cDNA sequences differed between the strains (Hooker & Brindley, 1996), as did paramyosin sequences (Hooker et al. 1995). Isoenzyme studies also indicated a high level of genetic divergence (Woodruff et al. 1987; McManus & Hope, 1993). There are biological differences between the strains. In particular, infectivity and pathogenicity to the definitive host differs between the different geographic strains, as does drug susceptibility (McManus & Hope, 1993 and references therein) and protective immunity after heterologous vaccination (Hope, Duke & McManus, 1996).

In this paper, we investigate the primary structure of the ribosomal first internal transcribed spacer (ITS1) in the *S. japonicum* species complex. The study is part of a larger phylogenetic study for which data has been presented by Blair et al. (1997). The 5′ and 3′ regions of ITS1, which are separated by a region of internal repeats, differ significantly (>5%) only between species. Here we present the
interesting variation of ITS1 within individuals and between strains and species, which is due principally to the presence of the internal repeats, and the implications of such variation, with particular reference to the Chinese and Philippine strains of *S. japonicum*.

**Materials and methods**

**DNA extraction**

Adult *S. malayensis* were raised in mice infected with cercariae, which were obtained from *Robertsiella* sp. snails in August 1993, from Baling, Malaysia (IMR, Kuala Lumpur). *S. mekongi* from Khong Island, Laos were maintained in *Neotricula operca* for 10 years at Mahidol University, Bangkok, Thailand. Strains of *S. japonicum* were collected from snails near Sorsogon (Philippines) in 1992 and maintained at QIMR, Brisbane, Australia. Naturally infected snails from Anhui (China) were routinely sent to Brisbane, where released cercariae were collected and used to infect Balb/c female mice. *S. mansoni* (Puerto Rico) DNA, from pooled worms only, was kindly donated by Paul Brindley (QIMR, Australia), for comparative purposes. Specimens used in this study had been stored in 80% ethanol at 4 °C. Prior to DNA extraction they were washed 3 times in extraction buffer (40 mM Tris pH 8.0, 10 mM EDTA, 200 mM NaCl) before being digested in digestion buffer (extraction buffer containing 1% SDS, 500 μg/ml Proteinase K) at 37 °C for 2 h. DNA was extracted from a lysed individual male worm and from pooled adult worms by standard techniques (Sambrook, Fritsch & Maniatis, 1989).

**DNA amplification**

The ITS1 region was amplified using primers BD1 (5′GTCGTAACAAGGTTTTCCGTA3′) and 4S (5′TCTAGATGCGTTCCGAA(G/A)TGTCGATG-3′) (Bowles & McManus, 1993). These primers have highly conserved annealing sites in the 18S rRNA and 5.8S rRNA respectively. PCR was done in a volume of 20 μl, as follows: 10–100 ng template and 10 pmol of each primer were added to 1 x *Taq* polymerase buffer (Promega), 400 μM of each dNTP, 3 mM MgCl₂ and 0.25 Units of *Taq* polymerase (Promega). After denaturation at 94 °C, 1 min; 50 °C, 30 sec and 72 °C, 30 sec on a Corbett Thermal cycler.

**Cloning**

PCR products were cloned according to the manufacturer’s instructions (TA cloning kit, Invitrogen). Positive recombinant clones were picked, grown in 5 ml of terrific broth (in the presence of 50 μg/ml ampicillin) overnight at 37 °C and plasmids were extracted using standard alkaline lysis procedures as described in the Prism Cycle Sequencing manual (Perkin Elmer). However, the PEG precipitation was omitted at the end of the procedure. Plasmids containing inserts of the correct size were confirmed by Eco RI digestion of an aliquot at 37 °C for 2 h and examined by ethidium bromide-stained 0.6% agarose gels.

**Sequencing**

Recombinant clones were sequenced in both orientations as follows. Eight μl of terminator premix (Prism Ready Reaction Dyedexoxy Terminator Cycle Sequencing Kit, Perkin Elmer) was added to 0.5 μg of double-stranded template and 3.2 pmoles of primer in a 20 μl reaction volume. Universal and PCR primers were used for sequencing as follows: 96 °C, 1 min followed by 25 cycles of 96 °C, 30 sec; 50 °C, 15 sec and 60 °C, 4 min. Direct sequencing was also done, in both orientations, for the *S. japonicum* single worm PCR products (both strains) and for the *S. mansoni* PCR products, by excising individual bands from 1% LMP agarose. Template was then prepared for sequencing using PCR MagicPrep, according to the manufacturer’s instructions (Wizard, Promega). Direct sequencing was done as above, but using only 30–100 ng template. Sequenced products were precipitated with 1/10 volumes of 3 M NaAcetate, pH 4.5 and 2.5 volumes of absolute ethanol overnight at 0 °C, pelleted and then washed in 70% ethanol before drying in a rotary evaporator (Speedvac). The pellet was reconstituted in 4 μl of 83 mM EDTA, 0.83 volumes deionized formamide and denatured at 90 °C for 2 min before loading onto an ABI 373 DNA sequencer (following the instructions in the User’s Manual). At least 2 individual clones were sequenced/克loned ITS1 size variant/species. The few random point differences observed were considered to be due to gel reading problems and misincorporation during PCR.

**Sequence alignment and analysis**

Sequences were aligned in the sequence editing program ESEE (Version 1-09) (Cabot & Beckenbach, 1989). DNAis (6th version 1988, Pharmacia) was used to construct Harr plots (dot plots) for identi-
280 bp of the 378 bp fragment.

Genomic digests and Southern blots

*S. japonicum* genomic DNA from 5 individuals was digested with *Hind* III, which cuts 200 bp upstream of the 3' end of 18S rDNA and Xba I, which cuts 133 bp downstream of the 5' end of ITS2, to remove complete ITS1 from the rest of the genome (with flanking 18S and 5.8S rDNA – partial ITS2 fragments, accounting for 483 bp) by standard techniques (Sambrook et al. 1989). Digestion should excise a fragment expected to be approximately 1000 bp, of which 500 bp is ITS1 only, in the smallest variant. *S. mansoni* genomic DNA was digested using *Alu* I and *Bln* I, each of which has a single cut site within the ITS1, to produce a 300 bp ITS1 fragment. Genomic digests of *S. japonicum* and *S. mansoni* were run on an 0.8% agarose gel. DNA was transferred to Hybond N+ membranes (Amersham) as described by Southern (1975). The filters were hybridized to an α-P32 dATP random-primed labelled (Boehringer Mannheim) 280 bp ITS1 fragment, which had been generated by PCR amplification of genomic template with internal primers (AsitsF and AsitsR). The probe did not contain any of the ITS1 repeat sequence, as the primers only amplify 3' of the repeats. Both membranes were hybridized and washed at 50 °C. Washes were done to a stringency of 0.5 × SSC, 0.1% SDS before exposure to a Phosphor Imager screen (Molecular Dynamics) and scanning. Relative band intensities were determined using ImageQuant (Molecular Dynamics).

**RESULTS AND DISCUSSION**

Internal repeats are located only near the 5' end of ITS1, in the *S. japonicum* species complex. The repeats are preceded by 23 bp of sequence at the 5' end and followed by 378 bp at the 3' end of ITS1 (the latter discussed by Blair et al. 1997). No repeats have been found elsewhere within the ITS1 of *S. japonicum* and related species, unlike the situation in *S. mansoni* and other African schistosomes (Kane et al. 1996), where an isolated copy of the internal
Fig. 3. Alignment of ITS1 repeats from Asian schistosomes. (A) ‘.’ indicates the same base as shown on upper line; ‘-’ indicates alignment gaps; SjP(a)ab from small *Schistosoma japonicum* (Philippines) clones, SjP(ab)a from larger *S. japonicum* (Philippines) clones, SjC(ab)aL1 from large *S. japonicum* (Chinese) clones, SjC(ab)aL2, from large *S. japonicum* (Chinese) with additional (T.A), Sma(a) is the ‘small’ *S. malayensis* clone, Sma(ab)a1 from ‘medium’ *S. malayensis* clones, copies numbered according to positional order within clone, Sma(ab)a1–4 are from ‘large’ *S. malayensis* clones, copies numbered according to positional order within clone, Sma(ab)a1–2 are from ‘small’ *S. mekongi* clones, copies numbered according to positional order within clone, Sma(ab)a1–2 are from ‘large’ *S. mekongi* clones, copies numbered according to positional order within clone. An ‘a’ indicates the start of a repeat and ‘b’ indicates start of b repeat in the alignment.

Repeat is sometimes present near the 3′ end of ITS1. Repeats within the *S. japonicum* species complex align well, as do such repeats among the African schistosomes studied to date, however, alignment is not possible between repeats of the African schistosomes and those of the *S. japonicum* species complex. Neither is alignment possible between these groups at the 5′ end of ITS1 (there is only 62-5% similarity), although *S. japonicum*, *S. malayensis* and *S. mekongi* differ by <10% in this region (data not shown). There is, however, good similarity between the 3′ tract (post-repeat) of ITS1 in both African and Asian groups (94% similarity), except for the 30 bp immediately prior to the 5.8S rRNA coding region.

Different numbers of the repeat element (in 2 parts: a ~ 90 nt in all 3 species, followed by b ~ 15 nt in *S. mekongi* and *S. malayensis* and 39 or 53 nt in *S. japonicum*) occur, commencing 24 nucleotides from the 5′ end of the ITS1. Length variants in the repeat region in a single individual can be demonstrated by PCR amplification (Fig. 1). ITS1 clones sequenced (from the DNA of pooled worms) were arbitrarily termed small (smallest sequenced clone), medium (intermediate size sequenced clone) or large (largest clone sequenced). These did not represent all the size variants present, as detected from genomic digests of DNA from pooled worms. In fact, the most abundant ITS1 variants present in the genome, were larger (ab)aL and (ab)L2 in the Philippine strain and (ab)b, (ab)bL in the Chinese strain (Fig. 2). Furthermore, the smaller variants sequenced were barely detectable by Southern blot
Table 1. Percentage pairwise differences of repeat a sequences in ‘large’ ITS1 clones in 3 Asian schistosomes

(S. mal, Schistosoma malayensis; S. mek, S. mekongi; S. jap, S. japonicum; SjP, S. japonicum (Philippine); SjC, S. japonicum (Chinese).)

<table>
<thead>
<tr>
<th></th>
<th>S. mal a1</th>
<th>S. mek a1</th>
<th>SjC a1</th>
<th>SjP a1</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. mal a1</td>
<td>—</td>
<td>2.5</td>
<td>7.5</td>
<td>8.75</td>
</tr>
<tr>
<td>S. mal a2</td>
<td>1.25</td>
<td>2.5</td>
<td>7.5</td>
<td>10.0</td>
</tr>
<tr>
<td>S. mal a3</td>
<td>2.5</td>
<td>3.75</td>
<td>10.0</td>
<td>11.25</td>
</tr>
<tr>
<td>S. mal a4</td>
<td>7.5</td>
<td>8.75</td>
<td>15.0</td>
<td>13.75</td>
</tr>
<tr>
<td>S. mek a1</td>
<td>2.5</td>
<td>—</td>
<td>8.75</td>
<td>10.0</td>
</tr>
<tr>
<td>S. mek a2</td>
<td>1.25</td>
<td>1.25</td>
<td>6.25</td>
<td>8.75</td>
</tr>
<tr>
<td>S. mek a3</td>
<td>2.5</td>
<td>2.5</td>
<td>8.75</td>
<td>10.0</td>
</tr>
<tr>
<td>S. mek a4</td>
<td>5.0</td>
<td>3.75</td>
<td>10.0</td>
<td>11.25</td>
</tr>
<tr>
<td>S. jap China a1</td>
<td>7.5</td>
<td>8.75</td>
<td>—</td>
<td>1.25</td>
</tr>
<tr>
<td>S. jap China a2</td>
<td>7.5</td>
<td>8.75</td>
<td>1.25</td>
<td>—</td>
</tr>
<tr>
<td>S. jap Philippine a1</td>
<td>8.75</td>
<td>10.0</td>
<td>1.25</td>
<td>—</td>
</tr>
<tr>
<td>S. jap Philippine a2</td>
<td>8.75</td>
<td>10.0</td>
<td>2.5</td>
<td>0</td>
</tr>
</tbody>
</table>

analysis, so they are much less abundant in the genome. PCR and cloning procedures are known to favour smaller fragments, which would explain why these were over-represented among the clones sequenced.

Variation among repeat sequences, both within and between clones of the same species, was 0–7.5%, with the greatest intra-individual variation occurring between a1 and the fourth copy a4 of the largest clone of S. malayensis sequenced (Fig. 3; Table 1). In S. malayensis divergence from the first copy of sequence a increases progressively with number of repeats in a particular clone from 1.25% (a2) to 2.5% (a3) to 7.5% (a4). A similar situation occurs in S. mekongi. Copies of a at the same position share more similarity across clones than they do within. Thus the first a repeat in the largest clone is more similar to the first a repeat in smaller clones from the same species than to the second third or fourth a repeats in the largest clone. The S. japonicum a region is somewhat divergent relative to the other 2 Asian species (Table 1).

Repeat element b is smaller and much less variable than a, where no variation is found either among S. malayensis clones or S. mekongi clones. The b repeat is identical in S. malayensis and S. mekongi. The b1 sequence in both strains of S. japonicum differs dramatically from the other 2 species, due to the presence of a 23 nucleotide insert containing a (TA)k microsatellite (SjP(ab)2-1) in Fig. 3 and 4/11 base differences prior to the insert mentioned above (i.e. a 29% variation from the b1 sequence of the other 2 species). The Chinese strain had the same (TA)k microsatellite (SjC(ab)2-1) in Fig. 3 and an additional b1 variant detected by direct sequencing (of a band excised from a gel), not observed in the Philippine strain, which contained an expanded (TA)13 microsatellite (SjC(ab)2L1 in Fig. 3).

The relative abundances of size variants of ITS1 were determined by Southern blots of genomic digests of pooled DNA from adult worms of S. japonicum (Chinese and Philippine strains) (Fig. 2) and of S. mansoni (data not shown). ITS1 was isolated from the rest of the genome by restriction digestion, as described in the Materials and Methods section. The most abundant ITS1 size variants in the Chinese strain were 1150 bp and 1020 bp, consistent with them representing (SjC(ab)a and (SjC(ab)a respectively. The relative abundance of each was 50.8 and 29.2% respectively, with a reduction in relative abundance as the number of repeats increases or decreases (Fig. 2). In a Philippine strain (Sorsogon), only 4 variants were detectable, of which 1020 bp and 760 bp were the most abundant, consistent with (SjP(ab)a and (SjP(ab)a respectively, (59.3 and 29.7% relative abundances respectively). S. mansoni from Puerto Rico was digested and probed in a similar manner as described above for the Asian schistosomes. As expected from published S. mansoni ITS1 data (Kane et al. 1996), there was only one ITS1 length class present in the genome. This appears to be the case for all African species for which ITS1 sequences have been published, as they lack intra-individual variation in the ITS1.

Many variant forms of ITS1 have been described, within and among individuals across a range of organisms, e.g. the dipterans: Simulium damnosum species complex (Tang et al. 1996), Drosophila (Schlötterer et al. 1994) and Aedes species (Wesson, Porter & Collins, 1992), the orthopteran Melanoplus species (Kuperus & Chapco, 1994), the coleopteran Cicindela dorsalis (Vogler & De Salle, 1994), the cestode Echinococcus (Bowles, Blair & McManus, 1995b), the trematodes Dolichosaccus species (Luton, Walker & Blair, 1992) and Schistosoma species (Kane & Rollinson 1994; Kane et al. 1996) and the mammal Homo sapiens (Gonzalez et al. 1990). The degree of sequence variation is a balance between processes.
generating variation and homogenizing processes (Schlötterer et al. 1994; Fritz et al. 1994). Variation can be generated by many factors, such as sexual reproduction (due to the formation of bivalents and chiasmata), dispersal of rDNA loci on various chromosomes (> 1 NOR) (Vogler & De Salle, 1994), interbreeding with sibling species and environmental pressure, which may select for minor alleles. Known homogenizing processes are those involved in concerted evolution (Dover, 1982; Hillis & Davis, 1988; Hillis et al. 1991; Linares, Bowen & Dover, 1994; Schlötterer et al. 1994; Tang et al. 1996).

The presence of intra-individual variation in ITS1 of the Asian schistosomes and absence of such variation in the African schistosomes may be explained by 2 factors, both of which may contribute to a reduction in concerted evolution processes in the Asian species. It is concerted evolution that generally eliminates variation in the multiple copies of ribosomal genes found within individuals and across populations of a species. (i) Putative ‘hot spots’ for recombination (Chi-like sites), present in the internal repeats of ITS1 of the African schistosomes (Kane & Rollinson, 1994; Kane et al. 1996), are absent from the internal repeats of the ITS1 of the S. japonicum species complex. (ii) The chiasma frequency in S. japonicum is extremely low (FXi = 3–0), whereas in S. mansoni FXi = 15–3 (Hirai et al. 1996). Similar levels of chiasma formation have been determined in mammals (FXi = 15–3 for Chinese hamsters and FXi = 17–3 for mice) as in S. mansoni (Hirai et al. 1996 and references therein). This suggests that genetic recombination due to crossing-over at chiasmata is reduced in S. japonicum. Both absence of recombination at ‘hot spots’ in the ITS1 and reduced genome-wide chiasma formation in the S. japonicum species complex would explain the apparent reduction in concerted evolution processes, which would in turn explain why there is increased intra-individual variation in the ITS1 of the S. japonicum species complex.

There appears to be a difference in the rates/patterns of concerted evolution of repetitive ribosomal genes within the genus Schistosoma. Such differences presumably occurred subsequent to the speciation event which resulted in the S. japonicum species complex, but prior to the speciation within the S. japonicum species complex. It would be most interesting to investigate the ITS1 sequence of another Asian schistosome species complex, the S. sinensis complex and identify whether internal repeats occur and if they do whether they are S. mansoni or S. japonicum-like.

Primers BD1, BD2 and 4S were designed by J. Bowles. H. Hirai (Primate Research Institute, Kyoto, Japan), T. Taguchi (Kochi Medical School, Japan), M. Hirata (Kurume University, Japan), S. Habe (Fukuoka University, Japan), K. Lai (IMR, Malaysia) and S. Upatham (Mahidol University, Thailand) kindly donated S. malayensis and S. mekongi samples. S. japonicum was kindly donated by Professor D. McManus (QIMR, Brisbane, Australia). Puerto Rican S. mansoni DNA was kindly provided by Paul Brindley (QIMR, Brisbane, Australia). Sequenced products were separated on ABI sequencers by Jenny Cassady and Emma Puttick at the DNA Sequence & Analysis Facility, CMCB, UQ. Dr R. Slade (QIMR, Brisbane, Australia) made some helpful suggestions after reading this manuscript. This research was funded by the Australian Research Council.

REFERENCES


ITS1 variants in Schistosoma japonicum

Schistosoma and Paragonimus (Trematoda), by using the chiasma distribution graph. *Genes and Genetic Systems* 71, 181–188.


