A Retrotransposon of the Non-Long Terminal Repeat Class from the Human Blood Fluke *Schistosoma mansoni*. Similarities to the Chicken-Repeat-1-like Elements of Vertebrates

Alexander C. Drew and Paul J. Brindley

Molecular Parasitology Unit, and Australian Centre for International and Tropical Health and Nutrition, Queensland Institute of Medical Research, Australia.

The genomes of representative species of fishes, amphibians, and reptiles contain non-long-terminal-repeat (non-LTR) retrotransposons showing strong sequence identity to the chicken repeat 1 (CR1) non-LTR retrotransposon from birds. These nonavian retroelements have been termed CR1-like elements. We have isolated sequences of a non-LTR retrotransposon from the human blood fluke *Schistosoma mansoni*. These schistosome sequences, which we have termed the SR1 family of non-LTR retrotransposons, contain regions of deduced amino acids characteristic of the CR1-like elements. SR1 elements possess atypical 3’ termini consisting of the tandem repeat (AACCATTTG), which are similar in structure to the imperfect tandem repeat of the 3’ termini of CR1. There are at least 200 copies of SR1 interspersed through the genome of *S. mansoni*. The structural and amino acid sequence similarities of SR1 with members of the CR1-like elements suggest that the SR1 family belongs to the CR1-like category of non-LTR retrotransposons. Although other non-LTR retrotransposons have been described in invertebrates, this is the first CR1-like element reported from a nonvertebrate taxon, suggesting that the phylogenetic distribution of CR1-like retrotransposons is not restricted to vertebrates.

**Introduction**

Eukaryotes contain a number of classes of mobile genetic elements which are capable of moving to new sites within their genomes, often in response to genetic or environmental stimuli. Transposition of mobile genetic elements within the genome can directly generate both genotypic and phenotypic variation and, moreover, the distribution of transposable genetic elements along the chromosomes facilitates recombination and rearrangements at loci harboring these elements (Lehrman et al. 1987; Finnegan 1989). Accordingly, possession of these kinds of elements contributes to the adaptive performance of natural populations, although instances of deleterious outcomes for the host genome as a result of movement of mobile genetic elements have been documented (Fawcett et al. 1986; Kazazian et al. 1988; Wallace et al. 1991).

Mobile genetic elements are classified broadly on their mechanisms of transposition. Elements belonging to class I undergo replicative transposition by a mechanism involving transcription, followed by reverse transcription and integration into a new position within the genome (Eickbush 1992). Class I transposable elements are comprised of the retrotransposons and the short interspersed elements or SINES. Retrotransposons are usually several kilobases in length and contain open reading frames (ORFs) for reverse transcriptase (RT) and other viral-like proteins (Boeke 1989). SINES, which include the human Alu elements, are shorter (less than 500 bp in length) do not encode proteins (Deininger 1989), and are transposed by proteins encoded by retrotransposons (Eickbush 1992). Two classes of retrotransposons are recognized: members of the first class, the LTR retrotransposons, are bounded by long terminal repeats (LTRs), whereas members of the second class lack LTRs and are termed the non-LTR retrotransposons. Non-LTR retrotransposons are also known as long interspersed sequences (LINES) or polyA retrotransposons (Boeke 1989). LTR retrotransposons are typified by the copia (Mount and Rubin, 1985) and gypsy (Marlor, Parkhurst, and Corces 1986) elements from *Drosophila melanogaster* while L1 elements (Singer and Skowronski 1985; Hutchison et al. 1989) from humans and I elements (Fawcett et al. 1986) of *D. melanogaster* are well-characterized non-LTR retrotransposons. Class II elements, in contrast to those from class I, transpose directly from DNA to DNA (Sentry and Kaiser 1992). Class II elements are characterized by inverted terminal repeats and encode a transposase responsible for their transposition. Well-known class II elements include P from *D. melanogaster* and mariner from *D. mauritiana* (Jacobson, Medhora, and Hartl 1986; Sentry and Kaiser, 1992).

Transposable genetic elements are usually present within a genome in multiple copies due to their ability to undergo replicative transposition. Indeed, they appear to constitute substantial percentages of the genomes of many eukaryotes (Finnegan 1989). The genome of the human blood fluke *Schistosoma mansoni* contains 40% repetitive DNA (Simpson, Sher, and McCutchan 1982). Although a family of SINE-like sequences, the SMe elements (Spotila et al. 1989; Drew and Brindley 1995), has been identified in schistosomes, no protein-encoding, transposable element has yet been reported from these worms. However, characterization of transposons from medically and economically important parasites such as schistosomes could facilitate the exploitation of molecular genetics to control these pathogens through,
for example, germ-line transformation (Warren and Crampton 1994).

This paper describes the isolation and characterization of a middle-repetitive DNA element present at more than 200 copies in the genome of *S. mansoni*. Analysis of the deduced amino acid sequence of the element showed that it contained an ORF encoding a reverse transcriptase and suggested that it was related to the chicken repeat 1 (CR1)-like, non-LTR retrotransposons previously described only from birds and other vertebrates (Burch, Davis, and Haas 1993; Vandergon and Reitman 1994). This retrotransposon, which we have named Sr1 (Schistosome Retrotransposon 1), is the first protein-encoding, mobile genetic element to be reported from any of the schistosomes.

**Materials and Methods**

Genomic DNA from Schistosomes.

Mixed-sex adult *S. mansoni* (Puerto Rican strain) and *Schistosoma japonicum* (Philippine strain) were perfused from mice 7–10 weeks after laboratory infection, washed, and stored in liquid N₂. Genomic DNAs from these worms were isolated by chromatography using QIAGEN (Chatsworth, Calif.) columns and reagents.

Isolation of Repetitive Sequences from the Genome of *S. mansoni*.

An *S. mansoni* HindIII-amplicon (an amplicon being a representative sample of a genome, as defined by Lisitsyn, Lisitsyn, and Wigler [1993]) was constructed as previously described (Drew and Brindley 1995) using the oligonucleotide primers RH24, (5' -AGCA-CTCTCCAGCCTCTACGCAG) and RH12, (5' -ACGTTCGCGTGA) (Lisitsyn, Lisitsyn, and Wigler 1993). One microgram of the *S. mansoni* amplicon was separated by electrophoresis through 1% agarose containing ethidium bromide and visualized with a transilluminator. Bands were evident in a smear of PCR products, and apparently represented repetitive sequences. In order to clone the DNA sequences visualized as bands, bands were isolated using the procedure of band-stab PCR (Bjourson and Cooper 1992) where a sterile 26-gauge hypodermic needle was inserted into discrete bands. Needles were subsequently agitated in 100 μl of PCR mix, after which the band-stabbed DNA was reamplified using the RH24 oligonucleotide as described (Lisitsyn, Lisitsyn, and Wigler 1993; Drew and Brindley 1995). Products of the reamplification were separated through agarose containing ethidium bromide and visualized with a transilluminator. Bands were evident in a smear of PCR products, and apparently represented repetitive sequences. In order to clone the DNA sequences visualized as bands, bands were isolated using the procedure of band-stab PCR (Bjourson and Cooper 1992) where a sterile 26-gauge hypodermic needle was inserted into discrete bands. Needles were subsequently agitated in 100 μl of PCR mix, after which the band-stabbed DNA was reamplified using the RH24 oligonucleotide as described (Lisitsyn, Lisitsyn, and Wigler 1993; Drew and Brindley 1995). Products of the reamplification were separated through agarose by electrophoresis, the bands excised, and the DNA purified by chromatography (Wizard, Promega). Inserts were labeled with 32P-dCTP by random priming and employed to probe the blots, using hybridization and washing conditions as described (Church and Gilbert 1984). For low-stringency washes, filters were washed in 2 × SSPE, 0.1% SDS for 10 min at 22°C, then 1 × SSPE, 0.1% SDS for 15 min at 65°C. High-stringency washes consisted of low-stringency washes followed by two washes in 0.1 × SSPE, 0.1% SDS for 10 min at 65°C.

Construction and Screening of Size-Selected, Genomic Libraries.

Five micrograms of *S. mansoni* genomic DNA was digested with endonuclease and electrophoresed through 1% agarose. The desired size range of restriction fragments was excised from the gel, purified using the Wizard System (Promega), ligated into pUC19, and used to transform *E. coli* as above. Transformed bacteria were cultured on LB-agar plates containing ampicillin, after which the colonies were transferred to nylon membranes. Inserts of plasmid clones were purified, radiolabeled, and employed to screen the nylon membranes, employing stringent washing conditions (Church and Gilbert 1984). Plasmid DNA was prepared from positive colonies as above.

Copy Number Estimation.

Five micrograms of restriction-endonuclease-digested *S. mansoni* genomic DNA was separated by agarose gel electrophoresis along with the equivalent of 1 to 1,000 copies per genome of reference clones BS6 and E3 (see Results), digested to release their inserts. The genome of *S. mansoni* weighs ~0.26 pg [Simpson, Sher, McCutchan 1982]). The DNA was Southern transferred to nylon membrane and hybridized to radiolabeled probes derived from the reference clones.

Nucleotide Sequencing and Sequence Analysis.

The nucleotide sequences of both strands of the inserts of clones and subclones in pUC19 were determined by the dyeoxy method using the ABI (Foster City, Calif.) *Taq* DyeDexxy terminator cycle sequencing system and an automated DNA sequencer (ABI). Oligomers for sequencing were prepared using a Beckman Oligo 1000 DNA synthesizer. Deduced amino acid sequences of clones were used to search the GenBank database using the MailBlast program (version 2.1) of the GCG Package (version 7.3.1-UNIX, 1993, Genetics Computer Group, Madison, Wis.). Amino acid alignments were produced using the PILEUP program of
GCG, and identities between sequences were analyzed with the GAP algorithm (GCG).

Results

A Repetitive, HindIII Fragment of the S. mansoni Genome Is Part of a Retrotransposon

By ligating oligonucleotide adaptors to HindIII-digested genomic DNAs, Lisitsyn and co-workers Lisitsyn, Lisitsyn, and Wigler (1993) have shown that restriction fragments of genomic DNA ranging in size up to 3 kb can be amplified using PCR. When we treated HindIII-digested genomic DNA of S. mansoni in this manner, the resulting products included at least 10 discrete bands superimposed on a smear of PCR products ranging in length from about 0.3 to 3 kb (fig. 1A). These discrete bands can be expected to be middle repetitive elements bearing at least two HindIII sites within 3 kb of each other. Subsequently, by employing the procedure of band-stab PCR to isolate single bands from the am-
plicon, we were able to reamplify a number of these bands. Reamplified bands were cloned, and their sequences were used to search the GenBank database using the BLAST algorithm. One clone, named BS6, which was derived from a ~800-bp band (fig. 1B), encoded an ORF of 272 deduced amino acid residues which exhibited strong identity to reverse transcriptase (RT) from non-LTR retrotransposons. (Sequences of other band-stab clones lacked identity to transposons and are not dealt with further here.) The best database matches to BS6 were with the CR1 non-LTR retrotransposon (41% identity over 134 aa) and with the T1Ag non-LTR retrotransposon from the mosquito Anopheles gambiae (51% identity over 60 aa). A modest match (37% over 64 aa) was found with the non-LTR retrotransposon BGR2 from Biomphalaria glabrata, the intermediate snail host of S. mansoni (Knight et al. 1992).

Seven blocks of conserved amino acid residues—termed domains 1–7—have been identified in RTs from various groups of retroelements (Xiong and Eickbush 1990; Eickbush 1992). BS6 encoded the RT conserved domains 4–7. An alignment of the deduced amino acid sequence of BS6 with these conserved domains of RT from representative non-LTR retrotransposons is presented in fig. 1C. Given the sequence homology of BS6 to non-LTR retrotransposons, it was clear that the repetitive HindIII fragment from the genome of S. mansoni was also part of a novel non-LTR retrotransposon.

The Schistosome Retrotransposon Has an Unusual 3' Terminus

To identify larger fragments of the non-LTR retrotransposon, a Southern blot of S. mansoni genomic DNA digested with various restriction enzymes was probed with the radiolabeled insert of BS6 (fig 2A). The probe hybridized strongly to a ~800-bp fragment in HindIII-digested genomic DNA (lane 1), as was expected since BS6 represented a repetitive HindIII fragment of 817 bp in length. In addition, strong bands of hybridization were apparent at 1.1 kb in HindIII-digested, 2 kb in EcoRI-digested (lane 2), and 0.6 and 1.4 kb in Pst I-digested genomic DNA (lane 5). Clone H1 (corresponding to the 1.1-kb HindIII band of hybridization) and clones E1, E2, and E3 (corresponding to the 2-kb EcoRI band of hybridization) were isolated from size-selected genomic libraries. Clone H1 was a 1,127-bp HindIII fragment resulting from the mutation of one of the HindIII sites flanking BS6; it consisted of the BS6 sequence and an additional 310 bp of downstream sequence. E1, E2, and E3 were 94%–96% identical, included the full sequence of BS6 and, in addition, encoded 392 aa amino-terminal to, and 66 aa residues carboxyl-terminal to those of BS6. H1 was employed to isolate three Pst I fragments, P1, P2, and P3, from size-selected genomic libraries, and these contained the 3' terminus of the retrotransposon (fig. 2B and C). The physical relationship of these fragments to BS6 is shown in figure 2B. P1, P2, and P3 shared strong sequence identity to each other (92%) up to a direct repeat of consensus (AACCATTGTG), after which the identity ended abruptly (fig. 2C), although there were a few nucleotides after the direct repeats of P1 and P3 which resembled part of the repeat motif. The structure of the 3' terminus of the schistosome non-LTR retrotransposon was unusual because elements of this class characteristically terminate in a simple, A-rich motif such as a poly(A) tract or (TAA), or (TGAAA), (Hutchison et al. 1989; Eickbush 1992). The only non-LTR retrotransposon reported with a tail diverging from these simple patterns is the CR1
element from birds, which terminates in the imperfect, octameric repeat (NATTCTRT)₂ (Silva and Burch 1989).

The members of retrotransposon families display sequence heterogeneity due to errors occurring during reverse transcription and during DNA replication after integration, and these errors can render individual elements functionally inactive (Hutchison et al. 1989; Preston 1996). In order to determine the sequence of a potentially active retroelement, a consensus sequence was constructed using the cloned fragments of schistosome DNA described above. As these fragments exhibited 92%-96% sequence identity, we were able to construct a consensus sequence for the schistosome retrotransposon in which the nucleotide at each position was present in at least two clones. This consensus sequence of 2,337 bp contained an ORF of 772 codons, containing all seven conserved domains of RT (Xiong and Eickbush 1989, 1990), a stop codon, 101 bp of 3' untranslated region, and the terminating tandem repeat (AACCATTTG)₂. We have termed this schistosome transposable element SR₁ (Schistosome Retrotransposon 1). SR₁ is the first retrotransposon to be reported from any of the schistosomes and, indeed, from any species of the phylum Platynhelminthes.

SR₁ Elements are Frequently 5'-Truncated and Heterogenous in Sequence

Most elements of families of non-LTR retrotransposons are truncated randomly at their 5' ends (Hutchison et al. 1989; Burch, Davis, and Haas 1993). This truncation is thought to be caused by frequent, premature termination during reverse transcription of the RNA intermediate (Luan et al. 1993). As a result, the copy number determination of non-LTR retrotransposons will vary according to the position of the nucleotide sequence employed for the estimation. Thus, the genome of the chicken has 1,000 copies of the CR₁ element which extend 0.8 kb from the 3' terminus but only 30 copies which extend as far as 2.2 kb (Burch, Davis, and Haas 1993). We determined the copy number of SR₁ at two positions along the sequence of SR₁. Around 200 copies of SR₁ extended as far as 1.1 kb from the 3' end of SR₁ when determined at low-stringency conditions, and this decreased to 50 copies at high stringency. About 20 copies extended as far as 2.2 kb (data not shown). These results demonstrated that the majority of SR₁ elements were 5' truncated. Moreover, because increasing stringency reduced the copy number of SR₁, there also appeared to be considerable sequence variation between individual members of the SR₁ family.

To determine whether a homolog of SR₁ was present in the genome of the Asian schistosome S. japonicum, a Southern blot of genomic DNA was probed with the labeled insert of BS6. No hybridization was evident to S. japonicum DNA even at low stringency (data not shown). Protein-encoding genes of the two species of schistosome usually cross hybridize under these conditions (e.g., Becker et al. 1995), indicating either that a homolog of SR₁ was absent from the genome of S. japonicum or that the SR₁ elements are more divergent than other schistosome protein-encoding sequences.

Similarity to the CR₁-Like Elements of Vertebrates

Given that SR₁ shared a higher identity with CR₁ than with other non-LTR retrotransposons, and that SR₁ and CR₁ elements have atypical 3' termini, a further analysis was undertaken of the sequence relationship between these two retrotransposons. When the sequences of the consensus SR₁ and CR₁ (Burch, Davis, and Haas 1993) were compared, the region of ~200 aa residues carboxyl-terminal to the RT conserved domains shared the same degree of identity (41%) as did the region of ~260 residues comprising the seven conserved domains of RT (40%). By contrast, the degree of sequence identity between SR₁ and CR₁ in the region of ~250 residues upstream of the RT domains was lower (27%). When the carboxyl-terminal region of SR₁ was compared using the GAP algorithm with the same region from other non-LTR retrotransposons, SR₁ showed a markedly higher degree of aa identity to CR₁ and CR₁-like elements (from nonavian vertebrates [Burch, Davis, and Haas 1993; Vandergon and Reitman 1994]) than to other non-LTR retrotransposons from either invertebrates or other organisms (fig. 3). Indeed, this degree of identity between SR₁ and CR₁ could only be found between elements from the same host genome: Q and T1Ag from A. gambiae and Doc and Jockey from D. melanogaster, both 39% aa identity (data not shown).

Further comparison of the ~200 carboxyl-terminal residues of CR₁ and SR₁ showed the presence of two regions of high sequence identity. We have designated these two regions A and B (see fig. 4A for their locations). CR₁ and SR₁ were 61% identical over 63 aa in region A and 52% identical over 51 aa of region B. T1Ag, which had the next highest degree of identity overall to SR₁, displayed 40% aa identity within region A but lacked any degree of conservation in region B. The sequence identity between SR₁ and other non-LTR retrotransposons was less than 26%. A high level of sequence identity with CR₁ in the regions termed A and B (see fig. 4) has been used to identify CR₁-like elements in the genomes of the torpedo ray and frog (Burch, Davis, Haas 1993) and region B has been used to identify CR₁-like elements in reptiles (Vandergon and Reitman 1994). When the sequences from these CR₁-like elements from nonavian vertebrates were compared with SR₁ and CR₁ (fig. 4), most residues conserved among CR₁-like elements were also conserved in SR₁. By contrast, T1Ag showed conservation of CR₁-like residues in region A but lacked any degree of conservation in region B, and Doc showed a very low degree of conservation in region A but totally lacked identity to SR₁ in region B. Taken together, our results presented in figures 3 and 4 showed that the schistosome element SR₁ exhibits closer sequence identity to the CR₁ and CR₁-like elements of vertebrates than to other non-LTR retrotransposons.

Discussion

We have here described a new transposable element from the human blood fluke Schistosoma mansoni. The
element, which we have named SR1, is a non-LTR retrotransposon and is apparently the first protein-encoding transposable genetic element to be described from schistosomes or indeed from any member of the phylum Platyhelminthes. SR1 is similar in structure and sequence to CR1 and CR1-like non-LTR retrotransposons from birds and other vertebrates, which indicates that the distribution and origins of CR1-like elements may be more diverse than previously thought (Burch, Davis, and Haas 1993; Vandergon and Reitman 1994). While the frequent 5' truncations of SR1 is longer than that available for CR1 (Burch, Davis, and Haas 1993), it is the only CR1-like element so far reported from an invertebrate. The probable presence a CR1-like retrotransposon in the genome of an invertebrate raises the issue of the origin of this element. Only the conserved domains of RT in both the sequences and the lengths of their repeat units. The number of repeats at the 3' terminus of a CR1 element can vary from one to four (Burch, Davis, and Haas 1993; Vandergon and Reitman 1994). The tandem repeats may either be derived from a larger bank of tandem repeats in a master element or result from slippage during retrotransposition. It has been suggested that the tandem repeats may prime reverse transcription of CR1 elements by hybridization to homologous sequences at nicked chromosomal sites (Burch, Davis, and Haas 1993). Whereas all three of the SR1 3'-terminal fragments characterized here exhibited two copies of the repeat unit, two of the three shared downstream sequences resembling part of the repeat motif. Whether these sequences are indeed truncated repeats or represent insertion site preferences directed by the tandem repeats, as suggested for CR1 (Burch, Davis, and Haas 1993), is not yet clear. However, characterization of insertion sites at both the 5' and 3' termini of individual elements should yield information about the mechanism of integration of SR1.

As well as similarity in the structure of 3' termini, SR1 and CR1 exhibit a remarkable sequence identity in the region of the RT ORF carboxyl-terminal to the seven core motifs. High levels of sequence identity in this region indicate that CR1-like elements may exist in all classes of vertebrates (Burch, Davis, and Haas 1993; Vandergon and Reitman, 1994). Given the sequence conservation common to other CR1-like elements (Burch, Davis, and Haas 1993). SR1 appears to be a member of the CR1-like group. If SR1 is indeed CR1-like, it is the only CR1-like element so far reported from an invertebrate. The probable presence a CR1-like retrotransposon in the genome of an invertebrate raises the issue of the origin of this element. Only the conserved domains of the RTs of non-LTR retrotransposons display strong sequence identity between elements from phylogenetically distinct hosts (Xiong and Eickbush 1990). However, the regions carboxyl-terminal to the conserved domains of RT in both CR1 and SR1 display the same strong degree of identity as do the RT conserved domains. If CR1 and SR1 have descended from a CR1-like element found in the ancestor of both vertebrates and platyhelminths, i.e., if they have been vertically transmitted, the carboxyl-terminal region of these two transposons would have had to be under a degree of selection pressure similar to that acting on the RT conserved domains. Such selection pressure could only have been maintained if the carboxyl-terminal region of CR1-like elements plays some pivotal role in retrotrans-
A

![Diagram](#)

**Region A**

**Region B**

**FIG. 4.**—Conservation of amino acids between CR1-like elements and SR1 in the region carboxyl-terminal to RT conserved domains. A, Diagrammatic representation of a non-LTR retrotransposon showing the position of two highly conserved blocks of amino acids, A and B. B, Multiple sequence alignment of the carboxyl-terminal regions of SR1 and CR1-like retrotransposons from vertebrates. Positions where four or more sequences have identical amino acids are boxed. Sequences: CR1-like: lizard (L3 1503) and snake (D13384) from Vandergon and Reitman (1994); ray (X56517) and frog (M24187) from Burch, Davis and Haas (1993); chicken (CR1) (L22152); Dot, D. melanogaster (X17551); Tl, A. gambiae (M93689). The full sequences for the CR1-like elements from lizard, snake, frog, and ray have not been reported.

While the carboxyl-region from other non-LTR retrotransposons contains an integrase or endonuclease domain with a nucleic-acid-binding motif (Xiong and Eickbush 1988; Besansky et al. 1992), no similar motif is evident in CR1-like elements. However, if the close relationship of CR1 to SR1 is indeed functional, then the role of the carboxyl-terminal region may be related to the atypical tails of CR1 and SR1.

Position. While the carboxyl-region from other non-LTR retrotransposons contains an integrase or endonuclease domain with a nucleic-acid-binding motif (Xiong and Eickbush 1988; Besansky et al. 1992), no similar motif is evident in CR1-like elements. However, if the close relationship of CR1 to SR1 is indeed functional, then the role of the carboxyl-terminal region may be related to the atypical tails of CR1 and SR1.

While vertical transmission may explain the presence of a CR1-like element within the genome of *S. mansoni*, an alternative explanation may be that SR1 transferred from a vertebrate to the schistosome genome. Horizontal transmission of mobile genetic elements is evident when elements from unrelated species show a greater degree of sequence identity than is consistent with the phylogenetic relationship between the host organisms (Huock et al. 1991; Kidwell 1992; Robertson 1993). Thus, recent horizontal transmission is obvious for the class II element *mariner* where closely related elements coexist in the genomes of distantly related species of insects (Robertson 1993). By contrast, none of the (albeit small number of) CR1-like elements so far
reported provides an obvious candidate for having undergone horizontal transmission to schistosomes. After horizontal transmission, an element like SR1 can be expected to have evolved independently within the two host genomes for some time. Indeed, since SR1 elements display considerable sequence heterogeneity, it is unlikely that SR1 is a recent arrival in the S. mansoni genome. Horizontal transfer of a CR1-like, non-LTR retrotransposon from a vertebrate may have been facilitated by the evolutionary history of schistosomes as blood vessel parasites of reptiles, birds, and mammals. Conversely, schistosomes may have been the source of horizontal transfer of CR1-like elements to the genomes of their vertebrate hosts.

The issue of horizontal transmission versus vertical transmission and conservation of SR1 is further complicated by the possibility that after an ancient horizontal transmission to schistosomes, a CR1-like element may have been subject to selection pressure which has conserved the carboxyl-terminal region. Study of the evolution of SR1 within the genome of S. mansoni and closely related species, coupled with a search for and characterization of CR1-like elements from both invertebrates and vertebrates, might clarify the evolutionary origin of SR1. Moreover, as non-LTR retrotransposons form a very diverse group of elements, the clarification of the origin of SR1 may provide insights into evolution of not just CR1-like elements but also non-LTR retrotransposons at large.

At a practical level, given that the schistosome genome is now the subject of a World Health Organization-sponsored genome project (Johnston, 1997), SR1 elements may soon find utility as sequence-tagged sites or other tools in PCR-based genome mapping strategies (Hodgkin, Plasterk, and Waterston 1995). Further, they may be of use in the identification of clinical isolates and in associated epidemiological studies.

Acknowledgments

We thank M. Duke for maintenance of the schistosomes and V. Mann, D. McManus, and R. Slade for comments on the manuscript. This research was supported by the National Health and Medical Research Council and by an Australian Postgraduate Award to A.C.D.

The sequences reported here have been deposited in the GenBank database: 2337 bp consensus sequence, accession no. U66331; individual SR1 elements used to derive the consensus sequence, accession nos. U66332–U66339.

Literature Cited


THOMAS EICKBUSH, reviewing editor

Accepted February 10, 1997