Comparison of European Trichobilharzia species based on ITS1 and ITS2 sequences

J. DVOŘÁK*, Š. VAŇÁČOVÁ, V. HAMPL, J. FLEGR and P. HORÁK

Department of Parasitology, Charles University, Viničná 7, CZ-12844 Prague 2, Czech Republic

(Received 12 May 2001; revised 14 September 2001; accepted 14 September 2001)

SUMMARY

Schistosomes are parasites of considerable medical and veterinary importance and, therefore, all aspects of their biology have been intensively studied. In contrast, our knowledge of species of the largest genus, Trichobilharzia, is insufficient. Because morphological characterization of Trichobilharzia species provides a limited number of criteria for species determination, molecular data are required. In the present paper, we sequenced internal transcribed spacers ITS1 and ITS2, and 58S ribosomal RNA (rRNA) genes of 3 European Trichobilharzia species (T. regenti, T. szidati and T. franki). We showed that ITS1 and ITS2 sequences can be used in species identification. Repetitive elements were found in ITS1 of all 3 Trichobilharzia species; their number and length varied depending on the species. Phylogenetic analysis showed that the visceral T. franki is more related to the nasal T. regenti, than to the visceral T. szidati. The newly designed primer, which is specific for T. regenti, might be used as a tool for diagnosis of this potential pathogen.

Key words: Trichobilharzia, schistosomes, ribosomal DNA, repeats, recombinational hot-spot site, chi-like site.

INTRODUCTION

Trichobilharzia represents the largest genus of the family Schistosomatidae and is found exclusively in birds. Although more than 40 Trichobilharzia species have been described (Blair & Islam, 1983; Horák, Kolářová & Dvořák, 1998), their taxonomic validity is questionable. In Central Europe, 2 clearly defined species T. franki Müller et Kimmig, 1994 and T. regenti Horák, Kolářová et Dvořák, 1998 occur. The third species T. szidati Neuhaus, 1952 has at present an unclear taxonomical position and its name probably represents a junior synonym of T. ocellata (La Valette, 1855) Brumpt (1931).

Generally, in order to determine trichobilharzians, a description of the entire life-cycle, developmental stages and host–parasite interactions needs to be obtained. This is, however, extremely difficult and contributes to the confusion in the systematics. Experiments on T. regenti revealed that it is an exceptional parasite within schistosomes. Miracidia are able to hatch from the eggs directly in the host tissue (nasal mucosa), waiting for contact with water. In water bodies, the larvae seem to be strictly specific for the snails of Radix peregra and R. ovata. After several weeks of intramolluscan development, the cercariae are released and they can penetrate the skin of birds and mammals. In both hosts, the migrating schistosomula exhibit a high affinity for the central nervous system (CNS) (Horák et al. 1999; Horák & Kolářová, 2001). CNS (spinal cord and brain) seems to serve as a major route to the nasal cavity of birds where maturation of worms occurs. In mammals, the parasite dies in the CNS without mating and offspring production. T. regenti also possesses interesting characters at the cytogenetical level. A sub-population of the parasites, both males and females, has an additional chromosome (B chromosome) of an unknown function (Špakulová, Horák & Dvořák, 2001).

Because of growing importance of T. regenti as a neuropathogen, and the necessity to confirm taxonomic validity of trichobilharzian species and their phylogenetical relationships, an analysis of rDNA has been started in several laboratories. At present, however, the molecular data on Trichobilharzia spp. are very scarce. There is only 1 published report on the 28S rRNA gene from T. ocellata; serving as a representative of the genus in the study of phylogenetical relationships within the family Schistosomatidae (Snyder & Loker, 2000). In order to obtain more data characterizing European trichobilharzians, and to clarify the unusual intrageneric position of the nasal T. regenti, rDNA sequencing was performed. For comparative purposes, the same sequence analysis was applied for 2 visceral species: 2 isolates of T. szidati and 2 isolates of T. franki, developing in L. stagnalis and R. auricularia, respectively.

MATERIALS AND METHODS

Parasites

Cercariae of 5 Trichobilharzia isolates were analysed: T. regenti (Tr; released from snails of the species R. ovata/R. peregra; maintained as a laboratory
strain for 3 years (Horák et al. 1998); T. franki (Tf; and Tf; 2 isolates from R. auricularia; naturally infected snails from 2 different localities in Southern Bohemia—Tf, from the pond of Schwanzenberg, Tf, from the pond Podkadowský); T. szidati (Ts, and Ts; 2 isolates from Lymnaea stagnalis; Ts, isolate kept as a laboratory strain for 5 years according to the method described by Meuleman, Huyer & Mooij (1984); Ts, isolate from naturally infected snails found in the pond Schwanzenberg, Southern Bohemia).

Isolation of genomic DNA

The DNA was extracted from 5000 cercariae by a modified assay for DNA isolation from adults of Schistosoma mansoni (Minchella et al. 1994). After shedding from snails, the cercariae were collected from water by centrifugation, fixed in 96% ethanol and stored at −18 °C. Prior to DNA extraction, the cercariae were centrifuged at 5000 g for 5 min and washed 3 times in PBS buffer (pH 7.4). The pellet was mixed with 2 volumes of lysis buffer (50 mM Tris–HCl, pH 8.0; 100 mM EDTA; 100 mM NaCl; 1% SDS) and proteinase K (200 µg/ml) and the mixture was incubated at 65 °C for 1 h. To remove proteins, the suspension was mixed for 30 sec with an equal volume of water-saturated phenol, pH 8.2, followed by centrifugation at 10000 g for 5 min. The top aqueous phase was recovered and treated with an equal volume of chloroform/isoamylalcohol (24:1): phenol (1:1) and centrifuged as previously. The DNA was precipitated from the aqueous phase by addition of 2.5 volumes of 96% ethanol and sodium acetate (final concentration 30 mM) at −20 °C overnight. The precipitated DNA was centrifuged at 10000 g for 15 min. The pellet was washed with 100 µl of 70% ethanol, centrifuged at 10000 g for 15 min, dried and resuspended in 50 µl of sterile H2O.

PCR amplification of ITS regions

Polymerase chain reactions were carried out in a reaction volume of 50 µl. Primarily the whole region including ITS1, 5.8S rRNA and ITS2 sequences was amplified with primers 5′-GGAAGT-AAAAGTCCGTAACAAGG-3′ complementary to the conserved region at the 3′ end of the 18S rRNA gene and its forward (5′-TCCCTCCGCTTAT-TGATATGC-3′) complementary to the conserved region at the 5′ end of the 28S rRNA gene. The reaction mixture consisted of 2 µl of the DNA template, 6 µM of each primer (Generi Biotech), 1 × Taq polymerase buffer (Promega), 3 mM MgCl2 (Promega), 400 µM dNTP (Promega) and 5 units of Taq polymerase (Promega). The amplification profile consisted of an initial denaturation step at 95 °C for 5 min followed by 35 cycles of 60 sec at 95 °C, 45 sec at 50 °C, and 120 sec at 72 °C after the last cycle) in a PTC-200 thermal cycler (MJ Research).

Subcloning and sequence analysis

The PCR products were separated by agarose gel electrophoresis; the major band of the expected size was excised from the gel and the DNA was purified by phenol/chloroform extraction. The PCR products were inserted by the TA overhang method into the pCR 2.1-TOPO cloning vector and transformed to and amplified in chemically competent E. coli cells by using the TOPO TA cloning kit (version K2, Invitrogen). The alkaline lysis method was used for the preparation of plasmid DNA. Sequencing was performed with the Big dye Terminator Cycle Sequencing Ready Reaction kit (Perkin Elmer) following the manufacturer’s protocol. Two clones derived from 2 independent amplification reactions of each Trichobilharzia species were selected for sequence analysis. The products of sequencing reactions were analysed on an automatic genetic analyser ABI Prism 310 (Perkin Elmer). Sequencing primers M13 forward, M13 reverse (Perkin Elmer, Sigma), complementary to sites located adjacent to the polylinker plasmid regions, were used. The internal regions were sequenced with PCR primers annealing directly to the DNA insert: its5Trem and its4Trem (see above), its2Trem (5′-GCTGCACTCTTTCATCGACGC-3′) and its reverse complementary primer its3Trem (5′-GGGCTCATGAAAGAGTGCA-3′) designed to the conserved region at the 3′ end of 5.8S rRNA gene, its6Trem (5′-AGGAGGGCAGTCAATAGATG-3′) specific for ITS1 of T. regenti, its11Trem (5′-GACAGGCTAGTACCCC-3′) and its12Trem (5′-CCTGTTCTCAGCCATC-3′) specific for ITS1 of T. franki. DNA sequences were processed using the DNASTar package (DNASTar Inc.) and Chromas 1.5 (Technelysium). The nucleotide sequences obtained in this study were submitted to NCBI/GenBank database under the following accession numbers: T. regenti – AF263829, T. szidati – AF263828, T. franki – AF356845.

Phylogenetic analysis

The identified ITS1 and ITS2 sequences of Trichobilharzia isolates were used to infer phylogenetic relationships among these species. The part of the sequences with repetitive elements was removed from the sequences prior to the analysis.

The sequence data were aligned using the clustalX program. Programs from the PHYLIP package 3.57c were used to construct phylogenetic trees and to estimate their robustness. The TREE-PUZZLE 5.0 was used to estimate transition–transversion ratio and compute base frequencies. The phylogenetic trees were rooted using the sequences of ITS1 and ITS2 of Schistosoma japonicum (GenBank accession
Molecular comparison of European Trichobilharzia species

Fig. 1. PCR products (amplification with its4Trem and its5Trem) of 5 Trichobilharzia isolates. The first line represents molecular weight markers. The following lines are PCR products: Tr, *T. regenti*; Ts, *T. szidati* laboratory strain; Tf and Tf', *T. franki*; 0, negative control (without any DNA). The sequenced products of parasite rDNA are always the main bands on the top of each line.

Table 1. The lengths of the analysed rDNA regions of *Trichobilharzia* species

<table>
<thead>
<tr>
<th>Sequenced region</th>
<th><em>T. regenti</em></th>
<th><em>T. szidati</em></th>
<th><em>T. franki</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total length of PCR products</td>
<td>1914 bp</td>
<td>1330 bp</td>
<td>1550 bp</td>
</tr>
<tr>
<td>ITS1</td>
<td>1325 bp</td>
<td>745 bp</td>
<td>966 bp</td>
</tr>
<tr>
<td>ITS2</td>
<td>322 bp</td>
<td>318 bp</td>
<td>322 bp</td>
</tr>
<tr>
<td>5S-8S</td>
<td>155 bp</td>
<td>155 bp</td>
<td>155 bp</td>
</tr>
<tr>
<td>18S (partial sequence)</td>
<td>53 bp</td>
<td>53 bp</td>
<td>53 bp</td>
</tr>
<tr>
<td>28S (partial sequence)</td>
<td>59 bp</td>
<td>59 bp</td>
<td>59 bp</td>
</tr>
</tbody>
</table>

numbers U82282 and U22167) as the outgroup. The phylogenetic trees were constructed by neighbor-joining (NJ), maximum parsimony (MP) and maximum likelihood (ML) methods. The NJ tree was constructed from the distance matrix corrected according to the Kimura two-parameter model for nucleotide change using the estimated transition-transversion ratio. The MP was conducted with the switch ‘jumble order for the taxa addition’ (repetition 10×). The ML tree was constructed using the estimated transition–transversion ratio and the calculated base frequencies. Global rearrangements and 10× random addition of taxa were used in the ML tree search. In total, 100 bootstrap replicates were carried out with each of these methods to estimate the tree robustness. The trees were drawn using the TreeView program.

Dot-plot analysis

Dot-plot is a method that visually compares 2 sequences and looks for regions of similarity between them. The sequences to be compared are arranged on the axes of the matrix. At every point in the matrix where the 2 sequences are identical or similar a dot is placed (i.e. at the intersection of every row and column that have the same letter or similar oligonucleotide in both sequences). In our study this method was used for search of repetitive elements within sequences. For this purpose the analysed sequence was put on both axes. The dot-plot was done with a filter in which 2 oligonucleotides are considered similar if at least 8 of their 10 bases match. When repetitive elements existed within the sequence, shorter lateral diagonals appear in addition to the central one. The same method was also used in the search for repetitive elements arranged in reverse, complement or reverse-complement order. The number of repetitive elements was counted by this method as \[N = Nd + 1\], where \(N\) is the number of repetitive elements and \(Nd\) is the number of lateral diagonals on one side of the central diagonal. Program Seaview was used to perform this analysis.

RESULTS

PCR products of different lengths were obtained by amplification of DNA from the 3 analysed *Trichobilharzia* species using the its5Trem and its4Trem

Fig. 2. Dot-plot matrices of pairwise alignments. This method was used for search of repetitive elements in the sequences. In each case (A, B, C) the examined sequence was arranged on both axes. The number of repetitive elements is \(N = Nd + 1\) (\(N\) is the number of repetitive elements, \(Nd\) is the number of lateral diagonals on one side of the central diagonal). (A) ITS1 + 5S-S rRNA + ITS2 of *Trichobilharzia regenti*; (B) ITS1 + 5S-S rRNA + ITS2 of *T. franki*; (C) ITS1 + 5S-S rRNA + ITS2 of *T. szidati*. 
primers. The products contained ITS1, ITS2, 5.8S and partial sequences of 18S and 28S rRNA genes. In the case of \textit{T. regenti}, the PCR products were 1914 bp in length, both isolates of \textit{T. szidati} yielded 1330 bp products and the 2 isolates of \textit{T. franki} gave 1555 bp products (Fig. 1). Based on sequence analysis with the its5Trem and its4Trem primers, PCR products from Tf1 isolate in length of 920 bp and from Tf2 in length of 857 bp were recognized as an algal contamination. Two minor PCR products obtained for \textit{T. regenti} of approximately 500 and 400 bp in length were not further analysed.

The 5.8S (155 bp) and partial 28S (59 bp) and 18S (53 bp) sequences of rRNA genes of the 3 \textit{Trichobilharzia} species were identical and they were similar to other schistosome sequences in NCBI/GenBank databases.

Sequencing of PCR products showed clear interspecific differences in the length and nucleotide composition of ITS1 and ITS2 (for details see Table 1). The ITS1 region showed a higher degree of interspecies variability in nucleotide sequence than the ITS2 region. The sequences of the 2 \textit{T. franki} isolates were identical and the comparison of the analysed sequences of the 2 \textit{T. szidati} isolates showed differences in only 2 nucleotide positions in ITS1.
In this work we focused on the analysis of ITS regions of species within the genetically poorly characterized genus *Trichobilharzia*. The sequence analysis of the laboratory strain of *T. szidati* and the wild strain of *Trichobilharzia* from *L. stagnalis* showed that these 2 isolates differed in only 2 nucleotide positions within the ITS1 region; both were named as *T. szidati*. The sequences of 2 isolates of *T. franki* were identical. The algal contaminations were probably from the snail shells and the method of material isolation did not eliminate this impurity. The sequence alignment of the 3 species revealed considerable differences in nucleotide sequences and lengths of both ITS regions; this was more pronounced in the ITS1 region.

Phylogenetic analysis, based on the sequence data of ITS1 and ITS2 regions, was performed. It confirmed previous determinations of *Trichobilharzia* species, based on the study of their life-cycles and morphology. The analysis showed a closer phylogenetic relationship of the visceral fluke *T. franki* with the nasal *T. regenti*, rather than with the visceral *T. szidati*. This result, however, corresponds to parasite specificity towards the snail intermediate hosts. *T. franki* and *T. regenti* infect the same snail genus (*Radix*) i.e. *T. regenti* parasitizes *R. peregra*/*R. ovata* (Horák et al. 1998) and *T. franki* develops in *R. auriculata* (Müller & Kimmig, 1994). Moreover, the shape of eggs of both species is similar. The phylogenetically distant *T. szidati* uses a different snail genus (*Lymnaea*) and differs also in egg morphology (Neuhaus, 1952). In view of these data, the neurotropic behaviour and final location of *T. regenti* in the nasal mucosa seem to be an apomorphic trait. However, the same location of adult flukes in the host nasal area has been recorded for 7 other *Trichobilharzia* species (see Horák et al. 1998 for review) and *S. nasale*; therefore, the polarity and number of the evolutionary change(s) within the schistosome’s phylogenetic tree is still unclear.

We detected repeating elements in ITS1 region of all 3 species. There was no significant homology between these repetitions and GenBank database sequences. Although these elements vary in length and slightly in nucleotide sequences, among the investigated trichobilharzian species, they contain a conserved sequence motif that could possibly serve for *Trichobilharzia* determination based on PCR probes. From the medical point of view, namely due to the neuropathogenicity of *T. regenti* (Horák et al. 1999), a finding of species-specific markers distinguishing *T. regenti* from the other species is desirable. We believe that the newly designed oligonucleotide primer its6Trem, specific to *T. regenti*, might serve for specific PCR determination of parasite cercariae or developing flukes in the nervous
tissue, as no cross-reaction was observed with the other 2 trichobilharzians. The existence of a dense smear in the electrophoretogram of _T. regenti_ DNA amplified with the primers its6Trem and its2Trem suggests that multiple copies of the repetitive motif may exist in the _T. regenti_ genome.

Repetitive sequences were previously reported from the ITS1 region of _Schistosoma_ species. These elements were variable in sequence patterns, and their number, depending on the species. The ITS1 region of _S. haematobium_ contains 2 repeats, _S. intercalatum_ 2, _S. mattheei_ 4 or 5, _S. spinale_ 7, _S. magrebowiei_ 4 and _S. mansoni_ 2 (Kane & Rollinson, 1994; Kane _et al._ 1996). In the _S. japonicum_ group of species, the number of tandemly repeated elements in the ITS1 region varies with as many as 7 repeats present (Van Herwerden, Blair & Agatsuma, 1998). However, there is no evidence for any biological function of these repetitions.

Chi-site sequences (5'-GCTGATTGG-3') were described as _recBCD_ enzyme-mediated recombinational hot-spot sequences in phage λ (Thaler & Stahl, 1988) and in the genome of _E. coli_ (Ponticelli _et al._ 1985; Thaler & Stahl, 1988). Also within repetitive elements of _schistosomes_ sequences resembling chi-site were detected. In detail, chi-like sequence 5'-GCTTGGTTGG-3' from repeats of _S. mattheei_, _S. haematobium_, _S. intercalatum_, _S. spinale_, 5'-GCTGATTGG-3' from _S. magrebowiei_ and 5'-GCTTGGTTGG-3' from _S. mansoni_ were reported (Kane _et al._ 1996). Kane _et al._ (1996) proposed that these sequences are inactive relics or they have a recombination-inducing function. Van Herwerden _et al._ (1998) hypothesized that the absence of the chi-like sites may cause a high number of intragenomic variants of ITS1 region of the _S. japonicum_ complex. We found similar chi-like sequences (e.g. 5'-GCTTGGTG-3', 5'-CTCGATTGG-3' and 5'-CTTGGTTGG-3') in repetitive elements of all 3 _Trichobilharzia_ species. The chi-like site 5'-GCTTGG-3' is present in each repetitive element of the investigated trichobilharzians. It is identical in 6 nucleotide positions with the chi-like site of _Schistosoma_ species and outside of this area, it contains an additional conservative doublet of nucleotides 5'-TG-3', the same as the doublet associated with chi-like site of _Schistosoma_ species. The chi-like sites (5'-CTCGATTGG-3' and 5'-CTTGGTTGG-3') identical in 8 nucleotide positions are present only in _T. franki_ repetitive elements.

In addition, we found other potential recombinational sequence hot-spots: 5'-AGCAC-3' in repetitive elements of all 3 _Trichobilharzia_ species and 5'-AGCTACT-3', in reverse complementary form, present exclusively in repetitive elements of _T. franki_. A similar recombinational site 5'-AGCAC-3' occurs in the genome of _N. crassa_ (Fraser, Koa & Chow, 1990). Again, the function of these sequences in trichobilharzian genome is unknown.

We conclude that our results may have practical implications. It has been shown that trichobilharzians are not only parasites of birds and agents of cercarial dermatitis of mammals. It is known that the schistosomula can even migrate to the mammalian lungs and survive there for a limited period (Horák & Kolářová, 2000, 2001). Moreover, the schistosomula of _T. regenti_ are neurotropic in birds and experimental mammals and sometimes cause neurologic disorders (Horák _et al._ 1999; Horák & Kolářová, 2001). From the medical viewpoint, it is therefore desirable to be able to diagnose these species. Unfortunately, cercariae of all 3 investigated species are indistinguishable by (electron) microscopy (Horák & Dvořák, unpublished observations). The main diagnostic criteria are to be found in adult worms. However, this usually requires experimental infections of birds, the number of which is restricted due to ethical reasons. From our results it is evident that ITS1 and ITS2 sequences can serve not only for

---

**Fig. 5.** Phylogenetic tree of 4 isolates of _Trichobilharzia_. The tree was constructed by the neighbor-joining method on the basis of ITS1 and ITS2 sequences. Sequences of _Schistosoma japonicum_ were used for rooting the tree. Bootstrap values for 100 resamplings are indicated on each node. The topologies and the bootstrap values of the trees constructed by the maximum parsimony and the maximum likelihood method were identical. Branch lengths are proportional to the estimated number of nucleotide changes. The analyses have been done after removal of repetitions from the sequences. Note: If schistosome sequences of ITS1 ( _S. mekongi_ , _S. malayensis_ , _S. mansoni_ and _S. japonicum_ ) and ITS2 ( _S. bovis_ , _S. haematobium_ , _S. mekongi_ , _S. mansoni_ , _S. japonicum_ , _S. intercalatum_ , _S. malayensis_ and _S. hippocotomai_ ) were used, the branching order of trichobilharzian species remained the same. The bootstrap support for the branch of _T. regenti_ and _T. franki_ never decreased below 50% and, in most cases, it was higher than 70%. This was independent of the tree construction method.
construction of phylogenetic trees and clarification of species relationships among trichobilharzians, but also for designing specific primers that can be used as species-specific markers, enabling diagnosis of the infection and/or determination of the parasite.

This work was supported by the Charles University (Grant No. 106/1998/B/BI) and the Czech Ministry of Education (Grant No. 1131-4). We wish to thank Dr S. D. Snyder for stimulating this work and subsequent advice on primer design, M. Šedinová for technical assistance and J. Rudolfová for providing \textit{T. franki} isolates.

REFERENCES


