Journal of Helminthology (2002) **76,** 7–19 DOI: 10.1079/JOH200191

Affinities between Asian non-human *Schistosoma* species, the *S. indicum* group, and the African human schistosomes

T. Agatsuma[1]*, M. Iwagami[1], C.X. Liu[2], R.P.V.J. Rajapakse[3], M.M.H. Mondal[4], V. Kitikoon[5], S. Ambu[6], Y. Agatsuma[7], D. Blair[8] and T. Higuchi[2†]

[1]Department of Environmental Health Science, and [2]Department of Physiology, Faculty of Medicine, Kochi Medical School, Oko, Nankoku 783-8505, Japan; [3]Department of Veterinary Para-clinical Studies, University of Peradenyia, Peradenyia, Sri Lanka; [4]Department of Veterinary Medicine, Bangladesh Agriculture University, Mymensingh, Bangladesh; [5]Department of Social and Environmental Medicine, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand; [6]Division of Medical Ecology, Institute for Medical Research, Kuala Lumpur, Malaysia; [7]Laboratory of Clinical Technology, Kochi Gakuen College, Asahi-Tenjin, Kochi 780-0955, Japan; [8]Department of Zoology and Tropical Ecology, James Cook University, Townsville, Queensland 4811, Australia

Abstract

Schistosoma species have traditionally been arranged in groups based on egg morphology, geographical origins, and the genus or family of snail intermediate host. One of these groups is the '*S. indicum* group' comprising species from Asia that use pulmonate snails as intermediate hosts. DNA sequences were obtained from the four members of this group (*S. indicum*, *S. spindale*, *S. nasale* and *S. incognitum*) to provide information concerning their phylogenetic relationships with other Asian and African species and species groups. The sequences came from the second internal transcribed spacer (ITS2) of the ribosomal gene repeat, part of the 28S ribosomal RNA gene (28S), and part of the mitochondrial cytochrome *c* oxidase subunit 1 (CO1) gene. Tree analyses using both distance and parsimony methods showed the *S. indicum* group not to be monophyletic. *Schistosoma indicum*, *S. spindale* and *S. nasale* were clustered among African schistosomes, while *S. incognitum* was placed as sister to the African species (using ITS2 and 28S nucleotide sequences and CO1 amino acid sequences), or as sister to all other species of *Schistosoma* (CO1 nucleotide sequences). Based on the present molecular data, a scenario for the evolution of the *S. indicum* group is discussed.

*Fax: +81 88 880 2535
E-mail: agatsuma@med.kochi-ms.ac.jp
†Present address: Department of Physiology, Faculty of Medicine, Fukui Medical University, Matsuoka, Fukui 910-1193, Japan.

Introduction

Evolutionary relationships among species of the medically important genus *Schistosoma* have been the focus of intense interest (e.g. Després *et al.*, 1992, 1995; Johnston *et al.*, 1993; Littlewood & Johnston, 1995; Bowles *et al.*, 1995; Barker & Blair, 1996; Blair *et al.*, 1997a; Agatsuma *et al.*, 2001). Nineteen species of *Schistosoma* are currently recognized, distributed among four groups on the basis of egg morphology, geographical origins and the identity of the snail intermediate hosts (Rollinson & Southgate, 1987). Three of the groups (the *S. mansoni* and *S. haematobium* groups, both primarily in Africa, and the *S. japonicum* group in east Asia) include the most important trematode parasites of man. The fourth group is the less well defined *S. indicum* group, (*S. indicum*, *S. spindale*, *S. nasale*, *S. incognitum*) which does not include any human parasites. Despite the volume of work on *Schistosoma* species, little molecular information is available on the *S. indicum* group (Johnston *et al.*, 1993; Littlewood & Johnston, 1995; Barker & Blair, 1996; Agatsuma *et al.*, 2001). In the present study, sequences from the second internal transcribed spacer (ITS2) of the ribosomal gene repeat, part of the 28S ribosomal RNA gene (28S) and of the mitochondrial cytochrome *c* oxidase subunit 1 (CO1) gene were obtained from all four members of the *S. indicum* group. 28S sequences were also obtained from *S. malayensis*, *S. mekongi* and *S. sinensium*. These sequences were used to determine whether the *S. indicum* group is monophyletic. The relationships of group members to other *Schistosoma* species and the evolution of the genus *Schistosoma* are discussed.

Materials and methods

Schistosomes

Adult worms of *S. indicum* and *S. nasale* were recovered from naturally infected cattle in abattoirs in Mymensingh, Bangladesh in 1997 and in Colombo, Sri Lanka in 1998 respectively. For *S. spindale*, naturally infected snails (*Indoplanorbis* sp.) were collected in Malaysia and adult worms recovered from an experimentally infected goat maintained in the laboratory at the Institute for Medical Research. Adult worms of *S. incognitum* were obtained from naturally infected *Bandicota indica* trapped in Phitsanulok, Thailand in 1997. For *Schistosoma malayensis*, *S. mekongi* and *S. sinensium*, the same DNA templates were used as in Agatsuma *et al.* (2001).

DNA extraction, PCR and sequencing

Genomic DNA was extracted from whole, individual adult worms for each species. Worms were incubated in extraction buffer containing proteinase K until the tissues were solubilized. The solubilized samples were treated with an equal volume of phenol (approximately pH 8.0) twice, and treated once with an equal volume of chloroform. Extracted DNAs were ethanol-precipitated. The second internal transcribed spacer (ITS2) of the ribosomal gene repeat, part of the 28S RNA gene (28S) including the three variable domains D1, D2 and D3 according to the description of Littlewood & Johnston (1995) and part of the mitochondrial cytochrome c oxidase subunit 1 (CO1) gene were amplified using the polymerase chain reaction (PCR). The PCR conditions were as follows: 94°C for 1 min, 50°C for 2 min, 72°C for 3 min, for 30 cycles. Amplification reactions were performed in a final volume of 50 μl containing primers (3.2 pmol), deoxynucleoside triphosphates (dNTPs, 0.2 mM), and Taq polymerase (1.75 U/reaction). Primers for the ITS2 are 5′-CGG TGG ATC ACT CGG CTC GT-3′ (3S, forward direction) and 5′-CCT GGT TAG TTT CTT TTC CTC CGC-3′ (A28, reverse diection) (Bowles *et al.*, 1995), for the 28S are 5′-ACC CGC TGA AYT TAA GCA-3′ (LSU5, forward direction), and 5′-TCC TGA GGG AAA CTT CGG-3′ (LSU3, reverse direction) (Littlewood & Johnston, 1995), and for the CO1 region are 5′-TTT TTT GGG CAT CCT GAG GTT TA-3′ (FH3, forward direction) and 5′-TAA AGA AAG AAC ATA ATG AAA ATA ATC-3′ (FH5, reverse direction) (Bowles *et al.*, 1993). The PCR products were purified using MicroSpin Columns (Pharmacia Biotech). Purified DNA was precipitated with ethanol, resuspended in 20 μl of distilled water and aliquots sequenced using the Bigdye kit (ABI). PCR primers were used as sequencing primers for all three regions. For the 28S region, two additional primers, 5′-GTA CCG TGA GGG AAA GTT G-3′ (TSD2, forward direction) and 5′-GTC CGT GTT TCA AGA CGG G-3′ (D4AR, reverse direction) were used as sequencing primers (Littlewood & Johnston, 1995). The reactions were purified according to the manufacturer's instructions (ABI) and applied to an ABI sequencer (373A).

Sequencing for the 28S region was performed using only a single male individual from each species, since the 28S gene is believed to be highly conserved. However, several individual worms in each species were sequenced for the other two regions. One male and one female were sequenced for both ITS2 and CO1 in *S. nasale*. Three males collected each from cattle, goat and sheep were sequenced for both of the two regions in *S. indicum*. Two males each were sequenced for both two regions in *S. incognitum* as well as *S. spindale*.

Alignment and tree analysis

Analyses for multiple sequence alignments were done using the programs CLUSTAL V (Higgins *et al.*, 1992) and GENETYXMAC ver. 6.0 (Software Development Co., Tokyo, Japan). The genetic code used was derived from Blair *et al.* (1999) and implemented in DNASIS ver. 3.2 (Hitachi Software Engineering Co., Japan 1994). Phylogenetic analysis was performed using distance and parsimony methods in MEGA (ver. 1.01) and PAUP (ver. 3.1.1), respectively. For distance analyses, the Kimura two-parameter method was used to construct the distance matrix and the tree inferred from this using the neighbour-joining (NJ) approach. Parsimony trees were inferred using the heuristic search option available in PAUP. Bootstrap resampling (1000 cycles) was performed for each method to assess tree topology. Other sequences used in this study, many of which have been reported elsewhere (Agatsuma *et al.*, 2001; Blair

https://doi.org/10.1079/JOH200191 Published online by Cambridge University Press

Table 1. Source of materials of species from four genera, *Schistosoma*, *Orientobilharzia*, *Schistosomatium* and *Paragonimus*.

Species	Countries surveyed	ITS2	28S	CO1
Schistosoma japonicum	–	U22167	Z46504	U82264
S. malayensis	–	U82398	Present study	U82262
S. mekongi	–	U82398	Present study	U82263
S. sinensium China	–	Agatsuma *et al.* (2000)	Present study	Agatsuma *et al.* (2000)
S. indicum	Bangladesh	Present study	Present study	Present study
S. spindale	Thailand	Present study	Present study	Present study
S. nasale	Sri Lanka	Present study	Present study	Present study
S. incognitum	Thailand	Present study	Present study	Present study
S. bovis	–	AF146035	–	–
S. intercalatum	–	U22166	–	U22160
S. haematobium	–	U22165	Z46521	U82266
S. mansoni	–	U22168	Z46503	U82265
S. hippopotami	–	Després *et al.* (1995)	–	–
Orientobilharzia turkestanicum	–	–	AF167092	–
Schistosomatium douthitti	–	Bowles *et al.* (1995)	AF167087	–
Paragonimus westermani	–	–	–	Blair *et al.* (1997a)
P. miyazakii	–	–	–	Blair *et al.* (1997a)

et al., 1997a; Bowles *et al.*, 1995; Després *et al.*, 1995) are given in table 1. Sequences from two other schistosomes of mammals, *Schistosomatium douthitti* and *Orientobilharzia turkestanicum*, were used as outgroups for the ITS2 and 28S data set (Bowles *et al.*, 1995; Snyder & Loker, 2000). Partial CO1 sequences of *Paragonimus* species (Blair *et al.*, 1997b) were used as outgroups for the CO1 data set.

Results

Alignments

Alignments of the ITS2, 28S and CO1 nucleotide sequences are shown in figs 1, 2 and 3a. Figure 3b shows the alignment of the CO1 amino acid sequences. Tables 2–4 show pairwise differences and transition/transversion ratios among ITS2, 28S and CO1 nucleotide sequences respectively. In all calculations, gaps were ignored. The 28S sequence from *S. spindale* obtained in the present study was identical to that reported in Littlwood & Johnston (1995).

ITS2 region

The alignment of the ITS2 region (308 bp for the *S. indicum* group) is shown in fig. 1. Table 2 provides pairwise differences and transition/transversion ratios among ITS2 nucleotide sequences (excluding gaps). No transversions were observed between *S. indicum* and *S. spindale* and only a few were observed between these species and *S. nasale* (Ts/Tv ratios 4.5–7.0). Pairwise differences among these three species were small (1.6–5.0%) and, taken together with the Ts/Tv ratios, suggest recent divergences. On the other hand, *S. incognitum* showed a low value for Ts/Tv ratios in comparison with other members of the group (mean 1.50) and relatively high pairwise differences (5.6–8.1%), showing that this species has diverged early within the *S. indicum* group. A good alignment was obtained between the *S. indicum* and the African group as well as within the *S. japonicum* group, with a few gaps. On the other hand, when the *Schistosomatium douthitti* sequences were included, numerous gaps, 1–24 bases long had to be inserted in the alignment. As shown in fig. 1, all four species of the *S. indicum* group shared more gaps with the African group than with the *S. japonicum* group.

28S region

The alignment of the 28S region (1252 bp for the *S. indicum* group, and approximately 1275 bp for the *S. japonicum* group) in the partial region of 28S rRNA gene is shown in fig. 2, including the three variable domains D1, D2 and D3 according to the description of Littlewood & Johnston (1995). There was no sequence difference present between *S. spindale* from Malaysia (present data) and from Sri Lanka (Littlewood & Johnston, 1995). Table 3 provides pairwise differences and transition/transversion ratios among the 28S nucleotide sequences (excluding gaps). There were no transversions observed between *S. indicum* and *S. spindale* and a few observed between these species and *S. nasale* (Ts/Tv ratios 7.7–8.0) as in ITS2. Pairwise differences among these three species were small (0.7–2.6%). These small values for pairwise differences and high ratios of Ts/Tv suggest recent divergences again. On the other hand, *S. incognitum* showed a low value for Ts/Tv ratios in comparison with other members of the group (mean 3.7) and relatively high pairwise differences (mean 4.2%), showing again as in ITS2 that this species has diverged early within the *S. indicum* group.

There were only few gaps between the *S. indicum* and the African group as well as within the *S. japonicum* group. On the other hand, a number of gaps (each 2–7 bases long) were observed in the alignment between the *S. indicum* and *S. japonicum* groups.

CO1 region

The alignments of the partial CO1 nucleotide (372 bp) and amino acid sequences are shown in fig. 3. Table 4 provides pairwise differences and transition/transversion ratios among CO1 nucleotide sequences. Ts/Tv ratios in the CO1 region form a striking contrast to those in the ITS2 region; most of the values were below 1.00,

https://doi.org/10.1079/JOH200191 Published online by Cambridge University Press

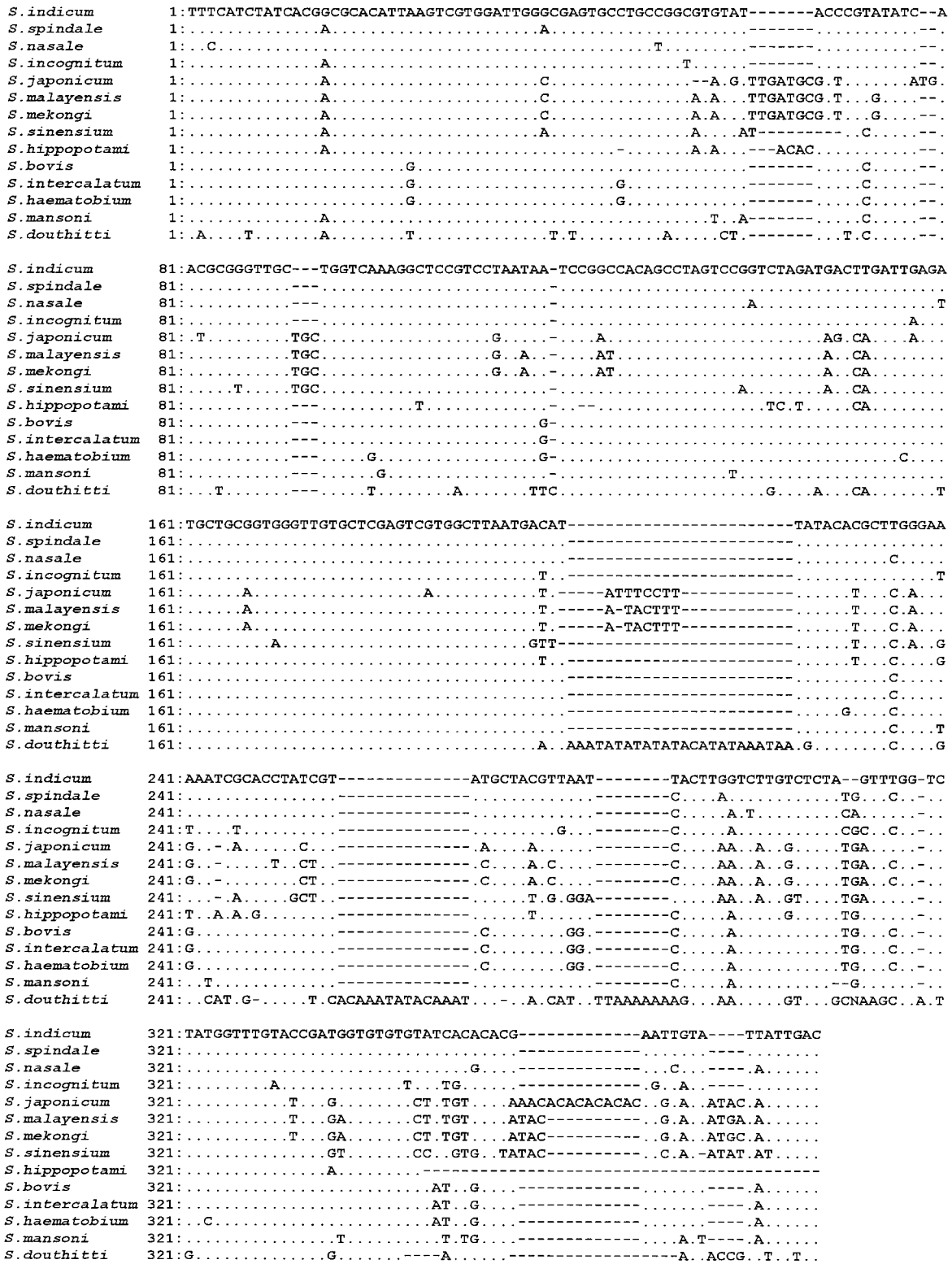

Fig. 1. Nucleotide sequence alignment of the second internal transcribed spacer (ITS2) of the ribosomal gene repeat among species in the genera *Schistosoma* and *Schistosomatium*. '.' indicates identity with the top sequence; '-' indicates an alignment gap.

https://doi.org/10.1079/JOH200191 Published online by Cambridge University Press

Table 2. Pairwise differences (%) and transitions/transversions in ITS2 sequences between species of two genera, *Schistosoma* and *Schistosomatium*.

Species	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. *S. japonicum*	–	15/6	14/5	26/14	31/12	29/12	36/12	24/19	24/8	32/12	34/13	37/13	27/17	40/33
2. *S. malayensis*	6.2	–	1/1	26/13	34/11	30/11	39/11	28/18	28/7	32/11	34/12	37/12	29/16	38/33
3. *S. mekongi*	5.6	0.6	–	26/12	33/11	29/11	38/11	27/18	27/7	31/11	33/12	36/12	28/16	36/34
4. *S. sinensium* China	12.4	12.0	11.7	–	33/13	32/13	38/13	32/20	23/9	34/13	35/14	38/14	29/13	32/33
5. *S. indicum*	13.7	14.3	14.0	14.8	–	5/0	9/2	11/7	16/6	13/2	14/3	17/3	11/7	26/27
6. *S. spindale*	13.0	12.9	12.6	14.4	1.6	–	14/2	11/7	17/6	14/2	15/3	18/3	12/7	25/27
7. *S.nasale*	15.2	15.8	15.5	16.3	3.6	5.0	–	17/9	19/7	16/2	17/3	20/3	13/7	29/27
8. *S. incognitum*	13.6	14.5	14.2	16.7	5.7	5.6	8.1	–	14/11	16/9	17/10	20/10	12/12	24/31
9. *S. hippopotami*	11.3	12.4	12.0	11.6	7.9	8.2	9.2	8.9	–	20/6	20/6	24/6	13/11	28/22
10. *S. bovis*	14.1	13.7	13.4	15.2	4.8	5.0	5.7	7.9	9.3	–	1/0	4/0	12/7	27/28
11. *S. intercalatum*	14.9	14.5	14.2	15.7	5.3	5.6	6.3	8.4	9.2	0.3	–	5/0	13/8	28/29
12. *S. haematobium*	15.9	15.5	15.1	16.6	6.3	6.6	7.2	9.4	10.6	1.3	1.6	–	16/8	30/29
13. *S. mansoni*	14.0	14.3	14.0	13.5	5.7	6.0	6.3	7.5	8.6	6.0	6.6	7.5	–	24/32
14. *Schistosomatium douthitti*	23.1	22.4	22.1	21.2	17.3	16.9	18.2	17.9	18.1	18.0	18.5	19.2	18.3	–

Values above the diagonal are transitions/transversions. Those below are pairwise differences (%).

Table 3. Pairwise differences (%) and transitions/transversions in 28S sequences between species of three genera, *Schistosoma*, *Orientobilharzia* and *Schistosomatium*.

Species	1	2	3	4	5	6	7	8	9	10	11	12
1. *Schistosoma japonicum*	–	27/3	30/3	35/18	52/29	55/29	55/27	45/28	55/26	51/30	68/35	88/45
2. *S. malayensis*	2.8	–	6/0	39/19	55/30	56/30	52/28	48/29	58/27	52/32	69/36	91/44
3. *S. mekongi*	3.1	0.6	–	43/19	54/30	55/30	51/28	47/29	57/27	50/32	69/36	90/44
4. *S. sinensium* China	5.0	5.5	5.8	–	55/29	57/29	53/29	47/30	55/26	54/29	63/32	79/46
5. *S. indicum*	7.8	8.2	8.1	8.1	–	7/0	24/3	33/9	19/2	30/8	42/26	74/49
6. *S. spindale*	8.1	8.3	8.2	8.3	0.7	–	23/3	36/9	22/2	33/8	45/26	76/49
7. *S. nasale*	7.9	7.7	7.6	7.9	2.6	2.5	–	35/10	27/3	30/9	46/25	76/50
8. *S. incognitum*	7.0	7.4	7.3	7.4	4.0	4.3	4.3	–	32/7	30/11	40/27	70/46
9. *S. haematobium*	7.8	8.2	8.1	7.8	2.0	2.3	2.9	3.8	–	34/5	46/23	70/47
10. *S. mansoni*	7.8	8.1	7.9	8.0	3.7	3.9	3.8	3.9	3.8	–	46/26	75/50
11. *Orientobilharzia turkestanicum*	9.9	10.1	10.1	9.2	6.6	6.8	6.9	6.5	6.7	7.0	–	79/47
12. *Schistosomatium douthitti*	12.8	13.0	13.0	12.0	11.8	12.0	12.1	11.2	11.3	12.1	12.1	–

Values above the diagonal are transitions/transversions. Those below are pairwise differences (%).

Table 4. Pairwise differences among CO1 sequences in species of the genera *Schistosoma* and *Paragonimus*.

Species	1	2	3	4	5	6	7	8	9	10	11	12
1. *S. japonicum*	–	24/29	20/28	31/38	34/53	36/46	28/47	27/53	29/49	31/45	28/47	60/52
2. *S. malayensis*	14.2	–	22/7	29/37	33/50	29/51	27/50	28/52	28/51	27/48	33/42	53/55
3. *S. mekongi*	13.0	7.8	–	22/38	31/48	33/49	22/48	29/49	27/50	22/47	29/48	53/54
4. *S. sinensium* China	18.5	17.7	16.2	–	38/47	37/50	24/51	37/55	38/50	33/51	37/55	58/52
5. *S. indicum*	23.4	22.3	21.3	22.8	–	23/23	29/30	34/52	26/41	29/40	29/42	60/57
6. *S. spindale*	22.0	21.5	22.1	23.4	12.4	–	26/41	32/61	26/36	22/39	26/41	65/58
7. *S. nasale*	20.2	20.7	18.9	20.2	15.9	18.0	–	33/52	24/31	18/34	23/44	61/63
8. *S. incognitum*	21.5	21.5	21.0	24.7	23.1	25.0	22.8	–	31/53	34/54	35/50	58/51
9. *S. intercalatum*	21.1	21.4	20.9	23.8	18.1	16.8	14.9	22.7	–	22/18	30/39	58/64
10. *S. haematobium*	20.4	20.2	18.6	22.6	18.5	16.4	14.0	23.7	10.8	–	26/44	63/67
11. *S. mansoni*	20.2	20.2	18.1	24.7	19.1	18.0	18.0	22.8	18.6	18.8	–	65/67
12. *Paragonimus westermani*	30.1	29.0	28.8	26.9	31.5	33.1	33.3	29.3	32.8	34.9	35.5	–

Values above the diagonal are transitions/transversions. Those below are pairwise differences (%).

suggesting saturation. Some exceptions included *S. indicum*/*S. spindale* (1.00), *S. malayensis*/*S. mekongi* (3.14), and *S. intercalatum*/*S. haematobium* (1.22), which correspond with lower values of percentage differences (7.8–12.4%) for each pairwise comparison. Generally speaking, saturation has occurred throughout most of all species level in the CO1 region. As shown in table 4, differences in amino acid substitutions showed that *S. indicum* and *S. spindale* are also closely related (2.4% in pairwise difference), corresponding to the degree between *S. malayensis* and *S. mekongi* (2.4% in pairwise difference). On the other hand, *S. incognitum* has large pairwise differences, 19.4% in average for the Pomatiopsidae-harbouring schistosomes, and 21.7% in average for

https://doi.org/10.1079/JOH200191 Published online by Cambridge University Press

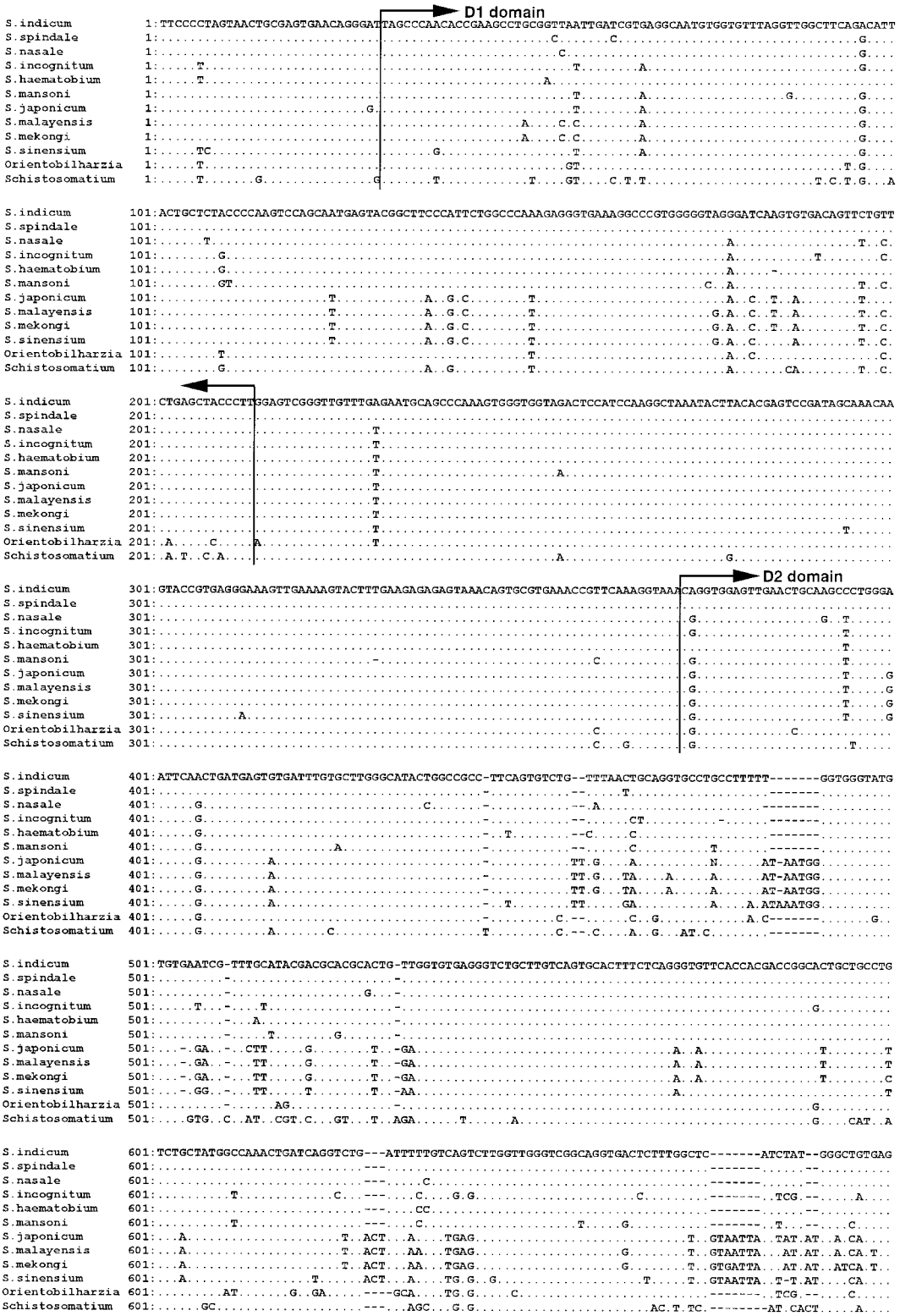

Fig. 2. Nucleotide sequence alignment of part of the 28S ribosomal RNA gene among species in three genera *Schistosoma*, *Orientobilharzia* and *Schistosomatium*. '.' indicates identity with the top sequence; '-' indicates an alignment gap.

https://doi.org/10.1079/JOH200191 Published online by Cambridge University Press

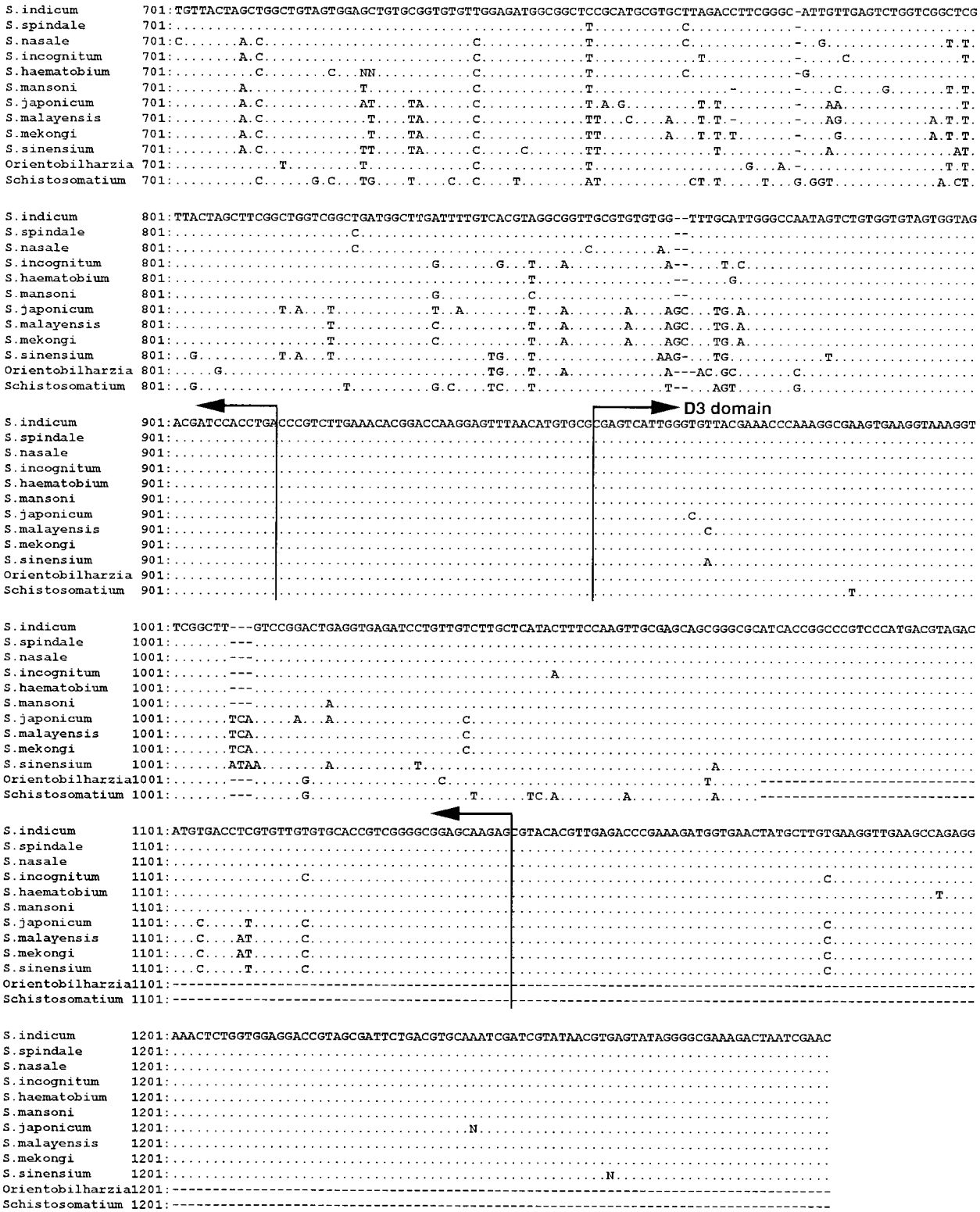

Fig. 2. (Continued)

https://doi.org/10.1079/JOH200191 Published online by Cambridge University Press

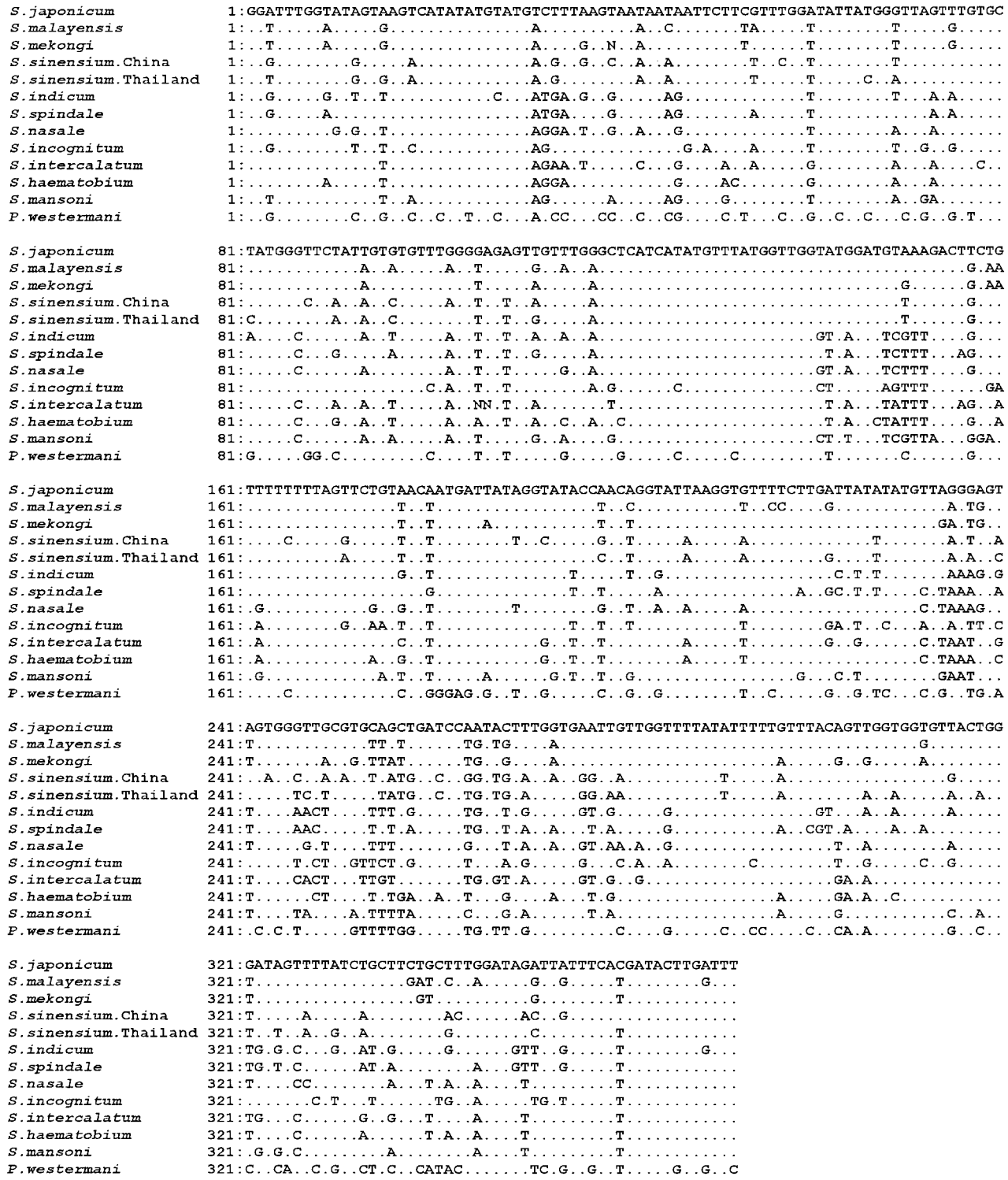

Fig. 3. (a) Nucleotide sequence alignment of part of the mitochondrial cytochrome c oxidase subunit 1 (CO1) gene among Asian and African species of *Schistosoma*. '.' indicates identity with the top sequence.

https://doi.org/10.1079/JOH200191 Published online by Cambridge University Press

```
S.japonicum             1:GFGIVSHICMSLSNNNSSFGYYGLVCAMGSIVCLGSVVWAHHMFMVGMDVKTSVFFSSVTMIIGIPTGIKVFSWLYMLGS
S.mekongi               1:..........T......L.................................AI.....................SG
S.malayensis            1:..........T......L.................................AI.................P.....SG
S.sinensium.China       1:...M......T.....................................A.....................F..S.
S.sinensium.Thailand    1:...M......T.....................................A.....................F..S.
S.incognitum            1:..............ST.....................G.......L.SL..I....I.............I..ISC
S.spindale              1:..........MI...D........I...A.................L.SL.A.....................F..N.
S.nasale                1:...V.......I...D........I...A.................L.SL.A.....................NG
S.indicum               1:..........MM...D........I...A.................L.SL.A.....................F..NG
S.intercalatum          1:...........I...D........I...A......X...V.......L.YL.AI.....................N.
S.haematobium           1:...........I...D.P......I...A.................L.YL.AI.....................N.
S.mansoni               1:..............D........I...A..........G.......F.SL.G.....I.....V............N.
P.westermani            1:..........T.T..D.L.......F...A.................L....A.......GV..M.........F...G

S.japonicum            81:SGLRAADPILWWIVGFIFLFTVGGVTGIVLSASALDSLFHDTWF
S.mekongi              81:C...VI..VV.....................S..........
S.malayensis           81:C...VV..VV.....................WS.........
S.sinensium.China      81:...HVM..VV..VI.....................N.......
S.sinensium.Thailand   81:.....M..VV..VI.....................................
S.incognitum           81:..S.FS...M...L.................S.V...V......
S.spindale             81:C.T.VS..V...L.......VI.....VA..S....V.......
S.nasale               81:C.V.VS......LI.............A....S..I.......
S.indicum              81:C.T.VS..V...L.......V......VA..S....V.......
S.intercalatum         81:C.T.LS..V...L........I.....VA....S..I.......
S.haematobium          81:C.S.VWE..V..L........I......A....S..I.......
S.mansoni              81:C.M.VL...V..L.............VA.......I.......
P.westermani           81:TR..FW..V....L.......M......I..S.M....L.....
```

Fig. 3. (b) Amino acid sequence alignment of part of the CO1 gene among Asian and African species of *Schistosoma*. '.' indicates identity with the top sequence.

the pulmonate-harboring schistosomes. In summary, the CO1 sequences exhibited Ts/Tv ratios strikingly different from those seen in the nuclear gene regions. Most values were below 1.00, whereas exceptions to this were always between very closely related species. Although this suggests substitutional saturation, it might also be partly due to the base composition of the gene. Any changes between a G and a T would be a transversion, and these two bases account for 68.9% of all bases. Because of the apparent saturation, we were particularly cautious in interpreting trees from the CO1 nucleotide sequences.

Description of trees

ITS2 region

As shown in fig. 4a, two large clades were formed. One includes *S. japonicum* and *S. sinensium* which formed a clade supported by high bootstrap values, as demonstrated previously (Agatsuma *et al.*, 2001). The other clade consists of the *S. indicum* group and all the African species. In this clade, members of the *S. indicum* group did not form a monophyletic subclade. Rather, one member of the group, *S. incognitum*, lay basal to all other African species except for *S. hippopotami*. The parsimony method (heuristic) using PAUP gave a different topology, yielding one most-parsimonious tree with a consistency index of 0.782 (0.648 excluding uninformative characters). The topology differed from the NJ tree in that *S. hippopotami* was placed at the basal position of a clade consisting of the *S. japonicum* group plus *S. sinensium* as shown in fig. 4b.

28S region

Two large clades were formed (fig. 5a). One includes the *S. japonicum* group and *S. sinensium* that formed a clade supported by high bootstrap values. The other clade consists of the *S. indicum* group and the two African species. In this clade, members of the *S. indicum* group did not form a monophyletic subclade. In this analysis, as in ITS2, one member of the group, *S. incognitum*, lay basal to all other African species. The parsimony (heuristic) analysis using PAUP also resulted in one most-parsimonious tree with a consistency index of 0.821 (0.717 excluding uninformative characters), the topology of which is identical to that of ITS2 (fig. 5b). Tree analysis for the most variable regions (D2, approximately 540 bp) also provided the same topology as that for the whole sequence by both the NJ method and the parsimony method (yielding one most-parsimonious tree with a consistency index of 0.813 (0.727 excluding uninformative characters). However, tree analyses for D1 (approximately 175 bp) and D3 (approximately 120 bp) gave unstable topologies, perhaps due to the short length.

CO1 region

As shown in fig. 6a, two large clades were formed to the exclusion of a cluster consisting of *Paragonimus ohirai* and *P. westermani* as outgroup taxa. One includes the *S. japonicum* group plus *S. sinensium*, as in the ITS2 and 28S trees. The second clade consists of two subclusters; one includes *S. indicum*, *S. spindale* and *S. nasale*, and all species of the African group, and another, only *S. incognitum*. In this tree, *S. indicum* and *S. spindale* were clustered into a single group with a high value of bootstrapping. But *S. nasale* has a closer affinity with a cluster of the *S. haematobium* group rather than to that of *S. indicum* and *S. spindale*. *Schistosoma mansoni* was placed outside those clusters. On the other hand, the parsimony method (heuristic) using PAUP

https://doi.org/10.1079/JOH200191 Published online by Cambridge University Press

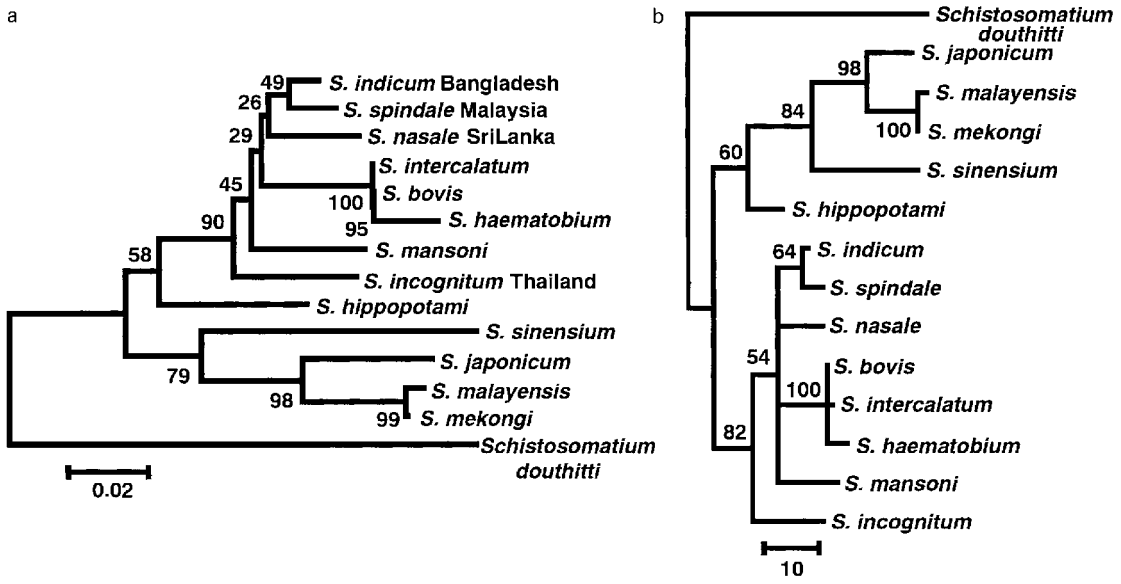

Fig. 4. (a) A phylogenetic tree rooted by outgroup (*Schistosomatium douthitti*), inferred from ITS2 sequences using the neighbour-joining method in MEGA. Scale bar indicates the proportion of sites changing along each branch. (b) A phylogenetic tree rooted by outgroup (*Schistosomatium douthitti*), inferred from ITS2 sequences using the parsimony method in PAUP. Scale bar indicates the number of changes inferred as having occurred along each branch.

gave a similar topology, yielding one most-parsimonious tree with a consistency index of 0.551 (0.501 excluding uninformative characters), as shown in fig. 6b. Phylogenetic trees constructed using amino acid sequences (124 amino acid) showed that *S. incognitum* is placed at the base of the African plus *S. indicum* group clade for both analyses using the NJ and parsimony methods. However, the NJ tree differed in the topology

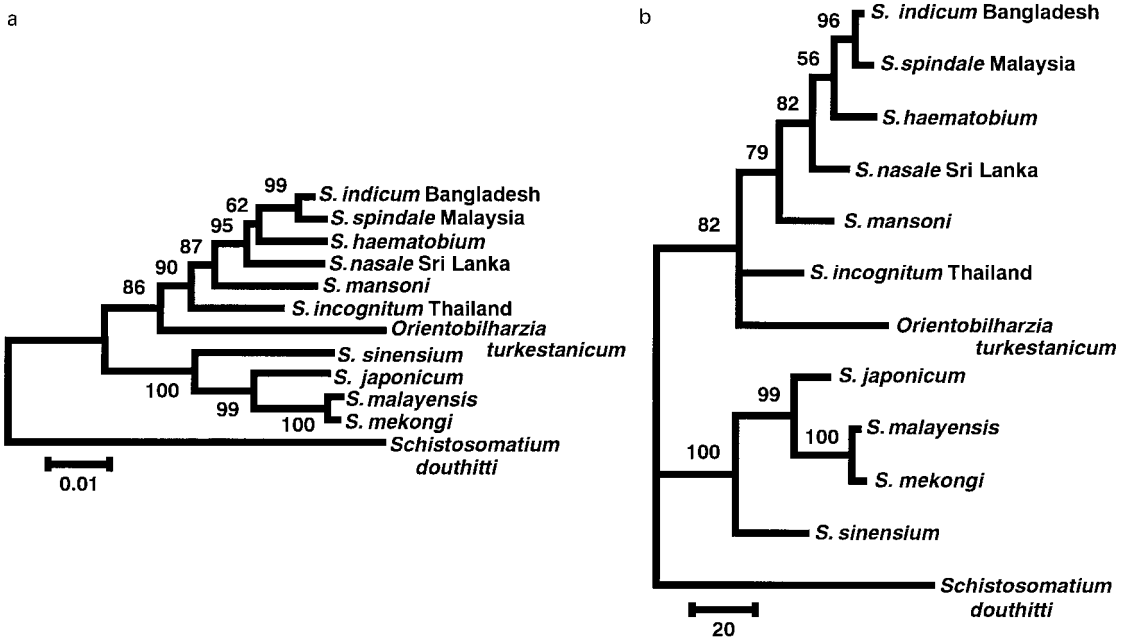

Fig. 5. (a) A phylogenetic tree rooted by outgroup (*Schistosomatium douthitti*), inferred from partial 28S sequences using the neighbour joining method in MEGA. Scale bar indicates the proportion of sites changing along each branch. (b) A phylogenetic tree rooted by outgroup (*Schistosomatium douthitti*), inferred from partial 28S sequences using the parsimony method in PAUP. Scale bar indicates the number of changes inferred as having occurred along each branch.

https://doi.org/10.1079/JOH200191 Published online by Cambridge University Press

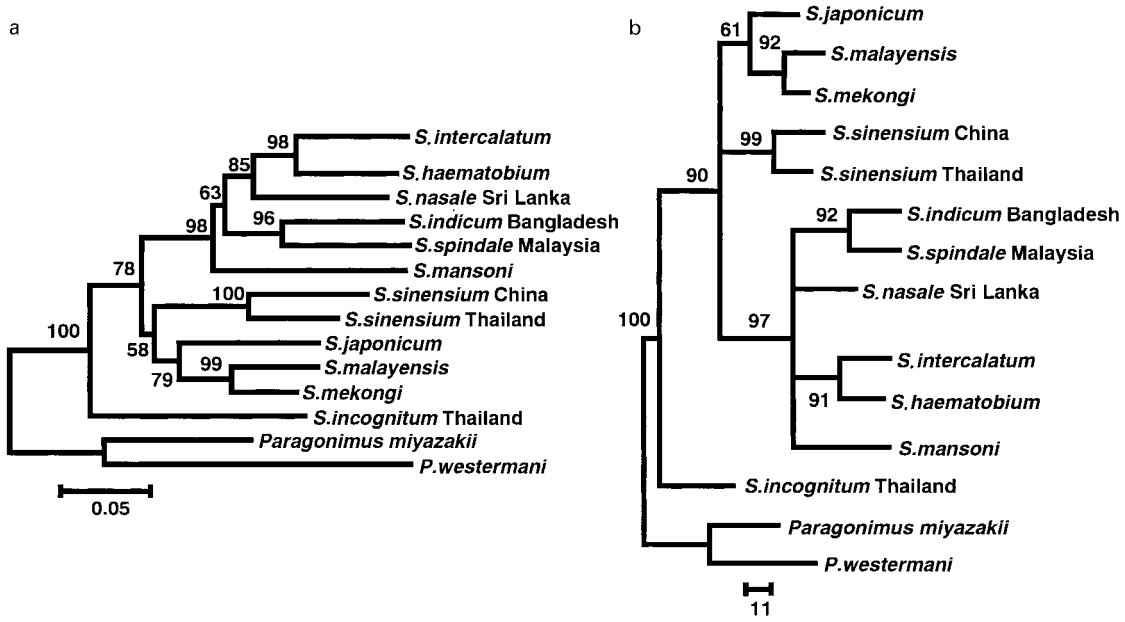

Fig. 6. (a) A phylogenetic tree rooted by outgroup (*Paragonimus miyazakii* and *P. westermani*), inferred from partial CO1 nucleotide sequences using the neighbour joining method in MEGA. Scale bar indicates the proportion of sites changing along each branch. (b) A phylogenetic tree rooted by outgroup (*Paragonimus miyazakii* and *P. westermani*), inferred from partial CO1 nucleotide sequences using the parsimony method in PAUP. Scale bar indicates the number of changes inferred as having occurred along each branch.

of a clade of the *S. japonicum* group plus *S. sinensium* (fig. 7a), and the parsimony tree gave a relatively poor resolution in the same clade, perhaps because of instability due to the small number of amino acids analysed (fig. 7b).

Discussion

Molecular data presented here did not support the monophyly of the *S. indicum* group in any analysis. It has been recognized previously that this group might be an

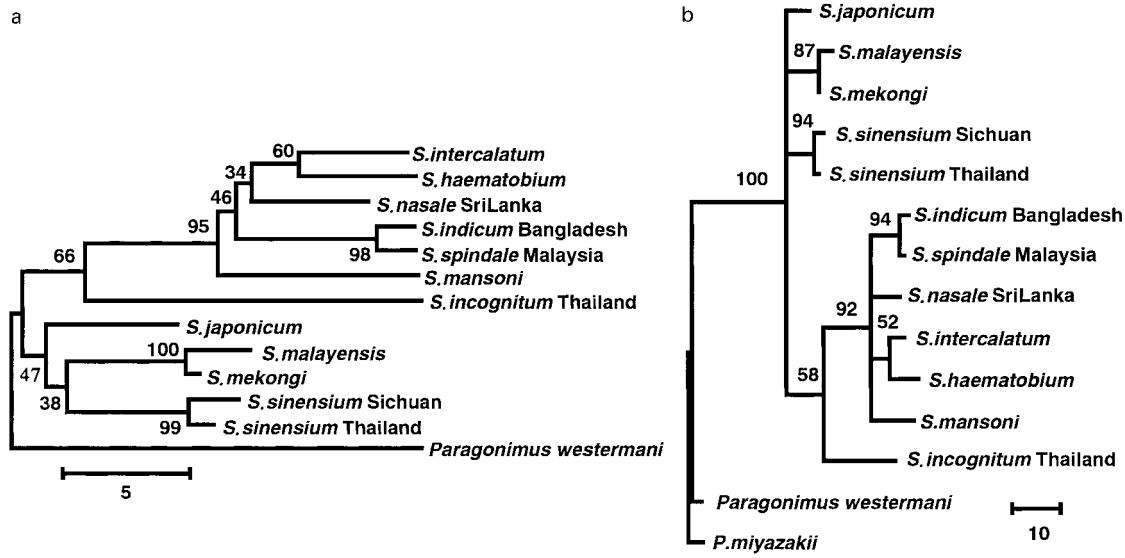

Fig. 7. (a) A phylogenetic tree rooted by outgroup (*Paragonimu westermani*), inferred from partial CO1 amino acid sequences using the neighbour joining method in MEGA. Scale bar indicates the proportion of sites changing along each branch. (b) A phylogenetic tree rooted by outgroup (*Paragonimus miyazakii* and *P. westermani*), inferred from partial CO1 amino acid sequences using the parsimony method in PAUP. Scale bar indicates the number of changes inferred as having occurred along each branch.

https://doi.org/10.1079/JOH200191 Published online by Cambridge University Press

artificial one (Rollinson & Southgate, 1987), and our data agree with this view. One species within the *S. indicum* group, *S. incognitum*, is always placed some distance from the other members. Trees constructed from CO1 nucleotide sequences (fig. 6) place this species as sister to the rest of the genus. However, those constructed from CO1 amino acid sequences (fig. 7) and ITS2 and 28S sequences (figs 4 and 5) place it as sister to all *Schistosoma* species other than those in the *S. japonicum* group (but *S. hippopotami* lies outside it in the ITS2 tree). Two other species within the *S. indicum* group, *S. indicum* and *S. spindale*, always cluster together. Judging by branch lengths on the trees, these two species are as closely related as are some members of the *S. haematobium* group. The fourth member of the *S. indicum* group, *S. nasale*, does not convincingly group with *S. indicum* plus *S. spindale*. Indeed, some analyses (CO1 nucleotide, NJ analysis) place it closer to the *S. haematobium* group, while others (CO1 nucleotide, parsimony analysis) fail to resolve its placement. We conclude that the *S. indicum* group is a paraphyletic assemblage, most closely related to the species currently found in Africa. A corollary of this is the conclusion that species currently occurring in Africa do not constitute a monophyletic lineage to the exclusion of species from other geographic regions.

Molecules and morphology conflict within the *S. indicum* group. *Schistosoma nasale* and *S. spindale* both have distinctive, long, eggs whereas *S. indicum* has subterminal-spined eggs. Sequence data suggest that *S. spindale* and *S. indicum* are closely related despite their differing egg shapes.

For some taxa, sequences were only available for a single gene region. The African species *S. hippopotami* was represented only by a partial ITS2 sequence. In the neighbour-joining ITS2 tree, this species was sister to all African species plus the *S. indicum* group. The basal placement of *S. hippopotami* in this clade was supported by a rather low bootstrap value. In the parsimony tree, it appeared as sister to the *S. japonicum* group, a position again supported by a rather low bootstrap value. The genus *Orientobilharzia* was represented only by the 28S sequence. This placed *O. turkestanicum* at the base of the clade containing the African species plus the *S. indicum* group, either emerging from the same node as *S. incognitum* (parsimony tree) or basal to *S. incognitum* (neighbour-joining tree).

Conflicting topologies mean that we cannot choose between two competing hypotheses concerning the geographical origin of the genus *Schistosoma*. Trees placing *S. incognitum* (Asia) as sister to all other species within the genus (CO1 nucleotide, parsimony and NJ analyses) suggest an Asian origin for the genus (fig. 6). The CO1 nucleotide parsimony tree implies a paraphyletic *S. japonicum* group radiating before the lineage containing the African and remaining *S. indicum* group species itself diversifies. Under this scenario, an Asian origin and radiation of *Schistosoma japonicum* and its relatives were followed by a later radiation further west and south in Asia, some members of which also reached and radiated in Africa. For reasons discussed above, we are wary of placing much faith in the CO1 nucleotide trees. However, further support for an Asian origin is given by trees (e.g. CO1 amino acid analyses) (fig. 7) in which *S. incognitum* is placed at the base of the African plus *S. indicum* group clade. In these trees, species living in Asia are basal in both main clades.

The second hypothesis proposes an African origin for the genus (e.g. Davis, 1980, 1992). None of the trees unambiguously supports this hypothesis and interpretations focus on the placement of *S. hippopotami* for which only ITS2 data were available. In the NJ tree, *S. hippopotami* is basal in the clade containing African species plus the *S. indicum* group. Members of the *S. japonicum* group constitute the other main clade. This tree provides no real evidence for either hypothesis, but does indicate that at least one species (*S. hippopotami*) has a long history in Africa, and that its sister lineage gave rise to species now found in Asia as well as Africa. The ITS2 parsimony tree is of little help. In this, *S. hippopotami* (Africa) is placed at the base of a clade otherwise containing only the *S. japonicum* group (Asia) and *S. incognitum* (Asia) lies at the base of the clade containing African and Asian species. All interpretations of this tree imply an unconvincing number of geographic shifts and indicate little about the geographic origins of the genus.

One possible problem with the ITS2 tree is the choice of the outgroup taxon. *Schistosomatium* is a reasonable choice for this and is the only other genus for which ITS2 data are available. Nevertheless, the choice of outgroup could be influencing tree topology and it would be desirable to repeat the analyses using other schistosome genera in this capacity.

The position of *Orientobilharzia* in the trees inferred from 28S data is particularly interesting and agrees with the findings of Snyder & Loker (2000) who included fewer species of *Schistosoma* in their analyses. These trees strongly suggest that *Orientobilharzia* is not distinct from *Schistosoma* and that it might lie relatively close to *S. incognitum*. *Orientobilharzia* species overlap the ranges of members of the *S. indicum* group and also extend further to the north and west, occurring in parts of the Middle East. They thus almost close the geographic gap between species of *Schistosoma* occurring in Asia and Africa.

A consideration of snail hosts utilized by *Schistosoma* species provides little help concerning the geographic origins of the genus. Both Davis (1980), using biogeographical inference and Snyder & Loker (2000) using molecular data, suggested that the ancestral *Schistosoma* species probably utilized a pulmonate snail. The *S. japonicum* group has made a host-switch to pomatiopsid snails whereas all other *Schistosoma* species (and *Orientobilharzia*) have retained the ancestral specificity for pulmonates. *Orientobilharzia* species and *S. incognitum* utilize lymnaeid snails whereas all other schistosomes in their clade utilize planorbids.

As for the timing of evolutionary events in the genus, these are hard to pin down. Davis (1980, 1992) has suggested extremely ancient divergences. Després *et al.* (1995) have likewise suggested ancient divergences. Consequently, all divergences noted probably occurred well before humans had an influence on the biosphere. For example, if the divergence rate for the ITS2 (0.3–0.8%/MY) proposed by Després *et al.* (1992) is accepted, *S. indicum* diverged from *S. mansoni* 7.1–19.0 MY. Consequently, the suggestion (Barker & Blair, 1996) that

S. spindale may have been brought to Asia from Africa by early humans and their livestock, is likely to be incorrect.

In conclusion, all data obtained from the three DNA regions indicate that the *S. indicum* group is paraphyletic, has a closer affinity with *Schistosoma* spp. infecting humans in Africa, and that the genus *Schistosoma* may have originated in Asia.

Acknowledgements

We would like to thank Mr W. D. Paranagama (University of Peradenyia) and Ms Y. Fusegawa (Kochi Medical School) for technical assistance. This study was partly supported by a Grant-in-Aid for International Scientific Research (T. Agatsuma: 10041188) from the Japanese Ministry of Education (1998).

References

Agatsuma, T., Iwagami, M., Liu, C.X., Saitoh, Y., Kawanaka, M., Upatham, S., Qui, D. & Higuchi, T. (2001) Molecular phylogenetic position of *S. sinensium* in the genus *Schistosoma*. *Journal of Helminthology* **75**, 215–221.

Barker, S.C. & Blair, D. (1996) Molecular phylogeny of *Schistosoma* species supports traditional groupings within the genus. *Journal of Parasitology* **82**, 292–298.

Blair, D., Agatsuma, T. & Watanobe, T. (1997a) Molecular evidence for the synonymy of three species of *Paragonimus*, *P. ohirai* Miyazaki, 1939, *P. iloktsuenensis* Chen, 1940 and *P. sadoensis* Miyazaki *et al.*, 1968. *Journal of Helminthology* **71**, 305–310.

Blair, D., van Herwerden, L., Hirai, H., Taguchi, T., Habe, S., Hirata, M., Lai, K., Upatham, S. & Agatsuma, T. (1997b) Relationships between *Schistosoma malayensis* and other Asian schistosomes deduced from DNA sequences. *Molecular and Biochemical Parasitology* **85**, 259–263.

Blair, D., Le, T.H., Després, L. & McManus, D.P. (1999) Mitochondrial genes of *Schistosoma mansoni*. *Parasitology* **119**, 303–313.

Bowles, J., Hope, M., Tiu, W.U., Liu, X. & McManus, D.P. (1993) Nuclear and mitochondrial genetic markers highly conserved between Chinese and Philippine *Schistosoma japonicum*. *Acta Tropica* **55**, 217–229.

Bowles, J., Blair, D. & McManus, D.P. (1995) A molecular phylogeny of the human schistosomes. *Molecular Phylogenetics and Evolution* **4**, 103–109.

Davis, G.M. (1980) Snail hosts of Asian *Schistosoma* infecting man: origin and coevolution. *Malacological Review* Supplement **2**, 195–238.

Davis, G.M. (1992) Evolution of prosobranch snails transmitting Asian *Schistosoma*; coevolution with *Schistosoma*: a review. *Progress in Clinical Parasitology* **3**, 145–204.

Després, L., Imbert-Establet, D., Combes, C. & Bonhomme, F. (1992) Molecular evidence linking hominid evolution to recent radiation of schistosomes (Platyhelminthes: Trematoda). *Molecular Phylogenetics and Evolution* **1**, 295–304.

Després, L., Kruger, F.J., Imbert-Establet, D. & Adamson, M.L. (1995) ITS2 ribosomal RNA indicates *Schistosoma hippopotami* is a distinct species. *International Journal for Parasitology* **25**, 1509–1514.

Higgins, D.G., Bleasby, A.J. & Fuchs, R. (1992) Improved software for multiple sequence alignment. *Computer Applications in the Biosciences* **8**, 189–191.

Johnston, D.A., Kane, R.A. & Rollinson, D. (1993) Small subunit (18S) ribosomal RNA gene divergence in the genus *Schistosoma*. *Parasitology* **107**, 147–156.

Littlewood, D.T. & Johnston, D.A. (1995) Molecular phylogenetics of the four *Schistosoma* species groups determined with partial 28S ribosomal RNA gene sequences. *Parasitology* **111**, 167–175.

Rollinson, D. & Southgate, V.R. (1987) The genus *Schistosoma*: a taxonomic appraisal. pp. 1–49 *in* Rollinson, D. & Simpson, A.J.G. (*Eds*) *The biology of schistosomes from genes to latrines*. London, Academic Press.

Snyder, S.D. & Loker, E.S. (2000) Evolutionary relationships among the Schistosomatidae (Platyhelminthes: Digenea) and an Asian origin for *Schistosoma*. *Journal of Parasitology* **86**, 283–288.

(Accepted 19 April 2001)