Molecular Ecology (2005) **14**, 3889–3902 doi: 10.1111/j.1365-294X.2005.02709.x

Origin and diversification of the human parasite *Schistosoma mansoni*

JESS A. T. MORGAN,* RANDALL J. DEJONG,* GRACE O. ADEOYE,‡ EBENEZER D. O. ANSA,§ CONSTANÇA S. BARBOSA,¶ PHILIPPE BRÉMOND,** ITALO M. CESARI,†† NATHALIE CHARBONNEL,‡‡ LYGIA R. CORRÊA,§§ GODEFROY COULIBALY,¶¶ PAULO SÉRGIO D'ANDREA,*** CECILIA PEREIRA DE SOUZA,††† MICHAEL J. DOENHOFF,‡‡‡ SHARON FILE,§§§ MOHAMED A. IDRIS,¶¶¶ R. NINO INCANI,**** PHILIPPE JARNE,†††† DIANA M. S. KARANJA,‡‡‡‡ FRANCIS KAZIBWE,§§§§ JOHN KPIKPI,¶¶¶¶ NICHOLAS J. S. LWAMBO,***** AMADOU MABAYE,††††† LUIZ A. MAGALHÃES,‡‡‡‡‡ ASANTELI MAKUNDI,§§§§§ HÉLÈNE MONÉ,¶¶¶¶¶ GABRIEL MOUAHID,¶¶¶¶¶ GERALD M. MUCHEMI,****** BEN N. MUNGAI,†††††† MARIAMA SÉNE,‡‡‡‡‡‡ VAUGHAN SOUTHGATE,§§§§§§ LOUIS ALBERT TCHUEM TCHUENTÉ,¶¶¶¶¶¶ ANDRE THÉRON,¶¶¶¶¶ FOUAD YOUSIF,******* ELIANA M. ZANOTTI-MAGALHÃES,‡‡‡‡‡ GERALD M. MKOJI†††††† and ERIC S. LOKER*

**Department of Biology, University of New Mexico, Albuquerque, New Mexico 87131 USA, ‡Department of Zoology, University of Lagos, Akoka, Yaba, Lagos, Nigeria, §CSIR Water Research Institute, Environmental Biology and Health Division, PO AH 38, Achimota, Accra, Ghana, ¶Centro de Pesquisas Aggeu Magalhães, FIOCRUZ, Campus da UFPE, Recife, PE, Brazil, **Equipe Schistosomoses, IRD, BP 434, 101 Antananarivo, Madagascar, ††Centro de Microbiología y Biología Celular, Instituto Venezolano de Investigaciones Científicas, Caracas, Venezuela, ‡‡Centre de Biologie et de Gestion des Populations, Campus International de Baillarguet, Montferrier s/Lez cedex, France, §§Departmento de Malacología, Instituto Oswaldo Cruz, Río de Janeiro, RJ, Brazil, ¶¶Institut National de Recherche en Santé Publique (INRSP), BP 1771, Bamako, Mali, ***Laboratório de Biologia e Controle da Esquistossomose, Departamento de Medicina Tropical, Instituto Oswaldo Cruz -Fiocruz, Av. Brazil 4365, 21045–900 Río de Janeiro, RJ, Brazil, †††Centro de Pesquisas René Rachou, FIOCRUZ, Belo Horizonte, Minas Gerais, Brazil, ‡‡‡School of Biological Science, University of Wales, Bangor LL57 2UW, Gwynedd, Wales, §§§Department of Biology, University of Puerto Rico, Río Piedras Campus, San Juan, Puerto Rico, ¶¶¶Department of Microbiology and Immunology, College of Medicine and Health Sciences, Sultan Qaboos University, Muscat, Oman, ****Departamento de Parasitología, Facultad de Ciencias de la Salud, Universidad de Carabobo, Valencia, Venezuela, ††††Centre d'Ecologie Fonctionelle et Evolutive–Centre National de la Recherche Scientifique, 1919 route de Mende, 34293 Montpellier Cedex 5, France, ‡‡‡‡Centre for Vector Biology and Control Research, Kenya Medical Research Institute, Kisumu, Kenya, §§§§Department of Medical Entomology and Parasitology, Vector Control Division, Ministry of Health, Kampala, Uganda, ¶¶¶¶Department of Zoology, University of Ghana, Legon, Ghana, *****National Institute for Medical Research, Mwanza, Tanzania, †††††Programme de Lutte Contre les Bilharzioses, Région Médicale de St Louis, Rue Thevanot 26, B.P. 394, St Louis, Senegal, ‡‡‡‡‡Departamento de Parasitología, Instituto de Biologia, Universidade Estadual de Campinas, Cidade Universitária, Caixa Postal 6109, 13083–970 Campinas, SP, Brazil, §§§§§Open University of Tanzania, PO Box 271, Babati, Tanzania, ¶¶¶¶¶UMR 5555 CNRS-UP, CBETM, Université de Perpignan, 52 AV Paul Alduy, 66860 Perpignan cedex, France, ******Department of Public Health, Pharmacology and Toxicology, University of Nairobi, PO Box 29053, Kabete, Kenya, ††††††Centre for Biotechnology Research and Development, Kenya Medical Research Institute, Nairobi, Kenya, ‡‡‡‡‡‡Institut de Recherche pour le Développement (IRD), Laboratoire de Paludologie, B.P. 17688 Dakar-Liberté, Dakar, Senegal, §§§§§§Department of Zoology, Biomedical Parasitology Division, Wolfson Wellcome Biomedical Laboratories, The Natural History Museum, Cromwell Road, South Kensington, London SW7 5BD, ¶¶¶¶¶¶Center for Schistosomiasis and Parasitology, PO Box 7244, Yaoundé, Cameroon, *******Schistosome Biological Supply Center, Theodor Bilharz Research Institute, PO Box 30 Imbaba, Egypt*

Correspondence: Eric S. Loker, Fax: 505-277-0304, Office: 505-277-2496; E-mail: esloker@unm.edu

Abstract

***Schistosoma mansoni* is the most widespread of the human-infecting schistosomes, present in 54 countries, predominantly in Africa, but also in Madagascar, the Arabian Peninsula, and the Neotropics. Adult-stage parasites that infect humans are also occasionally recovered from baboons, rodents, and other mammals. Larval stages of the parasite are dependent upon certain species of freshwater snails in the genus *Biomphalaria*, which largely determine the parasite's geographical range. How *S. mansoni* genetic diversity is distributed geographically and among isolates using different hosts has never been examined with DNA sequence data. Here we describe the global phylogeography of *S. mansoni* using more than 2500 bp of mitochondrial DNA (mtDNA) from 143 parasites collected in 53 geographically widespread localities. Considerable within-species mtDNA diversity was found, with 85 unique haplotypes grouping into five distinct lineages. Geographical separation, and not host use, appears to be the most important factor in the diversification of the parasite. East African specimens showed a remarkable amount of variation, comprising three clades and basal members of a fourth, strongly suggesting an East African origin for the parasite 0.30–0.43 million years ago, a time frame that follows the arrival of its snail host. Less but still substantial variation was found in the rest of Africa. A recent colonization of the New World is supported by finding only seven closely related New World haplotypes which have West African affinities. All Brazilian isolates have nearly identical mtDNA haplotypes, suggesting a founder effect from the establishment and spread of the parasite in this large country.**

Keywords: Africa, *Biomphalaria*, Brazil, phylogeography, schistosomiasis, slave trade

Received 20 May 2005; revision accepted 25 July 2005

Introduction

Schistosomiasis is one of the world's great neglected diseases, with six species of *Schistosoma* still infecting 200 million people (Crompton 1999; Chitsulo *et al*. 2000). One of the most common is *Schistosoma mansoni*, which infects more than 83 million humans in 54 countries, mostly in tropical Africa (Crompton 1999), causing intestinal schistosomiasis, which results in pathology, morbidity, and even death (Lambertucci *et al*. 2000). Adult worms living in veins surrounding the intestine produce eggs that either pass with the faeces to continue the life cycle, or become trapped in the tissues, provoking granulomatous reactions, fibro-obstructive disease in the liver and other organs and portal hypertension (Lambertucci *et al*. 2000).

The evolution of the genus *Schistosoma* has long fascinated parasitologists and evolutionary biologists (Davis 1980, 1992), and has garnered much attention by recent molecular studies (Desprès *et al*. 1992; Snyder & Loker 2000; Lockyer *et al*. 2003; Morgan *et al*. 2003a), but at least as many questions have been raised as have been answered. It is suspected that the genus has an Asian origin, as the most basal clade in the genus is Asian (Snyder & Loker 2000; Lockyer *et al*. 2003; Morgan *et al*. 2003a), but three other groups within the genus are entirely African. *S. mansoni*, along with its sister taxon *Schistosoma rodhaini*, form one of these groups. Both species utilize *Biomphalaria* snails as intermediate hosts, but the adult stages of *S. rodhaini* infect rodents, dogs or serval cats instead of humans and this species is confined to tropical Africa. Other *Schistosoma* taxa, from the hippopotamus, that were previously thought to be part of the *S. mansoni* and *S. rodhaini* group, have been shown by molecular analysis to comprise their own monophyletic lineage within the genus (Morgan *et al*. 2003a). It has been hypothesized that the separation of *S. rodhaini* and *S. mansoni* began 5–7 million years ago (Ma) with the emergence of hominins (Desprès *et al*. 1992), but this study was limited to a small amount of sequence data from a limited number of specimens. Also, *S. mansoni* and *S. rodhaini* have been shown to hybridize in the laboratory, and recently a naturally occurring hybrid was identified (Morgan *et al*. 2003b). How diverse *S. mansoni* really is, with such a large geographical range, and the ability to use several different mammalian and snail hosts, has not been well assessed. An accurate measure of within-species diversity should be informative for understanding the evolution of the species.

Additionally, the presence of *S. mansoni* in the New World has been hypothesized to be the result of introduction through the Atlantic slave trade within the last 500 years and data from snail–parasite strain compatibility studies (Files 1951), allozymes (Fletcher *et al*. 1981), and mitochondrial restriction enzyme sites (Desprès *et al*. 1993) are consistent with this hypothesis, but a final conclusion cannot be drawn from these studies. For example,

if *S. mansoni* originated in Africa several million years ago and colonized the New World with the slave trade, then much more diversity should be seen in Africa than in the New World. Desprès *et al*. (1993), found that the low level of genetic diversity (measured by restriction enzyme sites) amongst five New World samples was similar to that between New World samples and a single African sample, but because only one African sample was included, a comparison of diversity on each continent could not be made. Other evidence for or against this hypothesis is indirect. Colonization of the New World by *S. mansoni* is favoured by recent phylogenies of *Biomphalaria* which show that *Biomphalaria glabrata*, a New World species, is closely related to African *Biomphalaria* and thus may have been quite susceptible to *S. mansoni* coming from Africa during the slave trade (Campbell *et al*. 2000; DeJong *et al*. 2001, 2003). Yet, variability in host use, morphology, and behaviour have been described in the New World (Théron *et al*. 1997; D'Andrea *et al*. 2002; Freire *et al*. 2002), suggestive of genetic diversity. The genealogical information from DNA sequences from an extensive collection of *S. mansoni* specimens should be able to definitively confirm or deny a recent colonization event.

Control of schistosomes depends primarily on treatment with one drug, praziquantel, and a global initiative to reduce worm burden in endemic areas with this drug has begun (Crompton *et al*. 2003; see also www.who.int/wormcontrol/en/). Unfortunately, some isolates of *S. mansoni* have shown worrisome signs in recent years of praziquantel resistance (in Egypt and Senegal; Doenhoff *et al*. 2002), and it has been difficult to predict how and when resistance may develop, though one possibility is that an area of high genetic diversity is more likely to contain alleles that could contribute to drug resistance.

To address these issues, we examined DNA sequence variation for *S. mansoni* collected from geographically widespread localities, and from different intermediate and definitive hosts, and from two collections of *S. rodhaini*. We constructed phylogenies of observed haplotypes and tested whether host use or geographical separation has been important in the diversification of the parasite. Also, by collecting parasites from several infected snails, at a few localities where it was possible, we aimed to obtain a glimpse of within-locality diversity.

Materials and methods

Parasite specimens

We collected 143 specimens of *Schistosoma mansoni* from 53 geographically widespread localities in Africa, the Arabian Peninsula, Madagascar, South America, and the Caribbean region (Table 1). The number and geographical sampling of *S. mansoni* in this study is comparable to that of recent global genetic studies of human parasites (Joy *et al*. 2003; Leclerc *et al*. 2004). Most collections were from the field, but a few collections were made from laboratory stocks for which the original collection locality was known. Field-derived specimens of *S. mansoni* were obtained either from infected snails collected in the field, from anonymous human faecal samples provided by local health authorities, from baboon faecal samples collected in the field, or from the trapping, dissection, and perfusion of wild rodents. Though the percentage of field-collected snails found to be infected is often quite low (< 1%), this method is logistically simple and was the predominate method in our study (Table 1). To generate adult worms, the best source of DNA, field-collected snails were exposed to light to stimulate emergence of larval stages (cercariae) that were then used to infect laboratory mice. At 7 weeks post infection, adult worms were retrieved by dissection or perfusion from mice and preserved in ethanol. If specimens were obtained from faecal samples of humans or baboons, the process was similar, but instead began with hatching the larval stages (miracidia) from eggs in the faeces, and using these to infect laboratory-reared snails. Cercariae from these snails were then used to infect mice and adult worms were retrieved 7 weeks later. From a few localities, adult worms could not be obtained at all, so DNA was extracted from cercariae or sporocysts from individual snails.

An aspect of *S. mansoni* biology relevant to our sampling is that individual field-infected snails usually are infected by a single *S. mansoni* genotype (Eppert *et al*. 2002). All worms derived from such a snail are clones and cannot be treated as individual samples. Thus we attempted to find multiple infected snails within a locality, which enabled us to sample multiple worms within a few localities. Within-locality sampling and clone avoidance were also made possible for some locations when male and female worms were obtained, indicating multiple genotypes. Overall, within-locality sampling was very modest, and was not intended to thoroughly describe the within-locality variation in all parts of *S. mansoni*'s range. The number of worms analysed from a site varied from 1 to 13. Two specimens of *Schistosoma rodhaini* from separate localities in Kenya (Homa Bay and Kisumu) were also obtained.

Molecular data

DNA was extracted from whole worms as described previously (Morgan *et al*. 2003b). The primary data set for this study was obtained by amplifying and sequencing four partial regions of the mitochondrial genome to find point mutations to distinguish among globally collected *S. mansoni*: (COI, 617 bp; 16S-12S rDNA, 660–663 bp; cyt *b*-ND4L-ND4, 637–639 bp; ND1, 620 bp; total 2532–2535 bp). Polymerase chain reaction (PCR) and sequencing protocols have been described previously (Morgan *et al*. 2003b),

Table 1 Collection and haplotype information for *Schistosoma mansoni* samples

Locality*	Country	Latitude, longitude	Laboratory strain?†	Host collected from‡	Life stage used§	Haplotype¶
Old World						
Homa Bay, Lake Victoria (KE1)	Kenya	00°32′S, 34°27′E	No	*Biomphalaria sudanica*	Adult male	KE1a
				B. sudanica	Adult male	KE1b
				B. sudanica	Adult female	KE1c
				B. sudanica	Adult male	KE1d
				B. sudanica	Adult male	KE1e
				B. sudanica	Adult male	KE1f
Kisumu, Lake Victoria (KE2)	Kenya	00°06′S, 34°45′E	No	*B. sudanica*	Adult male	KE2a
				B. sudanica	Adult male	KE2b
Kitui (KE3)	Kenya	01°21′S, 38°01′E	No	*B. pfeifferi*	Adult male	KE3
Musilili stream (KE4)	Kenya	01°27′S, 37°15′E	No	*B. pfeifferi*	Adult male	KE4
Makueni (KE5)	Kenya	02°18′S, 37°50′E	No	*B. pfeifferi*	Adult male	KE5
Mtito River (KE6)	Kenya	02°41′S, 38°09′E	No	*B. pfeifferi*	Adult male	KE6
Machakos (KE7)	Kenya	01°32′S, 37°17′E	No	*B. pfeifferi*	Adult male	KE7a
				B. pfeifferi	Adult male	KE7b
				B. pfeifferi	Adult male	KE7c
				B. pfeifferi	Adult male	KE7d
				B. pfeifferi	Adult male	KE7e
				B. pfeifferi	Adult male	KE7f
				B. pfeifferi	Cercariae	KE7g
				B. pfeifferi	Cercariae	KE7g
				B. pfeifferi	Adult male	KE7h
Mwea (KE8)	Kenya	00°38′S, 37°28′E	No	*B. pfeifferi*	Adult male	KE8
Taveta (KE9)	Kenya	03°34′S, 37°46′E	No	*Homo sapiens*	Adult male	KE9
Kibwezi (KE10)	Kenya	02°29′S, 38°06′E	No	Baboon (*Papio cyanocephalus*)	Adult male	KE10a
				P. cyanocephalus	Adult male	KE10b
Kimamba (TZ1)	Tanzania	07°40′S, 36°05′E	No	*B. pfeifferi*	Adult male	TZ1a
				B. pfeifferi	Adult male	TZ1b
				B. pfeifferi	Adult male	TZ1c
Mwanza (TZ2)	Tanzania	02°31′S, 32°54′E	No	*B. sudanica*	Adult male	TZ2
Tunduma (TZ3)	Tanzania	03°34′S, 37°46′E	No	*B. pfeifferi*	Adult male	TZ3a
				B. pfeifferi	Adult female	TZ3b
				B. pfeifferi	Adult female	TZ3b
				B. pfeifferi	Adult male	TZ3c
				B. pfeifferi	Adult male	TZ3c
				B. pfeifferi	Adult male	TZ3d
				B. pfeifferi	Adult female	TZ3c
				B. pfeifferi	Adult male	TZ3e
Kaseni-Shuleni, Ukerewe Is., Lake Victoria (TZ4)	Tanzania	01°56′S, 32°51′E	No	*B. choanomphala*	Adult female	TZ4
Nansio, Ukerewe Is., Lake Victoria (TZ5)	Tanzania	02°07′S, 33°05′E	No	*B. sudanica*	Adult male	TZ5
Butiaba site 1, Lake Albert (UG1)	Uganda	01°49′N, 31°20′E	No	*B. sudanica*	Adult male	UG1a
				B. sudanica	Adult male	UG1b
Butiaba site 2, Lake Albert (UG2)	Uganda	01°49′N, 31°20′E	No	*B. sudanica*	Adult male	UG2
Hamukungu, Lake Albert (UG3)	Uganda	00°01′S, 30°05′E	No	*B. sudanica*	Adult female	UG3
				B. sudanica	Adult male	UG3
				B. sudanica	Adult female	UG3
				B. sudanica	Adult female	UG3
Runga (UG4)	Uganda	01°44′S, 31°18′E	No	*H. sapiens*	Adult male	UG4a
				H. sapiens	Adult female	UG4b
Kariba Dam (ZA1)	Zambia	16°31′S, 28°43′E	No	*H. sapiens*	Adult male	ZA1a
				H. sapiens	Adult female	ZA1b
				H. sapiens	Adult male	ZA1c
				H. sapiens	Adult female	ZA1d
Antsirabe (MA1)	Madagascar	19°51′S, 47°02′E	No	*B. pfeifferi*	Sporocyst	MA1
Ihosy (MA2)	Madagascar	22°23′S, 46°09′E	No	*B. pfeifferi*	Adult male	MA2
Manajary (MA3)	Madagascar	21°13′S, 48°20′E	No	*B. pfeifferi*	Adult male	MA3
Ihivoka (MA4)	Madagascar	22°28′S, 46°08′E	No	*B. pfeifferi*	Sporocyst	MA4

Table 1 *Continued*

Locality*	Country	Latitude, longitude	Laboratory strain?†	Host collected from‡	Life stage used§	Haplotype¶
Dhofar (OM1)	Oman	17°03′N, 54°07′E	No	*H. sapiens*	Adult male	OM1
				H. sapiens	Adult male	OM1
Alzahraa, Kafr El Sheikh (EG1)	Egypt	31°15′N, 30°50′E	No	*H. sapiens*	Adult male	EG1
Mansafire, Menia (EG2)	Egypt	28°00′N, 30°49′E	No	*B. alexandrina*	Adult male	EG2a
				B. alexandrina	Adult male	EG2a
				B. alexandrina	Adult female	EG2b
				B. alexandrina	Adult male	EG2a
				B. alexandrina	Adult male	EG2a
				B. alexandrina	Adult male	EG2b
				B. alexandrina	Adult female	EG2b
				B. alexandrina	Adult male	EG2c
SSCP lab isolate (EG3)	Egypt	Unknown	Yes	Unknown	Adult male	EG3
Ngoa Ekelle, Yaounde (CA1)	Cameroon	03°03′N, 11°13′E	No	*B. pfeifferi*	Adult female	CA1
Makenene (CA2)	Cameroon	04°47′N, 10°48′E	No	*H. sapiens*	Adult male	CA2a
				H. sapiens	Adult male	CA2b
				H. sapiens	Adult female	CA2c
				H. sapiens	Adult male	CA2d
				H. sapiens	Adult female	CA2e
Jos site 1 (NG1)	Nigeria	09°55′N, 08°54′E	No	*B. pfeifferi*	Adult male	NG1
Jos site 2 (NG2)	Nigeria	09°55′N, 08°54′E	No	*H. sapiens*	Adult male	NG2a
				H. sapiens	Adult female	NG2b
Accra site 1 (GH1)	Ghana	05°57′N, 00°36′W	No	*B. pfeifferi*	Adult male	GH1
Accra site 2 (GH2)	Ghana	05°57′N, 00°36′W	No	*B. pfeifferi*	Adult male	GH2a
				B. pfeifferi	Adult male	GH2b
				B. pfeifferi	Adult female	GH2c
				B. pfeifferi	Adult male	GH2d
				B. pfeifferi	Adult female	GH2e
				B. pfeifferi	Adult male	GH2e
				B. pfeifferi	Adult female	GH2f
				B. pfeifferi	Adult male	GH2e
				B. pfeifferi	Adult female	GH2g
				B. pfeifferi	Adult female	GH2f
				B. pfeifferi	Adult female	GH2b
				B. pfeifferi	Adult female	GH2f
1950s lab isolate (LB1)	Liberia	Unknown	Yes	Unknown	Adult male	LB1
				Unknown	Adult female	LB1
Bamako (MI1)	Mali	12°40′N, 07°59′W	No	*B. pfeifferi*	Adult male	MI1
Ndiangue, Richard Toll (SE1)	Senegal	16°28′N, 15°43′W	No	*B. pfeifferi*	Adult male	SE1
Ndombo, Richard Toll (SE2)	Senegal	16°26′N, 15°42′W	No	*B. pfeifferi*	Adult male	SE2a
				B. pfeifferi	Adult male	SE2b
				B. pfeifferi	Adult male	SE2c
				B. pfeifferi	Adult male	SE2c
Saint-Louis MB19 (SE3)	Senegal	16°01′N, 16°30′W	Yes	*H. sapiens*	Adult male	SE3
New World						
Belem, Para (BR1)	Brazil	01°27′S, 48°29′W	Yes	*B. glabrata*	Adult male	BR1
Corrego de Café, Minas Gerais (BR2)	Brazil	19°40′S, 44°10′W	No	*B. glabrata*	Adult male	BR2
Sabara, Minas Gerais (BR3)	Brazil	19°54′S, 43°46′W	No	*B. glabrata*	Adult female	BR3
				B. glabrata	Adult male	BR3
				B. glabrata	Adult female	BR3
				B. glabrata	Adult female	BR3
				B. glabrata	Adult male	BR3
				B. glabrata	Adult male	BR3
				B. glabrata	Adult female	BR3
				B. glabrata	Adult male	BR3
				B. glabrata	Adult female	BR3
				B. glabrata	Adult male	BR3
				B. glabrata	Adult male	BR3

Table 1 *Continued*

Locality*	Country	Latitude, longitude	Laboratory strain?†	Host collected from‡	Life stage used§	Haplotype¶
Recife, Pernambuco (BR4)	Brazil	08°06′S, 34°53′W	No	*H. sapiens*, from *B. straminea* area	Cercariae	BR4
Dionisio, Minas Gerais (BR5)	Brazil	19°49′S, 42°45′W	No	*B. glabrata*	Adult male	BR5a
				B. glabrata	Adult male	BR5a
				B. glabrata	Adult male	BR5a
				B. glabrata	Adult male	BR5a
				B. glabrata	Adult female	BR5a
				B. glabrata	Adult male	BR5a
				B. glabrata	Adult male	BR5a
				B. glabrata	Adult male	BR5a
				B. glabrata	Adult female	BR5a
				B. glabrata	Adult male	BR5a
				B. glabrata	Adult female	BR5a
				B. glabrata	Adult male	BR5b
				B. glabrata	Adult male	BR5a
Sumidouro, Rio de Janeiro (BR6)	Brazil	22°03′S, 42°41′W	No	Rodent (*Nectomys* sp.)	Adult male	BR6
Sumidouro, Rio de Janeiro (BR7)	Brazil	22°03′S, 42°41′W	No	*H. sapiens*	Adult male	BR7
Ceara-Mirim, Rio Grande del Norte (BR8)	Brazil	05°40′S, 35°13′W	Yes	Rodent (*Oryzomys* sp.)	Adult male	BR8
Grand Camp, Basse-Terre (GU1)	Guadeloupe	16°00′N, 61°43′W	Yes	*H. sapiens*	Adult male	GU1
Jacquot, Grand-Terre (GU2)	Guadeloupe	16°20′N, 61°22′W	No	Rodent (*Rattus rattus*)	Adult male	GU2a
				Rodent (*R. rattus*)	Adult female	GU2a
				Rodent (*R. rattus*)	Adult male	GU2b
				Rodent (*R. rattus*)	Adult male	GU2a
				Rodent (*R. rattus*)	Adult female	GU2a
				Rodent (*R. rattus*)	Adult male	GU2a
				Rodent (*R. rattus*)	Adult female	GU2a
				Rodent (*R. rattus*)	Adult male	GU2a
				Rodent (*R. rattus*)	Adult female	GU2a
Long-term laboratory isolate (PR1)	Puerto Rico	Unknown	Yes	Unknown	Adult male	PR1
Caguas Valley (PR2)	Puerto Rico	18°14′N, 66°02′W	No	*H. sapiens*	Adult male	PR2
Cagua, Aragua (VE1)	Venezuela	10°11′N, 67°27′W	Yes	*H. sapiens*	Adult male	VE1
Valencia (VE2)	Venezuela	10°11′N, 68°00′W	Yes	*H. sapiens*	Adult male	VE2

*Each locality is designated with a two-letter country code followed by a number.
†Yes if passaged in the laboratory more than one full life cycle.
‡The host from which material was collected in the field. Most specimens were collected from their snail hosts, and then passaged to mice in the laboratory to obtain adult worms.
§Adult worms were used for DNA extraction, unless noted.
¶Haplotypes are coded according to locality, and in some cases followed by a lowercase letter to differentiate among multiple haplotypes within the locality. Haplotype codes are used in Fig. 2.

and primer sequences and positions are available upon request. If both male and female worms were recovered from a locality, or if multiple infected hosts (snails or mammalian) were found within a locality, the most variable marker, ND1, was sequenced first to detect duplicate haplotypes. If duplicates were found within the same site, they were excluded from further sequencing, but for unique haplotypes the full complement of mitochondrial sequences was then obtained. The full complement of mitochondrial sequences was also obtained for one worm from each of 32 localities where within-locality sampling was not possible. GenBank Accession Nos for all data are AY446078–AY446326 and AY896574–AY896733.

We also sequenced nuclear markers, the rDNA internal transcribed spacers (ITS1, 460 nt; ITS2, 311 nt) for a large number of *S. mansoni* specimens but surprisingly, these varied at only three nucleotide sites, all located in ITS1. Desprès *et al*. (1992) also found a complete lack of variable sites in ITS2 of six specimens. The nuclear markers were therefore uninformative for *S. mansoni* comparisons, though they are of some use in comparing *S. mansoni* and *S. rodhaini*. We attempted to develop other nuclear markers for *S. mansoni*, but without success. There is clearly a need for single-copy nuclear markers for *S. mansoni*, as is the case for many organisms (Zhang & Hewitt 2003).

Genetic analyses

Phylogenetic relationships among haplotypes were inferred using maximum parsimony (MP), minimum evolution

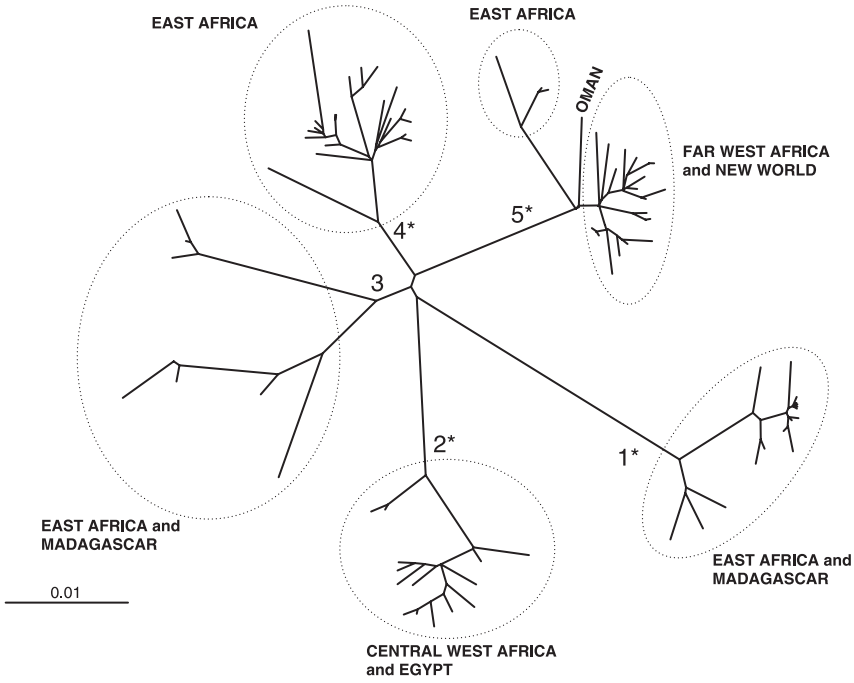

Fig. 1 Maximum-likelihood tree based on concatenated mitochondrial sequences. Numbers on main stems mark at least five lineages, of which four (1, 2, 4, and 5) had bootstrap or Bayesian support of 95–100% for all phylogenetic methods (symbolized by an asterisk). Scale bar indicates 1 substitution per 100 sites. Further detail of relationships and support within each lineage is provided in Fig. 2.

(ME), and maximum likelihood (ML) as implemented in PAUP*4.0b10 (Swofford 2004), with 1000 bootstrap replicates performed for maximum parsimony and minimum evolution, and 100 replicates for maximum likelihood. A bestfit ML model was chosen (GTR + I + Γ) using the Akaike criterion in MODELTEST 3.4 (Posada & Crandall 1998). Likelihood-ratio tests (Felsenstein 1981) were used to test the hypothesis that the data collected were consistent with a molecular clock. The ML phylogeny was estimated using the best-fit model, and then the likelihood of this phylogeny was recalculated while constraining the estimate to fit the molecular clock model. Additionally, we employed Bayesian phylogenetic inference using MRBAYES 3.0 (Ronquist & Huelsenbeck 2003), which can allow for multiple models of evolution for each DNA region (each model chosen using MODELTEST). Each individual mtDNA region was also analysed separately and no significant conflicts with the concatenated data set were found. Nested clade analysis (Templeton 1998) was not completed on our data set because minimum-spanning networks could not be constructed due to the high diversity of the sequences, even when each mitochondrial region was considered separately. There were simply too many missing theoretical haplotypes, as might be predicted from the long branch lengths leading to each lineage in the tree (Fig. 1).

Haplotype diversity (h) was calculated using DNASP 4.00 (Rozas *et al*. 2003). Measures of nucleotide diversity were made using the most complex substitution model conveniently available: nucleotide diversity (π) was calculated with Jukes–Cantor corrections using DNASP 4.00 and net nucleotide divergence (D_{xy}) between lineages was calculated with the Tamura–Nei gamma correction model using MEGA 3 (Kumar *et al*. 2004).

Tests of neutrality

We conducted two tests to determine whether selection appeared strong in our mitochondrial data. For protein-coding regions, the number of synonymous substitutions per synonymous site (K_S) was compared with the number of nonsynonymous substitutions per nonsynonymous site (K_N) and did not deviate from neutral expectations ($K_S > K_N$; $P < 0.0001$). Also, the McDonald–Kreitman test (McDonald & Kreitman 1991) for the two continuous protein-coding sequences revealed no differences in the ratios of nonsynonymous to synonymous changes within (R_P) *S. mansoni* and between (R_F) *S. mansoni* and *S. rodhaini* (Fisher's exact test: COI, $P = 0.37$; nd1, $P = 0.15$), indicating that strong selection is not acting on these genes.

Tests of host use

Shimodaira–Hasegawa (SH) tests were conducted in PAUP* to evaluate alternative trees with or without host specialization (Shimodaira & Hasegawa 1999). Trees constrained to monophyly of parasites using the same host were generated and used in the SH test compared to the unconstrained tree. For the test of snail host use, only samples obtained from snails were kept in the trees. For the definitive host tests, there were unfortunately few specimens obtained directly from the definitive hosts, rendering the test without

meaning. However, most specimens obtained in the field from snails, unless they were obtained from areas with few human hosts, can reasonably be assumed to be human-infecting. The definitive hosts test was conducted using this assumption.

Tests for founder effects

We also tested for bottleneck or founder effects where appropriate using the method of Galtier *et al*. (2000). This test assumes that historical events like bottlenecks, colonizations, and selective sweeps, through a quick increase of the coalescence rate, modify the shape of the genealogy of haplotypes, usually generating starlike genealogies. An advantage of the test is that the time (in generations) of the event is estimated using a maximum-likelihood approach, and a statistical inference is made. The test requires that no more than two states should be observed at any site, and that sequence and nucleotide sites give similar information (Galtier *et al*. 2000).

Results

Diversity and suitability of mtDNA

Mitochondrial diversity within *Schistosoma mansoni*, in contrast with nuclear ITS1 and ITS2, was substantial: the combined mitochondrial data set contains 376 variable sites (289 parsimony informative), and a total of 85 mitochondrial haplotypes. Measures of haplotype and nucleotide diversity for different regions and for within localities are presented in Tables 2 and 3, respectively. We tested for strong selection in protein-coding regions, and found no deviation from neutral expectations (methods described above). Likelihood-ratio tests suggested that in three individual mtDNA regions there has been a relatively constant rate of evolution (COI, $P = 0.263$; 12S-16S, $P = 0.652$; cyt *b*-ND4L-ND4, $P = 0.097$), although this was less true for the ND1 region ($P = 0.026$). These findings suggest that the mtDNA is suitable for tracing the evolutionary history of the parasite. A general 2–4% mitochondrial clock is often used to estimate divergence times and is appropriate here because the mtDNA regions we used are distributed throughout the mitochondrial genome and evolve at different rates. However, because three of the four regions used (COI, cyt *b*-ND4L-ND4, and ND1) evolve more quickly than the 12S-16S (Palumbi 1996), the 4% rate is used in this study (estimates based on a 2% rate are easily obtained by simply doubling the times given).

Phylogeographic structure in S. mansoni

The 85 mitochondrial haplotypes found in *S. mansoni* separated into five lineages, four of which were well supported by all phylogenetic methods (Fig. 1). These lineages were also present when each mitochondrial region was considered separately (not shown). Detail of each lineage is shown in Fig. 2. Lineages 1, 3 and 4 are comprised exclusively

Table 2 Sequence diversity (ND1) of *Schistosoma mansoni* within geographical regions

Region	*n*	u	*h*	Π
All	143	73	0.94 ± 0.014	0.0381 ± 0.00152
East Africa	52	39	0.99 ± 0.007	0.0386 ± 0.00218
Kenya	25	22	0.98 ± 0.022	0.0410 ± 0.00292
Tanzania	14	10	0.95 ± 0.045	0.0304 ± 0.00293
Uganda	9	6	0.89 ± 0.091	0.0176 ± 0.00522
Zambia	4	4	1.00 ± 0.177	0.0336 ± 0.01212
Madagascar	4	3	0.83 ± 0.222	0.0266 ± 0.00799
Oman	2	1	0.00 ± 0.000	0.0000 ± 0.00000
Egypt	10	5	0.76 ± 0.130	0.0055 ± 0.00135
Central West Africa	9	9	1.00 ± 0.052	0.0116 ± 0.00122
Cameroon	6	6	1.00 ± 0.096	0.0109 ± 0.00178
Nigeria	3	3	1.00 ± 0.272	0.0149 ± 0.00421
Far West Africa	22	14	0.96 ± 0.026	0.0064 ± 0.00062
Ghana	13	8	0.91 ± 0.056	0.0076 ± 0.00069
Liberia*	2	1	0.00 ± 0.000	0.0000 ± 0.00000
Mali	1	1	0.00 ± 0.000	0.0000 ± 0.00000
Senegal	6	4	0.87 ± 0.129	0.0025 ± 0.00069
Americas	44	6	0.49 ± 0.075	0.0034 ± 0.00065
Brazil	30	2	0.07 ± 0.061	0.0001 ± 0.00010
Guadeloupe	10	3	0.38 ± 0.033	0.0036 ± 0.00247
Puerto Rico*	2	2	1.00 ± 0.500	0.0065 ± 0.00323
Venezuela*	2	2	1.00 ± 0.500	0.0065 ± 0.00323

n, the number of worms sequenced; u, the number of unique haplotypes within the region; *h*, haplotype diversity ± standard deviation; Π, nucleotide diversity ± standard deviation.
*A majority of worms sampled were from laboratory strains.

Table 3 Within-locality sequence diversity (ND1) of *Schistosoma mansoni*

Locality*	*n*	u	*h*	Π
Homa Bay, Kenya (KE1)	6	6	1.00 ± 0.096	0.0116 ± 0.00330
Machakos, Kenya (KE7)	9	8	0.97 ± 0.004	0.0458 ± 0.00564
Tunduma, Tanzania (TZ3)	8	5	0.86 ± 0.108	0.0237 ± 0.00542
Mansafire, Egypt (EG2)	8	3	0.61 ± 0.164	0.0035 ± 0.00109
Accra site 2, Ghana (GH2)	12	7	0.89 ± 0.004	0.0017 ± 0.00066
Sabara, Brazil (BR3)	11	1	0.00 ± 0.000	0.0000 ± 0.00000
Dionisio, Brazil (BR5)	13	2	0.18 ± 0.144	0.0003 ± 0.00023
Jacquot, Guadeloupe (GU2)	9	2	0.22 ± 0.166	0.0004 ± 0.00027

n, the number of worms sequenced; u, the number of unique haplotype within a locality; *h*, haplotype diversity ± standard deviation; Π, nucleotide diversity ± standard deviation
*Only localities where six or more worms were sampled are shown.

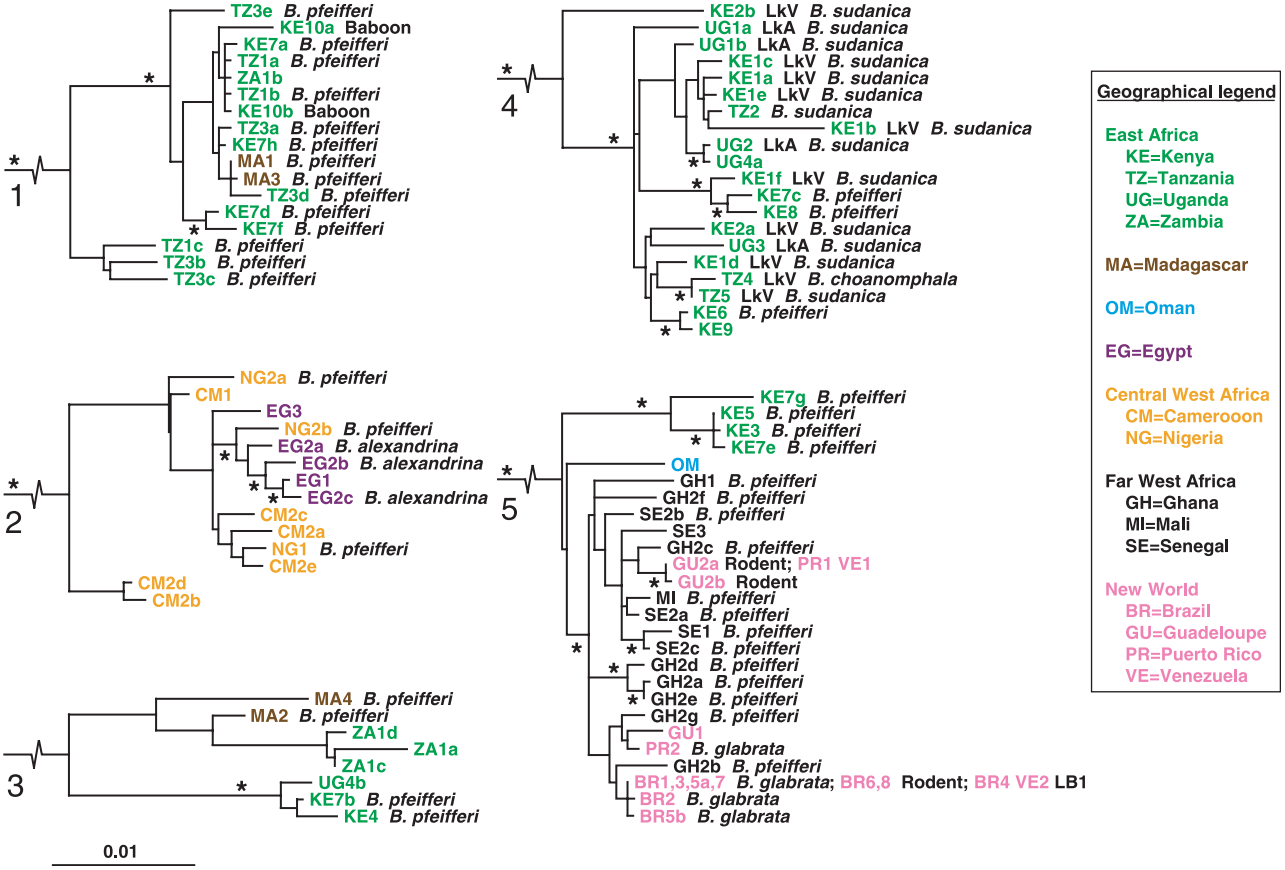

Fig. 2 Detailed view of five major lineages within the maximum-likelihood tree in Fig. 1. An asterisk indicates node support ≥ 80 for all methods used (parsimony, distance, and likelihood bootstrap analyses, and Bayesian posterior probabilities). Each terminal branch in the tree is marked as in Table 1, using a two-letter country code (also shown in the legend above), followed by a number indicating the locality and in some cases a lowercase letter differentiating haplotypes within a locality. Specimens collected from intermediate hosts (*Biomphalaria* snails) list the *Biomphalaria* species from which they were isolated. Specimens collected directly from humans do not have a host labelled for space reasons. Also, it is assumed that most of the specimens obtained from snail intermediate hosts infect humans as definitive hosts (see text). Specimens directly obtained from nonhuman definitive hosts (baboons or rodents) are also labelled. LkA, Lake Albert and LkV, Lake Victoria. Scale bar indicates 1 substitution per 100 sites.

of specimens from East Africa and Madagascar. Lineage 2 contains specimens from Central West Africa (Cameroon and Nigeria) and Egypt. The fifth and final lineage contains all of the specimens from South America, the Caribbean region, and Far West Africa (countries west of Nigeria), and at the base of this lineage lie a few East African samples. The presence of East African samples in four of the five clades is a revealing indicator of their diversity, which can also be seen in other measures (Table 2). All other areas were less diverse, with Far West Africa and the New World harbouring the least diversity. Within-locality sampling, where available, was consistent with this trend. The within-locality sampling is poor, but it is significant that the small amount of within-locality sampling easily uncovered diversity within East African localities, and almost none in New World localities (Table 3).

Host use

The maximum-likelihood tree represented in Figs 1 and 2 generally does not show much grouping of haplotypes by host use. The issue of snail host usage is somewhat confounded by geography, since many geographical regions have only one *Biomphalaria* species (e.g. *Biomphalaria alexandrina* in Egypt, *Biomphalaria pfeifferi* in West Africa). One potentially interesting area was East Africa, where several *Biomphalaria* species occur, and sometimes co-occur in a given locality (Bandoni *et al*. 2000). Lineage 4 is unique in having worms obtained from *Biomphalaria sudanica* and *Biomphalaria choanomphala*, but these are also the only parasites from Lakes Albert and Victoria, and more importantly, interspersed are a few parasites that use *B. pfeifferi* from locations very near the lakes (Fig. 2). An SH test rejected

the separation of lineage 4 haplotypes according to snail host use ($P < 0.001$). Shimodaira–Hasegawa tests rejected trees which were constrained to monophyly of baboon isolates ($P < 0.05$) or rodent isolates ($P < 0.001$).

Bottleneck tests

Due to the requirements of the test of Galtier *et al*. (2000), we were able to compute the test only for the following data sets: for Brazil, and for the New World (including Brazil). In both instances, a strong and highly significant reduction in diversity was inferred to have occurred very recently (estimates were 0 generations ago and $P < 0.01$ in both cases). We could not run the test to differentiate between demographic bottlenecks and selective sweeps since we essentially had only one marker (mtDNA).

Relationship to Schistosoma rodhaini

The two isolates of *S. rodhaini* we obtained had identical haplotypes, despite having been collected at different localities. We attempted to determine the root for the *S. mansoni* tree by using *S. rodhaini* as an outgroup, but the large genetic distances between *S. mansoni* and *S. rodhaini* caused different phylogenetic methods to determine different *S. mansoni* roots, even when third codon positions were removed (not shown). The molecular clock estimate for the divergence between *S. mansoni* and *S. rodhaini* is 2.8 ± 0.19 Ma ($D_{xy} = 0.224 \pm 0.015$), although having only a single haplotype from *S. rodhaini* overestimates this divergence (Edwards & Beerli 2000). A more recent divergence is supported by only six fixed nuclear ITS sequence differences between the two species.

Discussion

Schistosoma mansoni *exhibits high mitochondrial diversity*

The DNA sequences presented here, the first from an extensive and globally widespread set of *S. mansoni* specimens, show high mitochondrial diversity, both in haplotype and nucleotide diversity (Table 2). There were 73 haplotypes when only ND1 was considered (Table 2); 85 when all mitochondrial sequences were combined (Fig. 1).

Geographical structure vs. host usage

There is considerable geographical structure, at global and regional scales, in the haplotypes obtained in our study (Fig. 1), and we discuss these patterns below. In fact, it would appear that geographical separation over time has played a prominent role in the diversification of *S. mansoni*, far more important a role than host usage. We did not find any strong evidence that major or minor lineages within *S. mansoni* have specialized to different mammalian or snail hosts. The two baboon isolates in lineage 2 are close relatives, but several other haplotypes infecting humans are also closely related to these (Fig. 2). Also, an SH test rejected the separation of these haplotypes into monophyletic baboon- and human-infecting lineages ($P < 0.05$). Several rodent isolates from the New World are present in lineage 5, but also do not group together, and an SH test rejected rodent isolates as a monophyletic assemblage ($P < 0.001$). These results show that use of unusual mammalian hosts is not unique to a certain lineage, and that morphological or behavioural adaptations observed (Théron *et al*. 1997; D'Andrea *et al*. 2002; Freire *et al*. 2002) are likely recent and local. In the only plausible test for snail host use (see Results), an SH test rejected the separation of lineage 4 haplotypes according to snail host use ($P < 0.001$). A likely explanation for this pattern is that lineage 4 arose by geographical separation in the area of Lakes Victoria and Albert, and although *Biomphalaria sudanica* and its relatives such as *Biomphalaria choanomphala* (DeJong *et al*. 2001) are common there, the parasite lineage has retained the ability to also infect *Biomphalaria pfeifferi*. Snail, like mammalian, host use does not seem to comprise the barriers that have led to the diversification of *S. mansoni* seen in our data. Adaptation to local hosts probably does occur to some degree (e.g. the parasite may become more infective to local hosts after colonizing a new geographical area), but this adaptation does not preclude infection of a different host later in time. This has probably been particularly true in Africa because *Biomphalaria* has only been there for as little as 1.1 million years, and the African species of *Biomphalaria* are all quite closely related (DeJong *et al*. 2001).

East African origin

The conspicuous amount of diversity in East Africa, and the presence of East African haplotypes in four of the five lineages suggests an East African origin for *S. mansoni* (Figs 1 and 2). Although sampling is somewhat biased towards East Africa, the nucleotide diversity in East Africa is several times higher (3–10×) than all regions except Madagascar (Table 2). Diversity within even single sites in East Africa (Homa Bay 0.0116, Machakos 0.0458, and Tunduma 0.0237; Table 3) was 2–9× that in South America, the Caribbean region, and Far West Africa combined (0.00485 ± 0.00052; not shown in either table). It should also be noted that there is likely much East African (and perhaps Madagascaran) diversity yet unsampled, as evidenced by lineage 3 which contains some distantly related haplotypes and may represent two or more lineages (Fig. 2).

The estimate for the time of origin in East Africa is relatively recent, approximately 0.30–0.43 Ma (average D_{xy} between the five lineages 0.029 ± 0.005). These dates follow the arrival in Africa of *Biomphalaria* between 1.1 and 4.5 Ma

(Campbell *et al*. 2000; DeJong *et al*. 2001), suggesting that a new snail genus in Africa contributed to the formation of *S. mansoni* and *S. rodhaini*. It also has been suggested that the appearance of early hominins may have played a role in the divergence of *S. mansoni* from *S. rodhaini* (Desprès *et al*. 1992); however, our data indicate a more recent origin for the parasite, and *Homo ergaster*, the immediate ancestor to *Homo sapiens*, or *H. sapiens* itself, may have been the first hominin hosts. When we attempted to root the *S. mansoni* tree using *S. rodhaini*, we expected this might identify one of the East African lineages as the oldest *S. mansoni* lineage, but as stated above, we were not successful due to the large divergence of our *S. rodhaini* sequence. At this time it appears that a hypothetical ancestor at the 'centre' of the *S. mansoni* tree may be the most closely related to *S. rodhaini.*

Madagascar

The presence of Madagascar samples in two different East African lineages (Fig. 1) advocates multiple colonization events to and/or from the island. Although the isolation of Madagascar (Rabinowitz *et al*. 1983) could have led to a separate lineage of *S. mansoni*, none of the Madagascar isolates fall in basal positions within the two lineages. The simplest explanation is at least two separate and relatively recent colonizations of the island from East Africa.

Egyptian and Central West African affinities

The grouping of worms from Egypt and Central West Africa in lineage 2 was a surprising finding. Although the number of specimens and localities sampled is low, there appears to be more diversity in Central West Africa, which may suggest that *S. mansoni* may have colonized Egypt from Central West Africa. The net divergence between these two regions was 0.003 ± 0.001, indicating that such an event would have occurred within the last 25 000–50 000 years. It is interesting to note that lineage 2 is the only lineage without East African worms. More sampling could reveal East African worms that form the base of lineage 2, which would provide more support for the East African origin of *S. mansoni*, or conversely, more sampling could reveal further diversity in Central West Africa and/or Egypt, suggesting a long period of separation for this lineage. It is curious also that the worms from Oman on the Arabian Peninsula were not close relatives to those from Egypt, as might be expected according to geographical proximity, but fell into the complex lineage 5 which contains worms from East Africa, Far West Africa, as well as all the Neotropical specimens (Figs 1 and 2).

Low diversity in Far West Africa

A low diversity of isolates was seen in West Africa, suggesting only a recent existence there (though caution is warranted as more West African isolates are needed); they exhibit low nucleotide diversity and also a few East African specimens are basal to them in lineage 5 (Fig. 1). The D_{xy} (0.010 ± 0.002) between the basal East African specimens and the West African/New World worms in lineage 5 gives an estimate of 100 000–150 000 years ago for their divergence. This date follows the estimated time of the intermediate host *B. pfeifferi*'s spread to West Africa (DeJong *et al*. 2003). The absence of other *Biomphalaria* in West Africa prior to this time may have precluded *S. mansoni* from establishing there.

Colonization and founder effect in the New World

While a number of parasites that infect humans are thought to be post-Columbian colonists of the New World, this has been confirmed genetically only for the malaria agent *Plasmodium falciparum* (Joy *et al*. 2003), the liver fluke *Fasciola hepatica* (Mas-Coma *et al*. 2001), and the filarial worm *Onchocerca volvulus* (Zimmerman *et al*. 1994; Keddie *et al*. 1999). Previous studies of *S. mansoni* had found evidence consistent with a New World colonization (Files 1951; Fletcher *et al*. 1981; Desprès *et al*. 1993), but were not able to show the expected genealogical relationship between New World specimens and West African compared to East African specimens. In this study, the grouping of all West African, Caribbean and South American specimens into one well-supported lineage (derived part of lineage 5 in Figs 1 and 2), and the low number of haplotypes with low nucleotide diversity found in the New World, provides the strongest evidence to date that *S. mansoni* is a recent introduction to the New World, almost certainly with the African slave trade of the 15th to 19th centuries. The close relationship between New World and West African isolates is consistent with West African origins for many slaves (Lovejoy 1982).

The low number of haplotypes (7) in the New World is noteworthy given that 14 localities and a wide geographical range were sampled. It appears that despite multiple West African sources of slaves (Lovejoy 1982), genetic diversity of *S. mansoni* in the New World was reduced from that in Far West Africa, as was confirmed by the significant results from the test of Galtier *et al.* (2000). Even more remarkable was that worms from the eight geographically widespread localities from Brazil and one from Venezuela are nearly identical across the mitochondrial markers: one base substitution was present in the COI gene of one Brazilian specimen, and one substitution was present in the ND1 gene of another. Similarly, Desprès *et al.* (1993) also found very little variation among three Brazilian samples. In our study, the Galtier *et al.* (2000) test showed that a severe reduction in mitochondrial diversity occurred in the establishment of the parasite in Brazil. These reductions in diversity could have resulted from either a founder effect

(reduced number of parasites surviving the journey which slaves were forced to make) or from a selective sweep, imposed by a new environment and a new intermediate host, or both demographics and selection. We were unable to test for a selective sweep due to the lack of a second genetic marker, and consider the founder effect a plausible and simple explanation given the current data.

Few mutations would be expected in 500 years since the beginning of the slave trade, so the presence of multiple New World haplotypes suggests introductions from multiple West African sources, and one would expect to be able to find New World haplotypes in their West African region of origin. We found one such instance of a shared haplotype: Liberia, which had an identical haplotype to the most common Brazilian haplotype. However, we urge caution in interpreting this result because the Liberian isolate available to us was from a strain that has been maintained in the laboratory for more than 50 years (Gonnert 1955). We sampled from seven West African localities, but more robust West African sampling, especially from more regions which were sources of slaves and from which parasite material has been difficult to obtain, might yield haplotypes that match those in the New World.

Implications

Schistosoma mansoni shows high levels of mitochondrial differentiation among regions in Africa, and in fact, in East Africa, there is high genetic diversity within single localities. Though ITS sequences are conserved, there is probably some variation in the nuclear genome that correlates with mitochondrial variation, such that places with high mitochondrial diversity also have the highest nuclear diversity. Worms of different genetic backgrounds could respond differently to control measures such as drugs and vaccines. It would be prudent for researchers involved in the development and implementation of these tools against the disease to test them against *S. mansoni* strains from multiple geographical locations, including representatives from each of the lineages uncovered in this study. Also, as the complete genome sequence of *S. mansoni* is imminently completed (Chitsulo *et al*. 2004), it should be remembered that it is representative of a diverse parasite.

As a massive control effort is now underway, in which praziquantel, the current drug of choice to cure schistosomiasis, is being provided to the populations of whole countries (Crompton *et al*. 2003; www.who.int/wormcontrol/en/), signs of drug resistance need to be monitored. It is possible that high genetic diversity in worms, particularly in East Africa, might make development of drug resistance more likely, though the putative examples of praziquantel resistance thus far have been from areas (Egypt, Senegal; Doenhoff *et al*. 2002) that seem to have relatively lower diversity. It should be noted as well that worms from these areas are members of two different mtDNA lineages (lineages 2 and 5, respectively), so resistance appears not to have been confined to any one lineage.

It is interesting to consider that in East African localities where different lineages are present, worms from these lineages almost certainly are present together within individual human hosts. Do these genetic lineages mate with each other? Do they compete? Does the presence of multiple genetic lineages affect pathogenicity in the human host? A recent study (Gower & Webster 2005) found evidence for competition between *S. mansoni* strains of differing virulence co-infecting the intermediate host. These types of questions have not been addressed for human hosts, and the diversity within individual human hosts has only been assessed with one study (Curtis *et al*. 2002), in Brazil. Moderate microsatellite diversity was found within patients and genetic subdivision was found among them. Similar studies in areas with the other extreme of *S. mansoni* diversity, such as East Africa, may yield much different results, and it may be instructive to compare such findings. Indeed, *Schistosoma mansoni* research should include a substantial focus on East Africa, where the parasite originated and where it continues to thrive today.

Acknowledgements

We are grateful to Joseph Kinuthia, Geoffrey Maina, Timothy Kamau, Julius Andove of the Kenya Medical Research Institute (KEMRI), Jimmy Kihara and Francis Mungai of the Division of Vector-Borne Diseases, Ministry of Health, Kenya, for logistical support and assistance in the field. The Director of KEMRI is thanked for publication approval. We thank Sarah Tishkoff, Alain Froment, and Bernard Wood for feedback providing an anthropological perspective. We acknowledge technical support from the University of New Mexico's Molecular Biology Facility which is supported by NIH Grant number 1P20RR18754 from the Institute Development Award (IDeA) Program of the National Center for Research Resources. All authors collected *S. mansoni* isolates, and thank L. Mubila for providing isolates from Zambia, with the assistance of M. Shehata. M. Tanner and C. Grevelding are thanked for providing the Liberian isolates from each of their laboratories. Funding was provided by U.S. National Institutes of Health Grant AI44913 to G.M.M. and E.S.L.

References

Bandoni SM, Mulvey M, Loker ES (2000) Population structure and taxonomic discrimination among three species of *Biomphalaria* Preston, 1910 (Gastropoda: Planorbidae) from Kenya. *Zoological Journal of the Linnean Society*, **129**, 387–401.

Campbell G, Jones CS, Lockyer AE *et al*. (2000) Molecular evidence supports an African affinity of the Neotropical freshwater gastropod, *Biomphalaria glabrata*, Say 1818, an intermediate host for *Schistosoma mansoni*. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **267**, 2351–2358.

Chitsulo L, Engels D, Montresor A, Savioli L (2000) The global status of schistosomiasis and its control. *Acta Tropica*, **77**, 41–51.

Chitsulo L, Loverde R, Engels D *et al.* (2004) Schistosomiasis. *Nature Reviews Microbiology*, **2**, 12–13.

Crompton DWT (1999) How much human helminthiasis is there in the world? *Journal of Parasitology*, **85**, 397–403.

Crompton DWT, Engels D, Montresor A, Neira MP, Savioli L (2003) Action starts now to control disease due to schistosomiasis and soil-transmitted helminthiasis. *Acta Tropica*, **86**, 121–124.

Curtis J, Sorensen RE, Minchella DJ (2002) Schistosome genetic diversity: the implications of population structure as detected by microsatellite markers. *Parasitology*, **125**, S51–S59.

D'Andrea PS, Fernandes FA, Cerqueira R, Rey L (2002) Experimental evidence and ecological perspectives for the adaptation of *Schistosoma mansoni* Sambon, 1907 (Digenea: Schistosomatidae) to a wild host, the water-rat, *Nectomys squamipes* Brants, 1827 (Rodentia: Sigmodontinae). *Memorias Do Instituto Oswaldo Cruz*, **97**, 11–14.

Davis GM (1980) Snail hosts of Asian *Schistosoma* infecting man: evolution and coevolution. *Malcological Review*, **Suppl. 2**: 195–238.

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

DeJong RJ, Morgan JAT, Paraense WL *et al.* (2001) Evolutionary relationships and biogeography of *Biomphalaria* (Gastropoda: Planorbidae) with implications regarding its role as host of the human bloodfluke, *Schistosoma mansoni*. *Molecular Biology and Evolution*, **18**, 2225–2239.

DeJong RJ, Morgan JAT, Wilson WD *et al.* (2003) Phylogeography of *Biomphalaria glabrata* and *B. pfeifferi*, important intermediate hosts of *Schistosoma mansoni* in the New and Old World tropics. *Molecular Ecology*, **12**, 3041–3056.

Desprès 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.

Desprès L, Imbert-Establet D, Monnerot M (1993) Molecular characterization of mitochondrial-DNA provides evidence for the recent introduction of *Schistosoma mansoni* into America. *Molecular and Biochemical Parasitology*, **60**, 221–230.

Doenhoff MJ, Kusel JR, Coles GC, Cioli D (2002) Resistance of *Schistosoma mansoni* to praziquantel: is there a problem? *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **96**, 465–469.

Edwards SV, Beerli P (2000) Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution*, **54**, 1839–1854.

Eppert A, Lewis FA, Grzywacz C, Coura P, Caldas I, Minchella DJ (2002) Distribution of schistosome infections in molluscan hosts at different levels of parasite prevalence. *Journal of Parasitology*, **88**, 232–236.

Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution*, **17**, 368–376.

Files VS (1951) A study of the vector-parasite relationships in *Schistosoma mansoni*. *Parasitology*, **41**, 264–269.

Fletcher M, Loverde PT, Woodruff DS (1981) Genetic variation in *Schistosoma mansoni* — enzyme polymorphisms in poulations from Africa, southwest Asia, South America, and the West Indies. *American Journal of Tropical Medicine and Hygiene*, **30**, 406–421.

Freire N, Machado-Silva JR, Rodrigues-Silva R, Rey L (2002) Cercarial chaetotaxy and sex differentiation of *Schistosoma mansoni* deriving from humans and *Nectomys squamipes* (Muridae: Sigmondontinae) in Brazil. *Memorias Do Instituto Oswaldo Cruz*, **97**, 127–132.

Galtier N, Depaulis F, Barton NH (2000) Detecting bottlenecks and selective sweeps from DNA sequence polymorphism. *Genetics*, **155**, 981–987.

Gonnert R (1955) Schiztosomiasis studies. I. Contributions to the anatomy and histology of *Schistosoma mansoni*. *Zeitschrift fur Tropenmedizine und Parasitologie*, **6**, 18–33.

Gower CM, Webster JP (2005) Intraspecific competition and the evolution of virulence in a parasitic trematode. *Evolution*, **59**, 544–553.

Joy DA, Feng XR, Mu JB *et al.* (2003) Early origin and recent expansion of *Plasmodium falciparum*. *Science*, **300**, 318–321.

Keddie EM, Higazi T, Boakye D *et al.* (1999) *Onchocerca volvulus*: limited heterogeneity in the nuclear and mitochondrial genomes. *Experimental Parasitology*, **93**, 198–206.

Kumar S, Tamura K, Nei M (2004) MEGA 3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefings in Bioinformatics*, **5**, 150–163.

Lambertucci JR, Serufo JC, Gerspacher-Lara R *et al.* (2000) *Schistosoma mansoni*: assessment of morbidity before and after control. *Acta Tropica*, **77**, 101–109.

Leclerc MC, Durand P, Gauthier C *et al.* (2004) Meager genetic variability of the human malaria agent *Plasmodium vivax*. *Proceedings of the National Academy of Sciences, USA*, 14455–14460.

Lockyer AE, Olson PD, Ostergaard P *et al.* (2003) The phylogeny of the Schistosomatidae based on three genes with emphasis on the interrelationships of *Schistosoma* Weinland, 1858. *Parasitology*, **126**, 203–224.

Lovejoy PE (1982) The volume of the the Atlantic slave trade — a synthesis. *Journal of African History*, **23**, 473–501.

Mas-Coma S, Funatsu IR, Bargues MD (2001) *Fasciola hepatica* and lymnaeid snails occurring at very high altitude in South America. *Parasitology*, **123**, S115–S127.

McDonald JH, Kreitman M (1991) Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature*, **351**, 652–654.

Morgan JAT, DeJong RJ, Kazibwe F, Mkoji GM, Loker ES (2003a) A newly-identified lineage of *Schistosoma*. *International Journal for Parasitology*, **33**, 977–985.

Morgan JAT, DeJong RJ, Lwambo NJS *et al.* (2003b) First report of a natural hybrid between *Schistosoma mansoni* and *S. rodhaini*. *Journal of Parasitology*, **89**, 416–418.

Palumbi SR (1996) Nucleic acids II: the polymerase chain reaction. In: *Molecular Systematics* (eds Hillis DM, Moritz C Mable BK), pp. 205–248. Sinauer Associates, Sunderland, Massahusetts.

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.

Rabinowitz PD, Coffin MF, Falvey D (1983) The separation of Madagascar and Africa. *Science*, **220**, 67–69.

Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572–1574.

Rozas J, Sanchez-DeI, Barrio JC, Messeguer X, Rozas R (2003) DNASP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19**, 2496–2497.

Shimodaira H, Hasegawa M (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular Biology and Evolution*, **16**, 1114–1116.

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

Swofford DL (2004) PAUP*: *Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Sinauer Associates, Sunderland, Massachusetts.

Templeton AR (1998) Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology*, **7**, 381–397.

Théron A, Mouahid G, Moné H (1997) *Schistosoma mansoni*: cercarial shedding patterns from a mixed infection of *Biomphalaria glabrata* with two (early and late) chronobiological variants. *Parasitology Research*, **83**, 356–358.

Zhang DX, Hewitt GM (2003) Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Molecular Ecology*, **12**, 563–584.

Zimmerman PA, Katholi CR, Wooten MC, Langunnasch N, Unnasch TR (1994) Recent evolutionary history of American *Onchocerca volvulus*, based on analysis of a tandemly repeated DNA-sequence family. *Molecular Biology and Evolution*, **11**, 384–392.

This study is a key part of a project to assess the global diversity and the evolutionary history of *Schistosoma mansoni* and its snail hosts. Jess Morgan and Randall DeJong completed the molecular work, data analysis, and manuscript preparation and continue using molecular tools to investigate the interactions of parasitic organisms and their hosts. Gerald Mkoji and Eric (Sam) Loker continue to collaborate on projects on the interactions between genetic diversity and epidemiology of *Schistosoma mansoni* in Kenya.
