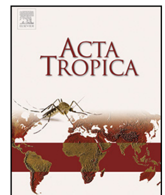
Acta Tropica 128 (2013) 261–274

Contents lists available at ScienceDirect

Acta Tropica

journal homepage: www.elsevier.com/locate/actatropica



Population genetic structure of *Schistosoma mansoni* and *Schistosoma haematobium* from across six sub-Saharan African countries: Implications for epidemiology, evolution and control

Charlotte M. Gower[a], Anouk N. Gouvras[a], Poppy H.L. Lamberton[a], Arminder Deol[a], Jaya Shrivastava[a], Polydor N. Mutombo[a], Judith V. Mbuh[a], Alice J. Norton[a,b], Bonnie L. Webster[a,c], J. Russell Stothard[c,d], Amadou Garba[e], Mariama S. Lamine[e], Curtis Kariuki[f], Charles N. Lange[g], Gerald M. Mkoji[g], Narcis B. Kabatereine[h], Albis F. Gabrielli[a], James W. Rudge[a], Alan Fenwick[a], Moussa Sacko[i], Robert Dembelé[i], Nicholas J.S. Lwambo[j,1], Louis-Albert Tchuem Tchuenté[k], David Rollinson[c], Joanne P. Webster[a,*]

[a] Department of Infectious Disease Epidemiology, School of Public Health, Imperial College Faculty of Medicine, St Mary's Hospital Campus, Norfolk Place, London W2 1PG, UK
[b] Wellcome Trust, 215 Euston Road, London NW1 2BE, UK
[c] Wolfson Wellcome Biomedical Laboratories, Department of Zoology, Natural History Museum, Cromwell Road, London SW7 5BD, UK
[d] Disease Control Strategy Group, Liverpool School of Tropical Medicine, Liverpool L3 5QA, UK
[e] Réseau International Schistosomoses, Environnement, Aménagement et Lutte (RISEAL-Niger), 333, Avenue des Zarmakoye, B.P. 13724, Niamey, Niger
[f] National Museums of Kenya, Dept of Invertebrate Zoology, PO Box 40658, Nairobi, Kenya
[g] Kenya Medical Research Institute, PO Box 58540, Mbagathi Road, Nairobi, Kenya
[h] Vector Control Division, Ministry of Health, P.O. Box 1661, Kampala, Uganda
[i] Institut National de Recherche en Santé Publique, Ministère de la Santé, Bamako, Mali; Service de Radiologie, Hôpital National du Point G, Bamako, Mali
[j] Mwanza Research Centre, National Institute for Medical Research, Mwanza, PO Box 1462, Tanzania
[k] Laboratoire de Parasitologie et Ecologie, Université de Yaoundé I, Yaoundé, Cameroon

ARTICLE INFO

Article history:
Available online 3 October 2012

Keywords:
Schistosoma mansoni
Schistosoma haematobium
Genetic diversity
Population structure
Evolution
Africa

ABSTRACT

We conducted the first meta-analysis of ten *Schistosoma haematobium* (one published and nine unpublished) and eight *Schistosoma mansoni* (two published and six unpublished) microsatellite datasets collected from individual schistosome-infected school-children across six sub-Saharan Africa countries. High levels of genetic diversity were documented in both *S. haematobium* and *S. mansoni*. In *S. haematobium* populations, allelic richness did not differ significantly between the ten schools, despite widely varying prevalences and intensities of infection, but higher levels of heterozygote deficiency were seen in East than in West Africa. In contrast, *S. mansoni* populations were more diverse in East than West African schools, but heterozygosity levels did not vary significantly with geography. Genetic structure in both *S. haematobium* and *S. mansoni* populations was documented, at both a regional and continental scale. Such structuring might be expected to slow the spread to new areas of anti-schistosomal drug resistance should it develop. There was, however, limited evidence of genetic structure at the individual host level, which might be predicted to promote the development or establishment of drug resistance, particularly if it were a recessive trait. Our results are discussed in terms of their potential implications for the epidemiology and evolution of schistosomes as well as their subsequent control across sub-Saharan Africa.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

There is considerable current interest and activity in the area of parasite population genetic dynamics, including that relating to those schistosome species of profound clinical and/or veterinary importance. Schistosomiasis infects over 206 million people, particularly the rural poor of the developing world, more than half of which are symptomatic and at least 20 million exhibit severe disease manifestations (King et al., 2005; Steinmann et al., 2006). Elucidating the genetic structure of natural schistsosome populations is important in terms of helping us undertand many aspects of disease epidemiology and transmission (Slatkin, 1987; Rudge et al., 2008; Agola et al., 2006, 2009; Steinauer et al., 2009;

* Corresponding author. Tel.: +44 20 7594 3636; fax: +44 207 402 3927.
E-mail address: joanne.webster@imperial.ac.uk (J.P. Webster).
[1] Deceased.

0001-706X/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.actatropica.2012.09.014

Lu et al., 2010). This may be particularly pertinent for predicting potential intervention-induced evolutionary change in parasite populations (Norton et al., 2010; Webster et al., 2008), especially perhaps given the recent move towards schistosome elimination rather than control in some areas (Knopp et al., 2013). Within the last few years we have obtained the sequences for the *Schistosoma mansoni* and *S. japonicum* genomes (Berriman et al., 2009; *Schistosoma japonicum* Genome Sequencing Functional Analysis Consortium, 2009; Webster et al., 2010), and also now that of the *S. haematobium* genome (Young et al., 2012). Furthermore, we now have initial microsatellite markers for these three major species (*S. mansoni* (Durand et al., 2000; Blair et al., 2001; Curtis et al., 2001); *S. japonicum* (Shrivastava et al., 2003); *S. haematobium* (Golan et al., 2007; Gower et al., 2011)). Schistosomes, however, present inherent logistical sampling problems for population genetic studies, since only egg or larval stages are directly available from living hosts. However, we now also have the methodology to collect, store and genotype larval schistosome samples directly from human (or animal) infections (Shrivastava et al., 2005; Gower et al., 2007, 2011), thereby avoiding the sampling, ethical and considerable logistical and temporal disadvantages and biases of laboratory passage.

One important aspect which has not been possible to investigate until now, however, is the extent and scale of any geographic structuring in schistosome populations on a continental level, and whether common features relating to patterns of schistosome genetic diversity can be identified in different geographic areas. Moreover, the two main causative agents of human schistosomiasis in sub-Saharan Africa (SSA), *S. mansoni* and *S. haematobium*, have important differences in their epidemiology, clinical manifestations and transmission – although understanding the extent of these differences and their subsequent influences on the impact of schistosomiasis treatment and control have been severely neglected (Rollinson, 2009). Population genetic studies, including those ideally incorporated into the routine monitoring and evaluation (M&E) of large-scale preventive chemotherapy programmes, would enhance understanding of the genetic diversity, presence and nature of population structure and gene flow in such human helminth populations over space and time. Furthermore, any potential changes in allele frequency or composition mediated by such large-scale drug administration can be monitored in order to determine the impact of treatment on the parasite population as a whole, inform of the size of likely *refugia*, as well as the likelihood of the spread of drug resistance, should it develop, in the population structure observed (Webster et al., 2008). Another potential population genetic parameter of interest is that of the effective number of breeding parasites in the population (*Nb*). This metric has the potential to measure the impact of treatment at a community level, to study the population dynamic processes by which elimination is achieved (or not), and/or the identification of areas where elimination might be most successful. These have not yet, however, been well studied with respect to schistosomes, or indeed parasites in general (Criscione et al., 2005).

We present here a meta-analysis of three published and 15 unpublished data sets obtained from six sub-Saharan Africa (SSA) countries, collected prior to the implementation of school-based national praziquantel (PZQ) treatment programmes. *S. mansoni* populations were collected from individual children in eight schools: two in Kenya, two in Tanzania (Norton et al., 2010), one in Uganda, two in Niger and one in Mali, while *S. haematobium* were collected from children at 10 schools: two in Kenya, two in Tanzania, one in Cameroon, three in Niger and two in Mali (Gower et al., 2011). Our primary aim was to use microsatellite markers to characterise the population genetic structure of *S. mansoni* and *S. haematobium* from multiple locations across the African continent, separated by thousands of kilometres, with varying water sources, in order to elucidate whether population structure occurs, and over what scales. Since it is necessary to use differing microsatellite markers for the different schistosome species, absolute measures of genetic diversity were not directly compared between species, but rather the nature of the patterns observed. We primarily report on analyses between the parasite populations of the different schools, but more detailed analyses of local variation at the level of parasite populations of individual hosts of the highly neglected *S. haematobium*, and associations with individual host characteristics such as age and gender, were conducted using *S. haematobium* samples collected in Cameroon as a key example. Finally, we also attempted to use genetic data to estimate the effective number of breeders in the populations using available statistical software. Our results are discussed with regard to their analytical, epidemiological and evolutionary implications and applications for human schistosomiasis in SSA.

2. Materials and methods

2.1. Collection of parasite material

Schistosome samples were collected from school-children aged 6–16 years in six SSA countries as shown in Fig. 1 and Table 1. All samples, with the exception of the Mali *S. mansoni* dataset, were collected in villages that had not, to the authors' knowledge, been previously treated with PZQ. The Mali *S. mansoni* dataset consisted of PZQ-naïve children (i.e. those who had not themselves been previously treated), but was collected two years following the introduction of their National Schistosomiasis Control Programme (as true baseline data were not available for this parasite in Mali). The number of children ranged from one to 53 children per school and a total of 144 and 183 parasite infrapopulations (here defined as the miracidia obtained from a single child; French et al., 2013) were collected for *S. haematobium* and *S. mansoni*, respectively.

Children were identified as schistosome-positive by urine filtration (*S. haematobium*) or duplicate Kato-Katz thick smears (*S. mansoni*). Epidemiological data (e.g. age, gender, infection intensity, coinfections with other helminths and morbidity parameters) were recorded for each child, wherever possible. All infected children were treated with PZQ at 40 mg/kg after sample collection. All collections were made after having obtained ethical approval from the responsible national committees (see below).

For *S. mansoni* isolation, stool samples from each infected child were prepared separately for miracidial hatching by breaking them up with a toothbrush and passing them through a 425 μm sieve with approximately 1 L bottled spring water. The resulting suspension was then poured through a Pitchford funnel consisting of two meshes; the inner mesh of 200 μm allowed eggs to pass through but collected any larger debris and the outer mesh of 40 μm collected the eggs. The outer mesh was washed through again with approximately 1 L bottled spring water and the eggs were released into a Petri dish. Similar filtration was conducted for *S. haematobium* isolation, using collected urine samples. Miracidial hatching from the eggs was stimulated by placing the Petri dish in indirect sunlight for up to 6 h (Lamberton et al., 2010).

For DNA storage, following successful hatching of eggs, individual miracidia were picked up using a P20 micropipette under the binocular microscope and transferred to Whatman FTA® indicator cards (Whatman plc; Maidstone, UK) and allowed to dry in the shade for one hour. Up to 100 miracidia were collected per child, depending on infection intensity.

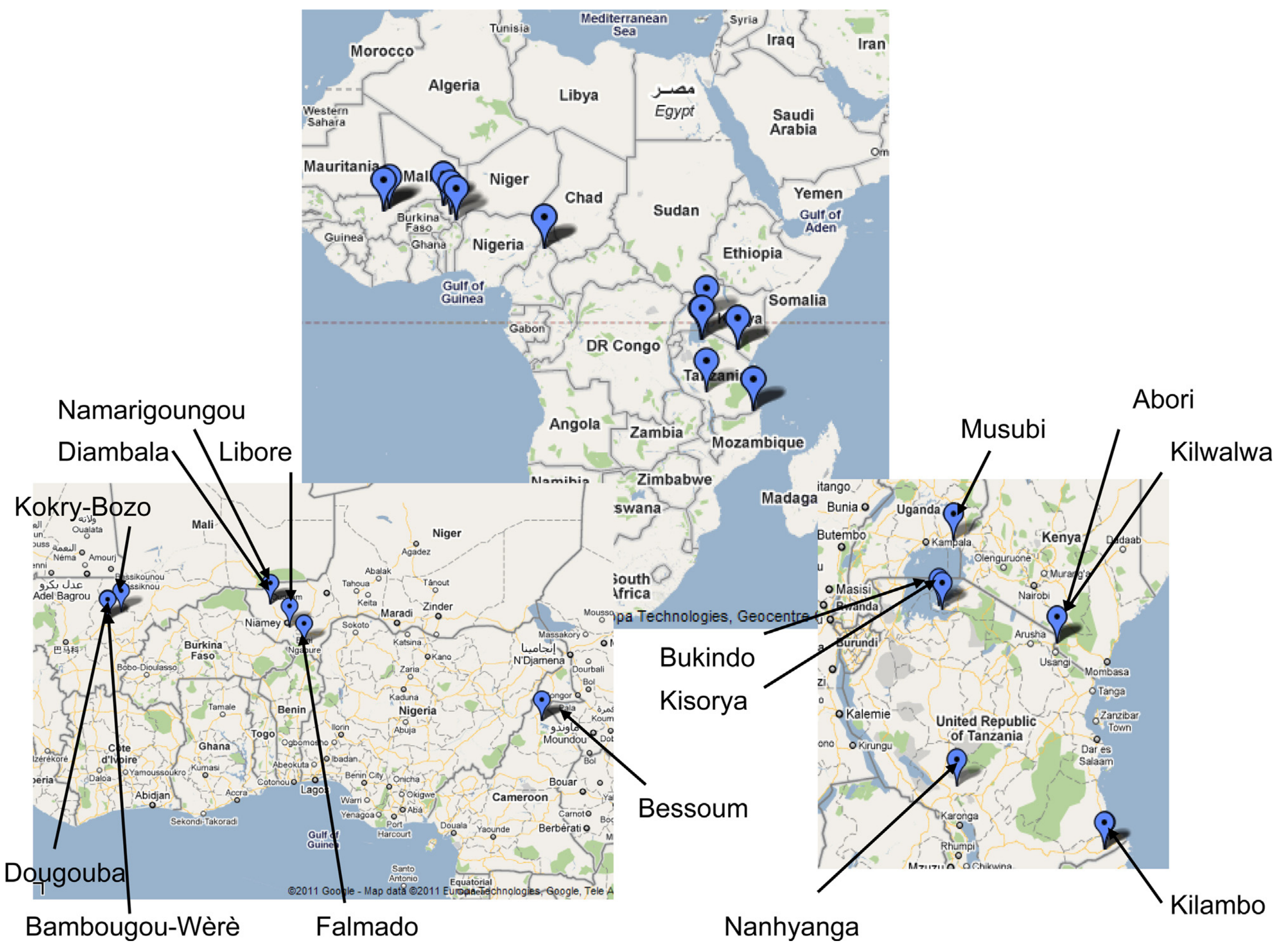


Fig. 1. Location of the eight schools across the six sub-Saharan African (SSA) countries from which *S. haematobium* and *S. mansoni* miracidia were collected.

2.2. Microsatellite genotyping of individual larvae

Individual miracidia were subject to DNA extraction and multiplex polymerase chain reaction (PCR) for seven *S. mansoni* [SMD28, SMDA28, SDM25, SMD89, CA11-1, SMS9-1 and SMU31768] (Gower et al., 2007; Norton et al., 2010) and eight *S. haematobium* microsatellite loci [C102, C11, C111, C116, C131, C146, C2 and D3] (Gower et al., 2011) using previously published protocols. However, C2 was recently identified as a duplicate of C111 and thus was excluded from all analyses here. Extensive previous empirical

Table 1
Location and features of *S. haematobium* and *S. mansoni* sampling sites.

Country	Place	School or village?	District	Geographic coordinates	Year	Mean age	Age range	No. of children	Prevalence	Major water course
Schistosoma haematobium										
Kenya	Abori	School	Taveta	−3.40, 37.68	2007	10.4	8 to 13	15	40.0	Lake Jipe
Kenya	Kiwalwa	School	Taveta	−3.40, 37.68	2007	9.7	7 to 12	14	40.0	Lake Jipe
Tanzania	Kilambo	School	Mtwara	−8.55, 33.80	2005	Unknown	Unknown	4	Unknown	River Rovuma
Tanzania	Nanhyanga	School	Mtwara	−10.80, 39.60	2005	Unknown	Unknown	1	Unknown	River Rovuma
Cameroon	Bessoum	School	Lagdo	9.13, 13.71	2007	10.5	6–16	33	87.0	River Benue
Niger	Falmado	Village	Dosso	21.51, −2.86	2007	8.49	6–10	13	99.0	Marshes fed by River Niger
Niger	Libore	School	Kollo	13.40, 2.19	2007	9.58	7–14	11	93.0	River Niger
Niger	Namarigoungou	School	Tillaberi	14.30, 1.30	2007	8.65	7–10	6	46.5	River Niger
Mali	Bambougou	School	Segou	13.60, −6.08	2005	9.04	6–13	22	59.0	River Niger
Mali	Dougouba	School	Segou	13.60, −6.11	2005	11.52	6–15	25	59.0	River Niger
Schistosoma mansoni										
Kenya	Abori	School	Taveta	−3.40, 37.68	2007	10.14	8–12	7	10.2	Lake Jipe
Kenya	Kiwalwa	School	Taveta	−3.40, 37.68	2007	10.60	7–15	5	10.2	Lake Jipe
Tanzania	Bukindo	School	Mwanza	−2.03, 33.10	2005	9.08	7–12	38	60.0	Lake Victoria
Tanzania	Kisorya	School	Mara	−2.14, 31.22	2005	8.15	7–11	42	91.0	Lake Victoria
Uganda	Musubi	School	Lake Victoria	0.35, 33.67	2007	8.4	8–9	5	75.7	Lake Victoria
Niger	Diambala	School	Tillaberi	14.31, 1.30	2007	12.4	10–14	5	62.4	River Niger
Niger	Namarigoungou	School	Tillaberi	14.30, 1.30	2007	9.0	7–13	28	66.2	River Niger
Mali	Kokry-bozo	School	Office du Niger	13.96, −5.51	2007	9.81	6–14	53	88.0	Irrigation channels fed by River Niger

work had confirmed the reliability of the multiplex reaction in assigning individual genotypes when amplified in this eight locus multiplex, and thus C2 was only excluded at the analytical stage. All genotyping of all samples took place in the same central laboratory (Imperial College, London) to ensure consistency and prevent bias.

The number of microsatellite loci used was limited by the number of alleles it was possible to amplify from a single miracidia at the time of collection and analyese. More recent work, such as the development of whole genome amplification techniques (Valentim et al., 2009) now allow repeated genotyping of individual larvae at a larger number of loci, and will be used in future studies. Nevertheless, separate statistical investigation of the *S. mansoni* assay, at least, have suggested that similar conclusions regarding population structure are robust to the removal of any 1, 2 or 3 of the microsatellite loci (French et al., 2013), thereby providing further support for the validity of the current analyses.

2.3. Data analysis

2.3.1. Between school characterisation and comparison of S. mansoni and S. haematobium

2.3.1.1. Diversity of infections. Allele sizes were calculated in Genemapper version 4 and manually checked. Analysis was restricted to miracidia with at least three allele calls (*S. mansoni*) or four allele calls (*S. haematobium*). Powermarker (Liu and Muse, 2005) was used to calculate the total number of alleles per locus, the major allele frequency, the expected and observed heterozygosity and the inbreeding coefficient for the parasite population of each child). Since the total number of alleles is affected by sample size, which varied between individual children due to differences in infection intensity, allelic richness – a measure of allelic diversity which rarefies to the smallest sample size and hence is independent of sample size – was calculated in FSTAT version 2.9.3.2 (Goudet, 2002). Analyses of allelic richness were conducted using the entire dataset for each schistosome species, such that the schistosome populations of individual children were rarefied to the same minimum sample size. The mean of within-child estimates was calculated for each school. In separate analyses, all data from a single school were pooled (see Component population; French et al., 2013) and the total number of alleles per locus, major allele frequency, expected and observed heterozygosity, inbreeding coefficient and allelic richness were calculated. A non-parametric permutation test, with 15,000 permutations, was also used to investigate any difference in the genetic diversity of parasite populations between geographic regions (FSTAT version 2.9.3.2). For *S. haematobium*, this consisted of assigning the parasite populations of individual schools into two groups: (i) East Africa, containing the parasite populations of all children from Kenya and Tanzania; and (ii) West Africa, containing all children from Mali and Niger, while for *S. mansoni*, Ugandan populations were also included in the East African group. The Cameroonian *S. haematobium* population was excluded from this analysis as it was the single representative of the Central African region.

2.3.1.2. Population structure. As a measure of genetic distance between the parasite populations of individual children, a matrix of Cavalli-Sforza and Edwards' chord distances (Cavalli-Sforza and Edwards, 1967) was estimated using Powermarker and visualised using a neighbour joining clustering algorithm (NJ). The reliability of phenograms was assessed by bootstrapping over loci with 100 replications using CONSENSE (Felsenstein, 2005). Two separate analyses were carried out for each schistosome species, one comparing the parasite populations of individual children, as above and one pooling all miracidia from a single school.

Hierarchical Wrights' Fst statistics measuring evidence of genetic differentiation between countries, schools and between children within schools were calculated for each of the two schistosome species using an analysis of molecular variance (AMOVA) procedure in Arlequin version 3.11 (Excoffier et al., 2005). *P*-values were calculated by 10,000 random permutations. A hierarchical analysis is necessary as it accounts for the potential relatedness of miracidia within individual children, which will tend to inflate estimates of classical Wrights Fst statistics (Rudge et al., 2008). Two separate analyses were conducted for each schistosome species due to limitations in the hierarchical levels, inherent within the software. Pairwise Fst statistics were calculated between the parasite populations of each of the 10 (*S. haematobium*) or eight (*S. mansoni*) schools (see Table 1 for list of schools). The correlation between the matrix of (Fst-1/Fst) genetic distance and a matrix of geographic distance was investigated for each schistosome species using Mantel tests. These analyses were conducted in XLSTAT-Pro 7.5 (Adinsoft Inc; New York, USA). The matrices of geographic distances between schools was calculated from geographic coordinates (latitude and longitude) using the haversine formula to calculate great-circle distance between two points, using the code available at http://www.movable-type.co.uk/scripts/latlong.html. Formulae used assumed a spherical earth (ignoring ellipsoidal effects), and thus errors of approximately 0.3% are expected in the distance calculations (or 3 m in every km).

2.3.1.3. Estimation of effective numbers of breeders (Nb). Two software programmes were compared here to estimate the *Nb* of the parasite population in each school, namely MLNe (Wang, 2001; Wang and Whitlock, 2003) which implements a pseudo-likelihood method for computing *Ne* based on allele frequencies and LDNe (Waples and Do, 2008), which estimates *Ne* based on linkage disequilibrium. (*Ne* is the average number of individuals in a population that contribute genes to succeeding generations, and thus is equivalent to *Nb* in an organism with overlapping generations). Miracidia were pooled at the school level and for MLNe, a maximal estimate of *Ne* of 32,000 was assigned, close to the upper limitation of the software. 95% confidence limits of estimates are reported. This software uses the temporal method for estimating *Ne* when there are two or more temporal samples, but also estimates *Ne* for a single isolated population as a special case (Wang and Whitlock, 2003). A minimal allele frequency of 0.01 was used in LDNe calculations. Estimated *Nb* for each village were investigated using general linear modelling, using schistosome species, geography (East Africa or West Africa), mean age of children, miracidial sample size, infection intensity and prevalence as the dependent variables. Correlations of the estimates of *Nb* using the two different software programmes, and the mean allelic richness, were investigated using Pearson correlation coefficients.

2.3.2. Additional individual level analyses of S. haematobium in Cameroon

Additional detailed analyses were conducted separately on the *S. haematobium* parasite samples collected in Bessoum School, Cameroon, which were also included in the meta-analyses of the datasets described above, as an example of the patterns of variation present in the parasite populations of individual children, and any associations with the epidemiological characteristics of individual hosts, such as age and gender. Measures of genetic diversity were calculated as for the meta-analyses described above and the parasite populations of individual children for children of different ages and sex were compared using general linear modeling Minitab Statistical in Software (Minitab Inc; State College, PA, USA), controlling for miracidial sample size. Similarly, population structure was investigated using Wright's Fst statistics and clustering analyses.

3. Results

For all *S. haematobium* analyses, data were examined both including and excluding locus D3, which was difficult to score in some of the studies. Similarly for *S. mansoni*, locus smu31768 was excluded in one set of analysis. Since the conclusions of these analyses were not divergent, data are presented using all 7 loci for each species

3.1. Between schools characterisation and comparison of S. mansoni and S. haematobium populations

3.1.1. Diversity of infections

As shown in Table 2, a high level of diversity was documented in African populations of both *S. haematobium* and *S. mansoni* with individual microsatellite loci detecting a mean of 36.1 alleles per locus (range 22–64). The mean number of alleles per loci was higher in *S. mansoni* (44.7) loci than *S. haematobium* (27.4) loci, but this might be expected since, as shown in Table 3, more *S. mansoni* were genotyped in these studies, and may be a characteristic of the set of microsatellite loci used. However, mean major allele frequency and expected heterozygosity were similar between the two species. All loci showed some evidence of heterozygote deficiency across the entire continental-scale datasets.

A total of 2,655 *S. haematobium* miracidia were analysed, (from a total of 2737 collected, and where samples with less than 4 loci amplifying were excluded from analyses), with a mean of 17.4 miracidia per child and of 266 miracidia per school. As shown in Table 3, the mean number of alleles per locus in each school was high, ranging from 3.88 to 18.49 and was strongly positively correlated with sample size (78% of variation in mean number of alleles could be explained by a linear relationship with the number of miracidia sampled). However, allelic richness, a measure of diversity independent of sample size, was remarkably similar in the ten schools where *S. haematobium* was sampled, despite widely varying prevalences and intensities of infection. Within children, the mean number of alleles ranged from 3.86 to 5.71 and both mean allelic richness and major allele frequency were similar in each of the schools. A permutation test comparing the mean allelic richness (per school) between East Africa (Kenya and Tanzania) and West Africa (Mali and Niger) was not significant ($P = 0.58$). Similarly high levels of expected heterozygosity were seen in all schools and there was no apparent consistent difference between them based on their geographical location ($P = 0.11$) or infection prevalence and intensity (Fig. 2 and Table 1). However, observed heterozygosity did differ, being lower in Kenya and Tanzania, intermediate in Cameroon and higher in West Africa and a permutation test confirmed that this was a significant difference ($P = 0.008$) (Fig. 2a). Thus higher levels of heterozygote deficiency were seen in East Africa than in West Africa. Estimates of allelic richness and major allele frequency and heterozygosity were similar when all data from a single school were pooled (Component population approach), and when diversities were calculated separately for the parasite populations of individual children (Infrapopulation approach) (Fig. 2 and Table 3).

A total of 3563 *S. mansoni* miracidia were analysed (from a total of 3743 collected, and where samples with less than 3 loci amplifying were excluded from analyses), with a mean of 20.5 miracidia per child and of 445 miracidia per school. As shown in Table 3, the mean number of alleles per locus in each school was high, ranging from 9.7 to 40.6 and was again strongly positively correlated with sample size. In common, with the *S. haematobium* dataset, allelic richness was much more similar between villages despite wide variation in prevalence and infection intensity (Tables 1 and 3), although the highest diversity was seen at Musubi Church of God School, Uganda, which had a particularly high mean intensity of infection. In general, and in contrast to *S. haematobium*, diversity was slightly higher in East than West Africa, and a permutation test comparing the mean allelic richness (per school) between East Africa (Kenya, Uganda and Tanzania) and West Africa (Mali and Niger) confirmed this was a significant difference ($P = 0.0002$). Similarly higher levels of expected heterozygosity and observed heterozygosity were seen in East than West Africa (He $P = 0.0002$; Ho $P = 0.001$) although there was no apparent consistent difference between them based on their infection prevalence or intensity (Fig. 3, Tables 1 and 3). Observed heterozygosity was lower than expected heterozygosity in all schools but Namarigoungou (Niger).

3.1.2. Population structure

Cavalli-Sforza and Edwards' chord distances between the *S. haematobium* populations of children at 10 schools are shown in Fig. 4a. Strong geographical clustering was evident, with East African (Kenyan and Tanzanian) populations clustering together, West African (Mali and Nigerian) samples in a separate cluster, with Cameroonian samples between them. Clustering of the parasite populations of individual children (Fig. 4b) demonstrated complete separation at the country level, but the populations of children from different schools within the same country were overlapping. Analysis of Molecular Variance (Table 4a) confirmed that approximately 16% of the total variation in the dataset was between the schools, and that this was largely due to differences between the five countries ($Fst = 0.173$, $P < 0.001$), with significant but only very low levels of differentiation ($Fst = 0.008$, $P < 0.001$) between schools within countries. There was significant differentiation between the parasite populations of different children within schools ($Fst = 0.017$, $P < 0.001$) but this only accounted for 1.4% of the

Table 2
Details of microsatellite loci used for analysis of 2,655 *S. haematobium* miracidia and 3,563 *S. mansoni* miracidia from untreated regions of six SSA countries, including locus name, Genbank accession number, repeat motif, allele range, number of alleles, frequency of most common allele, and expected (He) and observed (Ho) heterozygosity.

Species	Locus	GenBank accession no.	Repeat motif	Observed allele range	Major allele frequency	No alleles	He	Ho
S. haematobium	C102	HM856648	ATT	168–206	0.75	25	0.42	0.35
S. haematobium	C11	HM856644	TAA	84–126	0.67	33	0.51	0.30
S. haematobium	C111	HM856649	ATT	108–234	0.38	26	0.71	0.68
S. haematobium	C116	HM856652	TAA	263–425	0.15	26	0.91	0.66
S. haematobium	C131	EF608045	ATT	240–296	0.60	22	0.55	0.05
S. haematobium	C146	HM856657	TAA	138–186	0.27	35	0.84	0.72
S. haematobium	D3	EF608047	TAGA,TTGA	217–226	0.38	25	0.79	0.55
S. mansoni	SMDA28	AF325695	GATA	80–151	0.21	53	0.88	0.67
S. mansoni	SMD25	AF202965	CA	191–393	0.28	47	0.84	0.59
S. mansoni	SMD28	AF202966	CAA	140–250	0.62	31	0.58	0.24
S. mansoni	SMD89	AF202968	TC	135–286	0.58	33	0.61	0.33
S. mansoni	SMU31768	U31768	GAT	175–253	0.37	64	0.77	0.73
S. mansoni	CA11-1	AI068336	GA, GT	187–235	0.34	44	0.84	0.50
S. mansoni	SMS9-1	AF330106	GT	170–212	0.54	41	0.67	0.38

Table 3
Sample size, diversity values and estimates of effective population sizes for *S. haematobium* populations from 10 schools, and for *S. mansoni* populations from 8 SSA schools.

Schistosoma haematobium

Country	Place	Mean no miracidia/child	Total miracidia (*n*)	Mean infection intensity (samples)	Mean (±95% c.i.) across loci of parasite populations of individual children			Mean across loci of parasite populations of villages				
					Mean no. of alleles	Major allele frequency	Allelic richness (Ar)	Mean no. of alleles	Major allele frequency	Allelic richness (Ar)	Ne (±95% c.i.) LDNe	Ne[a] (95% c.i.) MNe
Kenya	Abori	19.7	296	223.21	5.06 ± 0.31	0.58 ± 0.01	3.38 ± 1.48	12.00	0.52	4.14	340.8 (188.1–1114.9)	233.0–32000+
Kenya	Kiwalwa	16.0	224	587.43	4.95 ± 0.41	0.57 ± 0.01	–	13.14	0.52	4.24	391.9 (188.6–6055.4)	155.0–32000+
Tanzania	Kilambo	13.8	55	Unknown	4.00 ± 0.88	0.61 ± 0.05	3.26 ± 1.52	6.14	0.58	3.51	50.4 (26.0–166.8)	30.2–32000+
Tanzania	Nanhyanga	10.0	10	Unknown	3.86	0.54	3.29 ± 1.39	3.88	0.53	3.29	37.4 (2.1–infinity)	–
Cameroon	Bessoum	17.2	568	Unknown	4.95 ± 0.55	0.56 ± 0.01	3.27 ± 1.42	18.49	0.51	4.18	465.2 (312.4–800.8)	482.0–32000+
Niger	Falmado	26.6	346	65.21	5.71 ± 0.34	0.51 ± 0.01	4.24 ± 1.53	10.48	0.49	3.86	333.5 (222.7–576.4)	237.0–32000+
Niger	Libore	20.9	230	110.3	5.51 ± 0.27	0.50 ± 0.02	3.96 ± 1.39	8.86	0.48	3.96	400.1 (216.4–1406.3)	117.0–32000+
Niger	Namarigoungou	14.2	85	0.43	4.50 ± 0.64	0.53 ± 0.02	4.59 ± 1.66	7.29	0.48	4.15	79.1 (52.4–137.1)	1.0–32000+
Mali	Bambougou-Wéré	12.5	276	218.4	4.07 ± 0.55	0.62 ± 0.03	3.45 ± 1.18	11.29	0.58	3.75	147.2 (107.6–214.6)	186.7–32000+
Mali	Dougouba	22.6	565	97.0	5.43 ± 0.28	0.58 ± 0.02	3.48 ± 1.18	13.71	0.56	3.87	85.3 (70.9–103.0)	409.6–32000+

Schistosoma mansoni

Country	Place	Mean no miracidia/child	Total miracidia (*n*)	Mean infection intensity (samples)	Mean (±95% c.i.) across loci of parasite populations of individual children			Mean across loci of parasite populations of villages				
					Mean no. of alleles	Major allele frequency	Allelic richness (Ar)	Mean no. of alleles	Major allele frequency	Allelic richness (Ar)	Ne (95% c.i.) LDNe	Ne (95% c.i.) MNe
Kenya	Abori	14.7	99	48.6	5.5 ± 0.41	0.51 ± 0.04	2.58 ± 0.09	12.00	0.47	7.14	47.3 (38.5–59.0)	108.0–32000+
Kenya	Kiwalwa	14.7	77	6.0	5.57 ± 1.21	0.47 ± 0.04	2.60 ± 0.10	10.52	0.43	7.10	35.5 (28.0–46.2)	59.9 –32000+
Tanzania	Bukindo	13.7	521	94.4	5.31 ± 0.53	0.46 ± 0.03	2.66 ± 0.05	14.71	0.39	6.46	213.7 (172.0–270.6)	462.0–32000+
Tanzania	Kisorya	12.1	509	502.7	5.22 ± 0.33	0.44 ± 0.02	2.66 ± 0.04	14.43	0.38	6.51	218.2 (169.7–289.5)	387.0–32000+
Uganda	Musubi	33.4	167	1650.5	10.34 ± 1.16	0.36 ± 0.03	2.91 ± 0.16	19.57	0.34	9.49	157.9 (123.1–212.7)	143.0–32000+
Niger	Diambala	23.1	33	208.0	3.77 ± 0.99	0.51 ± 0.10	2.58 ± 0.48	9.71	0.43	8.58	6.7 (3.6–10.3)	30.0–32000+
Niger	Namarigoungou	23.4	594	285.5	5.45 ± 0.22	0.50 ± 0.01	2.40 ± 0.03	15.86	0.49	5.40	163.6 (132.0–205.1)	554.0–32000+
Mali	Kokry-bozo	29.5	1563	511.6	6.61 ± 1.32	0.56 ± 0.03	2.56 ± 0.08	40.57	0.50	8.07	64.4 (58.4–70.8)	1791.0–32000+

[a] All point estimates of Ne were very close to maximum of 32,000.

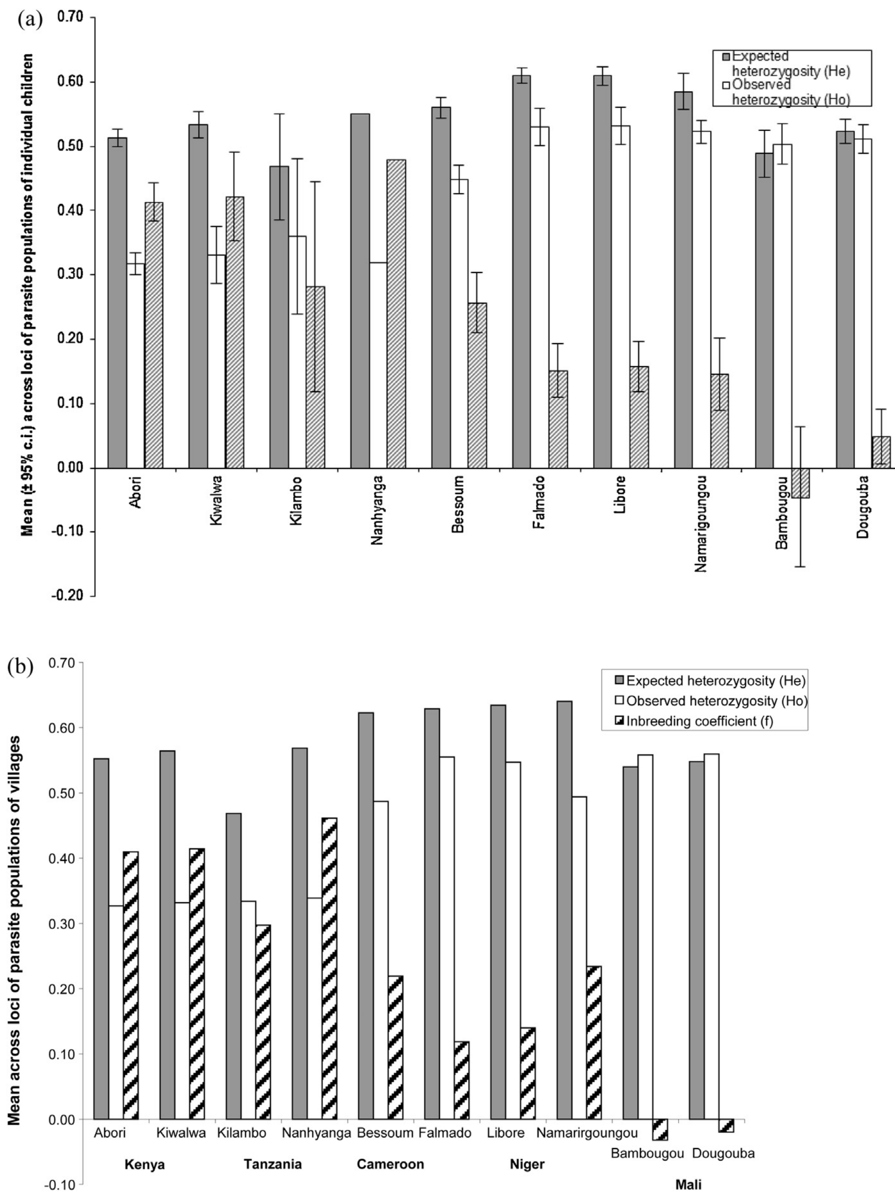


Fig. 2. Mean (across loci) expected (He), observed heterozygosity (Ho) and allelic richness representing (a) mean ± 95% of the 1–33 *S. haematobium* larval populations of individual children from and (b) combined *S. haematobium* larval population from 10 schools in five SSA countries.

total variation. Most of the variation (82.5%) was within the parasite populations of individual children. Geographic distances between schools varied between 0.2 km and 5,720 km and there was an overall significant positive correlation between the matrix of geographic and genetic (Fst/1-Fst) distances ($r = 0.47$, $P = 0.001$). The correlation was also positive within West Africa ($r = 0.94$, $P = 0.009$). Distances between schools ranged from 7 km to 979 km in West Africa, and from 122 km to 260 km in Niger. The correlation was not investigated in East Africa since the two Kenyan schools were located within the same town, and sample sizes in Tanzania were small.

Cavalli-Sforza and Edwards chord distances between the *S. mansoni* populations of children at eight schools are shown in Fig. 5a. In common with the *S. haematobium* dataset, there was evidence of strong geographical clustering with populations clustering together based on their geographic origins. Clustering of the parasite populations of individual children (Fig. 5b) showed the populations of children from different schools within the

Table 4a
AMOVA results for *Schistosoma haematobium* populations from five SSA countries.

Source of variation	Sum of squares	Variance components	Percentage of variation	Fst	*P* value
Between countries	1455.2	0.354	17.3	0.173	<0.001
Among schools within countries	34.6	0.013	0.62	0.008	<0.001
Within schools	8905.1	1.680	82.1	0.179	<0.001
Between schools	1498.8	0.324	16.1	0.161	<0.001
Among children within schools	360.3	0.028	1.4	0.017	<0.001
Between miracidia within children	8544.8	1.654	82.5	0.175	<0.001

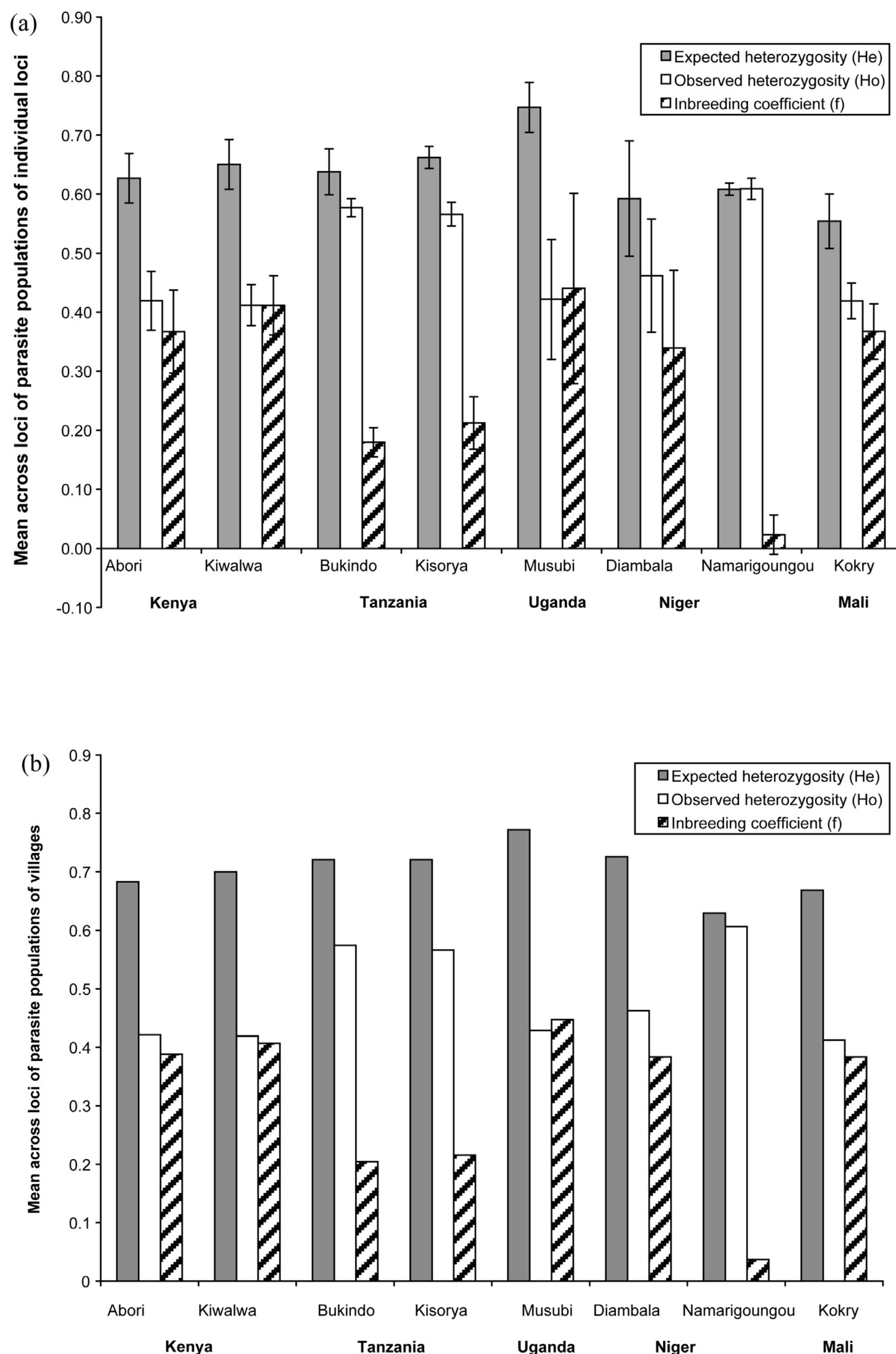


Fig. 3. Mean (across loci) expected (He), observed heterozygosity (Ho) and allelic richness representing (a) mean ± 95% of the 5–53 *S. mansoni* larval populations of individual children from and (b) combined *S. mansoni* larval population from eight schools in five SSA countries.

Table 4b
AMOVA results for *Schistosoma mansoni* populations from five SSA countries.

Source of variation	Sum of squares	Variance components	Percentage of variation	Fst	*P* value
Between countries	1156.4	0.238	10.8	0.108	0.006
Among schools within countries	4.0	−0.001	−0.07	−0.0007	0.972
Within schools	13733.0	1.950	89.2	0.108	<0.001
Between schools	1160.3	0.220	10.1	0.100	<0.001
Among children within schools	449.9	0.025	1.2	0.015	<0.001
Between miracidia within children	13283.4	1.936	88.7	0.113	<0.001

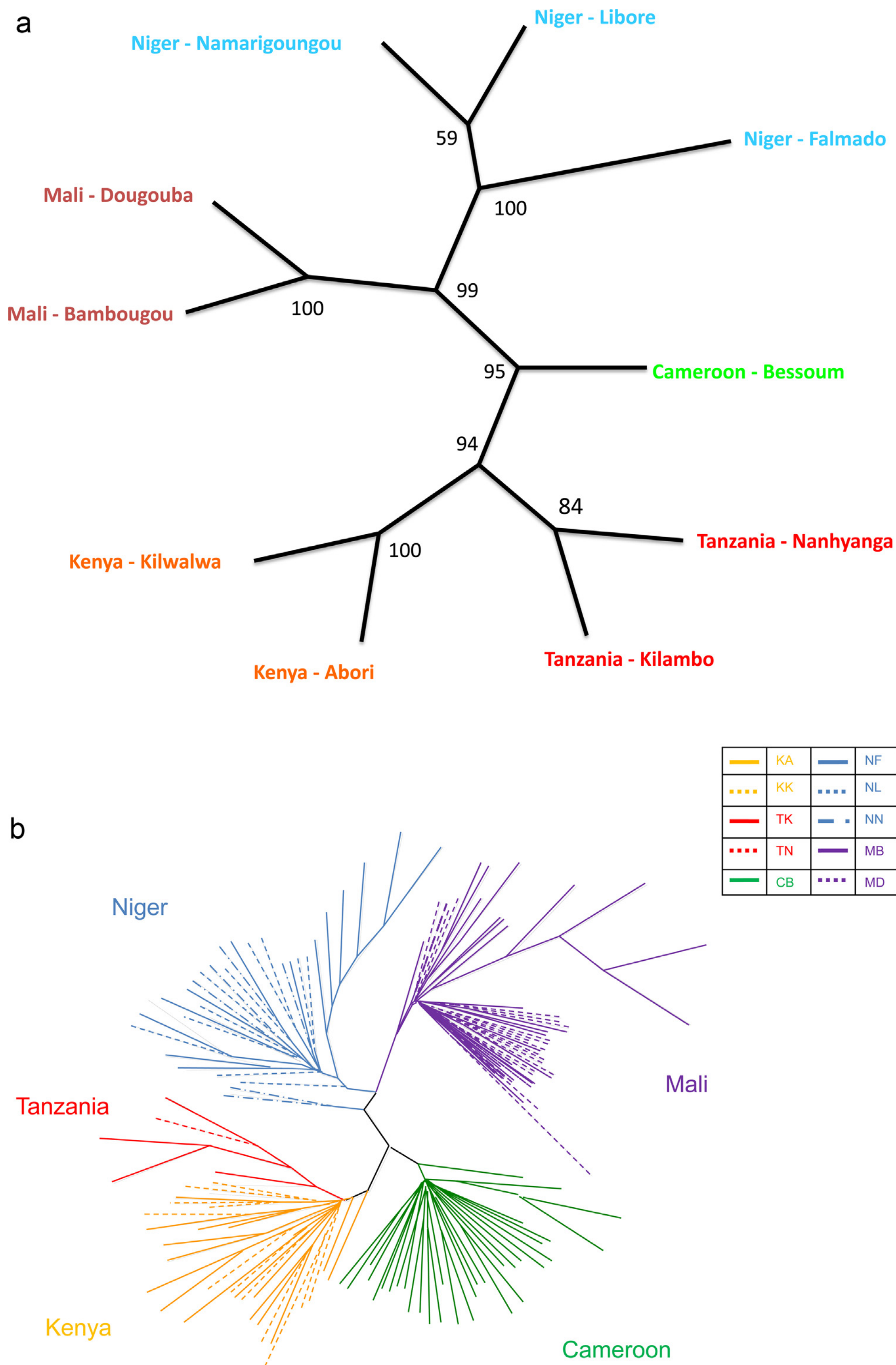


Fig. 4. (a) Neighbour-Joining clustering of Cavelli-Chord distances between pooled *S. haematobium* larval populations from 10 schools in five SSA countries. Bootstrap values from 100 repetitions are indicated. (b) Neighbour-Joining clustering of Cavelli-Chord distances between the individual *S. haematobium* larval populations of 144 children from 10 schools (KA, Abori; KK, Kiwalwa; TN, Nanhyanga; TK, Kilambo; CB, Bessoum; NF, Falmado; NL, Libore; NN, Namarigoungou; MB, Bambougou; MD, Dougouba) in five SSA countries.

same country were overlapping. In contrast to the *S. haematobium* dataset, there was substantial, but not complete separation, at the country level, with some overlap between villages in West Africa. Analysis of Molecular Variance (Table 4b) confirmed that approximately 10% of the total variation in the dataset was between the schools, and that this was largely due to differences between the five countries (Fst = 0.108, P = 0.006), with no significant differentiation (Fst = −0.0007, P = 0.972) between schools within countries. There was significant differentiation between the parasite populations of different children within schools (Fst = 0.053, P < 0.001) but this only accounted for 1.2% of the total variation. Most of the variation (88.7%) was within the parasite populations of individual

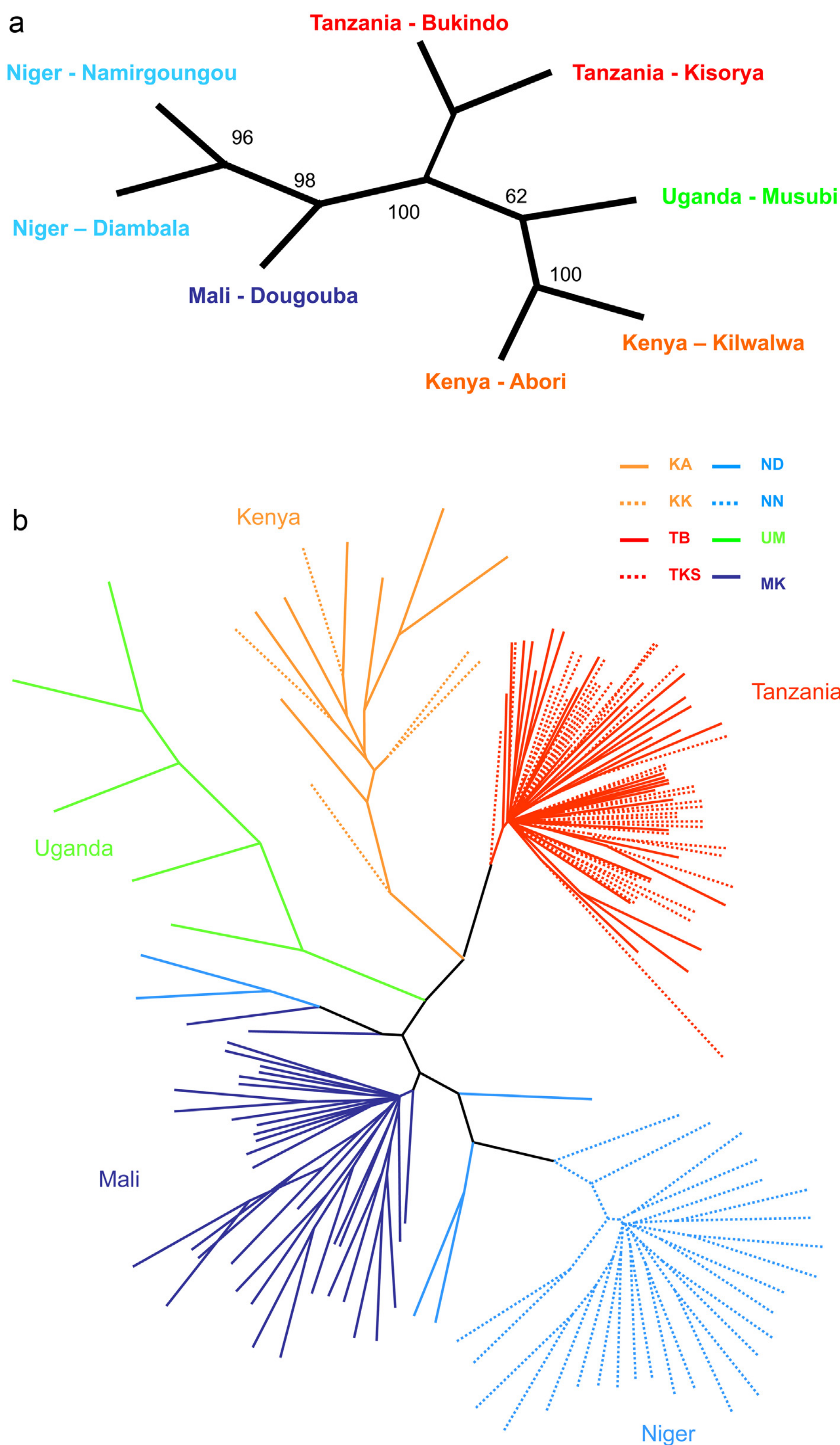


Fig. 5. (a) Neighbour-Joining clustering of Cavelli-Chord distances between pooled *S. mansoni* larval populations from eight schools in five SSA countries. Bootstrap values from 100 repetitions are indicated. (b) Neighbour-Joining clustering of Cavelli-Chord distances between individual *S. mansoni* larval populations of 183 children from eight schools (KA, Abori; KK, Kiwalwa; TB, Bukindo; TK, Kisorya; UM, Musubi; ND, Diambala; NN, Namarigoungou; MK, Kokry) in five SSA countrie.

children. Geographic distances between schools varied between 0.2 km and 5,138 km and there was an overall significant positive correlation between the matrix of geographic and genetic (Fst/1-Fst) distances ($r = 0.29$, $P = 0.03$). The correlation was also significantly positive within East Africa ($r = 0.64$, $P = 0.004$) but not within West Africa ($r = 0.43$, $P = 1.0$). Distances between schools ranged from 0.2 km to 612 km in East Africa, and in West Africa consisted of two neighbouring villages (1.1 km apart) in Niger and a village in Mali, at approximately 720 km distant. Fst values between villages in West Africa were very small.

3.1.3. Estimation of effective number of breeders (Nb)

Estimated effective number of breeders (*Nb*) for both *S. mansoni* and *S. haematobium* are shown in Table 3. There were orders of magnitude in difference in the estimated *Nb* for the same villages with the two methodologies used, with point estimates of *Nb* being near to the maximum set by the software (32,000) in all villages using the maximum likelihood analysis of allele frequencies implemented by MLNe (Wang, 2001), yet in the range of 37–466 using the linkage disequilibrium approach implemented by LDNe (Waples and Do, 2008). Moreover, there was little correlation between the estimated *Nb* for the same villages between the two methodologies (rs = −0.05), and between the estimated *Nb* and estimated genetic diversity of the populations, as measured by mean allelic richness (rs = 0.25 MLNe, rs = −0.34 LDNe). Lower 95% confidence limits of estimates using MLNE were strongly influenced by miracidial sample size ($F_{1,8}$ = 185.6, P = 0.0001) and were not related to schistosome species, geography nor the infection intensity, prevalence or mean age of the children sampled.

In contrast, estimates of *Nb* calculated using LDNe were not affected by miracidial sample size ($F_{1,8}$ = 0.06, P = 0.81). In common with estimates of *Nb* calculated in MLNe, they did not differ between the schistosome species ($F_{1,8}$ = 0.77, P = 0.41), between schools in East and West Africa ($F_{1,8}$ = 2.14, P = 0.18), between areas located near to lakes or rivers ($F_{1,8}$ = 0.55, P = 0.48) nor vary with the prevalence or intensity of infection ($F_{1,8}$ = 0.89, P = 0.37 or $F_{1,8}$ = 0.00, P = 0.97).

3.2. Additional individual host level analyses of S. haematobium populations from Bessoum School, Cameroon

3.2.1. Diversity of infections

There were populations which showed both an excess and a deficit of heterozygotes and considerable variation in the within population variability (FIS) between populations, but no patterns with child age (He: $F_{1,28}$ = 1.30, P = 0.24; Ho: $F_{1,28}$ = 1.14, P = 0.30; FIS: $F_{1,28}$ = 0.38, P = 0.54) or gender (He: $F_{1,28}$ = 0.24, P = 0.63; Ho: $F_{1,28}$ = 0.96, P = 0.34; FIS: $F_{1,27}$ = 0.03, P = 0.86). The allelic richness of the 33 *S. haematobium* populations did not differ between male and female children ($F_{1,28}$ = 3.53, P = 0.07) or vary with child age ($F_{1,28}$ = 0.95, P = 0.34).

3.2.2. Population structure

There was little genetic variation and limited bootstrap support for differentiation of the parasite populations of individual children, and no evidence of clustering based on age and gender (although the number of females sampled here was small) of the children surveyed. This was confirmed by a mean FST between populations of 0.020 (P = 1.0) and AMOVA results which showed that only 1.7% of the total variation was between the parasite populations of different children, with 98.3% of the variation being within children.

4. Discussion

In areas where schistosomiasis is highly endemic, the present goal is to mitigate the burden of the disease by controlling morbidity through preventive chemotherapy using PZQ and national control programmes have been established in several sub-Saharan African countries (Fenwick and Webster, 2006; Fenwick et al., 2009). This recent shift towards large-scale administration of PZQ may lead to intensive and prolonged new selection pressures on the parasite, including, but not exclusive to, that for drug resistant parasites (Doenhoff et al., 2009; Fallon and Doenhoff, 1994; Ismail et al., 1999). Parasites are not static, and our own previous laboratory studies show that selection can rapidly change schistosomes' infectivity and virulence phenotypes as well as population genetic structure within only a few generations (Davies et al., 2001; Webster et al., 2004, 2007; Gower and Webster, 2004). Yet, remarkably little is known about potential natural variations in human schistosome populations over space and time, nor how such variations may influence transmission dynamics, clinical impact or evolutionary responses to intervention-induced selection.

Population structure, the genetic subdivision of parasite populations, is vitally important to the potential emergence and spread of PZQ resistance. High levels of population structure between hosts, due perhaps to clumped transmission of related genotypes, may increase the likelihood of recessive alleles (which commonly confer drug resistance) coming together in the same host thus increasing the chance of a doubly recessive individual developing, while restricted gene flow between individuals and regions is likely to limit the spread of any such resistance should it develop (Gower et al., 2011). Patterns of population structure might be expected to vary between different regions and epidemiological settings (Steinauer et al., 2010). Our study was relatively consistent in both *S. mansoni* and, as one of the first major reports of geographic variation in the highly neglected *S. haematobium* (Doenhoff et al., 2009; Rollinson, 2009), in identifying that the majority of variation existed within the parasite populations of individual children. At this pre-PZQ 'baseline' stage, there was a small amount of genetic differentiation between the parasite populations of different children. Since we have attempted to correct for the potential relatedness of miracidia within individual hosts (who may be siblings) by using a hierachical FST analysis, this could additionally reflect genetic relatedness of parental worms, which might be expected given clumped transmission of clonally related adults at a single infection site. The very small Fst values, however, suggest that inbreeding is low, and that children are being exposed to genetically different infections, over space and time. Similarly, there was little evidence from Fst or clustering analyses of genetic divergence between the parasite populations of individual children of the same school, or between schools in the same country. This was further supported by little evidence of differentiation between the parasite populations of children of different ages and gender in more detailed analyses in Cameroon. However, analysis of molecular variance and clustering analyses demonstrated that there was strong isolation by distance and geographical divergence between the six SSA countries from which samples were collected. This is despite geographic distances (e.g. for *S. haematobium* Nanhyanga to Kisorya; Fig. 1) within countries sometimes exceeding those between them (e.g. Nanhyanga to Abori; Fig. 1), although in some instances (e.g. Kenya *S. haematobium* and *S. mansoni* and Tanzania *S. mansoni* collections), sampling sites within countries were geographically close to one another and on the same water course. Thus this study documented genetic population structure in both *S. haematobium* and *S. mansoni* populations, on both regional and continental scales, which might be expected to slow the spread of drug resistance should it develop, as well as limited evidence of structure at the individual host level that might promote its development. Interestingly, Steinauer et al. (2009) found limited genetic structuring in *S. mansoni* populations around a single watershed in Kenya, including in the Kenyan portion of Lake Victoria which stretches to 1800 km^2, in contrast to strong genetic structure detected in a previous study by the same authors over a wider geographic scale within Kenya encompassing different watersheds in the east, west and southwest portions of the country (Agola et al., 2006). One should acknowledge, however, these differences could also be influenced by variations in the sampling methodology used, since the latter study used animal-passaged schistosome samples (Steinauer et al., 2010). Our study documented similar and significant levels of genetic differentiation between the *S. mansoni* populations collected from Ugandan and Tanzanian schools on Lake Victoria, as between those collected from Uganda or Tanzania and schools on a different watershed in Kenya, with all

pairwise comparisons having very similar moderate/high Fst values. Thus, though there was evidence of genetic mixing occurring between populations on Lake Victoria within Kenya, larger scale mixing of parasite populations across the Lake does not appear to be occurring. This is not perhaps surprising given the large distances involved. Moreover, schistosomes live for, on average, 3–5 years in their human hosts and previous studies of schistosome diversity (Davies et al., 1999) have implicated geographical distance rather than closeness on water courses, and human rather than intermediate hosts, as the driving forces in schistosome dispersal and gene flow. Our results indicate that human movement within a country may be greater than across borders, but to geographically closer areas, with political boundaries having a stronger influence than absolute distance. In this study, the significant genetic variation between the *S. mansoni* populations collected in the Ugandan and Tanzanian areas of Lake Victoria would tend to support this. It would be interesting, and highly relevant to national control programmes, to know if national boundaries, due to their effect on the dispersal of people, act to limit schistosome gene flow in addition to the effects of geography, and warrants further studies in border regions. Our data also indicate that gene flow may be higher in *S. mansoni* than in *S. haematobium*, particularly in West Africa. Our study sites for *S. mansoni* in Niger were two neighbouring villages located in the Tillaberi department, north of the capital, Niamey. The Tillaberi region has for a long time been known as a *S. haematobium* focus although in 2001 an extensive malacological survey was conducted and *Biomphalaria pfeifferi*, a major intermediate host species for *S. mansoni*, was identified in five villages in this region (Labbo et al., 2003). Consequently a parasitological survey was conducted in 2002/2003 and *S. mansoni* was found in Namarigoungou (Garba et al., 2004). *S. mansoni* has now spread to Diambala and other villages in this area and appears to have overtaken *S. haematobium* as the dominant schistosome species in the area (Garba et al., 2004, 2011). Indeed our genetic data implicating higher gene flow across West African populations and/or reduced opportunities for local differentiation, supports this recent expansion in this area. Thus conclusions regarding gene flow and the implications for spread of drug resistance vary between the two schistosome species and in different regions.

Levels of genetic diversity, as opposed to genetic divergence between populations, may itself be an important component of the epidemiology of infection and disease, and a key indicator for monitoring the effects of selection imposed by drug treatment (Norton et al., 2010). In these studies, microsatellite analyses revealed high levels of genetic variability in both *S. mansoni* and *S. haematobium*. We also report differences between the schistosome species, in that *S. mansoni*, but not *S. haematobium*, were more variable in East than West Africa. Our *S. mansoni* results thus agree with a comprehensive intraspecific phylogeographical study using mitochondrial sequences which revealed that there are several divergent clades of *S. mansoni* across Africa, with the highest number of clades (3 of 5) located in East Africa (Morgan et al., 2005) suggesting an East African origin for *S. mansoni*, but comparable data are not currently available for *S. haematobium.* There was a marked deficiency of heterozygotes in our *S. haematobium* populations of East Africa, which was not present in the West African samples, nor in *S. mansoni* populations, indicating higher levels of inbreeding, the Wahlund effect, non-random mating and/or reduced gene flow in these populations. A further potential explanation for heterozygote deficiency is, of course, sampling bias – where such differences could occur if, for example, there was an increased frequency of longer alleles in these regions, although there is no *a priori* reason to expect such a bias in the current dataset. There are, nevertheless, multiple differences between the sampling sites here in addition to their geography. For instance, all East African *S. mansoni* samples were collected from Lake regions (Lake Victoria for Ugandan and Tanzanian samples, Lake Jipe for Kenyan samples) and West African samples were all collected along the Niger River, whereas in *S. haematobium*, both lake and rivers were sampled in East Africa. Likewise there were potential differences in snail intermediate host use and differences in infection prevalence and intensity between these study sites, and hence it is not possible to determine the exact causes of the reported differences at this stage.

A methodological difficulty in describing genetic diversity is also that of variable sample size, since larger sample sizes will tend to harbour more alleles. Many population genetic studies in other organisms address this by regulating the size of samples collected, but in schistosome collections the sample size of the parasite population of individual hosts is partly determined by infection intensity and is often highly variable between hosts. One sampling solution is to discard small samples, but this will tend to discriminate against low infection intensity infections, which may be of particular epidemiological interest, and would be particularly common following large-scale drug administration. Alternative solutions are to include sample size as a covariate (Norton et al., 2010) or, perhaps, most satisfactorily to rarefy to the smallest sample present (El Mousadik and Petit, 1996). Allelic richness is a measure of diversity which is rarefied in this way, but is not widely implemented in statistical packages, and care should be taken in its use since its value will vary depending on the minimum sample size with which it is being compared (within the same analysis). It may not be directly comparable between different species, or studies, depending on the characteristics of the microsatellite loci used, but is suitable, and highly informative, within studies or for meta-analyses such as described here. We also reported that conclusions regarding comparative genetic diversity were similar whether data were pooled at the village level (Component population approach), or when data were analysed separately at the level of the parasite populations of individual children (Infrapopulation approach). The latter is more satisfactory from an analytical point of view, since it not only allows comparison with epidemiological data such as host age, gender and infection intensity, but also effectively controls for potential relatedness of sampled miracidia, and is thus the methodology used within the reported studies.

Estimates of *Ne* or *Nb*, key parameters in population genetics have been poorly studied by parasitologists (Criscione et al., 2005), despite the information it may provide on the ability of parasite populations to respond to selection and, potentially a measure of the impact of treatment at the community level and on *refugia* (Webster et al., 2008) as well being as an important parameter for the success of local elimintion. The most commonly used methods for calculating *Nb*/*Ne* are variations on the temporal method (Waples and Yokota, 2007) whereby two samples are taken from the same population at two different time points. However, this methodology is of limited use in the monitoring of helminth control programmes, where temporal changes in *Nb*/*Ne* as a result of large-scale drug administration are the major interest and resampling untreated control groups is not possible due to ethical considerations. Published methodology and statistical programmes exist for calculating *Ne* from a sample taken at a single time point based on allele frequencies (MLNE; Wang and Whitlock, 2003) or linkage disequilibrium (LDNe; Waples and Do, 2008), which were investigated in this study, with the aim of assessing whether differences in *Nb* could be identified between schistosome species and geographical regions or epidemiological settings. The major conclusion of this meta-analysis was that the two methods used, which are themselves based on differing assumptions, provided widely differing estimates of *Nb*, with the methodology of Wang and Whitlock (2003) being strongly influenced by miracidial sample size. Estimates of *Nb* with LDNe seem low for entire schools, although there are many features of schistosome life cycles that will tend to reduce *Nb* below the population census size (Prugnolle et al., 2005).

Interestingly, there was a non-significant trend for *Nb* to be lower in schools where the mean age of children sampled was higher, which is consistent with the development of acquired immunity in older children. However, there were no obvious differences in *Nb* (or indeed between measures of larval diversity) between schools with widely differing prevalences and intensities of infections, which, if accurate, might suggest that traditional measures of community infection (prevalence among school-aged children is the indicator used by WHO and National Control Programmes to classify whole community risk) may not be good indicators of the size of the breeding population of parasites, and hence their susceptibility to the selective pressures imposed by large-scale drug administration. Prugnolle et al. (2005) conducted theoretical work attempting to account for the clonality and redundancy in the schistosome life-cycle in estimating population genetic parameters such as *Nb*. Further theoretical work (and implementation in statistical software) focussing on sampling larval schistosomes, and the impact of structured population, would certainly be a welcome addition to the field.

5. Conclusion

Recent methodological developments in schistosome population genetic studies have allowed investigation of the parasite populations of large numbers of infected individuals across the African continent. We report novel data on genetic variability in the two major African schistosome species, *S. mansoni* and *S. haematobium*, including the largest datasets to date on the highly neglected latter species. We report high levels of genetic diversity in both species, and population structure between countries in both schistosome species prior to the introduction of national control programmes for schistosomiasis in six SSA countries, on both regional and continental scales. Both the lack of genetic population structuring observed at the individual host level, and the presence of population structuring on regional levels for both species, would be expected to slow or discourage both the development, and subsequent spread, of drug resistance. Further samples are necessary to determine the extent of geographic differentiation within countries, and whether patterns of differentiation may vary between the two species and geographical regions. These data will thereby help serve as baseline upon which to measure the impact of large-scale PZQ administration on schistosome genetics in these regions.

5.1. Ethical considerations

Ethical approval was obtained from Imperial College Research Ethics Committee (ICREC), Imperial College London in the UK, in combination with the ongoing Schistosomiasis Control Initiative (SCI) activities. In Kenya, ethical approval was obtained from the Ethical Review Board of National Museums of Kenya/Kenya Medical Research Institute. In Uganda, the Ethical Review Board of the Ugandan National Council of Science and Technology approved the studies. In Tanzania, ethical approval was obtained from the Ethical Review Board of National Institute of Medical Research (NIMR). In Cameroon, ethical approval was obtained from the Commité National d'Ethiqué, Cameroon. In Niger, ethical clearance was obtained from the Niger National Ethical Committee and from NHS-LREC review board of St Mary's Hospital, Imperial College London. In Mali, official ethical clearance was not required because all aspects of monitoring and evaluation were carried out in the framework of the disease control activities implemented and approved by the Ministry of Health (MoH). Before conducting the study, the MoH-approved plan of action had been presented and adopted by regional and local administrative and health authorities.

In all studies, written informed consent for the schoolchildren to participate in longitudinal monitoring of the national control programmes for schistosomiasis were given by either parents, head teachers and/or village chiefs where there were no schools. The children and/or their parents/guardians verbal consent was recorded at school committees comprising of parents, teachers and community leaders after they received information regarding the study.

Acknowledgements

We are very grateful to the children and teachers of the 18 schools involved, to the technical teams of all national schistosomiasis control programme and research groups across all countries involved. Financial support was provided by the European Union (CONTRAST/EUINCO.Dev contract no: 032203), The Royal Society (CMG and JPW) and the Bill & Melinda Gates Foundation.

References

Agola, L.E., Mburu, D.N., DeJong, R.J., Mungai, B.N., Muluvi, G.M., Njagi, E.N.M., Loker, E.S., Mkoji, G.M., 2006. Microsatellite typing reveals strong genetic structure of *Schistosoma mansoni* from localities in Kenya. Infect. Genet. Evol. 6, 484–490.

Agola, L.E., Steinauer, M.L., Mburu, D.N., Mungai, B.N., Mwangi, I.N., Magoma, G.N., Loker, E.S., Mkoji, G.M., 2009. Genetic diversity and population structure of *Schistosoma mansoni* within human infrapopulations in Mwea, central Kenya assessed by microsatellite markers. Acta Trop. 111, 219–225.

Berriman, M., Haas, B.J., LoVerde, P.T., Wilson, R.A., Dillon, G.P., Cerqueira, G.C., Mashiyama, S.T., Al-Lazikani, B., Andrade, L.F., Ashton, P.D., Aslett, M.A., Bartholomeu, D.C., Blandin, G., Caffrey, C.R., Coghlan, A., Coulson, R., Day, T.A., Delcher, A., DeMarco, R., Djikeng, A., Eyre, T., Gamble, J.A., Ghedin, E., Gu, Y., Hertz-Fowler, C., Hirai, H., Hirai, Y., Houston, R., Ivens, A., Johnston, D.A., Lacerda, D., Macedo, C.D., McVeigh, P., Ning, Z., Oliveira, G., Overington, J.P., Parkhill, J., Pertea, M., Pierce, R.J., Protasio, A.V., Quail, M.A., Rajandream, M.A., Rogers, J., Sajid, M., Salzberg, S.L., Stanke, M., Tivey, A.R., White, O., Williams, D.L., Wortman, J., Wu, W., Zamanian, M., Zerlotini, A., Fraser-Liggett, C.M., Barrell, B.G., El-Sayed, N.M., 2009. The genome of the blood fluke *Schistosoma mansoni*. Nature 460, 352–358.

Blair, L., Webster, J.P., Barker, G.C., 2001. Isolation and characterisation of polymorphic microsatellite markers in *Schistosoma mansoni* from Africa. Mol. Ecol. Notes 1, 93–95.

Cavalli-Sforza, L.L., Edwards, A.W.F., 1967. Phylogenetic analysis: models and estimation procedures. Am. J. Hum. Genet. 19, 233–257.

Criscione, C.D., Poulin, R., Blouin, M.S., 2005. Molecular ecology of parasites: elucidating ecological and microevolutionary processes. Mol. Ecol. 14, 2247–2257.

Curtis, J., Sorensen, R.E., Page, L.K., Minchella, D.J., 2001. Microsatellite loci in the human blood fluke *Schistosoma mansoni* and their utility for other schistosome species. Mol. Ecol. Notes 1, 143–145.

Davies, C.M., Webster, J.P., Kruger, O., Munatsi, A., Ndamba, J., Woolhouse, M.E.J., 1999. Host-parasite population genetics: a cross-sectional comparison of *Bulinus globosus* and *Schistosoma haematobium*. Parasitology 119, 295–302.

Davies, C.M., Woolhouse, M.E.J., Webster, J.P., 2001. Tradeoffs in the evolution of virulence of schistosomes: macroparasites with an indirect lifecycle. Proc. R. Soc. Lond.: Ser. B 268, 251–257.

Doenhoff, M.J., Hagan, P., Cioli, D., Southgate, V., Pica-Mattoccia, L., Botros, S., Coles, G., Tchuem Tchuenté, L.A., Mbaye, A., Engels, D., 2009. Praziquantel: its use in control of schistosomiasis in sub-Saharan Africa and current research needs. Parasitology 136, 1825–1835.

Durand, P., Sire, C., Theron, A., 2000. Isolation of microsatellite markers in the digenetic trematode *Schistosoma mansoni* from Guadeloupe island. Mol. Ecol. 9, 997–998.

El Mousadik, A., Petit, R.J., 1996. High levels of genetic differentation for allelic richness among populations of the argan tree [*Argania spinosa* (L) *skeels*] endemic to Morocco. Theor. Appl. Genet. 92, 832–839.

Excoffier, L., Laval, G., Schneider, S., 2005. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. Evol. Bioinform. Online 1, 47–50.

Fallon, P.G., Doenhoff, M.J., 1994. Drug-resistant schistosomiasis: resistance to praziquantel and oxamniquine induced in *Schistosoma mansoni* in mice is drug specific. Am. J. Trop. Med. Hyg. 51, 83–88.

Felsenstein, J., 2005. PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Fenwick, A., Webster, J.P., Bosque-Oliva, E., Blair, L., Fleming, F.M., Zhang, Y., Garba, A., Stothard, J.R., Gabrielli, A.F., Clements, A.C.A., Kabatereine, N.B., Toure, S., Dembele, R., Nyandindi, U., Mwansa, J., Koukounari, A., 2009. The Schistosomaisis Control Initiative (SCI): rationale, development and implementation from 2002–2008. Parasitology 136, 1719–1730.

Fenwick, A., Webster, J.P., 2006. Schistosomiasis – challenges for control, treatment and drug resistance. Curr. Opin. Infect. Dis. 19, 577–582.

French, M.D., Churcher, T.S., Basáñez, M.-G., Lwambo, N.J.S., Norton, A.J., Webster, J.P., 2013. Reductions in genetic diversity of *Schistosoma mansoni* populations under chemotherapeutic pressure: the effect of sampling approach and parasite population definition. Acta Trop. 128, 196–205.

Garba, A., Labbo, R., Tohon, Z., Sidiki, A., Djibrilla, A., 2004. Emergence of *Schistosoma mansoni* in the Niger River Valley, Niger. Trans. R. Soc. Trop. Med. Hyg. 98, 296–298.

Garba, A., Lamine, M.S., Barkiré, N., Djibo, A., Sofo, B., Gouvras, A.N., Labbo, R., Sebangou, H., Webster, J.P., Fenwick, A., Utzinger, J., 2011. Efficacy and safety of two closely spaced doses of praziquantel against *Schistosoma haematobium* and *S. mansoni* in school-aged children and re-infection patterns in Niger. Acta Trop. 128, 334–344.

Golan, R., Gower, C.M., Emery, A.M., Rollinson, D., Webster, J.P., 2007. Isolation and characterization of the first polymorphic microsatellite markers for *Schistosoma haematobium* and their application in multiplex reactions of larval stages. Mol. Ecol. Notes 8, 647–649.

Goudet, J., 2002. FSTAT: a computer program to calculate F statistics. Version 2.9.3.2. http://www2.unil.ch/popgen/softwares/fstat.htm (accessed 2010).

Gower, C.M., Gabrielli, A.F., Sacko, M., Dembele, R., Golan, R., Emery, A.M., Rollinson, D., Webster, J.P., 2011. Population genetics of *Schistosoma haematobium*: development of novel microsatellite markers and their application to schistosomiasis control in Mali. Parasitology 138, 978–994.

Gower, C.M., Shrivastava, J., Lamberton, P.H.L., Rollinson, D., Emery, A.M., Webster, B.L., Kabatereine, N.B., Webster, J.P., 2007. Development and application of an ethical and epidemiologically appropriate assay for the multi-locus analysis of *Schistosoma mansoni*. Parasitology 134, 523–536.

Gower, C.M., Webster, J.P., 2004. Fitness of indirectly-transmitted pathogens: restraint and constraint. Evolution 58, 1178–1184.

Ismail, M., Botros, S., Metwally, A., William, S., Farghally, A., Tao, L.F., Day, T.A., Bennett, J.L., 1999. Resistance to praziquantel: direct evidence from *Schistosoma mansoni* isolated from Egyptian villagers. Am. J. Trop. Med. Hyg. 60, 932–935.

King, C.H., Dickman, K., Tisch, D.J., 2005. Reassessment of the cost of chronic helmintic infection: a meta-analysis of disability-related outcomes in endemic schistosomiasis. Lancet 365, 1561–1569.

Knopp, S., Stothard, J.R., Rollinson, D., Mohammed, K.A., Khamis, I.S., Marti, H., Utzinger, J., 2013. From morbidity control to transmission control: time to change tactics against helminths in Unguja, Tanzania. Acta Trop. 128, 412–422.

Labbo, R., Garba, A., Louboutin-Croc, J.P., Ernould, J.C., Sellin, B., Chippaux, J., Stothard, J.R., 2003. The spread of *Biomphalaria pfeifferi* in the Niger River Valley, Niger. Ann. Trop. Med. Parasitol. 97, 209–212.

Lamberton, P.H.L., Hogan, S.C., Katbatereine, N.B., Fenwick, A., Webster, J.P., 2010. In vitro praziquantel test capable of detecting reduced *in vivo* efficacy in *Schistosoma mansoni* human infections. Am. J. Trop. Med. Hyg. 83, 1340–1347.

Liu, K., Muse, S.V., 2005. Powermarker: integrated analysis environment for genetic marker data. Bioinformatics 21, 2128–2129.

Lu, D.B., Rudge, J.W., Wang, T.P., Donnelly, C.A., Fang, G.R., Webster, J.P., 2010. Transmission of *Schistosoma japonicum* in marshland and hilly regions of China: parasite population genetic and sibship structure. PLoS Negl. Trop. Dis. 4, e781.

Morgan, J.A.T., Dejong, R.J., Adeoye, G.O., Ansa, E.D.O., Barbosa, C.S., Bremond, P., Cesari, I.M., Charbonnel, N., Correa, L.R., Coulibaly, G., D'Andrea, P.S., Souza, C.P.D., Doenhoff, M.J., File, S., Idris, M.A., Incani, R.N., Jarne, P., Karanja, D.M.S., Kazibwe, F., Kpikpi, J., Lwambo, N.J.S., Mabaye, A., Magalhaes, L.A., Makundi, A., Mone, A., Mouahid, G., Mungai, C.M., Sene, M., Southgate, V., Tchuem Tchuente, L.A., Theron, A., Yousif, F., Zanotti-Magalhaes, E., Mkoji, G.M., Loker, E.S., 2005. Origin and diversification of the human parasite *Schistosoma mansoni*. Mol. Ecol. 14, 3889–3902.

Norton, A.J., Gower, C.M., Lamberton, P.H.L., Webster, B.L., Lwambo, N.J.S., Blair, L., Fenwick, A., Webster, J.P., 2010. Genetic consequences of mass human chemotherapy for *Schistosoma mansoni*: population structure pre- and post-praziquantel treatment in Tanzania. Am. J. Trop. Med. Hyg. 83, 951–957.

Prugnolle, F., Liua, H., de Meeus, T., Balloux, F., 2005. Population genetics of complex life-cycle parasites: an illustration with trematodes. Int. J. Parasitol. 35, 255–263.

Rollinson, D., 2009. A wake up call for urinary schistosomiasis: reconciling research effort with public health importance. Parasitology 136, 1801–1811.

Rudge, J.W., Carabin, H., Balolong, E.J., Tallo, V., Shrivastava, J., Lu, D.B., Basáñez, M.G., Olveda, R., McGarvey, S.T., Webster, J.P., 2008. Population genetics of *Schistosoma japonicum* within the Phillipines suggest high levels of transmission between humans and dogs. PLoS Negl. Trop. Dis. 2, e340.

Schistosoma japonicum Genome Sequencing Functional Analysis Consortium, 2009. The *Schistosoma japonicum* genome reveals features of host-parasite interplay. Nature 460, 345–351.

Shrivastava, J., Barker, G.C., Johansen, M.V., Zhou, X.N., Aligui, G.D., McGarvey, S.T., Webster, J.P., 2003. Isolation and characterisation of polymorphic DNA microsatellite markers from *Schistosoma japonicum*. Mol. Ecol. Notes 3, 406–408.

Shrivastava, J., Gower, C.M., Balolong, E.J., Wang, T.P., Qian, B.Z., Webster, J.P., 2005. Population genetics of multi-host parasites – the case for molecular epidemiological studies of *Schistosoma japonicum* using naturally sampled larval stages. Parasitology 131, 617–626.

Slatkin, M., 1987. Gene flow and the geographic structure of natural populations. Science 236, 787–792.

Steinauer, M.L., Hanelt, B., Agola, L.E., Mkoji, G.M., Loker, E.S., 2009. Genetic structure of *Schistosoma mansoni* in western Kenya: the effects of geography and host sharing. Int. J. Parasitol. 39, 1353–1362.

Steinauer, M.L., Blouin, M.S., Criscione, C.D., 2010. Applying evolutionary genetics to schistosome epidemiology. Infect. Genet. Evol. 10, 433–443.

Steinmann, P., Keiser, J., Bos, R., Tanner, M., Utzinger, J., 2006. Schistosomiasis and water resources development: systematic review, meta-analysis, and estimates of people at risk. Lancet Infect. Dis. 6, 411–425.

Valentim, C.L., LoVerde, P.T., Anderson, T.J., Criscione, C.D., 2009. Efficient genotyping of *Schistosoma mansoni* miracidia following whole genome amplification. Mol. Biochem. Parasitol. 166, 81–84.

Wang, J., 2001. A pseudo-likelihood method for estimating effective population size from temporally spaced samples. Genet. Res. 78, 243–257.

Wang, J., Whitlock, M.C., 2003. Estimating effective population size and migration rates from genetic samples over space and time. Genetics 163, 429–446.

Waples, R.S., Do, C., 2008. LDNE: a program for estimating effective population size from data on linkage disequilibrium. Mol. Ecol. Resour. 8, 753–756.

Waples, R.S., Yokota, M., 2007. Temporal estimates of effective population size in species with overlapping generations. Genetics 175, 219–233.

Webster, J.P., Shrivastava, J., Johnson, P., Blair, L., 2007. Is host-schistosome coevolution going anywhere? BMC Evol. Biol. 7, 91–100.

Webster, J.P., Gower, C.M., Blair, L., 2004. Do hosts and parasites coevolve? Empirical support from the *Schistosoma* system. Am. Nat. 164, S33–S51.

Webster, J.P., Gower, C.M., Norton, A.J., 2008. Evolutionary concepts in predicting and evaluating the impact of mass-chemotherapy schistosomiasis control programmes on parasites and their hosts. Evol. Appl. 1, 66–83.

Webster, J.P., Oliveira, G., Rollinson, D., Gower, C.M., 2010. Schistosome genomes: a wealth of information. Trends Parasitol. 26, 103–106.

Young, N.D., Jex, A.R., Li, B., Liu, S., Yang, L., Xiong, Z., Li, Y., Cantacessi, C., Hall, R.S., Xu, X., Chen, F., Wu, X., Zerlotini, A., Oliveira, G., Hofmann, A., Zhang, G., Fang, X., Kang, Y., Campbell, B.E., Loukas, A., Ranganathan, S., Rollinson, D., Rinaldi, G., Brindley, P.J., Yang, H., Wang, J., Wang, J., Gasser, R.B., 2012. Whole-genome sequence of *Schistosoma haematobium*. Nat. Genet. 44, 221–225.