Parasites & Vectors

RESEARCH   **Open Access**

The population genetic structure of *Biomphalaria choanomphala* in Lake Victoria, East Africa: implications for schistosomiasis transmission

Claire J Standley[1,2,4*], Sara L Goodacre[1], Christopher M Wade[1] and J Russell Stothard[3]

Abstract

Background: The freshwater snail *Biomphalaria* acts as the intermediate host of *Schistosoma mansoni*, a globally important human parasite. Understanding the population structure of intermediate host species can elucidate transmission dynamics and assist in developing appropriate control methods.

Methods: We examined levels of population genetic structure and diversity in 29 populations of *Biomphalaria choanomphala* collected around the shoreline of Lake Victoria in Uganda, Kenya and Tanzania, where *S. mansoni* is hyper-endemic. Molecular markers were utilized to estimate the degree to which snail populations are genetically differentiated from one another.

Results: High levels of snail genetic diversity were found coupled with evidence of geographically-determined population structure but low levels of local inbreeding. The data are consistent with an effect of schistosome infection on population structure of intermediate host snails, but other factors, such as habitat and historical demographic changes, could also be important determinants of the degree of population genetic structure in *Biomphalaria choanomphala*.

Conclusions: The low stratification of populations and high genetic diversity indicates potentially less local compatibility with intermediate snail populations than previously theorized, and highlights the importance of coordinated parasite control strategies across the region.

Keywords: *Biomphalaria choanomphala*, *Schistosoma mansoni*, Population structure, Population genetics

Background

Historically, the nature and dynamics of host-parasite relationships has attracted much scientific attention. Parasites of human populations are an area of particular research interest, given the medical interest in controlling their transmission. As such, the genetic diversity and population structure of hermaphroditic *Biomphalaria* freshwater snails of the family Planorbidae, which are the obligatory intermediate hosts of the globally important parasitic trematode of humans, *Schistosoma mansoni*, have been the subject of considerable study.

It is known that increased genetic diversity of the host *Biomphalaria* population reduces the overall parasite transmission rate whereas reduced host genetic diversity appears to benefit *S. mansoni* through an overall increase in transmission rate [1-5]. It has also been shown that schistosome infection in *Biomphalaria* reduces snail fitness, but resistance to infection is itself also associated with reduced offspring production, as often seen in other planorbids [6]. As such, it has been predicted that the natural *Biomphalaria* population structure would result in a tightly coupled system whereby parasites and snails co-adapt, resulting in localized compatibility and comparable population structures, but also driving local high diversity [7]. Of course, parasites other than *Schistosoma* could equally be implicated in driving diversity in such a system, as *Biomphalaria* are known to serve as intermediate hosts for a number of nematode and trematode worms [8,9].

* Correspondence: claire.standley@gmail.com
[1]School of Life Sciences, University of Nottingham, Nottingham NG7 2RD, UK
[2]Department of Zoology, Wolfson Wellcome Biomedical Laboratory, Natural History Museum, Cromwell Road, London SW7 5BD, UK
Full list of author information is available at the end of the article

Parasites such as *S. mansoni* are, however, not the only external factor known to influence their host. Selection driven by environmental factors, such as seasonality or habitat type, has also been shown to affect freshwater snail population differentiation [10-12]. There may also be anthropogenic influences, for example, removal of snails as part of disease control programs [13].

B. glabrata and *B. pfeifferi* are considered two of the most important intermediate host snail species of *S. mansoni* in the New and Old World, respectively, and most previous work on snail-schistosome population structure has focused on these species [14]. There are, however, other regions of the world where intestinal schistosomiasis is highly endemic, yet transmitted by other *Biomphalaria* species [15], and fewer studies have focused on these intermediate hosts. In this study we investigate the genetic diversity and population structure of *Biomphalaria choanomphala* from Lake Victoria, an important lacustrine environment in East Africa, and a well-known regional hot spot for transmission of *S. mansoni* [16-19]. The aim of the study is to determine the population structure of *Biomphalaria choanomphala* in this hyper-endemic region for schistosomiasis, at a large scale, to better understand factors that may contribute to parasite transmission.

Methods

The degree of population differentiation of *Biomphalaria choanomphala* in Lake Victoria was inferred from genetic variation in two mitochondrial genes (cytochrome oxidase sub-unit one (COI) and 16S ribosomal RNA (16S)), and four bi-parentally inherited nuclear microsatellite loci.

Sampling methodology and population selection

Snails were collected from 29 sites around the perimeter of Lake Victoria, using hand held scoops (Figure 1). Sampling was semi-quantified, with two collectors each surveying a 50 m length of shoreline for approximately 20 minutes and collecting all *Biomphalaria* found. Sites were selected to ensure an even geographical spread along the lakeshore, in all three countries and in both marsh and lake habitats; five localities were specifically included because they incorporated both habitat types in close proximity. All sites showed signs of human activity near the shoreline. In order to compare sites of similar population density, all sites selected had a high abundance of *Biomphalaria*, defined as 30 or more individuals collected, which accounted for the paucity of sites along the eastern shoreline of the lake in Tanzania, where *Biomphalaria* densities were generally low.

Snails were putatively identified as *B. choanomphala* or *B. sudanica*, based on shell morphology, as these are the species thought primarily to inhabit Lake Victoria [15]. However, recent taxonomic work has suggested that all the *Biomphalaria* in Lake Victoria can be considered a single species of which ecophenotypy of shell form gives rise to the previously considered taxonomic distinctions [20], and thus for the purposes of this study, individuals were simply classed as being one of the two morphotypes of *B. choanomphala*.

All collected snails (in total >30 per site, given the site selection criteria) were exposed, as a group, to sunlight for 2–4 hours to check for infection with *Schistosoma mansoni* or other parasite cercariae [21]. A random sample of snails was examined in this way on subsequent days, to ensure maximum reliability of infection testing. After shedding, snails were placed in glass tubes containing 95% ethanol solution.

DNA extraction, amplification and sequencing

Genomic DNA was extracted from between 10 and 12 individuals from each of the selected 29 populations, using a standard CTAB extraction [22] with a final re-suspension in pure water. The mitochondrial cytochrome oxidase sub-unit one (COI) gene and the mitochondrial 16S sub-unit of the ribosomal RNA gene were amplified using the Folmer 'universal' primers [23] and modified 16ar and 16br primers [24], respectively. The four microsatellite loci were Bpf1, Bpf2, Bpf3 and Bpf10 [25] – all nine of the primers described in referenced paper were tested, but the above four were the only ones which amplified successfully and were polymorphic in the Lake Victoria *Biomphalaria choanomphala* under examination here. Mitochondrial (COI and 16S) amplifications were carried out in 25 µl total volume, with 2.5 µl $MgCl_2$ (20 mM concentration), 2.5 µl 5 × buffer, 2.5 µl pre-mixed dNTPs (20 mM concentration), 1 µl each of forward and reverse primer (10 pmol concentration) and one unit of TAQ per reaction. Microsatellite reactions were amplified in 12.5 µl total volume, with half of the volumes stated above, and with forward primer concentration of 50 pmol and reverse (fluorescently labeled) primer concentration of 20 pmol. Cycling conditions in all cases followed published methods [23-25].

Positively amplified products were purified using Millipore PCR$_{96}$ Cleanup kits on a vacuum manifold (Millipore, Billerica, USA) as per manufacturer's instructions, using pure water for washing and re-suspension. Product concentration was quantified on a Nanodrop ND-1000 Spectrophotometer (Nanodrop Technologies Inc., Willington, USA), and sequencing reactions were performed on mitochondrial purified PCR products using an Applied Biosystems Big Dye Kit (version 1.1) and run on an Applied Biosystems 3730 DNA Analyzer (Applied Biosystems, Carlsbad, USA). Microsatellites were diluted 1:10 in HiDiformamide (Applied Biosystems, as above) and analyzed using an ABI3730 automated sequencer, using a GeneScan 500 LIZ size standard (both Applied Biosystems, as above).
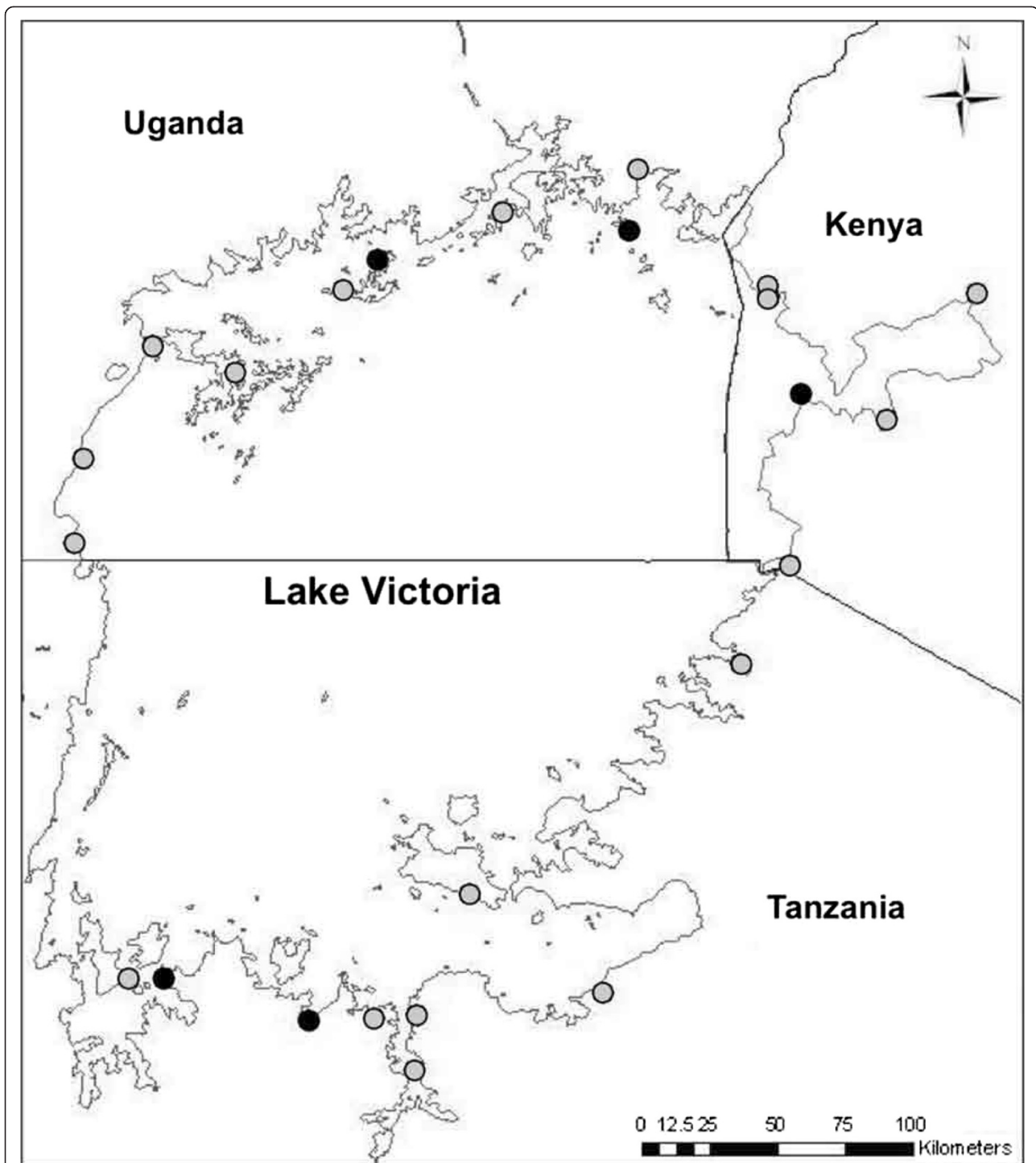
Figure 1 Map of the 29 study sites around Lake Victoria. "Single habitat" sites (marked by grey circle) denote an area with a single habitat (lake or marsh); "Joint habitat" sites (marked by black circle) indicate localities where a marsh and lake habitat were found together, and snails collected separately from each.

Population analyses

The overall framework of the population genetics analyses were designed to examine the population structure, diversity, demography, phylogeography and spatial correlations of the *B. choanomphala* populations surveyed from Lake Victoria [2].

COI and 16S mitochondrial sequences were aligned using MUSCLE on the web (http://www.ebi.ac.uk/Tools/

msa/muscle/) and edited visually in MacClade v 4 [26]. The aligned sequences were then compared and reduced to unique haplotypes, and the proportion of each haplotype *per* population was recorded. The 16S data contained many gap-rich regions, for which it was impossible to be assured of homology; any region with a gap was therefore removed from the alignment before calculating genetic distances amongst sequences. There was only one indel region in the COI alignment, of a single codon repeat, in three haplotypes, which was therefore easily alignable and retained for haplotype analysis.

Measures of population differentiation were calculated from mitochondrial sequence data using analysis of molecular variance (AMOVA), F-statistics, mismatch analysis, Tajima and Fu's tests for neutrality and a Mantel z-test for correlation by distance. Each AMOVA was conducted using "country" as the primary grouping, to investigate the level of variation partitioned between each country, versus between populations within each country and within each population. "Country" was deemed to be a potentially important geographical factor given that each country in the region undertakes different levels of schistosomiasis control initiatives, thus potentially influencing the level of parasitism experienced by local snails. All tests were executed in Arlequin v 3.2 [27].

A distance-based, neighbor-joining tree was created in Paup* 4.0 [28] for the unique haplotypes that remained when indels were removed from the alignment from each gene, using the GTR + gamma model of sequence evolution, determined to be the most appropriate model through likelihood testing, also in Paup* 4.0. Median-joining networks were constructed using Network 4.6 [29] based on the gap-free unique haplotypes.

The number and proportion of private microsatellite alleles per population was calculated using the web-based version of Genepop [30]. Calculations of F_{ST} and tests for Hardy-Weinberg equilibrium (through estimation of F_{IS} values per population and per locus) and linkage disequilibrium amongst loci were also done using Genepop. The loci were also tested for evidence of previous bottlenecking events by comparing the allelic diversity against the observed and expected heterozygosity, using the program Bottleneck [31].

Results

In total, over 300 *Biomphalaria choanomphala* were analyzed from the selected 29 populations. Both the mitochondrial markers and the microsatellites revealed high levels of genetic variation overall, but also significant population structuring throughout the lake as described below.
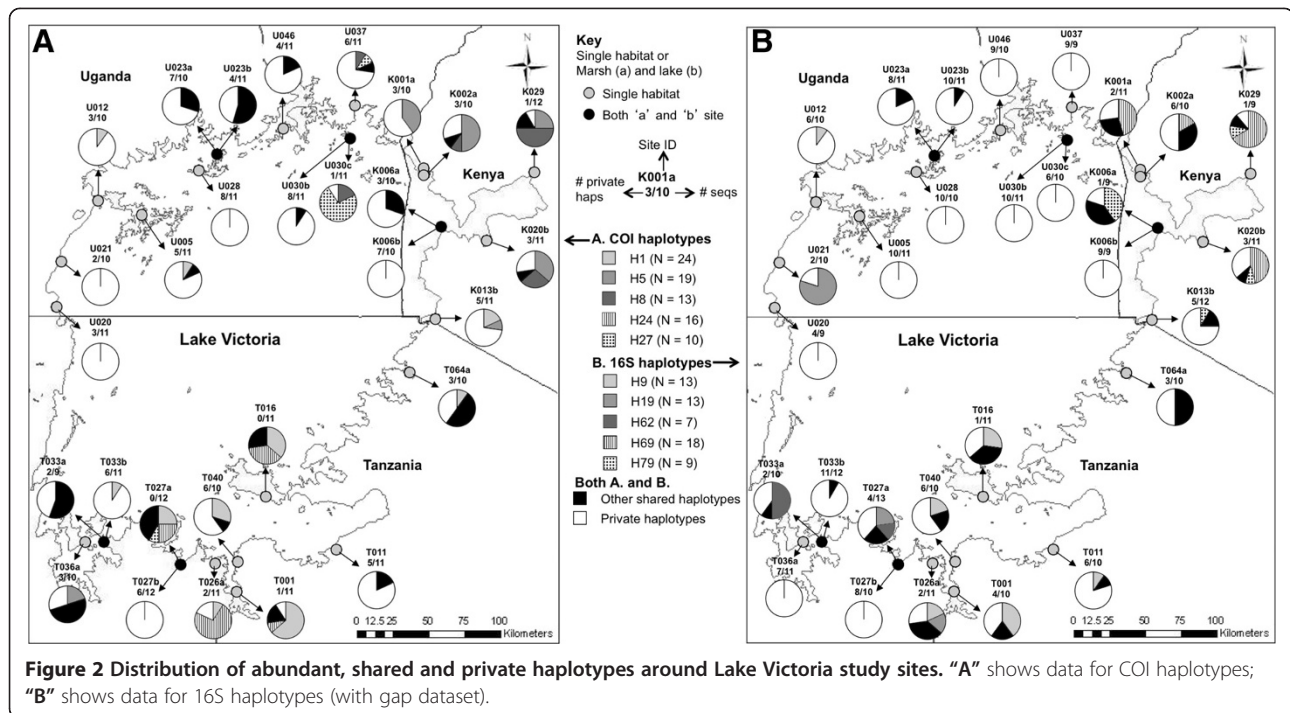
Mitochondrial data

The COI and 16S sequences were very genetically variable with 127/308 (CO1) and 181/300(16S) unique haplotypes/

total number of sequences respectively. Haplotypes were not highly divergent (0.002-0.058 pairwise distance for COI haplotypes and 0.000-0.047 for 16S). GenBank accession numbers for these datasets are HM769132-HM769258 for COI and HM768950-HM768980 and HM768982-HM769131 for 16S. When unalignable regions with gaps were removed from the analysis of the 16S sequences, the number of unique haplotypes was reduced from 181 to 64. Many common haplotypes were spread throughout the lake, but others were locally restricted and there was high variation in the haplotype richness of the sites: overall, intra-site haplotype diversity estimates varied from 0.200-0.982 for COI and 0.378-1.000 for 16S, where zero indicates complete homogeneity and one indicates maximum diversity.

Figure 2 shows the geographical distribution of the five most abundant COI and 16S ("A" and "B" maps in Figure 2, respectively) haplotypes (without removal of indels) per population across Lake Victoria, as compared to the frequency of other shared haplotypes and "private" haplotypes (i.e. haplotypes only observed at one site). The map reveals patterns of higher and lower haplotype sharing versus uniqueness within and between localities and countries; for example, Kenyan sites tended to share haplotypes whereas Ugandan sites, especially those in the central and western regions of the lakeshore, had high proportions of private haplotypes. Maps showing the distribution of all the haplotypes for COI and 16S, plus a full list of haplotype frequencies per site, can be found in the supplementary information (See Additional file 1: Figure S1, Additional file 2: Figure S2, Additional file 3: Table S1 and Additional file 4: Table S2 for COI and 16S respectively).

The phylogenetic trees for COI and 16S (Figure 3) were consistent, as would be expected from linked loci, and showed few deep divisions, but several more recent processes of divergence. These supported the hypothesis of geographical structuring, as divergent clades tended to be made up of haplotypes from the same country, or limited to two out of the three countries. The networks for each mitochondrial marker were difficult to analyze visually due to the large number of haplotypes (See Additional file 5: Figure S3 for COI and Additional file 6: Figure S4 for 16S). However there appeared to be evidence of geographical clustering by country; for both markers, the most abundant and geographical widespread haplotype (H1 for COI and H2 for 16S; the network software utilized the gap-free dataset) was a central node within each respective network.

The majority of the variation, for both the COI and the 16S data, was explained at the population level based on the analysis of molecular variance (AMOVA), which sought to determine how diversity was partitioned among and between the different sites. For these data, all three levels of structure (intergroup, intragroup and intrapopulation) were statistically significantly differentiated (see Table 1).

Figure 2 Distribution of abundant, shared and private haplotypes around Lake Victoria study sites. "A" shows data for COI haplotypes; "B" shows data for 16S haplotypes (with gap dataset).

When pair-wise fixation index (F_{ST}) values were calculated between each of the populations in turn, the majority demonstrated significant levels of differentiation, for both types of molecular marker. The range in values was 0.000-0.921 for COI and 0.000-0.845 for 16S. For the 16S data, all but one of the non-significant F_{ST} values were between sites in the same country, once again demonstrating geographically defined structuring of the populations for this gene. All but one paired site showed significant differentiation between the populations from lake versus marsh habitat, indicating even small-scale geographical segregation. Full tables of the pairwise F_{ST} values and their significance can be found in the supplementary information (See Additional file 7: Tables S3 and Additional file 8: Tables S4 respectively for COI and 16S).

All of the populations conformed to the assumptions of the neutral hypothesis, as tested for by Tajima's D statistic for the 16S data; one population, T001, deviated significantly (p <0.001) from assumptions of normality when the COI data was analyzed. However, this population also had a significant result for the sum of squares mismatch analysis (p <0.001), suggesting that demographic change, such as population expansion, may have biased the neutrality test. No other populations, for either marker, came out as significant in the mismatch analysis.
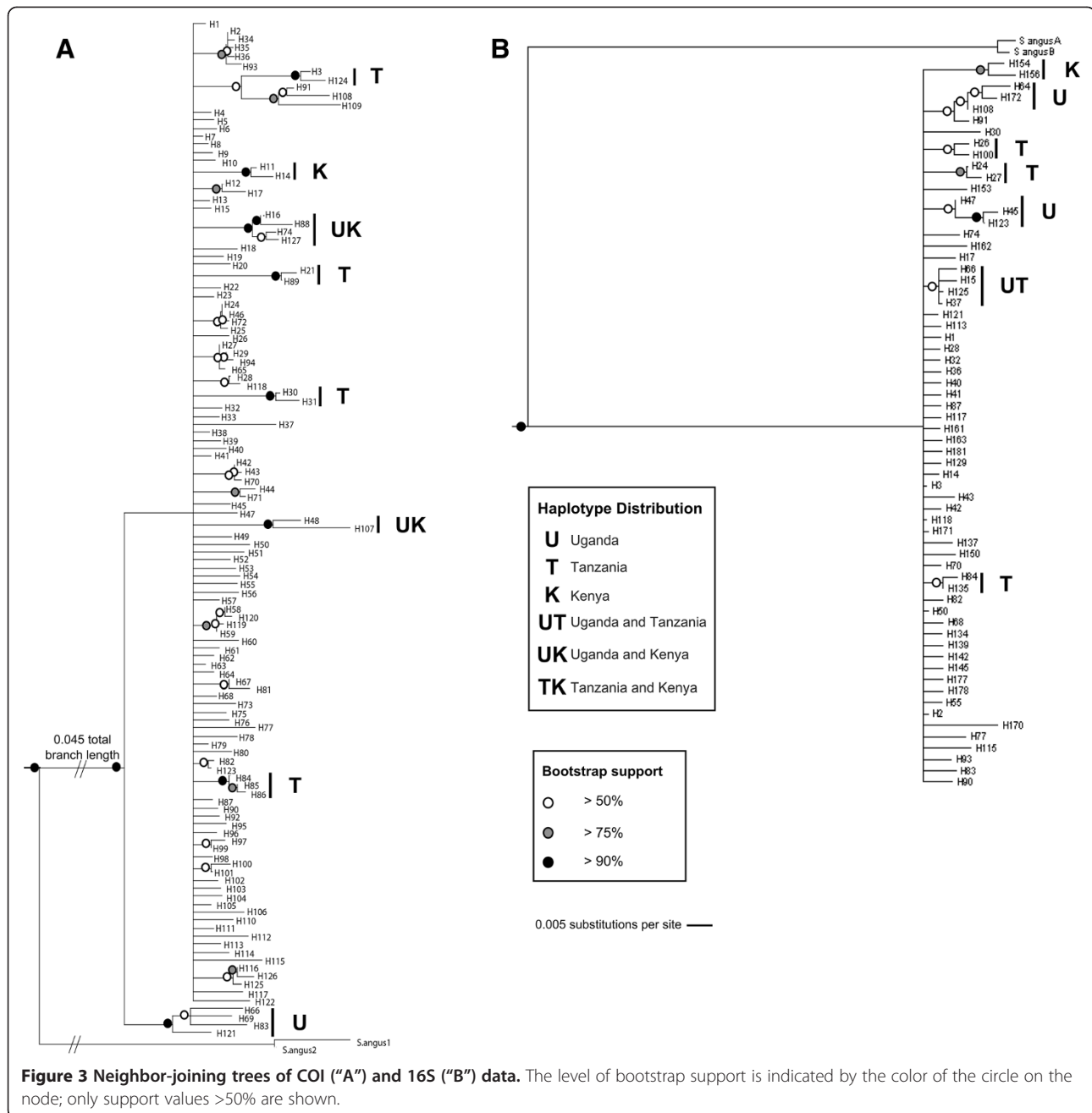
The relationship with geography was elucidated through the significance of the Mantel z-test for both the COI and 16S data (p =0.001, for both), meaning that the genetic and geographical distances between the populations was positively correlated; in other words, sites closer together were more likely to be genetically similar, even if the composition of distinct haplotypes differed per site (as seen in Figure 3 and Additional file 1: Figure S1). This demonstrated an overall relationship between genetic differentiation and geographic distance, on top of the localized differentiation already described, such as between marsh and lake habitats (Additional file 5: Figure S3 and Additional file 6: Figure S4).

Microsatellite data

The microsatellite data were tested for departures from Hardy-Weinberg equilibrium (HWE), linkage disequilibrium and for an excess of heterozygosity relative to allelic diversity, which would be consistent with populations having recently passed through a bottleneck.

As with the mitochondrial data, based on an AMOVA, the majority of the variation in microsatellite genotype was explained at the population level; values of differentiation at the intragroup and intrapopulation were both highly significant (explaining 10.55% and 89.13% of the total variation observed, respectively: Table 2). Much less variation was explained between country-level groups for the microsatellite data (0.31%) than for either of the mitochondrial genes, but instead there was a larger proportion of diversity within populations (>89%). Pairwise F_{ST} values ranged from 0.000-0.298, but fewer values were significant between populations, compared to the mitochondrial data. The full matrix of pairwise F_{ST} values and their significance can be found in the supplementary information (Additional file 9: Table S5). Supporting the F_{ST}

Figure 3 Neighbor-joining trees of COI ("A") and 16S ("B") data. The level of bootstrap support is indicated by the color of the circle on the node; only support values >50% are shown.

values in suggesting low gene flow between populations, 17 out of the 28 populations with microsatellite data had private alleles at one or more loci, and the proportions of private alleles across all populations and alleles was highly significant (Fisher's method, p = <0.001). The microsatellite data also supported the mitochondrial evidence for genetic spatial autocorrelation, as the Mantel z-test was significant (p = 0.017).

Intrapopulation tests however showed low levels of inbreeding within populations, based on the calculations of F_{IS} values per locus as well as across all loci. F_{IS} values,

across all loci, per population, ranged from −0.186 to 0.168 (Table 3), with a mean value across all populations of 0.019, indicative of random mating. More generally, all populations conformed to the assumptions of HWE in terms of observed and expected heterozygosity, supporting the hypothesis that there are low levels of selfing.

There was also no evidence for linkage disequilibrium between the loci, suggesting that the alleles considered are not closely associated in the genome and are mixing randomly during gamete production. Four populations (K013b, T001, T027b and U028), when tested for

Table 1 Analysis of molecular variance (AMOVA) results for COI and 16S data

	Source of variation	Degrees of freedom	Sum of squares	Variance components	Percentage of variation	p-value
COI sequences	**Among countries**	2	173.35	0.6646	12.75	< 0.0001
	Among populations, within countries	26	512.68	1.5774	30.26	< 0.0001
	Within populations	279	829.10	2.9717	57.00	< 0.0001
	Total	307	1515.12	5.2137	100.00	
16S sequences	**Among countries**	2	155.88	0.5893	10.00	< 0.001
	Among populations, within countries	26	524.77	1.5935	27.04	< 0.0001
	Within populations	271	1005.55	3.7105	62.96	< 0.0001
	Total	299	1686.20	5.8933	100.00	

evidence of a past bottleneck, showed heterozygote excess (p = 0.031 for all four), but this value was not significant once the Bonferroni correction for multiple tests was applied.

Discussion

Population structure of Biomphalaria choanomphala in Lake Victoria

The populations studied here were characterized by high intrapopulation diversity as well as high levels of population structure, with low levels of gene flow as inferred from molecular data and negligible inbreeding. High diversity is consistent with what would be expected if parasitism were strongly influencing population structure [3,7], and the low F_{IS} values and apparent random mating observed here further supports this hypothesis. The snails collected for this research were examined for infection with schistosomes and other parasites in the field as part of a different study; however, without more accurate and consistent methods such as molecular probes [32] for the detection of infected snails (and indeed, infection with potentially more than one parasite species [33]) and infected snail populations, it is difficult to determine whether parasitism is indeed a key driver of population structure in this system. Similarly, future studies should seek to use genetic markers which are known to correlate to parasite resistance or susceptibility, which were not available when this study was conducted, but have since been identified in several studies [34,35]. Even so, the challenges of modifying

markers for use on field populations of less well-studied African *Biomphalaria* remain.

We would also expect parasite populations to adapt to local snails, thus mirroring to some extent their population structure. We used a Mantel test [36] to compare the population structure of *B. choanomphala* from the six sites from which there were published data on schistosome population structure [19], and found no correlation. Notwithstanding the limitation of the small number of sites, this conforms to other studies where parasite populations were much less differentiated than those of the intermediate host, suggesting local compatibility in this context is being overshadowed by other factors such as high parasite migration [36,37]. Definitive host migration has been put forward as a factor that might homogenize parasite populations across a wide area, and this is compatible with the Lake Victoria context, where human populations are known to move widely throughout the region [18].

Role of other factors influencing population structure

Given the possibility of parasitism as a driver for population structure in Lake Victoria snail populations, it is important to consider what other factors might be influential in this system, and particularly how these differ from previous studies of *Biomphalaria* in other settings. Such considerations link back to parasitism, and particularly control of schistosomiasis in human populations, in terms of gaining a better understanding of drivers for distribution and abundance of host snails under different conditions. Studies of

Table 2 Analysis of molecular variance (AMOVA) results for microsatellite data

	Source of variation	Degrees of freedom	Sum of squares	Variance components	Percentage of variation	p-value
Microsatellite data	**Among countries**	2	9.51	0.0057	0.31	0.237
	Among populations, within countries	23	88.36	0.1900	10.55	< 0.0001
	Within populations	286	459.05	1.6051	89.13	< 0.0001
	Total	311	556.93	1.8008	99.99	

Table 3 Diversity estimates per site for COI, 16S and microsatellite data

Site	General information		COI sequences		16S sequences		Microsatellite data		
	Shedding	Habitat	Gene diversity	Nucleotide diversity	Gene diversity	Nucleotide diversity	Average gene div.	Private alleles	F_{IS}
K001a	N	Marsh	0.836	0.005	0.746	0.005	0.885	0.000	−0.010
K002a	N	Marsh	0.712	0.001	0.978	0.011	0.878	0.029	−0.001
K006a	N	Marsh	0.855	0.014	0.500	0.005	0.790	0.079	−0.077
K006b	N	Lake	0.889	0.026	1.000	0.043	0.843	0.000	0.119
K013b	N	Lake	0.909	0.014	0.879	0.012	0.710	0.139	0.025
K020b	N	Lake	0.933	0.004	0.800	0.006	0.825	0.000	0.083
K029	N	Lake	0.818	0.015	0.583	0.004	0.828	0.000	0.034
T001	Y	Marsh	0.978	0.017	0.867	0.017	0.754	0.016	−0.186
T011	N	Lake	0.927	0.009	0.933	0.016	0.738	0.015	0.073
T016	N	Marsh	0.964	0.016	0.727	0.008	0.743	0.000	−0.077
T026a	N	Marsh	0.964	0.009	0.891	0.005	0.636	0.019	−0.112
T027a	N	Marsh	0.782	0.010	0.923	0.018	0.848	0.014	−0.180
T027b	N	Lake	0.733	0.003	0.956	0.033	0.735	0.050	0.038
T033a	N	Marsh	0.778	0.002	0.711	0.014	0.738	0.000	0.026
T033b	Y	Lake	0.778	0.004	1.000	0.021	0.758	0.014	0.134
T036a	N	Marsh	0.618	0.008	0.818	0.016	0.744	0.013	0.092
T040	N	Lake	0.727	0.003	0.978	0.011	0.803	0.033	0.168
T064a	N	Marsh	0.491	0.002	0.889	0.022	0.783	0.000	0.112
U005	N	Lake	0.864	0.011	0.982	0.027	0.819	0.013	0.101
U012	N	Marsh	0.639	0.010	0.911	0.020	0.845	0.000	0.058
U020	N	Marsh	0.756	0.006	0.583	0.007	0.806	0.028	0.116
U021	N	Marsh	0.844	0.015	0.378	0.002	0.754	0.000	−0.119
U023a	N	Lake	0.644	0.004	0.978	0.033	0.818	0.000	−0.039
U023b	Y	Marsh	0.346	0.004	1.000	0.028	0.878	0.016	0.109
U028	Y	Lake	0.200	0.000	1.000	0.025	0.857	0.015	0.014
U030b	N	Marsh	0.982	0.016	0.982	0.028	0.877	0.000	−0.104
U030c	N	Lake	0.473	0.002	0.844	0.014	NA	NA	NA
U037	Y	Lake	0.836	0.005	0.978	0.019	0.833	0.012	0.091
U046	N	Lake	0.712	0.001	1.000	0.030	0.868	0.025	0.050

Information on infection status with schistosomes and the habitat type are also given. "N" stands for "No" and "Y" represents "Yes". "div."stands for diversity and "FIS" indicates the inbreeding co-efficient.

B. pfeifferi in Madagascar, for example, revealed high levels of interpopulation variation but, in contrast to what we observed here, consistently low levels of intra-population genetic differentiation [11]. This was explained as being due to the combination of low levels of migration combined with habitat stochasticity; *B. pfeifferi* are also known to be frequent in-breeders, which would assist in maintaining a metapopulation with seasonal local extinction events [38].

In contrast, Lake Victoria possesses very different environmental conditions; as a very large, permanent lake, it rarely experiences the kind of environmental perturbations that characterize temporally transient ponds and streams. Such homogeneity of habitat, with associated infrequent extinctions and opportunity for genetic drift, could account for the high levels of intrapopulation diversity that are maintained in Lake Victorian *B. choanomphala* [2]. This applies mainly to the lake habitats; indeed, this study showed significantly lower levels of gene and nucleotide diversity in the marsh habitats. However, in both lake and marsh populations, the tests for bottleneck events were uniformly negative,

suggesting that sudden demographical events are not overwhelmingly contributing to the patterns of diversity of Lake Victoria *Biomphalaria*.

Effect on transmission of intestinal schistosomiasis

Although the low population structure in *S. mansoni* observed previously in this region counters the hypothesis of highly specialized local parasite compatibility [19] for both host and parasite, certain *B. choanomphala* haplotypes were widespread around the lake. This suggests that, as seen in *B. pfeifferi* from elsewhere in Africa, there may be a risk of transmissive genotypes becoming widely successful, if there is selection for traits independent of parasite susceptibility [39].

Another factor to consider is human-mediated changes to parasite populations: all three countries bordering Lake Victoria have initiated efforts to control schistosomiasis in human populations through mass treatment campaigns, although these initiatives are not synchronized and differ in their degree of coverage. The low levels of population structuring seen in the parasite [19] are likely a function of high terminal host migration and rapid dispersal of the parasite into novel localities [36,40]; if some populations of *S. mansoni* experience more intense selection pressure based on frequent mass drug administration campaigns, and given the generally transmissive snail populations, there is a risk that these highly adapted forms of *S. mansoni* will be able to spread rapidly throughout the region. This emphasizes the need for coordinated control strategies, both at the level of human treatment as well as improved water access and sanitation, throughout lakeshore communities in the region.

Conclusions

In summary, our analysis observed high genetic diversity of *B choanomphala* snails, yet also high population structuring; the high levels of observed inter- and intrapopulation diversity are consistent with parasitism as an influencing factor, but further investigation is needed, utilizing new tools to detect infection and new markers directly associated with resistance or susceptibility, in order to confirm the role of parasitism in driving high diversity in this system. Moreover, other factors, such as environmental and demographical stability may also contribute to the observed population structure. From a public health perspective, population genetic surveys of intermediate hosts should seek to increase the scale of their focus, as parasite transmission is likely influenced by factors, such as human migration and national treatment campaigns, that act on a national or even regional level. Public health initiatives should take such issues of scale into consideration when designing control strategies.

Additional files

Additional file 1: Figure S1. Haplotype map of all COI haplotypes observed.

Additional file 2: Figure S2. Haplotype map of all 16S haplotypes observed.

Additional file 3: Table S1. List of all COI haplotypes and their frequencies, per site.

Additional file 4: Table S2. List of all 16S haplotypes and their frequencies, per site.

Additional file 5: Figure S3. Median-joining network of COI haplotypes.

Additional file 6: Figure S4. Median-joining network of 16S haplotypes (without-gap haplotypes).

Additional file 7: Table S3. Full table of pairwise F_{ST} values per site for COI marker. Non-significant pairwise distances are given in italics.

Additional file 8: Table S4. Full table of pairwise F_{ST} values per site for 16S marker. Non-significant pairwise distances are given in italics.

Additional file 9: Table S5. Full table of pairwise F_{ST} values per site for microsatellite markers. Non-significant pairwise distances are given in italics.

Author details
[1]School of Life Sciences, University of Nottingham, Nottingham NG7 2RD, UK. [2]Department of Zoology, Wolfson Wellcome Biomedical Laboratory, Natural History Museum, Cromwell Road, London SW7 5BD, UK. [3]Department of Parasitology, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool L3 5QA, UK. [4]Present address: Milken Institute School of Public Health, George Washington University, Washington, D.C. 20052, USA.

References
1. Hoffman JI, Webster JP, Ndamba J, Woolhouse MEJ: **Extensive genetic variation revealed in adjacent populations of the schistosome intermediate host *Biomphalaria pfeifferi* from a single river system.** *Ann Trop Med Parasitol* 1998, **92**:693–698.
2. Jarne P, Theron A: **Genetic structure in natural populations of flukes and snails: a practical approach and review.** *Parasitology* 2001, **123**:S27–S40.
3. Webster JP, Davies CM, Hoffman JI, Ndamba J, Noble LR, Woolhouse MEJ: **Population genetics of the schistosome intermediate host *Biomphalaria***

pfeifferi in the Zimbabwean highveld: implications for co-evolutionary theory. *Ann Trop Med Parasitol* 2001, **95**:203–214.

4. Sandland GJ, Foster AV, Zavodna M, Minchella DJ: Interplay between host genetic variation and parasite transmission in the *Biomphalaria glabrata Schistosoma mansoni* system. *Parasitol Res* 2007, **101**:1083–1089.

5. Sandland GJ, Wethington AR, Foster AV, Minchella DJ: Effects of host outcrossing on the interaction between an aquatic snail and its locally adapted parasite. *Parasitol Res* 2009, **105**:555–561.

6. Webster JP, Woolhouse MEJ: Cost of resistance: relationship between reduced fertility and increased resistance in a snail–schistosome host–parasite system. *Proc R Soc B Biol Sci* 1999, **266**:391–396.

7. Webster JP, Shrivastava J, Johnson PJ, Blair L: Is host-schistosome coevolution going anywhere? *BMC Evol Biol* 2007, **7**:91. doi:10.1186/1471-2148-7-91.

8. Anderson RC: *Nematode Parasites of Vertebrates: Their Development and Transmission. Volume 1.* Wallingford, UK: CABI; 2000:672.

9. Esteban J-G, Muñoz-Antoli C, Trelis M, Toledo R: *Biomphalaria Snails and Larval Trematodes.* New York, NY: Springer New York; 2011:127–157.

10. Webster JP, Davies CM, Ndamba J, Noble LR, Jones CS, Woolhouse MEJ: Spatio-temporal genetic variability in the schistosome intermediate host *Biomphalaria pfeifferi. Ann Trop Med Parasitol* 2001, **95**:515–527.

11. Charbonnel N, Angers B, Rasatavonjizay R, Bremond P, Debain C, Jarne P: The influence of mating system, demography, parasites and colonization on the population structure of *Biomphalaria pfeifferi* in Madagascar. *Mol Ecol* 2002, **11**:2213–2228.

12. Mavárez J, Pointier JP, David P, Delay B, Jarne P: Genetic differentiation, dispersal and mating system in the schistosome-transmitting freshwater snail *Biomphalaria glabrata. Heredity (Edinb)* 2002, **89**:258–265.

13. Lardans V, Dissous C: Snail Control Strategies for Reduction of Schistosomiasis Transmission. *Parasitol Today* 1998, **14**:413–417.

14. Morgan JAT, Dejong RJ, Snyder SD, Mkoji GM, Loker ES: *Schistosoma mansoni* and *Biomphalaria*: past history and future trends. *Parasitology* 2001, **123**:S211–S228.

15. Brown D: *Freshwater Snails of Africa and Their Medical Importance.* 2nd edition. London: Taylor and Francis; 1994.

16. Stothard JR, Webster BL, Weber T, Nyakaana S, Webster JP, Kazibwe F, Kabatereine NB, Rollinson D: Molecular epidemiology of *Schistosoma mansoni* in Uganda: DNA barcoding reveals substantive genetic diversity within Lake Albert and Lake Victoria populations. *Parasitology* 2009, **136**:1813–1824.

17. Stothard JR, Kabatereine NB, Tukahebwa EM, Kazibwe F, Mathieson W, Webster JP, Fenwick A: Field evaluation of the Meade Readiview handheld microscope for diagnosis of intestinal schistosomiasis in Ugandan school children. *Am J Trop Med Hyg* 2005, **73**:949–955.

18. Standley CJ, Adriko M, Alinaitwe M, Kazibwe F, Kabatereine NB, Stothard JR: Intestinal schistosomiasis and soil-transmitted helminthiasis in Ugandan schoolchildren: a rapid mapping assessment. *Geospat Health* 2009, **4**:39–53.

19. Standley CJ, Kabatereine NB, Lange C, Lwambo NJS, Stothard JR: Molecular epidemiology and phylogeography of *Schistosoma mansoni* around Lake Victoria. *Parasitology* 2010, **137**:1937–1949.

20. Standley CJ, Wade C, Stothard JR: A fresh insight into the transmission of schistosomiasis: A misleading tale of *Biomphalaria* in Lake Victoria. *PLoS One* 2011, **6**:e26563.

21. Asch HL: Rhythmic emergence of *Schistosoma mansoni* cercariae from *Biomphalaria glabrata*: Control by illumination. *Exp Parasitol* 1972, **31**:350–355.

22. Stothard JR, Rollinson D: Partial DNA sequences from the mitochondrial cytochrome oxidase subunit I (COI) gene can differentiate the intermediate snail hosts *Bulinus globosus* and *B. nasutus* (Gastropoda: Planorbidae). *J Nat Hist* 1997, **31**:727–737.

23. Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R: DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol* 1994, **3**:294–299.

24. Bonnaud L, Boucher-Rodoni R, Monnerot M: Phylogeny of decapod cephalopods based on partial 16S rDNA nucleotide sequences. *C R Acad Sci III* 1994, **317**:581–588.

25. Charbonnel N, Angers B, Razatavonjizay R, Bremond P, Jarne P: Microsatellite variation in the freshwater snail *Biomphalaria pfeifferi. Mol Ecol* 2000, **9**:1006–1007.

26. Maddison DR, Maddison WP: MacClade 4: Analysis of phylogeny and character evolution. 2005, http://macclade.org/downloads/ReadMe.pdf.

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

28. Swofford DL: PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods). In Version 4. Sunderland, Massachusetts: Sinauer Associates; 2002.

29. Bandelt HJ, Forster P, Röhl A: Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 1999, **16**:37–48.

30. Raymond M, Rousset F: Genepop (Version-1.2) - Population-Genetics Software for Exact Tests and Ecumenicism. *J Hered* 1995, **86**:248–249.

31. Cornuet JM, Luikart G: Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* 1997, **144**:2001–2014.

32. Caldeira RL, Jannotti-Passos LK, Carvalho OS: Molecular epidemiology of Brazilian *Biomphalaria*: a review of the identification of species and the detection of infected snails. *Acta Trop* 2009, **111**:1–6.

33. Born-Torrijos A, Poulin R, Raga JA, Holzer AS: Estimating trematode prevalence in snail hosts using a single-step duplex PCR: how badly does cercarial shedding underestimate infection rates? *Parasit Vectors* 2014, **7**:243.

34. Ittiprasert W, Miller A, Su X, Mu J, Bhusudsawang G, Ukoskit K, Knight M: Identification and characterisation of functional expressed sequence tags-derived simple sequence repeat (eSSR) markers for genetic linkage mapping of *Schistosoma mansoni* juvenile resistance and susceptibility loci in *Biomphalaria glabrata. Int J Parasitol* 2013, **43**:669–677.

35. Blouin MS, Bonner KM, Cooper B, Amarasinghe V, O'Donnell RP, Bayne CJ: Three genes involved in the oxidative burst are closely linked in the genome of the snail, *Biomphalaria glabrata. Int J Parasitol* 2013, **43**:51–55.

36. Sire C, Durand P, Pointier JP, Theron A: Genetic diversity of *Schistosoma mansoni* within and among individual hosts (*Rattus rattus*): infrapopulation differentiation at microspatial scale. *Int J Parasitol* 2001, **31**:1609–1616.

37. Dybdahl MF, Lively CM: The geography of coevolution: comparative population structures for a snail and its trematode parasite. *Evolution (N Y)* 1996, **50**:2264–2275.

38. Bandoni SM, Mulvey M, Koech DK, Loker ES: Genetic structure of Kenyan populations of *Biomphalaria pfeifferi* (Gastropoda, Planorbidae). *J Molluscan Stud* 1990, **56**:383–391.

39. Campbell G, Noble LR, Rollinson D, Southgate VR, Webster JP, Jones CS: Low genetic diversity in a snail intermediate host (*Biomphalaria pfeifferi* Krass, 1848) and schistosomiasis transmission in the Senegal River Basin. *Mol Ecol* 2010, **19**:241–256.

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