

## RESEARCH ARTICLE

# Genome-wide analysis of *Schistosoma mansoni* reveals limited population structure and possible praziquantel drug selection pressure within Ugandan hot-spot communities

Tushabe John Vianney<sup>1,2</sup>, Duncan J. Berger<sup>1</sup>, Stephen R. Doyle<sup>1</sup>, Geetha Sankaranarayanan<sup>1</sup>, Joel Serubanja<sup>2</sup>, Prossy Kabuubi Nakawungu<sup>2</sup>, Fred Besigye<sup>2</sup>, Richard E. Sanya<sup>2,3</sup>, Nancy Holroyd<sup>1</sup>, Fiona Allan<sup>4</sup>, Emily L. Webb<sup>5</sup>, Alison M. Elliott<sup>2,6\*</sup>, Matthew Berriman<sup>1<sup>aa</sup>\*, James A. Cotton<sup>1<sup>ab</sup>\*</sup></sup>

**1** Parasites and Microbes, Wellcome Sanger Institute, Hinxton, United Kingdom, **2** Immunomodulation and Vaccines Programme, Medical Research Council/Uganda Virus Research Institute and the London School of Hygiene & Tropical Medicine Uganda Research Unit, Entebbe, Uganda, **3** Health and Systems for Health Unit, African Population and Health Research Center, Nairobi, Kenya, **4** Department of Life Sciences, Natural History Museum, London, United Kingdom, **5** MRC International Statistics and Epidemiology Group, London School of Hygiene and Tropical Medicine, London, United Kingdom, **6** Department of Clinical Research, London School of Hygiene and Tropical Medicine, London, United Kingdom

<sup>aa</sup> Current address: Wellcome Center for Integrative Parasitology, School of Infection and Immunity, University of Glasgow, Glasgow, United Kingdom

<sup>ab</sup> Current address: Wellcome Center for Integrative Parasitology, School of Biodiversity, One Health and Veterinary Medicine, University of Glasgow, Glasgow, United Kingdom

\* [Alison.Elliott@shhtm.ac.uk](mailto:Alison.Elliott@shhtm.ac.uk) (AME); [matt.berriman@glasgow.ac.uk](mailto:matt.berriman@glasgow.ac.uk) (MB); [james.cotton@glasgow.ac.uk](mailto:james.cotton@glasgow.ac.uk) (JAC)



## OPEN ACCESS

**Citation:** Vianney TJ, Berger DJ, Doyle SR, Sankaranarayanan G, Serubanja J, Nakawungu PK, et al. (2022) Genome-wide analysis of *Schistosoma mansoni* reveals limited population structure and possible praziquantel drug selection pressure within Ugandan hot-spot communities. PLoS Negl Trop Dis 16(8): e0010188. <https://doi.org/10.1371/journal.pntd.0010188>

**Editor:** María-Gloria Basáñez, Imperial College London, Faculty of Medicine, School of Public Health, UNITED KINGDOM

**Received:** January 31, 2022

**Accepted:** July 5, 2022

**Published:** August 18, 2022

**Copyright:** © 2022 Vianney et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All sequence data are available from the European Nucleotide Archive, with accession numbers listed in [S1 Table](#). All other metadata is also included in the same table. Code is available at [https://github.com/jacotton/LaVIISWA\\_genomes](https://github.com/jacotton/LaVIISWA_genomes). The Lake Victoria Island Intervention Study on Worms and Allergy-related diseases (LaVIISWA) trial is registered in the ISRCTN repository with identifier

## Abstract

Populations within schistosomiasis control areas, especially those in Africa, are recommended to receive regular mass drug administration (MDA) with praziquantel (PZQ) as the main strategy for controlling the disease. The impact of PZQ treatment on schistosome genetics remains poorly understood, and is limited by a lack of high-resolution genetic data on the population structure of parasites within these control areas. We generated whole-genome sequence data from 174 individual miracidia collected from both children and adults from fishing communities on islands in Lake Victoria in Uganda that had received either annual or quarterly MDA with PZQ over four years, including samples collected immediately before and four weeks after treatment. Genome variation within and between samples was characterised and we investigated genomic signatures of natural selection acting on these populations that could be due to PZQ treatment. The parasite population on these islands was more diverse than found in nearby villages on the lake shore. We saw little or no genetic differentiation between villages, or between the groups of villages with different treatment intensity, but slightly higher genetic diversity within the pre-treatment compared to post-treatment parasite populations. We identified classes of genes significantly enriched within regions of the genome with evidence of recent positive selection among post-treatment and intensively treated parasite populations. The differential selection observed in post-

ISRCTN47196031 (<https://doi.org/10.1186/ISRCTN47196031>).

**Funding:** This work was funded by the Wellcome Trust via grants 206194 (MB) and 095778 (AME). <https://wellcome.org/>. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

treatment and pre-treatment parasite populations could be linked to any reduced susceptibility of parasites to praziquantel treatment.

## Author summary

Schistosomiasis is caused by parasitic helminths of the genus *Schistosoma*. *Schistosoma mansoni* is the primary cause of intestinal schistosomiasis, a devastating and widespread parasitic infection that causes morbidity, death and socio-economic impact on endemic communities across the world and especially sub-Saharan Africa. Using whole-genome sequencing, we were able to elucidate the parasite population within Lake Victoria island fishing communities in Uganda which are among the major hotspots for schistosomiasis. We further assessed genetic markers that might be linked to recent observations concerning reduced susceptibility to praziquantel, the major drug used in the treatment of this disease. Whole-genome data on the population genetics of *S. mansoni* in an African setting will provide a strong basis for future functional genomics or transcriptomic studies that will be key to identifying drug targets, improving existing drugs or developing new therapeutic interventions.

## Introduction

Schistosomiasis—also known as Bilharzia after its discoverer Theodor Bilharz [1]—is a neglected tropical disease that affects about 250 million people worldwide, most of whom live in low and middle-income countries (LMICs) [2]. To treat schistosomiasis, praziquantel (PZQ) is used for preventative chemotherapy by mass drug administration (MDA) [3] and has been used globally to treat schistosome infections since 1979 [4]. In Uganda, the ongoing use of PZQ in MDA started between 2002 and 2003 [3, 5]. The objective of MDA in these settings has historically been to reduce the prevalence and intensity of infection and hence pathology; cure and elimination are not expected in the absence of additional interventions such as improving sanitation and snail control [6, 7]. In the World Health Organisation 2021–2030 the goal has been set of reducing the proportion of people with high-intensity infections to < 1% and thereby to eliminate schistosomiasis as a public health problem in all countries in sub-Saharan Africa by 2030 [8]. The expectation is that this will be achieved primarily by increasing the frequency and coverage of treatment with PZQ—the sole drug commonly used for schistosomiasis MDA—which could inadvertently increase drug selection pressure on parasite populations.

There is a growing body of evidence that MDA programmes may affect how parasite populations respond to treatment, for example, through reduced efficacy of PZQ in lowering egg output in communities that have received multiple rounds of PZQ MDA [9, 10], but there is little evidence that this is a widespread phenomenon [11, 12]. Reduced genetic diversity of parasite populations has also been associated with reduced susceptibility of the parasites to PZQ [13], with reports from Senegal having earlier linked such outcomes to potential drug resistance [14]. The development of drug resistance in natural populations would be a major health concern. Furthermore, *in vitro* studies have shown that resistance to PZQ can be selected for in *S. mansoni* [13, 15–17]. There is growing interest in understanding the impact of continued PZQ monotherapy on the parasite genome in order to detect the potential development of resistance to this drug as early as possible [18, 19], and understand the mechanism(s) of resistance. One clue to resistance can come through understanding the mode of action of a drug.

The activity of PZQ has not been clearly understood, but recent findings suggest that the drug activates a schistosome  $\text{Ca}^{2+}$ -permeable transient receptor potential (TRP) channel (Sm. TRPMPZQ) [20], hence making it the primary target for PZQ action on schistosomes. Recently, a genetic cross involving a schistosome line experimentally selected for PZQ resistance identified this TRP channel as likely responsible [21], but it is not yet clear that this locus is involved in variation in PZQ efficacy in the field. Other candidate genes have been proposed, for example the *S. mansoni* P-glycoprotein (*smdr2*), which shows increased protein expression in male worms following exposure to sub-lethal doses of PZQ [16]. Susceptibility of the parasites to PZQ might involve multiple interactions between the drug, the parasite, and the respective host.

Collecting high-resolution genetic data from parasite populations under drug selection pressure may lead to new insights into the mode of action of PZQ or the mechanism of potential resistance to the drug. Furthermore, population genetic data from parasite populations will also give insights into the population biology of the parasite. This is vital for understanding schistosomiasis epidemiology, transmission, disease severity and why certain communities might respond better to treatment than others, especially within regions where drug selection pressure is being applied [22]. While lower-resolution markers have been extensively used (e.g. [23, 24]), much of our detailed understanding of schistosome population genetics has come from studies using microsatellite markers to describe the genetic structure of populations of *S. mansoni* [25–27] and other schistosome species [28, 29]. This work has revealed genetic differentiation between parasite populations that are geographically separated (e.g. [30–33]), but panmictic populations and very high within-host diversity within disease foci (e.g. [34, 35]). The population genetics of African schistosomes has recently been reviewed [36]. Microsatellite markers have also been employed to investigate both basic questions about parasite biology (e.g. [37]) as well as more applied, operational questions about schistosome control [22]. In particular, a few studies have shown changes in genetic diversity of schistosomes with praziquantel MDA [4, 38, 39], but other studies have failed to find this effect [40] particularly with longer-term follow-up [41] suggesting any genetic response to treatment may be only temporary [42].

With their high levels of polymorphism, microsatellite loci are powerful molecular markers, but inevitably represent only a small proportion of the parasite genome. There is an increasing amount of genome-scale data available for schistosome populations. A number of studies have used exome capture [43] to describe introgression between *Schistosoma* species [44] and to study the historical demography of schistosomiasis in the Americas [45]. Restriction site-associated sequencing (ddRAD-seq) has been used to demonstrate strong genetic structure in remaining endemic hot-spots of *S. japonicum* transmission [46, 47]. While providing high-resolution data in a cost-effective way, these reduced-representation sequencing approaches have some drawbacks, for example in identifying small haplotype blocks from ancient introgression [44]. Whole-genome data gives a more comprehensive picture of genetic variation, including non-coding variation, and so has the potential to provide more insights into understanding the population genetics of this species. While reference genome assemblies are available for a number of schistosome species [48–54], large-scale genome-wide variation data is only available from one *S. mansoni* population [55], with a number of other populations and other species most being represented by relatively few specimens [53, 56–58]. Efforts in elucidating the parasite population genetic structure have proven very helpful in understanding drug resistance or transmission mechanisms in other parasite species: most notably in the malaria parasite *Plasmodium falciparum* [59, 60] for which very extensive genome data is available [61].

Within Uganda, the Lake Victoria Island Intervention Study on Worms and Allergy-related diseases (LaVIISWA) was a cluster-randomised clinical trial [62] examining the impact of

intensive (quarterly) versus standard (annual) PZQ treatment. While the study was primarily designed to assess the impact of anthelmintic treatments on allergy-related outcomes, prevalence and intensity of *S. mansoni* was a secondary outcome with results suggesting a plateauing of infection, after an initial decline in intensively treated villages [63]. To assess this outcome, a pilot study in the fourth year of the LaVIISWA trial investigated cure rate and Egg Reduction Rate (ERR) [10]. A lower cure rate and ERR was seen among people receiving quarterly (intensive) treatment ( $n = 61$ ; cure rate 50.8%, 95% confidence interval (CI): 37.7% to 63.9%; ERR 80.6%, 95% CI: 43.8% to 93.7%) than in those receiving a single annual standard dose ( $n = 49$ , cure rate 65.3%, 95% CI: 50.4% to 78.3%; ERR 93.7%, 95% CI: 84.9% to 97.7%) [10]. The WHO recommends an ERR of 90% for effective PZQ treatment [9, 64]. While the sample size available precluded finding compelling statistical evidence, these results are suggestive of the first signs of reduced efficacy of PZQ treatment in the more intensively treated population, and that the plateau in reduction of infection during the intervention study could be due to PZQ resistance. These islands thus represent a 'hot spot' in which high baseline prevalence [62] of schistosomiasis has persisted despite multiple years of treatment [10, 65].

Here, we sought to establish genome-wide data on the population genetics of parasites present in this study population, with the ultimate goal of assessing the effects of MDA on parasite genome evolution. We take advantage of the opportunity to investigate these in the context of a randomised intervention trial within a defined geographical area, allowing us to compare the effects of geographical isolation and treatment intensity on genetic variation in this population. By comparing samples taken immediately before and after a treatment round at the end of the LaVIISWA study in the two treatment arms and for multiple villages (the level of randomisation in the study), we can investigate whether the genetic impact of a single treatment dose varies with history of drug exposure. Building on the evidence that there may be differences in treatment efficacy between treatment arms, we investigated whether the signatures of natural selection across the genome differ with previous drug exposure. We also compare these data with recently published genomic data from other Ugandan *S. mansoni* populations.

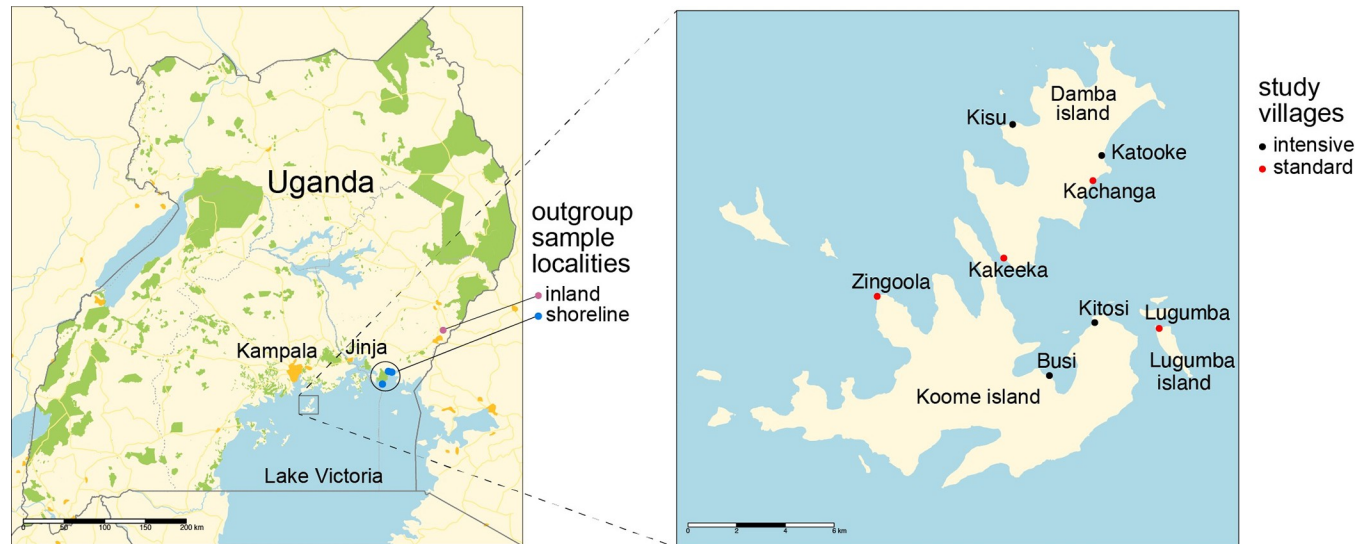
## Methods

### Ethics statement

This work was not expected to result in any harm to participants. Ethical approval was given by the Uganda Virus Research Institute (reference number GC127), the Uganda National Council for Science and Technology (reference number HS 1183) and the London School of Hygiene & Tropical Medicine (reference number 6187). As previously detailed [62], written informed consent was received from all adults and emancipated minors and from parents or guardians for children; additional assent was obtained from children aged  $\geq 8$  years.

### Sample selection and study site

Participants were selected from four villages each from the standard and intensive treatment arms from among the 27 study villages of Lake Victoria Island Intervention Study on Worms and Allergy-related diseases (LaVIISWA) trial [62, 63] at the end of its fourth year. The participants involved children and adults as previously described [62]. The LaVIISWA trial was registered as ISRCTN47196031 (<https://trialssearch.who.int/?TrialID=ISRCTN47196031>), and the results of the trial have already been reported elsewhere [63, 66]. The villages in the standard arm received PZQ once a year while those in the intensive arm received PZQ four times a year during the LaVIISWA trial period. The standard villages sampled were Kakeeka, Kachanga, Zingoola and Lugumba. The intensive villages were Busi, Kitosi, Kisu and Katooke (Fig 1). The map in Fig 1 was prepared using the R packages tmap v.3.3 [67] and sf v 1.0 [68] using



**Fig 1. Location of sample sites within Uganda.** Villages with black dots received standard (annual) intervention, those with red dots received intensive (quarterly) intervention. Outgroup samples were obtained from locations marked as inland and shoreline. Map data is from <https://data.humdata.org/dataset/cod-ab-uga>.

<https://doi.org/10.1371/journal.pntd.0010188.g001>

administrative boundaries from OCHA Eastern region from <https://data.humdata.org/> and public domain data on roads (v5.0.0), rivers (v5.0.0), lakes (v5.0.0) and urban areas (v4.0.0) from <https://www.naturalearthdata.com>.

Sample selection and collection was carried out as previously described in the parasitological survey [10]. The stool samples (collected from participants who tested positive for urine CCA) were processed for two Kato Katz slides as previously described [10] and miracidia hatching provided suitable material for DNA extraction. Participants were then treated, under observation, with a single dose of PZQ at 40 mg/kg (estimated by height pole), in accordance with the trial MDA procedures. Individuals whose pre-treatment sample tested positive for schistosome eggs by Kato Katz were followed up after four weeks and both Kato Katz and miracidia hatching were repeated. Miracidia hatching was carried out from each of these participants and the resultant miracidia were stored on Whatman FTA cards until DNA was extracted.

### Miracidia hatching

Miracidia hatching was carried out following previously described protocols [31]. In brief, the stool sample was homogenised through a metal sieve, then further washed and filtered using a Pitchford funnel assembly [69] consisting of a 40  $\mu\text{m}$  sieve placed inside a 200  $\mu\text{m}$  outer sieve. Stool samples were washed using deionised water (Rwenzori Bottling Company, Uganda). The concentrated *S. mansoni* eggs were transferred to a Petri dish in clean water and exposed to indirect sunlight to induce the hatching of miracidia. Hatching was performed in natural light (environmental conditions) with intervals of exposure to sunlight and cover depending on weather conditions. The time taken for miracidia to emerge varied between samples, so the Petri dishes were intermittently checked for the presence of miracidia for a maximum of 48 hours. Miracidia were picked in 1.5–5  $\mu\text{l}$  of water and then transferred to a second dish of deionised water to dilute bacterial contamination before being placed on Whatman indicating FTA cards (Qiagen) and left to dry. The FTA cards were wrapped in aluminium foil to keep

them away from continued direct light and placed in ziplock bags with silica gel in a cardboard drawer.

### Whatman FTA DNA Extraction

DNA was extracted using a modified CGP buffer protocol as previously described [70, 71]. The individual spots containing miracidia were punched from the FTA cards using a 2 mm Harris micro-punch and placed in 96-well plates. Protease buffer was prepared using Tris-HCl pH 8.0 (30 mM), Tween 20 (0.5%), IGEPAL CA-630 (0.5%), protease (1.25 µg/ml; Qiagen cat #19155) and water. Digestion was done by adding 32 µl of the protease buffer to each of the wells on the 96-well plate containing the punched spots from the FTA cards. The plate was vortexed to mix and spun down before incubation at 50°C for 60 min, 75°C for 30 min. Miracidia lysates containing DNA were transferred to a new labelled plate and stored at 4°C until used.

### Library preparation and sequencing

DNA sequencing libraries were prepared using a protocol designed for library preparation of Laser Capture Microdissected Biopsy (LCMB) samples using the Ultra II FS enzyme (New England Biolabs) for DNA fragmentation as previously described [71]. The LCMB library preparation method is optimised for uniform, low-input samples. A total of 12 cycles of PCR were used to amplify libraries and to add a unique 8-base index sequence for sample multiplexing. The LCMB library preparation protocol is optimised for uniform, low input samples. A total of 174 samples from 96 individuals in 8 villages were sequenced on two NovaSeq lanes, 108 on one lane and 66 on another lane. These 174 samples were chosen as having more than 10% of reads mapping to *S. mansoni* based on preliminary low-coverage genome sequencing of all 214 samples collected in the field.

### Mapping and SNP calling

The reads were mapped to the *S. mansoni* reference genome v7 (GCA\_000237925.3) [54] using the BWA-MEM algorithm in Burrows-Wheel Aligner software (BWA) (VN:0.7.15-r1140) to produce SAM files which were then converted to BAM format using Samtools v1.14. Scaffolds labelled as haplotype variants in the assembly were removed before read mapping to improve mapping accuracy, using the approach previously described [55]; see [55] also for a full list of removed scaffolds. PCR duplicate reads were identified using Picard v1.92 [72] and flagged as duplicates in the BAM file.

SNP variants were called using the GATK Haplotype Caller (v4.1.4.1) to find sites that differ from the *S. mansoni* reference genome followed by variant QC to remove low confidence SNPs and regions of consistently poor calls. The SNPs were hard-filtered in GATK to remove SNP calls with the following parameters: QD < 2.0; MQ < 40; FS > 60.0; SOR > 3.0; MQRankSum < -12.5; ReadPosRankSum < -8.0. The variants were further filtered using vcftools\_0.1.15 [73] to remove sites with high missingness (—max-missing 0.95), low minor allele frequency (—maf 0.01) and to retain only biallelic SNPs (—min-alleles 2 —max-alleles 2).

### Identification of population structure

The three islands on which the population structure was assessed were Koome, Damba and Lugumba in the Mukono district of Uganda. An outgroup made up of inland and shoreline samples was also included, consisting of 27 samples collected in a previous study [9] and for which whole-genome sequence data were recently published [55] from Tororo and Mayuge districts in Eastern Uganda. Tororo and Mayuge are approximately 120 km apart. Mayuge

district is a shoreline district located about 100 km from Mukono district. Both districts are located in south eastern Uganda, with Tororo being the inland district (Fig 1).

### Test for genetic differentiation

The fixation index ( $F_{ST}$ ) statistic was calculated between each of the villages across the different islands and treatment groups (standard, intensive, pre-treatment and post-treatment) to measure population differentiation due to genetic structure. The  $F_{ST}$  was calculated using vcfTools (version 0.1.15) [55] on the vcf file containing biallelic filtered SNPs. Mean  $F_{ST}$  was calculated from genome-wide weighted  $F_{ST}$  values with 99% symmetric bootstrap confidence intervals calculated using R version 3.5.1 (2018-07-02). We fitted a gravity model as

$$\log(N_m) = \log(g) + a \cdot \log(P_i \cdot P_j) - \gamma \cdot \log(D_{ij}) + \epsilon,$$

where  $N_m$  is an estimated number of migrants per generation, calculated from the  $F_{ST}$  between villages as:

$$N_m = 0.25((1/F_{ST}) - 1)$$

And  $D_{ij}$  is the linear distance between the villages. In  $(P_i \cdot P_j)$ ,  $P_i$  and  $P_j$  represent the population sizes of the two villages compared. Models were fitted using R version 4.0.2, with the MuMIn package v1.43 to assess model importance.

### Nucleotide diversity

Nucleotide diversity ( $\pi$ ,  $\pi$ ) was computed from high-confidence bi-allelic filtered SNPs using vcfTools 0.1.15 [73]. The genome-wide nucleotide diversity was calculated from a list of positions for each of the time points (pre- and post) and treatment groups (standard and intensive) using the option in vcfTools ‘—site-pi’ respectively. The average nucleotide diversity within each of the groups was calculated individually and the symmetric 99% bootstrap confidence intervals of the averages were estimated using R version 3.5.1. Statistical significance of differences between group means was assessed by whether the confidence interval for one mean was disjoint from the mean of other groups. Effective population size ( $N_e$ ) was estimated from nucleotide diversity using the relationship  $\pi = 4 \cdot N_e \cdot \mu$  [74] with the mutation rate  $8.1 \times 10^{-9}$  [57].

### Determination of rare allele sharing and kinship analysis

To identify the pairwise rare allele sharing we used a Perl script from Shortt et al. [47] available at <https://github.com/PollockLaboratory/Schisto>. We filtered for minor allele frequency  $\leq 0.1$  and sampled 500 SNP sites in 30 different generations. We then computed the mean value from the 30 generations for each pair. Allele-sharing scores were visualised in R version 4.0.2 using igraph v1.2.6 [75]. Significance of differences in mean allele sharing between groups were calculated against a non-parametric null distribution for each comparison generated by randomly permuting group labels 1000 times and calculating differences in mean allele sharing for each permutation. We also inferred relationships between miracidia samples using Sequoia version 2.3.5 [76] in R 4.0.2, using the GetMaybeRel command to identify related pairs with both the default genotyping error rate of  $1 \times 10^{-4}$  and a higher error rate of 0.05. Sequoia uses a computationally intensive likelihood procedure that limits the number of SNP variants that can be analysed, so we used a heavily thinned dataset of 3,829 variants such that all variants had allele frequencies of at least 0.1, were called in 80% or more of the samples and were at least 100,000 bp apart.

## Test for selection

To test for recent positive selection within the treatment arms and between pre-treatment and post-treatment, the cross-population extended haplotype homozygosity (XP-EHH) test [77] was performed. XP-EHH is designed to detect whether either an ancestral or derived allele is undergoing selection within a given population. The XP-EHH test has the power to detect weaker signals of selection as it compares two closely related populations giving a directional score. The XP-EHH detects selective sweeps in which the selected allele has approached fixation in one population but remained polymorphic in another population. A VCF file containing only bi-allelic SNPs was subset into respective chromosomes. A genomic linkage map for each of the chromosomes was computed for each individual chromosome using the adjusted map length in centimorgan (cM) for the respective chromosomes [78]. The haplotypes from each of the chromosomes were then phased separately with their respective genomic map using Beagle v5.0 [79]. The XP-EHH test was performed using Selscan v1.2.0a [80] and the output XP-EHH scores were normalised for subsequent analysis using the norm program distributed with v1.2.0a of Selscan. Functional enrichment was assessed using g:Profiler version (e99\_eg46\_p14\_f929183) [81] at a g:SCS threshold of 0.05 against a background of all annotated genes in *S. mansoni*, revealing genes showing significant purifying selection among the intensive and post-treatment parasite populations.

## Estimate of per-individual egg reduction rate (ERR) and association test

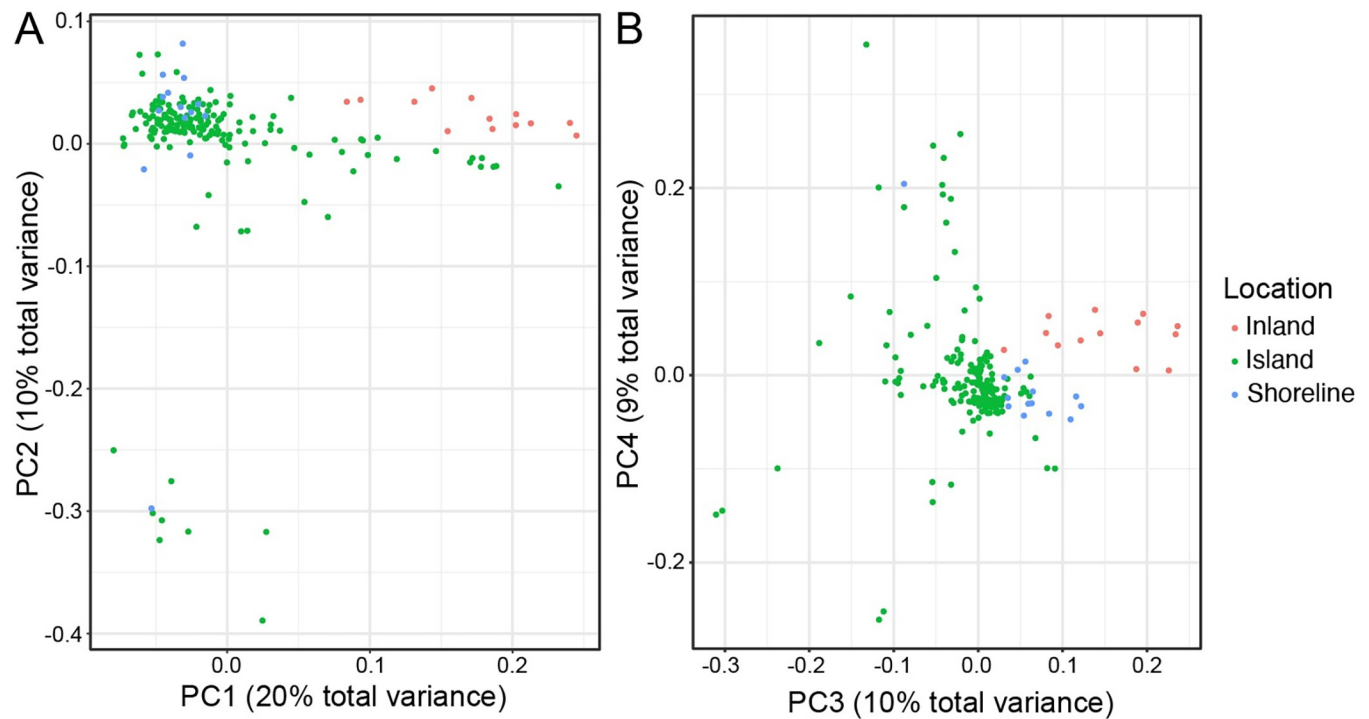
The posterior distribution for the ERR based on data from each individual for whom both pre- and post-treatment egg count data were available was estimated using a generalised linear mixed-effect model [82], incorporating nested random effects for treatment arm, village and individual. Means of the marginal posterior distribution per individual were used as quantitative phenotypes for an association study, in which a single miracidium was chosen at random for each individual for which an ERR was calculated ( $N = 44$ ) testing all 6,932,161 SNP sites that varied in this subset. The model used was a linear regression of each SNP genotype against mean ERR, using 20 principal components as covariates to control for population structure, calculated using the '—linear' and '—pca' flags in plink v1.9 [83]. Code for these analyses is available at [https://github.com/jacotton/LaVIISWA\\_genomes](https://github.com/jacotton/LaVIISWA_genomes).

## Results

### Population stratification

After filtering, 6,967,554 high-confidence SNPs were retained (out of 18,716,072 unfiltered SNPs) from 174 individual miracidia. Sample metadata and genetic variation data is available to view and analyse in an interactive web application [84] at <https://panoptes.schisto.watch/index.html>. Principal components analysis of these high confidence SNPs showed little genetic structure within the island parasite populations on the first four principal components, which together represent 49% of the genetic variation. In particular, we found no evidence of population stratification between the standard (annual treatment) and intensive arms (quarterly treatment) or between pre- and post-treatment samples from these fishing communities (S2 Fig). The shoreline samples (from Mayuge district) clustered more closely with the island (Mukono district) parasite populations as compared to the inland samples (from Tororo district) (Fig 2), but inland parasites were distinct from most island samples on principal component 3. The large-scale geographical pattern reflects the known genetic differentiation between inland and shoreline populations [55]. We also found that the island population is strikingly





**Fig 2. Principal components analysis of genetic variation within study samples and comparator Ugandan populations.** (A) Shows the first two principal components and (B) the third and fourth principal components. Each point represents a single miracidium, coloured by the population from which they are sampled, with 'Shoreline' samples from Mayuge district and 'Inland' from Tororo district.

<https://doi.org/10.1371/journal.pntd.0010188.g002>

more diverse than either of the other populations (Fig 2). While this is partly due to the larger number of samples included here, a larger sample of the shoreline and inland populations studied elsewhere also did not appear as diverse as the island population [55]. A number of miracidia appeared quite distinct from the main cluster of individuals on principal component 2. These divergent parasites were mostly (8 out of 9) from the islands and came from four different villages (Busi, Kakeeka, Zingoola, Katooke), with one from a shoreline village (Bwondha; S4 Fig). Although participants were all resident in the villages throughout the LaVIISWA trial for at least 3 years before this study, there is a great deal of migration to the islands from other parts of the shoreline of Lake Victoria, including Kenya and Tanzania; therefore, we suspect these miracidia represent parasites imported from other populations that we have not sampled here.

### Rare allele sharing and kinship analysis

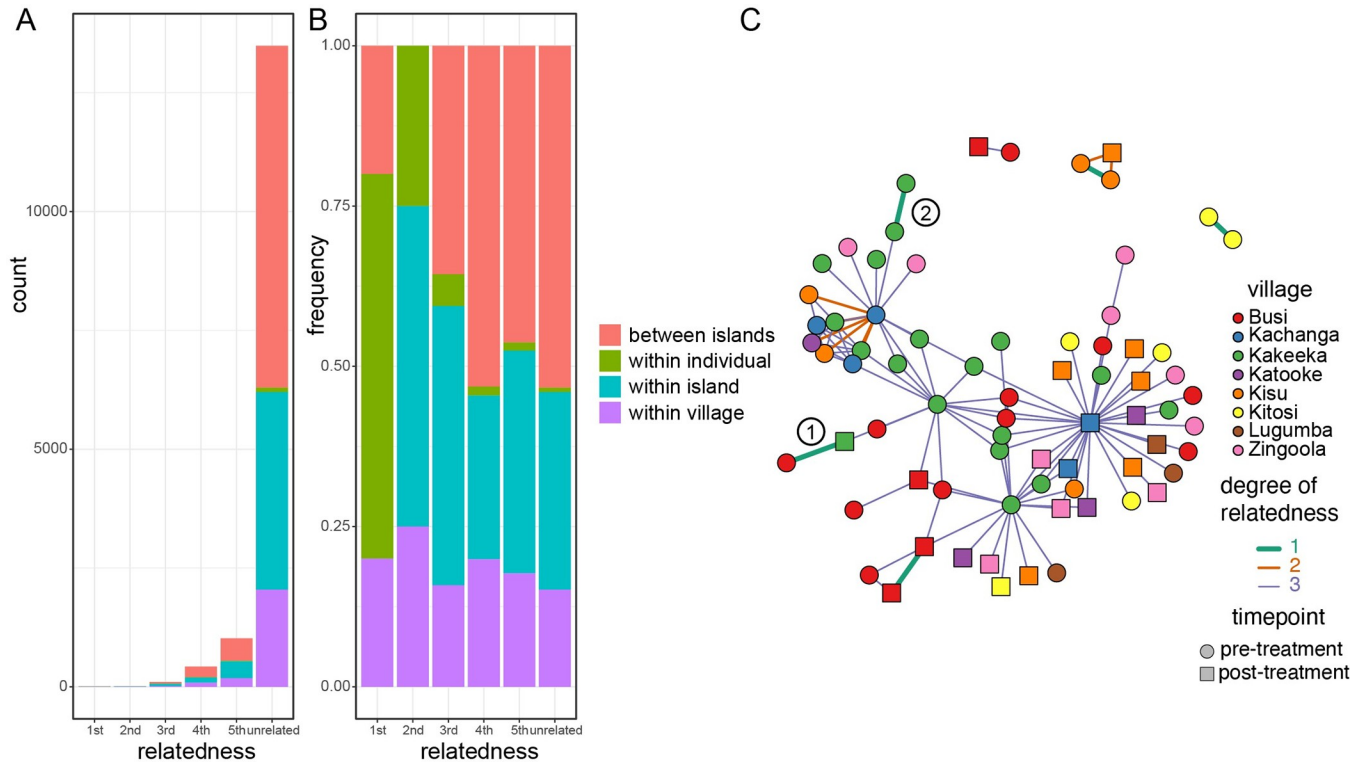
To investigate direct relatedness between individual parasites, we adopted an approach based on determining the level of sharing of rare alleles (defined by their population frequency being less than or equal to 10%) between samples [47]. This approach has recently been used to study *S. japonicum* populations in China with whole-genome data [58]. By definition, most unrelated individuals share very few rare alleles; here we found slightly higher average proportion of rare-allele sharing between pairs of miracidia isolated from the same individuals (0.1028) than in other comparisons (from the same village 0.0874, between villages on the same island 0.0862, between islands 0.0856). Differences between average allele sharing proportions were significant for comparing intrapopulation and within village groups (observed difference 0.0154, p-value from permutation test  $p < 0.001$ ) and within-village to within-island

groups (observed difference 0.0012,  $p \sim 0.002$ ) but only marginally so for within-island and between-island comparisons (observed difference 0.0006,  $p \sim 0.011$ ). These data suggest an increase in relatedness within populations and possibly some geographical signature of increased relatedness.

We found three miracidia collected from the same infected individual at the same time with rare allele sharing of at least 0.3, and used these to calculate the average allele sharing for first-degree relatives (full siblings or parent-offspring) of 0.403 (actual values 0.4, 0.4014 and 0.409). This value is slightly lower than the theoretically expected level of identity-by-descent of 0.5, but with only three observations it is not possible to exclude that this difference is due to chance. Similarly, the average rare allele sharing for pairs of miracidia from different islands was 0.086, which represents our best estimate for the level of sharing in unrelated individuals. The small number of miracidia available for putative first-degree relatives means the observed variance from our data is very small, and so our final classification is thus deterministic. We classified miracidia pairs sharing more than 0.3105 of these rare alleles as first-degree relatives, 0.1981–0.3104 as second-degree relatives, 0.1419–0.1980 as third-degree relatives, 0.1138–0.1418 as fourth-degree relatives and 0.0956–0.1138 as fifth-degree relative, while those with less than 0.0956 sharing were classified as unrelated.

While 24% of pairs of samples from the same individual were classified as being related, only 10% of other comparisons appeared related (Fig 3A and 3B;  $\chi^2 = 21.785$ , 1 df,  $p = 3.05 \times 10^{-6}$ ), and a similar pattern held for close relatives (Fig 3A and 3B; first and second degree relatives represented 4% of within-infrapopulation comparisons, but 0.05% of all comparisons;  $\chi^2 = 183.37$ , 1 df,  $p < 2.2 \times 10^{-6}$ ). There was no significant enrichment in related pairs of miracidia with either treatment intensity or for samples collected pre- and post-treatment. We found five pairs of first-degree relatives in total (Fig 3C); but one pair were from different islands (marked 1 on Fig 3C) and a second pair were from different individuals sampled on consecutive days in Kakeeka village (marked 2 on Fig 3C). On the face of it, this would imply that the same combination of clonal cercariae infected these people, which seems very unlikely—particularly for the geographically separated cases. We cannot exclude the possibility that either the high level of rare allele sharing is misleading in these cases, or errors in sample identification. The remaining three pairs of first-degree relatives were pairs of miracidia sampled from single individuals in Busi, Kitosi and Kisu villages on consecutive days. Interestingly, a miracidium sampled from the same individual in Kisu was a second-degree relative of the first-degree pair, but this was collected post-treatment 37 days later. This is one of only 8 pairs of second-degree relatives. This suggested that either an adult worm survived treatment but changed 'partners' (to produce a half-sibling or avuncular relationship) during this period, or two clonal worms with genetically distinct partners were present in this host at the two time-points and produced these miracidia.

We tested the accuracy of these inferred relationship using Sequoia, an independent likelihood-based method [76] that will identify first and second-degree relatedness, using a thinned subset of 3829 variants. With the standard genotyping error rate ( $1 \times 10^{-4}$ ) this approach identified only a small number (3) related miracidia, but with a more liberal error rate of 5% this approach identifies 30 pairs of related miracidia, and shows very good agreement with the inferred relatedness using the allele-sharing approach above. Sequoia identified all 5 first-degree relationships as full sibs, and 6 out of 8 second-degree relationships identified by allele sharing were also inferred to be second-degree relatives by Sequoia. A relatively small fraction of third-degree (15 out of 101) and fourth-degree (4 out of 427) relationship assignments from allele sharing were also assigned as second-degree relationships by Sequoia.

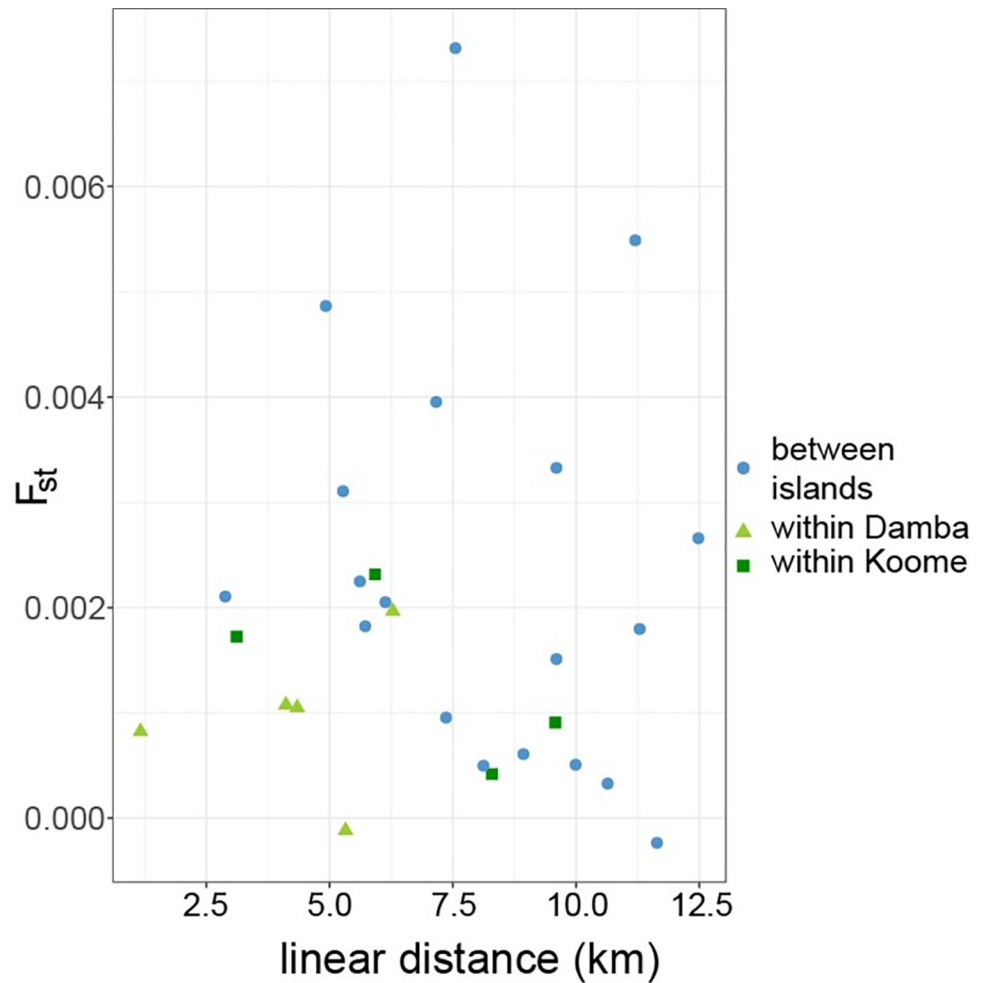


**Fig 3. Patterns of relatedness inferred from pairwise rare allele sharing.** (A) Number and (B) proportion of pairs of miracidia showing each degree of relatedness for miracidia sampled from the same individuals, villages or islands and for those on different islands. (C) Network representation of 1st, 2nd and 3rd degree relatedness. Vertices represent individual miracidia sampled, coloured by village and with a circle for samples taken pre-treatment and square for post-treatment samples. Edges join vertices inferred to share 1st, 2nd or 3rd degree relatedness, as indicated by both the width and colour of each edge. Numerical labels indicate two 1<sup>st</sup> degree relationships discussed in the text.

<https://doi.org/10.1371/journal.pntd.0010188.g003>

### Analysis of genetic differentiation between villages

To further investigate genetic structure within the island population, we calculated  $F_{ST}$  (the proportion of genetic variation explained by population structure) for each pair of villages. As we expected,  $F_{ST}$  between villages was very low (maximum 0.0067), indicating little or no geographic structure to our data. We observed higher genetic differentiation between villages on different islands compared to those within the same island, but the small number of pairwise comparisons ( $N = 8$  villages, 28 pairwise comparisons) meant that we did not have sufficient statistical power to detect any difference ( $p = 0.082$ , 1-way ANOVA of between/within village vs  $F_{ST}$ ). The villages were between 1 and 13 km apart, but there was no significant relationship between the distance between villages and  $F_{ST}$  (Fig 4). To explore the geographical structure in these data more fully, we also fitted a gravity model attempting to explain  $F_{ST}$  between each pair of villages by the distance between villages, the population of each village and a factor capturing the effect of being on the same island. In this model, none of the explanatory variables had a significant influence on  $F_{ST}$ , but the location of villages on the same island vs different islands was the most important variable with a likelihood weight in the best-fitting models of 0.48, while 0.31 for linear distance between villages and 0.23 for the product of village populations.



**Fig 4. Pairwise  $F_{ST}$  estimates do not vary with linear distances between villages.** Weir and Cockerham  $F_{ST}$  estimates used and distance measured in kilometres. Points show results of pairwise comparison between samples from different villages found on different islands or from different villages with Damba or Koome islands.

<https://doi.org/10.1371/journal.pntd.0010188.g004>

### Within-population genetic diversity

When comparing all pre- and post-treatment samples, we observed a very small but significant difference ( $p < 0.01$ ) in genetic diversity between samples taken before and after treatment, with the 99% confidence interval for the mean nucleotide diversity in pre-treatment samples not overlapping with the mean post-treatment nucleotide diversity. This is consistent with a

**Table 1. Genome-wide average nucleotide diversity ( $\pi$ ).**

Group	Average $\pi$	99% confidence interval	
Pre-treatment	$3.25 \times 10^{-3}$	$3.22 \times 10^{-3}$	$3.29 \times 10^{-3}$
Post-treatment	$3.20 \times 10^{-3}$	$3.16 \times 10^{-3}$	$3.23 \times 10^{-3}$
Pre-treatment standard	$3.27 \times 10^{-3}$	$3.23 \times 10^{-3}$	$3.32 \times 10^{-3}$
Pre-treatment intensive	$3.23 \times 10^{-3}$	$3.20 \times 10^{-3}$	$3.26 \times 10^{-3}$
Post-treatment standard	$3.24 \times 10^{-3}$	$3.21 \times 10^{-3}$	$3.27 \times 10^{-3}$
Post-treatment intensive	$3.16 \times 10^{-3}$	$3.12 \times 10^{-3}$	$3.19 \times 10^{-3}$

<https://doi.org/10.1371/journal.pntd.0010188.t001>

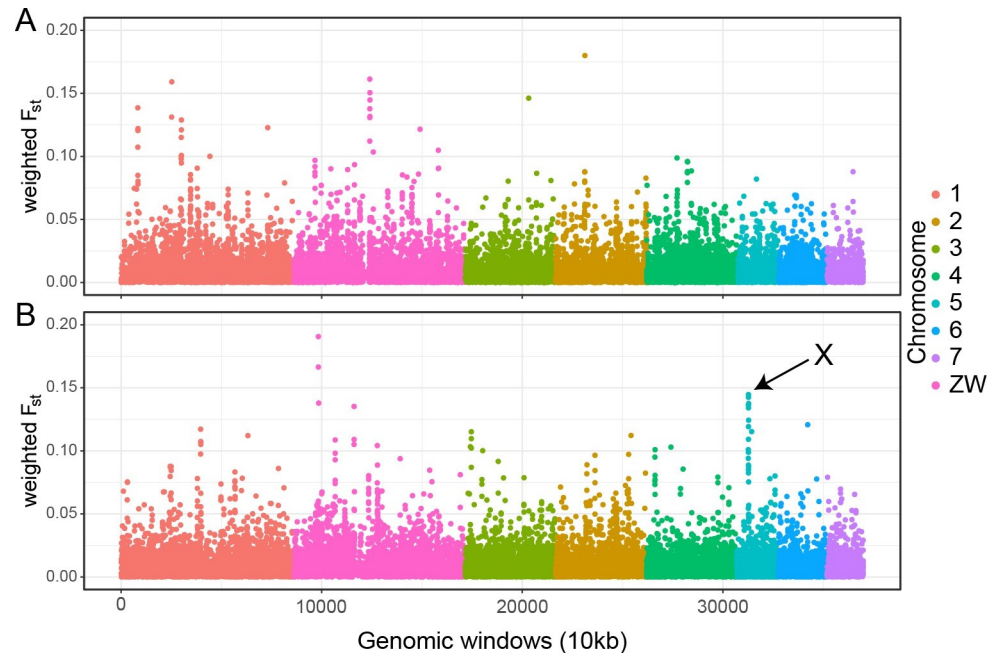
small effect of a single PZQ treatment round on the parasite population (Table 1). There was also lower diversity in parasites collected from villages in the intensive arm of the study than in the standard arm (Table 1), possibly reflecting a longer-term effect of more frequent PZQ treatment in these locations, despite the high levels of gene flow apparent between these locations implied by the very small levels of genetic differentiation we report. While this trend was consistent in both pre- and post-treatment samples, the difference between trial arms was most pronounced in post-treatment populations (Table 1). These diversity values are very similar to those observed in a recent study of the parasite populations on the lake shore and inland sites [55]. Using the mutation rate estimated previously [57], this implies an effective population size of around  $10^5$  individuals from this sample collection, just outside the upper confidence limit of the estimate for the East Africa population in the previous study ( $3.67\text{--}9.35 \times 10^4$ ) [57], and much higher than estimates from individual schools on the Lake Victoria shoreline ( $3.30\text{--}3.69 \times 10^4$ ) [55], highlighting the diversity of *S. mansoni* parasites present on the islands.

### Genetic differentiation with treatment between standard and intensive arms

Genome-wide average genetic differentiation was slightly higher (mean  $F_{ST} = 3.9 \times 10^{-4}$ ; bootstrap 99% CI:  $2.5 \times 10^{-4}\text{--}5.3 \times 10^{-4}$ ) between standard and intensive treatment populations post-treatment than before treatment (mean  $F_{ST} = 3.4 \times 10^{-4}$ ; 99% CI =  $1.8 \times 10^{-4}\text{--}5.0 \times 10^{-4}$ ), but these values did not differ significantly. We also find very low genetic differentiation between standard and intensive trial arms (mean  $F_{ST} = 5.6 \times 10^{-4}$ ; 99% CI =  $5.2 \times 10^{-4}\text{--}6.0 \times 10^{-4}$ ). There was also variation in these  $F_{ST}$  values across the genome. While much of this likely reflects sampling variation (Fig 5A), particularly striking was a region identified on chromosome 5 (Fig 5B) with a distinct peak of divergence among post-treatment parasite populations. This window spanned 1.21 Mb of genomic sequence (from SM\_V7\_5: 7.78–8.99 Mb) and contained 25 annotated protein-coding genes (S2 Table).

### Analysis of signatures of selection

We used the XP-EHH test to identify genomic regions under differing selection pressures in separate comparisons between standard and intensive treatment arms (Fig 6A) and between pre- and post-treatment samples (Fig 6B). Taking extreme XP-EHH scores of  $< -2$  or  $> 2$  as a cutoff, we identified 510 windows as outliers including 12.75 Mb or 3.1% of the genome in total and representing 123 contiguous regions. None of the windows from either comparison overlap the peak of differentiation between standard and intensive treatment populations on chromosome 5. We note that the Z chromosome was particularly enriched for windows with extreme XP-EHH scores, containing almost half of those found genome-wide (5.325 Mb). This could be a technical artefact caused by difficulty in mapping to a highly repetitive chromosome [54], or due to the smaller average population size or a stronger effect of selection on recessive alleles when hemizygous. There are also a number of reasons to expect sex-linked genes to frequently be under selection [85]. An increased variance in XP-EHH scores is apparent specifically in the Z-specific region (Fig 6) of the assembly scaffold representing the Z chromosome [54]. This region is not more repetitive than the autosomes or the pseudo-autosomal region shared by Z and W (see Table S12 of [54]), but is at lower copy number in the population as it is present in a single copy in female worms, so we suspect this enrichment of extreme XP-EHH scores represents a population genetic effect rather than a technical artefact. Some areas of the Z-specific region (6 contiguous windows, total length 300kb) show extreme scores in both comparisons between treatment arms and pre- and post-treatment samples.



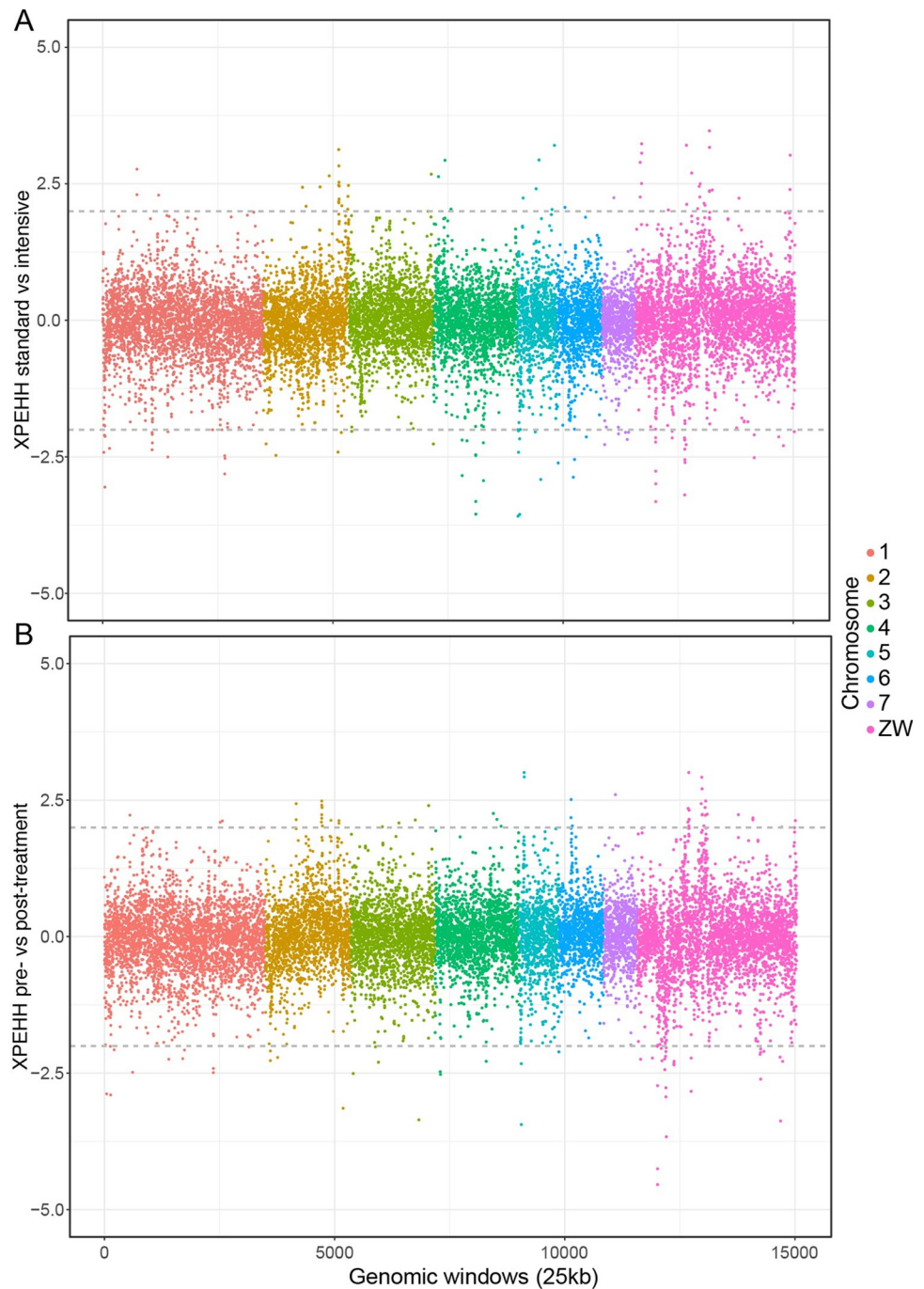
**Fig 5. Genome-wide genetic differentiation between standard and intensive populations.**  $F_{ST}$  calculated using pre-treatment (A) and post-treatment (B) samples. X marks the region of high post-treatment genetic differentiation discussed in the text. Each point represents the mean  $F_{ST}$  between genomic windows of 10 kb for all the called SNPs, with different coloured points representing SNPs on each chromosome.

<https://doi.org/10.1371/journal.pntd.0010188.g005>

There were 107 genes overlapping the outlier windows in the post-treatment samples, which were enriched for genes associated with seventeen GO terms for molecular function and biological processes (S3 Table; <https://biit.cs.ut.ee/gplink/l/wWE3Rp-ASq>). Only 53 of these genes were on autosomes (leaving 54 on the Z and/or W chromosomes), and no GO terms were enriched when considering just the autosomal gene subset. No statistically significant enrichment for any functional category was observed among the genes undergoing stronger selection in pre-treatment individuals. Functional profiling showed that the 132 genes (78 autosomal) under stronger selection in the intensive arm were significantly enriched for association with 10 GO terms (S3 Table; <https://biit.cs.ut.ee/gplink/l/1VaAMWpxQK>), which remained enriched in the autosomal subset. 88 genes were found in 46 autosomal windows with extreme XP-EHH values suggestive of stronger selection in the standard treatment arm include a pair of adjacent closely related genes likely to be a recent tandem duplication and possessing nucleoside deaminase activity on chromosome 4; these genes represent the only significantly enriched GO terms in this comparison (S3 Table; <https://biit.cs.ut.ee/gplink/l/4Fz7ZA3hTC>).

### Individual egg reduction rate phenotypes

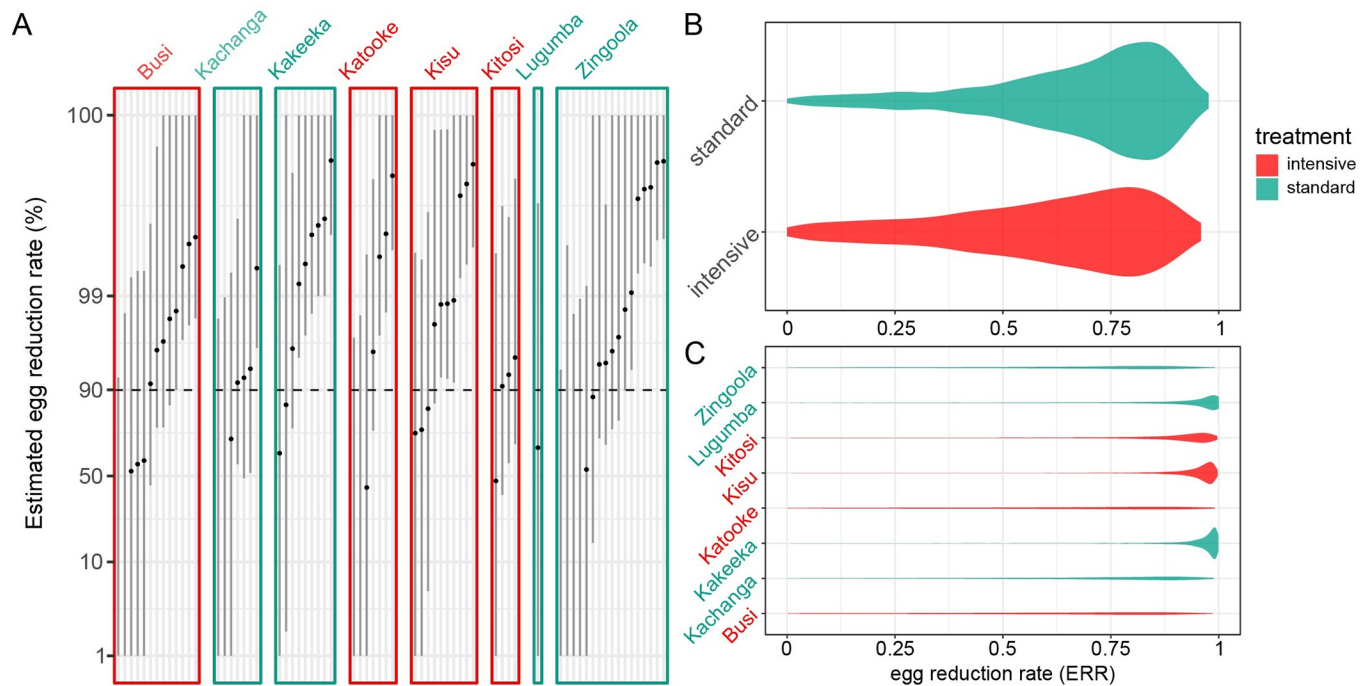
In an attempt to identify a phenotype for drug efficacy, we estimated the egg reduction rate (ERR) for 88 individuals for which genomic data was available and that had Kato-Katz egg counts taken both before and after treatment using a Bayesian linear mixed-effect model [82] that has previously been used to assess praziquantel efficacy [9, 86]. Previous analysis revealed a lower but not significant ERR in the intensive arm than the standard arm [10]. Similarly, we observe lower marginal ERR in samples collected in intensive than standard treatment villages, but with largely overlapping posterior distributions (Fig 7B), and while villages vary in ERR



**Fig 6. XP-EHH coloured by chromosome among treatment groups.** A. Comparison between standard and intensive treatment groups. B. Comparison between pre-treatment and post-treatment groups. Positive values in panel A represent windows under stronger selective pressure in annual vs quarterly treatment arms. In panel B, positive values represent windows under stronger selection in pre-treatment than post-treatment samples.

<https://doi.org/10.1371/journal.pntd.0010188.g006>

(Fig 7C), there were similar numbers of high- and low-clearance villages in the two arms. These differences were largely driven by a small number of individuals in some villages with very low (even negative—implying a higher egg count after treatment than before) ERR values (Fig 7A). Unlike in a previous study [9], no ERRs were significantly below the 90% threshold,



**Fig 7. Egg reduction rate (ERR) estimates.** (A) Posterior distributions of ERR for each individual for which pre- and post- egg count data were available. Lines indicate the 95% credible intervals (highest posterior density intervals) for each estimate, dots are the mean of the posterior distribution. Individuals are shown on an individual panel for each village, with panel headers coloured by treatment arm. Posterior distribution of average ERRs stratified by (B) treatment arm and (C) school were constructed by marginalizing over the fixed- and random-effects coefficients of the generalised linear mixed model.

<https://doi.org/10.1371/journal.pntd.0010188.g007>

probably because only duplicate counts were available before and after treatment here, so there was significantly less information to estimate ERR on a per-individual basis.

Despite the small sample size meaning our study is likely to be under-powered to detect any but the most highly penetrant high-frequency alleles, we attempted to identify genetic variants associated with differences in ERR, testing the 6.95 million high-quality SNPs found on the 7 autosomes or on the shared ZW scaffold. The smallest p-value for any SNP was  $1.393 \times 10^{-9}$ , which after adjustment for multiple testing represents an adjusted p-value of 0.01791 (S3A Fig). However, there was some evidence that p-values are systematically biased in this analysis (S3B Fig). Correcting for population structure based on PCA coordinates removed the significance of hits (lowest p-value =  $4.3 \times 10^{-8}$ ; adjusted p-value = 0.274). The most significant hit (SM\_V7\_7:1552004) is in an intron of Phosphatidylinositol 4-kinase (Smp\_130340) and has no obvious link to praziquantel mode-of-action. We thus conclude that there is no strong evidence linking any individual genetic variant in these data to variation in estimated ERR.

## Discussion

Schistosomiasis is second only to malaria in socio-economic impact among parasitic causes of morbidity and mortality [87–89]. MDA is the main method for schistosomiasis control, and there is currently an effort to expand the coverage of community-wide drug treatment to improve morbidity control [90] and address the persistence of schistosomiasis in some areas despite many years of PZQ distribution [91]. Understanding whether intensive treatment for individuals living in high transmission communities has an impact on parasite populations, potentially leading to drug resistance is of high importance for public health among schistosomiasis endemic communities in Africa. Determining the genetic basis of any drug resistance



that does emerge is also crucial for tracking the spread of resistance through schistosome populations and for future drug or vaccine development designed to circumvent resistance as has been demonstrated for oxamniquine resistance [92].

Here, we have taken advantage of a large-scale trial in which the entire communities of 26 fishing villages were regularly treated with PZQ. A number of features of the study made this an ideal place to detect an effect of PZQ treatment on parasite populations. Villages were assigned randomly to treatment arms, so treatment frequency was independent of morbidity, parasite prevalence or intensity. Treatment was given under direct observation, avoiding issues with drug compliance reported in other studies [5]. The four week follow-up interval post-treatment would minimise the possibility of diagnosing newly acquired infection after treatment based on the development time of *S. mansoni* [93], as a new infection would take longer than four weeks to result in egg production that could be detected by Kato Katz and microscopy [94]. The exception would be if, during the time of treatment, a patient had juvenile worms as these would not have been cleared by treatment [95]. We expected that as the study was based on a group of islands it might help isolate the parasite population and so allow us to detect drug-induced selection in this population without the confounding effect of high levels of gene flow from untreated populations. Only individuals who had lived in these villages for at least three years were included in the study to control for absenteeism and MDA compliance although it was still not possible to control for movement between villages and islands given that fishing is the main economic activity within these communities.

The population of parasites present on the islands is closely related to that recently described from communities on the shoreline of Lake Victoria, and as expected rather divergent from that inland from the lake (Fig 2A and 2B) [55]. This presumably reflects greater movement of people between the shoreline and island than with the inland populations, as well as that the inland population included here is further (approximately 160 km) from the shoreline than are the islands (approximately 80 km). We see little genetic differentiation between villages on the same island, as fishing villages are close to one another (1–13 km apart) and movement may be frequent among fishermen and village communities. Less expected was that we see little or no genetic differentiation between islands, with only a weak trend for greater genetic differentiation between villages on different islands than villages on the same island, albeit this is a larger effect than either the distance between villages or the size of village populations, perhaps suggesting that snail vector movement around the coasts of islands may play a role in parasite movement. The islands are separated by water that is deep enough [96] (primary data at <http://dataverse.harvard.edu/dataverse/LakeVicFish>) to prevent snails moving actively from one island to the next, but parasites could travel through movement of infected people or through infected snails being carried on fishermen's or conceivably by rafting on floating plants such as water hyacinth. Geographical conditions on these islands are similar except for Lugumba Island which has more rocky/stony shores compared to Koome and Damba which have more sandy shores and more vegetation, so we would expect snails to be able to establish similarly at most locations.

Despite seeing little or no genetic structure in the island parasite population, we see some evidence that PZQ treatment has had a small effect on the genetic diversity of the parasite population in this area. While we do not have baseline samples from before any PZQ treatment was administered as part of the LaVIISWA trial, we see very slightly higher genome-wide genetic diversity in the standard treatment arm than in the intensive arm, as would be expected if intensive treatment has been more effective at reducing the parasite population than the standard treatment regimen [10], although the effect we observe is very small and so maybe of limited biological relevance. Differences in the same direction were present when comparing subsets of samples taken before and after treatment separately, and was more pronounced in

the post-treatment populations. As we see only very few closely related parasites, and no significant enrichment in relatedness based on treatment arm or sampling time with respect to treatment, it seems that this effect is unlikely to be due to differences in the number of directly related miracidia. We observe little or no genetic differentiation between villages in the two study arms, and only very slightly higher differentiation between the arms in post-treatment than in pre-treatment samples.

Evidence that PZQ treatment has some effect on the parasite population led us to investigate whether particular variants might be related to exposure to PZQ and so potentially responsible for any reduced susceptibility of parasites to PZQ within MDA programs [9]. We identify several regions within the genome that were highly differentiated between samples from the standard and intensive arms of the study, including a particularly striking region on chromosome 5 that showed high differentiation between post-treatment samples from the two arms of the study. This region contained a number of genes with functions that could be potentially linked to PZQ drug action. These include an ATP-binding cassette (ABC) transporter-associated gene (Smp\_136310) that has previously been linked to helminth detoxification and drug resistance processes [97]. A gene with calcium dependent/modulatory functions (Smp\_347070) was also found in the enriched region on chromosome 5, which is of interest given that the mode of action of PZQ has long been linked to increased permeability of the cell membrane to calcium ions into the cells which then causes contraction, paralysis and eventual death of the worms [98]. We also investigated regions of the genome under different selective regimes either with treatment intensity or when comparing pre- and post-treatment samples. Among the genes under varying selection were purine-nucleoside phosphorylase activity associated genes (Smp\_197110 and Smp\_171620) which are involved in the nucleotide salvage pathway of *S. mansoni*. Given that *S. mansoni* depends entirely on the salvage pathway for its purine metabolism [99], there is a possibility that ongoing non-random selection within this gene might affect parasite metabolic processes and a potential future drug target. However, we note that many biological processes could be contributing to genetic variation between samples from natural populations apart from variation in drug susceptibility [100]. While this study has shed some light on possible drug resistance genetic markers, other approaches, such as genetic crosses between parasites [101, 102] from natural populations that vary in drug efficacy or from lines selected for resistance [21, 92], are likely to have more power to reveal the genetics of drug resistance and so enable more focused studies of the effect on treatment on parasite populations.

A limitation in this study was that we did not have parasite populations sampled several years apart since it has been observed in similar studies that differentiation occurs over time in a given community [38], so sampling over a longer time-span could provide stronger evidence of genetic change in the population. In particular, we would ideally have access to baseline samples from the same population taken prior to any large-scale PZQ treatment being administered. Despite the falling cost and rising throughput of nucleic acid sequencing, we were limited in the number of miracidia that we could sequence in this study. An additional limitation is the labour-intensive process of hatching and washing miracidia necessary to obtain high-quality data due to the non-selective nature of the whole-genome sequencing approach [70].

As control programmes expand and reduce pathogen populations, we would expect the genetic diversity of these populations to fall to reflect the reduced population size [13, 41, 103], and drug resistance to be reflected in particular genotypes being over-represented in samples collected after large-scale treatment has been applied. As in other recent studies [42, 55], we find evidence of at best a very limited effect of PZQ treatment on schistosome populations either post-treatment or over a longer time frame of intensive treatment with a large diverse parasite population and only minimal population structure. In most previous studies,

extensive refugia from treatment have been present in the community, as only school-age children are routinely treated in most areas, so it is instructive that we find similar results in this study despite community-wide treatment. Taken together, this body of work suggests that PZQ treatment may be some way from having a dramatic effect on schistosome transmission in areas of high infection prevalence, suggesting that additional interventions may be important in reducing transmission [91, 104, 105].

We see little evidence that PZQ treatment in these areas is strongly selecting for resistance. One possibility is that this study—and others—are underpowered to detect the early signs of selection, as might be expected if multiple resistance-conferring alleles are present in the population [106] or are present on many different genetic backgrounds in this large population, so that any selection for resistance is a very soft selective sweep [107] and so difficult to detect. While there is some evidence for reduced efficacy of PZQ in Uganda [9], other studies do not find this effect [11], including one study based on the same population as studied here [10]. Perhaps the most parsimonious explanation is that low efficacy, where reported, is not a sign of PZQ resistance, but due to host or other factors [91, 108]. If persistent schistosomiasis hot-spots are due to factors other than PZQ efficacy, this should reassure control programs that continued PZQ MDA will continue to play an important role—alongside other interventions—in reducing infection prevalence in these foci. Even in the absence of drug resistance emerging in natural populations, high-resolution genetic surveillance of African schistosome populations is ideally suited to detect changes in parasite population structure related to the impact of control measures [30], and could ultimately inform approaches to eliminate schistosome morbidity in remaining 'hot-spots' by helping us understand parasite transmission between hosts and between foci [91].

In summary, we demonstrate a small but significant effect of both short-term PZQ treatment intensity and a recent treatment episode on genome wide-diversity in a schistosome. This reduction in diversity does not appear to be associated with enrichment of closely related parasites, but rather could reflect ongoing non-random recent selection within these fishing communities in Uganda that might be under the influence of continued mass drug administration. We identify genomic windows that are either particularly differentiated following treatment or appear to be under differing selective regimes with different treatment intensity. These regions could include genes involved in drug response, but additional data is needed to prioritise candidates for further investigation.

## Supporting information

**S1 Fig. Mean  $F_{ST}$  between all pairs of villages.** Shading indicates levels of genetic differentiation between pairs of villages indicated on each row and column. Symbol shapes reflect pairwise comparisons of differentiation of populations samples between or within islands, and the area of each symbol is proportional to the linear distance between the villages. (EPS)

**S2 Fig. Principal component analysis of population structure by treatment and time point.** (TIF)

**S3 Fig.** (A) Manhattan plot of unadjusted  $-\log_{10}$  p-values for association of individual SNPs with per-individual mean egg-reduction rates. (B) QQplot of p-values from the same analysis against expectations under the null hypothesis. (TIFF)

**S4 Fig. Principal component analysis showing a cluster of nine distinct miracidia on principal component 2 in the lower left quadrant.**

(TIFF)

**S1 Table. Accession numbers and metadata for all samples included in analyses.**

(DOCX)

**S2 Table. Protein coding genes present in the region of highest genetic differentiation on chromosome 5. (region X on Fig 4B).**

(DOCX)

**S3 Table. GO terms significantly over-represented among genes overlapping regions of different natural selection.** Genes are from GO hierarchies for Molecular Function (MF), Biological Processes (BP), KEGG pathways and Cellular Component (CC) within the post-treatment (Tp2). Standard (Std) and Intensive (Int) groups with respective adjusted p-values ( $P_{adj}$ ). Asterisks mark terms that remain enriched in autosomal gene sets.

(DOCX)

## Acknowledgments

We thank the Cure Rates study team and participants, the Wellcome Sanger Institute sequencing team, and members of the Parasite Genomics group and Pathogen Informatics team in the Parasites and Microbes Programme at the Sanger Institute, in particular Tim Stickland for setting up the Panoptes web application. For the purpose of Open Access, the authors have applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

## Author Contributions

**Conceptualization:** Richard E. Sanya, Emily L. Webb, Alison M. Elliott, Matthew Berriman.

**Formal analysis:** Tushabe John Vianney, Duncan J. Berger, Stephen R. Doyle, Emily L. Webb, James A. Cotton.

**Funding acquisition:** Alison M. Elliott, Matthew Berriman.

**Investigation:** Tushabe John Vianney, Duncan J. Berger, Stephen R. Doyle, Geetha Sankaranarayanan, Joel Serubanja, Prossy Kabuubi Nakawungu, Fred Besigye, Richard E. Sanya, Fiona Allan.

**Methodology:** Fiona Allan, Emily L. Webb.

**Project administration:** Nancy Holroyd, Alison M. Elliott, Matthew Berriman.

**Resources:** Geetha Sankaranarayanan, Joel Serubanja, Richard E. Sanya, Fiona Allan.

**Supervision:** Alison M. Elliott, Matthew Berriman, James A. Cotton.

**Writing – original draft:** Tushabe John Vianney, James A. Cotton.

**Writing – review & editing:** Tushabe John Vianney, Duncan J. Berger, Stephen R. Doyle, Geetha Sankaranarayanan, Joel Serubanja, Prossy Kabuubi Nakawungu, Fred Besigye, Richard E. Sanya, Nancy Holroyd, Fiona Allan, Emily L. Webb, Alison M. Elliott, Matthew Berriman, James A. Cotton.

## References

1. Sandbach FR. The history of schistosomiasis research and policy for its control. *Med Hist*. 1976; 20(3):259–75. <https://doi.org/10.1017/s0025727300022663> PMID: 792584
2. Hotez PJ, Alvarado M, Basanez MG, Bolliger I, Bourne R, Boussinesq M, et al. The global burden of disease study 2010: interpretation and implications for the neglected tropical diseases. *PLoS Negl Trop Dis*. 2014; 8(7):e2865. <https://doi.org/10.1371/journal.pntd.0002865> PMID: 25058013
3. Kabatereine NB, Brooker S, Koukounari A, Kazibwe F, Tukahebwa EM, Fleming FM, et al. Impact of a national helminth control programme on infection and morbidity in Ugandan schoolchildren. *Bull World Health Organ*. 2007; 85(2):91–9. <https://doi.org/10.2471/blt.06.030353> PMID: 17308729
4. Blanton RE, Blank WA, Costa JM, Carmo TM, Reis EA, Silva LK, et al. *Schistosoma mansoni* population structure and persistence after praziquantel treatment in two villages of Bahia, Brazil. *Int J Parasitol*. 2011; 41(10):1093–9. <https://doi.org/10.1016/j.ijpara.2011.06.002> PMID: 21784077
5. Adriko M, Faust CL, Carruthers LV, Moses A, Tukahebwa EM, Lamberton PHL. Low Praziquantel Treatment Coverage for *Schistosoma mansoni* in Mayuge District, Uganda, Due to the Absence of Treatment Opportunities, Rather Than Systematic Non-Compliance. *Trop Med Infect Dis*. 2018; 3(4).
6. Krauth SJ, Balen J, Gobert GN, Lamberton PHL. A Call for Systems Epidemiology to Tackle the Complexity of Schistosomiasis, Its Control, and Its Elimination. *Trop Med Infect Dis*. 2019; 4(1).
7. Sanya RE, Tumwesige E, Elliott AM, Seeley J. Perceptions about interventions to control schistosomiasis among the Lake Victoria island communities of Kooro, Uganda. *PLoS Negl Trop Dis*. 2017; 11(10):e0005982. <https://doi.org/10.1371/journal.pntd.0005982> PMID: 28968470
8. Ending the neglect to attain the Sustainable Development Goals: a road map for neglected tropical diseases 2021–2030 2020.
9. Crellen T, Walker M, Lamberton PH, Kabatereine NB, Tukahebwa EM, Cotton JA, et al. Reduced Efficacy of Praziquantel Against *Schistosoma mansoni* Is Associated With Multiple Rounds of Mass Drug Administration. *Clin Infect Dis*. 2016; 63(9):1151–9. <https://doi.org/10.1093/cid/ciw506> PMID: 27470241
10. Tushabe JV, Lubyayi L, Sserubanja J, Kabuubi P, Abayo E, Kiwanuka S, et al. Does Intensive Treatment Select for Praziquantel Resistance in High-Transmission Settings? Parasitological Trends and Treatment Efficacy Within a Cluster-Randomized Trial. *Open Forum Infect Dis*. 2020; 7(4):ofaa091. <https://doi.org/10.1093/ofid/ofaa091> PMID: 32296727
11. Fukushige M, Chase-Topping M, Woolhouse MEJ, Mutapi F. Efficacy of praziquantel has been maintained over four decades (from 1977 to 2018): A systematic review and meta-analysis of factors influence its efficacy. *PLoS Negl Trop Dis*. 2021; 15(3):e0009189. <https://doi.org/10.1371/journal.pntd.0009189> PMID: 33730095
12. Mutapi F, Maizels R, Fenwick A, Woolhouse M. Human schistosomiasis in the post mass drug administration era. *Lancet Infect Dis*. 2017; 17(2):e42–e8. [https://doi.org/10.1016/S1473-3099\(16\)30475-3](https://doi.org/10.1016/S1473-3099(16)30475-3) PMID: 27988094
13. Coeli R, Baba EH, Araujo N, Coelho PM, Oliveira G. Praziquantel treatment decreases *Schistosoma mansoni* genetic diversity in experimental infections. *PLoS Negl Trop Dis*. 2013; 7(12):e2596. <https://doi.org/10.1371/journal.pntd.0002596> PMID: 24367712
14. Fallon PG, Sturrock RF, Niang AC, Doenhoff MJ. Short report: diminished susceptibility to praziquantel in a Senegal isolate of *Schistosoma mansoni*. *Am J Trop Med Hyg*. 1995; 53(1):61–2. PMID: 7625534
15. Lotfy WM, Hishmat MG, El Nashar AS, Abu El Einin HM. Evaluation of a method for induction of praziquantel resistance in *Schistosoma mansoni*. *Pharm Biol*. 2015; 53(8):1214–9. <https://doi.org/10.3109/13880209.2014.970289> PMID: 25609146
16. Messerli SM, Kasinathan RS, Morgan W, Spranger S, Greenberg RM. *Schistosoma mansoni* P-glycoprotein levels increase in response to praziquantel exposure and correlate with reduced praziquantel susceptibility. *Mol Biochem Parasitol*. 2009; 167(1):54–9. <https://doi.org/10.1016/j.molbiopara.2009.04.007> PMID: 19406169
17. Pinto-Almeida A, Mendes T, de Oliveira RN, Correa Sde A, Allegretti SM, Belo S, et al. Morphological Characteristics of *Schistosoma mansoni* PZQ-Resistant and -Susceptible Strains Are Different in Presence of Praziquantel. *Front Microbiol*. 2016; 7:594. <https://doi.org/10.3389/fmicb.2016.00594> PMID: 27199925
18. Blanton RE. Population Structure and Dynamics of Helminthic Infection: Schistosomiasis. *Microbiol Spectr*. 2019; 7(4). <https://doi.org/10.1128/microbiolspec.AME-0009-2019> PMID: 31325285
19. Vale N, Gouveia MJ, Rinaldi G, Brindley PJ, Gartner F, Correia da Costa JM. Praziquantel for Schistosomiasis: Single-Drug Metabolism Revisited, Mode of Action, and Resistance. *Antimicrob Agents Chemother*. 2017; 61(5).

20. Park SK, Gunaratne GS, Chulkov EG, Moehring F, McCusker P, Dosa PI, et al. The anthelmintic drug praziquantel activates a schistosome transient receptor potential channel. *J Biol Chem*. 2019; 294(49):18873–80. <https://doi.org/10.1074/jbc.AC119.011093> PMID: 31653697
21. Le Clec'h W, Chevalier FD, Mattos ACA, Strickland A, Diaz R, McDew-White M, et al. Genetic analysis of praziquantel response in schistosome parasites implicates a transient receptor potential channel. *Sci Transl Med*. 2021; 13(625):eabj9114. <https://doi.org/10.1126/scitranslmed.abj9114> PMID: 34936381
22. Webster JP, Neves MI, Webster BL, Pennance T, Rabone M, Gouvras AN, et al. Parasite Population Genetic Contributions to the Schistosomiasis Consortium for Operational Research and Evaluation within Sub-Saharan Africa. *Am J Trop Med Hyg*. 2020; 103(1\_Suppl):80–91. <https://doi.org/10.4269/ajtmh.19-0827> PMID: 32400355
23. Betson M, Sousa-Figueiredo JC, Kabatereine NB, Stothard JR. New insights into the molecular epidemiology and population genetics of *Schistosoma mansoni* in Ugandan pre-school children and mothers. *PLoS Negl Trop Dis*. 2013; 7(12):e2561. <https://doi.org/10.1371/journal.pntd.0002561> PMID: 24349589
24. Webster BL, Webster JP, Gouvras AN, Garba A, Lamine MS, Diaw OT, et al. DNA 'barcoding' of *Schistosoma mansoni* across sub-Saharan Africa supports substantial within locality diversity and geographical separation of genotypes. *Acta Trop*. 2013; 128(2):250–60. <https://doi.org/10.1016/j.actatropica.2012.08.009> PMID: 22935316
25. Blair L, Webster JP, Barker GC. Isolation and characterization of polymorphic microsatellite markers in *Schistosoma mansoni* from Africa. *Mol Ecol Notes*. 2001; 1:93–5.
26. Curtis J, Sorensen RE, Page LK, Minchella DJ. Microsatellite loci in the human blood fluke *Schistosoma mansoni* and their utility for other schistosome species. *Mol Ecol Notes*. 2001; 1:143–5.
27. Durand P, Sire C, Theron A. Isolation of microsatellite markers in the digenetic trematode *Schistosoma mansoni* from Guadeloupe island. *Mol Ecol*. 2000; 9(7):977–8. <https://doi.org/10.1046/j.1365-294x.2000.00939-4.x> PMID: 10886664
28. Golan R, Gower CM, Emery AM, Rollinson D, Webster JP. Isolation and characterization of the first polymorphic microsatellite markers for *Schistosoma haematobium* and their application in multiplex reactions of larval stages. *Mol Ecol Resour*. 2008; 8:647–9. <https://doi.org/10.1111/j.1471-8286.2007.02031.x> PMID: 21585859
29. Rudge JW, Carabin H, Balolong E, Tallo V, Shrivastava J, Lu DB, et al. Population genetics of *Schistosoma japonicum* within the Philippines suggest high levels of transmission between humans and dogs. *PLoS Negl Trop Dis*. 2008; 2(11):e340. <https://doi.org/10.1371/journal.pntd.0000340> PMID: 19030225
30. Gower CM, Gouvras AN, Lamberton PH, Deol A, Shrivastava J, Mutombo PN, et al. Population genetic structure of *Schistosoma mansoni* and *Schistosoma haematobium* from across six sub-Saharan African countries: implications for epidemiology, evolution and control. *Acta Trop*. 2013; 128(2):261–74. <https://doi.org/10.1016/j.actatropica.2012.09.014> PMID: 23041540
31. Gower CM, Shrivastava J, Lamberton PH, Rollinson D, Webster BL, Emery A, et al. Development and application of an ethically and epidemiologically advantageous assay for the multi-locus microsatellite analysis of *Schistosoma mansoni*. *Parasitology*. 2007; 134(Pt 4):523–36. <https://doi.org/10.1017/S0031182006001685> PMID: 17096873
32. Lawton SP, Hirai H, Ironside JE, Johnston DA, Rollinson D. Genomes and geography: genomic insights into the evolution and phylogeography of the genus *Schistosoma*. *Parasit Vectors*. 2011; 4:131. <https://doi.org/10.1186/1756-3305-4-131> PMID: 21736723
33. Blank WA, Reis EA, Thiong'o FW, Braghioroli JF, Santos JM, Melo PR, et al. Analysis of *Schistosoma mansoni* population structure using total fecal egg sampling. *J Parasitol*. 2009; 95(4):881–9. <https://doi.org/10.1645/GE-1895.1> PMID: 20049994
34. Thiele EA, Sorensen RE, Gazzinelli A, Minchella DJ. Genetic diversity and population structuring of *Schistosoma mansoni* in a Brazilian village. *Int J Parasitol*. 2008; 38(3–4):389–99. <https://doi.org/10.1016/j.ijpara.2007.07.011> PMID: 17825305
35. Van den Broeck F, Maes GE, Larmuseau MH, Rollinson D, Sy I, Faye D, et al. Reconstructing Colonization Dynamics of the Human Parasite *Schistosoma mansoni* following Anthropogenic Environmental Changes in Northwest Senegal. *PLoS Negl Trop Dis*. 2015; 9(8):e0003998. <https://doi.org/10.1371/journal.pntd.0003998> PMID: 26275049
36. Rey O, Webster BL, Huysse T, Rollinson D, Van den Broeck F, Kincaid-Smith J, et al. Population genetics of African *Schistosoma* species. *Infect Genet Evol*. 2021; 89:104727. <https://doi.org/10.1016/j.meegid.2021.104727> PMID: 33486128

37. Neves MI, Gower CM, Webster JP, Walker M. Revisiting density-dependent fecundity in schistosomes using sibship reconstruction. *PLoS Negl Trop Dis*. 2021; 15(5):e0009396. <https://doi.org/10.1371/journal.pntd.0009396> PMID: 33983965
38. Norton AJ, Gower CM, Lamberton PH, Webster BL, Lwambo NJ, Blair L, et al. Genetic consequences of mass human chemotherapy for *Schistosoma mansoni*: population structure pre- and post-praziquantel treatment in Tanzania. *Am J Trop Med Hyg*. 2010; 83(4):951–7. <https://doi.org/10.4269/ajtmh.2010.10-0283> PMID: 20889898
39. French MD, Churcher TS, Gambhir M, Fenwick A, Webster JP, Kabatereine NB, et al. Observed reductions in *Schistosoma mansoni* transmission from large-scale administration of praziquantel in Uganda: a mathematical modelling study. *PLoS Negl Trop Dis*. 2010; 4(11):e897. <https://doi.org/10.1371/journal.pntd.0000897> PMID: 21124888
40. Lelo AE, Mburu DN, Magoma GN, Mungai BN, Kihara JH, Mwangi IN, et al. No apparent reduction in schistosome burden or genetic diversity following four years of school-based mass drug administration in mwea, central kenya, a heavy transmission area. *PLoS Negl Trop Dis*. 2014; 8(10):e3221. <https://doi.org/10.1371/journal.pntd.0003221> PMID: 25299057
41. Gower CM, Gehre F, Marques SR, Lamberton PHL, Lwambo NJ, Webster JP. Phenotypic and genotypic monitoring of *Schistosoma mansoni* in Tanzanian schoolchildren five years into a preventative chemotherapy national control programme. *Parasit Vectors*. 2017; 10(1):593. <https://doi.org/10.1186/s13071-017-2533-6> PMID: 29197426
42. Faust CL, Crotti M, Moses A, Oguttu D, Wamboko A, Adriko M, et al. Two-year longitudinal survey reveals high genetic diversity of *Schistosoma mansoni* with adult worms surviving praziquantel treatment at the start of mass drug administration in Uganda. *Parasit Vectors*. 2019; 12(1):607. <https://doi.org/10.1186/s13071-019-3860-6> PMID: 31881923
43. Le Clec'h W, Chevalier FD, McDew-White M, Allan F, Webster BL, Gouvras AN, et al. Whole genome amplification and exome sequencing of archived schistosome miracidia. *Parasitology*. 2018; 145(13):1739–47. <https://doi.org/10.1017/S0031182018000811> PMID: 29806576
44. Platt RN, McDew-White M, Le Clec'h W, Chevalier FD, Allan F, Emery AM, et al. Ancient Hybridization and Adaptive Introgression of an Invadolysin Gene in Schistosome Parasites. *Mol Biol Evol*. 2019; 36(10):2127–42. <https://doi.org/10.1093/molbev/msz154> PMID: 31251352
45. Platt RN, Le Clec'h W, Chevalier FD, McDew-White M, LoVerde P, de Assis RR, et al. Genomic analysis of a parasite invasion: colonization of the Americas by the blood fluke, *Schistosoma mansoni*. *bioRxiv*. 2021:465783.
46. Shortt JA, Card DC, Schield DR, Liu Y, Zhong B, Castoe TA, et al. Whole Genome Amplification and Reduced-Representation Genome Sequencing of *Schistosoma japonicum* Miracidia. *PLoS Negl Trop Dis*. 2017; 11(1):e0005292. <https://doi.org/10.1371/journal.pntd.0005292> PMID: 28107347
47. Shortt JA, Timm LE, Hales NR, Nikolakis ZL, Schield DR, Perry BW, et al. Population genomic analyses of schistosome parasites highlight critical challenges facing endgame elimination efforts. *Sci Rep*. 2021; 11(1):6884. <https://doi.org/10.1038/s41598-021-86287-y> PMID: 33767307
48. Protasio AV, Tsai IJ, Babbage A, Nichol S, Hunt M, Aslett MA, et al. A systematically improved high quality genome and transcriptome of the human blood fluke *Schistosoma mansoni*. *PLoS Negl Trop Dis*. 2012; 6(1):e1455. <https://doi.org/10.1371/journal.pntd.0001455> PMID: 22253936
49. Oey H, Zakrzewski M, Gravermann K, Young ND, Korhonen PK, Gobert GN, et al. Whole-genome sequence of the bovine blood fluke *Schistosoma bovis* supports interspecific hybridization with *S. haematobium*. *PLoS Pathog*. 2019; 15(1):e1007513. <https://doi.org/10.1371/journal.ppat.1007513> PMID: 30673782
50. Stroehlein AJ, Korhonen PK, Chong TM, Lim YL, Chan KG, Webster B, et al. High-quality *Schistosoma haematobium* genome achieved by single-molecule and long-range sequencing. *Gigascience*. 2019; 8(9). <https://doi.org/10.1093/gigascience/giz108> PMID: 31494670
51. Luo F, Yin M, Mo X, Sun C, Wu Q, Zhu B, et al. An improved genome assembly of the fluke *Schistosoma japonicum*. *PLoS Negl Trop Dis*. 2019; 13(8):e0007612. <https://doi.org/10.1371/journal.pntd.0007612> PMID: 31390359
52. International Helminth Genomes C. Comparative genomics of the major parasitic worms. *Nat Genet*. 2019; 51(1):163–74. <https://doi.org/10.1038/s41588-018-0262-1> PMID: 30397333
53. Kincaid-Smith J, Tracey A, de Carvalho Augusto R, Bulla I, Holroyd N, Rognon A, et al. Morphological and genomic characterisation of the *Schistosoma* hybrid infecting humans in Europe reveals admixture between *Schistosoma haematobium* and *Schistosoma bovis*. *PLoS Negl Trop Dis*. 2021; 15(12):e0010062. <https://doi.org/10.1371/journal.pntd.0010062> PMID: 34941866
54. Buddenborg SK, Tracey A, Berger DJ, Lu Z, Doyle SR, Fu B, et al. Assembled chromosomes of the blood fluke *Schistosoma mansoni* provide insight into the evolution of its ZW sex-determination system. *bioRxiv*. 2021:456314.

55. Berger DJ, Crellen T, Lamberton PHL, Allan F, Tracey A, Noonan JD, et al. Whole-genome sequencing of *Schistosoma mansoni* reveals extensive diversity with limited selection despite mass drug administration. *Nat Commun.* 2021; 12(1):4776. <https://doi.org/10.1038/s41467-021-24958-0> PMID: 34362894
56. Young ND, Chan KG, Korhonen PK, Min Chong T, Ee R, Mohandas N, et al. Exploring molecular variation in *Schistosoma japonicum* in China. *Sci Rep.* 2015; 5:17345. <https://doi.org/10.1038/srep17345> PMID: 26621075
57. Crellen T, Allan F, David S, Durrant C, Huckvale T, Holroyd N, et al. Whole genome resequencing of the human parasite *Schistosoma mansoni* reveals population history and effects of selection. *Sci Rep.* 2016; 6:20954. <https://doi.org/10.1038/srep20954> PMID: 26879532
58. Nikolakis ZL, Hales NR, Perry BW, Schield DR, Timm LE, Liu Y, et al. Patterns of relatedness and genetic diversity inferred from whole genome sequencing of archival blood fluke miracidia (*Schistosoma japonicum*). *PLoS Negl Trop Dis.* 2021; 15(1):e0009020. <https://doi.org/10.1371/journal.pntd.0009020> PMID: 33406094
59. Shetty AC, Jacob CG, Huang F, Li Y, Agrawal S, Saunders DL, et al. Genomic structure and diversity of *Plasmodium falciparum* in Southeast Asia reveal recent parasite migration patterns. *Nat Commun.* 2019; 10(1):2665. <https://doi.org/10.1038/s41467-019-10121-3> PMID: 31209259
60. Taylor AR, Schaffner SF, Cerqueira GC, Nkhoma SC, Anderson TJC, Sriprawat K, et al. Quantifying connectivity between local *Plasmodium falciparum* malaria parasite populations using identity by descent. *PLoS Genet.* 2017; 13(10):e1007065. <https://doi.org/10.1371/journal.pgen.1007065> PMID: 29077712
61. MalariaGen, Ahouidi A, Ali M, Almagro-Garcia J, Amambua-Ngwa A, Amaratunga C, et al. An open dataset of *Plasmodium falciparum* genome variation in 7,000 worldwide samples. *Wellcome Open Res.* 2021; 6:42. <https://doi.org/10.12688/wellcomeopenres.16168.2> PMID: 33824913
62. Nampijja M, Webb EL, Kaweesa J, Kizindo R, Namutebi M, Nakazibwe E, et al. The Lake Victoria Island Intervention Study on Worms and Allergy-related diseases (LaVIISWA): study protocol for a randomised controlled trial. *Trials.* 2015; 16:187. <https://doi.org/10.1186/s13063-015-0702-5> PMID: 25902705
63. Sanya RE, Nkurunungi G, Hoek Spaans R, Nampijja M, O'Hara G, Kizindo R, et al. The Impact of Intensive Versus Standard Anthelmintic Treatment on Allergy-related Outcomes, Helminth Infection Intensity, and Helminth-related Morbidity in Lake Victoria Fishing Communities, Uganda: Results From the LaVIISWA Cluster-randomized Trial. *Clin Infect Dis.* 2019; 68(10):1665–74. <https://doi.org/10.1093/cid/ciy761> PMID: 30202872
64. Munisi DZ, Buza J, Mpolya EA, Angelo T, Kinung'hi SM. The Efficacy of Single-Dose versus Double-Dose Praziquantel Treatments on *Schistosoma mansoni* Infections: Its Implication on Undernutrition and Anaemia among Primary Schoolchildren in Two On-Shore Communities, Northwestern Tanzania. *Biomed Res Int.* 2017; 2017:7035025. <https://doi.org/10.1155/2017/7035025> PMID: 29094048
65. Kittur N, Binder S, Campbell CH, King CH, Kinung'hi S, Olsen A, et al. Defining Persistent Hotspots: Areas That Fail to Decrease Meaningfully in Prevalence after Multiple Years of Mass Drug Administration with Praziquantel for Control of Schistosomiasis. *Am J Trop Med Hyg.* 2017; 97(6):1810–7. <https://doi.org/10.4269/ajtmh.17-0368> PMID: 29016344
66. Sanya RE, Webb EL, Zziwa C, Kizindo R, Sewankambo M, Tumusiime J, et al. The Effect of Helminth Infections and Their Treatment on Metabolic Outcomes: Results of a Cluster-Randomized Trial. *Clin Infect Dis.* 2020; 71(3):601–13. <https://doi.org/10.1093/cid/ciz859> PMID: 31504336
67. Tennekes M. tmap: Thematic Maps in R. *Journal of Statistical Software.* 2018; 84:1–39.
68. Pebesma E. Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal.* 2018; 10:439–46.
69. Pitchford RJ, Visser PS. A simple and rapid technique for quantitative estimation of helminth eggs in human and animal excreta with special reference to *Schistosoma* sp. *Trans R Soc Trop Med Hyg.* 1975; 69(3):318–22. [https://doi.org/10.1016/0035-9203\(75\)90126-1](https://doi.org/10.1016/0035-9203(75)90126-1) PMID: 1179467
70. Doyle SR, Sankaranarayanan G, Allan F, Berger D, Jimenez Castro PD, Collins JB, et al. Evaluation of DNA Extraction Methods on Individual Helminth Egg and Larval Stages for Whole-Genome Sequencing. *Front Genet.* 2019; 10:826. <https://doi.org/10.3389/fgene.2019.00826> PMID: 31616465
71. Lee-Six H, Olafsson S, Ellis P, Osborne RJ, Sanders MA, Moore L, et al. The landscape of somatic mutation in normal colorectal epithelial cells. *Nature.* 2019; 574(7779):532–7. <https://doi.org/10.1038/s41586-019-1672-7> PMID: 31645730
72. Picard toolkit. Broad Institute; 2018.
73. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. *Bioinformatics.* 2011; 27(15):2156–8. <https://doi.org/10.1093/bioinformatics/btr330> PMID: 21653522



74. Kimura M. The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations. *Genetics*. 1969; 61(4):893–903. <https://doi.org/10.1093/genetics/61.4.893> PMID: 5364968
75. Csardi G, Nepusz T. The igraph software package for complex network research. *InterJournal*. 2006; *Complex Systems*:1695.
76. Huisman J. Pedigree reconstruction from SNP data: parentage assignment, sibship clustering and beyond. *Mol Ecol Resour*. 2017; 17(5):1009–24. <https://doi.org/10.1111/1755-0998.12665> PMID: 28271620
77. Sabeti PC, Varilly P, Fry B, Lohmueller J, Hostetter E, Cotsapas C, et al. Genome-wide detection and characterization of positive selection in human populations. *Nature*. 2007; 449(7164):913–8. <https://doi.org/10.1038/nature06250> PMID: 17943131
78. Criscione CD, Valentim CL, Hirai H, LoVerde PT, Anderson TJ. Genomic linkage map of the human blood fluke *Schistosoma mansoni*. *Genome Biol*. 2009; 10(6):R71. <https://doi.org/10.1186/gb-2009-10-6-r71> PMID: 19566921
79. Browning SR, Browning BL. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am J Hum Genet*. 2007; 81(5):1084–97. <https://doi.org/10.1086/521987> PMID: 17924348
80. Szpiech ZA, Hernandez RD. selscan: an efficient multithreaded program to perform EHH-based scans for positive selection. *Mol Biol Evol*. 2014; 31(10):2824–7. <https://doi.org/10.1093/molbev/msu211> PMID: 25015648
81. Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res*. 2019; 47(W1):W191–W8. <https://doi.org/10.1093/nar/gkz369> PMID: 31066453
82. Walker M, Churcher TS, Basanez MG. Models for measuring anthelmintic drug efficacy for parasitologists. *Trends Parasitol*. 2014; 30(11):528–37. <https://doi.org/10.1016/j.pt.2014.08.004> PMID: 25217844
83. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*. 2015; 4:7. <https://doi.org/10.1186/s13742-015-0047-8> PMID: 25722852
84. Vauterin P, Jeffery B, Miles A, Amato R, Hart L, Wright I, et al. Panoptes: web-based exploration of large scale genome variation data. *Bioinformatics*. 2017; 33(20):3243–9. <https://doi.org/10.1093/bioinformatics/btx410> PMID: 29028261
85. Johnson NA, Lachance J. The genetics of sex chromosomes: evolution and implications for hybrid incompatibility. *Ann N Y Acad Sci*. 2012; 1256:E1–22. <https://doi.org/10.1111/j.1749-6632.2012.06748.x> PMID: 23025408
86. Walker M, Mabud TS, Oliario PL, Coulibaly JT, King CH, Raso G, et al. New approaches to measuring anthelmintic drug efficacy: parasitological responses of childhood schistosome infections to treatment with praziquantel. *Parasit Vectors*. 2016; 9:41. <https://doi.org/10.1186/s13071-016-1312-0> PMID: 26813154
87. Hotez PJ, Molyneux DH, Fenwick A, Ottesen E, Ehrlich Sachs S, Sachs JD. Incorporating a rapid-impact package for neglected tropical diseases with programs for HIV/AIDS, tuberculosis, and malaria. *PLoS Med*. 2006; 3(5):e102. <https://doi.org/10.1371/journal.pmed.0030102> PMID: 16435908
88. Inobaya MT, Olveda RM, Chau TN, Olveda DU, Ross AG. Prevention and control of schistosomiasis: a current perspective. *Res Rep Trop Med*. 2014; 2014(5):65–75. <https://doi.org/10.2147/RRTM.S44274> PMID: 25400499
89. Semenya AA, Sullivan JS, Barnwell JW, Secor WE. *Schistosoma mansoni* infection impairs antimalaria treatment and immune responses of rhesus macaques infected with mosquito-borne *Plasmodium coatneyi*. *Infect Immun*. 2012; 80(11):3821–7. <https://doi.org/10.1128/IAI.00590-12> PMID: 22907811
90. French MD, Evans D, Fleming FM, Secor WE, Biritwum NK, Brooker SJ, et al. Schistosomiasis in Africa: Improving strategies for long-term and sustainable morbidity control. *PLoS Negl Trop Dis*. 2018; 12(6):e0006484. <https://doi.org/10.1371/journal.pntd.0006484> PMID: 29953454
91. Mawa PA, Kincaid-Smith J, Tukahebwa EM, Webster JP, Wilson S. Schistosomiasis Morbidity Hot-spots: Roles of the Human Host, the Parasite and Their Interface in the Development of Severe Morbidity. *Front Immunol*. 2021; 12:635869. <https://doi.org/10.3389/fimmu.2021.635869> PMID: 33790908
92. Valentim CL, Cioli D, Chevalier FD, Cao X, Taylor AB, Holloway SP, et al. Genetic and molecular basis of drug resistance and species-specific drug action in schistosome parasites. *Science*. 2013; 342(6164):1385–9. <https://doi.org/10.1126/science.1243106> PMID: 24263136
93. Colley DG, Bustinduy AL, Secor WE, King CH. Human schistosomiasis. *Lancet*. 2014; 383(9936):2253–64. [https://doi.org/10.1016/S0140-6736\(13\)61949-2](https://doi.org/10.1016/S0140-6736(13)61949-2) PMID: 24698483

94. Aragon AD, Imani RA, Blackburn VR, Cupit PM, Melman SD, Goronga T, et al. Towards an understanding of the mechanism of action of praziquantel. *Mol Biochem Parasitol.* 2009; 164(1):57–65. <https://doi.org/10.1016/j.molbiopara.2008.11.007> PMID: 19100294
95. Xiao SH, Sun J, Chen MG. Pharmacological and immunological effects of praziquantel against *Schistosoma japonicum*: a scoping review of experimental studies. *Infect Dis Poverty.* 2018; 7(1):9. <https://doi.org/10.1186/s40249-018-0391-x> PMID: 29409536
96. Aura CM, Musa S, Nyamweya CS, Ogari Z, Njiru JM, Hamilton SE, et al. A GIS-based approach for delineating suitable areas for cage fish culture in a lake. *Lakes and Reservoirs.* 2021; 26:e12357.
97. Kumkate S, Chunchob S, Janvilisri T. Expression of ATP-binding cassette multidrug transporters in the giant liver fluke *Fasciola gigantica* and their possible involvement in the transport of bile salts and anthelmintics. *Mol Cell Biochem.* 2008; 317(1–2):77–84. <https://doi.org/10.1007/s11010-008-9833-2> PMID: 18543082
98. Doenhoff MJ, Cioli D, Utzinger J. Praziquantel: mechanisms of action, resistance and new derivatives for schistosomiasis. *Curr Opin Infect Dis.* 2008; 21(6):659–67. <https://doi.org/10.1097/QCO.0b013e328318978f> PMID: 18978535
99. de Moraes MC, Cardoso CL, Cass QB. Immobilized purine nucleoside phosphorylase from *Schistosoma mansoni* for specific inhibition studies. *Anal Bioanal Chem.* 2013; 405(14):4871–8. <https://doi.org/10.1007/s00216-013-6872-7> PMID: 23535739
100. Doyle SR, Tracey A, Laing R, Holroyd N, Bartley D, Bazant W, et al. Genomic and transcriptomic variation defines the chromosome-scale assembly of *Haemonchus contortus*, a model gastrointestinal worm. *Commun Biol.* 2020; 3(1):656. <https://doi.org/10.1038/s42003-020-01377-3> PMID: 33168940
101. Doyle SR, Illingworth CJR, Laing R, Bartley DJ, Redman E, Martinelli A, et al. Population genomic and evolutionary modelling analyses reveal a single major QTL for ivermectin drug resistance in the pathogenic nematode, *Haemonchus contortus*. *BMC Genomics.* 2019; 20(1):218. <https://doi.org/10.1186/s12864-019-5592-6> PMID: 30876405
102. Doyle SR, Laing R, Bartley D, Morrison A, Holroyd N, Maitland K, et al. Genomic landscape of drug response reveals novel mediators of anthelmintic resistance. *bioRxiv.* 2021:465712.
103. Cotton JA, Berriman M, Dalen L, Barnes I. Eradication genomics—lessons for parasite control. *Science.* 2018; 361(6398):130–1. <https://doi.org/10.1126/science.aar6609> PMID: 30002241
104. Grimes JE, Croll D, Harrison WE, Utzinger J, Freeman MC, Templeton MR. The roles of water, sanitation and hygiene in reducing schistosomiasis: a review. *Parasit Vectors.* 2015; 8:156. <https://doi.org/10.1186/s13071-015-0766-9> PMID: 25884172
105. King CH, Bertsch D. Historical perspective: snail control to prevent schistosomiasis. *PLoS Negl Trop Dis.* 2015; 9(4):e0003657. <https://doi.org/10.1371/journal.pntd.0003657> PMID: 25905621
106. Chevalier FD, Le Clec'h W, McDew-White M, Menon V, Guzman MA, Holloway SP, et al. Oxamniquine resistance alleles are widespread in Old World *Schistosoma mansoni* and predate drug deployment. *PLoS Pathog.* 2019; 15(10):e1007881. <https://doi.org/10.1371/journal.ppat.1007881> PMID: 31652296
107. Hermisson J, Pennings PS. Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics.* 2005; 169(4):2335–52. <https://doi.org/10.1534/genetics.104.036947> PMID: 15716498
108. Stothard JR, Sousa-Figueiredo JC, Navaratnam AM. Advocacy, policies and practicalities of preventive chemotherapy campaigns for African children with schistosomiasis. *Expert Rev Anti Infect Ther.* 2013; 11(7):733–52. <https://doi.org/10.1586/14787210.2013.811931> PMID: 23879611