This is the Accepted Manuscript of the following article:


The final version is available online via http://dx.doi.org/10.1016/bs.apar.2014.12.006.

© 2015. This manuscript version is made available under the CC-BY-NC-ND 4.0 license http://creativecommons.org/licenses/by-nc-nd/4.0/.

The full details of the published version of the article are as follows:

TITLE: Modelling the Effects of Mass Drug Administration on the Molecular Epidemiology of Schistosomes

AUTHORS: Lamberton, PHL and Crellen, T and Cotton, JA and Webster, JP

JOURNAL TITLE: Advances in Parasitology

VOLUME/EDITION: 87

PUBLISHER: Elsevier

PUBLICATION DATE: 7 February 2015 (online)

DOI: 10.1016/bs.apar.2014.12.006
Running Title: Mathematical modelling for molecular epidemiology

Modelling the effects of mass drug administration on the molecular epidemiology of schistosomes

Lamberton, P.H.L.¹, Crenelle, T.¹², Cotton, J.A.², & Webster, J.P.¹³*

1. Department of Infectious Disease Epidemiology, School of Public Health, Imperial College Faculty of Medicine, St Mary’s Campus, London W2 1PG, UK
2. Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge. CB10 1SA, UK
3. Department of Pathology and Pathogen Biology, Centre for Emerging, Endemic and Exotic Diseases (CEEEED), Royal Veterinary College, University of London, AL9 7TA, UK

*corresponding author. (jowebster@rvc.ac.uk).
Abstract
As national governments scale up mass drug administration (MDA) programs aimed to combat neglected tropical diseases (NTDs), novel selection pressures on these parasites increase. To understand how parasite populations are affected by MDA and how to maximise the success of control programmes, it is imperative for epidemiological, molecular and mathematical modelling approaches to be combined. Modelling of parasite population genetic and genomic structure, particularly of the NTDs, has been limited through the availability of only a few molecular markers to date. The landscape of infectious disease research is being dramatically reshaped by Next Generation Sequencing (NGS) technologies and our understanding of how repeated selective pressures are shaping parasite populations is radically altering. Genomics can provide high resolution data on parasite population structure, identifying how loci may contribute to key phenotypes such as virulence and/or drug resistance. We discuss the incorporation of genetic and genomic data, focusing on the recently sequenced *Schistosoma* spp., into novel mathematical transmission models to inform our understanding of the impact of MDA and other control methods. We summarise what is known to date, the models that exist and how population genetics has given us an understanding of the effects of MDA on the parasites. We consider how genetic and genomic data have the potential to shape future research, highlighting key areas where data are lacking, and how future molecular epidemiology knowledge can aid understanding of transmission dynamics and the effects of MDA, ultimately informing public health policy makers of the best interventions for NTDs.

**Key words:** mathematical models, molecular epidemiology, genetics, genomics, transmission dynamics, population structure, genetic diversity, schistosomiasis
Introduction

Mass drug administration (MDA) is the recommended strategy of the World Health Organization (WHO) to control or, in certain cases, eliminate a sub-group of Neglected Tropical Diseases (NTDs) based on the annual distribution of inexpensive and/or donated drugs (Lammie et al., 2006, Webster et al., 2014). Through MDA, intensive and prolonged selective pressures are, and will be, placed on these parasites, which may have implications for the long-term success of campaigns (Webster et al., 2008, Webster et al., 2014). Policies to optimise success become crucial as programmes in selected countries shift from morbidity control toward elimination as a public health burden (WHO, 2012).

Mathematical models have the potential to explore a number of epidemiological features of both interest and importance in parasite populations (Levin et al., 1997, Levin, 1992, Anderson and May, 1991). Mathematical models can provide profound logistical, financial, temporal and biological benefits through, for instance, testing potential disease control strategies prior to final design and implementation (Rivers et al., 2014) or to improve current control methods (Turner et al., 2014, Luz et al., 2011). However, valuable models require accurate parameterisations and a reliable understanding of the disease dynamics under investigation and how these can, and do, respond to differential selection pressures. Unfortunately, to date, few strong predictive mathematical models exist for the majority of metazoan parasites under MDA pressure. This is linked to a wider neglect in the study of helminth biology and genetics (Prugnolle et al., 2005a) in comparison to protozoan and bacterial agents such as malaria and tuberculosis (Hotez and Pecoul, 2010). In the case of several of the NTDs, the extent to which mathematical models may be developed and influence public-health policy is limited by this absence of biological data on parasite population structure and genetic diversity. Substantial progress is occurring particularly with the recent establishment of the NTD modelling consortium (NTD_Modelling_Consortium). However, what remains to be addressed are the key data that these models require, particularly on the complex transmission dynamics of these often multi-host diseases.

The landscape of infectious disease biology is being radically restructured through the advancement of population genetic techniques, the development of Next Generation Sequencing (NGS) and the abundance of bioinformatics data generated from whole genome sequencing (WGS) (Luikart et al., 2003, Oleksyk et al., 2010, Vitti et al., 2013). Through applying NGS technology to parasite population studies, well designed research can reap the benefits of sequence data, such as locating regions of the genome under selection (Valentim et al., 2013), and understanding levels and directions of gene flow and infections within and between sub-populations (Kao et al., 2014). Such information will boost our capacity to model and predict the effects of MDA on disease transmission dynamics and selection and how best to maximise control success (Kim et al., 2014).

Within infectious disease epidemiology, bacteriology has been one of the disciplines to benefit earliest from the genomic revolution due to the low cost of sequencing smaller bacterial genomes. Recent work on drug resistant Mycobacterium tuberculosis, for instance, has used large datasets to identify key targets of selection (Farhat et al., 2013, Zhang et al., 2013), to show how antibiotic resistance spreads at different spatial scales (Harris et al., 2010), and quantifying how often resistant clones spread between animal and human populations (Mather et al., 2013). Applying WGS to eukaryotic parasitic organisms is complicated by the increased size and complexity of genomes. Population genomics approaches in protozoans such as Plasmodium falciparum (Cheeseman et al.,
2012, Takala-Harrison et al., 2013) and Leishmania donovani (Downing et al., 2011), have been driven by the need to identify the locus or loci responsible for drug resistance (Bright and Winzeler, 2013). Within parasitology, malaria research has paved the way in WGS and NGS (Volkman et al., 2012) since the P. falciparum genome was published over a decade ago (Gardner et al., 2002). For example, a major locus responsible for artemisinin resistance in P. falciparum (PF3D7_1343700) (Ariey et al., 2014), was identified using WGS of both a laboratory-induced resistant line and field isolates, and clear-cut examples of natural selection associated with resistance to artemisinin combination therapies in Asia have come from analysis of 91 P. falciparum genomes (Cheeseman et al., 2012).

Despite the recent advances in WGS and NGS, the challenges of genomic analyses of metazoan species are, however, still great. The S. mansoni genome of 380 megabases (Mb) (Protasio et al., 2012) is over an order of magnitude larger than the 23 Mb genome of P. falciparum (Gardner et al., 2002) and one order again than the 4 Mb genome of M. tuberculosis (Cole et al., 1998). Under the 50 Helminth Genomes Initiative of the Wellcome Trust Sanger Institute (WTSI) and other projects elsewhere (Blaxter et al., 2012), further helminth draft genomes are becoming available. While sequencing output is increasing, large genomes are fragmented (Parkhill, 2013) and stitching these genomes together into complete scaffolds took years (Holroyd and Sanchez-Flores, 2012). Nevertheless, changes in sequencing and mapping technologies are gradually making this process easier and more rapid (Chin et al., 2013, Dong et al., 2013), and the Schistosoma genome is starting to be used to study anthelminthic resistance. For example, WGS facilitated the finding of the locus responsible for oxamniquine resistance in Schistosoma mansoni (Smp_089320) by providing an easy means to generate SNPs between parental lines in a staged genetic cross (Valentim et al., 2013).

As genomic data becomes available and costs of WGS reduce, questions on how to maximise the information from such data become pertinent. Here we consider what is known today about the interdisciplinary approaches of molecular epidemiology and mathematical modelling. We discuss the current and newly developing techniques available in population genetics and genomics and how these could and should be incorporated into novel mathematical predictive models, primarily using Schistosoma spp. as a case study. Whilst many excellent mathematical models already exist on schistosome epidemiology, beyond the scope of this current article, here we focus on those incorporating genetics and genomics. We specifically address what information is urgently needed to parameterize mathematical models aimed to help us elucidate: i) key questions relating to the basic biological and transmission potential of schistosome populations under differing MDA pressures; ii) transmission cut-off points under repeated MDA informing on when treatments can be stopped, and/or scaled down without population infection recrudescing; iii) the potential gauge of when and if drug resistance will arise and how, where and how fast it may spread; iv) the role of non-human zoonotic reservoirs in maintaining transmission under differential selective pressures; and v) the potential for novel introgressed hybrid schistosomes to emerge and establish. We conclude by highlighting priority areas for future cross-disciplinary research in this emerging area and how findings can be translated into policy.

Schistosomes as a ‘model’ for anthelminthic MDA

Schistosomes, the causative agent of schistosomiasis, are amongst one of the major NTDs targeted by MDA. Schistosoma spp. are indirectly-transmitted dioecious trematode macroparasites, with an asexual stage in an
intermediate host snail and a sexual stage within a definitive mammalian host. Schistosomiasis infects >200 million people with >750 million at risk, of which over 90% are within sub-Saharan Africa (SSA) (Steinmann et al., 2006). Two forms of human schistosomiasis are recognised - urogenital and intestinal. Urogenital schistosomiasis, from infection by *Schistosoma haematobium*, is associated with haematuria, bladder damage and a risk of progression to kidney failure and bladder cancer. More than 150,000 people die annually from *S. haematobium*-related kidney failure. Intestinal schistosomiasis is predominantly caused by *S. mansoni* in SSA and *S. japonicum* in Asian foci, and is associated with bloody diarrhoea, hepato-splenomegaly and liver failure (Gryseels et al., 2006). Schistosomiasis is primarily treated and controlled through MDA with praziquantel (PZQ) to school-age children in endemic districts. The Schistosomiasis Control Initiative (SCI), based at Imperial College London, has been instrumental in providing over 100 million praziquantel (PZQ) treatments across parts of SSA from 2003-2014 (Fenwick et al., 2009, Webster et al., 2014). Vast numbers of the drug are donated through private-public partnerships including by pharmaceutical companies such as Merck-KGaA who have pledged to increase their donations and provide 250 million PZQ tablets annually for SSA by 2016. As treatment efforts are stepped-up, stronger selection pressures will be imposed on populations of *Schistosoma* spp. (Webster et al., 2008, Webster et al., 2014).

The genomes of the three major schistosome species of humans have been published - *S. mansoni* (Berriman et al., 2009), *S. haematobium* (Young et al., 2012) and *S. japonicum* (Schistosoma japonicum Genome and Functional Analysis, 2009), and have been made accessible through SchistoDB (Zerlotini et al., 2013). Unpublished draft genome data for a further eight *Schistosoma* spp. are also now available from the Wellcome Trust Sanger Institute ([http://www.sanger.ac.uk/research/initiatives/globalhealth/research/helminthgenomes/](http://www.sanger.ac.uk/research/initiatives/globalhealth/research/helminthgenomes/)). While these data have already been put to use in studies on functional (Protasio et al., 2013, Valentim et al., 2013) and comparative genomics (Tsai et al., 2013), work on population genomics is in its relative infancy.

Microsatellite studies have instead been the mainstay of recent *Schistosoma* evolutionary analysis, including population genetic studies examining potential changes to microsatellite diversity of *S. mansoni* before and after PZQ MDA. A key 2005-6 study in Tanzania (Norton et al., 2010), for instance, revealed a significant reduction in the allelic richness of *S. mansoni* following PZQ, within individuals treated, but most notably at the population level, as revealed by a similar reduction in genetic diversity observed in the youngest cohort of previously untreated children. Data from Senegal collected in 2007-8 in a similar study used microsatellites to analyse miracidia collected from 12 children pre and post treatment/s. The authors, however, found no significant changes to allelic richness or expected heterozygosity in *S. mansoni* before and after two rounds of treatment with PZQ amongst this small sample group (Huysse et al., 2013). In addition no reduction in genetic diversity following four years of MDA was observed in parasites from children in Western Kenya (L elo et al., 2014). Likewise a microsatellite study in Brazil (Blanton et al., 2011) which compared the similarity of *S. mansoni* populations that survived PZQ treatment with susceptible worms also found no significant difference according to the differentiation index (Jost, 2008). Such inconsistencies could be due to differences in sample sizes used or due to true biological differences. Future molecular and mathematical modelling analyses would help elucidate this, by, for instance, estimating the minimum number of miracidia per individual required and/or the number of individual host samples required to provide robust estimates of genetic diversity (French et al., 2012). It should also be acknowledged that the number of microsatellites needed to accurately infer an effect, varies greatly with the nature
of the question, the scale of the analyses, the specific parameters to be estimated as well as the sample size. Any apparent absence of differentiation or change at neutral markers is not, however, necessarily indicative of whole genome processes (Allendorf et al., 2010). A study on Atlantic cod (Pampoulie et al., 2006), for instance, showed that very little genetic differentiation \((F_{st})\) was observed over nine microsatellite loci, but that substantial variation was taking place at the \(Panl\) locus, which is known to be under natural selection. The level of detail that genomics can now provide, particularly in terms of parasite population structure, transmission dynamics, and genetic diversity and how these are affected by MDA, as well as the more specific effects of treatment on drug resistance and other genetic traits such as virulence, will greatly aid parameterisation of transmission models thereby helping policy makers formulate informed decisions on how to maximise control of these NTDs while minimising the risk of development and/or spread of drug resistance.

**Micro-evolutionary processes**

Some of the most genetically explicit models for multi-host parasites focus on basic population biology (Prugnolle et al., 2005a, Criscione et al., 2005). Analyses of genetic variation in parasites at different hierarchical levels enables elucidation of parameters such as gene flow, effective population sizes and breeding units, all information relevant for understanding the potential rate of spread of important traits such as drug resistance (Anderson and May, 1991). Schistosomes have, however, several rather unique aspects of their biology, separating them from other trematodes. Prugnolle and colleagues used an infinite island model to explore the alternation of sexual and asexual reproduction in monoecious trematodes on the partitioning of genetic variance among and within definitive hosts (Prugnolle et al., 2005a). Variation in reproductive success of clones was found to be important in shaping the distribution of the genetic variability both within and among definitive hosts (Prugnolle et al., 2005a, Prugnolle et al., 2005b) and \(F_{st}\) (a measure of inbreeding) increased with higher levels of self-fertilisation (Prugnolle et al., 2005a), limiting the scope for gene flow within these populations. Schistosomes are believed to have no or low levels of inbreeding (Basch and Basch, 1984, Prugnolle et al., 2005c, Huyse et al., 2009), and hence it might be predicted that \(F_{st}\) values would be lower for these organisms than for monoecious trematodes, with future models fitted to field molecular data helping to elucidate this.

The same group parameterised a model for dioecious trematodes, and compared their theoretical findings with empirical \(S. mansoni\) data from Guadeloupe (Prugnolle et al., 2005b). These models examined differential life-history traits such as sex biased dispersal or clonal reproductive success (Prugnolle et al., 2002). Such characterizing of host and parasite population genetic structure and estimating gene flow among populations is essential for understanding co-evolutionary interactions between hosts and parasites (Prugnolle et al., 2005c). Their models, however, assume non-overlapping generations. Adult schistosomes can live for, on average 6-15 and up to 40 years within a human host, and as hosts are repeatedly exposed over their lifetime, generations of the adult schistosome populations can and do overlap. It is unknown what the effects of overlapping generations are with regards to population structure among hosts and if overlapping generations might impact the genetic effects of drug treatment. Inclusion of over-lapping generations in mathematical models would undoubtedly increase the mathematical complexity (Prugnolle et al., 2005a), but may be in many cases the most biologically realistic scenario. In addition it is known that schistosome males can competitively mate (Tchuem Tchuenté et al., 1995, Tchuem Tchuenté et al., 1996, Webster et al., 2007) and therefore a male which survived treatment could mate with a surviving female, and/or including any female protected from the drug by a dying male within one
generation. New males from subsequent generations could also competitively mate with females from earlier generations, leading to pairs from these overlapping generations. Such competition among males will lead to an increased variance in male reproductive success. This could plausibly be higher with drug treatment, in comparison to natural death and generation turnover, as the new cohorts of males infecting the hosts post treatment may be fitter than those which have been exposed to the drugs, but could mate with previously exposed, but surviving females. The inclusion of overlapping generations in mathematical model design could thus have profound implications on the predicted spread of drug resistance via sexual reproduction. Work by Xu et al (2006) has indeed demonstrated the importance of incorporating mating structure into model design, where he showed the potential maintenance of drug-resistant strains of schistosomes where generations overlap in comparison to simpler models without mating structure. Furthermore, such mating structure models suggested that multiple strains of drug-resistant parasites are likely to be favoured as the treatment rate increases (Xu et al., 2006). Models produced to date have also suggested that the likelihood that these resistant strains will increase in frequency also depends on the interplay between their relative fitness, the costs of resistance, and the degree of selection pressure by the drug treatments (Feng et al., 2001).

**Macro-evolutionary processes**

Molecular models for macro-epidemiological processes directly relevant to public health policy makers are scarce, despite the fact that population genetics and epidemiology both extend basic biological processes at the individual level to the population level, and clearly come together to model drug resistance (Levin et al., 1997). The difference in the spread of drug resistant genes through a parasite population to that of drug resistant parasites through a host population, is that drug resistant alleles replace sensitive alleles depending on their relative fitness, whilst drug resistant parasite densities increase according to their absolute fitness (Paterson and Viney, 2000). This difference is of greater importance in complex parasite systems where high levels of genetic diversity are required to complete the life-cycle, such as schistosomes (Rollinson et al., 2009), with stronger constraining selective pressures acting on the parasites at different life-cycle stages than, for example, directly transmitted viral pathogens. Because the rate of increase of a parasite population (and therefore the density of the parasite population) depends on R0, the epidemiology of anthelmintic resistance cannot be determined without using a model which incorporates the underlying genetics of resistance (Paterson and Viney, 2000). Smith and colleagues used estimates of the absolute fitness of different parasite genotypes within an epidemiological framework to model the effects of under dosing, treatment strategies and mating probabilities on anthelmintic resistance (Smith et al., 1999). Such a model demonstrates how population genetics can help build theoretical models of infectious disease to understand patterns of transmission in the field.

**The role of genetics and genomics in mathematical models in order to:**

**Elucidate the basic biology and transmission potential of schistosomes**

*Population structure*
In many countries the majority of MDAs, particularly for schistosomiasis and soil-transmitted helminths (STH), target school-aged children. The effects that such treatment strategies have on parasite transmission dynamics, as well as the potential selection pressures, depend on a range of factors. These include at least three directly relating to refugia (the proportion of the parasite population not exposed to the drug): 1) the proportion of school-aged children who attend school and therefore receive treatment, 2) the proportion of the population that are school-aged, and 3) the proportion of the total parasite population harboured by this targeted treatment group. Additional key factors include drug efficacy, life-history costs/trade-offs of resistance, and parasite population diversity, structure and transmission dynamics, such as how different age groups are exposed to eggs or infective larvae produced by these school-aged children and vice versa (Anderson et al., 2013). These factors affect on-going transmission and re-infection of treated, and untreated, individuals as well as influencing the potential development, rate, and spread, of drug-resistant strains.

Models can aid understanding about potential benefits that school-based MDA may have on untreated groups. For example, if parasite population structure, elucidated through genetics and/or genomics, indicates that children and adults were found to be infected from a similar genetic pool of parasites (i.e. no genetic differentiation of parasites between children and adults), then treating children may potentially decrease infection in the untreated adult communities, recently termed the ‘herd impact’ of a treatment programme (Anderson et al., 2013). Conversely, if parasites in children only tend to circulate within their age groups, then infection intensities in adults will be unaffected by only treating the children.

Although measuring transmission rates in multi-host systems is difficult, data on genetic variation within an individual and population level hierarchy can also enable measurements of genetic structure and associated rates of gene flow (Weir and Cockerham, 1984). Genomic and gene flow data can likewise be used to build transmission trees, including those aimed to identify, for certain micro-parasites at least, the origins of any new infections (Kao et al., 2014). While these approaches may have lower resolution in metazoan parasites, due to their slower rates of molecular evolution, multi-locus and genomic approaches will allow us to approximate intraspecific phylogenies and elucidate transmission between hosts, as has been demonstrated for Ascaris (Criscione et al., 2010). Such data become vitally important as control programmes move towards elimination, highlighting key individuals or sub-populations driving re-infections. Examples of where genetic data has already informed public health policy include a study on Ascaris in pigs and humans in Guatemala, where even though infections were sympatric, there appeared to be, using the molecular tools available at the time, little gene flow between the parasite populations indicating no transmission between the two host species (Anderson et al., 1993). Conversely two other more recent studies have indicated cross transmission between the Ascaris lumbricoides and A. suum species, with, furthermore, up to 4% and 7% of Ascaris appearing as hybrids, which raises a number of potential implications for long-term evolutionary dynamics (Criscione et al., 2007). Models have already been used in directly transmitted pathogens, including sexually transmitted diseases, where contact tracing data may not be complete, but where genome data can inform on infection networks (e.g. HIV transmission in a dental practise (Ou et al., 1992) as well as theoretical pathogen models using evolutionary trees resulting from different evolutionary processes (Nee et al., 1994). In micro-parasites with rapid in-host evolution sensitivity of phylogenetic based networks may be reduced (Resik et al., 2007). Similar lack of contact tracing exists for indirectly transmitted pathogens such as STHs and schistosomes and micro-parasite models incorporating similar
genomic data may be developed for more complex indirect transmission networks (Gupta et al., 1996) and may be highly dependent on host immunological factors (Anderson et al., 1989).

At present mathematical models of indirectly transmitted parasites often assume that exposure to eggs or larvae across all age groups is random and independent of the relative contribution of infective stages from each age group and maintained transmission. However, the spatial structure of concomitant parasite transmission between age groups is unlikely to be random and even less so with MDA programmes targeting specific groups of individuals. Therefore, models should incorporate such heterogeneous mixing (Chan et al., 1994). An additional complication is that transmission dynamics and population biology are likely to change if MDA reduces parasite transmission significantly (Klepac et al., 2013). In addition, models for such organisms, with the exception of (Gurarie and King, 2005)), also assume that hosts sample from a common source of infection. However, even on a very local scale, parasites have been shown, using population genetics, to have focal transmission (Criscione et al., 2010) independent of the among age-group population structures, negating some of the classic models for parasite transmission.

Above, we have discussed how population genetics can inform on specific treatment scenarios, and be used to predict gene flow between certain individual hosts, or host groups. On a larger and more basic scale, population genetics can inform on the very basis of what types of models should be used for certain diseases. At present some network contact models estimate transmission patterns purely from host behaviour. The use of population genetics and genomics enables elucidation of contact network structures, even at the most basic level. Questions such as whether transmission is random or non-random, and therefore which model is appropriate to fit the data, can be answered, as has already been initiated as for A. lumbricoides in Nepal (Criscione et al., 2010). This is particularly important for multi-host parasites where several transmission mechanisms could result in a similar profile. To date, many of the theoretical challenges faced by epidemiologists and population geneticists were problems of scale, whilst the use of population genomics greatly diminishes these challenges.

**Transmission Rates**

Changes in the transmission rate of macroparasites may be inferred from longitudinal changes in descriptive statistics (prevalence, infection intensity), or modelled through the force-of-infection (FOI: the rate at which human hosts acquire parasites). Uganda, for instance, has been treating individuals with PZQ since 2003 with significant reductions in prevalence, infection intensity and FOI after only three rounds of MDA in low, moderate and high infection areas in both treated and untreated children (French et al., 2010). However, Basáñez and colleagues argue that presently there are very few models with the potential to inform on optimal methods at a clinical or epidemiological level to monitor such changes and they outline a number of areas for future model development. These include “the design of treatment efficacy and effectiveness studies; phenotypic characterisation of responses to treatment; and design of sampling protocols for the study of parasite genetic structure under treatment, thereby facilitating prompt detection of anthelmintic resistance” (Basáñez et al., 2012). Examples of modelling work already carried out in this field include studies examining the effect of density-dependent forces known to act on parasites generally, including schistosomes (Medley and Anderson, 1985) and the effect of human and parasite sample sizes on measuring reductions in genetic diversity in *S. mansoni* infections.
post treatment (French et al., 2012). Both of these factors are vitally important, particularly as control programmes progress and infection intensities reduce.

Genomic techniques can be employed to gain a better understanding of the nature of transmission in endemic settings, including those under differing MDA pressures (Volkman et al., 2012). Within a parasite population, the degree of outbreeding and recombination is linked to the proportion of people harbouring parasites of variable genotypes; this proportion is known to scale with the level of transmission (Anderson et al., 2000). A study on the Thai-Burma border where transmission of P. falciparum has declined from 2000-2010 associated a loss of heterozygosity across 96 SNPs with the sampling year (Nkhoma et al., 2013) using logistic regression, therefore multiple genotype infections had significantly declined over time in line with falling transmission. A study in Senegal found a similar reduction in the heterozygosity of P. falciparum populations as a consequence of public health interventions (Daniels et al., 2013). Nkhoma and colleagues argue that the linear relationship between the carriage of multiple parasite genotypes within human hosts and transmission intensity means that population genetics can be used as a reliable and inexpensive method of tracking temporal changes in transmission intensity (Nkhoma et al., 2013). Despite very different lifecycles, the effects seen in P. falciparum are likely to be replicated in Schistosoma spp. In both parasites, the total populations are subdivided within each definitive host (Anopheles in the case of P. falciparum and humans in the case of Schistosoma spp.), meaning that sexual reproduction for any individual parasite is possible only with a small fraction of organisms from the total population. This presents barriers to gene flow and outbreeding and so the reductions in heterozygosity are enhanced in both species. It may be expected to occur more slowly in schistosomes, due to a longer generation time, the relatively higher intensities of infections in humans than in the mosquitoes, and the different timing of sexual reproduction in the two lifecycles.

Spatial heterogeneity

Modelling spatial and demographic heterogeneity with the aim of understanding fundamental processes underlying infection dynamics provides a framework for evaluating potential control strategies for infectious diseases (Paterson and Viney, 2000). Using well-documented and commonly modelled systems, such as measles, social and geographical structure in contact networks have been deduced with heterogeneity in transmission rates within families, between children at school and between communities (Keeling et al., 1997). Models for macroparasites may also highlight key social and geographical groups where transmission is high. However, in indirect life-cycles, this is hard to do from host behaviour and contact monitoring. Parasite population genetics and genomics can ultimately inform on transmission structure. Heterogeneity in infection patterns are biological realities and must be incorporated into models (Paterson and Viney, 2000) improving their fit to empirical data. Kao et al (2014) review the use of WGS in contact tracing models to reveal points of control and predict directions of spread of diseases for microparasites (Kao et al., 2014). They discuss the complexities associated with inferring the epidemiological dynamics of multi-host pathogens, as is often the case for NTDs. Furthermore these authors explore the difficulties of when microparasite mutation rates are low in comparison to generation times, but that such situations may be resolved using within-host genetic variation to infer properties of between-host transmission (Stack et al., 2013). While these methods may be less powerful in metazoan parasites, whole-genome
data will certainly reveal significant details of the spatial genetic structure for these organisms (Archie et al., 2009).

**Evaluate transmission cut off points and estimating when to stop MDA**

Population genetic mathematical models can also inform on cut-off points for MDA through predicting the levels of infection where disease transmission should not recrudesce (Plaisier et al., 1997, Turner et al., 2013). This has been demonstrated for the NTD trachoma (*Chlamydia trachomatis*) in communities where MDA of antibiotic eye ointment are used (Lietman et al., 2011). This model included estimates of transmission parameters relating to reinfection from both within or outside the community. The values were robustly estimated from prevalence data at baseline and 24 month follow up. Nevertheless, future additional data on the level of gene flow over time and space and phylogenies of the infections appearing following treatment would enable control programmes to know the true influence of recrudescence and reinfection from within or outside the community. Such data would also inform on the distance that re-infection can occur from aiding policy makers in their decisions on when they can halt treatment in a central focus, depending on threshold infection levels in a range of surrounding foci. Population genomic data could also parameterise smaller spatial household models, identifying the extent of gene flow within and between infected sub-populations (Blake et al., 2009). Similar transmission potential parameters could be incorporated into other vector-borne/intermediate host disease elimination models, such as for *S. haematobium* on Zanzibar for example, where they are aiming toward elimination (Rollinson et al., 2013). Of particular value would be the use of genetic and genomic data to understand if any detected cases remaining or reappearing were from local transmission or imported cases from the mainland. Such genomic data has already enabled reconstruction of transmission trees for directly transmitted non NTDs, for example, understanding cross-species transmission in *Salmonella* (Mather et al., 2013).

Parasite diversity may also strongly affect transmission dynamics, reinfection rates post treatment and associated thresholds for ending MDA. Strain diversity in trachoma, for instance, is known to affect reinfection levels, as heterogeneity exists in part to evade the human immune system. If many strains become eliminated, then the remaining ones may not be able to reach pre-treatment levels due to independent factors such as strain specific host immunity (Zhang et al., 2004). Incorporating diversity into models predicting MDA cut off points, may thereby not only greatly enhance their accuracy, but also aid in the understanding of such interventions on the transmission dynamics of the infectious agent, as has recently demonstrated for dengue (Coudeville and Garnett, 2012).

**Elucidate the potential evolution and spread of praziquantel resistance**

Early detection of anthelminthic resistance is vital for controlling the spread of such genotypes. When (Churcher and Basáñez, 2009) and how (French et al., 2012) best to examine human helminth parasite populations post treatment, have recently been the subject of model-based studies. If PZQ resistance is recessive, drug resistant
alleles could spread through the population to relatively high levels before phenotypic manifestation (Churcher and Basáñez, 2009). There is also evidence for negative density-dependent fecundity in S. mansoni (Medley and Anderson, 1985) and models on helminth life-cycles show that density-dependent fecundity, in comparison to that in parasite establishment or mortality (Churcher and Basáñez, 2008) may facilitate the spread of resistance as parasite population intensities decrease with treatment. One of the most important aspects of mathematical models is their direct use to public health policy makers. A recent model based on antibiotic resistance and drug use and how to communicate the extent of problems to policy makers (Laxminarayan and Klugman, 2011) could be adapted for other infections to maximise the speed of implanting changes in control strategies should they be required.

Some modelling work to date has taken into account the life-cycle of schistosomiasis by using a time-delay function and found that the presence of this delay, i.e. the average time between two adult generations, makes it more likely for resistant strains, for example, to invade and persist in a parasite population (Castillo-Chavez et al., 2008). Furthermore, population genetic analysis of the structure of S. mansoni and S. haematobium diversity across Africa using microsatellite markers, found that, on the basis of population structuring and high genetic diversity, should drug resistance evolve it would be slow to spread through schistosome populations, at least across large scales (Gower et al., 2013), as there were low levels of gene flow observed between samples from different countries (although high levels of gene flow between samples within countries). As with many parasite species, the complex nature of the schistosome life-cycle violates assumptions of the Wright-Fisher (WF) model: with highly overlapping generations due to the long reproductive lifespan of adult schistosomes and non-random mating due to focal transmission and inability to mate with schistosomes in other hosts. These effects make it difficult to make use of genetic variation at neutral sites to understand population size and structure (Balloux and Lehmann, 2012) and to predict how effective natural selection will be in increasing the frequency of positively selected alleles in the face of random drift or constraining selective pressures elsewhere in the complex life-cycle. A population genetic model recently developed for malaria, which aids understanding of the emergence of resistance and its early spread (Kim et al., 2014), in this instance under combination drug therapy (something which is not currently available for schistosome treatment), will likely act as a key foundation for other indirectly transmitted parasite genetic models in the near future.

Recent mathematical models on large-scale PZQ administration have shown that the FOI of S. mansoni is reduced throughout communities; even in untreated PZQ naïve individuals, as control programmes progress (French et al., 2010). Such a FOI however is calculated from humans, through snails and back into humans, with little knowledge on the relative force between individual life-cycle stages. In S. japonicum, an Asian species which has multiple definitive hosts, the driving FOI is the transmission from snail to mammal (Riley et al., 2008), with different mammalian hosts maintaining transmission in different geographical regions (Lu et al., 2010, Rudge et al., 2013). In an S. mansoni focus in Guadeloupe, similar molecular epidemiology studies have indicated that parasite migration is primarily driven by the rodent hosts (Prugnolle et al., 2005c). Little however is known about the relative forces and spatial heterogeneities in S. mansoni and S. haematobium between sub-populations of parasites in humans or on the effect of PZQ. Future work needs to expand current FOI models, to include drug resistant parameters, and determine the driving forces of continued transmission despite repeated drug treatments. As
schistosome control programmes progress, empirical and theoretical data warrant further research on PZQ efficacy, and how best to monitor and evaluate (M&E) disease transmission, to maximise the life span of PZQ.

As some countries or regions in countries push towards elimination (Rollinson et al., 2013), bi-annual treatment with PZQ has been a suggested strategy to further reduce the FOI and halt transmission. One key area that mathematical models should evaluate is the effect of bi-annual MDA on the potential increased selection for resistance, versus the rate of reducing the FOI (French et al., 2010). Models should be fitted to known clearance data post treatment to elucidate the relative contribution of non-clearance versus reinfection in maintaining high infection intensities. These models should then incorporate the currently unknown parameters for PZQ-resistance, starting with a simple single locus model and progressing to multilocus genotypes. This would inform policy makers on the potential risks of increasing treatment frequencies, versus the potential benefits of reducing FOI, for a range of scenarios for potential genotypic resistance to the drug.

In the absence of any current molecular markers for monitoring potential PZQ resistance, or even a full understanding of the molecular mechanisms of PZQ action (Chan et al., 2013), phenotypic tests have been trialled with some success in both the laboratory and the field (Liang et al., 2001, Lamberton et al., 2010). Accurate phenotypic measures of drug efficacy are a vital requirement for comparison of genomic sequences to locate potential genes and/or regions associated with drug resistance. Complexities arise with several chronic macroparasite infections as methods for phenotypically detecting resistance are often not properly standardised. In schistosomes, the WHO designated phenotype for drug tolerance is lowered ‘cure rates’ and/or ‘egg reduction rates’ as measured by Kato-Katz thick smears (Katz et al., 1972) or urine filtration, depending on the species, before and 14-21 days following treatment (WHO 2013). Limitations in the sensitivity of such diagnostic tests (Lamberton et al., 2014) may restrict the inferences that can be drawn from any association testing between polymorphisms and reduced treatment efficacy. An important role for modelling will be to account for the limitations in diagnostic accuracy, which is increasingly performed through latent class analysis (Koukounari et al., 2013, Assefa et al., 2014). The most recent WHO manual for evaluating anthelmintic drug efficacy “tentatively” places the egg reduction rate at approximately 90% for the three main Schistosoma species infecting humans when individuals are treated with 40 mg/kg of PZQ (WHO 2013). On the basis of these criteria, treatment failures with PZQ, although rare, have been reported (Greenberg, 2013) with systematic non-clearers often observed in the field. Warning also comes from veterinary parasitology with anthelmintic resistance an inevitable consequence of mass anthelminthic treatment, with parasites in some regions being resistant to all major drug classes, leading to total anthelminthic failure (Kaplan and Vidyashankar, 2012, Webster et al., 2014), and threatening the profitability of whole sheep farming industries in Australia (Wolstenholme et al., 2004).

Care must, however, be taken in using genetic data from populations of larval, or egg, stages as a proxy for adult worm intensity or reproduction (Criscone and Blouin, 2005), or to measure transmission between hosts (Steinauer et al., 2010, Steinauer et al., 2013). Population genetics and/or genomics have been used to overcome these problems estimating the adult population sizes through kinship analysis to either partition miracidia into sibships or assign miracidia to parents of the parasites under investigation (Blouin, 2003, Jones and Ardren, 2003, Criscione and Blouin, 2005, Steinauer et al., 2013). This has also enabled the accurate incorporation of density-dependent factors into models by drawing comparisons between adult population sizes and egg counts at the individual level. \( N_e \) is an important parameter in evolutionary biology because it quantifies genetic drift, and
crucially with regard to MDA and potential drug resistance, the response to selection (Criscione and Blouin, 2005). $N_e$ has a large influence on the overall level of genetic diversity in populations and selection for drug-resistance alleles might be more efficient in parasite populations with a large $N_e$. The complex life-cycles of many NTDs affect the type of model chosen to estimate the effective population size ($N_e$) (Balloux and Lehmann, 2012) and make it more challenging to collect samples from appropriate life-cycle stages to infer genetic estimates, and to collect the necessary demographic data, such as generation time, that may be needed to augment these genetic estimates. For example, in many NTDs, such as schistosomes, the eggs from sexual reproduction are passed into the external environment, so that offspring from different infrapopulations are mixed every generation. Criscione and Blouin use a model which subdivides breeders into infrapopulations, nested within a component population, to demonstrate basic demographic factors that control $N_e$ in macroparasite species (Criscione and Blouin, 2005). They incorporate incomplete mixing, which increases reproductive success of some infrapopulations and discuss the effects of aggregation and crowding on per capita fecundity, both aspects vitally important in future models on the potential rate of spread of drug resistance. They also demonstrate a pronounced sex-ratio effect on $N_e$ due to separation of individuals among hosts. Such a model would be greatly enhanced by knowledge gained from kinship analysis through population genetics or genomics, briefly mentioned above (Blouin, 2003, Jones and Ardren, 2003), which would inform on reproductive success and density-dependent factors enabling the accurate estimation on the effective number of breeders in each infrapopulation, which could then be incorporated into models and would be even more important after MDA when density-dependent pressures may be reduced on potentially resistant strains.

Apart from details of the population structure and reproductive biology of the parasite population, the speed and extent of spread of drug-resistance alleles will depend on the detailed genetic architecture of drug resistance. Factors include how many loci are involved, how loci interact in determining resistance phenotypes, whether and to what extent the loci are genetically linked and whether resistant alleles are dominant, recessive or something in-between, the relative fitness of resistant genotypes in the presence of drug treatment, the cost of resistance (reduced fitness in the absence of drug treatment), and the degree of selection pressure exerted by treatment on different populations. Basáñez et al (2012) discuss models capturing some or all of these factors for some parasite species, including a number of schistosome-specific models, but nothing is known about loci underlying clinical failure of PZQ treatment or PZQ resistance, let alone any information to parameterise the other factors required for accurate, model-based prediction of the potential for drug resistance.

Using population genomics to understanding the mode of PZQ resistance

Circumstantial evidence suggests that PZQ acts on voltage-gated Ca2+ channels (Doenhoff et al., 2008, Chan et al., 2013). While these insights are important, whole-genome approaches enable an unbiased approach to understanding the genetics of drug resistance, avoiding potential biases from prior assumptions about mechanisms of action and/or resistance embodied in the choice of ‘candidate genes’ for more targeted genotyping approaches. Traditional ‘forward’ genetics approaches based on analysis of a genetic cross are possible in Schistosoma (Criscione et al., 2009) and have been successfully applied to understanding the genetics of drug resistance to Oxamniquine in one isolate (Valentim et al., 2013). Reverse genetic approaches such as transgenesis and RNAi are also available, or under active development for Schistosoma, and while not yet sufficiently reliable for high-throughput for ‘discovery’ of loci underlying resistance phenotypes, such techniques may be valuable in
confirming the importance of candidate loci discovered by other approaches. More fundamentally, these kinds of laboratory genetics approaches allow detailed investigation of causal loci in individual resistant isolates, but unless diverse isolates can be looked at – and relatively few schistosomes isolates are now maintained in laboratories, and once introduced from the field rapid genetic bottlenecks occur (Gower et al., 2007) – these loci may not be responsible for resistance observed in clinical conditions.

A complementary approach is to investigate genetic variation in natural populations. A conceptually simple approach is to look directly for genetic variation that is associated with a quantifiable phenotype, such as drug resistance (Borrman et al., 2013, Takala-Harrison et al., 2013, Zhang et al., 2013). The key difficulties for using such approaches in schistosomes at present lie in the lack of a diagnostic test for reduced PZQ efficacy, in part due to the complexity of the parasite genome and life-cycle and likely polygenic basis for such resistance. An alternative is to look for the impact of selection caused by drug treatment using genome-wide sequence data, an approach that implicitly assumes that chemotherapy is one of the most important selective forces on these genomes. Methods to detect recent selection are looking for the signals of a selective sweep (Smith and Haigh, 1974), in which a polymorphism under selection increases in frequency rapidly in a population, and the speed of this spread is sufficiently fast that an extended contiguous portion of a chromosome will spread through the population alongside the variant under selection, as there has not been enough time for this unit (haplotype) to be broken up through recombination. Different tests look for different patterns in the distribution of genetic variation across the genome and within populations generated by this process – these tests have been reviewed a number of times recently (Volkman et al., 2012, Oleksyk et al., 2010, Vitti et al., 2013). Examples of the patterns are enhanced linkage disequilibrium (LD) and depleted polymorphisms around a region, using tests such as the long-range haplotype test, which in malaria for example have detected drug-resistant loci for chloroquine (Wootton et al., 2002) and pyrimethamine (Nair et al., 2003), or differences in allele frequencies in populations under selection and those without selection. More sophisticated statistical tests build on these approaches e.g. XP-EHH (Sabeti et al., 2007, Grossman et al., 2010). These tests are often performed from in vitro cultures for drug resistance markers, with linkage mapping using laboratory crosses to correlate segregation patterns in the progeny that are associated with the drug resistant phenotype (Ferdig et al., 2004, Hayton and Su, 2008, Sanchez et al., 2011).

Knowledge on the effect of MDA on parasite genotypes, such as virulence and drug resistance, and their potential costs, are required to understand the risk of development and then spread. Understanding gene flow between subpopulations also informs on levels of refugia as touched upon previously, which is a key strategy for minimising the risk of drug resistance developing. It is imperative that we understand the complex interactions between minimising selective pressures, but treating the most heavily infected individuals, whilst maintaining an optimum level of refugia to maximise the short and long-term control of these NTDs.

Models developed for the demographics of human and malaria movement (Pindolia et al., 2013) and the potential spread of drug resistant malaria (Anderson, 2009, Lynch and Roper, 2011) indicate the type of predictive maps that can be produced for indirectly transmitted pathogens on regional (Pearce et al., 2009), national (Pindolia et al., 2013) and international (Lynch and Roper, 2011) scales. Given the often transitory migration lifestyles of many fishermen and their families in areas of high Schistosoma spp. endemicity, monitoring human movement, potentially through national migration statistics could be extrapolated to these metazoan parasite species.
Elucidate the potential role of non-human definitive hosts

Zoonotic reservoirs affect the transmission dynamics of a disease and can limit the extent to which a pathogen can be controlled or potentially eliminated (Taylor et al., 2001). Work on the zoonotic *S. japonicum* offers an example of how inter-disciplinary research with population genetics and modelling can be done successfully. *Schistosoma japonicum* is unique among schistosome species in that it can infect over 40 mammalian hosts and with zoonotic transmission an important factor in its epidemiology (He et al., 2001). Despite a concerted public health effort over many decades, the Chinese government has been unable to eradicate *S. japonicum* (Wang et al., 2008). Population genetic analyses found close phylogenetic relationships between strains of *S. japonicum* in human and rodent hosts in hilly parts of Anhui Province, China using microsatellite markers, whilst schistosomes in the humans in the marshland areas were more closely related to those in bovines, which appear to drive reinfection in those areas (Rudge et al., 2009). Bovines were generally assumed to be the primary reservoir of human infection (McManus et al., 2010), but population genetic analyses greatly increased our understanding of these complex interactions in zoonotic parasites. Knowledge of the genetic similarity and transmission potential between hosts was then used to develop a multi-host model of transmission for *S. japonicum* (Rudge et al., 2013). This showed that rodents were the only hosts with a basic reproductive number (*R₀*) > 1 in hilly regions of Anhui, China driving transmission and reinfection in treated humans. This explained the re-emergence of *S. japonicum* infection in some mountainous areas of China where the disease was thought to have been eliminated. In marshland areas, where bovines *R₀*>1, an earlier government organised cull of cattle may have aided transmission control, but similar extreme methods would not have helped in the hilly regions. Rudge and colleagues used a prevalence model framework which predicted that a reduction in rodent density by around 20% would lead to a 40-50% reduction in *S. japonicum* incidence in humans. This integration of techniques helps inform public health policy makers on how best to maximise control, through giving fixed targets which need to be reached to bring the overall *R₀* below 1 and to halt the re-emergence of the disease. Without such population genetic analyses of the parasite strains circulating in these regions, the source of human reinfection after these mass treatment campaigns would have remained misunderstood, dramatically reducing the cost benefit of specific control methods in these regions. Conversely, controlling infections in these host species may potentially induce atypical ‘refugia’ populations, such as wild animal reservoirs to become more epidemiologically important (Bockarie et al., 2013, Webster et al., 2014).

Elucidating the potential role of hybridisation and introgression

Hybridisation events can have important evolutionary outcomes on species and populations and may improve the fitness of resulting hybrids though the acquisition of adaptive traits. Schistosomatoidae are capable of trans-species mating within the definitive host. Schistosomes are known to hybridize frequently (Steinauer et al., 2010), and reports from Senegal have identified hybridisation between *S. haematobium/ S. mansoni* (Huyse et al., 2013), *S. haematobium /S. curassoni* and *S. haematobium / S. bovis* (Webster et al., 2013) and between *S. mansoni / S. rodhaini* in western Kenya (Steinauer et al., 2008). The implications of these events for MDA programs have yet to be established, but hybridisation may have implications for disease morbidity and drug tolerance and therefore represent a public health concern.
Continued high transmission of *S. haematobium* despite often high MDA coverage have recently been molecularly identified as being due to, at least in part hybrid infections with *S. bovis* (Huyse et al., 2009, Webster et al., 2013). Since these relatively new PZQ selective pressures have been imposed, any endemic equilibriums of co-infections and hybrids may now be changing. Whether or not these hybrids will maintain transmission and/or evolve into a potentially drug-resistant parasite species remains to be elucidated, however mathematical models for *S. haematobium* and *S. mattheei* indicate that such evolution is highly unlikely (Kruger, 1990), although this was prior to the introduction of MDA and did not take these extra selective pressures into account. Successful modelling work in the future will require both a better knowledge of the transmission dynamics of schistosomes under MDA pressure, taking into account factors such as hybridisation and zoonotic hosts, along with a better understanding of the genetic loci under positive selection for PZQ resistance.

Although not based on parasitic life-cycles, models on the origin of species by sympatric speciation (Dieckmann and Doebeli, 1999), with explicit description of genetic determinism, may act as starting points for the development of models which incorporate hybridization and/or drug resistant alleles. Dieckmann and Doebeli (1999) show that sympatric speciation is a likely outcome of competition for resources, where individuals mate preferentially with like individuals, a situation exacerbated in infections where parasites surviving treatment are more likely to mate with others also surviving either due to potentially harbouring resistant alleles, or increase resistance through hybrid vigour. Such scenarios could theoretically lead to increased drug resistance despite co-infections with susceptible parasites, and may ultimately lead to isolation of drug resistant strains within individuals harbouring only those parasites which survive treatment.

Research into adaptive evolution may also act as a useful springboard for future models discussed here. Models on plant parasites, for instance, have indicated that evolutionary divergence of parasite phenotypes can be driven by seasonal transmission and associated fitness trade-offs (Hamelin et al., 2011), such as may also occur with annual treatments and reduced density-dependent factors. One such recent model describes how competition explains intra-host diversification of parasites (Rascalou and Gourbiere, 2014). They show that parasite adaptive evolution is faster in highly fragments parasites populations and for weakly aggregated and virulent parasites, all factors which could be affected by drug selective pressures, hybridizations and associated trade-offs.

Genomics has the potential to shine a light on these discoveries to elucidate the timing and exact nature of these hybridisation events. An exemplary study on the plant pathogenic fungus *Zymoseptoria pseudotritici* (Stukenbrock et al., 2012) used the nuclear genomes of five individuals to investigate a recent hybridization event. This species shows an unusual pattern of genome-wide diversity, with the genome broken up into small (5.8kb) stretches of very high haplotype diversity interspersed between equally long sections showing no differentiation within the population. This mosaic pattern of haplotypes is indicative of a recent hybridization event followed by a population bottleneck. Estimates of recombination rate and mutation rate in this species allowed two different, approximately congruent estimates of the timing of this hybridization event, from data on the size of the haplotype blocks and the number of point mutations occurring in the genome. It is important to note that both understanding the pattern of hybridisation, and dating the founding event for these hybrids was only possible through a genome level analysis: if only individual genes had been isolated, then the global picture that this study has unearthed would have been lost. Through applying these kinds of methods to WGS data from hybrid populations it should be possible to gain a more complete picture of the nature and timing of introgression between *Schistosoma* species.
Understanding how often new populations of hybrids are founded, how long they persist and how their population genetics differs from parental populations will be important in understanding how these populations are likely to respond to MDA and the scope for PZQ resistance to spread within and between populations.

**Conclusions**

Accurately parameterised mathematical models incorporating genetic or genomic data can, for example, inform on rates of changes of schistosome phenotypes or genotypes associated with drug resistance, so that monitoring and evaluation studies can understand what they need to monitor, how and when is best to monitor, and can advise on optimal treatment strategies to maximise the gains from limited resources. Although the cost of NGS has fallen exponentially in recent years, the outlay of sequencing all 380Mb of the schistosome genome remains currently beyond the financial capacity of most institutions, although RAD-seq or exon-capture may represent a more cost-effective alternative (Gilabert and Wasmuth, 2013). When dealing with a disease that affects the very poorest in the world, the most useful public health interventions are those which cost the least and are broadly sustainable; therefore it is important that this technology is targeted to answer the most relevant public health questions. These NGS technologies, to complement population genetic data available from methodologies such as microsatellites, can thus be combined with mathematical models, and we anticipate that the next few years will represent a highly exciting and important era in such trans-disciplinary research of profound theoretical and applied importance.

**Acknowledgements**

We are very grateful to two anonymous referees for comments on the manuscript. PHLL is funded through an Imperial College Junior Research Fellowship; TC is funded through a Medical Research Council (MRC DPT PhD) studentship (supervised by JPW and JAC); JAC is funded by the Wellcome Trust through their support of the Wellcome Trust Sanger Institute (Grant 098051); JPW holds a Chair in Parasitic Diseases. RVC manuscript reference number PPB_00915.

**References**


Microsatellite markers reveal a spectrum of population structures in the malaria parasite Plasmodium falciparum. *Molecular Biology and Evolution*, 17, 1467-82.


WTSI. http://www.sanger.ac.uk/research/initiatives/globalhealth/research/helminthgenomes/


