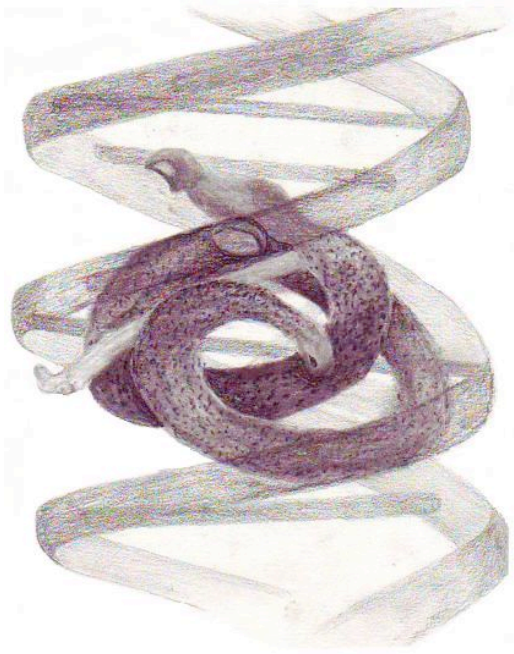
KU LEUVEN

ARENBERG DOCTORAL SCHOOL
FACULTY OF SCIENCE

Evolutionary potential of the human parasite *Schistosoma mansoni* in a changing world



Frederik Van den Broeck

Promoter:
Prof. Dr. Filip Volckaert

Co-Promoters:
Dr. Tine Huyse
Prof. Dr. Katja Polman

Dissertation presented in partial fulfilment of the requirements for the degree of Doctor in Science (Biology)

May 2014

To whom I owe a lot, my beloved grandparents
Mit Vertommen and Constant Van den Broeck
Nadine Saffers and William Denys

Laboratory of Biodiversity and Evolutionary Genomics
Section of Ecology, Evolution and Biodiversity Conservation
University of Leuven

Unit of Medical Helminthology
Department of Biomedical Sciences
Institute of Tropical Medicine

EVOLUTIONARY POTENTIAL OF THE HUMAN PARASITE *SCHISTOSOMA MANSONI* IN A CHANGING WORLD

Frederik Van den Broeck

Promoter:
Prof. F. Volckaert

Co-promoters:
Dr. T. Huyse
Prof. K. Polman

Members of the Examination Committee:
Dr. M. Larmuseau
Prof. T. Wenseleers
Prof. B. Goddeeris
Prof. D. Rollinson
Prof. JC. Dujardin

Dissertation presented in partial fulfilment of the requirements for the degree of Doctor in Sciences

May 2014

Cover illustration by Evelyne Denys
Two *Schistosoma* worms intertwined within a DNA helix: the long and slender female worm resides within the ventral groove of the male worm.

Dutch title
Evolutionair potentieel van de humane parasiet *Schistosoma mansoni* in een veranderende wereld

Uitgegeven in eigen beheer, FREDERIK VAN DEN BROECK, LEUVEN

ISBN 978-90-8649-714-0
D/2014/10.705/23

The research presented in this dissertation was funded by a VLADOC PhD scholarship of the Flemish Inter-University Council (VLIR-UOS).

Acknowledgements

This PhD would not have been possible were it not for the many people who supported me. The first small step was taken in the office of Prof. Filip Volckaert, where I sat as a naive master's student that keenly wanted to take part in an Erasmus program, preferably in a university where I could learn more about parasites. It was at this moment that I met parasitologist Dr. Tine Huyse, and soon after I had the privilege to start a master's and then a PhD under her supervision. Special thanks therefore to Tine for allowing me to participate in her parasitological research and for her sustained valuable guidance and advice. She was a strong mentor during the five years that I worked for her and if there is somewhere a good scientist in me now, then it is definitely thanks to her. I would also like to mention my gratitude to Prof. Filip Volckaert and Prof. Katja Polman for allowing me to join their research units. It was a pleasure to work in their groups and their flexibility allowed me to pursue my own research interests.

This thesis was also inspired by the many amazing colleagues and friends in the lab: Dirk Schaerlaekens with whom I shared many unforgettable moments in the lab (not to mention the unforgettable moments after the lab), Aurora Geerts for her invaluable friendship, Maarten Larmuseau for always believing in me, Bart Hellemans for his incredible support in the lab, Eveline Diopere and Sara Vandamme for their motherly warmth, Nele Boon for being a great colleague, Nellie Konijnendijk for her contagious positivism, Pascal Istvan Hablutsel, Maarten Vanhove and Maarten Vansteenberghe for the many lively conversations, Anurag Chaturvedi for his enthusiasm in bio-informatics, Joost Raeymaekers and Gregory Maes for raising the academic level in our lab to new heights and last but not least, Conny Coeckelberghs for her support throughout all these years.

I would like to thank my girlfriend Katia Stasinopoulou for her encouragement during the last years, her positivism and enthusiasm kept me going through the most difficult times. In addition, I cannot express enough my gratitude to my family – my siblings Nathalie and William and my parents Evelyne and Remi – for their love and support. In closing, I would also like to mention Pieter Albers, who was a great roommate for many years and with whom I shared many laughs, Niels Wynant and Pieter Vanwielendaele for the many lunches (and beers), Olivier Van der Brempt for being such a loyal friend, Enrico Borella for the delicious dinners, and Yuon-Man for serving me very good coffee with a welcoming smile!

Summary

The human helminth parasite *Schistosoma* spp. lives in the blood vessels of over 200 million people in Africa, Asia and South-America. The most serious infections are found in children and adolescents who eventually become physically and intellectually compromised, while people who have been infected chronically may develop liver damage, kidney failure or bladder cancer. Despite the availability of a cheap and effective drug, the disease has been largely neglected compared to other more 'deadly' diseases such as HIV/AIDS and malaria, both in terms of disease control and in terms of scientific research. Remedial action is therefore needed to gain insight into the distribution of these parasites at both a local (among humans and villages) and a global geographic scale (across regions or continents). It is essential to understand the factors and dynamics that shape transmission and the impact of selective pressures such as drug treatment. This will provide a theoretical framework that could represent a starting point for better disease control.

In this thesis we performed a population genetic study to reveal the distribution of *Schistosoma mansoni* parasites in the basin of the Senegal River (West Africa). Molecular markers such as microsatellites or single nucleotide polymorphisms serve as ideal tools to track the transmission of parasites and infer their ancestry. However, parasite worms cannot be used as a source for DNA as they are inaccessible within the human blood vessels. A protocol was therefore optimized that allowed the sampling, DNA-extraction and molecular analysis based on low quantities of DNA obtained from individual larval parasites (chapter 2). This optimization was an important first step that enabled us to perform multiple analyses on the same individual parasite with low genotyping error rates.

The case study in Northwest Senegal represents one of the most intense epidemic foci of *S. mansoni*. The parasite invaded this area in 1986 after the construction of two dams on the Senegal River. Parasites were obtained at four time points over a period of 14 years (1993-2007) and from several regions (Northwest Senegal, Southeast Senegal and Mali) to study the nature of the *S. mansoni* epidemic (chapter 3). Typing of nuclear and mitochondrial markers revealed that parasites from Northwest Senegal have a West-African origin, that they harbor moderate to high levels of genetic diversity, that they increase in population size and that some parasites are genetically different from others (chapter 3). These results suggest that the epidemic of schistosomiasis in Northwest Senegal was probably not elicited

by a few *S. mansoni* parasites, but that the colonization history is much more complex. It is most likely that a multitude of parasites successfully colonized the local human population. Furthermore, most of the parasite genetic variation observed in the region was found within individual human hosts (chapter 4), suggesting that they accumulate a wide range of parasites during their lifetime. However, children appeared to be much more infected by related parasites than adult hosts, who tended to be infected by more genetically diverse parasites. The age-dependent recruitment of genetically diverse parasite infections may be explained by (1) genotype-dependent 'concomitant immunity' that leads to selective recruitment of genetically unrelated worms with host age, and/or (2) the 'genetic mixing bowl' hypothesis, where older hosts have been exposed to a wider variety of parasites than children.

Currently the cheap and effective drug praziquantel is used to treat schistosomiasis. However, this could lead to serious bottlenecks in *S. mansoni* populations, possibly leading to increased inbreeding, low levels of genetic diversity and the random fixation of (possibly deleterious) alleles. The effect of treatment on the genetic composition of *S. mansoni* populations was therefore studied using data obtained from simulations of an island model at equilibrium (chapter 5) and from naturally collected parasites before and after treatment (chapter 6). Both studies showed that treatment has only a limited effect on the genetic diversity of schistosome populations. Complementary simulations revealed that only a sustained treatment policy could decrease schistosome population sizes and therefore drive the success of a control program.

This thesis has shed new light on the factors that shape the distribution of *S. mansoni* parasites. The main conclusion is that *S. mansoni* populations harbor substantial levels of genetic diversity, and hence are able to cope with strong selection pressures such as chemotherapeutic treatment. This large evolutionary potential will hamper attempts either to control or to eliminate these parasites, as well as complicate the development of new drugs or vaccines.

Samenvatting

De menselijke parasiet *Schistosoma* spp. leeft in de bloedvaten van meer dan 200 miljoen mensen in Afrika, Azië en Zuid-Amerika. De hevigste infecties komen voor bij kinderen en jonge volwassenen die hierdoor een fysische en intellectuele achterstand oplopen, terwijl chronische infecties leiden tot nierfalen, schade aan de lever of blaaskanker. Ondanks de beschikbaarheid van een goedkoop en doeltreffend geneesmiddel is de ziekte grotendeels verwaarloosd in vergelijking met andere meer dodelijke infectieziekten zoals HIV/AIDS en malaria, en dit zowel inzake ziektebestrijding als wetenschappelijk onderzoek. Het is daarom des te belangrijker om nieuwe inzichten te verwerven in de verspreiding van deze parasieten op een lokale (tussen mensen en dorpen) en globale schaal (tussen regio's en continenten). Het begrijpen van de factoren die de verspreiding van deze parasieten bepalen alsook het begrijpen van de effecten van selectieve druk zoals een behandeling met geneesmiddelen is essentieel, en laten toe een theoretisch kader te ontwerpen dat een startpunt kan vormen voor een betere bestrijding van de ziekte.

In deze thesis voerden we een populatiegenetische studie uit om de verspreiding van *Schistosoma mansoni* parasieten in het Senegal Rivier Bekken (West-Afrika) in kaart te brengen. Moleculaire merkers zoals microsatellieten of Unieke-Nucleotide Polymorfismen zijn ideale middelen om de transmissie van parasieten te traceren. Het is echter heel moeilijk om de wormen als bron van DNA te gebruiken aangezien zij onbereikbaar zijn in het menselijk bloedvatenstelsel. We hebben daarom eerst enkele protocollen geoptimaliseerd om individuele parasietenlarven te verzamelen in het veld, er DNA uit te extraheren en deze te gebruiken voor moleculaire analyses. De optimalisatie van deze protocollen was een eerste belangrijke stap dat toeliet meerdere moleculaire analyses uit te voeren op één en dezelfde individuele parasiet met slechts lage genotyperingsfouten.

Het studiesysteem in Noord-Senegal betrof één van de meest intense epidemieën van *S. mansoni* die ooit werden waargenomen (hoofdstuk 3). De parasiet daagde op in dit gebied vanaf 1986 na de bouw van twee stuwdammen op de Senegal rivier. Om de kolonisatiegeschiedenis van *S. mansoni* in dit gebied te achterhalen, werden parasieten bekomen op vier verschillende tijdstippen gedurende een periode van 14 jaar (1993-2007) en uit verschillende regio's (Noord-Senegal, Zuid-Senegal en Mali). De genetische karakterisatie van deze parasieten op nucleaire en mitochondriale merkers gaf aan dat de

Noord-Senegalese *S. mansoni* parasieten een West-Afrikaanse oorsprong kenden, dat ze relatief veel genetische variatie vertoonden, dat ze een groei in populatiegrootte kenden en dat sommige *S. mansoni* parasieten genetisch sterk verschilden van andere parasieten (hoofdstuk 3). Deze resultaten suggereren dat de epidemie van schistosomiase in dit gebied complex is, waarbij wellicht veel verschillende *S. mansoni* parasieten de gastheerpopulatie koloniseerden in dit gebied. Bovendien werd de meeste genetische variatie van parasieten geobserveerd binnen de menselijke gastheren (hoofdstuk 4). Kinderen bleken hierbij meer geïnfecteerd te zijn door verwante parasieten dan volwassenen, die op hun beurt geïnfecteerd waren met genetisch meer diverse parasieten. Deze waarnemingen kunnen verklaard worden door (1) een genotype-afhankelijke 'concomitante immuniteit' dat leidt tot selectieve verwerving van genetisch niet verwante parasieten in oudere gastheren, en/of (2) de 'genetische mengkom' hypothese, waarbij oudere gastheren blootgesteld zijn aan een grotere variëteit aan parasieten dan kinderen.

Momenteel wordt het goedkope en doeltreffend geneesmiddel praziquantel gebruikt om schistosomiase te bestrijden. Dergelijk grootschalig gebruik zou kunnen leiden tot flessenhalzen in *S. mansoni* populaties, dat kan leiden tot verhoogde inteelt, lage genetische diversiteit en de willekeurige fixatie van allelen. Het effect van de behandeling op de genetische samenstelling van *S. mansoni* populaties werd daarom bestudeerd door simulaties van een eilandmodel in evenwicht (hoofdstuk 5) en door een veldstudie waarbij parasieten werden verzameld voor en na behandeling (hoofdstuk 6). Beide studies toonden aan dat behandeling slechts een klein effect heeft op de genetische diversiteit van deze parasietenpopulaties. Onze simulaties toonden verder aan dat enkel een intense behandelingscampagne een reductie in genetische diversiteit op lange termijn kan veroorzaken.

Deze thesis heeft nieuwe inzichten gebracht in de factoren die de verspreiding van *S. mansoni* parasieten beïnvloeden. De belangrijkste conclusie is dat *S. mansoni* populaties voldoende divers zijn om selectieve drukken zoals een chemotherapeutische behandeling te weerstaan. Dergelijk sterk evolutionair potentieel zal elke poging tot controle of eliminatie van deze parasieten bemoeilijken, alsook de ontwikkeling van nieuwe geneesmiddelen en vaccins compromitteren.

List of abbreviations and list of symbols

AR	allelic richness
ANOVA	analysis of variance
BIC	bayesian information criterion
BLAST	basic local alignment tool
blastn	nucleotide-nucleotide BLAST
bp	base pair
CI	confidence interval
***cox*1**	cytochrome oxidase subunit 1
D	number of dying parasites
ΔK	second order rate change of LnP(D)
DMS1	data microsatellites 1
DMS2	data microsatellites 2
DNA	deoxyribonucleic acid
DSEQ	data sequences
e_{obs}	multilocus genotype error rate
e_l	mean error rate per locus
EDTA	ethylene diamine tetra-acetic acid
EPG	eggs per gram
EtOH	ethanol
f	F_{IS} as estimated following Weir and Cockerham (1984)
F_{IS}	inbreeding coefficient
F_{ST}	fixation index
FCA	factorial correspondence analysis
GLM	general linear model
h	haplotype diversity
Hs	unbiased expected heterozygosity
Ho	observed heterozygosity
HIV/AIDS	human immunodeficiency virus / acquired immune deficiency syndrome
HWE	Hardy Weinberg Equilibrium
IAM	infinite allele model
K	number of clusters
K_c	carrying capacity
***k*-means**	distance based clustering algorithm
LnP(D)	log-likelihood
m	proportion of newly produced parasites within a given infrapopulation that infect other infrapopulations
m_g	number of multilocus genotypes including at least one allelic mismatch
m_l	number single-locus genotypes including at least one allelic mismatch

MCMC	Markov Chain Monte Carlo
MDS	multidimensional scaling
MERL	mean error rate per locus
MGER	multilocus genotype error rate
MLG	multilocus genotype
MLH	multilocus heterozygosity
mtDNA	mitochondrial DNA
na	not applicable
Ne	effective population size
$\sum N_i$	component population size
N_i	infrapopulation size
n	number of infrapopulations
nt	number of replicated single- or multilocus genotypes
NGS	next-generation sequencing
NTD	neglected tropical disease
П	nucleotide diversity
p	*p*-value
PCR	polymerase chain reaction
PZQ	praziquantel
r	correlation coefficient
R	software package R
R	per capita infection rate
R_{exp}	expected relatedness
R_{obs}	observed relatedness
R_{ST}	measure for genetic differentiation following Slatkin (1995)
RD-PCR	rapid diagnostic PCR
SCI	Schistosomiasis Control Initiative
SD	standard deviation
sib	sibling
SRB	Senegal River Basin
SMM	stepwise mutation model
SNP	single nucleotide polymorphism
STH	soil-transmitted helminthiasis
θ	F_{ST} following Weir and Cockerham (1984)
Taq	*Thermus aquaticus*
Tm	melting temperature
TPM	two-phase model
WHO	World Health Organization

Table of contents

ACKNOWLEDGEMENTS I

SUMMARY III

SAMENVATTING V

LIST OF ABBREVIATIONS AND LIST OF SYMBOLS VII

TABLE OF CONTENTS IX

CHAPTER 1 1

GENERAL INTRODUCTION

CHAPTER 2 17

OPTIMAL SAMPLE STORAGE AND EXTRACTION PROTOCOLS FOR RELIABLE MULTILOCUS GENOTYPING OF THE HUMAN PARASITE *SCHISTOSOMA MANSONI*

CHAPTER 3 31

INVASION GENETICS OF *SCHISTOSOMA MANSONI* IN NORTHWEST SENEGAL REVEALS SIGANTURES OF POPULATION EXPANSION

CHAPTER 4 51

INBREEDING WITHIN HUMAN *SCHISTOSOMA MANSONI*: DO HOST-SPECIFIC FACTORS SHAPE THE GENETIC COMPOSITION OF PARASITE INFRAPOPULATIONS?

CHAPTER 5 75

FIGHTING A LOSING BATTLE? A SIMULATION-BASED APPROACH TO THE IMPACT OF COMMUNITY-BASED DRUG TREATMENT ON THE GENETIC DIVERSITY OF SCHISTOSOME POPULATIONS

CHAPTER 6 89

REGULAR TREATMENTS OF PRAZIQUANTEL DO NOT IMPACT ON THE GENETIC MAKE-UP OF *SCHISTOSOMA MANSONI* IN NORTHWEST SENEGAL

CHAPTER 7 101

GENERAL DISCUSSION

GLOSSARY 119

REFERENCES 123

LIST OF PUBLICATIONS 137

CHAPTER 1

General introduction and aims

1.1 This wormy world: parasites and parasitism in perspective

Parasites are organisms that live in or on other living organisms (i.e. the host) and obtain part or all of their organic nutrients from this host (Goater et al., 2013). This definition describes parasites in the broad sense and comprises both the macroparasitic (e.g. helminths, arthropods) and the microparasitic organisms (e.g. bacteria, viruses, Protozoa). Parasitism is a form of a symbiosis, which is an intimate interaction between two organisms of different species that live together. Symbiotic interactions are often classified into four groups based on the type of interaction (exploitation, mutualism, commensalism and phoresy), although the exact definitions and boundaries remain vague (Goater et al., 2013).

To most people, the word parasite only rings a bell when they take their pet to the veterinarian for the annual deworming, or when they take pills and receive vaccines to protect themselves against diseases that they may acquire during their 'exotic' holidays. Many people know some of the most deadly diseases such as AIDS, malaria, tuberculosis and the bubonic plague (also known as Black Death, which killed around 1/3 of the European population in the fourteenth century). Few people however, realize that parasites still plague the vast majority of the world's population, especially in (sub-) tropical regions. Furthermore, they do not realize that during the relatively short history of man, humans have acquired over 70 species of Protozoa and about 340 species of helminth worms, not to mention all the fungi and bacteria (Ashford and Crewe, 1998). Although many of these may be rare, humans still harbour about 90 relatively common parasitic species. For instance, it is estimated that about 20% of the world's population (i.e. >1.4 billion people) is currently infected with the roundworm *Ascaris lumbricoides* (Crompton, 1999). Humans and parasites have exhibited a considerable period of engagement in mutual co-adaptation and selection such that parasitized hosts may have an advantage over uninfected individuals in some contexts (Thomas et al., 2000; Dunne and Cooke, 2005; Dunn, 2011). For instance, experimental studies in rodents have shown that infection with the helminth parasite *Schistosoma mansoni* prevents the onset of some autoimmune disorders such as type 1 diabetes (Cooke et al., 1999; La Flamme et al., 2003; Nagayama et al., 2004).

The huge number of helminths infecting humans was first coined by Norman Stoll in 1947 (Stoll, 1947). Stoll titled his text 'This Wormy World' to draw attention to the global public health impact of helminth infections, a title that has later been borrowed by many others (Bundy and de Silva, 1998; Chuan et al., 2010). Stoll noted that *"Helminthiases do not have the journalistic value of great pandemics like flu or plague. They do not, for most part, present dramatic clinical cases, but to make up for their lack of drama, they are unremittingly corrosive. If you were aroused by the sufferings of, say, ten thousand service men with filariasis and schistosomiasis, what can your imagination do with ten thousand upon ten thousand natives in endemic areas – who have no homeside relatives to write letters to their congressmen?"* and finishes with *"What we need are worm treatments that are effective and well-tolerated as phenothiazine in sheep; and from the greatness of the need, we ought to have a hundred workers seeking them, instead of a handful in desultory effort".* Of the 342 helminth species infecting humans nowadays, Stoll was mainly concerned with 25 that according to him merit global attention and require major control programs (Stoll, 1947). Although a large variety of chemotherapeutic drugs have now been developed and commercialised, large-scale prevention and treatment remain a global crisis (Hotez and Kamath, 2009). All major helminthiases are therefore still classified as neglected tropical diseases (NTDs), which are a group of chronic, disabling conditions that are widespread among the poor in sub-Saharan Africa (Hotez et al., 2007). Among the most common NTDs are soil-transmitted helminth (STH) infections, lymphatic filariasis (LF), trachoma, onchocercariasis and schistosomiasis that together affect more than 500 million people worldwide (Molyneux et al., 2005). In this study we will focus our research on the helminth *Schistosoma mansoni* that is the causative agent of the disease schistosomiasis (also called 'bilharzia' after Theodor Bilharz who discovered the species *Schistosoma haematobium* in 1852).

1.2 The human parasite *Schistosoma* spp.

1.2.1 Taxonomy, phylogeny and phylogeography

The genus *Schistosoma* belongs to the Phylum Platyhelminthes, Classis Trematoda, Subclassis Digenea, Ordo Strigeatida and Familia Schistosomatidae. Platyhelminthes (from Greek *platy-* 'flat' + *helminth-* 'worm') is a diverse group of soft-bodied invertebrates

comprising both free-living and parasitic organisms. The phylum is divided into the mostly nonparasitic Turbellaria and the three entirely parasitic Cestoda, Trematoda and Monogenea. The Digenea is a Subclassis comprising about 11,000 parasitic flatworms and together with the Monogenea they are often called flukes. Well-known examples of non-schistosome Digenea are *Fasciola hepatica* infecting sheep and cattle, and *Fasciola magna* infecting deer. With almost no exception, digeneans have a vertebrate final host and a snail as first intermediate host. The name Digenea (from Greek *di-* 'twice' + *genea-* 'generation, race') refers to the alternation of sexually reproducing adults and asexually reproducing larval stages. Unlike other trematodes, schistosomes are not hermaphroditic but dioecious, forming two separate sexes (Goater et al., 2013).

The genus *Schistosoma* comprises 23 species, all infecting mammals (Lawton et al., 2011). The species of *Schistosoma* were initially classified into four groups based on the egg morphology and their intermediate host specificity (Rollinson and Simpson, 1987). The groups were named after the most important species within that group: *S. haematobium* (Bilharz, 1852), *S. mansoni* (Sambon, 1907), *S. japonicum* (Katsurada, 1904) and *S. indicum* (Montgomery, 1906). Recent phylogenetic analyses investigating the interrelationships of the Schistosomatidae using molecular markers recognised these four groups as monophyletic clades, but revealed two additional groups named the proto-*S. mansoni* clade and the *S. hippopotami* clade (Barker and Blair, 1996; Lockyer et al., 2003; Morgan et al., 2003; Webster et al., 2006; Lawton et al., 2011). The distribution of the six parasite groups is closely linked to the geography of its obligate intermediate snail host species (Morgan et al., 2001; Agatsuma, 2003). The *S. japonicum* group and the *S. indicum* group are primarily found in Asia, while both the *S. mansoni* and *S. haematobium* groups are found throughout Africa (Figure 1.1). *Schistosoma mansoni* is the only species that can be found in south-America, more specifically in Venezuela, Surinam, the Caribbean and some parts of Brazil.

It has been suggested that the genus *Schistosoma* would have an African origin before the separation of the super continent Gondwanaland (Davis, 1993). However, the basal position of the *S. japonicum* group on the tree suggests that *Schistosoma* most likely originated in Asia (Figure 1.1). By combining mitochondrial data with cytogenetic data, Lawton and colleagues (2011) were able to reconstruct the history of *Schistosoma* and to confirm an Asiatic origin. The genus *Schistosoma* probably arose from avian schistosomatids and

radiated in rodents approximately 60-70 million years ago in China and Southeast Asia. Schistosomes then colonized Africa with the widespread mammal migration 15-20 million years ago (late Miocene). The African schistosomes diverged approximately 1-4 million years ago and gave rise to the *S. mansoni* and *S. haematobium* groups, while the ancestors of the *S. indicum* group emerged and dispersed to India, again mediated by an extensive mammal migration into Asia during Plio-Pleistocene (Morgan et al., 2001; Lawton et al., 2011).

A total of eight schistosome species are known to infect humans (*S. japonicum*, *S. malayensis*, *S. mekongi*, *S. mansoni*, *S. mattheei*, *S. intercalatum*, *S. haematobium* and *S. guineensis*). Human schistosomiasis probably originated three times independently (Webster et al., 2006), once in the *S. japonicum* clade, once in the *S. mansoni* clade and once in the *S. haematobium* clade (Figure 1.1). The most recent event may have been promoted by the domestication of cattle as *S. mattheei* is usually found in domestic stock and wild ungulates. In the *S. haematobium* lineage, the human host preference was lost again with *S. curassoni* and *S. bovis* primarily using domestic stock as final hosts (Figures 1.1 and 1.3). It is also interesting to note that the recently described *S. kisumuensis* that was isolated from three murid rodent species in the Lake Victoria Basin (Hanelt et al., 2009) falls in the middle of a clade consisting of human schistosomes (Figure 1.1). Although *S. kisumuensis* has not been found in humans yet, further research testing human populations specifically for infection with this new species is warranted.

An extensive phylogeographic study on the human parasite *S. mansoni* revealed considerable within-species mtDNA diversity with 85 haplotypes grouped into five divergent lineages across Africa (Morgan et al., 2005). The highest diversity of *S. mansoni* was reported in East-African lineages suggesting an East-African origin for the parasite 0.30 – 0.43 M years ago, a time frame similar to the estimated arrival of the *Biomphalaria* snail host (DeJong et al., 2003). The data furthermore indicated a recent New World colonization (likely during the slave trade) that originated from multiple West-African countries (Morgan et al., 2005).

1.2.2 Schistosome lifecycle

Humans become infected with *Schistosoma* parasites when they are in contact with snail-infested fresh water in the African, American and Asian continents. Schistosomes form five different development stages: eggs, miracidia, sporocysts, cercariae and adult worms.

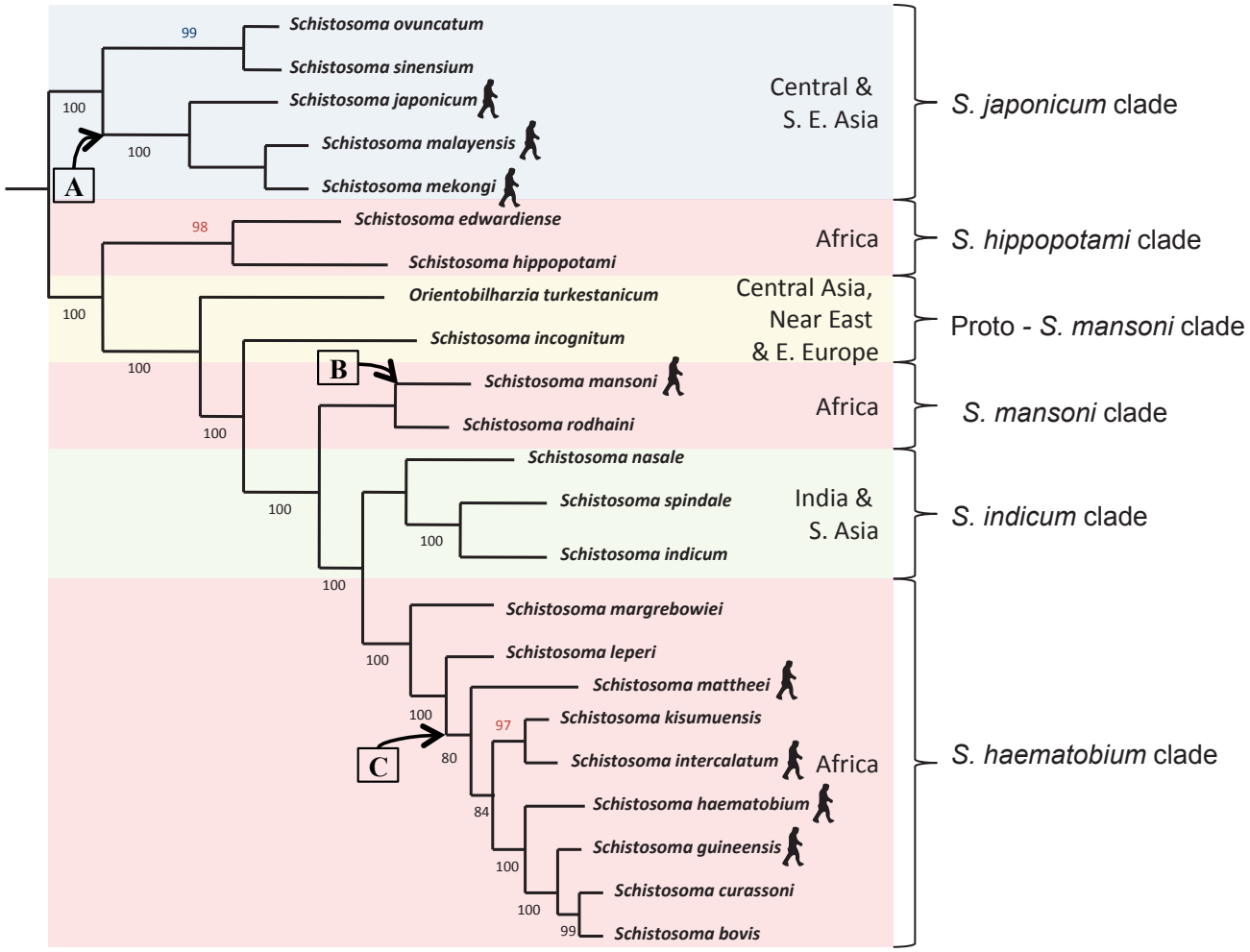


Figure 1.1 Phylogeny of the members of the *Schistosoma* genus estimated with a Bayesian analysis of combined partial lsrDNA, complete ssrDNA and partial *cox*1 (mtDNA). The robustness of each node was inferred using 2000 bootstraps from maximum parsimony analysis. The tree depicts the four historically recognised species groups *S. japonicum, S. mansoni*, *S. indicum* and *S. haematobium*, with addition of two new clades (*S. hippopotami* and proto-*S. mansoni*) as suggested by Lawton et al. (2011). The tree furthermore illustrates the basal nature of Asian schistosomes that are ancestral to the African schistostomes and the three independent events (A, B, C) that gave rise to human schistosomes. This figure was adapted from Webster et al. (2006) and Lawton et al. (2011).

The adult schistosomes reside within the veins surrounding the urinary bladder (*S. haematobium*) or intestines (*S. mansoni* and *S. japonicum*). The long and slender female worm (7 – 20 mm) can only mature upon copulation in the ventral groove of the shorter but fatter male worm (sexual reproduction phase). On a daily basis, the females deposit eggs (200-3,000 eggs per day) in the small venules of the portal and perivesical systems. Due to excretion of lytic products, the eggs penetrate through the gut and bladder wall and are evacuated in urine or faeces. If deposited in freshwater (and if enough sunlight), the egg will hatch into a free-swimming miracidium that will actively seek a suitable intermediate host for up to 24 hours (Figure 1.2). *Schistosoma* species show narrow host specificity for snails:

S. haematobium is transmitted by the *Bulinus* spp., *S. mansoni* infects *Biomphalaria* spp. and *S. japonicum* is transmitted by *Oncomelania* spp. (Figure 1.2). Once the parasite has invaded the soft tissues of the snail it will transform into a primary (mother) sporocyst, usually near the site of penetration. Every primary sporocyst then produces several secondary (daughter) sporocysts 2-6 weeks after infection. The secondary sporocysts migrate to other organs in the snail and produce thousands of cercariae that are released into the water. Note that all cercariae from one miracidium are (nearly-) identical clones from each other as a result of asexual reproduction and thus have the same sex. The cycle continues when a cercaria attaches and actively penetrates the skin of the final host, after which it transforms into a schistosomulum (i.e. 'little schistosomes'). After migration through skin and lungs, schistosomula reach the portal vessels in the liver where they develop in about three weeks into an adult worm. The lifecycle is complete when young female and male worms pair and migrate to the veins surrounding the gut or bladder where egg production starts from 4-8 weeks after initial infection. Adult worms are long-lived and can survive for 2 to 5 years with an average of 4.5 years, although some may last as long as 30 years (Rollinson and Simpson, 1987).

1.2.3 Interspecific hybridization

The special life cycle of the dioecious schistosome parasites necessitates an obligatory reproduction between male and female worms within the final mammal host. Once schistosomula have developed into adult worms within the liver, young male and female worms have to localise each other in the circulatory system in order to pair. Pairing is for most schistosome species highly important for the body growth and sexual differentiation of the female worm and for the migration from the hepatic portal veins to the egg-laying sites (Armstrong, 1965). *In vivo* experiments on schistosome mating behaviour have shown that in mixed infections these pairings occur readily between worms of different species (Tchuem Tchuenté et al., 1993). However, intraspecific attraction appears to be stronger than interspecific attraction, meaning that whenever a choice is possible mating will preferentially – but not exclusively – occur between partners of the same species (Tchuem Tchuenté et al., 1993; Norton et al., 2008).

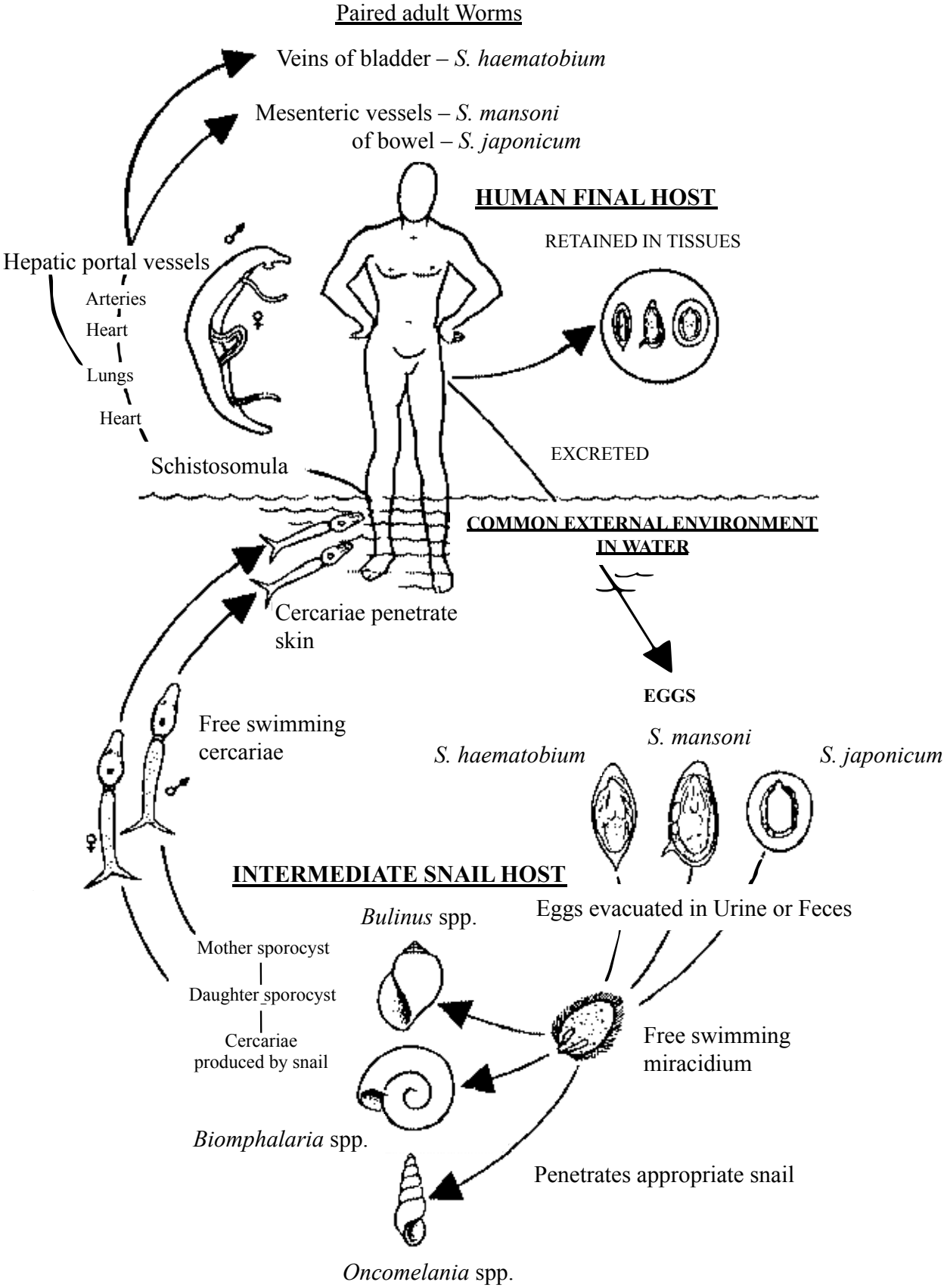


Figure 1.2. Life cycle of *Schistosoma* spp. See text for details.

Although it has always been assumed that male and female worms remain together in an intimate and permanent association throughout their life, further experiments showed that a change of mate could occur in mixed or sequential infections (Tchuem Tchuenté et al., 1995), also between pairs of the same species (Pica-Mattoccia et al., 2000; Beltran et al., 2008).More specifically, heterospecific worm pairs will change partners to become conspecific pairs (Tchuem Tchuenté et al., 1995). Interestingly it was shown that, next to mate choice and mate change, there is also a strong mating competition. When unpaired *S. mansoni* male worms arrived in an established *S. intercalatum* infection, the stronger *S. mansoni* males will pull away female *S. intercalatum* from male *S. intercalatum* (Tchuem Tchuenté et al., 1993, 1995). Such interspecific competition could have consequences for parasite epidemiology. Studies suggest for instance that the exclusion of *S. intercalatum* by *S. mansoni* and by *S. haematobium* may be an important factor explaining the limited distribution of *S. intercalatum* in Africa (Southgate, 1978; De Clercq et al., 1994; Tchuem Tchuenté et al., 1996a).

Altogether, these experimental studies showed that the mating behaviour between schistosomes may be much more dynamic than initially thought (Taylor, 1970; Southgate et al., 1982, 1995, 1998; Rollinson et al., 1990). Pairing between schistosome worms of different species is thus possible, but depending on the phylogenetic distance between them the pairing will lead to either hybridisation or parthenogenesis (Taylor, 1970). Hybridisation, with the production of viable fertile offspring, will usually occur when both species belong to the same species group. It has been documented within laboratory or natural conditions for species belonging to the *S. mansoni* clade (e.g. human *S. mansoni* x rodent *S. rodhaini*), the *S. haematobium* clade (e.g. bovine *S. bovis* and human *S. haematobium*; Figure 1.3) and the *S. japonicum* clade (e.g. human *S. mekongi* x rodent *S. malayensis*) (Tchuem Tchuenté et al., 1996b, 1997; Southgate et al., 1998; Webster and Southgate, 2003; Steinauer et al., 2008b; Huyse et al., 2009; Webster et al., 2013a). Hybridization could be of major epidemiological importance because it potentially leads to the formation of new hybrid pathogens that show a higher fitness than the parental species. When species belong to a different species group, pairing will usually result in parthenogenetic haploid offspring that are often viable but not fertile (Southgate et al., 1998). Such crossings have been shown for *S. mansoni* females paired with *S. douthitti* males

(Basch and Basch, 1984) and with *S. intercalatum* males (Tchuem Tchuenté et al., 1994), and for *S. mansoni* worms paired with *S. japonicum* worms (Imbert-Establet et al., 1994).

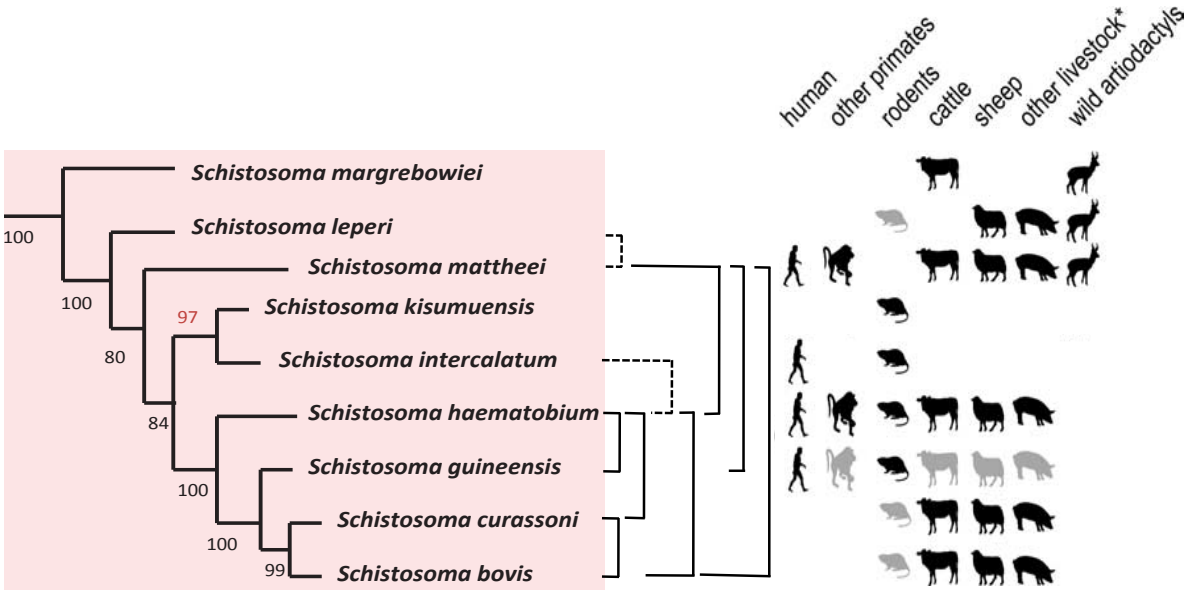


Figure 1.3. Phylogenetic tree of members of the *Schistosoma haematobium* species group. Vertebrate hosts for each species are given on the right side (black icons = data from the wild; grey icons = experimental evidence). All pairings of species known to hybridise are given (solid lines = in the wild or experimentally; broken lines = suspected or demonstrated for a limited number of generations). This figure was adapted from Webster et al. (2006) and Lawton et al. (2011).

1.2.4 Schistosomiasis

One of the implications of the complex schistosome life cycle is the fact that both final and intermediate host species need to be present in order for transmission to occur. Within developing countries, humans are very dependent on the availability of freshwater for their daily activities (agricultural workers, fishermen, women during their domestic activities…). When the respective snail species are present at the water contact sites, the schistosome life cycle can be completed. Over 200 million people within 74 development countries in Africa, Asia and South-America are affected by the disease schistosomiasis (Figure 1.4), from which 83 million are infected by *S. mansoni*. Twenty million people suffer from a severe form of the disease, 120 million people are symptomatic, and about 600 million people are at risk (Chitsulo et al., 2000; van der Werf et al., 2003). The disease continues to (re-) emerge in new areas due to increasing population sizes and movement (Chitsulo et al., 2000). For instance, refugee movements and population displacements in the Horn of Africa have introduced intestinal schistosomiasis to Somalia and Djibouti (Figure 1.4). Furthermore,

with increasing human demands for the world's supplies of freshwater, serious changes have been made in tropical freshwater habitats that influence the distribution of snails and therefore the spread of schistosome species. The construction of the Low and High dams at Aswan in Egypt for instance, resulted in a shift of inundation to perennial irrigation, probably creating sufficiently stable water bodies to allow proliferation of *Biomphalaria alexandrina* (Malone et al., 1997; El-Khoby et al., 2000). This change in habitat is thought to be one of the underlying reasons why *S. mansoni* has completely replaced *S. haematobium* in the Nile Delta. Another striking example of the rapidity to which water development projects can have serious consequences towards the emergence of schistosomiasis is Senegal (see section 2.1). In 1985, the Diama dam was built at the mouth of the Senegal River to promote agriculture of rice and reduce salinity, which led to the unforeseen effect of increasing pH and creating permanent water bodies, which in turn favored the colonization by both *Biomphalaria* and *Bulinus* snails (Southgate, 1997). By 1988 *S. mansoni* infections were reported for the first time in the town of Richard Toll (Talla et al., 1990). By 1989, almost 50% of the patients were infected with *S. mansoni* and by 1994 the mean prevalence of village around Richard Toll were 72%, thereby presenting one of the world's most intense foci of *S. mansoni* (Picquet et al., 1996).

Intense treatment (see section 1.2.5) in South-America, Asia and Egypt (Figure 1.4) has significantly lowered the burden of schistosomiasis in most of the regions. However, in sub-Saharan Africa the burden of this disease remains enormous: *S. haematobium* is estimated to cause haematuria in 70 million people, major bladder wall pathology in 18 million people and hydronephrosis in 10 million people. Annual mortality is estimated to be about 280,000 due to non-functioning kidneys (*S. haematobium*) and portal hypertension (*S. mansoni*) (van der Werf et al., 2003). The heaviest infections are found in children and young adults resulting in physically and intellectually compromised school-aged children (Chitsulo et al., 2000; van der Werf et al., 2003). Besides the medical importance of this disease, it is also of major veterinary importance with an estimated 165 million infected cattle worldwide (De Bont and Vercruysse, 1997). Within endemic regions, infection is usually first noticed at the age of 3-4, with a strong increase in both prevalence and infection intensity to a maximum when children reach the age of 15-20 years and a decrease thereafter (Stelma et al., 1993; Meurs et al., 2012). This pattern could be explained by immunity that is acquired over age

and/or differences in exposure (water contact), skin composition or hormones between children and adults (Fulford et al., 1996, 1998; Kabatereine et al., 1999).

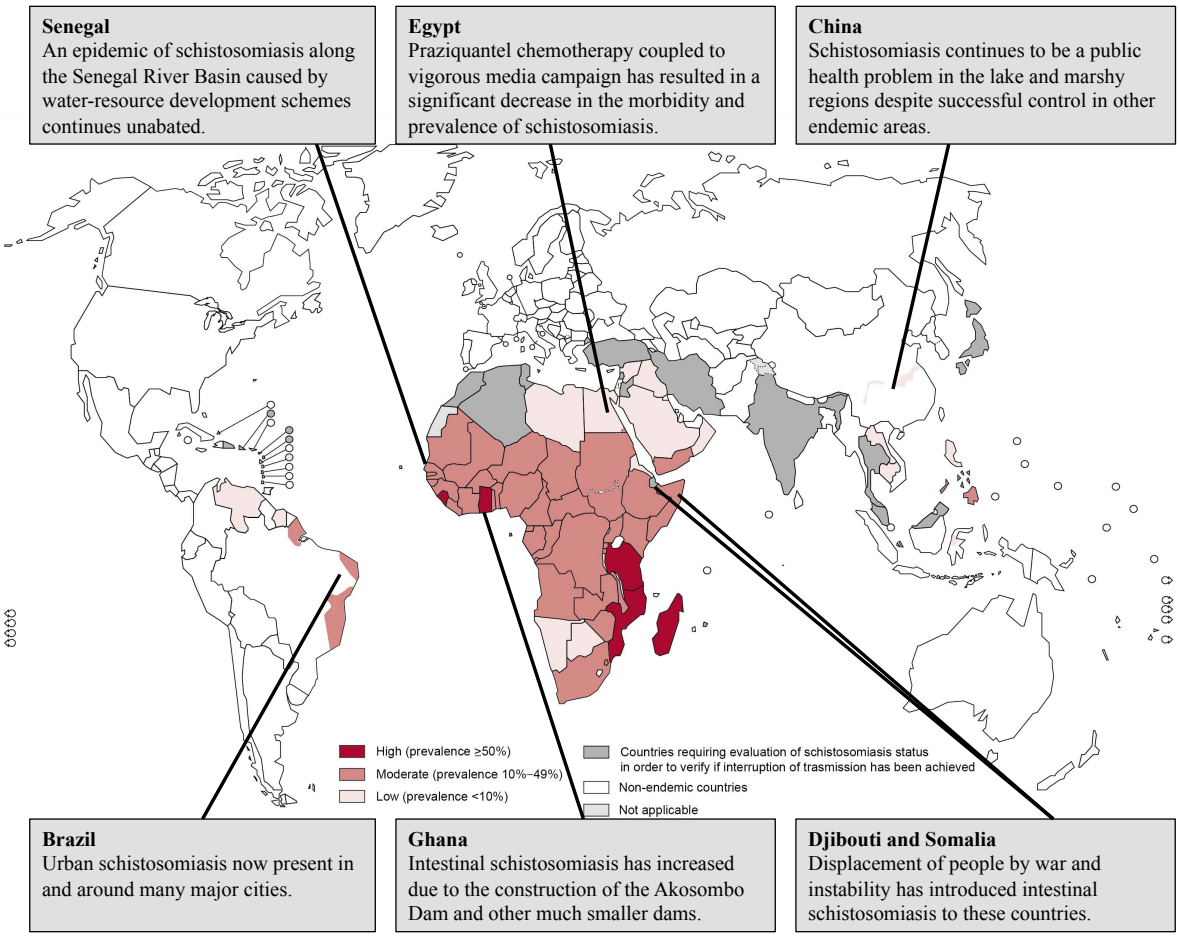


Figure 1.4. Global distribution of schistosomiasis (adapted from WHO, 2011).

1.2.5 Chemotherapy of schistosomiasis and the possibility of drug resistance

The treatment and control of schistosomiasis relies almost exclusively on a single drug, praziquantel (PZQ) (Fenwick et al., 2003). Due to its activity against all schistosome species, the excellent pharmacological properties and the substantial reduction in price, PZQ has become the recommended drug for almost 30 years now to treat schistosomiasis at both community level and in individual practice (WHO, 2006). The other antischistosomal drug available on the market, oxamniquine, has an excellent record of efficacy and safety for the treatment of infections caused by *S. mansoni* (Cioli et al., 1995). Oxamniquine is however not effective against other human schistosomes and its use is therefore almost entirely restricted to the New World where only *S. mansoni* occurs (Coura and Amaral, 2004; Katz and Coelho, 2008). The antimalarial drug artemisinin is known to have activity against *S.*

japonicum and *S. mansoni* and is of particular interest because it is more active against immature worms than PZQ and oxamniquine. Proposals for the use of artemisinins in areas where *Plasmodium* spp. and schistosomes coexist however, raised concerns about inducing drug-resistance in the former (Keiser and Utzinger, 2007; Utzinger et al., 2007). Many other drugs are active against schistosomes, but are not (anymore) in use (Cioli et al., 1995).

Although some effects of PZQ on worm physiology (massive influx of calcium, contraction of musculature) and morphology (disruption of tegument) are described, the precise molecular target(s) of PZQ remains unclear (reviewed in Cioli and Pica-Mattoccia, 2003). It is shown that an unusual variant of the voltage-gated calcium-channel β subunits would render schistosome cells sensitive to PZQ, although binding of PZQ to the calcium-channel β subunits has not been demonstrated (Kohn et al., 2001, 2003; Greenberg, 2005). A recent study, however, showed that the accumulation of calcium by itself is not crucial in the antischistosomal activity of PZQ *in vitro* (Pica-Mattoccia et al., 2008), challenging the hypothesis of Ca^{2+} involvement in the activity of PZQ. Alternatively, the target of PZQ might be other cellular factors that can regulate intracellular levels of calcium, such as receptors that regulate the uptake of adenosine, an essential metabolite schistosomes cannot produce themselves (Angelucci et al., 2007).

It is generally accepted that PZQ-resistant schistosomes do exist, but thus far their clinical relevance in the field is probably limited as in most countries normal cure rates (60%-90%) are obtained (Doenhoff and Pica-Mattoccia, 2006). There have been several alarming reports within *S. mansoni* endemic communities in the early ‘90s, notably in the Nile-delta region of Egypt and the Senegal River basin (Fallon and Doenhoff, 1994; Gryseels et al., 1994; Stelma et al., 1995; Ismail et al., 1996), and recently also in Kenya (Melman et al., 2009). Subsequent experimental studies *in vivo* and *in vitro* have shown that isolates from these foci were indeed less susceptible to PZQ than control strains and that PZQ-R strains can be selected for in laboratory maintained *S. mansoni* isolates (Fallon and Doenhoff, 1994; Fallon et al., 1995; Ismail et al., 1999; William et al., 2001; Cioli et al., 2004; Melman et al., 2009). However, a recent study within the same community in Egypt did not find anymore an indication of resistant strains after 10 years of chemotherapy (Botros et al., 2005). The extremely low cure rates reported from Senegal can probably be explained, albeit in part, by epidemiological factors (Cioli, 2000).

In contrast to PZQ, the mechanism of action of oxamniquine is well understood: it has to be activated by a parasite sulfotransferase and resistant schistosomes lack the enzymatic activity (Pica-Mattoccia et al., 2006). Resistance of schistosomes to oxammniquine has been documented and confirmed by both *in vivo* and *in vitro* studies (Cioli et al., 1993; Valentim et al., 2013). The drug is being replaced by PZQ (Cioli et al., 1993).

1.3 Molecular epidemiology of animal macroparasites

Parasitism is thought to be the most common lifestyle on Earth (Price, 1980; Poulin and Morand, 2004) with at least 60 independent evolutionary transitions from a free-living to a parasitic lifestyle (Poulin and Morand, 2004). It is increasingly regarded as one of the driving forces of evolution. This awareness has stimulated biologists to study the ecology and evolution of parasites and infections, with its first formalization about 40 years ago (Price, 1980). In the past 10 years, interest in parasite ecology and evolution has reached an unprecedented level, resulting in some major advances in our understanding of parasite biodiversity, transmission dynamics, patterns of speciation, host-parasite co-evolution and the evolution of host specificity and drug resistance (Poulin, 2007). These achievements were aided thanks to the explosion of applying molecular methods in parasitological research (Blouin et al., 1995; Nadler, 1995; Tibayrenc, 1995; Criscione et al., 2005).

The emergence of molecular techniques and its application to parasitic organisms furthermore stimulated collaborations between population geneticists and epidemiologists, giving rise to a new subspecialty named ‘molecular epidemiology’ (Paterson and Viney, 2000; Foxman and Riley, 2001; de Meeus et al., 2007; Archie et al., 2009). The term was first coined in 1973 by Kilbourne in an article entitled “The molecular epidemiology of influenza” (Kilbourne, 1973). Although many definitions of molecular epidemiology were put forward (reviewed in Foxman and Riley, 2001), the name *sensu stricto* refers to both “molecular”, the use of molecular techniques to characterize nucleic acid- or amino acid-based content and “epidemiology”, the study of the distribution and determinants of disease occurrence in (human) host populations (Foxman and Riley, 2001). The reason for the integration of both disciplines is because of the clear analogy between the transmission of genes (heritable information) from one generation to the next and the transmission of parasites (infection) from one host to another (Paterson and Viney, 2000). Molecular genetic data therefore complement and enhance the resolution of classic epidemiological methods because they

improve the ability to identify parasite origins and (risk) factors that influence their spread and transmission (Foxman and Riley, 2001; Archie et al., 2009).

Although population genetic theories can be readily applied to animal macroparasites such as schistosomes, an additional introductorion is needed to clarify some terminology and concepts. First of all, what is a parasite population? Ideally, populations are defined as a group of randomly mating organisms of the same species that occupy the same space in time and comprise a unique gene pool (Hartl and Clark, 2007). However, as parasites are restricted to their host, other definitions of populations are needed. As such, Esch et al. (1975) developed the concept of parasite infrapopulation that includes all parasites of a given species in one host individual, and parasite suprapopulation that includes all the parasites of a given species in all stages of development within all hosts in an ecosystem (Esch et al., 1975). The terminology was extended by Bush et al. (1997) that defined component population as all the infrapopulations of a given species in an ecosystem which is similar to the term metapopulation that is used by many population ecologists (Goater et al., 2013). An important characteristic of schistosomes is that infrapopulations are formed by recruitment (immigration) from the suprapopulation and not as a result of birth within (or on) the host (Nadler, 1995). This special characteristic has many consequences towards the (interpretation of) population dynamics and thus the genetics of animal macroparasites. Although infrapopulations may represent the group of breeding individuals, the genetic composition of such a group may change by recruitment during the life span of the host. This recruitment is affected by factors that limit infrapopulation density (e.g. concomitant immunity) and by the dispersal capability of their most mobile host(s) (Blouin et al., 1995; Jarne and Theron, 2001; McCoy et al., 2003a; Criscione and Blouin, 2004; Prugnolle et al., 2005c). Studying the distribution of parasite genotypes among infrapopulations can thus be very useful for understanding the ecology of transmission (Anderson et al., 1995; Nadler, 1995; Jarne and Theron, 2001; Criscione et al., 2005). For instance, parasite genotypes may be clustered within households, within social host groups or may occur randomly throughout the whole village, potentially revealing insights into the factors that shape parasite transmission.

Although *F*-statistics are the most commonly used estimates to describe the distribution of genetic variability within and among given populations, it is important to note that there are

several concerns when applying and interpreting those estimates to infrapopulations. First, most studies rely on the genotyping of the offspring because the adult worms are inaccessible for sampling (de Meeus et al., 2007). Sampling a large number of related offspring from a limited number of adults could inflate F_{ST} estimates between infrapopulations (also known as the Allendorf-Phelps effect) and lead to biased estimates of F_{IS} and linkage disequilibrium (Steinauer et al., 2013). Biased and imprecise estimation is furthermore complicated when dealing with small sample sizes (Waples, 1998), which is sometimes unavoidable due to low infection intensities or the highly aggregated distribution of some parasite species.

1.4 Outline and aims

In view of the introduction of schistosomiasis into new areas and the increased artificial selection pressures invoked by control programs, it is critical to understand the factors that shape the transmission of *Schistosoma* parasites at a local and regional geographic scale. Remarkably little is known about the factors that control levels of genetic diversity, genetic drift and gene flow among populations of schistosome parasites. In this thesis a molecular epidemiological approach was used to study the distribution of *Schistosoma mansoni* parasites in Northwest Senegal and assess how colonization history, host-specific factors and drug treatment affect the demography of these parasites.

As with many other parasites, studying the genetic variability of *Schistosoma mansoni* populations is a challenging endeavour. Direct observation of infrapopulations infecting humans is impossible because of their small body size and the site of infection (blood vessels). Their population biology is therefore mainly studied indirectly through their offspring (miracidia, eggs or cercariae). Sampling and preserving thousands of microscopically small larvae under subtropical conditions in such a way that reliable molecular analyses can be performed requires special attention. Before addressing biological questions, a sampling protocol was optimized that allowed reliable and cost-effective genotyping of many individual *S. mansoni* parasites (**chapter 2**).

In **chapter 3** a macro-epidemiological study was performed to reconstruct the disease outbreak of *S. mansoni* since its introduction in Northwest Senegal about thirty years ago. Despite many epidemiological studies that have emphasized the special nature of this highly epidemic focus of *S. mansoni*, little is known about its colonization history. Genetic variation

was therefore studied within and among *S. mansoni* parasites obtained from Northwest Senegal, Southeast Senegal and Mali. More specifically, we wanted to assess whether the sudden epidemic of *S. mansoni* was triggered by a limited number of strains (i.e. founder effect) or by many different strains, whether the *S. mansoni* population size has changed since its introduction and how the genetic variability of *S. mansoni* was distributed among villages.

Once the distribution of the introduced *S. mansoni* strains in Northwest Senegal and neighbouring regions was clarified, three villages in Northwest Senegal were studied to reveal the distribution of these parasites among individual hosts (**chapter 4**). The main aim of this study was to investigate to what extent the allocation of *S. mansoni* parasites among hosts is influenced by host-specific factors such as age and gender. As population genetic analyses were performed at the infrapopulation level, additional analyses were performed to avoid biases due to sampling and Wahlund effects.

In the last two chapters, the effect of community-based drug treatment on levels of genetic diversity in *S. mansoni* populations was explored. Although treatment is expected to induce huge genetic bottlenecks due to the strong decline in population sizes, no theoretical framework exists that explores the possible impact of treatment on schistosome populations. An island model at equilibrium was used to simulate the effect of treatment on genetic diversity in schistosome populations. This was done for different scenarios regarding the amount of treated hosts within a community, the effectiveness of treatment, the frequency of treatment and the pre-treatment infection intensities (**chapter 5**). To assess changes in parasite population diversity, size and structuring in a natural setting, *S. mansoni* parasites were collected and genotyped before and after they were exposed to the drug praziquantel (**chapter 6**).

All these studies provided insight into how colonization history, host-specific factors and drug treatment might shape the (distribution of) genetic variability of *S. mansoni* populations. It demonstrated that the amount of genetic diversity is crucial for both long-term and short-term dynamics of populations, allowing them to adapt to changing conditions. This thesis therefore increased the understanding of the evolutionary potential of these parasites in a dynamic world and allowed us to make some assumptions as to why they are such a successful parasite.

CHAPTER 2

Optimal sample storage and extraction procotols for reliable multilocus genotyping of the human parasite *Schistosoma mansoni*

F Van den Broeck, S Geldof, K Polman, FAM Volckaert & T Huyse

Published in Infection, Genetics and Evolution 2011, 11 (6): 1413-1418

Abstract

Genotyping individual larval stages and eggs of natural parasite populations is complicated by the difficulty of obtaining reliable genotypes from low quantity DNA template. A suitable storage and extraction protocol, together with a thorough quantification of genotyping errors are therefore crucial for molecular epidemiological studies. Here we test the robustness, handling time, ease of use, cost effectiveness and success rate of various fixation (Whatman FTA® Classic and Elute Cards, 70% EtOH and RNAlater®) and subsequent DNA extraction methods (commercial kits and proteinase *K* protocol). None of these methods require a cold chain and are therefore suitable for field collection. Based on a multiplex microsatellite PCR with nine loci the success and reliability of each technique is evaluated by the proportion of samples with at least eight scored loci and the proportion of genotyping errors. If only the former is taken into account, FTA® Elute is recommended (83% success; 44% genotyping error; 0.2€/sample; 1h20m handling time). However, when also considering the genotyping errors, handling time and ease of use, we opt for 70% EtOH with the 96-well plate technology followed by a simple proteinase *K* extraction (73% success; 0% genotyping error; 0.2€/sample; 15m handling time). For eggs we suggest 1) to pool all eggs per person in 1.5ml tubes filled with 70% EtOH for transport and 2) to identify each egg to species level prior to genotyping. To this end we extended the Rapid Diagnostic PCR developed by Webster et al. (2010) with a *S. mansoni*-specific primer to discriminate between *S. mansoni*, *S. haematobium* and *S. bovis* in a single PCR reaction. The success rate of genotyping eggs was 75% (0% genotyping error). This is the first study to incorporate genotyping errors through re-amplification for the evaluation of schistosome sampling protocols and the identification of error-prone loci.

2.1 Introduction

The use of molecular data in epidemiological studies allows researchers to elucidate parasite transmission patterns, intermediate host specificity and interactions between closely related parasite species. By studying the population genetic structure of parasites, factors influencing the spread of infectious agents can be inferred and the gene flow between parasite populations can be estimated (Archie et al., 2009). This information is crucial to estimate if and how resistance or virulence alleles can spread among and between populations. Genetic structure analyses are therefore increasingly recognised as powerful tools in epidemiological and evolutionary research of many parasites, as is the case for the endoparasites *Schistosoma* spp. (Platyhelminthes, Digenea). These parasites are the underlying cause of schistosomiasis, a tropical disease of profound medical and veterinary importance affecting about 200 million humans in 76 developing countries (Steinmann et al., 2006). The schistosome lifecycle consists of an obligatory alternation of generations between a mammalian definitive host, in which fertilized female worms produce eggs that leave the body and hatch as miracidia upon contact with water, and a molluscan intermediate host, in which miracidia undergo an asexual reproduction that develop into cercariae.

Studying the genetic diversity of natural *Schistosoma* populations is complicated by the fact that adult worms are inaccessible in the blood circulatory system of the mammalian host, and the larval stages are very small (<200 µm; Rollinson and Simpson, 1987). From 2005 onwards, several protocols were developed to collect larval stages and eggs in the field (Shrivastava et al., 2005; Sorensen et al., 2006), thereby circumventing the ethical, technical and biological disadvantages of laboratory passage (Curtis and Minchella, 2000). These protocols require however a cooling chain (-20 °C), which is less amenable for field sampling. A significant breakthrough was achieved through FTA® Cards (Whatman), allowing the long-term storage of field-collected larval stages at room temperature and easy transport (Gower et al., 2007). Despite this important step forward, samples could only be analyzed once (Gower et al., 2007), which is a major disadvantage if one wants to perform different genetic analyses or test for repeatability. Recently, RNA*later*® has been tested for fixation of both miracidia and eggs at room temperature (Webster, 2009). The quality of the DNA collected in RNA*later*® was verified by PCR and sequencing but not by microsatellite analysis, which is

more sensitive to low quantities of template DNA.

These recent improvements in sampling strategies, and the availability of several microsatellite markers for *S. mansoni* (Durand et al., 2000; Blair et al., 2001; Curtis et al., 2001; Rodrigues et al., 2002; Silva et al., 2006) have enabled the direct large-scale sampling and exhaustive genotyping of individual larval stages. However, the limited amount of template DNA obtained from schistosome larval stages can lead to potential genotyping errors (i.e. observed genotype differs from true genotype), which can seriously affect the final conclusions of a study (e.g. Constable et al., 2001). As these errors can be generated at every step of the genotyping process from sampling to data analysis (Bonin et al., 2004), it is very important to evaluate sampling protocols not only by the success rate, but also by the quantification of genotyping errors.

Using a multiplex microsatellite PCR with nine loci, we compare four sampling and extraction protocols in order to find the most practical, time and cost effective method that guarantees a maximum number of scored loci (genotyping success) that are also reliable (genotyping error). This is the first study for schistosomes that 1) compares such a wide variety of fixating agents and extraction protocols, 2) uses nine multiplexed loci and 3) evaluates loci and protocols by the quantification of genotyping errors. Special attention has been paid to the processing of eggs and the rapid diagnosis up to species level by means of an extended diagnostic multiplex PCR.

2.2 Materials and methods

All extractions and PCR preparations have been executed in pre-PCR conditions. This is important as PCR-products (contamination) present in post-PCR facilities can interfere with down-stream analyses when dealing with low quantities of template DNA.

2.2.1 Ethical aspects

This study is part of a larger investigation of schistosomiasis epidemiology, transmission and control in Senegal, for which approval was obtained from the ethical committees of the Ministry of Health in Dakar (Senegal) and the Institute of Tropical Medicine in Antwerp (Belgium). Oral consent was obtained from all parents and teachers for urine and stool examination and the data were analyzed anonymously. All schistosomiasis positive inhabitants were treated with a single dose of praziquantel at 40 mg/kg of bodyweight.

2.2.2 Strategies used for sampling larval stages and eggs in the field

The respective protocols were chosen because they allow large-scale sampling and do not require a cooling chain and/or an equipped DNA laboratory, making them ideal for field sampling. First, we optimised the protocol described by Gower et al. (2007) by processing the FTA® Classic cards with a commercial DNA extraction kit (Nucleospin® Tissue Kit). This allowed multiple PCR analyses per sample. The need to process cards with an extraction kit is automatically avoided in the more recently developed FTA® Elute technology with Whatman extraction, being also much cheaper and faster. FTA® cards, however, can only be used for the storage of larval schistosome stages (miracidia or cercariae), but not for eggs (personal data). RNA*later*® is suitable for the fixation of both eggs and larvae (Webster, 2009) but contains EDTA (a PCR-inhibitor), and thus requires a commercial extraction kit (Nucleospin® Tissue Kit). Finally, we sought a new sampling and extraction protocol for both eggs and larval stages that is inexpensive, uses 96-well technology and requires short handling times. The protocol is based on DNA fixation in ethanol (EtOH), followed by a crude proteinase *K* DNA extraction (Zietara et al., 2000).

2.2.3 Sample collection and DNA fixation

Schistosoma mansoni eggs were filtered from human stool samples obtained from four inhabitants of the village Ndieumeul (Northwest Senegal; January 2010). Per inhabitant, 24 samples were collected for each fixating agent to exclude variation among inhabitants (in total 96 samples per fixating agent and 144 samples per inhabitant; Figure 2.1). Filtered eggs were collected in Petri dishes with bottled spring water, individually fixed into 96-well plates containing 10 µl RNA*later*® or 40 µl EtOH (70%) per sample, and also pooled by inhabitant in 1.5 ml tubes filled with EtOH (70%). After the remaining eggs hatched, miracidia were individually collected onto FTA® Elute (in 3 µl volume of water), or in 96-well plates containing either 10 µl RNA*later*® or 40 µl EtOH (70%) per sample (Figure 2.1).

In addition, snails of the genus *Biomphalaria* were collected in the village Nder (Northwest Senegal; March 2007) and Ndieumeul (January 2010) and they were exposed to sunlight for 5mins. Released cercariae from Nder (2007) were individually collected onto FTA® Classic Cards (in 3 µl volume of water) and stored for three years at room temperature. Cercariae

from Ndieumeul (2010) were fixed in 96-well plates containing 40 µl EtOH (70 %) (Figure 2.1). All samples were stored and transported at room temperature.

2.2.4 DNA extraction

For each protocol, handling time and processing time were recorded and cost per sample was calculated. We adapted all protocols to a 96-well plate technology.

Whatman extraction

A 3.0 mm disc was removed with a Harris Micro Punch from the FTA® Elute Cards at the centre of where the sample was loaded. An initial washing step with 200 µl MilliQ H_2O was performed, followed by the addition of 30 µl MilliQ H_2O and heating for 30 minutes at 95 °C. DNA was eventually collected by centrifuging for 30 seconds; the disc was removed from the sample.

Nucleospin® extraction

The 3.0 mm discs excised from the FTA® Classic Cards were treated like the miracidia and eggs stored into RNA*later*® using the Nucleospin® Tissue kit (Macherey-Nagel). Apart from the washing buffers, all other reagents have been downscaled by ¼ to fit the 96-well PCR plates that were originally used for sampling. This significantly reduced the handling and processing time. Samples were eluted in 100 µl elution buffer, vacuum dried using a Univapo 150 ECH (Sanyo Biomedical Equipment) and re-suspended in a volume of 20 µl MilliQ H_2O to obtain a more concentrated final solution.

Proteinase K *extractions*

Pooled eggs in EtOH were, after transport, individually transferred to a 96-well plate and then treated like the individually collected eggs, miracidia and cercariae. The 96-well plates were centrifuged for 1 min to collect the samples at the bottom of the wells. The EtOH was evaporated by placing the plate in the oven at 82 °C for 1 hour. Then 10 µl MilliQ H_2O and 10 µl lysis buffer (1x PCR buffer, 0.45% (v/v) Tween 20, 0.45% (v/v) NP 40 and 60 µg/ml of proteinase *K*; Zietara et al., 2000) was added, followed by an incubation period at 65 °C for 25 mins and 10 mins at 96 °C to denature the proteinase *K*.

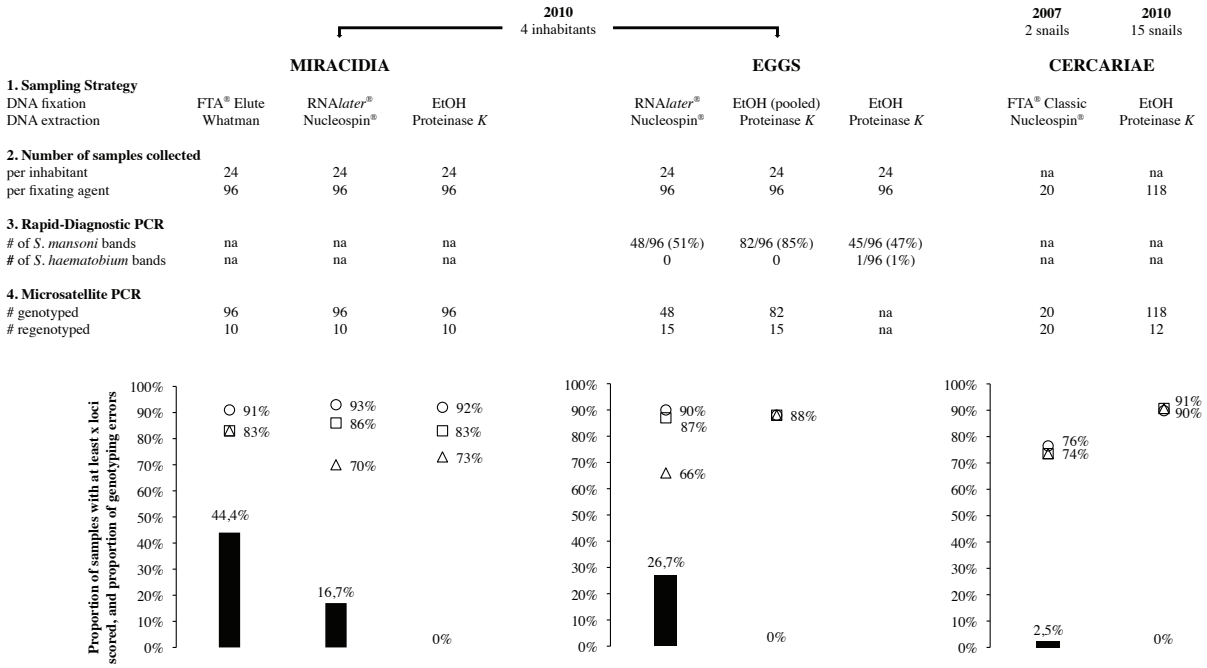


Figure 2.1 Summary of the experimental set-up of this study. Type and number of samples and the outcome for the RD-PCR and microsatellite genotyping are shown. Genotyping success was calculated as the proportion of samples with at least six (circles), seven (squares) or eight (triangles) scored loci and the proportion of genotyping errors (black bars).

2.2.5 Rapid diagnostic multiplex PCR (RD-PCR)

Because eggs isolated from stool samples do not always belong to *S. mansoni* (heterologous pairing between *S. mansoni* and *S. haematobium* (Southgate et al., 1998) or a spill-over in case of high infection intensities), we chose to diagnose all eggs to species level, prior to microsatellite analysis. For this we optimized the previously described RD-PCR amplifying partial cytochrome oxidase I (mtDNA) (Webster et al., 2010) by designing a new reverse primer specific for *S. mansoni* (5'-TGCAGATAAAGCCACCCCTGTG-3'), which amplifies a fragment of 375 bp. This was done based on an alignment of several *S. mansoni* isolates from Cameroon, Senegal, Kenya, Mali, Egypt and Tanzania (unpublished data) with MacVector® 9.5.2 (Accelrys) using the following primer settings: length 18-30 bp, GC% 30-55, Tm (°C) 55-80, (MacVector takes several features into account including self-duplexing, hairpins, specificity and mismatches). PCR amplifications were done in 25 µl reactions, each containing 0.5 units of SilverStar DNA Taq polymerase (Eurogentec), 1x reaction buffer (Eurogentec), 1.5 mM $MgCl_2$, 0.2 mM dNTP's, 1.6 µM of the universal forward primer and 0.8 µM of each of the three species-specific reverse primers (*S. bovis*, *S. haematobium* and *S. mansoni*). The following conditions were used: 3 min at 95 °C, followed by 35 cycles of 30

sec at 94 °C, 1.5 min at 58 °C, 1.5 min at 72 °C with a final extension of 7 min at 72 °C. The specificity of the primer was tested using adult worm DNA from *S. haematobium, S. bovis, S. mansoni, S. curassoni* and *S. rodhaini*. The extended RD-PCR was used on eggs as a diagnostic tool (species identification) and as a test for DNA quality control (amplification success). Prior knowledge on species status and DNA quality is desirable because the microsatellite multiplex PCR is costly and specific to *S. mansoni* only. The RD-PCR was also used to compare the amplification success of pooled *versus* individually fixed eggs in EtOH.

2.2.6 Microsatellite analysis

To determine the success of the different extraction protocols, we genotyped all samples using nine previously characterized microsatellite markers in a single multiplex PCR reaction, namely *SMDA28*, *SMD43*, *Ca11-1*, *SMS9-1*, *SMD28*, *L46951*, *SMD25*, *SMD89* and *SMD11* (Durand et al., 2000; Blair et al., 2001; Curtis et al., 2001). Loci were fluorescently labelled using the dye's 6-FAM (blue), VIC (green), NED (black) and PET (red). PCR amplifications were performed in 10 µl reactions with 2 µl DNA template, 5 µl QIAGEN Multiplex PCR Master Mix (HotStarTaq® DNA Polymerase, Multiplex PCR Buffer and dNTP Mix), 2.58 µl MilliQ H_2O, 0.11 µl of primer L46951 (20 µM) and 0.04 µl of each of the other primers (20 µM). Thermal cycling was conducted under the following conditions: 15 min at 95 °C, followed by 45 cycles of 30 sec at 94 °C, 1.5 min at 58 °C, 1.5 min at 72 °C with a final extension of 30 min at 60 °C. Products were analysed using an ABI 3130 Genetic Analyser (Applied Biosystems) and GeneScan™ 500 LIZ™ as Size Standard. Allele sizes were manually verified using GENEMAPPER v4.0 (Applied Biosystems). TANDEM v1.07 was used for an automated binning of allele lengths (Matschiner and Salzburger, 2009).

2.2.7 Quantification of genotyping errors

Genotyping errors can have various causes such as the failure to amplify an allele due to primer-site mutations (null alleles), low template quantity or quality (allelic dropouts), human errors (e.g. scoring errors) or a combination of these three (Pompanon et al., 2005). First, the software package MICRO-CHECKER v2.2.3 (Van Oosterhout et al., 2004) was used to check our microsatellite data for scoring errors, allelic dropouts and null alleles (i.e. non-amplified allele due to mutation in primer target sequence). Next, genotyping errors were quantified by re-amplifying at least 10% of all samples, randomly chosen per sampling

protocol, which is recommended by Bonin et al. (2004) for an accurate measure of error rates. Mismatches between the two replicates were identified and classified according to four categories: 1) Scoring errors due to stuttering; 2) Multiple peaks (i.e. more than 2 peaks of similar height); 3) Allelic dropout (i.e. only one of the two alleles present at a heterozygous locus was amplified) or False allele (an allele-like PCR-generated artefact) or Mistaken allele (i.e. an allele that does not correspond to the true allele, excluding the null allele, allelic dropout and false allele); and 4) No amplification. As only two replicates were obtained, it was not possible to assign a reference genotype, making it impossible to discriminate among allelic dropout, false alleles and mistaken alleles with certainty. Several formulae, valid for co-dominant markers and based on the mismatches observed between replicates, were used to quantify error rates. The observed error rate per multilocus genotype (Multilocus Genotype Error Rate or MGER) was calculated as following: $e_{obs} = m_g / nt$, with m_g the number of multilocus genotypes including at least one allelic mismatch, and *nt* the number of replicated multilocus genotypes (Pompanon et al., 2005). MGER reflects the reliability of the obtained genotypes and is thus useful to assess the quality of the DNA sample. To identify error-prone loci, an additional 70 miracidia fixed in EtOH were re-genotyped yielding a total of 140 replicates. The Mean Error Rate per Locus (MERL) was quantified using the formula $e_l = m_l / nt$, with m_l the number of single-locus genotypes including at least one allelic mismatch, and *nt* the number of replicated single-locus genotypes (Pompanon et al., 2005).

2.3 Results

2.3.1 Validation of the multiplex RD-PCR analysis

Figure 2.2 (lanes 3-6) shows that the multiplex PCR always generated clear single PCR bands when the following adult DNA was used: *S. haematobium* (543 bp), *S. bovis* (306 bp), *S. mansoni* (375 bp) and *S. curassoni* (306 bp). For *S. rodhaini* no amplification was observed (Figure 2.2, lane 7). When a mix of adult DNA (*S. haematobium*, *S. bovis* and *S. mansoni*) was used, three clear bands were obtained (Figure 2.2, lane 8). The RD-PCR amplification success of eggs fixed in RNA*later®* and 70% EtOH was used as a quality control before genotyping. On average, 51% and 47% of all individually fixed eggs in RNA*later®* and EtOH respectively

yielded successful amplification (Figure 2.1). The amplification success of pooled fixed eggs in EtOH was 85% (Figure 2.1).

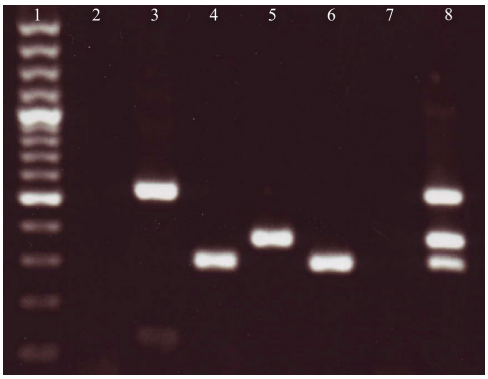


Figure 2.2: Rapid Diagnostic Multiplex PCR on *Schistosoma* spp.
Lane 1: 1500 bp size standard
Lane 2: Blanco
Lane 3: *S. haematobium* (543 bp)
Lane 4: *S. bovis* (306 bp)
Lane 5: *S. mansoni* (375 bp)
Lane 6: *S. curassoni* (306 bp)
Lane 7: *S. rodhaini*
Lane 8: Mixed DNA from *S. haematobium* (543 bp), *S. mansoni* (375 bp) and *S. bovis* (306 bp)

2.3.2 Validation of the multiplex microsatellite analysis

Table 2.1 summarizes the Mean Error Rates per Locus (MERL) and the results obtained from MICRO-CHECKER. Locus *SMD43* showed an error rate of 11%, while the other loci ranged between 0% for *SMD89* to 4.29% for *L46951* (Table 2.1). Locus *SMD11* and *L46951* proved to be weak amplifiers compared to the other loci (1.43%-2.86%), while scoring problems due to stuttering were observed for loci *CA11-1* and *SMDA28* (1.43%-2.27%) and allelic dropout, false alleles or mistaken alleles were observed for loci *CA11-1* and *SMD43* (1.42%-2.14%; Table 2.1). Null alleles were present at locus *SMD11* and *SMD43* (Table 2.1). MICRO-CHECKER indicated no problems regarding stutters or allelic dropout (results not shown).

2.3.3 Processing and handling time, cost per sample

Table 2.2 summarizes the practical specifications of all sampling strategies (fixating agent + extraction protocol). Processing (≈8h), handling time (≈2h) and cost per sample (€ 2.5) were similar when Nucleospin® was used in combination with RNA*later*® or FTA® Classic Cards (Table 2.2). The processing of eggs, miracidia and cercariae stored in EtOH using proteinase *K* extractions took 1h50m, the handling time was 15 m and the cost per sample was € 0.2 (Table 2.2). FTA® Elute showed intermediate values for handling time (1h20m) and the cost per sample was € 0.2 (Table 2.2).

Table 2.1 Evaluation of each of the nine microsatellite markers used, based on the Mean Error Rate per Locus (MERL; in %) and the results of MICRO-CHECKER (0 = not present; 1 = might be present). In order to calculate MERL, 70 miracidia fixed in EtOH were re-genotyped, yielding a total of 140 replicates.

Locus	*L46951*	*CA11-1*	*S9-1*	*SMD11*	*SMD25*	*SMD28*	*SMD43*	*SMD89*	*SMDA28*
Repeat motif	3	2	2	4	2	3	4	2	4
Range	160-230	190-230	180-225	290-420	260-312	225-250	126-180	130-185	90-125
1. Mean Error Rate per Locus									
Stutters	0.00	1.43	0.71	0.00	0.71	0.00	0.00	0.00	2.27
Multiple peaks	0.71	0.71	0.00	1.43	0.00	0.00	8.57	0.00	0.00
A.D / F.A. / M.A.	0.71	1.42	0.71	0.71	0.00	0.00	2.14	0.00	0.00
No amplification	2.86	0.00	0.00	1.43	0.71	0.71	0.71	0.00	0.71
e_I	**4.29**	**3.57**	**1.42**	**3.57**	**1.42**	**0.71**	**11.42**	**0.00**	**2.98**
2. MICRO-CHECKER results									
Null Alleles	0	0	0	1	0	0	1	0	0

A.D. = Allelic Dropout; F.A. = False Allele; M.A. = Mistaken Allele

Table 2.2 Processing time (96 samples), handling time (96 samples) and cost (1 sample; VAT excluded) for each sampling strategy (fixating agent + extraction protocol).

	Processing time	Handling time	Cost / sample
RNA*later*® + Nucleospin®	8h00m	2h00m	€ 2.5
FTA® Classic + Nucleospin®	8h30m	2h30m	€ 2.5
FTA® Elute + Whatman	1h50m	1h20m	€ 0.2
EtOH + proteinase *K*	1h50m	15m	€ 0.2

2.3.4 Sample quality

Figure 2.1 shows the genotyping success of the different sampling and extraction protocols based on two criteria (success and reliability). For miracidia, the average proportion of samples with at least six, seven or eight loci amplified ranged between 91%-93% for six loci, 83%-86% for seven loci and 70%-83% for eight loci (Figure 2.1). Eggs stored individually in RNA*later*® resulted in 90% successfully genotyped eggs when the threshold was six loci, while for pooled eggs in EtOH a success rate of 88% was obtained (Figure 2.1). When the threshold was set to eight loci, 66% and 88% was obtained for RNA*later*® and EtOH respectively. The genotyping success of cercariae stored for three years on FTA® Classic

Cards was 76% for six loci, and 74% for seven and eight loci amplified, while 90% of the cercariae stored individually in EtOH had eight loci amplified (Figure 2.1).

The observed Error Rate per Multilocus Genotype (MGER) was 44.4% for miracidia stored on FTA® Elute Cards, 16.7% for RNA*later*®, 0% for EtOH and MGER ranged between 0% for eggs stored as a pooled sample in EtOH and 26.7% when stored in RNA*later*® (Figure 2.1). About 97.5% and 100% (2.5% and 0% MGER) of the replicate genotypes obtained from cercariae stored on FTA® Classic Cards and EtOH respectively were identical (Figure 2.1).

2.4. Discussion

We compared and evaluated four sampling and extraction protocols (EtOH – proteinase *K*; RNA*later*® - Nucleospin®; FTA® Elute - Whatman; FTA® Classic - Nucleospin®) by 1) the proportion of samples with eight successfully scored loci (i.e. amplification success), and 2) the proportion of genotyping errors based on mismatches between replicates of the same DNA sample. We aimed for the most practical, time and cost effective method, yielding a high amplification success and the least number of genotyping errors.

2.4.1 Validation of multiplex RD-PCR and multiplex microsatellite PCR

In order to diagnose all eggs to species level prior to microsatellite analysis, we extended the multiplex PCR developed by Webster et al. (2010) by adding a new reverse primer specific to *S. mansoni.* One universal forward primer and three species-specific reverse primers amplified a partial *cox*1 fragment of 543 bp for *S. haematobium*, 375 bp for *S. mansoni* and 306 bp for *S. bovis* and *S. curassoni*. The newly developed primer specific to *S. mansoni*, did not amplify in *S. bovis*, *S. haematobium, S. curassoni* or *S. rodhaini* (Figure 2.2). Thus a single PCR reaction followed by fragment separation by gel electrophoresis allows a rapid and reliable species diagnosis.

Our microsatellite multiplex of nine loci contains seven loci that have been validated for potential errors like allelic dropout, false alleles and lack of amplification using a direct pedigree analysis (Steinauer et al., 2008a). As EtOH fixation and subsequent proteinase *K* DNA extraction proved to be the most reliable protocol (see section 4.2), only miracidia fixed in EtOH were used to quantify the Mean Error Rate per Locus (MERL). Locus *SMD43* showed the highest error rate (11.42% MERL), mainly due to the generation of multiple peaks, and null alleles were present (Table 2.1). We therefore chose to remove this locus from

subsequent analysis. All other loci ranged between 0% and 4.3% MERL (Table 2.1), which is about half the error rates obtained by Steinauer et al. (2008) (0.9% - 6.9%). Since both methods are quite different in principle, a meaningful comparison of error rates is precluded and we therefore underline the importance of a consensus strategy to quantify genotyping errors (Pompanon et al., 2005).

2.4.2 Sampling and extraction protocols evaluated by genotyping success and reliability

For miracidia, FTA® Elute + Whatman extraction appeared the most successful protocol when the threshold was eight successfully scored loci (83%), but the most reliable genotypes were obtained with EtOH + proteinase *K* extraction (0% MGER; Figure 2.1). As the genotyping success of samples stored in EtOH does not differ significantly from those stored on FTA® Elute, we recommend the use of EtOH as a fixating agent for miracidia as it guarantees reliable genotypes.

Eggs that were transported as a pooled sample in 1.5 ml tubes filled with 70% EtOH scored significantly better (85%) in the RD-PCR than those that were individually transported in EtOH (47%) (Figure 2.1). This difference might be due to the fact that, unlike miracidia, eggs do not immediately absorb EtOH. It is therefore important that eggs stay completely immersed during transport, which is more difficult to achieve in the multi-well plates that are only filled with 40 µl EtOH. A dilution of PCR inhibitors might be another/additional explanation. We did not test this pooling design for RNA*later*® but suspect a similar outcome. Of the 85% successfully extracted eggs, 88% resulted in 8 successfully scored microsatellite loci (Figure 2.1). This is a better result (0.88*85=75%) than Beltran et al. (2008), who obtained a lower success rate (52%) with a much less stringent threshold (1/5 loci scored). Furthermore, repeatability testing showed that 100% of all replicated eggs yielded exactly the same genotype as the first replicate (0% MGER; Figure 2.1).

The success rate (threshold 8/8 loci) of cercariae stored for three years at room temperature on FTA® Classic Cards was still 74% and the reliability was high (2.5% MGER; Figure 2.1). Although cercariae stored in EtOH yielded better results both in terms of success (90%) and reliability (0% MGER; Figure 2.1) than FTA® Classic, we cannot compare these two protocols due to a difference in storage time.

2.4.3 Towards an optimal sampling strategy

Our study clearly showed that EtOH fixation and subsequent crude proteinase *K* extraction yielded the best results in terms of reliability (Figure 2.1), cost and handling time (Table 2.2) for eggs, miracidia and cercariae. The 96-well technology allows for an efficient sampling, transport and extraction protocol, with a cost of €0.2 per sample, and about a quarter of an hour handling time to process 96 larvae and eggs. One drawback of EtOH is that transport by air is less convenient, as specific procedures must be followed when packaging and shipping hazardous material. This problem is avoided when using RNA*later*® or FTA® Elute cards, but these sampling strategies yielded higher genotyping errors. FTA® Classic fixation with subsequent Nucleospin® extraction is a successful and reliable alternative, albeit more expensive and labour intensive, and it is not suitable for eggs.

2.5 Conclusions

The best strategy for fixating larval schistosome stages is 96-well plates containing 70% EtOH. For eggs we suggest 1) to pool all eggs per person in 1.5ml tubes filled with 70% EtOH before transport and 2) to test for DNA quality and species status using the extended RD-PCR before genotyping. All loci but *SMD43* showed low error rates. As such, the current sampling and extraction protocol, together with the above multiplex assay should guarantee a thorough and reliable population genetic analysis of natural *S. mansoni* populations.

CHAPTER 3

Invasion genetics of *Schistosoma mansoni* in Northwest Senegal reveals signatures of population expansion

F Van den Broeck, G Maes, D Rollinson, B Hellemans, K Vereecken, I Talla, I Sy, F Volckaert, K Polman & T Huyse

Abstract

The construction of two dams in the Senegal River Basin (SRB) in the mid '80s led to the introduction and spread of *Biomphalaria pfeifferi* snails, the intermediate host of the human parasite *Schistosoma mansoni*. This allowed the parasite to successfully colonize human populations in the Delta and part of the Middle Valley of the SRB within just a few years. In this study molecular markers were used in an attempt to reconstruct the invasion of *S. mansoni* parasites in Northwest Senegal. More specifically, the spatio-temporal genetic diversity and structure was estimated at microsatellite markers and a partial *cox*1 fragment of samples obtained from several localities in Northwest Senegal over a 14-year time frame (1993-2007), from one locality in Southeast Senegal in 2011 and from one locality in Southwest Mali in 2008. Genetic diversity of *S. mansoni* in Northwest Senegal was high at the mitochondrial level compared to that of *S. mansoni* in other African countries, and the diversity at nuclear level was similar to those found in neighboring West African countries. Parasites sampled in 1993 at the onset of the epidemic showed similar levels of diversity as those sampled 14 years later in 2007. Furthermore, results revealed that parasites from Northwest Senegal experienced a population expansion, that they had a West-African origin and that some parasites were genetically different from others (chapter 3). These findings suggest that the epidemic of schistosomiasis in Northwest Senegal was probably not elicited by a few *S. mansoni* parasites, but that the colonization history is much more complex. It is most likely that a multitude of parasites successfully colonized the local human population. Results of the genetic structure of *S. mansoni* parasites were remarkably similar to previously published results of the genetic structure of its intermediate snail host *B. pfeifferi* in the same region. Our study therefore implies that the distribution of intermediate snail hosts could be an important factor determining the invasion success of human *S. mansoni* parasites.

3.1 Introduction

Environmental change and increasing movements of people and animals lead to species introductions into new areas. The colonization, establishment and the success of the introduced species depend on a number of biotic and abiotic factors (Mack et al., 2000; Kolar and Lodge, 2001; Sakai et al., 2001; Suarez and Tsutsui, 2008). In the case of parasitic organisms, the life cycle is paramount in determining the success of colonization (Torchin et al., 2003). Parasites with a direct life cycle (use of single host species) can readily invade new areas together with their host while parasites with a complex life cycle need the presence of one or more intermediate host species in order to establish successfully. This is exemplified by the epidemic outbreak of human intestinal schistosomiasis in Northwest Senegal in 1986. This debilitating disease is caused by the digenean *Schistosoma mansoni* that cycles through two hosts each generation: a human final host and a snail intermediate host of the species *Biomphalaria pfeifferi* (Rollinson and Simpson, 1987). As the Senegal River Basin (SRB) suffered from severe droughts during the 1970s and 1980s (Verheye, 1995), two dams were build to improve the agricultural conditions for rice production: the Diama dam near the mouth of the Senegal River and the Manantali dam upstream in Mali on the Bafing River (Southgate, 1997). Subsequent agricultural and hydrological changes were accompanied by 1) strong agro-industrial developments at Richard Toll, resulting in a massive immigration of agricultural workers from neighboring regions in Senegal, Mali and Mauritania to the SRB (Talla et al., 1990; Handschumacher et al., 1992), and 2) major ecological changes such as lower salinity levels and the formation of open and permanent water bodies, favoring the growth and spreading of *B. pfeifferi* snails (Vercruysse et al., 1994). These factors promoted the invasion of *S. mansoni* into the area followed by one of the most severe outbreaks of intestinal schistosomiasis ever described (Talla et al., 1990, 1992; Gryseels et al., 1994; Verle et al., 1994; Picquet et al., 1996; Southgate, 1997).

Before the construction of the two dams, human population densities were relatively low in Northwest Senegal and they were concentrated around Saint-Louis, Ross Bethio and Richard Toll. The parasite *S. mansoni* was absent and the intermediate snail host *B. pfeifferi* was only reported in low densities (< 1%) with its distribution being restricted to the city Saint-Louis, Lake Guiers and the village Pakh (Chaine and Malek, 1983; Vercruysse et al., 1985). *Schistosoma mansoni* was first reported in 1988 in Richard Toll, the supposed epicenter of

the epidemic outbreak (Talla et al., 1990), where *B. pfeifferi* represented 70% of all collected snails and 44% of them were infected with *S. mansoni* (Diaw et al., 1991). The number of cases of intestinal schistosomiasis increased rapidly to epidemic proportions (Talla et al., 1992; Sow et al., 2002), and soon after *S. mansoni* and its intermediate host *B. pfeifferi* colonised much of the Lower and part of the Middle Valley of the SRB (Picquet et al., 1996). Population genetic studies on the intermediate snail host *B. pfeifferi* revealed very low levels of genetic diversity and differentiation in the region of Richard Toll (Campbell et al., 2010). These results were explained by a rapid expansion of the most fecund snails that displaced the less fecund ones (Campbell et al., 2010), with fecundity being a cost to resistance (Webster and Woolhouse, 1999). In addition, it was shown that sympatric combinations of *B. pfeifferi* and *S. mansoni* isolates from SRB showed extraordinarily high vectorial capacities, with higher snail longevity and higher frequency of patent infections compared to allopatric combinations (Southgate et al., 2000b). The high degree of parasite/snail compatibility together with the occurrence of dense human populations were probably the main factors explaining the spread of the parasite, the intensity of transmission and the prevalence of infection since its introduction in the SRB.

The epidemic of *S. mansoni* in the SRB presents a unique system to study the evolutionary genetics of invasive species. Introduced populations may be founded by a small number of individuals and may thus be genetically less diverse than the source population from which it originated, i.e. founder effect (Cornuet & Luikart 1996; Sakai et al., 2001; Kolbe et al. 2004; Vrijenhoek & Graven 1992). On the other hand, a high number of founding individuals, multiple introductions from disparate source populations or high gene flow between introduced and source populations may alleviate the loss of genetic variation (Sakai et al., 2001; Suarez and Tsutsui, 2008). Selectively neutral microsatellite markers are ideal tools to study these dynamics due to their high level of polymorphism, providing a tool for detection of population divergence and recent population size changes (Cornuet & Luikart 1996). *Schistosoma mansoni* parasites were genotyped at nine microsatellite markers and a partial mitochondrial *cox*1 fragment to reveal new insights into the colonization history of these parasites since their epidemic outbreak in Northwest Senegal almost 30 years ago. More specifically, we tested whether the sudden epidemic of *S. mansoni* was triggered by a limited number of parasites, whether the *S. mansoni* population

size changed since its introduction and how the genetic variability of *S. mansoni* was partitioned among localities and years.

3.2 Material & Methods

3.2.1 Study sites and data collection

Schistosoma mansoni parasites were collected from human stool samples in 2007 from seven villages (Diadiam, Rhonne, Theuss, Mbodjene, Ndieumeul, Nder and Gaya) in the Delta and part of the Middle Valley of the Senegal River Basin (SRB) and in 2011 from one village (Assoni) in Southeast Senegal near the city Kédougou (Figure 3.1). Ethical approval was obtained from the ethical committees of the Ministry of Health in Dakar (Senegal) and the Institute of Tropical Medicine in Antwerp (Belgium). Stool samples were randomly obtained from school age children. *Schistosoma mansoni* infections were microscopically diagnosed for the presence of eggs by duplicate 41.7mg Kato Katz per stool sample, which is a method for preparing human stool samples prior to searching for parasite eggs. Eggs from positive stool samples were isolated after filtration, hatched and miracidia were individually pipetted onto Whatman FTA® indicator cards in a volume of 3µl of water as described in Van den Broeck et al. (2011). All schistosomiasis positive children were treated with a single dose of praziquantel at 40 mg/kg bodyweight. In schools or classes where the percentage of *S. haematobium* or *S. mansoni* infections were more than 50%, mass treatment of all children was carried out at the end of the study.

In addition, adult worms were obtained from the Schistosomiasis Collection at the Natural History Museum in London (Emery et al., 2012). Worms were collected after one laboratory passage of naturally collected miracidia and/or cercariae from two villages in Senegal (Richard Toll in 1993 and 1994 and Ndombo in 1997 and 2006) and from two villages in Southwest Mali (Wayowayanko and Farako, both in 1993) (Figure 3.1).

Finally, we obtained part of previously published *S. mansoni* genotypes of miracidia that were collected in the village Kokry-Bozo in Southwest Mali in 2007 (Figure 3.1; see Gower et al. (2013) for details on sampling and genotyping). The data was provided by the Imperial College in London as raw genotyping chromatogram files, which were used to manually score allele sizes in GENEMAPPER v4.0 (Applied Biosystems). These genotypes and the genotypes generated in this study (see section below) were then imported into ALLEOGRAM v2.2 (Morin et al., 2009) for binning of allele lengths.

Figure 3.1 Map showing Senegal, Mauritania, Mali, Guinea and Guinea-Bissau. Dots represent sampling locations used in this study (nine villages in Northwest Senegal, two villages in Southeast Senegal and three villages in Southwest Mali). Shaded area shows the Senegal River Basin. Map was made with Natural Earth (www.naturalearthdata.com).

3.2.2 Molecular analyses

Genomic DNA extractions of adult worms and naturally collected miracidia were performed with the Nucleospin Tissue kit (Macherey Nagel). For miracidia, 3 mm discs containing the whole miracidium were excised from the FTA® cards and for worms the whole sample was used as DNA source. DNA was extracted using the Nucleospin® Tissue kit (Macherey-Nagel) following the manufacturer's standard protocols.

All individual *S. mansoni* parasites (both naturally obtained miracidia and lab-derived worms) were genotyped using nine microsatellite loci (*L46951*, *SMD11*, *S9-1*, *CA11-1*, *SMD25*, *SMD28*, *SMD43*, *SMD89*, *SMDA28*; Durand et al., 2000; Blair et al., 2001; Curtis et al., 2001) as described in Van den Broeck et al. (2011). All PCR products were analyzed using an ABI 3130 Genetic Analyser (Applied Biosystems) and GeneScan™ 500 LIZ™ as Size Standard. Allele sizes were manually verified using GENEMAPPER v4.0 (Applied Biosystems). As adult worms may be genetically identical (i.e. clones), genotypes obtained from worms were visually inspected and when identical multilocus genotypes (MLGs) were found within a sample, they were removed from the dataset.

Sequences of the mitochondrial *cox*1 gene (450 bp) were obtained for a part of the miracidia collected in Assoni (Southeast Senegal) and for the adult worms originating from Northwest Senegal and Southwest Mali. Only adult worms with a unique MLG as inferred from the microsatellite genotyping were used for sequencing. This was done using primers Asmit-1 and Schisto-3' (Bowles et al., 1992; Lockyer et al., 2003) in 25µL PCR reactions, each containing 2 µl of DNA template, 0.5 units of Platinum Taq DNA polymerase (Life Technologies), 1x reaction buffer (Life Technologies), 2 mM $MgCl_2$, 0.2 mM dNTPs and 0.8 µM of each primer. PCR conditions were the following: denaturation for 3 min at 95°C, followed by 35 cycles of 45s at 94°C, 45s at 49°C, 45s at 72°C with a final extension of 10 min at 72°C. PCR products were visualized on a 1% agarose gel to check for amplicons, which were sequenced using a Big Dye Chemistry Cycle Sequencing Kit v1.1 in a 3130 Genetic Analyser (Applied Biosystems) using the forward primer Asmit-1. When the quality of the sequence was insufficient, the fragment was also sequenced using the reverse primer Schisto 3'. All *cox*1 sequences were manually edited and aligned using Geneious R6 (http://www.geneious.com/) and species identity was confirmed using BLAST (http://blast.ncbi.nlm.nih.gov/).

3.2.3 Analyzed datasets

A total of three datasets were analyzed. The first dataset, hereafter referred to as DMS1, comprised the microsatellite data that was generated within this study from naturally collected miracidia and from worms after laboratory passage. More specifically, this dataset comprised the microsatellite data from Northwest Senegal over all sampled time periods (i.e. 1993, 1994, 1997 and 2007) and from the village Assoni in Southeast Senegal sampled in 2011. Note that these samples were typed at nine microsatellite markers.

The second dataset, hereafter referred to as DMS2, includes the same microsatellite data as those of DMS1, but complemented with the genotypes obtained from Kokry-Bozo in Southwest Mali (Gower et al., 2013) and reduced to the six microsatellite markers that were shared between both studies (i.e. *CA11-1*, *S9-1*, *SMD25*, *SMD28*, *SMD89* and *SMDA28*).

Finally, the third dataset (DSEQ) comprised the *cox*1 sequences that were generated in this study from samples collected in Northwest Senegal, Southeast Senegal and Southwest Mali, complemented with previously published *cox*1 sequences (Webster et al., 2013b) obtained from Senegal in 2007 (villages Temey and Nder) and 2009 (village Kolda) (Figure 3.1), from

seven other countries in Africa (Niger, Nigeria, Cameroon, Tanzania, Coastal Kenya, Uganda and Zambia) and from Brazil (Accession numbers: JQ289587-JQ289617, JQ289622-JQ289640, JQ289643-JQ289650, JQ289655-JQ289673, JQ289678-JQ289715, JQ289721-JQ289741). Sequences of cercariae and worms from Webster et al. (2013b) were not included as they might be clones from each other, possibly introducing a bias in estimates of diversity. In contrast, sequences generated from worms in this study were included because the microsatellite genotyping allowed us to identify clones and subsequently remove them from the dataset. All sequences were aligned using Muscle as implemented in GENEIOUS.

3.2.4 Phylogeographic analysis of partial cox1 sequences

Genetic diversity at the *cox*1 fragment (DSEQ) were quantified per region and per village by estimating the number of haplotypes (i.e. unique sequences), the number of polymorphic sites, the nucleotide diversity *Π* and the haplotype diversity *h* (Nei, 1987) in DNA-SP v5.10.1 (Librado and Rozas, 2009).

The genealogical relationships between all sequences were explored by constructing two networks based on statistical parsimony (Templeton et al., 1992) in the R package 'pegas' (Paradis, 2010). Haplotypes were first identified using the function *haplotype* and used to construct a network with the function *haploNet.* The number of sequences that represented a given haplotype was logarithmically transformed to narrow high and small values and used to determine the size of its corresponding pie diagram. A first network included all the sequences from the DSEQ dataset. A second network included only sequences from Northwest Senegal, Southeast Senegal and Southwest Mali that were either obtained during this study or that were previously published in Webster et al. (2013b).

3.2.5 Population genetic analysis of microsatellite markers

As natural selection may shape genetic variation differentially in different populations, the neutrality of each locus was tested with the selection detection workbench LOSITAN (Antao et al., 2008). Analyses were performed with 50,000 simulations for the Infinite-Allele-Model (IAM) and the Stepwise-Mutation-Model (SMM), both with the options 'Neutral' mean F_{ST} and force mean F_{ST}. The nominal level for multiple testing was set to 0.05. Analyses were done for both DMS1 and DMS2 using region as level of subdivision. When outlier loci were found, they were removed from any further analyses to avoid bias in microsatellite variation

due to selective effects. In addition, sequences of all microsatellite primers were subjected to a BLAST search (http://blast.ncbi.nlm.nih.gov) using the blastn suite and the nucleotide database v2.2.23 to verify if microsatellites are located within functional regions of the genome.

The observed heterozygosity (Ho), unbiased expected heterozygosity (Hs) and the inbreeding coefficient F_{IS} as estimated by f (Weir and Cockerham, 1984) were quantified in GENETIX v4.05 (Belkhir et al., 1996-2004). The significance of f was tested using 10,000 permutations, which were corrected for multiple testing using Bonferroni corrections. Allelic richness (AR) was estimated using the function *allelic.richness* in the R package 'hierfstat' (Goudet, 2005). Analyses were done per region and per village for DMS1 and DMS2 and per year for DMS1.

Genetic structure was first analyzed using a Factorial Corresponance Analysis (FCA) as implemented in GENETIX to explore the distribution of genotypes in hyperspace. Here, genotypes were pooled according to village and year for both DMS1 and DMS2.

The equivalent θ (Weir and Cockerham, 1984) of the standardized allelic variation F_{ST} and its analogue R_{ST} (Slatkin, 1995) were estimated at each locus and over all loci and this pairwise between regions, villages and years for DMS1 and DMS2. Significant population differentiation was tested for all estimates by 1,000 permutations of individuals among localities. To assess the contribution of genetic drift versus stepwise mutation on genetic differentiation, allele sizes were permuted (1,000) among allelic states under the null hypothesis that allele sizes do not contribute to population differentiation (Hardy et al., 2003). Note that the allele permutation test is expected to remain robust with respect to violations of the mutation-drift equilibrium, an assumption that might be at stake for colonizing species. All computations were done in SPAGEDI v1.4 (Hardy and Vekemans, 2002). Pairwise estimates of F_{ST} between villages were visualized with classical multidimensional scaling (CMDS) plots using the R software. Only samples containing at least 10 genotypes were kept for visualization in order to minimize biases due to sample size.

The ancestry of individual parasites was inferred using a Bayesian Markov chain Monte Carlo (MCMC) clustering analysis as implemented in STRUCTURE v2.2.3 (Pritchard et al., 2000). The number of clusters K was derived assuming the admixture model and correlated

allele frequencies for the two datasets DMS1 and DMS2. Sampling locations were included in the model as a prior (LOCPRIOR = 1), as they could assist clustering when the amount of genetic markers is low (Hubisz et al., 2009). Note that the LOCPRIOR model will not falsely identify genetic structure when none is present and will ignore sampling information when the ancestry of individuals is uncorrelated with sampling locations (Hubisz et al., 2009). For both DMS1 and DMS2 two analyses were run, one with village as sampling location and one with region as sampling location, resulting in a total of four different STRUCTURE jobs. As suggested by the authors, STRUCTURE was also run using the original models without prior information on sampling locations (i.e. LOCPRIOR = 0) to check whether there were major discrepancies with the results of the first model. Three replicate runs were initiated for each predefined *K* (ranging from 1 to 10); each run consisted of 100,000 MCMC chains, initiated by 10,000 burn-in steps. The optimal *K* value was identified by the highest loglikelihood value LnP(D) and its second order rate change ΔK (Evanno et al., 2005). The program DISTRUCT v1.1 was used to visualize the estimated individual membership coefficients (Rosenberg, 2004).

The demographic history of the parasite was assessed using the software Bottleneck v1.2.02 (Cornuet and Luikart, 1996). When a population suffers from changes in effective population size, there will be a correlative change of the allelic diversity and heterozygosity. However, the change in allelic diversity will occur faster than the change in heterozygosity, and therefore a transient excess (bottleneck) or deficiency (expansion) in heterozygosity will be observed when compared to the heterozygosity estimated at mutation-drift equilibrium. Tests for heterozygosity excess and deficiency were performed within each village and region for DMS1 and DMS2 and within each year for DMS1. This was done using 1000 iterations for the 'stepwise mutation model' (SMM) and the 'two-phase model' (TPM) as these models should reflect best the way microsatellites evolve, with TPM providing the most realistic picture (Di Rienzo et al., 1994; Piry et al., 1999). The TPM was run with a variance among multiple steps of 12% and multistep mutation events of 5%, as suggested by Piry et al., (1999). The Wilcoxon test was used as it has been proven to be statistically powerful for a low number (i.e. less than 20) of microsatellite markers (Piry et al., 1999).

3.3 Results

3.3.1 Datasets

The DMS1 dataset comprised a total of 542 *S. mansoni* parasites that were successfully amplified for all nine microsatellite loci, from which 152 from the village Assoni in Southeast Senegal and 388 from several villages in Northwest Senegal (Table 3.1). Sample sizes for Northwest Senegal ranged between 5 genotypes in Ndombo in 2006 and 98 genotypes in Rhonne in 2007 (Table 3.1).

The DMS2 dataset comprised a total of 758 *S. mansoni* parasites that were successfully genotyped for all six microsatellite loci. A total of 73 out of 104 genotypes were successfully scored for the Kokry-Bozo sample from Southwest Mali that were obtained from Imperial College in London. Sample sizes for Northwest Senegal ranged between seven in Diadiam in 2007 and Ndombo in 2006 and 152 in Nder (Table 3.1).

The DSEQ dataset comprised a total of 671 *cox*1 sequences of which 124 were generated in this study from samples collected in the villages Richard Toll (1993 and 1994), Ndombo (1997 and 2006), Assoni (2011) and the villages Wayowayanko and Farako in Southwest Mali (1993). After alignment and trimming, sequence fragments of 420bp long were obtained for further analysis.

3.3.2 Phylogeographic analysis

Twenty unique *cox*1 haplotypes were found in Northwest Senegal (Table 3.2), which is about one fifth of the total amount of haplotypes found in Africa (i.e. 105). Nineteen of these haplotypes were also found within a single village only, namely Nder (2007). Haplotype diversity of all parasites found in Northwest Senegal (h = 0.847) was high compared with other regions in Africa, ranging between 0.573 in Niger and 0.927 in Tanzania (Table 3.2). The level of haplotype diversity found in Richard Toll in 1993 at the onset of the epidemic was the highest diversity found in this study. Similarly, nucleotide diversity of all parasites sampled in Northwest Senegal (Π = 0.0081) was high compared with other regions in Africa; only parasites sampled in Zambia, Coastal Kenya and the village Kolda showed higher levels of nucleotide diversity (Table 3.2), probably because these populations comprised divergent haplotypes as revealed by the statistical parsimony network (Figure 3.2). Within Northwest Senegal, high levels of haplotype and nucleotide

diversities were found at the onset of the epidemic (i.e. in 1993) and about 14 years later in 2007. Levels of diversity were lowest in Ndombo in 1997 and 2006 (Table 3.2).

The statistical parsimony network shows that the sequences from Northwest Senegal clustered together with the haplotypes found in Southeast Senegal, Southwest Mali, Niger and Brazil. Parasites from the other regions in Africa were grouped into divergent phylogeographic clades, which were separated from the West-African clade by many unsampled or extinct haplotypes (Figure 3.2). The second network reveals the extent of the diversity found in Senegal, showing many divergent haplotypes (Figure 3.2). Hapotypes found in Senegal did not cluster according to village or year of sampling (Figure 3.2).

Table 3.1 Genetic diversity as estimated per village, per region and per year for *Schistosoma mansoni* samples typed at nine (DMS1 – 542 samples in total) or six (DMS2 – 758 samples in total) microsatellites markers.

					DMS1					DMS2				
Region	**Village**	**Year**	**Sample**	**Study**	**$N_{\mu sat}$**	**Hs**	**Ho**	**AR[#]**	**F_{IS}**	**$N_{\mu sat}$**	**Hs**	**Ho**	**AR[##]**	**F_{IS}**
Mali	Kokry-Bozo	2008	Miracidia	GS	na	na	na	na	na	73	0.45	0.42	3.31	0.06*
S. Senegal	Assoni	2011	Miracidia	TS	154	0.50	0.45	3.55	0.12**	168	0.35	0.32	2.85	0.09**
N. Senegal	pooled	na	na		388	0.54	0.52	3.74	0.04**	517	0.38	0.36	2.90	0.05**
N. Senegal	Richard Toll	1993	Worms[S]	TS	7	0.54	0.51	3.83	0.07	7	0.37	0.38	2.83	-0.03
		1994	Worms[S]	TS	12	0.55	0.56	3.71	0.03	22	0.38	0.42	2.94	-0.11*
	Ndombo	1997	Worms[S]	TS	53	0.49	0.48	3.41	0.03	62	0.35	0.33	2.72	0.05
		2006	Worms[S]	TS	5	0.46	0.49	3.00	-0.06	7	0.33	0.33	2.17	-0.01
	Theuss	2006	Miracidia	TS	7	0.52	0.44	3.63	0.16**	18	0.37	0.36	2.68	0.01
		2007	Miracidia	TS	67	0.54	0.52	3.72	0.04*	68	0.39	0.38	2.94	0.03
	Diadiam	2007	Miracidia	TS	6	0.55	0.52	3.73	0.06	8	0.42	0.44	2.93	-0.05
	Gaya	2007	Miracidia	TS	11	0.57	0.60	3.74	0.06	14	0.42	0.42	2.96	-0.001
	Mbodjene	2007	Miracidia	TS	18	0.54	0.60	3.43	-0.11*	21	0.39	0.44	2.72	-0.13*
	Nder	2007	Miracidia	TS	89	0.54	0.53	3.71	0.01	152	0.38	0.36	2.85	0.06*
	Rhonne	2007	Miracidia	TS	98	0.55	0.50	3.74	0.06**	121	0.38	0.35	2.90	0.07*
	Ndieumeul	2007	Miracidia	TS	15	0.54	0.47	3.77	0.12*	17	0.37	0.33	2.94	0.11*

na: not applicable. TS: obtained in this study. GS: obtained from the study of Gower et al. (2013). Worms[S]: worms obtained from SCAN. $N_{\mu sat}$: number of successfully genotyped parasites. Hs: unbiased expected heterozygosity. Ho: observed heterozygosity. AR: Allelic richness. [#]: minimum of 10 alleles used for rarification. [##]: minimum of 14 alleles used for rarification. Statistical significant F_{IS} values are given with * for the nominal level of 0.05 and with ** for the nominal level of 0.001.

3.3.3 Population genetic analysis

Results from the BLAST analyses indicated that three out of nine microsatellite markers are probably linked or adjacent to a gene (Table 3.3). None of the microsatellite loci were however under positive or balancing selection according to the LOSITAN analyses (results not shown). All microsatellites were therefore retained for further analyses.

Parasite population diversity as estimated by unbiased expected heterozygosity (Hs) and allelic richness (AR) was rather uniform across all villages sampled in Northwest Senegal (for

DMS1 Hs = 0.54 – 0.57 and AR = 4.43 – 3.77; for DMS2 Hs = 0.37 – 0.42; AR = 2.72 – 2.96), with the exception of parasites that were sampled in Ndombo in 2006 that showed lower values of diversity (Table 3.1). The levels of *S. mansoni* diversity in Northwest Senegal were higher than the diversity in the village Assoni in Southeast Senegal (for DMS1 Hs = 0.50 and AR = 3.96; for DMS2 Hs = 0.35 and AR = 2.85; Table 3.1), but lower compared to the village Kokry-Bozo in Southwest Mali (for DMS2 Hs = 0.45 and AR = 3.31; Table 3.1).

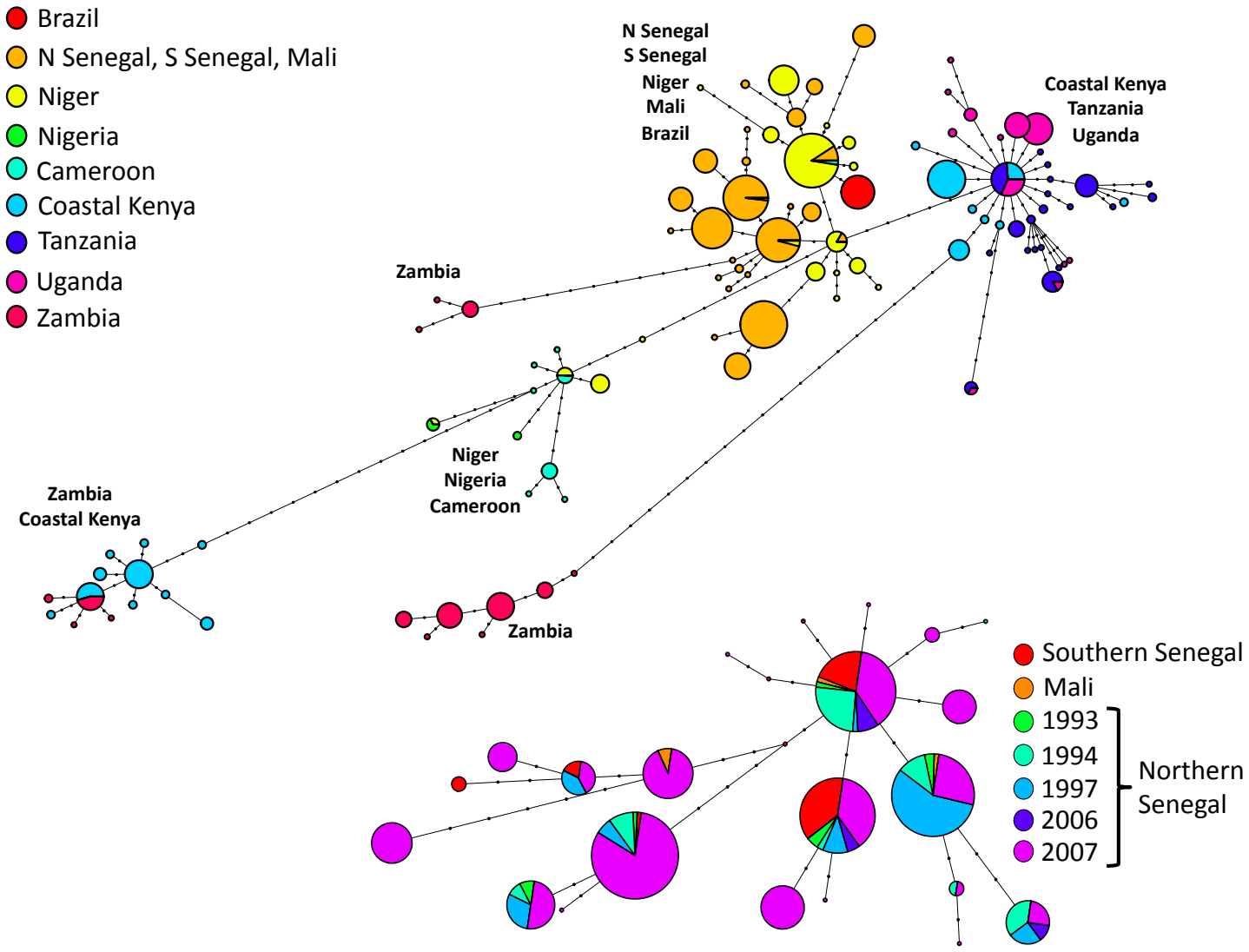


Figure 3.2 Haplotype networks based on statistical parsimony (Templeton et al., 1992) of partial cytochrome oxidase subunit 1 sequences. The network above comprises all sequences obtained from 10 different African countries and Brazil. The network below comprises sequences obtained from different villages in Northwest Senegal (1993-2007), Southeast Senegal (2011) and Southwest Mali (2007). Each pie diagram represents a haplotype (i.e. unique sequence) and dots represent haplotypes that were either not sampled or went extinct. The sizes of the pie diagrams are in relation to the log transformed number of sequences that represent the respective haplotypes, and the colors indicate the location or year of sampling. Sequences were obtained during this study or the previous study of Webster et al. (2013b).

Factorial Correspondence analyses (FCA) revealed that most of the parasites from Northwest Senegal sampled in different villages from 1993 – 2007 always clustered together and differed strongly from parasites sampled in Assoni in 2011 (Southeast Senegal) and Kokry-Bozo in 2007 (Southwest Mali) (Figure 3.3). For DMS1, samples taken in Mbodjene (2007), Ndombo (1997 and 2006) and Gaya (2007) differed slightly from the main cluster, although the second and third axis only explained 9.84% and 6.75% resp. of the total observed variation (Figure 3.3). For DMS2, the only divergent population was the sample taken in Nder in 2007 (Figure 3.3). The distribution of all individual parasites in the FCA plot showed a slight overlap between all three regions (graphs not shown).

Table 3.2 Genetic diversity at a partial cytochrome oxidase subunit 1 fragment in *Schistosoma mansoni* populations as estimated per village, per region and per year.

Region	Village	Study	N_{seq}	N_{hap}	N_{pol}	*h* (SD)	*Π* (SD)
Northwest Senegal			241	20	23	0.847 (0.012)	0.0081 (0.0001)
	Richard Toll 1993	TS	8	6	9	0.929 (0.084)	0.0079 (0.0017)
	Richard Toll 1994	TS	30	7	10	0.772 (0.003)	0.0060 (0.0009)
	Ndombo 1997	TS	46	7	12	0.563 (0.007)	0.0054 (0.0010)
	Ndombo 2006	TS	7	3	4	0.667 (0.160)	0.0032 (0.0014)
	Nder 2007	WS	81	19	22	0.906 (0.014)	0.0087 (0.0005)
	Temey 2007	WS	69	10	14	0.679 (0.059)	0.0078 (0.0007)
Southeast Senegal			31	8	12	0.705 (0.060)	0.0045 (0.0012)
	Assoni 2011	TS	27	6	7	0.638 (0.068)	0.0025 (0.0007)
	Kolda 2009	WS	4	3	9	0.833 (0.222)	0.0127 (0.0034)
Mali		TS	3	3	4	1.000 (0.074)	0.0064 (0.0024)
Cameroon		WS	11	7	9	0.873 (0.089)	0.0074 (0.0010)
Coastal Kenya		WS	85	18	32	0.860 (0.029)	0.0234 (0.0008)
Niger		WS	164	20	32	0.573 (0.046)	0.0071 (0.0011)
Uganda		WS	43	12	23	0.806 (0.043)	0.0056 (0.0009)
Tanzania		WS	44	20	24	0.927 (0.021)	0.0073 (0.0009)
Zambia		WS	46	14	44	0.884 (0.025)	0.0321 (0.0043)
Brazil		WS	16	1	0	0	0

TS: obtained in this study. WS: obtained from the study of Webster et al. (2013b). N_{seq}: number of sequences. N_{hap}: number of unique haplotypes. N_{pol}: number of polymorphic sites. *h*: haplotype diversity. *Π*: nucleotide diversity. SD: standard deviation.

Multilocus estimates of genetic differentiation among regions were F_{ST} = 0.043 and R_{ST} = 0.216 for DMS1 and F_{ST} = 0.036 and R_{ST} = 0.037 for DMS2 (Table 3.3). Multi- and single locus estimates were highly significant (i.e. $p < 0.001$), except for locus *SMD89* for DMS1 and DMS2 and locus *SMD28* for DMS1 (Table 3.3). Estimates of R_{ST} were much higher than estimates of F_{ST} for loci *SMD25* and *SMD11*, and were significant for the random permutation test of allele sizes among allelic states, indicating a mutational component to genetic differentiation (Table 3.3).

Pairwise estimates of F_{ST} and R_{ST} between regions for DMS2 were highest between Southwest Mali and Northwest Senegal (F_{ST} = 0.064; R_{ST} = 0.112) and Southwest Mali and Southeast Senegal (F_{ST} = 0.056; R_{ST} = 0.075). The lowest estimates were found between Northwest Senegal and Southeast Senegal (F_{ST} = 0.044; R_{ST} = 0.032). All pairwise estimates were highly significant (i.e. $p < 0.001$) following permutations of genotypes among regions. Significant allele permutation tests were found for pairwise R_{ST} between Southwest Mali and Northwest Senegal for DMS2 (pR_{ST} = 0.042) and between Northwest Senegal and Southeast Senegal for DMS1 (F_{ST} = 0.044; R_{ST} = 0.133; pR_{ST} = 0.013).

Table 3.3 Genetic structure per microsatellite locus. Estimates of F_{ST} and R_{ST} were obtained at the regional level, i.e. Northwest Senegal and Southeast Senegal for DMS1 and Northwest Senegal, Southeast Senegal and Southwest Mali for DMS2.

			DMS1			DMS2		
Locus	**AC**	**BLAST result**	F_{ST}	R_{ST}	pR_{ST}	F_{ST}	R_{ST}	pR_{ST}
L46951	L46951	c-GMP dependent protein kinase	0.060**	-0.001	0.133	na	na	na
CA11-1	AI068335	B-cell receptor-associated protein	0.072**	0.135**	0.215	0.036**	0.055**	0.303
S9-1	AF330106	/	0.027**	0.006	0.417	0.021**	0.053**	0.290
SMD11	AF325698	/	0.032**	0.158**	0.029	na	na	na
SMD25	AF202965	/	0.023**	0.147**	0.000	0.034**	0.119**	0.031
SMD28	AF202966	STATc protein putative mRNA	0.001	0.002	0.499	0.062**	0.065**	0.518
SMD43	AF325697	/	0.035**	0.012*	0.878	na	na	na
SMD89	AF202968	/	0.002	0.002*	0.917	-0.0028	-0.001	0.553
SMDA28	AF325695	/	0.060**	0.003	0.369	0.041**	0.015**	0.183
Overall	na		0.044**	0.133**	0.025	0.036**	0.037**	0.956
Jackknifed estimators (over loci)								
Mean	na	na	0.043	0.216	na	0.036	0.019	na
SD	na	na	0.007	0.107	na	0.003	0.034	na

na = not applicable. AC = Accession number. SD = standard deviation. pR_{ST} = *p*-value obtained after permutation of allele sizes among allelic states. * significant for permutation of genotypes among regions at the nominal level of 0.05. ** significant for permutation of genotypes among regions at the nominal level of 0.001.

Table 3.4 summarizes the pairwise F_{ST} and R_{ST} between villages for DMS1 (results for DMS2 were not shown but were similar). The table shows that villages Assoni (Southeast Senegal), Ndombo 1997 and Mbodjene 2007 (both Northwest Senegal) were always significantly differentiated from the other samples when genotypes were permuted among villages. Some of these comparisons were also significant following the allele permutation test, suggesting contribution of mutation rather than genetic drift to divergence between these samples (Table 3.4). In contrast, genetic structure among the other samples in Northwest Senegal was low and often insignificant (Table 3.4). Classical multidimensional scaling

(CMDS) plots based on pairwise F_{ST} between villages confirmed this pattern (Figure 3.3).

STRUCTURE analysis revealed two genetic clusters for DMS1 as indicated by ΔK and maximum three genetic clusters as revealed by LnP(D) (Figure 3.4). For K = 2, parasites from all villages sampled in Northwest Senegal were assigned to one genetic cluster, while parasites from the village Assoni in Southeast Senegal were assigned to a second genetic cluster. For K = 3, some parasites sampled in Ndombo in 1997 were assigned to the third genetic cluster (Figure 3.4). Similar results were obtained for DMS2 with two exceptions, namely that parasites from Southwest Mali were assigned to the third genetic cluster and that the Ndombo sample from 1997 was not assigned to a separate genetic cluster as found in the DMS1 dataset (Figure 3.4).

Table 3.4 Pairwise F_{ST} (above diagonal) and R_{ST} (below diagonal) between samples from one village in Southeast Senegal (Assoni '11) and eight villages in Northwest Senegal (i.e. DMS1).

		1	2	3	4	5	6	7	8	9
1	**Assoni '11**		0.038**	0.073**	0.045**	0.068**	0.046**	0.028**	0.039**	0.049**
2	**Richard Toll '94**	0.035		0.026*	0.009	0.031*	0.014*	-0.007	0.003	0.004
3	**Ndombo '97**	0.049*	-0.013		0.038*	0.047**	0.020**	0.016*	0.021**	0.021**
4	**Gaya '07**	0.056*	-0.012	-0.022		0.030*	0.020*	0.009	0.012	0.007
5	**Mbodjene '07**	0.295**#	0.141*	0.192**#	0.176*#		0.015*	0.029*	0.013*	0.013*
6	**Nder '07**	0.167**#	0.023	0.059**	0.041	0.033*		0.001	0.002	0.003
7	**Ndieumeul '07**	0.087*	-0.027	-0.001	-0.013	0.092*	-0.004		-0.0002	0.001
8	**Rhonne '07**	0.119**#	-0.007	0.031*	0.019	0.078*	0.003	-0.017		0.001
9	**Theuss '07**	0.164**#	0.021	0.062*	0.046	0.046*	-0.005	-0.007	0.0003	

= significant for permutation of allele sizes among allelic states at the nominal level of 0.05. * = significant for permutation of genotypes among villages at the nominal level of 0.05. ** = significant for permutation of genotypes among villages at the nominal level of 0.001 (i.e. Bonferroni corrected).

Signatures of population expansion (heterozygosity deficiency) were detected under the TPM and SMM for all three regions (Table 3.5). Most of the villages sampled in 2007 and the sample from the village Ndombo in 1997 showed a heterozygosity deficiency for both DMS1 and DMS2 (Table 3.5). Parasites sampled in 1993 and 1994 in Richard Toll, as well as parasites from Mbodjene 2007, Gaya 2007 and Diadiam 2007 did not show deviations from mutation-drift equilibrium (Table 3.5).

Table 3.5 Tests of mutation-drift equilibrium in *Schistosoma mansoni* populations following the two-phase mutation model (TPM) and the stepwise mutation model (SMM) as implemented in the BOTTLENECK software (Cornuet and Luikart, 1996). *P*-values of the two tails Wilcoxon tests are always given (P2), while *p*-values of the one tail Wilcoxon test (P1) were only given when the two tails test was significant or borderline significant.

			DMS1				DMS2			
			TPM		SMM		TPM		SMM	
Region	Village	Year	P2	P1	P2	P1	P2	P1	P2	P1
Mali	Kokry-Bozo	2008	na	na	na	na	0.031	0.016 (D)	0.016	0.008 (D)
S. Senegal	Assoni	2011	0.004	0.002 (D)	0.002	0.001 (D)	0.031	0.016 (D)	0.016	0.008 (D)
N. Senegal	na	na	0.027	0.014 (D)	0.006	0.003 (D)	0.016	0.008 (D)	0.016	0.008 (D)
N. Senegal	Richard Toll	1993	0.156		0.156		1.000		1.000	
		1994	0.910		0.652		0.109		0.109	
	Ndombo	1997	0.020	0.010 (D)	0.014	0.007 (D)	0.016	0.008 (D)	0.016	0.008 (D)
		2006	0.578		0.469		0.875		1.000	
	Teuss	2006	0.687		0.813		0.312	0.156 (D)	0.016	0.008 (D)
		2007	0.359		0.027	0.014 (D)	0.031	0.016 (D)	0.016	0.008 (D)
	Diadiam	2007	0.813		0.813		1.000		1.000	
	Rhonne	2007	0.203		0.020	0.010 (D)	0.016	0.008 (D)	0.016	0.008 (D)
	Gaya	2007	0.570		0.496		0.109		0.110	
	Ndieumeul	2007	0.074	0.037 (D)	0.039	0.020 (D)	0.031	0.016 (D)	0.031	0.016 (D)
	Nder	2007	0.164		0.020	0.010 (D)	0.016	0.008 (D)	0.016	0.008 (D)
	Mbodjene	2007	0.375		0.055	0.027 (D)	0.125		0.125	

D: heterozygote deficiency

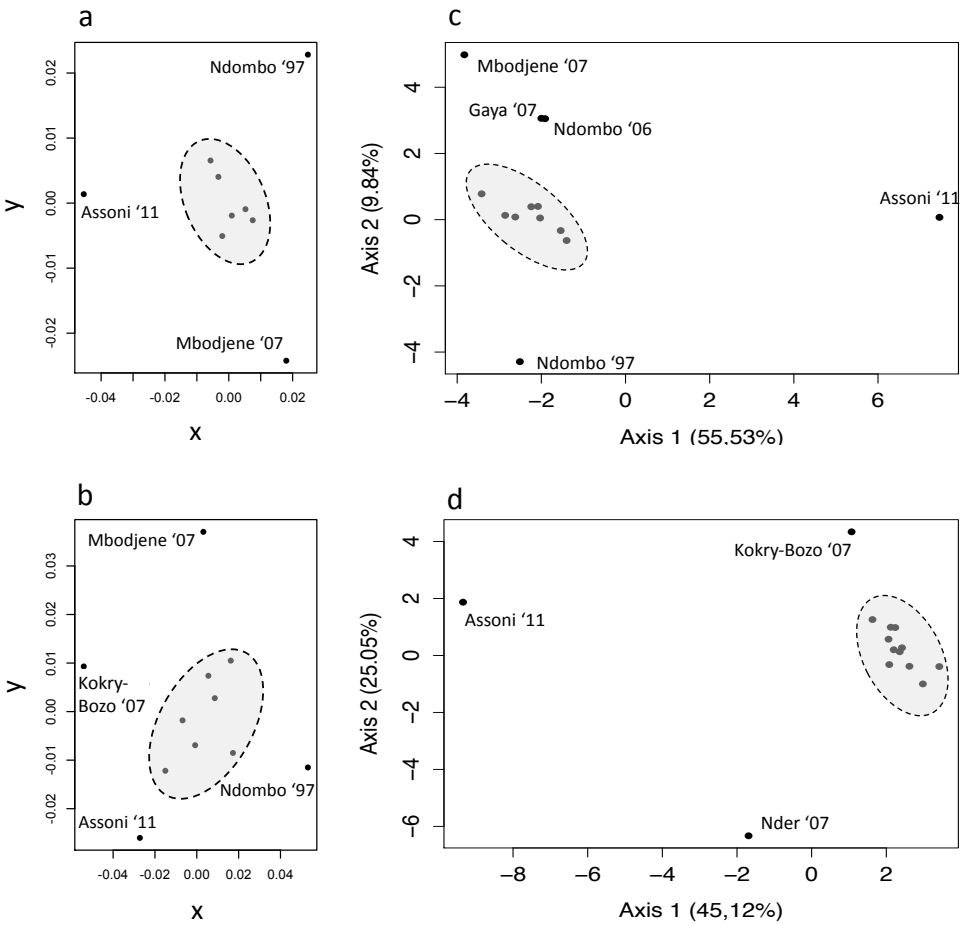


Figure 3.3 Classical multidimensional scaling plots of pairwise F_{ST} based on microsatellites for dataset DMS1 (a) and dataset DMS2 (b). Results of the Factorial Correspondence Analysis based on microsatellites for DMS1 (c) and DMS2 (d). Shaded areas comprise samples from Northwest Senegal that were collected in 1993 – 2007 while samples differentiated from this cluster were labeled.

3.4 Discussion

The construction of two dams in 1988 within the Senegal River Basin led to a massive outbreak of intestinal schistosomiasis, a debilitating disease that was until then absent in the region. The rate at which the epidemic expanded over Northwest Senegal was devastating: within a few years following the construction of the dams, *S. mansoni* parasites were found within most villages in the Delta and part of the Middle Valley (Picquet et al., 1996). Although several mechanisms may "filter" out parasites during an invasion process (Torchin et al., 2002), no signals of founder effects were found (i.e. loss of genetic variation that occurs when a new population is established by a very small number of individuals). Despite the fact that mitochondrial genes due to their haploid state with uniparental inheritance are particularly prone to losing diversity after founder effects (Avise, 2004), levels of nucleotide and haplotype diversities at the *S. mansoni cox*1 marker were similar to the ones found in other African countries (Table 3.2). Genetic diversity at the nuclear level was higher than in Southeast Senegal, and was not substantially lower than the one found in Southwest Mali (Table 3.1). We furthermore found evidence that the *S. mansoni* population in Northwest Senegal experienced an increase in population size rather than a population bottleneck (Table 3.5). A technical problem related to the Bottleneck analyses however, is that this method cannot distinguish between a population bottleneck and (subsequent) rapid population growth, where the latter could wipe out the genetic signal of the former (Bonhomme et al., 2008; Lawler, 2008). However, the diversity of *S. mansoni* strains collected at the beginning of the epidemic in 1993 and 1994 in the epicenter Richard Toll was similar to the diversity found 15 years later in 2007 (Table 3.1) and no signals of bottleneck were found in these samples either (Table 3.5).

Our data furthermore indicated a complex colonization history, most likely with multiple introductions from disparate source populations. First, levels of nucleotide and haplotype diversities in Northwest Senegal were relatively high compared to other African countries, which could be explained from an admixture event between independent introductions (Table 3.2). Second, analyses of genetic structure showed that samples from the villages Ndombo 1997 and Mbodjene 2007 were significantly differentiated from the other samples in Northwest Senegal (Figures 3.3 & 3.4). Stepwise mutation, rather than genetic drift, explained these levels of differentiation (Table 3.4). This result suggests that these samples

are probably anciently differentiated from the other samples in Northwest Senegal and that they possibly represent signatures of different independent introductions. Third, multidimensional scaling of pairwise estimates of F_{ST} and R_{ST} revealed that most of the samples from Northwest Senegal occupied a central position between the divergent populations Ndombo 1997 and Mbodjene 2007 (both Northwest Senegal), Assoni (Southeast Senegal) and Kokry-Bozo (Southwest Mali) (Figure 3.3). Such a pattern could be expected if most of the invaded populations of Northwest Senegal represent an admixture of different source populations (in this case Southeast Senegal, Southwest Mali and the source populations of Ndombo 1997 and Mbodjene 2007). Multiple introductions could be expected due to the substantial immigration from agricultural workers from many different regions, such as Mauritania, Mali and southern Senegal (Handschumacher et al., 1992). An interesting finding in this respect is that *S. mansoni* parasites from the Mbodjene sample (2007) in the Lampsar region were significantly differentiated from most of the other samples taken around Richard Toll and the Senegal River. Similarly, the *B. pfeifferi* snail populations in the Lampsar region showed high inter- and intra population genetic variation compared to populations in the region around Richard Toll that were largely homogenous (Campbell et al., 2010). The authors explained these results by a different ecology within both settings, with the Senegal River and Lake Guiers representing stable, permanent water bodies while the Djeuss and the Lampsar River represent transient habitats, characterized by droughts and annual flooding. Increased intrapopulation similarity in permanent water bodies such as in the Senegal River and Lake Guiers has been reported previously for *B. pfeifferi* in other settings (Hoffman et al., 1998; Webster et al., 2001), while habitat instability such as in the Lampsar region can cause extreme fluctuations in population size and may restrain gene flow. The fact that both *B. pfeifferi* and *S. mansoni* populations in the Lampsar region show different patterns than the other samples imply that the distribution of intermediate snail hosts could be an important factor determining the invasion success of human *S. mansoni* parasites. Further research including additional samples from *S. mansoni* from the Lampsar region and including data on the distribution and genetics of snail populations is warranted to explore these hypotheses.

The invaded parasites in Northwest Senegal do have, not surprisingly, a West-African origin. All *cox*1 haplotypes sampled in Northwest Senegal belonged to the same phylogeographic clade as those sampled in Southeast Senegal, Southwest Mali and some of Niger (Figure

3.2), confirming results of previous phylogeographic analysis (Webster et al., 2013b). A higher resolution was obtained from the microsatellite data (Figure 3.4). Structure analysis revealed that parasites sampled in Northwest Senegal, Southeast Senegal and Southwest Mali (1) were significantly differentiated from each other and (2) that stepwise mutations explained the observed levels of genetic differentiation. These results suggest that parasites from Southeast Senegal and Southwest Mali (or at least the villages that were sampled in this study) were most likely not the source for *S. mansoni* parasites in Northwest Senegal. On the other hand, one could argue that parasites from Northwest Senegal diverged from Southeast Senegal or Southwest Mali since their introduction at the onset of the epidemic. In this case however, we would expect temporal differentiation between samples taken in 1993 and samples taken 14 years later in 2007, a pattern that was not observed (Figures 3.3, 3.4 and Table 3.4). We therefore consider the timeframe since invasion to be too short to explain these levels of genetic differentiation and assume that parasites from Northwest Senegal most likely originated from another source population than those sampled in this study.

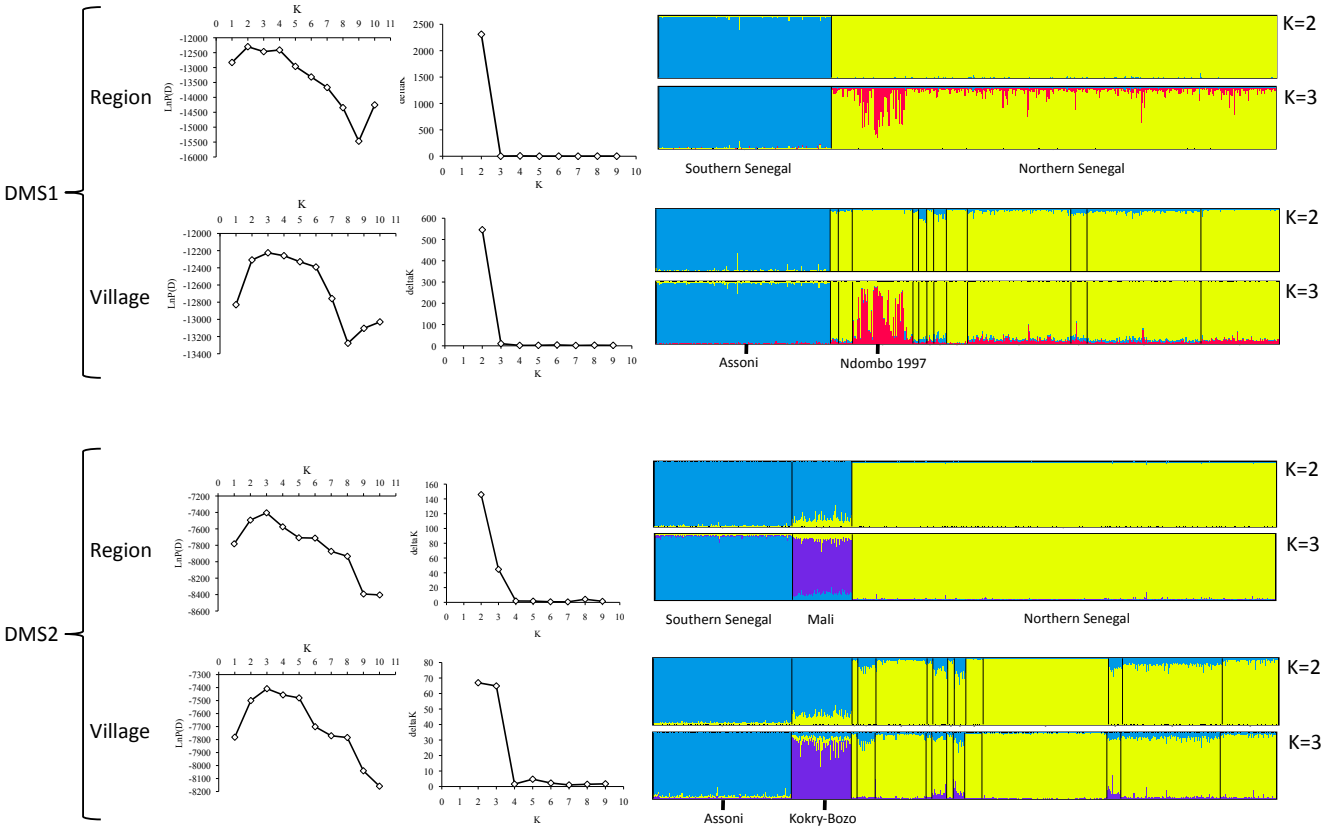


Figure 3.4 Results of STRUCTURE analysis for DMS1 and DMS2. Graphs on the left show the loglikelihood LnP(D) and its second order rate change *ΔK* for a given number of clusters *K*. Barplots on the right show the membership probabilities of each individual parasite to a given cluster *K*.

3.5 Conclusions

This study revealed some new insights into the epidemic of human intestinal schistosomiasis in Northwest Senegal. Our results revealed a complex colonization history and suggest that multiple introductions of *S. mansoni* parasites from disparate source populations could have occurred. We furthermore found no signals of founder effects as levels of genetic diversity at the onset of the epidemic were similar to those found in neighboring regions. Our results on *S. mansoni* genetic structure showed some similarities with those found during a previous study on the indermediate snail host *B. pfeifferi*, suggesting an important role of the latter in the invasion success of the former. Further research incorporating more genetic markers as well as the demography and spatial structure of *B. pfeifferi* is needed to reveal the factors that boosted *S. mansoni* epidemic in Northwest Senegal.

CHAPTER 4

Inbreeding within human *Schistosoma mansoni*: do host-specific factors shape the genetic composition of parasite infrapopulations?

F Van den Broeck, L Meurs, JAM Raeymaekers, N Boon, T Dieye, FAM Volckaert, K Polman & T Huyse

Published in Heredity 2014, early publication, art. nr. 10.1038/hdy.2014.13

Abstract

The size, structure and distribution of host populations are key determinants of the genetic composition of parasite populations. Despite the evolutionary and epidemiological merits, there has been little consideration of how host heterogeneities affect the evolutionary trajectories of parasite populations. We assessed the genetic composition of natural populations of the parasite *Schistosoma mansoni* in Northwest Senegal. A total of 1346 parasites were collected from 14 snail and 57 human hosts within three villages and individually genotyped using nine microsatellite markers. Human host demographic parameters (age, gender and village of residence) and co-infection with *Schistosoma haematobium* were documented and *S. mansoni* infection intensities were quantified. *F*-statistics and clustering analyses revealed a random distribution (panmixia) of parasite genetic variation among villages and hosts, confirming the concept of human hosts as 'genetic mixing bowls' for schistosomes. Host gender and village of residence did not show any association with parasite genetics. Host age however, was significantly correlated with parasite inbreeding and heterozygosity, with children being more infected by related parasites than adults. The patterns may be explained by 1) genotype-dependent 'concomitant immunity' that leads to selective recruitment of genetically unrelated worms with host age, and/or 2) the 'genetic mixing bowl' hypothesis, where older hosts have been exposed to a wider variety of parasite strains than children. This study suggests that host-specific factors may shape the genetic composition of schistosome populations, revealing important insights into host-parasite interactions within a natural system.

4.1 Introduction

As parasites and their hosts are closely associated, Price (1980) predicted that parasite infrapopulations (i.e. all parasites of the same species within an individual host; Bush et al., 1997) would represent isolated patches, resulting in low genetic diversity and high levels of inbreeding. This seemed particularly true for recurrent generations of parasites infecting a single host (e.g. phytophagous insects; Price, 1980). However, the majority of animal macroparasites (e.g. flatworms, nematodes) release their offspring into the environment or multiply within one or more intermediate host species, mostly resulting in high genetic diversity levels similar to those reported in free-living organisms (Bush et al., 2001). It has been suggested that the high genetic diversity of macroparasites within individual hosts may reflect the tendency of these hosts to sample a number of transmission sites (Anderson et al., 1995; Nadler, 1995), thereby becoming 'genetic mixing bowls' for parasite genes (Curtis and Minchella, 2000; Curtis *et al.*, 2002). The result is that parasite offspring are well mixed in the environment and that parasite genetic variation is randomly distributed between hosts (Criscione et al., 2005). As emphasized by Nadler (1995), many factors contribute to patterns of genetic diversity in parasites, such as the mating system (selfing, clonal, outcrossing), the number of intermediate host species, and whether parasites transmit on land, in the water or both (e.g. Criscione & Blouin, 2004). Studying the factors that affect the distribution of parasite genotypes among infrapopulations represents a powerful tool for understanding transmission dynamics. The joint analysis of the population genetic structure of the parasite with characteristics of the host such as age, gender and mobility can generate important insights in parasite evolution (factors that shape local adaptation and speciation) as well as epidemiology (determining whether transmission is focal or not; Criscione, 2008).

The digenean trematode *Schistosoma mansoni* has a complex two-host life cycle with an asexual amplification stage in the snail intermediate host, yielding thousands of clonal cercariae that infect the human final host. The cercariae develop into dioecious worms that reproduce sexually, resulting in offspring (miracidia) upon hatching of worm eggs that are released into the external environment through feces. Population genetic studies have shown that the geographic scale of differentiation of *S. mansoni* varies substantially between study areas. For example, in Melquiades (Brazil), a village with a complex

hydrosystem of ridges and valleys, high differentiation was found between *S. mansoni* populations from different households along different water bodies (Thiele et al., 2008). Low differentiation was found in Virgem das Gracas (Brazil) where households were situated along the same water body (Thiele et al., 2008). The authors argued that the complex geography of Melquiades probably restricted host movement for water usage, resulting in focal points of transmission and thus high genetic differentiation. These data suggest that the gene flow of *S. mansoni* may be determined by human host mobility and that gene flow may occur across large geographical distances in the absence of boundaries between transmission sites (Steinauer et al., 2010). Similar conclusions were drawn from studies in Kenya (Agola et al., 2009) and for non-human foci (*Rattus rattus*) of *S. mansoni* (Sire et al. 2001; Prugnolle et al., 2005c).

Apart from host movement for water usage (influenced by geographical entities), other host-specific factors could, directly or indirectly, affect the distribution of *S. mansoni* strains within and among hosts. For instance, the genetic diversity of *S. mansoni* infrapopulation within male *Rattus rattus* hosts was higher than within female hosts (Caillaud et al., 2006). The authors explained this by differences in immunocompetence or water contact behavior between male and female rat hosts. Similar explanations (resistance and/or exposure to infection) were put forward to explain the age and gender-dependent infection intensities observed within human schistosomiasis endemic communities (Gryseels 1994; Kabatereine et al. 1999). Nevertheless, studies on the impact of host-specific factors on the genetic structure of schistosome populations are scarce (Thiele et al., 2008; Gower et al., 2011). This is mainly due to the methodological limitations of collecting and genotyping larval field stages. So far, only one study on human *S. mansoni* infections incorporated host-specific factors (gender and age) in their parasite genetic analyses (Thiele et al., 2008). No relationship between parasite population structure and host-specific factors was detected (Thiele et al., 2008), possibly because *S. mansoni* genotypes were inferred after laboratory passage in mice. This strategy most likely induced a bias due to genetic bottlenecking or differential immune responses of laboratory animals.

In this study, we investigated host-specific factors that may shape the genetic composition of natural *S. mansoni* parasite populations in Northwest Senegal. In total, 1346 parasites were collected from 14 snail and 57 human hosts from three villages. By genotyping larval

schistosomes directly isolated from humans, the need for laboratory passage in mice was avoided. This allows an improved analysis of the interaction between human host-specific factors and schistosome population genetics.

4.2 Material and methods

4.2.1 Ethical statement

This study is part of a larger investigation of the epidemiology, transmission and control of schistosomiasis in Senegal (SCHISTOINIR: www.york.ac.uk/res/schistoinir). Ethical approval was obtained from 'Le Comité National d'Ethique de la Recherche en Santé' in Senegal, the review board of the Institute of Tropical Medicine in Belgium, and the ethical committee of the Antwerp University Hospital in Belgium. Informed and written consent was obtained from all participants before the start of the study. All inhabitants were treated after the study according to WHO guidelines (WHO, 2006) with a single dose of praziquantel (40 mg/kg of body weight) and mebendazole (500 mg) for schistosomiasis and soil-transmitted helminthiasis respectively.

4.2.2 Description of the study sites and epidemiology

Parasites and host demographic data were obtained from the villages Pakh (16°24'12"N 15°48'42"W; 790 inhabitants), Diokhor Tack (16°11'24"N 15°52'48"W; 984 inhabitants) and Ndieumeul (also known as Thiekène; 16°13'12"N 15°51'36"W; 384 inhabitants), all located on the western side of Lake Guiers (Northwest Senegal). Diokhor Tack and Ndieumeul lie in close proximity (4 km) on the Peninsula Nouk Pomo, situated about 22 km from Pakh that lies next to a large irrigation canal close to the lake (Figure 4.1). Almost all inhabitants (99%) in Ndieumeul and Diokhor belong to the ethnic group Wolof, while in Pakh they are Wolof (63%) and Peul (30%). To our knowledge, there have been no anthelminthic treatment programs in these villages prior to the study.

The prevalences and intensities of *S. mansoni* infection increased up to the second decade of life, with a subsequent decrease in adults (Meurs et al., 2012). *Schistosoma mansoni* prevalences were 16% in Pakh, 75% in Ndieumeul and 55% in Diokhor. *Schistosoma haematobium* was co-endemic in all the villages (57% in Pakh, 66% in Ndieumeul and 44% in

Diokhor) with mixed infection prevalences of 12% in Pakh, 55% in Ndieumeul and 32% in Diokhor (Meurs et al. 2012; unpublished data).

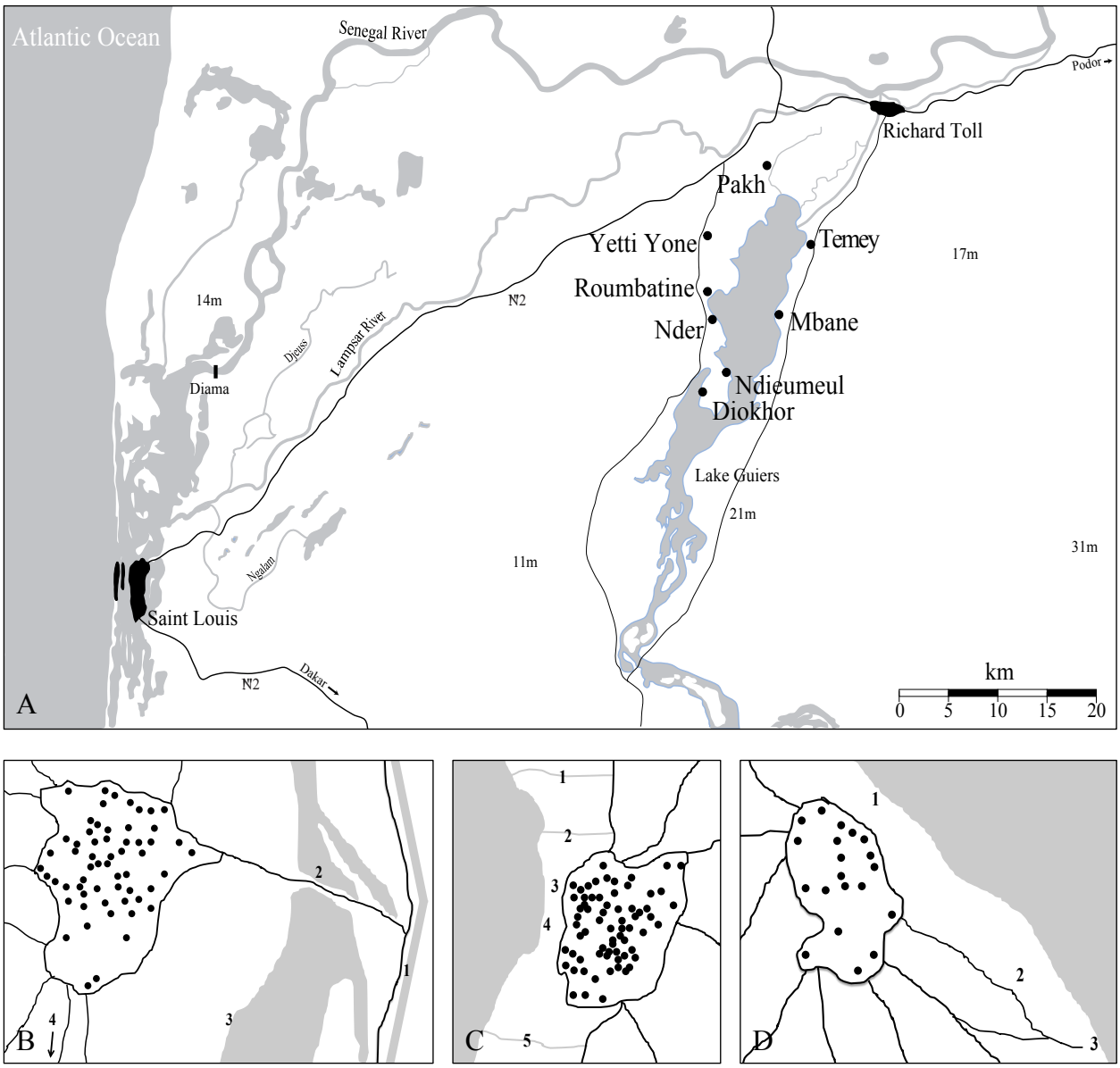


Figure 4.1. A) Map of Northwest Senegal showing the three major water bodies (Lake Guiers, Senegal River and its tributaries Lampsar River and Djeuss), the national road N2, cities Saint-Louis and Richard Toll and the location of the study villages (Pakh, Yetti Yone, Roumbatine, Nder, Diokhor Tack, Ndieumeul, Mbane and Temey). While snails were collected in all villages, human samples were only collected in Pakh, Diokhor and Ndieumeul. The elevation of some geographic points are indicated in meters above sea level. Detailed maps of Pakh (B), Diokhor (C) and Ndieumeul (D): dots are households, black lines are 'roads', grey areas are water and numbers indicate the known transmission sites within each village.

4.2.3 Data collection from humans and snails

In total, three surveys were conducted within less than a year during February 2009 (Pakh), August 2009 (Pakh, Diokhor Tack and Ndieumeul) and January 2010 (Ndieumeul). Stool samples were collected from all participants and processed by the Kato-Katz technique as described elsewhere (Meurs et al., 2012). *Schistosoma mansoni* infection intensity was expressed as the number of eggs detected per gram of faeces (EPG). From each participant, data were collected on age, gender and village of residence. *Schistosoma mansoni* eggs were filtered from positive stool samples by homogenizing each sample with water and then passing it through a metal sieve of 212 μm pore size to remove any larger debris. Eggs were then concentrated in Petri dishes by passing the remaining aqueous solution through a homemade Pitchford and Visser funnel. Eggs were either pooled per inhabitant in 1.5 ml tubes filled with EtOH (70%), or hatched for miracidia that were individually pipetted onto Whatman FTA® indicator cards in a volume of 3 μl of water.

Snails of the species *Biomphalaria pfeifferi* were collected within a two-year period during August 2008 (Mbane, 16°16'15"N 15°48'07"W; Temey, 16°19’45"N 15°46’04"W; Nder, 16°16'00"N 15°52'28"W; Yetti Yone, 16°20'56''N 15°53'4''W; Roumbatine, 16°17'19''N 15°52'58''W and Pakh), February 2009 (Pakh), August 2009 (Diokhor, Ndieumeul and Pakh) and January 2010 (Ndieumeul and Pakh; Figure 4.1). Not all transmission sites could be sampled at all times as some dried out or were inaccessible due to vegetation or inundation (see Supplementary Table 4.1 for details on which sites were sampled). Each transmission site was thoroughly sampled by two researchers for at least 15 minutes and longer when many snails were found. Snails were stimulated to shed cercariae by transferring them in containers with bottled water and exposing them for 5-10 min to direct sunlight. Cercariae were individually fixed in 96-well plates containing 40 μl EtOH (70%). All samples were stored and transported at room temperature.

4.2.4 Molecular analysis

Genomic DNA extraction and genotyping was performed as described before (Van den Broeck et al., 2011). In short, individual *S. mansoni* parasites were genotyped using nine putatively neutral microsatellite markers (*L46951*, *CA11-1*, *S9-1*, *SMD11*, *SMD25*, *SMD28*, *SMD43*, *SMD89*, *SMDA28*; Durand et al., 2000; Blair et al., 2001; Curtis et al., 2001). PCR

products were analyzed using an ABI 3130 Genetic Analyser (Applied Biosystems) and GeneScan™ 500 LIZ™ as size standard. Allele sizes were manually verified using GENEMAPPER v4.0 (Applied Biosystems). Miracidia that were successfully genotyped for less than seven of the nine loci were excluded from the analyses. As more hosts, rather than more miracidia per host, leads to a more robust estimate of parasite population genetic diversity (French et al., 2012), all hosts were included for data-analysis except those that were genotyped for only one parasite. For each infected snail host, eight cercariae were genotyped, unique multi-locus genotypes (MLGs) were identified and used for data-analysis while the identical MLGs were removed.

4.2.4 Genetic diversity

For each village and for each host, parasite observed (Ho) and Nei's unbiased expected heterozygosity (Hs) were calculated using GENETIX v4.05 (Belkhir et al. 1996-2004). The inbreeding coefficient F_{IS}, defined as the probability that two alleles at a locus are identical by descent, was estimated in GENETIX at the village and at the host level using *f* (Weir and Cockerham, 1984), tested for significance using 10,000 permutations and corrected for multiple testing using sequential Bonferroni corrections. Parasite allelic richness (AR; which corrects the number of alleles per locus for unequal sample sizes) was estimated based on a minimum of 2 alleles per host and per village using the R package HIERFSTAT v0.04-6 (Goudet, 2005). The same analyses were also performed for the cercarial population of Ndieumeul, which was the only sample with a sufficient amount of infected snails (see results).

4.2.5 Temporal and spatial genetic structure

To assess the relative importance of the temporal to the spatial genetic variation, differences in allele frequencies were sought between sampling times within one village. To this end, the function *varcomp.glob* as implemented in the R package HIERFSTAT was used to test whether alleles were correlated within one sampling time relative to the total observed variation within each village independently (for Ndieumeul and Pakh, as Diokhor Tack was surveyed only once). This nested design was suggested as a solution for crossed factors (i.e. different villages were sampled at the same sampling time; de Meeus and

Goudet, 2007). Significance was tested by 10,000 permutations of host samples among sampling times in each village independently (function *test.between*).

Parasite gene flow within and between hosts and villages was assessed using a hierarchical analysis of genetic structure that quantified the subdivision of parasites ("Par") at two levels: within hosts ("Pop") and within villages ("Vil"). The terminology for the classical *F*-statistics was modified as follows: $F_{Vil/Total}$ reflects the correlation of alleles within villages relative to the total; $F_{Pop/Vil}$ reflects the correlation of alleles within hosts relative to the village and $F_{Par/Pop}$ reflects the correlation of alleles within parasites relative to the host. Each of these hierarchical estimates was computed using the R package HIERFSTAT, both overall (function *varcomp.glob*) as for each locus separately (function *varcomp*). Overall F_{ST} estimates were standardized by dividing the observed F_{ST} values calculated from the data by the maximum F_{ST} value given the data (Hedrick, 2005) using RecodeData v.0.1 (Meirmans, 2006). Significance of all *F*-values was tested with 10,000 permutations. Using FSTAT v2.9.3 (Goudet, 2005), alleles were permuted among parasites within hosts ($F_{Par/Pop}$), while HIERFSTAT was used for testing the significance of $F_{Pop/Vil}$ (parasites were permuted among hosts but kept within their village of origin; function *test.within*) and of $F_{Vil/Total}$ (whole host samples were permuted among villages; function *test.between*). All *p*-values were corrected for multiple testing. Finally, 10,000 bootstraps were performed over loci in HIERFSTAT to obtain confidence intervals for every *F*-statistic (function *boot.vc*).

Population structure was further analyzed with a Bayesian Markov chain Monte Carlo (MCMC) clustering analysis as implemented in STRUCTURE v2.2.3 (Pritchard et al., 2000). This method assigns individuals probabilistically to *K* populations, or jointly to two or more populations if their genotypes indicate they are admixed, in such a way that loci within populations are in Hardy-Weinberg and linkage equilibrium. The number of clusters *K* was derived assuming the admixture model and correlated allele frequencies; three replicate runs were run for each predefined *K* (ranging from 1 to 10). Each run was initiated by 100,000 burn-in steps and consisted of 1,000,000 MCMC steps. The optimal *K* value was identified by the highest lnP(D) value and by estimating ΔK, which is based on the second order rate change of the posterior lnP(D) (Evanno et al., 2005).

4.2.6 Impact of Wahlund effect, inbreeding and biased sampling on variations in FIS

F-statistics revealed variable levels of parasite F_{IS} estimates at the host level, ranging from some negative (heterozygote excess) to many positive deviations (heterozygote deficiency) from Hardy-Weinberg Equilibrium (see results). Excluding the possibility of scoring errors (see Van den Broeck et al. 2011), a more in depth analysis of three competing mechanisms related to the biology of schistosomes were performed following a similar approach as presented by Castric et al. (2002).

First, heterozygote deficiencies could result from a subdivision of the local population into isolated and differentiated reproductive units. When hosts are infected with parasites from spatially or temporally separated gene pools (i.e. cryptic population structure), fewer heterozygotes are expected than under random mating, resulting in a Wahlund effect. Tests for Wahlund effects were performed using STRUCTURE, as described above, which is a method that aims at unraveling cryptic population structure by assigning individuals to *K* populations as such that the populations are in Hardy-Weinberg and linkage equilibrium. When the optimal *K* value was two or more, than this means the presence of population structure, which could possibly lead to a Wahlund effect.

A second cause for heterozygote deficiencies is mating between close relatives, i.e. inbreeding. When paired adult worms are genetically related, the chance that two alleles in the offspring are identical by descent will be higher than expected under random mating. Inbred offspring will be homozygous for most loci and will therefore show lower individual heterozygosity than offspring resulting from random mating. We therefore estimated the proportion of heterozygous loci per parasite (i.e. multi-locus heterozygosity; MLH) and compared the mean observed MLH per host (MLH_{obs}) with the expected under random mating (MLH_{exp}). The expected distribution was obtained from 1000 pseudosamples in which alleles were randomly associated within individuals using GENETIX. Significance was tested by estimating *p*-values as the probability of observing lower MLH than expected under random mating. To test whether deviations in HWE were due to inbreeding, an association was sought between F_{IS} and individual heterozygote deficiency as estimated by $(MLH_{obs} - MLH_{exp})/MLH_{exp}$.

A final mechanism that could explain variations in F_{IS} is the non-random sampling of offspring from a limited number of families. As adult worms reside in the human blood

vessels, genetic analyses could only be performed on offspring genotypes without having an idea about the number of adult worms that contributed to this pool of offspring (Steinauer et al., 2013). As some hosts might harbor less adult worms than others, the probability of sampling two related offspring parasites is higher than expected in a randomly mating population. Such a non-random sampling of offspring from a limited number of families will result in heterozygote excesses or at least a downward biased estimation of F_{IS}. We therefore assessed the genetic relatedness of every pair of offspring within a sample and tested whether the mean observed relatedness was higher than expected. Pairwise relatedness between each individual parasite was estimated according to the identity coefficient as it has a smaller variance than other estimates of relatedness (Belkhir et al., 2002). The expected distributions were generated by 1000 permutations of genotypes in IDENTIX v1.1 (Belkhir et al., 2002) and significance was tested by estimating *p*-values as the probability of observing higher relatedness than expected.

4.2.7 Link between human host-specific factors and S. mansoni parasite genetics

Associations between parasite population genetic statistics (Ho, Hs, AR and F_{IS}) and human host demographic variables (age, gender and village of residence), infection intensity (EPG) and co-infection with *S. haematobium* were analyzed using Statistica v11 (StatSoft, Tulsa, OK, USA). A General Linear Model (GLM) was constructed for each parasite population statistic in order to test their dependency on individual host demographic parameters, infection intensity and co-infection with *S. haematobium*. Host age and infection intensity were included in the model as continuous covariates, and host gender (male, female), co-infection (absent, present) and village of residence (Pakh, Diokhor Tack and Ndieumeul) as categorical variables. All interactions were tested, but excluded from the model when not significant.

Two additional tests were finally performed to investigate the robustness of the statistical outcome. First, to test the effect of low sample sizes, hosts that were genotyped for less than 13 parasites were removed, resulting in a dataset comprising only 45 human hosts. The same GLM models as described above were then repeated. Second, the effect of individual loci was assessed following a jack-knife approach: each locus was individually removed, parasite population statistics were re-estimated using the remainder of the loci and univariate statistics were then performed for each jack-knifed estimate.

4.3 Results

4.3.1 Dataset

A total of 2891 *S. mansoni* parasites from 63 hosts were subjected to genotyping, which is on average 45.9 parasites per host (min = 2, max = 157, median = 46). Of these, 1345 genotypes (i.e. 46.5%) were successfully scored for at least seven out of nine loci and were used for further analyses. The failure rate depended heavily on the type of sample: eggs were genotyped with an average success of 44.2%, while miracidia showed almost double the success rate (i.e. 81.2%). The final dataset consisted of 57 human hosts, from which 35 female and 22 male hosts; 10 hosts originated from the village Diokhor Tack, 11 from Pakh and 36 from Ndieumeul (Table 4.1). Host age ranged between 4 and 50 and differed significantly between the three villages (one-way ANOVA; F = 8.02; p = 0.0009), in particular because of the young age of the Ndieumeul sample (mean age was 23 in Diokhor Tack, 29 in Pakh and 14 in Ndieumeul). Each host harbored on average 24 parasites (min = 2, max = 74; Table 4.1). All parasites presented unique multilocus genotypes (MLGs).

In total, 283 *B. pfeifferi* snails were collected in eight villages (Supplementary Table 4.1). Infected snails were only found in three villages: in Temey and Nder during August 2008 and in Ndieumeul during January 2010. Snail *S. mansoni* prevalences were 3.5% (1/29) in Nder, 9.62% (13/135) in Ndieumeul and 14.3% in Temey (1/7). Microsatellite genotypes were successfully obtained from all snails, except for one snail from Ndieumeul. Two snails from Ndieumeul were each infected with two different *S. mansoni* strains (i.e. two different MLGs) while all the other snails were infected with one parasite strain only, resulting in a final cercarial population of 14 MLGs. In Nder and Temey, only one MLG per infected snail was found. No snail harbored near-identical genotypes (i.e. genotypes that differed at one or a very few loci from a frequently observed MLG).

4.3.2 Genetic diversity

All measures of genetic diversity were relatively uniform among the different sampling sites, although *S. mansoni* populations from Pakh (AR = 1.55; Hs = 0.55) were slightly more diverse than those from Diokhor Tack (AR = 1.54; Hs = 0.54) and Ndieumeul (AR = 1.53; Hs = 0.53; Table 4.1). The genetic diversity of the cercarial population in Ndieumeul (AR = 1.57; Hs =

0.49) was similar to the parasite genetic diversity found within each village and each host (Table 4.1).

4.3.3 Temporal and spatial genetic structure

The genetic variation of *S. mansoni* within the three communities showed no spatial or temporal structure. The variance explained by the factor time in both Pakh ($F_{SamplingTime/Pakh}$ = 0.001; CI: -0.004-0.005; p = 0.177) and Ndieumeul ($F_{SamplingTime/Ndieumeul}$ = 0.008; CI: -0.001-0.025; p = 0.088) was low and insignificant. Although significant, the estimates of spatial differentiation among villages relative to the whole sampling area were very low ($F_{Vil/Total}$ = -0.003; CI: 0.002-0.004; p < 0.001; Table 4.2). Similar results were obtained at the host level ($F_{Pop/Vil}$ = 0.004; CI: 0.002-0.006; p < 0.001; Table 4.2).

Hierarchical analyses showed that alleles were highly correlated within parasites relative to the host ($F_{Par/Pop}$ = 0.102; CI: 0.074-0.135; p < 0.001; Table 4.2), which were on average a hundred times higher than the estimates at the host and the village level (Table 4.2). This result was confirmed when host samples were considered independently: about 48 hosts harbored parasite populations showing positive F_{IS} values, from which 27 were highly significant (p < 0.01; Table 4.1). Only eight hosts harbored parasite populations showing negative F_{IS} values; none of them were significant.

Analysis in STRUCTURE v2.2.3 revealed no subdivision of populations as no optimal K value could be identified; the highest lnP(D) value was obtained for K = 1 (Supplementary Figure 4.1). The membership probabilities (Q-values) for K = 2 and higher were uniform across clusters (Supplementary Figure 4.1), indicating that individuals were inconclusively assigned to the clusters, which furthermore supports the results that K = 1.

4.3.4 Evidence for inbreeding

Both the Wahlund effect and biased sampling were rejected as systematic explanations for departures of HWE at the host level. STRUCTURE analyses revealed a homogenized population, thereby suggesting a minimum impact of the Wahlund effect (Supplementary Figure 4.1). Mean pairwise relatedness between parasites from each host did not depart significantly from its expected distribution for almost all host samples (Supplementary Table 4.2). Only host P46 potentially suffered from biased sampling and was therefore removed from subsequent statistical analyses (Supplementary Table 4.2). On the contrary, evidence

was found of a shift to lower individual heterozygosity (MLH) than expected under random mating (Supplementary Table 4.2) for many hosts (33) and variations in F_{IS} were strongly associated with variations in individual heterozygote deficiencies (MLH_{obs} - MLH_{exp}) / MLH_{exp} (Pearson's correlation test, $p < 0.001$; $r = -0.97$). Altogether, these results suggest that the variation observed in F_{IS} estimates between hosts is most likely due to variations in levels of parasite inbreeding.

Table 4.1. Genetic diversity of *Schistosoma mansoni* populations as estimated independently per host and per village.

Village	Host	Sampling Time	EPG	*Sh.*	Host age	#	AR	Hs	Ho	F_{IS}
Diokhor						162	1.54	0.54	0.46	0.16**
	P01	Aug09	560	0	50	17	1.52	0.52	0.52	0.00
	P02	Aug09	710	1	12	2	1.48	0.48	0.44	0.13**
	P03	Aug09	750	1	12	16	1.48	0.48	0.33	0.35**
	P04	Aug09	800	0	29	22	1.55	0.55	0.44	0.20*
	P05	Aug09	1050	0	37	9	1.54	0.54	0.54	0.00
	P06	Aug09	560	1	9	33	1.57	0.57	0.46	0.22**
	P07	Aug09	/	0	30	20	1.55	0.55	0.49	0.15*
	P08	Aug09	6470	1	35	6	1.55	0.55	0.55	0.00
	P09	Aug09	1000	1	13	6	1.52	0.52	0.43	0.19*
	P10	Aug09	20	0	4	31	1.54	0.54	0.43	0.21**
Pakh						280	1.55	0.55	0.51	0.08**
	P11	Feb09	340	0	37	51	1.57	0.57	0.56	0.02
	P12	Feb09	4590	0	29	10	1.51	0.51	0.53	-0.04
	P13	Aug09	4360	0	13	37	1.54	0.54	0.50	0.09*
	P14	Feb09	290	1	40	3	1.61	0.61	0.59	0.03
	P15	Feb09	3950	0	26	38	1.55	0.55	0.47	0.15**
	P16	Feb09	130	1	14	10	1.52	0.52	0.50	0.05
	P17	Aug09	430	0	6	20	1.52	0.52	0.37	0.31**
	P18	Feb09	10	1	35	2	1.65	0.65	0.61	0.08
	P19	Feb09	150	0	26	35	1.56	0.56	0.58	-0.05
	P20	Feb09	10	0	44	38	1.53	0.53	0.44	0.17**
	P21	Aug09	240	0	45	36	1.57	0.57	0.54	0.06
Ndieumeul						903	1.53	0.53	0.48	0.11**
	P22	Jan10	650	0	10	74	1.52	0.52	0.46	0.12**
	P23	Jan10	1570	1	10	17	1.50	0.50	0.46	0.08
	P24	Jan10	1010	0	16	64	1.50	0.50	0.47	0.05*
	P25	Jan10	110	0	16	24	1.54	0.54	0.46	0.17**
	P26	Aug09	370	1	10	35	2.15	0.54	0.48	0.13**
	P27	Jan10	290	0	9	20	1.53	0.53	0.50	0.07
	P28	Jan10	3950	1	20	67	1.52	0.52	0.48	0.08*
	P29	Aug09	770	1	12	19	1.47	0.47	0.31	0.35**
	P30	Jan10	480	1	10	25	1.56	0.56	0.43	0.23**
	P31	Jan10	190	0	8	21	1.51	0.51	0.45	0.12*
	P32	Jan10	330	1	15	21	1.59	0.59	0.51	0.14*
	P33	Jan10	1230	1	14	25	1.53	0.53	0.49	0.08*
	P34	Aug09	2020	1	5	22	1.48	0.48	0.34	0.33**

P35	Jan10	470	1	5	27	1.52	0.52	0.48	0.09*
P36	Aug09	760	1	5	37	1.47	0.47	0.33	0.30**
P37	Jan10	1050	0	45	44	1.53	0.53	0.49	0.07*
P38	Jan10	920	1	10	14	1.51	0.51	0.40	0.22*
P39	Jan10	440	1	5	21	1.53	0.53	0.56	-0.04
P40	Jan10	200	0	50	18	1.53	0.53	0.56	-0.04
P41	Jan10	390	1	14	15	1.56	0.56	0.53	0.05
P42	Jan10	180	1	22	17	1.55	0.55	0.57	-0.03
P43	Jan10	1260	1	15	16	1.55	0.55	0.57	-0.02
P44	Jan10	230	1	15	3	1.51	0.51	0.39	0.30*
P45	Jan10	980	1	16	31	1.52	0.52	0.49	0.05*
P46	Jan10	290	0	12	2	1.39	0.39	0.17	0.67**
P47	Jan10	520	1	8	5	1.51	0.51	0.29	0.48*
P48	Jan10	4140	1	5	29	1.55	0.55	0.51	0.08*
P49	Jan10	1260	0	30	17	1.51	0.51	0.38	0.27**
P50	Jan10	850	1	15	17	1.51	0.51	0.50	0.01
P51	Jan10	1980	1	11	64	1.50	0.50	0.50	0.01
P52	Jan10	200	1	8	18	1.51	0.51	0.38	0.27**
P53	Jan10	330	1	6	25	1.54	0.54	0.55	-0.01
P54	Jan10	70	1	16	2	1.41	0.41	0.28	0.41*
P55	Jan10	200	1	11	16	1.54	0.54	0.49	0.09*
P56	Jan10	130	1	6	13	1.54	0.54	0.46	0.15*
P57	Jan10	540	1	10	18	1.53	0.53	0.54	-0.01
Cercariae	Jan10	na	na	na	14	1.62	0.51	0.45	0.11*

EPG: eggs per gram. *Sh*: co-infection with *S. haematobium*. #: number of successfully genotyped parasites. AR: Allelic Richness. Hs: unbiased expected heterozygosity. Ho: observed heterozygosity. F_{IS}: inbreeding coefficient. Significant F_{IS} values were indicated with * when $p < 0.05$ and ** when $p < 0.001$ (i.e. Bonferroni corrected). The information on gender of the host was excluded from this table to preserve anonymity.

Table 4.2. Hierarchical analysis of parasite genetic diversity overall and per locus at three levels: within villages ($F_{Vil/Total}$), within hosts ($F_{Pop/Vil}$) and within individual parasites ($F_{Par/Pop}$). Standardized estimates of genetic differentiation are given for $F_{Vil/Total}$ and $F_{Pop/Vil}$ after recoding the data following Meirmans (2006). Note that recoding the data does not change the within population diversities; the estimate of $F_{Par/Pop}$ therefore remains the same. Number of alleles (# Alleles) are given per locus and overall.

Marker	# Alleles	$F_{Vil/Total}$	$F_{Pop/Vil}$	$F_{Par/Pop}$
L46951	7	0.002	0.006*	0.078**
CA11-1	7	0.003*	-0.001	0.085**
S9-1	5	-0.001	0.002	0.082**
SMD11	31	0.002**	0.007**	0.109**
SMD25	8	0.005*	0.004*	0.090**
SMD28	5	0.002	0.008*	0.015
SMD43	15	0.003*	0.004**	0.179**
SMD89	5	0.006	-0.004	0.059*
SMDA28	11	0.004**	0.001	0.067**
Overall	10.4	0.003**	0.004**	0.102**
(CI: 2.5% - 97.5%)		(0.002 - 0.004)	(0.002 - 0.006)	(0.074 - 0.135)
Overall (standardized)	NA	0.007	0.009	NA

Significance was indicated with * when $p < 0.05$ and ** when $p < 0.002$ (i.e. Bonferroni corrected).

4.3.5 Link between human host-specific factors and S. mansoni parasite genetics

Model building revealed that host age was the only predictor explaining the variance in some parasite population statistics (Table 4.3). In a GLM with host age, gender, infection intensity (EPG) and co-infection with *S. haematobium*, *S. mansoni* inbreeding (F_{IS}) was negatively associated with age of the host (β = -0.004, p = 0.018), while a significant increase in observed heterozygosity with host age was found (β = 0.003, p = 0.005). Host age explained 11% of the variation in parasite inbreeding (adjusted R^2 = 0.112, p = 0.007) and 17% of the variation in observed heterozygosity (adjusted R^2 = 0.169, p = 0.001). No significant interactions between age and other covariates were found (results not shown). Results were similar when GLMs were repeated excluding hosts that harbored less than 13 genotyped parasites (Table 4.3) and jackknifing over loci revealed that all loci equally contributed to these patterns (Table 4.4). These results suggest that the observed associations did not suffer from sampling bias in terms of number of hosts, number of genotyped parasites and number of markers. Expected heterozygosity (Hs) was significantly higher in the village Pakh than in Ndieumeul (β = 0.020; p = 0.024), while Hs increased borderline significantly with age of the host (β = 0.0005, p = 0.072; Table 4.3). No significant associations with Hs were found however when excluding hosts that were genotyped for less than 13 parasites (Table 4.3). Allelic richness (AR) did not reveal any significant effects (Table 4.3).

Table 4.3 Results from general linear models examining the association of parasite population statistics with host demographic variables, infection intensity (EPG) and co-infection with *Schistosoma haematobium*.

			Minimum 2 parasites		Minimum 13 parasites	
Parasite	**Host**		**Estimate β (95% CI)**	***p*-value**	**Estimate β (95% CI)**	***p*-value**
F_{IS}		Intercept	0.219		0.219	
	Village	Ndieumeul	ref		ref	
		Diokhor	0.032 (-0.030, 0.093)	0.308	0.058 (-0.014, 0.131)	0.112
		Pakh	-0.013 (-0.076, 0.050)	0.690	-0.009 (-0.085, 0.065)	0.790
	Gender	Female	ref		ref	
		Male	-0.004 (-0.039, 0.031)	0.806	-0.020 (-0.057, 0.016)	0.261
	Co-infection	Present	ref		ref	
		Absent	0.001 (-0.039, 0.041)	0.948	0.009 (-0.036, 0.055)	0.679
	Age		-0.004 (-0.007, -0.001)	0.018*	-0.004 (-0.007, -0.0008)	0.016*
	EPG		-2.10^{-5} (-4.10^{-5}, 1.10^{-5})	0.207	-1.10^{-5} (-4.10^{-5}, 3.10^{-5})	0.689
Ho		Intercept	0.416		0.425	
	Village	Ndieumeul	ref		ref	
		Diokhor	-0.024 (-0.061, 0.014)	0.208	-0.034 (-0.079, 0.011)	0.132
		Pakh	0.026 (-0.012, 0.064)	0.177	0.021 (-0.026, 0.067)	0.370
	Gender	Female	ref		ref	
		Male	0.005 (-0.016, 0.026)	0.626	0.012 (-0.011, 0.034)	0.291
	Co-infection	Present	ref		ref	
		Absent	-0.009 (-0.034, 0.014)	0.418	-0.009 (-0.038, 0.019)	0.497
	Age		0.003 (0.0008, 0.005)	0.005**	0.002 (0.0003, 0.004)	0.023*
	EPG		4.10^{-6} (-1.10^{-5}, 2.10^{-5})	0.552	0.000 (-2.10^{-6}, 2.10^{-5})	0.971
Hs		Intercept	0.520		0.531	
	Village	Ndieumeul	ref		ref	
		Diokhor	-0.008 (-0.025, 0.010)	0.381	-0.004 (-0.022, 0.014)	0.668
		Pakh	0.020 (0.003, 0.038)	0.025*	0.014 (-0.004, 0.033)	0.118
	Gender	Female	ref		ref	
		Male	0.002 (-0.007, 0.012)	0.631	0.004 (-0.005, 0.012)	0.418
	Co-infection	Present	ref		ref	
		Absent	-0.007 (-0.018, 0.005)	0.238	-0.003 (-0.015, 0.008)	0.547
	Age		0.001 (-0.00007, 0.0017)	0.072	0.0004 (-0.0004, 0.001)	0.360
	EPG		-2.10^{-6} (-9.10^{-6}, 5.10^{-6})	0.532	-2.10^{-6} (-1.10^{-5}, 6.10^{-6})	0.587
AR		Intercept	1.53		1.55	
	Village	Ndieumeul	ref		ref	
		Diokhor	-0.011 (-0.059, 0.036)	0.640	-0.014 (-0.083, 0.055)	0.690
		Pakh	0.018 (-0.031, 0.067)	0.456	0.022 (-0.049, 0.093)	0.534
	Gender	Female	ref		ref	
		Male	-0.009 (-0.036, 0.018)	0.510	-0.121 (-0.046, 0.022)	0.482
	Co-infection	Present	ref		ref	
		Absent	-0.014 (-0.045, 0.018)	0.389	-0.014 (-0.057, 0.029)	0.521
	Age		0.001 (-0.002, 0.003)	0.653	-6.10^{-5} (-0.003, 0.003)	0.967
	EPG		-6.10^{-6} (-3.10^{-5}, 1.10^{-5})	0.505	-1.10^{-5} (-4.10^{-5}, 1.10^{-5})	0.364

Significance was indicated with * when $p < 0.05$ and ** when $p < 0.006$ (i.e. Bonferroni corrected).

Table 4.4 Jackknifed estimates of Pearson's correlation tests between human host age on the one hand and *Schistosoma mansoni* inbreeding and heterozygosity on the other hand.

	host age - parasite F_{IS}		host age - parasite Ho	
	r	***p***	***r***	***p***
L46951	-0.316	0.018*	0.405	0.005*
CA11-1	-0.377	0.004*	0.423	0.003*
S9-1	-0.350	0.008*	0.383	0.008*
SMD11	-0.347	0.009*	0.437	0.002**
SMD25	-0.339	0.011*	0.415	0.004*
SMD28	-0.357	0.007*	0.385	0.006*
SMD43	-0.392	0.003*	0.455	0.001**
SMD89	-0.361	0.006*	0.398	0.006*
SMDA28	-0.315	0.018*	0.423	0.001**
Overall	-0.350	0.008*	0.429	0.001**

Significance was indicated with * when $p < 0.05$ and ** when $p < 0.003$ (i.e. Bonferroni corrected).

4.4 Discussion

Schistosomiasis has a complex epidemiology with a substantial variation among human hosts in infection intensity and pathology. Within schistosomiasis-endemic communities, infection intensities are related to host age, with an increase in intensity in early childhood, a peak in adolescence (8 – 15 years) and a decrease thereafter (Stelma et al. 1993; Meurs et al. 2012). This pattern could be explained by immunity that is acquired over age (Kabatereine et al., 1999), by differential water contact behavior with children being more exposed to infected water than adults (Fulford et al., 1996), and/or by other factors (e.g. skin composition, hormones) that vary with age (Fulford et al., 1998). Here we found a negative correlation between human host age and *S. mansoni* inbreeding levels, and a positive correlation between host age and parasite heterozygosity (Figure 4.3). These results indicate that older human hosts are to a greater extent infected by genetically unrelated parasites, resulting in heterozygous offspring. Such age-related differences could result from a number of mutually non-exclusive scenarios, which will be discussed below.

4.4.1 Host immunity selects for unrelated S. mansoni parasites

The concomitant immunity hypothesis states that adult schistosomes invoke a partial immunoprotective host response against new incoming larval schistosomes (Terry, 1994; Brown & Grenfell, 2001). This host protective immunity may be parasite genotype-dependent, with the infectivity rate of incoming schistosomes being lower when genetically

more similar to the (immunizing) adult schistosomes (Beltran et al., 2011). Such a genotype-dependent antigenicity could explain the observed decline in parasite inbreeding with host age: the first infecting schistosome genotypes are recognized by the host immune system that, upon super-infection, eliminate genetically similar genotypes more efficiently than unrelated schistosomes (Galvani, 2005). This would lead to an increased recruitment of genetically dissimilar worms with host age.

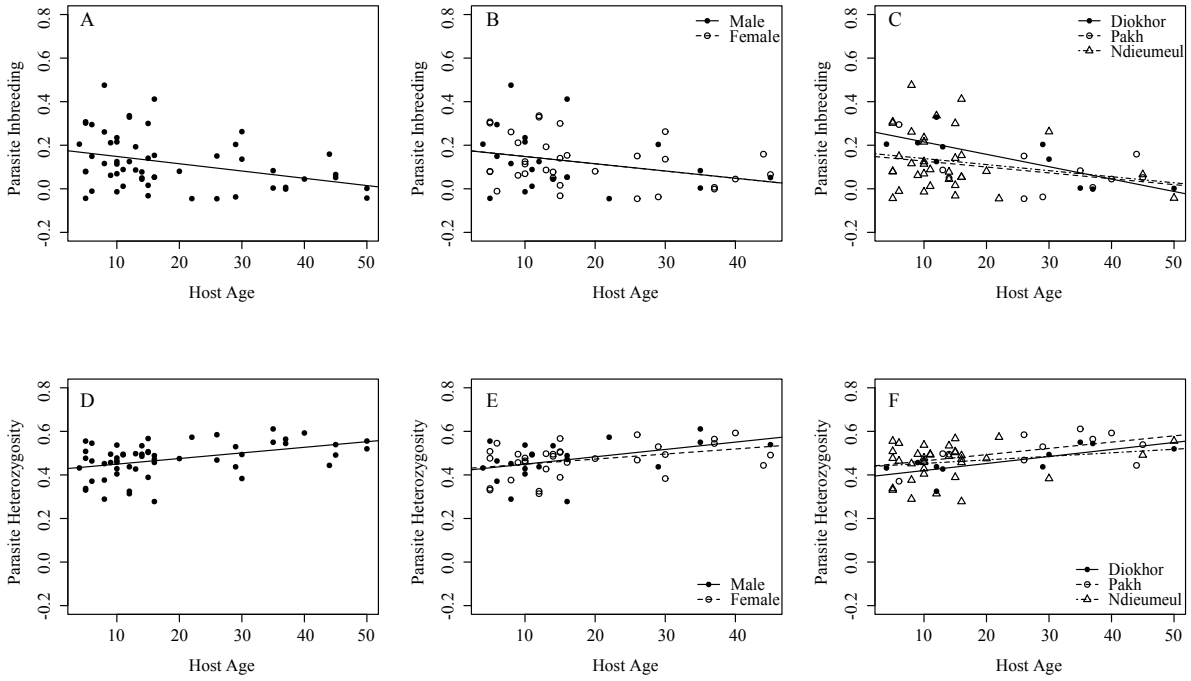


Figure 4.3. Linear regressions between human host age on the one hand and *Schistosoma mansoni* inbreeding (A) and heterozygosity (D) on the other hand. B and E: according to host gender. C and F: according to village.

4.4.2 Hosts serve as genetic mixing bowls for S. mansoni parasites

Given the long life expectancy of schistosome worms, older hosts may have acquired genetically unrelated parasites in space (if the action radius of the host increases with its age) and in time (if cercariae at a given transmission site are genetically dissimilar between time points). Human hosts then become 'genetic mixing bowls', where the accumulation of unrelated parasites during their lives results in an increase in the diversity of the respective parasite population (Curtis and Minchella, 2000).

On the other hand, children may be more exposed to related parasites than adults due to differences in water contact behavior. Children in Senegal (Scott et al., 2003; Sow et al., 2011) and elsewhere (e.g. Fulford et al. 1996) showed more and longer water contact than adults, and a study in Kenya showed that children visit fewer transmission sites than adults

(Kloos et al., 1997). Children that visit the same transmission site frequently would have an increased risk of re-infection by *S. mansoni* strains that are related to the ones that previously contributed to the gene pool (i.e. sib transmission; Anderson et al., 1995; Criscione et al., 2005). This will increase the chance for biparental inbreeding of *S. mansoni* parasites resulting in an increase in F_{IS}. If sib transmission were strong, however, most of the genetic variation would be found among (groups of) hosts (e.g. Mulvey et al., 1991), while models have shown that gene flow of unrelated parasites and the mixing of larvae before or after asexual reproduction erode the effects of sib transmission (Prugnolle et al., 2005b). This study revealed low genetic differentiation between *S. mansoni* parasite populations at the village and the host level, indicating high parasite gene mixing (Table 4.2). Furthermore, the mean pairwise relatedness between the cercariae from Ndieumeul did not differ significantly from its expected distribution, whereas the contrary would be expected if the pool consisted of siblings (Supplementary Table 4.2). Altogether, these results suggest that the impact of sib transmission on variations in F_{IS} estimates within our study area (or at least within the transmission site of Ndieumeul) is probably minimal. This implies that the suggested role of age-dependent water contact behavior in schistosome inbreeding may be of little relevance within our study area. Preliminary data also showed that the majority of the inhabitants within our study area tend to visit the same transmission site regardless of their age, suggesting little differences in water contact behavior between age groups. It can however not be excluded that we missed water contact sites that are separated from the main transmission sites along Lake Guiers.

Similarly, a study on *S. haematobium* within two schools in Mali (Gower et al., 2011) found higher numbers of unique adult worms (as inferred from larval genotypes) in a school where the average age of the children was also higher than in the other school. A clear association between host age and parasite diversity was however not reported. This might be due to the fact that only parasites from children were sampled, which narrows the age range under study. However, a higher number of unique adult worms were found in the parasite populations recovered from boys in one of the two schools (Gower et al., 2011). A study in Zimbabwe also found higher levels of *S. haematobium* diversity in boys than in girls, and this was in accordance with the observation that males had more water contact and more intense infections than females (Brouwer et al., 2003). We however did not find any indication that gender would have an impact on the genetic constitution of *S. mansoni*

infrapopulations (Table 4.3). An earlier epidemiological survey within the same villages found no differences between gender and risk of *S. mansoni* or *S. haematobium* infection or infection intensity (Meurs et al. 2012). Only older women showed higher risk for infection than older men, but the difference was not significant. Altogether, these results suggests that within our study area there might be little differences between gender or age in water contact, or that gene flow is large enough to break down any association. Note that the fact that we reject the day-to-day water contact behavior as a potential explanation for higher levels of inbreeding within children does not necessarily reject the genetic mixing bowl hypothesis, as at a larger time scale older hosts could still have been exposed relatively more to unrelated parasites than children.

4.4.3 Host mobility and Wahlund effect

The fact that parasite genotypes are randomly distributed among hosts and villages (Table 4.2) may be explained by high host mobility. Previous studies hypothesized that the lack of spatial differentiation could be due to the absence of boundaries to water bodies, thereby facilitating host movement among transmission sites (Thiele et al., 2008; Agola et al., 2009; Steinauer et al., 2009). Indeed, the open and flat Sahel region within our study area may impose few restrictions on human host movement. In addition, all transmission sites are situated along (or near) the same water body (Lake Guiers), most likely facilitating both human and snail mobility among sites. Such a human host movement could result in a Wahlund effect when transmission sites represent spatially or temporally separated gene pools (i.e. presence of cryptic population structure). This Wahlund effect has often been used as an argument to explain positive deviations from Hardy-Weinberg Equilibrium (HWE) within natural schistosome populations (e.g. Agola et al., 2006; Gower et al., 2011). Here, we showed that the impact of the Wahlund effect within our sampling area could be considered minimal due to the absence of population structure, and that positive deviations from HWE within natural schistosome populations could also result from non-random mating patterns that are linked with host-specific factors.

4.5 Conclusions

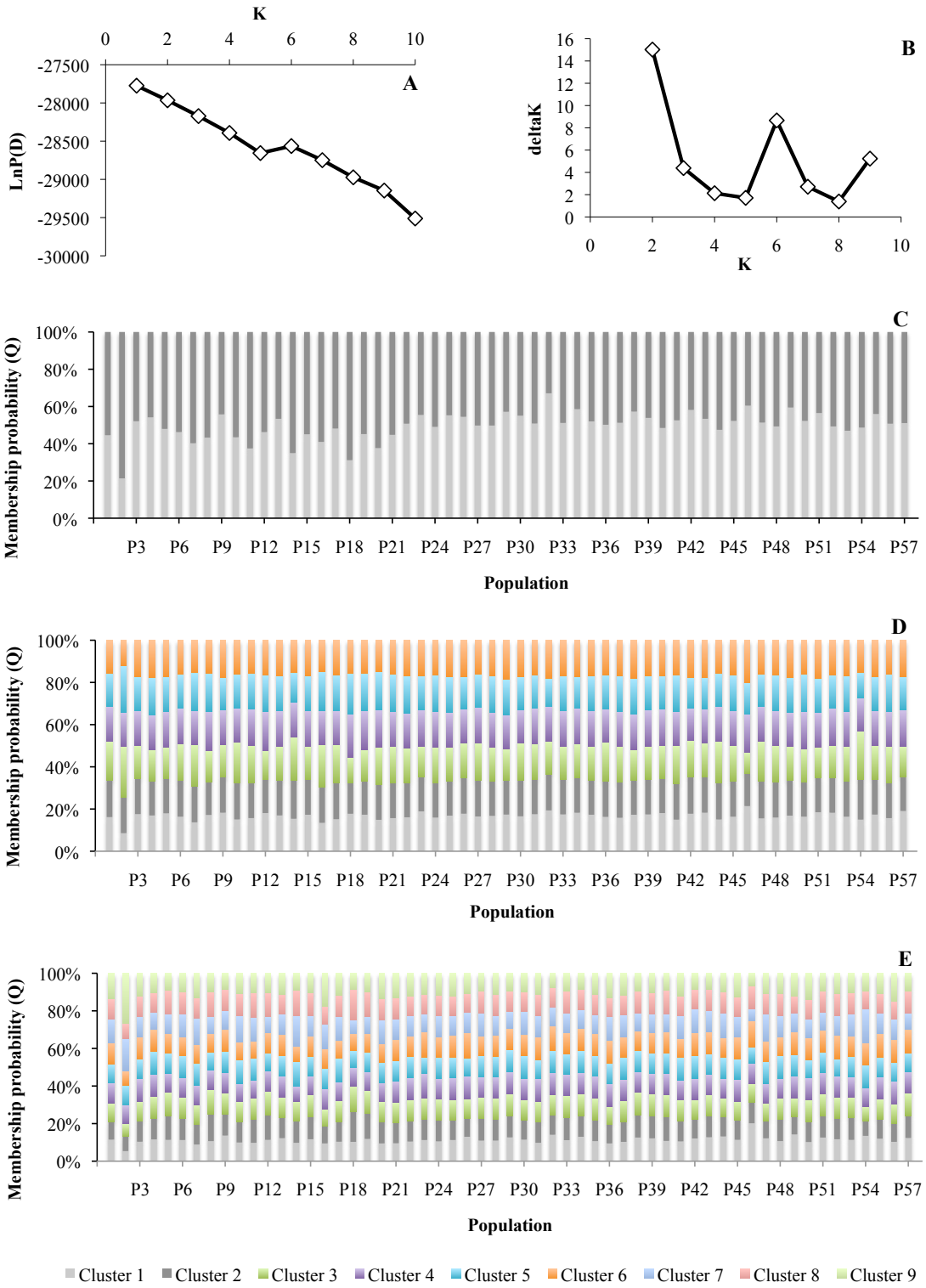
A significant association between host age on the one hand and parasite inbreeding and heterozygosity on the other hand was found. Our results show that young hosts are infected by related parasites while there is an accumulation of unrelated strains as the host ages. We argue that these patterns could be explained by 1) genotype-dependent 'concomitant immunity', leading to selective recruitment of genetically unrelated worms with host age, and/or 2) the 'genetic mixing bowl' hypothesis, where older hosts may have been exposed to a wider variety of schistosome strains than children. Regardless of the underlying mechanism, we have shown that host age may shape the genetic diversity of parasites and therefore their evolutionary potential. More field-based data are needed to confirm these results and to assess the relative contribution of each of these suggested mechanisms to the structuring of natural parasite populations.

Acknowledgements

We would like to thank the villagers for their cooperation, the Senegalese field team (M. Diop, A. Fall, N. Sy, M. Wade, A. Yague) for assisting with the fieldwork, Dr. M.H.D. Larmuseau for many discussions on population genetics, D. Schaerlaekens for his comments on the manuscript and F.A. Kidane for her help with sampling and genotyping. FVDB is a doctoral fellow of the Flemish Interuniversity Council (VLADOC). TH is a postdoctoral fellow of the Research Foundation – Flanders (FWO-Vlaanderen). Research benefited from a Research Grant of Research Foundation – Flanders (contract G.0552.10) and from travel grants awarded by the Flemish Interuniversity Council (VLIR) to FVDB and NB.

Data Archiving

R scripts, parasite genotypes and human host-specific factors were deposited in the Dryad repository: doi:10.5061/dryad.6nm40. Note that the information on gender was excluded from the raw data to preserve anonymity.



Supplementary Figure 4.1 Estimates of the posterior probability LnP(D) of the data and its second order rate change *ΔK* for a given *K* (1-10) calculated with Structure v2.2.3. The maximum estimate of LnP(D) is given for *K* = 1 (A) and *ΔK* revealed no clear change with increasing *K* (B). Figures C, D and E denote the average membership probabilities of each parasite infrapopulation to two, five or ten clusters respectively. The numbering of the hosts is the same as in Table 4.1.

Supplementary Table 4.1. Abundance and *Schistosoma mansoni* prevalence of collected *Biomphalaria pfeifferi* snails.

Village	Date	Site	N sampled	N infected	*S. mansoni* prevalence
Mbane	Aug-08	1	81	0	0.00%
		3	0		
Temey	Aug-08	2	7	1	14.29%
Nder	Aug-08	1	29	1	3.45%
Roumbatine	Aug-08		0		
Yetti Yone	Aug-08		0		
Diokhor	Aug-09	1	1	0	0.00%
Ndieumeul	Aug-09	3	0		
	Jan-10	1	135	13	9.62%
		2	0		
		3	0		
Pakh	Aug-08	1	0		
		2	1	0	0.00%
	Feb-09	1	0		
		3	15	0	0.00%
	Aug-09	1	0		
		2	1	0	0.00%
		3	0		
		4	0		
	Jan-10	1	0		
		4	13		0.00%

Supplementary Table 4.2. *P*-values for individual relatedness and multilocus heterozygosity (MLH) after 1000 randomizations. $R_{obs} > R_{exp}$ reflects the probability of observing higher pairwise relatedness than expected, suggesting biased sampling of (many) siblings from a limited number of adults. $MLH_{obs} < MLH_{exp}$ reflects the probability of observing a lower individual heterozygosity than expected, suggesting parasite inbreeding. The pattern of individual heterozygosity follows best the pattern of observed F_{IS} as estimated per host. Significance was indicated with * when $p < 0.05$ and ** when $p < 0.01$.

Host	$R_{obs} > R_{exp}$	$MLH_{obs} < MLH_{exp}$	F_{IS}
P01	0.709	0.438	0.00
P02	0.507	0.297	0.13**
P03	0.431	0.000**	0.33**
P04	0.585	0.000**	0.20**
P05	0.097	0.650	0.00
P06	0.647	0.000**	0.21**
P07	0.831	0.108	0.14*
P08	0.831	0.483	0.00
P09	0.556	0.011*	0.19*
P10	0.38	0.000**	0.20**
P11	0.327	0.356	0.01
P12	0.746	0.910	-0.04
P13	0.041*	0.206	0.09*
P14	0.775	0.266	0.04

P15	0.835	0.000**	0.15**
P16	0.318	0.405	0.05
P17	0.332	0.000**	0.29**
P18	0.266	0.106	0.08
P19	0.679	0.958	-0.05
P20	0.309	0.000**	0.16**
P21	0.099	0.098	0.05
P22	0.047*	0.000**	0.12**
P23	0.063	0.072	0.07
P24	0.328	0.031*	0.05*
P25	0.472	0.000*	0.15**
P26	0.453	0.000*	0.11**
P27	0.711	0.073	0.06
P28	0.473	0.003**	0.08**
P29	0.893	0.000**	0.34**
P30	0.752	0.000**	0.23**
P31	0.600	0.001**	0.12**
P32	0.519	0.000**	0.14**
P33	0.680	0.018*	0.08
P34	0.715	0.000**	0.31**
P35	0.729	0.008**	0.08**
P36	0.234	0.000**	0.30**
P37	0.145	0.010*	0.07*
P38	0.616	0.000**	0.22**
P39	0.288	0.856	-0.04
P40	0.126	0.794	-0.04
P41	0.220	0.117	0.04
P42	0.165	0.726	-0.05
P43	0.905	0.627	-0.03
P44	0.144	0.014*	0.30*
P45	0.923	0.030*	0.05*
P46	0.000**	0.021*	0.67**
P47	0.052	0.000**	0.48**
P48	0.181	0.006*	0.08*
P49	0.484	0.000**	0.26**
P50	0.216	0.450	0.02
P51	0.558	0.344	0.01
P52	0.362	0.000**	0.26**
P53	0.706	0.519	-0.01
P54	0.127	0.078	0.41*
P55	0.124	0.023*	0.09*
P56	0.775	0.001**	0.15**
P57	0.948	0.475	-0.01
Cerc.	0.791	0.027*	0,12*

CHAPTER 5

Fighting a losing battle? A simulation-based approach to evaluate the impact of community-based drug treatment on the genetic diversity of schistosome populations

F Van den Broeck, J Vanoverbeke, FAM Volckaert, K Polman & T Huyse

Abstract

The shift towards community-based chemotherapy in sub-Saharan Africa may lead to intensive and prolonged selection pressures on schistosome populations, possibly leading to genotypic and phenotypic changes in traits such as fecundity, infectivity or drug resistance. In the absence of molecular markers for detecting adaptive traits, selectively neutral markers are ideal tools to study recent population size changes, allowing to fully understand and predict their effect on current and future epidemiological dynamics. In this study we performed simulations to assess the impact of population size bottlenecks on levels of neutral genetic diversity of schistosome parasites. A population model with a logistic growth model in an island model at equilibrium was used to assess the effect of treatment coverage, effectiveness and frequency on changes in parasite genetic diversity at both infrapopulation and component population level. Results showed that levels of genetic diversity would not decrease under all treatment scenarios. More specifically, the genetic composition was only affected to a small extent in heavily infected individuals or when treatment was only administered once. Coverage was the only factor affecting levels of component population genetic diversity, which would only decrease when all hosts were treated within a given community. Finally, increasing the effectiveness and the frequency of treatment increased the severity of the genetic bottleneck. While our simulations are a simplified representation of community-based treatments within a natural setting, we predict that control programs under current treatment policies are not expected to seriously affect levels of *Schistosoma* genetic diversity. High levels of sustained genetic diversity could compromise the success of control programs as they allow the rapid selection of epidemiological relevant traits such as drug resistance.

5.1 Introduction

About 200 million people worldwide are infected with schistosome parasites (Steinmann et al., 2006). These parasitic trematodes cause schistosomiasis, a chronic debilitating disease that affects the physical and mental health of children (Chitsulo et al., 2000; van der Werf et al., 2003). Although it kills about 280,000 people each year in Sub-Saharan Africa alone (van der Werf et al., 2003), the disease has been largely neglected throughout the history within this region and is overshadowed by the three 'lethal' diseases malaria, tuberculosis and HIV/AIDS (Fenwick and Webster, 2006). Since 2003 some sub-Saharan countries implemented national control programs using the current drug of choice praziquantel (PZQ), aiming to mitigate the burden of disease by killing adult worms and reducing egg production (Kabatereine et al., 2007). In six sub-Saharan countries, the Schistosomiasis Control Initiative (SCI) assisted the development and implementation of these morbidity control programs, resulting in the treatment of more than 44 million people over a period of six years (Fenwick et al., 2009). With the increasing use of a single drug on a large scale, concerns were raised about the possible emergence of drug resistance (Doenhoff et al., 2002; Fenwick and Webster, 2006), and the possible impact on the evolution of other adaptive traits such as virulence, infectivity and fecundity (Webster et al., 2008). Parasite evolution could indeed become a key obstacle in the development of any effective disease control program (Webster et al., 2008) and as such, knowledge on how the genotype and phenotype of parasite populations may change in response to such selective pressure is very important.

Besides the obvious biological and biomedical relevance of directly assessing the adaptive evolution of parasites in response to drug treatment, an equally important step towards effective mass drug administration may be to investigate its impact on the parasite population. Treatment is expected to result in a substantial decline in effective population size (at least locally). Such population bottlenecks may thus lead to increased inbreeding and low levels of genetic diversity and could result in the random fixation of (possibly deleterious) alleles and thus hamper the evolutionary potential of parasites to adapt to new conditions (such as further chemotherapy), driving the success of mass treatment programs. Yet, schistosomes typically live in large populations characterized by a high degree of genetic variability (Gower et al., 2013), and reduced population sizes after treatment may therefore not necessarily result in a (strong) decrease in genetic diversity. Moreover, treatment

programs might have been too restricted in space (or time), resulting in a large proportion of the (meta)population left untreated (i.e. *refugia*) and allowing for the treated area or (sub)population to be quickly re-colonized by untreated parasites. Assessing and understanding the impact of treatment on population genetic composition of schistosome populations could therefore not only provide insights into the evolutionary potential of the parasite but also support the implementation, monitoring and evaluation of mass human chemotherapy programs.

Here we focus on the effects of treatment on a set of neutral markers as indicators of genome-wide changes in genetic diversity. We use a finite island model to explore the effects of community-based treatment on the genetic variability within parasite populations. We defined the model in such a way that it allows to investigate how genetic variability is affected by 1) the effective population size before treatment, 2) the coverage of treatment (i.e. the proportion of human hosts receiving chemotherapy), 3) the frequency of treatment (i.e. the number of treatments administered within a given time frame) and 4) the effectiveness of treatment (i.e. the proportion of parasites killed within a given host).

5.2 Methods

5.2.1 Description of the model

The digenean trematode *Schistosoma* spp. has a complex two-host life cycle with an asexual amplification stage in the snail intermediate host, yielding thousands of clonal cercariae that infect the human final host. The cercariae develop into dioecious worms that reproduce sexually within the veins, producing eggs that are expelled into the external environment through urine or feces. Upon contact with water, eggs will hatch into free-swimming miracidia that infect the intermediate snail host. Note that the parasite cycles obligatory through two hosts each generation and that population growth within the human host can only happen due to new infections (Figure 5.1).

A schistosome population thus comprises adult parasites within the human host where genetic recombination occurs and larval parasites that reproduce asexually within the snail host. Here we will simplify our model by excluding the asexual amplification stage and focus on the adult parasites in human hosts. A study on *S. mansoni* infecting wild rat *Rattus rattus* showed that the asexual amplification within snails did not play an important role in the

total genetic diversity of the adult stages (Theron et al., 2004). Models furthermore confirmed that the effect of the clonal amplification phase on the adult population genetic structure is expected to be small when there are high levels of parasite gene flow (Prugnolle et al., 2005a, 2005b). This has been reported within most natural settings of human schistosomes (Agola et al., 2009; Steinauer et al., 2009).

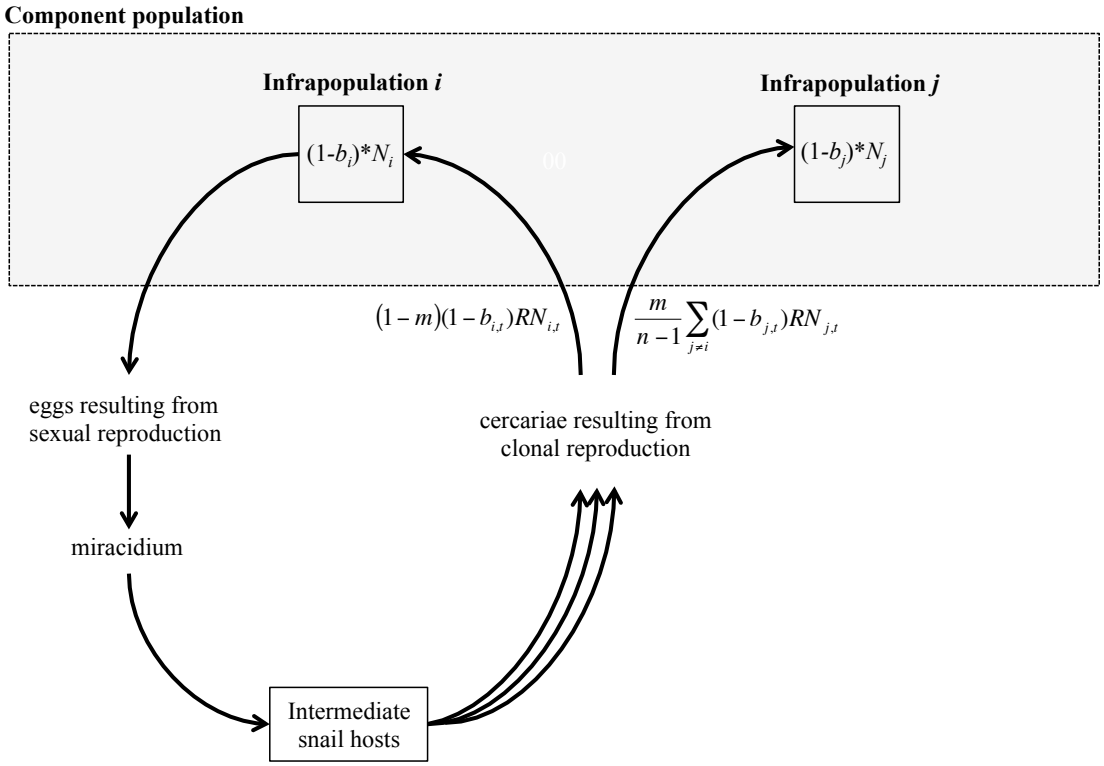


Figure 5.1 Life cycle of *Schistosoma* spp. with indication of the component population (i.e. all adult worms within one host population) and infrapopulation (i.e. all adult worms within one host individual). Arrows indicate the direction of the life cycle. See text for details on the life cycle and the model.

We consider a population of $\sum N_i$ adult worms of the same species within a given host population, i.e. component population (Figure 5.1) (Bush et al., 1997). The component population is subdivided into *n* subpopulations (finite island model), each with the same carrying capacity K_c. A subpopulation *i* represents a group of N_i randomly mating adult worms of the same species within one individual host, i.e. infrapopulation (Figure 5.1) (Bush et al., 1997). Note that the boundaries of the component population are not explicitly defined: it could represent a social host group (e.g. according to host gender or ethnicity), a village or a region/country. According to the island model, each reproduction cycle a proportion *m* of the newly produced individuals within a given infrapopulation will randomly

infect other infrapopulations, while a proportion 1-*m* infects the infrapopulation of origin. Note that we did not include explicit spatial structure: gene flow is the same among all infrapopulations and *m* represents the total migration from one infrapopulation to all other infrapopulations. Within each infrapopulation, population growth follows a discrete logistic growth model with per capita infection rate *R*, per capita mortality b_i and carrying capacity K_c.

$$N_{i,t+1} = (1-b_{i,t})N_{i,t} + \left[(1-m)(1-b_{i,t})N_{i,t} + \frac{m}{n-1}\sum_{j\neq i}(1-b_{j,t})N_{j,t}\right]R\left(1-\frac{(1-b_{i,t})N_{i,t}}{K_c}\right)$$

The population size of a schistosome infrapopulation *i* at time *t+1* is thus composed of the individuals that survive treatment and natural death, $(1-b_{i,t})N_{i,t}$, plus the individuals that result from new infections by larval stages from parents of the same infrapopulation, $(1-m)(1-b_{i,t})N_{i,t}$ and from parents from all other infrapopulations after migration, $\frac{m}{n-1}\sum_{j\neq i}(1-b_{j,t})N_{j,t}$. Note that an adult infrapopulation *i* does not decline in size as a result of emigration but only as a result of death, as only larvae migrate and not the adults. Note also that *R* does not represent the basic per capita reproduction number of infective larval stadia, but the per capita number of larvae that successfully infect a human host within the component population.

Mortality rate $b_{i,t}$ comprises both the background mortality and the mortality due to treatment with praziquantel (PZQ). The parameter $b_{i,t}$ can vary over infrapopulations and time to reflect the presence or absence of treatment-related mortality at given time *t* in infrapopulation *i*. It is defined as a matrix containing mortality rates per infrapopulation *i* (rows) per time unit *t* (columns).

In our model, we do not explicitly model individuals but rather the set of alleles that represents each infrapopulation. For an infrapopulation of *N* individuals we thus only keep track of the allelic identity of 2*N* alleles without taking into account the genotypic structure of the population. As a consequence, we assume that our infrapopulations are in Hardy-Weinberg and Linkage equilibrium at all times.

5.2.2 Implementation of the model

Simulations were initialized by generating alleles in an island model at equilibrium using the function *sim.genot* in the R (R Development Core Team, 2013) package 'hierfstat' (Goudet, 2005). This function generates alleles following the K-allele mutation model (KAM), and will therefore yield a number of alleles that is consistent for both microsatellites and SNPs. As most studies on schistosome population genetics use microsatellite markers, a total of 20 genetic markers were simulated each mutating to a maximum of 10 possible allelic states. The mutation rate was set to 10e-4, which is the estimated mutation rate for microsatellites in schistosomes (Valentim et al., 2009). The proportion of migration among islands was set to m as used in our subsequent simulations. We used a matrix of 100 islands, a matrix large enough not to result in an artificial loss of regional genetic diversity. The island population size equaled the carrying capacity K_c of our infrapopulations in the subsequent simulations.

Once our model was initialized with the generated set of alleles, during each timestep (generation), a number D of dying individuals was chosen in each infrapopulation based on a random binomial distribution with b_i probability of mortality. Subsequently, $2D$ alleles were randomly removed from the infrapopulation. After mortality, the number of successful new infections within each infrapopulation was calculated. For each new infection the infrapopulation of origin was drawn from a random multinomial distribution with probabilities based on N_i and dispersal probabilities, and two alleles were randomly drawn from that infrapopulation based on the current allele distributions.

5.2.3 Rationale of simulated scenarios

Policies on the frequency and target group of PZQ treatment as defined by the World Health Organization (WHO) depend on the Schistosoma prevalence that is measured among school-aged children (WHO, 2006). A community is classified as a high-risk community when this prevalence exceeds 50%, and all school-aged children and adults considered at risk (e.g. women in their domestic tasks, fishermen, farmers, irrigation workers) should be treated once a year. Within moderate-risk communities (i.e. 10% ≤ prevalence <50%), treatment should only be administered once every 2 years (WHO, 2006). We therefore incorporated the **effect of repeated treatments** by implementing scenarios where every 4 generations

one treatment is administered, which is equal to about one treatment every year if a generation time of 3 months is assumed.

The effectiveness of treatment relates to how well the treatment works in the field, as opposed to efficacy that measures how well it works in clinical trials or laboratory studies. Although PZQ should have an efficacy of 99-100%, its effectiveness is expected to be lower within natural conditions for several reasons. Some clinical studies that assessed the therapeutic efficacy of the recommended PZQ dose of 40 mg/kg in schistosomiasis patients that were not exposed to re-infection (mostly tourists) indicated that full cure may only be achieved in 40-60% of the cases (van Lieshout et al., 1994, 1997). The drug is furthermore not effective against immature worms present within the host and is less effective in lightly infected patients that show less robust immune responses (Kumar and Gryseels, 1994). Finally, there is some variability between different schistosome strains (at least in *S. mansoni*) in PZQ susceptibility (Cioli et al., 2004). Effectiveness of PZQ within natural settings could therefore be much lower than 99-100% because parasite infrapopulations could partly survive treatment because of the above reasons. We therefore simulated the **effect of effectiveness** on the genetic composition of schistosome populations by changing the mortality parameter as such that either 80% or 95% of all worms within a given infrapopulation are killed. Note that the effectiveness (the proportion of worms killed within a human host) is different from the cure rate (the proportion of hosts cured after treatment).

Control programs focus their treatment campaigns mainly on school-aged children, often leaving adults and pre-school children untreated (Odogwu et al., 2006). In addition, it is impossible to reach all school-aged children because of the difficult conditions met in the field, such as a low school enrollment rate (Toure et al., 2008). Lightly infected individuals will furthermore act as a continued reservoir for transmission as they are often left untreated. To illustrate, the SCI program aimed at reaching at least 75% of all school-aged children within a given country, and often targeted only those regions that were known to be highly endemic (Fenwick et al., 2009). It is therefore unlikely that a program will reach 100% of all infected hosts within a given village. We therefore tested what the **effect** is **of treatment coverage** on levels of parasite genetic diversity. This was done by administering treatments to 1%, 25%, 50%, 75% or 100% of all hosts.

Population size bottlenecks are expected to result in a decrease in genetic diversity when population sizes are small, but when populations are substantially large even a 99% reduction in size will not result in a concomitant decrease in genetic diversity. We therefore assessed the **effect of pre-treatment population sizes** (infection intensities) on the genetic composition of parasite populations post-treatment.

Parasite refugia are the group of parasites that are not reached during a control program and are therefore left untreated (Webster et al., 2008). The **effect of refugia** is not explicitly implemented in the model, but is indirectly estimated through the effects of **coverage** (human hosts that are (not) treated) and **effectiveness** (the number of worms that are (not) killed within a human host). Other possible refugia are larval stages within infested water or intermediate snail hosts and alternative reservoir host species such as rodents or baboons that are not reached during a control program (Duplantier and Sene, 2000). Our model however, does not allow inferring the effect of these types of refugia as a different transmission model is required when including more host species.

Gene flow m was set to 0.99 (i.e. m = 1-1/n) to mimic the scenario of a panmictic population, i.e. every host has an equal chance to become infected by parasites from any other host within the community. This means that our model simulates groups of hosts / communities that share the same gene pool, which is a relatively realistic assumption since most studies found very low levels of genetic differentiation between parasite infrapopulations from the same household (e.g. Thiele et al., 2008) or village (e.g. chapters 3 and 4). Note however that the group of infrapopulations is assumed to be completely isolated, i.e. there is no incoming gene flow from other groups. The different parameter values used to simulate the scenarios outlined above are given in Table 5.1. Scenarios investigating number of treatments, treatment coverage and effectiveness were implemented by changing the parameter $b_{i,t}$, which is defined by a two-dimensional matrix with columns representing populations and rows representing time-steps (generations). Infrapopulations that were not treated experienced the same natural mortality $b_{natural}$ of 0.03, meaning that 3% of each untreated infrapopulation at each time-step will die due to natural death. This number was chosen because in the absence of infection a natural mortality of 0.03 would eradicate an infrapopulation within 33 generations. The average lifetime of a schistosome worm is about 4-6 years (Rollinson and Simpson, 1987), which is

equal to about 12-30 generations (average 21 generations) as the generation time of *Schistosoma* is about 3-5 months. Infrapopulations that were treated experienced a mortality b_{PZQ} of 0.80 or 0.95 in order to test the effect of PZQ effectiveness. In these infrapopulations, the mortality due to natural death was considered negligible in comparison to the mortality due to PZQ. The frequency of treatment was implemented by defining b_{PZQ} for a given (set of) population(s) at generation 5 (one treatment only) or at generations 5-9-13-17 (four treatments, one every four generations). The coverage of treatment was implemented by defining b_{PZQ} for 1/100, 25/100, 50/100, 75/100 or all infrapopulations (i.e. 100/100). Finally, the effect of pre-treatment population sizes was assessed by varying the parameter K_c. Although direct quantification of *Schistosoma* worms within endemic settings is impossible, models predicted that worm burdens over 100 would be abundantly present within all endemic regions (Gryseels and De Vlas, 1996). Within a highly endemic focus at least 80% of the human population could be infected with more than 1,000 worms and 8% with at least 10,000 worms (Gryseels and De Vlas, 1996). We therefore implemented the following parameter range for K_c: 50, 100, 500 and 1000.

As the aim is to assess the effect of treatment on short time-scales, simulations were run for 50 generations. Knowing that *Schistosoma* have a generation time of 3-5 months depending on the species (Rollinson and Simpson, 1987), this means that 50 generations comprise about 10 to 17 years. The parameter R was set to 0.5. This value was chosen based on the lifetime reproduction number R_0 that was estimated to be 1-5 for schistosomes (Anderson and May, 1985; Woolhouse et al., 1996). For schistosomes, R_0 can be interpreted as the number of mated female schistosomes that were produced by one mated female schistosome during its lifetime. With an average lifetime of *Schistosoma* around 4-6 years (Rollinson and Simpson, 1987), reproductive output per generation R should be between 0.22-0.9.

Population genetic statistics were calculated using the allele frequencies after mortality, but prior to new infections. The number of alleles and gene diversity were estimated at both the infrapopulation level and the component population level. For each estimate the mean and standard error over all runs was calculated per generation.

Table 5.1 Parameters used to test each treatment scenario. The coverage of treatment is the proportion of infrapopulations that were treated (1%, 25%, 50%, 75% or 100%). The frequency of treatment is the number of times that a given (set of) infrapopulation(s) were treated: either one treatment or one treatment every 4 generations (for a total of 4 treatments). The effectiveness is the amount of parasites that are killed (80% or 95%) within an individual host. Each scenario of treatment was combined with four levels of K_c (50, 100, 500 and 1000) in two replicate runs, resulting in a total of 160 simulations.

Scenario	Coverage	Frequency	Effectiveness
1	1%	1 treatment	80%
2			95%
3		4 treatments	80%
4			95%
5	25%	1 treatment	80%
6			95%
7		4 treatments	80%
8			95%
9	50%	1 treatment	80%
10			95%
11		4 treatments	80%
12			95%
13	75%	1 treatment	80%
14			95%
15		4 treatments	80%
16			95%
17	100%	1 treatment	80%
18			95%
19		4 treatments	80%
20			95%

5.3 Results & Discussion

This study investigated the effects of community-based drug treatment on the genetic diversity of schistosome populations. A group of infrapopulations was simulated within an isolated panmictic component population, which means that every host has an equal chance to become infected by parasites from any other host within the community. A total of 160 simulations were performed that explored 20 different scenarios of treatment coverage, frequency and effectiveness (Table 5.1) in combination with different levels of pre-treatment population sizes. Our results allow us to make some general predictions on the control of schistosomiasis.

First, control programs mostly focus on regions that are known to be highly endemic, and on school-aged children that are the most heavily infected individuals within a community (WHO, 2006; Fenwick et al., 2009). In other words, they focus on those people that need treatment the most. Studies confirmed the success of such programs in terms of reductions

in prevalence and infection intensity as a result of worm death (Koukounari et al., 2006a, 2006b, 2007). Despite the decline in worm burden, our simulations showed that a concomitant substantial reduction in parasite genetic diversity is not necessarily observed, especially within those hosts that are the most heavily infected (Figure 5.2). A single treatment did not result in a strong decrease in levels of genetic diversity (less than 1 in 10 alleles) when hosts were infected with at least 1,000 worms (Figure 5.2). Note that a highly endemic focus at least 80% of the human population could be infected with more than 1,000 worms (Gryseels and De Vlas, 1996). The results are not surprising, as even an effectiveness of 95% would still leave sufficient parasites within the most heavily infected host individuals to sustain comparable levels of parasite genetic diversity as the ones before treatment. High pre-treatment levels of infection were put forward previously to explain low cure rates observed in Senegalese villages (Danso-Appiah and De Vlas, 2002).

Second, control programs do not reach all infected hosts within a given country (e.g. 75% were on average reached during the SCI program; Fenwick et al., 2009) or a given community (e.g. due to low school enrollment rate; Touré et al. 2008). Many parasites therefore escape treatment and remain in refugia. The coverage of treatment had a strong effect on the recovery of parasite populations and the estimates of genetic diversity at component population level. Only when all hosts were treated, a long-lasting effect was observed at both the infrapopulation and component population level (Figures 5.2 and 5.3). However, genetic diversity would recover relatively quickly in just a few generations when a proportion of the hosts were left untreated (Figures 5.2 and 5.3). This is because parasites in untreated hosts represent a source of genetic diversity from which new infections will occur in the next generation. Note that genetic diversity at the component population level is only affected when all hosts were treated, which is a consequence of the population being panmictic: almost all genetic variability present within the component population will be found within each individual infrapopulation, as such that treatment of all hosts is needed in order to affect the total amount of genetic diversity.

Third, the frequency and effectiveness of treatment were important determinants towards the severity of the genetic bottleneck. Several rounds of treatment resulted in a stepwise impoverishment of the allelic composition. Only after the third round of treatment, a clear reduction in genetic diversity was also seen at the component population level (Figure 5.3).

A genetic bottleneck was not observed within hosts infected with 500 strains when the effectiveness was only 80%. In contrast, genetic bottlenecks were observed even within infrapopulations containing 1000 worms when the effectiveness was increased to 95% (Figure 5.2). Although praziquantel is expected to kill 99% of the worms within a given host, lower cure rates (75-90%) are generally observed within most settings due to intense transmission, resulting in rapid re-infection and a high number of immature worms that are tolerant to the drug praziquantel.

Coverage	Effectiveness	K_c	5th generation	6th generation	7th generation	8th generation
1	80	50	−0.33	−0.09	−0.06	−0.03
		100	−0.15	−0.04	−0.01	−0.01
		500	−0.03	0	−0.01	−0.01
		1000	−0.01	−0.01	0	0
	95	50	−0.7	−0.08	−0.01	−0.01
		100	−0.4	−0.06	−0.01	−0.01
		500	−0.15	−0.01	−0.01	0
		1000	−0.09	−0.01	0	0
50	80	50	−0.28	−0.11	−0.05	−0.03
		100	−0.18	−0.06	−0.03	−0.01
		500	−0.03	−0.01	0	0
		1000	−0.02	−0.01	0	0
	95	50	−0.62	−0.19	−0.08	−0.04
		100	−0.49	−0.11	−0.05	−0.02
		500	−0.17	−0.02	−0.01	0
		1000	−0.09	−0.01	0	0
100	80	50	−0.26	−0.19	−0.13	−0.09
		100	−0.16	−0.11	−0.07	−0.05
		500	−0.03	−0.02	−0.01	−0.01
		1000	−0.02	−0.02	−0.01	0
	95	50	−0.62	−0.59	−0.5	−0.4
		100	−0.5	−0.42	−0.33	−0.25
		500	−0.17	−0.12	−0.09	−0.07
		1000	−0.09	−0.06	−0.04	−0.03

Figure 5.2 Heatmap showing the relative reductions in number of alleles in response to a single treatment, and this for different levels of coverage, effectiveness and pre-treatment population sizes (K_c). Estimates of relative reductions are obtained by comparing the number of alleles before treatment (i.e. 4th generation) with the number of alleles after treatment, but before re-infection (i.e. 5th generation), after the first re-infection (i.e. 6th generation), after the second re-infection (i.e. 7th generation) and after the third re-infection (i.e. 8th generation). Results for unbiased expected heterozygosity were similar (heatmap not shown).

To what extent these simulations and interpretations can be generalized towards the field remains an open question. We argue that it is safe to assume that treatment will not necessarily result in a decrease in parasite genetic diversity under all scenarios, but that in some cases sustained levels of parasite diversity will be observed. A few field studies that found no decrease in *S. mansoni* genetic diversity after treatment (Blanton et al., 2011; Huyse et al. 2013) corroborate these findings. A decrease in genetic diversity was however observed in a setting in Tanzania, despite the high endemicity of this focus (Norton et al.,

2010). Clearly, more studies on naturally collected parasites are needed to fully predict the impact of community based drug treatment on schistosome populations, as well as understand the factors that shape their outcome (Webster et al., 2008). French and colleagues used a stochastic resampling approach based on microsatellite genotypes obtained from naturally collected miracidia to explore the effects of various field sampling approaches on estimates of parasite genetic diversity (French et al., 2012). Their results indicate that sampling more hosts rather than more miracidia per host will lead to more robust estimates of parasite diversity. Future studies on naturally collected data should therefore aim to include as many hosts as possible in their population genetic analyses, including hosts that were left untreated, in order to fully predict the impact of treatment on the component population of schistosomes.

The main aim of most control programs such as the SCI is to mitigate the burden of disease by killing adult worms and reducing egg production (WHO, 2006; Fenwick et al., 2009). Control programs are therefore successful on the short-term as they often reduce the prevalence or infection intensity after treatment (Koukounari et al., 2006a, 2006b, 2007) and thus instantly relieve the patient from the burden of its disease. However, our simulations and the small amount of field-based data indicate that current treatment practices will most likely not seriously affect schistosome component populations on the long-term. This could compromise control programs, as sustained levels of parasite genetic diversity after treatment allow the evolution of epidemiological relevant traits such as drug resistance.

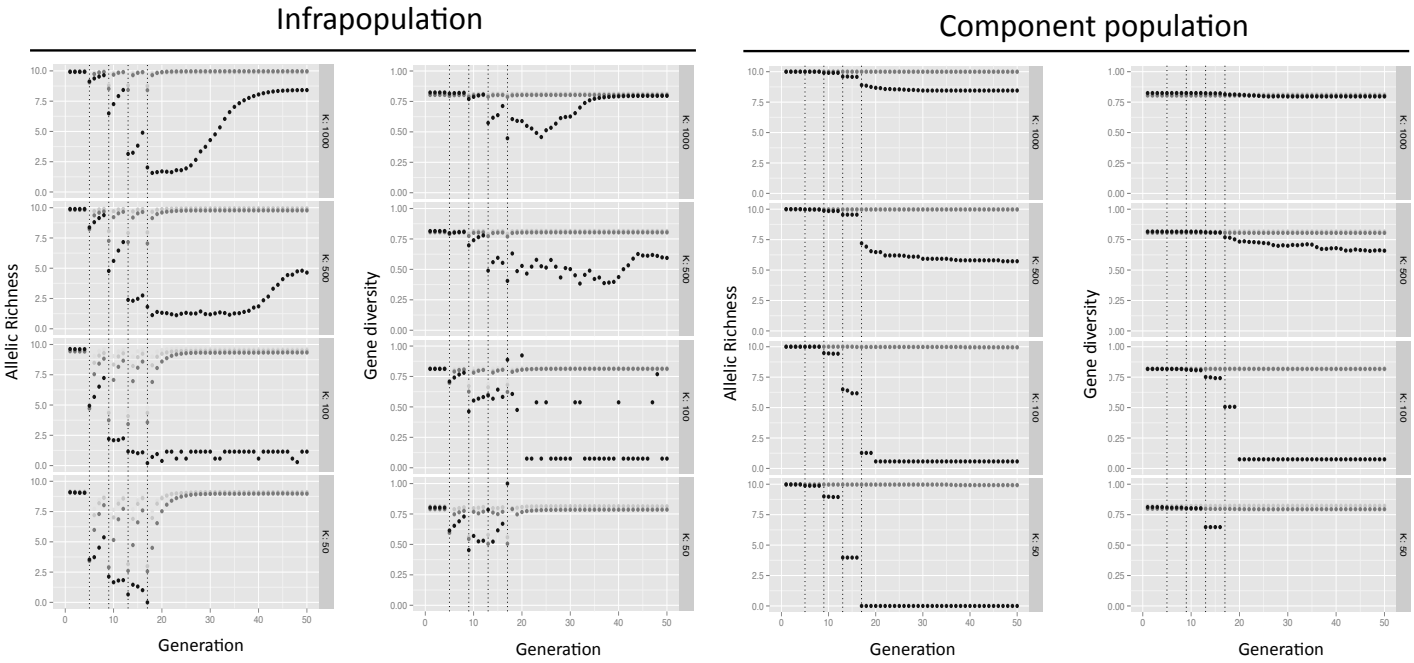


Figure 5.3 Results for the effect of treatment coverage on the genetic diversity (AR and Hs) of infrapopulations and component populations for different combinations of pre-treatment population sizes (K_c). The dashed lines show the time points when treatments were administered (four treatments in total) with an effectiveness of 95%. Dots represent the mean over all treated infrapopulations. Dots are colored in different shades of gray according to the coverage: 25% (light gray) until 100% (dark gray).

CHAPTER 6

Regular treatments of praziquantel do not impact on the genetic make-up of *Schistosoma mansoni* in Northwest Senegal

T Huyse, F Van den Broeck, T Jombart, B Webster, O Diaw, F Volckaert, F Balloux, D Rollinson & K Polman[1]

Published in Infection, Genetics and Evolution 2013, 18: 100-105

Abstract

The Senegal River Basin (SRB) experienced a major epidemic of intestinal schistosomiasis in the early nineties, after the construction of a dam for irrigation purposes. Exceptionally low cure rates following praziquantel (PZQ) treatment at the onset of the epidemic raised concerns about PZQ resistant strains of *Schistosoma mansoni*, although they could also be attributed to the intense transmission at that time. A field study in the same region more than 15 years later found cure rates for *Schistosoma mansoni* still to be low, whereas *S. haematobium* responded well to treatment. We collected *S. mansoni* miracidia from children at base-line prior to treatment, six months after two PZQ treatments and two years after the start of the study when they had received a total of five PZQ treatments. In total, 434 miracidia from 12 children were successfully genotyped with at least six out of nine DNA microsatellite loci. We found no significant differences in the genetic diversity of, and genetic differentiation between parasite populations before and after repeated treatment, suggesting that PZQ treatment does not have an impact on the neutral evolution of the parasite. This is in stark contrast with a similar study in Tanzania where a significant decrease in genetic diversity was observed in *S. mansoni* miracidia after a single round of PZQ treatment. We argue that PZQ resistance might play a role in our study area, although rapid re-infection cannot be excluded. It is important to monitor this situation carefully and conduct larger field studies with short-term follow-up after treatment. Since PZQ is the only general schistosomicide available, the possibility of PZQ resistance is of great concern both for disease control and for curative use in clinical practice.

[1] Author Contributions. Performed the sampling: TH BW OD DR. Molecular analyses: FVDB. Analyzed the data: FVDB TJ FB TH. Wrote the paper: TH FVDB TJ BW FV FB DR KP

6.1 Introduction

Schistosomiasis or bilharzia is a parasitic disease that mainly occurs in tropical and subtropical regions of the world and is caused by blood flukes of the genus *Schistosoma* (subclass Digenea); over 200 million people are infected, of which more than 90% live in Africa (Steinmann et al., 2006). *Schistosoma* species have a two-host life cycle with an asexual stage within a freshwater snail host and a sexual stage within the definitive mammalian host; parasite eggs are voided in the urine (eg *Schistosoma haematobium*) or faeces (e.g. *S. mansoni*). Despite the availability of adequate tools for diagnosis and treatment, schistosomiasis remains a major public health concern (Savioli et al., 2004). Due to alterations of the environment and increasing migration of man and their livestock, schistosomiasis continues to (re-) emerge. A dramatic example is the outbreak in Northwest Senegal in the early nineties. The Diama dam on the Senegal River was constructed in 1985 to produce fresh water for rice and sugar cane agriculture and water supply for municipal use in Dakar. The subsequent ecological changes favored the spread of freshwater snails, followed by a major outbreak of intestinal schistosomiasis (Talla et al., 1990). Soon after, the restricted urinary schistosomiasis foci of the lower delta spread upstream (Verle et al., 1994), and many children can now be found with both urinary and intestinal schistosomiasis. Praziquantel (PZQ) is the drug of choice to treat schistosomiasis because of the few side effects, the low cost and it is the only drug that is effective against all human schistosome species (Doenhoff et al., 2002). Whereas cure rates for *S. mansoni* usually lie between 70-90% (Gryseels et al., 2006), the observed cure rate at the onset of the Senegalese epidemic reached only 18-32% (Stelma et al., 1995). Such a low figure had never been reported elsewhere before and the emergence of resistance was feared. Several alternative explanations have been put forward related to intense transmission and/or the recent nature of the focus, e.g. rapid re-infection, immunological naivety of the human population, and a high number of immature worms (Gryseels et al., 1994, 2001), which are tolerant to PZQ. A meta-analysis including PZQ treatment studies from various endemic countries showed that cure rates from Senegal were consistently lower than expected, even when initial infection intensity, follow-up time and sensitivity of diagnosis were accounted for (Danso-Appiah and De Vlas, 2002). Laboratory experiments showed that *S. mansoni* isolated from snails in the epicentre of the *S. mansoni* epidemic were significantly less responsive to

PZQ as compared to Kenyan and Puerto Rican strains. They were however fully responsive to the drug oxamniquine (Fallon et al., 1995, 1997), supporting the possibility of PZQ resistance in these Senegalese strains (Fallon et al., 1997). Conclusive evidence for any of the above scenarios has not been obtained so far.

It has been suggested that cure rates may not be a good proxy for drug efficacy against schistosomiasis and soil-transmitted helminths (Gryseels et al., 1994; Montresor, 2011). The standard Kato Katz technique for the diagnosis of *S. mansoni* is not sufficiently sensitive to detect light infections, and cure rates are dependent on baseline/pre-treatment infection intensities. We now have molecular tools to genetically characterize parasite populations. By quantifying neutral genetic variation, we can infer changes in parasite population diversity, size and structuring. Observed variations between pre- and post-treatment populations could in turn be linked to drug pressure and therefore serve as a proxy for intervention efficacy. Here we specifically test with neutral microsatellite markers if and how natural schistosome populations within human hosts change when exposed to repeated PZQ treatments. Genetic diversity of miracidial offspring sampled from each individual host was quantified and used as a proxy for the genetic diversity of the adult worms within that host. We hereby assume that if treatment were effective in eliminating (most of the) adult worms, a significant decrease in genetic diversity of the offspring will be observed.

6.2 Material and Methods

6.2.1 Ethics Statement

This study was part of the EU-FP6 CONTRAST study looking at re-infection rates post-treatment, for which approval was obtained from the ethical committees of the Ministry of Health in Dakar, Senegal, and the NHS-LREC of Imperial College London, England. All parents and teachers gave oral consent for urine and stool examination and the data were anonymized prior to analysis. All schistosomiasis positive children were treated with praziquantel (40mg/kg) throughout the study (Table 6.1) even if they were not included in the study cohort. Treatment of all children in the village was carried out one year and two years after baseline (i.e. S4 and S5 respectively; Table 6.1).

Table 6.1. *Schistosoma mansoni* infection intensity (eggs per gram) of the children enrolled in the study followed through time. Shaded columns indicate from which time points schistosome populations have been sampled and the darker squares the samples genotyped.

	Timing					
Child ID	**S0**	**S1**	**S2**	**S3**	**S4**	**S5**
1	70	0.0	1200.0	220.0	447.0	1332.0
3	26.7	6.7	600.0	7.0	13.0	84.0
9	146.7	0.0	646.7	0.0	187.0	852.0
11	53.3	0.0	673.3	7.0	193.0	660.0
53	246.7	0.0	320.0	20.0	360.0	132.0
65	366.7	0.0	733.3	13.0	13.0	24.0
49	100	0.0	1126.7	27.0	570.0	2698.0
15	400	33.3	247.0	53.0	107.0	12.0
45	146.7	0.0	944.0	160.0	1053.0	660.0
46	1360	6.7	247.0	93.0	547.0	36.0
73	13.3	0.0	420.0	20.0	600.0	480.0
31	20	0.0	N/A			
85	70	0.0	1200	220	447	1332

S0 = baseline survey (survey and double treatment)
S1 = six weeks post-baseline (survey only)
S2 = six months post-baseline (survey and double treatment)
S3 = six weeks post-S2 (survey only)
S4 = one year post-baseline (survey and single treatment)
S5= two years post-baseline (survey and single treatment)

6.2.2 Treatment and data collection

The village of Nder is situated on the western side of Lake Guiers, about 30 km from Richard Toll, and counts about 500 inhabitants. They mainly depend on the lake for their water-related activities. The study started in April 2007, with the collection of urine and stool samples from 107 children aged 5-15 years on three consecutive days, followed by two PZQ treatments three weeks apart (S0; Table 6.1). Follow-up surveys and treatments were conducted on the same cohort of children (see Table 6.1 for treatment and survey regime). The children received a maximum of five treatments in total over a period of 13 months. *Schistosoma mansoni* infections were diagnosed using the Kato Katz technique (Katz et al., 1972) with duplicate thick smears for each stool sample collected on three consecutive days; *S. haematobium* infections by filtration of 10mls of urine sampled on three consecutive days. The initial prevalence was 100% for *S. mansoni* and 97% for *S. haematobium*, with geometric mean infection intensities of 102 eggs per gram (EPG) and 14 eggs/10ml, respectively. For detailed infection data and study design see Webster et al. (2013c).

6.2.3 Parasite collection and molecular analysis

For this molecular study *S. mansoni* eggs were filtered from positive stool samples using a Pitchford and Visser funnel (Pitchford and Visser, 1975), concentrated and hatched in bottled mineral water. Using a binocular microscope individual miracidia were pipetted onto Whatman FTA® classic cards in a volume of 3 µl of water. The cards were allowed to dry and transported to the lab for molecular analysis. A 3.0 mm disc was removed with a Harris Micro Punch from the Whatman cards at the center of which the sample was loaded and the DNA was either purified using the manufacturer's instructions (only the samples from 2007) or eluted and purified using the Nucleospin® Tissue kit (Macherey-Nagel). The latter samples were eluted in 100 µl elution buffer, vacuum dried using a Univapo 150 ECH (Sanyo Biomedical Equipment) and re-suspended in a volume of 20 µl MilliQ H_2O. This latter procedure allows for multiple analyses on a single sample, in contrast to the classical FTA assay where a single FTA punch can only be used once (Van den Broeck et al., 2011). Samples were genetically characterized using a multiplex microsatellite PCR with nine loci (Van den Broeck et al., 2011). Allele sizes were manually verified using GENEMAPPER v4.0 (Applied Biosystems).

6.2.4 Quality control

The software package MICRO-CHECKER v2.2.3 (Van Oosterhout et al., 2004) was used to test for scoring errors, allelic dropouts (i.e. only one of the two alleles present at a heterozygous locus was amplified) and null alleles (i.e. non-amplified allele due to mutation in primer target sequence). Genotyping errors were quantified by re-amplifying at least 10% of all samples (Van den Broeck et al., 2011). Since *SMD43* and *SMD11* appeared to suffer from null alleles, analyses were performed with and without these loci. All other loci had an estimated error rate ranging between 0% for *SMD89* to 4.29% for L46951.

6.2.5 Theoretical expectations

While it has been recognized that gene dynamics in schistosomes is complex due to their indirect life-cycle (Prugnolle et al., 2005b), we here focus on a simplified case, namely the sexually generated offspring from a dioecious population of size *N*, sampled prior to migration. We assume that praziquantel randomly kills worms in the treated population and that there is no immigration in the population, so that N_2 (the size of the reproducing

population after treatment) is a random subset of N_1 (the population before treatment), with $N_2 \leq N_1$ and $\Delta N = N_1 - N_2$. A small N_2 (successful treatment) will result in a correlative reduction of the allelic diversity and expected heterozygosity (Cornuet and Luikart, 1996). While the observed heterozygosity is expected to be rather insensitive to ΔN, this is not the case for the inbreeding coefficient F_{IS} (in the offspring population prior to migration). Under random mating, we expect $F_{IS} = 1/(-2N-1)$ (equation 29 in Balloux (2004), with the number of subpopulations n set to 1). When N_2 is small, the effect of size reduction ΔN is reflected in F_{IS}. Since N_1 and N_2 are effective population sizes, excess variance in reproductive success and deviations from equal sex ratio, which has been described for *S. mansoni* (Webster et al., 1999), will further increase the possibility to detect a bottleneck using F_{IS}. Low effective population sizes will also result in alleles becoming fixed randomly (i.e. genetic drift), causing genetic differentiation (F_{ST}) to increase between parasite populations after treatment, and between pre- and post-treatment populations. Re-infection following treatment could also result in higher F_{ST} values between pre-and post-treatment populations if it occurs from genetically differentiated source populations. We therefore assess the impact of PZQ treatment on the following population statistics: allelic richness (AR), unbiased expected heterozygosity (Hs), inbreeding coefficient (F_{IS}) and genetic differentiation (F_{ST}).

6.2.6 Data analysis

Parasite genetic diversity per host was computed as the unbiased expected heterozygosity (Hs) and the allelic richness (AR) using FSTAT v2.9.3 (Goudet, 1995). Paired t-tests were performed to compare these parameters estimated per locus for each host before (S0) and after treatment (S2) (STATISTICA v9.0). We furthermore tested whether AR, Ho, Hs, F_{ST}, and F_{IS} estimated per host differed between sampling times (i.e. before (S0), six months after (S2) and 2 years after treatment (S5); Table 6.2; two-sided *p*-values were obtained after 2000 permutations). This was done in FSTAT using the option "comparisons among groups of samples" where miracidia from each host were treated as a sample and each sampling time as a group.

Pairwise differentiation between hosts was estimated using pairwise F_{ST} following Weir and Cockerham (Weir and Cockerham, 1984) in FSTAT (4000 permutations). *K*-means clustering coupled with Bayesian Information Criterium (BIC) (Jombart et al., 2010) as implemented in the adegenet package (Jombart, 2008) for R (R Development Core Team, 2013) was used to

study the differentiation among hosts and among surveys. The principal component analysis (PCA) was not scaled, the 60 first PCs and 4 discriminant functions were retained; the proportion of conserved variance was 0.997. Ten independent runs of *K*-means were used. The number of clusters was assessed by means of successive *K*-means clustering with increasing number of clusters. The 'optimal' number of clusters was selected on the basis of the lowest associated BIC (i.e. after which the BIC increases or decreases by a negligible amount).

6.3 Results

After the second round of double treatment (time point S3), *S. mansoni* prevalence and infection intensities remained high (67%; 9.8 EPG) with a cure rate of 34.1%, while *S. haematobium* was fully cleared. Detailed results on *S. haematobium* and *S. mansoni* infection following each treatment have been described by (Webster et al., 2013c). Here we present the data for the subset of children from which *S. mansoni* populations have been genotyped (Table 6.1).

6.3.1 Dataset

Only miracidia with at least six successfully scored loci were included in the analysis, leading to a total of 434 miracidia (91%) divided in 17 samples collected from 12 children at different time points (see Table 6.1). The sample size ranged between 14-45 miracidia per child per time point (mean 26). Nine samples were collected at S0 (214 miracidia; baseline), five at S2 (140 miracidia; six months post-baseline) and three at S5 (80 miracidia; two years post-baseline).

6.3.2 Genetic diversity before and after treatment

We detected no significant effect of treatment on the genetic diversity of *S. mansoni* populations. This was true for all summary statistics investigated (AR, Hs, F_{IS} and F_{ST}; Table 6.2), which showed no significant difference between the three sampling times S0, S2 and S5 (day 0; six months later; two years later). Paired t-tests comparing AR, Hs, F_{IS} per locus and per child separately before (S0) and after treatment (S2) were not significant either. The number of private alleles was higher in populations from S5 (8) than those of S0 (2) and S2 (0).

Table 6.2. Statistical comparison of genetic diversity indices of parasite populations from S0 (baseline), S2 (six months post-baseline), and S5 (two years post-baseline) using FSTAT (Goudet, 2001; two-sided *p* test; 2000 permutations).

Sampling time	AR	Hs	F_{IS}	F_{ST}
S0	2.38	0.40	0.06	0.001
S2	2.47	0.41	-0.001	0.017
S5	2.52	0.43	0.08	-0.007
***p*-value**	0.12	0.06	0.25	0.16

AR: allelic richness. Hs: unbiased expected heterozygosity.
F_{IS}: inbreeding coefficient. F_{ST}: fixation index.

6.3.3 Genetic differentiation before and after treatment

K-means clustering coupled with BIC found 11 distinct clusters in the complete dataset without prior boundary definition (overall F_{ST} between clusters was 0.17); these clusters could not be assigned to the individual children. Each child harbored parasite genotypes from almost all clusters. There was no significant shift in the genetic composition of the parasite population at the host level or survey level before and after treatment as indicated by the chi-square test (all *p* values > 0.01). This is also illustrated in Figure 6.1 showing the assignment of the parasite genotypes collected from child ID 49 to the eleven clusters inferred by *K*-means clustering, for S0, S2 and S5. When comparing all 17 samples (host level) with each other, there were only four pairwise F_{ST} values significant after Bonferroni correction, two between children from the same survey (S0), and two between children from S0 and S2. Parasite populations collected from the same child before and after treatment were never significantly different.

6.4 Discussion

More than two decades after the outbreak of intestinal schistosomiasis in Northwest Senegal, we now have new tools to study the impact of treatment on *S. mansoni* populations. Microsatellite markers allow to study the population genetic structure of schistosomes, and to infer the demographic fluctuations through time. We genotyped parasite populations from twelve children sampled at different time points with nine microsatellite loci. Special care was taken to maximize data quality by means of re-genotyping and detailed quality control (Van den Broeck et al., 2011). Thorough data-analysis demonstrated no significant change in the genetic diversity and structure of

parasite populations after repeated PZQ treatment. Pooling parasites according to child or survey did not influence this outcome.

These results are in stark contrast with those reported by Norton et al. (2010) who compared *S. mansoni* populations from two Tanzanian schools before and after treatment, using seven DNA microsatellite markers (of which six have also been used in this study). They found a significant decrease in genetic diversity six months after a single round of PZQ treatment, and the parasite populations before and after treatment were significantly differentiated. The latter was suggested to be the result of re-infection. A similar reduction in genetic diversity was observed in parasite populations from the untreated pre-school children, demonstrating that PZQ can have a strong and long-lasting effect on *S. mansoni* population structure. So why do we not find a similar impact of PZQ in this study?

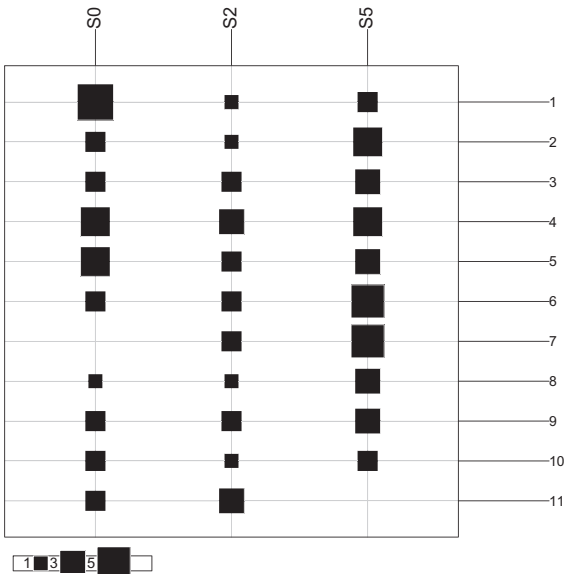


Figure 6.1. Assignment of the parasite genotypes collected from child ID 49 to the eleven clusters inferred by *K*-means clustering, for survey 0, 2 and 5. The sample sizes are represented by black squares.

6.4.1 Re-infection or PZQ resistance?

Drug misuse can be excluded because of the successful elimination of *S. haematobium.* Additionally, the second treatment three weeks later should have eliminated the immature worms that may have survived the first treatment (Renganathan and Cioli, 1998). As such, our results can either be explained by rapid re-infection (intense transmission), or by

resistance of the Senegalese *S. mansoni* strains to PZQ, or a combination of these two. Of note, here we consider a population resistant when it is significantly less responsive to treatment than a fully susceptible population, following Coles (2006).

In the first scenario, re-infection should have been very fast and intense to restore the genetic diversity within six months after the first two treatments. Despite the high transmission in the Tanzanian study, the decrease in genetic diversity was still clearly detectable six months after a single treatment (Norton et al., 2010). *F*-statistics in this study showed that parasite populations from the same child sampled at the start and six months later (S2), were not significantly different from each other. The number of private alleles only increased in the last survey (S5), two years after the start of the study, suggesting that re-infection might be mainly important on a longer timescale. The absence of new alleles in S2, together with the continued high diversity, could suggest that (part of) the parasite population from S2 survived double treatment.

The possibility of PZQ resistant strains in Northwest Senegal has been raised before, as an alternative explanation for the low cure rates at the onset of the epidemic in the early nineties (Stelma et al., 1995; Ernould et al., 1999; Danso-Appiah and De Vlas, 2002). At the time, no conclusion could be reached due to many confounding factors of intense transmission and/or the recent nature of the focus (see Introduction). Today, the epidemiological situation has changed. The infection intensities of *S. mansoni* have decreased considerably, with current figures in Nder about 5-fold lower compared to those in 1996 (Picquet et al., 1996). Snail abundance and snail infection are also much lower (personal data) compared to the year-round high numbers at that time (De Clercq et al., 1999). As such, the above confounding factors are less likely to play a prominent role in the current context.

6.4.2 Additional factors

We observed a drastic drop in egg production six weeks after treatment followed by a rapid increase six months later (Table 6.1). Aside from rapid re-infection, it is possible that this could be explained by a temporary cessation of egg production induced by PZQ (Webster et al., 2013c; Polman et al., 2002). Other factors that can be involved are treatment history and the high number of mixed infections. The village Nder has been involved in several longitudinal studies with mass treatments in e.g. 1996 and 1997 (Picquet et al., 1996, 1998),

and more recent treatments in 2003 and 2006 (our studies). The intense treatment in Northwest Senegal might have imposed a selection for PZQ-resistant parasites.

Mixed infections can lead to direct competition and mating interaction between schistosome species (Southgate et al., 1998). Such interactions have already been documented in Senegal, with ectopic elimination of eggs (Ernould et al., 1999; Huyse et al., 2009; Meurs et al., 2012), and the occurrence of hybrids between human and animal schistosome species (Huyse et al., 2009). Ernould et al. (1999) found cure rates to be much lower in the Senegalese village with mixed infection compared to villages with single infections. The high prevalence of ectopic *S. mansoni* eggs in urine samples (31%) indicated heterologous pairing between *S. mansoni* and *S. haematobium*. Ten months after treatment, *S. haematobium* infection remained low, while *S. mansoni* egg excretion was seven times higher than at the start of the study. Besides the possibility of PZQ resistance, the authors hypothesized that this was due to heterologous pairing: elimination of *S. haematobium* after treatment 'released' the *S. mansoni* females for mating with the unpaired *S. mansoni* males that survived treatment. A similar scenario might occur here, as the number of mixed infections in Nder increased from 23% in 1996 (De Clercq et al., 1999) to 97.2% in the current study (Webster et al., 2013c).

6.4.3 Implications and future perspectives

For almost a decade several African countries have been enrolled in mass treatment initiatives aiming at the broad scale control of morbidity due to schistosomiasis. These campaigns provide annual PZQ treatment of all school children. However, even four treatments in a single year were unable to control *S. mansoni* infections in Nder, suggesting that a single treatment per year might not be sufficient. Comparative studies are needed to find out whether Senegal represents a special, isolated case. A meta-analysis by King et al. (2011) showed that the observed cure rates for *S. mansoni* were higher after two treatments with PZQ compared to a single treatment, but the optimal timing interval for the second treatment remains uncertain. This might depend on local parameters such as the transmission season and the maturation rate of the specific strains. Laboratory experiments showed that the Senegalese isolates have a slower maturation rate than the isolates from Kenya and Puerto Rico (with a prepatent period of up to 10 days longer (Fallon et al., 1997).

It might therefore be an option to administer the second treatment 4-8 weeks rather than 3 weeks after the initial treatment.

In order to better quantify the role of re-infection in the continued high genetic diversity of the *S. mansoni* populations after repeated PZQ treatment, future studies should include a larger study cohort together with shorter follow-up times and in depth snail surveys and cercariae genotyping. Increased genomic coverage will provide a better insight in the impact of PZQ treatment on the genetic make-up of schistosome populations, while it can also identify genomic regions that are potentially under selection.

6.5 Conclusions

We could not find an effect of repeated PZQ treatment on the genetic diversity and population structure of *S. mansoni* in Senegal. Besides the possibility of rapid re-infection, this could suggest that some strains may survive repeated PZQ treatment. More field data from the SRB coupled with in depth molecular studies are needed to confirm the results, together with laboratory experiments to assess the possibility of PZQ resistance in this region. Repeated praziquantel dosing as suggested by King et al. (2011) should be explored.

Acknowledgements

TH is currently a Postdoctoral Fellow of the Research Foundation - Flanders (FWO-Vlaanderen). We would like to thank the Senegalese field team (M. Diop, A. Fall, N. Sy, M. Wade and A. Yague) for their great help with the fieldwork, and the Senegalese children for their enthusiastic participation in this project. We thank one anonymous referee for useful comments on a previous version of this paper. The molecular part of this work was financed by a project grant of Research Foundation - Flanders (G0552.10) and the sample collection was organized through the EU-FP6 project “CONTRAST”: A multidisciplinary alliance to optimize schistosomiasis control and transmission surveillance in sub-Saharan Africa (FP6 STREP contract no: 032203, http://www.eu-contrast.eu). TJ and FB were supported by the European Research Council (ERC) grant 260801-BIG_IDEA; TJ is currently supported by the NIGMS MIDAS initiative and the Medical Research Council. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. .

CHAPTER 7

General discussion

7.1 The evolutionary potential of parasites

Genetic diversity is crucial for both long-term and short-term population dynamics and determining the rate of evolutionary change (Wiese, 2008). Populations harbouring very low levels of genetic variation may be unable to adapt to changing conditions. Maintaining the genetic diversity is thus important to guarantee the evolutionary potential, as the loss of genetic variation can adversely affect a population by reducing individual fitness. In terms of public health however, the evolutionary potential of parasites could present a 'risk'. Parasite populations with a high evolutionary potential are more likely to circumvent the attack of the host immune system or to counteract control methods such as chemotherapy than parasite populations with a low evolutionary potential (McDonald and Linde, 2002). Knowledge of the evolutionary response of parasite populations to new conditions such as novel host resistance genes or vaccines is thus central to many fundamental and applied issues such as (the management of) the spread of drug resistance genes (McDonald and Linde, 2002).

Evolutionary forces that generally promote high levels of diversity are large effective population sizes, gene flow and natural selection, while genetic drift is expected to induce loss of genetic diversity (Hartl and Clark, 2007). Small populations suffer from genetic drift that tends to decrease the evolutionary potential through the random fixation of (possibly deleterious) alleles. In large populations natural selection will have a stronger effect than genetic drift and could promote the evolutionary potential of parasites when advantageous alleles are transmitted to the next generation, leading to the fixation of alleles that are beneficial under local conditions (i.e. local adaptation). Parasites suffering from reductions in population size (bottlenecks or founder effects) are thus less diverse and will adapt slower than parasites that maintain large population sizes. High levels of parasite gene flow will distribute new alleles across populations, increasing the effective population size and resulting in high levels of within-population genetic diversity (Table 7.1). In contrast, the more populations are subdivided, the more parasites will be vulnerable to environmental stochasticity.

Because parasite species show a range of transmission modes and life-history strategies, many factors might promote or hamper their evolutionary potential (McDonald and Linde, 2002; Barrett et al., 2008). For instance, parasites that obligatory infect a single host species are more likely to experience frequent local extinction and recolonisation events, while parasites that are able to infect multiple hosts are unlikely to regularly experience such extreme levels of population stochasticity (Table 7.1). While the former should promote loss of genetic diversity within parasite populations and generate among-population differences (Ericson et al., 1999; Lajeunesse and Forbes, 2002), the latter should maintain comparatively higher levels of genetic variation (Zhan et al., 2003). Another life history trait that may shape the evolutionary potential of parasite populations is the mode of reproduction, where parasites showing sexual reproduction usually exhibit a higher degree of genetic diversity than pathogens that undergo inbreeding or asexual reproduction (Table 7.1). In addition, most parasites that depend on their hosts for long-term survival, as such that the size, spatial structure and distribution of host populations may shape parasite genetic composition (Table 7.1). When hosts are short-lived for instance, parasite populations are more likely to experience regular extinctions and colonisations than when hosts are long-lived and provide a perennial resource. Although these examples apply to differences between parasites species, many of these factors could also be translated to dynamics within a single species.

Table 7.1 Host and parasite life history traits that are likely to shape the genetic structure and sizes of parasite populations (from Barrett et al., 2008).

	Factors that generally increase effective population size	**Factors that generally decrease effective population size**
Host exploitation and specificity	Opportunistic and/or generalist species, multiple host species	Specialised parasite, single host species
Mode of pathogen reproduction	Sexual	Clonal or inbreeding
Pathogen dispersal	Long-distance dispersal	Restricted, local dispersal
Environmental stochasticity	Stable environment and host population dynamics	Frequent population extinction and recolonisation, short-lived hosts
Host longevity; ephemerality of tissues attacked	Perennial or long-lived host	Annual or ephemeral hosts
Host population size and structure	Large, interconnected host populations	Small, fragmented host populations
Epidemiological dynamics	Endemic, systemic	Epidemic, boost and bust

In this thesis we studied levels of genetic variation and its distribution within and among populations of the human parasite *Schistosoma mansoni* in Northwest Senegal. We assessed the effect of the colonization history of the parasite, host-specific factors and treatment on (the partitioning of) levels of parasite genetic diversity. In the following sections we will discuss these results in greater detail and interpret their relevance with respect to the evolutionary potential of *S. mansoni*.

7.2 Parasite colonization history

The demographic and evolutionary dynamics of schistosomes is closely tied to that of freshwater snails that serve as an obligatory intermediate host. The transmission of schistosome parasites is therefore restricted to freshwater habitats where susceptible snail species are present. Schistosome parasites show a high specificity for snail intermediate hosts, as such that most species can only develop successfully in a single snail species (Lockyer et al., 2004). The interaction is called compatible, i.e. the parasite recognises, penetrates and reproduces in the snail, while in incompatible interactions the larval trematode fails to do so and is often destroyed by the snail internal defence system (Lockyer et al., 2004). Such a compatibility does not only apply at the interspecific level, but also at the intraspecific level it has been shown that sympatric schistosome-snail combinations can be more compatible than allopatric combinations (Tchuem Tchuenté et al., 1999; Southgate et al., 2000a), resulting in less snail mortality and morbidity (Gower and Webster, 2005).

The fact that a compatible snail host should be present in order for the parasite to successfully colonize a new region is exemplified by our study system in Northwest Senegal where the construction of two dams resulted in the introduction and spread of the snail *Biomphalaria pfeifferi* and subsequently the invasion of the human parasite *S. mansoni*, until then absent in the region (**chapter 3**). The initial compatibility between snail and schistosome populations at the onset of the epidemic must have been extremely high as malacological field studies demonstrated that the overall *S. mansoni* prevalence in *B. pfeifferi* was 44% while generally 0-10% is found (Diaw et al., 1991). Experimental studies confirmed that 1) Senegalese (sympatric) *S. mansoni* strains were more compatible with Senegalese *B. pfeifferi* snails compared to (allopatric) *S. mansoni* strains from Cameroon (Tchuem Tchuenté et al., 1999; Southgate et al., 2000a), and that 2) schistosome-snail compatibility could evolve very fast within just one or two generations (Webster and

Woolhouse, 1999). The high degree of schistosome-snail compatibility is probably one of the key factors explaining the devastating spread of the *S. mansoni* parasite, the intensity of transmission and the prevalence of infection since its introduction in the Senegal River Basin (Southgate et al., 1998).

Colonisation events such as seen for *S. mansoni* and *B. pfeifferi* in Northwest Senegal are generally expected to result in low genetic diversity within the introduced population because of founder effects, which can lead to high levels of genetic differentiation due to the effects of genetic drift (Vrijenhoek and Graven, 1992; Slatkin, 1995; Cornuet and Luikart, 1996; Kolbe et al., 2004). Such a pattern of low genetic diversity was indeed observed for *B. pfeifferi* snail populations sampled in the region surrounding Richard Toll, suggesting a rapid expansion of a highly fecund snail strain from a single source population (Campbell et al., 2010). Although the lack of *B. pfeifferi* heterozygosity in other settings could be explained by high selfing rates (Charbonnel et al., 2000, 2002; Angers et al., 2003), self-fertilization is not likely to be the principle factor causing loss of genetic variation in the Senegalese *B. pfeifferi* populations (Campbell et al., 2010). In contrast to the low level of genetic diversity found within the snail populations, moderate levels of genetic diversity were found within the *S. mansoni* component populations (**Tables 3.1 and 4.1**). While six of the nine loci were monomorphic in *B. pfeifferi* populations with an overall Hs and Ho of 0.04 and 0.01 resp. (Campbell et al., 2010), all loci were polymorphic in *S. mansoni* populations with an overall Hs and Ho of 0.54 and 0.52 resp. and one locus harboring up to 31 alleles (**chapters 3 and 4**). Such levels of parasite diversity are probably explained by 1) schistosome and human longevity, both allowing the long-term survival of schistosome infrapopulations and 2) the possible introduction of a wide array of strains by infected immigrant workers from neighboring countries. Similarly, a study on the invasion of two digenean trematode parasites and their Asian mud snail host into North America found low levels of diversity in the snail host and one of the trematode species while high levels of diversity were found in the other trematode species (Miura et al., 2006). High genetic diversity within one of the introduced parasite populations were explained 1) by high gene flow between introduced and native populations mediated by birds that can carry more than 1,000 worms and 2) by parasite longevity that allowed the parasite to survive for more than one year (Miura et al., 2006). Both these and our own results therefore show that parasite introductions into new areas are not necessarily followed by population bottlenecks, but that their evolutionary

potential could be sustained within introduced populations, possibly through high parasite gene flow, parasite longevity and/or host longevity.

7.3 Parasite population subdivision and effective population sizes

The spatial scale of parasite gene flow will affect the component population effective sizes, thereby changing the level and distribution of genetic diversity and altering the opportunities for response to selection and adaptive evolution. Parasites that are characterised by high levels of gene flow should have higher within-population diversities than those showing lower levels of gene flow (Barrett et al., 2008). High gene flow will furthermore counteract the effects of genetic drift and homogenise adjacent parasite populations, thereby increasing the spatial area encompassed by a deme (i.e. a local randomly mating population of organisms that share a distinct gene pool). The question of what constitutes a deme has been raised repeatedly for macroparasites (Nadler, 1995; Huyse et al., 2005; Criscione et al., 2011). Although adult schistosomes within infrapopulations may represent the actively mating group of parasites, they may still be randomly distributed among (groups of) infrapopulations. An infrapopulation could therefore represent a single deme when there is low gene flow, but it could also be part of a deme in the case of high gene flow. The amount of parasite dispersal is thus very important for their evolution as the deme constitutes the unit at which selection or drift operates (Nadler, 1995; Criscione et al., 2005). In this thesis we genetically characterised *S. mansoni* parasites sampled from hosts within several villages in Northwest Senegal (**chapters 3, 4 and 5**). Analyses of genetic structure showed very low levels of parasite genetic differentiation between hosts and villages (**Table 3.4 and Table 4.2**). Bayesian inference of their ancestry also revealed that most of the parasites sampled in Northwest Senegal belonged to the same genetic cluster, while parasites sampled in Southeast Senegal and Mali showed an independent ancestry (**Figure 3.5**). These results corroborate previous findings on the genetic structure of *S. mansoni* and *S. haematobium* parasites sampled from six different countries across Africa that showed low genetic differentiation between schistosome samples from the same country but high differentiation between component populations from different countries (Gower et al., 2013). These observations suggest that the spatial area occupied by a schistosome deme could easily encompass vast regions and is mostly not restricted to a single host or village. Some studies showed strong local structure within

human schistosome infrapopulations, indicating that the schistosome deme may be restricted to households (Curtis et al., 2001; Thiele et al., 2008). Such observations remain however rather rare.

A key factor determining parasite gene flow and therefore the genetic boundaries of a deme is host movement, as free-living larval stages in general have low dispersal capabilities (Blouin et al., 1995; McCoy et al., 2003b). Gene flow of parasites with a complex life cycle should hereby closely mimic that of the host with the highest dispersal rate (Criscione and Blouin, 2004; Blasco-Costa et al., 2012). Specifically for schistosomes, this means that parasites are more dependent on the final host for their dispersal than on their snail host (Davies et al., 1999; Jarne and Theron, 2001; Prugnolle et al., 2005c). The strong levels of schistosome genetic structure found across countries (**chapter 3**; Gower et al., 2013) could thus indicate that human host movement is greater within a region/country than across borders and suggest that political boundaries could have a greater impact on schistosome dispersal than absolute distance. Interesting in this respect is that high levels of gene flow were found between *S. mansoni* populations collected along the Kenyan shores of Lake Victoria that stretches about 1800 km^2 (Steinauer et al., 2009), while low levels of gene flow were found between samples collected in the east, west and southwest portions of Kenya that encompass different water bodies (Agola et al., 2006). Similarly, in our study we found low levels of genetic differentiation between samples collected along the shores of Lake Guiers in Senegal, while higher levels of genetic structure were found among different water bodies (i.e. Lampsar River versus Senegal River and Lake Guiers; **chapter 3**). These results suggest that the subdivision of schistosome component populations could be determined not only by political boundaries, but possibly also by water bodies. The latter is most likely coupled to the snail distribution and their population dynamics (see section 7.2).

Although the effective population size (*Ne*) is an important parameter that could provide insight into the ability of the parasite to respond to selection pressures such as drug treatment, estimates of *Ne* for *Schistosoma* and other parasites have so far been poorly studied (Criscione and Blouin, 2005; Criscione et al., 2005). Several methods based on allele frequencies and linkage disequilibrium were applied to estimate *Ne* for schistosome populations, but failed to yield consistent results (Gower et al., 2013). Further theoretical work on *Ne* estimation in macroparasite populations as well as its implementation into

statistical software is therefore necessary. Let's however assume that the prevalence of infection is positively correlated with the effective population size (assuming that the component population is the deme). In this respect it is interesting to note that the highest levels of parasite diversity were found in the village of Pakh that has the lowest *S. mansoni* prevalence (16%) compared to the other two villages (75% in Ndieumeul and 55% in Diokhor; **chapter 4**). Furthermore, we could not find any infected *B. pfeifferi* snail within the four known transmission sites in Pakh over a two-year survey (**Supplementary Table 4.1**), suggesting low intensity of (current) *S. mansoni* transmission within this village. Although these results indicate that the *S. mansoni* component population size in Pakh could be limited because of the relatively low prevalence, we controversially found high levels of parasite diversity compared to other villages that show higher levels of prevalence in humans and snails (**chapter 4**). These results might be explained by the fact that 1) on average we sampled older individuals in Pakh (29) compared to Ndieumeul (14) and Diokhor (23) with host age being positively correlated with parasite heterozygosity (**Figure 4.3**), 2) the action radius of the inhabitants of Pakh is larger compared to Ndieumeul and Diokhor that remain 'isolated' on the Peninsula, 3) hosts in Pakh have recently been infected outside Pakh or at unknown transmission sites, or that 4) they acquired these infections a longer time ago (e.g. *S. mansoni* infection levels in Pakh were much higher in 2006; unpublished data). Irrespective of the underlying explanation, the findings indicate that continuous transmission and high prevalence is not indispensable for maintaining high levels of parasite diversity.

7.4 Parasite reproductive modes and mating behaviour

The mode of reproduction (asexual vs sexual) and mating systems (inbreeding, outcrossing or selfing) will strongly influence the genetic structure and evolutionary potential of populations (Charlesworth, 2003). In general, organisms that undergo sexual recombination are expected to exhibit higher levels of diversity than organisms that undergo inbreeding, selfing or asexual reproduction. Schistosome species show an obligatory alternation of sexual and asexual reproduction each generation, with every reproductive stage expected to affect levels of population diversities in different ways.

Sexual reproduction will result in genetic recombination, increasing the chance for heterozygous offspring. Heterozygous offspring will generally have a higher fitness than

inbred individuals to counter host resistance or other selective pressures. In this thesis miracidia (offspring) were genotyped using nine microsatellite markers. Of all the miracidia that were genotyped, none were genetically identical to another miracidium. Furthermore, no signal of linkage disequilibrium was found (results not shown), suggesting random mating. These results confirm that generation after generation of sexual reproduction between schistosomes is a strong factor promoting the genetic diversity of their offspring. It also shows how the negative effects of asexual reproduction in the snail host are counteracted by sexual reproduction in the final host. In addition to sexual reproduction, sexual selection could increase the genetic benefits of their offspring. An experimental study showed that schistosome females will switch mate for genetically more dissimilar males when the opportunity arrives, suggesting female choice for genetically unrelated males (Beltran et al., 2008). It was also shown that larger males showed a higher reproductive success, which indicates that larger males compete for higher quality females or that females may compete for larger males (Steinauer, 2009). More research in this field is needed to understand what factors drive mate choice and reproduction in schistosomes, and how these dynamics shape their reproductive output and benefit the fitness of their offspring. An interesting hypothesis in this respect is the role of the host immune system in 'promoting' schistosome genetic diversity (Beltran et al., 2011). Such genotype-dependent antigenicity could favour genetic variation through accumulation of genetically unrelated worms during the lifetime of the host. The positive correlation between parasite heterozygosity and host age that was found in this thesis (**chapter 4**) could support this hypothesis.

While sexual reproduction is expected to increase levels of schistosome diversity, the asexual phase could counteract this. Asexual multiplication could lead to the synchronous/clumped transmission of thousands of genetically identical parasites to the same infrapopulation, thereby resulting in local scale genetic structure (as for example found in *Fascioloides magna*, Mulvey et al., 1991) or an increase in selfing rates (as for example in *Plasmodium falciparum*, Anderson et al., 2000). Likewise, if a definitive host deposits parasites' offspring into a given area, it is likely that siblings are co-transmitted to a given infrapopulation, leading to an increased chance for biparental inbreeding (Anderson et al., 1995; Nadler, 1995). These effects will be much stronger when the effective population sizes are small (Anderson et al., 1995). It has therefore been hypothesized that some

trematodes keep a second intermediate host in their life cycle that collects different cercarial genotypes over time before ingestion in a definitive host in order to avoid inbreeding mechanisms (Rauch et al., 2005). In our study no evidence was found for clumped or sib transmission. As explained in detail in **chapter 4**, local scale genetic structure was absent, indicating that the impact of clonal amplification in our study area is negligible. Also, the relatedness of cercariae within the transmission site of Ndieumeul did not deviate from its expected distribution, indicating that sib transmission does not occur (at least not within the transmission site of Ndieumeul). Similar results were found in a natural lake in Kenya where the relatedness was investigated between *S. mansoni* strains within the same snail host (Steinauer et al., 2009). Results showed that *S. mansoni* parasites were not more or less related than expected, suggesting the absence of sib transmission (Steinauer et al., 2009). This and our study therefore show that snails within the studied areas acquire multiple infections by unrelated schistosomes (that are potentially transmitted together to the same final host), which could lead to increased outcrossing rates within the final host (Minchella et al., 1995; Eppert et al., 2002). High gene flow probably explains the absence of clumped or sib transmission, resulting in the mixing of larvae before or after asexual reproduction (**chapter 4**) (Prugnolle et al., 2005a). In conclusion, the generally high levels of schistosome diversity found within component populations with low genetic structure suggest that their evolutionary potential is not hampered by clonal amplification.

7.5 Host-specific factors

In section 7.3 we discussed the importance of human host dispersal in determining both schistosome dispersal and the boundaries of a parasite deme. Besides host dispersal, heterogeneities in host resistance or host-specific factors (age, gender, ethnic group, social status, residence, etc.) could potentially result in differential parasite recruitment, thereby affecting the genetic composition of schistosome infrapopulations. In this thesis we were able to address this issue for human hosts by incorporating human host age, gender, residence and ethnic group in our population genetic analyses of *S. mansoni*. In contrast to other studies on human schistosomes (Thiele et al., 2008; Gower et al., 2011; Barbosa et al., 2013), this was the only study that demonstrated possible effects of such host-specific factors on levels of schistosome genetic diversity. More specifically we showed that levels of parasite diversity were higher within older hosts, which could be explained by differential

water contact behaviour, by genotype-dependent concomitant immunity or by the fact that older hosts have acquired more infections over their lifetime (**chapter 4**). There are several possible reasons why the other studies could not find such an association. The study of Thiele and colleagues (2008) based their population genetic analyses on *S. mansoni* adult worms that were obtained after laboratory passage of naturally collected miracidia, which could potentially break down the association with human host-specific factors. The study of Gower and colleagues (2011) used *S. haematobium* samples collected from school-aged children only, such that the age-range was probably too narrow. Finally, the Barbosa et al. (2013) study used a pooled approach by genotyping filtered stool samples (instead of individual parasite samples). A pooled design does not allow the estimation of observed heterozygosity and the inbreeding coefficient, which were the two estimates that were found to be significantly associated with host age in our study (**Figure 4.3**). Our study used genotypes from individual parasites collected from 57 host individuals ranging in age between 4 and 50 years. However, future studies should increase the number of human hosts and the number of genotyped miracidia per host in order to make the statistical analyses more robust. Meanwhile, it is clear that schistosomes benefit at the evolutionary level from infecting humans because it allows schistosome infrapopulations to increase their diversity in time through the accumulation of genetically diverse infections.

7.6 Drug treatment

So far we have discussed natural processes that are inherently linked to the biology of the parasite and the host and that could influence levels of genetic diversity within and among infra- and component populations. An anthropogenic factor, and probably the most important factor that could compromise levels of diversity within schistosome infrapopulations is drug treatment as it induces massive population bottlenecks. Experimental infections in mice showed that levels of *S. mansoni* genetic diversity indeed decreased after seven rounds of PZQ treatment while inbreeding increased, clearly suggesting that selection for decreased PZQ susceptibility selects for genetically less diverse parasites (Coeli et al., 2013). While the study summarizes the possible effects of treatment on levels of genetic diversity and structure under controlled conditions (i.e. without gene flow from refugia), less is understood for natural settings. Therefore, in **chapter 5** we created a theoretical framework to obtain insight into what could be expected from

community-based drug treatment based on simulations using an island model at equilibrium. Results of these simulations showed that a prolonged or sustained treatment regime with at least one treatment per year and high coverage would be needed to substantially decrease the genetic diversity of schistosome component populations. High pre-treatment infection intensities, low effectiveness (e.g. due to the presence of immature worms) and a low frequency of treatment will all corroborate the severity of the genetic bottleneck after treatment. The coverage (i.e. the relative number of hosts treated) was the only factor affecting levels of component population genetic diversity, which would only decrease when all hosts were treated within a given community. In **chapter 6** we assessed the impact of treatment on schistosome infrapopulations within a natural setting by genotyping parasites that were collected before and after (repeated) treatment. Although the sampling range was limited (i.e. 12 children within a single village), results were very clear in that the infrapopulation diversities did not decrease, not even after two rounds of treatment within six months. Possible factors such as drug resistance, high pre-treatment worm burden, the presence of immature worms and even the possible effects of interspecific competition were put forward to explain the maintenance of genetic diversity after treatment (**chapter 6**). Similar results were obtained from field data in Brazil where treatment had no effect on levels of genetic diversity (Blanton et al., 2011). However, contrasting results were found in two villages (Kisorya and Bukindo) in Tanzania where a significant decrease in genetic diversity of infrapopulations was found one year after a single round of drug treatment (Norton et al., 2010). The coverage within one of the districts was as high at 90%, possibly explaining why a concomitant decrease in diversity was observed. There are however several findings in their study that raise questions. First, while in most national control programs a clear reduction in prevalence and intensity of schistosome infections is observed (Fenwick et al., 2009), there were no changes in prevalence or intensity in the Kisorya villages. In the Bukindo village, a decrease in infection intensity was observed, but an increase in prevalence. The authors explained this by high levels of transmission and re-infection that is typically observed around Lake Victoria (Norton et al., 2010). Based on our simulations and findings in Senegal however, it is hard to believe that under such conditions a reduction in genetic diversity would be observed one year after a single round of treatment. Second, parasites were also collected before and after treatment from a group of seven-year old children that were not treated because they were not of

school age during baseline sampling and therefore act as a control group for non-chemotherapy induced changes (Norton et al., 2010). Surprisingly, genetic diversity also decreased in this control group. The authors argued that community-based drug treatment had a regional effect where treatment within one cohort also decreased levels of diversity in the second cohort where no treatment was administered (Norton et al., 2010). However, a decrease in genetic diversity in the control group rather suggests that another factor than treatment explains the observed decrease in levels of *S. mansoni* diversity in that region. As the few studies to date yield conflicting insights, clearly more field-based studies are needed to understand the impact of treatment on schistosome populations.

The observation that schistosome populations do not appear strongly affected by treatment suggests that the selective pressure imposed by current treatment practices may not be sufficiently large to select for drug resistance, or at least not large enough to promote its spread. The only assumption here is that there is a cost to resistance that hampers the spread of resistant alleles through a population. This means that only when selection pressure by drug treatment is sufficiently large, will resistant worms show a higher fitness than sensitive worms, despite the cost that accompanies resistance. An observation that could support this hypothesis is the fact that resistance to the drug oxamniquine remained restricted to sporadic foci in Brazil without any apparent spread throughout the human population (Cioli et al., 1993; Secor and Colley, 2005). A cost to resistance has been experimentally described for oxamniquine resistance with resistant worms being less viable than their sensitive counterparts at all life stages (Cioli et al., 1992). The few experimental studies performed on PZQ-resistant worms suggest contrasting results depending on the life stage. On one hand, PZQ-resistant schistosomes exhibited reduced cercarial output compared with control strains (Liang et al., 2001b; William et al., 2001). One experimental study also encountered difficulties to maintain PZQ-resistant strains as only female worms were recovered after 11 treatments (Coeli et al., 2013). On the other hand, PZQ-resistant isolates from Senegal showed a higher infectivity of cercariae to snails, a longer prepatent period within the snails, a higher longevity of snails infected with PZQ-R isolates and significantly more eggs were found within the faeces and tissues of mice infected with PZQ-resistant isolates (Liang et al., 2001b). Although these observations should be interpreted with care and much more research is needed, they suggest that a fitness cost of resistance at a certain point in the life cycle might be counterbalanced by fitness benefits elsewhere in

the cycle (Vanaerschot et al., 2013). Such tradeoffs have also been described in bacteria, although more at a molecular level (Andersson and Hughes, 2011). The cost of resistance in schistosome parasites in both the vertebrate host and invertebrate vector could hamper the spread of resistance when selective pressures are not high (Vanaerschot et al., 2013), and could explain why praziquantel drug resistance has not been reported yet in China despite 20 years of intense treatment (Liang et al., 2001a).

7.7 Evolutionary potential of *Schistosoma mansoni*

Although the scale of genetic differentiation differed between studies (reviewed in Steinauer et al., 2010), *S. mansoni* diversity is randomly distributed among hosts and villages in most settings (**chapters 3 and 4**; Gower et al., 2013). This means that individual infrapopulations harbor almost all the genetic variation that is present within the component population. Throughout the thesis we have studied several factors (colonization history, host-specific factors and treatment) that could influence the genetic composition of *S. mansoni* populations and finally discussed how these parasite- and host-related factors promote the diversity of *S. mansoni* infra- or component populations, rather than compromising them. Besides the factors discussed in this thesis, there are possibly many more factors such as interspecific hybridization that could potentially increase levels of diversity because it unites divergent genomes (e.g. Collado-Romero et al., 2010). Results of this thesis and other studies have shown that levels of genetic diversity of *S. mansoni* are relatively high and are similar to the levels of diversity seen within free-living organisms (Figure 7.1). If we compare the genetic diversity of the two most important schistosome species infecting man in Africa, we can see that the mean expected heterozygosity of *S. mansoni* (Hs = 0.70) is larger than the one of *S. haematobium* (Hs = 0.54) (Figure 7.1). What is even more striking is that the mean expected heterozygosity in *S. mansoni* is equal to the one of *Homo sapiens* (Hs = 0.69), which is regarded as the most successful free-living species on Earth. Although comparing levels of genetic diversity between species that were typed at different genetic markers is cumbersome, they suggest that *S. mansoni* harbors a considerable amount of genetic diversity.

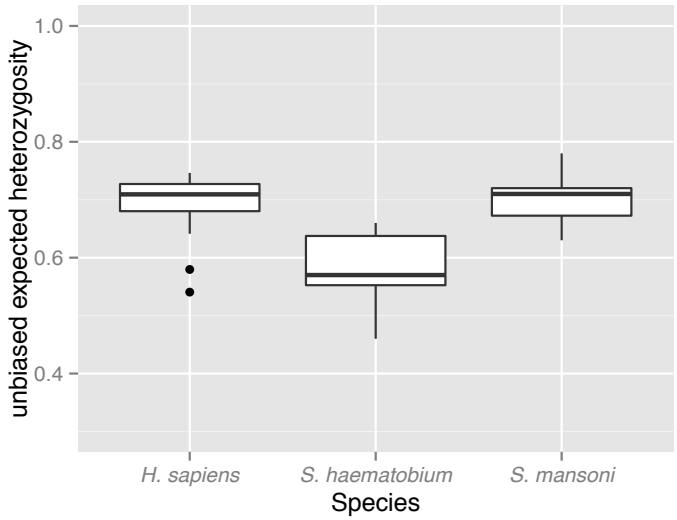


Figure 7.1 Boxplots summarizing estimates of unbiased expected heterozygosity for *Homo sapiens* and the two major human schistosome species in Africa, *Schistosoma haematobium* and *Schistosoma mansoni.* Estimates of expected heterozygosity for *Homo sapiens* were calculated per country (53 in total from all continents) based on 783 microsatellite markers (data from Ramachandran et al., 2005). Estimates of expected heterozygosity for *S. haematobium* and *S. mansoni* were calculated per country (5 in total from Africa, from which 4 shared between the two species) based on 7 microsatellite markers for each of the two species (data from Gower et al., 2013).

These findings suggest that the evolutionary potential of *S. mansoni* in response to selective pressure should not be underestimated. A small number of studies using naturally collected data showed that treatment, possibly the strongest selection factor, had only little or no effect on levels of diversity, confirming that the evolutionary potential of *S. mansoni* is huge. Large population sizes are furthermore a strong driving force for adaptation and high levels of genetic diversity present a genetic basis for adaptation in response to selective pressures (Conover et al., 2006). This could have serious epidemiological complications such as the quick development of adaptive traits such as virulence, higher fecundity and drug resistance. Experimental studies have indeed shown that selection can rapidly change the infectivity and virulence phenotypes of schistosomes (Davies et al., 2001; Webster et al., 2004, 2007). High levels of genetic variation could furthermore complicate the design of drugs and vaccines and tackle the success of control programs (Cupit et al., 2011).

7.8 The way forward in a genomic era

This thesis highlights the need for further research. Studies on neutral genetic markers should be extended 1) with a larger coverage of the genome and 2) with adaptive genetic markers. There is a growing body of evidence suggesting that patterns of variation and divergence in adaptive traits are not well reflected by neutral markers (Pfrender et al., 2000; Gomez-Mestre and Tejedo, 2004). The power and resolution of the microsatellites used in this study should be increased towards the use of genome wide Single Nucleotide Polymorphism (SNP) markers. Although individual SNP markers may be less informative than individual microsatellite markers, Next-Generation Sequencing (NGS) technology enables the generation of a high number of SNPs across the genome (Luikart et al., 2003). Library preparations for NGS purposes could aim to capture genetic diversity across the whole genome or could be restricted to the exome (Majewski et al., 2011) or randomly derived SNPs through Genotyping-By-Sequencing approaches (Narum et al., 2013). When such genome-wide SNP markers are used in a population genetic framework (population genomics), they offer an extremely powerful tool to study both neutral demographic processes and adaptive divergence. The two main principles of population genomics are that neutral loci across the genome will be similarly affected by demography and the evolutionary history of populations, and that loci under selection will often behave differently and therefore reveal 'outlier' patterns of variation. Such outlier patterns can be identified through selective sweep mapping (i.e. detecting selection based on the concept of genetic hitchhiking) or trait mapping (i.e. finding loci underlying phenotypic traits) (Ellegren, 2014). These advances will reveal important new insights into the transmission dynamics of *S. mansoni* as well as reveal genes underlying important adaptive traits. Understanding which epidemiological relevant traits and how fast they evolve or spread in natural populations of schistosomes, whether or not in response to artificial selection, will ultimately help the control of schistosomiasis.

7.9 Implications for control of schistosomiasis

In light of the recent technological and theoretical progress that has revolutionized genetics, there is a need to assess the ways in which genetic research (often restricted to the academic world) may be embedded in control programs of schistosomiasis. Current programs are focused on the large-scale administration of praziquantel to school-age

children (Fenwick et al., 2009; Rollinson et al., 2013), especially in sub-Saharan Africa where the disease continues to be a public health problem. The main aim of these school-based treatments is to control morbidity, and lower the burden of schistosomiasis in part of the human population. Putting the short-term benefits of these morbidity control measures aside, treatment with praziquantel does not prevent reinfection. The question therefore arises how treatment campaigns can be optimized in order to have a long-term effect on the transmission of the disease. The results of this thesis advocate the use of genetics as a tool for optimizing control programs.

First, higher levels of *S. mansoni* genetic diversity were found within older human hosts than within children (**chapter 4**). Control programs that perform school-based treatments will thus miss the older hosts that harbor the genetically most diverse infections, which will result in a sustained level of genetic diversity within the parasite component population after treatment. In this respect it would be desirable to improve the coverage of treatment and shift from a school-based treatment towards a community-based treatment that includes older hosts as well as children that are not enrolled in school.

Second, only maximum coverage will reduce schistosome genetic diversity at component population level (**chapter 5**). High coverage remains one of the greatest challenges of today, especially because of the slow socio-economic development in countries in sub-Saharan Africa (Fenwick et al., 2009). Despite huge efforts to scale up schistosomiasis treatment by the Schistosomiasis Control Inititiative, the United States Agency for International Development, the Department for International Development and Merck Serono which donates 250 million praziquantel tablets a year (WHO, 2012), the number of people treated in 2009 in 21 (out of 76) endemic countries was only 8.2% of the estimated number of people infected with schistosomes (WHO, 2011; Rollinson et al., 2013). The national treatment coverage in 2010 in African countries ranged between 0.01% in Sudan to a maximum of 42% in Mali (Rollinson et al., 2013). Coverage is therefore far too low to compromise the genetic diversity of schistosome component populations. It could therefore be much more cost-effective to decentralize large-scale treatment campaigns into smaller geographic units that allow independent management (i.e. biologically relevant management units, a term borrowed from the fisheries management; Waples & Naish, 2009). These units could be determined with genetics that reveal the distribution of

schistosome genetic variation and identify the genetic boundaries of parasite component populations. Gower and colleagues (2013) showed that parasites clustered according to country, but also within countries there is evidence of sub-structuring (e.g. according to water body; section 7.3). In Senegal for instance, parasites from Northwest Senegal were genetically highly differentiated from those from Southeast Senegal (region of Kédougou), suggesting very little gene flow between them (**chapter 3**). These two regions should therefore be tackled independently. The ultimate goal is to obtain a high-resolution map of biologically relevant management units that allow a country to organize its treatment campaigns according to smaller geographic entities, which is logistically and financially more efficient than organizing nation-wide control programs.

Glossary

Allele[1]: an alternative form of a gene. Each allele represents a DNA sequence with slight differences from each other.

Asexual reproduction[1]: reproduction that does not involve formation and fusion of gametes and results in progeny with an identical genetic consitution to the parent and to each other.

Bottleneck[1]: sudden decrease in population density with a resulting decrease in genetic variability within a population.

Component population[3]: all of the individuals of a specified life history phase at a particular place and time; all of the infrapopulations in a single host species in an ecosystem.

Concomitant immunity[3]: resistance to re-infection of a host by a specific parasite when the host is currently infected with that parasite.

Definitive host[3]: that host in a parasite's life cycle in which the parasite reaches sexual maturity.

Deme[2]: breeding group or subpopulation.

Effective population size (*Ne*)[2]: the number of individuals in an idealized, randomly mating population with an equal sex ratio that would exhibit the same rate of heterozygosity loss over time as an actual population with a particular census (total adult number) size.

Epidemiology[1]: the study of the occurrence of infectious diseases, their origins and pattern of spread through a population.

Epidemic[1]: an outbreak of epidemic disease; affecting a large number of individuals at the same time.

Expansion[1]: sudden increase in population density with a resulting increase in genetic variability within a population.

Founder effect[2]: the genetic consequences of starting a new population with a small number of individuals, and thus only a subsample of the genetic diversity present in the original population. Comparable with the dominating effects of genetic drift in small populations.

[1] Lawrence (2005), [2] Beebee and Rowe (2008) [3] Goater et al. (2013)

***F*-statistics[2]:** statistics designed to estimate the partitioning of heterozygosity among individuals, subpopulations and full populations. Widely used to quantify genetic differentation among subpopulations.

GenBank[1]: the main publicly accessible electronic database of DNA sequences, containing many millions of sequences.

Gene[2]: a defined sequence of DNA that is transcribed into RNA.

Genetic[1]: anything involving, caused by, or pertaining to genes.

Gene flow[1]: the spread of particular alleles within and between populations, typically resulting from the dispersal of individuals.

Genetic distance[1]: a measure of the difference between two DNAs from different species, which is used in the construction of phylogenetic trees. In its crudest form it is the percentage of nucleotide differences between the two DNAs.

Genetic diversity[2]: the amount of genetic variation present in a population, often quantified as expected heterozygosity or allelic richness (codominant markers), and gene or haplotype diversity (dominant or haploid markers).

Genetic drift[1]: random changes in allele frequency in small isolated populations owing to factors other than natural selection, such as sampling of only a small number of gametes in each generation, *alternative* Sewall Wright effect.

Genetic marker[1]: a gene or other piece of DNA whose properties, and sometimes position on the chromosome, are known and which may be used to identify particular cells or organisms, or as a point of reference in a genetic mapping experiment.

Genotype[2]: the genetic constitution of an organism at one, many or all genetic loci.

Genotyping[1]: the determination of the detailed genetic make-up, or genotoype, of an individual, usually in respect to particular genes or sets of genes.

Haploid[2]: cell or organism containing a single copy of each genetic locus.

Hardy-Weinberg Equilibrium[2]: the proportions of homozygotes and heterozygotes expected in a large, randomly mating (panmictic) population when overall allele frequencies are known. Assuming no migration, mutation or selection the Hardy-Weinberg law states that allele frequencies should remain unchanged from generation to generation.

Haplotype: a unique sequence.

[1] Lawrence (2005), [2] Beebee and Rowe (2008) [3] Goater et al. (2013)

Helminth[1]**:** parasitic flatworm (fluke and tapeworm) or roundworm.

Heterozygote[2]**:** diploid individual with two different alleles at a genetic locus.

Heterozygosity[2]**:** the observed or expected (under Hardy-Weinberg equilibrium) proportion of heterozygotes in a population.

Inbreeding[1]**:** matings between related individuals.

Inbreeding coefficient[1]**:** in population genetics, a measure of the reduction in heterozygosity as a result of inbreeding.

Infrapopulation[3]**:** all of the parasites of a single species within a single host.

Intermediate host[3]**:** that host in a parasite's life cycle required by the parasite to complete its life cycle, and in which some morphological change or development occurs.

Linkage equilibrium[2]**:** situation, promoted by recombination, in which genetic loci segregate independently of one another during reproduction. Disequilibrium occurs when alleles at two loci segregate together.

Locus (genetic)[2]**:** a defined sequence of DNA on a chromosome. May or may not be a gene.

Macroparasite[3]**:** a parasite that is usually visible with the naked eye, e.g. helminth, arthropods.

Metapopulation[2]**:** population subdivided into multiple demes at least some of which occasionally go extinct and are subsequently repopulated by immigrants from other demes.

Microsatellite[2]**:** genetic locus with a simple sequence (usually, di, tri, or tetranucleotide) repeated multiple times.

Mutation[2]**:** alteration of the nucleotide sequence in DNA.

Network[2]**:** a diagram of (usually) haplotype relationships that can show multiple possible mutation pathways between them.

Panmictic[2]**:** a population with random mating among all the individuals present (=panmixia).

Parasite[1]**:** organism that for all or some part of its life derives its food from a living organism or another species (the host). It usually lives in or on the body or cells of the host, which is usually harmed to some extent by the association.

Parthenogenesis[1]**:** reproduction from a female gamete without fertilization by a male gamete.

[1] Lawrence (2005), [2] Beebee and Rowe (2008) [3] Goater et al. (2013)

PCR[2]: polymerase chain reaction, a method for amplifying large quantities of a DNA sequence using oligonucleotide primers and a thermostable DNA polymerase.

Population genetics[1]: the study of how genetic principles apply to groups of interbreeding individuals (a population) as a whole.

Primer[2]: short oligonucleotide (typically 15-25 nucleotides long) complementary to a DNA sequence and which can be used in PCR amplifications.

Refugium[1]: an area or population that has remained unaffected by external influences.

Sexual reproduction[1]: reproduction involving the formation and fusion of two different kinds of gametes to form a zygote, usually resulting in progeny with a somewhat different genetic constitution from either parental type and from each other.

SNP[2]: single nucleotide polymorphism, the occurrence of alleles with different nucleotide bases at a specific point in a DNA sequence.

Strain: an individual representing a unique genetic variant.

Suprapopulation[1]: population that includes all the developmental phases of a species at a given time.

Symbiosis[1]: close and usually obligatory association of two organisms of different species living together, not necessarily to their mutual benefit.

Wahlund effect[2]: reduction in heterozygosity relative to Hardy-Weinberg expectations that arises when genotype data from multiple subpopulations are pooled for analysis.

[1] Lawrence (2005), [2] Beebee and Rowe (2008) [3] Goater et al. (2013)

References

Agatsuma, T., 2003. Origin and evolution of *Schistosoma japonicum*. Parasitol. Int. 52, 335–340.

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

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

Anderson, R.M., May, R.M., 1985. Helminth infections of humans: mathematical models, population dynamics, and control. Adv. Parasitol. 24, 1–101.

Anderson, T.J., Romero-Abal, M.E., Jaenike, J., 1995. Mitochondrial DNA and *Ascaris* microepidemiology: the composition of parasite populations from individual hosts, families and villages. Parasitology 110, 221–229.

Anderson, T.J., Haubold, B., Williams, J.T., Estrada-Franco, J.G., Richardson, L., Mollinedo, R., Bockarie, M., Mokili, J., Mharakurwa, S., French, N., Whitworth, J., Velez, I.D., Brockman, A.H., Nosten, F., Ferreira, M.U., Day, K.P., 2000. Microsatellite markers reveal a spectrum of population structures in the malaria parasite *Plasmodium falciparum*. Mol. Biol. Evol. 17, 1467–1482.

Andersson, D.I., Hughes, D., 2011. Persistence of antibiotic resistance in bacterial populations. FEMS Microbiol. Rev. 35, 901–911.

Angelucci, F., Basso, A., Bellelli, A., Brunori, M., Pica Mattoccia, L., Valle, C., 2007. The anti-schistosomal drug praziquantel is an adenosine antagonist. Parasitology 134, 1215–1221.

Angers, B., Charbonnel, N., Galtier, N., Jarne, P., 2003. The influence of demography, population structure and selection on molecular diversity in the selfing freshwater snail *Biomphalaria pfeifferi*. Genet. Res. 81, 193–204.

Antao, T., Lopes, A., Lopes, R.J., Beja-Pereira, A., Luikart, G., 2008. LOSITAN: a workbench to detect molecular adaptation based on a F_{ST}-outlier method. BMC Bioinformatics 9, 323.

Archie, E.A., Luikart, G., Ezenwa, V.O., 2009. Infecting epidemiology with genetics: a new frontier in disease ecology. Trends Ecol. Evol. 24, 21–30.

Armstrong, J., 1965. Mating behaviour and development of schistosomes in the mouse. J. Parasitol. 51, 605–616.

Ashford, R.W., Crewe, W., 1998. The parasites of *Homo sapiens*. Liverpool School of Tropical Medicine, Liverpool, UK.

Avise, J.C., 2004. Molecular Markers, Natural History and Evolution, 2nd Edition. Sinauer Associates Inc., Sunderland, USA.

Balloux, F., Lugon-Moulin, N., 2002. The estimation of population differentiation with microsatellite markers. Mol. Ecol. 11, 155–165.

Balloux, F., 2004. Heterozygote excess in small populations and the heterozygote-excess effective population size. Evolution 58, 1891–1900.

Barbosa, L.M., Silva, L.K., Reis, E.A., Azevedo, T.M., Costa, J.M., Blank, W.A., Reis, M.G., Blanton, R.E., 2013. Characteristics of the human host have little influence on which *Schistosoma mansoni* population are acquired. PLoS Negl. Trop. Dis. e2572.

Barker, S.C., Blair, D., 1996. Molecular phylogeny of *Schistosoma* species supports traditional groupings within the genus. J. Parasitol. 82, 292–298.

Barrett, L.G., Thrall, P.H., Burdon, J.J., Linde, C.C., 2008. Life history determines genetic structure and evolutionary potential of host-parasite interactions. Trends Ecol. Evol. 23, 678–685.

Basch, P.F., Basch, N., 1984. Intergeneric reproductive stimulation and parthenogenesis in *Schistosoma mansoni.* Parasitology 89, 369–376.

Beebee, T., Rowe, G., 2008. An introduction to molecular ecology. 2nd Edition. Oxford University Press, NY, USA.

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N., Bonhomme, F., 1996-2004 GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France).

Belkhir, K., Castric, V., Bonhomme, F., 2002. IDENTIX, a software to test for relatedness in a population using permutation methods. Mol. Ecol. Notes 2, 611–614.

Beltran, S., Cezilly, F., Boissier, J., 2008. Genetic dissimilarity between mates, but not male heterozygosity, influences divorce in schistosomes. PLoS One 3, e3328.

Beltran, S., Gourbal, B., Boissier, J., Duval, D., Kieffer-Jaquinod, S., Pierce, R.J., Grunau, C., Theron, A., Mitta, G., 2011. Vertebrate host protective immunity drives genetic diversity and antigenic polymorphism in *Schistosoma mansoni*. J. Evol. Biol. 24, 554–572.

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

Blanton, R.E., Blank, W.A., Costa, J.M., Carmo, T.M., Reis, E.A., Silva, L.K., Barbosa, L.M., Test, M.R., Reis, M.G. 2011. *Schistosoma mansoni* population structure and persistence after praziquantel treatment in two villages of Bahia, Brazil. Int. J. Parasitol. 41 (10): 1093-9

Blasco-Costa, I., Waters, J.M., Poulin, R., 2012. Swimming against the current: genetic structure, host mobility and the drift paradox in trematode parasites. Mol. Ecol. 21, 207–217.

Blouin, M.S., Yowell, C.A., Courtney, C.H., Dame, J.B., 1995. Host movement and the genetic structure of populations of parasitic nematodes. Genetics 141, 1007–1014.

Bonhomme, M., Blancher, A., Cuartero, S., Chikhi, L., Crouau-Roy, B., 2008. Origin and number of founders in an introduced insular primate: estimation from nuclear genetic data. Mol. Ecol. 17, 1009–1019.

Bonin, A., Bellemain, E., Eidesen, P.B., Pompanon, F., Brochmann, C., Taberlet, P., 2004. How to track and assess genotyping errors in population genetics studies. Mol. Ecol. 13, 3261–3273.

Botros, S., Sayed, H., Amer, N., El-Ghannam, M., Bennett, J.L., Day, T.A., 2005. Current status of sensitivity to praziquantel in a focus of potential drug resistance in Egypt. Int. J. Parasitol. 35, 787–791.

Bowles, J., Blair, D., McManus, D.P., 1992. Genetic-Variants within the Genus *Echinococcus* Identified by Mitochondrial DNA Sequencing. Mol. Biochem. Parasitol. 54, 165–174.

Brown, S.P., Grenfell, B.T., 2001. An unlikely partnership: parasites, concomitant immunity and host defence. Proc. R. Soc. London Ser. B-Biological Sci. 268, 2543–2549.

Bush, A.O., Lafferty, K.D., Lotz, J.M., Shostak, A.W., 1997. Parasitology meets ecology on its own terms: Margolis et al revisited. J. Parasitol. 83, 575–583.

Bush, A.O., Fernandez, J. c, Esch, G.W., Seed, J.R., 2001. Parasitism: The diversity and ecology of animal parasites. Cambridge University Press, Cambridge, UK.

Caillaud, D., Prugnolle, F., Durand, P., Theron, A., de Meeus, T., 2006. Host sex and parasite genetic diversity. Microbes Infect. 8, 2477–2483.

Campbell, G., Noble, L.R., Rollinson, D., Southgate, V.R., Webster, J.P., Jones, C.S., 2010. Low genetic diversity in a snail intermediate host (*Biomphalaria pfeifferi* Krass, 1848) and schistosomiasis transmission in the Senegal River Basin. Mol. Ecol. 19, 241–256.

Castric, V., Bernatchez, L., Belkhir, K., Bonhomme, F., 2002. Heterozygote deficiencies in small lacustrine populations of brook charr *Salvelinus Fontinalis* Mitchill (Pisces, Salmonidae): a test of alternative hypotheses. Heredity. 89, 27–35.

Chaine, J.P., Malek, E.A., 1983. Urinary schistosomiasis in the Sahelian region of the Senegal River Basin. Trop. Geogr. Med. 35, 249–256.

Charbonnel, N., Angers, B., Razatavonjizay, R., Bremond, P., Jarne, P., 2000. Microsatellite variation in the freshwater snail *Biomphalaria pfeifferi*. Mol. Ecol. 35, 249–256.

Charbonnel, N., Angers, B., Rasatavonjizay, R., Bremond, P., Debain, C., Jarne, P., 2002. The influence of mating system, demography, parasites and colonization on the population structure of *Biomphalaria pfeifferi* in Madagascar. Mol. Ecol. 11, 2213–28.

Charlesworth, D., 2003. Effects of inbreeding on the genetic diversity of populations. Philos. Trans. R. Soc. London Ser. B-Biological Sci. 358, 1051–1070.

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

Cioli, D., Pica-Mattoccia, L., Moroni, R., 1992. *Schistosoma mansoni*: hycanthone/oxamniquine resistance is controlled by a single autosomal recessive gene. Exp. Parasitol. 75, 425–432.

Cioli, D., Pica-Mattoccia, L., Archer, S., 1993. Drug resistance in schistosomes. Parasitol. Today 9, 162–166.

Cioli, D., Pica-Mattoccia, L., Archer, S., 1995. Antischistosomal drugs: past, present ... and future? Pharmacol. Ther. 68, 35–85.

Cioli, D., 2000. Praziquantel: is there real resistance and are there alternatives? Curr. Opin. Infect. Dis. 13, 659–663.

Cioli, D., Pica-Mattoccia, L., 2003. Praziquantel. Parasitol. Res. 90 Supp 1, S3–9.

Cioli, D., Botros, S.S., Wheatcroft-Francklow, K., Mbaye, A., Southgate, V., Tchuente, L.A., Pica-Mattoccia, L., Troiani, A.R., El-Din, S.H., Sabra, A.N., Albin, J., Engels, D., Doenhoff, M.J., 2004. Determination of ED50 values for praziquantel in praziquantel-resistant and -susceptible *Schistosoma mansoni* isolates. Int. J. Parasitol. 34, 979–987.

Coeli, R., Baba, E.H., Araujo, N., Coelho, P.M.Z., Oliveira, G., 2013. Praziquantel treatment decreases *Schistosoma mansoni* genetic diversity in experimental infections. PLoS Negl. Trop. Dis. 7, e2596.

Coles, G., 2006. Drug resistance and drug tolerance in parasites. Trends Parasitol. 22, 348.

Collado-Romero, M., Jiménez-Díaz, R.M., Mercado-Blanco, J., 2010. DNA sequence analysis of conserved genes reveals hybridization events that increase genetic diversity in *Verticillium dahliae*. Fungal. Biol. 114 (2-3): 209-18

Conover, D.O., Clarke, L.M., Munch, S.B., Wagner, G.N., 2006. Spatial and temporal scales of adaptive divergence in marine fishes and the implications for conservation. J. Fish Biol. 69, 21–47.

Constable, J.L., Ashley, M. V, Goodall, J., Pusey, A.E., 2001. Noninvasive paternity assignment in Gombe chimpanzees. Mol. Ecol. 10, 1279–1300.

Cooke, A., Tonks, P., Jones, F.M., O'Shea, H., Hutchings, P., Fulford, A.J., Dunne, D.W., 1999. Infection with *Schistosoma mansoni* prevents insulin dependent diabetes mellitus in non-obese diabetic mice. Parasite Immunol. 21, 169–176.

Cornuet, J.M., Luikart, G., 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144, 2001–2014.

Coura, J.R., Amaral, R.S., 2004. Epidemiological and control aspects of schistosomiasis in Brazilian endemic areas. Mem. Inst. Oswaldo Cruz 99, 13–19.

Criscione, C.D., Blouin, M.S., 2004. Life cycles shape parasite evolution: comparative population genetics of salmon trematodes. Evolution 58, 198–202.

Criscione, C.D., Blouin, M.S., 2005. Effective sizes of macroparasite populations: a conceptual model. Trends Parasitol. 21, 212–217.

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

Criscione, C.D., 2008. Parasite co-structure: Broad and local scale approaches. Parasite-Journal La Soc. Fr. Parasitol. 15, 439–443.

Criscione, C.D., Vilas, R., Paniagua, E., Blouin, M.S., 2011. More than meets the eye: detecting cryptic microgeographic population structure in a parasite with a complex life cycle. Mol. Ecol. 20, 2510–2524.

Crompton, D.W.T., 1999. How much human helminthiasis is there in the world? J. Parasitol. 85, 397–403.

Cupit, P.M., Steinauer, M.L., Tonnessen, B.W., Eric Agola, L., Kinuthia, J.M., Mwangi, I.N., Mutuku, M.W., Mkoji, G.M., Loker, E.S., Cunningham, C., 2011. Polymorphism associated with the *Schistosoma mansoni* tetraspanin-2 gene. Int. J. Parasitol. 41, 1249–1252.

Curtis, J., Minchella, D.J., 2000. Schistosome population genetic structure: when clumping worms is not just splitting hairs. Parasitol. Today 16, 68–71.

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

Curtis, J., Sorensen, R.E., Minchella, D.J., 2002. Schistosome genetic diversity: the implications of population structure as detected with microsatellite markers. Parasitology 125, S51–S59.

Danso-Appiah, A., De Vlas, S.J., 2002. Interpreting low praziquantel cure rates of *Schistosoma mansoni* infections in Senegal. Trends Parasitol. 18, 125–129.

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

Davies, C.M., Webster, J.P., Woolhouse, M.E.J., 2001. Trade-offs in the evolution of virulence in an indirectly transmitted macroparasite. Proc. R. Soc. London Ser. B-Biological Sci. 268, 251–257.

Davis, G., 1993. Evolution of Prosobranch snails transmitting Asian *Schistosoma*; coevolution with *Schistosoma:* a review. Prog. Clin. Parasitol. 3, 145-204

De Bont, J., Vercruysse, J., 1997. The epidemiology and control of cattle schistosomiasis. Parasitol. Today.

De Clercq, D., Rollinson, D., Diarra, A., Sacko, M., Coulibaly, G., Landouré, A., Traoré, M., Southgate, V.R., Kaukas, A., Vercruysse, J., 1994. Schistosomiasis in Dogon country, Mali: Identification and prevalence of the species responsible for infection in the local community. Trans. R. Soc. Trop. Med. Hyg. 88, 653–656.

De Clercq, D., Vercruysse, J., Picquet, M., Shaw, D.J., Diop, M., Ly, A., Gryseels, B., 1999. The epidemiology of a recent focus of mixed *Schistosoma haematobium* and *Schistosoma mansoni* infections around the "Lac de Guiers" in the Senegal River Basin, Senegal. Trop. Med. Int. Health 4, 544–550.

De Meeus, T., Goudet, J., 2007. A step-by-step tutorial to use HierFstat to analyse populations hierarchically structured at multiple levels. Infect. Genet. Evol. 7, 731–735.

De Meeus, T., McCoy, K.D., Prugnolle, F., Chevillon, C., Durand, P., Hurtrez-Bousses, S., Renaud, F., 2007. Population genetics and molecular epidemiology or how to "débusquer la bête". Infect. Genet. Evol. 7, 308–332.

DeJong, R.J., Morgan, J.A.T., Wilson, W.D., Al-Jaser, M.H., Appleton, C.C., Coulibaly, G., D'Andrea, P.S., Doenhoff, M.J., Haas, W., Idris, M.A., Magalhaes, L.A., Mone, H., Mouahid, G., Mubila, L., Pointier, J.P., Webster, J.P., Zanotti-Magalhaes, E.M., Paraense, W.L., Mkoji, G.M., Loker, E.S., 2003. Phylogeography of *Biomphalaria glabrata* and *B. pfeifferi*, important intermediate hosts of *Schistosoma mansoni* in the New and Old World tropics. Mol. Ecol. 12, 3041–3056.

Di Rienzo, A., Peterson, A.C., Garza, J.C., Valdes, A.M., Slatkin, M., Freimer, N.B., 1994. Mutational processes of simple-sequence repeat loci in human populations. Proc. Natl. Acad. Sci. USA. 91, 3166–3170.

Diaw, O.T., Vassiliades, G., Seye, M., Sarr, Y., 1991. Epidemiology of intestinal schistosomiasis with *Schistosoma mansoni* in Richard-Toll (Delta of Senegal River) - Malacological Survey. Bull. La Soc. Pathol. Exot. 84, 174–183.

Doenhoff, M.J., Kusel, J.R., Coles, G.C., Cioli, D., 2002. Resistance of *Schistosoma mansoni* to praziquantel: is there a problem? Trans R. Soc. Trop. Med. Hyg. 96, 465–469.

Doenhoff, M.J., Pica-Mattoccia, L., 2006. Praziquantel for the treatment of schistosomiasis: its use for control in areas with endemic disease and prospects for drug resistance. Expert. Rev. Anti. Infect. Ther. 4, 199–210.

Dunn, R., 2011. The wildlife of our bodies: predators, parasites and partners that shape who we are today. HarperCollins Publishers, New York, USA.

Dunne, D.W., Cooke, A., 2005. A worm's eye view of the immune system: consequences for evolution of human autoimmune disease. Nat. Rev. Immunol. 5, 420–426.

Duplantier, J.M., Sene, M., 2000. Rodents as reservoir hosts in the transmission of *Schistosoma mansoni* in Richard Toll, Senegal, West Africa. J. Helminthol. 74, 129–135.

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

El-Khoby, T., Galal, N., Fenwick, A., Barakat, R., El-Hawey, A., Nooman, Z., Habib, M., Abdel-Wahab, F., Gabr, N.S., Hammam, H.M., Hussein, M.H., Mikhail, N.N., Cline, B.L., Strickland, G.T., 2000. The epidemiology of schistosomiasis in Egypt: summary findings in nine governorates. Am. J. Trop. Med. Hyg. 62, 88–99.

Ellegren, H., 2014. Genome sequencing and population genomics in non-model organisms. Trends Ecol. Evol. 29, 51-63

Emery, A.M., Allan, F.E., Rabone, M.E., Rollinson, D., 2012. Schistosomiasis collection at NHM (SCAN). Parasit. Vectors 5, 185.

Eppert, A., Lewis, F.A., Grzywacz, C., Coura-Filho, P., Caldas, I., Minchella, D.J., 2002. Distribution of schistosome infections in molluscan hosts at different levels of parasite prevalence. J. Parasitol. 88, 232–236.

Ericson, L., Burdon, J.J., Müller, W.J., 1999. Spatial and temporal dynamics of epidemics of the rust fungus *Uromyces valerianae* on populations of its host *Valeriana salina*. J. Ecol. 87, 649–658.

Ernould, J.C., Ba, K., Sellin, B., 1999. Increase of intestinal schistosomiasis after praziquantel treatment in a *Schistosoma haematobium* and *Schistosoma mansoni* mixed focus. Acta Trop. 73, 143–52.

Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14, 2611–2620.

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

Fallon, P.G., Hamilton, J. V, Doenhoff, M.J., 1995. Efficacy of treatment of murine *Schistosoma mansoni* infections with praziquantel and oxamniquine correlates with infection intensity: role of host antibody. Parasitology 111, 59–66.

Fallon, P.G., Mubarak, J.S., Fookes, R.E., Niang, M., Butterworth, A.E., Sturrock, R.F., Doenhoff, M.J., 1997. *Schistosoma mansoni*: maturation rate and drug susceptibility of different geographic isolates. Exp. Parasitol. 86, 29–36.

Fenwick, A., Savioli, L., Engels, D., Bergquist, N.R., Todd, M.H., 2003. Drugs for the control of parasitic diseases: current status and development in schistosomiasis. Trends Parasitol. 19, 509–515.

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

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

Foxman, B., Riley, L., 2001. Molecular epidemiology: focus on infection. Am. J. Epidemiol. 153, 1135–1141.

French, M.D., Churcher, T.S., Basanez, M.G., Norton, A.J., Lwambo, N.J., Webster, J.P., 2012. Reductions in genetic diversity of *Schistosoma mansoni* populations under chemotherapeutic pressure: The effect of sampling approach and parasite population definition. Acta Trop. 128, 196–205.

Fulford, A.J., Ouma, J.H., Kariuki, H.C., Thiongo, F.W., Klumpp, R., Kloos, H., Sturrock, R.F., Butterworth, A.E., 1996. Water contact observations in Kenyan communities endemic for schistosomiasis: methodology and patterns of behaviour. Parasitology 113, 223–241.

Fulford, A.J., Webster, M., Ouma, J.H., Kimani, G., Dunne, D.W., 1998. Puberty and age-related changes in susceptibility to schistosome Infection. Parasitol. Today 14, 23–26.

Gaggiotti, O.E., Lange, O., Rassmann, K., Gliddon, C., 1999. A comparison of two indirect methods for estimating average levels of gene flow using microsatellite data. Mol. Ecol. 8, 1513–1520.

Galvani, A.P., 2005. Age-dependent epidemiological patterns and strain diversity in helminth parasites. J. Parasitol. 91, 24–30.

Goater, T.M., Goater, C.P., Esch, G.W., 2013. Parasitism - The Diversity and Ecology of Animal Parasites. 2nd Edition. Cambridge University Press, New York, USA.

Gomez-Mestre, I., Tejedo, M., 2004. Contrasting patterns of quantitative and neutral genetic variation in locally adapted populations of the natterjack toad, *Bufo calamita*. Evolution 58, 2343–2352.

Goudet, J., 1995. FSTAT (Version 1.2): A computer program to calculate *F*-statistics. J. Hered. 86, 485–486.

Goudet, J., 2005. HIERFSTAT, a package for R to compute and test hierarchical *F*-statistics. Mol. Ecol. Notes 5, 184–186.

Gower, C.M., Webster, J.P., 2005. Intraspecific competition and the evolution of virulence in a parasitic trematode. Evolution. 59, 544–553.

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

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

Gower, C.M., Gouvras, A.N., Lamberton, P.H., Deol, A., Shrivastava, J., Mutombo, P.N., Mbuh, J. V, Norton, A.J., Webster, B.L., Stothard, J.R., Garba, A., Lamine, M.S., Kariuki, C., Lange, N.C., Mkoji, G.M., Kabatereine, N.B., Gabrielli, A.F., Rudge, J.W., Fenwick, A., Sacko, M., Dembele, R., Lwambo, N.J., Tchuente, L.A., Rollinson, D., Webster, J.P., 2013. Population genetic structure of *Schistosoma mansoni* and *Schistosoma haematobium* from across six sub-Saharan African countries: Implications for epidemiology, evolution and control. Acta Trop. 128(2): 261-74.

Greenberg, R.M., 2005. Ca^{2+} signalling, voltage-gated Ca^{2+} channels and praziquantel in flatworm neuromusculature. Parasitology 131 Suppl, S97–108.

Gryseels, B., 1994. Human resistance to *Schistosoma* infections: age or experience? Parasitol. Today 10, 380–384.

Gryseels, B., Stelma, F.F., Talla, I., van Dam, G.J., Polman, K., Sow, S., Diaw, M., Sturrock, R.F., Doehring-Schwerdtfeger, E., Kardorff, R., et al., 1994. Epidemiology, immunology and chemotherapy of *Schistosoma mansoni* infections in a recently exposed community in Senegal. Trop. Geogr. Med. 46, 209–219.

Gryseels, B., De Vlas, S.J., 1996. Worm burdens in schistosome infections. Parasitol. Today 12, 115–119.

Gryseels, B., Mbaye, A., De Vlas, S.J., Stelma, F.F., Guisse, F., Van Lieshout, L., Faye, D., Diop, M., Ly, A., Tchuem-Tchuente, L.A., Engels, D., Polman, K., 2001. Are poor responses to praziquantel for the treatment of *Schistosoma mansoni* infections in Senegal due to resistance? An overview of the evidence. Trop. Med. Int. Heal. 6, 864–873.

Gryseels, B., Polman, K., Clerinx, J., Kestens, L., 2006. Human schistosomiasis. Lancet 368, 1106–1118.

Handschumacher, P., Dorsainville, R., Diaw, O.-T., Hebrard, G., Niang, M., Herve, J.P., 1992. Constraints climatiques et amenagements hydroliques à propos de l'epidemie de bilharziose intestinale de RIchard-

Toll (Sénégal) ou la modification des risques sanitaire en milieu sahelian. Risques Pathol. Rythm. Paroxysmes Clim. 287–295.
Hanelt, B., Brant, S. V, Steinauer, M.L., Maina, G.M., Kinuthia, J.M., Agola, L.E., Mwangi, I.N., Mungai, B.N., Mutuku, M.W., Mkoji, G.M., Loker, E.S., 2009. *Schistosoma kisumuensis* n. sp. (Digenea: Schistosomatidae) from murid rodents in the Lake Victoria Basin, Kenya and its phylogenetic position within the *S. haematobium* species group. Parasitology 136, 987–1001.
Hardy, O.J., Vekemans, X., 2002. SPAGEDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Mol. Ecol. Notes 2, 618–620.
Hardy, O.J., Charbonnel, N., Freville, H., Heuertz, M., 2003. Microsatellite allele sizes: a simple test to assess their significance on genetic differentiation. Genetics 163, 1467–1482.
Hartl, D., Clark, A., 2007. Principles of population genetics, 4th ed. Sinauer Associates, Inc. Publishers, Sunderland, Massachusetts, USA.
Hedrick, P.W., 2005. A standardized genetic differentiation measure. Evolution 59, 1633–1638.
Hoffman, J.I., Webster, J.P., Ndamba, J., Woolhouse, M.E., 1998. Extensive genetic variation revealed in adjacent populations of the schistosome intermediate host *Biomphalaria pfeifferi* from a single river system. Ann. Trop. Med. Parasitol. 92, 693–698.
Hotez, P.J., Molyneux, D.H., Fenwick, A., Kumaresan, J., Sachs, S.E., Sachs, J.D., Savioli, L., 2007. Control of neglected tropical diseases. N. Engl. J. Med. 357, 1018–1027.
Hotez, P.J., Kamath, A., 2009. Neglected tropical diseases in sub-saharan Africa: review of their prevalence, distribution, and disease burden. PLoS Negl. Trop. Dis. 3, e412.
Hubisz, M.J., Falush, D., Stephens, M., Pritchard, J.K., 2009. Inferring weak population structure with the assistance of sample group information. Mol. Ecol. Resour. 9, 1322–1332.
Huyse, T., Poulin, R., Theron, A., 2005. Speciation in parasites: a population genetics approach. Trends Parasitol. 21, 469–475.
Huyse, T., Webster, B.L., Geldof, S., Stothard, J.R., Diaw, O.T., Polman, K., Rollinson, D., 2009. Bidirectional introgressive hybridization between a cattle and human schistosome species. PLoS Pathog. 5, e1000571.
Huyse, T., Van den Broeck, F., Jombart, T., Webster, B., Diaw, O., Volckaert, F., Balloux, F., Rollinson, D., Polman, K., 2013. Regular treatments of praziquantel do not impact on the genetic make-up of *S. mansoni* in Northwest Senegal: evidence for drug tolerance? Infect. Genet. Evol. 18: 100-5
Imbert-Establet, D., Xia, M., Jourdane, J., 1994. Parthenogenesis in the genus *Schistosoma*: electrophoretic evidence for this reproduction system in *S. japonicum* and *S. mansoni*. Parasitol. Res. 80, 186–191.
Ismail, M., Metwally, A., Farghaly, A., Bruce, J., Tao, L.F., Bennett, J.L., 1996. Characterization of isolates of *Schistosoma mansoni* from Egyptian villagers that tolerate high doses of praziquantel. Am. J. Trop. Med. Hyg. 55, 214–218.
Ismail, M., Botros, S., Metwally, A., William, S., Farghally, A., Tao, L.F., Day, T.A., Bennett, J.L., 1999. Resistance to praziquantel: direct evidence from *Schistosoma mansoni* isolated from Egyptian villagers. Am. J. Trop. Med. Hyg. 60, 932–935.
Jarne, P., Theron, A., 2001. Genetic structure in natural populations of flukes and snails: a practical approach and review. Parasitology 123 Suppl, S27–40.
Jombart, T., 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24, 1403–1405.
Jombart, T., Devillard, S., Balloux, F., 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet. 11, 49
Kabatereine, N.B., Vennervald, J.B., Ouma, J.H., Kemijumbi, J., Butterworth, A.E., Dunne, D.W., Fulford, A.J.C., 1999. Adult resistance to schistosomiasis mansoni: age-dependence of reinfection remains constant in communities with diverse exposure patterns. Parasitology 118, 101–105.
Kabatereine, N.B., Brooker, S., Koukounari, A., Kazibwe, F., Tukahebwa, E.M., Fleming, F.M., Zhang, Y., Webster, J.P., Stothard, J.R., Fenwick, A., 2007. Impact of a national helminth control programme on infection and morbidity in Ugandan schoolchildren. Bull. World. Heal. Organ. 85, 91–99.
Katz, N., Chaves, A., Pellegrino, J., 1972. A simple device for quantitative stool thick-smear technique in *Schistosomiasis mansoni*. Rev. Inst. Med. Trop. Sao Paulo 14, 397–400.
Katz, N., Coelho, P.M.Z., 2008. Clinical therapy of schistosomiasis mansoni: The Brazilian contribution. Acta Trop. 108, 72-78
Keiser, J., Utzinger, J., 2007. Artemisinins and synthetic trioxolanes in the treatment of helminth infections. Curr. Opin. Infect. Dis. 20, 605–612.
Kilbourne, E., 1973. The molecular epidemiology of influenza. J. Infect. Dis. 127, 478–87.

King, C.H., Olbrych, S.K., Soon, M., Singer, M.E., Carter, J., Colley, D.G., 2011. Utility of repeated praziquantel dosing in the treatment of schistosomiasis in high-risk communities in Africa: a systematic review. PLoS Negl. Trop. Dis. 5.

Kloos, H., Fulford, A.J.C., Butterworth, A.E., Sturrock, R.F., Ouma, J.H., Kariuki, H.C., Thiongo, F.W., Dalton, P.R., Klumpp, R.K., 1997. Spatial patterns of human water contact and *Schistosoma mansoni* transmission and infection in four rural areas in Machakos District, Kenya. Soc. Sci. Med. 44, 949–968.

Kohn, A.B., Anderson, P.A. V, Roberts-Misterly, J.M., Greenberg, R.M., 2001. Schistosome calcium channel beta subunits - Unusual modulatory effects and potential role in the action of the antischistosomal drug praziquantel. J. Biol. Chem. 276, 36873–36876.

Kohn, A.B., Roberts-Misterly, J.M., Anderson, P.A., Greenberg, R.M., 2003. Creation by mutagenesis of a mammalian Ca(2+) channel beta subunit that confers praziquantel sensitivity to a mammalian Ca^{2+} channel. Int. J. Parasitol. 33, 1303–1308.

Kolar, C.S., Lodge, D.M., 2001. Progress in invasion biology: predicting invaders. Trends Ecol. Evol. 16, 199–204.

Kolbe, J.J., Glor, R.E., Schettino, L.R.G., Lara, A.C., Larson, A., Losos, J.B., 2004. Genetic variation increases during biological invasion by a Cuban lizard. Nature 431, 177–181.

Koukounari, A., Fenwick, A., Whawell, S., Kabatereine, N.B., Kazibwe, F., Tukahebwa, E.M., Stothard, J.R., Donnelly, C.A., Webster, J.P., 2006a. Morbidity indicators of *Schistosoma mansoni*: relationship between infection and anemia in Ugandan schoolchildren before and after praziquantel and albendazole chemotherapy. Am. J. Trop. Med. Hyg. 75, 278–286.

Koukounari, A., Gabrielli, A.F., Toure, S., Bosque-Oliva, E., Zhang, Y., Sellin, B., Donnelly, C.A., Fenwick, A., Webster, J.P., 2007. *Schistosoma haematobium* infection and morbidity before and after large-scale administration of praziquantel in Burkina Faso. J. Infect. Dis. 196, 659–669.

Koukounari, A., Sacko, M., Keita, A.D., Gabrielli, A.F., Landoure, A., Dembele, R., Clements, A.C., Whawell, S., Donnelly, C.A., Fenwick, A., Traore, M., Webster, J.P., 2006b. Assessment of ultrasound morbidity indicators of schistosomiasis in the context of large-scale programs illustrated with experiences from Malian children. Am. J. Trop. Med. Hyg. 75, 1042–1052.

Kumar, V., Gryseels, B., 1994. Use of praziquantel against schistosomiasis: a review of current status. Int. J. Antimicrob. Agents 4, 313–320.

La Flamme, A.C., Ruddenklau, K., Bäckström, B.T., 2003. Schistosomiasis decreases central nervous system inflammation and alters the progression of experimental autoimmune encephalomyelitis. Infect. Immun. 71, 4996–5004.

Lajeunesse, M.J., Forbes, M.R., 2002. Host range and local parasite adaptation. Proc. Biol. Sci. 269, 703–710.

Lawler, R.R., 2008. Testing for a historical population bottleneck in wild Verreaux's sifaka (*Propithecus verreauxi verreauxi*) using microsatellite data. Am. J. Primatol. 70, 990–994.

Lawrence, E., 2005. Henderson's dictionary of biology. 13th Edition. Pearson Education Limited, Essex, England.

Lawton, S.P., Hirai, H., Ironside, J.E., Johnston, D.A., Rollinson, D., 2011. Genomes and geography: genomic insights into the evolution and phylogeography of the genus *Schistosoma.* Parasit. Vectors 4, 131.

Liang, Y.S., Dai, J.R., Ning, A., Yu, D.B., Xu, X.J., Zhu, Y.C., Coles, G.C., 2001a. Susceptibility of *Schistosoma japonicum* to praziquantel in China. Trop. Med. Int. Health. 6 (9): 707-14.

Liang, Y.S., Coles, G.C., Dai, J.R., Zhu, Y.C., Doenhoff, M.J., 2001b. Biological characteristics of praziquantel-resistant and -susceptible isolates of *Schistosoma mansoni*. Ann. Trop. Med. Parasitol. 95, 715–723.

Librado, P., Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25, 1451–1452.

Lockyer, A.E., Olson, P.D., Ostergaard, P., Rollinson, D., Johnston, D.A., Attwood, S.W., Southgate, V.R., Horak, P., Snyder, S.D., Le, T.H., Agatsuma, T., McManus, D.P., Carmichael, A.C., Naem, S., Littlewood, D.T., 2003. The phylogeny of the Schistosomatidae based on three genes with emphasis on the interrelationships of *Schistosoma* (Weinland, 1858). Parasitology 126, 203–224.

Lockyer, A.E., Jones, C.S., Noble, L.R., Rollinson, D., 2004. Trematodes and snails: an intimate association. Can. J. Zool. 82: 251-269.

Luikart, G., England, P.R., Tallmon, D., Jordan, S., Taberlet, P., 2003. The power and promise of population genomics: From genotyping to genome typing. Nat. Rev. Genet. 4, 981–994.

Mack, R.N., Simberloff, D., Mark Lonsdale, W., Evans, H., Clout, M., Bazzaz, F.A., 2000. Biotic invasions: causes, epidemiology, global consequences, and control. Ecol. Appl. 10, 689–710.

Majewski, J., Schwartzentruber, J., Lalonde, E., Montpetit, A., Jabado, N., 2011. What can exome sequencing do for you? J. Med. Genet. 48, 580–589.

Malone, J.B., Abdel-Rahman, M.S., El Bahy, M.M., Huh, O.K., Shafik, M., Bavia, M., 1997. Geographic information systems and the distribution of *Schistosoma mansoni* in the Nile delta. Parasitol. Today 13, 112–119.

Matschiner, M., Salzburger, W., 2009. TANDEM: integrating automated allele binning into genetics and genomics workflows. Bioinformatics 25, 1982–1983.

McCoy, K.D., Boulinier, T., Tirard, C., Michalakis, Y., 2003a. Host-dependent genetic structure of parasite populations: Differential dispersal of seabird tick host races. Evolution. 57, 288–296.

McCoy, K.D., Tirard, C., Michalakis, Y., 2003b. Spatial genetic structure of the ectoparasite *Ixodes uriae* within breeding cliffs of its colonial seabird host. Heredity. 91, 422–429.

McDonald, B.A., Linde, C., 2002. Pathogen population genetics, evolutionary potential, and durable resistance. Annu. Rev. Phytopathol. 40, 349–379.

Meirmans, P.G., 2006. Using the AMOVA framework to estimate a standardized genetic differentiation measure. Evolution 60, 2399–2402.

Melman, S.D., Steinauer, M.L., Cunningham, C., Kubatko, L.S., Mwangi, I.N., Wynn, N.B., Mutuku, M.W., Karanja, D.M., Colley, D.G., Black, C.L., Secor, W.E., Mkoji, G.M., Loker, E.S., 2009. Reduced susceptibility to praziquantel among naturally occurring Kenyan isolates of *Schistosoma mansoni*. PLoS Negl. Trop. Dis. 3, e504.

Meurs, L., Mbow, M., Vereecken, K., Menten, J., Mboup, S., Polman, K., 2012. Epidemiology of mixed *Schistosoma mansoni* and *Schistosoma haematobium* infections in Northwest Senegal. Int. J. Parasitol. 42, 305–311.

Minchella, D.J., Sollenberger, K.M., Pereira de Souza, C., 1995. Distribution of schistosome genetic diversity within molluscan intermediate hosts. Parasitology 111 (2), 217–220.

Miura, O., Torchin, M.E., Kuris, A.M., Hechinger, R.F., Chiba, S., 2006. Introduced cryptic species of parasites exhibit different invasion pathways. Proc. Natl. Acad. Sci. 103, 19818–19823.

Molyneux, D.H., Hotez, P.. J., Fenwick, A., 2005. "Rapid-impact interventions": how a policy of integrated control for Africa's neglected tropical diseases could benefit the poor. PLoS Med. 2, e336.

Montresor, A., 2011. Cure rate is not a valid indicator for assessing drug efficacy and impact of preventive chemotherapy interventions against schistosomiasis and soil-transmitted helminthiasis. Trans R. Soc. Trop. Med. Hyg. 105, 361-363

Morgan, J.A., Dejong, R.J., Snyder, S.D., Mkoji, G.M., Loker, E.S., 2001. *Schistosoma mansoni* and *Biomphalaria*: past history and future trends. Parasitology 123 Suppl, S211–S228.

Morgan, J.A.T., DeJong, R.J., Lwambo, N.J.S., Mungai, B.N., Mkoji, G.M., Loker, E.S., 2003. First report of a natural hybrid between *Schistosoma mansoni* and *S. rodhaini.* J. Parasitol. 89, 416–418.

Morgan, J.A., Dejong, R.J., Adeoye, G.O., Ansa, E.D., Barbosa, C.S., Bremond, P., Cesari, I.M., Charbonnel, N., Correa, L.R., Coulibaly, G., D'Andrea, P.S., De Souza, C.P., Doenhoff, M.J., File, S., Idris, M.A., Incani, R.N., Jarne, P., Karanja, D.M., Kazibwe, F., Kpikpi, J., Lwambo, N.J., Mabaye, A., Magalhaes, L.A., Makundi, A., Mone, H., Mouahid, G., Muchemi, G.M., Mungai, B.N., Sene, M., Southgate, V., Tchuente, L.A., Theron, A., Yousif, F., Zanotti-Magalhaes, E.M., Mkoji, G.M., Loker, E.S., 2005. Origin and diversification of the human parasite *Schistosoma mansoni*. Mol. Ecol. 14, 3889–3902.

Morin, P.A., Manaster, C., Mesnick, S.L., Holland, R., 2009. Normalization and binning of historical and multi-source microsatellite data: overcoming the problems of allele size shift with allelogram. Mol. Ecol. Resour. 9, 1451–1455.

Mulvey, M., Aho, J.M., Lydeard, C., Leberg, P.L., Smith, M.H., 1991. Comparative population genetic structure of a parasite (*Fascioloides magna*) and its definitive host. Evolution 45, 1628–1640.

Nadler, S.A., 1995. Microevolution and the genetic structure of parasite populations. J. Parasitol. 81, 395–403.

Nagayama, Y., Watanabe, K., Niwa, M., McLachlan, S.M., Rapoport, B., 2004. *Schistosoma mansoni* and alpha-galactosylceramide: prophylactic effect of Th1 Immune suppression in a mouse model of Graves' hyperthyroidism. J. Immunol. 173, 2167–2173.

Narum, S., Buerkle, C., Davey, J., Miller, M., Hohenlohe, P., 2013. Genotyping-by-sequencing in ecological and conservation genomics. Mol. Ecol. 22, 2841–2847.

Nei, M., 1987. Molecular Evolutionary Genetics. Columbia University Press, New York, USA.

Norton, A.J., Webster, J.P., Kane, R.A., Rollinson, D., 2008. Interspecific parasite competition: mixed infections of *Schistosoma mansoni* and *S. rodhaini* in the definitive host. Parasitology 135, 473–484.

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

Odogwu, S.E., Ramamurthy, N.K., Kabatereine, N.B., Kazibwe, F., Tukahebwa, E., Webster, J.P., Fenwick, A., Stothard, J.R., 2006. *Schistosoma mansoni* in infants (aged < 3 years) along the Ugandan shoreline of Lake Victoria. Ann. Trop. Med. Parasitol. 100, 315–326.

Paradis, E. 2010. pegas: an R package for population genetics with an integrated-modular approach. Bioinformatics 26 (3), 419-20.

Paterson, S., Viney, M.E., 2000. The interface between epidemiology and population genetics. Parasitol. Today.

Pfrender, M.E., Spitze, K., Hicks, J., Morgan, K., Latta, L., 2000. Lack of concordance between genetic diversity estimates at the molecular and quantitative-trait levels. Conserv. Genet. 1, 263–269.

Pica-Mattoccia, L., Carlini, D., Guidi, A., Cimica, V., Vigorosi, F., Cioli, D., 2006. The schistosome enzyme that activates oxamniquine has the characteristics of a sulfotransferase. Mem. Inst. Oswaldo. Cruz. 101 Suppl, 307–312.

Pica-Mattoccia, L., Moroni, R., Tchuem Tchuente, L.A., Southgate, V.R., Cioli, D., 2000. Changes of mate occur in *Schistosoma mansoni.* Parasitology 120 (5), 495–500.

Pica-Mattoccia, L., Ruppel, A., Xia, C.M., Cioli, D., 2008. Praziquantel and the benzodiazepine Ro 11-3128 do not compete for the same binding sites in schistosomes. Parasitology 135, 47–54.

Picquet, M., Ernould, J.C., Vercruysse, J., Southgate, V.R., Mbaye, A., Sambou, B., Niang, M., Rollinson, D., 1996. The epidemiology of human schistosomiasis in the Senegal River Basin. Trans. R. Soc. Trop. Med. Hyg. 90, 340–346.

Picquet, M., Vercruysse, J., Shaw, D.J., Diop, M., Ly, A., 1998. Efficacy of praziquantel against *Schistosoma mansoni* in Northwest Senegal. Trans. R. Soc. Trop. Med. Hyg. 92, 90–93.

Piry, S., Luikart, G., Cornuet, J., 1999. BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. J. Hered. 90, 502–503.

Pitchford, R.J., Visser, P.S., 1975. Simple and rapid technique for quantitative estimation of helminth eggs in human and animal excreta with special reference to *Schistosoma* spp. Trans. R. Soc. Trop. Med. Hyg. 69, 318–322.

Polman, K., Stelma, F.F., Le Cessie, S., De Vlas, S.J., Falcão Ferreira, S.T.M., Talla, I., Deelder, A.M., Gryseels, B., 2002. Evaluation of the patterns of *Schistosoma mansoni* infection and re-infection in Senegal, from faecal egg counts and serum concentrations of circulating anodic antigen. Ann. Trop. Med. Parasitol. 96, 679–689.

Pompanon, F., Bonin, A., Bellemain, E., Taberlet, P., 2005. Genotyping errors: Causes, consequences and solutions. Nat. Rev. Genet. 6, 847–859.

Poulin, R., 2007. Evolutionary ecology of parasites, 2nd ed. Princeton University Press, Princeton, USA.

Poulin, R., Morand, S., 2004. Parasite biodiversity. Smithsonian Books, Washington DC, USA.

Price, P.W., 1980. Evolutionary biology of parasites. Monogr. Popul. Biol. 15, 1–237.

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945–959.

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

Prugnolle, F., Roze, D., Theron, A., DE Meeus, T., 2005b. *F*-statistics under alternation of sexual and asexual reproduction: a model and data from schistosomes (platyhelminth parasites). Mol. Ecol. 14, 1355–1365.

Prugnolle, F., Theron, A., Pointier, J.P., Jabbour-Zahab, R., Jarne, P., Durand, P., de Meeus, T., 2005c. Dispersal in a parasitic worm and its two hosts: consequence for local adaptation. Evolution 59, 296–303.

Rauch, G., Kalbe, M., Reusch, T.B.H., 2005. How a complex life cycle can improve a parasite's sex life. J. Evol. Biol. 18, 1069–1075.

Ramachandran, S., Deshpande, O., Roseman, C.C., Rosenberg, N.A., Feldman, M.W., Cavalli-Sforza, L.L., 2005. Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. P. Natl. Acad. Sci. USA 102: 15942-15947.

Renganathan, E., Cioli, D., 1998. An international initiative on praziquantel use. Parasitol. Today 14, 390–391.

Rodrigues, N.B., Coura, P., de Souza, C.P., Passos, L.K.J., Dias-Neto, E., Romanha, A.J., 2002. Populational structure of *Schistosoma mansoni* assessed by DNA microsatellites. Int. J. Parasitol. 32, 843-851.

Rollinson, D., Simpson, A.J.G., 1987. The biology of schistosomes. From genes to latrines. Academic Press Ltd, London, UK.

Rollinson, D., Southgate, V.R., Vercruysse, J., Moore, P.J., 1990. Observations on natural and experimental interactions between *Schistosoma bovis* and *S. curassoni* from West Africa. Acta Trop. 47, 101–114.

Rollinson, D., Knopp, S., Levitz, S., Stothard, J.R., Tchuem Tchuenté, L.-A., Garba, A., Mohammed, K.A., Schur, N., Person, B., Colley, D., Utzinger, J., 2013. Time to set the agenda for schistosomiasis elimination. Acta Trop. 128: 423-440.

Rosenberg, N.A., 2004. DISTRUCT: a program for the graphical display of population structure. Mol. Ecol. Notes 4, 137–138.

Sakai, A.K., Allendorf, F.W., Holt, J.S., Lodge, D.M., Molofsky, J., With, K.A., Baughman, S., Cabin, R.J., Cohen, J.E., Ellstrand, N.C., McCauley, D.E., O'Neil, P., Parker, I.M., Thompson, J.N., Weller, S.G., 2001. The Population Biology of Invasive Species. Annu. Rev. Ecol. Syst. 32, 305–332.

Savioli, L., Engels, D., Roungou, J.B., Fenwick, A., Endo, H., 2004. Schistosomiasis control. Lancet. 363, 658.

Scott, J.T., Diakhate, M., Vereecken, K., Fall, A., Diop, M., Ly, A., De Clercq, D., de Vlas, S.J., Berkvens, D., Kestens, L., Gryseels, B., 2003. Human water contacts patterns in *Schistosoma mansoni* epidemic foci in Northwest Senegal change according to age, sex and place of residence, but are not related to intensity of infection. Trop. Med. Int. Heal. 8, 100–108.

Secor, W.E., Colley, D.G., 2005. Schistosomiasis, World class parasites. Springer, New York, USA.

Shrivastava, J., Qian, B.Z., McVean, G., Webster, J.P., 2005. An insight into the genetic variation of *Schistosoma japonicum* in mainland China using DNA microsatellite markers. Mol. Ecol. 14, 839–849.

Silva, L., Liu, S., Blanton, R.E., 2006. Microsatellite analysis of pooled *Schistosoma mansoni* DNA: an approach for studies of parasite populations. Parasitology 132, 331–338.

Sire, C., Durand, P., Pointier, J.P., Théron, a, 2001. Genetic diversity of *Schistosoma mansoni* within and among individual hosts (*Rattus rattus*): infrapopulation differentiation at microspatial scale. Int. J. Parasitol. 31, 1609–16.

Slatkin, M., 1995. A measure of population subdivision based on microsatellite allele frequencies. Genetics 139, 457–462.

Sorensen, R.E., Rodrigues, N.B., Oliveira, G., Romanha, A.J., Minchella, D.J., 2006. Genetic filtering and optimal sampling of *Schistosoma mansoni* populations. Parasitology 133, 443–451.

Southgate, V.R., 1978. Factors Possibly Restricting Distribution of *Schistosoma intercalatum* (Fisher, 1934). Zeitschrift Fur Parasitenkunde-Parasitology Res. 56, 183–193.

Southgate, V.R., Rollinson, D., Ross, G.C., Knowles, R.J., 1982. Mating-behavior in mixed infections of *Schistosoma haematobium* and *Schistosoma intercalatum*. J. Nat. Hist. 16, 491–496.

Southgate, V.R., Tchuente, L.A.T., Vercruysse, J., Jourdane, J., 1995. Mating-Behavior in Mixed Infections of *Schistosoma haematobium* and *Schistosoma mattheei*. Parasitol. Res. 81, 651–656.

Southgate, V.R., 1997. Schistosomiasis in the Senegal river basin: Before and after the construction of the dams at Diama, Senegal and Manantali, Mali and future prospects. J. Helminthol. 71, 125–132.

Southgate, V.R., Jourdane, J., Tchuente, L.A.T., 1998. Recent studies on the reproductive biology of the schistosomes and their relevance to speciation in the Digenea. Int. J. Parasitol. 28, 1159–1172.

Southgate, V.R., de Clercq, D., Sene, M., Rollinson, D., Ly, A., Vercruysse, J., 2000a. Observations on the compatibility between *Bulinus* spp. and *Schistosoma haematobium* in the Senegal River basin. Ann. Trop. Med. Parasitol. 94, 157–164.

Southgate, V.R., Tchuente, L.A.T., Theron, A., Jourdane, J., Ly, A., Moncrieff, C.B., Gryseels, B., 2000b. Compatibility of *Schistosoma mansoni* Cameroon and *Biomphalaria pfeifferi* Senegal. Parasitology 121, 501–505.

Sow, S., de Vlas, S.J., Engels, D., Gryseels, B., 2002. Water-related disease patterns before and after the construction of the Diama dam in Northwest Senegal. Ann. Trop. Med. Parasitol. 96, 575–586.

Sow, S., de Vlas, S.J., Stelma, F., Vereecken, K., Gryseels, B., Polman, K., 2011. The contribution of water contact behavior to the high *Schistosoma mansoni* Infection rates observed in the Senegal River Basin. BMC Infect. Dis. 11, 198.

Steinauer, M.L., 2009. The sex lives of parasites: Investigating the mating system and mechanisms of sexual selection of the human pathogen *Schistosoma mansoni*. Int. J. Parasitol. 39, 1157–1163.

Steinauer, M.L., Agola, L.E., Mwangi, I.N., Mkoji, G.A., Loker, E.S., 2008a. Molecular epidemiology of *Schistosoma mansoni*: A robust, high-throughput method to assess multiple microsatellite markers from individual miracidia. Infect. Genet. Evol. 8, 68–73.

Steinauer, M.L., Hanelt, B., Mwangi, I.N., Maina, G.M., Agola, L.E., Kinuthia, J.M., Mutuku, M.W., Mungai, B.N., Wilson, W.D., Mkoji, G.M., Loker, E.S., 2008b. Introgressive hybridization of human and rodent schistosome parasites in western Kenya. Mol. Ecol. 17, 5062–5074.

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

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

Steinauer, M.L., Christie, M.R., Blouin, M.S., Agola, L.E., Mwangi, I.N., Maina, G.M., Mutuku, M.W., Kinuthia, J.M., Mkoji, G.M., Loker, E.S., 2013. Non-invasive sampling of schistosomes from humans requires correcting for family structure. PLoS Negl. Trop. Dis. 7, e2456.

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

Stelma, F.F., Talla, I., Polman, K., Niang, M., Sturrock, R.F., Deelder, A.M., Gryseels, B., 1993. Epidemiology of *Schistosoma mansoni* infection in a recently exposed community in Northwest Senegal. Am. J. Trop. Med. Hyg. 49, 701–706.

Stelma, F.F., Talla, I., Sow, S., Kongs, A., Niang, M., Polman, K., Deelder, A.M., Gryseels, B., 1995. Efficacy and side effects of praziquantel in an epidemic focus of *Schistosoma mansoni*. Am. J. Trop. Med. Hyg. 53, 167–170.

Stoll, N.R., 1947. This wormy world. J. Parasitol. 33, 1–18.

Suarez, A. V, Tsutsui, N.D., 2008. The evolutionary consequences of biological invasions. Mol. Ecol. 17, 351–360.

Talla, I., Kongs, A., Verle, P., Belot, J., Sarr, S., Coll, A.M., 1990. Outbreak of intestinal schistosomiasis in the Senegal River Basin. Ann. Soc. Belg. Med. Trop. 70, 173–180.

Talla, I., Kongs, A., Verle, P., 1992. Preliminary study of the prevalence of human schistosomiasis in Richard-Toll (the Senegal River Basin). Trans. R. Soc. Trop. Med. Hyg. 86, 182.

Taylor, M.G., 1970. Hybridization experiments on five species of African schistosomes. J. Helminthol. 44, 253–314.

Tchuem Tchuenté, L.A., Imbert-Establet, D., Delay, B., Jourdane, J., 1993. Choice of mate, a reproductive isolating mechanism between *Schistosoma intercalatum* and *S. mansoni* in mixed infections. Int. J. Parasitol. 23, 179–185.

Tchuem Tchuenté, L.A., Imbertestablet, D., Southgate, V.R., Jourdane, J., 1994. Interspecific stimulation of parthenogenesis in *Schistosoma intercalatum* and *Schistosoma mansoni*. J. Helminthol. 68, 167–173.

Tchuem Tchuenté, L.A., Southgate, V.R., Imbertestablet, D., Jourdane, J., 1995. Change of mate and mating competition between males of *Schistosoma intercalatum* and *Schistosoma mansoni*. Parasitology 110, 45–52.

Tchuem Tchuenté, L.A., Morand, S., ImbertEstablet, D., Delay, B., Jourdane, J., 1996a. Competitive exclusion in human schistosomes: The restricted distribution of *Schistosoma intercalatum*. Parasitology 113, 129–136.

Tchuem Tchuenté, L.A., Southgate, V.R., Combes, C., Jourdane, J., 1996b. Mating behaviour in schistosomes: Are paired worms always faithful? Parasitol. Today 12, 231–236.

Tchuem Tchuenté, L.A., Southgate, V.R., Jourdane, J., Kaukas, A., Vercruysse, J., 1997. Hybridisation between the digeneans *Schistosoma haematobium* and *S. mattheei*: Viability of hybrids and their development in sheep. Syst. Parasitol. 36, 123–131.

Tchuem Tchuenté, L.A., Southgate, V.R., Theron, A., Jourdane, J., Ly, A., Gryseels, B., 1999. Compatibility of *Schistosoma mansoni* and *Biomphalaria pfeifferi* in Northwest Senegal. Parasitology 118, 595–603.

R Development Core Team, 2013. R: A language and environment for statistical computing. Vienna, Austria: the R Foundation for Statistical Computing.

Templeton, A.R., Crandall, K.A., Sing, C.F., 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. Genetics 132, 619–633.

Terry, R.J., 1994. Human immunity to schistosomes: concomitant immunity? Parasitol. Today 10, 377–378.

Theron, A., Sire, C., Rognon, A., Prugnolle, F., Durand, P., 2004. Molecular ecology of *Schistosoma mansoni* transmission inferred from the genetic composition of larval and adult infrapopulations within intermediate and definitive hosts. Parasitology 129, 571–585.

Thiele, E.A., Sorensen, R.E., Gazzinelli, A., Minchella, D.J., 2008. Genetic diversity and population structuring of *Schistosoma mansoni* in a Brazilian village. Int. J. Parasitol. 38, 389–399.

Thomas, F., Poulin, R., Guegan, J.F., Michalakis, Y., Renaud, F., 2000. Are there pros as well as cons to being parasitized? Parasitol. Today. 16, 553-536

Tibayrenc, M., 1995. Population genetics and strain typing of microorganisms: how to detect departures from panmixia without individualizing alleles and loci. C. R. Acad. Sci. III. 318, 135–139.

Torchin, M.E., Lafferty, K.D., Kuris, A.M., 2002. Parasites and marine invasions. Parasitology 124 Suppl, S137–51.

Torchin, M.E., Lafferty, K.D., Dobson, A.P., McKenzie, V.J., Kuris, A.M., 2003. Introduced species and their missing parasites. Nature 421, 628–630.

Toure, S., Zhang, Y., Bosque-Oliva, E., Ky, C., Ouedraogo, A., Koukounari, A., Gabrielli, A.F., Bertrand, S., Webster, J.P., Fenwick, A., 2008. Two-year impact of single praziquantel treatment on infection in the national control programme on schistosomiasis in Burkina Faso. Bull. World Heal. Organ. 86, 780–7.

Utzinger, J., Xiao, S.H., Tanner, M., Keiser, J., 2007. Artemisinins for schistosomiasis and beyond. Curr. Opin. Investig. Drugs 8, 105–116.

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

Valentim, C.L., Cioli, D., Chevalier, F.D., Cao, X., Taylor, A.B., Holloway, S.P., Pica-Mattoccia, L., Guidi, A., Basso, A., Tsai, I.J., Berriman, M., Carvalho-Queiroz, C., Almeida, M., Aguilar, H., Frantz, D.E., Hart, P.J., LoVerde, P.T., Anderson, T.J., 2013. Genetic and molecular basis of drug resistance and species-specific drug action in schistosome parasites. Science 342, 1385–9.

Vanaerschot, M., Huijben, S., Van den Broeck, F., Dujardin, J.C., 2013. Drug resistance in vector-borne parasites: multiple actors and scenarios for an evolutionary arms race. FEMS Microbiol. Rev. Ahead of print, art.nr. 10.1111/1574-6976.12032.

Van den Broeck, F., Geldof, S., Polman, K., Volckaert, F.A.M., Huyse, T., 2011. Optimal sample storage and extraction procotols for reliable multilocus genotyping of the human parasite *Schistosoma mansoni.* Infect. Genet. Evol. 11, 1413–1418.

van der Werf, M.J., de Vlas, S.J., Brooker, S., Looman, C.W.N., Nagelkerke, N.J.D., Habbema, J.D.F., Engels, D., 2003. Quantification of clinical morbidity associated with schistosome infection in sub-Saharan Africa. Acta Trop. 86, 125–139.

Van Lieshout, L., de Jonge, N., el-Masry, N., Mansour, M.M., Bassily, S., Krijger, F.W., Deelder, A.M., 1994. Monitoring the efficacy of different doses of praziquantel by quantification of circulating antigens in serum and urine of schistosomiasis patients., Parasitology 108: 519-26.

Van Lieshout, L., Polderman, A.M., Visser, L.G., Verwey, J.J., Deelder, A.M., 1997. Detection of the circulating antigens CAA and CCA in a group of Dutch travellers with acute schistosomiasis. Trop. Med. Int. Health 2, 551–557.

Van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M., Shipley, P., 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol. Ecol. Notes 4, 535–538.

Vercruysse, J., Southgate, V.R., Rollinson, D., 1985. The epidemiology of human and animal schistosomiasis in the Senegal River Basin. Acta Trop. 42, 249–259.

Vercruysse, J., Southgate, V.R., Rollinson, D., Declercq, D., Sacko, M., Debont, J., Mungomba, L.M., 1994. Studies on transmission and schistosome interactions in Senegal, Mali and Zambia. Trop. Geogr. Med. 46, 220–226.

Verheye, W.H., 1995. Impact of climate and soil-conditions on conception and implementation of irrigation schemes in the Senegal River Basin. Agric. Water Manag. 28, 73–94.

Verle, P., Stelma, F., Desreumaux, P., Dieng, A., Diaw, O., Kongs, A., Niang, M., Sow, S., Talla, I., Sturrock, R.F., Gryseels, B., Capron, A., 1994. Preliminary study of urinary schistosomiasis in a village in the Delta of the Senegal River Basin, Senegal. Trans. R. Soc. Trop. Med. Hyg. 88, 401–405.

Vrijenhoek, R.C., Graven, M.A., 1992. Population genetics of Egyptian *Biomphalaria alexandrina* (Gastropoda, Planorbidae). J. Hered. 83, 255–261.

Waples, R.S., 1998. Separating the wheat from the chaff: Patterns of genetic differentiation in high gene flow species. J. Hered. 89, 438–450.

Waples, R.S., & Naish, K., 2009. Genetic and evolutionary considerations in fishery management: research needs for the future. In: The future of fisheries science in North America. Springer, The Netherlands.

Webster, B.L., Southgate, V.R., Tchuente, L.A.T., 1999. Mating interactions between *Schistosoma haematobium* and *S. mansoni*. J. Helminthol. 73, 351–356.

Webster, B.L., Southgate, V.R., 2003. Mating interactions of *Schistosoma haematobium* and *S. intercalatum* with their hybrid offspring. Parasitology 126, 327–338.

Webster, B.L., Southgate, V.R., Littlewood, D.T., 2006. A revision of the interrelationships of *Schistosoma* including the recently described *Schistosoma guineensis*. Int. J. Parasitol. 36, 947–955.

Webster, B.L., 2009. Isolation and preservation of schistosome eggs and larvae in RNA*later*® facilitates genetic profiling of individuals. Parasit. Vectors 2, 50.

Webster, B.L., Rollinson, D., Stothard, J.R., Huyse, T., 2010. Rapid diagnostic multiplex PCR (RD-PCR) to discriminate *Schistosoma haematobium* and *S. bovis.* J. Helminthol. 84, 107–114.

Webster, B.L., Diaw, O.T., Seye, M.M., Webster, J.P., Rollinson, D., 2013a. Introgressive hybridization of *Schistosoma haematobium* group species in Senegal: species barrier break down between ruminant and human schistosomes. PLoS Negl. Trop. Dis. 7, e2110.

Webster, B.L., Webster, J.P., Gouvras, A.N., Garba, A., Lamine, M.S., Diaw, O.T., Seye, M.M., Tchuem Tchuenté, L.-A., Simoonga, C., Mubila, L., Mwanga, J.R., Lwambo, N.J.S., Kabatereine, N.B., Lange, C.N., Kariuki, C., Mkoji, G.M., Rollinson, D., Stothard, J.R., 2013b. DNA "barcoding" of *Schistosoma mansoni* across sub-Saharan Africa supports substantial within locality diversity and geographical separation of genotypes. Acta Trop. 128, 250–60.

Webster, B.L., Diaw, O., Seye, M., Faye, D., Stothard, J., Figueiredo, J., Rollinson, D., 2013c. Praziquantel treatment of school children from single and mixed infection foci of intestinal and urogenital schistosomiasis along the Senegal River Basin: monitoring treatment success and re-infection patterns. Acta Trop. 128 (2): 292-302.

Webster, J.P., Woolhouse, M.E.J., 1999. Cost of resistance: relationship between reduced fertility and increased resistance in a snail-schistosome host-parasite system. Proc. R. Soc. London Ser. B-Biological Sci. 266, 391–396.

Webster, J.P., Davies, C.M., Ndamba, J., Noble, L.R., Jones, C.S., Woolhouse, M.E.J., 2001. Spatio-temporal genetic variability in the schistosome intermediate host *Biomphalaria pfeifferi*. Ann. Trop. Med. Parasitol. 95, 515–527.

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

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

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

Weir, B.S., Cockerham, C.C., 1984. Estimating *F*-Statistics for the analysis of population structure. Evolution. 38, 1358–1370.

WHO, 2006. Preventive chemotherapy in human helminthiasis: coordinated use of anthelminthic drugs in control interventions: a manual for health professionals and programme managers. World Health Organization, Geneva.

WHO, 2011. Schistosomiasis: number of people treated worldwide in 2009. Weekly Epidemiological Record 86: 73-80.

WHO, 2012. Accelerating work to overcome the global impact of neglected tropical diseases: a roadmap for implementation. World Health Organization, Geneva.

Wiese, R.J., 2008. Conservation of wildlife populations: demography, genetics, and management. J. Hered. 99, 441–442.

William, S., Sabra, A., Ramzy, F., Mousa, M., Demerdash, Z., Bennett, J.L., Day, T.A., Botros, S., 2001. Stability and reproductive fitness of *Schistosoma mansoni* isolates with decreased sensitivity to praziquantel. Int. J. Parasitol. 31, 1093–1100.

Woolhouse, M.E.J., Hasibeder, G., Chandiwana, S.K., 1996. On estimating the basic reproduction number for *Schistosoma haematobium*. Trop. Med. Int. Heal. 1, 456–463.

Zhan, J., Pettway, R.E., McDonald, B.A., 2003. The global genetic structure of the wheat pathogen *Mycosphaerella graminicola* is characterized by high nuclear diversity, low mitochondrial diversity, regular recombination, and gene flow. Fungal Genet. Biol. 38, 286–297.

Zietara, M.S., Arndt, A., Geets, A., Hellemans, B., Volckaert, F.A.M., 2000. The nuclear rDNA region of *Gyrodactylus arcuatus* and *G. branchicus* (Monogenea: Gyrodactylidae). J. Parasitol. 86, 1368–1373.

List of publications

IN PREPARATION

Van den Broeck F, Vanoverbeke J, Volckaert F, Polman K & Huyse T. Genetic diversity of schistosome parasites in response to drug treatment: a simulation-based approach.

Van den Broeck F, Maes G, Rollinson D, Hellemans B, Vereecken K, Talla I, Sy I, Volckaert F, Polman K & Huyse T. Invasion genetics of *Schistosoma mansoni* into Northwest Senegal reveals signatures of expansion and high gene flow.

PUBLISHED

Van den Broeck F, Meurs L, Raeymaekers JAM, Boon N, Tandakha N, Volckaert FAM, Polman K & Huyse T 2014. Inbreeding within human *Schistosoma mansoni*: do host-specific factors shape the genetic composition of parasite populations? Heredity. Ahead of print, art. nr. 10.1038/hdy.2014.13

Meurs L, Mbow M, Boon N, **Van den Broeck F**, Vereecken K, Dieye T, Abatih E, Huyse T, Mboup S, Polman K 2013. Micro-Geographical Heterogeneity in *Schistosoma mansoni* and *S. haematobium* Infection and Morbidity in a Co-Endemic Community in Northwest Senegal. PLoS Neglected Tropical Diseases, 7 (12), e2608.

Vanaerschot M, Huijben S, **Van den Broeck F**, Dujardin JC, 2013. Drug resistance in vector-borne parasites: multiple actors and scenarios for an evolutionary arms race. FEMS Microbiology Reviews. Ahead of print, art.nr. 10.1111/1574-6976.12032.

Huyse T, **Van den Broeck F**, Jombart T, Webster B, Diaw O, Volckaert F, Balloux F, Rollinson D & Polman K 2013. Regular treatments of praziquantel do not impact on the genetic make-up of *S. mansoni* in Northwest Senegal: evidence for drug tolerance? Infection, Genetics and Evolution. 18: 100-5

Huyse T, **Van den Broeck F**, Hellemans B, Volckaert F, Polman K, 2013. Hybridization between the two major African schistosome species of humans. International Journal of Parasitology. 43 (8): 687-9

De Waele V, **Van den Broeck F**, Huyse T, McGrath G, Higgins I, Speybroeck N, Berzano M, Raleigh P, Mulcahy GM & Murphy TM 2013. Panmictic structure of the *Cryptosporidium parvum* population in Irish calves: influence of prevalence and host movement. Applied and Environmental Microbiology. 79 (8): 2534-41

Raeymaekers JAM, Lens L, **Van den Broeck F**, Van Dongen S & Volckaert FAM 2012. Quantifying population structure on short timescales. Molecular Ecology 21(14): 3458-73

Adaui V, Maes I, Huyse T, **Van den Broeck F**, Talledo M, Kuhls K, De Doncker S, Maes L, Llanos-Cuentas A, Schönian G, Arevalo J & Dujardin JC 2011. Multilocus genotyping reveals a polyphyletic pattern among naturally antimony-resistant *Leishmania braziliensis* isolates from Peru. Infection, Genetics and Evolution 11(8): 1873-1880

Van den Broeck F, Geldof S, Polman K, Volckaert FAM & Huyse T 2011. Optimal sample storage and extraction protocols for reliable multilocus genotyping of the human parasite *Schistosoma mansoni*. Infection, Genetics and Evolution 11(6): 1413-1418

FACULTY OF SCIENCE
DEPARTMENT OF BIOLOGY
ECOLOGY, EVOLUTION AND BIODIVERSITY CONSERVATION DIVISION
Charles Deberiotstraat 32
B-3001 HEVERLEE, BELGIUM
www.kuleuven.be