Schistosoma mansoni AND *Biomphalaria* SNAILS IN LAKE VICTORIA: DISTRIBUTION, GENETICS AND ECOLOGICAL DYNAMICS

Claire J. Standley

MA (Cantab), MSc (Oxon)

Thesis submitted to the University of Nottingham

for the degree of Doctor of Philosophy

OCTOBER 2010

i. Abstract

Ex Africa semper aliquid nova

- Pliny the Elder

Background

Intestinal schistosomiasis, caused by the trematode parasite *Schistosoma mansoni*, is a disease of major public health importance in the Lake Victoria region. Accurate information pertaining to the disease's distribution can greatly assist in the maintenance and realignment of existing control strategies. Rapid mapping of disease prevalence is reliant on diagnostic technologies; in the case of intestinal schistosomiasis, traditional stool-based methods are beginning to be complimented with new, rapid diagnostic tools. However, these require extensive validation, in a variety of settings, to determine their efficacy and field-utility. Similarly, diagnostic tests are influenced by other factors, such as the level of endemicity of the parasite, or even intraspecific factors such as genetic diversity. Preliminary research has shown *S. mansoni* to have high levels of genetic diversity throughout East Africa, although its population dynamics within Lake Victoria have never been explored in detail.

Local transmission of the parasite is determined by the presence of compatible snails of the genus *Biomphalaria*, which act as the intermediate host for *S. mansoni*. As such, basic distribution mapping of the presence and abundance of *Biomphalaria* around Lake Victoria will better reveal potential risk areas for transmission; these data could also lead to a greater understanding of the specific environmental conditions preferred by *Biomphalaria*, allowing for predictive mapping of suitable transmission environments. Sophisticated geostatistical tools have facilitated this process, although new, Bayesian models, which have yielded dividends in parasite mapping, have yet to be applied to intermediate host distribution mapping.

There is evidence for species-level differences in compatibility with *S. mansoni*; the taxonomy of *Biomphalaria* is confused and confusing, with two suspected species being present in Lake Victoria, but these populations have never been subject to detailed classification scrutiny. Similarly, there may be intraspecific effects on transmission compatibility; population level analysis would elucidate patterns of genetic variation across Lake Victoria.

Taking both the genetics and ecological preferences of *Biomphalaria* into consideration, there may be an effect of scale, which should be explored. Patterns that are found to be significant at the level of Lake Victoria may not hold true at a local scale, crucial for the local transmission biology of the parasite. For example, research has suggested that the presence of other gastropods could reduce the abundance of *Biomphalaria*, through competitive effects. Diversity may influence the transmission of schistosomiasis at the level of the terminal host as well; it is well known that rodents are capable of acting as reservoirs for *S. mansoni*, as are numerous non-human primates. However, within the context of Lake Victoria, non-human primates have rarely been comprehensively surveyed for the disease, despite the obvious implications for conservation, as well as the potential effect on maintaining local transmission cycles even in the face of treatment campaigns.

Findings

The research presented in the following thesis sought to address the above themes, relating to the dynamics of *S. mansoni* and *Biomphalaria* in Lake Victoria. Through a series of four field expeditions to the Ugandan, Tanzanian and Kenyan shorelines of Lake Victoria, data were collected pertaining to the distribution of *S. mansoni* in school-age children; 27 schools in Uganda were first surveyed for prevalence and intensity of infection. Questionnaires revealed high levels of migration among the school-children, with high levels of itinerancy also associated with increased risk of being infected with *S. mansoni* and of missing school-based treatment with praziquantel. Later surveys in Tanzania and Kenya compared prevalence of infection as diagnosed by Kato-Katz stool thick smears against a novel, urine-based, rapid diagnostic called the cathodic cirulating antigen (CCA) dipstick. The CCA performed well, with good agreement against the Kato-Katz diagnostic and high sensitivity and specificity in this high-endemicity

environment. However, when the CCA tests were used alongside Kato-Katz thick smears in the Ssesse Islands in Uganda, the agreement was not so good, suggesting that different transmission environments may affect the efficacy of diagnostic tests, and moreover, may result in widely discordant treatment recommendations being put forward.

Throughout all of these field expeditions, stool samples were also used for hatching of schistosome eggs and miracidia were collected and stored on Whatman® FTA cards for later molecular analysis. Combined with DNA sequences obtained from cercariae from shedding snails and adult worms passaged in the laboratory, genetic material was obtained from 25 sites along the shoreline of Lake Victoria, representative of all three countries. Using both cytochrome oxidase sub-unit 1 (COI) and microsatellite DNA markers, population genetics analyses revealed extremely high genetic diversity, consistent with previous research, but also low levels of population structuring, with no clear geographical patterns. This contrasted with earlier work which had suggested that the populations of *S. mansoni* in Lake Albert and Lake Victoria appeared segregated; the greater and more extensive sampling presented in this thesis revealed a cross-over of haplotypes between the two lakes, potentially associated again with human migration throughout the region, but also suggesting that adaptation and compatibility with local *Biomphalaria* populations for successful transmission might be less crucial than previously assumed.

In order to investigate the distribution and genetics of *Biomphalaria*, as well as relate these patterns to the transmission of *S. mansoni*, malacological surveys were carried out at 223 sites across the Lake Victoria shoreline. Two forms of *Biomphalaria*, based on shell characteristics, were observed; *B. sudanica*-like snails were commonly associated in marsh-like habitats whereas *B. choanomphala*-type snails were more usually found in the lake proper. These observations, together with a number of other environmental variables, were statistically tested using multivariate models within a Bayesian inference framework, a method which has never previously been used on snail intermediate host distributions. The models revealed different factors as being significant predictors of *B. sudanica* versus *B. choanomphala* presence, reinforcing the hypothesis of habitat segregation between the two forms.

Closer examination of the two forms of *Biomphalaria* found suggested intermediate shell forms as well as the more easily recognised *B. choanomphala*-like and *B. sudanica*-like forms. In order to elucidate the taxonomy of these various forms, a detailed molecular and morphological assessment was carried out on 7 populations of *Biomphalaria.* Morphologically, principal component analysis (PCA) of shell measurements, aperture outlines and internal anatomy measurements revealed overlapping groupings, supporting the earlier observation of intermediate forms. Crucially, there were no clear species-level divisions in the genetic data; the groupings that were observed were not significantly associated with those according to morphology. These data combined to imply that the *Biomphalaria* in Lake Victoria should be considered one species, with two ecophentoypic variants: *B. choanomphala* var. *choanomphala* and *B. choanomphala* var. *sudanica.* Population level analysis of a total of 29 populations revealed significant levels of population structuring with strong geographical patterns; given these local variations, yet also considering the very different patterns observed for *S. mansoni,* these findings could suggest that local compatibility is less of a restriction on the spread of the parasite than previously thought.

At a micro-scale, surveys conducted on Ngamba and Kimi Islands in Uganda, revealed complex patterns of *Biomphalaria* genetic structuring mirroring those seen at a lake-wide level. However, the Bayesian statistical models used for predicting the distribution of the snails and their abundance showed a marked difference to the large scale models, with few, if any, significant environmental predictors. The one exception was species diversity; the micro-scale surveys also provided an opportunity to examine closely the gastropod species assemablages on the two islands, and their effect on the abundance of *Biomphalaria* as well as transmission of *S. mansoni.* Gastropod species diversity, surprisingly, proved to be a positive predictor of *Biomphalaria* abundance at a micro-scale, as it had been for *Biomphalaria* presence throughout Lake Victoria. The role of anthropogenic disturbance on gastropod diversity was also investigated; while human influence did not have a statistically significant effect on species assemblage diversity, more snails (of any genus) were found infected with parasites on Kimi Island, which is characterised by high human and livestock population density. Ngamba Island, on the other hand, is a chimpanzee sanctuary; the micro-scale surveys were followed up by a

parasitological investigation of children on Kimi island as well as staff and chimpanzees resident on Ngamba Island and found *S. mansoni* in all. Molecular analysis on miracidia hatched from these wild-born, semi-captive chimpanzees, never previously attempted, revealed haplotypes common throughout Lake Victoria, indicating shared or even anthropozoonotic transmission. Snails shedding *S. mansoni* were further found in the waters around Ngamba Island; a later survey also observed a snail shedding *S. rodhaini*, which is an often-overlooked, and potentially zoonotic, schistosome.

Conclusions

Overall, the research presented in this thesis used an interdisciplinary, multi-faceted approach to explore the prevalence, distribution and dynamics of *S. mansoni* and *Biomphalaria* in Lake Victoria. As well as noting the high levels of the disease in human communities, a key element of the research revealed the importance of gastropod biodiversity in predicting *Biomphalaria* distributions. At the terminal host level, biodiversity was also seen to be important; chimpanzees were observed to be infected with common 'human' forms of *S. mansoni*, indicating their potential role as reservoirs for the disease, but also demonstrating how wild primate populations, including threatened species such as chimpanzees, can be vulnerable to schistosomiasis as well as other parasitic infections. As such, this Ph.D. research has shed light on a very important avenue of future investigation: the relationship between biodiversity and the spread of infectious diseases, particularly in the context of transmission between humans, non-human primates, and potentially also other reservoir species. While East Africa would be an ideal location to continue this kind of work on schistosomiasis, it also provides the opportunity to extend the research to other parasitic infections, within a broader context of studying the evolution and molecular epidemiology of the interface between medical and veterinary diseases.

ii. List of published papers

STANDLEY, C. J., MUGISHA, L., VERWEIJ, J. J., ADRIKO, M., ARINAITWE, M., ROWELL, C., ATUHAIRE, A., BETSON, M., HOBBS, E., VAN TULLEKEN, C. R., KANE, R. A., VAN LIESHOUT, L., AJAROVA, L., KABATEREINE, N. B. & STOTHARD, J. R. 2011. Confirmed infection with intestinal schistosomiasis in semi-captive wild born chimpanzees on Ngamba Island, Uganda. *Vector-borne and Zoonotic Diseases*, 11, 169-176.

STANDLEY, C. J., POINTIER, J-P., ISSIA, L., WISNIVESKY-COLLI, C. & STOTHARD, J. R. 2011. The identification and characterisation of *Biomphalaria peregrina* (Orbignyi, 1835) from Agua Escondida in northern Patagonia, Argentina. *Journal of Natural History*, 45, 347-356.

STANDLEY, C. J., KABATEREINE, N. B., LANGE, C. N., LWAMBO, N. J. S. & STOTHARD, J. R. 2010. Molecular epidemiology and phylogeography of *Schistosoma mansoni* around Lake Victoria. *Parasitology*, 137 (13), 1937-1949.

STANDLEY, C. J. & STOTHARD, J. R. 2010. Towards defining appropriate strategies for targeted NTD control. *Tropical Medicine & International Health*, 15, 772-773.

STANDLEY, C. J., ADRIKO, M, ARINAITWE, M., ATUHAIRE, A., KAZIBWE, F., FENWICK, A., KABATEREINE, N. B. & STOTHARD, J. R. 2010. Epidemiology and control of intestinal schistosomiasis on the Ssesse Islands, Uganda: integrating malacology and parasitology to tailor local treatment recommendations. *Parasites & Vectors*, 3: 64, available online 27 July, 2010.

STANDLEY, C. J., LWAMBO N. J. S., LANGE C. N., KARIUKI H. C., ADRIKO M. & STOTHARD J. R. 2010. Performance of circulating cathodic antigen (CCA) urine-dipsticks for rapid detection of intestinal schistosomiasis in schoolchildren from shoreline communities of Lake Victoria. *Parasites & Vectors*, 3:7, available online 5 February, 2010.

STANDLEY, C. J., ADRIKO, M., ALINAITWE, M., KAZIBWE, F., KABATEREINE, N. B. & STOTHARD, J. R. 2009. Intestinal schistosomiasis and soil-transmitted helminthiasis in Ugandan schoolchildren: a rapid mapping assessment. *Geospatial Health*, 4, 39-53.

STOTHARD, J. R., SOUSA-FIGUEIREDO, J. C., STANDLEY, C. J., VAN DAM, G. J., KNOPP, S., UTZINGER, J., AMERI, H., KHAMIS, A. N., KHAMIS, I. S., DEELDER, A. M., MOHAMMED, K. A. & ROLLINSON, D. 2009. An evaluation of urine-CCA strip test and fingerprick blood SEA-ELISA for detection of urinary schistosomiasis in schoolchildren in Zanzibar. *Acta Tropica,* 111, 64-70.

iii. Acknowledgements

The research presented in this thesis would not have been possible without the support of a number of different people, in ten different countries. First of all, I would like to extend my heartfelt appreciation to the staff at the Vector Control Division in Kampala, the National Medical Research Institute in Mwanza and the National Museum of Kenya in Nairobi for all the help and support they provided during fieldwork. Specifically, I would like to thank: Andrew 'Fiddi' Rugigana, Moses Adriko, Moses Arinaitwe, Aaron Atuhaire, Candia Rowell, Francis Kazibwe and Narcis Kabatereine from VCD; James Kubeja, Anna Rufutu and Nicholas Lwambo from NIMR; and Charles Lange from the NMK. A number of other people participated in field work and I would like to thank Allen Nalugwa, Sylvester Nyakaana (both of Makerere University), Curtis Kariuki (Ministry of Health, Kenya) and above all, Aslak Jørgensen (of the Danish Bilharzia Laboratory) for their invaluable assistance. Aslak, in particular, across three field expeditions, provided much-needed support, both malacological and psychological! Liz Humphries and her charity, HUG (http://www.helpuganda.org.uk/) were enormously generous in allowing us to use their boat for both Ssesse Island surveys, and she and Judy were invaluable team members in the field to boot. None of the work on Ngamba Island would have been possible without the generosity and involvement of Lilly Ajarova and Lawrence Mugisha, plus all the staff on Ngamba Island who assisted with sample collection and making our stay such a memorable one. Cristoffer van Tulleken and Emma Hobbs also deserve special thanks for their hard work on Ngamba Island, as does Chloe McKeon who assisted with analysis of the Ngamba *Biomphalaria* data as part of her BSc project.

For other training, I must thank Gabriel Mushinge of the University of Lusaka and Christopher Saarnak of DBL for teaching me how to use ArcGIS. Jean-Pierre Pointier of the University of Perpignan provided me with much-needed instruction on morphological dissections of *Biomphalaria*, and also advice and information throughout the rest of my PhD. Hugh Sturrock, of the London School of Tropical Medicine and Hygiene, assisted with spatial statistics.

Once back at the NHM, I was lucky to receive instruction and advice from several experts in their field. One who deserves special mention is Professor Norm MacLeod, who very generously gave of his time and expertise regarding morphometric analysis. He also provided some of the software utilised in these analyses, namely the 'xy->phi converter' and 'Eigenvector'; please contact him (n.macleod@nhm.ac.uk) for more information or to obtain copies of the programmes. Gary Jones and later Stanislav Strekopytov in the EMMA lab assisted with the water chemistry analysis and their help was much appreciated. Others who assisted me beyond the call of their duty included Bonnie Webster, Richard Kane and Aidan Emery for molecular work, Pete Olson and Cia Anderung for phylogenetics, Steve Ansell for population genetics and Jose da Sousa Figueiredo for his statistical expertise. Jose, along with Martha Betson, Scott Lawton and Fiona Allen, deserve a resounding thank you for being fantastic friends throughout my PhD, as well as for their scientific advice. Jayne King and Mike Anderson provided much appreciated assistance and expertise in the maintenance of laboratory isolates and Julia Llewellyn-Hughes and Claire Griffin were invaluable with their help in the Sequencing Facility. David Rollinson always was on hand for sage words and encouragement.

Other parts of the analyses and personnel involved in the project were based away from the NHM; for these I would like to thank Jaco Verweij and Lisette van Lieshout, both of Leiden University. Cambridge University provided FTA cards for inclusion in this research, and the surveys on the Ssesse Islands in 2010 would not have been possible without the support of Alan Fenwick, from SCI/Imperial College. Penelope Vounatsou and Laura took two full days out of their schedules to teach me Bayesian spatial statistics, for which I am deeply grateful. Above all, I would like to thank the whole CONTRAST team for advice and comments during the three years of my PhD; Thomas Kristensen and Juerg Utzinger have been particularly encouraging throughout and I thank them for their support.

This research would have not been possible without a variety of funding sources. The PhD studentship was generously provided by the NHM and many fieldwork costs were covered by the EU grant CONTRAST (FP6 STREP contract no: 032203, http://www.eu-contrast.eu; held by Russell Stothard and David Rollinson) and/or SIMI (Wellcome Trust; held by Russell Stothard).

Field costs for the Ssesse surveys were provided by the Global Network for Neglected Tropical Diseases (GNNTD; held by Alan Fenwick and Narcis Kabatereine). Additional travel funding was given by the Malacological Society, the British Society of Parasitology, the Graduate School of the University of Nottingham and the School of Biology, also at the University of Nottingham.

Finally, it remains for me to extend my deepest gratitude to the clutch of people who really made it all possible. These include all my friends and family, but particularly Paul and my parents, who have borne the brunt of my stress and panic. Sara Goodacre has been a fantastic and supportive advisor, providing important scientific and pastoral advice along the way. Of course, my final thanks are to my two supervisors: Chris Wade and Russell Stothard. Chris Wade, at Nottingham, gave me crucial and timely advice and instruction, particularly with molecular analysis, and provided a link to malacologists at the NHM. Last, but certainly not least, I am deeply indebted to all the on-going help, advice and support given to me by Russell Stothard, both at the NHM and in the field, where he was always on hand with solutions and a Jetboil whenever things got rough. Thank you!

iv. Table of contents

v. List of Figures

vi. List of Tables

vii. Abbreviations and acronyms

°C – degrees Centigrade

µl – microlitres

DNA - deoxyribonucleic acid

dNTPs – deoxynucleotide triphosphates

EPG – eggs per gramme

g/L – grammes per litre

GTR – General time reversible (model of nucleotide evolution)

HKY – Hasegawa-Kishino-Yano (model of nucleotide evolution)

hr - hour

km – kilometers

LDA – linear discriminant analysis

m – metres

min – minute

ml – millilitres

NHM – Natural History Museum (London)

PCA – principal

PCR – polymerase chain reaction

ppm – parts per million

PZQ – praziquantel

s – second

VCD – Vector Control Division (part of Ministry of Health, Kampala, Uganda)

1 General introduction

This chapter introduces *Schistosoma mansoni* parasites and *Biomphalaria* snails, placing them in the context of transmission of intestinal schistosomiasis across Africa. The impact and dynamics of this disease, specifically in East Africa, is described, and this ensuing chapter concludes with a review of the current state of knowledge regarding intestinal schistosomiasis and its intermediate host snails around and in Lake Victoria. This overview sets the scene and leads on to introduce the main research questions of this Ph.D. thesis.

1.1 *Schistosoma mansoni*

This section introduces *S. mansoni* as a parasite of global public health importance. Though describing the current state of knowledge about its evolutionary history, distribution and genetics, the gaps in our understanding are revealed, specifically regarding the dynamics of the disease in the Lake Victoria.

1.1.1 Classification and evolutionary history

Schistosomes are flatworms (Kingdom Animalia; Phylum Platyhelminthes) that are of biomedical importance as parasites of humans and a variety of other vertebrates. The Phylum was originally divided into four classes: the Turbellaria, the Monogenea, the Trematoda and the Cestoda. However, a detailed reclassification of the Platyhelminthes divided the Turbellaria into multiple orders, some of which are no longer considered true Platyhelminthes (Ehlers, 1985). Those Turbellaria still classified as platyhelminthes can be roughly divided into two groups, the Catenulida and the Rhabidora (Littlewood *et al.*, 2004); this latter group also contains a new sub-order called the Neodermata, which now groups together the Trematoda, Monogenea and Cestoda (see Figure 1.1).

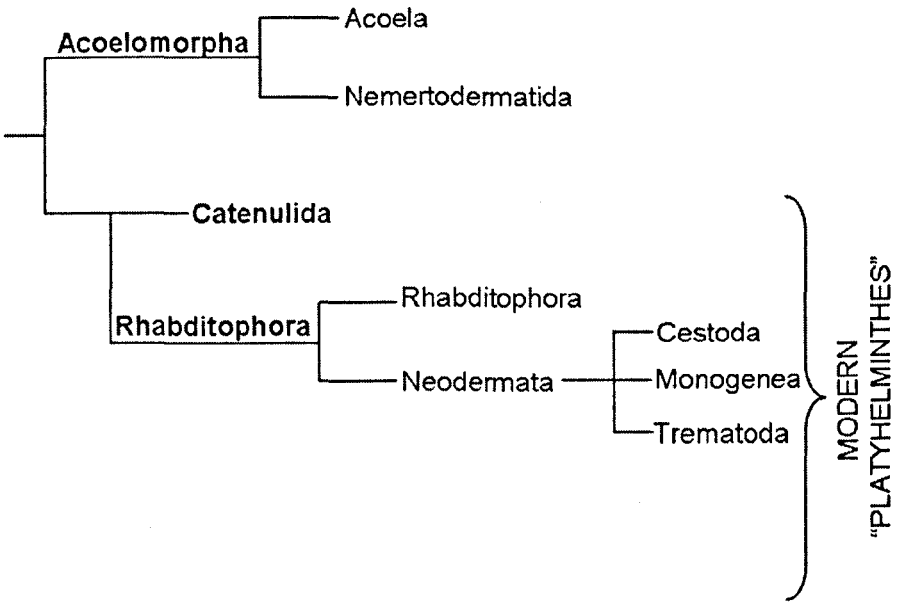

Figure 1.1 – Schematic of the new classification of the Platyheminthes
Redrawn following Ehlers (1985) and Littlewood *et al.* (2004)

Of these, schistosomes belong to the Trematoda, which includes two groups of obligate parasitic worms, the Digenea and the Aspidogastrea (Littlewood and Bray, 2000). The Digenea are defined by the presence of two suckers, one oral and one posterior ventral, and includes the schistosomes. Based on the presence of penetration glands and the fork-tailed nature of the cercarial larval stage, the schistosomes are further classified as being part of the order Strigeidida. The majority of the Strigeidida are, however, hermaphroditic; schistosomes, being dioecious, are virtually unique in their order, and as such are classified within their own family, the Schistosomatidae. This family is further characterised by being specialist parasites of the venous system, and contains up to 14 genera: 8 primarily infect birds, 5 are adapted to mammals and one preferentially infects crocodiles.

S. mansoni belongs to the genus *Schistosoma*, which comprises of 22 described species, divided into four groups, based primarily on egg morphology and the type of intermediate and terminal hosts exploited (Rollinson and Southgate, 1987). The *S. haematobium* group is restricted to Africa and the Middle East, while the *S. mansoni* group is also distributed throughout Latin America and the Caribbean; the *S. indicum* group and *S. japonicum* groups are primarily found

in East Asia, with a few singular species in India. The genus as a whole is not generally considered monophyletic, with some molecular analyses reporting that *Orientobilharzia* clusters within *Schistosoma* (Zhang *et al.*, 2001), but the four groups as defined above seem to hold up well under molecular scrutiny.

The monophyly of the *S. mansoni* group, which also contains *S. rodhaini*, *S. hippopotami* and *S. edwardiense* was initially supported by molecular methods (Bowles *et al.*, 1995, Barker and Blair, 1996). However, more recently, *S. hippopotami* has been shown to be basal to the *S. mansoni* group, using DNA sequencing (Morgan *et al.*, 2003, Webster *et al.*, 2006). These techniques have also broadly elucidated the evolutionary history of *Schistosoma*; based on changes in the mitochondiral gene order, as well as direct sequencing, the genus is currently believed to have originated in East Asia, later spreading west to India and finally into Africa (Snyder and Loker, 2000, Lockyer *et al.*, 2003, Littlewood *et al.*, 2006). It is thought that the *S. haematobium* group and the *S. mansoni* group are derived from an early schistosome which colonised Africa, transmitted by pulmonate snails, as both *S. haematobium* and *S. mansoni* utilise such snails today (Morgan *et al.*, 2001). Whereas the intermediate hosts of *S. haematobium* are *Bulinus* snails, an ancient lineage in Africa, *S. mansoni* requires snails of the genus *Biomphalaria*, which are thought to have only dispersed across to Africa, from the Neotropics, between 4 and 2.5 million years ago (Campbell *et al.*, 2000). As such, *S. mansoni* is a relatively 'young' species, believed to have evolved in East Africa soon after this introduction, which also corresponds closely to the evolutionary origins of humans and suggests a very ancient association between *S. mansoni* and *Homo sapiens* in this region (Despres *et al.*, 1992).

1.1.2 Life cycle and transmission biology of S. mansoni

The particular life history traits of *S. mansoni* have been alluded to briefly above, in that along with most other trematodes, schistosomes possess a complicated life cycle involving multiple host species.

A schematic of the life cycle of *S. mansoni* can be seen in Figure 1.2. The parasite, in the form of a cercaria, enters the body of the definitive mammal host by burrowing directly through the

skin during exposure to freshwater inhabited by the free-swimming cercariae. *S. mansoni* is capable of using a number of different mammals as its terminal host, but it appears to be highly adapted to infecting humans, perhaps due to its close evolutionary history. *S. mansoni* is known to infect other primate species, such as baboons and chimpanzees (Nelson, 1960, Fenwick, 1969, Bakuza and Nkwengulila, 2009), and rodents are also effective definitive hosts; indeed, today transmission cycles are primarily maintained by rats in the Caribbean and some parts of Latin America (Alarcon de Noya *et al.*, 1997, Modena *et al.*, 2008). Rats and rodents have also been shown to be infected with *S. mansoni* in Africa (Schwetz, 1954, Hanelt *et al.*, 2010) although their role as reservoir hosts has been much less well studied in this geographical context.

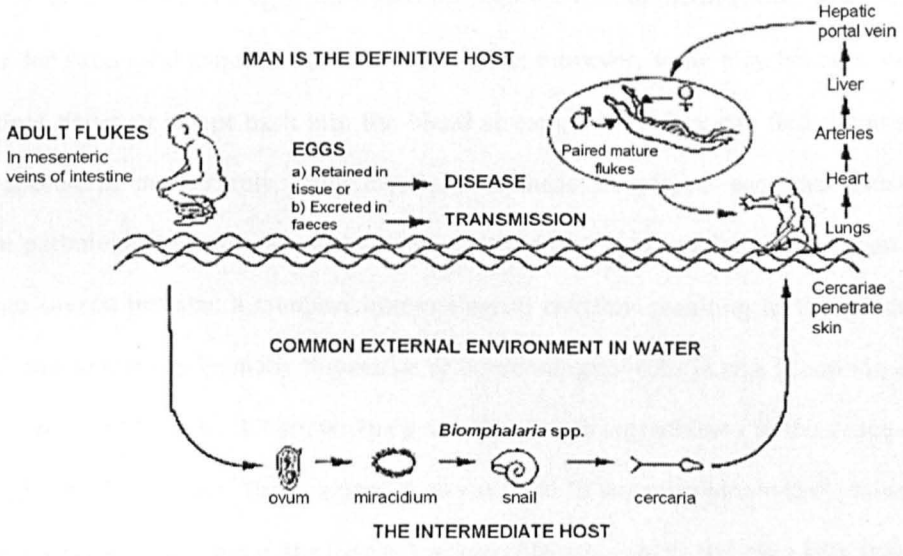

Figure 1.2 – Life cycle of *S. mansoni* (Jordan *et al.*, 1993)

The penetration of the host's skin is undertaken in three stages: the cercaria first attaches to the skin, then searches for a suitable entry point, before finally entering the epidermis (Haas and Schmitt, 1982b, Haas and Schmitt, 1982a). At this stage, the cercaria loses its tail and undergoes other major physiological maturation, becoming a schistosomulum. The schistosomulum eventually passes through the layers of the dermis and then searches for a

blood vessel; the entire penetration process can take up to several days. Once in the blood vessel, the schistosomulum is swept along with the blood flow to the heart and then the lungs.

Further developmental changes occur in the lungs, before the schistosomulum migrates, again via the blood vessels, to the hepatic portal vein in the liver. The parasites may flow through the circulatory system multiple times before reaching the liver. Here, the parasite becomes an immature adult, including the development of the copulatory system. At this stage, the male and female worms pair up, with the female fitting in to the male's gynacophoric canal, initiating further maturation, and migrate once again, this time to the mesenteric veins surrounding the intestines. Here, the pair feed on plasma and red blood cells and the female starts producing eggs, usually between 1 and 3 at a time for this species and approximately 350 per day per pair (Cheever *et al.*, 1994). The eggs must pass through the wall of the intestine, to be voided in the faeces, for successful continuation of the life cycle; however, some may become lodged in the intestinal tissue or swept back into the blood stream, where they can find themselves in the liver, spleen, or more rarely, other organs. It is these 'misplaced' eggs that cause the most severe pathological symptoms of the disease. When the eggs are lodged in organ tissue, the damage caused initiates a complex immunological reaction resulting in the production of a granuloma, which can be many thousands of immunological cells in size (Coutinho *et al.*, 2003, Lenzi *et al.*, 2006). Figure 1.3 shows the presence of such a granuloma in the tissue of a mouse infected with *S. mansoni*. These granulomas can lead to organosplenomegaly, where affected organs increase in size, and in the case of the liver, fibrosis, cancer and even liver failure.

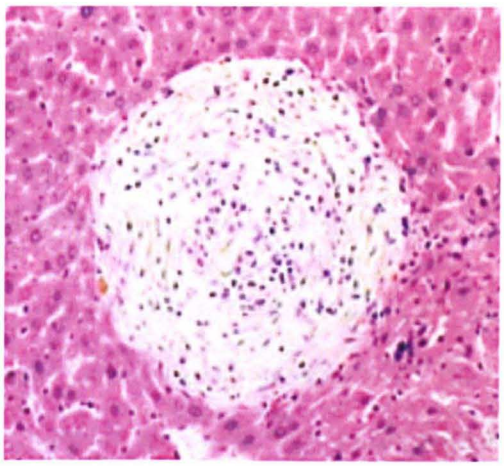

Figure 1.3 – Image of a liver granuloma, caused by *S. mansoni*, in a mouse
Note the egg in the centre of the image, surrounded by immune cells. Image taken from Coutinho *et. al* (2003).

Those eggs that do successfully pass through the intestinal membrane and into the gut lumen are voided in the faeces. The exact mechanism of egg hatching is under dispute, but is known to be influenced by osmotic potential as well as the movement of the miracidium inside the egg (Kusel, 1970, Xu and Dresden, 1990). It may also be that the miracidium produces an enzyme which breaks down the egg at a particular point (Samuelson *et al.*, 1984, Xu and Dresden, 1986). The miracidium is the second free-swimming parasite larval form, which actively seeks out a snail intermediate host; in the case of *S. mansoni*, these are snails of the genus *Biomphalaria*, which will be described in further detail in section 1.2 below. As the life cycle will only continue given the exposure of miracidia to these snails, transmission is most likely in areas without formal sanitation or latrines, and thus where people defecate in such a way that their faeces can wash directly into bodies of fresh water. The specificity of host-choice by *Schisotoma* is under debate; research on *S. haematobium* has shown that miracidia preferentially infect susceptible *Bulinus globosus* snails, and uninfected ones over previously infected ones (Allan *et al.*, 2009). On the other hand, studies on *S. mansoni* have shown that certain isolates of *S. mansoni* preferentially select susceptible snails, based on chemical cues in the water, whereas others show no difference in host choice between *Biomphalaria* and other sympatric pulmonates (Kalbe *et al.*, 1996, Hassan *et al.*, 2003). Differences in signalling glycoconjugate proteins may provide the proximate chemical cue for miracidial behaviour (Haberl and Haas, 1992, Haas *et al.*, 1995, Kalbe *et al.*, 1996), though ultimately the differences

seen between isolates could be the result of selection based on variations in local snail assemblages. Overall, schistosome-pulmonate susceptibility is a complex issue, which will only be touched upon briefly in this overview, but of which several comprehensive reviews exist (for example, El-Ansary & Al-Daihan, 2006, Webster & Davies, 2001).

The miracidium is only fully viable in freshwater for between 4-6 hours after hatching, as it is unable to feed during this time, possessing no intestinal tract. The larva is phototrophic and geotrophic, in both cases resulting in swimming towards the surface of the water, where the intermediate hosts are more likely to be found. Once the miracidium detects a snail host, it attaches to the mantle or foot using a modified sucker called the terebratorium and bores directly into the tissue. Cross-infection experiments in the lab have shown that even when the miracidicium penetrates the correct genus of snail, there are still a number of intrageneric factors that contribute to the ability of parasite to develop, leading to differences in compatibility between species and even populations of snails (Basch, 1976). Once inside the snail, the miracidium develops into a sporocyst; in resistant snails, this sporocyst is attacked by haemocytes and destroyed, whereas in susceptible snails, the sporocyst evades the snail's immune system, possibly by secreting products that modulate nitric oxide production in the host haemocytes (Zahoor *et al.*, 2009). Other humoral factors may also influence susceptibility, or indeed antagonism between sporocysts acquired at different times; some laboratory strains of *B. glabrata*, shown to be susceptible to an isolate of *S. mansoni*, were later observed to be resistant to subsequent infection with the same isolate (Sire *et al.*, 1998).

The primary sporocysts engender a second generation via clonal reproduction; these secondary or 'daughter' sporocysts, migrate rapidly towards the hepatopancreatic region of the snail and there become cercariogenous whilst continuing to duplicate (Théron and Jourdane, 1979, Théron and Touassem, 1989). Approximately 20 days after initial miradicial exposure, the cercariae emerge directly through the mantle, once again into fresh water. An individual snail can produce thousands of cercariae on a daily basis, with periods of emergence dictated by circadian rhythms, selected for by the life history of the terminal host (Théron, 1984). For example, in areas where humans are the main definitive host for *S. mansoni*, cercarial

emergence is at its peak at midday, when the sun is at its hottest and people are more likely to be immersed in water. Given these environmental cues for emergence, it has further been demonstrated that cercariae, like miracidia, respond to a number of stimuli with regards to host attraction. Phototaxis is one, but turbulence, chemotaxis and thermotaxis have also been postulated as stimulatory agents; consensus as to the relative importance between each of these has not yet been reached (McKerrow and Salter, 2002, Curwen and Wilson, 2003). Also like miracidia, cercariae are non-feeding and short-lived, and so must find a suitable host to penetrate within approximately 12 hours after emerging from the snail. Once the host has been located, the cercariae attach and the cycle begins again.

1.1.3 Distribution and public health impact of intestinal schistosomiasis

Schistosoma mansoni is the most widely distributed of its genus, found in over 54 countries worldwide, the bulk of which are in Africa but also in Latin America (mainly Surinam, Brazil and Venezuela), the Caribbean (Puerto Rico, the Dominican Republic and the Antilles) and parts of the Middle East (Morgan *et al.*, 2005). The species is thought to have evolved in Africa and secondarily re-invaded the Middle East; the introduction to Latin America probably occurred as a result of the slave trade in the 18th and 19th centuries (Fletcher *et al.*, 1981, Despres *et al.*, 1993), although the intermediate host snails were present in the area previously.

This global presence means that, after malaria, schistosomiasis is considered the world's second most important parasitic infection, with nearly 700 million people at risk of exposure and an estimated 83 million or so actively infected with *S. mansoni* (Chitsulo *et al.*, 2000, Stothard *et al.*, 2009a). Despite this global presence, the main burden of the disease is in sub-Saharan Africa, due to low levels of sanitation and plumbing, with the majority of people still reliant on lakes, rivers and streams for their daily water needs.

Demographically, due to their propensity to swim and play in the water, pre-pubescent and pubescent boys are most at risk of high intensity infections (Rudge *et al.*, 2008, Kapito-Tembo *et al.*, 2009) but also because later in adulthood immune responses appear to provide some resistance to infection. The mechanism for this change in susceptibility has long been under

research scrutiny, but a consensus has not yet been reached; some studies point to the importance of different cytokines in adults and children with varying infection intensities (Scott *et al.*, 2004), whereas others suggest that other immunoglobins may be responsible (Caldas *et al.*, 2000) or indeed levels of hormones (Fulford *et al.*, 1998) and even skin quality (Stirewalt, 1956), which can differ greatly between age groups. It may be in fact that all these different responses work synergistically in ways that are not yet fully understood, combining age-mediated and exposure-related mechanisms of resistance. More recently, pre-school-age children have been shown to be infected with *S. mansoni* as well, implying that the incidence of severe clinical signs in children may be related to having been exposed to the parasite at a much younger age than previously suspected (Odogwu *et al.*, 2006, Sousa-Figueiredo *et al.*, 2010).

1.1.4 Clinical manifestation of intestinal schistosomiasis

An initial symptom of infection with schistosomiasis can be the development of 'swimmer's itch', otherwise known as cercarial dermatitis, which is caused by an immune response to antigens on the cercarial outer coat (Boros, 1989). This manifests itself within hours of contact with infected water; while repeated exposure to non-human schistosomes can result in severe symptoms such as papules, erythema and even ruptured vesicles (Cort, 1950, Baird and Wear, 1987), the dermatitis caused by *S. mansoni* cercariae is usually mild. More clinically significant for infection with human schistosomes (though rarely for *S. mansoni*), there can also be an acute manifestation, in the form of Katayama fever, brought on by an immune response to the initial egg production of maturing worms, although innate variation in pathological consequences may also contribute to the patient's symptoms (Boros, 1989, Lambertucci *et al.*, 2000). Probably due to exposure to schistosome antigens from an early age and even *in utero* (Carlier *et al.*, 1980), acute schistosomiasis is relatively uncommon in areas endemic for the parasite, and is more usually manifested in travellers and tourists visiting such areas. Where people, and particularly children, are repeatedly exposed to infectious waters, chronic schistosomiasis is much more likely to represent the greatest portion of the burden of the disease on the community.

Chronic schistosomiasis is manifested in a spectrum of clinical signs and patient symptoms. In a mild form, the infection is characterised by abdominal pain, blood in stool, diarrhoea and often anaemia, due to the consumption of red blood cells by the adult worms (Boros, 1989, van der Werf *et al.*, 2003). Lethargy and malnutrition are often also observed, with studies suggesting that intestinal schistosomiasis (along with other forms of schistosomiasis and gut helminths) could be a major contributing factor to poor growth and decreased school performance in infected children although other research has been inconclusive on this subject (de Clercq *et al.*, 1998, Nokes *et al.*, 1999, Dickson *et al.*, 2000). More serious and damaging effects can be observed when schistosome eggs fail to pass into the intestinal lumen, are caught up in the bloodstream and become lodged in internal organs such as the spleen and liver. It was estimated in 2003 that more than 8 million people in sub-Saharan Africa alone are likely to be affected by hepatomegaly (van der Werf *et al.*, 2003), with renal manifestation another serious consequence of infection with *S. mansoni* in about 10-15% of those who already have critical hepatosplenic symptoms (Andrade and Van Marck, 1984). In total, the number of deaths from intestinal schistosomiasis is estimated at approximately 150,000 per year (van der Werf *et al.*, 2003). As in many cases there appears to be a direct relationship between intensity of infection with *S. mansoni* and the severity of the clinical manifestation (Sukwa *et al.*, 1986, Balen *et al.*, 2006), current control interventions have focused on monitoring infection intensity as a means of morbidity reduction in endemic communuties. This, along with an overview of other forms of schistosomiasis control, will be discussed in the next section.

1.1.5 Overview of global schistosomiasis control initiatives

The first concerted efforts to study the nature and medical implications of schistosomiasis are closely tied to the expansion of European colonial powers into schistosome-endemic regions, such as Egypt and other parts of North Africa in the 1880s. However, it was primarily the First World War which brought recognition of schistosomiasis as a public health problem to public attention, as European troops stationed in North Africa and the Middle East were diagnosed with the disease (Sandbach, 1976, Sturrock, 2001). The creation of the World Health Organisation (WHO) after World War II allowed for the creation of coordinated research and

community control programmes, as part of an associated emergence of interest in tropical medicine and improvement of public health, particularly in ex-colonial developing countries (Savioli *et al.*, 1997, Sturrock, 2001). For example, Venezuela, Puerto Rico, Japan and Israel all initiated control programmes at this time (Sandbach, 1976). However, even given the desire to curb the impact of the disease, it was not known how best to control the transmission of the disease, and numerous methods were tried in different settings: elimination of the intermediate host, education and sanitation and chemotherapy have all been used over the course of the last half-century, to greater or lesser effect.

Prior to the elucidation of the parasite's life cycle, in the 20th century, behavioural and hygienic methods for prevention of infection were paramount in control policies. However, once effective drugs were discovered and the intermediate host of the parasite described, the emphasis on sanitation decreased, partially due to a prevailing view of tropical diseases as being determined by climate rather than behaviour, which was deemed unshakeable (Bryant, 1972). The first treatment campaigns were carried out as early as the First World War era but using drugs which were expensive, unreliable and often highly toxic (Jordan, 1968, Jordan, 2000). At the same time, the increased trend for rapid measures of disease control further put behavioural and sanitation-based projects out of favour, moving instead towards efforts to eliminate the intermediate host snails. As early as the 1920s, the efficacy of copper sulphate as a molluscicidal agent was realised and was rolled-out in a growing number of campaigns through the following four or five decades; research was even undertaken to demonstrate how aerial spraying of such chemicals provided an efficient means of wide-scale extermination of snails in regions that were difficult to reach by other means (Sturrock and Barnish, 1973).

The importance of hygiene and sanitation in the transmission of schistosomiasis resurfaced in the 1960s, given new strength by increasing concern with the effect of molluscicidal chemicals on aquatic ecosystems (Sandbach, 1976). Moreover, large-scale irrigation and water management plans had mushroomed in the post-war years, significantly increasing the area of suitable habitat for snails and moreover, exacerbating transmission of schistosomiasis; intermediate host control even with molluscicides was futile in the face of such expansion,

although it was still deemed appropriate in certain settings (Malek, 1976, McCullough *et al.*, 1980).

The greatest single step forward in the modern fight against schistosomiasis was the discovery of praziquantel (PZQ), an anthelmintic drug that proved very successful in killing adult schistosome worms, with few severe side effects at normal dosage (Katz *et al.*, 1979). In 1984, the World Health Organisation's Expert Committee on the Control of Schistosomiasis officially advocated a strategy of morbidity control, based on chemotherapy to reduce worm burdens (Magnussen, 2003). In response, a number of countries developed national control programmes based on mass drug administration, often administered through schools; as a result, a number of countries have greatly reduced the prevalence, and more importantly, morbidity, of schistosomiasis. Modern control strategies have sought to integrate socio-economic and demographical information into control programmes, to provide education and elicit behavioural change as well as providing treatment in communities (Kloos, 1995, Stothard *et al.*, 2006b). Similarly, much attention has been focused on the possibility of integrating treatment needs for other neglected tropical diseases into schistosomiasis control programmes, for maximum public health benefit and efficiency of resource use (Utzinger *et al.*, 2009).

A key aspect to the principle of control programmes based on chemotherapy is the need for accurate diagnostic information (Magnussen, 2003). For maximum efficacy and efficiency, treatment should be tailored to the particular needs of a community, based on the prevalence and intensity of infection as revealed by baseline surveys. The WHO has published treatment protocols based on levels of prevalence in school-based surveys; prevalence of over 50% warrants annual treatment, under 50% but greater than 10% requires treatment every other year, whereas schools with prevalence lower than 10% will get a one-off dose of PZQ (WHO, 2002). The same report also sets guidelines for categorising intensity of infection, based on the number of eggs observed per gram of faeces (EPG); fewer than 100 eggs is considered a light infection, 100-399 eggs is regarded as medium intensity infection and anything above 400 EPG is a heavy infection. However, the need for accurate, rapid assessments of on-the-ground prevalence has led to the development of a number of different diagnostic methods, not all of

which rely on stool samples for detection of infection; these will be reviewed in the following section.

1.1.6 Diagnostic tools for detection of intestinal schistosomiasis

Hand in hand with the scale up of control has been the need for better disease distribution maps at local and regional scales. Diagnostic tests are used in order to ascertain the levels of disease burden, and therefore treatment needs, in a particular locale; traditionally, these diagnostic tests were based on collecting stool samples, although more recently other diagnostic tests have been introduced, some still stool-based but others requiring the collection of urine or blood.

1.1.6.1 Stool-based diagnostics

Up until recently, virtually all intestinal schistosomiasis monitoring and mapping surveys have relied on a single diagnostic: microscopic examination of a stool sample, to examine for the presence of eggs. The most well-known microscopy technique involves the preparation of Kato-Katz thick smears, utilising a standardised volume of faeces ($1/24^{th}$ of a gramme) and a glycerin-infused cellophane cover slide, which can be stained with malachite green for better contrast of the egg against the background (Katz *et al.*, 1972). Eggs have been shown to be distributed randomly in the stool, so even a single thick smear can give a representative picture of the number of eggs present in the stool as a whole (Martin and Beaver, 1968). Kato-Katz thick smears are additionally useful for the diagnosis of soil-transmitted helminthiasis and so are useful for integrated control strategies. Requiring little in the way of equipment apart from sieves, slides and the glycerin-infused cover slips, they are very cheap to produce and have been shown to be cost-effective as a means of diagnosis, for example compared to treatment based on clinical questionnaires or symptoms (Carabin *et al.*, 2000). However, Kato-Katz thick smears are time-consuming to make and also require expertise and care when reading, and so can be costly in terms of personnel during a survey. Moreover, recent research has shown that single Kato-Katz thick smears, in other words from a single stool, are unreliable owing to sporadic egg excretion, and that it is only with three consecutive faecal samples that an

accurate egg count can be obtained (Kongs *et al.*, 2001, Booth *et al.*, 2003). Three stool samples are often difficult to obtain, both logistically in the framework of a rapid survey but also due to compliance issues; as such, other diagnostic tools have been investigated as alternatives to this approach.

Some of the other diagnostics that have been researched also involved stool samples. For example, stool can be used for hatching of the eggs contained within it; if eggs or miracidia are observed, then the sample can be considered positive. A modified Pitchford funnel (Pitchford, 1959) is often used for this process, as it removes large debris but allows eggs to pass through; in the final elute, eggs can be seen and thus used as part of the diagnosis. However, the procedure can be time-consuming, requires a large supply of filtered water and is not adapted to processing large numbers of samples (Yu *et al.*, 2007). Another stool-based technique that has shown very promising sensitivity for diagnosis of strongyloidiasis is the FLOTAC method, which uses flotation to separate eggs from the surrounding matrix (Knopp *et al.*, 2009). It is also more sensitive than ether concentration, a sedimentation technique, and is comparable to the 'gold standard' triple Kato-Katz thick smear, although its efficacy specifically for diagnosis of *S. mansoni* infection is not well established (Utzinger *et al.*, 2008, Glinz *et al.*, 2010). The FLOTAC protocol does, furthermore, require centrifugation and thus is challenging to use in a field context. The separation of eggs from faeces using a Percoll gradient has been shown to be highly sensitive, particularly when processing large amounts of faeces, but again requires centrifugation, as well as extensive preparation (Eberl *et al.*, 2002).

With the advent of molecular technologies, PCR-based approaches have also been used on stool for diagnosis of *S. mansoni*. These tend to be very effective and sensitive (Pontes *et al.*, 2003, ten Hove *et al.*, 2008), but at best only semi-quantifiable and again, difficult to use in a field context. Moreover, PCR machines are expensive and require expertise, and the amplification takes time; these factors, while not negating the value of PCR as a laboratory diagnostic, reduce its usefulness for on-going monitoring and on-the-ground mapping initiatives.

1.1.6.2 The options for new diagnostic approaches

In terms of rapid diagnostics for use in the field, blood and urine-based methods seem to have come closest to rivaling the dominance of Kato-Katz thick smears. Fingerprick blood can be used in a soluble egg antigen enzyme-linked immunosorbent assay (SEA-ELISA), which is a very fast, sensitive and easy to use method for detecting the presence of schistosome egg antibody in a patient's blood (Stothard *et al.*, 2009b). Although the presence of eggs are required at some point during the infection in order for antibodies to be produced, the SEA-ELISA kits have the advantage of not relying on the adult worms producing any eggs at that particular time-point for positive diagnosis, unlike the stool-based procedures described above. Unfortunately, the kits require cool storage, plus centrifugation as part of the protocol, and are expensive. Moreover, schistosome antibodies are known to stay in the blood stream even after treatment, and once the infection has been cleared, making them potentially problematic in areas already covered by control programmes. However, in low prevalence areas, and those removed from existing mass drug administration, the SEA-ELISA holds promise as a very sensitive measure of infection detection and as such should be evaluated further, in different endemicity settings (van Gool *et al.*, 2002, Stothard *et al.*, 2009b).

The main urine-based test that is commercially available is the schistosome circulating cathodic antigen (CCA) dipstick. It works in a similar fashion to a pregnancy test, by containing beads labelled with a coloured antibody conjugate that bind to circulating cathodic antigen, produced in the vomit of adult schistosomes and voided in the urine. The beads adsorb to the dipstick and migrate towards a fixed band on the test to which they bind, indicating a positive result. Easy to use and quick, one of the disadvantages of the CCA test is the cost, which currently is prohibitive for large-scale roll-out, particularly in developing countries, although this may change with subsequent manufacture (Stothard *et al.*, 2009b). The sensitivity of the test has also been questioned in certain geographical localities, particularly in low endemicity settings, and is known to be poor for detection of *S. haematobium* (Stothard *et al.*, 2006a, Obeng *et al.*, 2008). This regional variation suggests that different populations of schistosomes, with different genetic diversity and recent evolutionary history, may have circulating antigens that

are more or less able to be picked up by the CCA test due to differential binding ability. Genetic variation could also have an effect on the success of the diagnosis and could also be a factor affecting the SEA-ELISA test, as well as stool-based PCR procedures. Given these factors, the genetics of *S. mansoni* should be considered when developing and testing new diagnostics. More generally, the evolutionary history and genetic structure of *S. mansoni* may well have implications for the epidemiology and clinical manifestation of the disease, and so these topics are crucial research directions.

1.1.7 On-going thematic research: genetics and molecular epidemiology

For the development and success of a control programme, detailed information is required regarding the distribution and burden of *S. mansoni* in the target region. Furthermore, as has also been discussed, different diagnostic tools have advantages and disadvantages in terms of cost, convenience and reliability; moreover, there may be regional factors which affect the outcomes in a particular locality. One such factor could be the genetic variation observed within the parasite from different regions; this section overviews some of the existing research relating to the genetics of *S. mansoni* on a global as well as population level.

Molecular tools were instrumental in elucidating the interrelationships within the *Schistosoma* genus. Many of these same methods have also been utilised to explore the genome of *S. mansoni* at an intraspecific level, although usually different specific markers have been investigated, appropriate to the level of query. Although molecular analysis of this species has been on-going for some years, it is particularly timely now, as the genome of *S. mansoni* has just been published, allowing researchers a detailed insight into the organisation and functionality of its genetic structure (Berriman *et al.*, 2009). The publication of the genome is too recent to have resulted in large amounts of subsequent research as of yet, but given the use of the mitochondrial genome in determining evolutionary relationships within *Schistosoma*, it is expected that the full genome will also provide a plethora of opportunities for detailed evolutionary study (Littlewood *et al.*, 2006, Gentile and Oliveira, 2008).

Although the deeper evolutionary history of *S. mansoni* is clearly of great interest, in the context of disease transmission and public health, local population genetics are likely to have more impact. Early work in the 1980s investigated variation in patterns of enzyme electrophoresis, for example, to show evidence of potential selection due to different terminal hosts (LoVerde *et al.*, 1985) and also intraregional variation (Fletcher *et al.*, 1981). PCR-based methods began to be used in the 1990s; some focused on the amplification of random fragments of DNA, which could be compared across populations to look for variation, although the results were not replicable and thus the conclusions difficult to corroborate (Dias Neto *et al.*, 1993). Direct sequencing was more effective than RAPD analysis at spotting point mutations and could be repeated even between labs, and as such has become one of the main methods for population genetics analyses; both mitochondrial and nuclear gene fragments have been successfully amplified and sequenced for the analysis of variation between and within *Schistosoma* populations, in a variety of settings (Bowles *et al.*, 1993, Morgan *et al.*, 2003, Steinauer *et al.*, 2010). For example, analysis on *S. mansoni* at a global level revealed five distinct lineages, shedding light on the evolutionary history of the spread of the parasite, as well as creating new questions as to the specific differences between and within these lineages, which can be sympatric (Morgan *et al.*, 2005). The development of microsatellites specific to *S. mansoni* further provided scientists with the ability to look at intrapopulation variation at a very high resolution (Durand *et al.*, 2000, Blair *et al.*, 2001, Rodrigues *et al.*, 2002).

The majority of the population genetics analyses on *S. mansoni* are on South American and Caribbean populations, the latter of which differ significantly from those in Africa by the transmission cycle being almost exclusively maintained by rodents. On both continents, the population structure within the intermediate as well as terminal host has been investigated; in both, the parasite is seen to be overdispersed, although genetic diversity in the snail is often much lower than in the definitive host (Minchella *et al.*, 1995, Sire *et al.*, 2001). Genetic structure between infrapopulations (i.e. between hosts) also tends to be low, in snails, rats and also in humans, with high gene flow and little if any indication of isolation by distance (Sire *et al.*, 2001, Théron *et al.*, 2004, Agola *et al.*, 2006, Thiele, 2007, Steinauer *et al.*, 2009), suggesting that definitive host migration and dispersal may account for genetic mixing (Prugnolle *et al.*,

2005). There may be wider geographical patterns to population structure; recent research on two of the Great Lakes, Lake Albert and Lake Victoria, have suggested a possible separation of genetic types between the two, which could be attributed to host selection, possibly at the level of local snail compatibility (Stothard *et al.*, 2009c). Throughout East Africa, high levels of genetic diversity within individual human hosts may have unknown effects on the pathology and clinical manifestation of the disease in affected individuals. As such, initiatives for determining the prevalence and intensity of *S. mansoni* infections in this region should be combined with genetic analysis; determining the population genetic diversity and structure in this setting will shed light on the transmission dynamics and molecular epidemiology of the parasite.

1.2 *Biomphalaria*

Investigating dynamics of parasite transmission also requires scrutiny of the snails that act as intermediate hosts for *S. mansoni*. These snails, of the genus *Biomphalaria*, pose some different challenges than schistosomes to their study, including taxonomic difficulties, but also benefit from the study of their distribution and genetic structure. These can be achieved through an integrated and multifaceted analysis combining environmental data, spatial analysis and molecular and morphological tools.

1.2.1 Snail classification and evolutionary history

Biomphalaria snails are small (usually < 2cm) freshwater pulmonates (Kingdom Animalia; Phylum Mollusca). The Gastropoda class, comprising of all snails, slugs, limpets and nudibranchs, contains the highest species diversity among Mollusca as well as the widest geographical distribution, with members exploiting freshwater, land and the sea. The class is under constant taxonomic revision, with continual changes even to the nomenclature of the divisions. The currently accepted classification (Grande *et al.*, 2008) is based on 'unranked clades' for divisions above superfamily level but within Gastropoda as a whole; within this system, the old order (and previously a sub-class) of 'Pulmonata', which includes *Biomphalaria*, is re-described as an 'informal group', owing to its suspected paraphyly, and is placed in the Heterobranchia unranked clade (Bouchet and Rocroi, 2005). Within Pulmonata, the Basommatophora, among which *Biomphalaria* are included, are another informal group, characterised by the placement of the eye spots and almost always a lack of an operculum. Within the monophyletic Hygrophila clade, the superfamily Planorboidea has recently been analysed by molecular methods, which resulted in a further reorganisation of the many genera contained within the two main families, the Planorbidae and Physidae (Albrecht *et al.*, 2007). *Biomphalaria* are part of the Planorbidae, also known as ramshorn snails, which are identified as discoidal air-breathing snails that inhabit freshwater, and consist of a number of 'tribes', as per the classification of Bouchet and Rocroi (2005); the *Biomphalaria* belong to the

Biomphalariini tribe, of which *Biomphalaria* is the type genus, and was described in the early years of the 20th century (Preston, 1910).

Despite the description of the genus in 1910, a number of species of *Biomphalaria* had already been described by the end of the 19th century, mostly categorised as the synonym *Taphius* or alternatively as *Planorbis*. The majority of 34 currently recognised species are found in Latin America, with only approximately 12 species present in the Old World. Originally, it was thought that the presence of *Biomphalaria* on both sides of the Atlantic was an ancient relict of the separation of Gondwanaland, approximately 95-106 million years ago. However, more recent molecular analyses revealed an affinity between *B. glabrata*, a New World species, and the African clade, suggesting that the origin of the genus was in fact in the Neotropics, followed by a more recent colonisation of Africa (Campbell *et al.*, 2000). This view is now generally well-accepted; it is theorised that a proto-*B. glabrata* crossed the Atlantic, either by means of a vegetation raft or attached to migratory birds, between 2 and 5 million years ago, and from there radiated eastwards and southwards, and eventually into the Middle East as well (Campbell *et al.*, 2000, Morgan *et al.*, 2001).

Original classification of the African species were based on morphological characters, such as conchological measurements, the shape of the radula and various measurements related to the copulatory organs, namely the shape and relative sizes of the preputium and penis sheath and the number and form of the prostate diverticula (Mello, 1972); see Figures 1.4-1.6.

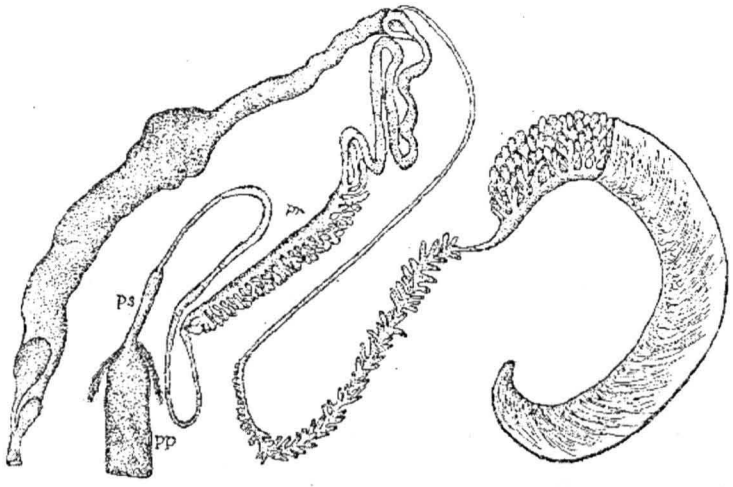

Figure 1.4 – Drawing of the copulatory structures of *B. sudanica*

'pp' = preputium; 'ps' = penis sheath; 'pr' = prostate. The lack of scale is retained from the original published sketch. Note the relative shortness of the penis sheath compared to the preputium and the large number of short prostate vesicles. From Mello *et al.* (1972).

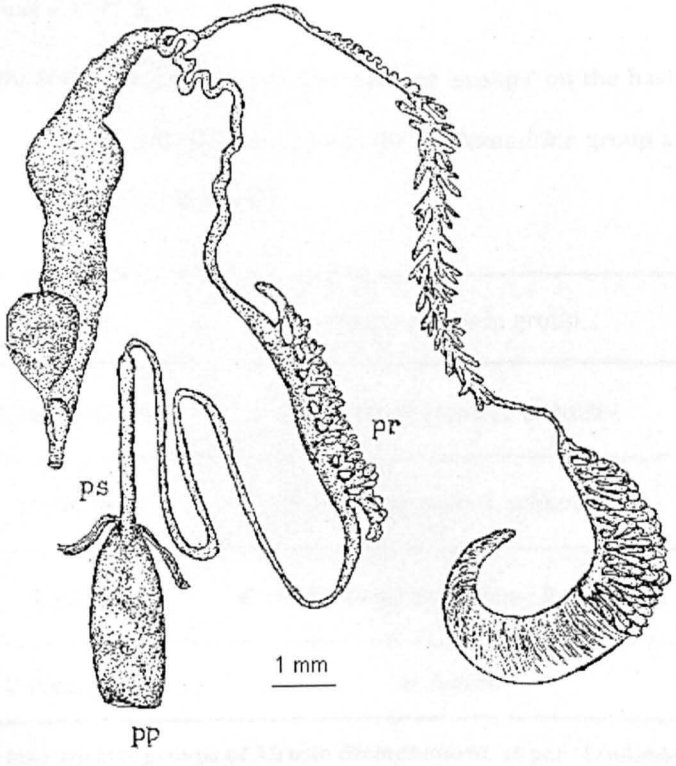

Figure 1.5 – Drawing of the copulatory structures of *B. choanomphala*

'pp' = preputium; 'ps' = penis sheath; 'pr' = prostate. Scale redrawn from original published sketch. Note the long penis sheath as compared to the preputium and the small number of long, densely branched prostate diverticula. From Mello *et al.* (1972).

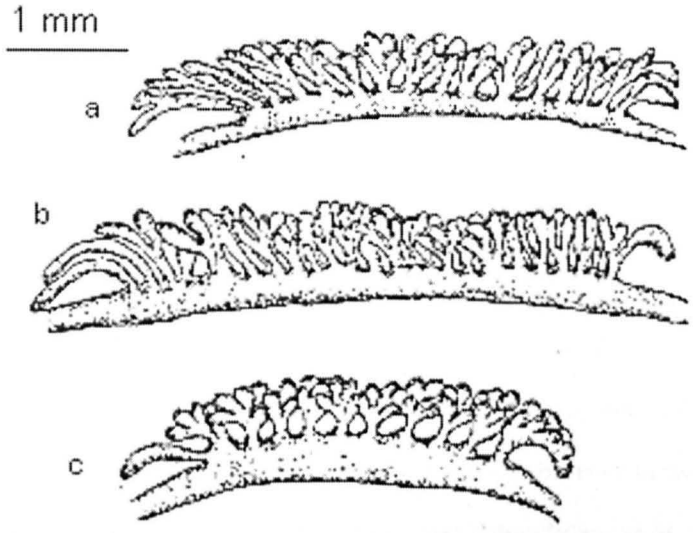

Figure 1.6 – Comparative drawing of the prostate gland and associated diverticula of three *Biomphalaria* species from Africa

a = *B. alexandrina* (Alexandrina, Egypt); b = *B. sudanica* (Kisumu, Kenya); c = *B. choanomphala* (Mwanza, Tanzania). Scale redrawn from original published sketch. From Mello *et al.* (1972).

African *Biomphalaria* were categorised into four species 'groups' on the basis of morphology: the *B. choanomphala* group, the *B. sudanica* group, the *B. alexandrina* group and the *B. pfeifferi* group (Mandahl-Barth, 1957); see Table 1.1.

Group	Other species in group
B. choanomphala	*B. smithi; B. stanleyi; B. barthi*
B. sudanica	*B. camerunensis; B. salinarum*
B. pfeifferi	*B. rhodesiensis; B. arabica; B. rueppelli*
B. alexandrina	*B. boissyi*

Table 1.1 – The four species groups of African *Biomphalaria*, as per Mandahl-Barth (1957)

Later attempts to add further characters to better define the groups were unsuccessful; similarly, early molecular analyses involving enzyme electrophoresis failed to lend weight to the original groupings (Hendriksen and Jelnes, 1980, Brown, 1994). More recent molecular

analysis, involving direct sequencing of three gene fragments (mitochondrial and nuclear) completely contradicted the earlier morphological divisions; instead, it was shown that *B. choanomphala*, *B. sudanica* and *B. alexandrina* were closely related, in a clade termed 'the Nilotic species complex', with *B. pfeifferi* more closely linked to *B. stanleyi*, previously part of the *B. chonaomphala* group as it has a similar shell morphology and is also preferentially lacustrine.

Additional molecular work has confirmed the existence and validity of the Nilotic species complex, although the 'species' divisions within the clade are less robust and often paraphyletic (Jørgensen *et al.*, 2007). Moreover, inconsistencies have been observed between morphological and molecular forms, particularly in the Ugandan Great Lakes (Plam *et al.*, 2008); simply put, identification within the *Biomphalaria*, and especially within the Great Lakes of East Africa, is complex and challenging. Given the biomedical importance of this genus and the potentially differential roles played by various species in the transmission of *S. mansoni*, the taxonomy of *Biomphalaria* in this region urgently requires more attention.

1.2.2 Snail ecology and life history

As mentioned previously, *Biomphalaria* are of biomedical importance; although they are capable of transmitting a variety of parasites, it is as their role as the intermediate hosts of *S. mansoni* that they garner the most research attention. Crucially, it is the distribution of susceptible *Biomphalaria* which largely determines the endemic zone of *S. mansoni*-induced intestinal schistosomiasis in man.

Biomphalaria have been shown to exhibit optimum growth at water temperatures of between 22-27°C, although they can likely tolerate temperatures of between 18-32°C, and even perhaps lower at high altitudes or latitudes (Appleton, 1976, Kristensen *et al.*, 2001). The chemical conditions of the water in which they live has also be shown to vary widely, even across a relatively small spatial scale (Sturrock, 1974). Specific ecological conditions required by *Biomphalaria* are to a certain extent determined by the particular species in question; for example, *B. pfeifferi* is generally found in small ponds and streams with slight turbidity and low

flow whereas *B. stanleyi* and *B. choanomphala* are thought exclusively to inhabit large lakes, often in quite deep water (Brown, 1994). Factors such as temperature, using the ranges as described above, rainfall and Normalized Differences Vegetation Index (NDVI) have been used for large-scale mapping initiatives to predict the distribution of *Biomphalaria* on the ground (Kristensen *et al.*, 2001), but little detailed research has gone into integrating these broad climatic predictors with the more heterogeneous changes in water conditions which may also be important for the persistence of local-level populations of the snails.

The relationship between vegetative cover and the distribution of *Biomphalaria* is not a straightforward one. In the context of predicting large-scale distribution of the snails, NDVI is a proxy for water availability, thus merely conveying the existence of freshwater bodies and the association of vegetation with marshy habitats. This is likely why rainfall has also been correlated with the abundance of *Biomphalaria* in several studies (Goll, 1982, Erko *et al.*, 2006). On the other hand, in larger, permanent water systems, aquatic vegetation, which is not registered by satellite imagery and therefore unaccounted for by NDVI values, may act as important habitat niches for *Biomphalaria*. Vegetation provides food as well as a substrate for laying eggs and avoiding predation, and as such, habitats with reed beds or other species of aquatic plant life might be predicted to hold higher numbers of *Biomphalaria* (Pimentel, 1957, van Schayck, 1985). *Vallisneria spiralis*, an aquatic macrophyte, in particular is noted for being a typical habitat substrate for *Biomphalaria*, as well as other snails (Kazibwe *et al.*, 2006). However, there are reports that in some localities, such as in the southern portion of Lake Victoria, *Biomphalaria* can be found living directly on exposed rock faces along sandy-bottomed portions of the shoreline (Magendantz, 1972). These apparent contradictions in terms of the ecological preferences suggest that for a particular species and region, a more concerted, and potentially fine-scale, investigation would be useful to be able to predict more accurately the presence of *Biomphalaria*, and thus sites which might be transmissive for *S. mansoni*.

Factors other than environmental variables and habitat choice will also affect the demographics and life history of *Biomphalaria*. For example, little is known about the pathogens that may infect *Biomphalaria* and limit population distribution and abundance. Population growth itself

has been shown to be density dependent, with negatively exponential egg production correlated with increasing population density; this is probably due to competitive interactions (Woolhouse and Chandiwana, 1990). In addition, infection with *S. mansoni* has been shown to lead to an immediate increase in egg-laying; whether this is stimulated by the parasite, in order to generate more susceptible snails, or by the snail as compensation for expected future reduction in reproductive capacity is not known. At the same time, if exposed to the parasite but not successfully parasitised, the snails may experience a decrease in fecundity later in their lives (Minchella and Loverde, 1981).

The reproductive strategy of *Biomphalaria* is replete with such trade-offs; as hermaphrodites, different environmental and demographical situations can influence the incidence of selfing versus out-breeding. For example, it has been shown experimentally that social facilitation may occur with *Biomphalaria*, whereby presence of others of the same species induce preferential cross-breeding, with ultimately greater reproductive output (Vernon, 1995). Similarly, habitat heterogeneity can result in selective changes to reproductive mode; for example, in Madagascar, *B. pfeifferi* from transient, yet isolated, habitats may have evolved pure selfing as a means of protecting against frequent local extinction events (Charbonnel *et al.*, 2005). High levels of inbreeding have been reported in other *Biomphalaria* species, such as wild *B. glabrata* in Brazil (Wethington *et al.*, 2007), although in other settings *B. glabrata* has also been shown to out-cross, again emphasising how environment can affect reproductive strategy even within a species (Vernon *et al.*, 1995, Mavárez *et al.*, 2002). The fitness of offspring from out-crossed gametes is thought to be greater, based on predictions of the theories regarding the evolution of sexual reproduction and also, in the case of *Biomphalaria*, through experimental observation of breeding success under various scenarios (Tian-Bi *et al.*, 2008, Sandland *et al.*, 2009). Considering this fitness advantage, it could be hypothesised that the stability of the environment in Africa's Great Lakes would provide a suitable opportunity for high levels of out-breeding among the *Biomphalaria* populations; however, this theory has yet to be tested.

1.2.3 Biomphalaria *distribution in the Americas and Africa*

Based on the optimum temperature ranges stipulated in the previous section, *Biomphalaria* are primarily restricted to the tropics and subtropics, in Africa, the Americas and the Middle East. However, modern introductions have led to populations being established in East Asia and Europe. Moreover, anthropogenic changes to the landscape, and particularly large-scale water management projects, have significantly increased the amount of available habitat able to be exploited by *Biomphalaria*, thus increasing their distribution in a number of regions.

Biomphalaria are thought to have evolved in the Neotropics (Campbell *et al.*, 2000), and in the Americas the genus extends from the southern United States to Patagonia in Argentina, with the most southerly population reported from 42° south of the equator, in Mendoza Province (Paraense, 2001, Ciocco *et al.*, 2008); see also Appendix 13.5 for a recent publication describing the identification of *Biomphalaria* from this province (Standley *et al.*, 2011). In addition, *Biomphalaria* can be found throughout the islands of the Caribbean. In Africa, *Biomphalaria* can be found throughout the continent apart from in the Sahara Desert and at altitudes where night time or seasonal temperatures drop below their developmental minimum; this of course will vary with latitude, and *Biomphalaria* have been found at altitudes of greater than 1500m close to the equator, both in Uganda and in Ethiopia (Erko *et al.*, 2006, Rubaihayo *et al.*, 2008).

The first reports of the presence of *Biomphalaria* snails in East Asia stem from the early 1970s, when populations of *B. straminea* were observed in Hong Kong. Thought to have been brought to the island in aquatic vegetation used in the aquarium trade, the snails have since spread to mainland China (Woodruff *et al.*, 1985a). Molecular research on the population genetics of these populations has subsequently suggested that there may have been two separate introduction events, with the latter occurring in the early 1980s (Woodruff *et al.*, 1985b). Other human-mediated events resulted in the introduction of *B. tenagophila* populations in the Congo (Pointier *et al.*, 2005) and the escape of laboratory strains of *B. glabrata* in Egypt, where they have subsequently hybridised with *B. alexandrina* (Kristensen *et al.*, 1999, Lotfy *et al.*, 2005). These events have all occurred in frost-free subtropical zones, in several cases where other species of *Biomphalaria* were already present. More recently, populations of *B. tenagophila*

have been identified from Romania, very much a temperate climate and further north than any previous reports of the genus' distribution (Majoros *et al.*, 2008). Although currently minimum winter temperatures are likely to be too low for the development of *S. mansoni*, if the parasite were to be introduced, for example by recent visitors to Africa or South America, the possibility of seasonal transmission during the summer months should not be overlooked. Similarly, in the face of imminent and on-going climate change, it is likely that conditions in southern and central Europe will become more favourable for the development of subtropical and tropical parasites, such as *S. mansoni*.

Other anthropogenic activities have also contributed to the increased and currently very widespread distribution of *Biomphalaria* snails. Introductions of modern irrigation systems, involving construction of canals, ditches and dams, have in many locations increased the available habitat for *Biomphalaria* and other intermediate host snails, thus increasing the risk of schistosomiasis transmission. For example, the creation of the Aswan Dam in Egypt has been heavily implicated in the resurgence of schistosomiasis as a public health problem in Egypt; this effect was exacerbated by a cessation in the annual closure of smaller side canals, resulting in stable, year-long populations of snails (Malek, 1976). In Senegal, the construction of the Diama dam prevented the backflow of estuarine water into the Senegal River; it is thought that the influx of brackish water had limited the suitability of the river as a habitat for *Biomphalaria*. After the dam was built, populations of *Biomphalaria* increased rapidly, with corresponding rises in the prevalence of *S. mansoni* throughout the Senegal River Basin (Southgate *et al.*, 2001). Given the concurrent trends in many of these countries of increasing urbanisation and an associated decrease in effective sanitation, it is no surprise that incidence of *S. mansoni* has also been on the rise, despite various on-going control initiatives (Callisto *et al.*, 2005).

The above two examples demonstrate the effect of large-scale, regional changes in the distribution of intestinal schistosomiasis due to anthropogenic influences. However, these effects can also be seen on a much smaller scale. For example, very recent malacological surveys in Cuba have demonstrated that man-made dams and artificial ponds are the most significant sources of *Biomphalaria* on the island (Vazquez Perera *et al.*, 2010). Similarly, in

other host-parasite systems, it has been shown that anthropogenic changes to community diversity may favour the persistence and success of generalist species, which often are those that transmit disease (Resh *et al.*, 1988, Keesing *et al.*, 2006). As such, a detailed investigation into the environmental factors, anthropogenically induced as well as natural, that influence *Biomphalaria* distribution should include an explicit consideration of the varying scales at which these factors operate.

1.2.4 Biomphalaria *genetics and population structure*

The identification of the human-assisted introductions of *Biomphalaria* described above was assisted by the advent of molecular tools, which have greatly expanded the understanding of the evolutionary history and interrelatedness of the various species of *Biomphalaria*. Similarly, as mentioned earlier, the taxonomy of *Biomphalaria* in many cases is confused and confusing, and molecular markers are crucial tools being used to try to elucidate gradually the species boundaries, particularly for the less-well understood African groups. This is particularly important given that some species are susceptible and others appear refractory to infection by the parasite. However, as with *S. mansoni*, in terms of the local dynamics of transmission, compatibility and public health impact of the disease, it is the genetics of the intermediate host snail at the population level that holds the most interest.

Early studies on the population structure of *Biomphalaria* relied heavily on allozyme electrophoresis as a proxy for genetic diversity. Most studies reported low levels of intrapopulation variation, with heterozygote deficiency, with greater differentiation between populations, indicative of low migrations levels, non-random mating and potentially high levels of inbreeding, particularly in *B. pfeifferi* (Bandoni *et al.*, 1990, Vrijenhoek and Graven, 1992, Jarne and Stadler, 1995, Bandoni *et al.*, 2000). *B. choanomphala* and *B. sudanica*, unlike *B. pfeifferi*, showed lower levels of inbreeding, but also insignificant evidence for separate gene pools, suggesting they may belong to a single species, although unfortunately this theory was not explored further at the time (Bandoni *et al.*, 2000). Developments in PCR methods led to the use of RAPD markers for population genetics analyses; these tended to corroborate with the enzymatic findings of high selfing, low intrapopulation diversity but well structured population

with limited gene flow (Webster *et al.*, 2001a, Webster *et al.*, 2001b). However, limitations with RAPD analysis encouraged the application of direct sequencing of variable gene regions, such as the cytochrome oxidase sub-unit one (COI) mitochondrial gene, to population studies. This region was to become the gene of choice for the DNA 'barcoding' movement, which will be discussed in more detail in section 1.4.3. Although designed primarily for species-level identification and taxonomy (Hebert *et al.*, 2004), barcoding also led to the increased usage of COI as a population marker (Hajibabaei *et al.*, 2007, Frezal and Leblois, 2008).

Despite this rise in popularity in other taxa, few population studies have been carried out in *Biomphalaria* using DNA sequence data. Those that did on the whole supported the findings of the enzyme analyses; within-species differentiation is greater for South American species than African, for example, ostensibly due to the more recent evolutionary origin of the latter (DeJong *et al.*, 2003). However, even within African species, similar phylogeographic and population-level studies also revealed high intraspecific genetic diversity in these sequences, indicating their potential usefulness as a population level marker (Angers *et al.*, 2003, Jørgensen *et al.*, 2007, Plam *et al.*, 2008).

However, direct sequencing of a single gene, and particularly a maternally inherited mitochondrial one, does not incorporate consideration of allelic diversity and thus limits the analyses which can be undertaken. The development of primers to amplify co-dominant microsatellite loci in *Biomphalaria* allowed for investigation of more detailed demographical factors, such as Hardy-Weinberg equilibrium, levels of inbreeding and the history of past bottlenecks (Jones *et al.*, 1999, Charbonnel *et al.*, 2000, Mavárez *et al.*, 2000). Again, most studies were in congruence with the allozyme data, revealing highly structured population, with high interpopulation but low intrapopulation diversity and in many cases, strong evidence for inbreeding, either caused by rapid exploitation of a favourable new habitat or transient habitat patches and frequent extinction events (Charbonnel *et al.*, 2002a, Charbonnel *et al.*, 2002b, Wethington *et al.*, 2007, Campbell *et al.*, 2010). This is in contrast, for example, with research on *Bulinus* species, which are intermediate hosts for *S. haematobium* and *S. bovis*, which showed high intrapopulation diversity as well, perhaps due to greater dispersal ability

between populations, although they also showed high levels of inbreeding (Stothard and Rollinson, 1997, Gow *et al.*, 2004, Gow *et al.*, 2007).

It is worth emphasising that virtually all of the above examples of past studies on the population genetics of African *Biomphalaria* have focused on *B. pfeifferi*, as the most widespread species in Africa and arguably the most important in terms of transmission of *S. mansoni*, on a continental scale. However, at a regional level, it is expected that other species, which are currently very understudied in terms of their taxonomy as well as population genetic structure, are locally more important intermediate hosts. One such example of this is likely the Great Lakes of East Africa, as large, stable aquatic ecosystems with varied, and sometimes endemic, *Biomphalaria* species. Moreover, intestinal schistosomiasis is rife throughout the region, and given the high socio-economic importance and human population density around lakes such as Victoria, these are crucial settings in which to investigate more closely the dynamics of *S. mansoni*, *Biomphalaria* and the prospects for control of intestinal schistosomiasis.

1.3 Study site: Lake Victoria

Lake Victoria is the largest freshwater tropical lake in the world and also the largest lake in Africa. Situated in East Africa, between approximately 0.4°N and 2.5°S and 31.6° and 34.5°E, the lake is shared politically between Uganda in the north, Kenya in the east and Tanzania in the south (Figure 1.7). The unique geological history of the lake, together with its cultural and socio-economic importance in the region, make Lake Victoria an ideal natural laboratory for exploring the evolutionary and ecological dynamics of *Biomphalaria* and in turn the transmission of intestinal schistosomiasis.

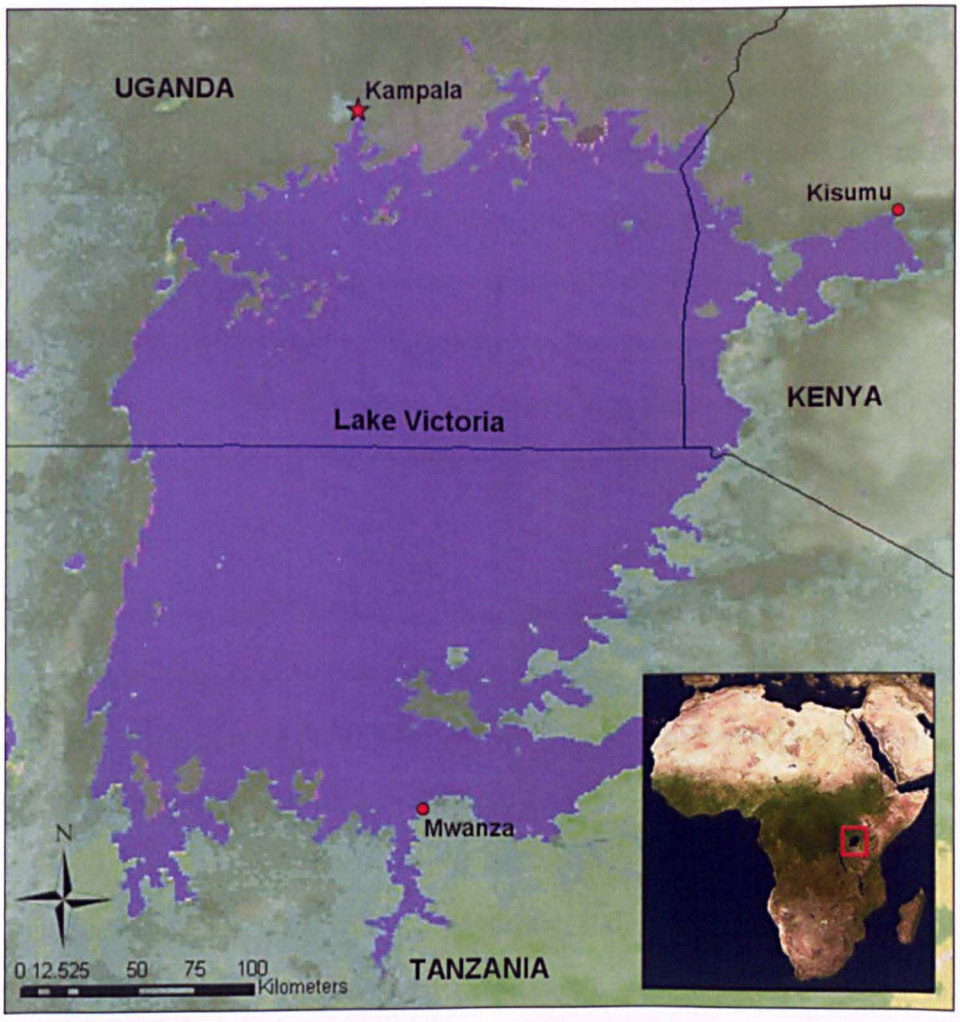

Figure 1.7 – Map of Lake Victoria, showing location within Africa and major political boundaries

1.3.1 Geological and anthropological history

Lake Victoria is unique among the East African Great Lakes in being a comparatively shallow, young lake, not formed as part of the tectonic separation which created the Rift Valley. During the gradual sinking of the Rift Valley, over the last 14 million years, the Kagera and Katonga Rivers flowed eastwards, down into the valley. About 400,000 years ago, a tectonic event pushed up the eastern edge of the rift into a ridge, damming the flow of these rivers and creating the first manifestation of Lake Victoria (Johnson *et al.*, 2009b). Eventually, more uplift caused backflow of this water along the Katonga valley and into Congo, via a gap in the Rwenzori mountains known as the Beni Gap. This gap closed, again due to uplift, about 100,000 years ago, forcing the creation of the Albert Nile as the water drained north, through Lake Albert and into present-day Sudan. Further uplift, potentially as recently as 30,000 years ago, caused the desiccation of the Katonga river and changed the direction of flow of the Kagera; this led to the formation of Lake Kyoga between Albert and Victoria, and also created the present day 'source of the Nile' at Jinja, on the northern shores of Lake Victoria (Beuning *et al.*, 1997, Roberts, 2007). Studies of sediment cores from the lake bottom have suggested several periods of near or complete desiccation throughout this history, with the most recent coinciding with the last global glaciation, between 10,000-15,000 years ago (Johnson *et al.*, 1996, Beuning *et al.*, 1997).

Archaeological evidence such as rock art suggests ancient habitation of Lake Victoria's shoreline and its surroundings by hominids and early *Homo sapiens* (Chaplin, 1974). Indeed, Olduvai Gorge, the site of some of the most famous finds of early hominid fossils, is less than 200km from the present shores of Lake Victoria, across the fertile plains of the Serengeti, where the ancestors to modern humans may have taken their first upright steps (Leakey *et al.*, 1964). Since the last glaciation and until about 500 years of the Common Era (CE), tribes in and around eastern Lake Victoria, in modern day Kenya, are thought to have mainly been hunter-gatherers who also engaged in subsistence fishing, and may have followed migratory herds across short distances for hunting purposes (Dale and Ashley, 2010). Towards the west, pygmy

communities are believed to have inhabited forested regions as early as 40,000 ago, although they left little by way of concrete archaeological evidence.

The first major anthropological change in the Lake Victoria region was the arrival of Bantu tribes from Central Africa around 500 CE. The Bantu were agriculturalists, and lived in permanent settlements; the introduction of the banana as a domestic crop at around this time led to the clearance of huge tracts of rainforest to create ever larger harvests (Roberts, 2007). The arrival of Nilotic pastoralist tribes from Sudan and northern Kenya occurred around 1000 CE (Johnston *et al.*, 1913); although elsewhere many have retained the semi-nomadic lifestyle of their ancestors, by Lake Victoria several of the Nilotic tribes became fisherman, a tradition still carried on by their descendants today.

Currently, the mix of tribes and linguistic groups living around Lake Victoria is extremely varied and heterogeneous. Bantu tribes such as the Haya and the Sukuma inhabit the western and southern shores of Lake Victoria in Tanzania, with the Luo, a Nilotic tribe, the dominant group along the eastern edge of the lake, in Tanzania and Kenya (Ogot, 1967, Maddox *et al.*, 1996). In Uganda, the central portion of the lakeshore is predominantly occupied by the Buganda, the largest tribe in Uganda, who are also Bantu agriculturalists (Roberts, 2007). The urbanisation of towns such as Mwanza in Tanzania, Kisumu in Kenya and of course Kampala, the capital of Uganda, as well as Entebbe further south, has added even greater diversity to the population as people from elsewhere in the region move into the cities.

The first written records of Lake Victoria stem from medieval Arab sources; Swahili and Arab traders used trade routes from the coast into the interior of the continent for exchange of slaves and ivory (Holmes, 1971). The Al-Idrisi map, dated to approximately 1160, shows Lake Victoria in remarkably accurate form, considering a full cartographical circumnavigation of the lake would only be carried out some 700 years later, by Henry Morton Stanley in 1875 (Stanley, 1876, Jeal, 2007). The first European to set eyes on Lake Victoria was John Hanning Speke in 1858, whilst on an expedition with Richard Burton to discover the source of the Nile. Burton did not accompany Speke at this time, and Speke's conclusion that Lake Victoria was in fact the Nile's source was disputed by Burton owing to lack of evidence, causing an enormous rift

between the two men. Speke later returned with James Augustus Grant in 1860 to add further validation to his claim, but in the end, it was Stanley's explorations that finally proved Speke to have been correct (Baker, 1944).

The perambulations of these early explorers were often also accompanied by an interest in the natural history of the plants and animals encountered on their expeditions, who sent many of their finds back to Europe for identification and classification. For example, it was as part of the explorations of Emin Pasha (born Eduard Carl Oscar Theodor Schnitzer, and actually rescued from what is now the Congo during a Mahdist revolt in 1886 by a relief force led by none other than Stanley) to the African interior that the German zoologist Dr Franz Stulhmann first collected gastropod specimens from Lake Victoria. One of these, described by Martens at the Museum fur Natürkunde in Berlin in 1879, became the type specimen of *B. choanomphala. B. sudanica* had already been described (also by Martens) in 1870 from a type locality in Sudan, though Stulhmann brought back other specimens from Lake Victoria which were considered variants of this species and thus awarded lectotype status; this type material is still available for study at the Museum für Naturkunde.

During the colonial period, Lake Victoria was divided between British control in the north and German occupation in the south, in what was then German East Africa. The outcome of the First World War placed the larger portion of the colony into British control, at the same time reverting to its earlier name of Tanganyika. The building of the Mombasa-Kampala railway in 1901 was a major event in the history of the Lake Victoria region, as it enabled much faster, more reliable and greater volumes of trade to flow between the lake and the coast (Roberts, 2007). Although originally artisanal and limited to two species of tilapia, by 1905 fishing had become a major industry in Lake Victoria, due to the introduction of gill nets. Despite the establishment of fishing commissions to attempt to regulate the industry, tilapia fishing all but collapsed in the 1960s, pushing fishermen to catch smaller species instead. The introduction of the Nile Perch and four new tilapia species in the 1950s and 1960s were catastrophic environmentally (see the following section) but allowed for a five-fold increase in the fishery output (Ogutu-Ohwayo, 2001). Fishing is still a major economic activity along the Lake Victoria

shoreline, with large factories for export competing with local, subsistence communities (Abila and Jansen, 1997).

The declarations of independence of Tanganyika, Uganda and Kenya in 1961, 1962 and 1963 respectively opened a new chapter in the history of the Lake Victoria region; the 20th century as a whole has brought significant changes to the socio-economic structure of communities around Lake Victoria as well as the fauna and flora of the lake itself. As the following sections will show, these changes combined to have widespread effects on the dynamics and persistence of intestinal schistosomiasis.

1.3.2 Ecology and natural history

Despite its young age, Lake Victoria has a rich and diverse fauna (Darwall *et al.*, 2005). While not exceptionally speciose, the gastropod assemblages in the lake contain several endemic species but moreover, several groups of biomedical importance. However, recent anthropogenic influences have caused catastrophic damage to the lake's aquatic communities, which in turn may affect the distribution and abundance of intermediate snail species as the system dynamically changes.

Lake Victoria was famous for its incredible species flock of endemic cichlid fishes, most of the genus *Haplochromis* (and thus also often referred to as haplochromids). Given the most recent suggested desiccation of the lake less than 20,000 years ago, this radiation is all the more remarkable for having occurred so rapidly (Johnson *et al.*, 1996). Over 500 different species were described, many inhabiting extremely specific habitat niches; some even specialised in feeding on molluscs, with jaws designed to crush gastropod shells (Slootweg, 1987, Hulsey, 2006). The flora of Lake Victoria has been reported to be similarly diverse, both aquatic and terrestrial, with up to 20 families of freshwater macrophytes alone represented (Omondi and Mwendi, 2006). The islands of Lake Victoria too contain high levels of diversity; for example, the Ssesse Islands, in the western portion of Uganda's segment of Lake Victoria, have been shown to be floristically rich, with 146 genera in total (Ssegawa and Nkuutu, 2006). In terms of invertebrate diversity, Lake Victoria is one of the hot-spots for decapod species richness in East

Africa, and its eastern catchment rivers harbour very high numbers of freshwater and terrestrial molluscs (Darwall *et al.*, 2005).

The number of freshwater mollusc species inhabiting Lake Victoria itself is rather less than in other East African lakes such as Malawi or Tanganyika. 13 genera are thought to be found within the lake, comprising of 28 species, of which 13 are considered endemic (Brown, 1994). The ecological interactions of these different snail species is poorly understood; it has been suggested that *Pila ovata*, an ampullariid prosobranch, may consume *Biomphalaria* egg masses, thus potentially being an effective, and native, form of biological control (Hofkin *et al.*, 1991, Mkoji *et al.*, 1998). It has been theorised also that population abundance of pulmonate snails may be related to the absence of other gastropod species, either due to suppression of reproduction from excretory products or competitive interactions (Jordan *et al.*, 1980). However, these effects have never been formally investigated in the field; the dynamics of the Lake Victorian gastropod assemblages have not been researched, despite the obvious implications for the spread of *Biomphalaria.*

If competitive interactions do exist, then factors that affect the structure of the gastropod community may also impact on the transmission potential for schistosomiasis. For example, it has been shown that snail diversity is lower in human-disturbed sites in Lake Victoria, versus non-disturbed sites (Lange, 2005). This could have potential impact on the transmission of *S. mansoni* (or indeed other schistosomes) if the reduction in diversity allows for greater abundance of *Biomphalaria* and other schistosome host snail, through relaxation of the constraints mentioned above. Similarly, vector species are often considered to be generalists, and able to exploit disturbed environments that are no longer suitable for specialist species; again, if *Biomphalaria* follow this trend being an ecological generalist as well as an intermediate host genus, then they too might be expected to abound in anthropogenically mediated habitats, also increasing the potential for transmission of *S. mansoni.* Given the devastating history of anthropogenic influence on the biota of Lake Victoria, further investigations on the effects of human actions on mollusc community assemblages should certainly be undertaken.

Several anthropogenically-mediated events have contributed to the degradation of the Lake Victoria ecosystem. The introduction of the common water hyacinth (*Eichhornia crassipes*) to Africa probably first occurred in the early 20th century, but it did not become established in Lake Victoria until the 1980s, when it spread rapidly and prolifically (Opande *et al.*, 2004). The plant was assisted in its increase by eutrophication, caused by intensification of agriculture, deforestation and sewage run-off from on-going urbanisation, which increased primary productivity (Verschuren *et al.*, 2002). The overuse of the hydroelectric dams at Nalubaale and Kiira of the Victoria Nile has been reported to have lowered the water level of Lake Victoria by up to two meters between 2000 and 2006, intensifying the effects of eutrophication (Kiwango and Wolanski, 2008). While the extent of water hyacinth cover has been associated with a decrease in the diversity and abundance of some invertebrates, they provide a perfect habitat for *Biomphalaria* snails, which utilise the undersides of the floating leaves as a substrate for laying eggs (Masifwa *et al.*, 2001, Plummer, 2005).

Vertebrate introductions have also had a large impact on the ecology of Lake Victoria. Several species of fish were introduced into the lake in the mid-20th century to revitalise the flagging fishing industry. Of these, the Nile perch had the most significant impact, as a voracious predator of the smaller, native cichlid species (Ogutu-Ohwayo, 1990, Witte *et al.*, 1991, Goldschmidt *et al.*, 1993). The lowered water levels of the lake caused by the hydroelectric dams further dried out many of the papyrus swamps that fringed the lake and provided nursery habitats for cichlid fry (Kiwango and Wolanski, 2008). Recent changes in fishing policy, such as the use of larger-mesh gill nets have reduced predation levels, allowing some recovery of the cichlid flock, although many species were already completely wiped out (Kitchell *et al.*, 1997). There have also been signs of survival of the genus through hybridisation, reducing the number of species but increasing the ability to inhabit more than one habitat niche and thus be more resilient to change. There is also evidence of on-going rapid adaptive change in the surviving species of cichlid, a reflection perhaps of the remarkably fast speciation of the genus' evolutionary origins (Witte *et al.*, 2000). It is worth emphasising again that a key impact of the reduction in cichlid species could be the knock-on effect on gastropods, given that certain haplochromids were specialist molluscivores; in Lake Malawi, over-fishing of natural snail

predators was observed to lead directly to increased numbers of *Bulinus* snails, with a concurrent rise in prevalence of *S. haematobium* (Stauffer *et al.*, 2006). Although difficult to test changes in species composition so long after the Nile perch was introduced, an analysis of the effect of gastropod species assemblage and richness on the presence of *Biomphalaria*, or indeed level of infection with *S. mansoni*, could elucidate the role of ecology in the transmission of the disease.

1.3.3 Intestinal schistosomiasis in Lake Victoria

Schistosomiasis was first brought to attention as a public health problem in the Lake Victoria basin in the 1950s, although the first observation of its occurrence in the region, at least in Uganda, was in 1902 (Nelson, 1958, Odongo-Aginya and Ekkehard, 2008). Schistosomiasis was reported from the Tanzanian shoreline of Lake Victoria, based on more comprehensive surveying, as early as the 1960s (McClelland and Jordan, 1962); the observation that fisherman, given their high levels of water contact, were likely to contract and spread the disease was similarly prescient (Bradley, 1968). Other reports followed throughout the 1960s and 1970s (Webbe and Jordan, 1966, McCullough, 1972); at this time, however, more emphasis was placed on controlling the intermediate host snail rather than community-based public health initiatives. The emergence of PZQ in the 1970s as a safe and effective drug for treatment of schistosomiasis instigated the first concerted surveys for the disease around Lake Victoria, with a view to using the information for control purposes (Katz *et al.*, 1979, Lakwo and Odongoaginya, 1990).

The adoption of morbidity control as a key tenet in the battle against schistosomiasis by the WHO in the early 1980s marked the beginning of the modern era of schistosomiasis surveys (Magnussen, 2003), although the price of PZQ was prohibitive at this stage, at approximately $2 per dose. The expiry of the Bayer-E. Merck patent from 1989-1994 led to a rapid decrease in the cost of the drug, as generic forms were produced. Since then, private-public partnerships with the major pharmaceutical companies producing PZQ made it additionally more affordable, and allowed for the realisation of government-led control initiatives in developing countries, based on mass drug administration with PZQ. In Lake Victoria, Uganda and Tanzania both have

national control programmes, although Uganda's is the more comprehensive. Initialised in 2003 with assistance from the Schistosomiasis Control Initiative (SCI), the programme is run by the Vector Control Division of the Ministry of Health, and is currently operational in 23 districts, including most of those bordering on Lake Victoria (Kabatereine *et al.*, 2004, Kabatereine *et al.*, 2007). Alongside mass drug administration of praziquantel, treatment with albendazole for soil-transmitted helminthiasis was also included. In 2007, the programme was expanded to include greater geographic coverage of treatment by placing additional districts under national control (Kabatereine *et al.*, 2006a). In Tanzania, SCI has also helped to set up a control programme for schistosomiasis and soil-transmitted helminthiasis, in conjunction with the Ministry of Health and the National Institute of Medical Research in Mwanza, although efforts thus far have focused more heavily on urinary, rather than intestinal, schistosomiasis (Clements *et al.*, 2008). To date, Kenya has no national or even district-level programme for control of schistosomiasis; however, localised surveys in particular regions have led to treatment delivery in certain areas, including along the Lake Victoria shoreline (Handzel *et al.*, 2003).

Although morbidity control and prevalence reduction are currently the main aims of control measures in Lake Victoria, elimination of schistosomiasis should be the long-term aspiration. However, one aspect of schistosomiasis transmission dynamics is often overlooked in an African context, and yet may prove very important for control initiatives: the role of reservoir hosts. In the Neotropics, rats are the most common definitive host of *S. mansoni*, and have been critical in maintaining the existence of the parasite in the face of improved public health and sanitation (Martins, 1958, Théron *et al.*, 2004, Gentile and Oliveira, 2008). In Lake Victoria, rodents have also been shown to be infected with the disease (Hanelt *et al.*, 2010), but also with other, potentially zoonotic, species of schistosomes, such as *S. rodhaini* and the newly discovered *S. kisumuensis*. Although *S. rodhaini* is not generally thought to be able to be transmitted through humans, it is known to hybridise naturally with *S. mansoni*, producing a potential zoonotic agent. In general, the distribution and prevalence of *S. rodhaini* has been understudied in Lake Victoria, and given the possibility of transmission of both it and its sister species *S. mansoni*, warrants more research effort.

Similarly, research on other potential terminal hosts of intestinal schistosomiasis in Lake Victoria has been extremely limited. Baboons and chimpanzees have both been shown to be experimentally susceptible to infection with *S. mansoni*, and in a limited number of settings, have also been observed with natural infections, although never from this region (Fenwick, 1969, Sadun and Gore, 1970, Nutter, 1993, Bakuza and Nkwengulila, 2009). Lake Victoria moreover contains two islands which are inhabited by chimpanzees: Rubondo Island in Tanzania and Ngamba Island in Uganda. The former, containing a wild, semi-habituated population, has been surveyed for parasitic nematodes but no observations of infection with *S. mansoni* were made at that time (Petrželková *et al.*, 2010). Ngamba Island functions as a sanctuary for wild-born, semi-captive chimpanzees; these animals have never been surveyed for intestinal schistosomiasis, despite reports of frequent water contact (L. Mugisha, pers. comm.) and being located in a highly endemic, high-risk area for the disease (Kabatereine *et al.*, 2004, Clements *et al.*, 2006, Brooker, 2007). Without a thorough examination of the potential role of non-human definitive hosts for schistosomiasis, efforts towards the control and eventual elimination of intestinal schistosomiasis in Lake Victoria, and elsewhere, are likely to be thwarted.

1.4 New opportunities for combined research into *Biomphalaria* and *S. mansoni*

The opportunity to add to the existing data on *S. mansoni* from a highly endemic region such as Lake Victoria, as well as increase our understanding of the taxonomically-confusing *Biomphalaria* snails in the lake, should not be overlooked. On a basic note, detailed maps of the distribution of *S. mansoni* and *Biomphalaria* have never been produced for Lake Victoria as a whole; country-level efforts tend to leave geographical gaps in sampling, and moreover are difficult to directly compare due to differences in methodology, expertise and research objective. Similarly, population structures of *S. mansoni* and *Biomphalaria* have not previously been considered in conjunction with each other in this region, let alone in Africa as a whole; the relatively recent development of certain molecular aids and tools has greatly facilitated the prospects of accomplishing such a study. This section outlines some of the spatial, molecular and morphological tools that can be combined for a powerful multidisciplinary approach to pushing forward research on the dynamics of *S. mansoni* and *Biomphalaria* in Lake Victoria.

1.4.1 Geographical information systems (GIS) and spatial statistical tools for mapping disease and intermediate host distributions

While maps of both *S. mansoni* prevalence and *Biomphalaria* distribution have been made in the past (de Kock *et al.*, 2004, Brooker *et al.*, 2009), rarely have the two been explicitly considered together, and in these few examples, the aim is usually to predict distribution of parasites based on known presence of snails, or even just suitable habitat for snails (Kristensen *et al.*, 2001, Stensgaard *et al.*, 2006). Predictions are usually accomplished through statistical models incorporating environmental data, usually obtained through remote sensing satellites, which are then correlated to known snail or parasite distributions. The models can be created using a variety of different approaches; a common, yet powerful, tool is the use of geographical information systems (GIS) software, such as ArcGIS (ESRI (UK) Ltd., Aylesbury, UK). The benefit of this programme is its visual representation of the data; prevalence or distribution points can be plotted onto a georeferenced landscape, where geographical features such as political

boundaries, river systems, elevation contours and even vegetative cover can also be seen. These 'layers' can then be used for multivariate statistical analysis to determine which factors are correlated with the data values. If environmental variables are known for the whole area, the best-fit model of the data, including those environmental factors that were found to be significant, can be applied to the whole landscape, creating a predictive map. These maps are particularly important in data deficient regions, and can be used by policy-makers and practitioners to estimate the transmission risk and coordinate control efforts at a particular locality (Brooker and Michael, 2000, Abdel-Rahman *et al.*, 2001, Carvalho *et al.*, 2010). They have also been used, though less commonly, to predict distributions and abundances of intermediate host snails in several regions of Africa (Kristensen *et al.*, 2001, Stothard *et al.*, 2002, Stensgaard *et al.*, 2005, Stensgaard *et al.*, 2006, Simoonga *et al.*, 2009). If environmental variables are not known across the whole landscape, one of the variables can be weighted and used for non-linear, non-parametric interpolation in a process known as 'kriging', also producing a predictive map based on variation in those data values. A disadvantage with this type of interpolation is that the goodness of fit of the resultant model cannot be calculated, nor the error range of the predictions, and it relies on an assumption of autocorrelation between the data points (Oliver and Webster, 1990, Guimaraes *et al.*, 2010).

Testing for spatial autocorrelation is important for investigating the connectivity of the landscape, and also can enhance understanding of the scale at which environmental and ecological factors are influencing prevalence of disease or distribution of intermediate hosts. Spatial autocorrelation can be tested for in many statistical software programmes, by comparing every pair of data points with every other, and seeing whether there is a relationship between geographical distance and similarity between the values of the data points, be they a measure of prevalence of intestinal schistosomiasis or abundance of snails. Semivariograms are a useful graphic for presenting these data; the asymptote of the line represents the distance at which point the data become decoupled spatially, that is to say they lose autocorrelation. While testing for autocorrelation has been applied reasonably extensively to parasite prevalences in various parts of East Africa (Clennon *et al.*, 2006, Brooker, 2007, Brooker and Clements, 2009), the particular geography of Lake Victoria, with its myriad

archipelagos of islands and convoluted shoreline, presents a unique opportunity to examine spatial correlation in a heterogeneous landscape. Moreover, spatial autocorrelation has rarely been explicitly tested for in *Biomphalaria* distributions, and not at all, in the author's knowledge, in Africa (Guimaraes *et al.*, 2009, Guimaraes *et al.*, 2010).

Another method involving spatial statistics that has become popular both with theoreticians as well as field workers has been the application of Bayesian inference probability testing of spatial statistics. The basic approach is similar to that described above for basic mapping and kriging; environmental variables, usually derived from remote sensing data, are associated with prevalence or distribution data points and the two are correlated to look for significant values. However, the methodology behind the calculations is based on Bayesian probability theory; this works by creating a posterior probability of a particular distribution, based on a prior estimation of the probability taken together with the data points and any number of potential correlates, such as environmental data. These models allow for the estimation of parameters even for complex models including many covariates, while also explicitly integrating spatial structure and allowing also for a random effects parameter. The levels of uncertainty in the predictive output can also be calculated, allowing for evaluation of the success of the model. These models use extensive processing power and as such it is only in the last decade that computers have been capable of running these complex models for high numbers of data points; as it stands, there remain theoretical constraints to analysing greater than 1000 sites within a Bayesian framework, which is providing a rich vein of cutting edge statistical research (P. Vounatsou, pers. comm.).

In the Great Lakes, data from surveys over a number of years have been successfully incorporated into Bayesian explicative as well as predictive maps of schistosomiasis prevalence; however, due to data deficiencies, the maps' coverage for large portions of the Lake Victoria lakeshore is poor (Clements *et al.*, 2006, Clements *et al.*, 2010). These models, as with those developed through conventional multivariate and GIS models, require ground-truthing and validation, through follow-up parasitological surveys in areas covered by the maps, in order to test their accuracy (Standley and Stothard, 2010). Moreover, as with spatial

autocorrelation, no studies have yet applied Bayesian spatial statistics to the distribution of *Biomphalaria*, despite the obvious benefits this would have for creating accurate predictive maps of these intermediate host snails. As such, detailed mapping of schistosomiasis prevalence and *Biomphalaria* distributions in Lake Victoria and its environs could be analysed within the context of some of these relatively new and, in the case of the snails, under-utilised geographical tools, to validate and greatly improve the existing data from the region.

1.4.2 Advent of Whatman® FTA indicator cards for storage of S. mansoni *larval stages* in situ *and later molecular analysis*

Another important technological development, within the context of combined molecular analysis of *S. mansoni* and *Biomphalaria*, was the introduction of commercially available FTA indicator cards. Produced by Whatman plc (Maidstone, UK), a subsidiary of GE Healthcare, they allow for the storage of even small amounts of DNA for long periods of time, without need for refrigeration. In this way, for the first time, larval stages of parasites could be individually collected in the field, at their point of observation, and later sequenced.

Previously, the molecular analysis of wild isolates *S. mansoni* DNA had relied almost exclusively on the successful passage of material through a laboratory system (Sorensen *et al.*, 2006); miracidia would be hatched in the field and used for infecting susceptible snails, which were then transported back to the lab. Once infections were patent, the cercariae could be used to infect laboratory mice, which could later be euthanised to collect adult worms for molecular analysis (Smithers and Terry, 1965). Similarly, if snails were found shedding in the field, these could be directly brought back to infect mice. Worms were stored in ethanol or liquid nitrogen, depending on the type of molecular analysis require, such as enzyme electrophoresis in early studies, replaced later by direct sequencing and microsatellite analysis (Fletcher *et al.*, 1981, LoVerde *et al.*, 1985, Minchella *et al.*, 1995, Thiele *et al.*, 2007). Alternatively, DNA could be extracted from whole snails and then used in enzyme electrophoresis, which could detect the presence of parasites and give an indication as to the species and even population (Wright *et al.*, 1979). There were various problems associated with these methodologies; in the case of lab passage, the process of subsequent infection of snails and mice was suspected to influence the

genetic structure and diversity of the parasite sequences through applying selective pressures. Transport of snails requires extensive permits for both export and import, and snail mortality *en route* can greatly reduce the number of isolates that finally are established in the laboratory. Finally, although suitable for detection of infections, enzyme electrophoresis of snail DNA can only give limited detail with regards to the population diversity and structure of the parasite.

The use of FTA cards can overcome all of these obstacles. The cards are impregnated with a chemical which lyses cell membranes and suspends the larva's DNA within a stable matrix. This allows the sample to be stored, at ambient temperature, for long periods of time (years and even decades, according to Whatman's FTA card datasheet: www.whatman.com/References/51613rerevised.pdf.) Moreover, the breakdown of the cell membranes and the stability of the matrix allows for more efficient extraction, so that even a single larva can be amplified and sequenced. Finally, because the cards are light, easy to transport and resistant to changes in temperature, they are easily utilised in the field (see Figure 1.8).

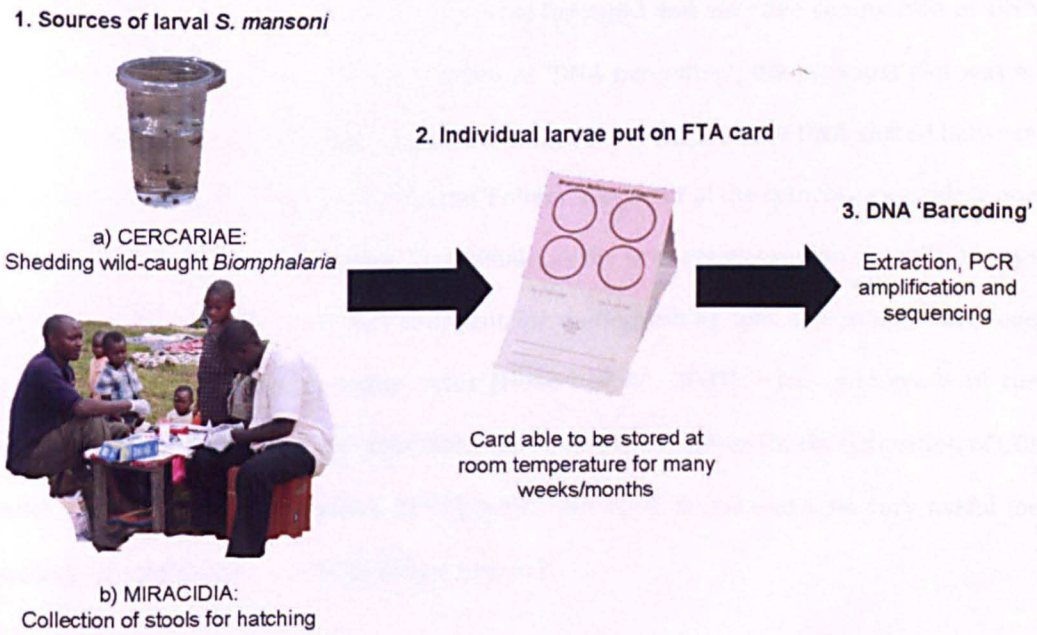

Figure 1.8 – Schematic showing the mode of use of an FTA card, from field to laboratory

This means that cercariae from snails and miracidia and eggs from stool can be stored *in situ*, and thus associated from the outset with a particular locality without later changes to

population structure or diversity through the impact of passage. FTA cards were first developed over ten years ago and successfully tested under laboratory conditions for extraction and processing of DNA (Hsaio *et al.*, 1999). The method was rapidly exploited by parasitologists; the cards gained popularity within the malaria research community first, before being adopted by scientists working on trypanosomes as well (Picozzi *et al.*, 2002). Their aim was primarily diagnostic; PCR of a blood spot on an FTA card with parasite-specific primers proved to be a very sensitive indicator of infection. However, in the context of schistosomiasis, FTA cards were used from the outset for analyses of diversity and population genetics; early testing showed they were a viable, and ethically preferable, alternative to using only lab passaged material (Gower *et al.*, 2007). Since then, the body of work relying on FTA material has steadily grown, and it remains a crucial innovation in on-going field-based research on *Schistosoma* molecular epidemiology.

1.4.3 'DNA barcoding': A methodology for rapid and effective molecular comparisons

The emergence of FTA technology for collecting larval schistosomes was particularly timely as it coincided with the emergence of new protocols for rapid and effective comparison of DNA sequences between and within species. Known as 'DNA barcoding', the principal aim was to establish a method of easily identifying species, using a short sequence of DNA shared between all, or at least most, organisms. In practice, the 'Folmer' fragment of the cytochrome oxidase one gene was identified as a suitable region for animals, as the primers worked on virtually all taxa tested and the level of variation was sufficient for distinguishing and demarcating between even closely related species, in many cases (Hebert *et al.*, 2003). The emergence of the barcoding movement also led to the establishment of online databases for the deposition of COI barcodes (Ratnasingham and Hebert, 2007), which are open access and thus very useful for comparison of contemporary data with past research.

A number of different studies have looked at different fragments of the COI gene in *S. mansoni* and *Biomphalaria*, including the Folmer fragment (Campbell *et al.*, 2000, Jørgensen *et al.*, 2007), although for *S. mansoni* it has been shown that even a shorter fragment, called the ASMIT fragment, is sufficient given its high genetic variability (Morgan *et al.*, 2005, Stothard *et al.*,

2009c). However, a key benefit of barcoding is the ability to sequence the same fragment of DNA from various life stages of an organism; given the availability of *S. mansoni*'s various life stages with the introduction of FTA cards, barcoding would be an ideal approach to take advantage of these collections. Moreover, in this case, barcoding provides an opportunity to compare the same gene between *S. mansoni* and *Biomphalaria*, as a way of elucidating their concurrent molecular dynamics from the same region. Figure 1.9 demonstrates how the barcoding procedure can be applied to these two taxa, from the identification of an appropriate fragment of the COI gene to amplification and sequencing.

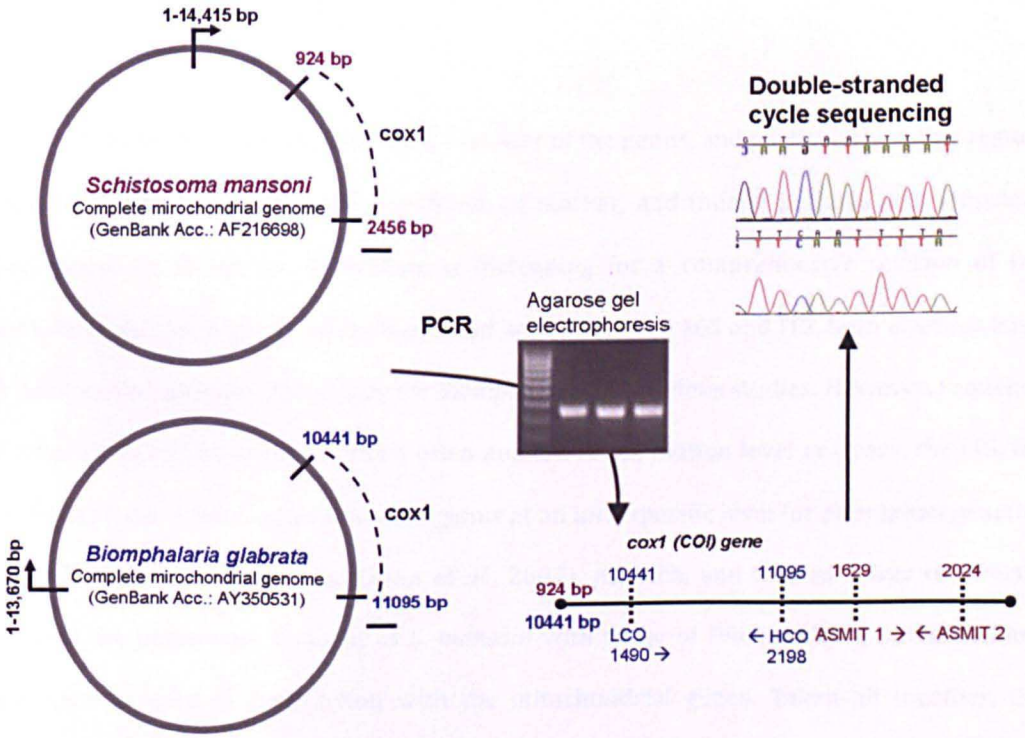

Figure 1.9 – Schematic showing the DNA barcoding methodology, as it could be carried out for *S. mansoni* in conjunction with *Biomphalaria*

Redrawn partially from Stothard *et al.* (2009c).

This could be a particularly important methodology for *Biomphalaria*, as an explicit aim of the barcoding rationale is to be able to distinguish between potentially cryptic species (Hebert *et al.*, 2004); if this is the case for *Biomphalaria*, then barcoding may go some way to resolving the taxonomic difficulties of the genus in Lake Victoria.

This is not to say that other markers should be overlooked when considering molecular analysis of *S. mansoni* and *Biomphalaria*. Indeed, other genetic regions are likely to be highly useful in supporting data obtained by the COI analysis; likewise, the limitations of mitochondrial DNA can be complemented by also analysing nuclear markers. For *S. mansoni*, the existing COI database for samples from East Africa and specifically the Great Lakes region is reasonably comprehensive; moreover, there are no pressing taxonomic questions regarding this species, and so COI can stand by itself as a mitochondrial marker. Rather, the research questions regarding *S. mansoni* in Lake Victoria are more closely aligned to population level analyses; as such, including microsatellite data alongside COI would be a way of strengthening the results.

For *Biomphalaria*, the known taxonomic difficulties of the genus, and particularly in this region, suggest that sequencing a second mitochondrial marker, and indeed perhaps also a nuclear region, would be beneficial in addition to barcoding for a comprehensive revision of the species. Examples of genetic regions that could be selected are 16S and ITS, both of which have both been used reasonably frequently for *Biomphalaria* in previous studies. However, sequence data from these regions have not been often applied to population level analyses; the ITS, for example, is rarely variable enough in this genus at an intraspecific level for population genetics analysis (Vidigal *et al.*, 2004, Jørgensen *et al.*, 2007). As such, and also as a way of directly comparing the population patterns of *S. mansoni* with those of *Biomphalaria*, microsatellites should also be used in conjunction with the mitochondrial genes. Taken all together, the barcoding methodology holds great promise for molecular analyses of *S. mansoni* and *Biomphalaria*, particularly if other markers are also integrated into the analysis.

1.4.4 *Addition of morphology for taxonomic clarification of* Biomphalaria

While fortunately the taxonomy of *S. mansoni* in Lake Victoria is not under question, the same cannot be said for *Biomphalaria*. Molecular advances, and particularly the barcoding approach, have provided an opportunity to use molecular data to resolve even morphologically cryptic species. However, without even considering the morphological it might be difficult to interpret certain molecular findings; for example, genetically similar populations of organisms might

have recently diverged morphologically, creating reproductive barriers which will eventually lead to speciation. Alternatively, a continuum of morphological types might be observed, suggesting instead that there is plasticity in the characters being analysed, and so further research might look into possible drivers of this plasticity, such as environmental or anthropogenic change (Millien *et al.*, 2006, Teske *et al.*, 2007, Hendry *et al.*, 2008). Figure 1.10 demonstrates the ways in which molecular and morphological data could work in synergy to resolve the taxonomy of cryptic samples.

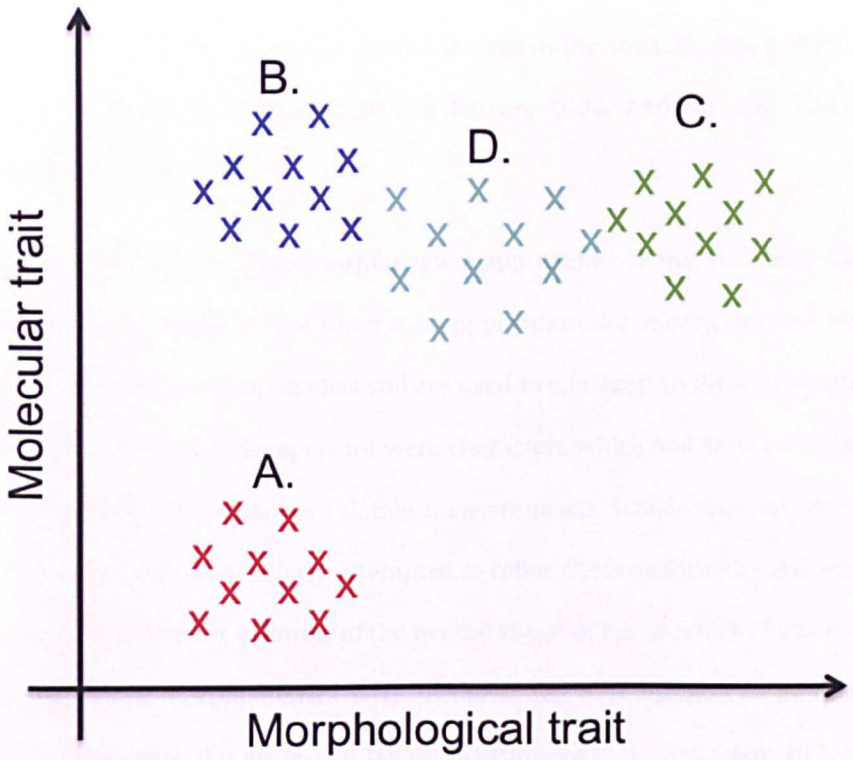

Figure 1.10 – Schematic showing how molecular differences can resolve morphological similarities and vice versa

A and B might look indistinguishable morphologically yet are quite distinct genetically; these are examples of 'cryptic' species. Conversely, B and C may not be widely different in terms of a given molecular trait but have functional morphological differences which restrict interbreeding and therefore are good species; it would be expected for these species to diverge genetically over time. However, if further sampling revealed a group of individuals similar to group D, which are morphologically intermediate to B and C and genetically similar to both, then it could be that the samples represent one species with a range of morphological forms, perhaps manifested by environmental or habitat cues.

Regarding the taxonomy and classification of *Biomphalaria*, traditional morphology did have some success in originally separating out suspected species based on shell shape. Internal anatomy was also considered a useful character; in some species, the presence or absence of a particular character or even organ has indeed been a valuable aid to classification (Paraense, 1981). However, for relative measurements, the mode of specimen preparation may have a large effect, particularly on the preservation of the soft tissues and reducing the reliability of these measurements (Mandahl-Barth, 1957). Other traditional gastropod characters such as the shape and organisation of the radula have not been extensively analysed for *Biomphalaria*, and moreover, are generally less commonly used among modern taxonomists problems with variation relating to confounding factors such as the age of the snail, its diet, and the substrate on which it lives (Peters, 1978, Nybakken and Perron, 1988, Andrade and Solferini, 2006, Andrade and Solferini, 2009).

The weaknesses of these traditional morphological approaches to the questions surrounding *Biomphalaria* taxonomy suggests that there is an opportunity for testing out new methods on this genus, even if these novel approaches still are used to consider 'traditional' characters. For example, the shape and size of the aperture were characters which had been considered useful in the past, but relied on numerous relative measurements, which can run into problems relating to scale. Later researchers have attempted to refine these analyses by also adding angle measures, to try to get a better estimate of the overall shape of the aperture (Plam *et al.*, 2008). Modern innovations in morphometrics have included the development of outline analysis, which implicitly combines the angle and length measurements of the aperture into a single measurement. While originally developed for closed shapes, the emergence of eigenshape analysis as one of the main methodologies within the umbrella of outline analysis has also allowed for open outlines, making such approaches suitable for *Biomphalaria* apertures (MacLeod, 1999). Eigenshape analysis is based on setting a number of landmarks, starting from a fixed point; the deviation of each landmark to the next is compared to that of a perfect circle, creating a different pattern of variation from the circle for each sample (Lohmann and Schweitzer, 1990). This removes any distortion of the outcomes based on variation in size between specimens as well as rotation, leaving the analysis applied purely to shape alone. This

type of analysis has never been attempted before for *Biomphalaria*; taken together with the recent developments in geostatistical programmes as well as molecular methods, the opportunity to apply these tools to the questions surrounding the taxonomy and population structure of these snails as well as *S. mansoni* in Lake Victoria should not be overlooked.

1.5 Research aims

This general review introducing the biology of *S. mansoni*, *Biomphalaria* snails and the dynamics of diagnosis, treatment and control in Africa has emphasised a number of key areas where multidisciplinary and integrated research would greatly enhance our understanding. Here, the research aims of this Ph. D., as described in this thesis, are laid out, on a chapter-by-chapter basis. Figure 1.11 demonstrates how the various subject areas brought together by this research are combined in different ways, producing each of the individual chapters, which in a wider context can be used for informing health policy and control initiatives.

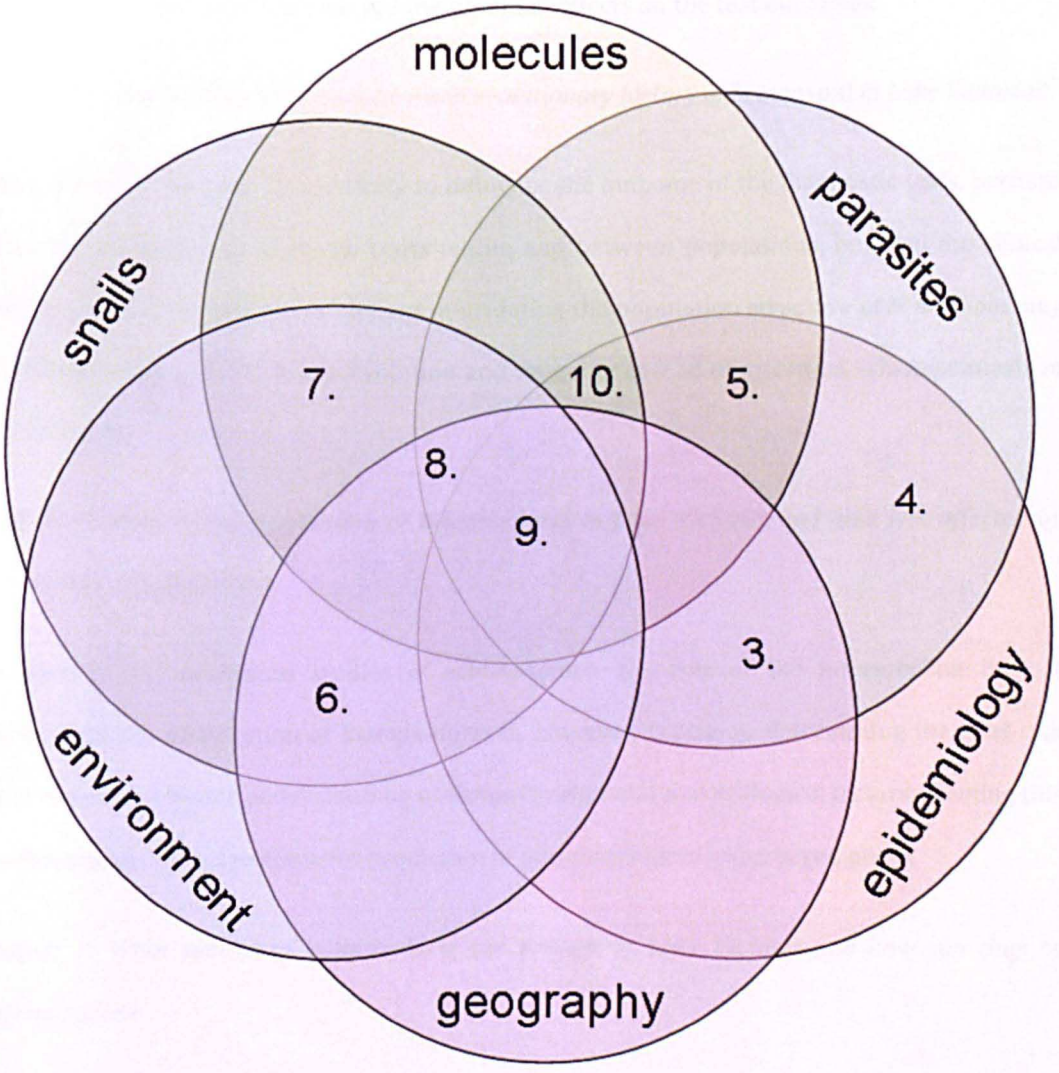

Figure 1.11 – Schematic showing the six main themes relating to the research in this PhD
The numbers refer to the different chapters of the thesis, positioned to demonstrate the combination of disciplines brought together to produce those results.

Chapter 3: What is the present distribution of intestinal schistosomiasis in Lake Victoria?

A better appreciation of the exact distribution of the disease is required, especially to tailor local treatment to community needs for maximum efficiency of the control initiative. In addition, these data can assist in identifying locations where transmission may be elevated.

Chapter 4: How do different diagnostic tools affect rapid mapping efforts?

Different diagnostic tools may result in widely varying estimates of levels of disease prevalence and intensity of infection; as mentioned above, this will impact on the perceived treatment needs of a community. Tests should also be trialed in multiple locations, given different levels of local endemicity of the disease and the potential effects on the test outcomes.

Chapter 5: What is the genetic structure and evolutionary history of S. mansoni *in Lake Victoria?*

The genetics of the parasite are likely to influence the outcome of the diagnostic tests, perhaps based on the spread of heritable traits within and between populations, but also the clinical manifestation of the disease in humans; elucidating the population structure of *S. mansoni* may additionally give clues as to the evolution and on-going spread of intestinal schistosomiasis in Lake Victoria.

Chapter 6: What is the distribution of Biomphalaria *in Lake Victoria, and how is it affected by environmental conditions?*

So often in parasitological studies of schistosomes, the role of the intermediate host is overlooked; the distribution of *Biomphalaria* is, however, crucial to determining the local risk for *S. mansoni.* A better understanding of the environmental and ecological factors affecting this distribution is needed to allow for prediction of potential risk in unsurveyed areas.

Chapter 7: What species of Biomphalaria *are present in Lake Victoria and how can they be differentiated?*

The current taxonomic status of *Biomphalaria* in East Africa as a whole, and specifically in Lake Victoria, is confused and confusing. Considering anecdotal evidence of variations in

susceptibility to *S. mansoni*, solving the species problem could shed light on the relative importance of the different species as intermediate hosts.

Chapter 8: What is the population genetic structure of Biomphalaria sudanica/choanomphala *in Lake Victoria, and how does this relate to parasite transmission?*

The population structure of *Biomphalaria* could yield important data as to the impact of infection with *S. mansoni* on host diversity and natural history, within a host-parasite evolutionary context.

Chapter 9: How do micro-scale environmental factors, including anthropogenic disturbance, affect the abundance and genetic structure of Biomphalaria*?*

The importance of scale in both parasitological and environmental research is well appreciated; as such, do the patterns observed at the level of Lake Victoria hold at a local scale? In this setting, can the effects of human influence on the landscape also be detected in terms of the abundance of *Biomphalaria*?

Chapter 10: Are non-human primates (i.e. chimpanzees) at risk from anthropozoonotic transmission of S. mansoni*?*

Virtually all the research attention on intestinal schistosomiasis focuses on humans as definitive hosts. However, other animals may be at risk of infection as well; in the case of endangered animals, this has conservation implications. Moreover, non-human hosts may act as reservoirs for the disease, maintaining zoonotic cycles that further endanger human communities.

These inter-related, yet diverse, research aims are best approached using a multidisciplinary arsenal of methods, combining those most often utilised in parasitology, malacology, spatial geography, statistics, population genetics and phylogenetics. In particular, the geographical patterns of surveys for parasites and snails will be explored in more detail using spatial statistical tools as well as molecular methods. All of the different procedures and protocols

which contributed to the acquisition of the results presented in the later chapters are described in detail in the following Methods section.

2 Methods

This section gives details for the majority of the methods common throughout the research presented in subsequent chapters. As such, details of these procedures will not be repeated in full, as the relevant sections of this chapter will be referred to. Where necessary, additional information will be provided later, for example to distinguish changes in methodology between earlier field surveys and subsequent ones.

2.1 Field work, sampling design and survey methods

Fieldwork was conducted on four separate missions to Lake Victoria: February-March 2008 (Uganda), June 2008 (Tanzania), January-February 2009 (Tanzania and Kenya) and January-February 2010 (Uganda). These field expeditions will be referred to throughout the thesis by the country visited and the year in which the survey took place: i.e., 'Uganda 2010 mission'. Each trip amounted to between three to four weeks of surveying and had specific objectives with a common theme of sampling and collection of biological material specifically for this research study. Figure 2.1 demonstrates the combination of methods performed at each site, and how intra- and intersite comparisons were made using molecular analyses as well as through spatial statistics. All these specific procedures will be described in full in the following sections.

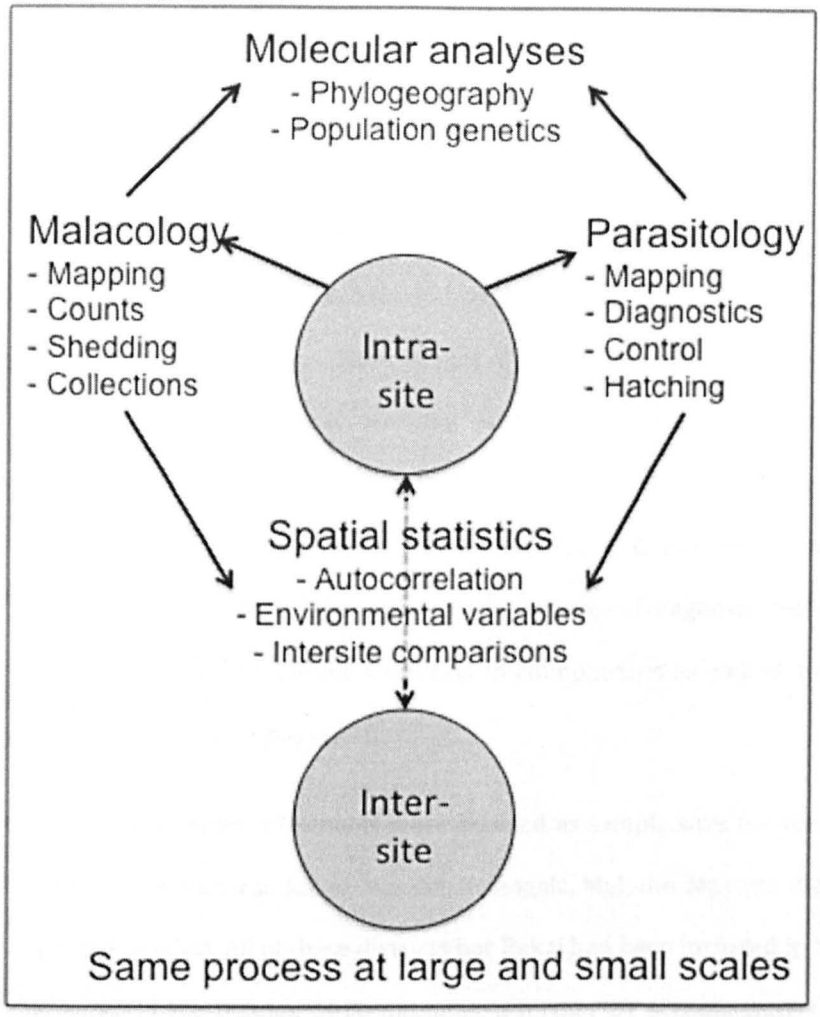

Figure 2.1 – Schematic of intra- and intersite methods, combining malacology and parasitology

2.1.1 Epidemiological and parasitological surveys

Parasitological surveys in all three countries surrounding Lake Victoria were undertaken in lakeshore schools and communities and conducted with assistance from local health officers and community mobilisers. To avoid bias, schools and landing sites were chosen on an *ad hoc* basis, primarily to prevent selection interference that might be based on *a priori* performance or knowledge of the school or locality on the part of the district health officer. At each selected site, in accordance with modified WHO rapid assessment protocols (Montresor *et al.*, 1998, Brooker *et al.*, 2005), approximately 15 boys were selected by the head teacher from primary school classes 1, 2 and 3. This corresponded roughly to ages 6 through 10. Boys were typically targeted to optimise detection of schistosomiasis, as they are more likely to be infected than girls (Kapito-Tembo *et al.*, 2009), but where there were insufficient number of boys to make up

the desired 15 children, girls were also accepted for the survey on an *ad hoc* basis. Each child provided a single stool sample and single urine sample for processing, as well as fingerprick blood for the Uganda 2010 survey. Please refer to section 2.1.3 of this chapter for information on consent and ethical clearance for the surveys.

The only exception to the procedure of school and community-based surveys was on Ngamba Island, during the Uganda 2010 mission. Here, as part of a comparative study between Ngamba and Kimi Islands, adult staff members working for the Chimpanzee Sanctuary Wildlife Conservation Trust (CSWCT) were examined for schistosomiasis and other parasitic infections as well as given a full clinical consultation by a medical doctor. Chimpanzees on this island were also surveyed for the same parasitic infections using a range of diagnostic techniques. For full details of the methodology for sample collection in chimpanzees as well as the results of this comparative study please see Chapter 10.

During the Uganda 2008 mission, 27 schools were selected as sample sites for study, from six districts bordering on Lake Victoria: Rakai, Masaka, Kalangala, Mukono, Mayuge and Busia (see Chapter 3 for map and results). All of these districts bar Rakai had been included in the national treatment programme. In the Uganda 2010 missions, a further 61 schools were surveyed as part of a larger Vector Control Division/Global Network for Neglected Tropical Diseases (GNNTD) survey of the Ssesse Islands, in Kalangala District (see Chapter 4 for map and results). Kimi Island, in Mukono District, having been visited in the Uganda 2008 mission, was resurveyed in the Uganda 2010 mission as well. In the Tanzania/Kenya 2009 mission, a total of 11 schools across the eastern Lake Victoria shoreline were visited: five of these were in northern Tanzania (in the Mara region) and the remaining six were in the Nyanza province of western Kenya (refer again to Chapter 4 for map and results). In 2008, 11 sites were surveyed across the southernmost shoreline of the Lake, in Tanzania, although owing to logistical limitations patient sample sizes were smaller and so the results are presented separately (Chapter 4). All sites were georeferenced with handheld geographical positioning system (GPS) devices (Garmin Ltd., Kansas City, USA) for later spatial analysis; see section 2.2.

2.1.2 Demographical surveys

In addition to the collection of stool and urine samples, each of the surveyed children was asked to answer a short questionnaire intended to provide key demographical data (Figure 2.2 shows the questionnaire process, in a school during the Uganda 2008 field mission).

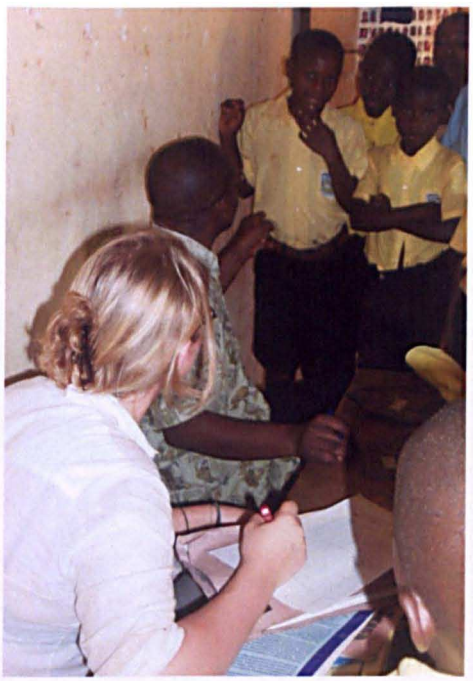

Figure 2.2 – Demographical questionnaires conducted in a school in Uganda (right)

The questionnaire consisted of five simple questions, posed to the children in their local language. These were: 1) age; 2) place of birth; 3) length of time in the district; 4) previous treatment with praziquantel (PZQ) and; 5) understanding/knowledge of intestinal schistosomiasis (also referred to as 'bilharzia' in the questions if the children were not familiar with the term 'schistosomiasis'). During the Uganda 2010 mission, a further three questions were added involving ownership of a mosquito bed net, previous treatment history with artemisinin-based combination therapy (ACT) for malaria and previous treatment history with ivermectin for strongyloidiasis. Ngamba Island staff members participated in a more comprehensive survey which included water contact behaviour questions and clinical symptom evaluation (see Chapter 10).

2.1.3 Ethical considerations and review

For all surveys involving school-age children, the head teacher at the school, community chairperson or the child's parent was asked for written consent on behalf of the child to undertake the surveys and sample collection. Each child was further asked to give oral consent for the procedures and could withdraw at any time during the survey. The sample collection procedure and purpose for the surveys was explained to the children, in their local language, by

a VCD technician or local community mobiliser who also assisted the teachers in marshalling the children (see Figure 2.3 for an example).

Figure 2.3 – Image of a community mobiliser explaining the survey process to a class of children in Uganda (right)

In the schools, a teacher remained in the room with the pupils at all times. Staff on Ngamba Island signed written consent forms prior to taking part in any of the surveys. Each child was treated with praziquantel (40mg/kg, dosed using a standard dosage height pole) and albendazole (400mg) upon completion of the survey, regardless of infection status. Where malaria Paracheck© rapid diagnostic tests were employed (see below: section 2.2.2), all children who tested positive for *Plasmodium falciparum* malaria were offered ACT treatment on site. All medication was administered by a community nurse or health officer.

Ethical review and approval for the surveys involving schoolchildren was granted by National Health System-Local Research Ethics Committee (NHS-LREC) at St Mary's Hospital in London, the Uganda National Council for Science and Technology (UNCST) in Kampala, the Commission for Science and Technology (COSTECH) in Tanzania and the National Council for Science and Technology in Kenya. In addition, permission for the surveys on Ngamba Island was granted by UNCST and the Uganda Wildlife Authority (UWA).

2.1.4 Malacological sampling

The sampling design for the snail surveys was intended to cover the majority of the perimeter of Lake Victoria and to include as many islands located within its interior as accessible. The aim was to locate sites at approximately equal distances along the lakeshore, between five and

fifteen kilometers apart. However, this ideal was constrained by local access to the shoreline, which had to be achieved either by car or boat, and so in many cases sites were selected that did not conform exactly to the distance prerequisite. 154 sites were examined for snails during the Uganda 2008 mission, 40 in Tanzania 2008 mission, a further 30 sites in Tanzania and 33 in Kenya during the Tanzania/Kenya 2009 mission and finally 14 in the Uganda 2010 mission. Four of these latter 14 sites had been visited previously, on the Uganda 2008 mission. Figure 2.4 summarises the information in this section as well as subsequent sections, regarding the collection, processing and analysis of the snail samples.

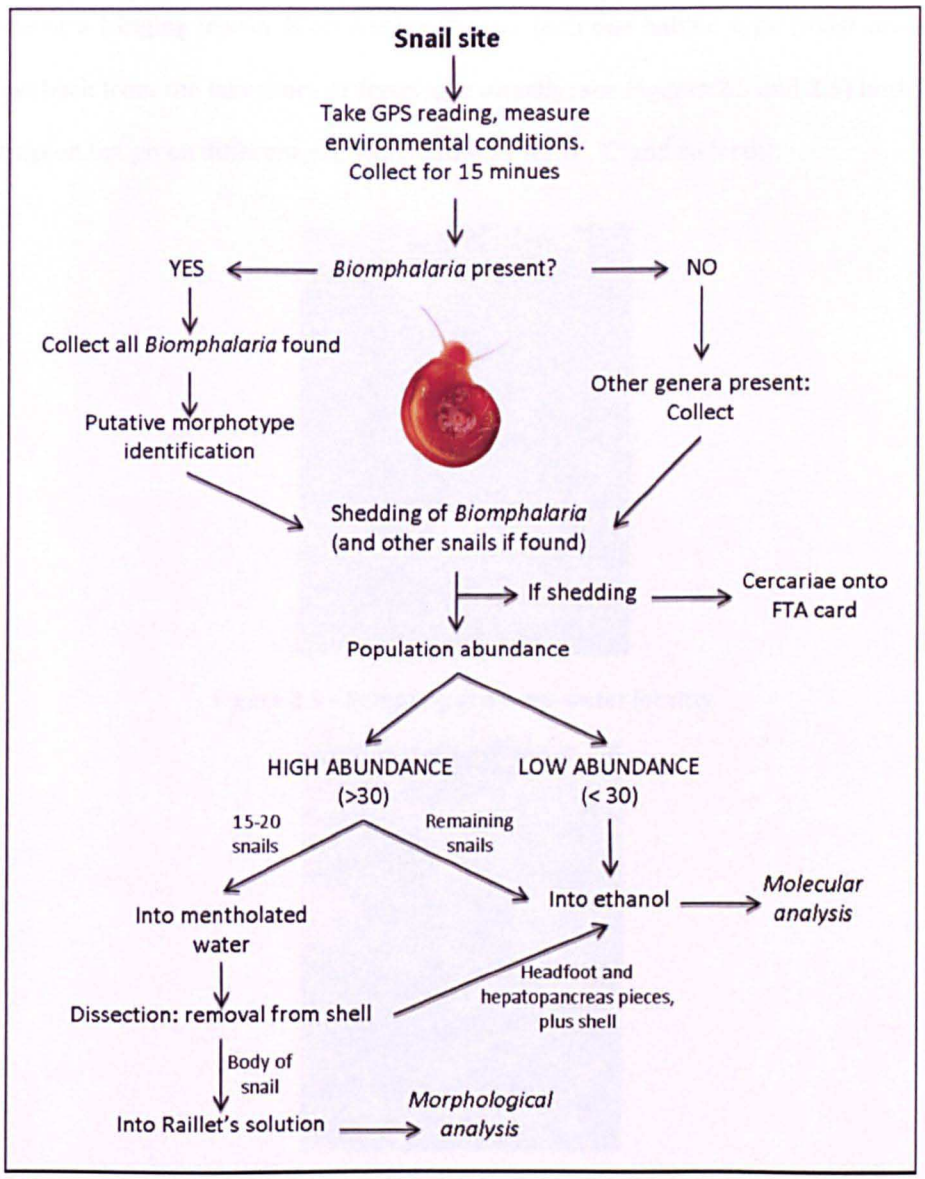

Figure 2.4 – Flow chart schematic of snail collecting, processing and analysis

As part of the Uganda 2010 mission a detailed micro-scale survey was carried out on Kimi and Ngamba Islands, investigating 20 pre-selected sites on each. Using aerial photographs and images from Google Earth (http://earth.google.com), each island was divided up into 20 equal 'zones', which were separated by visible landmarks for later identification on the ground. Each sector was also geo-referenced for cross-checking with hand-held GPS figures. The process of surveying for snails was identical to other surveys (as described below), apart from when stated otherwise. The results of this micro-scale survey can be found in Chapter 9.

At each locality, the habitat composition was first inspected, for example if it was fully lacustrine or a fringing marsh. If consisting of more than one habitat type (most commonly, a marsh set back from the lakeshore or fringing it directly; see Figures 2.5 and 2.6) both habitats were sampled but given different sub-designations ('A', 'B', 'C' and so forth).

Figure 2.5 – Sampling at a deep-water locality

Figure 2.6 – Sampling in a marsh habitat

No more than three sub-localities were identified per site. The date, time, weather and GPS coordinates were recorded for each site, along with water chemistry variables such as water temperature (°C), microconductivity (measured in micro-Siemens, or μS), total dissolved solids (TDS, measured in grammes per litre [g/L]), salinity (g/L) and pH. These water chemistry variables were measured using a HI9813 handheld portable water meter (Hanna Instruments®, Inc., Woonsocket, USA). The habitat type, substrate, depth and wave action were also noted, the latter two on a relative scale (shallow-medium-deep and low-moderate-high respectively). An aliquot of water (approximately 15ml) was also collected from each site and stored in plastic universal tubes to bring back to the NHM for detailed chemical analysis (see section 2.5.1).

Figure 2.7 – *Biomphalaria* found attached directly on shallow, submerged rock
In this case, scoops were ineffective, and so snails were removed by hand, using forceps.

Snails were collected using metal mesh paddle scoops, which were either short- or long-handled depending on the depth of the site. At sites where snails were found directly on submerged rocks, they were picked off by hand using forceps; see Figure 2.7. Sampling was

semi-quantitative, with two collectors scooping for approximately 15 minutes duration. The presence or absence of all genera of gastropods was noted; based on Brown (1994), there are 14 genera present in Lake Victoria, namely *Bellamya, Biomphalaria, Bulinus, Burnupia, Ceratophallus, Cleopatra, Ferrissia, Gabbiella, Gyraulus, Lentorbis, Lymnaea, Melanoides, Pila* and *Segmentorbis. Segmentorbis* and *Lentorbis* are difficult to distinguish without detailed anatomical dissection; the same is the case for *Gyraulus* and *Ceratophallus* and so these four genera were collectively classified in the field as 'small planorbids'. At each site, a rough estimate of the abundance of *Biomphalaria* snails was made; abundance was noted as either being absent, less than 10 individuals, between 10 and 30 individuals, or greater than 30 individuals. *Biomphalaria* were inspected and putatively identified as either *B. sudanica, B. choanomphala* or *B. pfeifferi* based on the shell shape and size before being placed in plastic jars filled with water for later processing. During the microscale surveys on Kimi and Ngamba Islands as part of the Uganda 2010 mission, exact counts of *Biomphalaria* and all other gastropod species were recorded.

2.2 Samples collected from people: processing and diagnosis

Stool, urine and fingerprick blood samples were coded sequentially, so that each sample could be identified to the appropriate child and their specific questionnaire responses. Each sample was used for a number of different diagnostic and parasitological tests (see below). All diagnostic results were tabulated in Excel, codified and then analysed in R statistical package, version 2.8.1 (Ihaka and Gentleman, 1996) for prevalence and intensity estimates, as well as modeling with other variables to detect risk factors. Spatial autocorrelation was also tested for in R, using omnidirectional and directional semivariograms, with varying bin numbers, and using a spherical fit model (McBratney and Webster, 1986, Jian *et al.*, 1996). Details of the exact statistical tests used will be given fully in the relevant data chapters.

2.2.1 *Stool samples*

A portion of each stool sample was used to create a double smear Kato-Katz microscope slide, according to standard methods (Katz *et al.*, 1972). These were typically read by technicians in the field (see Figure 2.8) and eggs counts of the following parasites were recorded: *Schistosoma mansoni* (see Figure 2.9), *Ascaris lumbricoides, Trichuris trichiura* and hookworm (*Necator americanus* and *Ancloystoma duodenale*).

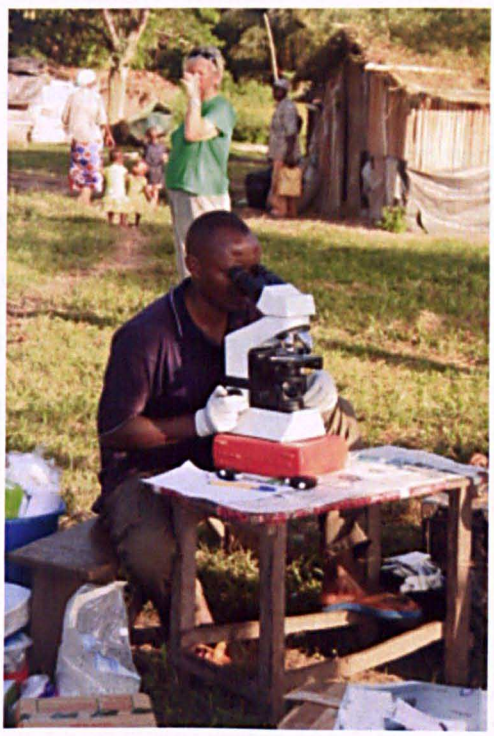

Figure 2.8 – Moses Arinaitwe, a VCD technician, reads Kato-Katz slides in the field (right)

Other parasite species were also noted if present. Slide reads were cross-checked in the UK by CJS and in some cases additionally by JRS. As Kato-Katz slides contain 1/24th of a gram of faeces, counts were averaged across the two smears then multiplied by 24 to obtain an egg per gram (EPG) value. Stool samples were also used for other parasitological techniques, such as Percoll

sedimentation, during the Uganda 2010 mission to Ngamba and Kimi Islands, which are detailed in full in the chapter devoted to that study (please refer to section 10.3 in Chapter 10).

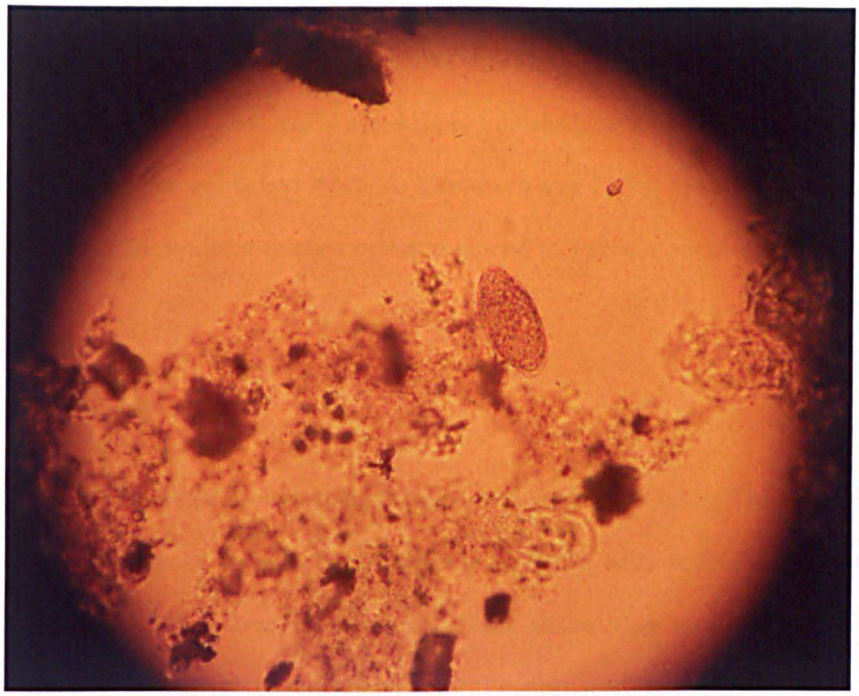

Figure 2.9 – An egg of *S. mansoni*, as seen on a Percoll slide under a compound microscope

If the slide reads showed a high abundance of *S. mansoni* eggs, the remainder of the stool was used for concentration and hatching of schistosome eggs using a modified Pitchford-Visser funnel (Pitchford, 1959, Stothard *et al.*, 2009c). The stool was homogenised in bottled mineral water, passed through a 212µm mesh sieve, then repeatedly washed through a homemade Pitchford funnel (made by Russell Stothard) which collected the eggs and removed all excess debris. The remaining liquid from the funnel was released into a series of Petri dishes and examined for the presence of eggs. If found, the dishes were exposed to direct sunlight to encourage hatching of the eggs into miracidia. Once hatched (usually after 2-4 hours), the miracidia were individually harvested in 2.5µl of water and placed on Whatman FTA® indicator cards. Where individual slide reads did not show high levels of eggs, stool samples were pooled by site to produce sufficient miracidia for collection.

In cases where a large number of miracidia were observed after hatching, some were used to infect individual laboratory-bred *Biomphalaria glabrata* snails taken to the field from the NHM laboratory, in order to obtain a future laboratory isolate of the parasite. For each isolate,

between 10 and 15 snails were exposed to five miracidia for one hour. In some cases, locally collected snails were also infected, as part of on-going compatibility experiments. All snails were then returned to the laboratory at the NHM and monitored for infection status; when found to be shedding these cercariae were used to infect several laboratory mice (150-200 cercariae per mouse), which were euthanised three months later and the adult worms collected through perfusion, as per published methods (Smithers and Terry, 1965). Paired adults were separated by dissection under a microscope and placed in 100% ethanol until later molecular analysis.

2.2.2 Urine samples

Urine samples were collected and used for two separate diagnostic tests. The first investigated the presence of blood in urine, which can indicate infection with *Schistosoma haematobium*. Microhaematuria, which is blood in urine too dilute to be seen with the naked eye, was examined for using Hemastix© (Bayer Healthcare, Leverkusen, Germany), as seen in Figure 2.10.

Figure 2.10 – Hemastix© dipsticks used to test for microhaematuria (left)

The urine was also visually inspected for macrohaematuria and turbidity. The second test was used in all field missions apart from Uganda 2008 and is called the circulating cathodic antigen (CCA) dipstick (Rapid Medical Diagnostics, Pretoria, South Africa). This is a relatively novel rapid diagnostic for schistosomiasis based schistosome circulating cathodic antigen, which is excreted in the vomit of adult worms and binds to monoclonal antibodies present in the dipstick.

The dipstick has undergone a revision to its formulation over the course of this PhD research: the first iteration was used on the Tanzania 2008 mission (see Figure 2.11) but all subsequent CCA testing was done with the second formulated dipstick (Figure 2.12).

Figure 2.11 – Image of the first formulation (carbon colloidal Ig-conjugate) of CCA urine dipstick, as used in Tanzania in 2008

Figure 2.12 – Image of the second formulation (colloidal gold Ig-conjugate) of CCA test used on all subsequent surveys

The original 'dipstick' formulation was replaced with a lateral flow strip.

In both cases, the test is fast and simple to use in the field; a small drop (approximately 15-25µl) of host urine is taken from the collected sample and a few drops (75µl) of buffer solution added. In the case of the first type of test, the buffer contained carbon-bound monoclonal antibodies that bind to CCA. The dipstick strip was then placed in the well containing the buffer and urine; as the solution was adsorbed up the dipstick during the 40 minute incubation period the captured CCA was bound to a particular zone of the dipstick, indicating presence of cathodic antigens in the urine sample. A control line was also present to determine if the test functioned properly. Control solutions were provided with this kit, allowing for quantitative analysis of antigen levels. In the second formulation, the principal remained identical but the conjugate was bound to colloidal gold and the dipstick was integrated into a lateral flow 'pregnancy test'-like format with a much shorter (5-10min) incubation time, allowing for even more rapid diagnosis. However, a range of control solutions was not included and so intensity estimates could only be semi-quantitative, based on relative band density compared to the internal control.

2.2.3 Fingerprick blood samples

Fingerprick blood samples (Figure 2.13 shows the process of collecting fingerprick blood) were primarily used for diagnosis of schistosomiasis using soluble egg antigen (SEA) enzyme-linked immunosorbent assays (ELISA) kits (IVD Inc.; Carlsbad, USA).

Figure 2.13 – Image of a child having fingerprick blood taken by Moses Adriko, a VCD technician (left)

These kits detect antibodies to SEA directly in the blood of patients infected with *Schistosoma* parasites and are semi-quantitative, as the results can be compared against a control solution of known concentration. The kits were used according to manufacturer's protocols, at a serum dilution of 1:40 (see Figure 2.14). Serum was collected through

centrifugation of the fingerprick blood samples, which, in the absence of electricity in the field, had to be done by hand. Positive and negative control dilutions of antigen were supplied in the kits and used to check that the test had functioned adequately (see Figure 2.15).

Figure 2.14 – Field preparation of the SEA-ELISA plates

Sera were collected by hand-centrifugation of fingerprick blood samples.

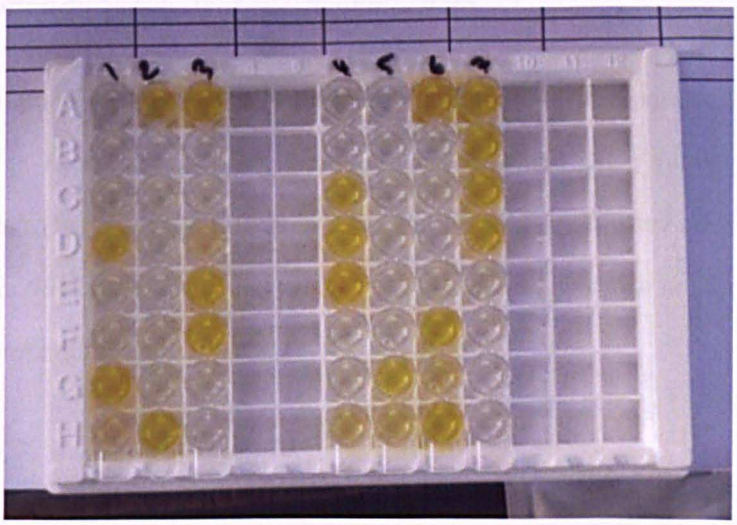

Figure 2.15 – Image of a post-reaction SEA-ELISA plate used in the field

The yellow indicates a positive result whereas colourless indicates a negative.

In addition, during the Uganda 2010 mission, fingerprick blood samples were measured for haemoglobin levels using an Hb 201 portable haemoglobin reader (HemoCue AB, Ängelholm, Sweden) and also used in detection of *Plasmodium falciparum* by a malaria rapid diagnostic (Paracheck-Pf®, Orchid Biomedical Systems, Goa, India). Samples from an individual child were numbered so that diagnostic results could be used to provide a comprehensive diagnostic profile for that child, as well as linked to questionnaire responses (Figure 2.16).

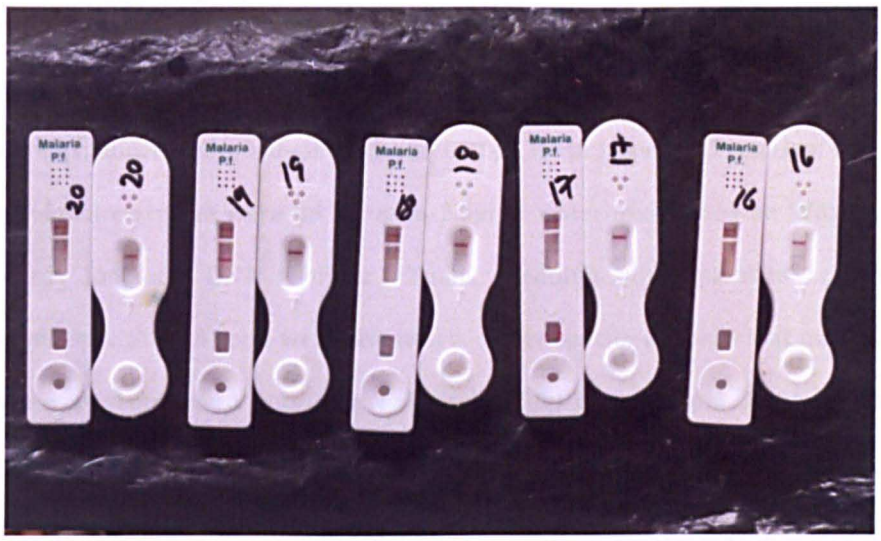

Figure 2.16 – Paracheck® and CCA urine tests paired by child

This demonstrates again how samples from each child were matched to produce a comprehensive diagnostic profile per individual as well as site-level prevalence estimates.

2.2.4 Tabulation and analysis of questionnaire data

Questionnaire data were electronically entered, combined with the parasitological data and then tabulated in Excel. Once coded, analyses were executed using version 2.8.0 of the R statistical package (Ihaka and Gentleman, 1996). For all models, adjusted odds ratios (OR) are given along with 95% confidence intervals (CI). Statistical significance in all cases was set at less than 0.05 probability.

2.3 Processing of snail collections

After a typical day's collecting, snails were removed from the plastic jars and sorted per site. Figure 2.4 in the section 2.1.4 further shows the steps taken in processing the snail samples. All *Biomphalaria* were put up for shedding individually; this involved placing individuals in trays containing 3ml wells, filled with bottled water, and exposing the tray to sunlight for 2-3 hours. Ideally, the snails were shed during the middle of the day, as this has been reported to be the hours of peak cercarial production for *S. mansoni* (Kazibwe *et al.*, 2010). Following exposure the wells were examined under a dissecting microscope for presence of *S. mansoni* cercariae. If present, individual cercariae were taken up in 2.5µl of water and placed on Whatman FTA® indicator cards for later DNA analysis. Snails shedding other parasite cercariae (i.e. *Paramphistoma* spp., *Strigea* spp.) were also noted. Other species of snails that may have been collected were shed in group pots (approximately 25ml of water per 10 snails) and any emerging parasites noted.

If abundance of *Biomphalaria* at a particular site was low (< 30 individuals), then all individuals were placed directly into 100% ethanol, in copper-topped glass universals, and labeled per site. In cases when more than 30 *Biomphalaria* had been collected, 15-20 were placed in pots of bottled water containing a few crystals (0.5g) of crushed menthol crystals for between 4 and 8 hours. The remainder of the snails from that site was placed into 100% ethanol. The mentholated water gradually relaxed the snails so that they can be quickly killed in hot (60°C) water without retraction into their shells. Once dead, the snails were placed in a Petri dish in about a centimeter's depth of water and the bodies carefully pulled out from the shells using two forceps. At this point, a small piece of the headfoot was cut, along with the last tip of the hepatopancreas, and both placed (together with the empty shell) into an individual, labeled 2ml plastic tube containing 1.5ml of 100% ethanol. These pieces of tissue could later be used for genetic analysis. The rest of the body was placed into another 2ml tube containing 1.5ml modified Raillet's solution, which preserves the internal organs without distorting their shape and size. The solution is made according to the following formula: 930ml distilled water, 6.0g sodium chloride, 50ml formalin and 20ml acetic acid.

As mentioned previously, if there was a particularly high abundance of snails at a site, combined with a strong miracidial hatch from stool collected that day, then local snails might also be exposed alongside *B. glabrata* to the hatched miracidia, for compatibility testing but also to bring that locality's isolate back to the NHM.

2.4 Molecular DNA techniques: Lab protocols

Molecular methods were employed for both schistosomes and snails to generate genetic data. New procedures and methodologies, such as those introduced by the advent of DNA 'barcoding' and the use of Whatman FTA® indicator cards, were utilised to create genetic datasets for both *Biomphalaria* and *S. mansoni*, and from point-of-collection larval stages for the latter. Analyses were carried out at both phylogenetic and population genetics levels to explore different patterns within the data.

2.4.1 *DNA extraction*

Methods for DNA extraction varied depending on the way the material had been stored. For schistosomes, the majority of specimens had been preserved in the field on Whatman FTA® indicator cards and were extracted by punching out the disc using a Whatman 1.2mm punch and washing three times with 200µl of Whatman FTA® Purification Reagent followed by three washes with pH 7.4 TE buffer (the recipe can be found in Appendix 13.1). Once dried thoroughly, the disc could be used directly in a PCR.

Adult worms and snail tissue stored in ethanol were extracted using the same standard CTAB (hexadecyltrimethylammonium bromide) protocol. The tissue sample was homogenised in 500ml of 2 x CTAB extraction buffer (for CTAB formula see Stothard and Rollinson, 1997), then 20µl of proteinase K (20g/ml concentration) added and the mixture agitated at 55°C for four to eight hours to encourage complete lysis. 600µl of chloroform-isoamyl alcohol were then added and the mixture spun for 5 minutes at 13200 rpm. The supernatant was taken off using wide-bore tips to avoid shearing of the DNA, and placed into pre-labeled tubes containing 1ml of 100% ethanol. The sample was left to precipitate in the ethanol, at 4 °C, for 4-6 hours before being spun again at 13200, this time for 20 minutes. All the ethanol was then removed, leaving the DNA pellet, and 1ml of 70% ethanol (diluted with distilled water) added. After a final spin of 5 minutes, the liquid was again removed and the pellet allowed to dry completely at 60 °C. Once all excess ethanol had evaporated, 50-100µl of pure water were added to each sample,

depending on the size of the visible pellet. The sample was placed at 4°C and allowed to rehydrate before being used in a PCR.

2.4.2 Selection of molecular DNA markers

Genetic markers were selected on the basis of providing enough variation for phylogenetic as well as population-level signal, for both the schistosomes as well as the snails. Repeatability of results was desirable, so sequence typing methods were preferred to random amplified polymorphic DNA (RAPD) or fingerprinting (Jarne and Théron, 2001).

2.4.2.1 Schistosomes

For the *S. mansoni* samples, the aim was to investigate population genetic structure rather than deeper phylogenetic divisions, and so a marker that provided sufficient variation at the population level was required. The cytochrome oxidase sub-unit 1 (COI) gene had already been shown to be highly variable between populations, and moreover, there was evidence of local structuring, for example between regions or lakes (Morgan *et al.*, 2005, Stothard *et al.*, 2009c). As a mitochondrial marker, it is furthermore only maternally inherited and considered non-recombinant, allowing for simplified analyses. A 396 base pair fragment of the COI gene, called the ASMIT fragment (Bowles *et al.*, 1992), was selected in order to compare these data to those of previous studies (Stothard *et al.*, 2009c).

For corroboration at a population genetics level, it was also decided to add a nuclear, biparentally inherited marker, which would also show high levels of variation: for this, a set of five microsatellite loci were selected. These were: smd89, smd28, smd25, sms9-1 and cal1-1. All were previously published and fluorescently labelled with 6-FAM, PET, NED and VIC (Durand *et al.*, 2000, Blair *et al.*, 2001, Gower *et al.*, 2007).

2.4.2.2 Snails

With regards to the snails, markers were required that would both distinguish between species, in order for phylogenetic analysis of the species found in Lake Victoria, as well as population genetics analyses of groups within the lake. COI was again chosen as a suitable marker, but in

this case the larger Folmer fragment (Folmer *et al.*, 1994) was amplified (655-658 base pairs), as this is the 'barcoding' fragment, proposed as a suitable marker for identifying species (see section 1.4.3). COI was also thought to be sufficiently variable to act as a population level marker. In order to corroborate the COI data, given the large number of samples to be analysed, an additional mitochondrial marker, ribosomal 16S, was also selected. As a gene coding for the mitochondrial large ribosomal RNA subunit, 16S is often used in molluscan phylogenies and so provides a useful comparison with the COI data, particularly in terms of elucidating the phylogenetic structure of the *Biomphalaria* from Lake Victoria (Bonnaud *et al.*, 1994, Wollscheid-Lengeling *et al.*, 2001, Jørgensen *et al.*, 2007).

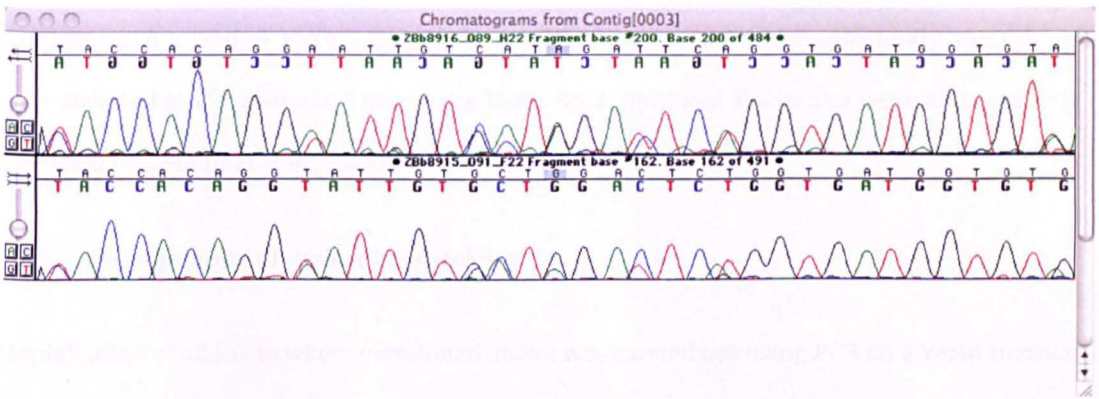

Figure 2.17 - Image of an actin sequence chromatogram file from a *Biomphalaria*
Note the presence of clear double peaks; these suggest the presence of different alleles of the gene, or multiple copies within an individual.

Nuclear markers were also investigated, to add a biparentally inherited marker to the phylogenetic and population genetics analyses. Primers coding for exon 2 of an actin gene were tested on 4 snails but the resulting sequences contained double peaks (Figure 2.17), indicating the possible presence of multiple copies of the gene; in fact, it is thought that *B. glabrata* has five copies of the actin gene (Adema, 2002), and as it would be impossible to determine whether the same copy was amplified for all individuals and the sequences were thus homologous, this marker was discarded. The first internal transcribed spacer region (ITS1) was also tested on 13 individual snails, using ETTS17 and ETTS2 (Plam *et al.*, 2008), but all were found to have double peaks at several points within the sequence, suggesting that there may be intraindividual polymorphisms in this region and so was proven unsuitable for further use on these taxa. The entire ITS region (including the 5.8S rDNA gene together with the two

flanking ITS spacers) was then tested using the primers ETTS1 and ETTS2 (Kane and Rollinson, 1994); while unsuitable in this case for direct sequencing due to multi-peak and intraindividual variation problems mentioned above, this region has been shown to reveal species-level differences in related pulmonates such as *Bulinus* through enzyme digest (Stothard *et al.*, 1996) and so it was decided to include this marker in the phylogenetic analysis, though in the form of a restriction enzyme digest, as being more amenable for analysis.

As with the schistosomes, microsatellite loci were also selected as nuclear markers for the population genetics analyses. Unfortunately, none had been developed specifically for *B. choanomphala* or *B. sudanica*, and so 13 primer pairs that were developed for *B. pfeifferi* and *B. glabrata* were tested. 4 were successfully amplified and found to be sufficiently variable so were included in the analysis. These were Bpf1, Bpf2, Bpf3 and Bpf10 and were all previously published (Charbonnel *et al.*, 2000).

2.4.3 Amplification of molecular markers

Amplification of all the markers mentioned above was carried out using PCR on a Veriti thermal cycler (Applied Biosystems Inc., Foster City, CA, USA) and all products were visualised on a 1% agarose gel stained with ethidium bromide (prior to October 2008) or GelRed™ (Hayward, CA, USA). Microsatellite products were visualised on a 2% agarose gel. Primer names, sequences, references and annealing temperatures can be found in Table 2.1.

Primer	Sequence (5' – 3')	DNA region	Reference	Tm (°C)
ASMIT1/ ASMIT2	F: TTTTTTGGGCATCCTGAGGTTTAT R: TAAAGAAAGAACATAATGAAAATG	COI gene	Bowles et al (1992)	40
smd89	F: AGACTACTTTCATAGCCC R: TTAAACCGAAGCGAGAAG-6FAM	Microsat.	Durand et al (2000)	48
smd28	F: CATCACCATCAATCACTC R: TATTCACAGTAGTAGGCG-6FAM	Microsat.	Durand et al (2000)	48
smd25	F: GATTCCCAAGATTAATGCC R: GCCATTAGATAATGTACGTG-NED	Microsat.	Durand et al (2000)	48
sms9-1	F: ATTACGATTGCACAGATACTTTTG R: TTTCAGAAATTTGTTTCCTCCTC	Microsat.	Blair et al (2001)	50
cal1-1	F: TTCAAAACCATGAGCAATAGATAC R: CAACAAACAAGAAGGCTGATTAG-6FAM	Microsat.	Blair et al (2001)	51
LCO1490/ HCO2198	F: GGTCAACAAATCATAAAGATATTGG R: TAAACTTCAGGGTGACCAAAAAATCA	COI gene	Folmer et al (1994)	45
16sar/16sbr	F: CGCCTGTTTATCAAAAACAT R: CCGGTCTGAACTCTGATCAT	16S gene	Bonnaud et al (1994)	47
ETTS1/ETTS2	F: TGCTTAAGTTCAGCGGGT R: TAACAAGGTTTCCGTAGGTGAA	ITS region	Kane and Rollinson (1994)	47
Bpf1	F: TCCTATCCTTGTAACTTCTCCAC R: CGAAACCATGCAAATCAG-6FAM	Microsat.	Charbonnel et al (2000)	52
Bpf2	F: GCAGCTTCATTCAACATTCC R: AAATTAACATTTCGCTGAAACAG-6FAM	Microsat.	Charbonnel et al (2000)	52
Bpf3	F: CAGTATTTCCCGTGACTGCTC R: CTTTCCCTTGTTCGATACCATAC-HEX	Microsat.	Charbonnel et al (2000)	52
Bpf10	F: TGTCCAGCATGTCCAGTTC R: CAGAGATGATATTGCAGTCAGG-HEX	Microsat.	Charbonnel et al (2000)	52

Table 2.1 – Primer details

All schistosome DNA amplifications were performed using Ready-To-Go PCR beads (GE Healthcare, Chalfont St Giles, UK) in a 25µl total reaction volume, with 10pmol of forward and reverse primer in each reaction for the COI amplifications and 1pmol for the microsatellites. Ready-To-Go PCR beads were also used for the snail ITS amplifications, using 10pmol of each primer. All other snail amplifications were carried out using Promega Go-TAQ (Promega Corporation, Madison, UK) liquid TAQ. Mitochondrial (COI and 16S) amplifications were done in 25µl total volume, with 2.5µl $MgCl_2$ (20MM concentration), 2.5µl 5 x buffer, 2.5µl pre-mixed dNTPs (20MM concentration), 1µl each of forward and reverse primer (10pmol concentration) and one unit of TAQ per reaction. Microsatellite reactions were amplified in 12.5µl total

volume, with half all of the volumes stated above, and with a forward primer concentration of 50pmol and reverse (fluorescently labeled) primer concentration of 20pmol.

The PCR cycling conditions varied slightly for all the markers but were all based on the following conditions: An initial denaturation stage, followed by a set number of cycles each with a denaturation stage, an annealing stage and an elongation stage, finished with a final (optional) elongation stage. Cycling details are as follows:

ASMIT1/ASMIT2:

5 minutes (min) initial denaturation at 95°C, followed by 40 cycles of 30 seconds (s) at 95°C, 30s 40°C and 2min at 72°C, with a final elongation of 7min at 72°C.

smd89, smd25, smd28, sms9-1 and cal1-1 (individual PCR reactions):

15min initial denaturation at 95°C, followed by 40 cycles of 30s at 94°C, 1.5min at the T_m for that primer (see Table 2.1) and 1min at 72°C, with a final elongation of 30min at 60°C.

LCO1490/HCO2198:

2min initial denaturation at 96°C, followed by 35 cycles of 1min at 96°C, 45s at 45 °C and 2min at 72 °C, followed by an extra 5min elongation stage at 72°C.

16sar/16sbr:

2min initial denaturation at 96°C, followed by 35 cycles of 1min at 96°C, 45s at 47°C and 2min at 72°C, followed by an optional 5min extra elongation stage at 72°C.

ETTS1/ETTS2:

2min initial denaturation at 96°C, followed by 35 cycles of 1 minute at 96°C, 45 seconds at 47°C and 2 minutes at 72°C, followed by an extra 5 minute elongation stage at 72°C.

Bpf1, Bpf2, Bpf3 and Bpf10 (individual PCR reactions):

2min initial denaturation at 95°C, followed by 30 cycles of 30s at 94°C, 30s at the primer's T_m (see Table 2.1) and 45s at 72°C.

Bead-amplified PCR products were purified using a QIAQuick PCR Purification Kit (QIAGEN Ltd, Crawley, UK) whereas TAQ-amplified products were purified using a Millipore PCR_{96} Cleanup kits on a vacuum manifold (Millipore, Billerica, USA) as per manufacturer's instructions, using pure water for washing and resuspension. Product concentration was quantified on a Nanodrop ND-1000 Spectrophotometer (Nanodrop Technologies Inc., Willington, USA), and sequencing reactions were performed on mitochondrial purified PCR products using an Applied Biosystems Big Dye Kit (version 1.1) and run on an Applied Biosystems 3730 DNA Analyzer (Applied Biosystems, Carlsbad, USA). Microsatellites were not sequenced, but diluted 1:10 in HiDi formamide (Applied Biosystems, as above) and analysed using ABI3730 automated sequencer, using a GeneScan 500 LIZ size standard (both Applied Biosystems, as above).

2.4.4 Enzyme digest of ITS fragment

The ITS fragment amplified in the snails was not sequenced, but rather used in two separate enzyme digests (*AluI* and *HaeIII*). 2.4µl of unpurified PCR product were added to a mixture containing 2µl (2 units) of the restriction enzyme, 2.4µl of buffer solution appropriate to that enzyme and 17.6µl pure water. The mixture was left to digest overnight at 37°C, and then 15µl of it was run on a 4% agarose gel stained with GelRed™.

2.5 Molecular DNA techniques: Analysis

The DNA sequences and microsatellite genotypes for both *S. mansoni* and *Biomphalaria* were used for a variety of analyses. As all of the schistosome samples were assumed to be of the same species, the analysis of the *S. mansoni* COI and microsatellite data focused on population genetics analyses, such as networks and statistical tests for population structure and divergence, although a phylogenetic tree was also built to investigate evolutionary relationships between the haplotypes. Due to the taxonomic difficulties known in *Biomphalaria* from Lake Victoria, a sub-set of 7 populations, consisting of approximately 70 snails, was first used to examine whether there were indeed species-level divisions within the snails sampled; this dataset is hereon referred to as the 'phylogenetic dataset'. This was done using phylogenetic trees and networks on both COI and 16S sequences; this sub-set was also subjected to morphometric analyses (see section 2.6). The full *Biomphalaria* dataset of COI and 16S haplotypes and microsatellite genotypes from 29 populations (referred to as the 'population dataset' was then used for population-level analysis, again using networks, statistical tests and phylogenetic trees. All of these methods of analysis will be described in full below.

2.5.1 Sequence alignment and microsatellite allele determination

Sequences were assembled and edited by eye using Sequencher v 4.8 (Gene Codes Corporation, Ann Arbor, Michigan, USA: http://www.genecodes.com). Examples of these chromatograms, for *S. mansoni* and *Biomphalaria*, can be seen in Figures 2.18 and 2.19, respectively. Individual sequences were visually aligned in MacClade 4.05 (Sinauer Associates, Sunderland, Massachusetts, USA: http://macclade.org/macclade.html) and polymorphic positions checked back to the original chromatograms in Sequencher. For 16S, where lots of gaps were present, alignment was automated using Muscle (EMBL-EBI: www.ebi.ac.uk/Tools/muscle/) and then re-imported back into MacClade for visual re-inspection of the alignment. Species identification of each sequence was checked by performing a BLAST (Basic Local Alignment Search Tool;

http://blast.ncbi.nlm.nih.gov/Blast.cgi) search. Unique haplotypes were identified using Collapse v 1.2 (http://darwin.uvigo.es/software/collapse.html).

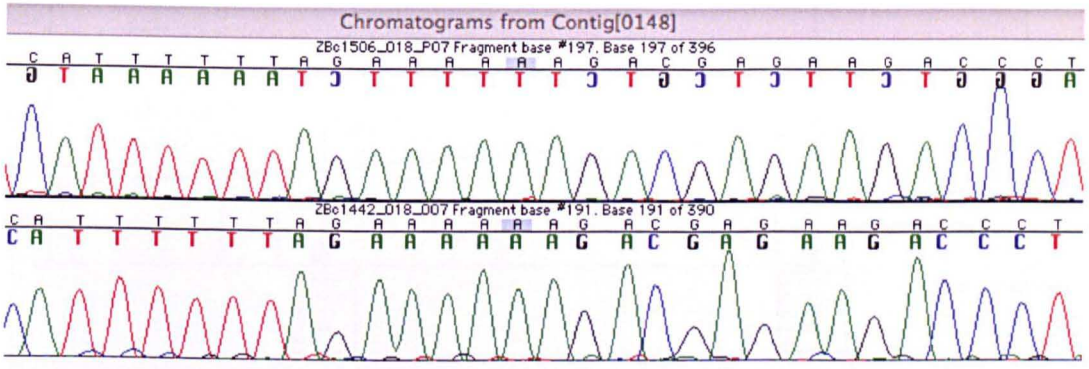

Figure 2.18 – Image of a chromatogram sequence obtained from sequencing the ASMIT fragment of the COI gene in a Lake Victorian *S. mansoni*

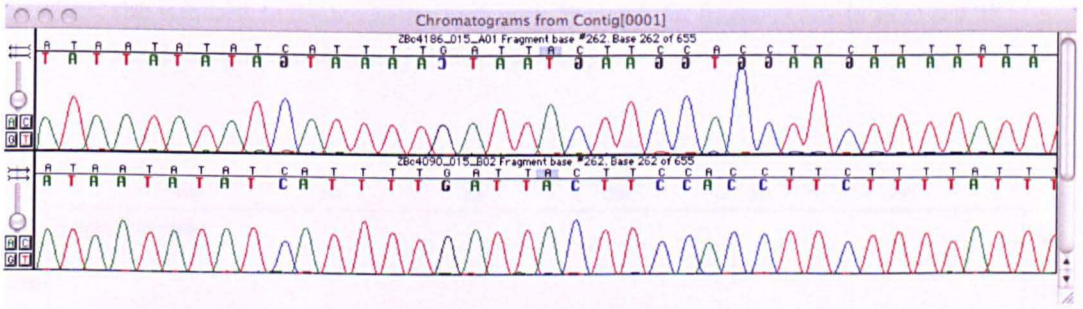

Figure 2.19 – Image of a chromatogram sequence obtained from sequencing the Folmer fragment of the COI in a Lake Victorian *Biomphalaria*

Microsatellite allele sizes were calculated using PeakScanner v 1.0 (Applied Biosystems, as above; see Figures 2.20 and 2.21).

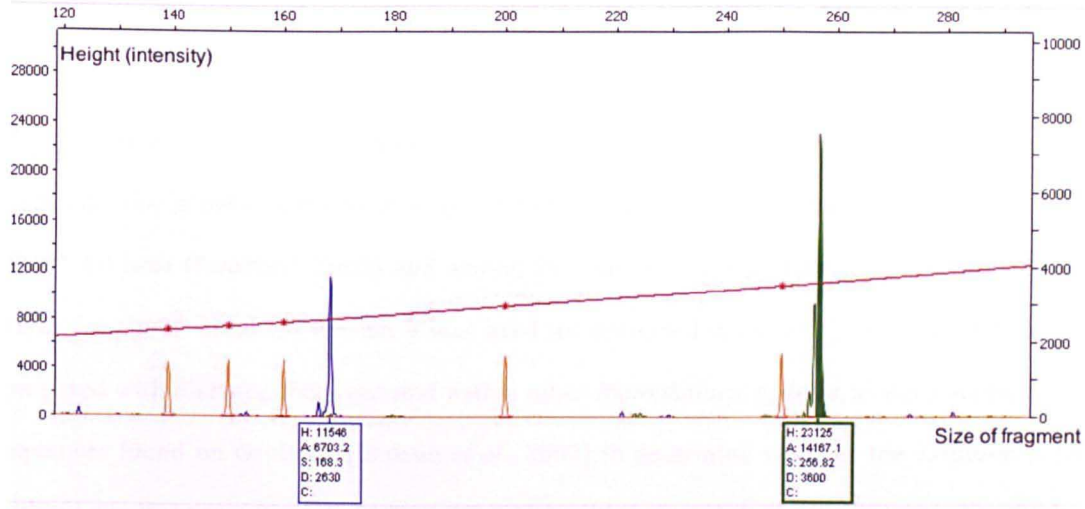

Figure 2.20 – Example of a snail homozygous at loci Bpf 2 and Bpf 10

Image shows loci Bpf 2 and Bpf 10 (blue and green, respectively) for a *Biomphalaria* from population T011. This snail is homozygous for both loci, as can be seen by the single peaks of size 168 and 256, respectively. The orange peaks represent the internal size standard. In the box below each peak, 'S' stands for fragment size (x axis) and 'H' stands for height, or intensity (left-hand y axis).

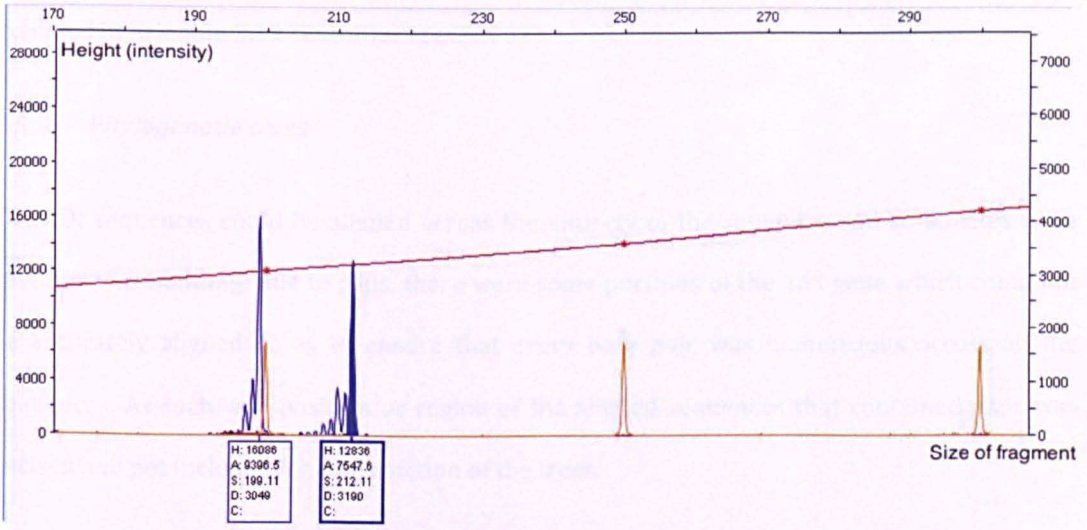

Figure 2.21 – Example of a snail heterozygous at locus Bpf 1

Image shows locus Bpf 1 for a *Biomphalaria* snail from population T033a. The allele sizes can be seen to be 199 and 212; the smaller peaks before the main peak represent stutter peaks, which result from PCR slippage during amplification and sequencing. The orange peaks represent the internal size standard. In the box below each peak, 'S' stands for fragment size (x axis) and 'H' stands for height, or intensity (left-hand y axis).

2.5.2 Distance calculations

Uncorrected and corrected (see below for information on the model of nucleotide substitution used) pairwise genetic distances among and between the individual snails were calculated in PAUP* 4.0 beta (Swofford, 2003) and among and within group distances were calculated in MEGA (as above, although version 5 was used for corrected distances). These could then be compared with distances between and within other *Biomphalaria* species, as determined from sequences found on GenBank (Benson *et al.*, 2009) to determine whether the variation seen within Lake Victorian *Biomphalaria* is consistent with being part of a single, versus multiple, species.

For the correlations between pairwise genetic distances of *S. mansoni* and *Biomphalaria*, uncorrected p-distances were used for the COI data, as the best-fitting model of nucleotide substitution differed between the two data sets (see following section). This was done in MEGA (as above). Microsatellite distance matrices were calculated using the infinite allele model, and executed in Arlequin 3.11 (Excoffier *et al.*, 2005).

2.5.3 Phylogenetic trees

The COI sequences could be aligned across the entirety of the sequence and so all sites were used for tree-building; due to gaps, there were some portions of the 16S gene which could not be accurately aligned so as to ensure that every base pair was homologous across all the sequences. As such, any position or region of the aligned sequences that contained gaps was excised and not included for construction of the trees.

For the schistosome COI data, the optimal model of nucleotide substitution was calculated as being the Tamura-Nei 4-parameter model, using ModelTest v 3.7: http://darwin.uvigo.es/software/modeltest.html). A bifurcating phylogenetic tree was built using neighbour-joining branch-addition based on haplotype distances, with branch support provided by bootstrapping (500 replicates, only probabilities above 50% were included). The tree was built using MEGA 4.0 (Tamura *et al.*, 2007).

For the *Biomphalaria* COI and 16S sequence data, testing for the best-fitting model of nucleotide substitution was done differently. In this case, four neighbour-joining distance trees were built using different models of nucleotide substitution, namely HKY, HKY+gamma, GTR and GTR+gamma. The likelihood of each of these trees was estimated and then compared sequentially using a Chi-squared test. For both 16S and COI, GTR+gamma proved to be significantly the most likely model ($p<0.05$) and so was used for all subsequent analyses. The various parameters within the model were estimated in PAUP across successive iterations, until the estimates converged on the values representing maximum model likelihood.

For the phylogenetic dataset of 7 *Biomphalaria* populations, trees were built for all the individual sequences (not haplotypes) for both COI and 16S using four methods: pairwise distance (neighbour-joining), maximum parsimony, maximum likelihood and Bayesian inference. The first two were done using PAUP, the third in PhyML 3.0 (Guindon and Gascuel, 2003) and the last in Mr Bayes (Ronquist and Huelsenbeck, 2003). For the parsimony tree, 109 characters out of 655 base pairs were parsimony-informative for the COI sequences, whereas 76 characters out of 385 total base pairs were parsimony-informative for the 16S data. In the construction of the Bayesian inference tree, the model was run for sufficient generations for the model to converge. Trees were sampled every 100 generations and the last 10% saved, outputted as a consensus and used to calculate posterior probabilities.

All trees produced had roughly the same topography; nodes that had a posterior probablity of over 50% were also tested for support based on the other tree-building techniques through bootstrapping (1000 replicates in each case). Bootstrap values were then added to the original Bayesian inference tree at the relevant node, for direct comparison of the support for that clade across the different techniques.

Constrained neighbour-joining trees were also created in PAUP, which enforced monophyly on particular taxa. The likelihood estimates of these constrained trees could then also be compared, via Shimodaira-Hasegawa and Kishino-Hasegawa tests (both one-tailed RELL-resampling tests) for significance versus the unconstrained neighbour-joining tree (Kishino and Hasegawa, 1989, Shimodaira and Hasegawa, 1999).

For the *Biomphalaria* population dataset, trees of the unique haplotypes, for both COI and 16S were constructed using distance methods and neighbour-joining branch addition, again using the GTR+gamma model of nucleotide substitution, and with parameters set according to repeated iterations of estimated values until maximum model likelihood was achieved. Node support was again determined through bootstrapping (1000 replicates) and the bootstrap values added to the original tree.

2.5.4 Networks

When dealing with intraspecific or closely related taxa, the small genetic distances can result in a number of possible phylogenetic trees; the use of multiple tree-building methods and bootstrapping to test for node support can help resolve this problem. However, it is often also useful to build a network from the data instead, which explicitly demonstrates the possibility of different evolutionary histories by linking different sequences or haplotypes by multiple paths. In this case, two different methods of drawing networks were utilised, one for the *S. mansoni* COI data and the other for the *Biomphalaria* COI and 16S sequences, for both the phylogenetic and population datasets.

For the schistosome COI sequences, the unique haplotypes were used to create a minimum spanning network using TCS (http://darwin.uvigo.es/software/tcs.html), with the connection limit set at 95%. The method is based on statistical parsimony (Templeton *et al.*, 1992) and each node represents a single base-pair change and so more related sequences have fewer nodes separating them.

The *Biomphalaria* phylogenetic dataset were also used to create a network using TCS, with the connection limit set to 95%. This will not connect into the network taxa that are more than 5% distant from each other, and so is a good proxy for species-level divides, although the exact percentage for true species-level distance will vary between genera and even species. For the *Biomphalaria* population genetics dataset, networks were created, again for both COI and 16S, using the median-joining function in the software programme Network 4.516 (Bandelt *et al.*,

1999). This is also a maximum-parsimony based approach which requires non-recombinant genetic data and so is assumed to be suitable for mitochondrial markers.

2.5.5 Population genetics statistics

Population genetics tests were also carried on sequence data for schistosomes and snails, with the microsatellite markers also included as a nuclear marker. Unless otherwise stated, all haplotypic (i.e. COI and 16S) analyses were carried out in Arlequin 3.11 (Excoffier *et al.*, 2005). Microsatellite analyses were either calculated in Arlequin, Genepop on the web (Raymond and Rousset, 1995, Rousset, 2008) or FSTAT (Goudet, 1995), also unless otherwise stated. Where multiple tests were performed simultaneously, Bonferroni corrections were applied to significance estimates.

The population genetics tests for *S. mansoni* were applied to two separate datasets. The first consisted of COI sequences: only miracidia were included, and only those from hosts where greater than 10 samples had been successfully sequenced, to ensure sufficient statistical power. In all, 163 COI sequences from 10 children, from 6 sites across the three countries, were included in this sub-set. The second dataset consisted 36 adult worms from four isolates of lab-passaged worms, which were used for microsatellite analysis.

The *Biomphalaria* population dataset was also analysed for mitochondrial (COI and 16S) and microsatellite markers. In this case, gaps were not excluded from the 16S analyses as they represent true differences between individuals in a population. Pie charts showing haplotype diversity per location were created in Excel (Microsoft® Office, Microsoft Corporation, Redmond, WA, USA).

For both schistosomes and snails, tests included analysis of molecular variance (AMOVA), Tajima's and Fu's tests of neutrality, a mismatch test of population expansion and calculation of F_{ST} pairwise distances between and within groups (Weir and Cockerham, 1984, Tajima, 1989, Excoffier *et al.*, 1992, Rogers and Harpending, 1992, Fu, 1997). AMOVA examines the proportion of the variation within the data explained by the various groupings, such as within a population, between populations in a group (i.e. in a country) and between groups. Tests for

neutrality are important because many tests assume that the data are evolving neutrally and are not under selective pressure. However, there are circumstances in which data can fail to be deemed neutral using these tests, but it is due to factors other than selection occurring; population expansion is one such factor, which is why mismatch tests are useful as a cross-check. F_{ST} pairwise distances determine the level of gene flow between populations, or in other words, how structured they are: An F_{ST} value of 0 means complete population mixing, whereas a value of 1 implies total segregation. For the schistosomes, analyses were carried out at host, site and country level; for the snails, comparisons were made between groups defined by location and by country.

A Mantel z-test (Mantel, 1967) was used to investigate correlation between genetic and geographical distances: geographical distances were measured in ArcView (ArcGIS, http://www.esri.com/software/arcgis/). In addition, the microsatellite data were tested for Hardy-Weinberg equilibrium (HWE; again due to assumptions of HWE implicit in many population genetics tests), linkage disequilibrium (to ensure independence of each locus as a marker), allelic diversity and a history of past bottlenecks. Bottleneck testing was executed using Bottleneck v 1.2.02 (Cornuet and Luikart, 1996).

2.6 Morphological techniques

Morphological methods were used on the snail samples collected in the field, as well as the shells of type specimens at the Museum für Naturkunde in Berlin, which were made available for study via a SYNTHESYS Access Grant. A large proportion of the type material had been described by Karl Eduard von Martens in the 1870s, based on material collected from Africa by explorers such as Emin Pasha and Franz Stulhmann. Shell measurements, aperture outlines and copulatory organ measurements were each done on individual snails from the 7 populations included in the phylogenetic analyses above (plus the Berlin specimens, where the measurements were possible). In addition, all snails in the population genetics database had shell measurements taken.

In all cases, principal component analysis was performed on the data to determine groupings, either in R (Ihaka and Gentleman, 1996) or Eigenvector (see Acknowledgements). Principal component analysis (PCA) transforms a number of data points into a smaller set of uncorrelated variables, called 'eigenvalues', where the largest proportion of the variance in the data is accounted for by the first variable. In this way, if the majority of the variation in the data is accounted for in the first two eigenvalues, the data can be plotted in a 2-dimensional graph, to see clearly the way in which variation is partitioned throughout the data. Linear discriminant analysis (LDA) explicitly uses a separate, categorical variable to model the variance in terms of the groups as defined by the categories. In this case, putative field identification of the *Biomphalaria* morphotypes was used as the classifier. For the shell and internal anatomy measurements, PCA was performed in R using the 'princomp' function, and for all three sets of measurements LDA was also performed in R (using the 'lda' function), on the first two principal components to test whether the same putative species identifications were corroborated by the PCA groupings.

The procedure was slightly different for the aperture outlines. In this case, the phi factors were inputted into an eigenshape analysis, which effectively does the same job as a PCA, by looking for the vectors which account for the most variation (axis 1), the second most variation (axis 2)

and so on. The programme also reduced the number of phi functions to the minimum number of landmarks required to match 97.5% of the original curve, which was 12. The output of the eigenshape analysis is also in the form of eigenvalues, as with standard PCA, which can then be used for further analyses such as LDA and matrix correlations. Plots of the first two principal component axes were drawn in Excel; tables comparing the putative species identifications from the field to the groupings calculated by the LDA were created in Word.

For all the morphological datasets, correlations with the genetic data were investigated using Mantel r-tests for matrix correlation (Mantel, 1967), executed in R (Ihaka and Gentleman, 1996). Mantel tests are non-parametric so robust even when applied to non-normally distributed data. Average Euclidian pair-wise distances were calculated amongst and within the morphogroups and used to create a 4 x 4 matrix for each measurement; average uncorrected pairwise p-distances were calculated for the genetic data in MEGA (as above; see section 2.3.5) and also put into a 4 x 4 matrix, for direct comparison with the morphological data. 999 iterations were used for all matrix correlations.

2.6.1 Conchology I - Linear shell measurements

As there is no typical, definitive set of measurements for morphological analysis of this genus, 5 measurements were taken on all shells. These measurements were carried out on the specimens in the population genetics dataset as well as additional shells studied at the Museum für Naturkunde in Berlin. These measurements were: maximum height of shell, maximum length of the shell, depth of the shell below the aperture, depth of shell at highest point and the number of whorls (measured on the apex side). Figure 2.22 shows the position of these various measurements. These measurements were chosen to compare with previous conchological studies (Mandahl-Barth, 1962) but were slightly modified. For example, whorl number was counted on the apex side rather than the umbilical side, as in the author's experience, the umbilical side is more likely to be damaged or obstruct counting in some other way. Similarly, as two depth measurements were taken in this survey instead of one previously (described as 'height' in Mandahl-Barth's terminology), it was decided to move the location of the measurement around the shell slightly so as to provide a direct contrast between the two sides

of the shell, and also to make the measurements more homologous between shells. It is worth mentioning that the orientation of the shell, as seen in Figure 2.22, is traditionally associated with dextral shells. *Biomphalaria* are sinistral but as with other discoidal forms, the shell is carried inverted and so has the superficial appearance of being dextral. Mandahl-Barth (1957) referred to this condition as 'ultradextral'; he also recommended measuring the number of whorls from the convex side, which appears apical but is actually homologous to the umbilicus in other sinistral snails. As such, he presented pictures of *Biomphalaria* shells with the convex side facing out, and this convention is followed here, as in Figure 2.22.

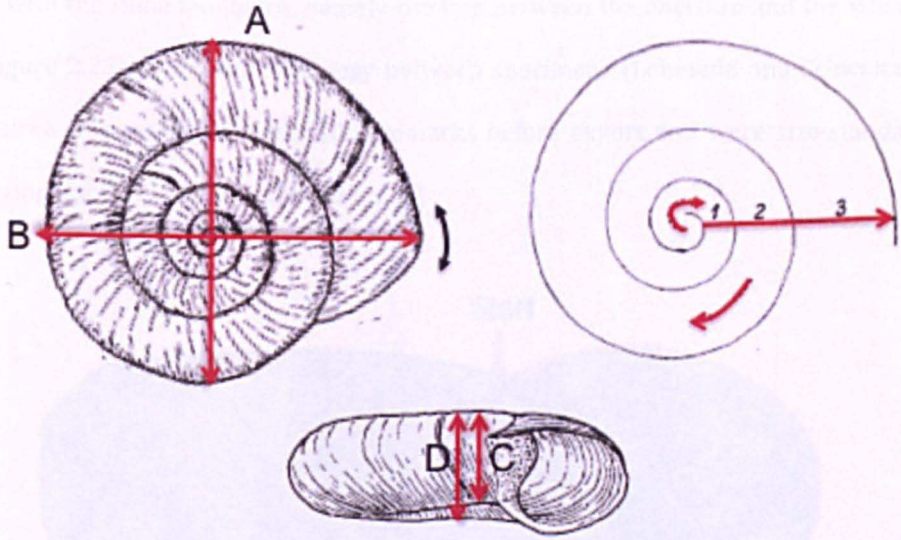

Figure 2.22 – Shell measurements

'A' refers to maximum height, 'B' is the maximum length, 'C' is the depth of the shell below the aperture and 'D' is the depth of the shell at the highest point (measure from the opposite side of the shell). The outline of the shell on the right shows the method of counting whorls, with the apex facing the viewer. Diagrams modified from Mandahl-Barth (1962).

All shells were measured using digital calipers accurate to the nearest tenth of a millimeter. Numbers of whorls were counted by eye under a dissecting microscope. Measurements were tabulated in Excel (as above), log-transformed to control for size variation and PCA/LDA executed in R (Ihaka and Gentleman, 1996).

2.6.2 Conchology II – Aperture outline analysis

A subset of the shells, consisting of those from the 7 populations analysed in the phylogenetics section above, plus the Berlin specimens, also had their aperture shapes investigated using outline analysis. This was done using photographs of each shell, oriented so the aperture was directly facing the camera, and including a ruler for later scale correction. In this way, inconsistencies in the rate of twist between different shells were avoided. The photographs were taken with a Leica DFC420 camera attachment (Leica Camera Ltd., Milton Keynes, UK) on a dissecting microscope. The dataset was compiled using tpsUtil (Rohlf, 2007b) and then each outline drawn, by hand, on the computer using tpsDig2 (Rohlf, 2007a). Each curve began and ended with the same landmark, namely the join between the aperture and the whorl beneath (see Figure 2.23), to ensure homology between specimens (Lohmann and Schweitzer, 1990). Each curve was resampled with 50 landmarks before export and were size-standardised by calculating a scale factor for each individual.

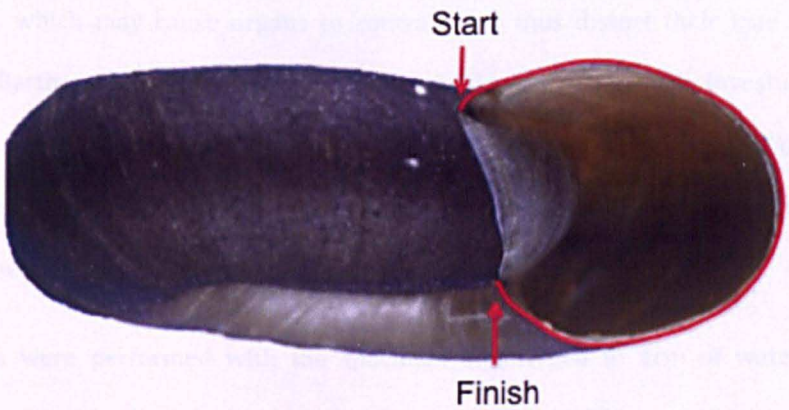

Figure 2.23 – Aperture outline methodology
The line in red represents the curve drawn by hand over the aperture.

Once drawn, the 50 x-y coordinates were converted into phi functions, which describe the process of angle change to follow the curve through each landmark, by the programme 'xy->phi converter' (see Acknowledgements for reference). This process is designed to remove variation due to orientation, scale and position and leave just 'shape' from subsequent analysis. The next step was to put the phi factors through eigenshape analysis, using Eigenshape software (see Acknowledgements for reference). Eigenshapes are effectively principal components of

outlines. Using a covariance matrix, the programme looks for the vectors among the specimens that account for the most variation (axis 1), and then sequentially all other vectors that account for variation (axes 2, then 3, etc). The programme also selected, through recursive estimation, the minimum number of landmarks required to match 97.5% of the original curve, which was 12 (MacLeod, 1999). The eigenvalues associated with a particular axis could then be plotted to look for grouping, as in a standard PCA.

2.6.3 Copulatory organ measurements

Snails that had been preserved in modified Raillet's solution were dissected to examine the relative lengths of the internal copulatory organs. Particularly in South American *Biomphalaria*, the length of the preputium, the penis sheath and the number of prostate diverticula have been considered good species characters (Pointier *et al.*, 2005, Estrada *et al.*, 2006). Sketches of these organs can be seen in the Introduction, section 1.4.4. Analyses of these characters have also been attempted on African species, but usually on ethanol-preserved specimens, which may cause organs to contract and thus distort their true shape and size (Mandahl-Barth, 1958). However, it may be that a more thorough investigation of these characters will reveal their value as identifying characters for Lake Victorian *Biomphalaria*, and so individuals from the 7 populations analysed phylogenetically (see above, section 2.3.5) were dissected to reveal their internal anatomy.

Dissections were performed with the specimen submerged in 1cm of water and under a dissecting microscope to reveal the preputium, penis sheath (these were joined together) and the prostrate, the latter of which was separated from the others for easier visualization. The number of prostate diverticula was evaluated by carefully cutting apart each diverticulum from the next and counting the total. The lengths of the preputium and penis sheath were measured by sketching the outline of the two organs using a camera lucida attachment to the dissecting microscope. A scale bar was included under the microscope so that relative length measurements could be converted to true lengths. The length measurements were log transformed to control for size variation between specimens before being submitted to PCA and LDA.

2.7 Environmental data

The environmental data collected at each site were used for descriptive comparisons as well as statistical analysis. Figure 2.24 shows how water samples were collected and measured *in situ*, in this example from a location in the middle of the lake.

Figure 2.24 – Measuring water samples from a site mid-lake, in Uganda

2.7.1 Water data

194 water samples that had been collected at sites all around Lake Victoria were analysed by support technicians in the EMMA wet chemistry laboratories at the Natural History Museum. Anions such as chloride, fluoride, sulphate, phosphate, nitrate, bromide and nitrite were detected by ion chromatography. Prior to the analysis all samples were filtered through Whatman Anotop® IC or LC grade 0.2 µm syringe filters directly into the 10 ml autosampler vials. The instrument used was an ICS-3000 system (Dionex Inc., Sunnyvale, USA) and the determination was done in conductivity mode after separation on a 2mm AS-19 analytical column and matrix removal with an ASRS® 300 suppressor. Both analytical column and conductivity cell were kept at 35 °C. 10 µl of the sample was automatically injected using autosampler AS. Eluent was produced in-line by a Reagent-Free Ion Chromatography™ module (Dionex Inc, as above) using an EluGen® KOH cartridge (also Dionex Inc.). Gradients were

created by the RFIC™ module (see Table 2.2). Before each batch of samples the instrument was calibrated using a set of six standards (see Table 2.3). Peaks with area smaller that 0.02μS*min were ignored.

Retention (min)	KOH concentration (mM)
-5	10
2	10
10	20
15	20
22	50
33	Acquisition stopped

Table 2.2 - RFIC™ module gradients for anion quantification

Standards	F^-	Cl^-	NO_2^-	Br^-	NO_3^-	SO_4^{2-}	PO_4^{3-}
1	0.2	1	1	1	1	1	2
2	0.4	2	2	2	2	2	4
3	0.8	4	4	4	4	4	8
4	2	10	10	10	10	10	20
5	4	20	20	20	20	20	40
6	20	100	100	100	100	100	200
Retention time (min)	5.1	8.1	9.4	10.6	11.7	15.8	23.5

Table 2.3 - Standards used for anion quantification and typical retention times

Experimental standards were prepared by dilution from a stock 7-anion standard

Cations such as calcium (Ca), potassium (K), magnesium (Mg) and sodium (Na) ions were determined by inductively coupled plasma atomic emission spectroscopy (ICP-AES). The instrument used was a Varian Vista Pro axially-viewed ICP-AES with a CCD detector, with a Varian SPS-5 autosampler (Varian, Inc., Palo Alto, USA). Wavelengths used were: Ca-317.933 nm, K-766.491 nm, Mg-285.213 nm and Na-589.592 nm. Multi-element calibration standards were made form single element standards in 0.3 M nitric acid.

In addition, rainfall data were extracted from the African Data Dissemination Service (http://earlywarning.usgs.gov/adds/) at 8 km^2 spatial resolution for the locations of each of these 194 sites and 10 days temporal resolution for the month that the sites were visited as well as the month previous; both months were combined to create a total rainfall value for each location. The original rainfall data were processed using IDRISI software and extracted at the data locations

using a program written in Fortran. The data were then combined with the water chemistry variables that had been collected on site, plus qualitative data on the habitat type, substrate, depth, turbidity and presence/abundance of *Biomphalaria* and other snails to create a comprehensive dataset for statistical analysis.

A separate set of water samples from 40 sites, twenty each from Ngamba and Kimi Islands, were also analysed using the above methods, and combined with detailed qualitative data. However, the sites were too close together to allow for extraction of rainfall data for each site so this was not included in the analysis.

2.7.2 Statistical analysis

The dataset containing site and snail data was analysed using progressively more sophisticated and highly parameterised models, with a view to investigate the significance of the variables within a spatial context. The outcomes used for the models related to presence/absence or abundance of *Biomphalaria* with the environmental and water chemistry data used as explanatory variables. For the Kimi/Ngamba microscale survey, the observation of snails shedding cercariae was also used as a model outcome. Results from all of the multivariate models can be seen in Chapters 6 and 9.

Each variable was first run in a univariate logistical regression model in STATA v 11 (StataCorp, 2009), with 'presence/absence' of *Biomphalaria* as the outcome. All were then included in a multivariate logistical regression model, and those with a p-value of less than 0.15 were removed. Variables were also checked for correlation with each other; if correlation was found between two variables, one was removed from the multivariate model. The data were then exported into WinBugs 1.4.3 (Lunn *et al.*, 2000) for modeling using Bayesian inference. Three models were compared: the first was a non-spatial model without random effects (in other words, a standard multivariate logistical regression, similar to the one carried out before), the second was a non-spatial 'exchangeable' model which included random effects and the last was a spatial model which also included random effects. Parameter estimates from the first were used as initial conditions for the second, more complex, model, and parameter estimates

calculated in the second model were used as initial conditions for the third. This process was repeated separately for presence/absence of *B. sudanica* and *B. choanomphala.*

A similar methodology was applied to counts of *Biomphalaria choanomphala* from the 20 sites on Kimi Island and the 20 sites on Ngamba Island, with the Shannon Index, Simpson's Index and Fisher's alpha included as variables in the univariate analysis as well. Again, only variables approaching significance ($p < 0.15$) were included in the subsequent multivariate models. In this case, the outcome was 'counts' rather than 'presence/absence', and so the distribution was inferred to be negative binomial (based on the variance being much greater than the mean of the counts) and so a negative binomial, rather than logistical, regression model was used. The presence/absence of snails shedding trematode cercariae was used as a separate outcome, modelled with a probability binomial distribution.

Where spatial autocorrelation was observed, the snail abundance data could be interpolated by kriging, to create an approximation of a predictive map. Kriging is a geostatistical technique that uses known values of a chosen variable (often referred to as the 'random field') to interpolate unknown values at nearby geographical locations. Ordinary kriging was used, which assumes a constant but unknown mean of the random field (also referred to as the expectation of the random field) and a known variogram. In this case, a spherical model was used to create the semivariogram, with a variable search radius along 12 points. A map of the variance of the kriging was also calculated. All kriging and associated statistical measurements were executed in ArcGIS (ESRI (UK) Ltd., Aylesbury, UK).

3 Intestinal schistosomiasis and soil-transmitted helminthiasis in Ugandan schoolchildren: a rapid mapping assessment along the Lake Victorian shoreline

3.1 Abstract

Even with a national control programme, intestinal schistosomiasis continues to be a major public health problem in school-aged children and other community members in Uganda. This is especially the case in the environments around the Great Lakes, where disease transmission is high, such as Lake Victoria's shoreline. Moreover, in the most remote areas, some schools might periodically miss large-scale drug administrations owing to their inaccessibility. To provide contemporary monitoring and surveillance data, 27 schools along the lakeshore were surveyed with a rapid assessment protocol to determine both prevalence and intensity of *S. mansoni* and soil-transmitted helminth infections. In total, 25 (92.6%) of schools were positive for *S. mansoni*, and average prevalence across the surveyed children was 42% with an average infection intensity of 634 eggs per gram of faeces. Mean prevalence of *Trichuris trichiura*, *Ascaris lumbricoides* and hookworm was 12.9%, 9.3% and 2.4% respectively. Results from questionnaire data revealed a high level of itinerancy among the children, and a total of 38.2% reported to have never received treatment for schistosomiasis, despite 96% living in districts targeted by the national control programme. A birthplace outside of Uganda was a significant predictor for increased risk of schistosomiasis infection (odds ratio [OR] = 9.6), and being resident at a school for less than a year was significantly associated with absence of praziquantel treatment (OR = 0.3). Univariate regression analysis showed a trend of increasing prevalence of schistosomiasis towards the eastern region of Uganda, while semivariograms of infection prevalence demonstrated a range of spatial autocorrelation of ~78 km. Soil-transmitted helminth infections were more common in the Western region. Our results emphasise how social and demographic variables such as migration may affect epidemiological trends and confound the impact of existing treatment regimes.

3.2 Contributions of the author

The field work for the data presented in this chapter was carried out during the Uganda 2008 field mission. The team consisted of the author, Dr Russell Stothard, Dr Aslak Jørgensen, Moses Adriko, Moses Arinaitwe and Dr Francis Kazibwe. Of these, the author and Moses Arinaitwe were present for the full duration of the field mission. As such, the author was present at all school visits. Questionnaires were administered by the author, with linguistic assistance from one of the technicians (Moses Adriko or Moses Arinaitwe) or a local health officer/community mobiliser. Parasitological slides were made by all members of the team and reads were done in the field by either Moses Adriko or Moses Arinaitwe, and additionally by the author. All statistical analyses, including the geospatial variograms, were done by the author.

3.3 Introduction

Schistosomiasis was first detected in north-west Uganda in 1902, and subsequent surveys also discovered the disease in Lake Victoria (Emmanuel and Doering, 2008). Lakeshore communities are typically dependent on water from these lakes for various daily activities, including cooking, bathing and washing clothes. Fishing is usually a major occupation of the men of these communities, further bringing them into contact with the water, and sometimes creating itinerant communities that move to follow productive fishing zones (Reynolds and Kitakule, 1991). Since the 1950s, the Nile perch fishery in Lake Victoria has grown into a multi-million dollar business, and thus provides a lucrative means of earning cash rather than just subsistence. These circumstances dramatically increase infection risk with *S. mansoni* for people living close to the lakeshore (Stothard *et al.*, 2005).

To combat schistosomiasis at the national level, efforts to control the disease are underway in areas considered high-risk, based on previous baseline surveys, snail distributions and more recently, environmental predictors extracted from remote sensing and satellite information (Brooker *et al.*, 2001, Stensgaard *et al.*, 2006). Chemotherapy with praziquantel (PZQ) is currently the mainstay of control, which is available at a low cost and also donated to control

programmes by pharmaceutical companies manufacturing the drug (Doenhoff *et al.*, 2008, Fleming *et al.*, 2009). Uganda has had a national control programme in place since 2003, instigated in partnership with the Schistosomiasis Control Initiative (SCI) (Kabatereine *et al.*, 2006b). Schools were targeted for PZQ mass drug administration (MDA) based on the increased benefits of reducing infection burdens in children compared to adults and also due to the ease of providing treatment (Kabatereine *et al.*, 2007, Zhang *et al.*, 2007). The programme was expanded in 2007 to include new districts and also to integrate treatment with albendazole against common soil-transmitted helminth infections such as hookworm, *Ascaris lumbricoides* and *Trichuris trichiura*; this drug had previously been administered separately (Kabatereine *et al.*, 2006a). The five year initial run of the programme finished in 2008 (Fenwick *et al.*, 2009); however, the government is committed to maintaining the programme, assisted now by new funding from the United States Agency for International Development (USAID: http://www.rti.org/page.cfm/Neglected_Tropical_Diseases), the Global Network for Neglected Tropical Diseases (GNNTD: www.globalnetwork.org) as well as SCI.

Despite these efforts to measure precisely geographical patterns of schistosomiasis risk, many of the variables used for predicting prevalence of the disease have large margins of error, and may fall short of accurate 'ground-truthing', the description of the situation as actually observed in the field. For example, if a school is located in a generally low-transmission zone, it may be precluded from regular PZQ administration, despite actually being highly affected by schistosomiasis. As such, traditional monitoring at a local level is vital in ensuring efficiency and maximum cost-effectiveness within the control programme. It is within this context that as control is fully rolled out, further efforts at developing and executing new monitoring schemes are needed.

This chapter presents data using a rapid assessment methodology to examine the prevalence and intensity of *S. mansoni* and soil-transmitted helminth infection in schools around the Ugandan perimeter of Lake Victoria. The aim was to shed light on the infection status of these schools, and thus simultaneously to assess outreach of the national control programme within the six districts examined. Also, the survey was designed to integrate demographical data and

thus elucidate other factors that might contribute to continued high levels of infection and/or re-infection with these intestinal parasites, such as place of origin, past treatment history and time of residence in a particular community. As this was the first field mission associated with this Ph.D. research, snails and parasite larvae were also collected for later spatial and molecular analyses.

3.4 Methods

3.4.1 Study design and data collection

The data presented here were collected over a three-week period in February and March of 2008. 27 schools were targeted, on an *ad hoc* basis, for inclusion in the study, from 6 districts across the shoreline of Lake Victoria (see Figure 3.1).

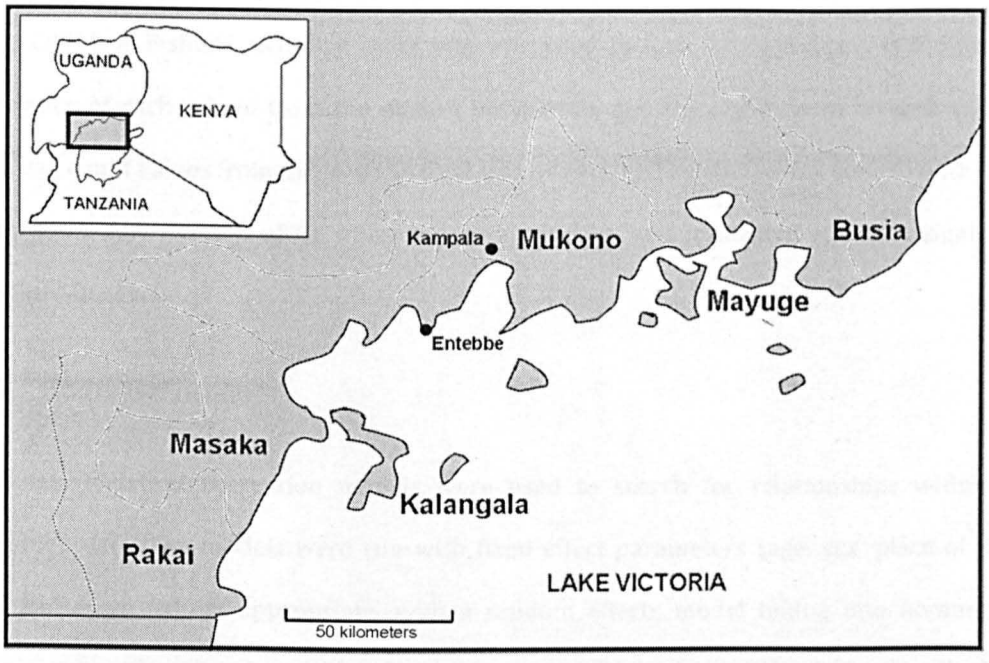

Figure 3.1 – Map of the shoreline of Lake Victoria in Uganda, showing the 6 districts surveyed.

Children were selected for the surveys as described in section 2.1.1 of the Methods chapter, and single stool samples collected. Samples were processed to produce double smear Kato-Katz slides (section 2.2.1) and all children were also asked to complete a short, structured questionnaire regarding their place of birth, length of time at the school, previous treatment history and knowledge of intestinal schistosomiasis (see section 2.1.2). Information on consent and ethical clearance can be found in section 2.1.3.

3.4.2 Parasitology

Slides were inspected for presence of eggs of various parasite species: *S. mansoni*, *Ascaris lumbricoides*, *Trichuris trichiura* and hookworm (section 2.2.1). Counts were averaged across

the double smear and multiplied by 24 to create an 'egg per gram of faeces' value (EPG). Intensities of infection were calculated according to WHO guidelines (WHO, 2002), whereby an infection is designated as 'light' when EPG is <99, 'moderate' if 100–399 and 'heavy' if >400.

The geometric means of all the egg counts per school were calculated, as were the arithmetic means of all positive egg counts per school ('1' was added to all egg counts when calculating geometric means). As the egg count data were not normally distributed, an exact binomial was used to calculate prevalence and confidence intervals (CI) per school and per district (Armitage *et al.*, 2001). A Fisher's two-way exact test was used to look for significant differences in prevalence of each school from the overall prevalence, and the significance of each school's mean egg count values from those of the total was determined by comparing each CI with those of the total. Any overlap of CI when comparing means was indicative of a non-significant difference in value.

3.4.3 Statistical analyses

Univariate logistical regression models were used to search for relationships within the assembled data. The models were run with fixed effect parameters (age, sex, place of birth, etc.), and again, where appropriate, with a random effects model taking into account the schools as a confounding factor. Significance ($p<0.05$), odds ratios (OR) and CI were calculated for each. As each site was georeferenced using handheld GPS devices, it was also possible to analyse the prevalence data for spatial autocorrelation: this examines whether there is greater commonality between sites that are closer together than those that are further away from each other, and at what scale this relationship acts over. Omnidirectional and directional semivariograms were calculated to test for spatial autocorrelation: covariance parameters were estimated using a parametric spherical model and fitted with weighted least squares to compare between different experimentally-derived semivariograms for best fit. All statistical tests were performed using R (Ihaka and Gentleman, 1996).

3.5 Results

3.5.1 Infection with S. mansoni

Prevalence of *S. mansoni* infection, geometric mean of the overall EPGs and the arithmetic EPG mean for infected children were calculated per school/site and overall, for each the Western and Eastern regions (Tables 3.1 and 3.2, respectively).

School (Map ID) *GPS coordinates (dec.°)*	District	N	n	% Prev (95% CIs)	Geometric mean (95% CIs)	Arithmetic mean (95% CIs)
Bukakata (2) *S0.273267 E32.026517*	Masaka	15	6	40.0 (16.3-67.7)	7.31 (5.99-8.64)	130.00§ (33.71-226.29)
Bridge of Hope (9) *S0.322220 E32.284466*	Kalangala	18	5	27.8 (9.7-53.5)	2.14§ (1.18-3.10)	230.40 (0.00-595.41)
Bubeke (13) *S0.320250 E32.575500*	Kalangala	18	3	16.7 (3.6-41.4)	0.65§ (0.10-1.20)	24.00§ (0.48-47.52)
Bugoma (3) *S0.257183 E32.066383*	Kalangala	16	9	56.3 (29.9-80.2)	10.23* (9.06-11.41)	148.00§ (39.72-256.28)
Bunyama (10) *S0.363533 E32.294883*	Kalangala	13	4	30.7 9.1-61.4)	2.34§ (1.30-3.37)	54.00§ (29.52-78.48)
Buyange (11) *S0.351583 E32.571533*	Kalangala	13	0	0.0§ (0.0-24.7)	0.00§ (0.00-0.00)	NA
Jaana (12) *S0.233800 E32.575083*	Kalangala	16	6	37.5 (15.2-64.6)	5.35§ (4.06-6.63)	328.00 (0.00-664.74)
Kazira (5) *S0.322849 E32.193616*	Kalangala	19	2	10.5§ (1.3-33.1)	0.45§ (0.00-0.96)	36.00§ (12.48-59.52)
Kibanga (8) *S0.323950 E32.284250*	Kalangala	16	5	31.3 (11.0-58.7)	2.52§ (1.56-3.49)	72.00§ (20.47-123.53)
Mulabana (6) *S0.438650 E32.227900*	Kalangala	17	9	52.9 (27.8-77.0)	6.41 (5.37-7.44)	118.67§ (0.00-255.27)
St Kizito (4) *S0.300666 E32.141233*	Kalangala	18	4	22.2 (6.4-47.6)	1.20§ (0.48-1.92)	48.00§ (0.00-96.01)
St Thereza (7) *S0.417300 E32.228283*	Kalangala	18	5	27.8 (9.7-53.5)	2.43§ (1.48-3.37)	86.40§ (62.41-110.39)
Goma (1) *S0.915350 E31.767300*	Rakai	18	1	5.5§ (0.1-27.3)	0.26§ (0.00-0.70)	60.00§ (only 1 inf.)
Total for western region	--	**232**	**67**	**28.9§ (23.1-35.2)**	**2.47§ (2.20-2.73)**	**180.54§ (100.50-260.57)**
Overall Total	--	**456**	**190**	**41.7 (37.1-46.3)**	**7.30 (7.17-7.44)**	**634.25 (466.25-802.25)**

Table 3.1 – Prevalence and intensity of infection with *S. mansoni* in schoolchildren from Western districts

Geometric means were calculated for all egg counts, whereas arithmetic means were calculated from positive counts only. 'dec. °' stands for decimal degrees. 'N' = number of children surveyed; 'n' = number positive; § = significantly lower value than overall

School (Map ID) *GPS coordinates(dec.°)*	District	N	n	% Prev (95% CIs)	Geometric mean (95% CIs)	Arithmetic mean (95% CIs)
Maduwa (27) *N0.253067 E33.988933*	Busia	18	2	11.1§ (1.4-34.7)	0.68§ (0.00-1.41)	234.00 (0.00-645.60)
Majanji (26) *N0.263167 E33.984750*	Busia	17	0	0.0§ (0.0-19.5)	0.00§ (0.00-0.00)	NA
Bugoto (23) *N0.318117 E33.627033*	Mayuge	17	10	58.8 (32.9-81.6)	10.56* (9.46-11.68)	121.20§ (38.64-203.76)
Bumba (24) *N0.031083 E33.644650*	Mayuge	18	13	72.2* (46.5-90.3)	29.80* (28.68-30.92)	228.92§ (68.61-389.23)
Jagusi (20) *N0.156183 E33.566017*	Mayuge	18	4	22.2 (6.4-47.6)	1.52§ (0.70-2.35)	66.00§ (43.48-88.52)
Kaaza (21) *N0.111767 E33.602150*	Mayuge	17	10	58.9 (32.9-81.6)	18.96* (17.58-20.33)	488.40 (63.68-913.12)
Sagitu (25) *N0.003033 E33.658600*	Mayuge	17	14	82.4* (56.6-96.2)	84.96* (83.80-86.13)	456.00 (178.40-733.60)
Serinyabi (22) *N0.052350 E33.605833*	Mayuge	17	12	70.6* (44.0-89.7)	31.63* (30.47-32.80)	214.00§ (94.67-333.33)
Busagazi (16) *N0.240300 E33.137083*	Mukono	17	9	52.9 (27.8-77.0)	7.65 (6.20-9.11)	406.00 (40.03-771.97)
Kasimizi (19) *N0.185583 E33.215217*	Mukono	18	18	100.0* (81.5-100.0)	1337.43* (1336.81-1338.04)	2165.65* (1430.62-2900.68)
Kimi (14) *S0.086383 E32.652200*	Mukono	18	8	44.4 (21.5-69.2)	9.94* (8.58-11.30)	549.00 (61.12-1036.88)
Kisu (15) *N0.014767 E32.767167*	Mukono	16	14	87.5* (61.7-98.4)	309.89* (308.49-311.28)	1842.00* (843.94-2840.06)
Makonge (17) *N0.270217 E33.205883*	Mukono	16	5	31.3 (11.0-58.7)	5.86 (4.40-7.31)	552.00 (230.54-873.46)
Namatale (18) *N0.173200 E33.183967*	Mukono	17	12	70.6* (44.0-89.7)	102.16* (100.50-103.83)	1698.55 (384.37-3012.72)
Total for eastern region	--	**224**	**123**	**54.9*** **(48.1-61.5)**	**19.79*** **(19.38-20.20)**	**891.86** **(643.84-1139.89)**
Overall Total	--	**456**	**190**	**41.7** **(37.1-46.3)**	**7.30** **(7.17-7.44)**	**634.25** **(466.25-802.25)**

Table 3.2 - Prevalence and intensity of infection with *S. mansoni* in schoolchildren from Eastern districts

Geometric means were calculated for all egg counts, whereas arithmetic means were calculated from positive counts only. 'dec. °' stands for decimal degrees. 'N' = number of children surveyed; 'n' = number positive; § = significantly lower value than overall; * = significantly higher value than overall.

Total mean prevalence was 41.7% (95% CI = 37.1-46.3%), geometric mean of all EPG values was 7.30 (95% CI = 7.17-7.44) and the arithmetic mean of infected cases was 634.3 EPG (95% CI = 466.3-802.3). At the unit of the school, there were six schools with significantly higher overall prevalence, 10 schools with significantly higher geometric or arithmetic mean egg

counts, five schools with lower prevalence and 19 schools with one or both of its mean egg counts lower than the total average. Four schools had mixtures of significantly higher and lower results, when compared to the total sample. These data were then used to map the prevalences and mean arithmetic intensity of infection together onto the locations of each of the schools (Figure 3.2).

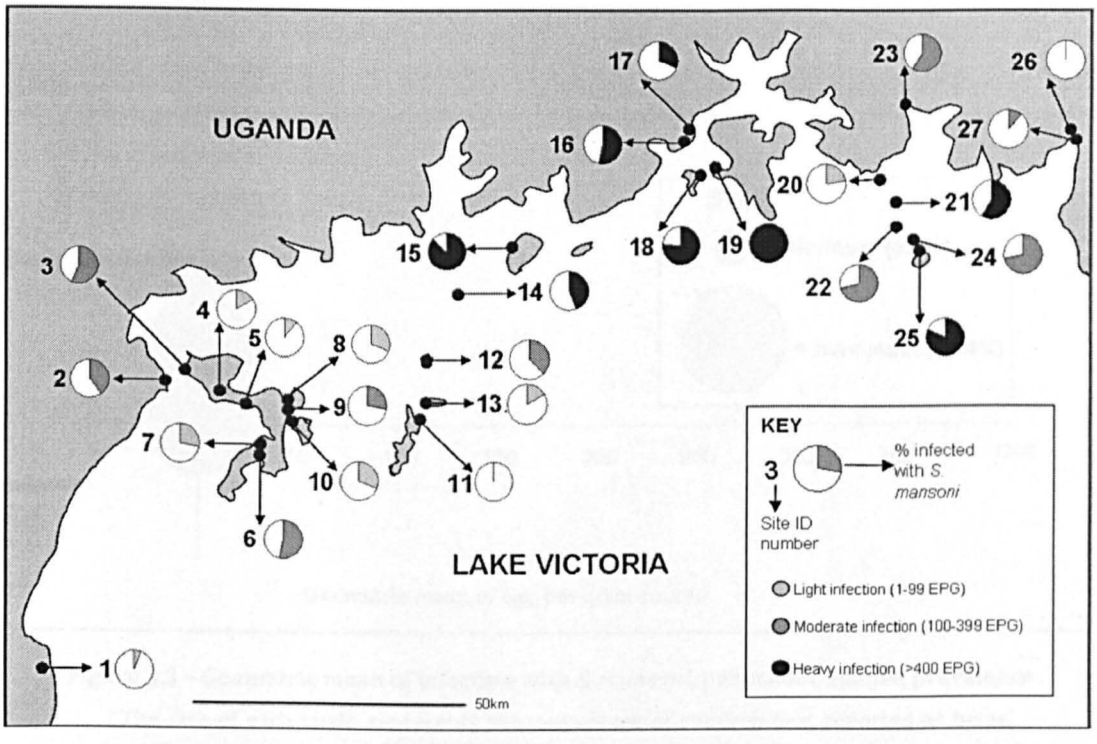

Figure 3.2 – Map of prevalence and intensity of infection across all 27 schools
The prevalence of infection at each locality is shown by the size of the pie slice; intensity of infection (based on WHO categories) is demonstrated by the shading of the slice. Site numbers correspond to those in Tables 3.1 and 3.2.

Dividing the geographic range of the surveyed schools into 'Eastern' and 'Western' regions based on longitude, the districts of Rakai, Kalangala and Masaka were considered 'Western', whereas Mukono, Mayuge and Busia were 'Eastern'. The Western districts had, on average, significantly lower prevalence, geometric mean of egg counts and arithmetic mean of positive egg counts compared to the overall total, whereas the Eastern districts had significantly higher prevalence and geometric egg count. Geometric means of EPG for each school were also plotted against prevalence, showing an asymptotic relationship (Figure 3.3). Place of birth data were also included in this analysis (see next section for further demographic results).

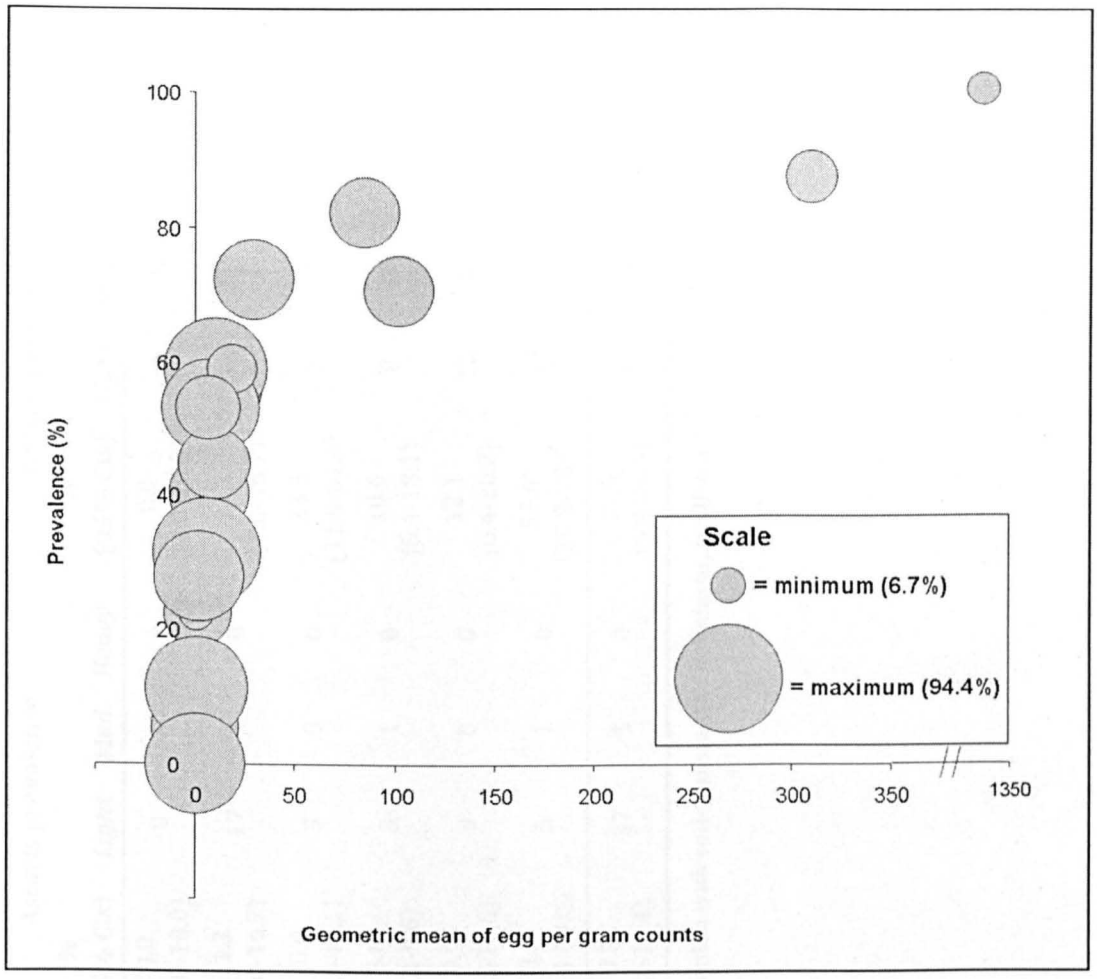

Figure 3.3 – Geometric mean of infection with *S. mansoni*, per school, against prevalence

The size of each circle represents the percentage of children that reported as being born in that district; maximum and minimum values are shown in the scale.

3.5.2 Infections with soil-transmitted helminths

At a district level, soil-helminths were also widespread, but showed strong heterogeneity (Table 3.3).

Four out of six districts (66.7%) were negative for hookworm, and the overall prevalence was low (2.4%; 95% CI = 1.2-4.3%). Only one district (Busia) was negative for both *A. lumbricoides* and *T. trichiura*, and average prevalence in the other districts was 9.3% (95% CI = 6.8-12.4%) and 12.9% (95% CI = 9.9-16.3%) respectively. Soil-transmitted helminth infections, stratified by infection intensity, are presented in Table 3.3.

School (Map ID)	Hookworm prevalence				*Ascaris* prevalence				*Trichuris* prevalence			
	% (95% CIs)	*Light*	*Mod.*	*Heavy*	% (95% CIs)	*Light*	*Mod.*	*Heavy*	% (95% CIs)	*Light*	*Mod.*	*Heavy*
Busia	0.0 (0.0-10.0)	0	0	0	0.0 (0.0-10.0)	0	0	0	0.0 (0.0-10.0)	0	0	0
Kalangala	3.4 (1.2-7.2)	5	0	1	11.2 (7.0-16.7)	17	3	0	11.2 (7.0-16.7)	19	1	0
Masaka	0.0 (0.0-21.8)	0	0	0	20.0 (4.3-48.1)	3	0	0	33.3 (11.8-61.6)	5	0	0
Mayuge	0.0 (0.0-3.5)	0	0	0	3.8 (1.1-9.6)	3	1	0	10.6 (5.4-18.1)	8	3	0
Mukono	0.0 (0.0-3.7)	0	0	0	9.1 (4.2-16.6)	9	0	0	12.1 (6.4-20.2)	12	0	0
Rakai	27.8* (9.7-53.5)	5	0	0	33.3* (13.3-59.0)	5	1	0	55.6* (30.8-78.5)	8	2	0
TOTAL	2.4 (1.2-4.3)	10	0	1	9.3 (6.8-12.4)	37	5	0	12.9 (9.9-16.3)	52	6	0

Table 3.3 – Prevalence and intensity of infection with soil-transmitted helminths, by district

3.5.3 Questionnaire results

The survey responses revealed a large number of children (48.9%) born outside the district in which they currently lived (Table 3.4).

District (*N*)	District of birth (%)				Length of stay (%)			
	This District	*Other L.V.*	*Other UG*	*KE / TZ*	*Whole Life*	*< 1 year*	*1-3 yrs*	*> 3 yrs*
Busia (*36*)	91.7	0.0	8.3	0.0	2.8	2.8	2.8	91.7
Kalangala (*189*)	40.8	43.9	14.8	0.5	15.7	0.8	21.5	62.0
Masaka (*14*)	57.1	28.6	14.3	0.0	0.0	27.3	18.2	54.5
Mayuge (*105*)	44.8	23.8	20.0	11.4	18.6	10.8	22.5	48.0
Mukono (*92*)	59.8	9.8	28.2	2.2	4.4	15.4	14.3	57.1
Rakai (*18*)	66.7	27.8	5.5	0.0	5.6	11.1	27.8	55.6
TOTAL (*454*)	**51.1**	**27.8**	**17.8**	**3.3**	**5.5**	**17.9**	**15.0**	**59.4**

Table 3.4 – Demographical questionnaire responses, by district

'L.V.' stands for Lake Victoria, 'KE' stands for Kenya and 'TZ' stands for Tanzania.

Despite existing expectations of considerable movement of communities within Lake Victoria, 17.8% of children were from a Ugandan district that did not border on the lake. Similarly, length of residence in a particular area varied considerably, with 23.4% of children having only lived in that district for less than three years, and 5.5% less than one year (Table 4). Treatment history also varied widely across the six districts surveyed, from 0% to almost 100% (Figure 4); overall, 61.8 % (95% CI = 57.1-66.2%) of children reported having received PZQ at some point in their memory. Treatment history was closely correlated with knowledge of the disease (Figure 3.4, R^2=0.88).

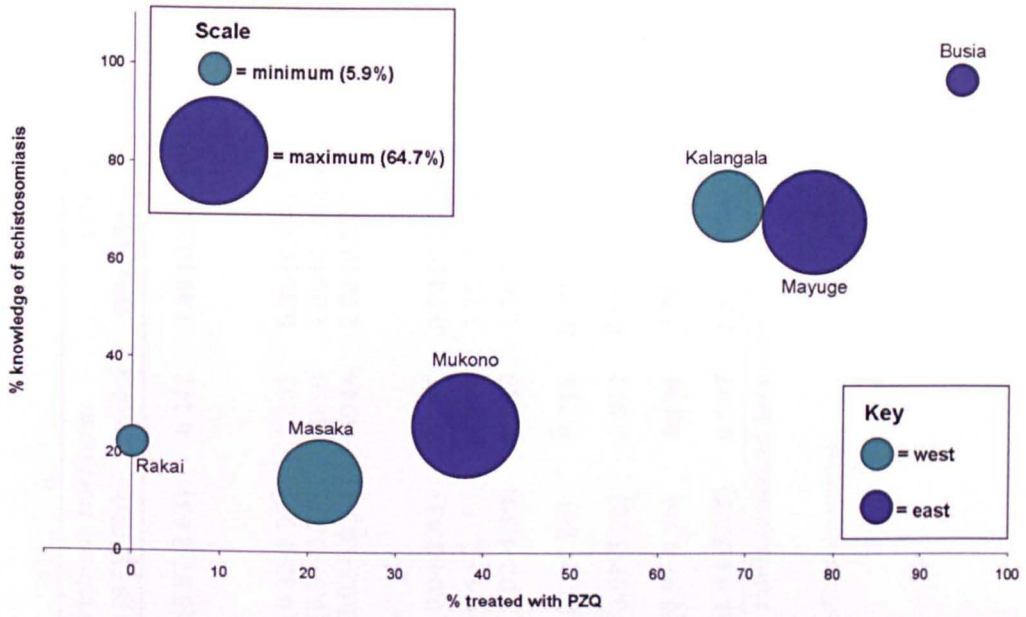

Figure 3.4 – Correlation of treatment with PZQ against knowledge of schistosomiasis
The area of each bubble represents the prevalence of *S. mansoni* recorded in that district and the colour represents the location of each district.

3.5.4 Univariate regression modeling

Univariate fixed effect models were used to test for predictors of infection status (Table 3.5). For schistosomiasis, sex and age were not significant predictors of infection status, and nor was previous treatment history. Significant risk factors were place of birth outside Uganda (OR = 9.6, 95% CI = 2.1-43.7, p = 0.003) compared to local birth, and region of current residence; children in the Western region were only a third as likely to be infected with schistosomiasis compared to children in the Eastern region (OR = 0.33, 95% CI = 0.23-0.49, p < 0.001). Soil-transmitted helminth infections showed a reverse relationship for this factor, with an increased likelihood of infection in children from the Eastern region (OR = 2.85, 95% CI = 1.75-4.64, p < 0.001). Infection was also negatively associated with treatment with PZQ (OR = 0.49, 95% CI = 0.31-0.78, p = 0.003).

Variable (Baseline; Factor)	*Schistosoma mansoni*		*Ascaris lumbricoides*		*Trichuris trichiura*		**STHs overall**	
	ORs (95% CIs)	p-value	ORs (95% CIs)	p-value	ORs (95% CIs)	p-value	ORs (95% CIs)	p-value
Infection with *S.mansoni* (No; Yes)	NA	NA	1.20 (0.64-2.28)	0.569	1.51 (0.87-2.63)	0.142	1.46 (0.93-2.30)	0.10
Place of Birth (Local;								
Other Lake district	0.74 (0.46-1.18)	0.208	1.19 (0.54-2.64)	0.664	0.66 (0.32-1.37)	0.261	0.86 (0.49-1.52)	0.608
Other Uganda district	1.30 (0.76-2.20)	0.336	1.36 (0.56-3.28)	0.498	1.07 (0.51-2.27)	0.850	1.08 (0.57-2.02)	0.818
Other country	**9.60 (2.11-43.68)**	**0.003**	2.93 (0.75-11.39)	0.121	**3.22 (1.03-10.11)**	**0.044**	2.53 (0.86-7.49)	0.093
Length of residence (Whole life;								
< 1 year	0.97 (0.39-2.44)	0.951	1.19 (0.54-2.63)	0.664	0.31 (0.04-2.43)	0.267	0.74 (0.21-2.62)	0.635
1 year - < 3 years	1.67 (0.97-2.90)	0.066	1.36 (0.56-3.28)	0.498	0.97 (0.44-2.18)	0.949	1.75 (0.93-3.29)	0.085
3 years - < whole life	1.70 (0.94-3.06)	0.081	2.93 (0.75-11.39)	0.121	**2.09 (1.02-4.30)**	**0.045**	1.76 (0.90-3.47)	0.100
PZQ treatment (No; Yes)	1.21 (0.82-1.79)	0.347	**0.30 (0.15-0.60)**	**<0.001**	0.72 (0.4-1.26)	0.245	**0.49 (0.31-0.78)**	**0.003**
Sex (Male; Female)	1.16 (0.57-2.37)	0.687	**4.24 (1.89-9.51)**	**<0.001**	1.30 (0.52-3.26)	0.578	**2.09 (1.03-4.27)**	**0.042**
Age (continuous; +1)	1.02 (0.95-1.08)	0.632	0.95 (0.85-1.06)	0.345	1.06 (0.97-1.16)	0.194	1.03 (0.96-1.11)	0.446
Region (East; West)	**0.33 (0.23-0.49)**	**<0.001**	**2.63 (1.31-5.27)**	**0.007**	**1.84 (1.04-3.25)**	**0.037**	**2.85 (1.75-4.64)**	**<0.001**

Table 3.5 - Univariate logistical regression models of infection with *S. mansoni* and/or soil-transmitted helminths

'CI' stands for confidence intervals and 'OR' stands for odds ratios. Significant findings are in bold.

A random effects model controlling for variation at the school level was added to the univariate analysis for factors which were considered independent of school, normally distributed and not highly skewed (age, length of residence and region). This model revealed increased age as a predictor for an infection with *A. lumbircoides* (OR = 0.91, 95% CI = 0.84-0.99, p = 0.037), and an *S. mansoni* infection as a predictor for soil-transmitted helminths (OR = 1.79, 95% CI = 1.07-2.97, p = 0.027) but no other changes in significance occurred.

Treatment was also tested for covariates among the data (Table 3.6). The model statistically confirmed the significance of knowledge of the disease as a correlate of PZQ treatment (OR = 25.3, 95% CI = 15.2-41.8, p<0.001). It also showed that children who had been resident in an area for less than one year were significantly more likely not to have ever received treatment (OR = 0.3, 95% CI = 0.1-0.7, p = 0.006). Children in the Eastern region were more likely overall to have received treatment compared to in the Western region (OR = 0.6, 95% CI = 0.4-0.8, p = 0.004).

Variable (Baseline; Factor)	***Treatment with PZQ***	
	OR (95% CI)	p-value
Place of Birth (Local;		
Other Lake district	0.7 (0.5-1.2)	0.192
Other Uganda district	**0.5 (0.3-0.9)**	**0.015**
Other country	2.0 (0.5-7.3)	0.294
Length of residence (Whole life;		
< 1 year	**0.3 (0.1-0.7)**	**0.006**
1 year - < 3 years	0.6 (0.3-1.1)	0.092
3 years - < whole life	1.2 (0.6-2.4)	0.522
Knowledge of Bilharzia (No; Yes)	**25.3 (15.2-41.8)**	**<0.001**
Age (continuous; +1)	**1.2 (1.1-1.3)**	**<0.001**
Region (East; West)	**0.6 (0.4-0.8)**	**0.004**

Table 3.6 - Univariate logistical regression models of incidence of treatment with PZQ

'CI' stands for confidence intervals; 'OR' stands for odds ratios.

3.5.5 Spatial analysis

A number of experimentally derived variograms were calculated, their lines tested for best fit to the data and the two best models selected (Figure 3.5). The first was omnidirectional, 10 bin and fitted with a 'sphere' correlation model, and the other was 12 bin, directional for 60°, and also with a 'sphere' correlation model.

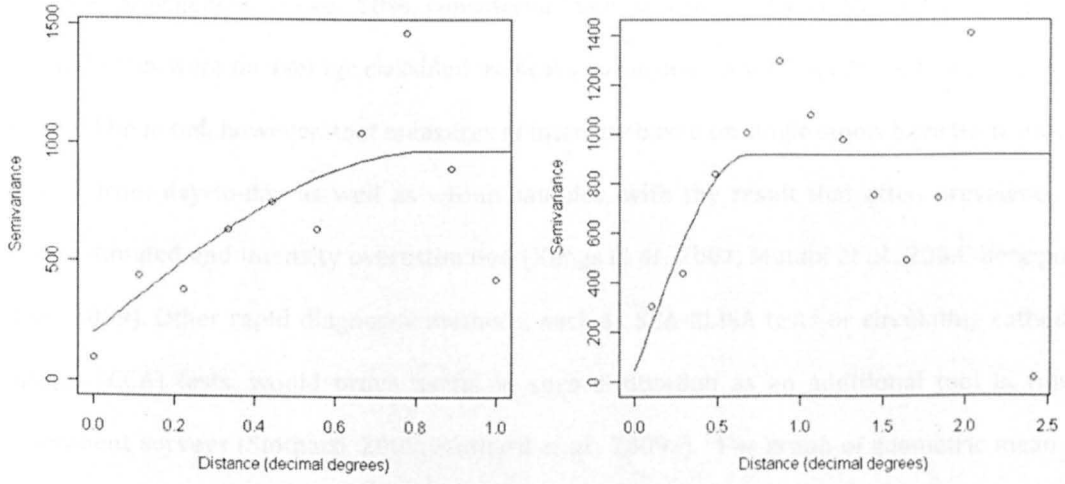

Figure 3.5 – Experimental semivariograms of *S. mansoni* prevalence data
The left-hand figure is omnidirectional while the right-hand figure is directional for 60°: both were fitted with a 'sphere' correlation model.

The omnidirectional and 60° model showed spatial autocorrelation existing within the data up to a range of 0.7 and 0.8 decimal degrees, respectively. As the study region falls directly across the equator, this corresponds to spatial autocorrelation of between 77.9 and 89.1 km.

3.6 Discussion

3.6.1 Prevalence and intensity of infection with S. mansoni

The survey found that intestinal schistosomiasis is still ubiquitous around the Lake Victorian shoreline of Uganda, with prevalence and intensity of infection significantly higher in the Eastern region. Eleven schools, which were in districts covered by large-scale administration of PZQ, had prevalences of over 50%, considered 'high prevalence' by WHO (WHO, 2002), and infected cases were on average classified as 'heavy' according to WHO guidelines (WHO, 2002). It should be noted, however, that measures of intensity based on single stools have been shown to vary from day-to-day as well as within samples, with the result that often prevalence is underestimated and intensity overestimated (Kongs *et al.*, 2001, Mutapi *et al.*, 2003, Bergquist *et al.*, 2009). Other rapid diagnostic methods, such as SEA-ELISA tests or circulating cathodic antigen (CCA) tests, would prove useful in such a situation as an additional tool in rapid assessment surveys (Stothard, 2009, Stothard *et al.*, 2009a). The graph of geometric mean of EPGs against prevalence for each school showed a negative binomial (asymptotic) shape, which is consistent with expected observations indicating overdispersal of heavy infection in a few individuals (Anderson and May, 1991).

3.6.2 Soil-transmitted helminths

The prevalence of soil-transmitted helminth infections were much lower than those of *S. mansoni* overall, and were also differently geographically distributed. The distribution of soil-transmitted helminth infections went exactly opposite to that of *S. mansoni*, with significantly lower prevalences of infection in the Eastern region as compared to the Western region, although this effect was not seen at the district level. The regional scale corresponds to previous studies which also indicated a higher prevalence of *A. lumbricoides* in the south-west of Uganda (Kabatereine *et al.*, 2005). This study also found a high incidence of soil-transmitted helminths on islands in Lake Victoria, known to be hot-spots for transmission but which are difficult to reach from the mainland by large-scale drug administration programmes due to the

prohibitive logistical and access costs. The implications of these findings are that monitoring at different scales could produce different conclusions as to the significance of a particular parasite in a given community, district or region (Brooker and Clements, 2009), and perhaps lend further weight to calls to increase funds for mobilisation of treatment to less readily accessible communities.

Although treatment with albendazole was not specifically surveyed for in this study, it can be reasonably assumed to be closely correlated with PZQ administration, as the two were combined in the national control programme in 2003. As such, it was encouraging to find a strong positive relationship between history of PZQ treatment and lack of soil-transmitted helminth infection, as this could be seen as an indication of the success of treatment interventions. However, treatment levels were lower in the Western region of the country, which is also where soil-transmitted helminths showed the highest prevalence; these areas should become of greater importance for future treatment interventions using albendazole.

3.6.3 Human migration and demography

Most noticeably, the results of the survey showed high levels of migration throughout the Lake Victoria region of Uganda, and thus may influence the geographical patterns of disease observed. This aspect of human behaviour is often overlooked in parasitological surveys, which are more concerned with the current state of infection in a particular locality, although migration is commonly cited as a significant factor in the spread of other infectious diseases (Lurie *et al.*, 2003, Weiss and McMichael, 2004). The implications of these findings are manifold. First of all, although the low sample size (n=15) weakens the finding, the significantly higher chance of children born outside Uganda being infected with *S. mansoni* could hint at a major obstacle facing the success of a national control programme emphasizing PZQ treatment; if children are coming in from countries without large-scale drug administration, they could provide a continuous source of parasite population in the communities to which they move. They, along with in-country migrants, may also be at higher risk of missing annual treatments, through frequent changes in school. This hypothesis is strengthened by the finding that children who had been resident in a location for less than a year were significantly less likely to

have ever received treatment, and the p-value for the group who had been resident between one and three years was also approaching significance. Frequent movement may also act as a mixing effect among communities, thus contributing to the spatial autocorrelation shown in the data. To emphasise this point, a follow-up survey at two schools (Kisu and Kimi), conducted 10 months after the initial field work, only succeeded in re-finding two of the 34 children previously surveyed at these sites, demonstrating once again the high level of itinerancy among lakeshore communities.

3.6.4 Antihelminthic treatment

In this study, treatment history was not found to be significantly associated with reduced prevalence of schistosomiasis, potentially due to a number of reasons. First of all, perhaps annual treatment is not enough, and in high transmission zones, biannual treatment may be necessary to have a more enduring impact on disease prevalence. It is also worth recalling that due to low perceived thresholds of prevalence and intensity of *S. mansoni*, Rakai has not been included in Uganda's national treatment programme for large-scale drug administration in schools, and Masaka only included as of 2007. This may partially account for the findings of this survey of lower treatment incidence in the Western region. Secondly, it suggests that perhaps coverage across districts is not as thorough as it could be, and that certain isolated or difficult-to-access schools are being missed by treatment interventions. The study was designed specifically to target schools on an *ad hoc* basis as an 'ambush' strategy, to prevent visiting only schools which are commonly surveyed or well-known by local health officers. By doing so, some of the more overlooked schools may have been sampled here, which correspondingly had lower incidences of treatment, irrespective of their prevalence levels. Thirdly, other factors may be involved which were not considered in this survey, such as differences in socio-economic status, health education, local access to health centres and water contact behaviours (Kloos, 1995, Akogun and Akogun, 1996, Raso *et al.*, 2005). One factor which was investigated and proved significant in some cases was place of birth, which should be explored further, perhaps with a targeted survey focusing on the Eastern region and aiming to compare schistosomiasis prevalence and intensity in local schoolchildren with migrants. The finding that

recent migrants are less likely to have received treatment immediately suggests a potentially improvement to the treatment methodology; in areas of high community motility, schools could be encouraged to treat all new registrants upon starting at that school, to reduce the chance of missing treatment if that child is going to move again. This might also benefit children who register at a school, but subsequently fail to attend regularly. Finally, the local geographical heterogeneities seen in the data suggest that district-wide or even sub-county level analyses are too broad for effective and efficient targeting of schistosomiasis control in schoolchildren, and thus more detailed observation methods are required for fine-scale reporting and monitoring. Rapid assessment protocols, such as that used in this study, are effective for covering numbers of schools in a short amount of time, and using limited resources, and thus are ideally suited for this level of coverage (Brooker and Clements, 2009).

3.6.5 Spatial analysis

The spatial analysis of our data showed two main spatial trends; first of all, regionally there was a significant difference in the prevalence of *S. mansoni* between the Eastern and Western regions of the lakeshore. The exact reasons for this cline remain unclear, although the high prevalence and intensity of infection found in the Eastern region corroborates with previous schistosomiasis surveys, such as those undertaken by Stothard et al. (2005) at the inception of the national control programme. Secondly, locally there appeared to be evidence for spatial autocorrelation at roughly a 75-85 km scale. This was an unexpected finding, as although climatic conditions are influential at that kind of scale, microdistribution of suitable snail habitats would suggest that autocorrelation, if present at all, would operate on a much smaller spatial scale. Previous work on the spatial distribution of *S. mansoni* has shown limits of autocorrelation that are much lower, which perhaps reflect snail micro-distribution, or perhaps socio-economic factors that also operate on a small scale (Xu *et al.*, 2004, Raso *et al.*, 2005). The error margins on the semivariograms should also be considered; a larger sample size of schools would be recommended in future surveys to add resolution to these data.

3.7 Conclusions

In summary, it has been shown that this rapid epidemiological assessment has been effective in mapping variation in the distribution of *S. mansoni* and soil-transmitted helminths and revealed some potentially unique demographic trends at a local level. However, inter-site heterogeneity demonstrated that it may not be viable to extrapolate local findings to the district level. Similarly, distributions of schistosomiasis and soil-transmitted helminthiasis may not follow the same patterns for distributions, complicating efforts at delivering efficient co-ordinated treatment for both infections simultaneously; this has implications for the way in which national treatment programmes should target risk areas for future anthelminthic drug administration. Furthermore, high levels of movement of some lakeshore communities may confound attempts at complete treatment coverage. Further study, explicitly considering migration as a risk factor and possibly comparing its value as a predictor for risk across different regions of East Africa, should be a priority on the research agenda. Of course, a key factor when considering the distribution of *S. mansoni* is the distribution of its intermediate host snails; these patterns, and their implications on the transmission of *S. mansoni*, will be explored in later chapters.

4 Rapid assessment of intestinal schistosomiasis in Lake Victoria using multiple diagnostics: Comparison between high and low endemicity settings

4.1 Abstract

Rapid diagnostic tests (RDTs) are becoming increasingly more popular for mapping neglected tropical diseases (NTDs) and for schistosomiasis, the circulating cathodic antigen (CCA) urine-dipstick has proven useful for detection of *S. mansoni* infections. Before being more widely adopted, the CCA dipstick requires validation, particularly in remote and under-surveyed disease endemic regions, and also between areas of different transmission levels. Here, we compared the CCA test against a 'gold-standard' of double Kato-Katz faecal smears, in a pilot expedition followed by two more comprehensive surveys of school-age children, all from under-reported areas of Lake Victoria. The pilot test of the CCA dipstick surveyed 63 children at 11 sites along the southern shoreline of Lake Victoria, in Tanzania. After this field validation, the main comparative studies were undertaken: the first comprising of 171 children from 11 schools along the eastern lakeshore of Tanzania and Kenya and the second a survey of 950 children from 61 schools on the Ssesse Islands in Uganda, with 45 of these schools also surveyed with CCA. For further arbitration, SEA-ELISA on fingerprick blood was used as an additional diagnostic at 5 schools in the latter survey. As expected, prevalence by both Kato-Katz microscopy and CCA urine dipstick were higher in the first survey in the suspected high-transmission zone than in the second, low-transmission setting (Kato-Katz: 68.6% versus 31.4%; CCA: 71.3% versus 46.8%), but diagnostic scores between the two tests also varied between the study sites. Difficulties in 'calling' precisely the CCA test result were noted in both surveys, especially in discrimination of borderline trace/positive reactions. Upon balance, we conclude the CCA tests hold promise for field diagnosis and rapid mapping of intestinal schistosomiasis although careful consideration of their cost-effectiveness and further validation in low-endemicity environments is needed for optimal use.

4.2 Contributions of the author

This chapter combines data collected from the Tanzania 2008 field mission, the Tanzania/Kenya 2009 field mission and the Uganda 2010 field mission. The author was present for the entire duration of all three field missions, and in the case of Tanzania 2008 and Tanzania/Kenya 2009, visited every school and site. During the Ssesse Island surveys during the Uganda 2010 mission, three teams were sent to survey the archipelago, given the vast number of islands needed to be visited. The author was part of one of the teams (together with Dr Russell Stothard, Moses Adriko and Moses Arinaitwe) which surveyed Mazinga sub-county, the most remote portion of the island group in the far south of the archipelago. The other teams also collected parasitological data, which were shared with the author. Across all the field missions, as for Uganda 2008, duties were shared evenly between team members, with the author leading the parasitological surveys and administrating the questionnaires with assistance in translation from one of the technicians. Slides were created primarily by the technicians and read by the author as well as the technicians. All data handling and statistical analyses were done by the author.

4.3 Introduction

Traditional diagnostics used for assessment of intestinal schistosomiasis prevalence have relied on stool microscopy, and most commonly, Kato-Katz smears. However, slide preparation is time-consuming, egg identification requires training and it can be difficult to obtain stool samples on demand if surveys are time-limited. While these and other factors, such as the insensitivity of single stool samples to detect low egg count infections, are known shortcomings of Kato-Katz methodology (Kongs *et al.*, 2001), it has long been recommended by WHO for general disease surveillance largely owing to lack of any other pragmatic alternatives (WHO, 2002). Recently, there has been growing interest in the use of new rapid diagnostic tests (RDTs) for on-going monitoring and mapping purposes as potential alternative methodologies (Stothard, 2009).

One such RDT is the circulating cathodic antigen (CCA) urine dipstick test (Rapid Medical Diagnostics, Pretoria, South Africa), as described in section 2.2.2. Preliminary studies showed low sensitivity to certain species/populations of *Schistosoma*, and *S. haematobium* in particular, suggesting that validation in other regions should be carried out as well (Stothard *et al.*, 2009b). Specifically, the diagnostic capability of the CCA test has never been directly compared between regions thought to have different base levels of prevalence, for example an area of high endemicity versus one with generally lower transmission and prevalence.

This chapter describes the results of three separate surveys to Lake Victoria shoreline. The first was a pilot study in Tanzania, where the first formulation CCA urine dipsticks were tested. This was followed by two more comprehensive surveys, one in Tanzania/Kenya (a high prevalence setting) and the second in the Ssesse Islands in Uganda, long considered to have a relatively low prevalence of intestinal schistosomiasis. In both these surveys, the second formulation of the CCA urine flow tests were compared with double Kato-Katz smears from a single stool and, for the Ssesse Island surveys, were further corroborated against SEA-ELISA serological testing. The aims were to evaluate the performance of the CCA tests in the field, both in terms of diagnostic capability but also ease of use. In addition, using the Ssesse Islands as an example, the prevalence results were compared against the Kato-Katz and SEA-ELISA findings (where available), to determine whether treatment recommendations at any particular location would change through use of different diagnostics. Treatment recommendations are based on the following prevalence categories: when prevalence is <10%, no mass drug administration is recommended; a one-off treatment is recommended where prevalence is between 10% and 50%; treatment every year is recommended when prevalence exceeds 50% (WHO, 2002). As such, even small variations in diagnostic power can have a large impact on the treatment delivered to a particular locality.

4.4 Methods

The CCA urine dipstick, in its first colloidal carbon formulation, was tested during a small pilot survey in June 2008 along the southern shoreline of Lake Victoria, in Tanzania. 63 children, in 11 locations, participated in the survey, which aimed to test the viability of using the dipstick in the field (and was also part of a larger test of the CCA urine dipstick for cases of *Schistosoma haematobium* infection: see Stothard *et al.*, 2009b). Parasitological surveys were carried out in schools as per section 2.1.1 of the Methods chapter but aiming for only 5 and 10 children per school. The first comparative survey was carried out during the Tanzania/Kenya 2009 field mission in January/February, comprising of 171 children from 11 schools. The second comparative survey was of 950 children from 61 schools on the Ssesse Islands in Uganda, during the Uganda 2010 field mission; of these 45 were also surveyed with CCA urine dipsticks. In all surveys, children were asked to participate in a demographical questionnaire, and methods of child selection and sample collection followed that as described in section 2.1.1, with information on consent and ethical matters described in section 2.1.3.

In the pilot as well as both subsequent surveys, single urine and stool samples were requested from each participant. Stools were used to create double smear Kato-Katz microscope slides (see section 2.2.1) and the urine was used both for the CCA tests and to check for micro- and macrohaematuria, using visual inspection and Hemastix© (see section 2.2.2). In the latter comparative survey, on the Ssesse Islands, fingerprick blood was also collected by a community nurse for use in SEA-ELISA serological tests for *Schistosoma* infection at five schools. The fingerprick blood was also used to check haemoglobin levels and prevalence of malaria, the latter using an RDT called Paracheck © (see also section 2.2.2).

For the second formulation of the CCA urine dipstick, the test result was semi-quantitatively called for intensity on the basis of relative band strength. Kato-Katz slides were read in the field as well as cross-checked at the NHM by CJS and JRS. SEA-ELISA scores were also semi-quantitatively called by comparing the test strength to the control dilution and cross-checked

with a second observer. Prevalences per site were calculated for all three diagnostics using R (Ihaka and Gentleman, 1996).

4.5 Results

Across all three field missions, only two children (one in Tanzania and one in Kenya) tested positive for *S. haematobium*; one had severe macrohaematuria (extremely bright red urine) and the other microhaematuria, detected using Hemastix. Both also tested positive for *S. mansoni*, through eggs in stool, but in order not to confound the CCA diagnostic results for *S. mansoni*, these individuals were excluded from later analyses.

The CCA urine dipstick proved workable in the field on the pilot study. Prevalence was generally very high by both diagnostics, though the small sample sizes should be taken into consideration at individual sites (see Table 4.1).

Village (*GPS coordinates, dec.* °)	Kato-Katz examination			CCA urine dipsticks		
	N	Prevalence (95% CI)	Intensity L / M / H	N	Trace as pos. (95% CI)	Trace as neg. (95% CI)
Nyanghomango (*S2.64253;E32.95962*)	3	100.00 (29.24-100.00)	2 / 1 / 0	3	66.67 (9.43-99.16)	66.67 (9.43-99.16)
SembaA (*S2.41402;E32.94063*)	5	80.00 (28.36-99.49)	2 / 2 / 0	5	60.00 (14.66-94.73)	20.00 (5.05-71.64)
Nyashimba (*S2.45220;E33.51693*)	NA	NA	NA	5	100.00 (47.82-100.00)	100.00 (47.82-100.00)
Mwabarugu (*S2.26215;E33.80880*)	14	79.00 (49.20-95.34)	5 / 6 / 0	14	85.71 (57.19-98.22)	85.71 (57.19-98.22)
Nyehunge (*S2.53780;E32.23285*)	5	20.00 (5.05-71.64)	1 / 0 / 0	5	20.00 (5.05-71.64)	20.00 (5.05-71.64)
Nkome (*S2.49390;E31.98610*)	5	100.00 (47.82-100.00)	0 / 1 / 4	5	100.00 (47.82-100.00)	100.00 (47.82-100.00)
Kahunda (*S2.40525;E32.05935*)	5	60.00 (14.66-94.73)	0 / 3 / 0	5	80.00 (28.36-99.49)	80.00 (28.36-99.49)
Maisome (*S2.34837;E32.04040*)	5	100.00 (47.82-100.00)	2 / 2 / 1	5	100.00 (47.82-100.00)	80.00 (28.36-99.49)
IzumacheliA (*S2.38317;E31.96692*)	5	100.00 (47.82-100.00)	3 / 1 / 1	5	100.00 (47.82-100.00)	80.00 (28.36-99.49)
IzumacheliB (*S2.40662;E31.94512*)	5	100.00 (47.82-100.00)	1 / 2 / 2	5	100.00 (47.82-100.00)	60.00 (14.66-94.73)
Nyakaliro (*S2.43602;E32.41142*)	3	100.00 (29.24-100.00)	2 / 0 / 1	NA	NA	NA
TOTAL	**55**	**81.82 (69.10-90.92)**	**18 / 18 / 9**	**57**	**82.46 (70.09-91.25)**	**70.18 (56.60-81.57)**

Table 4.1 – Results of pilot study of CCA urine dipstick compared to Kato-Katz microscopy
'N' is the number of school-aged children surveyed at each location. '95% CI' stands for 95% confidence intervals. Intensity classes are as per WHO (2002) recommendations. 'dec. °' stands for decimal degrees.

Overall prevalence based on Kato-Katz microscopy was 81.82% (95% CI = 69.10-90.92%, N = 55), whereas for the CCA urine dipsticks it was 82.46% (95% CI = 70.09-91.25%, N = 57). However, difficulties were noted with calling 'trace' dipstick results, whereby a very weak band was seen. The manufacturer's instructions suggested these would be positive test results, but prevalence was also calculated if 'trace' results actually represented a negative test. In this case, prevalence would be 70.18% (95% CI = 56.60-81.57%, N=57). Intensities of infection varied with site, but were divided across the whole sample as 40% light infections (<100 EPG), 40% moderate infections (100 – 400 EPG) and 20% heavy infections (>400 EPG).

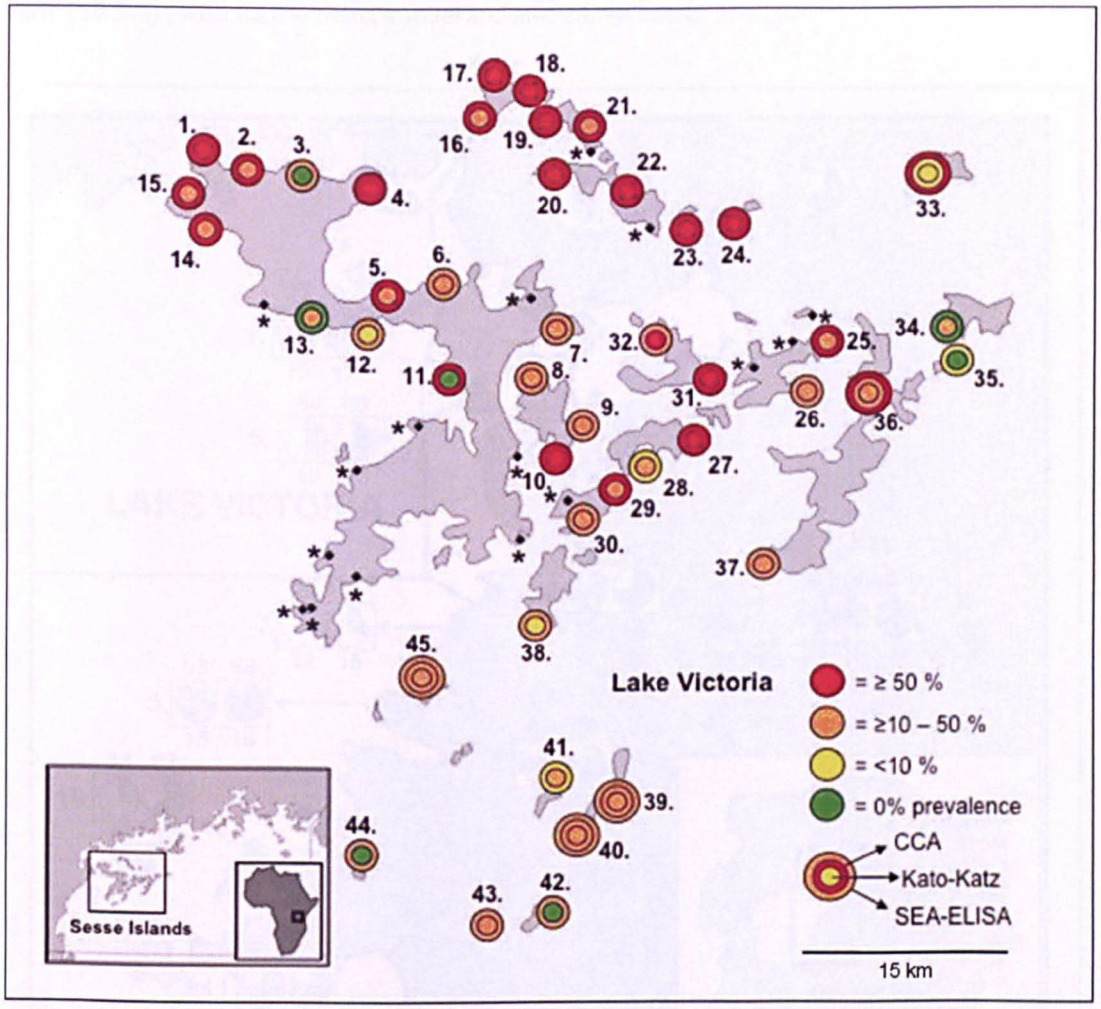

Figure 4.1- Map of the schools surveyed for intestinal schistosomiasis in the Ssesse Islands using CCA tests

Colour of the concentric circles indicates prevalence of schistosomiasis infection as diagnosed by CCA urine tests, comparing when 'trace' results were considered positive against when 'trace' calls were considered indicative of no infection.

In the comparative surveys, prevalence varied within and between the two regions and the diagnostic tests performed very differently as well (Tables 4.2 and 4.3; Figures 4.1 and 4.2). Overall prevalence of infection for the Tanzania/Kenya survey was 68.6% (95% CIs = 60.7-75.7, N [number of children surveyed] = 159) as determined by Kato-Katz double smear slide reads. As before, prevalence by CCA was calculated both for 'trace' test results indicating a positive and a negative. For the former, prevalence was calculated to be 94.2% (95% CIs = 89.5-97.2%, N=171) and when trace was negative prevalence was 71.4% (95% CIs = 63.9-78.0, N = 171). The majority of infection intensities, as classified by WHO guidelines (WHO, 2002), were 'light' (50.3%) , with 25.8% being 'moderate' and 23.9% being 'heavy'.

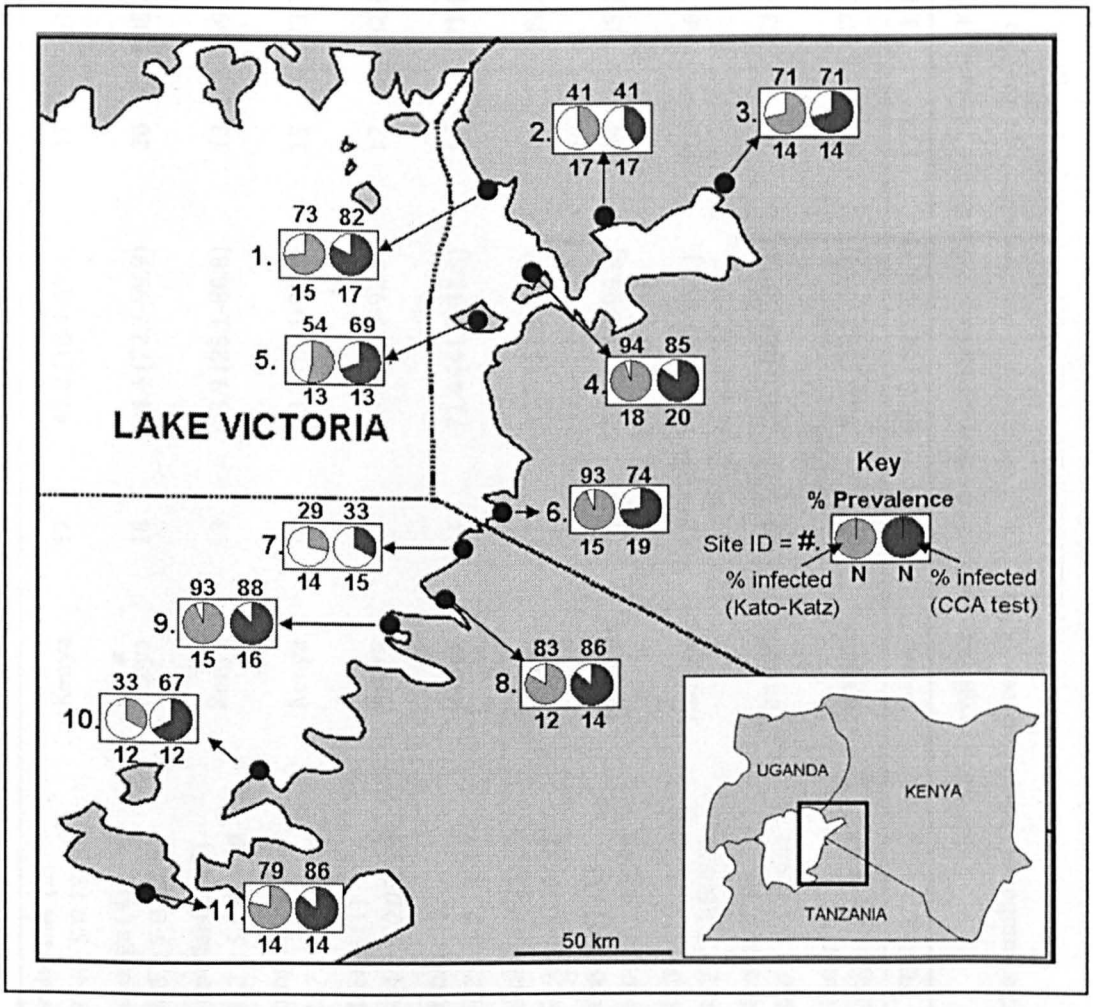

Figure 4.2 – Map of the schools surveyed for intestinal schistosomiasis in Tanzania/Kenya in 2009

Pie charts indicate prevalence of schistosomiasis infection as diagnosed by Kato-Katz double smear slides or urine-CCA tests; prevalence value is indicated by the number above each pie chart. The 'N' value below each pie chart indicates the surveyed sample size.

School (Map ID) *GPS coordinates (dec. °)*	Country	*Kato-Katz faecal slides*		*CCA tests*		
		N	Prevalence % (95% CI)	N	Prevalence % (95% CI) 'trace' as negative	Prevalence % (95% CI) 'trace' as positive
Asembo Bay (2) *E34.38747; S-0.18737*	Kenya	17	41.2 (18.4-67.1)	17	41.2 (18.4-67.1)	82.3 (56.6-96.2)
Kolunga (4) *E34.13857; S-0.42781*	Kenya	18	94.4 (72.7-99.9)	20	85.0 (57.2-98.2)	95.0 (75.1-99.9)
Mfangano Island (5) *E34.04121; S-0.48410*	Kenya	13	53.9 (25.1-80.8)	13	69.2 (38.6-90.9)	100.0 (75.3-100.0)
Shining Star (6) *E34.12677; S-1.01244*	Kenya	15	93.3 (68.1-99.8)	19	73.7 (48.8-90.9)	94.7 (74.0-99.9)
Usengi (1) *E34.05849; S-0.07313*	Kenya	15	73.3 (44.9-92.2)	17	82.4 (56.6-96.2)	100.0 (80.5-100)
Usoma (3) *E34.71829; S-0.10519*	Kenya	14	71.4 (41.9-91.6)	14	71.4 (41.9-91.6)	100.0 (76.8-100)
Busanga (8) *E33.95499; S-1.30631*	Tanzania	12	83.3 (51.6-97.9)	14	85.7 (57.2-98.2)	100.0 (76.8-100)
Hamuyebe (11) *E33.06409; S-2.12397*	Tanzania	14	78.6 (49.2-95.3)	14	85.7 (57.2-98.2)	85.7 (57.2-98.2)
Majita (10) *E33.40522; S-1.80939*	Tanzania	12	33.3 (9.9-65.1)	12	66.7 (34.9-90.1)	100.0 (73.5-100)
Mkoma (7) *E34.03001; S-1.14983*	Tanzania	14	28.6 (8.3-58.1)	15	33.3 (11.8-61.6)	80.0 (51.9-95.7)
Ruhu (9) *E33.82505; -1.35198*	Tanzania	15	93.3 (68.1-99.8)	16	87.5 (61.7-98.4)	100.0 (79.4-100)
TOTAL	------	**159**	**68.6 (60.7-75.7)**	**171**	**71.4 (63.9-78.0)**	**94.2 (89.5-97.2)**

Table 4.2 – Prevalence of intestinal schistosomiasis per site for the Tanzania/Kenya 2009 survey

'N' refers to the number of children surveyed, and '95% CI' refer to 95% confidence intervals. The Map ID can be cross-referenced with Figure 4.1 to see the location of each site. 'dec. °' stands for decimal degrees.

Map ID (GPS coordinates, dec.°)	Village	Subcounty	Prevalence by Kato-Katz (N, 95% CI)	Prevalence by CCA 'Trace' = pos. (N, 95% CI)	Prevalence by CCA 'Trace' = neg (N, 95% CI)
1 (S0.22048, E32.07748)	Kibanga	Mugoye	68.8 (16, 41.3-89.0)	100.0 (16, 79.4-100.0)	68.8 (16, 41.3-89.0)
2 (S0.22820, E32.10802)	Buziga	Mugoye	46.2 (13, 19.2-74.9)	56.3 (16, 29.9-80.2)	25.0 (16, 7.3-52.4)
3 (S0.22986, E32.14310)	Bungo	Mugoye	0.0 (12, 0.0-26.5)	15.4 (13, 1.9-45.4)	0.0 (13, 0.0-24.7)
4 (S0.23807, E32.18468)	Kasekulo	Mugoye	50.0 (16, 24.7-75.3)	62.5 (16, 35.4-84.8)	43.8 (16, 19.8-70.1)
5 (S0.31496, E32.19963)	Kagolomola	Mugoye	14.3 (14, 1.8-42.8)	100.0 (14, 76.8-100.0)	50.0 (14, 23.0-77.0)
6 (S0.30533, E32.23756)	Kasenyi	Mugoye	16.7 (12, 2.1-48.4)	13.3 (15, 1.7-40.5)	6.7 (15, 0.2-31.9)
7 (S.32837, E32.31014)	Mwena	Mugoye	21.4 (14, 4.7-50.8)	28.6 (14, 8.4-58.1)	7.1 (14, 0.2-33.9)
8 (S0.36381, E32.29543)	Kaagonya	Mugoye	25.0 (8, 3.2-65.1)	12.5 (8, 0.3-52.7)	12.5 (8, 0.3-52.7)
9 (S0.39600, E32.32784)	Kisujju	Mugoye	28.6 (14, 8.4-58.1)	46.2 (13, 19.2-74.9)	46.2 (13, 19.2-74.9)
10 (S0.41439, E32.30922)	Kivunza	Mugoye	77.8 (9, 40.0-97.2)	66.7 (9, 30.0-92.5)	66.7 (9, 30.0-92.5)
11 (S0.36485, E32.23876)	Njoga	Mugoye	0.0 (14, 0.0-23.2)	100.0 (15, 78.2-100.0)	0.0 (15, 0.0-21.8)
12 (S0.33256, E32.18855)	Banga	Mugoye	6.7 (15, 0.2-31.9)	11.8 (17, 1.5-36.4)	11.8 (17, 1.5-36.4)
13 (S0.32341, E32.14947)	Mutambala	Mugoye	10.0 (10, 0.3-44.5)	0.0 (14, 0.0-23.2)	0.0 (14, 0.0-23.2)
14 (S0.26678, E32.07942)	Banda1	Mugoye	28.6 (14, 8.4-58.1)	64.3 (14, 35.1-87.2)	42.9 (14, 17.7-71.1)
15 (S0.24923, E32.06704)	Luku	Mugoye	47.1 (17, 23.0-72.2)	58.8 (17, 32.9-81.6)	35.3 (17, 14.2-61.7)
16 (S0.19225, E32.25963)	Luwungulu	Bufumira	30.8 (13, 9.1-61.4)	50.0 (14, 23.0-77.0)	21.4 (14, 4.7-50.8)
17 (S0.16808, E32.27124)	Kammesse	Bufumira	73.3 (15, 44.9-92.2)	93.3 (15, 68.1-99.8)	66.7 (15, 38.4-88.2)
18 (S0.17993, E32.29384)	Kachanga	Bufumira	75.0 (16, 47.6-92.7)	100.0 (15, 78.2-100.0)	73.3 (15, 44.9-92.2)
19 (S0.18954, E32.30225)	Kaaya	Bufumira	66.7 (18, 41.0-86.7)	100.0 (15, 78.2-100.0)	60.0 (15, 32.3-83.7)
20 (S0.22574, E32.30759)	Kibibi	Bufumira	57.1 (7, 18.4-90.1)	57.1 (7, 18.4-90.1)	42.9 (7, 9.9-81.6)
21 (S0.19813, E32.33113)	Misonzi	Bufumira	46.9 (32, 29.1-65.3)	73.3 (15, 44.9-92.2)	53.3 (15, 26.6-78.7)
22 (S0.24036, E32.35637)	Bosa	Bufumira	52.9 (17, 27.8-77.0)	71.4 (14, 41.9-91.6)	35.7 (14, 12.8-64.9)
23 (S0.25941, E32.39413)	Banda	Bufumira	73.3 (15, 44.9-92.2)	93.3 (15, 68.1-99.8)	86.7 (15, 59.5-98.3)
24 (S0.26246, E32.42848)	Kitobo	Bufumira	69.2 (26, 48.2-85.7)	80.0 (15, 51.9-95.7)	60.0 (15, 32.3-83.7)
25 (S0.34038, E32.48982)	Namasengere	Bufumira	37.5 (16, 15.2-64.6)	92.9 (14, 66.1-99.8)	50.0 (14, 23.0-77.0)
26 (S0.37258, 32.47824)	Mukaka	Bufumira	46.7 (15, 21.3-73.4)	33.3 (15, 11.8-61.6)	6.7 (15, 0.2-31.9)
27 (S0.40281, E32.40434)	Kaazi	Bufumira	86.7 (15, 59.5-98.3)	80.0 (15, 51.9-95.7)	53.3 (15, 26.6-78.7)

Continued on the following page

Map ID *(GPS coordinates)*	Village	Subcounty	Prevalence by Kato-Katz (N, 95% CI)	CCA Prevalence 'Trace' = pos. (N, 95% CI)	CCA Prevalence 'Trace' = neg (N, 95% CI)
28 *(S0.42411, E32.36667)*	Kyankolokolo	Bufumira	20.0 (15, 4.3-48.1)	6.7 (15, 0.2-31.9)	6.7 (15, 0.2-31.9)
29 *(S0.43715, E32.35111)*	Lwabaswa	Mugoye	18.2 (11, 2.3-51.8)	84.6 (13, 54.6-98.1)	76.9 (13, 46.2-95.0)
30 *(S0.45702, E32.32824)*	Kasisa	Mugoye	25.0 (12, 5.5-57.2)	45.5 (11, 16.7-76.6)	36.4 (11, 10.9-69.2)
31 *(S0.36561, E32.41432)*	Lulindi	Bufumira	55.6 (18, 30.8-78.5)	66.7 (12, 34.9-90.1)	50.0 (12, 21.1-78.9)
32 *(S0.33807, E32.37603)*	Ssemawundo	Bufumira	59.4 (32, 40.6-76.3)	46.7 (15, 21.3-73.4)	40.0 (15, 16.3-67.7)
33 *(S0.22757, 32.55733)*	Jaana	Jaana	5.0 (20, 0.1-24.9)	8.7 (23, 1.1-28.0)	4.3 (23, 1.1-21.9)
34 *(S0.33335, E32.56855)*	Bubeke	Jaana	23.1 (13, 5.0-53.8)	0.0 (16, 0.0-20.6)	0.0 (16, 0.0-20.6)
35 *(S0.35153, E32.57128)*	Buyange	Kyamuswa	0.0 (17, 0.0-19.5)	6.3 (16, 0.2-30.2)	0.0 (16, 0.0-20.6)
36 *(S0.37463, E32.51683)*	Lwenabatwa	Kyamuswa	10.0 (10, 0.3-44.5)	40.1 (23, 19.7-61.5)	13.0 (23, 2.8-33.6)
37 *(S0.48870, E32.44622)*	Nakibanga	Kyamuswa	26.7 (15, 7.8-55.1)	13.3 (15, 1.7-40.5)	13.3 (15, 1.7-40.5)
38 *(S0.52895, E32.29731)*	Mawala	Mazinga	7.7 (13, 0.2-36.0)	20.0 (15, 4.3-48.1)	20.0 (15, 4.3-48.1)
39 *(S0.64666, E32.35232)*	Kuuye	Mazinga	20.0 (20, 5.7-43.7)	25.0 (20, 8.7-49.1)	15.0 (20, 3.2-37.9)
40 *(S0.67167, E32.32633)*	Kachungwa	Mazinga	10.0 (20, 1.2-31.7)	13.0 (23, 2.8-33.6)	4.3 (23, 0.1-21.9)
41 *(S0.63200, E32.31278)*	Kirugu	Mazinga	14.3 (14, 1.8-42.8)	5.9 (17, 0.1-28.7)	5.9 (17, 0.1-28.7)
42 *(S0.71890, E32.30853)*	Nkose	Mazinga	0.0 (13, 0.0-24.7)	15.4 (13, 1.9-45.4)	7.7 (13, 0.2-36.0)
43 *(S0.73168, E32.26802)*	Miyana	Mazinga	12.5 (8, 0.3-52.7)	28.6 (7, 3.7-71.0)	0.0 (7, 0.0-41.0)
44 *(S0.68592, E32.18170)*	Lujjabwa	Mazinga	0.0 (14, 0.0-23.2)	35.3 (17, 14.2-61.7)	0.0 (17, 0.0-19.5)
45 *(S0.56909, E32.22253)*	Butulume	Mazinga	13.3 (15, 1.7-40.5)	26.7 (15, 7.8-55.1)	0.0 (15, 0.0-21.8)
TOTAL	--	--	**34.6 (683, 31.0-38.3)**	**46.5 (666, 42.7-50.4)**	**28.1 (666, 24.7-31.7)**

Table 4.3 - Prevalence of intestinal schistosomiasis per site for the Ssesse Island 2010 survey

'N' refers to the number of children surveyed, and '95% CI' refer to 95% confidence intervals. The Map ID can be cross-referenced with Figure 4.2 to see the location of each site. 'dec. °' stands for decimal degrees.

For the Ssesse Island survey, 61 sites were visited in total, and 45 were surveyed using both Kato-Katz and CCA urine lateral flow tests (Figure 4.1). Five of these were further tested with SEA-ELISA kits. Overall prevalence of intestinal schistosomiasis for the 61 sites, as observed via positive Kato-Katz smears, was 31.4% (95% CIs = 28.4-34.5%; N = 905). When considering only the 45 sites that were also included in the CCA test surveys, the prevalence by Kato-Katz slide read was 34.6% (95% CIs = 31.0-38.3, N= 683). When measured by CCA urine lateral flow test, it was 46.5% (95% CIs = 42.7-50.4%; N = 666); this was when 'trace' calls were considered positive, as per the manufacturer's instructions. To compare with previous CCA efficacy studies, the prevalence was also calculated for when 'trace' calls were considered negative, and in this case, the value was 28.4% (95% CIs = 25.1-32.0%; N = 666). Figure 4.3 directly compares the prevalence at each site when 'trace' results are infection-positive as opposed to infection-negative. Prevalence ranged widely across the islands, with high prevalence (over 50%) localities predominant in the northern islands of Bufumira sub-county.

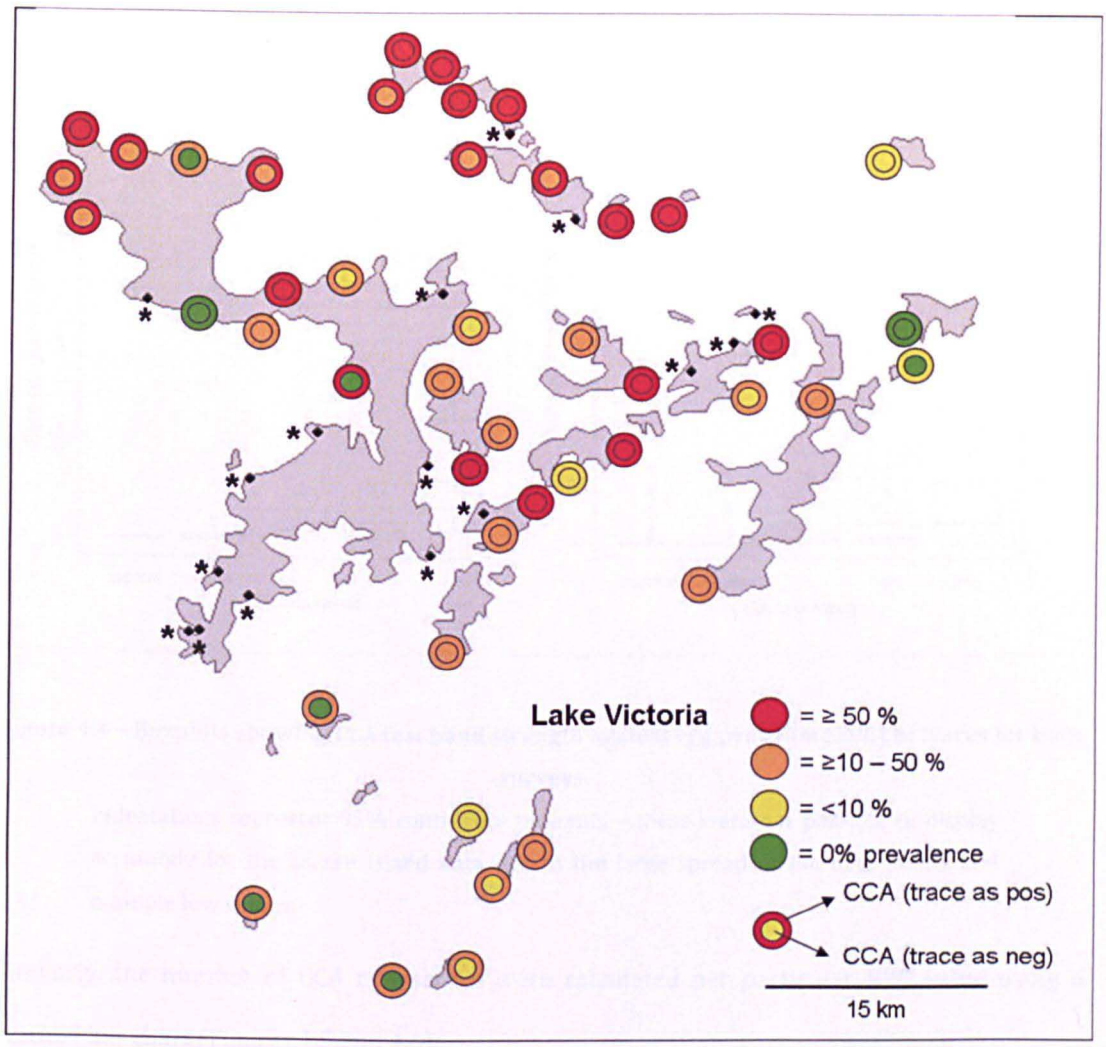

Figure 4.3 - Map of the schools surveyed for intestinal schistosomiasis in the Ssese Islands in 2010

Colour of the concentric circles indicates prevalence of schistosomiasis infection as diagnosed by urine-CCA tests, comparing 'trace' results considered positive with when considered negative. '*' indicates sites that were not surveyed using CCA urine dipsticks; data from these sites will not be presented here.

For the five villages surveyed with Kato-Katz, CCA and SEA-ELISA, overall prevalence by each of the three diagnostics was 11.7 % (95% CIs = 5.8-20.6%; N = 85), 22.1 (95% CIs = 14.6-31.3%; N = 104) and 39.2% (95% CIs = 29.7-49.4%; N = 102) respectively (Table 4.4). In all sites bar one, SEA-ELISA resulted in the highest prevalence. Prevalence varied widely between when traces were considered negative as opposed to positive, even at the same site. The relationship between Kato-Katz smear result and CCA was investigated further by directly comparing CCA test scores against EPG (Figure 4.4).

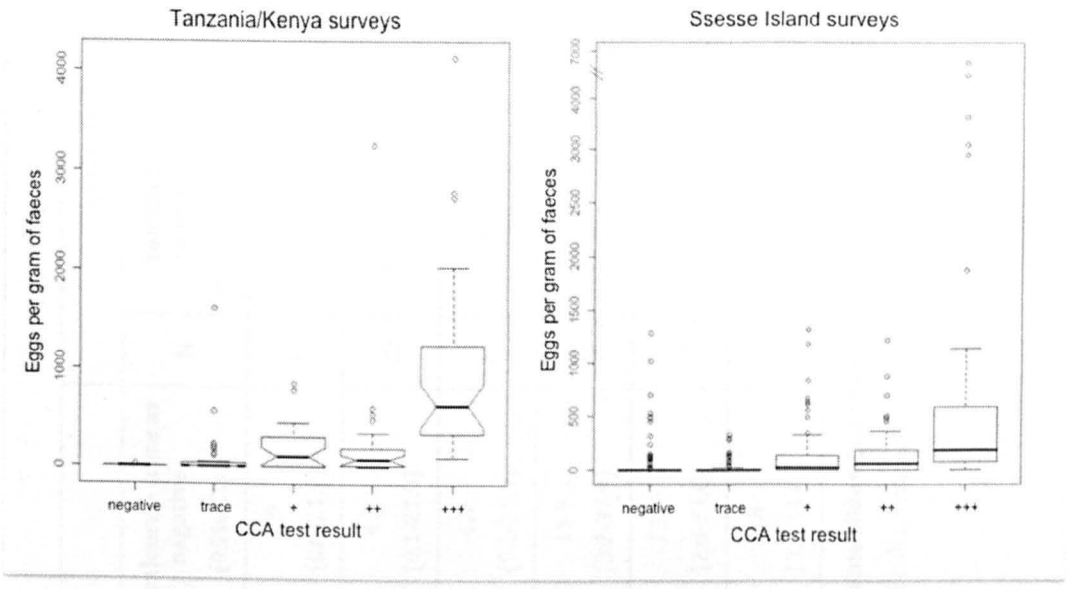

Figure 4.4 – Boxplots showing CCA test band strength against egg per gram (EPG) of faeces for both surveys

Indentations represent 95% confidence intervals – these were not possible to display accurately for the Ssesse Island data due to the large spread of the data points and multiple low values.

Similarly, the number of CCA test scores were calculated per particular EPG value using a stacked bar chart (Figures 4.5 and 4.6).

Site (Map ID)	DIAGNOSTIC TEST									
	Kato-Katz			CCA				SEA-ELISA		
	N	Infection intensity L / M / H	Prevalence (95% CIs)	N	Infection intensity tr / + / ++ / +++	Prevalence – trace as positive (95% CIs)	Prevalence – trace as negative (95% CIs)	N	Infection intensity tr / + / ++ / +++	Prevalence (95% CIs)
Butulume	15	2 / 0 / 0	13.3 (1.7-40.5)	15	4 / 0 / 0 / 0	26.7 (7.8-55.1)	0.0 (0.0-21.8)	15	3 / 0 / 0 / 0	20.0 (4.3-48.1)
Jaana	20	1 / 0 / 0	5.0 (0.1-24.9)	23	1 / 1 / 0 / 0	8.7 (1.1-28.0)	4.3 (0.1-21.9)	22	0 / 6 / 2 / 4	54.5 (32.2-75.6)
Kachungwa	20	2 / 0 / 0	10.0 (1.2-31.7)	23	2 / 1 / 0 / 0	13.0 (2.8-33.6)	4.3 (0.1-21.9)	23	2 / 1 / 4 / 2	39.1 (19.7-61.4)
Kuuye	20	3 / 1 / 0	20.0 (5.7-43.7)	20	2 / 2 / 1 / 0	25.0 (8.7-49.1)	15.0 (3.2-37.9)	20	1 / 1 / 3 / 1	30.0 (11.9-54.3)
Lwenabatwa	10	1 / 0 / 0	10.0 (0.3-44.5)	23	6 / 2 / 1 / 0	39.1 (19.7-61.5)	13.0 (2.8-33.6)	22	2 / 2 / 3 / 6	59.1 (36.4-79.3)
TOTAL	85	9 / 1 / 0	11.7 (5.8-20.6)	104	15 / 6 / 2 / 0	22.1 (14.6-31.3)	7.8 (3.3-14.6)	102	8 / 10 / 12 / 13	39.2 (29.7-49.4)

Table 4.4 – Prevalence of intestinal schistosomiasis in five villages in the Ssese Islands, surveyed using Kato-Katz slides, CCA urine dipstick tests and SEA-ELISA kits

'tr', '+', '++' and '+++' refer to trace, single, double and triple positive test calls, respectively

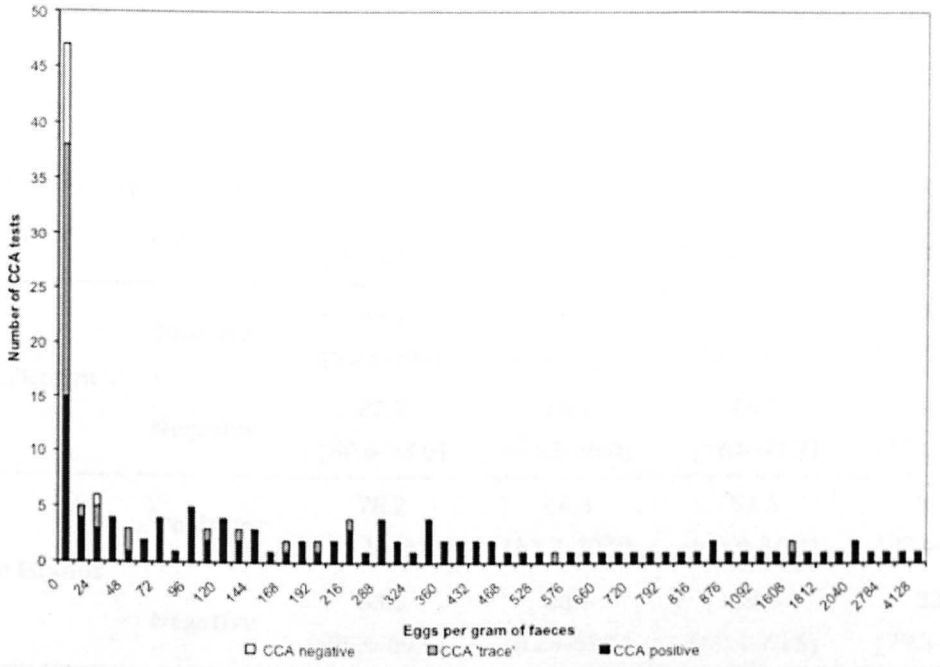

Figure 4.5 – Stacked bar chart showing number of CCA test results per EPG value for the Tanzania/Kenya 2009 survey

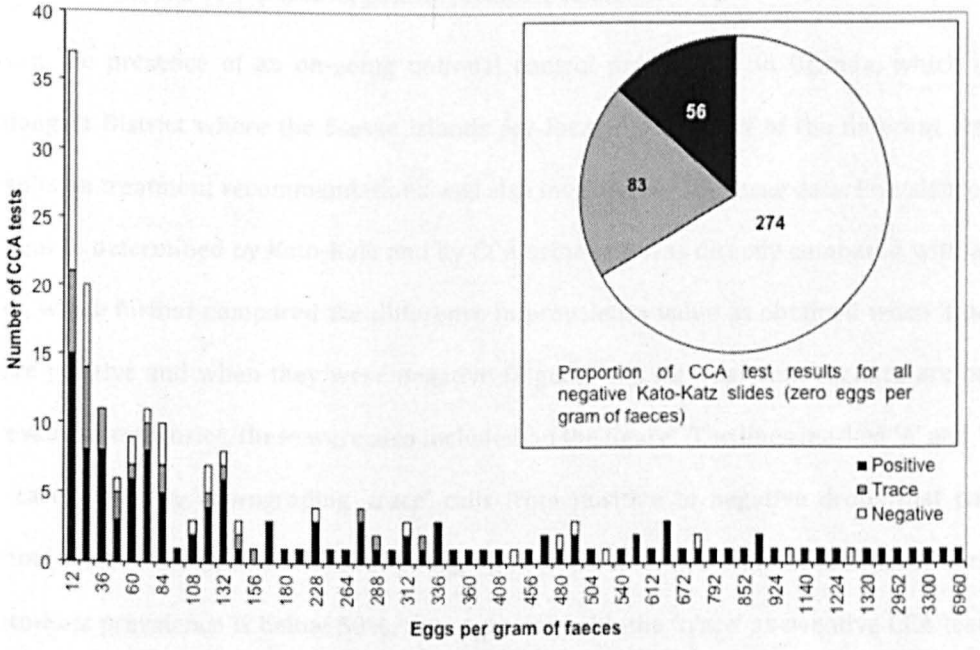

Figure 4.6 - Stacked bar chart showing number of CCA test results per EPG value for the Ssesse Island 2010 survey

The breakdown of number of CCA test results for when EPG was zero is given in the pie chart (inset).

Diagnostic scores were also calculated for both surveys, comparing the CCA urine dipstick against the 'gold standard' of a double smear Kato-Katz slide read. Scores were calculated both for when traces were considered positive as well as negative (Table 4.5).

Survey	Trace call	SS (95% CIs)	SP (95% CIs)	PPV (95% CIs)	NPV (95% CIs)
Tanz./Kenya	Positive	99.1 (95.6-100)	19.1 (9.8-31.7)	73.4 (65.8-80.2)	90.0 (63.2-99.4)
	Negative	87.7 (80.6-93.0)	68.1 (54.3-80.0)	86.1 (78.8-91.7)	71.1 (57.2-82.8)
Ssesse Islands	Positive	78.2 (72.1-83.5)	66.3 (61.7-70.8)	52.6 (46.9-58.2)	86.4 (82.4-89.9)
	Negative	63.5 (56.6-69.9)	86.4 (82.9-89.5)	69.1 (62.1-75.5)	83.2 (79.5-86.5)

Table 4.5 – Diagnostic scores comparing the CCA urine dipstick against a double Kato-Katz slide read

'SS' stands for sensitivity, 'SP' stands for specificity, 'NPV' stands for negative predictive value and 'PPV' stands for positive predictive value.

Given the presence of an on-going national control programme in Uganda, which includes Kalangala District where the Ssesse Islands are located, the effect of the different diagnostic results on treatment recommendations was also investigated for these data. Prevalence at each school as determined by Kato-Katz and by CCA urine test was directly compared with a scatter plot, which further compared the difference in prevalence value as obtained when 'trace' calls were positive and when they were negative (Figure 4.7). As treatment regimes are based on prevalence categories, these were also included on the figure. The lines marked 'A' and 'B' refer to cases whereby downgrading 'trace' calls from positive to negative drops that particular school under 50% prevalence, thus changing its recommended treatment regime. For 'A', the Kato-Katz prevalence is below 50%, thus agreeing with the 'trace' as negative CCA test result; however, for 'B', Kato-Katz prevalence was above 50%. Table 4.6 shows a breakdown of the recommended treatment regimes for each school, based on these surveys, and further demonstrates the number of schools that would change treatment regimes either between the Kato-Katz and CCA results, or even between calling 'trace' results as positive and calling them as negative.

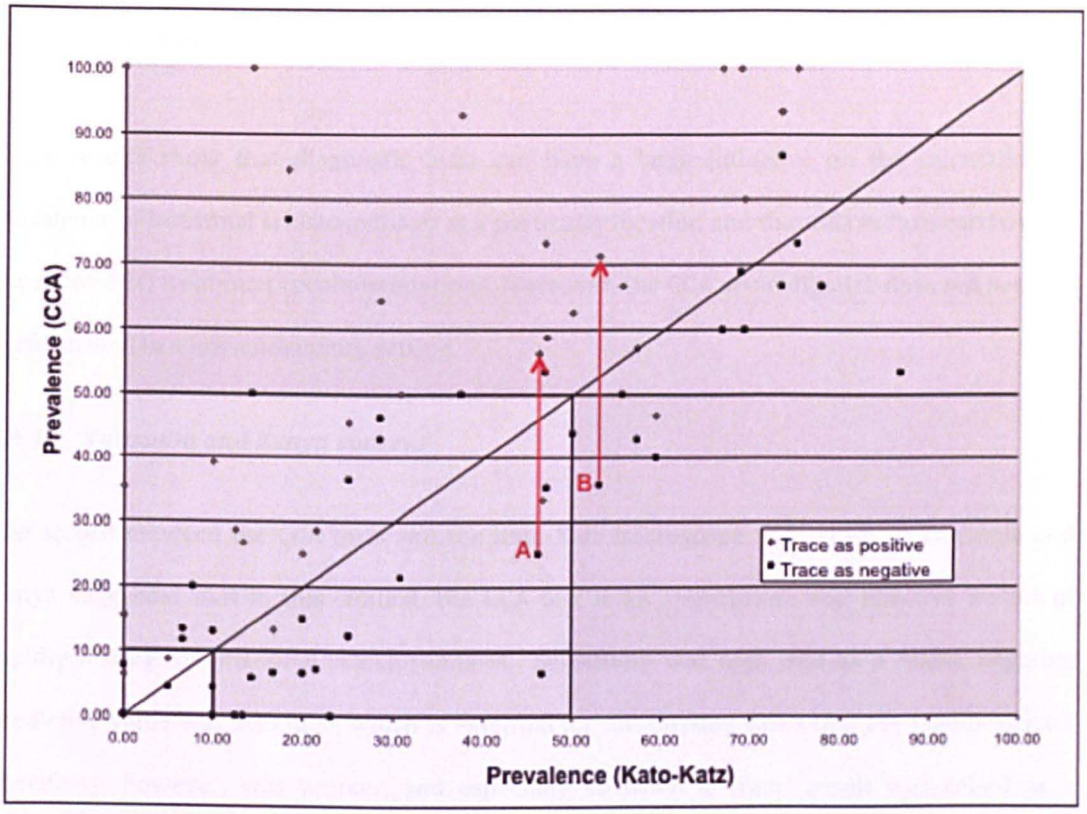

Figure 4.7 – Scatter plot comparing prevalence by Kato-Katz slide against prevalence by CCA, for the Ssese Island survey data

Treatment recommendations are based on WHO protocols (2002), whereby no treatment is recommended when prevalence is <10%, a one-off treatment is recommended where prevalence is between 10% and 50% and treatment every year is recommended when prevalence exceeds 50%.

Treatment classes	Number of schools per treatment class and per diagnostic test		
	Kato-Katz	CCA (trace = positive)	CCA (trace = negative)
No treatment (prevalence < 10%)	9	6	16
Single treatment (≥ 10 - 50%)	23	18	16
Annual treatment (≥ 50%)	13	21	13

Table 4.6 – Number of schools requiring each level of recommended treatment, based on the different diagnostic tests

Treatment recommendation categories are as described in Figure 4.7

4.6 Discussion

These results show that diagnostic tests can have a large influence on the calculation of prevalence of intestinal schistosomiasis at a particular location and that this in turn can have an impact on PZQ treatment recommendations. Moreover, the CCA urine dipstick does not seem to perform well in a low-endemicity setting.

4.6.1 Tanzania and Kenya surveys

The accord between the CCA tests and the Kato-Katz microscope slide reads in Tanzania and Kenya suggested that in this context, the CCA test is an appropriate and effective means of rapidly testing for intestinal schistosomiasis. Sensitivity was high, and as a result, negative predictive value was also high, which is essential for not missing cases that are true positives. Specificity, however, was weaker, and especially so when a 'trace' result was called as a positive; this could arguably be due to the Kato-Katz double smear being an inaccurate 'gold standard' against which to compare, despite WHO guidelines that it is sufficient for rapid testing. Previous surveys have shown that Kato-Katz sensitivity for a single stool sample can be low (Kongs *et al.*, 2001), and that it is only with three consecutive stools that false negatives are better avoided (Booth *et al.*, 2003). However, three daily stool samples is not feasible for rapid assessment protocols in terms of cost, time and survey team logistics, whereas the convenience afforded by a single urine sample (in itself easier to obtain than a faecal sample) is important when conducting rapid mapping or on-going monitoring.

It is worth emphasizing here that, regardless of diagnostic, the prevalence of intestinal schistosomiasis in this region of Lake Victoria was generally found to be very high. Much of lakeshore northern Tanzania has not, to the author's knowledge, been extensively formally surveyed (Clements *et al.*, 2006, Clements *et al.*, 2010), and so rapid assessment surveys are a crucial means of building information about the extent of the disease in this area. Such information is important to support the nascent national control programme that is active in other parts of the country. Indeed, in this survey, it was found that children in Tanzania were already significantly more likely to have received praziquantel than children in Kenya. This

reflects Kenya's lack of an active national control programme, and such surveys could be helpful in assisting and advocating for the development of such an initiative which is urgently required.

4.6.2 Ssesse Island surveys

From a mapping perspective, far from being a uniformly low transmission zone, as had been assumed in the past, the Ssesse Islands display significant heterogeneities in terms of the prevalence of *S. mansoni*, with obvious implications for on-going treatment campaigns. Moreover, the CCA urine lateral flow tests did not perform as well in this setting, in terms of diagnostic scores as compared to the Kato-Katz smears, as had been the case in the Tanzania and Kenya survey.

Across the district, a number of sites had significant differences in the prevalence as ascertained by one diagnostic or another. For example, Site 11 would have been considered high prevalence on the basis of CCA results, and yet not a single positive Kato-Katz slide was observed. Conversely, two sites would have been considered absent of *S. mansoni* had only CCA tests been used, and yet had Kato-Katz prevalence of 23.1% and 10.0% respectively.

The variability of the CCA test in this setting is the cause of some concern, and indeed the diagnostic scores are generally lower than given in other studies (Stothard *et al.*, 2006a, Legesse and Erko, 2007). Moreover, the results could be explained by the particular transmission dynamics of schistosomiasis in these islands. Though the data are not presented here, snail surveys were also undertaken in some of the parasitological survey locations in the Ssesse Islands, and few snails were found as compared to other parts of the Ugandan shoreline (see Chapter 6 for details). If the Ssesse Islands themselves are not a high transmission zone, children may only be exposed to the parasite infrequently or even only once, resulting in long-term infection from single or few exposure events. As these worms age, their fecundity could decrease, reducing egg output intensity; in this scenario, a CCA test would be positive whereas the Kato-Katz smear might not reveal eggs. Unfortunately, the high-prevalence islands in the north of the archipelago were not surveyed; this should be a priority for further surveys to

determine the true transmission status of these islands (a field mission in November 2010 has been planned to address this).

Another factor to consider is that the CCA test is thought to be sensitive to differences in antigenicity of the parasite (Stothard, 2009), which may be related to genetic diversity. This is important because current research on the molecular epidemiology of *S. mansoni* has revealed very high levels of genetic diversity within Lake Victoria (Stothard *et al.*, 2009c, Standley *et al.*, 2010); so it could be plausible that changes in genotype across the lakeshore may result in altered performance of the CCA test, based on changes to the structure or form of the antigens produced by the schistosomes. Similarly, if exposure is low, the diversity of the genotypes present within an individual child may also be lower than in other regions, which too could alter the way in which the test reacts.

The SEA-ELISA results potentially reveal a further complication; the effect of prior treatment. As the egg antibodies tested for in the ELISA may persist after the parasite is eliminated, for example post treatment with praziquantel, individuals may test seropositive despite not having an active infection. In the Ssesse Islands, while treatment is far from ubiquitous, the district has been involved in the national control programme since 2005, and mass drug administration has been underway since then, particularly on the main island. In this survey, 49 out of 226 children (21.7%) reported having received praziquantel at some point. As such, whereas SEA-ELISA testing is an extremely sensitive and effective way of testing baseline prevalence in an un-surveyed area, and is particularly useful in confirming negative infection status, its lack of discrimination between historical and current infections make it unsuitable for on-going monitoring, particularly where a treatment programme is in place.

4.6.3 Implications for PZQ treatment recommendations

It is worth bearing in mind that the main aim in measuring prevalence of schistosomiasis in these various locations is to use the information to provide recommendations to the Vector Control Division's Ministry of Health, in order to maximise the efficiency and efficacy of their national treatment programme. Under-estimating the prevalence of intestinal schistosomiasis

at a locality could result in a diversion of precious drug resources to other sites, but similarly, in order to be cost-effective it is important not to overestimate wildly the prevalence of the disease.

Here, we have shown that SEA-ELISA is likely to cause this latter problem, particularly in areas with a previous treatment history. However, we also show that CCA tests can also overcall prevalence, at least in the context of the Ssesse Island data and when compared against Kato-Katz smears from a single stool sample. For example, the costs of a blanket annual mass drug administration programme can be estimated (based on an average dose of 2.7 tablets of praziquantel per child, $0.05 per tablet, and a population of an estimated 250 children per village, over a five-year period) at $7593.75. Simply by doing a local-scale rapid assessment survey such as this one, costs can be more than halved. Based on Kato-Katz prevalence, 13 schools require annual treatment; if the other 32 are all given a single dose, the cost is just $3273.75, for the five-year period. This value holds true for if 'trace' CCA calls are considered negative, as well. If 'trace' is considered positive, the amount needed rises to $4353.75, but this is still much less than the cost of blanket annual treatment. Of course, this does not take into account the cost of training, transport, monitoring and other supplementary activities, but the point stands: regardless of diagnostic technique, implementing local-scale monitoring into control programme surveys can greatly reduce the cost of the intervention, by tailoring treatment regimes to local needs (Stothard, 2009).

4.6.4 Further considerations

The cost of the tests themselves is a factor that needs to be considered prior to their widespread implementation; currently batches of less than 10000 tests cost $2.85 per test. This is likely to be reduced to $1.61 for batches over 10000, but is still too expensive for mass usage by developing countries which should be approaching $1 per test to be cost-effective (Stothard *et al.*, 2009a). However, one must also consider the labour and time costs of making and reading the Kato-Katz slides; whereas CCA tests can be set up and read in parallel, Kato-Katz slides must be prepared and read in series. Taking urine and stool samples from 15 children at a school takes approximately an hour; from that point on, the CCA tests require

about half an hour to be set up and read, and can be managed easily by one person. On the contrary, 15 Kato-Katz slides, take about an hour to make, require large amounts of water and detergent to clean the materials, and take about 45 minutes to read, which further requires a compound microscope. All told, two or three technicians are usually required. Finally, stool samples are usually only readily collected from pre-adolescent children, whereas urine samples can be more easily gathered from a range of ages and are also essential for detection of urinary schistosomiasis. The rapidity, simplicity and efficacy of the new model of CCA test lead us to conclude that they are, at the least, a very useful addition to Kato-Katz slides for mapping the prevalence of *S. mansoni* infections. However, the Kato-Katz methodology is used in diagnosing soil-transmitted helminthiasis, and for this reason should not be overlooked if control of schistosomiasis and other helminthiasis are to be combined.

4.7 Conclusions

These surveys have shown that prevalence varies widely between and within different regions of Lake Victoria, and furthermore, diagnostic tests can have varying efficacies on different populations that are surveyed. This was reflected in the different diagnostic test scores between the two regions surveyed here. In many cases on the Ssesse Islands, different diagnostics furthermore could have resulted in different treatment recommendations. It is also worth bearing in mind that heterogeneity in snail distributions may create focality of transmission of *S. mansoni*, subtly changing the dynamics of transmission at the local level and impacting diagnostic performance; snail distributions around Lake Victoria and their effect on distribution of *S. mansoni* will be explored in later chapters. Considering the need to determine exactly what local needs are, in the face of varying results from different tests, more research is needed, for example in comparing the CCA tests against triple Kato-Katz smears from successive stools, as a more accurate gold measure. This way, it might be possible to create a correction factor for the CCA test, to be used in rapid assessment settings where triple Kato-Katz smears are impractical. For the meantime, we urge researchers to continue to test the CCA urine lateral flow test alongside Kato-Katz stool smears, in different epidemiological settings. For example, it would be useful to explore in greater detail the effect of genetic diversity on the

efficacy and strength of the CCA tests; as our knowledge base grows regarding the molecular epidemiology of *S. mansoni* in Lake Victoria, we may be better placed to judge its effect on diagnostic tools. The next chapter takes up this theme and presents the results of a molecular epidemiological analysis of *S. mansoni* from the full extent of Lake Victoria.

5 Molecular epidemiology and phylogeography of *S. mansoni* around Lake Victoria

5.1 Abstract

Intestinal schistosomiasis continues to be a major public health problem in sub-Saharan Africa, and is endemic in communities around Lake Victoria. Interest is growing in the molecular evolution and population genetic structure of *S. mansoni* and we describe a detailed analysis of the molecular epidemiology and phylogeography of *S. mansoni* from Lake Victoria. In total, 388 cytochrome oxidase 1 (COI) sequences were obtained from parasite material retrieved from 25 sites along the Ugandan, Tanzanian and Kenyan shorelines of Lake Victoria, and 122 unique haplotypes were identified; 9 corresponded to previously discovered haplotypes from Lakes Victoria and Albert. A subset of the data, composed of COI sequences from miracidia from 10 individual children, was used for population genetics analyses; these results were corroborated by microsatellite analysis of 4 isolates of lab-passaged adult worms. Overall, 12 haplotypes were found to be shared across all 3 countries, whereas the majority occurred singly and were locally restricted. The population genetics analyses were in agreement in revealing high diversity at the level of the human host and negligible population structuring by location. The lack of correlation between genetic distance and geographical distance in these data may be attributed to the confounding influence of high intraindividual diversity as well as human migration between communities.

5.2 Contributions of the author

All of the schistosome genetic material used in this chapter was collected in the field by the author, bar the Lake Albert material, which was brought back from a separate field expedition but maintained in the lab by the author (with assistance from the animal house and snail room staff; see Acknowledgements section). All of the DNA extractions (FTA card as well as adult worm) were done by the author, as were all of the PCR amplifications and column purifications.

From 2007 until the summer of 2009, the author was also setting up all the schistosome sequencing reactions; from summer 2009 onwards, these were done by the Sequencing Facility staff, as part of a wider change to the standard operating procedures of the facility. The Sequencing Facility ran all DNA sequence and microsatellite samples, while all sequence/microsatellite editing and analyses were done by the author.

5.3 Introduction

The previous two chapters have described the distribution and prevalence of *S. mansoni* around Lake Victoria, with a view to evaluating the best diagnostic tools for rapid mapping and use the information to tailor treatment interventions in endemic areas. However, such surveillance rarely looks further than measuring point prevalence and intensity of infection, even though molecular tools could potentially inform, with greater accuracy, data regarding re-infection rates, emerging resistance to PZQ, and theories of local snail-parasite compatibility.

Current developments of molecular tools, and in particular advances in DNA sequencing, have made genetically-focused research all the more timely and DNA 'barcoding' approaches are being more widely adopted (Rollinson *et al.*, 2009). For example, using such methods, a new species of schistosome, from the *S. haematobium* group and believed to be a sister species to *S. intercalatum*, has been described from a site near Kisumu, in the Lake Victoria basin (Hanelt *et al.*, 2009), highlighting that there are still exciting opportunities for molecular characterization research.

Other recent work in East Africa has shown that *S. mansoni* from the region was included in four out of five global 'groups' of the parasite, by looking at the cytochrome oxidase 1 (COI) gene, but sampling was reported at a district, if not country, level, and haplotypes from Lake Victoria all clustered together in Group 4 (Morgan *et al.*, 2005). Attempts at greater population sampling in later surveys, such as that by Stothard *et al.* (2009c) have shown a high diversity of haplotypes in the 'ASMIT' fragment of the COI gene, and also an apparent divide between haplotypes collected from Lake Victoria versus those of neighbouring Lake Albert, suggesting population segregation, perhaps due to differences in snail-parasite compatibility between the

two lakes. If this were the case, then it might also be expected to find differentiation between the haplotypes found in various parts of Lake Victoria, spanning as it does over 3000km of shoreline. Work on microsatellites in *S. mansoni* from Kenya and Uganda (Lake Albert region) has revealed high allelic diversity across numerous loci, which also revealed a surprising absence of heterozygotes, perhaps due to population structuring within human host (Agola *et al.*, 2006, Gower *et al.*, 2007).

The aims of this study were to sample *S. mansoni* as extensively as possible, taking advantage of FTA card technology, from Lake Victoria by collecting widely across the shoreline environments, to determine whether diversity is high throughout the lake and if so, to consider possible reasons for this. Using multiple genetic markers, this variation could also be examined at regional, local and host levels. Furthermore, the data were analysed spatially, to look for geographical trends that would either support or counter the conclusions drawn by previous genetic studies of *S. mansoni* in this region where significant geographical patterning was inferred.

5.4 Methods

5.4.1 Sample collection

Miracidia and cercariae were collected from 25 different sites around the Ugandan, Tanzanian and Kenyan lakeshore of Lake Victoria (see Figure 5.1), using the methods of hatching from human stool (see Methods section 2.2.1) and shedding wild-caught snails (section 2.3).

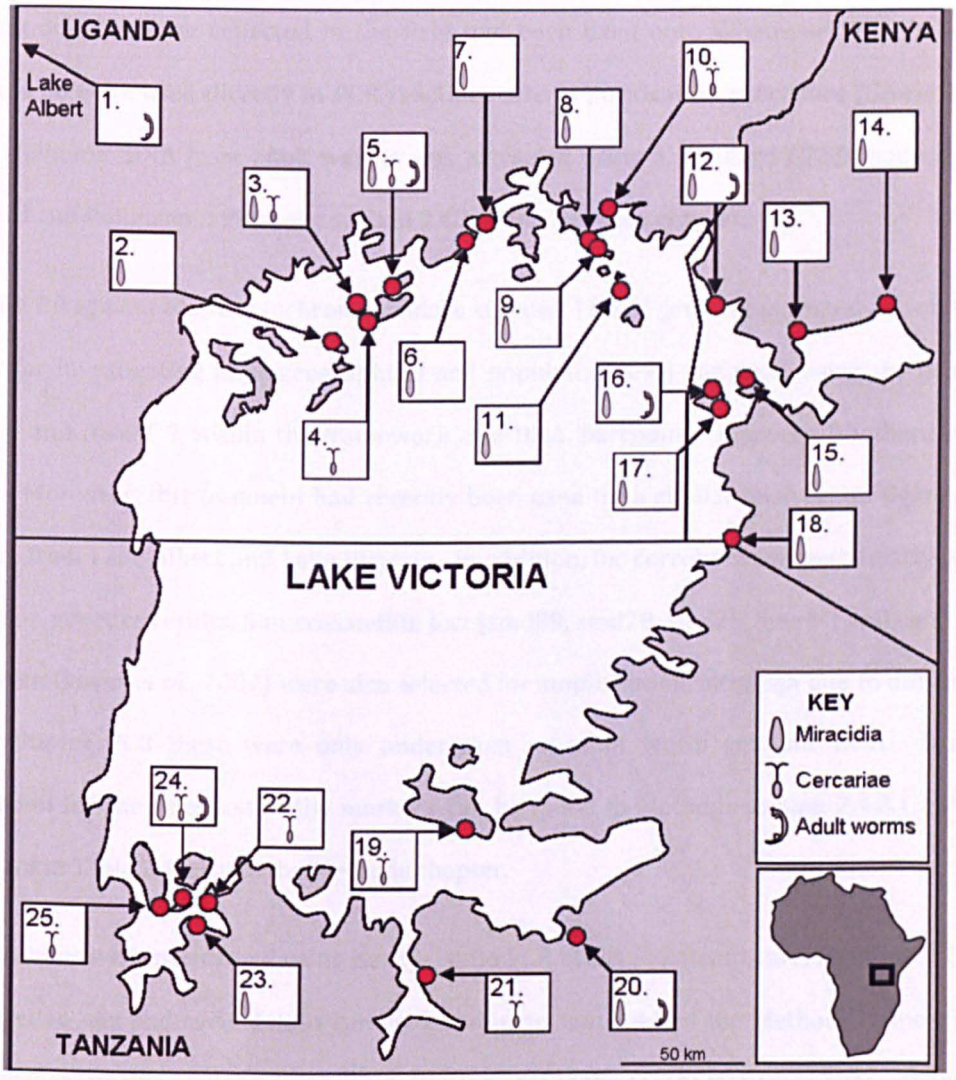

Figure 5.1 – Map of collection sites for *S. mansoni* genetic material, denoting which forms of the parasite were collected at each location.

Adult worms were collected from laboratory-infected mice, which had been exposed to cercariae produced by snails (of various species, see Table 5.1) infected from miracidia from the sites shown here.

Where snails had been experimentally infected with miracidia in the field, emerging cercariae were used to infect laboratory mice, allowing for later collection of adult worms (section 2.2.1). The material was collected over three missions, each lasting three to four weeks: Uganda 2008, Tanzania 2008 (Mwanza region and further west), and finally Tanzania/Kenya 2009 (Musoma region and north across the Kenyan border up to the Ugandan border).

5.4.2 Extraction, amplification and sequencing

Miracidia and cercariae collected in the field had been fixed onto Whatman© FTA Indicator cards and so were used directly in PCR reactions after a purification procedure (Gower *et al.*, 2007). Genomic DNA from adult worms was extracted using a standard CTAB methodology (Stothard and Rollinson, 1996); see section 2.4.1 of the Methods chapter.

The ASMIT fragment of the cytochrome oxidase sub-unit I (COI) gene was selected as a suitable marker for investigating phylogeographical and population level variation, using the primers ASMIT 1 and ASMIT 2 within the framework of a DNA 'barcoding' approach (Stothard *et al.*, 2009c). Moreover, this fragment had recently been used for a similar analysis on Ugandan *S. mansoni*, from Lake Albert and Lake Victoria. In addition, for corroboration particularly of the population genetics results, 5 microsatellite loci (smd89, smd28, smd25, sms9-1 and cal1-1, all taken from Gower *et al.*, 2007) were also selected for amplification, although due to difficulties with multiplex PCR these were only undertaken on adult worm genomic DNA. Further justification for the selection of the markers can be found in Methods section 2.4.2.1; primer details are in Table 2.1, also in the Methods chapter.

Amplifications were performed using Ready-To-Go PCR beads (GE Healthcare, Chalfont St Giles, UK); reaction mix and cycle details can be found in section 2.4.3 of the Methods chapter. PCR products were purified, quantified and then, in the case of the COI ASMIT fragments, sequenced (section 2.4.3). Purified microsatellite fragments were diluted and analysed on an automated sequencer. Sequences were aligned and visually inspected to check all base calls against the chromatogram trace; similarly allele sizes of microsatellite amplifications were individually checked (section 2.5.1).

5.4.3 Phylogeographical and population analyses

Once aligned, the COI sequences were evaluated for the number of unique haplotypes and the location and frequency of each of these haplotypes was noted. The haplotypes were also compared against an existing dataset of ASMIT *S. mansoni* fragments from Uganda (Lake Albert as well as Victoria; Stothard *et al.* 2009c) to check for duplicates. All novel haplotypes found were submitted, per geographical location of origin, to Genbank (accession numbers GQ415163-GQ415316). The haplotypes were used to create a phylogenetic tree and a minimum-spanning network; these are described more fully in sections 2.5.3 and 2.5.4 of the Methods chapter.

The population genetics tests were applied to two separate datasets (see section 2.5.5). The first consisted of COI sequences: only miracidia were included, and only those from hosts where greater than 10 samples had been successfully sequenced, to ensure sufficient statistical power. The second dataset comprised of the microsatellite alleles amplified from the adult worm isolates. Population genetics tests aimed to look for population structure at the host, site and regional level, for demographic change such as population expansion and for the neutrality of the data. Spatial correlations were also investigated. For the microsatellites, tests were also performed to ensure that the loci were independent of each other and that the populations conformed to Hardy-Weinberg equilibrium. Full details of the software used for the tests as well as further explanations of the tests can be found in Methods section 2.4.6.

5.5 Results

The results showed a very high level of genetic diversity, low population structuring and little evidence for any geographical trends.

5.5.1 COI haplotype data

Overall, 388 sequences were obtained: 60 from adult worms, 57 from cercariae and 272 from miracidia (see Table 5.1, over two pages).

Country	Site Name (Map ID) GPS coordinates (dec. °)	Life stage	Source material: pooled or individual?	Number of sequences	Barcodes found
UGANDA	**Bugoigo, Lake Albert** (1) *N1.90522;E31.40728*	adult worms	Pooled children, pooled snails[g. 1]	8	1, 8, 81, 84, 85
	Jaana (2) *S0.23380;E32.57508*	miracidia	Child # 9	4	1, 33, 79
			Pooled children	2	1, 19
	Kimi Island (3) *S0.08638;E32.65220*	cercariae	Snail # A[c]	7	1, 24, 25
	Ngamba Island (4) *S0.09987;E32.65303*	cercariae	Pooled snails[c]	6	26, 27
	Kisu (5) *N0.01477;E32.76717*	miracidia	Pooled children	5	4, 8, 14, 20, 34
		cercariae	Snail # C[c]	5	23
		adult worms	Pooled children, pooled snails[g. 2]	16	47, 72, 73
	Busagazi (6) *N0.24030;E33.13708*	miracidia	Pooled children	2	1, 14
	Kiyindi (7) *N0.27132;E33.15325*	miracidia	Pooled children	1	22
	Bwondha (8) *N0.17063;E33.56510*	miracidia	*Child # 551*	12	2, 10, 14, 20, 23, 54, 120, 121, 122, 123, 124
			Child # 622	13	1, 17, 79, 100, 114, 115, 116, 117, 118, 119
			Child # 624	20	1, 14, 17, 35, 36, 42, 47, 125, 126, 127
			Pooled children	28	1, 2, 10, 16, 17, 20, 42, 124, 128, 129, 130, 131
	Kaaza (9) *N0.11177;E33.60215*	miracidia	Pooled children	2	21, 30
	Bugoto (10) *N0.31812;E33.62703*	cercariae	Snail # 30[c]	3	8
			Snail # 9[c]	3	28, 29
			Pooled snails[s]	3	16
	Sagitu (11) *N0.00303;E33.65860*	miracidia	Pooled children	3	1, 31, 32
	Mayuge District (Pooled 8+10) *See above*	miracidia	Pooled children	1	18
	Mukono District (Pooled 3+5) *See above*	miracidia	Pooled children	11	2, 8, 10, 16, 41, 42, 45, 54
KENYA	**Usengi** (12) *N0.07313;E34.05849*	miracidia	*Child # 1*	13	4, 16, 36, 42, 45, 65, 66, 80, 82, 83
			Child # 3	3	2, 35, 67
			Child # 13	1	18
			Child # 15	1	16
			Pooled children	1	35
		adult worms	Pooled children, pooled snails[g. 1]	11	14, 87, 91, 92, 93, 94, 95, 96
	Asembo Bay (13) *S0.18737;E34.38747*	miracidia	Child # 20	7	1, 36, 53, 64
			Pooled children	2	51, 52
	Usoma (14) *S0.10519;E34.71829*	miracidia	Child # 12	3	1, 54, 55
			Child # 13	11	1, 31, 42, 59, 67, 68, 69, 70, 86
			Pooled children	3	31, 36
	Kolunga (15) *S0.42781;E34.13857*	miracidia	*Child # 1*	24	1, 10, 17, 23, 36, 41, 48, 49, 50, 60, 61, 62, 74, 75, 76, 78
			Child # 18	16	1, 10, 16, 23, 36, 43, 103, 110, 111, 112, 113

			Child # 20	1	63
			Child # 21	3	1, 36
			Pooled children	2	43, 44
	Yokia (16) *S0.43738;E34.01513*	miracidia	Pooled children	5	1, 10, 16, 56, 57
		adult worms	Pooled children, pooled snails[g, 1]	10	16, 65, 87, 88, 89, 90
	Kitawi (17) *S0.53805;E34.16403*	miracidia	Pooled children	4	1, 14, 36
	Migori (18) *S1.01244;E34.12677*	miracidia	Pooled children	3	1, 42
TANZANIA	**Hamuyebe** (19) *S2.12400;E33.06497*	miracidia	Child # 6	1	16
			Child # 7	17	1, 10, 16, 17, 35, 39, 40, 97, 98, 99, 100, 101
			Child # 10	19	1, 10, 14, 16, 18, 36, 75, 102, 103, 104, 105
		cercariae	Pooled snails[g,1]	3	37
	Nyashimba (20) *S2.45220;E33.51693*	miracidia	Pooled children	6	10, 23, 36, 46, 71
		adult worms	Pooled children, pooled snails[g, 1]	5	46, 58
	Kigongo (21) *S2.71347;E32.89392*	cercariae	Snail # 1[s]	8	1, 16
	Kahunda (22) *S2.40525;E32.05935*	cercariae	Pooled snails[c]	8	1
	Nkome (23) *S2.50673;E32.01462*	miracidia	Pooled children	4	1, 16, 41, 67
	Izumacheli (24) *S2.38317;E31.96692*	miracidia	*Child # 4*	18	1, 10, 16, 17, 18, 31, 36, 75, 77, 106, 107, 108, 109
		cercariae	Pooled snails[s], exposed to miracidia from Child # 4	3	10, 35
			Pooled snails[g], exposed to miracidia from Child # 4	3	38
		adult worms	Child # 4, pooled snails[s, 1]	10	1, 36, 77
	Rubondo Island (25) *S2.41502;E31.92260*	cercariae	Pooled snails[c]	2	36

Table 5.1 – Schistosome sampling sites, life stage of parasite collected and source, number of individuals sequenced and haplotypes found at each site

Wild caught snails were preliminarily identified by shell morphology, and *Biomphalaria glabrata* used for laboratory infections: [c] = *Biomphalaria choanomphala*; [s] = *B. sudanica*; [g]= *B. glabrata*. [1, 2] = refers to the passage number of the isolate through laboratory mice. The sequences recovered from children in *italics* were used in the COI sub-dataset for population genetics analyses. 'Map ID' refers to the location of each site in Figure 5.1; 'dec. °' stands for decimal degrees.

These were resolved into 122 unique haplotypes; of these, nine corresponded to haplotypes already described by Stothard *et al.* (2009c), who also found a further nine haplotypes which were not recovered in this survey, creating a combined haplotype dataset of 131 haplotypes. The majority of haplotypes were recovered from miracidia collected from individual children (66.4%).

The percentage occurrence of each haplotype varied widely, with the majority (66.4%) of the collected haplotypes being only found once; seven haplotypes were recovered ten or more times. However, the abundant haplotypes were also generally widely geographically distributed, occurring in all three countries (see Figure 5.2).

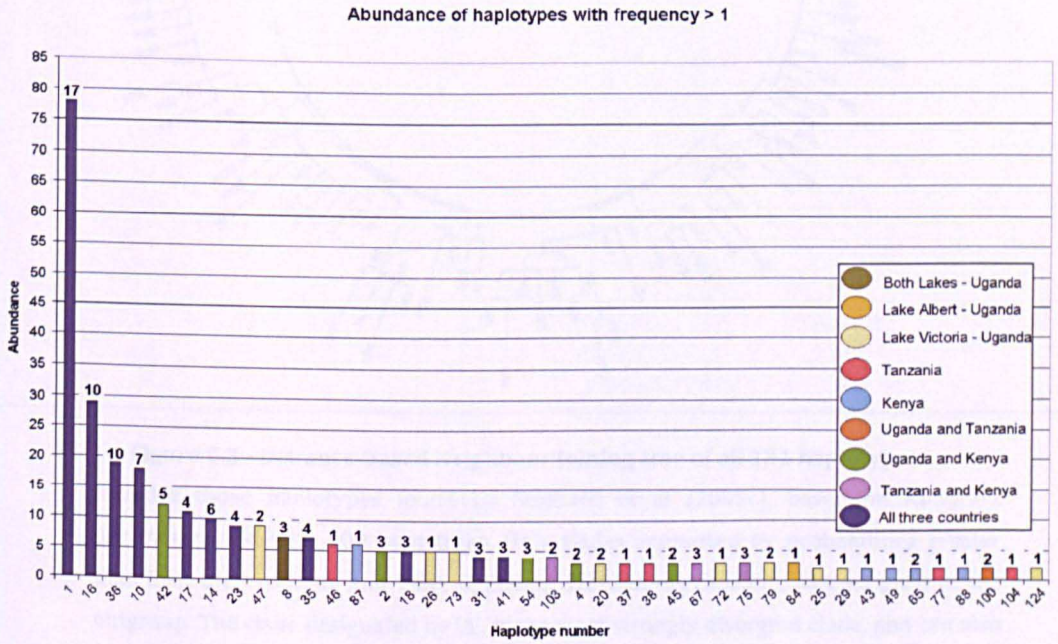

Figure 5.2 - Haplotype abundance bar chart, for those haplotypes recovered more than once.

The colour of each column refers to the countries and/or lake in which the haplotype was found; the number above each bar is the number of sites, out of the total of 25, at which the haplotype was found.

5.5.2 Phylogenetic tree

The bifurcating tree supported showed a majority of haplotypes showing low differentiation from each other, with the exception of one clade, which was reasonably well supported (Figure 5.3).

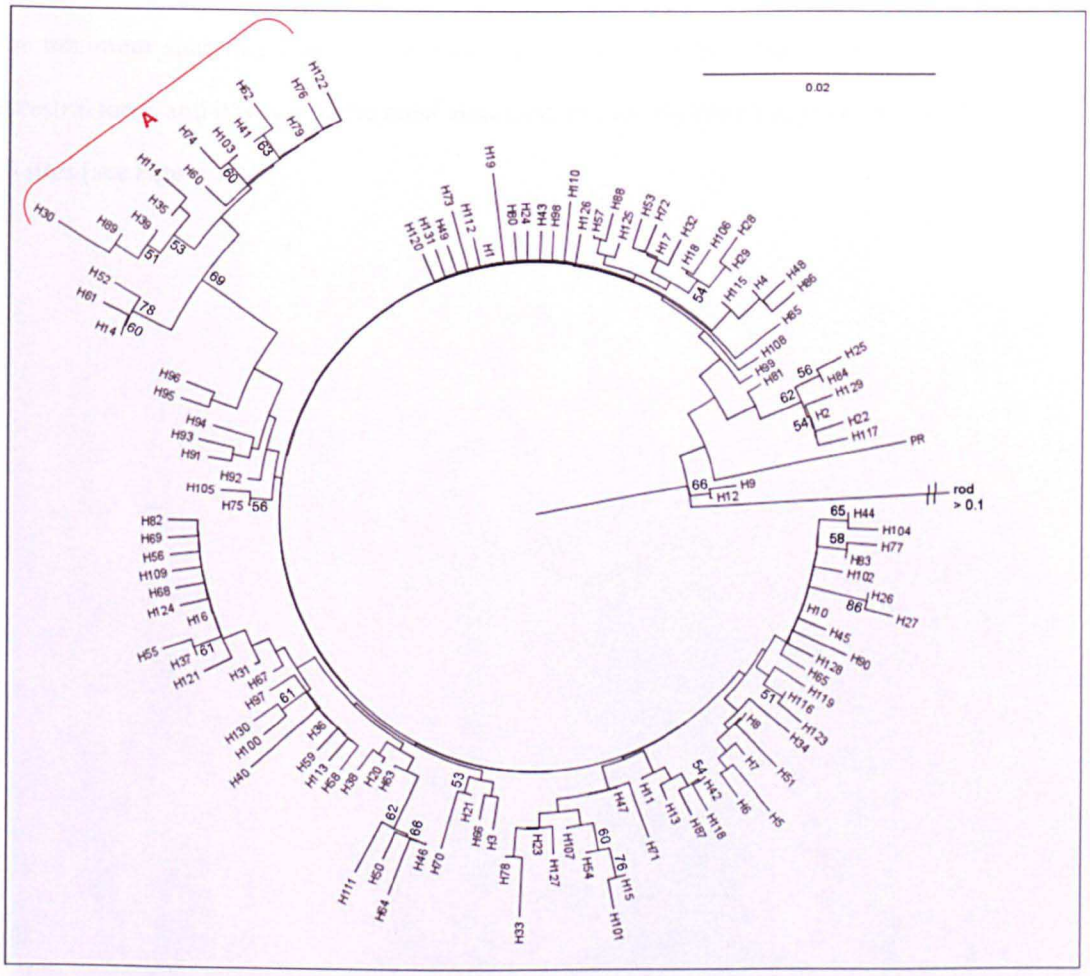

Figure 5.3 - Distance-based Neighbour Joining tree of all 131 haplotypes

Includes those haplotypes found by Stothard et al (2009c), based on Kimura-2 parameters and with 500 bootstraps. Only clades supported by probabilities greater than 50% are reported, and 'rod', signifying *Schistosoma rhodaini*, was assigned as the outgroup. The clade designated by 'A', is the most strongly divergent clade, and can also be seen as a more distant grouping on the minimum spanning network in Figure 5.4.

Two haplotypes, 9 and 12, which had not been observed in the Lake Victoria schistosomes in this study but had been collected by Stothard and colleagues (2009c) were observed to cluster most closely with the Puerto Rican sample, and indeed even basal to it, indicating their distinctiveness from the rest of the Great Lakes haplotypes.

5.5.3 Network

The minimum spanning haplotype network suggested that haplotype 1 was the most likely ancestral form, and it was also the most abundant and widely distributed, occurring at 17 of the 25 sites (see Figure 5.4).

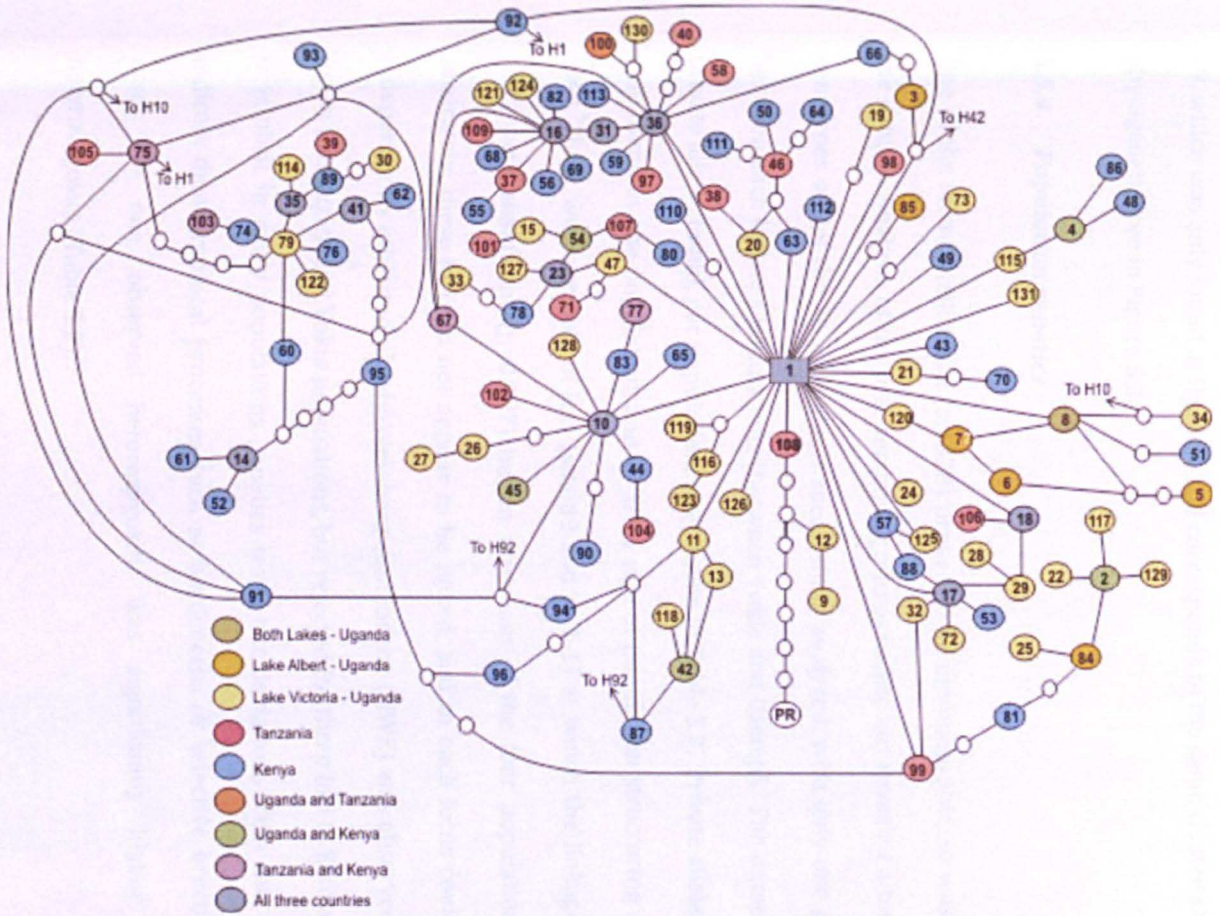

Figure 5.4 - Minimum spanning haplotype network

Includes those haplotypes found by Stothard *et al.* (2009c). The colour of each circle refers to the countries and/or lakes in which that haplotype was found (note that size of circle is not proportional to abundance). Each line represents a single base change; small white circles represent unsampled, or perhaps extinct, haplotypes. 'PR' refers to an isolate of *S. mansoni* from Puerto Rico, sequenced for the same 'ASMIT' region (reference as above).

In addition, the network showed that central nodes tended to be widely distributed and occur in all three countries. The hypothesised presence of some non-sampled haplotypes was evident by the occurrence of missing steps between haplotypes on the network. Two branches were emphasised by longer haplotype-free stepwise substitutions: one culminated in the Puerto Rican *S. mansoni* out group sample, but the other centred around haplotype 79, which was only found in Uganda, and corresponds to the same cluster as observed in the phylogenetic tree in Figure 5.3.

5.5.4 Population genetics

One of the microsatellite loci (smd28) proved to be monomorphic, so was discarded from the dataset. Analysis of the four remaining microsatellite loci revealed a total of 22 distinct genotypes out of the 36 individuals successfully analysed, with only one genotype shared between each isolate's populations (between Yokia and Usengi). The number of alleles and private alleles found, per population, can be seen in Table 5.2. Private alleles indicate alleles only found in one population, and so are a sign of population structuring and lack of gene flow. There was one pair of loci (smd89 and cal1-1) at which the linkage disequilibrium tests were significant (p=0.007) but in three out of the four populations, when tested separately, these loci did not appear to be linked, and so each locus could be considered independently inherited. Hardy-Weinberg equilibrium (HWE) was observed for all four loci in the Izumacheli and Yokia populations, but rejected for three loci in Kisu and Usengi. HWE is implicit in many populations genetics tests, but deviations from this equilibrium can indicate demographical processes such as bottlenecks or selective breeding. In all these cases bar one, observed heterozygosity was significantly higher than expected heterozygosity (Table 5.2).

Site (Map ID)	N	B	G	smd89			smd25			s9-1			cal1-1			smd28	
				A	Ap	H_O (H_E)	A	Ap	H_O (H_E)	A	Ap	H_O (H_E)	A	Ap	H_O (H_E)	A	Ap
Kisu (5)	10	3	3	3	0	1.00 (0.68)	4	0	*1.00 (0.70)**	2	0	*0.91 (0.52)**	4	0	*0.55 (0.68)***	1	0
Usengi (12)	8	5	5	3	0	0.43 (0.38)	3	0	*1.00 (0.60)**	2	0	*1.00 (0.54)**	4	0	*1.00 (0.73)**	1	0
Yokia (16)	10	7	7	4	0	0.50 (0.44)	5	1	0.90 (0.73)	2	0	0.44 (0.52)	4	0	0.88 (0.66)	1	0
Izumacheli (24)	8	3	7	3	0	0.63 (0.69)	3	0	0.57 (0.62)	2	0	0.25 (0.50)	4	0	0.80 (0.71)	1	0
TOTAL	**36**	**17**	**22**	**4**	--	--	**6**	--	--	**2**	--	--	**5**	--	--	**1**	--

Table 5.2 - Microsatellite allelic diversity, from four populations of lab-passaged adult worms

'N' = number of adult worms analysed, 'B' = number of different COI haplotypes observed among these adult worms, 'G' = number of different microsatellite genotypes observed, 'A' = number of alleles observed at that locus, 'Ap' = number of alleles unique to that population ('private' alleles), H_O= observed heterozygosity and H_E= expected heterozygosity. Size ranges observed for the five loci were: smd89 = 154-160; smd28 = 241 (only one allele observed); smd25 = 284-298; s9-1 = 198-200; cal1-1 = 205-214. '*' signifies expected heterozygosity was significantly lower than observed heterozygosity; '**' corresponds to the opposite.

Tests for population bottlenecks showed that two populations (Kisu and IZ4) may have passed through a recent genetic bottleneck, as there was significant heterozygotic excess compared to allelic richness (p=0.031 in both cases). The average microsatellite F_{ST} value across the populations was 0.06, indicating very low genetic differentiation, although differentiation was significant (and greater than 0.10) between Kisu and Usengi and Izumacheli and Usengi, respectively. Differentiation was also significant between Yokia and Kisu, but was 0.047, indicating high cross-over of alleles. Overall, the results of the AMOVA showed that 93.9% of the total variation seen could be explained at the level of individual children rather than between hosts or sites, suggesting high infrahost genetic diversity (Table 5.3).

Source of variation	COI 'Asmit' region				Four microsatellite loci			
	d.f.	Sum of squares	Variance components	% of variation	d.f.	Sum of squares	Variance components	% of variation
Among groups	5	2.715	-0.002 Va	-0.42	2	4.120	0.048 Va	5.45
Among populations, within groups	4	2.399	0.008 Vb	1.67	1	0.933	0.006 Vb	0.62
Within populations	153	71.684	0.469 Vc	98.75	68	57.280	0.842 Vc	93.93
Total	162	76.798	0.474	100.00	71	62.333	0.897	

Table 5.3 - AMOVA results for COI 'Asmit' region population sub-set and microsatellite four loci population subset

'd.f.' = degrees of freedom. Note that these were two different sub-sets of the total dataset (see section 5.3.3).

Similar tests were performed for the COI region on the sub-set of the data containing sequences retrieved from individual children. These showed an F_{ST} fixation index for the sub-dataset as whole of 0.033 (not significant, once Bonferroni correction for multiple tests was applied), suggesting negligible levels of genetic differentiation, and the same was true for the pairwise F_{ST} values calculated between populations. The AMOVA analysis, performed with groups defined by site, again showed that the vast majority of the variation was explained at the intra-host level (Table 5.3). The analysis was executed a second time, with the groups defined by country, and the results remained similar, with 98% of the variation explained by variation within each host.

The Mantel Z-test, which was used to look for spatial autocorrelation within the genetic distance data, was insignificant (p = 0.888), and the correlation coefficient was negative (-

The Mantel Z-test, which was used to look for spatial autocorrelation within the genetic distance data, was insignificant ($p = 0.888$), and the correlation coefficient was negative (-0.374), which suggests increasing genetic similarity with greater geographical distance. A plot of the mean pairwise genetic distances against the pairwise geographical distance similarly showed no clear trend, apart from a slight negative correlation (see Figure 5.5).

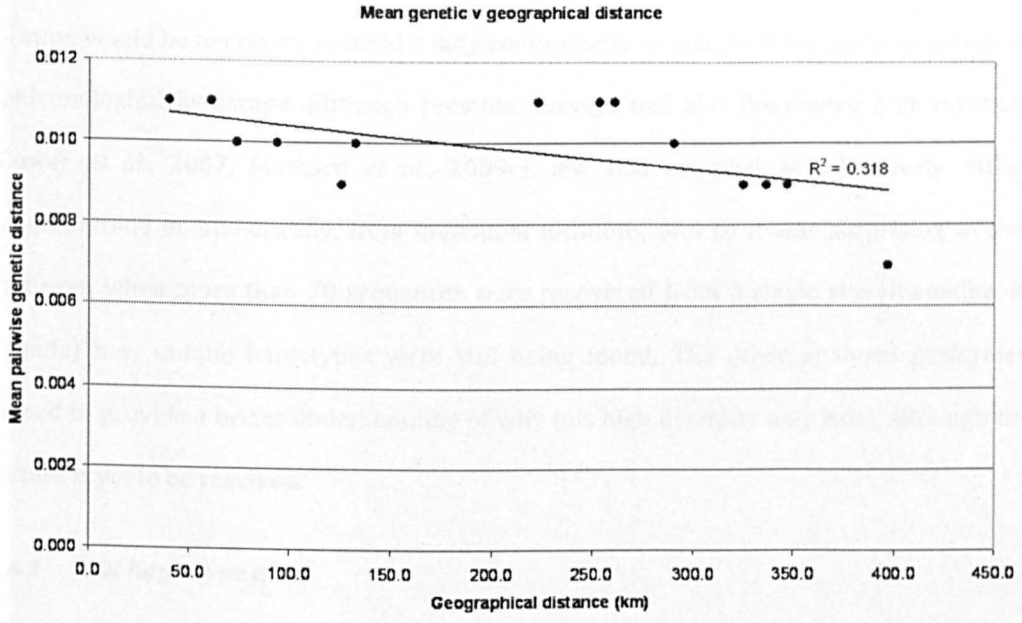

Figure 5.5 - Graph showing pairwise geographical distance (km) plotted against mean pairwise genetic distance.

Tajima's neutrality test was performed on each of the populations in the smaller sub-set of data, and the p-values produced from the resultant *D* values were all found to be significant (bar Child # 4 from Izumacheli), which rejects neutrality (Tajima, 1989). Fu's Fs test also produced values which were all significant ($p < 0.02$), also suggests rejecting neutrality (Fu, 1997). A mismatch test was performed on the dataset, and both the sum of squared deviation from simulated data and Harpenden's Raggedness index were not significant, thus preserving the null hypothesis of population expansion, which could also explain the lack of neutrality as observed both by Tajima's *D* and Fu's Fs tests.

5.6 Discussion

This survey has revealed that the extent of genetic variation of *S. mansoni* in Lake Victoria is much higher than previously envisaged. Most importantly, the vast amount of diversity is found within individual children and not at the geographical level from which they were sampled (see Table 5.1); this immediately highlights that increased sampling per child by location would be necessary to build a fully comprehensive picture of the parasite across its epidemiological landscape. Although previous surveys had also discovered high diversity (Gower *et al.*, 2007, Stothard *et al.*, 2009c), few had sampled as extensively, either geographically or numerically, from individual locations, and so it was surprising to find that even when more than 70 sequences were recovered from a single site (Bwondha, in Uganda) new unique haplotypes were still being found. The other analyses performed helped to provide a better understanding of why this high diversity may exist, although the picture is yet to be resolved.

5.6.1 COI haplotype data

The haplotype frequency bar chart showed that abundant haplotypes were also geographically widespread and found in multiple sites. Rare haplotypes may be examples of recent, local mutations, contributing to the earlier suggestion that haplotypes may be 'philopatric' to a particular lake or site (Stothard *et al.*, 2009c). Our data, in contrast, demonstrate how this type of site-based geographical interpretations can be misleading; as can be seen in the minimum spanning network, rare, geographically restricted haplotypes appear to radiate from widespread, common ones. There was, however, sampling bias intrinsic in the methodology, in that it was more difficult to collect many haplotypes from some sites than others, and some populations were sequenced more successfully; again, greater sampling effort at more locations might increase the abundance, and perhaps geographical spread, of even some of the 'rare' haplotypes found here (cross-ref Fig. 5.1 with Table 5.1). The mismatch test demonstrated that population expansion could also have occurred, which adds support for these results, and explains the non-neutrality of the data.

The Neighbour-joining (NJ) distance tree further supported the hypothesis of recent sequence radiation, as the majority of the bifurcations were only weakly supported (bootstrap probability < 50%). However, one clade, also identified on the minimum spanning network as being multiple steps away from the main, ancestral haplotype, was strongly supported by the tree, suggesting this divide may have occurred less recently. Morgan *et al.* (2005) showed that East African *S. mansoni* is represented in four out of the five major lineages recovered, although all Lake Victorian and Albertine samples were in one group (Group 4). Within this group, two differentiated subgroups were strongly supported (>80%) by all phylogenetic methods used. However, there appeared to be no obvious national or even lacustrine clustering within these groups, with haplotypes from Lake Victoria (Kenya and Tanzania) and Lake Albert (Uganda) appearing in both clusters. This is consistent with the tree structure found with our data; while most of the samples were poorly differentiated, a few sub-groupings were reasonably well-supported, though without clear geographical patterns. Conversely, Stothard *et al.* (2009c) found that when the separate 'MORGAN' section of the COI gene was sequenced for their Lake Victorian and Albertine data, their samples clustered within two out of the five major lineages of Morgan *et al.* (2005); haplotype 2 was found to be in a separate lineage to the other 17 haplotypes they found. As haplotype 2 was also recovered in this survey, the possibility that this pattern could be seen again with these data, given analysis of different genetic fragments, should not be discounted, especially as divergence of haplotype 2 was reasonably well-supported in the analysis presented here (Fig 4).

5.6.2 Population genetics

It is worth emphasising that despite some apparent differentiation between specific haplotypes, the sub-sets of the data analysed from a population genetics perspective showed very little structure, indicating that describing collections of parasites from a particular site or child as a 'population' may not be valid. This low level of population structuring is consistent with other research on allogenic parasites that have both terrestrial and aquatic hosts (Criscione and Blouin, 2004, Keeney *et al.*, 2009). The greater population structuring seen in microsatellite markers has been seen in other studies of

schistosome populations and has been in some cases attributed to human migration (Curtis *et al.*, 2002, Steinauer *et al.*, 2009). These studies also found significantly different levels of observed and expected heterozygosity, which have been attributed to effects of lab passage (Rodrigues *et al.*, 2002) and in this study, resulted in higher observed heterozygosity than expected. This may be due to enhanced selection for genetically diverse individuals (such as heterozygotes) in a laboratory setting, or a bottleneck event while the isolate is being established in the laboratory (Gower *et al.*, 2007). However, these findings also contrast with other recent studies of East African *S. mansoni*, where much stronger population structuring was evident (Agola *et al.*, 2006). The Mantel Z-test added statistical weight to the observation that there appeared to be no significant relationship between genetic distance and geographical distance among the COI sub-dataset.

5.6.3 Effect of human demography

Evidence highlighting the potential importance of human migration in creating the results described above comes from examining questionnaires, carried out at the same time as the parasite collection. These revealed a very high level of itinerancy among the communities being surveyed, as described in Chapter 3 (Standley *et al.*, 2009). Given a large parasite population size, terminal host migration may be implicated in the low levels of population structuring, as observed in this dataset (Blouin *et al.*, 1995, Dybdahl and Lively, 1996, Jarne and Théron, 2001). Furthermore, if migration carries particular parasite genotypes to new areas, transmission may still be able to occur, if compatible snails are also widely distributed. As part of this research, snail taxonomy has also been investigated (see Chapter 7), and genetic diversity found to be very high, which means that perhaps these common haplotypes of *S. mansoni* are compatible with a wider diversity of snails than previously thought. Even without the snails, adult worms may live for years within a terminal host, so people may carry genotypes of the parasite from a home village to areas far from the original infection site. As it happened, this effect was encountered within the data: Child # 4 from Izumacheli, in the south-western region of Lake Victoria, was found to have been born in Kisumu District in Kenya, near the Usoma site, and thus could have carried parasites from the region across the Lake to Tanzania. This raises the suggestion that the diversity of each

child's parasitological burden may act as a historical 'signature' of infection events, across different water contact sites and through time, depending on where that child has been living. Snail movement throughout Lake Victoria could also be implicated in de-coupling geographical and genetic patterns, as is the case in China with *S. japonicum* (Shrivastava *et al.*, 2005), although in this case further research on the snails revealed strong geographical structuring which in turn suggests limited dispersal (see Chapter 8 for full details of population genetics results on *Biomphalaria*).

5.6.4 Implications

The implications of these genetic findings are manifold. First of all, it would seem that sampling is still not sufficient even at the levels we have attempted to make a comprehensive assessment of diversity. Moreover, the existence of temporal variation in diversity patterns should also not be ruled out; despite exhaustive sampling, in this survey five of the haplotypes found in Mayuge District (Uganda) by Stothard *et al.* (2009c) were not recovered, despite revisiting two out of four of the same villages. This could be the result of treatment in the intervening time period, but also potentially the effect of variation in haplotypes between individuals, even from the same locality or population turnover. It is worth emphasising that the population structuring seen in these data is likely mirrored in other systems, where the parasite has high diversity and its human host may be highly migratory (Kumar *et al.*, 2002).

Secondly, the data also show that children have very diverse individual parasite loads; the implication of this finding on the resultant pathology of the disease is not known, nor how this relates to treatment performance, although there is evidence from studies of *S. haematobium* that genetic type may have implications for clinical manifestation of the disease (Brouwer *et al.*, 2003). Morbidity has been shown to vary between individual children, even from the same locality, independent of infection intensity, and therefore it is highly likely there could be a parasite genetic component to disease pathology (Balen *et al.*, 2006). Again, there are further implications to high parasite diversity; the children in this survey were shown to have unique profiles of haplotype diversity, so if morbidity is linked

to genetic type, then child-to-child comparisons of infections will be rendered dubious, as morbidity could only be evaluated or predicted at an individual level, by looking at that child's particular parasitological haplotype profile. Diversity may also impact re-infection and treatment dynamics (Dumont *et al.*, 2007). It is hypothesised that a wide genetic base for selection to act upon may increase the rate of resistance to treatment developing, though this would result in a decline in diversity over time to a few, resistant genotypes (Feng *et al.*, 2001). Continued monitoring of *S. mansoni* diversity in individual children will provide information regarding the efficacy of on-going treatment campaigns and warning signals were resistance to emerge.

Finally, our results have shown very high levels of genetic diversity of *S. mansoni,* which defy straightforward attempts to assign simple geographical patterns. Furthermore, temporal variation has not been explicitly considered here; on-going research has shown that in areas targeted for mass drug administration in Uganda, children entering school at the age of 7 have a completely separate suite of genotypes, which form a discrete phylogenetic clade, as compared to the previous year group (J. Webster, unpublished data). These temporal changes may have a strong impact on the patterns of population structure observed, based on molecular data. As such, rather than be tied to the concept of a particular school, or even site, at a specific time as the basis for a parasite 'population', these data suggest that future research on the genetic patterns of *S. mansoni* in this region should potentially consider Lake Victoria as one, inter-mixing parasite population, with a sub-structure of individual children who are highly mobile across the Lake and differential snail compatibilities. Further surveys, across more of the region, could determine at what scale this mixing occurs: for example, how parasitologically disjointed, or rather, how connected, is Lake Albert from Lake Victoria? This is particularly important in the context of on-going national control programmes, which might benefit from a more regionally integrated outlook.

5.7 Conclusions

Molecular characterisation of *S. mansoni* from Lake Victoria revealed very high levels of genetic diversity, even within a single human host. Little population structure was observed, with common and abundant haplotypes spread across all three countries and sites not spatially autocorrelated in terms of genetic distance. It is hypothesised that human migration may be implicated in mixing parasite populations, though snail-parasite dynamics should also be examined further. Lab passage appeared to affect the Hardy-Weinberg equilibrium of the microsatellite alleles, but otherwise results corroborated with the COI data. The implications of such high parasite genetic diversity on pathology, reinfection dynamics and emergence of resistance to PZQ should be investigated thoroughly in the future.

6 Patterns of presence and abundance of *Biomphalaria* in Lake Victoria: Using Bayesian spatial statistics to test for environmental predictors of distributions

6.1 Abstract

Understanding the distribution of intermediate host species is crucial to monitoring the transmission risk of parasitic diseases, such as intestinal schistosomiasis. Moreover, information on the environmental and ecological conditions that determine presence or absence of *Biomphalaria* can be used to make predictions of potential transmission to data-deficient regions. Here, 223 sites were visited along the shoreline of Lake Victoria and malacological surveys undertaken in each. Concurrently, environmental measurements were recorded and water samples taken for determination of anion and cation concentrations. The results were modeled in three separate, Bayesian, multivariate models: two non-spatial (one with a random effects parameter) and one spatial, also with random effects. Presence of *B. choanomphala* and presence of *B. sudanica* were modeled as separate outcomes. Results showed that chloride, nitrate, sulphate and the presence of other species were significant predictors of *B. choanomphala* presence whereas marsh habitat, depth, pH and also sulphate were predictors for *B. sudanica*. Variance in the models was high, indicating other factors, such as perhaps climatic variables or habitat composition details, should be considered in the future. The range of spatial autocorrelation was also large (572.9 km for *B. choanomphala* and 175.3 km for *B. sudanica*). The predictive kriging map created by interpolating abundance data revealed three potential hot spots of *Biomphalaria* abundance: one of these is in a remote, under-surveyed region and so should be a priority target for future expeditions, which would further assist with validation of the model. Overall, the predictive power of these models was limited by the difficulty of remotely measuring the variables found to be significant, and so on-the-ground malacological work remains extremely important in determining the transmission risk of intestinal

schistosomiasis based on the presence of its host snails, especially in the context of changing seasonal patterns or longer term trends caused by climate change.

6.2 Contribution of the author

The water samples and *in situ* water chemistry measurements were taken during the course of all the different field missions to Lake Victoria, and the author was present for the entirety of each one, and visited every site sampled. The author usually took water readings personally, but on occasion other team members would assist, as was also the case with the malacological scooping. The author was responsible for the sorting, indentification and shedding of the snails at each site. Water chemistry analyses at the NHM were conducted by staff members of the EMMA facility (see the Acknowledgements section), although all data management and analysis was done by the author. The Bayesian geospatial models were executed by the author, with instruction and assistance from Dr Penelope Vounatsou and Dr Laura Gosoniu (see again the Acknowledgements section).

6.3 Introduction

The transmission potential of intestinal schistosomiasis is limited to areas where *Biomphalaria* snails, the parasite's intermediate host, are found. As such, a sound knowledge of the distribution and abundance of these snails will assist in identifying potential high-risk transmission zones, although not all areas with *Biomphalaria* are necessarily hot-spots for prevalence of the disease.

Malacological surveys, on the ground, are a simple method for gathering information about the distribution of snails. However, they are time-consuming and logistically challenging, particularly when trying to survey areas with little or no road access, where surveys must be conducted by boat, via suitable landing sites. In addition, snail populations are known to fluctuate due to seasonal and climatic effects, such as changes in water flow/level, desiccation of temporary water-bodies and variation in temperature (Woolhouse and Chandiwana, 1990). In response to these issues, attempts have been made to model factors

that may be indicative of *Biomphalaria* abundance, and use the information to predict the presence of the snails in areas that have not been surveyed on the ground. This is important in the context of shedding light on parasitological data in the same localities.

This approach requires initial data about snail distributions collected from malacological surveys, which can then be added to environmental or other data and statistically investigated together to look for significant trends. Often, the environmental data used have been extracted from remote sensing data from satellite images of the region. As remote sensing data are readily accessible and extensive, they are increasingly being utilized to derive environmental information using geographical information systems (GIS) software (Brooker and Michael, 2000, Simoonga *et al.*, 2009). Combining GIS to the statistical analysis can include an explicit spatial element to the model, allowing for tests of autocorrelation within the data; the range of autocorrelation in the data determines the extent of interpolation that is possible later, when creating predictive maps from the data. This approach has been used directly to predict schistosomiasis prevalence in a number of settings, based on factors such as temperature, rainfall and demographical questionnaires (Clennon *et al.*, 2004, Raso *et al.*, 2005, Clements *et al.*, 2008); more recently, snail habitat suitability has also been included in these attempts (Stensgaard *et al.*, 2006). However, until recently, the spatial resolution of remote sensing data has only been sufficient for reasonably large scale models; newer satellite technology has reduced resolution from the level of square kilometers to the levels of meters, but few of these data are freely available for researchers.

The presence of snail intermediate hosts has also been modeled directly based on these methods (Kristensen *et al.*, 2001). Here, specific ranges of temperature and Normalized Differences Vegetation Index (NDVI, used often as a proxy of water availability) were determined to be strong predictors of true snail distributions, although this study relied on historical collections of snails rather than concurrent on-the-ground malacological surveys. Again, scale is a problem, as the resolution of remote sensing data is insufficient for detecting local-scale heterogeneities in environmental conditions.

This chapter presents a new approach to modeling snail distributions, in an attempt to increase the number of environmental variables under consideration as well as by-pass issues of remote sensing spatial resolution. By measuring environmental conditions directly on-the-ground, while at the same time surveying for *Biomphalaria* snails, the aim was to create a comprehensive and accurate model of conditions which can be used to predict the distribution of *S. mansoni* intermediate hosts.

6.4 Methods

6.4.1 Site selection and sample collection

Sites were selected to try to sample comprehensively around the majority of the shoreline of Lake Victoria. Snails were sampled semi-quantitatively at each location, by always using two collectors and sampling for approximately 15 minutes (section 2.1.4). If more than one habitat type was present at a given locality, both were inspected for presence of snails, and given a sub-designation ('A' and 'B', for example). Any *Biomphalaria* snails observed were putatively identified based on shell characters as *B. choanomphala* or *B. sudanica* and the abundance noted in categorical form (0, <10, 10-30 or >30). Collected *Biomphalaria* were placed in plastic jars filled with lake water for later processing. The presence of all other gastropod genera was also recorded (see Methods section 2.1.4 for a full list of gastropod genera expected to be present in Lake Victoria, from Brown (1994)).

Environmental conditions of pH, water temperature, microconductivity (μS), total dissolved solutes (TDS), salinity, habitat type, substrate type, depth and turbulence were noted *in situ* for each location and for each sub-designation within a site. Water samples were also collected in 15ml plastic universals for detailed chemical analysis in the laboratory at the NHM. The water samples collected were analysed for anion presence and concentration by ion chromatography and for cation presence and concentration by inductively coupled plasma atomic emission spectroscopy (ICP-AES). Details of the methods of water chemistry analysis can be found in section 2.7.1 of the Methods chapter.

6.4.2 Sample and data processing

At the end of each day, snails were sorted per site and put up for shedding the next day in individual shedding wells filled with 2-3ml of bottled water. After being exposed to sunlight for 1-3 hours, the wells were inspected with a dissecting microscope for emergence of cercariae (see Methods section 2.3).

As the surveys had been carried out in different portions of the lake at different times of year and across a period of nearly three years, seasonal variation needed to be controlled for. As it was hypothesized that rainfall would be the most important seasonal environmental change for freshwater molluscs, rainfall data were extracted for each of the months of the survey as well as the month before and added together to create a rainfall total, which would be a proxy for seasonal change and could be added to the other environmental data. Details of the methods of rainfall extraction can be found in the Methods chapter, section 2.6.1.

6.4.3 Modelling

The full dataset containing GPS coordinates for each site, presence or absence of *B. choanomphala* and *B. sudanica*, number of other species observed at the site and all the water chemistry and environmental variables was analysed using a variety of different statistical models. Three multivariate logistical regression models were compared: the first was a standard, non-spatial multivariate model which did not include random effects, the second was a non-spatial 'exchangeable' model which included random effects and the last was a spatial model which also included random effects. Median values, 95% confidence intervals and odds ratios were calculated; results were significant if the 95% confidence intervals did not span zero. These models were applied separately to presence of *B. choanomphala* and *B. sudanica*. Kriging, using the abundance categories at each stage, interpolated values to the entire Lake Victoria basin to predict areas of high *Biomphalaria* abundance. Details of the models and software used can be found in section 2.7.2 of the Methods.

6.5 Results

In total, 174 sites were visited along the Lake Victoria shoreline, including on many of the islands in the lake. Figures 6.1 and 6.2 display the abundance of *Biomphalaria* at each of the visited sites; the collections from lake and marsh sites are presented separately so as not to obscure results from different habitats present at the same location.

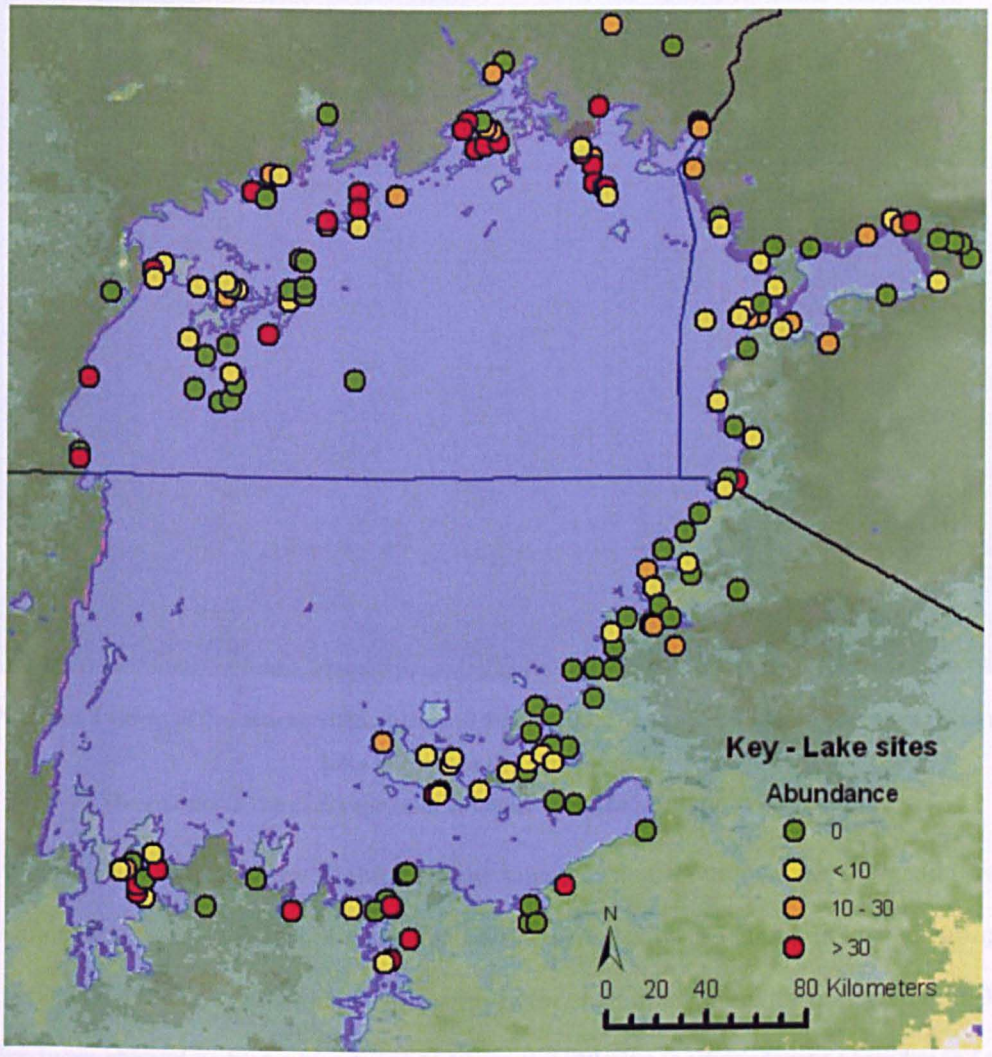

Figure 6.1 – Map of the sites surveyed for *Biomphalaria* snails, for single habitat sites or, if multiple habitats were present, for the lake habitat

The colours of the dots represent abundance of *Biomphalaria* at that location.

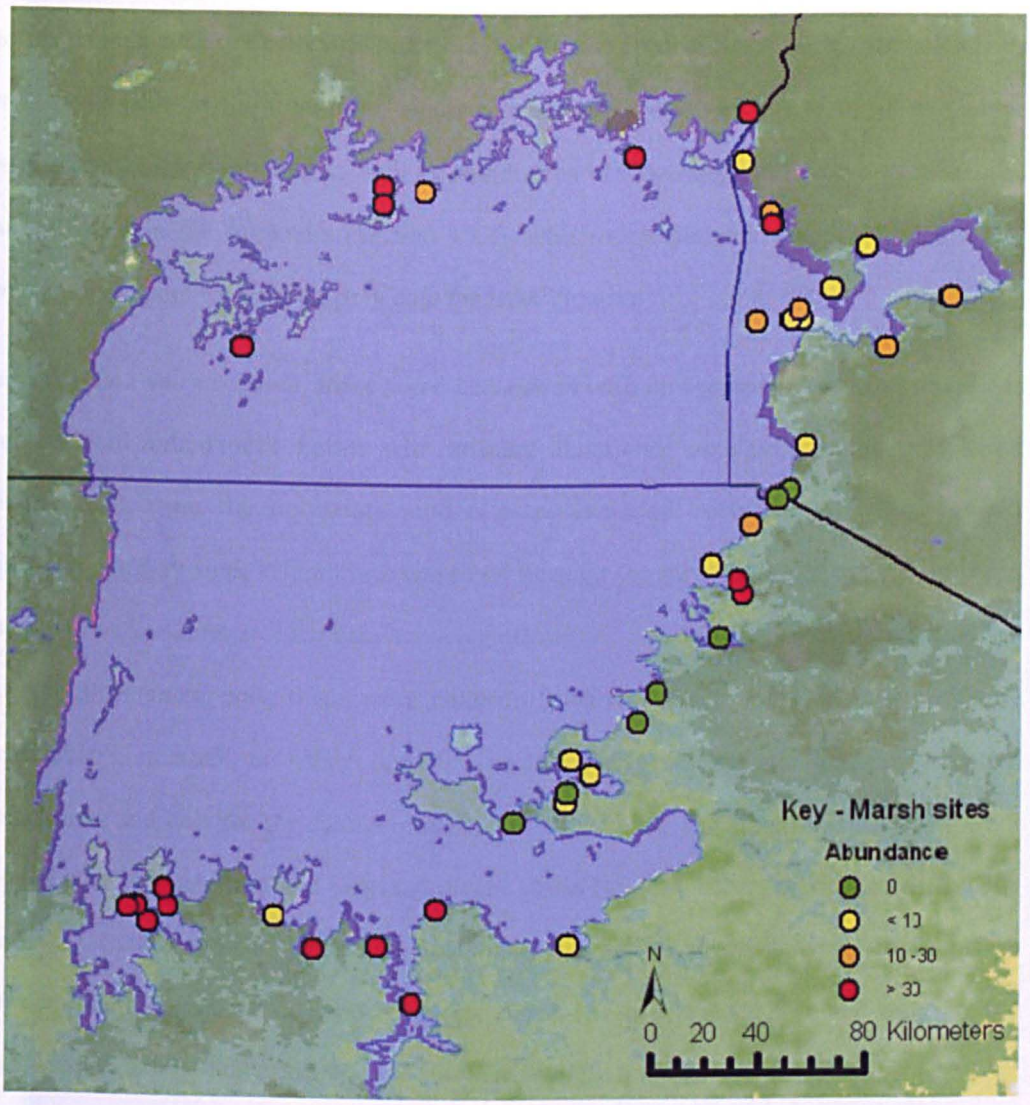

Figure 6.2 - Map of the marsh sites surveyed for *Biomphalaria* snails, where there was also a lake habitat at the same locality

The colours of the dots represent abundance of *Biomphalaria* at that location.

42 sites were surveyed for two habitats at the same location. Four localities in Uganda were re-sampled at a later survey; in addition, 3 water samples were taken mid-lake, creating a total dataset of 223 points which were examined for snails and/or environmental data. 49 of these were classed as high abundance (>30 snails), 30 sites had moderate abundance (10-30 snails), 55 sites had low abundance (<10 snails) and *Biomphalaria* snails were absent from 79 sites (abundance was not recorded at 10 sites). Snails putatively identified as *B. choanomphala* were observed at 61 sites and *B. sudanica* at 57 sampling points; both were observed together in a habitat only 4 times, but were often present in different habitats at the same locality. Visually, small scale heterogeneities in *Biomphalaria* abundance can be seen, particularly between the marsh and the lake habitats that are separated by only a few

metres (compare Figure 6.1 with Figure 6.2), and these patterns were tested statistically as well. A full table of the minimum, median, maximum and mean values of all the water chemistry variables obtained as well as summaries of the categorical variables measured can be found in the Appendix (section 13.2), which also contains a table of all the raw environmental and water chemistry data for Lake Victoria.

Once the data were collated, there were 223 points with environmental measurements as well as snail abundances. Points with missing abundance or significant environmental variable data from the univariate models were excluded, resulting in a final dataset comprising of 149 sites. When these were run through the three different models, for both species, four variables in each case were significant for one or more of the models. In the case of *B. choanomphala*, these were chloride (positive association), sulphate (negative association), number of other species (positive association) and nitrate (negative association, and only for the spatial model). For *B. sudanica*, the significant variables were pH (negative association, and only significant when random effects were included in the model), sulphate (unlike for *B. choanomphala*, this was a positive association), marsh habitat (positive association) and moderate depth (positive association, and only for the non-spatial model). Tables 6.1 and 6.2 summarise these findings.

Factor	Model					
	Non-spatial, no random effects		Non-spatial, with random effects		Spatial, with random effects	
	Median (95% CI)	OR (95% CI)	Median (95% CI)	OR (95% CI)	Median (95% CI)	OR (95% CI)
Chloride	0.200 (0.052, 0.384)	1.221 (1.053, 1.469)	0.188 (0.049, 0.368)	1.207 (1.050, 1.445)	1.850 (0.394, 4.095)	6.360 (1.483, 60.04)
Sulphate	-1.344 (-2.649, -3.95E-01)	0.261 (0.071, 0.674)	-0.978 (-1.920, -0.265)	0.376 (0.147, 0.767)	-11.85 (-23.49, -3.998)	7.14E-06 (6.29E-11, 0.018)
# Other species	1.257 (0.773, 1.888)	3.516 (2.165, 6.610)	0.993 (0.617, 1.490)	2.700 (1.854-4.435)	9.017 (4.657, 15.32)	8.25E+03 1.05E+02, 4.52E+06)
Nitrate	Not sig.	Not sig.	Not sig.	Not sig.	-6.943 (-18.69, -0.1854)	9.65E-04 (7.65E-09, 0.831)

Table 6.1 – Significant environmental variables for predicting presence of *B. choanomphala* in one or more of the models

The 'Median' columns denote the median value of the regression coefficient. 'CI' signify confidence intervals and 'OR' stands for odds ratios. CI that do not overlap zero are significant, as are OR that do not cross 1.

Factor	Model					
	Non-spatial, no random effects		Non-spatial, with random effects		Spatial, with random effects	
	Median (95% CI)	OR (95% CI)	Median (95% CI)	OR (95% CI)	Median (95% CI)	OR (95% CI)
pH	Not sig.	Not sig.	-0.686 (-1.664, -0.222)	0.504 (0.189, 0.801)	-1.124 (-2.320, -0.378)	0.325 (0.098, 0.685)
Sulphate	1.220 (0.628, 2.015)	3.385 (1.873, 7.502)	1.066 (0.510, 2.218)	2.904 (1.666, 9.188)	0.996 (0.362, 1.982)	2.709 (1.436, 7.255)
Marsh habitat	5.256 (2.357, 8.759)	191.3 (10.54, 6395)	1.854 (0.245, 3.375)	6.383 (1.277, 29.22)	1.844 (0.073, 3.518)	6.321 (1.075, 33.71)
Moderate depth	2.436 (0.078, 5.282)	11.45 (1.082, 197.1)	Not sig.	Not sig.	Not sig.	Not sig.

Table 6.2 - Significant environmental variables for predicting presence of *B. sudanica* in one or more of the models

The 'Median' columns denote the median value of the regression coefficient. 'CI' signify confidence intervals and 'OR' stands for odds ratios. CI that do not overlap zero are significant, as are OR that do not cross 1.

The parameters of the data were also recorded for both models, which included random effects ('Tables 6.3 and 6.4, for *B. choanomphala* and *B. sudanica* respectively). These gave an indication as to the non-spatial variance and precision of the model and in the case of the spatial model, the spatial variance and precision as well as the range of the spatial effect.

Other parameters	Non-spatial w random	Spatial w random
	Median (95% CIs)	
sigma	0.751 (0.196, 6.537)	493.9 (122.5, 1890)
tau	1.332 (0.153, 5.092)	0.002 (5.29E-04, 0.008)
rho	NA	0.581 (8.03E-08, 86.30)
sigma.sp	NA	0.6189 (0.1878, 4.375)
tau.sp	NA	1.616 (0.229, 5.325)

Table 6.3 - Regresison coefficient estimates of the two random effects models for *B. choanomphala*
'Sigma' is a measure of the non-spatial variance of the data; 'tau' is calculated simply as 1/sigma and is a measure of precision of the data; 'rho' is the decay parameter which determines the range at which it is reasonable to consider spatial autocorrelation; 'sigma.sp' and 'tau.sp' are the spatial equivalents of sigma and tau.

Other parameters	Non-spatial w random	Spatial w random
	Median (95% CIs)	
sigma	1.894 (0.240, 41.80)	0.571 (0.180, 4.237)
tau	0.528 (0.024, 4.162)	1.752 (0.236, 5.552)
rho	NA	1.899 (0.277, 10.54)
sigma.sp	NA	13.78 (0.902, 121.2)
tau.sp	NA	0.073 (0.008, 1.109)

Table 6.4 - Regression coefficient estimates of the two random effects models for *B. sudanica*
'Sigma' is a measure of the non-spatial variance of the data; 'tau' is calculated simply as 1/sigma and is a measure of precision of the data; 'rho' is the decay parameter which determines the range at which it is reasonable to consider spatial autocorrelation; 'sigma.sp' and 'tau.sp' are the spatial equivalents of sigma and tau.

In this case, for the non-spatial model, the *B. choanomphala* outcome had lower variance and higher precision than the *B. sudanica* outcome. However, for the spatial model, sigma was very large for the *B. choanomphala* outcome with correspondingly low precision whereas it was low for *B. sudanica*; conversely, the spatial variance was higher for the *B. sudanica* outcome than for *B. choanomphala.* The extent at which it is sensible to consider spatial autocorrelation can be calculated by dividing 3 by the rho value; this gives the number of decimal degrees, which as Lake Victoria straddles the equator, can be multiplied

by 110.947 (calculated by averaged the numbers of kilometers per degree of latitude and the kilometers per degree of longitude at the equator) to get the range of spatial autocorrelation in kilometers. Here, the value of rho for *B. choanomphala* was calculated to be 0.581, which corresponds to a range of 572.9 kilometers. Rho for *B. sudanica* was 1.899, resulting in a range of spatial autocorrelation of 175.3 kilometers. It is worth drawing attention to the wide confidence intervals for both rho values, indicating that the range estimates will also have large error margins.

Given the large range of the spatial autocorrelation in the data, kriging could be used as an informal method of extrapolating abundance data to try to predict levels of *Biomphalaria* abundance to unsampled areas of the lakeshore (Figure 6.3). It should be noted that this predictive map does not distinguish between species and so is for *Biomphalaria* as a whole, and does not incorporate environmental data.

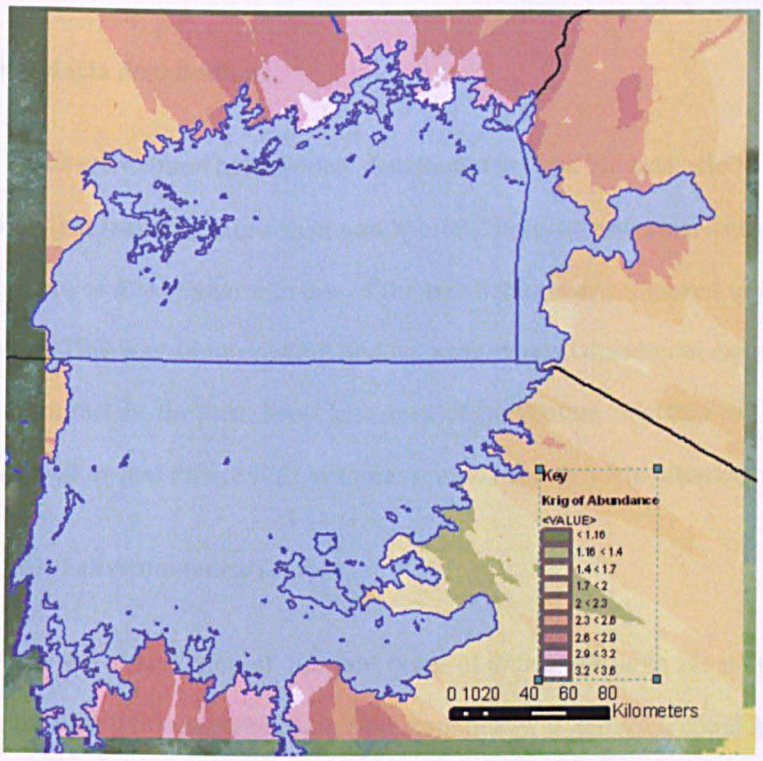

Figure 6.3 – Kriging extrapolation of the abundance data for all *Biomphalaria*
Environmental variables were not included in this analysis; extrapolations are based on known spatial autocorrelation (from the spatial models above) and abundance categories as recorded in the field.

6.6 Discussion

The data here show that not only are *Biomphalaria* snails widely found around the shoreline of Lake Victoria, their distribution is heterogeneous on a local scale. However, as environmental variables also vary widely throughout the lake, some of these conditions can be seen to be predictors of *Biomphalaria* presence or absence, and moreover, the significant factors vary between *Biomphalaria* species. Extrapolating the abundances of the *Biomphalaria* snails found can create basic predictive maps, although future work should focus on including other environmental variables, such as aquatic vegetation type, quantitative measures of wave action and also perhaps investigation into levels of snail predation or competition in different habitat types. These could be used to create more detailed and locally accurate predictions with a view to using the information to identify regions at high risk for intestinal schistosomiasis transmission.

6.6.1 Biomphalaria *distributions*

Biomphalaria snails were found to be widely distributed in Lake Victoria, and in a variety of habitats. At sites that included more than one habitat, there was no clear-cut relationship between abundance of *Biomphalaria* in one of the two habitats as compared to the other in the same location. This was an interesting finding, as previous educational campaigns have targeted water contact in the lake itself as a way of preventing infection with intestinal schistosomiasis (Dalton and Pole, 1978), whereas marsh habitats were often overlooked.

6.6.2 Significant environmental factors

The importance of the observation of different types of *Biomphalaria* in the marsh habitats compared to other habitat types was confirmed statistically. *B. sudanica* could generally be found in marshy habitats set back from the lakeshore (and marsh habitat was a statistically significant predictive variable), whereas *B. choanomphala* was exclusively found in the lake proper. Given the potentially transient nature of marsh habitats across different times of the year and the fact that sampling took place over several field trips, it could be assumed that seasonal variation would contribute to the observed patterns., and indeed, further sampling

at different times of the year would be recommended for future investigations. However, it would appear that rainfall (as a proxy for season) is not predicting the presence of *B. sudanica*, thus implying that seasonality of habitat is not important in the persistence of this species, at least in the context of Lake Victoria.

B. sudanica was also significantly negatively associated with pH, indicating that it is more likely to be found in habitats with lower, more acidic, pH values. Marsh habitats were more likely to have low pH (although the association was not significant), so it would be interesting to ascertain whether there was a causal relationship; for example, whether *B. sudanica* are predicted to live in marshes which happen to have low pH, or whether they in fact thrive in low pH environments, which in the context of Lake Victoria (a highly alkaline environment) these conditions are almost completely confined to marsh habitats. Further, quantified, investigation of the factors that distinguish marshes from lake habitats (for example, turbulence, vegetation levels, oxygen saturation) may be able to distinguish between these hypotheses.

A significant association was seen between presence of *B. choanomphala* and several anions, but of these, only sulphate was significant for *B. sudanica*. Intriguingly, the relationship was negative for *B. choanmphala* but positive for *B. sudanica*, meaning that the former is associated with lower-sulphate habitats than the latter. Elevated sulphate levels in freshwater are most often caused by sulphate-containing fertilizers, atmospheric deposition (including acid rain) and oxidation of pyrite minerals in the soil. Interestingly, this oxidation can be caused by nitrate leaching (Smolders *et al.*, 2010); although weak, there was also a negative association between *B. choanomphala* presence and nitrate concentration, suggesting perhaps that in areas where nitrate leaching is leading to increased sulphate concentrations, *B. choanomphala* is less prevalent. Anthropogenic addition of nitrates and sulphates through fertilizer run-off could increase the concentrations of these ions in Lake Victoria; while this might reduce the presence of lake-dwelling *B. choanomphala*, the positive association seen between sulphate and *B. sudanica* could mean that marsh-forms of *Biomphalaria* would then grow in number, potentially exacerbating transmission risk in marsh habitats and particularly in those close to human influences.

Perhaps the most ecologically relevant variable that was a significant predictor of the presence of *B. choanomphala* was the number of other species also observed at the site. This is contrary to the suggestion in other studies that density dependent effects such as competition might create a negative relationship between intermediate host snails and other species (Woolhouse and Chandiwana, 1990, Mkoji *et al.*, 1998). Moreover, the on-going global loss of biodiversity has prompted a number of studies investigating its effect on disease transmission, and particularly parasitic diseases; however, much of this work has focused on the 'dilution effect' of increased diversity of terminal hosts (Ostfeld and Keesing, 2000, Dobson *et al.*, 2006). It is equally important to look at the effect of biodiversity loss directly on the population dynamics of the intermediate host species, either through competition, predation or indeed another factor. For example, in numerous settings, it has been shown that habitat loss or fragmentation reduces carrying capacity for predator species, resulting in population booms of vector organisms and subsequently, elevated disease transmission (Allan *et al.*, 2003, Ostfeld and Holt, 2004, Carlson *et al.*, 2009).

More directly, the loss of predators, through over-fishing, has led to greater numbers of *Bulinus* in Lake Malawi, with correspondingly higher prevalence of urinary schistosomiasis in the region (Stauffer *et al.*, 2006). Similarly, vector species, as they are often generalists, have been shown to be able to exploit changes in habitat (including anthropogenically modified), potentially through reduced competition with other species that do not adapt as well (Johnson and Thieltges, 2010). As such, it might be expected that *Biomphalaria* in Lake Victoria thrive in locations which are unsuitable for other mollusc species, where they also are not in direct competition. However, this study showed the exact opposite, with *B. choanomphala* more likely to inhabit sites that also had a high number of other species. However, laboratory-based research, on *B. glabrata*, has shown that community diversity can act on the level of cercarial production and incidence of infection (Johnson *et al.*, 2009a); this effect of biodiversity should be explored further in the context of Lake Victoria. The exact composition of the malacological community and the relationship to infection with *S. mansoni* was not tested statistically here, but is likely to be important; this will be explored further in Chapter 9.

There are several variables worth mentioning for their lack of significance in predicting the presence of *Biomphalaria* in this study. One such was calcium ion concentration, despite suggestions of a significant relationship, and indeed a crucial 'minimum' calcium concentration, in previous studies (Alves, 1958, Utzinger *et al.*, 1997). Magnesium and sodium have also been observed to be predictors of *Biomphalaria* presence in previous studies (Schutte and Frank, 1964, Harrison *et al.*, 1966), although neither was significant in any of the multivariate models run here. Finally, rainfall, as mentioned above, was not significant, despite NDVI (as a proxy for water availability) being used successfully in previous attempts to model *Biomphalaria* habitat suitability and distribution. It is possible that the spatial resolution (approximately 8 km) for the rainfall satellite data was too large to be accurate at the level of analysis carried out here. Alternatively, it may be that water availability is insignificant in the context of a permanent, large water body like Lake Victoria, while still important for transient pond and stream ecosystems elsewhere.

6.6.3 Implications of the model parameters

The high values for sigma and/or spatial sigma in the models for both *B. choanomphala* and *B. sudanica* indicate weaknesses in the fit of the models. This is likely due to variation that was not captured by the factors measured in this study, and so opens up the possibility of including further variables in future surveying work. One option could be to integrate more satellite-derived remote sensing alongside on-the-ground measurements; however, too many parameters can create problems, particularly if the dataset is too small to calculate the values of each of the parameters correctly. Alternatively, the number of sites could be increased; however, as mentioned previously, the logistical challenges and resource needs of surveying in remote areas should also be taken into consideration.

The level of autocorrelation suggested by both models was also quite high. As might be expected, the range was lower for *B. sudanica*, which is consistent with being more commonly found in marshes with low, if any, connectivity between sites. Although there is biological rationale for the idea that sites close together will also share ecological characteristics, it is also possible that different variables will be correlated across different

scales. For example, whereas rainfall and water temperature are likely to be similar across a scale of kilometers, substrate and habitat type, for example, may change across a matter of meters. The relationship between terrestrial and aquatic variables could also be important; for example, whereas rainfall might be an 'umbrella' variable, and similar across a wide area, it could influence terrestrial run-off into the lake, which would affect water chemistry values at the scale of local variations in soil composition or fertiliser use. These could be explanations for the wide confidence intervals seen for both rho values; multiple tests, using variables with a hypothesized similar scale of influence, could improve the accuracy of the spatial autocorrelation but would also require making *a priori* assumptions.

6.6.4 *Kriging: Interpolation to undersurveyed regions*

Given the large scale of spatial autocorrelation as determined by the models, the *Biomphalaria* abundance data were kriged, to create a pilot predictive map of the abundance of the snails in Lake Victoria. Ultimately, it would be robust to create further, more formal, predictive maps, using significant environmental variables, within a Bayesian inference framework. Unfortunately, environmental data were not available for sites that were not visited on the malacological surveys, and the only variable for which lake-wide data were available (rainfall) was not a significant predictor of *Biomphalaria* presence and so could not be used. Ideally, if ways were developed to determine other environmental factors, such as substrate, habitat type, turbulence or even water chemistry, perhaps through more sophisticated analysis of spectral patterns, these should be incorporated into future surveying and modeling attempts.

As is, the abundance krig map clearly highlighted three regions as potential hot-spots of *Biomphalaria* abundance (the white areas on Figure 6.3): the southern-most corresponds exactly with Mwanza, a town of more than 500,000. The other two potential high risk areas are in Uganda: one of these is in Mukono district, which also has high population density, and known to have extremely elevated levels of intestinal schistosomiasis in school children (Stothard *et al.*, 2005, Standley *et al.*, 2009). The other hot-spot is in a very remote portion of Mukono district, which has, to the author's knowledge, never been formally surveyed for

Biomphalaria snails; if resources allowed, this should be a priority on future expeditions. Having said this, even if the snails are highly abundant in these regions, it does not automatically mean that prevalence of schistosomiasis will correspondingly be high, although as mentioned this is certainly the case in Mukono district; not only does the transmission of the parasite depend on additional abiotic requirements for development, but there are a myriad of demographical, social and biotic factors which would need to be taken into consideration. Above all, predictions should be validated by ground-truthing, through malacological and parasitological surveys, to avoid errors created by the model of extrapolation, for example when based on too few data points (Standley and Stothard, 2010). In other settings, risk maps based on kriging have proved to be reasonably robust predictors of true snail distributions after *post hoc* surveys had been carried out; however, it is more justified to describe these maps as 'auxiliary tools' in the arsenal against mollusc, and ultimately schistosomiasis, control (Guimaraes *et al.*, 2009).

6.7 Conclusions

This was the first attempt at a Lake Victoria-wide synthesis of *Biomphalaria*, and also the first time Bayesian statistical methods have ever been applied to snail distributions. Overall, it is clear that *Biomphalaria* are widely spread throughout Lake Victoria, but moreover, their distribution is heterogeneous. It also appears that *B. sudanica* and *B. choanomphala* inhabit separate ecological niches, based on the observation that different variables are significant predictors of one or the other. This is an important observation, as it is supporting evidence that what is currently considered separate 'species' may in fact be two ecophenotypes adapted to very different habitat conditions; this will be explored in more detail in the following chapter. Unfortunately, these variables can only be measured *in situ* at this point, making predictive mapping difficult. However, the data used for generating these local-scale models could later be taken up by large-scale modellers as well, thus extending the value of this work from the local to the regional scale. Further surveys should also seek to quantify more variables that might be important to snail ecological dynamics, and in particular,

investigate in more detail the relationship between *Biomphalaria* abundance and biodiversity, as the results here are contrary to what had been expected.

7 Characterisation of *Biomphalaria* (Preston, 1910) in Lake Victoria using molecular and morphological tools: A new appraisal with implications for transmission of intestinal schistosomiasis

7.1 Abstract

Although known to be a hot-spot for intestinal schistosomiasis, the *Biomphalaria* species of the Lake Victoria shoreline of Uganda, Tanzania and Kenya have long been subject to taxonomic confusion. Currently, the accepted view is that two species, *B. choanomphala* and *B. sudanica,* inhabit the lake; both have been implicated in transmission in different parts of the lake, although their relative importance varies by location. Despite this biomedical importance, the relationship and taxonomy of these species, as found in Lake Victoria, has never been explored in detail. Here, seven field-caught groups of snails, from all three countries and consisting of both *B. choanomphala* and *B. sudanica* morphotypes, were analysed using molecular and morphological tools. Two mitochondrial genes (16S and COI) and one nuclear region (ITS) were amplified from approximately 70 snails; the 16S and COI genes were sequenced and the ITS subjected to restriction enzyme digestion. In addition, three sets of morphological measurements were taken: shell measurements, shell aperture outlines and of internal anatomical characters. The data were analysed separately, using trees and networks for the molecular data and principal component analysis (PCA) and analysis of variance tests (ANOVA) for the morphological data, as well as being directly compared to look for correlation. The results suggested that the molecular groupings were not consistent with morphological divisions, which were in fact overlapping, indicating the presence of intermediate morphological forms. Habitat proved to be a significant factor in determining morphotype, suggesting that the different forms of *Biomphalaria* seen in Lake Victoria might be ecophentoypes of one, highly morphologically and molecularly diverse, species, all populations of which are potentially compatible with *S. mansoni.* The nomenclature should be revised to reflect this finding; the names *B. choanomphala* var. *choanomphala* and *B. choanomphala* var. *sudanica* are proposed.

7.2 Contribution of the author

As with the previous chapter, the author was present on all the field missions which resulted in the snail collections used in this research, and furthermore selected the sites to be included in this analysis. The author was also responsible for field identification of the *Biomphalaria* snails collected. All DNA extractions and amplifications were performed by the author; until the summer of 2009, all purifications and sequencing reactions were also carried out by the author. After this time, as with Chapter 5, changes to the standard operating procedure of the Sequencing Facility resulted in it being more efficient and cheaper for the facility to carry out the purifications and sequencing reaction. Sequence editing and analysis as well as analysis of ITS digest data and microsatellite fragments was done by the author. The author carried out all morphological dissections (after instruction by Dr Jean-Pierre Pointier in Perpignan; please see Acknowledgements section) as well as shell measurements, both of the field-collected samples and those from the collections at the Museum für Naturkunde in Berlin. All morphometric analyses were executed by the author, with guidance from Prof Norm MacLeod (see Acknowledgements).

7.3 Introduction

As the intermediate host for the parasite *Schistosoma mansoni*, which causes intestinal schistosomiasis in humans, *Biomphalaria* snails (Basommatophora, Planorbidae) have long been the subject of intense scientific scrutiny. They are widely distributed in tropical and sub-tropical regions across the world, but with the highest number of different species in South and Central America and Africa. Attempts to characterise the different species of this genus have been on-going since the 19th century, but received renewed attention in the latter half of the 20th century, as large scale surveys began to reveal the public health importance of schistosomiasis (Liang *et al.*, 2006, Odongo-Aginya and Ekkehard, 2008) and interest in third world development became part of the international agenda (Thomas, 2000). It was noted early on that different species of *Biomphalaria* appeared to transmit *S.*

mansoni more or less successfully (Files and Cram, 1949, DeWitt, 1954), which encouraged a more detailed investigation into the taxonomy and species characterisation of this genus.

The Neotropics are considered the centre of evolution for *Biomphalaria* (Campbell *et al.*, 2000, DeJong *et al.*, 2001), and to this day contain the most number of species, generally considered to be around 22 (DeJong *et al.*, 2001). There are approximately 12 species in Africa, of which none are known to be completely refractory to infection in the same way that some of the New World species are (Malek, 1985, Brown, 1994). Moreover, the African species have not been awarded the same level of research attention as their American related species, and there has long been confusion as to the status of various putative species and sub-species. Reasons behind this disorder have included problems of sampling, with samples selected from too few or not geographically representative sites; a frequent lack of local expertise has also contributed to the uncertainties of matching species names to specific field-collected samples. The species question is particularly difficult with *Biomphalaria* due to their hermaphroditic status; they can reproduce without copulating, or engage in copulation that does not result in the transfer of sperm. Moreover, there appear to be environmental factors that affect the reproductive behaviour of populations, with some forms acting as one species in a particular area while maintaining separate 'races' in another location (Mandahl-Barth, 1957).

A concerted attempt to rectify this taxonomic tangle was made in the late 1950s by Georg Mandahl-Barth, who used a combination of morphological characters to re-classify African *Biomphalaria*. Based on radula shape, shell measurements and the dimensions of the copulatory organs, Mandahl-Barth proposed classifying the African *Biomphalaria* into four groups: the *B. pfeifferi* group, the *B. choanomphala* group, the *B. sudanica* group and the *B. alexandrina* group, each with numerous sub-'species' (Mandahl-Barth, 1957). Mandahl-Barth explicitly acknowledges that anatomical and morphological differences between African *Biomphalaria* are very small, and in fact could be grouped a single, diverse species, were it not for the observation of distinct, sympatric forms living in apparent reproductive isolation, such as *B. sudanica* and *B. stanleyi* in Lake Albert, or indeed *B. choanomphala* and

B. sudanica in Lake Victoria, although the reproductive habits of these two species have not been studied in detail in this setting (Mandahl-Barth, 1958)

Early collectors of molluscs in Lake Victoria noted an abundance of forms, in keeping with the diversity of *Biomphalaria* in the rest of Africa. Sub-species of *B. sudanica*, such as *B. s. minor* and *B. s. major*, are marked as lectotypes in the collection of the Museum für Naturkunde in Berlin, where the type specimens of *B. sudanica* and *B. choanomphala* are also housed. The Museum also posseses a holotype of a sub-species of *B. choanomphala*, described as *B. choanomphala* var. *basinulacatus*, which is described as being from 'Victoria Nyansa', an antiquated name for Lake Victoria. Later, Mandahl-Barth proposed the presence of three species of *Biomphalaria* in Lake Victoria: *B. sudanica*, *B. rüppellii* and *B. choanomphala*, the last of which could be further divided into two sub-types of *B. c. choanomphala* and *B. c. elegans* (Mandahl-Barth, 1958). Brown (1994) dismisses this separation, and further suggests that *B. rüppellii* is not present in Lake Victoria, thus limiting the number of species of *Biomphalaria* found in the lake to just two: *B. choanomphala* and *B. sudanica.*

More recently, molecular tools have been used to try to elucidate the taxonomy of *Biomphalaria.* DeJong and colleagues (2001) created a molecular phylogeny that completely contradicted Mandahl-Barth's four species groups, by showing that *B. choanomphala*, *B. sudanica* and *B. alexandrina* were actually more closely related to each other than to the other species in their respective 'species groups' as defined by Mandahl-Barth. This cluster of *B. sudanica, B. choanomphala, B. alexandrina* and *B. smithi* was termed the 'Nilotic species complex'. More recent molecular work on East African *Biomphalaria* has supported this clade, although some inconsistencies have been observed, particularly with snails that show a typical *B. pfeifferi* morphology clustering closely with the Nilotic species complex rather than the well-supported *B. pfeifferi* clade, which in some analyses also includes *B. stanleyi* (Jørgensen *et al.*, 2007, Plam *et al.*, 2008). It has been shown that the shell shape of *B. pfeifferi* is affected by environmental and habitat conditions (Dupouy *et al.*, 1993), so this may be responsible for some of the difficulties in making accurate field identifications, both for this species as well as other *Biomphalaria.*

The desire to identify accurately the species of *Biomphalaria* in Lake Victoria is linked directly to the need for greater understanding of the transmission patterns of *S. mansoni* in the lake. Intestinal schistosomiasis is rife in communities living by the lakeshore (see Chapters 3 and 4) and so by understanding more about the dynamics of the snails responsible for transmission it will be possible to identify possible transmission hot-spots, and direct treatment and education interventions to these areas. Currently, in Uganda, *B. choanomphala* is considered the main transmitter of the parasite in Lake Victoria, as no *B. sudanica* have ever been found naturally infected in the Ugandan portion of the lake, and indeed, early laboratory studies suggested they might be refractory to infection with *S. mansoni* (Prentice, 1972). However, transmission by *B. sudanica* is common in the Kenyan and Tanzanian portions of Lake Victoria, suggesting that perhaps the earlier evidence from Uganda is misleading in rejecting the importance of *B. sudanica* as an intermediate host. Moreover, the relationship between the two species as found in the lake has never been investigated in detail, and certainly not since the advent of molecular tools.

This chapter presents research directly aimed at tackling the question of the characterisation of *Biomphalaria* in Lake Victoria, through molecular and morphological examination of seven sites widely distributed around the perimeter of the lake. Specifically, the study aimed to combine molecular and morphological data for a comprehensive revision of the taxonomy of *B. choanomphala* and *B. sudanica*, to determine whether they are indeed separate species. With new evidence as to the relationship and relatedness of these species will come a greater appreciation of the relative role of the various forms of *Biomphalaria* in transmission of *S. mansoni* in Lake Victoria, and will provide policy-makers with best-evidence for developing recommendations for control.

7.4 Methods

Molecular data from two mitochondrial markers (COI and 16S) and one nuclear region (ITS) were amplified from snails from 7 sites around Lake Victoria. The snails were specifically preserved so that morphological analyses could also be undertaken on those same individuals, for a detailed insight into the connection between the genotype and phenotype of the samples.

7.4.1 Sample collection and site selection

7 locations (Figure 7.1) were chosen for inclusion in the taxonomic study on the basis of geography, habitat, abundance of snails and putative field species.

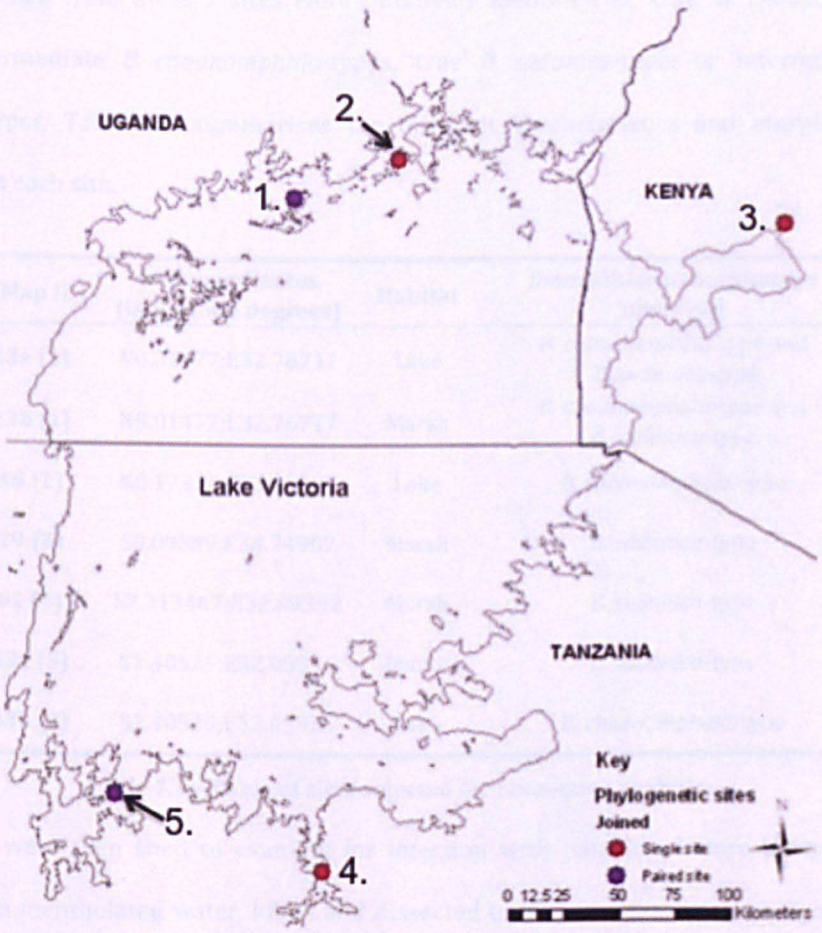

Figure 7.1 – Map of the sites included in the taxonomic and phylogenetic analyses
The 'paired' sites are indicated by a purple circle whereas single sites are marked by a red circle. Paired sites consisted of a marsh habitat separated from the lake proper, whereas single sites only had one habitat that was sampled.

The Uganda sites were surveyed during the Uganda 2008 field mission, the Tanzania sites during the Tanzania 2008 field mission and the Kenya sites during the Tanzania/Kenya 2009 field mission. The sites were intended to be well-spread geograpically, with each country represented; similarly an effort was made to include marsh and lake sites of roughly equal number; this was assisted by the discovery of 'paired' sites, whereby due to the resolution of the GPS device, a single location possessed two non-contiguous habitats, containing separate groups of snails, that were situated so close together that they were indistinguishable by their GPS coordinates. As such, they were grouped as the same geographical location, but considered separate sites. The abundance of snails was important to ensure sufficient sample size of individuals, and finally, 3 of the sites chosen contained *B. sudanica*-type snails, 2 contained *B. choanomphala*-like snails and 2 a mixed group of both. Snails collected from these 7 sites were putatively identified as 'true' *B. choanomphala*-types, 'intermediate' *B. choanomphala*-types, 'true' *B. sudanica*-types or 'intermediate' *B. sudanica*-types. Table 7.1 summarises the different characteristics and morphogroups observed at each site.

Site (Map ID)	GPS coordinates (in decimal degrees)	Habitat	*Biomphalaria* morphotype observed
U023a (1)	N0.01477;E32.76717	Lake	*B. choanomphala*-type and *B. sudanica-type*
U023b (1)	N0.01477;E32.76717	Marsh	*B. choanomphala*-type and *B. sudanica-type*
U046 (2)	N0.17320;E33.18397	Lake	*B. choanomphala*-type
K029 (3)	S0.09589;E34.74907	Marsh	*B. sudanica*-type
T001 (4)	S2.713467;E32.89392	Marsh	*B. sudanica*-type
T033a (5)	S2.40525;E32.05935	Marsh	*B. sudanica*-type
T033b (5)	S2.40525;E32.05935	Lake	*B. choanomphala*-type

Table 7.1 - Table of sites selected for taxonomic analysis

The snails were then shed to examine for infection with parasites before being relaxed overnight in mentholated water, killed and dissected to preserve the whole shell, a portion of tissue in alcohol and the bulk of the body in Raillet's solution (see section 2.3 of the Methods chapter). 10-12 individuals per site were included in both the molecular and morphological analyses.

7.4.2 Molecular methods

Two mitochondrial genes were amplified by PCR then sequenced for phylogenetic analysis between the two morphotypes. In addition, a nuclear region was amplified and used in a restriction enzyme digest to investigate variation between the specimens. Phylogenetic methods used included drawing of networks as well as analysis of trees built using a variety of models and methods.

7.4.2.1 DNA extraction, amplification and sequencing

Genomic DNA was extracted from between 10 and 12 snails from each of the designated 7 sites, using a standard CTAB protocol as described in section 2.4.1 of the Methods (Stothard and Rollinson, 1997). Separate PCRs were performed to amplify fragments of the cytochrome oxidase sub-unit 1 (COI) gene, the 16S ribosomal RNA gene (16S) and the whole internal transcribed spacer region of the nuclear genome, using published primers (Bonnaud *et al.*, 1994, Folmer *et al.*, 1994, Kane and Rollinson, 1994). Details on the primers can be found in Table 2.1 and information on the cycling conditions in section 2.4.3. Amplifications were done using Promega Go-Taq (Promega Corporation, Madison, UK) in a 25µl reaction. All successful amplifications were purified and the COI and 16S fragments sequenced – see section 2.4.3 for details on PCR reaction mix, purification and sequencing. The ITS fragment was also sequenced for a few individuals (specifically, the ITS1 segment was sequenced), but was not very successful and showed potential for having multiple copies in a single individual; in this instance it is difficult to determine whether sequences between individuals are homologous and can be compared. As such, instead of direct sequencing, the entire amplified ITS fragment from all the samples was used in a random fragment length polymorphism restriction enzyme digest, to look for point mutations this way. Each purified PCR product was digested with the restriction enzymes *AluI* and *HaeIII* and run on a 4% agarose gel stained with GelRed™ to compare banding patterns (see section 2.4.4).

7.4.2.2 Alignment of sequences

Sequences were assembled, edited and visually checked against their chromatogram data, before being aligned. Unalignable regions were removed from the alignment. For COI, 655 base pairs were available for analysis, and for 16S, 385 base pairs could be used once gaps had been excised. The programmes used for these processes are detailed in section 2.5.1. analysis.

7.4.2.3 Distance calculations

The mean uncorrected as well as corrected genetic distances were calculated for each of the sites used in the analysis, as well as each of the morphogroups; between site and between morphogroup mean distances were also obtained. The choice of model used to correct the distance calculations is discussed in section 2.5.3. The genetic distance across the whole dataset was also evaluated and compared to that within other *Biomphalaria* species, from sequences retrieved from GenBank (Methods section 2.5.2); mean distances were also calculated between the Lake Victoria dataset and other species, again to ascertain whether levels of genetic distance seen were comparable to intra- or interspecific variation.

7.4.2.4 Tree building

Four different trees were built, using the GTR+gamma model of nucleotide substitution (see section 2.5.3 for method of model choice as well as details on tree building). Four separate methods for determining relationships between taxa were used: distance (neighbour-joining), maximum likelihood, maximum parsimony and Bayesian inference. The last 10% of trees were used for calculation of posterior probabilities on the Bayesian inference tree and node support was calculated by bootstrapping using the other three methods (1000 replicates for each).

In addition, two more trees were created for each of the two sequence datasets. These were used for hypothesis testing, to see whether the resultant tree was more or less likely if monophyly of the morphogroups was enforced (section 2.5.3). The first tree was built with a constraint that all specimens which had been putatively identified as *B. choanomphala*-like

be put together in a monophyletic group; the second one's constraint was that all *B. sudanica*-like snails would be monophyletic. The likelihood values of these trees could then be statistically compared to that of the unconstrained COI and 16S trees, to see if there was statistical evidence of either morphotype being a monophyletic grouping.

7.4.2.5 Networks

The alignments of sequences for both COI and 16S were used to create two minimum spanning distance networks, with a connection limit set to 95%, thus separating out sequences with greater than 5% disparity, which is often considered the level of differentiation between species within a genus (see section 2.5.4). The abundance of each unique haplotype and the proportion of morphotypes (based on putative field identification) at each node was noted and included in the final network for each gene.

7.4.3 Morphological methods

Three separate sets of morphological measurements were taken to compare with the genetic findings. All length measurements were log-transformed prior to further analysis.

7.4.3.1 Shell measurements

Shell height, width, depth below the umbilicus and depth at the highest point were all measured using digital calipers, accurate to 0.01 mm. Photos of each shell were taken using a camera attachment to a light microscope and the number of whorls counted on-screen, at high magnification. Figure 2.1 in the Methods chapter demonstrates where on the shell the measurements were taken, and methods are described fully in section 2.6.1.

In addition to the field collected snails from the sites described in Figure 7.1, a number of type specimens from the Berlin Museum für Naturkunde were also measured and included in the analysis. For comparative purposes, field-collected *B. pfeifferi* (from Zambia) and *B. sudanica* from Lake Albert were also included as putative outgroups.

7.4.3.2 Shell aperture outlines

Photographs of the shells' apertures were also taken, oriented so that the aperture was directly facing the camera, regardless of twist in the shell as a whole. The outline of each aperture was marked out with 50 equidistant points, which were converted into phi functions to describe the shape of the curve. As with the shell measurements, additional shells, from field sites in Zambia and Lake Albert and type specimens from Berlin, were included in the analysis. Details of this method can be found in section 2.6.2.

7.4.3.3 Copulatory organs

Using specimens that had been removed from their shells in the field and their bodies placed in Raillet's solution, further dissections could be made in order to compare internal anatomy. Samples were dissected under a light microscope to reveal the copulatory structures, and in particular, the penis sheath/preputium and the prostate gland. The lengths of the penis sheath and the preputium were noted using a camera lucida microscope attachment, and the prostate dissected further to count the number of diverticula. Details of these procedures can be found in section 2.6.3.

7.4.4 Principal component analysis, linear discriminant analysis and ANOVA

Principal component analysis (PCA) was performed on all three morphological datasets (see section 2.6). Linear discriminant analysis (LDA) was also executed, with putative field identification as either '*B. choanomphala*', '*B.choanomphala*-intermediate', '*B. sudanica*' or '*B. sudanica*-intermediate' (plus *B. pfeifferi*, in the case of the field-caught Zambian samples) as the classifier (see also section 2.6). ANOVAs were run to test directly for significance between various measurements and the putative species identifications as well as the habitat type of each site. In this case, to control for size, ratios were calculated between shell measurements before use in the ANOVA. These were: height/width, depth1/depth2 and depth 2/width. Raw, unlogged measurements were used for ANOVAs with the copulatory organ measurements, but the ratio of the penis sheath/preputium also used, together with number of diverticula as well. Tukey's Honestly Significant Difference (HSD) test was used to test statistically between the means of the various groupings, also in R (Hayter, 1984). As

there were no discrete variables to compare, aperture outlines were not included in the ANOVA testing.

7.4.5 Matrix correlations

In order to look for correlation between the various different types of data and even the different types of measurement, distance matrices were constructed and then statistically tested for correlation. Because genetic data were only available for the seven Lake Victoria populatons under investigation, the other field-collected samples as well as the specimens from Berlin were excluded from this part of the analysis.

For each putative morphological group ('*B. choanomphala*', '*B.choanomphala*-intermediate', '*B. sudanica*' or '*B. sudanica*-intermediate'), the mean eigenvalue score for the first two principal components was calculated. This was done by averaging the PCA coordinates from all of the individuals within a morphogroup, and was done separately for each morphological analysis. Pair-wise, Euclidean distances between each of the means of the groups was then determined, in Excel, creating a matrix, which could be compared in a pair-wise fashion to a similar matrix of genetic distance data, using a Mantel r-test (see section 2.4 of the Methods chapter).

7.5 Results

B. sudanica-like snails were observed mostly at marsh sites whereas *B. choanomphala*-like snails were collected almost exclusively from the lake proper. Molecular findings did not separate the two morphogroups; morphometric analysis did result in rough groups but with substantial overlap and individuals classified differently depending on the type of measurement made.

7.5.1 Molecular results

Low levels of variation were seen in the ITS banding patterns. Of the 655 base pairs of the COI gene used in the analysis, 135 positions were variable and 109 were parsimony-informative. 36 unique COI haplotypes were recovered from 75 samples. For 16S, 88 sites were variable and 76 were parsimony-informative. 23 unique 16S haplotypes were observed from 71 sequences. All haplotypes were compared to the existing database of haplotypes as calculated for the entire set of Lake Victoria *Biomphalaria* sequences (see the following chapter) and matched accordingly. A full list of haplotypes, with GenBank accession numbers, for both COI and 16S can be found in Chapter 8.

7.5.1.1 Restriction enzyme digest

The ITS restriction enzyme digest revealed only small variations in banding pattern (Figures 7.2 and 7.3).

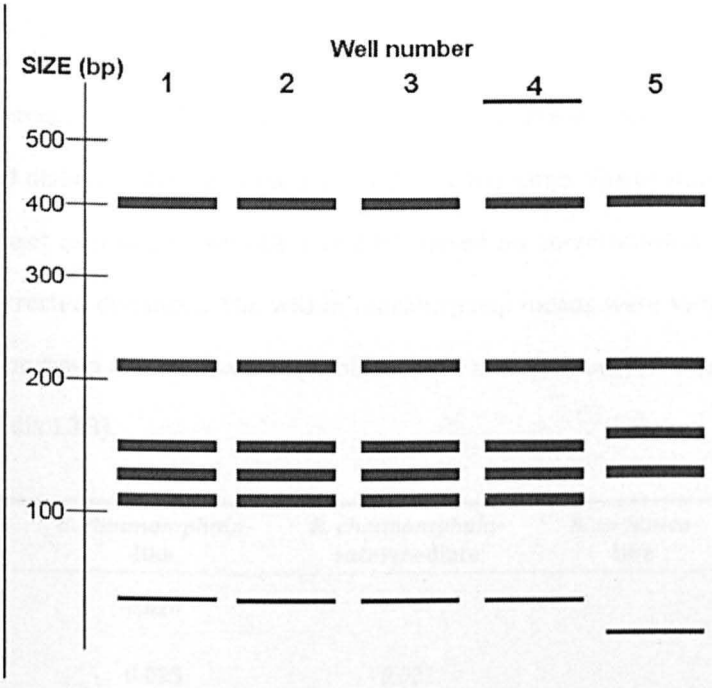

Figure 7.2 – Schematic representation of the gel image of the ITS *Hae*III digest

Lane 1 represents digestion of the ITS fragment amplified from a *B. choanomphala* snail, Lane 2 from a *B. sudanica*, Lanes 3 and 4 from intermediate forms and Lane 5 from an Argentinian *B. peregrina*, as a putative outgroup. Note that Lane 5 is the only one showing clear distinction from the others.

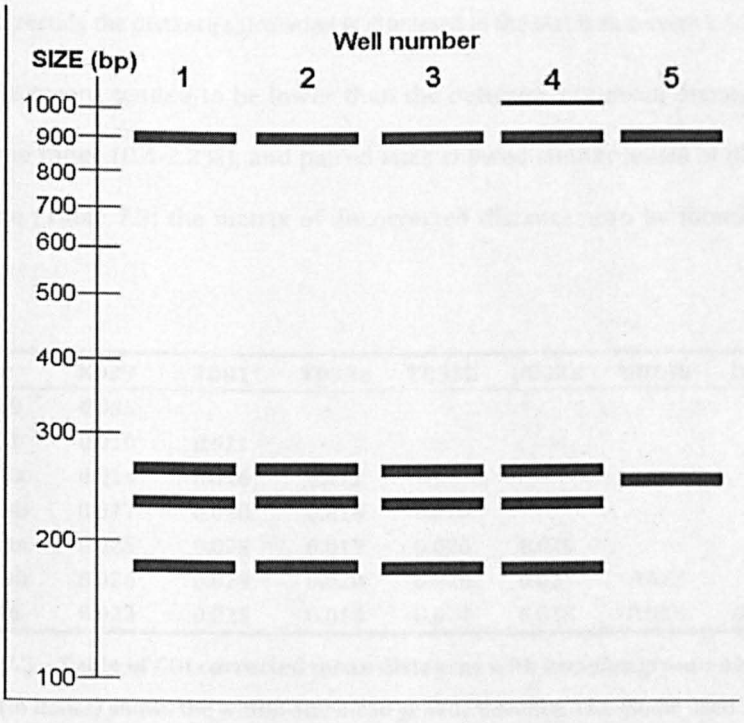

Figure 7.3 – Schematic representation of the gel image of the ITS *Alu*I digest

Lane 1 represents digestion of the ITS fragment amplified from a *B. choanomphala* snail, Lane 2 from a *B. sudanica*, Lanes 3 and 4 from intermediate forms and Lane 5 from an Argentinian *B. peregrina*, as a putative outgroup. Note that Lane 5 is the only one showing clear distinction from the others.

7.5.1.2 *Genetic distances*

For the COI sequences, the corrected genetic distances were greater across the samples than the uncorrected distances, although the patterns were the same. The overall mean distance across the dataset as a whole for COI was 2.0% based on corrected distances and 1.6% based on uncorrected distances. The within-morphogroup means were very similar to the between-morphogroup mean distances (Table 7.2; the matrix of uncorrected values can be found in Appendix 13.3).

Morphogroup	*B. choanomphala*-like	*B. choanomphala*-intermediate	*B. sudanica*-like	*B. sudanica*-intermediate
***B. choanomphala*-like**	*0.020*			
***B. choanomphala*-intermediate**	0.023	*0.022*		
***B. sudanica*-like**	0.023	0.026	*0.020*	
***B. sudanica*-intermediate**	0.018	0.021	0.018	*0.015*

Table 7.2 – Table of COI corrected mean distances within samples grouped by morphogroup
The diagonal (in italics) shows the within-morphogroup mean genetic distance. The model used for correcting the distance calculation is discussed in the Methods section 2.5.3.

The within-site means tended to be lower than the between-site mean distances, although within the same range (0.4-2.2%), and paired sites showed similar levels of distance to the overall average (Table 7.3; the matrix of uncorrected distances can be found in Appendix 13.3).

Site	K029	T001	T033a	T033b	U023a	U023b	U046
K029	*0.005*						
T001	0.010	*0.011*					
T033a	0.014	0.016	*0.012*				
T033b	0.017	0.020	0.018	*0.018*			
U023a	0.025	0.028	0.019	0.026	*0.020*		
U023b	0.026	0.029	0.020	0.028	0.021	*0.022*	
U046	0.023	0.025	0.016	0.024	0.018	0.018	*0.013*

Table 7.3 – Table of COI corrected mean distances with samples grouped by site
The diagonal (in italics) shows the within-site mean genetic distance. The model used for correcting the distance calculation is discussed in the Methods section 2.5.3.

For 16S, there were very few variations between the corrected and the uncorrected distances. Overall uncorrected distance for 16S across the whole dataset was 0.7% for both

corrected and uncorrected distances; within-morphogroup distances were again very similar to between-morphogroup distances (Table 7.4; the matrix of uncorrected distances can be found in Appendix 13.3)

Morphogroup	*B. choanomphala*-like	*B. choanomphala*-intermediate	*B. sudanica*-like	*B. sudanica*-intermediate
***B. choanomphala*-like**	*0.007*			
***B. choanomphala*-intermediate**	0.008	*0.007*		
***B. sudanica*-like**	0.008	0.008	*0.008*	
***B. sudanica*-intermediate**	0.007	0.006	0.007	*0.005*

Table 7.4 – Table of 16S corrected mean distances with samples grouped by morphogroup
The diagonal (in italics) shows the within-morphogroup mean genetic distance. The model used for correcting the distance calculation is discussed in the Methods section 2.5.3.

As with COI, although within-site mean distances were generally lower than the between-site means, the values were all within the same range (of 0.4-1.0%); again, as with COI, the paired sites had mean genetic distances which were the same as between non-paired sites (see Table 7.5. The table of uncorrected distances can be found in Appendix 13.3).

Sites	K029	T001	T033a	T033b	U023a	U023b	U046
K029	*0.003*						
T001	0.008	*0.004*					
T033a	0.007	0.005	*0.004*				
T033b	0.008	0.005	0.005	*0.005*			
U023a	0.010	0.009	0.007	0.008	*0.008*		
U023b	0.009	0.008	0.006	0.007	0.007	*0.007*	
U046	0.010	0.009	0.008	0.008	0.009	0.008	*0.008*

Table 7.5 – Table of 16S corrected mean distances with samples grouped by site
The diagonal (in italics) shows the within-site mean genetic distance. The model used for correcting the distance calculation is discussed in the Methods section 2.5.3.

The average genetic distance within the samples was comparable to that within other *Biomphalaria* species. For COI, there was 1.6% difference across the Lake Victorian samples, as compared to 1.8% between African-wide samples of *B. pfeifferi* and 3.4% difference between three sequences of *B. glabrata*, from South America as well as laboratory strains. For the 16S data, the values were 0.7% for the Lake Victorian samples, 0% for *B. pfeifferi* (sequences from different individuals were used) and also 0% for *B. glabrata*. *B. peregrina* and *B. tenagophila* both had 16S distances of 0.8% for this fragment of the 16S gene.

Distances between the Lake Victorian samples and other species were also comparable to inter-specific distances between other *Biomphalaria* species. Difference compared to *B. alexandrina* was lowest, at 1.7% for both markers, but otherwise for the other African *Biomphalaria* it ranged from 2.9% to 3.7%. South American species, excluding *B. glabrata*, were between 4.8% and 6.8% different.

7.5.1.3 Trees

The phylogenetic trees show similar results of genetic differentiation of certain groups, but that do not conform to the putative species identifications from the field. In the COI tree, the samples are split into two main groups, with both *B. choanomphala*-like and *B. sudanica*-like morphotypes in both groups. Some of the more derived branching, towards the end of the clades, separated samples into their respective sites (see Figure 7.4).

The 16S tree was less structured than the COI tree, but again showed cross-over of morphotypes within clades (Figure 7.5).

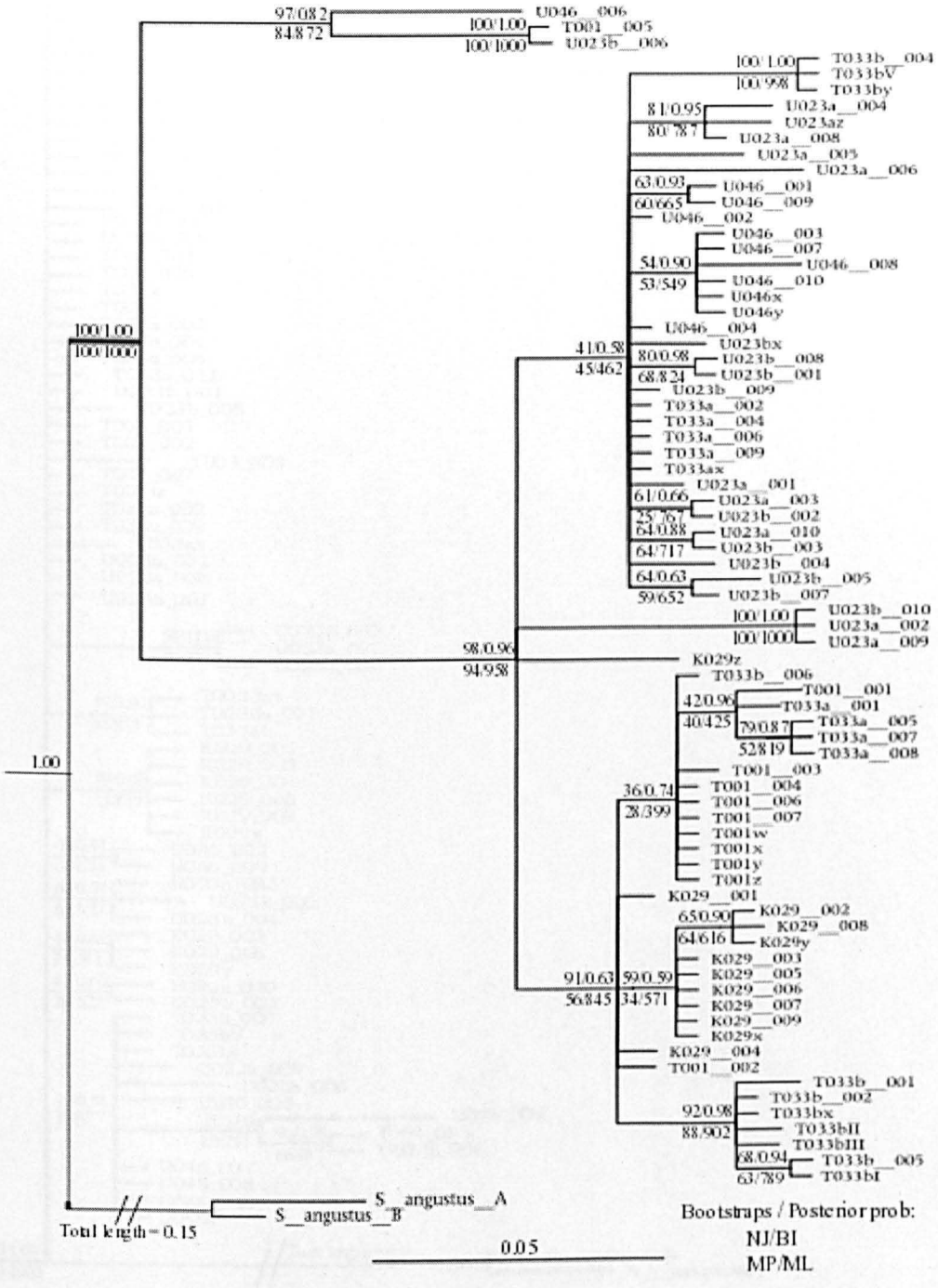

Figure 7.4 – Phylogenetic tree of COI sequence data

The branch lengths and general topology are based on the Bayesian inference (BI) tree; node support values are those that were >50% based on the Bayesian posterior probabilities. Neighbour-joining (NJ), maximum likelihood (ML) and maximum parsimony (MP) bootstrap support values have also been added, as per the legend, with NJ to the top left, BI to the top right, MP to the bottom left and ML to the bottom right of each node. Two sequences of *Segmentorbis angustus*, also from Lake Victoria, were used as the outgroup (designated S_angustus_A and S_angustus_B).

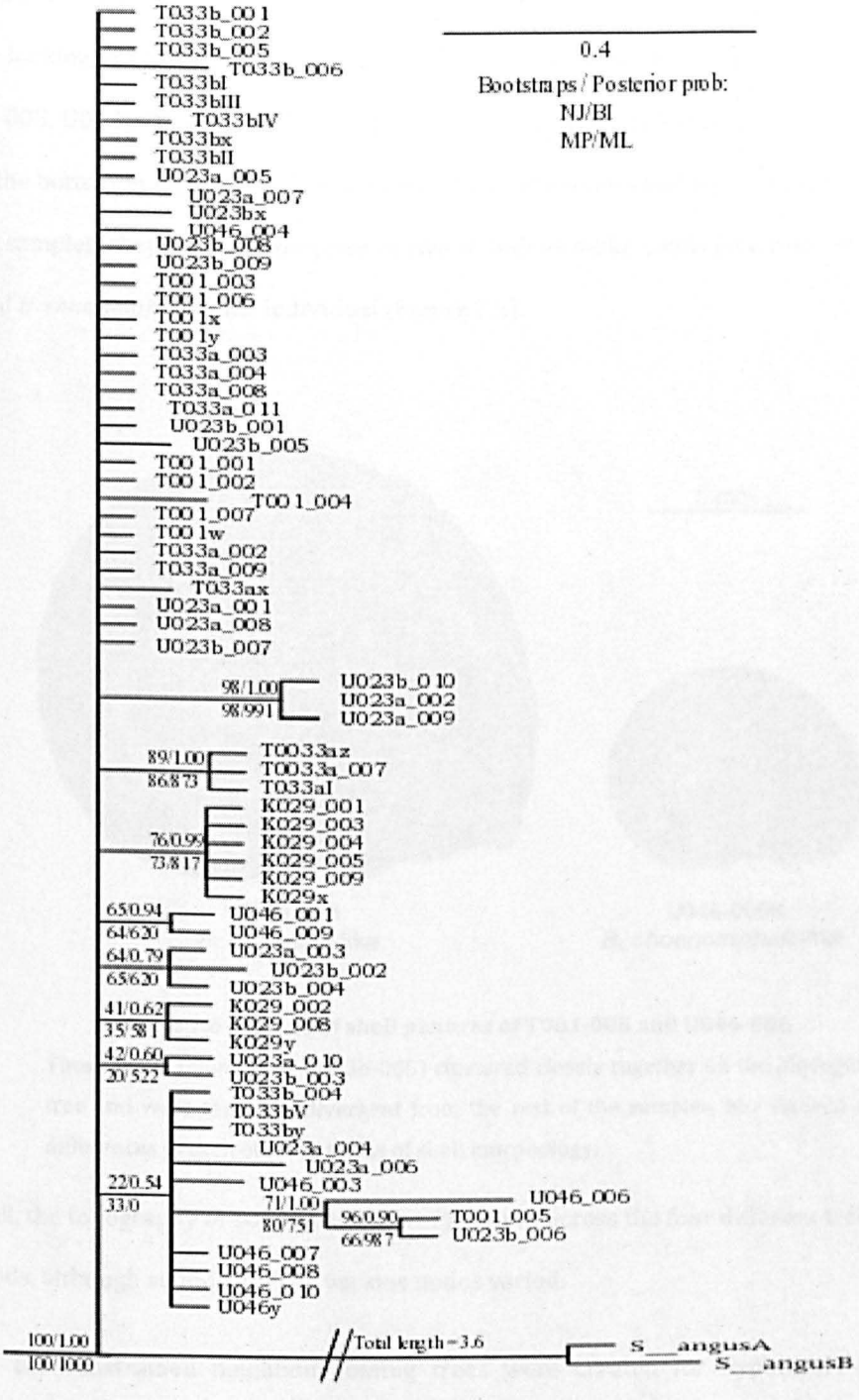

Figure 7.5 – Phylogenetic tree of 16S sequence data

The branch lengths and general topology are based on the Bayesian inference tree; node support values are those that were >50% based on the Bayesian posterior probabilities. Neighbour-joining (NJ), maximum likelihood (ML) and maximum parsimony (MP) bootstrap support values have also been added, as per the legend, with NJ to the top left, BI to the top right, MP to the bottom left and ML to the bottom right of each node. Two sequences of *Segmentorbis angustus*, also from Lake Victoria, were used as the outgroup (designated S_angusA and S_angusB).

For both trees, the example of cross-over of morphotypes within clades is clearly shown when looking at the very well-supported, well-differentiated clade consisting of sequences T001-005, U023b-005 and U046-006 (the top-most branch in Figure 7.4 and in the centre near the bottom in Figure 7.5). Despite their close relatedness and high divergence from the other samples, they actually comprise of two *B. sudanica*-like snails (the first two) and one typical *B. choanomphala*-like individual (Figure 7.6).

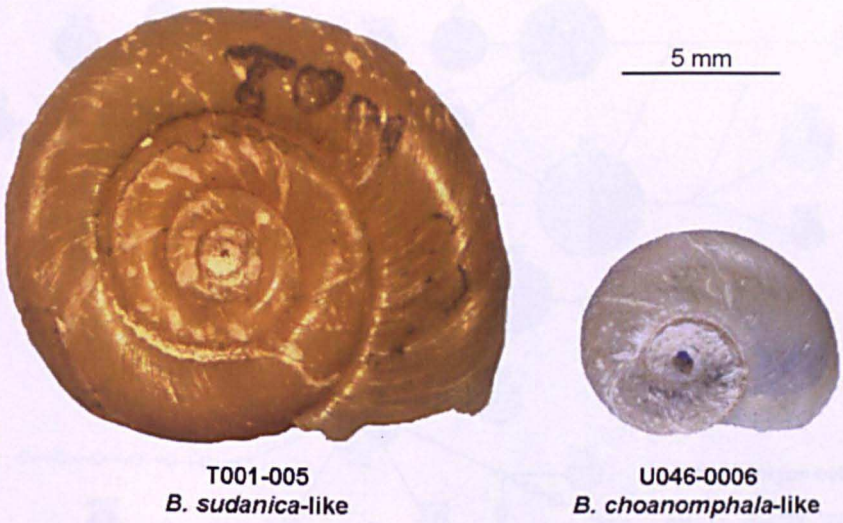

Figure 7.6 – Image of shell pictures of T001-005 and U046-006
These snails (along with U023b-005) clustered closely together on the phylogenetic tree and were the most divergent from the rest of the samples, but showed clear differences to each other in terms of shell morphology.

Overall, the topography of both trees was very similar across the four different tree-building methods, although support for the various nodes varied.

When the constrained neighbour-joining trees were created for hypothesis testing of monophyly, the trees with enforced monophyly for *B. sudanica*-type snails were significantly less likely than the unconstrained trees, for both genes (for COI, $p < 0.001$; for 16S, $p = 0.002$, for both the KH test and the SH test). The *B. choanomphala*-type constraint for monophyly was also significantly less likely for the COI data ($p < 0.001$ for both tests) and less likely, but not significantly so for the 16S data (KH test: $p = 0.071$; SH test: $p = 0.05$). These results statistically support the presence of mixed morphotype clades that had been visually identified on the trees.

7.5.1.4 *Networks*

The COI network showed two main groups, with one clustering around H1 and H8 and the other around H32 (Figure 7.7). However, both groups have snails of *B. choanomphala* and *B. sudanica* morphotypes. In fact, in several cases, unique haplotypes were even shared by different morphotypes. The network also shows that abundant haplotypes tend to be central nodes with less common haplotypes radiating out from these abundant forms.

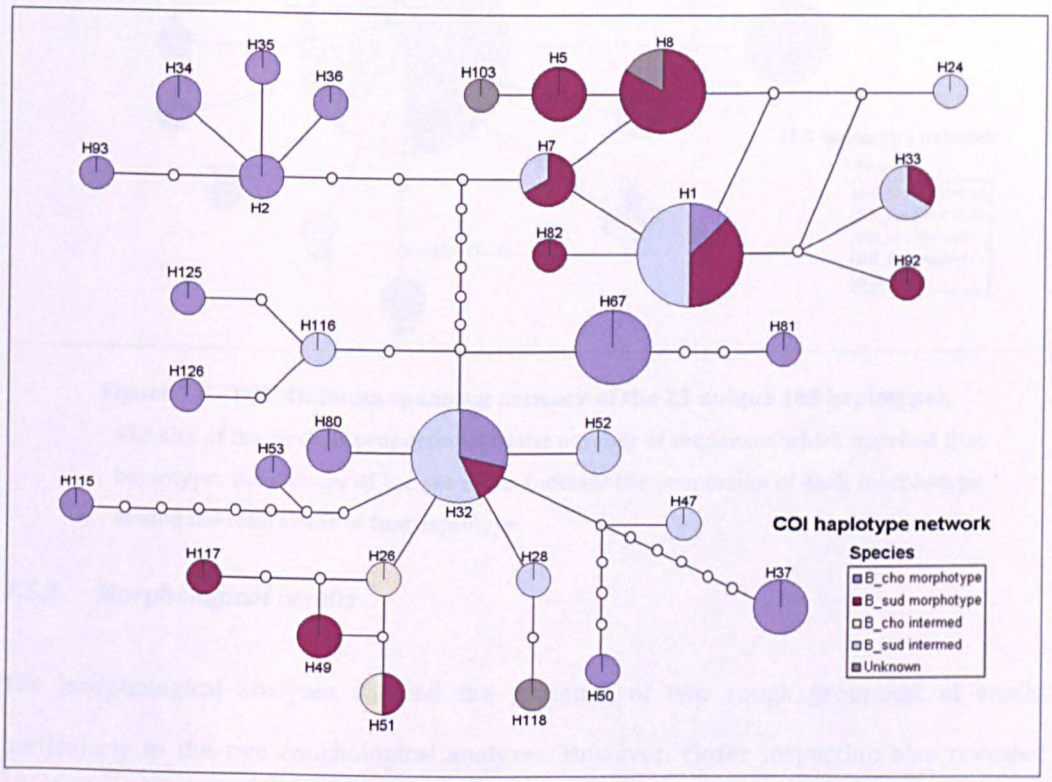

Figure 7.7 – TCS Minimum spanning network of the 37 unique COI haplotypes.
The size of the circle is proportional to the number of sequences which matched that haplotype; the colours of the pie slices indicate the proportion of each morphotype among the total snails of that haplotype.

The minimum spanning network for 16S again shows abundant haplotypes as being the central nodes of radiation towards rarer haplotypes, although there is less evidence of two separate groups as compared to the COI data (Figure 7.8).

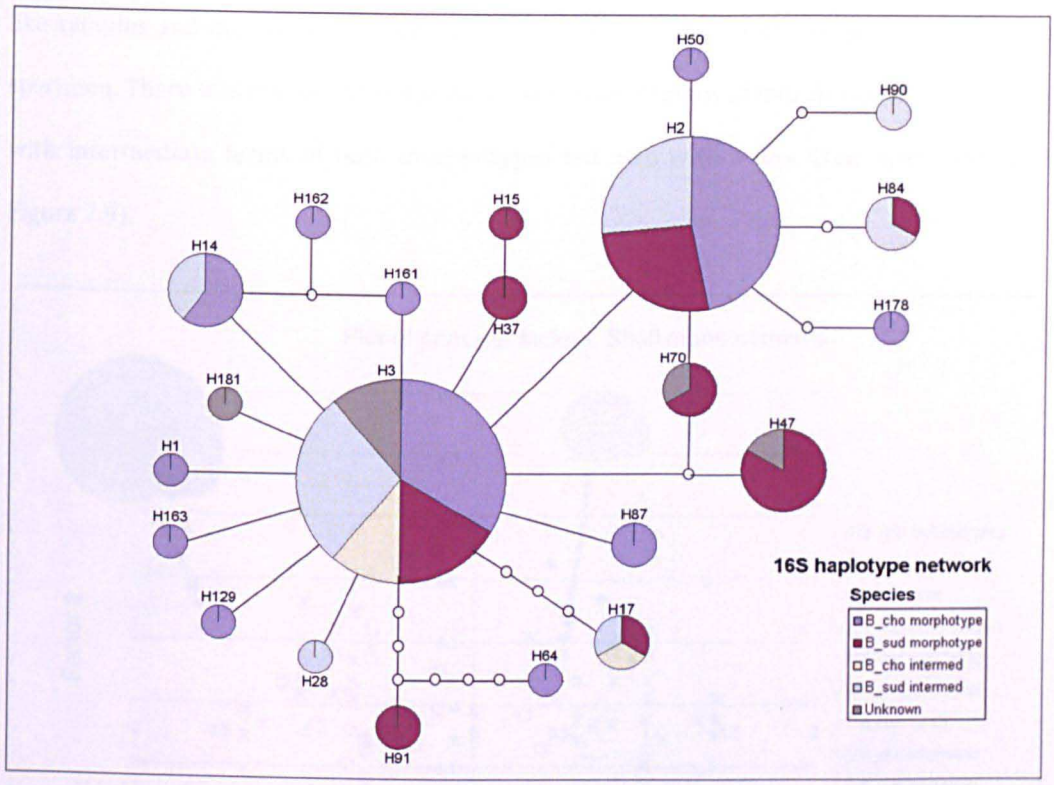

Figure 7.8 - TCS Minimum spanning network of the 23 unique 16S haplotypes. The size of the circle is proportional to the number of sequences which matched that haplotype; the colours of the pie slices indicate the proportion of each morphotype among the total snails of that haplotype.

7.5.2 Morphological results

The morphological analyses showed the presence of two rough groupings of snails, particularly in the two conchological analyses. However, closer inspection also revealed overlaps between the groupings, and particularly with the intermediate forms, whereby individuals were classified into different morphotypes from one analysis to the next. Tables of the mean measurements, plus standard deviation and variance, can be found in Appendix 13.3.

7.5.2.1 PCA and LDA

The PCA of the shell measurements showed a rough division between the *B. choanomphala*-types and the *B. sudanica*-types, with the *B. pfeifferi* samples clustered within the *B. choanomphala* morphospace, apart from one outlier. The 'type' material from Berlin clustered centrally within the rough groupings of morphospecies, apart from one *B. sudanica* var. *minor* type specimen, which was quite distant from the rest of the *B. sudanica*-

like samples and moreover was very close to the *B. choanomphala* var. *basinulacatus* type specimen. There was cross-over between the two main regions of morphospace, particularly with intermediate forms of both morphotypes but also with a few 'true' specimens (see Figure 7.9).

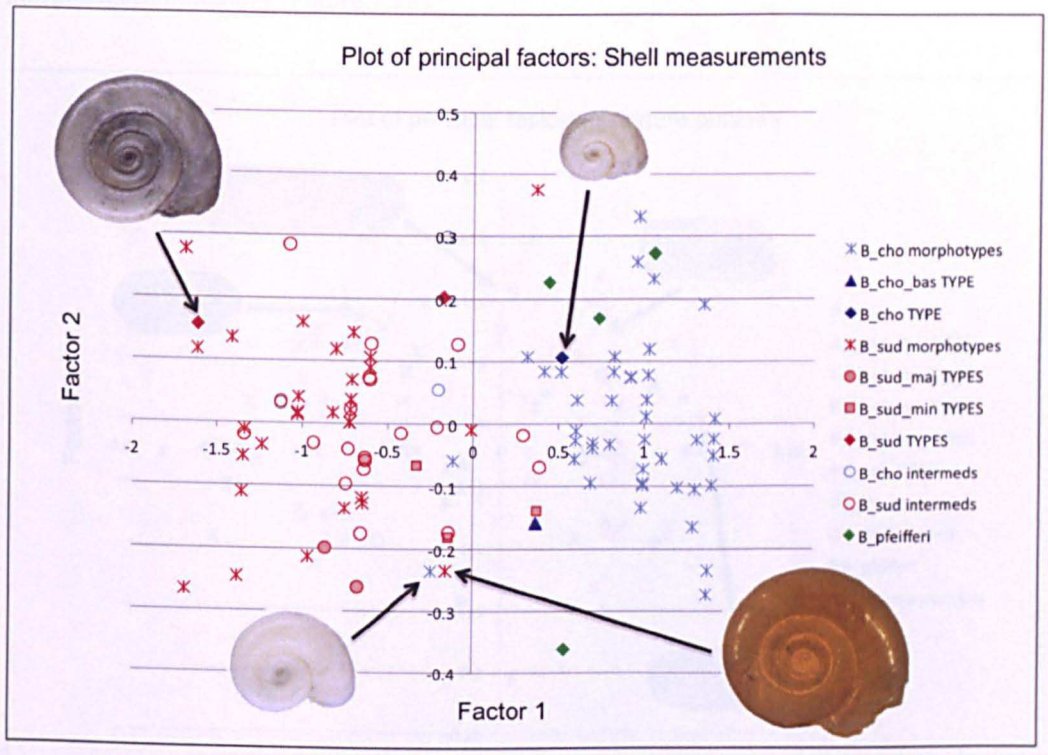

Figure 7.9 – PCA plot showing the first two eigenvalues ('Factor 1' and 'Factor 2') for the shell measurement analysis

These two factors explained 99.8% of the variation in the data. The shell pictures demonstrate the visual form of the shell corresponding to that particular coordinate on the plot. The various different morpho-groups (Lake Victoria samples versus 'type' material from Berlin versus field-caught *B. pfeifferi*) are indicated as per the key. The abbreviations for the species names are as follows: 'B_cho' = *B. choanomphala*; 'B_cho_bas' = *B. choanomphala* var. *basinulacatus* 'B_sud' = *B. sudanica*; 'B_sud_maj' = *B. sudanica* var. *major*; 'B_sud_min = *B. sudanica* var. *minor*; 'intermeds' = intermediate forms.

Note the proximity of morphospace shared by the bottom two shell pictures; although grouped closely together, visually the shells appear very different and clearly characteristic of a *B. choanomphala*-like (on the left) and a *B. sudanica*-like snail (on the right).

The aperture outline PCA plot showed a similar picture, with rough groupings blurred by cross-over, particularly of intermediate forms. It is worth pointing out here that the type specimens of *B. choanomphala* and *B. sudanica* are towards the outer edges of their

respective groups, indicating perhaps they are extreme examples of this aspect of shell morphology. In contrast to this, the *B. choanomphala* var. *basinulacatus* type specimen is found well within the *B. sudanica*-morphospace. The *B. pfeifferi* specimens here are not as clustered as for the shell measurements, but again span the *B. choanomphala/B. sudanica* morphospace boundary (Figure 7.10).

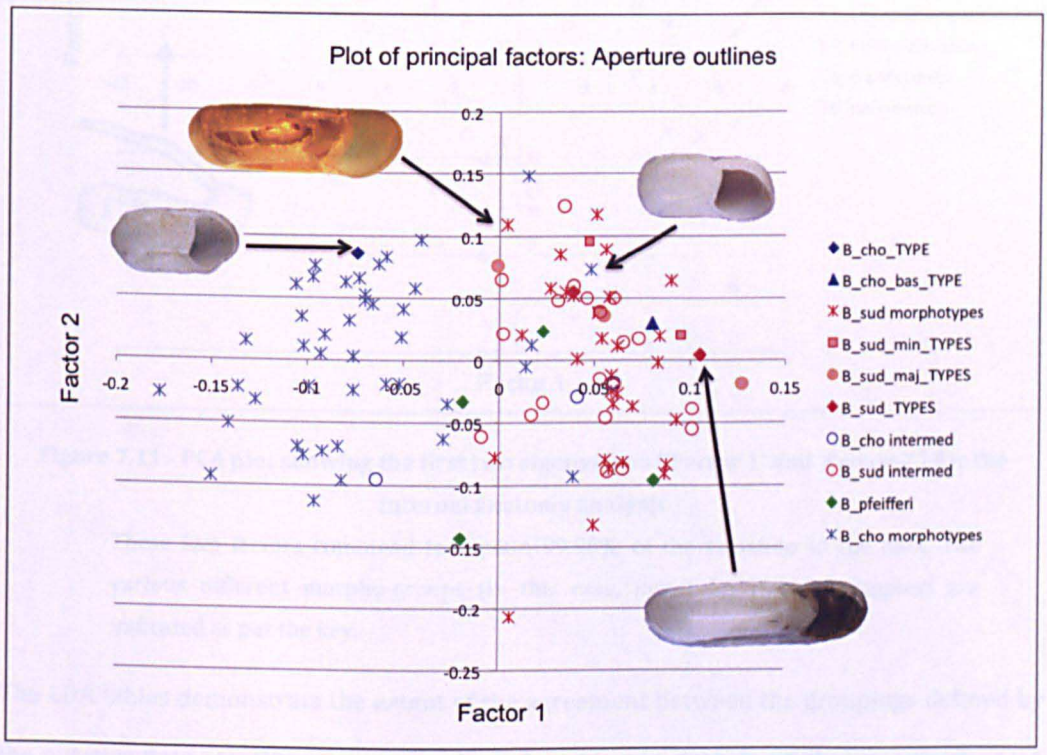

Figure 7.10 - PCA plot showing the first two eigenvalues ('Factor 1' and 'Factor 2') for the aperture outline analysis

These two factors account for 61.45% of the variation in the data. The shell pictures demonstrate the visual form of the aperture corresponding to that particular coordinate on the plot. The various different morpho-groups (Lake Victoria samples versus 'type' material from Berlin versus field-caught *B. pfeifferi*) are indicated as per the key. The abbreviations for the species names are as follows: 'B_cho' = *B. choanomphala*; 'B_cho_bas' = *B. choanomphala* var. *basinulacatus* 'B_sud' = *B. sudanica*; 'B_sud_maj' = *B. sudanica* var. *major*; 'B_sud_min = *B. sudanica* var. *minor*; 'intermeds' = intermediate forms.

The internal anatomy PCA plot showed the least differentiation between the morphotypes, with extensive cross-over of both intermediate and 'true' specimens. However, the most extreme examples of cross-over, for both *B. choanomphala* and *B. sudanica*, are intermediate forms, apart from one 'true' *B. sudanica* sample (Figure 7.11).

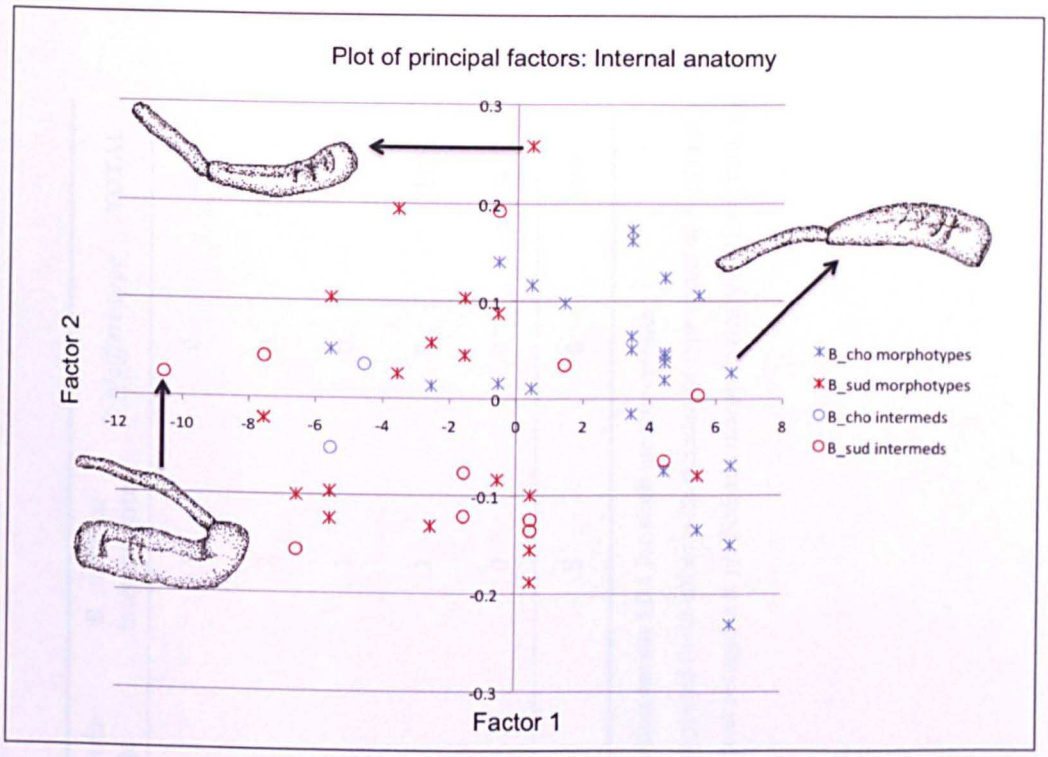

Figure 7.11 - PCA plot showing the first two eigenvalues ('Factor 1' and 'Factor 2') for the internal anatomy analysis

These two factors combined to explain 99.98% of the variation in the data. The various different morpho-groups (in this case, just Lake Victoria samples) are indicated as per the key.

The LDA tables demonstrate the extent of the agreement between the groupings defined by the putative field identification versus those found by the PCA. In all three cases, the LDA analysis removed *B. choanomphala*-intermediates as a separate group, placing them either with the *B. choanomphala*-types or one of the two *B. sudanica* morphogroups. For the shell measurement LDA, *B. pfeifferi* was also not considered a separate group, with those individuals instead being classified as *B. choanomphala*-type (Table 7.6). Overall, fewer individuals were classified as intermediates based on the LDA than had been defined in the field, but there was cross-over between the morphotypes. For example, an individual classified as being part of one morphogroup in shell measurements could later be grouped into a different morphogroup on the basis of aperture analysis, indicating difficulties in precise identification (Table 7.7).

Method of ID		LDA re-grouping					
	Morpho-groups	*B. choanomphala*-type	*B. sudanica*-type	*B. choanomphala*-intermediate	*B. sudanica*-intermediate	*B. pfeifferi*-type	**TOTAL**
Putative field ID	*B. choanomphala*-type	43	1	0	1	0	**45**
	B. sudanica-type	2	34	0	2	0	**38**
	B. choanomphala-intermediate	0	2	0	1	0	**3**
	B. sudanica-intermediate	2	16	0	1	0	**19**
	B. pfeifferi-type	4	0	0	0	0	**4**
	TOTAL	**51**	**53**	**0**	**5**	**0**	**109**

Table 7.6 – Comparison of the morpho-groups as determined by field identification versus LDA for shell measurements

The LDA analysis removed all individuals from the *B. choanomphala*-intermediate morphogroup, but placed them either with *B. sudanica*-type snails or *B. sudanica*-intermediates. All of the *B. pfeifferi* snails were classified as *B. choanomphala*-type snails; the LDA did not recognize *B. pfeifferi* as a morphologically distinct grouping.

Method of ID		LDA re-grouping					
	Morpho-groups	*B. choanomphala*-type	*B. sudanica*-type	*B. choanomphala*-intermediate	*B. sudanica*-intermediate	*B. pfeifferi*-type	**TOTAL**
Putative field ID	*B. choanomphala*-type	39	6	0	0	0	**45**
	B. sudanica-type	1	36	0	0	1	**38**
	B. choanomphala-intermediate	1	2	0	0	0	**3**
	B. sudanica-intermediate	0	20	0	0	0	**20**
	B. pfeifferi-type	1	2	0	0	1	**4**
	TOTAL	**42**	**66**	**0**	**0**	**2**	**110**

Table 7.7 - Comparison of the morpho-groups as determined by field identification versus LDA for aperture outline analysis

The LDA analysis removed all individuals from the *B. choanomphala*-intermediate and the *B. sudanica*-intermediate morphogroups; in this case all the *B. sudanica*-intermediates were re-classified by the analysis as 'true' *B. sudanica*-type snails. The *B. pfeifferi* samples were split between the three morphotypes.

Method of ID		LDA re-grouping				
	Morpho-groups	*B. choanomphala*-type	*B. sudanica*-type	*B. choanomphala*-intermediate	*B. sudanica*-intermediate	**TOTAL**
Putative field ID	*B. choanomphala*-type	21	3	0	0	**24**
	B. sudanica-type	3	15	0	1	**19**
	B. choanomphala-intermediate	0	2	0	0	**2**
	B. sudanica-intermediate	4	7	0	0	**11**
	TOTAL	**28**	**27**	**0**	**1**	**56**

Table 7.8 - Comparison of the morpho-groups as determined by field identification versus LDA for internal anatomy measurements

The LDA analysis removed all individuals from the *B. choanomphala*-intermediate and the *B. sudanica*-intermediate morphogroups; in this case all the *B. sudanica*-intermediates were re-classified by the analysis as 'true' *B. sudanica*-type snails. The *B. pfeifferi* samples were split between the three morphotypes.

The internal anatomy measurements had the greatest number of changes between groupings, suggesting that these characters are not reliable for taxonomic analysis within this species complex (Table 7.8).

7.5.2.2 ANOVA/Tukey HSD tests

The ANOVA tests and associated Tukey HSD supported the PCA findings of an overlap between morphotypes. They also strongly associated particular morphological features with certain habitat types.

The significance tests between the means of the shell measurements for the various species groups showed that the *B. choanomphala* and *B. pfeifferi* have quite similar ratios between their dimensions, except for the depth ratio, in which they are significantly different (Table 7.9).

Ratio	Comparison	Difference	Lower CI	Upper CI	adjusted p-value
Height / Width	cho / sud	0.049	0.031	0.068	0.000
	cho / cho int	0.071	0.022	0.121	0.001
	cho / sud int	0.047	0.024	0.070	0.000
	sud / pfei	-0.059	-0.103	-0.015	0.003
	cho int / pfei	-0.081	-0.144	-0.017	0.006
	sud int / pfei	-0.056	-0.102	-0.011	0.008
	lake / marsh	0.032	0.014	0.050	0.000
	stream / marsh	-0.068	-0.119	-0.016	0.007
Depth 1 / Depth 2	cho / sud	0.110	0.081	0.139	0.000
	cho / sud int	0.109	0.073	0.145	0.000
	cho /pfei	0.082	0.013	0.151	0.012
	lake / marsh	0.058	0.025	0.091	0.000
Depth 2 / Width	cho / sud	-0.133	-0.160	-0.106	0.000
	cho / cho int	-0.102	-0.174	-0.029	0.002
	cho / sud int	-0.117	-0.150	-0.083	0.000
	sud / pfei	0.125	0.060	0.189	0.000
	sud int / pfei	0.108	0.041	0.175	0.000
	lake / marsh	-0.069	-0.103	-0.034	0.000
	stream / marsh	0.112	0.015	0.209	0.020

Table 7.9 – Significant results for Tukey HSD tests on shell measurements

'cho' stands for *B. choanomphala*-type snails, 'cho int' stands for *B. choanomphala*-intermediates, 'sud' stands for *B. sudanica*-type snails, 'sud int' stands for *B. sudanica*-intermediate, and 'pfei' stands for *B. pfeifferi* snails.

B. pfeifferi is also significantly different from all the other morphogroups for the height/width ratio, and from both *B. sudanica* types (the 'true' form and the intermediate forms) in terms of depth 2/width ratio. Otherwise, it is the 'true' *B. choanomphala*-types that were most often significantly different from the other morphotypes. *B. choanomphala* intermediates were not significantly different from the *B. sudanica* intermediates for any of the measurement ratios, and nor were 'true' *B. sudanica* and the *B. sudanica* intermediates. The most striking result came in terms of testing for significance based on habitat: all three ratios were significantly different between snails from the lake habitats versus those from the marsh habitat. The stream habitat was also significantly different from the marsh habitat, when testing the height/width and the depth 2/width ratio.

Fewer species combinations were significant when considering the copulatory organ measurements and ratios (Table 7.10).

Measurement	Comparison	Difference	Lower CI	Upper CI	adjusted p-value
# Diverticula	cho / sud	5.625	2.659	8.591	0.000
	cho / cho int	8.125	0.916	15.334	0.021
	cho / sud int	3.471	0.098	6.844	0.041
	lake / marsh	3.673	1.387	5.960	0.002
Preputium	cho / sud	0.617	0.200	1.035	0.001
	cho / sud int	0.637	0.155	1.119	0.005
	lake / marsh	0.650	0.373	0.928	0.000
Penis sheath	lake / marsh	0.518	0.304	0.732	0.000

Table 7.10 – Significant results for Tukey HSD tests on internal anatomy measurements
Abbreviations for morphogroup as above, in Table 7.9.

However, the pattern was similar; the 'true' *B. choanomphala* snails were significantly different from all three other morphotypes in terms of the number of diverticula, and from both of the *B. sudanica* morphotypes in terms of preputium length. No species combinations were significant, however, for the ratio of the penis sheath to the preputium. For all three raw measurements, there was a significant difference between the means of the marsh versus the lake samples. However, once again, the ratio of the penis sheath/preputium was not significantly different between the two habitat types.

7.5.3 Matrix correlations

The matrix correlation tests between the average pair-wise distances of the morphological measurements and the two different genetic markers were not significant. This means that the groupings seen in the genetic data cannot be significantly correlated to differences in the morphological characters that were analysed, or vice versa.

7.6 Discussion

The combination of the molecular and morphological data suggest that the accepted view of *B. choanomphala* and *B. sudanica* as separate species in Lake Victoria needs to be revised within the context of this alternative taxonomic appraisal, while the significance of habitat points to a strong environmental influence on phenotype. While the genetic data showed some differentiation, it was not consistent with visually identified morphological divisions, and moreover, was not at a level consistent with species-level distinctions in other *Biomphalaria*. The morphological data suggested a continuum of morphological types within a framework of three main groups: *B. choanomphala*-type, *B. sudanica*-type and an intermediate form. The morphological means were moreover more strongly significant when tested against habitat type rather than putative species identification, suggesting that the two 'species' of *Biomphalaria* in Lake Victoria are more likely ecophenotypes of a highly plastic and adaptable species, with morphological variation driven by selection pressures or environmental conditions from different habitats.

7.6.1 Molecular data

The high levels of variation in both the 16S and COI gene were in marked contrast to the negligible variation seen in the ITS region. Nuclear genes tend to accumulate changes more slowly than mitochondrial genes, which is why the former tend to be more commonly used for deeper phylogenies, for example between families or orders rather than at the population level. Moreover, ITS digests have been used successfully for species identification with South American *Biomphalaria* (Spatz *et al.*, 2000, Vidigal *et al.*, 2004). As such, in this case, the almost uniform nature of the banding patterns for the ITS digest lends weight to the hypothesis that the samples do not represent two distinct species.

Although highly variable, the level of genetic variation seen in the Lake Victorian *Biomphalaria* was not extraordinary in the context of other gastropods, which are often highly diverse (Thaewnon-ngiw *et al.*, 2003, Zhao *et al.*, 2010). Moreover, the extent of the variation in the 16S and COI genes was not sufficient in the author's opinion to warrant a

species-level division. The total genetic distance within the samples was very similar to the intraspecific difference for other *Biomphalaria* species, and much less than that for *B. glabrata*. When the mean distance was compared across other *Biomphalaria* species, the distances were also comparable; as expected the Lake Victorian samples were more similar to other African *Biomphalaria* than to most of the South American *Biomphalaria*. The genetic proximity to *B. glabrata* was more similar to that of African species than South American, once again giving support to the hypothesis that *Biomphalaria* in Africa are derived from a *B. glabrata*-like ancestor that originated in South America (Campbell *et al.*, 2000, DeJong *et al.*, 2001).

7.6.2 Morphological data

The morphological measurements revealed some evidence of groupings that were consistent with the putative field identification as either a *B. sudanica*- or *B. choanomphala*-type snail, although the small number of measurements used should be taken into consideration. Even so, there were also overlaps in morphospace between the two groups, even from type specimens, as in the case of *B. choanomphala* var. *basinulacatus* in the shell measurements and aperture outline PCAs. Overlap of *B. pfeifferi* in morphospace with other *Biomphalaria* types has been reported previously (Plam *et al.*, 2008), but generally *B. sudanica*-type shells are considered easy to distinguish from the rapidly-whorling forms. However, the overlap seen in the shell measurement and aperture outline PCAs between both *B. sudanica*-types and *B. choanomphala* types suggest that statistically the two forms are difficult to differentiate, and in fact intermediate forms may be found on a continuum between the extremes exemplified by the type specimens. Although the LDA rarely grouped individuals as intermediates, different individuals were incorrectly reclassified by the analysis, indicating inconsistency in the morphometric groupings; this was further supported by the matrix correlations, which did not find significant associations between the groups.

A very interesting finding was that many of the differences in the average measurements were significant between habitat types. Although causation is far from certain, this evidence

could tentatively suggest that habitat might be a factor in driving the morphological plasticity seen in the dataset. Whereas shell shape is known to be strongly environmentally determined in many snail species (Wellington and Kuris, 1983, Lam and Calow, 1988, Irie, 2006, Hoverman and Relyea, 2007), it was somewhat surprising also to see the copulatory measurements being significantly different between habitats. However, it is worth noting that the ratio between the penis sheath and preputium lengths was not significantly different; given that size of the snail overall is likely to be autocorrelated to the shape, size and organisation of the internal organs, this suggests that shell size may be the driving factor behind variation in the relative proportions of the internal anatomy.

7.6.3 A case for ecophenotypic variation

It has long been hypothesized that there may be ecophenotypes of *B. choanomphala* in Lake Victoria (Mandahl-Barth, 1957); indeed, the presence of the *basinulacatus* variety in the collection of the Berlin Museum für Naturkunde points to this, although, as mentioned previously, Mandahl-Barth originally only proposed two sub-species (*B. c. choanomphala* and *B. c. elegans*). Similarly, the description of varieties such as *B. sudanica minor* and *B. s. major* explicitly acknowledged the existence of a range of morphological forms. However, the status of *B. choanomphala* and *B. sudanica* from Lake Victoria as separate species has never before been seriously questioned; investigative taxonomy within African *Biomphalaria* has traditionally been thwarted by a paucity of reliable, informative characters.

Here, the majority of *B. sudanica*-type snails were found in marshy habitats, whereas *B. choanomphala*-type snails were located in the lake proper. Of course, many 'good' species are confined to a particular habitat type, and so this finding is not that surprising; however, the genetic evidence showing complete cross-over between these morphotypes encourages a deeper look at the supporting morphological evidence. Here, we see individuals possessing characters which do not consistently group with one morphotype or another, and moreover, that specific morphological attributes are significantly correlated with habitat. As such, it is tempting to suggest that perhaps the *Biomphalaria* from Lake Victoria

should be considered one, highly diverse species, both in molecular and morphological terms. The 'species' as they appear from putative field identifications could in fact be adaptations, based on some inherent plasticity in morphology, to environmental cues (Lam and Calow, 1988, Irie, 2006). It has been hypothesized that this has occurred in Lake Albert, as well as elsewhere in Uganda, where snails putatively identified as *B. pfeifferi*, based on shell characters, were in fact molecularly part of the Nilotic species complex (Jørgensen *et al.*, 2007); similarly a snail with a typically *B. stanleyi*–like morphology from Lake Albert had a DNA sequence that clustered closely with *B. pfeifferi* (Plam *et al.*, 2008). In this latter case, it was suggested that the rapidly increasing whorls of the *B. stanleyi* morphotype could be an adaptation to lacustrine conditions; *B. choanomphala*-type snails have a very similar form and as such the same process could be occurring in Lake Victoria. In this case, the genetic distance between the Lake Victorian samples and *B. sudanica* from elsewhere in Africa suggests that while they are all part of the Nilotic species complex, along with *B. alexandrina*, the *B. sudanica* from Lake Victoria are more closely related to their sympatric *B. choanomphala* than other *B. sudanica*, and so Lake Victorian *Biomphalaria* should perhaps be considered as a separate group.

The high levels of genetic distance observed between *B. sudanica* elsewhere in Africa as compared with the Lake Victorian samples analysed here has implications for the revision of the nomenclature for these *Biomphalaria*. *B. sudanica* was actually described in 1870 by von Martens, 9 years before he also described *B. choanomphala*. As such, based on the rules applied by the International Commission for Zoological Nomenclature (ICZN; www.iczn.org), '*B. sudanica*' should take precedence over '*B. choanomphala*'. However, given that *B. sudanica* elsewhere is likely a true species, whereas *B. choanomphala* is considered endemic to Lake Victoria, here it is proposed that '*B. choanomphala*' should take priority, with the two morphotypes, as identified by the main morphological groupings observed by the shell measurement and aperture analyses, denoted as variants. As such, *B. choanomphala*-type snails, and close intermediate forms, should be referred to as *B. choanomphala* var. *choanomphala* whereas *B. sudanica* and its intermediate forms should be described as *B. choanomphala* var. *sudanica*.

The main implication of this result concerns the ability of *Biomphalaria* in Lake Victoria to transmit *S. mansoni*. In most parts of the lake, *B. choanomphala* is considered to be the main transmitter of the parasite, whereas *B. sudanica* is not considered as successful an intermediate host; for example, no naturally infected *B. sudanica* have ever been found along then Ugandan shoreline of Lake Victoria (F. Kazibwe, pers. comm.). However, the molecular and morphological evidence presented above suggests that they are, in fact, the same species, and as such, both should be capable of transmitting *S. mansoni*, given other factors such as local compatibility. In fact, snails from three of the sites considered in this investigation were found naturally shedding *S. mansoni* cercariae; two were *B. choanomphala*-type but one, from site T001 in Tanzania, was *B. sudanica*-like. Moreover, informal compatibility experiments were conducted in the field, whereby a local isolate of *S. mansoni* miracidia were used to infect snails from sites T033a (*B. sudanica*-like, found in a marsh set back approximately 10 m from the lakeshore) and T033b (*B. choanomphala*-like snails, from a rocky and relatively exposed site in the lake proper). Snails from both sites proved to be compatible with the parasite; after 19-25 days, individuals from both T033a and T033b started shedding cercariae, which were able to infect laboratory mice, thus proving their viability. Further population analyses of snails from Lake Victoria would be beneficial, to explore in detail the connectivity, structure and relatedness of the different morphotypes.

If *B. sudanica* and *B. choanomphala* are two forms of the same species and both able to transmit *S. mansoni*, then there are also implications for the control of intestinal schistosomiasis in people. To date, many national control programmes have included educational messages alongside mass chemotherapy, in order to change water contact behaviours and reduce exposure to the disease (Stothard *et al.*, 2006b, Fenwick *et al.*, 2009). On the whole, the message has been for people to reduce their contact with water collected directly from the lake, for example through standing it for 24 hours prior to use, or, as a more long-term strategy, by installing water bores and pumps. However, if *B. sudanica*-type snails are also capable of transmitting, then the habitats that these forms frequent should be considered a high-risk zone for transmission. This includes marsh areas set back from the

lake, which are often crossed while going to and from the lake, and where livestock are often grazed. Adding information about the risks of these marsh areas would assist in controlling exposure risk in this 'new' transmission environment. Similarly, in some cases it may be possible and desirable to drain marsh areas, given that they can also be important habitats for mosquito larval development. Moves such as this would be an immediate, and highly beneficial, public health consequence of this research.

7.7 Conclusions

Although the taxonomy of African *Biomphalaria* is generally confused, combining molecular and morphological tools is a powerful method for resolving the muddle. Here, *B. choaonomphala*-type and *B. sudanica*-type snails are shown to be highly genetically diverse, but not divergent, with well-supported clades containing mixtures of morphotypes. Overall, habitat was a strong factor in determining morphology, which, when combined with the molecular data, let to the conclusion that rather than separate species, *B. choanomphala* and *B. sudanica* from Lake Victoria could be ecophenotypes of the same species. As such, it is proposed to combine the morphotypes in Lake Victoria into a single species with two ecophenotypic variants, namely *B. choanomphala* var. *choanomphala* and *B. choanomphala* var. *sudanica*, based on the main morphological groupings seen in the shell and aperture data. By referring to *B. choanomphala*-type and *B. sudanica*-type snails as *B. cho.* var. *choanomphala* and *B. cho.* var. *sudanica*, the genetic congruence between the ecophenotypes is observed, whilst the morphological differences are also recognised. The revision of this genus is of crucial importance, given the far-reaching implications for the transmission, and on-going control, of schistosomiasis in the region. Further work, presented in the next chapter, takes this genetic work further by sampling more widely across Lake Victoria and incorporating microsatellite markers with the mitochondrial data for an explicitly population-level analysis.

8 Population genetics of *Biomphalaria* snails in Lake Victoria: Implications for transmission of *S. mansoni*

8.1 Abstract

Host-parasite relationships involve a complicated variety of interactions, which can impact the genetic diversity of each. The situation for *Biomphalaria* and *S. mansoni* is of particular scientific interest given the biomedical importance of the parasite as a pathogen of man. Investigating the population structure and genetic diversity of the intermediate host, and comparing it to that of the parasite, can elucidate the evolutionary history and direction of the relationship. To achieve these aims, over 300 snails from 29 populations widely spread across the perimeter of Lake Victoria were studied using molecular markers. Two mitochondrial gene regions were used together with four microsatellite loci for population genetics and phylogeographical analyses. Results showed very high levels of inter- and intrapopulation genetic diversity as well as significant population structuring, the latter being positively correlated with increasing geographical distance. However, pairwise genetic distance values were not correlated with those from *S. mansoni* from the same localities; similarly, genetic diversity of the snails was not significantly correlated with infection status or susceptibility to *S. mansoni*. The observation of different population structures between intermediate host and parasite is typical of other systems, but the levels of diversity in the *Biomphalaria* surveyed here were very high, and perhaps indicative of the stable, homogeneous environment that Lake Victoria provides. While high host diversity is often considered detrimental to parasite development and success, Lake Victoria is a highly endemic zone for intestinal schistosomiasis, suggesting that this is a topic that should be investigated further. Moreover, while intermediate hosts clearly provide a fruitful avenue of research in terms of understanding host-parasite relationships, the effect of the diversity of terminal hosts should also be considered in future analyses.

8.2 Contribution of the author

As described in Chapters 6 and 7, the author was present throughout the duration of all the field missions used for making the snail collections used in this research, and moreover selected the sites to be included in the analysis (as described in the Methods section below). As with Chapter 7, the author was responsible for all DNA extractions and amplifications, as well as all purifications and sequencing reactions until the summer of 2009, upon which time the Sequencing Facility at the NHM took over the purification of PCR products and the sequencing reactions. All sequences/microsatellite fragments were analysed by the author, and all statistical and genetics tests were also performed by the author.

8.3 Introduction

Historically, both theoretical and experimental work has been performed on the nature of the host-parasite relationship. The advent and growth of molecular studies has revolutionised this field, allowing scientists to test predictions about the effect of selection and evolution on various aspects of the host-parasite system. For example, it is hypothesised that inbreeding and the subsequent loss of heterozygosity and/or genetic variation may contribute to host susceptibility to parasite infection (Hedrick *et al.*, 2001, Spielman *et al.*, 2004, Pearman and Garner, 2005), and enhanced parasite gene flow can further increase its ability to adapt to local hosts (Gandon *et al.*, 1996, Kaltz and Shykoff, 1998). Rare genotypes may be maintained in the host population if parasite populations are less able to develop in those individuals, which in turn may drive genetic diversification in the host (van Valen, 1973, Koskella and Lively, 2009).

As such, the genetic diversity and population structure of *Biomphalaria* has received a significant amount of research attention in recent years. It has been demonstrated, in numerous settings, that increased genetic diversity in the snail intermediate host reduces the success of the parasite, whereas host homogeneity can result in greater parasite transmission (Hoffman *et al.*, 1998, Jarne and Théron, 2001, Webster *et al.*, 2001a, Sandland *et al.*, 2009). In some cases, the success of snail populations that are highly susceptible to *S.*

mansoni infection may be due to stronger selective forces acting on other aspects of the snail's physiology and life history, for example fecundity (Campbell *et al.*, 2010). These situations allow for the rapid range expansion of compatible snail genotypes and subsequent increases in the distribution of intestinal schistosomiasis in the region. This effect can be exacerbated by preferential in-breeding, as observed in certain *Biomphalaria* species, which can maintain low diversity whilst increasing the number of susceptible individuals within a population (Bandoni *et al.*, 1990).

Most previous research has focused on *B. glabrata* and *B. pfeifferi*, considered the two most important intermediate host snail species of *S. mansoni* in the New and Old World, respectively. However, there are other regions of the world where intestinal schistosomiasis is highly endemic, yet transmitted by other *Biomphalaria* species. For example, no previous investigations have focused on or indeed included geographically representative samples from Lake Victoria. As a local and regional hot spot for *S. mansoni* (Stothard *et al.*, 2005, Odogwu *et al.*, 2006, Standley *et al.*, 2009), *Biomphalaria* from this lake could provide insights into the local dynamics of the transmission of the parasite, through looking at their population structure. Comparing the genetic patterns of *Biomphalaria* with those of *S. mansoni* in the same region may elucidate the role of intermediate host compatibility in driving the high genetic diversity and widespread distribution of the parasite, as demonstrated earlier in this thesis.

In the previous chapter, the taxonomic difficulties associated with *Biomphalaria* were described in detail. Based on molecular and morphological evidence it was shown the snails found in Lake Victoria appear to be consistent with belonging to a single species. As such, it was possible to conduct a detailed and thorough population genetics analysis of *Biomphalaria* from populations from all around the shoreline of Lake Victoria, the results of which are presented here. Lake Victoria differs ecologically to the locations where previous *Biomphalaria* population studies have been carried out; the effect of its size and stability as a habitat may be reflected in the population structure of the snails. As such, the aims of the research were to ascertain the levels of population structuring of *Biomphalaria* from Lake Victoria, which have never before been investigated, and to look for geographical patterns

which might reflect different levels of exposure and susceptibility to *S. mansoni*. Correlation with *S. mansoni* population structure and diversity could also be compared directly, using the schistosome genetic data presented in Chapter 5.

8.4 Methods

The population structure of *Biomphalaria* in Lake Victoria was tested using two mitochondrial genes as well as four co-dominant nuclear microsatellite loci.

8.4.1 Sampling methodology and population selection

Snails were collected from around the perimeter of Lake Victoria, using hand held scoops and a semi-quantitative sampling methodology, as described in section 2.1.4 of the main Methods chapter. All snails were shed for infection with *S. mansoni* or other parasite cercariae by exposing individuals to sunlight for several hours (see Methods section 2.3). 29 sites were selected for inclusion in the population genetics study (Figure 8.1).

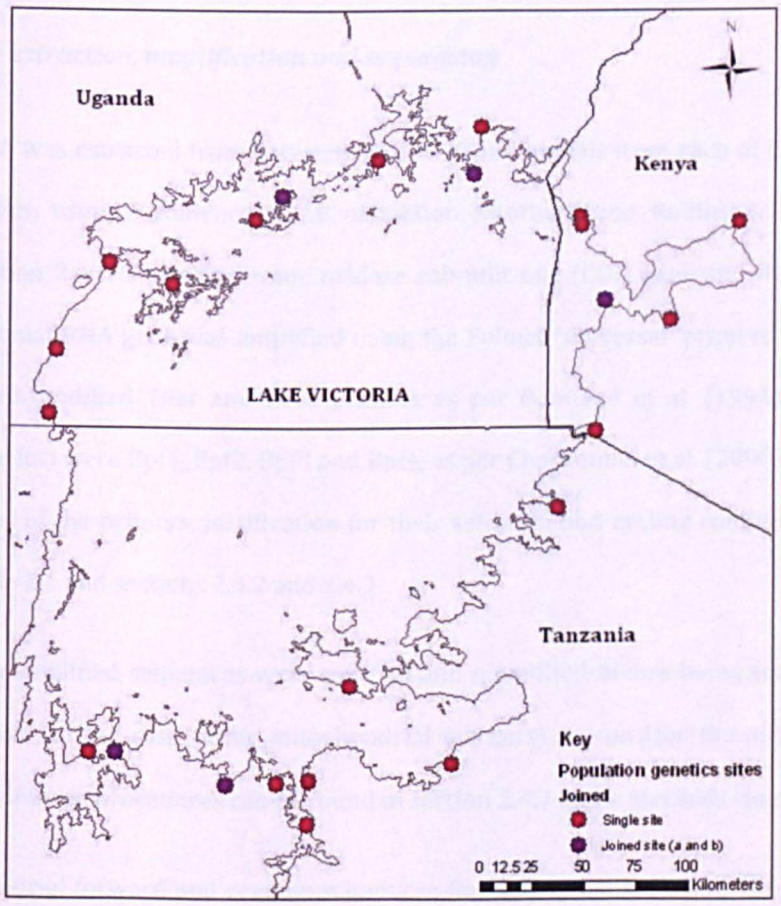

Figure 8.1 – Sites selected for the population genetics analyses

Sites in purple represent 'paired' localities, where snails were collected from two, unconnected habitats at the same geographical point. Sites in red are single habitat sites.

The first criterion for selection was the abundance of *Biomphalaria* found at the site; any site with fewer than 20 snails was not included. Sites were then chosen for geographical spread across the shoreline of the lake, in all three countries, and to represent both marsh and habitat types; five localities were specifically included because they incorporated both a marsh and lake site in close proximity.

Snails were putatively identified as *B. choanomphala* or *B. sudanica*, based on shell morphology, as these are the species thought primarily to inhabit Lake Victoria (Brown, 1994). However, as mentioned above, on-going taxonomic work has suggested that all the *Biomphalaria* in Lake Victoria can be considered a single species (see Chapter 7), and thus for the purposes of this study species-level distinctions were not made between the two morphotypes.

8.4.2 DNA extraction, amplification and sequencing

Genomic DNA was extracted from between 10 and 12 individuals from each of the selected 29 populations, using a standard CTAB extraction (Stothard and Rollinson, 1997); see Methods section 2.4.1. The cytochrome oxidase sub-unit one (COI) gene and the 16S sub-unit of ribosomal RNA gene was amplified using the Folmer 'universal' primers (Folmer *et al.*, 1994) and modified 16ar and 16br primers as per Bonnaud et al. (1994). The four microsatellite loci were Bpf1, Bpf2, Bpf3 and Bpf4, as per Charbonnel et al. (2000). Details of the sequences of the primers, justification for their selection and cycling conditions can be found in Table 2.1 and sections 2.4.2 and 2.4.3.

All positively amplified sequences were purified and quantified before being sequenced in both directions (in the case of the mitochondrial markers) or run (for the microsatellite loci). Details of these procedures can be found in section 2.4.3 of the Methods chapter.

The mitochondrial forward and reverse sequences for each individual were combined and edited before being aligned with the rest of the samples (section 2.5.1). Regions with gaps were visually aligned as best as possible, and retained when reducing the sequences into unique haplotypes, as the gaps in this case represent true step-wise changes to the gene sequence. However, when the data were analysed phylogenetically, unalignable gap regions

were removed as they could not be ensured to represent homology across each position. The sizes of each of the alleles for the microsatellite loci was calculated using Peakscanner.

8.4.3 Population genetics analyses

The overall framework of the population genetics analyses were designed to test specific hypotheses relating to the population structure, diversity, demography, phylogeography and spatial correlations of the *Biomphalaria* populations surveyed from Lake Victoria (Jarne and Théron, 2001).

The COI and 16S sequences were compared and reduced to unique haplotypes, and the proportion of each haplotype per population was recorded (section 2.5.1). Pie charts of these data were created in Excel and assembled on a map using ArcGIS and Power Point. A distance-based, neighbour-joining tree was also created in Paup* 4.0 (Swofford, 2003) for the unique haplotypes from each gene, using the GTR+gamma model of sequence evolution, previously determined to be the most appropriate model (see Chapter 7, and also section 2.5.3 for details of model choice and tree-building). Median-joining networks were constructed based on the unique haplotypes (section 2.5.4); due to the methodology of the network-building, this excluded gaps and unaligned regions from the analysis and so reduced the number of unique haplotypes.

Population genetics tests included AMOVAs between populations, calculated both with groups by locality (thus combining habitat types at the same geographical point) and by country, F-statistics, mismatch analysis, Tajima and Fu's tests for neutrality and a Mantel z-test for correlation by distance (section 2.5.5); all tests were executed in Arlequin v 3.2 (Excoffier *et al.*, 2005). The microsatellite loci were also analysed using a variety of tests, and included all those mentioned above apart from the neutrality tests. In this case, F_{ST} calculations were based on numbers of different alleles. In addition, the alleles were tested for Hardy-Weinberg equilibrium (through calculation of F_{IS} values per population and per locus) as well as linkage disequilibrium; these tests were performed in Genepop for the Web as well as in Arelquin, as a cross-check. The number and proportion of private alleles per population was counted in Genepop. The loci were also tested for previous bottlenecking

events by comparing the allelic diversity against the observed and expected heterozygosity, using the programme Bottleneck.

8.4.4 *Comparison with* S. mansoni *data*

In order to investigate the co-evolution of *S. mansoni* and *Biomphalaria* in Lake Victoria, pair-wise genetic distances between the two species were compared statistically (Sire *et al.*, 2001). A matrix of uncorrected pairwise genetic distances for *S. mansoni* (see Chapter 5) was compared to a matrix of *Biomphalaria* genetic distances (also uncorrected p-distances; see section 2.5.2 for details) from the same locality, using a Mantel matrix correlation test in R (see Methods section 2.5.5 and 2.6). This was done for the COI data and microsatellite data; no 16S sequence data were available for the *S. mansoni* populations. In cases where the collection site of the schistosome isolate was not exactly the same as a snail collection site, the nearest snail site was used.

As all snails had been shed for infection with *S. mansoni*, plus some populations had been used in compatibility experiments in the lab, informal information as to the transmission ability of each population could be tested against genetic and nucleotide diversity. This was done using a two-tailed, Student's t-test (in Excel); snails observed as shedding in the wild and 'susceptible' snails ('susceptible' included laboratory shedders as well as wild-caught shedding populations) were tested separately.

8.5 Results

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

8.5.1 Mitochondrial data

8.5.1.1 Distribution and abundance of haplotypes

The COI and 16S sequences were very genetically variable, with 127 and 181 unique haplotypes out of 308 and 300 sequences, respectively (Genbank accession numbers HM769132-HM769258 for COI and HM768950-769131 for 16S). A full table of the name and frequency of the haplotypes found at each site is in the Appendix, section 13.4. The haplotypes were reasonably well spread throughout the lake, although some pockets of lower diversity were observed, such as along the western shoreline of the lake. Similarly, some areas appeared to be particularly diverse, across both genes, such as the central portion of the lake in Uganda. For the COI data, several haplotypes could be observed to be locally common, and found in several sites, but restricted to a particular country (Figure 8.2). This pattern was particularly evident in Kenya (for example haplotypes 5 and 8) and the central-southern shoreline of Tanzania (such as haplotypes 23 and 24).

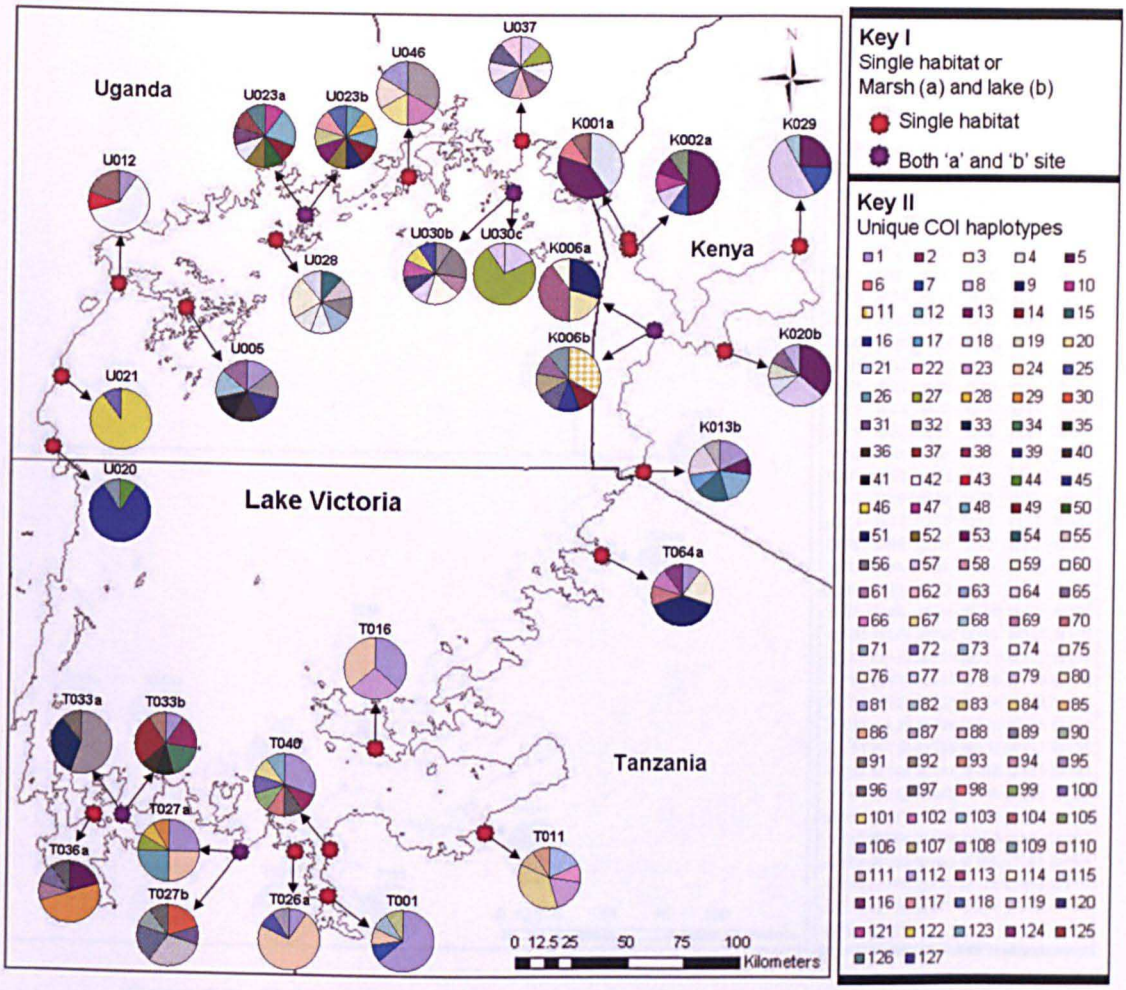

Figure 8.2 – Pie charts of the 127 unique COI haplotypes, per population

The figure shows the names of each of the 29 sites used for population genetics analysis; the site name is printed above each pie chart. As in Figure 8.1, the colour of the point denoting the site indicates whether it was 'paired', containing two habitats, or single (see Key I). The different COI haplotypes are labelled as per Key II, with a different colour representing each unique haplotype. The haplotype numbers are consistent throughout the chapter. A full list of the haplotypes found at each site is in the Appendix, section 13.4.

The 16S data showed a similar geographical pattern in terms of haplotype frequency and distribution (Figure 8.3). While in this case the central-southern shoreline of Tanzania was more diverse than for the COI sequences, again shared haplotypes could be observed, such as haplotype 62, which was not found in either Uganda or Kenya. The observation of shared COI haplotypes in Kenya was reinforced by the 16S data, which showed 69 to be the dominant haplotype in this part of the lake, shared by four sites but not observed further afield.

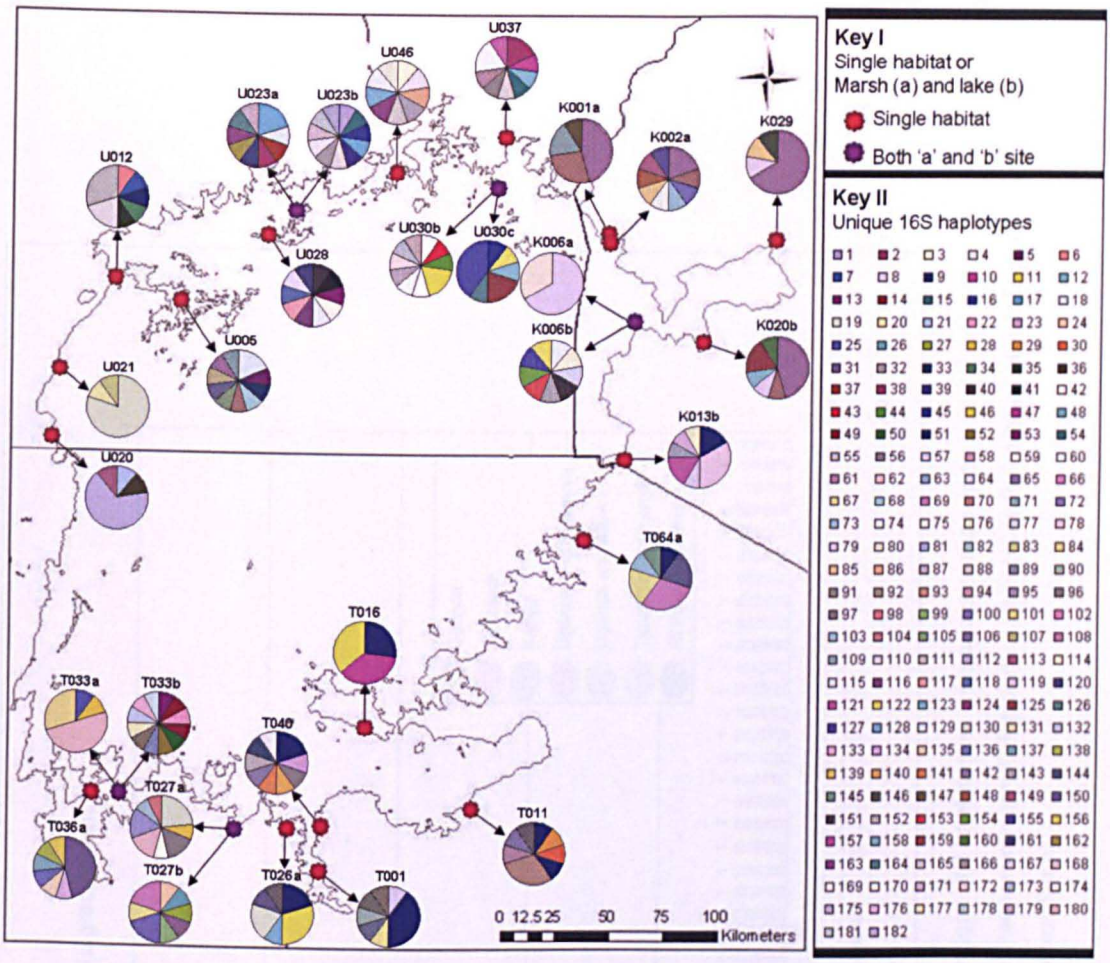

Figure 8.3 – Pie charts of the 181 unique 16S haplotypes, per population

The figure shows the names of each of the 29 sites used for population genetics analysis; the site names are situated above each pie chart. As in Figure 8.1, the colour of the point denoting the site indicates whether it was 'paired', containing two habitats, or single (see Key I). The different 16S haplotypes are labelled as per Key II, with a different colour representing each unique haplotype. The haplotype numbers are consistent throughout the chapter. A full list of the haplotypes found at each site is in the Appendix, section 13.4.

Both the COI and 16S haplotypes appeared to be reasonably locally restricted, with only one COI haplotype and no 16S haplotypes found across all three countries surrounding Lake Victoria, despite high abundance of common haplotypes. However, for the COI data, the most abundant haplotype was also the only one found in all three countries (haplotype 1; see Figure 8.4). In contrast, for the 16S data, haplotype 69 was very abundant, and found in multiple sites, but was restricted to Kenya, hinting that there may be geographically dependent population structuring in these snails (Figure 8.5).

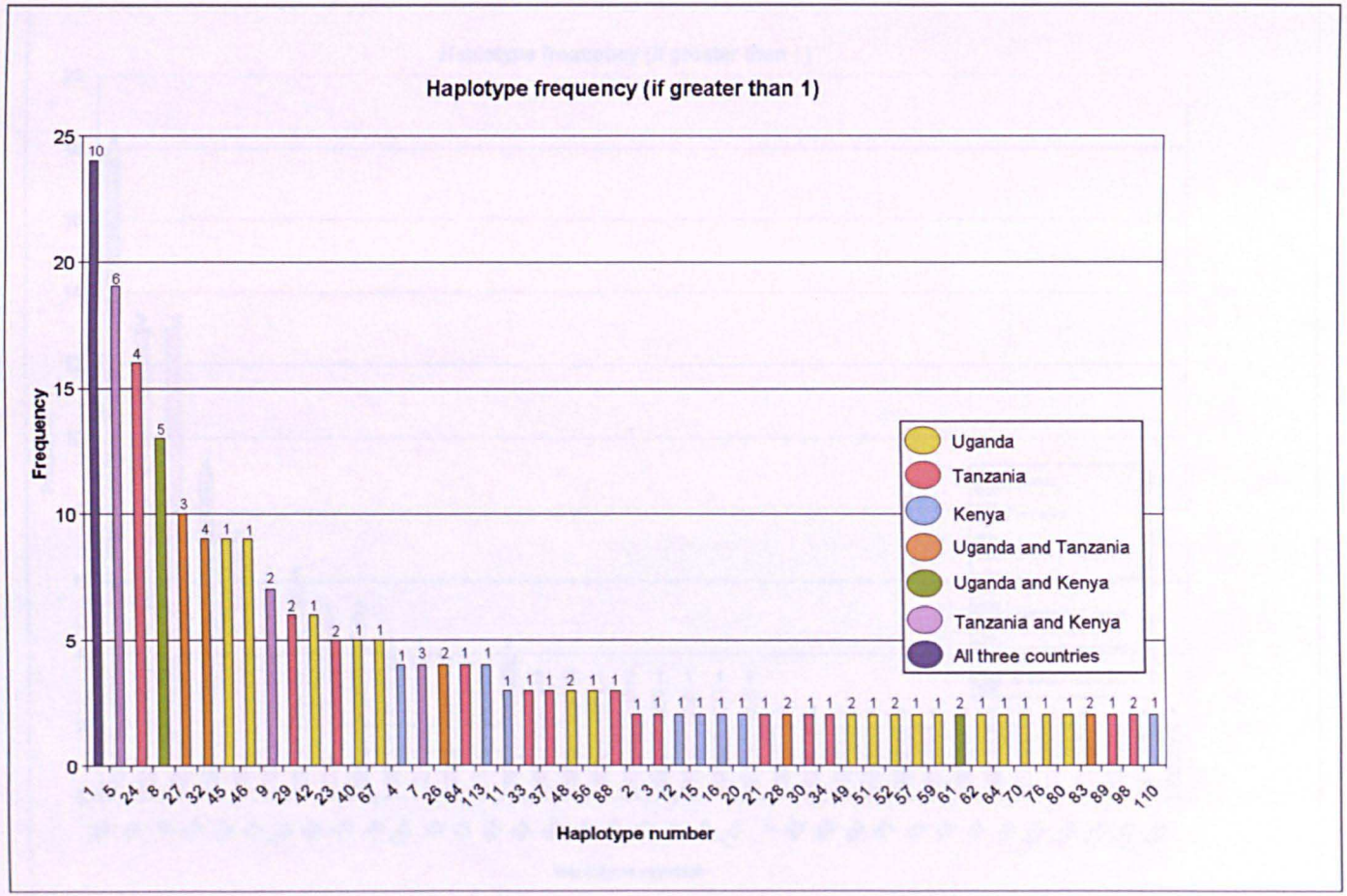

Figure 8.4 – Bar chart of the frequency of unique COI haplotypes, if found more than once

Each bar is coloured according to the country or countries in which that haplotype was observed. Haplotype numbers are the same as seen in Figure 8.2 above. The number above each bar represents the number of sites at which that haplotype was found.

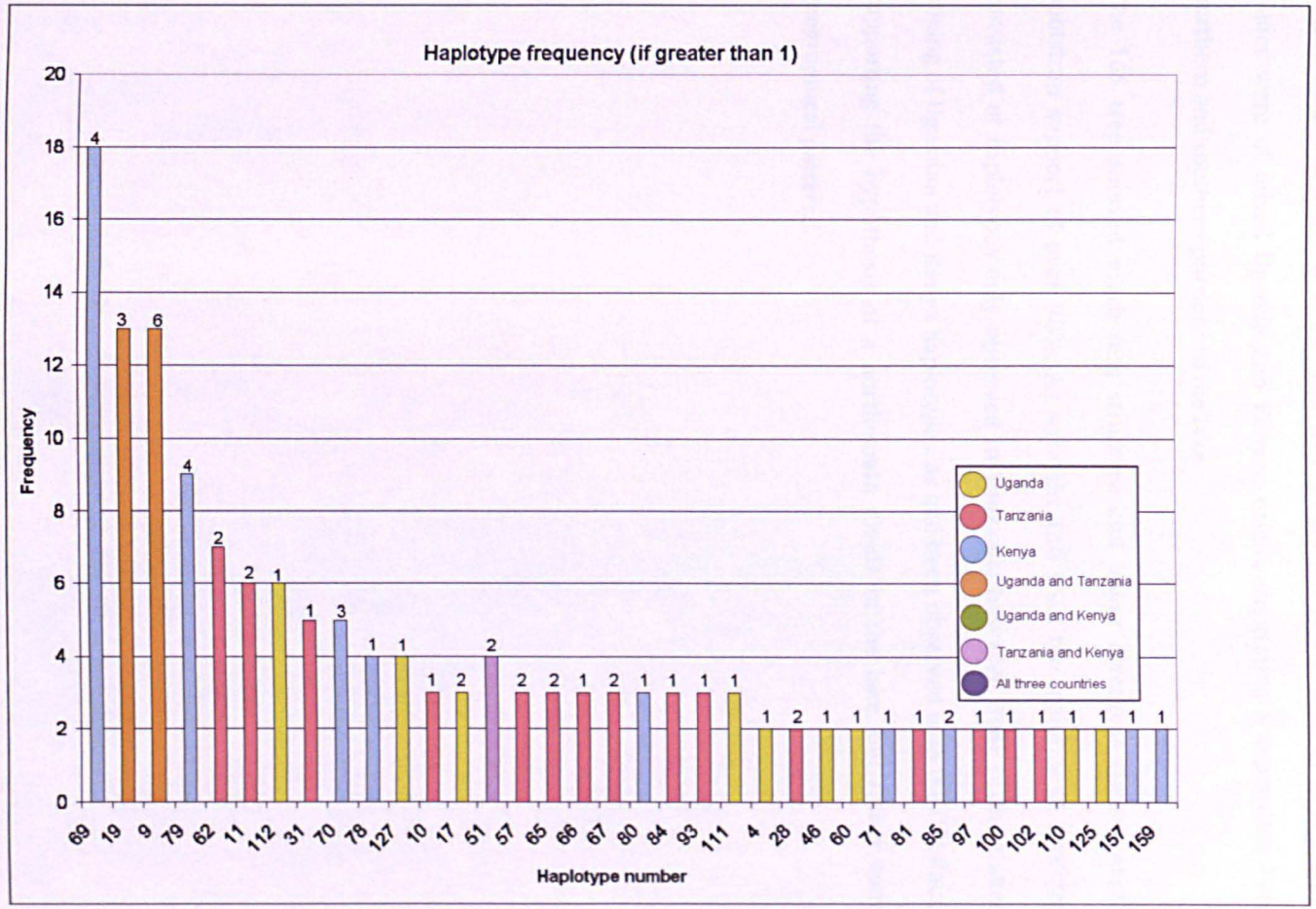

Figure 8.5 - Bar chart of the frequency of unique 16S haplotypes, if found more than once

Each bar is coloured according to the country or countries in which that haplotype was observed. Haplotype numbers are the same as seen in Figure 8.3 above. The number above each bar represents the number of sites at which that haplotype was found.

8.5.1.2 Trees

The COI neighbour-joining phylogenetic tree showed some structuring with a few well-supported groups that tended to include haplotypes from the same population or country (Figure 8.6). The majority of the divergent clades were of Tanzanian origin, although two clades were of mixed Uganda and Kenyan origin, suggesting a separation between the northern and southern portions of the lake.

The 16S tree showed much less structure and fewer divergent clades; only one had bootstrap support of over 90%. As with the COI tree, the majority of divergent clades consisted of haplotypes only observed in Tanzania; however, none of the clades showed mixing of Ugandan and Kenya haplotypes, as had been observed with the COI data, thus not supporting the hypothesis of a north-south divide in the lake, but rather country-level geographical patterns.

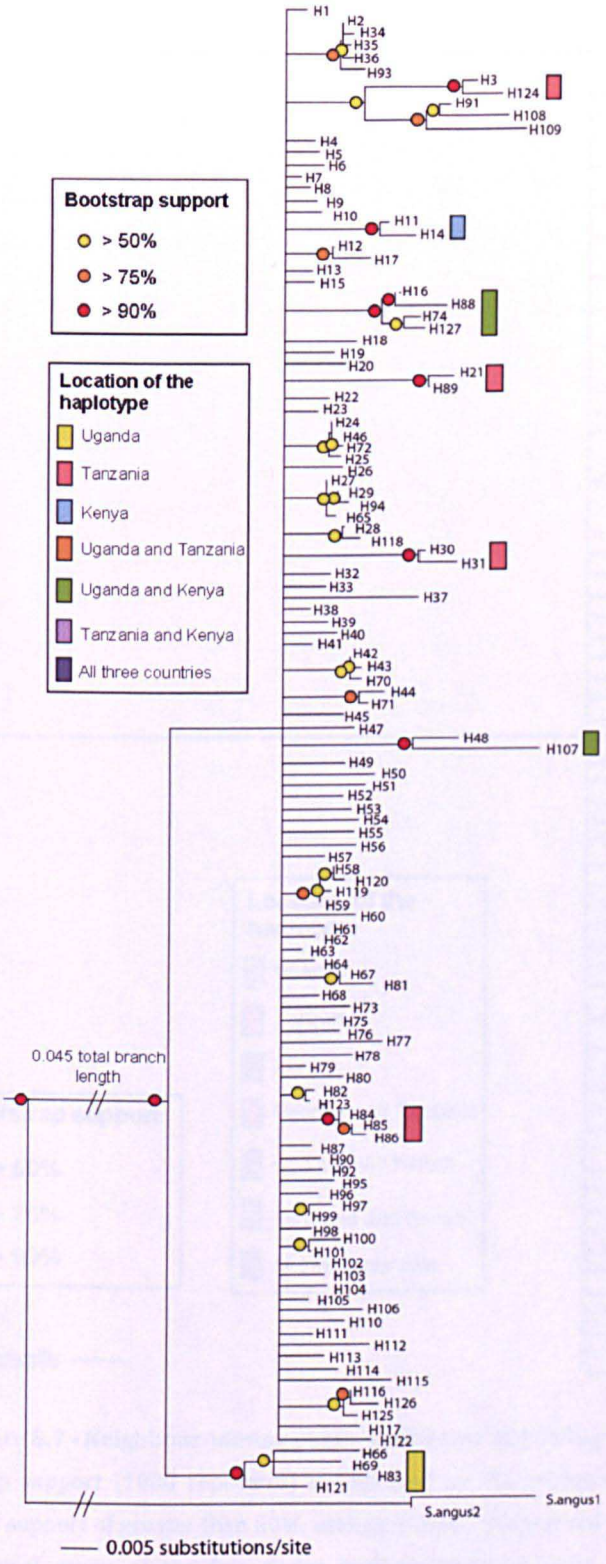

Figure 8.6 – Neighbour-joining phylogenetic tree of COI haplotypes

Bootstrap support (1000 replicates) is indicated by the coloured circles; yellow indicates support of greater than 50%, orange is over 75% and red is over 90%. The geographical origin of the haplotypes is also indicated by a coloured bar, for the groups that have bootstrap support of over 90%. S.angus1 and 2 are *Segmentorbis* snails used as an outgroup.

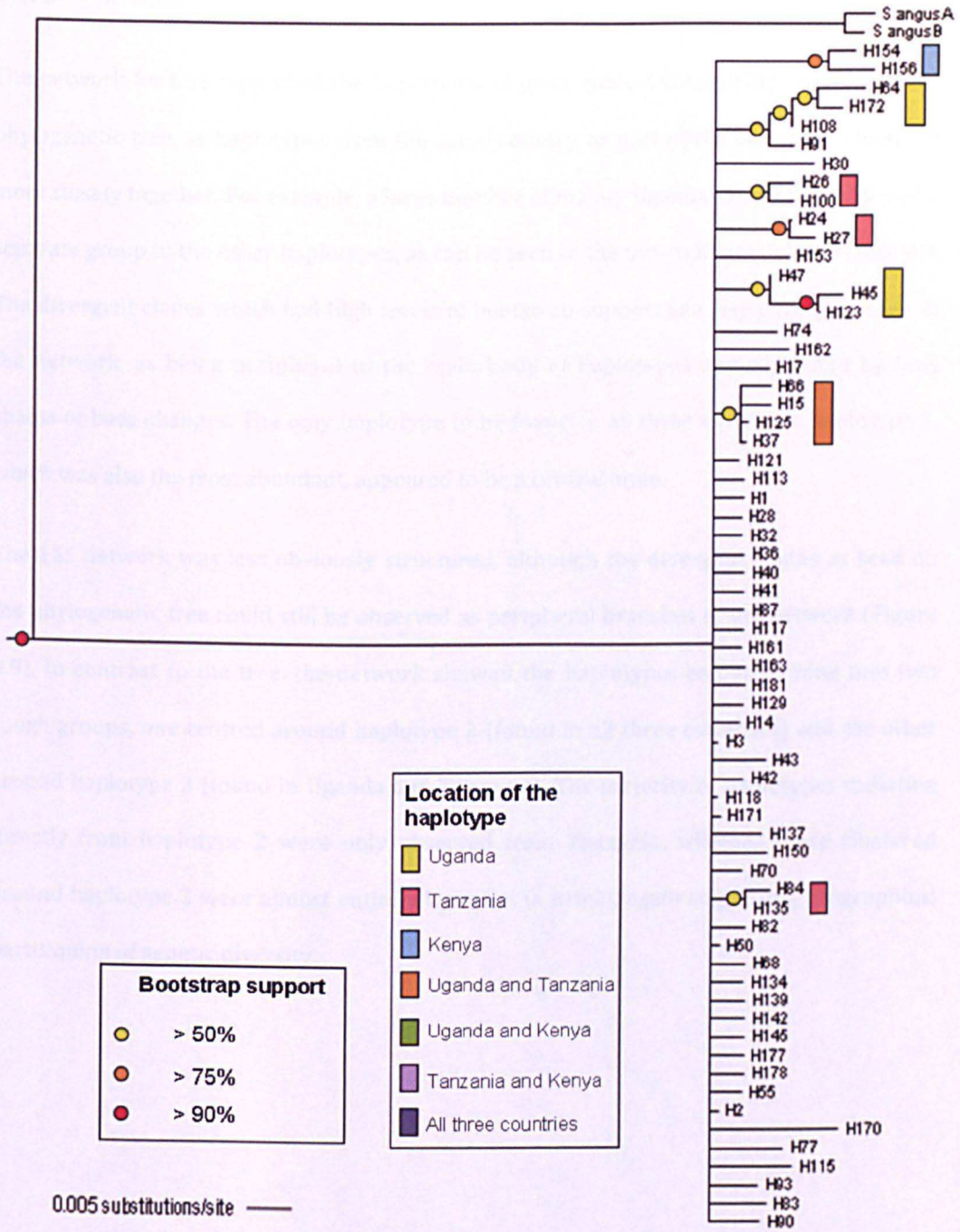

Figure 8.7 - Neighbour-joining phylogenetic tree of 16S haplotypes

Bootstrap support (1000 replicates) is indicated by the coloured circles; yellow indicates support of greater than 50%, orange is over 75% and red is over 90%. The geographical origin of the few clades with bootstrap support of over 50% is indicated by the coloured bars. The numbering of the haplotypes is consistent with earlier figures except that here the haplotypes are gap-free and therefore represent a greater number of sequences; this accounts for any differences in the colour of the haplotype here as compared to Figure 8.5, which displays haplotypes with gaps.

8.5.1.3 Networks

The network for COI supported the hypothesis of geographical structuring, as seen on the phylogenetic tree, as haplotypes from the same country or part of the lake were clustered more closely together. For example, a large number of mainly Ugandan haplotypes formed a separate group to the other haplotypes, as can be seen in the top-right portion of Figure 8.8. The divergent clades which had high levels of bootstrap support can also clearly be seen in the network, as being peripheral to the main body of haplotypes and separated by long chains of base changes. The only haplotype to be found in all three countries, haplotype 1, which was also the most abundant, appeared to be a central node.

The 16S network was less obviously structured, although the divergent clades as seen on the phylogenetic tree could still be observed as peripheral branches of the network (Figure 8.9). In contrast to the tree, the network showed the haplotypes as partitioning into two rough groups, one centred around haplotype 2 (found in all three countries) and the other around haplotype 3 (found in Uganda and Tanzania). The majority of haplotypes radiating directly from haplotype 2 were only observed from Tanzania, whereas those clustered around haplotype 3 were almost entirely Ugandan in origin, again suggesting geographical partitioning of genetic diversity.

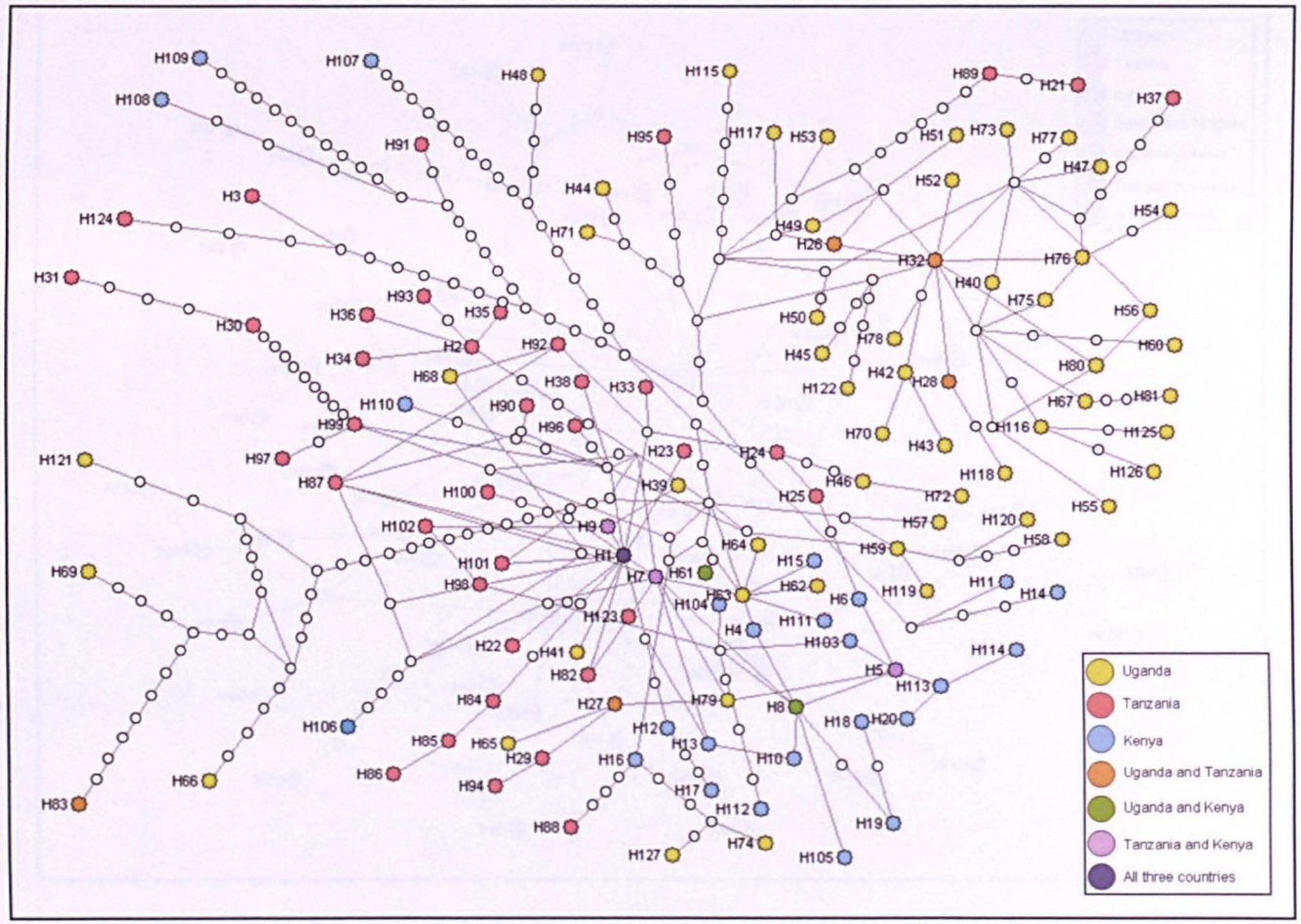

Figure 8.8 – Median-joining network of unique COI haplotypes

Each white circle represents a single base change, and so could also be thought of as indicating haplotypes which were not sampled. The larger, coloured circles represent COI haplotypes and the colour reflects the country or countries of origin. Haplotype numbering is consistent with that in previous figures.

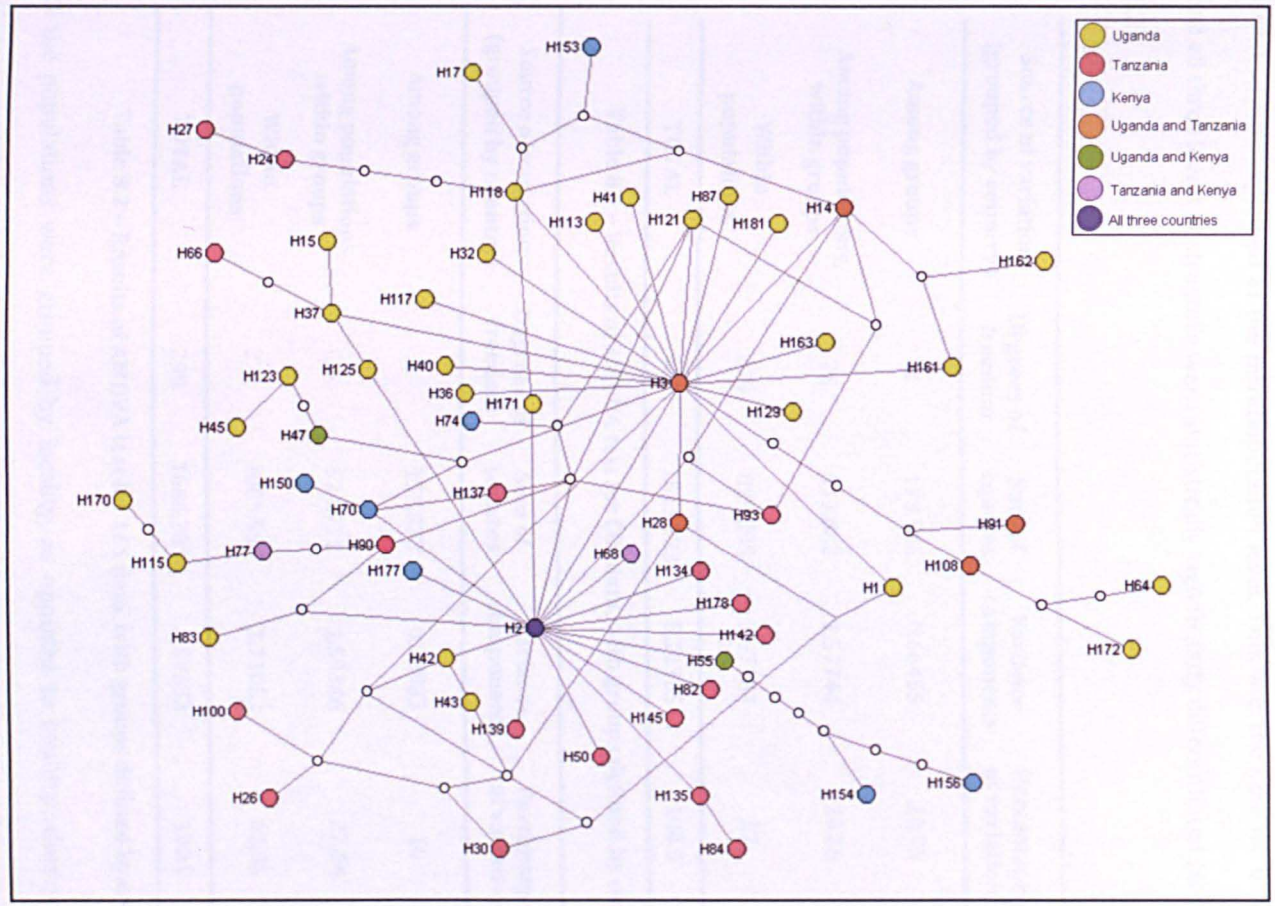

Figure 8.9 – Median-joining network of unique 16S haplotypes (gaps excluded)

Each white circle represents a single base change, and so could also be thought of as indicating haplotypes which were not sampled. The larger, coloured circles represent 16S haplotypes and the colour reflects the country or countries of origin. Haplotype numbering is consistent with that in the phylogenetic tree as for this analysis gaps were also removed resulting in a smaller dataset of unique haplotypes.

8.5.1.4 *Population genetics tests*

The population genetics tests confirmed statistically the presence of high levels of population structuring. For the AMOVA tests, which explore the percentage of the molecular variance that can be explained at the intergroup, intragroup and intrapopulation level, the majority of the variation was explained at the intrapopulation level. This was the case for both COI and 16S, and all three levels of structure were statistically significantly differentiated (see Tables 8.1 and 8.2).

Source of variation (grouped by country)	Degrees of freedom	Sum of squares	Variance components	Percentage of variation	p -value
Among groups	2	173.346	0.66455	12.75	< 0.0001
Among populations, within groups	26	512.682	1.57744	30.26	< 0.0001
Within populations	279	829.095	2.97167	57	< 0.0001
TOTAL	307	1515.123	5.21365	100.0	

Table 8.1 – Results of AMOVA test for COI data, with groups defined by country

Source of variation (grouped by country)	Degrees of freedom	Sum of squares	Variance components	Percentage of variation	p -value
Among groups	2	155.878	0.58933	10	< 0.001
Among populations, within groups	26	524.771	1.59346	27.04	< 0.0001
Within populations	271	1005.554	3.71053	62.96	< 0.0001
TOTAL	299	1686.203	5.89333	100.0	

Table 8.2 – Results of AMOVA test for 16S data, with groups defined by country

If the populations were grouped by locality, as opposed to country, thus pairing different habitats at the same geographical point, the proportion of variation explained by comparisons among the groups dropped to 10.25% and 2.81% for COI and 16S respectively, and was no longer significant. This demonstrates that overall, snails from marsh and lake habitats at the same geographical point were as, if not more, differentiated from each other as non-paired sites. When pair-wise F_{ST} values were calculated between each of the populations in turn,

virtually all demonstrated significant levels of differentiation between the populations, for both markers. This is a sign of restricted gene flow or movement of individuals between populations. For the 16S data, all but one of the non-significant F_{ST} values were between sites from the same country, once again suggesting geographically defined structuring of the populations for this gene. Only one paired site (U023a and U023b) had non-significant differentiation for both genes; T033a/T033b did not show significant population structuring for the 16S gene but did for COI. Full tables of the pairwise F_{ST} values and their significance can be found in the Appendix (section 13.4).

All of the populations conformed to the assumptions of the neutral hypothesis, as tested for by Tajima's D statistic, for the 16S data; one population, T001, deviated significantly from assumptions of normality when the COI data were analysed. Assumptions of neutrality are implicit in many population tests, although the absence of neutrality may indicate demographic influences such as selective breeding or recent population expansion. In this case, the T001 population also had a significant result for the sum of squares mismatch analysis, suggesting that indeed population expansion may have occurred, thus causing the result of the neutrality test. No other populations, for either marker, came out as significant in the mismatch analysis. The full results of the neutrality tests and mismatch analysis can be seen in the tables in Appendix 13.4, which also contain information about the molecular diversity indices of both COI and 16S.

The relationship with geography was elucidated through the significance of the Mantel z-test for both the COI and 16S data, meaning that the genetic and geographical distances between the populations was positively correlated; in other words, sites closer together were more likely to be genetically similar as well. This added statistical weight to the casual observation of geographical structuring in the haplotype frequency bar charts, trees and networks (Figures 8.4 to 8.9).

8.5.2 Microsatellite data

The microsatellite data broadly agreed with the mitochondrial data in terms of the population genetics test results, but were also able to be used for further tests such as for Hardy-Weinberg

equilibrium (HWE), linkage and whether the populations had recently passed through a bottleneck.

As with the mitochondrial data, the majority of the variation in microsatellite genotype was explained at the intrapopulation level; values of differentiation at the intragroup and intrapopulation level were both highly significant (Table 8.3).

Source of variation (grouped by country)	Degrees of freedom	Sum of squares	Variance components	Percentage of variation	p -value
Among groups	2	9.512	0.00567	0.31	0.237
Among populations, within groups	23	88.363	0.19001	10.55	< 0.0001
Within populations	286	459.051	1.60507	89.13	< 0.0001
TOTAL	311	556.926	1.80075	99.99	

Table 8.3 - Results of AMOVA test for the microsatellite data, with groups defined by country

When the groups were defined by locality, the among group differentiation was not significant for the microsatellite data, whether the groups were defined by locality or country. Overall, much less variation was explained between groups for the microsatellite data (0.31%) than for either of the mitochondrial genes. Compared to the populations for the mitochondrial data, fewer of the pairwise F_{ST} values were significant between populations for the microsatellite data. The full matrix of pairwise F_{ST} values for the microsatellite data and their significance can be found in Appendix section 13.4. The microsatellite data also supported the mitochondrial evidence for genetic spatial autocorrelation (in other words, of a positive correlation between geographical and genetic distance, with populations distributed closer together being more likely to also be genetically more similar), as the Mantel z-test was significant.

The extra tests (namely testing for Hardy-Weinberg equilibrium, linkage disequilibrium and evidence for bottleneck events) applied to the microsatellite data revealed no evidence for deviation from Hardy Weinberg equilibrium (HWE) in any of the populations, for any of the loci, nor any significant linkage between the loci. HWE is assumed in many microsatellite analyses, but deviations from this equilibrium may indicate important biological effects, such as

non-random mating or migration, although in this case the assumption of equilibrium could not be rejected for any of the populations. Some populations, when tested for evidence of a past bottleneck, showed heterozygote excess, but this was not significant once the Bonferroni correction was applied. Moreover, the increased observed heterozygosity was not seen in conjunction with low allele diversity, thus suggesting the populations had not been through a recent severe depletion in population numbers or loss of allele richness.

8.5.3 Comparison with schistosome data

The matrix correlations of genetic distance between pairs of schistosome isolates and pairs of *Biomphalaria* populations were not significant, for either COI or the microsatellite data, suggesting that the patterns of population structure seen in *Biomphalaria* do not match that observed for *S. mansoni*. This is intuitive given the low levels of population structure seen in *S. mansoni* from Lake Victoria (see Chapter 5) versus the strongly significant and geographically differentiated populations of *Biomphalaria* as demonstrated above.

The relationship of shedding or susceptibility to *S. mansoni* was investigated for an association with gene and/or nucleotide diversity. The only significant finding ($p = 0.005$) was for the 16S data, where populations found naturally infected with *S. mansoni* had higher gene diversity than those not found naturally infected. However, the relationship became insignificant once susceptible snails (i.e. those that were proven to be transmissive in the laboratory) were included. There was no relationship between infection status of the snail and gene or nucleotide diversity for the COI data, nor with proportion of private alleles for the microsatellite data.

8.6 Discussion

The results found here support previous findings relating to the population structure of intermediate host snails, in the observation of significant population structuring and also as compared to *S. mansoni*. However, some of the findings, such as the high levels of intrapopulation diversity (considering each of the 29 sites as a population), contrast with those for other species of *Biomphalaria*, suggesting that the specific conditions of Lake Victoria have created a system with its own particular population dynamics.

8.6.1 Explanations for genetic diversity patterns of Biomphalaria

The levels of genetic diversity seen within the *Biomphalaria* snails in Lake Victoria were very high, for all the markers used, although still consistent with that seen in other gastropod species and systems (Thaewnon-ngiw *et al.*, 2003, Zhao *et al.*, 2010; Rollinson *et al.*, 2009. However, the patterns of diversity in Lake Victoria are less obvious than seen in other studies. For example, other Nilotic species have been shown to lose genetic diversity with dispersal, hypothesized to be due to selection for successful genotypes, which in the case of hermaphroditic snails like *Biomphalaria*, can result in an increase in selfing rather than out-breeding (Vrijenhoek and Graven, 1992). This does not seem to have occurred with the snails from the 29 Lake Victorian populations studied here as they appear more or less universally diverse. However, this theory of how dispersal affects diversity could explain some of the pockets of low diversity observed, such as site U021 and U020. This would imply a centre of expansion based around the eastern/central portions of the lake, and moving westwards; given that the eastern part of the lake is deeper than the western part, this region could have acted as a refuge for *Biomphalaria*, for example during periods of desiccation of Lake Victoria, for example approximately 17,000 years ago, during the last glacial maximum (Stager and Johnson, 2008).

An alternative hypothesis is that diversity could have been acquired through the 'flush' phase of colonisation (Woodruff and Mulvey, 1997), whereby a new environment encourages out-breeding and increased diversity. This would indicate the western portions of the lake as the

ancestral habitat, prior to rapid, recent radiation eastward. This is consistent with hypotheses relating to the original introduction of *Biomphalaria* to Africa via a proto-*B. glabrata* which arrived into West Africa between 2.3 and 4.5 million years ago (Woodruff and Mulvey, 1997, Campbell *et al.*, 2000). It would also suggest that population differentiation should be greatest between populations on the western shore of the lake to those in the east, furthest away; indeed, K029, the most easterly population, is significantly more differentiated from U012, U020 and U021 (the three most western populations) than from all the other populations in the lake ($p = < 0.0001$). However, this could also simply be a case of diversity as a function of isolation by distance (Jarne and Théron, 2001), and indeed, the Mantel z-tests of all three molecular markers were significant, indicating that this is a valid hypothesis in this case.

Other studies of *Biomphalaria*, for example of *B. pfeifferi* in Madagascar, have revealed high levels of interpopulation variation but only very low levels of intrapopulation genetic differentiation (Charbonnel *et al.*, 2002a, Charbonnel *et al.*, 2002b). This was explained as being due to low levels of migration but also habitat stochasticity in the stream and pond systems inhabited by *B. pfeifferi*, which led to frequent population bottlenecks or indeed local extinctions and re-colonisations. *B. pfeifferi* are also known to be frequent in-breeders, which would assist in maintaining a metapopulation with seasonal local extinction events. In general, it is thought that habitat patchiness and dispersal ability will overwhelmingly define the population structure of freshwater molluscs (Stadler and Jarne, 1997). Lake Victoria posseses very different environmental conditions; as a very large, permanent lake, it rarely experiences the kind of environmental perturbations that characterise temporally transient ponds and streams. This homogeneity of habitat, with associated infrequent extinctions and opportunity for on-going genetic drift, could account for the high levels of intrapopulation diversity seen among the Lake Victoria *Biomphalaria* (Jarne and Théron, 2001). This applies mainly to the lake habitats; the marsh habitats seen along the fringes of Lake Victoria are likely more transient and so could conform more closely to the earlier *B. pfeifferi* findings; however, for these data, there was no significant difference in gene or haplotype diversity between lake and marsh sites, implying that extinction events are rare or not, in fact, driving changes to diversity. Similarly, the tests for bottleneck events were uniformly negative, suggesting that demographical events

are not overwhelmingly contributing to the patterns of diversity as seen in the Lake Victoria populations of *Biomphalaria.* The combination of high allele diversity together with general excess of heterozygotes might rather be attributable to out-crossing; little is known about the reproductive preferences of Nilotic *Biomphalaria,* and as such should be a future research direction.

8.6.2 Effect on transmission of intestinal schistosomiasis

The implications for transmission of *S. mansoni* are manifold. First of all, the high levels of intra- and interpopulation diversity may indicate very specialised local parasite compatibility. However, certain haplotypes were widespread around the lake; as seen in *B. pfeifferi* from elsewhere in Africa. If common haplotypes spread due to selection for traits independent of parasite susceptibility, the risk is that a highly transmissive genotype could become widespread in a particular locality (Campbell *et al.,* 2010). However, in this case, on-going informal compatibility tests have shown that isolates of *S. mansoni* from all three surrounding countries are compatible with snails from populations in all three countries. It is possible that detailed, quantified analysis of the shedding patterns from these different populations would reveal differential compatibility; however, the implication for public health is that *S. mansoni* is capable of being transmitted across a wide geographical range in Lake Victoria. Indeed, the low levels of population structuring seen in the parasite, as demonstrated in Chapter 5, is likely a function of terminal host migration and dispersal of the parasite into novel localities.

Indeed, the patterns of genetic distance between parasite and snail populations, as tested by the Mantel r-tests for the COI and microsatellite data, failed to show any significant correlation. This pattern of differential structure between trematode and snail host has been seen in other speices such as *B. glabrata* (Sire *et al.,* 2001), and generally the intermediate host has been shown to have stronger population partitioning than the parasite (Dybdahl and Lively, 1996), as seen in Lake Victoria. One difficulty with a survey of the magnitude of Lake Victoria is that temporal variation is inevitable; a smaller scale, micro survey of *Biomphalaria* could perhaps reveal more accurately the relationship between parasite and snail population structures.

One particularly interesting finding related to the transmission of schistosomiasis was the observation that there was no significant relationship between lower gene diversity in the snails and the incidence of naturally shedding or laboratory-susceptible populations. In fact, for the 16S gene, shedding populations were significantly associated with increased gene diversity. This is in contrast to some predictions made by evolutionary theory (van Valen, 1973); one hypothesis could be that the individual snails found shedding had low genetic diversity, although the populations as a whole were more diverse. On the other hand, in naturally shedding populations, frequent exposure to the parasite could act as a driver for increased diversity, as per the Red Queen hypothesis, if in some way rare genotypes were less able to be infected by *S. mansoni* (Webster *et al.*, 2001a). This emphasises the need to consider the genetics of individual snails as well as population-level analyses; unfortunately, the ability to achieve this is limited by the rarity of finding snails naturally infected with *S. mansoni*. With low sample sizes, statistically robustness is compromised, and thus generally population-level surveys are all that can be carried out with confidence.

8.6.3 Future research directions

The effect of intermediate host variation on schistosome genetic diversity and transmission has been considered in detail, in theoretical and experimental terms. However, little work has utilized wild populations for these studies; with the information on population genetic structure afforded by this chapter, it would be interesting to test directly what relevance the population differentiation observed would have on transmission of *S. mansoni*, if any. An ideal way of testing this would be to undertake a series of compatibility experiments, using isolates of *S. mansoni* from various sections of the lake and wild-caught populations of *Biomphalaria*, which, following the experiment, could be sequenced to measure any relationship between transmission success and population of origin or genetic diversity.

The question of how dispersal and diversity might be linked is also worth further scrutiny. In the case of the Lake Victoria populations here, one method might be to sample further from outside the margins of the lake, such as *B. sudanica* from Lake Albert and central Kenya, as this may reveal which region is ancestral to the Lake Victoria *Biomphalaria*. Another method could

be to look at other genetic markers, perhaps ones associated with reproductive success and fecundity, and see whether selection is indeed acting upon these markers in regions of particularly low diversity, such as the western shore of the lake, in Uganda. Significant levels of selection on these markers would provide support for the hypothesis that these populations are a recent introduction, and that low diversity is an effect of increased in-breeding in a favourable new habitat; observations or experiments on reproductive behavior could also test the validity of this theory, although ethological investigations on wild snails are difficult to undertake, and laboratory equivalents generally disrupted by confounding influences not present in the snails' wild environment.

On another note, although clearly also key to the on-going transmission of *S. mansoni*, the role of the terminal host in driving patterns of parasite distribution and genetic variability is much less commonly investigated than that of the intermediate host. If selection acts on a parasite based on host choice, then different genotypic patterns may emerge when faced with one host versus an environment where a variety of suitable terminal hosts abound. The latter case could be another mechanism for driving diversity in the parasite, with knock-on effects on the diversity and susceptibility of the intermediate host (Jarne and Théron, 2001). Terminal hosts have been observed to maintain much higher levels of infrapopulation genetic diversity than intermediate hosts (Théron *et al.*, 2004, Standley *et al.*, 2010); as schistosomes reproduce sexually in the terminal host but only asexually in the intermediate host, this could lead to a much more rapid accumulation of variation in the terminal host, with associated implications for the pathology of the infection. Similarly, a stable polymorphism in susceptibility could arise, whereby schistosomes are divided based on their compatibility with one, but not the other, host (Combes and Théron, 2000). Recent investigations into the schistosome species harboured by rodents in the Lake Victoria region have revealed that multiple terminal hosts should be something to consider in the future. Likewise, chimpanzees living on an island in Lake Victoria have been shown to be infected with 'human' forms of *S. mansoni* (see Chapter 10); the effect of anthropozoonotic and zoonotic transmission of intestinal schistosomiasis on the diversity of *S. mansoni* deserves attention.

8.7 Conclusions

This study revealed high levels of genetic diversity but also population structure among *Biomphalaria* in Lake Victoria, with a strong correlation between genetic distance and geographical distance between the populations. These findings differ from previous investigations by noting significant intra- as well as interpopulation variation, with significant spatial structuring correlated to geographical distance. Although this diversity can be explained by a number of different theories, it is only with further analysis that these various hypotheses can be tested. The main implications of the findings concern the patterns of evolution and transmission of *S. mansoni*; patterns of parasite population structure were not correlated with those of the intermediate host snails, nor was there a relationship between *Biomphalaria* population diversity and shedding or susceptibility to infection with *S. mansoni*. Further sampling, perhaps at a smaller geographical and temporal scale, could elucidate the relationship between infection and intermediate host diversity further. Similarly, the influence on the diversity and availability of different terminal hosts should be integrated into future research on the population genetics and evolutionary dynamics of *S. mansoni* in this region.

9 Micro-scale surveys of mollusc community diversity on Ngamba and Kimi Islands, Lake Victoria, Uganda: Population genetics of *Biomphalaria* and effect on transmission of *S. mansoni*

9.1 Abstract

Based on the preceding results of a Lake Victoria-wide survey for *Biomphalaria* snails, the same methods were applied to a study of two small islands in the lake, to investigate the effect of scale on environmental predictors of abundance. 40 sites from Kimi and Ngamba Islands (20 on each) were surveyed for *Biomphalaria* and all other freshwater gastropods; other environmental conditions were also recorded. Snails were examined for trematode infections and emerging *S. mansoni* cercariae DNA 'barcoded'. Snails from four populations from each island were also sequenced for population genetics analysis. Levels of phosphate concentrations were much higher on Kimi Island, confirming greater anthropogenic disturbance. However, community diversity was not significantly different between the islands, nor affected probability of transmission of *S. mansoni* or other parasites at that location. The same multivariate models that had been used for the large-scale survey of Lake Victoria were again applied to the data here, but in this case, no factors were found to be significant in the spatial model, although species diversity was a positive predictor in both non-spatial models. The genetics of *S. mansoni* and the population genetic structure of *Biomphalaria* followed similar patterns to that seen on a lake-wide basis. These findings suggest that smaller scale studies may prove useful as proxies for regional level investigations, with reduced logistical and resource output required. However, further research should also include surveys of terminal host parasite burden, as these will affect significantly even micro-scale dynamics of parasite-intermediate host interactions, and in their own right, terminal host parasite burdens may be important to public health.

9.2 Contribution of the author

The author was present throughout the duration of the Uganda 2010 field mission, during which time the data presented in this chapter were collected. The author selected and georeferenced the sites on each island, and was involved with all snail collections, shedding of snails and snail identifications., together with a team consisting of Dr Russell Stothard, Moses Adriko, Dr Emma Hobbs and a boat captain. The author also measured the water samples *in situ* and did all statistical analyses; water ion analysis was performed by the staff at the NHM's EMMA facility, as for Chapter 6 (see Acknowledgements). The *S. mansoni* extractions, amplifications, purifications and sequencing reactions were all performed by the author. The *Biomphalaria* extractions for Ngamba Island sites, as well as for all other snails across the two islands, were performed by the author. The extractions for the Kimi Island *Biomphalaria* were performed by Chloe McKeon, under the author's supervision. Miss McKeon also amplified, purified and sequenced the *Biomphalaria* population genetics samples and ran the analyses presented, again under the author's supervision (see Acknowledgements section). The figures presenting these findings in this chapter were drawn by the author.

9.3 Introduction

Research on the dynamics of disease transmission is often conducted on a national or regional scale, which is relevant to the level at which policy decisions regarding treatment interventions are taken. Similarly, when investigating risk factors for treatment or transmission, the environmental factors often brought into consideration act on a very large scale, which in turn ensures that research has a correspondingly broad perspective. These issues were explored in detail in Chapter 6, which dealt with the patterns of distribution of *Biomphalaria* at the regional scale of Lake Victoria, and identified statistically significant risk factors at this level. Of particular import was the finding that the presence of other species, and thus gastropod community assemblage, was a significant predictor of *Biomphalaria* presence at a lake-wide scale.

However, Chapter 6 also noted the presence of heterogeneities at a much smaller scale, both in terms of the presence of *Biomphalaria* but also in environmental conditions within the habitats in and around the lake. Moreover, it was noted in Chapter 7 that these environmental heterogeneities may have led to significant morphological differentiation within *Biomphalaria* in Lake Victoria; this observation was further supported by the high levels of population structuring reported in Chapter 8. Given the known importance of ecological and environmental conditions in shaping patterns of distribution and population dynamics in snail populations (Loreau and Baluku, 1987, Rollinson *et al.*, 2001), a more detailed investigation of variation at a smaller scale was warranted.

Firstly, there are factors that may vary widely at a micro-scale, or similarly which might not be detectable across larger distances (Turner, 1989). The importance of scale is well recognised, particularly when considering community composition and diversity (Collier, 2004); the freshwater gastropods that inhabit Lake Victoria are small, and as such interact on a scale of meters rather than kilometers. Similarly, the habitats which these different species exploit, such as the face of a submerged rock or the underside of a water hyacinth leaf, are specific focalities within a landscape. The synergies between habitat type, species composition and environmental factors can additionally produce variations in abiotic measurements such as water chemistry. Nor should the influence of human activity be overlooked; although capable of inflicting large scale ecological changes on aquatic landscapes, for example through the introduction of Nile perch to Lake Victoria in the 1960s (Witte *et al.*, 1991, Goldschmidt *et al.*, 1993), human communities can also have much more focalised, small scale effects (Resh *et al.*, 1988). For example, a lake site used primarily for washing clothes will have a very different chemical composition of the water, across a scale of a dozen or so metres along the shoreline, than one that is commonly used for docking boats with diesel or petrol engines. Similarly, in communities where basic sanitation facilities are lacking, certain portions of the shorelines may be designated 'latrine' spots, which not only may affect the environmental conditions of the water but will have a highly important impact on transmission of parasites to compatible intermediate host snails (Rudge *et al.*, 2008); in some cases water chemistry may additionally have a direct effect on the ability of the parasite to be infective or to develop (Esch and

Fernandez, 1994). The use of lakeshore sites for watering cattle will likewise impact on transmission dynamics, and again will vary between areas of the shoreline a few dozen meters away.

The acknowledgement of the presence of small scale changes in environmental conditions and the distribution of *Biomphalaria*, as detected during a large scale study, prompted a detailed investigation of the dynamics of *Biomphalaria* at a micro-scale level. Specifically, would heterogeneities of *Biomphalaria* and other gastropod distribution still be evident? Would the same environmental factors statistically predict the presence of absence of *Biomphalaria* and the incidence of infection with *S. mansoni* or other parasites? From a landscape genetics point of view, would the same population structure be seen in the snails at a micro-scale level, and indeed how would this relate to the genetic structure of *S. mansoni* within the same setting? Finally, would the influence of anthropogenic actions be detectable in either the environmental conditions observed at this scale or indeed in the community dynamics of freshwater gastropods, with associated parasitic infections? In this chapter, two small islands in Lake Victoria were surveyed in fine detail in order to throw some light on the questions posed above.

9.4 Methods

9.4.1 Study site selection and sampling protocol

Two islands, Ngamba and Kimi, were selected as the study sites for the micro-scale survey, based on their equivalently matched sizes, close geographical location to the mainland and each other, and ecological history. The islands are both approximately 100 acres in area, easily accessible from the mainland and reasonably close together, to control for large-scale environmental stochasticity but also for logistical ease of surveying between the two. The two islands, however, have contrasting ecological histories; although both were originally inhabited by humans, Ngamba Island only supported a small fishing community on its northern shore, who selectively logged the secondary growth forest but did not remove all the large trees (L. Mugisha, pers. comm.). In 1997, the fishing community was re-located and the island became a sanctuary dedicated to the rehabilitation of semi-captive, wild-born chimpanzees. As such, the

island and its surrounding waters, extending 200 metres from the shoreline, are now environmentally protected, with human alteration of the landscape tightly controlled. Kimi, on the other hand, currently supports two itinerant fishing communities, located at either end of the island, with an estimated population of approximately 2300 inhabitants, based on the number of households and an average Ugandan household size of 5 people (Uganda National Household Survey, 2002/2003; Uganda Bureau of Statistics, www.ubos.org).

Once the two islands had been identified, each was divided up into 20 assigned sampling 'zones', based on Google Earth images and aerial photographs, taken in 2008 and 2009. The centre of each zone was geo-referenced to allow for cross-checking with hand-held GPS equipment in the field (Figure 9.1).

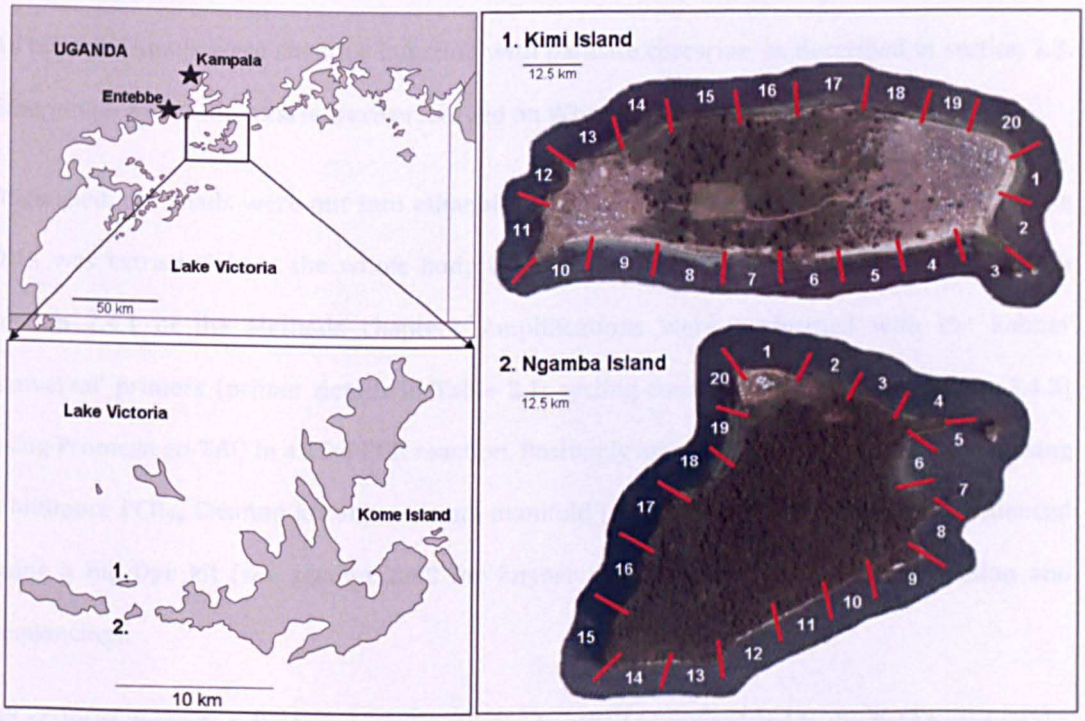

Figure 9.1 – Map showing locations of Ngamba and Kimi Islands and the 20 'zones' on each island
The top left map shows the general location of the Kome Island group (part of Mukono District) in Lakw Victoria; the bottom left map shows the relative positions of Kimi Island (1.) and Ngamba Island (2.) in the Kome Island group. The pictures on the right show the 20 georeferenced surveying zones on each island.

A section of each of the twenty zones was surveyed for a period of 15 minutes, with two collectors, as described in section 2.1.4 of the Methods chapter. Sites from Ngamba Island were given the designation 'NG' whereas sites from Kimi Island were designated 'KM', each followed

by the site number as denoted in Figure 9.1. All gastropods found in the sampling time were counted and all *Biomphalaria*, *Bulinus* and *Lymnaea* retained; a representative sub-sample of other gastropods from each island was also kept. Environmental conditions such as water temperature, microconductivity (μS), pH, habitat type, substrate type, depth and relative turbulence were also recorded for each zone. If more than one habitat type was present within a zone, all were sampled for 15 minutes, with both collectors, and given a sub-designation of 'A', then 'B', and so on. A water sample was also collected from each for laboratory analysis of cation and anion concentrations (see section 2.6.1. for more details on these methods of analysis).

9.4.2 Sample processing

All collected snails were shed for infection with parasite cercariae, as described in section 2.3. Emerging *S. mansoni* cercarie were preserved on Whatman FTA® indicator cards.

Once shed, the snails were put into ethanol for transportation back to the laboratory, where DNA was extracted from the whole body using a standard CTAB extraction procedure (see section 2.4.1 of the Methods chapter). Amplifications were performed with the Folmer 'universal' primers (primer details in Table 2.1; cycling conditions in Methods section 2.4.3) using Promega go-TAQ in a 25μl PCR reaction. Positively amplified samples were purified using a Millipore PCR_{96} Cleanup kit on a vacuum manifold (Millipore, Billerica, USA) and sequenced using a Big Dye kit (see section 2.4.3 for further details on amplification, purification and sequencing).

All of the gastropods collected were putatively identified to genus level in the field based on the key in Brown (1994). Those that could not be identified in this way had DNA extracted and were DNA 'barcoded', also using the Folmer primers; all methods were as above and in the relevant Methods sections.

9.4.3 *Statistical analyses*

The counts of the different gastropod species abundances per location were used to calculate three diversity indices for each sampling zone: Shannon's Index, Simpson's Index and Fisher's alpha. All calculations were performed in R (Ihaka and Gentleman, 1996).

The collated environmental, *Biomphalaria* count and snail diversity data were used in a series of statistical models to determine predictive factors for *Biomphalaria* abundance. Sites where snails were found shedding human and/or non-human cercariae were also used as an outcome; count data were modelled using a negative binomial regression due to the high variance of the count data (apart from the random effects model, incorporation site, where Poisson distribution was used as the closest alternative) and shedding/not shedding data were modelled using a binomial regression. Variables were first put through a univariate model; if the p-value was 0.15 or lower, that variable was then included in the later multivariate models. Three multivariate models were run within a Bayesian inference framework; the first was a basic regression with no random effects, the second also a regression but with a parameter which accounted for random effects (called an 'exchangeable' model) and the third was an explicitly spatial model, incorporating location data along with the covariates and a parameter to account for random effects. Details of these models and the software used to calculate them can be found in section 2.6.2 of the Methods chapter.

9.4.4 *Extraction and sequencing of* S. mansoni

S. mansoni DNA was extracted from the Whatman FTA® indicator cards as per the method described in section 2.4.1. A fragment of the COI gene was amplified using the ASMIT primers (primer details and cycling conditions also in Table 2.1 and Methods section 2.4.3 respectively) and sequenced in both forward and reverse directions. Both snail and schistosome sequences were screened against the GenBank online database using the BLAST website (www.blast.ncbi.nlm.nih.gov). *S. mansoni* sequences were also compared to the existing database of ASMIT sequences, as collated by Stothard et al. (2009c) and Standley et al. (2010); see also Chapter 5 of this thesis.

9.4.5 Biomphalaria *population genetics analyses*

8 localities were selected for population genetics analysis on the COI sequences; on Ngamba Island, these were NG01, NG03, NG11 and NG18. On Kimi Island, due to low numbers of snails, KM05, KM06 and KM07 were grouped into a single 'population', as were KM12, KM13, KM14, KM15, KM16, KM17 and KM18. KM02 and KM03 were also included. Ten snails from each were sequenced, as described above, and the resulting sequences used in a variety of analyses, including a neighbour-joining distance-based phylogenetic tree (section 2.4.5), a median-joining network (see 2.4.6 of Methods chapter) and a series of population genetics tests, all executed in Arlequin v 3.11 (Excoffier *et al.*, 2005). Tests included analysis of molecular variance (AMOVA), population structure analysis (including calculating of pairwise F_{ST} values), calculation of Tajima's D and Fu's F_S for neutrality and a Mantel test for correlation between genetic and geographical distance. Distances were calculated based as uncorrected p-distances, in MEGA v 4.0. Details of these tests can be found in section 2.4.6 of the Methods chapter.

9.5 Results

9.5.1 Snail diversity surveys

In total, 9 different types of gastropod were found, comprising of 5 pulmonate groups and 4 prosobranch groups: *Bellamya, Biomphalaria, Bulinus, Burnupia, Gabbiella, Lymnaea, Melanoides, Pila* and small planorbids. Of these, 6 could be identified to species level, based on Brown (1994): *Biomphalaria choanomphala, Lymnaea natalensis, Melanoides tuberculata, Gabbiella humerosa, Pila ovata* and *Burnupia stuhlmanni.* The *Bulinus* species were suspected to be *Bulinus trigonus* or *Bulinus transversalis* (chromosomal counts to distinguish between these two species could not be performed), and the *Bellamya* were likely to be *Bellamya unicolor.* Due to difficulties with the taxonomy of several of these genera (including *Biomphalaria*; see Chapter 7) it was decided to only attempt identification to the generic level. The small planorbids posed further taxonomic difficulties; they were morphologically divisible into two groups, which, taking into account the DNA sequence data and habitat information, likely corresponded to *Gyraulus/Ceratophallus* and *Segmentorbis.* Unfortunately, the size of the snails and paucity of the samples collected *in situ* prevented detailed anatomical dissection of the snails from which DNA had been amplified, which prevented a higher resolution of identification. As such, it was decided not to split the small planorbids for subsequent analyses.

Most notably, the distribution and relative abundance of each of these gastropod groups varied between the islands as well as between the site. No *Burnupia*, for example, were found on Kimi Island, whereas Ngamba Island possessed all types of gastropop (see Figures 9.2 and 9.3).

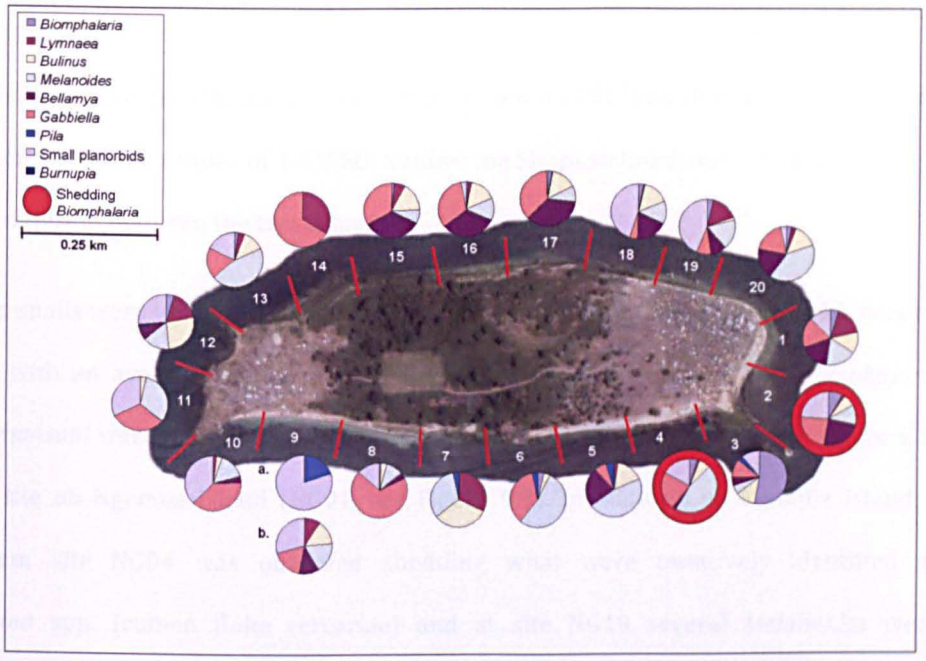

Figure 9.2 – Map of relative abundance of freshwater gastropods on Kimi Island

The proportion of each gastropod genus found can be seen in the pir charts; the genus represented by each colour can be found in the key in the top left of the figure. The red circles outlining the pie charts indicate sites where *Biomphalaria* shedding *S. mansoni* were collected.

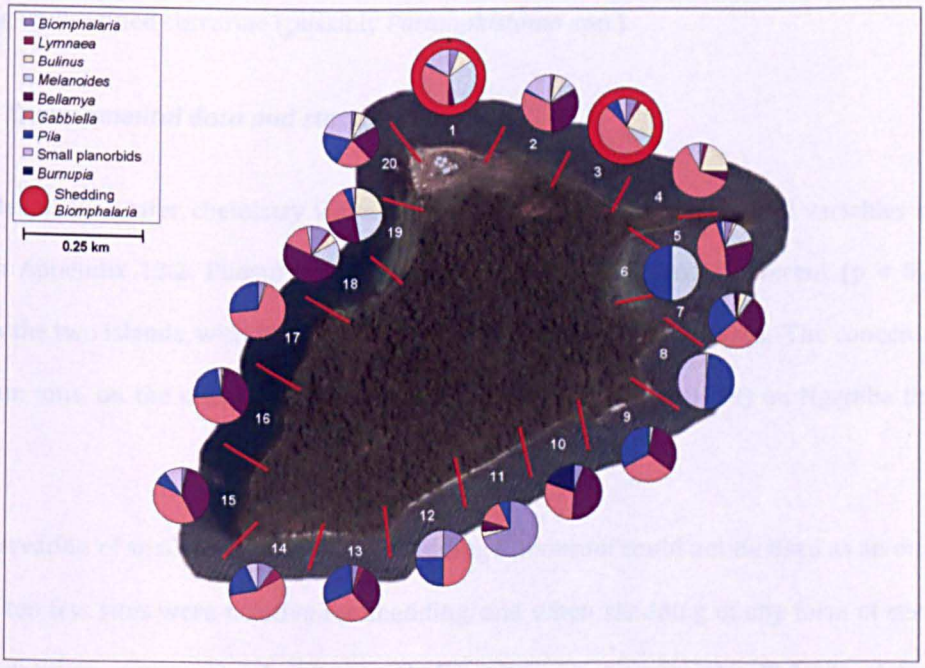

Figure 9.3 – Map of relative abundance of freshwater gastropods on Ngamba Island

The proportion of each gastropod genus found can be seen in the pir charts; the genus represented by each colour can be found in the key in the top left of the figure. The red circles outlining the pie charts indicate sites where *Biomphalaria* shedding *S. mansoni* were collected.

However, despite this, Kimi Island had a mean diversity (as measured by the Shannon Index) of 1.53977, which was greater (though not significantly: $p = 0.315$) than that of Ngamba Island, which had a mean Shannon Index of 1.43880. Neither the Simpson Index nor Fisher's alpha was significantly different between the two islands.

Biomphalaria snails were found at 16 of the 20 sites on Ngamba Island and 14 of the 21 sites on Kimi Island, with an average abundance of 7.2 and 5.1 per site, respectively. *Biomphalaria* shedding *S. mansoni* were found at two sites on Kimi Island (KM02 and KM04, see Figure 9.2) and on one site on Ngamba Island (NG01, see Figure 9.3). In addition on Ngamba Island, a *Lymnaea* from site NG04 was observed shedding what were putatively identified as *Paramphistoma* spp. (rumen fluke cercariae) and at site NG19 several *Melanoides* were shedding unidentified, non-human cercariae. Kimi Island had many more sites with snails shedding non-*Schistosoma* cercariae: *Bulinus* from KM02, KM04, KM07, KM08 and KM20 were shedding *Paramphistoma* spp.; *Biomphalaria* from KM02 were shedding *Strigea* spp; *Melanoides* from KM10 were shedding *Strigea*-like cercariae and *Gabbiella* from KM14 were shedding unidentified cercariae (possibly *Paramphistoma* spp.).

9.5.2 Environmental data and statistical models

Full tables of the water chemistry values and the categorical environmental variables can be found in Appendix 13.2. Phosphate levels and pH were significantly different ($p < 0.0001$) between the two islands, with Kimi having much higher levels in both cases. The concentration of calcium ions, on the other hand, was significantly higher ($p = 0.0179$) on Ngamba than on Kimi.

The observation of snails shedding or not shedding *S. mansoni* could not be used as an outcome because too few sites were positive for shedding, and when shedding of any form of cercariae was used as an outcome, only gastropod diversity, as measured by Fisher's alpha, was significant for the exchangeable model ($p = 0.030$). However, the odds ratio was extremely high with wide confidence intervals (OR = 8213.0, 95% CI = 3.3-20276446.9), suggesting a poor fit of the model.

The only variable that was significant across all three of the multivariate models used was Fisher's alpha, which was a positive predictor of *Biomphalaria* abundance in the exchageable model (OR = 1.918, 95% CI = 1.120-3.295). The parameters of variance were relatively similar between the exchangeable and spatial models, with values of sigma of 0.119 and 0.544 respectively. The parameter for spatial variance, sigma.sp, was 0.595 for the spatial model. The value of rho, the decay parameter, in the spatial model was 2184.0; this corresponds to a limit of autocorrelation of 0.152km.

9.5.3 DNA sequencing of S. mansoni

In total, 12 partial COI sequences were obtained from cercariae emerging from snails at sites on Kimi Island, with a further 2 sequences from cercariae from Ngamba Island. These 14 sequences corresponded to 6 unique haplotypes (labelled H1, H14, H15, H38, H100 and H138; see Standley et al., 2010), one of which had not been observed previously in Lake Victoria (H138; see Table 9.1).

Sequence name / Snail ID / Site ID	Base number (base 1480 = base 1 of ASMIT fragment)															
	1556	1562	1603	1609	1621	1636	1646	1696	1717	1720	1729	1759	1760	1771	1801	1864
H1	T	A	A	A	T	C	T	T	A	T	G	C	A	A	T	A
Snail A / KM04	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Snail A / KM04	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Snail A / KM04	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Snail B / KM02	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
H14	T	A	A	G	C	C	T	C	G	T	A	T	A	A	T	G
Snail A / NG01	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
H15	C	A	G	A	T	C	T	T	A	T	G	C	A	G	C	A
Snail A / KM04	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Snail A / KM04	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Snail A / KM02	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Snail A / KM02	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
H38	T	A	A	A	T	C	T	T	A	A	G	C	G	A	T	A
Snail A / KM02	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
H100	T	G	A	A	T	C	C	T	A	A	G	C	A	A	T	A
Snail A / NG01	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
H138*	T	A	A	A	T	T	T	T	A	T	G	C	A	A	T	A
Snail A / KM02	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Snail A / KM02	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Snail B / KM02	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Table 9.1 – Base changes between the 6 unique haplotypes collected from cercariae from Ngamba and Kimi Islands

'-' indicates the same base pair as the sequence above. '*' denotes the previously unrecorded haplotype of *S. mansoni*.

GenBank accession numbers for the sequences are as follows: H1 (GQ415163), H14 (GQ415176), H15 (FJ750538), H38 (GQ415215), H100 (GQ415284) and H138 (HQ122388). H1 had previously been observed in cercariae and miracidia (from human stool) in Lake Victoria, whereas H38 had only previously been collected from shedding *Biomphalaria* and H14, H15 and H100 had only been collected from human stool samples. There were 16 base changes across all 6 haplotypes. There was no cross-over between Kimi and Ngamba Island in terms of the *S. mansoni* COI sequences; Kimi Island's haplotypic diversity was made up of 4 sequences of H1, 4 sequences of H15, one sequence of H38 and 3 of H138. H14 and H100 were each observed once on Ngamba Island. There was, however, mixing between the sites where cercariae were collected on Kimi Island; both H1 and H15 were shared between KM02 and KM04, and two different snails from KM02 were found shedding H138.

9.5.4 Population genetics of Biomphalaria

In total, 43 unique COI haplotypes were recovered from 78 sequences. One sequence was obtained from a small planorbid and was a further unique haplotype, after BLAST-ing (www.blast.ncbi.nlm.nih.gov), it was discovered to match closely with *Gyraulus* and was included in the phylogenetic analysis as an out-group. When compared to the Lake Victoria-wide dataset of *Biomphalaria* COI haplotypes, it was found that none of the Ngamba and Kimi sequences corresponded to existing haplotypes, and so were labelled H128-H170, sequentially on from the earlier dataset (Genbank accession numbers HM768907-HM768949). Apart from H140, which was found at sites NG01 and KM03, all of the haplotypes were completely divided between the two islands (Figures 9.4 and 9.5).

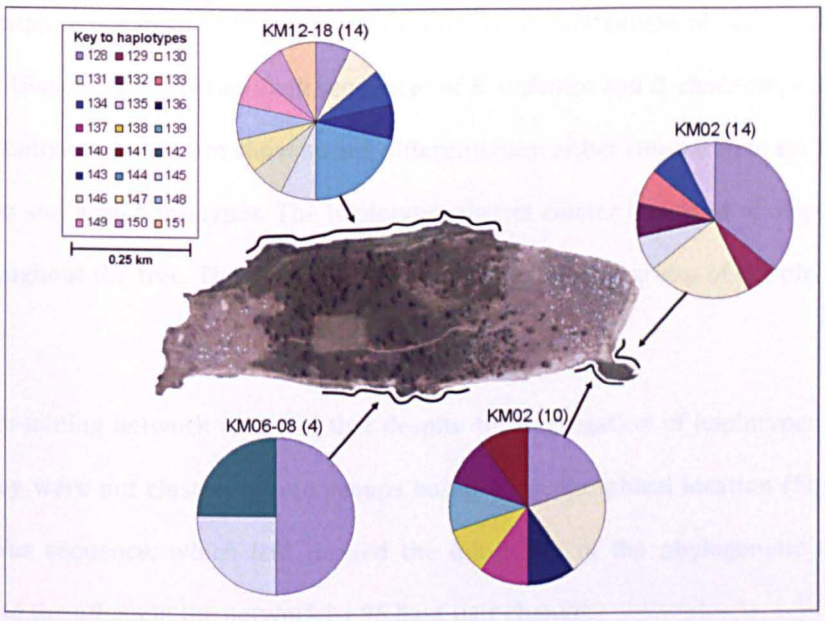

Figure 9.4 – Map of *Biomphalaria* haplotype diversity on Kimi Island

Haplotypes H128 – H151 were found on Kimi Island. The black lines demonstrate the location of the sites or combination of sites for each population; the text above each pie chart lists the sites by name and the figure in parentheses denotes the sample size of snails sequenced.

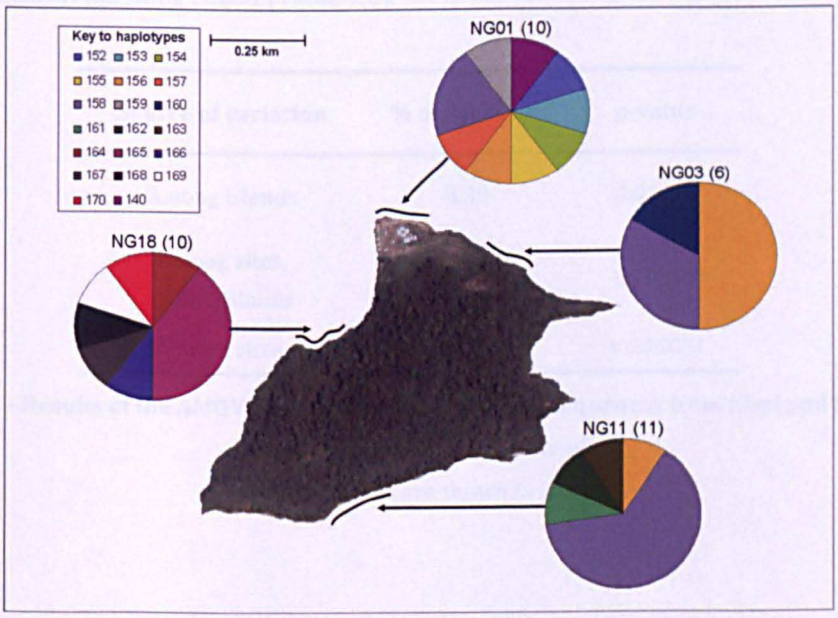

Figure 9.5 – Map of *Biomphalaria* haplotype diversity on Ngamba Island

Haplotypes H152 – H170 were found on Ngamba Island, as well as H140 which was also found on Kimi. The black lines demonstrate the location of the sites for each population; the text above each pie chart lists the sites by name and the figure in parentheses denotes the sample size of snails sequenced.

The phylogenetic tree showed some structure with a few more divergent clades but the low bootstrap support for most of the clades shows the close relatedness of many of the other haplotypes (Figure 9.6). The GenBank sequences of *B. sudanica* and *B. choanomphala* cluster within the samples, and do not show strong differentiation either compared to each other or the Ngamba and Kimi haplotypes. The haplotypes do not cluster by island of origin but are mixed throughout the tree. The *Gyraulus* sequence forms the out group of the phylogenetic tree.

The median-joining network revealed that despite the segregation of haplotypes between islands, they were not clustering into groups based on geographical location (Figure 9.6). The *Gyraulus* sequence, which had formed the out-group of the phylogenetic tree, was connected to the others in the network by 96 base pair changes.

The population genetics tests further explored the population structure between and within the studied localities. When the sites were grouped by island, the AMOVA test attributed significant levels of variation to the island and the individual population level, but not between sites on the same island (Table 9.2).

Source of variation	% of variation	p-value
Among islands	8.39	*0.02346*
Among sites, within islands	3.14	0.10264
Within sites	88.47	*< 0.00001*

Table 9.2 - Results of the AMOVA test for *Biomphalaria* COI sequences from Kimi and Ngamba Islands, when grouped by island

Significant values are shown in italics.

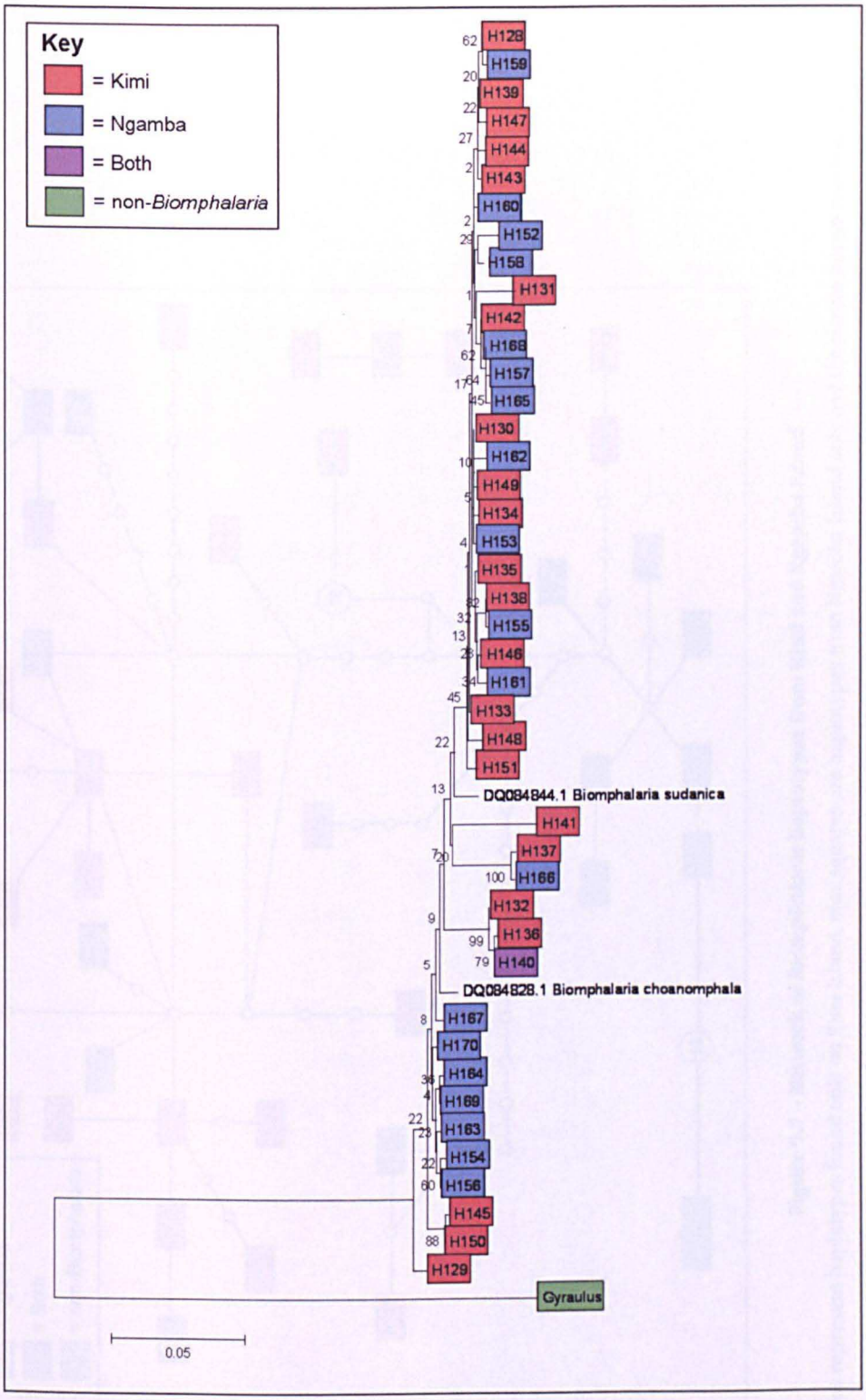

Figure 9.6 – Distance-based neighbour-joining tree of the 44 unique COI haplotypes observed in *Biomphalaria* from Kimi and Ngamba Islands

The model used was the GTR+gamma, based on earlier *Biomphalaria* model testing (see Chapter 7) and 500 bootstrap replicates. Haplotypes found only on Kimi are enclosed in a red box, whereas those from Ngamba are in blue, and H140, the haplotype found on both islands, is in purple. The tree is unrooted.

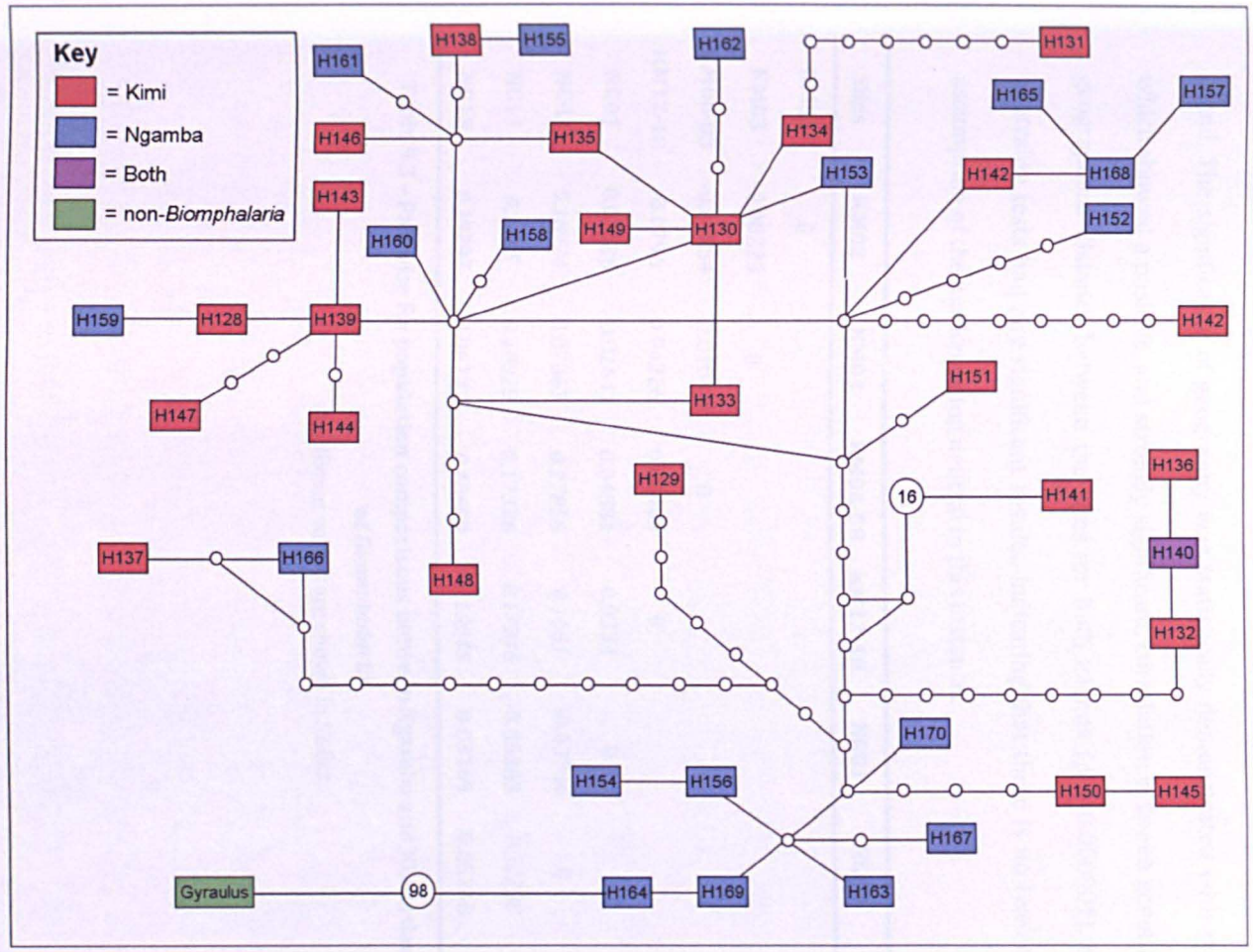

Figure 9.7 – Network of *Biomphalaria* haplotypes from Kimi and Ngamba Island

Red squares represent haplotypes found only on Kimi Island, blue squares are haplotypes from Ngamba Island only and the purple square represents the only haplotype shared between the two. Each circle represents a 'missing' haplotype, or can otherwise be thought of as a single step-wise base pair change.

The F_{ST} pairwise population comparisons reflected the results of the AMOVA, with most of the Ngamba populations being significantly differentiated from the Kimi populations (Table 9.3). The Kimi populations were less differentiated from each other than the Ngamba populations, with no significant F_{ST} values. The single significant pairwise comparison within the Ngamba populations was between NG03 and NG11, on opposite sides of the island. The significance of geography was statistically demonstrated with the Mantel test, which showed a positive, and strongly significant, correlation between genetic distance and geographical distance between the sites for both islands ($p < 0.000001$). Neither of the neutrality tests had any significant results, indicating that there is no reason to reject the assumption of the marker being neutral in this instance.

Sites	KM02	KM03	KM06-08	KM12-18	NG01	NG03	NG11	NG18
KM02	0							
KM03	-0.00223	0						
KM06-08	-0.04854	-0.01044	0					
KM12-18	-0.0796	0.04226	0.00405	0				
NG01	0.03186	-0.02641	0.04881	0.0334	0			
NG03	*0.18924*	0.07067	*0.27856*	*0.1561*	-0.03746	0		
NG11	*0.1645*	*0.10029*	0.17326	*0.17946*	0.06363	*0.22238*	0	
NG18	*0.19297*	0.06773	*0.18659*	*0.18045*	0.03169	0.05248	0.07704	0

Table 9.3 – Pairwise F_{ST} population comparisons between Ngamba and Kimi Island populations of *Biomphalaria*

Significant values are shown in *italics*.

9.6 Discussion

The aim of the study was to use a multidisciplinary approach to investigate micro-scale transmission biology of intestinal schistosomiasis on Kimi and Ngamba Islands. The results showed high levels of diversity and heterogeneity, across the myriad types of data collected and analysed. This demonstrates the complications inherent in studying the dynamics of this disease, even at a very small scale; however, it also provides important information as to how small scale surveys may reflect similar patterns at a regional level, and thus be a useful tool for on-going research in this field.

9.6.1 Environmental predictors of Biomphalaria *presence*

A key reason for measuring environmental data and community diversity metrics was to model these variables see if they could be used to predict the presence or absence of *Biomphalaria* snails. In this case, all the *Biomphalaria* found were of the *B. choanomphala* morphotype; although probably not a true species-level distinction, as earlier molecular studies and on-going research have shown (Jørgensen *et al.*, 2007), the morphotype does have implications for habitat selection and thus environmental predictors.

Across a Lake Victoria-scale, sulphate concentration, nitrate concentration, pH and the number of other species observed at a site were all significant predictors of *B. choanomphala* presence in all models tested (see Chapter 6). Interestingly, the only one of these which was also significant at the micro-scale level, as examined here, was the influence of gastropod diversity; Fisher's alpha was a positively significant predictor of *B. choanomphala* abundance on Kimi and Ngamba, for the non-spatial and exchangeable models. This contrasts with previous work on other intermediate host snail species, whereby it was theorised that competitive interactions with other gastropods or reproductory inhibition by their excretory products would result in a negative density dependent relationship (Appleton, 1976, Jordan *et al.*, 1980, Woolhouse and Chandiwana, 1990).

Further to the influence of gastropod community assemblages, an observation that might have implications for the distribution of *Biomphalaria* was the presence of significantly more *Pila ovata* on Ngamba than on Kimi; mean abundance on Ngamba was 34.5 whereas it was only 0.9 on Kimi ($p = 0.001$). Lab experiments have shown that presence of *Pila* could act as a control on the multiplication of *Biomphalaria* populations (Mkoji *et al.*, 1998). However, in this case there was no significant association, either positive or negative, between the presence or abundance of *Biomphalaria* and *Pila*, suggesting that outside of experimental conditions, where for example food supply and macrophytic cover were controlled, other factors are more important in determining the distribution and structure of gastropod assemblages.

The lack of significant environmental factors in the spatial model could be an effect of the small scale of study; given the large amount of heterogeneity in environmental conditions, even at this level, it might be that there was too much 'noise' in the model to distinguish significant variables. It could also be that there are temporal variations in the environmental conditions, which influence *Biomphalaria* abundance but only after a lag period. The shoreline could be in a state of constant dynamic flux; if snails are sampled just before or after conditions change, the environmental variables measured may not accurately reflect the extent to which the community has adapted to these changes.

9.6.2 Spatial and temporal landscape heterogeneity

As with the survey of Lake Victoria as a whole, this micro-scale survey revealed high levels of environmental and landscape heterogeneity. Habitat type, substrate, turbidity and depth varied on a scale of meters around the perimeter of each island; perhaps even more surprisingly, water chemistry composition also changed at a similar scale. As such, it was not unexpected to discover the wide variations in gastropod species distributions and relative abundances between the two islands but also even between sites.

Another aspect of heterogeneity which was not formally measured in this survey but which has been observed through repeated visits to the two study islands is temporal change. For the purposes of this investigation, all surveys were conducted across only three days, in

order to remove time or seasonality as a variable. However, the aerial photographs and Google Earth images reveal substantial changes in habitat type and even in the shape of the shoreline, across the two years' worth of visits to the island group (Figure 9.8). The transient nature of water bodies has been noted as a defining feature of many freshwater snail populations (Rollinson *et al.*, 2001, Charbonnel *et al.*, 2002a), but this is not directly the case for Lake Victoria. Rather, in this setting, a longitudinal study of temporal change to the depth, habitat types, substrate and water chemistry of Lake Victoria would be recommended, to further investigate the population dynamics and influence of environment on *Biomphalaria* snails and the transmission of *S. mansoni*.

Figure 9.8 - Series of aerial images showing the changes in shoreline, aquatic vegetation cover and water depth at the eastern end of Kimi Island

Note the lower water level in the October 2009 pictures, as seen by the wider beach and more exposed vegetation. The isthmus to the south of the island is also completely submerged in May 2009 and June 2010 but exposed in October 2009. The pictures also show changes in the number of fishing boats operating across different portions of the shoreline. Pictures taken by Russell Stothard (unpublished).

9.6.3 Anthropogenic influence on environmental variables

Despite the clear impact of human influence on the forestation and shoreline activity between the two islands, virtually none of the water chemistry values showed any significant differences. One reason behind this could be the averaging effect of only investigating differences at an island level; despite their small sizes, Ngamba and Kimi both have sites which are more frequented by humans or animals than others. Therefore, a more detailed analysis could observe levels of water usage and contact behaviour, and separate sites based on these results for a more direct comparison of the effect of human influence on environmental conditions. However, one factor was highly significant, and likely relates directly to anthropogenic effects; this was the concentration of phosphate in the water,

which was much higher on Kimi than on Ngamba ($p < 0.00001$). Sources of phosphate in freshwater can be from naturally occurring minerals, liberated for example by sulphuric acid from acid rain, but more commonly from waste-water, laundry detergents and fertilizer run-off. Given the poverty and lack of large-scale agriculture on Kimi Island, fertilizer is unlikely to be the major cause of elevated phosphate levels in the surrounding waters; more likely, the combination of lack of sanitation facilities and the frequent use of the lakeshore for washing clothes has resulted in the concentrations reported here. This finding is important because it shows a direct way in which anthropogenic influence can be measured, which, if found to be a factor predicting the presence of any intermediate host or vector species, could result in a link between human activity and disease transmission.

The effect of human influence might have been expected to have an effect on the gastropod community assemblage. For example, in a study comparing disturbed and undisturbed sites along the Kenyan shoreline of Lake Victoria, greater species diversity was recorded at the undisturbed sites (Lange, 2005). However, in this case, by all diversity metrics, diversity was higher on Kimi Island, albeit not significantly so. One difficulty with examining the community diversity of gastropods in Lake Victoria is the relative species paucity and the many taxonomic difficulties that plague even generic-level identifications. Lake Malawi, with its extensive list of endemic Prosobranchs and Lake Tanganyika, containing a suite of dozens of endemic Thalassoid species, would be ideal comparative study sites for a more detailed analysis of the effect of gastropod community assemblage on the presence and abundance of schistosome intermediate host snails.

9.6.4 Association between environment and snail infection status

Although not statistically linked to any environmental variables that were measureed in this survey, it was interesting to observe that more sites contained snails infected with some form of parasite on Kimi Island than on Ngamba. Given the high density of humans and livestock on Kimi Island, this is not particularly surprising, but it does add to the growing body of evidence reporting on the parasitological implications of human communities, particularly when encroaching into new, potentially pristine, environments. Unfortunately,

due to the low number of sites at which *Biomphalaria* were found to be shedding *S. mansoni*, it was not possible to determine whether any environmental variables were correlated with transmission of the parasite. Although some work has focused on this in the past, it has mostly focused on basic, remotely-sensed data such as temperature and humidity limits for development, and little has been done to investigate new parameters in recent years (Stirewalt, 1954, Wagner and Moore, 1959, Brooker, 2007). This would be a fruitful avenue of further, field-based study; a research programme dedicated to elucidating environmental factors associated with schistosome transmission in the wild would greatly improve our knowledge of the ecological dynamics and risk factors associated with the parasite.

9.6.5 Genetics of S. mansoni *and* Biomphalaria

This survey allowed for the direct comparison of genetic diversity and population structure of snails and parasites at a micro-scale as opposed to a regional scale. One factor which immediately differentiates the two is sample size; by virtue of visiting fewer locations, across a smaller area, fewer sequences were included in the micro-scale dataset. Similarly, this survey only considered schistosome haplotypes collected from shedding snails and not from terminal hosts. However, even despite these caveats, a high level of diversity was seen in the schistosome sequences collected, which corresponded to that seen throughout Lake Victoria. Moreover, the haplotypes found matched those previously observed in the lake, as far afield as Ukerewe Island in Tanzania and in the environs of Kisumu, in Kenya.

These data support the finding, as reported in Chapter 5 (and in Standley *et al.*, 2010) of widespread mixing of haplotypes of *S. mansoni* in Lake Victoria, which was theorised to be heavily influenced by patterns of human migration and itinerancy. While this holds for Kimi Island (where previous surveys have demonstrated high levels of human movement; see for example Chapter 3), it is interesting that widespread schistosome haplotypes are also being found in Ngamba Island snails. As a chimpanzee sanctuary, there is no permanent human habitation on the island, and only a small number of staff reside directly on the island. This could indicate how even low levels of human movement can spread genetic types of parasites across a large region; alternatively, it may indicate more extensive levels of human

contact with Ngamba Island, for example through fishermen defecating near the shoreline, than previously suspected.

As with the schistosomes, the population structuring of *Biomphalaria* around Ngamba and Kimi matched the larger patterns seen in Lake Victoria as a whole. First of all, the snails exhibited high levels of genetic diversity, with a large number of unique haplotypes. Secondly, and in direct contrast to the patterns of schistosome population structure, the populations of *Biomphalaria* here were well differentiated between sites, with low genetic flow between islands. However, there was one shared haplotype between the islands; there are several possible evolutionary scenarios that might explain the observed distribution of the haplotypes between the islands. For example, it could be that there is a deep-water population of *Biomphalaria* in the channel between the two islands, which has 'seeded' the shoreline of both Kimi and Ngamba at the nearest point, with populations differentiating as they move further around the shoreline of the two islands. Alternatively, one or the other island might have been colonized first, with the sites that share the common haplotype the point of cross-over from one to the other; finally, it might be that both islands were colonized separately, with recent crossing-over of snails between the island, for example via floating aquatic vegetation or boat traffic. Given the low differentiation and lack of genetic divergence between haplotypes from the two islands, the last scenario is probably the least likely, but sampling of snails from neighbouring islands, as well as the channel between Kimi and Ngamba, might further shed light on which is the correct explanation for the observed results. Figure 9.10 pictorally explains these three different scenarios.

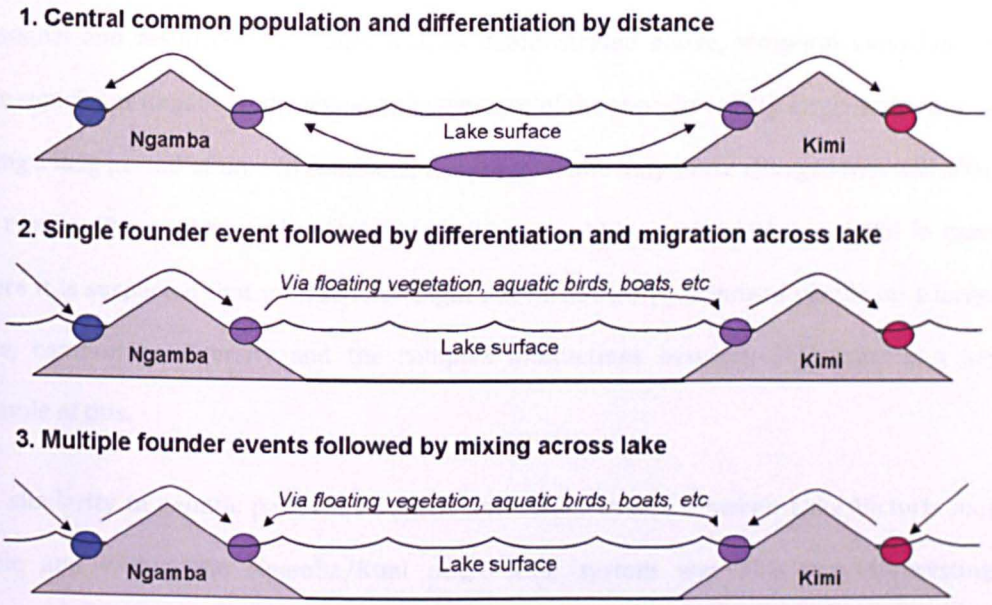

Figure 9.9 – Three possible evolutionary scenarios to explain *Biomphalaria* haplotype distribution on Kimi and Ngamba Islands

Given the evidence of low differentiation and lack of geographical separation of halotypes on the phylogenetic tree and the networks, scenario 3 is the least likely. Further sampling may elucidate whether scenario 1 or 2 is more likely.

The presence of several shedding snails, found on both islands, suggests that different snail genotypes can still be compatible with *S. mansoni*, lending weight to the hypothesis that local adaptation to the snail host may not be such a powerful evolutionary force in this system as has been previously thought. Similarly, although it is hypothesised that parasitism may drive the continued existence of rare, resistant genotypes, and thus genetically homogeneous populations may be more susceptible, in this case there was no relationship between the genetic diversity of a site and whether it was found to contain snails infected with *S. mansoni* or indeed any other kind of cercariae.

9.6.6 Implications

The most poignant finding from the above results is how the regional patterns observed in Lake Victoria as a whole have remained significant at a micro-scale level. The statistical models also reinforced the importance of gastropod diversity as a predictor of *Biomphalaria* distributions, matching the result found at a lake-wide scale. Regional or national surveys are logistically challenging and time-consuming; these results suggest that perhaps smaller

investigations could be used to infer patterns on a larger scale, without the associated personnel and resource costs. Similarly, as demonstrated above, temporal variation can have significant impact on the shape and structure of the shoreline; with large-scale surveys taking a long period of time to complete, temporal factors may cause changes that will affect the results. Using micro-scale surveys as a proxy could be particularly powerful in cases where it is suspected that interactions might not be as strong or indeed visible on a larger scale; community diversity and the complex interactions between organisms is a key example of this.

The similarity in genetic patterns of snails and schistosomes between Lake Victoria as a whole and within the Ngamba/Kimi micro-scale system was also very interesting. Moreover, it implies that attention should be focused on also surveying terminal host parasite prevalence and genetic diversity on both of these islands, where the question of scale also comes into play. Although most surveys are on a regional, national or district level, due to those being the scale at which policy is implemented, smaller scale surveys can show interesting patterns as well and provide more detail, for example the differential dynamics of different parasites (Brooker *et al.*, 2006).

This could also be case with regards to different terminal hosts; Kimi Island has been surveyed several times as part of on-going monitoring and control initiatives, but Ngamba Island, as it has no permanent human habitants, has been entirely overlooked. This is despite the presence of chimpanzees on the island, who are known to engage in water contact on a regular basis, and so could frequently be exposed to infection. Although the role of non-human primates as reservoirs of schistosomiasis was studied reasonably extensively in the 1960s and 1970s, few recent studies have looked at these animals as potential hosts, or indeed victims of anthropozoonotic transmission. Ngamba Island provides an ideal natural laboratory for such an investigation, which will be described in the next chapter. However, it should not be overlooked that there are numerous other locations where non-human primates come into contact with potentially infected water; Lake Tanganyika may be such an example, and given the high mollusc diversity found in the lake, could be a

particularly rewarding research site for exploring the dynamics between schistosome transmission, terminal host diversity as well as intermediate host community assemblage.

9.7 Conclusion

The micro-scale survey of Ngamba and Kimi Islands revealed substantial levels of heterogeneity, both in terms of the community assemblages found and the environmental conditions. The impact of human influence could clearly be seen in the comparisons of water chemistry values between the two islands and the incidence of infected snails. Despite few environmental factors being seen as statistically significant in this survey as compared to lake-wide investigations, the importance of gastropod species diversity on the distribution of *Biomphalaria* was reinforced, and should be included in future research efforts. Considering the small number of samples, a remarkable genetic diversity of *S. mansoni* sequences were also observed; populations of *Biomphalaria* were similarly genetically diverse. Moreover, the population structure of the snails matched that seen at a regional level, suggesting that small scale patterns could be useful models for understanding host population dynamics. A suggested further investigation should include survey of the terminal hosts on these two islands; given that non-human primates are found on Ngamba, this further has the potential to reveal zoonotic elements to the transmission of intestinal schistosomiasis in Lake Victoria.

10 Confirmed anthropogenic transmission of intestinal schistosomiasis to semi-captive wild born chimpanzees on Ngamba Island, Lake Victoria, Uganda

10.1 Abstract

Using a variety of stool, urine and serological diagnostic methods, 39 semi-captive wild born chimpanzees and 37 staff members at Ngamba Island Chimpanzee Sanctuary, Uganda, were examined for *S. mansoni* infection. Chimpanzees were unequivocally shown to be infected with intestinal schistosomiasis with seroprevalence in excess of 90% and three egg-positive cases in chimpanzees were detected, although the sensitivity of the diagnostic tests varied likely due to earlier prophylactic praziquantel (PZQ) treatment. Genetic 'barcoding' of miracidia hatched from chimpanzee stool revealed three DNA haplotypes also commonly found in humans living throughout Lake Victoria, including staff on Ngamba Island, as well as a novel haplotype. *Biomphalaria* snails were collected from four sites around the island and screened using PCR (polymerase chain reaction) for presence of *S. mansoni* DNA with 57.1% testing positive, indicating potentially high levels of exposure to the parasite and putative local transmission. Clearly, the anthropozoonotic potential of intestinal schistosomiasis on Ngamba Island is greater than previously thought. Moreover, the ability of chimpanzees to void schistosome eggs capable of hatching into viable miracidia further suggests these non-human primates may be capable of currently maintaining a local, presumably zoonotic, transmission of schistosomiasis independently of humans. The implications for management of captive and wild primate populations at risk of exposure are discussed.

10.2 Contribution of the author

The author was present on Ngamba Island for the duration of the main study presented in this chapter; however, the pilot study, in 2009, was conducted by Dr Russell Stothard and Dr Martha Betson, with assistance from Dr Lawrence Mugisha. The stool and urine collections from the chimpanzees were made by the trained keepers on Ngamba Island; parasitological techniques were carried out by a large team, consisting of the author, Dr Betson, Dr Stothard, Candia Rowell, and Moses Adriko. Serology of the chimpanzee frozen samples was carried out by the author and Dr Emma Hobbs. Real-time PCR of stool samples was executed by Dr Jaco Verveij of Leiden University (see Acknowledgements section). The extraction, amplification and analysis of the schistosome FTA material was done by the author; the extractions of *Biomphalaria* DNA used for the PCR screen for infection status had also been done by the author (as described in Chapter 9), as was the screen itself (with assistance from Ms Chloe McKeon and Mr Richard Kane of the NHM; see Acknowledgements). Clinical examinations of Ngamba staff members were performed by Dr Christoffer van Tulleken, Aaron Atuhaire and Dr Emma Hobbs.

10.3 Introduction

Zoonotic transmission of diseases is currently of interest to science, the media and policy makers. However, less commonly discussed are instances of 'anthropozoonoses', where so-called 'human' diseases are transmitted to wild animals, either in captivity or when living in close proximity to human communities which may have encroached upon previously 'pristine' habitat. As expanding human populations come into ever-closer contact with wild animals, and as conservation measures are forced to maintain wildlife populations *ex situ*, the potential for disease transmission from humans requires careful monitoring (Epstein and Price, 2009). As shown in chapter 9, the proximity of chimpanzees to areas of shoreline with known infective potential for intestinal schistosomiasis warrants a closer look at the possibility of anthropozoonotic transmission occurring in this context.

Previous reports on the incidence of intestinal schistosomiasis in non-human primates are almost exclusively confined to either the laboratory or past decades. The role of non-human primates, especially baboons, as natural reservoir hosts for the disease was investigated extensively in the 1960s and 1970s (Miller, 1960, Nelson, 1960, Fenwick, 1969), but little attention has been paid to this topic since the introduction of molecular tools. In 2000, *S. mansoni* eggs were found in the faeces of a baboon in Gombe National Park in Tanzania (Murray *et al.*, 2000), but chimpanzees were not found to be infected, which is in contrast to a prior study where two infected chimpanzees were identified (Nutter, 1993). Other reports of schistosomiasis in chimpanzees are mainly confined to experimental infections, where they have been shown to be permissive hosts of *Schistosoma*, including *S. mansoni* (Sadun *et al.*, 1966, Sadun *et al.*, 1970). It should be noted that past surveys have exclusively relied on stool sampling and microscopy as a diagnostic, which, based on a single sample, may not accurately reflect infection status as typically seen with humans, where single samples result in under-diagnosis (Booth *et al.*, 2003). Molecular and serological methods could further revise these diagnostic appraisals (Bergquist *et al.*, 2009).

Ngamba Island in Lake Victoria is a sanctuary for chimpanzees, which have either been found orphaned in the wild, confiscated at border crossings or rescued from poor living conditions in captivity. Created in 1998 by the Chimpanzee Sanctuary & Wildlife Conservation Trust (CSWCT), it is located approximately 23 kilometres from Entebbe in an area of Mukono District now known to be highly endemic for schistosomiasis (Standley *et al.*, 2009). Most of the 44 chimpanzees are free to roam daily throughout the 100-acre forested island and have access to the shoreline at many points. Observations of chimpanzee water contact have been made on numerous occasions, so it is hypothesised that these chimpanzees will have had some exposure to *S. mansoni*. This appeared more plausible after a malacological survey in 2008 confirmed the presence of infected *Biomphalaria* snails, the intermediate host of the parasite, around the island (see Chapter 6).

Based on the observation of infected snails on Ngamba Island in 2008, an informal spot check for schistosomiasis involving three different diagnostics was carried out on ten chimpanzees resident on the island. Four individuals tested positive by one or more

diagnostic, which resulted in blanket prophylactic treatment with praziquantel for all the chimpanzees in February 2009 for immediate animal welfare improvement and initiated the detailed research investigation reported here.

This pilot study was followed up by a more comprehensive survey in February 2010, which applied a range of specialised diagnostic tools, many never before used on chimpanzees, to determine the prevalence and intensity of infection with *S. mansoni* as accurately as possible. By also surveying the island for infected snail hosts, and the staff of the island for the disease, a comprehensive picture of the status and transmission potential of intestinal schistosomiasis on Ngamba Island was created. Genotyping parasite eggs and/or miracidia recovered from the chimpanzees and staff allowed evaluation of whether chimpanzees were infected with the same genetic lineages of *S. mansoni* as found in people, thus testing the hypothesis of anthropozoonotic transmission. The intention was to improve the health and welfare of the chimpanzees and staff living there, as well as provide recommendations for re-homed and wild primate conservation.

10.4 Methods

Following an initial malacological survey in 2008 revealing infected snails, an informal parasitology spot-check was carried out in April 2009, whereby sera from six chimps and stool and urine samples from 11 chimps were tested by SEA-ELISA, stool microscopy and circulating cathodic antigen (CCA) urine lateral flow tests respectively for infection with *S. mansoni*. The findings of this pilot study (the results of which are included below) demonstrated the possibility of infection of schistosomiasis in the chimpanzees and were used to inform the methodology of the resultant full survey, which took place in January-February 2010.

10.4.1 Sample collection

Blood samples had been collected during the annual health checks of the chimpanzees in September-October 2009. These had been centrifuged to isolate sera, which were then frozen at -20°C and stored at the Entebbe Virus Research Institute on the mainland, where they were later analysed.

Urine and faeces from chimpanzees were collected non-invasively by Ngamba Island staff caregivers, primarily during the morning feeding session. Urine was collected directly from individual animals, by catching it in named collection containers, attached to long poles, held through the bars of the sleeping enclosure. Stool produced by a particular individual was noted and its position recorded; it was gathered once the animals had been moved out of the sleeping enclosure for their daily free-roaming of the island. These samples were transported directly to the laboratory facilities for processing.

Staff working on Ngamba Island were also invited to submit stool, urine and finger prick blood samples, which were collected by a Vector Control Division (Ministry of Health) community nurse. All participants gave written informed consent prior to sample collection.

10.4.2 Analysis of the samples

A variety of diagnostic tests were carried out on the serum, stool, and urine samples collected. Sera were defrosted and used in SEA-ELISA tests for *Schistosoma* (IVD Inc., Carlsbad, USA) at a dilution of 1:40, as per the manufacturers' instructions. Details of the majority of the methods, unless otherwise indicated, can be found in sections 2.2.1 (for stool samples) and 2.2.2 (for urine and fingerprick blood samples) of the Methods chapter.

Urine samples were tested for haematuria using commercially available Hemastix® (Bayer, UK), to assess for the possibility of cross-reaction with *Schistosoma haematobium* infections (French *et al.*, 2007), although this parasite is not known to be endemic to the immediate region of Ngamba Island. A further 15µl of urine were used in a CCA urine lateral flow test. Results were scored as 'trace', 'single positive', 'double positive' or 'triple positive' based on band intensity, and cross-checked by a second observer.

Stool samples were divided and used to create double smear Kato-Katz slides according to published methods (Katz *et al.*, 1972). Where indicated by positive CCA tests or Kato-Katz slides, the remainder of the stool was used for Percoll separation and examination under the microscope for *S. mansoni* and other eggs (Eberl *et al.*, 2002) and for hatching using a Pitchford funnel (Pitchford, 1959). Schistosome miracidia which hatched from the eggs collected by the Pitchford funnel were collected in 2.5µl of water and placed on Whatman® FTA indicator cards for later extraction of DNA.

Approximately 0.5 g of sieved stool was suspended in 1ml 100% ethanol and used for a quantitative polymerase chain reaction (real-time PCR) detection method for *Schistosoma.* This method was not used elsewhere during the course of this thesis and so is described in detail in this section. For DNA isolation, 200 µl of faeces suspension was centrifuged and the pellet was washed with 1 ml of PBS twice. After centrifugation the pellet was resuspended into 200 µl of 2 percent polyvinylpolypyrolidone (PVPP) (Sigma) suspension and heated for 10 minutes at 100 °C. After sodium-dodecyl sulphate-proteinase K treatment (2 hours at 55 °C), DNA was isolated using QIAamp DNA-easy 96-well plates (QIAgen, Hilden, Germany). In

each sample, 10^3 PFU/ml Phocin Herpes Virus 1 (PhHV-1) was added within the isolation lysis buffer, to serve as an internal control (Verweij et al., 2001, Niesters, 2002).

Schistosoma real-time PCR including PhHV-1 as an internal control was performed using primers and probes as described previously (Obeng et al., 2008). The PCRs showed a high sensitivity and 100% specificity when tested against a panel (n=150) of parasite control DNA samples and DNA samples derived from well defined faecal samples positive for a range of parasitic infections and samples from cases without a known history of parasitic infections (Verweij *et al.*, 2007, Obeng *et al.*, 2008, Verweij *et al.*, 2009).

10.4.3 Barcoding of schistosome eggs and miracidia

Genomic DNA was extracted from the Whatman® FTA indicator cards according to standard protocol (Gower *et al.*, 2007). PCR amplifications were performed on a portion of the cytochrome oxidase sub-unit 1 (COI) gene using the 'ASMIT 1' and 'ASMIT 2' primers (Bowles *et al.*, 1992); amplification, visualisation, purification and sequencing were all carried out as per section 2.4.3 of the Methods chapter. Sequences were aligned and compared to the existing database of *S. mansoni* sequences to look for unique haplotypes (see Chapter 5).

10.4.4 *PCR screening of* Biomphalaria

Following on from the micro-scale survey of *Biomphalaria* from Ngamba Island as described in Chapter 9, the four sites that had been included in the population genetics analysis were screened for pre-patent infection with *Schistosoma*, using a novel assay (R. Kane, in prep.). These four sites (NG01, NG03, NG11 and NG18) were chosen due to their relatively high abundance of *Biomphalaria* snails and, in the case of the latter 3, possibility of contact with chimpanzees (Figure 10.1).

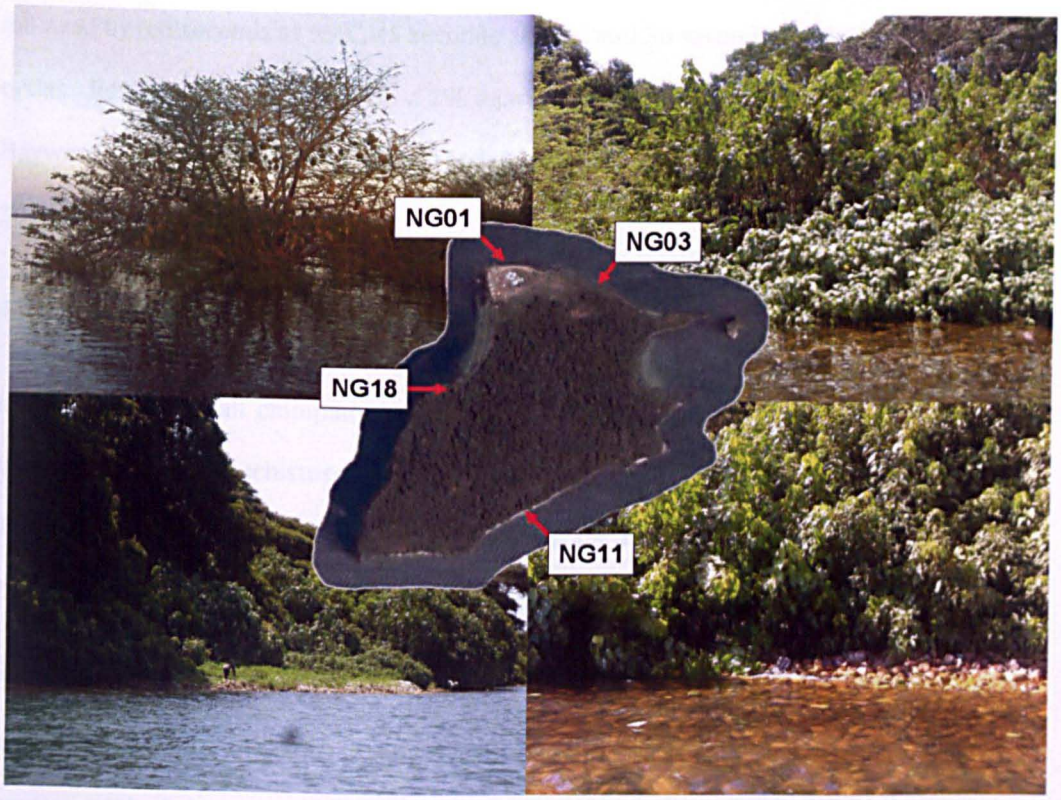

Figure 10.1 – Location map and photographs of the four sites from which *Biomphalaria* were screened for infection with *S. mansoni*

Sites NG03, NG11 and NG18 are accessible to the chimpanzees, and at each of these sites, chimpanzees were observed near the shoreline at the time of the photographs displayed here (in the bush, for site NG03).

Ten snails collected from site NG01 in 2008 were also included in the screen; one of the snails collected in 2008 had been patently infected and had shed schistosome cercarie, and so was used as a positive control. In total, 81 snails were examined. As this screen was not used elsewhere during the snail surveys of Lake Victoria, the full protocol is described below.

The whole snail was used to extract genomic DNA following a standard CTAB extraction protocol with a chloroisoamyl alcohol purification step and ethanol precipitation (Stothard and Rollinson, 1996), followed by resuspension in 50-100µl pure water (see section 2.4.1). 1µl of DNA template was used in a PCR amplification reaction using *Schistosoma*-specific primers that have recently been developed by Kane *et al.* (in preparation). Called RAKqIGSF and RAKqIGSR, they amplify a partial ribosomal intergenic spacer (IGS) amplicon for *Schistosoma* under the following cycling conditions: one cycle of 3.5 minutes at 95°C,

followed by ten seconds at 95°C, 45 seconds at 55°C and 30 seconds at 72°C, repeated for 55 cycles. Results were visualised on a 2% agarose gel, stained with GelRed™ (Biotium, Inc., Hayward, USA); snails infected with *Schistosoma* were identified by the presence of a band 309 base pairs in size.

10.4.5 Treatment and clinical examinations

In February 2009, all chimpanzees were treated with albendazole and praziquantel (the latter specifically for schistosomiasis) and then again with praziquantel in February 2010. The animals were also treated with ivermectin, for nematode infections, in September-October 2009, respectively.

All Ngamba Island staff members who volunteered samples were treated with albendazole, and those who tested positive by any diagnostic for schistosomiasis were also treated with praziquantel. Prior to treatment, each participant was also invited to take part in a full clinical consultation with a medical doctor (Christoffer van Tulleken) and community nurse (Aaron Atuhaire), who recorded reports of symptoms and past medical and treatment histories. As part of a wider community health outreach, malaria testing and treatment were also provided (data not shown).

10.4.6 Ethical clearance

Ethical clearance to take samples from the chimpanzees was granted by the Uganda Wildlife Authority and the Uganda National Council for Science and Technology (UNCST). Permission to survey and treat human adults was given by UNCST and the National Health System-Local Research Ethics Committee at St. Mary's Hospital in London as part of the national control programme for bilharzia and intestinal worms.

10.5 Results

Intestinal schistosomiasis was found in the chimpanzees by all the diagnostic tests used, both in the pilot study and in the comprehensive follow-up in 2010. Both sets of results can

be seen in Table 10.1; subsequent referrals to prevalence values or test results are based on the 2010 survey.

Diagnostic	Target group	Number of positives	Total number sampled	% Prevalence (95% CI)
Schistosoma ELISA	Chimpanzees (pilot)	4	6	66.7 (22.3-95.7)
	Chimpanzees (2010)	29	31	93.5 (78.6-99.2)
	Staff	14	23	60.9 (38.5-80.3)
CCA test	Chimpanzees (pilot)	1	11	9.1 (0.2-41.3)
	Chimpanzees (2010)	10	20	50.0 (27.2-72.8)
	Staff	15	34	44.1 (27.2-62.1)
Kato-Katz double smear	Chimpanzees (pilot)	1	11	9.1 (0.2-41.3)
	Chimpanzees (2010)	1	13	7.7 (0.2-36.0)
	Staff	5	27	18.5 (6.3-38.1)
Percoll	Chimpanzees (2010)	1	9	11.1 (0.3-48.2)
	Staff	3	25	12.0 (2.5-31.2)
Real-time PCR *S. mansoni*	Chimpanzees (2010)	13	24	54.2 (32.8-74.4)
	Staff	10	24	41.7 (22.1-63.3)

Table 10.1 – Table of diagnostic test results for chimpanzees and staff on Ngamba Island, including the pilot survey in 2009

'CI' stands for confidence intervals

Of the five diagnostic techniques, the highest prevalence for *Schistosoma* was observed using the SEA-ELISA for both the chimpanzees (93.5%) and the staff (60.9%) in 2010 (Table 10.1). Upon titration of the chimpanzee serum, a sample judged to be a 'strong positive' at 1:40 dilution could typically still be observed at a dilution of 1:600, indicating high antibody titres. Circulating cathodic antigen (CCA) urine lateral flow tests were less sensitive than the SEA-ELISA for detecting schistosomiasis, but the results followed the same pattern as the serum observations. In terms of intensity of the CCA test reaction, seven chimps had 'trace' readings, two were 'single positives' and one was a 'double positive'. The staff had five 'trace' readings, nine 'single positive' readings and one 'triple positive' reading.

The microscope-based diagnostics on stool (Kato-Katz and Percoll) detected one egg-patent chimpanzee, and two further individuals produced eggs which were collectable via the Pitchford funnel methodology and subsequently hatched into miracidia, proving the viability of this larval stage. All three of these chimpanzees had positive CCA tests and two had strong positive SEA-ELISA results; serum was not available for the third. Real-time PCR detection of *S. mansoni* DNA in stool revealed similar prevalence levels to those observed with the CCA tests, in both humans and chimpanzees (41.7% and 54.2%, respectively); seven chimpanzees and five humans had high levels of DNA in the stool (Ct < 30), whereas three chimpanzees and three humans had moderate amounts of DNA (Ct 30-35) and three chimpanzees and two humans had low amounts of DNA (Ct > 35). Of the individuals who were egg-positive (either by Kato-Katz or Pitchford), one chimpanzee had high levels of *Schistosoma* DNA as detected by real-time PCR, another had moderate levels and the third's faeces was not archived in ethanol and so could not be tested. In the staff, three had high levels of *Schistosoma* DNA in their faeces, and two had low levels. Only two chimpanzees and two humans had positive CCA tests yet no DNA detected in the real-time PCR; all four of these individuals had 'trace' CCA reactions.

The miracidia collected from the chimpanzees were genotyped for variation in the ASMIT fragment of the mitochondrial cytochrome oxidase sub-unit one (COI) gene. The sequences were compared to the database of haplotypes or 'barcodes' established in 2009 (Stothard *et al.*, 2009c) and enlarged in 2010 (Standley *et al.*, 2010), which had included previous analysis of cercariae from field-caught snails from Ngamba Island in 2008. The 12 haplotypes recovered from chimpanzee miracidia (11 from one individual and one haplotype from another individual) corresponded to three previously known haplotypes: H1, H16 and H36 (Genbank acquisition numbers GQ415163.1, GQ415179.1 and GQ415211.1 respectively), as well as two novel ones, labelled 'H132' and 'H133' (Genbank acquisition numbers HM031081 and HM055377). In total, there were six variable positions within the 396 base pair fragment between these 12 samples (Table 10.2).

Barcode Chimp sequence	Base number					
	1696	1706	1844	1868	1908	1935
H1 S3,S5,S8,S10,S11	C	T	T	T	A	T
H16 S7,K1	C	T	C	A	A	G
H36 S1,S2	C	T	T	A	A	T
H132* S4,S5	T	T	C	A	A	G
H133* S9	C	C	T	A	G	T

Table 10.2 – Table of base changes between haplotypes observed in miracidia from chimpanzees, compared to previously observed haplotypes

Haplotypes are labeled as per Chapter 5 and the GenBank accession information; sequences labeled 'S' came from a chimp called Sunday, whereas the 'K' sequence came from Kalema. '*' denotes a previously unobserved haplotype of *S. mansoni*.

Twenty miracidia collected from two staff members were also sequenced and corresponded to haplotypes H1 (the most common locally), H16, H17, H35, H38 and a novel haplotype, named H137 (Genbank acquisition number HM055378). There were thirteen variable positions between these six haplotypes (Table 10.3).

Although only one snail had been found shedding *S. mansoni* cercariae (see Chapter 9), based on the PCR screen results mean prevalence of schistosome DNA in *Biomphalaria* from the four sites, including snails from 2008, was 57.1% (Figure 10.2). The two snails that had shed cercariae, one from 2008 and the other from 2010, both tested positive for the screen, further supporting these findings.

Barcode Human sequence	Base number												
	1703	1754	1769	1844	1865	1868	1877	1890	1908	1935	1986	1995	2012
H1 G2, G8, G9, G10, E3, E4, E8, E10	T	A	T	T	A	T	G	A	A	T	T	C	A
H10 G4, G6, E5	T	A	T	C	A	T	G	A	A	T	T	C	A
H16 E1, E9	T	A	T	C	A	A	G	A	A	G	T	C	A
H17 E2	T	A	T	T	A	T	G	A	A	T	T	T	A
H35 G1, G3, G5	C	G	C	T	G	T	A	A	A	T	C	C	G
H38 G7	T	A	T	T	A	A	G	A	G	T	T	C	A
H137* E6, E7	T	A	T	T	A	T	G	G	A	T	T	C	A

Table 10.3 - Table of base changes between haplotypes observed in miracidia from Ngamba Island staff members, compared to previously observed haplotypes

Haplotypes are labeled as per Chapter 5 and the GenBank accession information; sequences labeled 'G' came from one (unnamed to preserve confidentiality) staff member; sequences labeled 'E' came from a different staff member. '*' denotes a previously unobserved haplotype of *S. mansoni*.

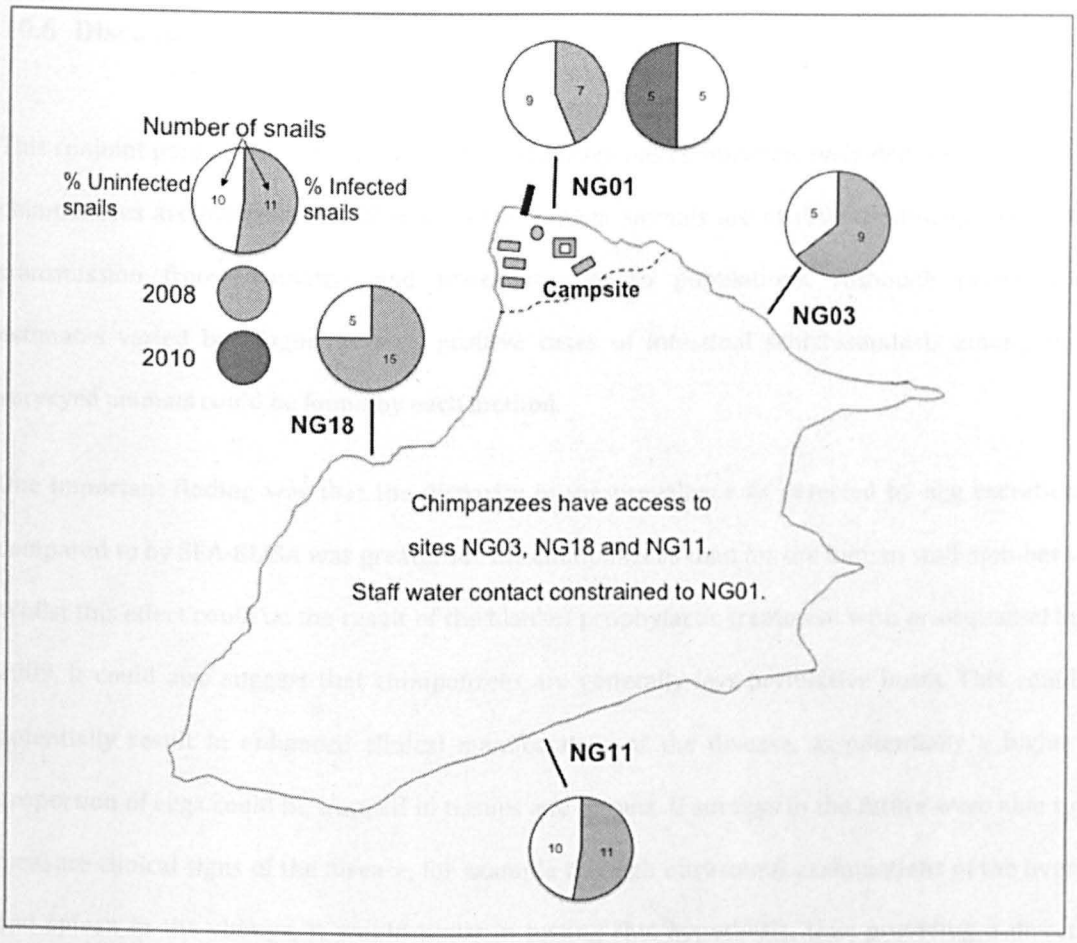

Figure 10.2 – Results of the PCR screen for *Schistosoma* DNA from *Biomphalaria* collected at sites NG01 (in 2008 and 2010), NG03, NG11 and NG18

The proportion of grey in the pie charts represents the percentage of snails that tested positive for the presence of *Schistosoma* DNA; the shade of grey denotes the year of sampling. The number inside the halves of the pie chart indicates the sample size.

Clinical consultation of staff revealed that only 55.6% of those surveyed had previously taken praziquantel. Eight of twenty-seven staff members who answered questions about their symptoms reported to have suffered abdominal pain, diarrhoea or other gastrointestinal symptoms in the last month. Two of these staff members were also found to be egg-patent for *S. mansoni* based on Kato-Katz double smears, and a third was CCA positive.

10.6 Discussion

This conjoint parasitological and malacological survey has comprehensively demonstrated that chimpanzees are infected with *S. mansoni* and these animals are at risk of anthropozoonotic transmission from sympatric and proximate human populations. Although prevalence estimates varied by diagnostic test, positive cases of intestinal schistosomiasis among the surveyed animals could be found by each method.

One important finding was that the disparity in the prevalence as detected by egg excretion compared to by SEA-ELISA was greater for the chimpanzees than for the human staff members. Whilst this effect could be the result of the blanket prophylactic treatment with praziquantel in 2009, it could also suggest that chimpanzees are generally less permissive hosts. This could potentially result in enhanced clinical manifestation of the disease, as potentially a higher proportion of eggs could be trapped in tissues and organs. If surveys in the future were able to measure clinical signs of the disease, for example through ultrasound examinations of the liver and spleen in the chimps, it would assist in testing this hypothesis, thus providing a direct health benefit as a result of the research. As an alternative, in humans there is evidence that immunosuppression, in patients co-infected with schistosomiasis and human immunodeficiency virus (HIV), can result in reduced egg excretion and thus complicate diagnosis, although these studies are contested (Karanja et al., 1997, Kallestrup et al., 2005). Retroviruses in chimpanzees might have a similar effect; on Ngamba Island, none of the chimpanzees surveyed are infected with simian immunodeficiency virus (SIV) but around 70% are infected with simian foamy virus (SFV), a closely related retrovirus (Mugisha *et al.*, 2010); it is not known what effect this underlying infection might have on egg excretion patterns of schistosomes or indeed susceptibility to contracting schistosomiasis, and should be borne in mind.

The staff members on Ngamba showed high seroprevalence of schistosomiasis and treatment with praziquantel was reported from over half of the participants. This could be responsible for the reduced prevalence seen in egg-detection diagnostic tests, although the relatively high number of positive CCA tests indicates the likely persistence of adult worms.

This study is the first to report genetic testing of the schistosome miracidia recovered from chimpanzees or in fact from any other non-human primate with 'naturally-acquired' infections. Common haplotypes were found in both chimpanzees and staff; the same haplotypes were also found abundantly in people throughout Lake Victoria. These results have important implications: chimpanzees can clearly act as hosts for forms of *S. mansoni* that are capable of infecting humans, and *vice versa*. From a conservation medicine perspective, this proves that in regions where wild or semi-wild chimpanzees come into close contact with human communities that carry the infection, they too are at considerable risk from contracting the disease contingent on water contact where intermediate host snails are present. Considering that other primate species have also been shown to be permissive to experimental schistosome infections (Sadun *et al.*, 1966), these findings should be taken on board by primate conservationists, who should integrate monitoring and, if the need arises, treatment, into existing disease management. Furthermore, it should be noted that if the Ngamba Island chimpanzees were to be reintroduced into the wild, they should be treated with praziquantel prior to release, to avoid establishment of the disease in a wild setting.

From the perspective of investigating the micro-scale transmission dynamics on Ngamba Island, the PCR screens of local *Biomphalaria* revealed a high level of non-patent exposure, from three sites accessed very rarely by humans and one site which is regularly entered by staff members as well as tourists. It should be mentioned that a later, additional malacological survey in June 2010 found a *Biomphalaria* snail actively shedding schistosomes at site NG03, which can be directly accessed by the chimpanzees. When sequenced, a portion of these cercariae were found to be *S. rodhaini*; as the screen does not distinguish between *Schistosoma* species, it could be that exposure to *S. rodhaini* contributed to the high PCR screen findings. No chimpanzees have yet been observed to be producing eggs that correspond to *S. rodhaini*, though their infection status for this parasite should be monitored carefully.

The general paucity of patent infections, by any *Schistosoma* species, observed across both surveys (in 2008 and 2010, the latter as described in Chapter 9) demonstrates how traditional measures of determining infection status may greatly underestimate the risk of a particular locality, especially given potential seasonal patterns and temporal fluctuations. As sanitation

facilities for staff on the island are very good, it is more likely that it is chimpanzee faecal contamination of the shoreline that has exposed these snails to infection, although 'hygienic bathing' from staff or visitor water contact could also play a role (Sow *et al.*, 2008). Moreover, although fisherman are required to stay at least 200 meters from the island at all times, there are reports of this regulation being ignored and boats approaching closer and even landing. Human faecal material is furthermore sometimes used as a fish attractant or bait. It is also worth noting that neighbouring communities have been shown to have high levels of schistosomiasis; 42.3% of children on Kimi Island, about 1.5 kilometres away from Ngamba, were found to be egg-positive for *S. mansoni*. When barcoded, several of the genotypes of parasites from Kimi corresponded to those collected from staff, chimps and/or snails on Ngamba, namely H1, H14, H16, H36 and H38.

As mentioned, the Ngamba Island chimpanzees have regular access to the water at three of the four sites that were screened for non-patent infection in the snails. One direct result of this research, from an animal welfare perspective, could be to reduce the water access available to the chimpanzees at points where snails are most abundant. Further testing of other populations could refine these restrictions; no snails were found at four sites, which could be retained as places where the chimps could access the water, should they so wish. Likewise, measures can be taken to prevent staff and visitors from exposing themselves to the disease through unnecessary water contact, or else they could be encouraged to take post-immersion measures (Ramaswamy *et al.*, 2003). A further immediate consequence of these surveys has been the addition of praziquantel, to treat for schistosomiasis, to the annual regime of de-worming as administered by the veterinarian on the island. This will immediately improve the health of those already suffering from the disease; a further line of research could be to investigate appropriate dosage of the drug for chimpanzees for maximum efficacy.

Above all, these results emphasise the need to carry out small scale parasitological surveys that include non-human primates as well as human communities, particularly in areas which are highly endemic for intestinal schistosomiasis. The findings also stress the importance of considering captive animals, whose health can be monitored more readily, as a model for understanding the dynamics of disease between humans and wildlife and molecular

epidemiology more generally. Given the ever-increasing importance of research into zoonotic diseases, our study presents the benefits of also examining animals for signs of emerging infection with human pathogens, with obvious conservation and welfare implications.

10.7 Conclusions

Chimpanzees on Ngamba Island were conclusively shown to be actively infected with *S. mansoni*, as well as to have had high levels of exposure to the parasite. The multiple diagnostics assisted in ascertaining the effects of previous treatment on infection status, for staff as well as chimpanzees. Given the high proportion of exposed *Biomphalaria*, Ngamba Island could be considered a high-risk environment for transmission of *Schistosoma*, and prevention measures accordingly integrated, for example reducing chimpanzee access to certain parts of the shoreline and encouraging tourists not to swim. A more detailed investigation of the prevalence of *S. rodhaini* on the island should also be carried out. The survey explicitly recognises the possibility of anthropozoonotic transmission of *S. mansoni* to non-human primates in other settings, and thus encourages conservation medicine practitioners to include schistosomiasis in their monitoring of wild primate populations. Finally, these findings highlight the need to investigate parasite prevalence at a variety of scales in order to gain a full appreciation of the dynamics of a disease in a particular location.

11 General Discussion

The research presented in this thesis sought to elucidate comprehensively the dynamics of intestinal schistosomiasis and its transmission in Lake Victoria using a multidisciplinary approach. Bringing together diagnostic testing, prevalence surveys and malacological observations, geographical/spatial modelling and molecular analyses, the overall conclusions can be presented within the framework of three, broad categories, which form the basis of projected future research: 'Schistosome prevalence and population dynamics', 'Intermediate host dynamics' and 'Zoonotic potential'. Each of these is explained in terms of the putative causal factors that shape the patterns reported on in the previous chapters.

11.1 Schistosome prevalence and population dynamics

Although Lake Victoria is known to be highly endemic for intestinal schistosomiasis, the surveys undertaken as part of this research have demonstrated unequivocally that the distribution of the disease is highly heterogeneous. As such, the scale of the survey can be crucial in determining the exact treatment needs or risk of exposure of a particular region, which is confounded further by diagnostic test accuracy. The examination of different diagnostic methods also emphasised the difficulty in identifying exact community treatment requirements. Finally, the population genetic studies of the parasite suggested that human demographical factors can decouple local prevalence within a community and the risk of exposure in that same location; the spread of different genotypes throughout Lake Victoria hint at more widespread compatibility with intermediate hosts than previously assumed.

Areas of high prevalence of intestinal schistosomiasis were found in all three countries surrounding Lake Victoria. However, as shown in Chapters 3 and 4 there were also pockets of much lower infection prevalence, and even absence of the disease. Interpretation of causation of patterns was difficult; some of the reduced prevalence, for example in Busia District in Uganda, is likely due to rigorous involvement in the national control programme, as the school surveyed there had actually been involved in the SCI roll-out of mass praziquantel

administration as early as the pilot stages (Kabatereine *et al.*, 2006b). However, other locations, such as some of the communities on the Ssesse Islands, did not have as comprehensive a history of treatment, and also had low prevalence of the disease. In the example of the Ssesse Islands, the lack of compatible snail populations may also be a factor; however, there may also be other environmental, behavioural and even genetic reasons why these patterns are seen.

Another factor brought to the fore in this thesis is the influence of different diagnostic tests on the prevalence and intensity estimates. While surveys have long used microscopy as the 'gold standard' for parasitological examinations, the paradigm of rapid mapping assessments has precipitated a move towards quicker, less cumbersome diagnostic techniques. Non-invasive methods, such as the CCA urine dipstick reported on in Chapter 4, proved a valuable addition to the diagnostic arsenal. However, questions regarding its reliability, especially in low endemicity settings, suggest that there is still a place for Kato-Katz slides in schistosomiasis surveys. If CCA diagnostic reliability drops with decreases in infection intensity or disease endemicity, the implication is that on-going treatment will reduce the CCA test's efficacy as a diagnostic in those settings. The Schistosomiasis Consortium for Operational Research and Evaluation (SCORE; http://score.uga.edu) has recently taken up this line of research, by funding a study to examine the changes in diagnostic score of the CCA urine test as treatment interventions are rolled-out in high prevalence areas, and thus emphasising the value of these initial findings. In addition to the SCORE research, a host of new funding has recently been secured, by donors such as the Wellcome Trust and the UK Department for International Development (DFID) in order to carry on researching the impact and efficacy of control of intestinal schistosomiasis in various settings in East Africa. Perhaps most promisingly, funding from SCORE has also been allocated to the NHM and other partners to attempt elimination of urinary schistosomiasis on Zanzibar; if successful, this will pave the way for elimination in other contexts, and will be a very positive step forward with regards to reducing the burden of schistosomiasis in Africa.

The serological SEA-ELISA tests were extremely powerful at detecting host antibodies (IgG/IgM) against *Schistosoma* antigens in the blood, but have three drawbacks; first of all, they are invasive, requiring fingerprick blood, which can be difficult to obtain for compliance and bio-safety reasons. Secondly, as host antibodies are likely to persist even once the infection has

been cured, the SEA-ELISA tests tend to overpredict prevalence, especially in the context of on-going treatment campaigns. Finally, the SEA-ELISA methodology is time-consuming and logistically challenging in the field, as centrifugation and cold storage are required. Given limited resources for treatment interventions, it is crucial for on-going monitoring campaigns to be able to identify accurately the treatment needs of a community as well as follow reductions in morbidity. One technique which has recently been field-trialled for morbidity at the community level is the faecal occult blood (FOB) test, which examines for blood in stool. While not appropriate as an individual diagnostic, due to other conditions commonly causing intestinal bleeding as well as schistosomiasis, it has been shown to provide a reasonable indication of community-level morbidity associated with schistosomiasis, thus potentially providing policy makers with better-evidence for monitoring putative reductions in community morbidity in areas without frequent treatment interventions (Betson *et al.*, in press).

A key factor that came to light during the surveys presented in Chapter 3 was the role of human migration in influencing patterns of schistosomiasis prevalence and transmission. Specifically, extremely high levels of migration and itinerancy were observed, both in the children being surveyed and the communities as a whole. Although the epidemiological importance of migratory fishing communities had been recognised (Odongo-Aginya and Mugisa, 1987), as far as the author is aware no parasitological studies have attempted to quantify human movement, as was done in Chapter 3; indeed some authors have dismissed such research as too difficult to achieve (Clements *et al.*, 2010). Moreover, it was shown that itinerancy was a statistical predictor of infection risk as well as also being an indicator of no previous treatment history with praziquantel. If children were moving schools and missing annual treatment, or indeed were coming from a region or district which does not have an organised control programme, then they were not only at risk of continued infection themselves, but could also seed transmission cycles in new regions. Given that point prevalence at a particular locality may be influenced by high numbers of migrant children, who actually acquired their infection elsewhere in the region or even in another country, a better understanding of these patterns and levels of itinerancy would be of a huge benefit to the understanding of schistosomiasis transmission in Lake Victoria (Standley and Stothard, 2010). Furthermore, the comparative

study on Kimi and Ngamba Islands showed how micro-scale movement of people between the islands may be assisting in maintaining a zoonotic reservoir of intestinal schistosomiasis in this area.

The question of where infection is acquired versus where it is measured can also be seen at a micro-scale. Given the heterogeneity of the shoreline and its habitats (as will be discussed in detail in the next section), then the exact locations where the water was accessed may determine whether exposure occurs or not. This was noted in Chapter 10, during the parasitological surveys of chimpanzees on Ngamba Island; certain locations showed higher levels of *Schistosoma* exposure, indicating these zones may be higher risk for transmission. One explanation for this could be variations in the frequency and strength of wave action onto the shoreline at various points; the qualitative measures of turbulence used in the statistical models presented in Chapter 9 were not statistically significant, but it might be that a more detailed, quantitative analysis would yield different results. An attempt at quantitatively measuring wave action on Ngamba and Kimi Islands is underway, using TinyTag (Gemini Data Loggers [UK] Ltd., Chichester, UK) motion sensor monitors which were suspended under the water at four sites on Ngamba (NG01, NG03, NG11 and NG18, the same sites used for snail population genetics and schistosome infection screening) and two on Kimi (KM20 and KM08). Preliminary results are promising; although marred by significant amounts of 'noise', the motion traces show clear differences between the sites deemed to have high levels of wave exposure versus those that were more sheltered. Further optimisation will certainly result in data that can be integrated into future statistical models for predicting infection risk at particular localities.

More broadly, a detailed understanding of the exact water contact behaviours is important. For example, with the chimps this could allow interventions that limit their exposure risk; the same could be applied to children and mothers in high endemicity areas. Indeed, recent research (SIMI project/Russell Stothard, unpublished data) has sought to use small, mobile GPS devices which can be attached on the wrists or arms of children and their mothers. These devices then monitor the movements of the pair on a minute-by-minute basis, allowing researchers to see exactly where and how often they are in contact with water. Continual monitoring of their

parasitological status over time may reveal at which point they were most likely to be infected, thus re-connecting point prevalence and transmission. Although potentially more difficult to achieve with the chimpanzees given their tendency to rip off such devices, gaining a better appreciation of their exact shoreline movements could directly provide recommendations for improving their welfare. Another method for achieving this could be the use of camera traps, set up at suspected water contact sites. Triggered by movement in the day and by infrared sensors at night, these cameras could provide a quantitative measure of frequency of water contact; the cameras are also able to take video footage in daylight, which would allow for estimates of water contact duration as well. The cameras might be additionally useful for capturing images of the rodents inhabiting Ngamba Island; as potential reservoirs for *S. mansoni* and other schistosomes, further research into the population size and species diversity of small mammals on the island is also warranted (as will be discussed further in a later section).

The genetics of *S. mansoni* in Lake Victoria reflected some of the survey findings. For example, despite the very high levels of genetic diversity, there was virtually no genetic structure inherent within the populations, be it at host, site or even country level. The lack of correlation between genetic and geographical distance has been observed in other parasite systems and is also consistent with the migratory nature of the human communities around the lake. It is worth emphasising another aspect of scale here; high parasite diversity was even found at the level of an individual child, suggesting that sampling efforts thus far have not even begun to capture the true levels of inter- and intrahost diversity, even at one location. In many parasitological settings, it has been shown that parasite genetic type can influence the clinical manifestation of the disease (Macedo and Pena, 1998, Brouwer *et al.*, 2003, Vega *et al.*, 2003, Morrison *et al.*, 2010); it is worth considering, therefore, what potentially synergistic effect the presence of multiple haplotypes/parasite diversity might have on the pathology of infection with *S. mansoni.*

The child pictured in Figure 11.1 was from Bwondha, in Uganda: out of 22 miracidia sequenced 10 different haplotypes were observed. The majority were common and widespread but four

were unique to that individual child, at least in the context of those individuals that were screened, indicating once more the challenges of sampling such high levels of diversity.

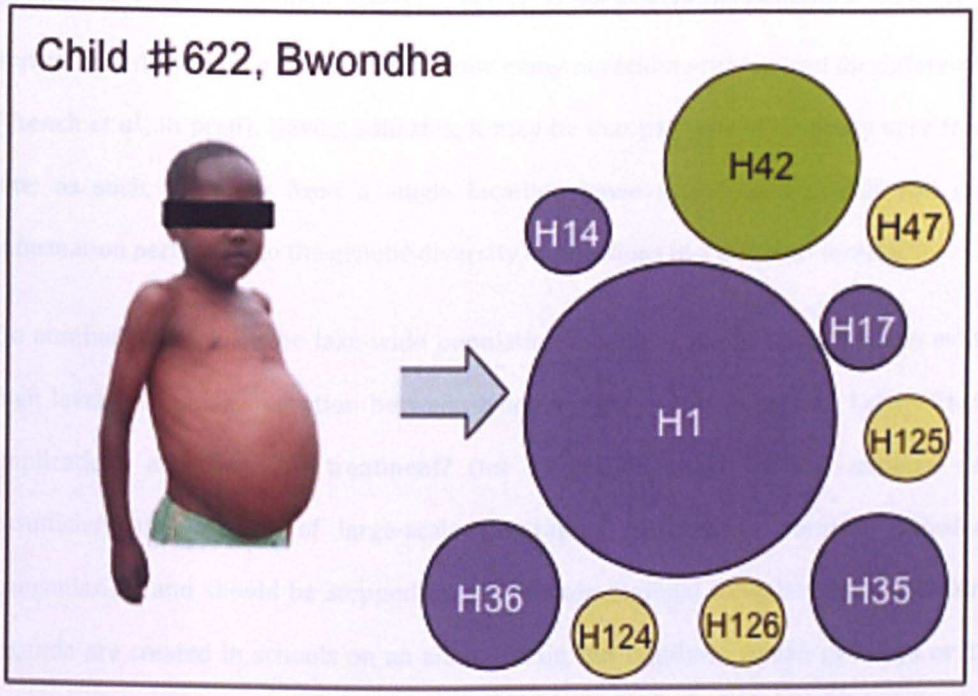

Figure 11.1 – Schematic of the high levels of parasite diversity found within an individual child
As in the figures in Chapter 5, the purple colour represents haplotypes found in all three countries; green represents haplotypes recovered from Uganda and Tanzania whereas the yellow circles represent haplotypes only observed in Uganda. The size of the circle is proportional to the number of sequences of that haplotype found in that particular child's stool (out of a total of 22 sequences).

Children were often observed to have several sequences of an abundant, widespread haplotype, but then would also have rarer, more locally restricted haplotypes, including some which were unique to that child and not observed elsewhere (in this case, haplotypes H124, H125 and H126). However, it is worth remembering that the diversity of the miracidia sampled may not directly reflect the genetic make-up of the adult worms residing within that particular patient.

One implication suggested by the extremely high levels of diversity observed both at the host and population level is how future research projects should design their sampling strategy to maximise collection of diversity, given inevitable restrictions in resources. For example, is it best to sample a few miracidia widely from the lake or better to focus on a single site or even individual and sequence many miracidia? Of course, the answer will depend somewhat on the aims of the research, but given the amount of lake-wide mixing reported in Chapter 5, yet the

ever-increasing number of unique haplotypes from individual children, it could be argued that more intensive sampling at single locations might actually be a more efficient means of cataloguing *S. mansoni* genetic diversity. In fact, at the level of the individual child, attempts are currently underway to quantify exactly how many miracidia are required for different analyses (French *et al.*, in prep). Having said this, it may be that patterns of diversity vary from site to site; as such, sampling from a single location, however intensively, still will not reveal information pertaining to the genetic diversity of infections in a different locality.

On another note, given the lake-wide population mixing of the parasite and the evidence for high levels of human migration between neighbouring countries around Lake Victoria, what implications are there for treatment? One suggestion could be that national control is insufficient in the face of large-scale geographic patterns in parasite prevalence and transmission, and should be stepped-up to integrate regional considerations. If children from Uganda are treated in schools on an annual basis, but regularly return to Kenya or Tanzania, where mass drug administration is more sporadic, they will continually replenish transmission of the parasite in Ugandan waters. Similarly, interventions such as improved sanitation and health education need to be rolled-out across the Lake Victoria region, to prevent local hot spots maintaining sources of infection that can be transported subsequently elsewhere in the lake region. It could also be that the most effective scale of intervention is even wider; the observed cross-over of genetic types between Lake Albert and Lake Victoria, which were previously considered genetically separate, demonstrates that perhaps multiple lake systems need to be considered in parallel for maximum impact of morbidity and prevalence control measures. Although the emphasis is currently on national maps of helminth prevalence and infection intensity, a resource such as the Global Atlas of Helminth Infections (GAHI; www.thiswormyworld.org, accessed August 2010) could be useful in bringing together data collected by health workers in different countries, in order to understand treatment needs and facilitate control at a regional or international level.

A key factor in terms of creating or negating transmission barriers between lakes may be compatibility between the parasite and its intermediate hosts. Despite the high genetic variety of *S. mansoni* in Lake Victoria, and the well-mixed nature of the populations, the parasite seems

compatible with whatever snails it comes into contact with and does not show the highly coupled compatibility observed in other schistosome-snail interactions, such as in past studies in Senegal (Southgate *et al.*, 2000). Although not analysed formally here, a more detailed look at some of the compatibility experiments carried out over the course of the Lake Victoria surveys might reveal some signs of differential transmission ability, although all snail populations tested were able to shed viable cercariae, based on infection with isolates of *S. mansoni* from all three countries. This is a timely reminder of the importance of snails in this system; elucidating the patterns of distribution and population structure of *Biomphalaria* is a crucial step in understanding its relationship and influence on *S. mansoni.*

11.2 Intermediate snail host dynamics

Biomphalaria intermediate host snails were, like the prevalence of the parasite, also found to be heterogeneously distributed throughout Lake Victoria. This focality was likewise observed at different scales of investigation; patterns which were seen at the Lake Victoria-wide level were mirrored, in terms of population genetic structure and influence of gastropod diversity, during the micro-scale surveys of Kimi and Ngamba Island, as described in Chapters 6 and 9. Moreover, Lake Victoria provided a unique environment for investigating the taxonomy and population genetics of these biomedically important snails, resulting in a finding which could have large implications for the transmission of schistosomiasis in the region.

One of the main initial observations relating to the distribution of *Biomphalaria* in Lake Victoria was the split between lake and marsh habitats. Not only were different morphotypes of snail found in each, but they appeared de-coupled, even at this same scale, in terms of abundance or presence of *Biomphalaria.* The statistical tests revealed that indeed, the presence of different morphotypes appeared to correlate with separate environmental factors, which hinted at the types of cues which may have led to the morphological differences. Overall, the outcome of the taxonomic investigation was to suggest a revision of the nomenclature of *Biomphalaria* from Lake Victoria. By referring to *B. choanomphala*-type and *B. sudanica*-type snails as *B. cho.* var. *choanomphala* and *B. cho.* var. *sudanica,* the genetic congruence between the ecophenotypes is observed, whilst the morphological differences are also recognised. It remains important to

acknowledge the two morphological variants of the species, as collectors in the field can use this nomenclature to test further some of the assumptions of combining the ecophenotypes into a single species, for example whether they are indeed equally compatible with *S. mansoni.* Similarly, this revision should not close the door on additional taxonomic work on *Biomphalaria* in the Great Lakes region; certainly, taxonomic confusion remains regarding the status of *B. stanleyi* and *B. pfeifferi* in Lake Albert. Moreover, the relationship between *B. sudanica* from Lake Albert and the newly-named *B. cho.* var. *sudanica* in Lake Victoria should be examined, as this might shed light on larger scale evolutionary patterns within the genus.

Another interesting finding was the relationship between the presence of other species of freshwater gastropod and the distribution of *B. choanomphala* var. *choanomphala* snails; this finding contradicted previously published suggestions that intracommunity competition with other gastropods could limit the success of intermediate host species at a given location (Woolhouse and Chandiwana, 1990). Moreover, the micro-scale survey showed high gastropod community diversity at sites that were heavily influenced by human activity, as seen in the water chemistry composition; if all the Lake Victoria molluscan fauna thrive under anthropogenically altered habitat conditions, then *B. choanomphala* too may abound, resulting in increased transmission of *S. mansoni.* A similar situation is seen in Lake Malawi, where overfishing of molluscivorous fishes led to an increase in the abundance of *Bulinus* snails (Stauffer *et al.*, 2006). Given the current interest in the effects of human disturbance on biodiversity and how this impacts on the spread of infectious diseases, further research into the synergies between other gastropods, *Biomphalaria* and incidence of *S. mansoni* should be pursued, in Lake Victoria as well as in other lacustrine settings.

Looking at transmission in this way, the genetic diversity of each child's *S. mansoni* burden is a 'snapshot' reflecting the cumulative history of the precise time and place in which s/he has been infected in the past. It will also reflect the effect of treatment interventions, and subsequent re-infection events. Figure 11.2 demonstrates how exposure to different snails, shedding different haplotypes of *S. mansoni* in different proportions, can result in a unique and highly diverse infection within a human host, even if just exposed to water at one site. Moreover, it is likely that over time the shedding profiles of the snails will change, resulting in

infection with a 'new' haplotype or haplotypes. Finally, given that people are likely to access the water at multiple points, across different times of the day, it becomes apparent that disentangling the exact exposure profile of any particular infected person is hugely complex, even on a very local scale. As such, the distribution of snails at a micro-scale level also has implications for the dynamics of transmission, and particularly when considering the high genetic diversity of *S. mansoni* infections observed across Lake Victoria.

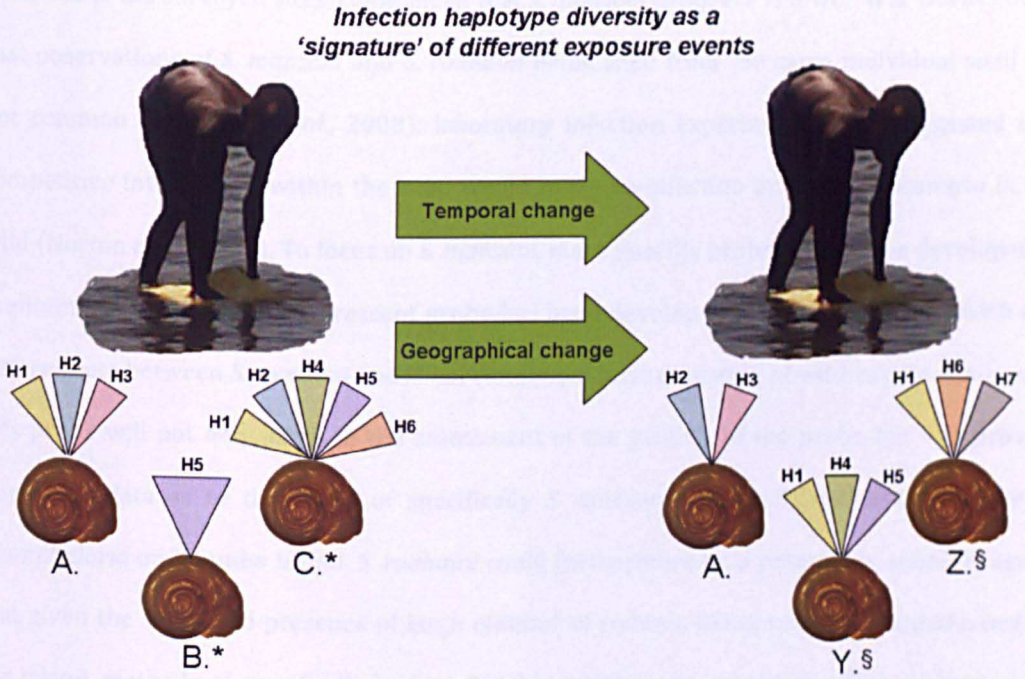

Figure 11.2 – Schematic showing the temporal and spatial variation in cercarial output which might influence infection genetic diversity

H1-H7 represent possible haplotypes of *S. mansoni* being shed by various *Biomphalaria*; the schematic demonstrates that while some snails may only shed one haplotype, in other cases they may be infected by more than one miracidium, resulting in emerging cercariae with different haplotypes. Similarly, shedding snails may die (as marked by '*' in the schematic) whereas at a later time point new snails may start shedding cercariae (indicated by §). As such, a person collecting water could be exposed to a number of different genetic types of *S. mansoni*; factoring in temporal and geographical change exposes how complicated the relationship between transmission and infection diversity really is, even at a micro-scale level.

As it was, shedding patterns were difficult to ascertain thoroughly due to the low prevalence of snails actively shedding *S. mansoni* cercariae. If more locations were found, a more detailed analysis of the environmental conditions which might facilitate transmission could be undertaken. Instead, molecular methods have been developed to try to determine presence of schistosome DNA in snails which are not actively shedding cercariae. In this case, when tested on snails from Ngamba Island, more than 50% were found to be positive, despite only 2 snails (approximately 2%) shedding cercariae across both surveys (in 2008 and 2010). However, this probe was designed to pick up all schistosomes, and given the later identification of *S. rodhaini* from one of the surveyed sites, could mean that *S. mansoni* exposure is lower. It is worth noting that observations of *S. mansoni* and S. *rodhaini* being shed from the same individual snail are not common (Steinauer *et al.*, 2008); laboratory infection experiments have suggested that competitive interactions within the snail would make co-infection an unlikely scenario in the wild (Norton *et al.*, 2008). To focus on *S. mansoni*, more specific probes need to be developed. A preliminary real-time PCR fluorescent probe has been developed (R. Kane, in prep) which can differentiate between *S. mansoni* and *S. haematobium*; further testing of wild-caught snails with this probe will not only assist in the assessment of the validity of the probe but will provide important data as to the levels of specifically *S. mansoni*, versus *S. rodhaini*, exposure in *Biomphalaria* on Ngamba Island. *S. rodhaini* could furthermore be a potentially zoonotic agent, and given the suspected presence of large number of rodents living relatively undisturbed on the island, methods of specifically looking for this schistosome should also be investigated for use in this context.

The existence of separate morphotypes of *Biomphalaria* in Lake Victoria was noted above, and was explored in detail in Chapter 7. Molecular evidence revealed high levels of genetic diversity, but none that was correlated with morphotype. Similarly, although morphological analyses demonstrated visually observable groupings, these groups overlapped at the edges and moreover were not statistically consistent across the various analyses. A key aspect of this analysis was to include the type specimens of the two 'species' thought to inhabit Lake Victoria; in some cases the type material were clear extremes of a particular morphotype, whereas in other instances, the type specimens clustered in the overlapping regions, indicating statistically

insignificant separation. Habitat was, as had been observed, a significant variable in accounting for the morphological differences, which lent weight to the hypothesis that the different shell morphologies were the result of inherent plasticity being acted on by environmental differences in the snails' surroundings. A similar situation seems to exist with *B. pfeifferi* and *B. stanleyi* in Lake Albert (Jørgensen *et al.*, 2007, Plam *et al.*, 2008), and further studies on that system might come to the same conclusion.

The differences between the population structure of *Biomphalaria* and *S. mansoni* can be seen in the comparison between the frequency and abundance of haplotypes (Figure 11.3); *S. mansoni* has many more geographically widely-spread haplotypes, which were moreover also very abundant. The strong population structuring of *Biomphalaria*, as statistically shown to be significant through the population genetics tests, is demonstrated by only a single haplpotype shared across the whole lake, with most restricted to a single country. The population genetics of the *Biomphalaria* in Lake Victoria were very different to those of the schistosomes; indeed, the lack of correlation between the snails and the schistosomes in terms of population structure was also statistically demonstrated, and shed light on the unique ecological and evolutionary conditions of the lake.

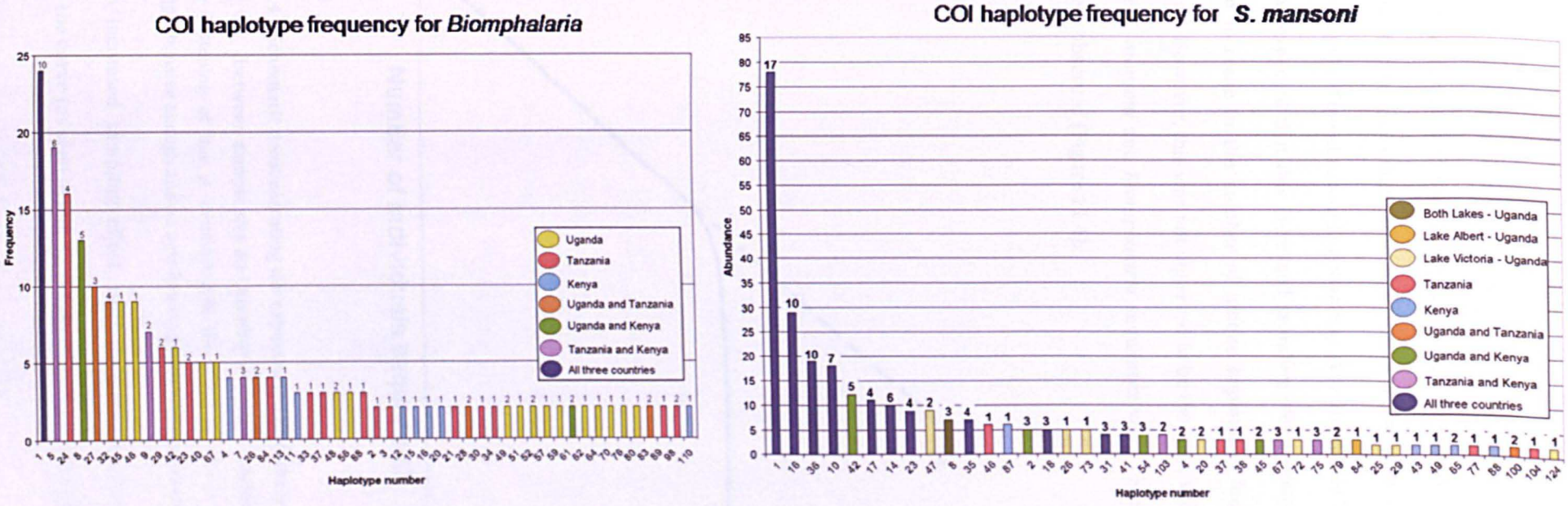

Figure 11.3 – Comparison of the abundance and frequency of COI haplotypes between *S. mansoni* and *Biomphalaria*

These figures were taken directly from Chapters 8 and 5 respectively, and show the very different pattern of geographical distribution of abundant haplotypes in *Biomphalaria* (left hand graph) and *S. mansoni* (right hand graph). The colour coding is consistent between the two figures as well as in the earlier chapters.

One similarity with the genetic structure of *S. mansoni*, however, was the high levels of diversity seen in all the markers used for characterising *Biomphalaria*. For both parasite and intermediate host, this somewhat confounded expectations; this research achieved more comprehensive levels of genetic sampling than had ever previously been carried out in Lake Victoria. It was expected that this increased sampling effort would eventually result in a much reduced increase in the number of unique haplotypes found, or indeed no new haplotypes at all. However, this was not observed to be the case; with every additional field mission, more *S. mansoni* and *Biomphalaria* sequences were retrieved and more unique haplotypes were observed (Figure 11.4).

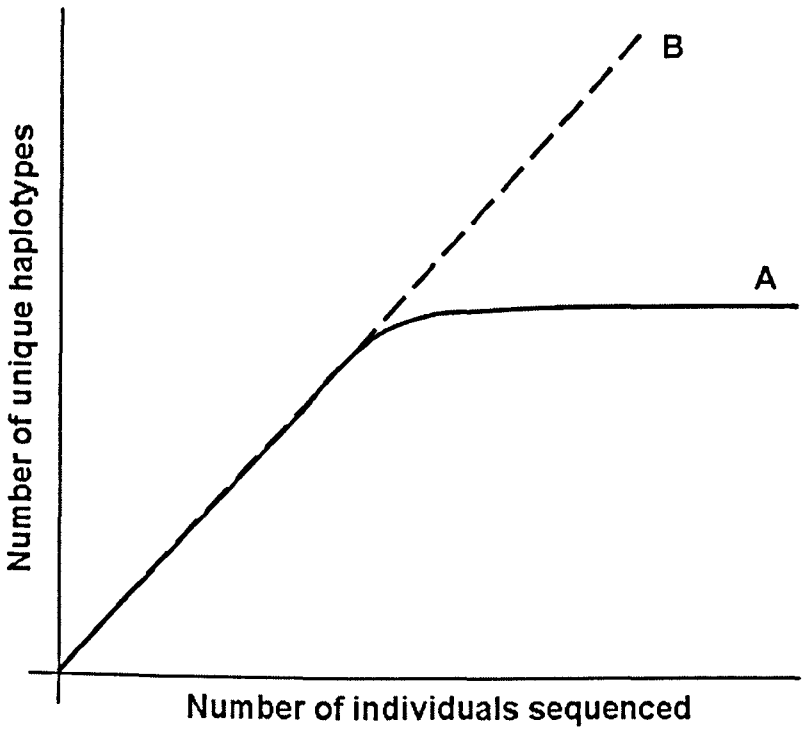

Figure 11.4 – Schematic demonstrating the expected (A) and observed (B) relationships between sample size and number of unique haplotypes

The flattening of line A demonstrates the point at which sampling has been comprehensive enough and no new haplotypes are being observed.

Usually, with increased sampling effort, total diversity is eventually observed, at the asymptote of the curve (as seen in line A). However, for both *Biomphalaria* and *S. mansoni* in

Lake Victoria, the observed pattern was more similar to B, whereby increased sampling always resulted in new, unique haplotypes.

Based on the Red Queen hypothesis, one theory would be that *S. mansoni* is so highly prevalent in Lake Victoria that snails are constantly exposed to new forms, continually driving higher levels of diversity as part of the on-going host-parasite arms race. However, it should be also noted that gastropods in general are often characterized by having high levels of population-level genetic diversity, and so the abundance of haplotypes observed for *Biomphalaria* here is actually not inconsistent with findings for other snail species, such as apple snails (Thaewnon-ngiw *et al.*, 2003) or *Cepaea nemoralis* (Davison, 2000). As such, it could be argued that the drivers of the high diversity seen in Lake Victoria *Biomphalaria* could be similar to those for other snails, rather than directly a result of the host-parasite relationship as observed with *S. mansoni.*

The implications of the high levels of intra- and interpopulation genetic diversity are also apparent when it comes to theories relating to transmission biology of *S. mansoni.* It would be expected that the genetic structure of the intermediate host would be a driver for localised compatibility, but as mentioned above, this does not seem to be the case in this system, as parasite types were mixed throughout the lake, and informal compatibility experiments showed susceptibility across a large geographical scale. It would be interesting to look more closely at the relationship between genotype and shedding patterns of different snail populations when exposed to different parasite isolates; however, while possible in the laboratory, it would be logistically challenging to attempt such a study in the field, where the environmental conditions would however be more realistic.

11.3 Zoonotic potential of schistosomiasis

The third, and arguably the most exciting, aspect of this research was the discovery of the zoonotic potential of intestinal schistosomiasis in Lake Victoria. Chimpanzees on Ngamba Island were unequivocally shown to be infected with *S. mansoni*, with individuals producing eggs that hatched into viable miracidia. Moreover, this was the first time that miracidia from

non-human primates were genotyped, proving their human origin. As such, the research presented in Chapter 10 represents the first observation of naturally-acquired intestinal schistosomiasis as a confirmed anthropozoontic disease.

This discovery is particularly timely given the recent renewed interest in investigating the schistosome burdens of non-human terminal hosts in East Africa. For example, in 2009, a new species of schistosome, named *S. kisumuensis*, was discovered during a detailed examination of rodents in Kisumu, Kenya (Hanelt *et al.*, 2009). Similarly, there is on-going study into the presence of schistosomiasis in baboons in the Coast province of Kenya (J. Cooper, pers. comm.) and also with baboons and chimpanzees in Gombe Stream National Park in Tanzania (J. Bakuza, pers. comm.), but these are isolated research projects and not indicative of a concerted effort to investigate the dynamics of transmission of schistosomiasis between humans and non-human hosts.

On a basic level, the observation of schistosomiasis in chimpanzees on Ngamba Island demonstrates they must come into contact with infected water, and indeed, informal reports of chimpanzees playing or wading through the island's shallows have been made. However, the Ngamba Island chimpanzees are human-habituated and many have been in captivity for most of their lives; whether these behaviours are as common among wild populations, or indeed among other primate species, is not known. Quantified behavioural observations of chimpanzee water contact on Ngamba Island could be a valuable avenue of future research; levels could then be compared to 'wild' populations which may have been more extensively studied from an ethological perspective, to see whether other populations may also be at risk of exposure. As mentioned earlier, camera traps or CCTV could be an alternative method for quantifiably measuring levels of exposure and water contact. These could be particularly useful for making observations of wild primates that are not habituated for field observations.

The Ngamba Island surveys also observed the presence of *S. rodhaini* in a snail from around the island's edge. Although rats are the usual terminal host for this schistosome species, it is known to hybridise with *S. mansoni* in the wild as well as in the laboratory (Brémond *et al.*, 1989, Morgan *et al.*, 2003), which might influence its ability to infect and cause clinical

symptoms in other animals. It has not yet been observed in the chimpanzees, perhaps owing to low levels of egg excretion, but molecular analysis will prove critical in further monitoring and identification of schistosome infections in these chimpanzees. Analysis of multiple genetic markers, including both a mitochondrial gene and a nuclear gene should be included to test for instances of hybridisation (Morgan *et al.*, 2003), as in the research presented above, only mitochondrial markers were used. On Ngamba Island, the population density and infection status of rats is currently unknown although they are pests; capture-mark-recapture surveys, together with parasitological examinations of a sub-set of individuals, could assist in determining the role that rodents may play in maintaining schistosome populations on Ngamba Island. Similarly, infrared camera traps, as mentioned above, could be useful for quantifiably estimating the rats' levels of water contact and exposure to the disease, as they are known to be strong swimmers but usually nocturnal, and therefore their behaviour is difficult to observe.

In a wider context, the importance of *S. rodhaini* as a potentially zoonotic infection has been generally considered low, due to experimental infections of baboons suggesting that primates are unlikely to serve as viable hosts unless co-infected with *S. mansoni* (Nelson and Teesdale, 1965). However, considering this research has not been revisited for 40 years and moreover bearing in mind the high prevalence of *S. mansoni* in the region, greater vigilance and research focus on the distribution of *S. rodhaini* in primates is warranted. In addition, *S. rodhaini* is known to be transmittable through canines; given that domestic dogs have been heavily implicated in spreading diseases such as rabies to wild carnivores in East Africa's national parks, it could be that they are also reservoirs for schistosomes, with the potential to cause health issues for endangered wild animals.

11.4 Future work

The research described above, while providing a comprehensive picture as to the status of the dynamics and distribution of *S. mansoni* and *Biomphalaria* in Lake Victoria, has also revealed a number of avenues for further investigation.

A particularly interesting direction for future research is the relationship between biodiversity and the transmission of infectious diseases, including schistosomiasis. Molecular tools are a powerful addition to more traditional ecological methods in the analysis of biodoiversity; similarly, by including spatial analysis to molecular data more meaningful conclusions can be drawn regarding the patterns of transmission of diseases. In the context of schistosomiasis, the influence of multiple terminal hosts adds an extra dimension to the influence of biodiversity on the dynamics of the system.

Specifically, it would be interesting to study the effect of anthropogenic influence on the landscape and the transmission of parasitic infections. An ideal location for this research would be at the boundaries of national parks; a comparison of study sites inside and outside the park could reveal the effects of human activity on the environmental conditions for intermediate hosts and vector species, such as snails or mosquito larvae. National parks also provide a wide diversity of potential terminal hosts; will this presence of other mammals provide a 'dilution effect', thus protecting human communities just outside the park from disease vectors, or will they act as reservoirs for other infections? On the other hand, will the presence of fringing human communities and game wardens inside the reserves put the, often endangered, animals residing in the parks at risk from anthropozoonotic infections? A particularly fruitful vein of research could be to study the habituated 'wild' animals which frequently pass through the park boundaries, and provide a link between human communities and the protected wildlife; such species include baboons, jackals and rodents. Above all, the emphasis throughout this Ph.D. has been on bringing together a range of methodologies to create a multidisciplinary research process, and it is this approach that would be required for a closer look at the relationship between biodiversity and infectious diseases, particularly in the context of parasites and molecular epidemiology.

12 References

ABDEL-RAHMAN, M. S., EL-BAHY, M. M., MALONE, J. B., THOMPSON, R. A. & EL BAHY, N. M. 2001. Geographic information systems as a tool for control program management for schistosomiasis in Egypt. *Acta Tropica,* 79, 49-57.

ABILA, R. O. & JANSEN, E. G. 1997. From local to global markets: the fish exporting and fishmeal industries of Lake Victoria: structure, strategies, and socio-economic impacts in Kenya. *Socio-economics of the Lake Victoria Fisheries Report.* UCN Eastern Africa Regional Office.

ADEMA, C. M. 2002. Comparative study of cytoplasmic actin DNA sequences from six species of Planorbidae (Gastropoda: Basommatophora). *Journal of Molluscan Studies,* 68, 17-23.

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. *Infection, Genetics and Evolution,* 6, 484-490.

AKOGUN, O. B. & AKOGUN, M. K. 1996. Human behaviour, water usage and schistosomiasis transmission in a small settlement near Yola, Nigeria. *Annals of Tropical Medicine and Parasitology,* 90, 303-311.

ALARCON DE NOYA, B., POINTIER, J. P., COLMENARES, C., THÉRON, A., BALZAN, C., CESARI, I. M., GONZALEZ, S. & NOYA, O. 1997. Natural *Schistosoma mansoni* infection in wild rats from Guadeloupe: parasitological and immunological aspects. *Acta Tropica,* 68, 11-21.

ALBRECHT, C., KUHN, K. & STREIT, B. 2007. A molecular phylogeny of Planorboidea (Gastropoda: Pulmonata): insights from enhanced taxon sampling. *Zoologica Scripta,* 36, 27-39.

ALLAN, F., ROLLINSON, D., SMITH, J. E. & DUNN, A. M. 2009. Host choice and penetration by *Schistosoma haematobium* miracidia. *Journal of Helminthology,* 83, 33-38.

ALVES, W. 1958. Chemical constituents of surface water in Southern Rhodesia, with special reference to the molluscan vectors of bilharziasis. *Bulletine of the World Health Organisation,* 18, 1071.

ANDERSON, R. M. & MAY, R. M. 1991. *Infectious Diseases of Humans,* Oxford, Oxford University Press.

ANDRADE, S. C. S. & SOLFERINI, V. N. 2006. Transfer experiment suggests environmental effects on the radula of *Littoraria flava* (Gastropoda: Littorinidae). *Journal of Molluscan Studies,* 72, 111-116.

ANDRADE, S. N. C. S. & SOLFERINI, V. N. 2009. The influence of size on the radula of *Littoraria angulifera* (Gastropoda: Littorinidae). *Malacologia,* 49, 1-5.

ANDRADE, Z. A. & VAN MARCK, E. 1984. Schistosomal glomerular disease. *Memorias do Instituto Oswaldo Cruz,* 79, 499-506.

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. Genetics Research,* 81, 193-204.

APPLETON, C. C. 1976. *The influence of abiotic factors on the distribution of B. pfeifferi (Krauss, 1848) (Planorbidae: Mollusca) and its life-cycle in South-Eastern Africa.* M.Sc., Rhodes University.

ARMITAGE, P., MATTHEWS, J. N. S. & BERRY, G. 2001. *Statistical Methods in Medical Research,* Oxford, Blackwell.

BAIRD, J. K. & WEAR, D. J. 1987. Cercarial dermatitis: the swimmer's itch. *Clinical Dermatology,* 5, 88-91.

BAKER, J. N. L. 1944. Sir Richard Burton and the Nile sources. *The English Historical Review,* 59, 48-61.

BAKUZA, J. & NKWENGULILA, G. 2009. Variation over time in parasite prevalence among free-ranging chimpanzees at Gombe National Park, Tanzania. *International Journal of Primatology,* 30, 43-53.

BALEN, J., STOTHARD, J. R., KABATEREINE, N. B., TUKAHEBWA, E. M., KAZIBWE, F., WHAWELL, S., WEBSTER, J. P., UTZINGER, J. & FENWICK, A. 2006. Morbidity due to *Schistosoma mansoni*: an epidemiological assessment of distended abdomen syndrome in Ugandan school children with observations before and 1-year after anthelminthic chemotherapy. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 100, 1039-1048.

BANDELT, H. J., FORSTER, P. & ROHL, A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution,* 16, 37-48.

BANDONI, S. M., MULVEY, M., KOECH, D. K. & LOKER, E. S. 1990. Genetic structure of Kenyan populations of *Biomphalaria pfeifferi* (Gastropoda, Planorbidae). *Journal of Molluscan Studies,* 56, 383-391.

BANDONI, S. M., MULVEY, M. & LOKER, E. S. 2000. Population structure and taxonomic discrimination among three species of *Biomphalaria* Preston, 1910 (Gastropoda : Planorbidae) from Kenya. *Zoological Journal of the Linnean Society,* 129, 387-401.

BARKER, S. C. & BLAIR, D. 1996. Molecular phylogeny of *Schistosoma* species supports traditional groupings within the genus. *The Journal of Parasitology,* 82, 292-298.

BASCH, P. F. 1976. Intermediate host specificity in *Schistosoma mansoni. Experimental Parasitology,* 39, 150-169.

BENSON, D. A., KARSCH-MIZRACHI, I., LIPMAN, D. J., OSTELL, J. & SAYERS, E. W. 2009. GenBank. *Nucleic Acids Research,* 37, D26-31.

BERGQUIST, R., JOHANSEN, M. V. & UTZINGER, J. 2009. Diagnostic dilemmas in helminthology: what tools to use and when? *Trends in Parasitology,* 25, 151-156.

BERRIMAN, M., HAAS, B. J., LOVERDE, P. T., WILSON, R. A., DILLON, G. P., CERQUEIRA, G. C., MASHIYAMA, S. T., AL-LAZIKANI, B., ANDRADE, L. F., ASHTON, P. D., ASLETT, M. A., BARTHOLOMEU, D. C., BLANDIN, G., CAFFREY, C. R., COGHLAN, A., COULSON, R., DAY, T. A., DELCHER, A., DEMARCO, R., DJIKENG, A., EYRE, T., GAMBLE, J. A., GHEDIN, E., GU, Y., HERTZ-FOWLER, C., HIRAI, H., HIRAI, Y., HOUSTON, R., IVENS, A., JOHNSTON, D. A., LACERDA, D., MACEDO, C. D., MCVEIGH, P., NING, Z., OLIVEIRA, G., OVERINGTON, J. P., PARKHILL, J., PERTEA, M., PIERCE, R. J., PROTASIO, A. V., QUAIL, M. A., RAJANDREAM, M. A., ROGERS, J., SAJID, M., SALZBERG, S. L., STANKE, M., TIVEY, A. R., WHITE, O., WILLIAMS, D. L., WORTMAN, J., WU, W., ZAMANIAN, M., ZERLOTINI, A., FRASER-LIGGETT, C. M., BARRELL, B. G. & EL-SAYED, N. M. 2009. The genome of the blood-fluke *Schistosoma mansoni. Nature,* 460, 352-358.

BETSON, M., SOUSA-FIGUEIREDO, J. C., ROWELL, C., KABATEREINE, N. & STOTHARD, J. R. In press. Intestinal schistosomiasis in mothers and youn children in Uganda: Investigating field-applicable markers of bowel morbidity. *American Journal of Tropical Medicine and Hygiene*

BEUNING, K. R. M., KELTS, K., ITO, E. & JOHNSON, T. C. 1997. Paleohydrology of Lake Victoria, East Africa, inferred from $^{18}O/^{16}O$ rations in sediment cellulose. *Geology*, 25, 1083-1086.

BLAIR, L., WEBSTER, J. P. & BARKER, G. C. 2001. Isolation and characterization of polymorphic microsatellite markers in *Schistosoma mansoni* from Africa. *Molecular Ecology*, 1, 93-95.

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.

BONNAUD, L., BOUCHER-RODONI, R. & MONNEROT, M. 1994. Phylogeny of decapod cephalopods based on partial 16S rDNA nucleotide sequences. *Comptes rendus de l'Académie des sciences. Série III*, 317, 581-588.

BOOTH, M., VOUNATSOU, P., N'GORAN, E. K., TANNER, M. & UTZINGER, J. 2003. The influence of sampling effort and the performance of the Kato-Katz technique in diagnosing *Schistosoma mansoni* and hookworm co-infections in rural Cote d'Ivoire. *Parasitology*, 127, 525-531.

BOROS, D. L. 1989. Immunopathology of *Schistosoma mansoni* infection. *Clinical Microbiology Reviews*, 2, 250-269.

BOUCHET, P. & ROCROI, J.-P. 2005. Classification and nomenclator of gastropod families. *Malacologia: International Journal of Malacology*, 47, 397pp.

BOWLES, J., BLAIR, D. & MCMANUS, D. P. 1992. Genetic variants within the genus *Echinococcus* identified by mitochondrial DNA sequencing. *Molecular and Biochemical Parasitology*, 54, 165-173.

BOWLES, J., HOPE, M., TIU, W. U., LIU, X. & MCMANUS, D. P. 1993. Nuclear and mitochondrial markers highly conserved between Chinese and Philippine *Schistosoma japonicum*. *Acta Tropica*, 55, 217-229.

BOWLES, J., BLAIR, D. & MCMANUS, D. P. 1995. A molecular phylogeny of the human schistosomes. *Molecular Phylogenetics and Evolution*, 4, 103-109.

BRADLEY, D. J. 1968. The epidemiology of fishermen as migrants: Observations on Small Island. *East African Medical Journal*, 45, 254-262.

BRÉMOND, P., THÉRON, A. & ROLLINSON, D. 1989. Hybrids between *Schistosoma mansoni* and *S. rodhaini*: characterization by isoelectric focusing of six enzymes. *Parasitology Research*, 76, 138-145.

BROOKER, S. & MICHAEL, E. 2000. The potential of geographical information systems and remote sensing in the epidemiology and control of human helminth infections, *Advances in Parasitology*, 47, 245-288.

BROOKER, S., HAY, S. I., ISSAE, W., HALL, A., KIHAMIA, C. M., LWAMBO, N. J. S., WINT, W., ROGERS, D. J. & BUNDY, D. A. P. 2001. Predicting the distribution of urinary schistosomisis in Tanzania using satellite sensor data. *Tropical Medicine & International Health*, 6, 998-1007.

BROOKER, S., KABATEREINE, N. B., MYATT, M., STOTHARD, J. R. & FENWICK, A. 2005. Rapid assessment of *Schistosoma mansoni*: the validity, applicability and cost-effectiveness of the Lot Quality Assurance Sampling method in Uganda. *Tropical Medicine & International Health*, 10, 647-658.

BROOKER, S., ALEXANDER, N., GEIGER, S., MOYEED, R. A., STANDER, J., FLEMING, F., HOTEZ, P. J., CORREA-OLIVEIRA, R. & BETHONY, J. 2006. Contrasting patterns in the small-scale

heterogeneity of human helminth infections in urban and rural environments in Brazil. *International Journal for Parasitology,* 36, 1143-1151.

BROOKER, S. 2007. Spatial epidemiology of human schistosomiasis in Africa: risk models, transmission dynamics and control. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 101, 1-8.

BROOKER, S. & CLEMENTS, A. C. A. 2009. Spatial heterogeneity of parasite co-infection: Determinants and geostatistical prediction at regional scales. *International Journal for Parasitology,* 39, 591-597.

BROOKER, S., KABATEREINE, N. B., SMITH, J. L., MUPFASONI, D., MWANJE, M. T., NDAYISHIMIYE, O., LWAMBO, N. J., MBOTHA, D., KARANJA, P., MWANDAWIRO, C., MUCHIRI, E., CLEMENTS, A. C., BUNDY, D. A. & SNOW, R. W. 2009. An updated atlas of human helminth infections: the example of East Africa. *International Journal of Health Geographics,* 8 (42), published online 9 July, 2009.

BROUWER, K. C., NDHLOVU, P. D., WAGATSUMA, Y., MUNATSI, A. & SHIFF, C. J. 2003. Urinary tract pathology attributed to *Schistosoma haematobium*: Does parasite genetics play a role? *American Journal of Tropical Medicine and Hygiene,* 68, 456-462.

BROWN, D. S. 1994. *Freshwater snails of Africa and their Medical Importance,* London, Taylor & Francis.

BRYANT, J. 1972. *Health and the Developing World,* New York, Cornell University Press.

CALDAS, R., CORREA-OLIVIERA, R., COLOSINO, E., CARVALHO, O. S., MASSARA, C. L., COLLEY, D. G. & GAZZINELLI, G. 2000. Susceptibility and resistance to *Schistosoma mansoni* re-infection: parallel cellular and isotypic immunologic assessment. *American Journal of Tropical Medicine and Hygiene,* 62, 57-64.

CALLISTO, M., MORENO, P., GONÁALVES, J. F., JR., FERREIRA, W. R. & GOMES, C. L. Z. 2005. Malacological assessment and natural infestation of *Biomphalaria straminea* (Dunker, 1848) by *Schistosoma mansoni* (Sambon, 1907) and *Chaetogaster limnaei* (K. von Baer, 1827) in an urban eutrophic watershed. *Brazilian Journal of Biology,* 65, 217-228.

CAMPBELL, G., JONES, C. S., LOCKYER, A. E., HUGHES, S., BROWN, D., NOBLE, L. R. & ROLLINSON, D. 2000. Molecular evidence supports an African affinity of the Neotropical freshwater gastropod, *Biomphalaria glabrata,* Say 1818, an intermediate host for *Schistosoma mansoni. Proceedings of the Royal Society of London Series B-Biological Sciences,* 267, 2351-2358.

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. *Molecular Ecology,* 19, 241-256.

CARABIN, H., GUYATT, H. & ENGELS, D. 2000. A comparative analysis of the cost-effectiveness of treatment based on parasitological and symptomatic screening for *Schistosoma mansoni* in Burundi. *Tropical Medicine & International Health,* 5, 192-202.

CARLIER, Y., NZEYIMANA, H., BOUT, D. & CAPRON, A. 1980. Evaluation of circulating antigens by a sandwich radioimmunoassay, and of antibodies and immune complexes, in *Schistosoma mansoni*-infected African parturients and their newborn children. *American Journal of Tropical Medicine and Hygiene,* 29, 74-81.

CARLSON, J. C., DYER, L. A., OMIRM F. X. & BEIER, J. C. 2009. Diversity cascades and malaria vectors. *Journal of Medical Entomology,* 46, 460-464.

CARVALHO, O. S., SCHOLTE, R. G., GUIMARAES, R. J., FREITAS, C. C., DRUMMOND, S. C., AMARAL, R. S., DUTRA, L. V., OLIVEIRA, G., MASSARA, C. L. & ENK, M. J. 2010. The Estrada

Real project and endemic diseases: the case of schistosomiasis, geoprocessing and tourism. *Memorias do Instituto Oswaldo Cruz,* 105, 532-536.

CHAPLIN, J. H. 1974. The prehistoric rock art of the Lake Victoria region. *Azania,* 9, 1-50.

CHARBONNEL, N., ANGERS, B., RAZATAVONJIZAY, R., BREMOND, P. & JARNE, P. 2000. Microsatellite variation in the freshwater snail *Biomphalaria pfeifferi. Molecular Ecology,* 9, 1006-1007.

CHARBONNEL, N., ANGERS, B., RASATAVONJIZAY, R., BREMOND, P., DEBAIN, C. & JARNE, P. 2002a. The influence of mating system, demography, parasites and colonization on the population structure of *Biomphalaria pfeifferi* in Madagascar. *Molecular Ecology,* 11, 2213-2228.

CHARBONNEL, N., ANGERS, B., RASATAVONJIZAY, R., BREMOND, P. & JARNE, P. 2002b. Evolutionary aspects of the metapopulation dynamics of *Biomphalaria pfeifferi,* the intermediate host of *Schistosoma mansoni. Journal of Evolutionary Biology,* 15, 248-261.

CHARBONNEL, N., RASATAVONJIZAY, R., SELLIN, E., BREMOND, P. & JARNE, P. 2005. The influence of genetic factors and population dynamics on the mating system of the hermaphroditic freshwater snail *Biomphalaria pfeifferi. Oikos,* 108, 283-296.

CHEEVER, A. W., MACEDONIA, J. G., MOSIMANN, J. E. & CHEEVER, E. A. 1994. Kinetics of egg production and egg excretion by *Schistosoma mansoni* and *S. japonicum* in mice infected with a single pair of worms. *American Journal of Tropical Medicine and Hygiene,* 50, 281-295.

CHITSULO, L., ENGELS, D., MONTRESOR, A. & SAVIOLI, L. 2000. The global status of schistosomiasis and its control. *Acta Tropica,* 77, 41-51.

CIOCCO, N., STOR, F. & SCHEIBLER, E. E. 2008. Malacofauna of the littoral benthos of a saline lake in southern Mendoza, Argentina. *Fundamental and Applied Limnology / Archiv für Hydrobiologie,* 172, 87-98.

CLEMENTS, A. C. A., MOYEED, R. & BROOKER, S. 2006. Bayesian geostatistical prediction of the intensity of infection with *Schistosoma mansoni* in East Africa. *Parasitology,* 133, 711-719.

CLEMENTS, A. C. A., BROOKER, S., NYANDINDI, U., FENWICK, A. & BLAIR, L. 2008. Bayesian spatial analysis of a national urinary schistosomiasis questionnaire to assist geographic targeting of schistosomiasis control in Tanzania, East Africa. *International Journal for Parasitology,* 38, 401-415.

CLEMENTS, A. C. A., DEVILLE, M. A., NDAYISHIMIYE, O., BROOKER, S. & FENWICK, A. 2010. Spatial co-distribution of neglected tropical diseases in the East African Great Lakes region: revisiting the justification for integrated control. *Tropical Medicine & International Health,* 15, 198-207.

CLENNON, J. A., MUNGAI, P. L., MUCHIRI, E. M., KING, C. H. & KITRON, U. 2006. Spatial and temporal variations in local transmission of *Schistosoma haematobium* in Msambweni, Kenya. *American Journal of Tropical Medicine and Hygiene,* 75, 1034-1041.

COMBES, C. & THÉRON, A. 2000. Metazoan parasites and resource heterogeneity: constraints and benefits. *International Journal for Parasitology,* 30, 299-304.

COLLIER, K. J. 2004. Invertebrate community dynamics in soft-bottomed streams of northern New Zealand: a spatio-temporal hierarchy. *New Zealand Journal of Marine and Freshwater Research,* 38, 1-18.

CORNUET, J. M. & LUIKART, G. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics,* 144, 2001-14.

CORT, W. W. 1950. Studies on schistosome dermatitis: XI. Status of knowledge after more than twenty years. *American Journal of Epidemiology,* 52, 251-307.

COUTINHO, E. M., BARROS, A. F., BARBOSA JR, A., OLIVEIRA, S. A., SILVA, L. M., ARAUJO, R. E. & ANDRADE, Z. A. 2003. Host nutritional status as a contributory factor to the remodeling of schistosomal hepatic fibrosis. *Memorias do Instituto Oswaldo Cruz,* 98, 919-925.

CRISCIONE, C. D. & BLOUIN, M. S. 2004. Life cycles shape parasite evoluton: Comparative population genetics of salmon trematodes. *Evolution,* 58, 198-202.

CURWEN, R. S. & WILSON, R. A. 2003. Invasion of skin by schistosome cercariae: some neglected facts. *Trends in Parasitology,* 19, 63-66.

CURTIS, J., SORENSEN, R. E. & MINCHELLA, D. J. 2002. Schistosome genetic diversity: the implications of population structure as detected with microsatellite markers. *Parasitology,* 125, 551-559.

DALE, D. & ASHLEY, C. Z. 2010. Holocene hunter-fisher-gatherer communities: new perspectives on Kansyore Using communities of Western Kenya. *Azania,* 45, 24-48.

DALTON, P. R. & POLE, D. 1978. Water-contact patterns in relation to *Schistosoma haematobium* infection. *Bulletin of the World Health Organization,* 56, 417-26.

DARWALL, W., SMITH, K., LOWE, T. & VIE, J.-C. 2005. The Status and Distribution of Freshwater Biodiversity in Eastern Africa. *In:* COMMISSION, T. I. S. S. (ed.) *Occasional Paper of the IUCN Species Survival Commission.* IUCN – The World Conservation Union.

DAVISON, A. 2000. An East–West distribution of divergent mitochondrial haplotypes in British populations of the land snail, *Cepaea nemoralis* (Pulmonata). *Biological Journal of the Linnean Society,* 70 (4), 697-706.

DE CLERCQ, D., SACKO, M., BEHNKE, J., GILBERT, F. & VERCRUYSSE, J. 1998. The relationship between *Schistosoma haematobium* infection and school performance and attendance in Bamako, Mali. *Annals of Tropical Medicine and Parasitology,* 92, 851-858.

DEJONG, R. J., MORGAN, J. A. T., PARAENSE, W. L., POINTIER, J. P., AMARISTA, M., AYEH-KUMI, P. F. K., BABIKER, A., BARBOSA, C. S., BREMOND, P., CANESE, A. P., DE SOUZA, C. P., DOMINGUEZ, C., FILE, S., GUTIERREZ, A., INCANI, R. N., KAWANO, T., KAZIBWE, F., KPIKPI, J., LWAMBO, N. J. S., MIMPFOUNDI, R., NJIOKOU, F., PODA, J. N., SENE, M., VELASQUEZ, L. E., YONG, M., ADEMA, C. M., HOFKIN, B. V., MKOJI, G. M., LOKER, E. S. 2001. Evolutionary relationships and phylogeography of *Biomphalaria* (Gastropoda: Planorbidae) with implications regarding its role as host of the human bloodfluke, *Schistosoma mansoni. Molecular Biology and Evolution,* 18, 2225-2239.

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. *Molecular Ecology,* 12, 3041-3056.

DE KOCK, K. N., WOLMARANS, C. T. & BORNMAN, M. 2004. Distribution and habitats of *Biomphalaria pfeifferi,* snail intermediate host of *Schistosoma mansoni,* in South Africa. *Water South Africa,* 30, 29-36.

DESPRES, L., IMBERT-ESTABLET, D., COMBES, C. & BONHOMME, F. 1992. Molecular evidence linking hominid evolution to recent radiation of schistosomes (Platyhelminthes: Trematoda). *Molecular Phylogenetics and Evolution,* 1, 295-304.

DESPRES, L., IMBERT-ESTABLET, D. & MONNEROT, M. 1993. Molecular characterization of mitochondrial DNA provides evidence for the recent introduction of *Schistosoma mansoni* into America. *Molecular and Biochemical Parasitology,* 60, 221-229.

DEWITT, W. B. 1954. Susceptibility of snail vectors to geographic strains of *Schistosoma japonicum. The Journal of Parasitology,* 40, 453-456

DIAS NETO, E., DE SOUZA, C. P., ROLLINSON, D., KATZ, N., PENA, S. D. & SIMPSON, A. J. 1993. The random amplification of polymorphic DNA allows the identification of strains and species of schistosome. *Molecular and Biochemical Parasitology,* 57, 83-8.

DICKSON, R., AWASTHI, S., WILLIAMSON, P., DEMELLMEEK, C. & GARNER, P. 2000. Effects of treatment for intestinal helminth infection on growth and cognitive performance in children: systematic review of randomised trials. *British Medical Journal,* 320, 1697-1701.

DOBSON, A, CATTADORI, I., HOLT, R. D., OSTFELD, R. S., KEESING, F., KRICHBAUM, K, ROHR, J. R., PERKINS, S.E. & HUDSON, P. J. 2006. Sacred cows and sympathetic squirrels: The importance of biological diversity to human health. *PLoS Medicine,* 3, e231.

DOENHOFF, M., J., CIOLI, D. & UTZINGER, J. 2008. Praziquantel: Mechanisms of action, resistance and new derivatives for schistosomiasis. *Current Opinion in Infectious Diseases,* 21, 659-667.

DUMONT, M, MONE, H., MOUAHID, G., IDRIS, M. A., SHABAN, M. & BOISSIER, J. 2007. Influence of pattern of exposure, parasite genetic diversity and sex of degree of protection against reinfection with *Schistosoma mansoni. Parasitology Research,* 101, 247-252.

DUPOUY, J., ROUSSEAU, D., DUSSART, G., LIAUD, M. V. & NASSI, H. 1993. Correspondence analysis of shell morphology in the African fresh-water snail *Biomphalaria pfeifferi* (Kraus, 1848) (Pulmonata, Gastropoda). *Biological Journal of the Linnean Society of London,* 50, 329-338.

DURAND, P., SIRE, C. & THÉRON, A. Isolation of microsatellite markers in the digenean trematode *Schistosoma mansoni* from Guadeloupe island. *Molecular Ecology,* 9, 997-998.

DYBDAHL, M. F. & LIVELY, C. M. 1996. The geography of coevolution: comparative population structures for a snail and its trematode parasite. *Evolution,* 50, 2264-2275.

EBERL, M., AL-SHERBINY, M., HAGAN, P., LJUBOJEVIC, S., THOMAS, A. W. & WILSON, R. A. 2002a. A novel and sensitive method to monitor helminth infections by faecal sampling. *Acta Tropica,* 83, 183-7.

EHLERS, U. 1985. Phylogenetic relationships within the Platyhelminthes. *In:* CONWAY MORRIS, S., GEORGE, J. D., GIBSON, R. & PLATT, H. M. (eds.) *The Origins and Relationships of Lower Invertebrates.* Oxford: Clarendon Press.

EL-ANSARY, A. & AL-DAIHAN, S. Important aspects of *Biomphalaria* snail-schistosome interactions as targets for antischistosome drug. *Medical Science Monitor,* 12, 282-292.

EMMANUEL, I. O. A. & DOERING, E. 2008. Epidemiology, of bilharzias (schistosomiasis) in Uganda from 1902 until 2005. *African Health Sciences,* 8, 239-243.

EPSTEIN, J. H. & PRICE, J. T. 2009. The significant but understudied impact of pathogen transmission from humans to animals. *Mount Sinai Journal of Medicine: A Journal of Translational and Personalized Medicine,* 76, 448-455.

ERKO, B., BALCHA, F. & KIFLE, D. 2006. The ecology of *Biomphalaria sudanica* in Lake Ziway, Ethiopia. *African Journal of Ecology,* 44, 347-352.

ESCH, G. W. & FERNANDEZ, J. C. 1994. Snail-trematode interactions and parasite community dynamics in aquatic systems: A review. *The American Midland Naturalist,* 131, 209-237.

ESTRADA, V. E., VELASQUEZ, L. E., CALDEIRA, R. L., BEJARANO, E. E., ROJAS, W. & CARVALHO, O. S. 2006. Phylogenetics of South American *Biomphalaria* and description of a new species (Gastropoda: Planorbidae). *Journal of Molluscan Studies*, 72, 221-228.

EXCOFFIER, L., SMOUSE, P. E. & QUATTRO, J. M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, 131, 479–491.

EXCOFFIER, L., LAVAL, G. & SCHNEIDER, S. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, 1, 47-50.

FENG, Z., CURTIS, J. & MINCHELLA, D. J. 2001. The influence of drug treatment on the maintenance of schistosome genetic diversity. *Journal of Mathematical Biology*, 43, 52-68.

FENWICK, A. 1969. Baboons as reservoir hosts of *Schistosoma mansoni*. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 63, 557-567.

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., 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 (13), 1719-1730.

FILES, V. S. & CRAM, E. B. 1949. A study on the comparative susceptibility of snail vectors to strains of *Schistosoma mansoni*. *The Journal of Parasitology*, 35, 555-560.

FLEMING, F. M., FENWICK, A., TUKAHEBWA, E. M., LUBANGA, R. G. N., NAMAWANGYE, H., ZARAMBA, S. & KABATEREINE, N. B. 2009. Process evaluation of schistosomiasis control in Uganda, 2003 to 2006: perceptions, attitudes and constraints of a national programme. *Parasitology*, 136, 1759-1769.

FLETCHER, M., LOVERDE, P. T. & WOODRUFF, D. S. 1981. Genetic variation in *Schistosoma mansoni*: Enzyme polymorphisms in populations from Africa, Southwest Asia, South America, and the West Indies. *American Journal of Tropical Medicine and Hygiene*, 30, 406-421.

FOLMER, O., BLACK, M., HOEH, W., LUTZ, R. & VRIJENHOEK, R. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, 3, 294-299.

FRENCH, M. D., ROLLINSON, D., BASÁÑEZ, M.-G., MGENI, A. F., KHAMIS, I. S. & STOTHARD, J. R. 2007. School-based control of urinary schistosomiasis on Zanzibar, Tanzania: Monitoring micro-haematuria with reagent strips as a rapid urological assessment. *Journal of Pediatric Urology*, 3, 364-368.

FRÉZAL, L. & LEBLOIS, R. 2008. Four years of DNA barcoding: Current advances and prospects. *Infection, Genetics and Evolution*, 8, 727-736.

FU, Y. X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, 147, 915-925.

FULFORD, A. J. C., WEBSTER, M., OUMA, J. H., KIMANI, G. & DUNNE, D. W. 1998. Puberty and age-related changes in susceptibility to schistosome infection. *Parasitology Today*, 14, 23-26.

GANDON, S., CAPOWIEZ, Y., DUBOIS, Y., MICHALAKIS, Y. & OLIVIER, I. 1996. Local adaptation and gene-for-gene coevolution in a metapopulation model. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 263, 1003-1009.

GENTILE, R. & OLIVEIRA, G. 2008. Brazilian studies on the genetics of *Schistosoma mansoni*. *Acta Tropica*, 108, 175-178.

GLINZ, D., SILUE, K. D., KNOPP, S., LOHOURIGNON, L. K., YAO, K. P., STEINMANN, P., RINALDI, L., CRIGNOLI, G., N'GORAN, E. K. & UTZINGER, J. 2010. Comparing diagnostic accuracy of Kato-Katz, Koga agar plate, ether-concentration and FLOTAC for *Schistosoma mansoni* and soil-transmitted helminths. *PLoS Neglected Tropical Diseases*, 4, 3754.

GOLDSCHMIDT, T., WITTE, F. & WANINK, J. 1993. Cascading effects of the introduced Nile Perch on the detritivorous/phytoplanktivorous species in the sublittoral areas of Lake Victoria. *Conservation Biology*, 7, 686-700.

GOLL, P. H. 1982. Seasonal changes in the distribution of *Biomphalaria sudanica* (Martens) in Lake Ziway, Ethiopia. *Annals of Tropical Medicine and Parasitology*, 76, 159-164.

GOUDET, J. 1995. Fstat version 1.2: a computer program to calculate Fstatistics. *Journal of Heredity*, 86, 485-486.

GOW, J. L., NOBLE, L. R., ROLLINSON, D., MIMPFOUNDI, R. & JONES, C. S. 2004. Breeding system and demography shape population genetic structure across ecological and climatic zones in the African freshwater snail, *Bulinus forskalii* (Gastropoda, Pulmonata), intermediate host for schistosomes. *Molecular Ecology*, 13, 3561-3573.

GOW, J. L., NOBLE, L. R., ROLLINSON, D., TCHUENTE, L. A. T. & JONES, C. S. 2007. Contrasting temporal dynamics and spatial patterns of population genetic structure correlate with differences in demography and habitat between two closely-related African freshwater snails. *Biological Journal of the Linnean Society*, 90, 747-760.

GOWER, C. M., SHRIVASTAVA, J., LAMERTON, P. H. L., 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.

GRANDE, C., TEMPLADO, J. & ZARDOYA, R. 2008. Evolution of gastropod mitochondrial genome arrangements. *Bmc Evolutionary Biology*, 8 (61), published online 26 February, 2008.

GUIMARAES, R. J. D. P. S., FREITAS, C. C., DUTRA, L. V., SCHOLTE, R. G. C., MARTINS-BEDE, F. T., FONSECA, F. R., AMARAL, R. S., DRUMMOND, S. C., FELGUEIRAS, C. A., OLIVEIRA, G. C. & CARVALHO, O. S. 2010. A geoprocessing approach for studying and controlling schistosomiasis in the state of Minas Gerais, Brazil. *Memorias do Instituto Oswaldo Cruz*, 105, 524-531.

GUIMARAES, R. J. P. S., FREITAS, C. C., DUTRA, L. V., FELGUEIRAS, C. A., MOURA, A. C. M., AMARAL, R. S., DRUMMOND, S. C., SCHOLTE, R. G. C., OLIVEIRA, G. & CARVALHO, O. S. 2009. Spatial distribution of *Biomphalaria* mollusks at Sao Francisco River Basin, Minas Gerais, Brazil, using geostatistical procedures. *Acta Tropica*, 109, 181-186.

GUINDON, S. & GASCUEL, O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, 52, 696-704.

HAAS, W. & SCHMITT, R. 1982a. Characterization of chemical stimuli for the penetration of *Schistosoma mansoni* cercariae. I. Effective substances, host specificity. *Zeitschrift für Parasitenkunde*, 66, 293-307.

HAAS, W. & SCHMITT, R. 1982b. Characterization of chemical stimuli for the penetration of *Schistosoma mansoni* cercariae. II. Conditions and mode of action. *Zeitschrift für Parasitenkunde*, 66, 309-319.

HAAS, W., KÖRNER, M., HUTTERER, E., WEGNER, M. & HABERL, B. 1995. Finding an recognition of the snail intermediate hosts by 3 species of echinostome cercariae. *Parasitology*, 110, 133-142.

HABERL, B. & HAAS, W. 1992. Miracidium of *Schistosoma mansoni*: a macromolecular glycocongugate as signal for the behaviour after contact with the snail host. *Comparative biochemistry and physiology: Comparative physiology*, 101, 329-333.

HAJIBABAEI, M., SINGER, G. A. C., HERBERT, P. D. N. & HICKEY, D. A. 2007. DNA barcoding: How it complements taxonomy, molecular phylogenetics and population genetics. *Trends in Genetics*, 23, 167-172.

HANDZEL, T., KARANJA, D. M. S., ADDISS, D. G., HIGHTOWER, A. W., ROSEN, D. H., COLLEY, D. G., ANDOVE, J., SLUTSKER, L. & SECOR, W. E. 2003. Grographic distribution of schistosomiasis and soil-transmitted helminths in Western Kenya: Implications for anthelmintic mass treatment. *American Journal of Tropical Medicine and Hygiene*, 69, 318-323.

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* group. *Parasitology*, 136, 987-1001.

HANELT, B., MWANGI, I. N., KINUTHIA, J. M., MAINA, G. M., AGOLA, L. E., MUTUKU, M. W., STEINAUER, M. L., AGWANDA, B. R., KIGO, L., MUNGAI, B. N., LOKER, E. S. & MKOJI, G. M. 2010. Schistosomes of small mammals from the Lake Victoria Basin, Kenya: new species, familiar species, and implications for schistosomiasis control. *Parasitology*, 137, 1109-1118.

HARRISON, A. D., NDUKU, W. & HOOPER, A. S. 1966. The effects of a high magnesium-to-calcium ratio on the egg-laying rate of an aquatic planorbid snail, *Biomphalaria pfeifferi*. *Annals of Tropical Medicine and Parasitology*, 60, 212-214.

HASSAN, A. H. M., HABERL, B., HENTEL, J. & HAAS, W. 2003. Miracidia of an Egyptian strain of *Schistosoma mansoni* differentiate between sympatric snail species. *The Journal of Parasitology*, 89, 1248-1250.

HAYTER, A. J. 1984. A proof of the conjecture that the Tukey-Kramer multiple comparisons procedure is conservative. *The Annals of Statistics*, 12, 61-75.

HEBERT, P. D. N., RATNASINGHAM, S. & DEWAARD, J. R. 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society B: Biological Sciences*, 270, S96-99.

HEBERT, P. D. N., PENTON, E. H., BURNS, J. M., JANZEN, D. H. & HALLWACHS, W. 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 14812-14817.

HEDRICK, P. W., KIM, T. J. & PARKER, K. M. 2001. Parasite resistance and genetic variation in the endangered Gila topminnow. *Animal Conservation*, 4, 103-109.

HENDRY, A. P., FARRUGIA, T. J. & KINNISON, M. T. 2008. Human influences on rates of phenotypic change in wild animal populations. *Molecular Ecology*, 17, 20-29.

HENRIKSEN, U. B. & JELNES, J. E. 1980. Experimental taxonomy of *Biomphalaria* (Gastropoda: Planorbidae): I. Methods for experimental taxonomic studies on *Biomphalaria* carried out by horizontal starch gel electrophoresis and staining of twelve enzymes. *Journal of Chromatography A*, 188, 169-176.

HOFFMAN, J. I., WEBSTER, J. P., NDAMBA, J. & WOOLHOUSE, M. E. J. 1998. Extensive genetic variation revealed in adjacent populations of the schistosome intermediate host *Biomphalaria pfeifferi* from a single river system. *Annals of Tropical Medicine and Parasitology*, 92, 693-698.

HOFKIN, B. V., STRYKER, G. A., KOECH, D. K. & LOKER, E. S. 1991. Consumption of *Biomphalaria glabrata* egg masses and juveniles by the ampullariid snails *Pila ovata, Lanistes carinatus* and *Marisa cornuarietis*. *Acta Tropica,* 49, 37-44.

HOLMES, C. F. 1971. Zanzibari influence at the southern end of Lake Victoria: The lake route. *African Historical Studies,* 4, 477-503.

HOVERMAN, J. T. & RELYEA, R. A. 2007. How flexible is phenotypic plasticity? Developmental windows for trait induction and reversal. *Ecology,* 88, 693-705.

HSAIO, K. M., LIN, H. M., PAN, H., LI, T. C., CHEN, S. S., JOU, S. B., CHIU, Y. L., WU, M. F., LIN, C. C. & LI, S. Y. 1999. Application of FTA sample collection and DNA purification system on the determination of CTG trinucleotide repeat size by PCR-based Southern blotting. *Journal of Clinical Laboratory Analysis,* 13, 188-193.

HULSEY, D. C. 2006. Function of a key morphological innovation: fusion of the cichlid pharyngeal jaw. *Proceedings of the Royal Society B: Biological Sciences,* 273, 669-675.

IHAKA, R. & GENTLEMAN, R. 1996. R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics* 5, 299-314.

IRIE, T. 2006 Geographical variation of shell morphology in *Cypraea annulus* (Gastropoda: Cypraeidae). *Journal of Molluscan Studies,* 72, 31-38.

JARNE, P. & STÄDLER, T. 1995. Population genetic structure and mating system evolution in freshwater pulmonates. *Cellular and Molecular Life Sciences,* 51, 482-497.

JARNE, P. & THÉRON, A. 2001. Genetic structure in natural populations of flukes and snails: a practical approach and review. *Parasitology,* 123, S27-S40.

JEAL, T. 2007. *Stanley: The Impossible Life of Africa's Greatest Explorer* London, Faber & Faber.

JIAN, X., OLEA, R. A. & YU, Y.-S. 1996. Semivariogram modeling by weighted least squares. *Computers & Geosciences,* 22, 387-397.

JOHNSON, P. T. J., LUND, P. J., HARTSON, R. B. & YOSHINO, T. P. 2009a. Community diversity reduces *Schistosoma mansoni* transmission, host pathology and human infection risk. *Proceedings of the Royal Society B: Biological Sciences,* 276, 1657-1663.

JOHNSON, T. C., SCHOLZ, C. A., TALBOT, M. R., KELTS, K., RICKETTS, R. D., NGOBI, G., BEUNING, K., SSEMMANDA, I. & MCGILL, J. W. 1996. Late Pleistocene desiccation of Lake Victoria and rapid evolution of cichlid fishes. *Science,* 273, 1091-1093.

JOHNSON, T. C., KELTS, K. & ODADA, E. 2009b. The Holocene history of Lake Victoria. *Ambio,* 29, 2-11.

JOHNSON, P. T. J. & THIELTGES, D. W. 2010. Diversity, decoys and the dilution effect: how ecological communities affect disease risk. *Journal of Experimental Biology,* 213, 961-970.

JOHNSTON, H. H., TORDAY, E., JOYCE, T. A. & SELIGMANN, C. G. 1913. A survey of the ethnography of Africa: And the former racial and tribal Migrations in that continent. *The Journal of the Royal Anthropological Institute of Great Britain and Ireland,* 43, 375-421.

JONES, C. .S., LOCKYER, A. E., ROLLINSON, D., PIERTNEY, S. B. & NOBEL, L. R. 1999. Isolation and characterization of microsatellite loci in then freshwater gastropod, *Biomphalaria glabrata,* an intermediate host for *Schistosoma mansoni. Molecular Ecology,* 8, 2149-2151.

JORDAN, P. 1968. Chemotherapy of schistosomiasis. *Bulletin of the New York Academy of Medicine,* 44, 245-258.

JORDAN, P., CHRISTIE, J. D. & UNRAU, G. O. 1980. Schistosomiasis transmission with particular reference to possible ecological and biological methods of control - a review. *Acta Tropica,* 37, 95-135.

JORDAN, P., WEBBE, G. & STURROCK, R. F. 1993. *Human schistosomiasis,* Wallingford, CAB INTERNATIONAL.

JORDAN, P. 2000. From Katayama to the Dakhla Oasis: the beginning of epidemiology and control of bilharzia. *Acta Tropica,* 77, 9-40.

JØRGENSEN, A., KRISTENSEN, T. K. & STOTHARD, J. R. 2007. Phylogeny and biogeography of African *Biomphalaria* (Gastropoda: Planorbidae) with emphasis on endemic species of the great east African lakes. *Zoological Journal of the Linnean Society,* 151, 337-349.

KABATEREINE, N. B., BROOKER, S., TUKAHEBWA, E. M., KAZIBWE, F. & ONAPA, A. W. 2004. Epidemiology and geography of *Schistosoma mansoni* in Uganda: implications for planning control. *Tropical Medicine & International Health,* 9, 372-380.

KABATEREINE, N. B., TUKAHEBWA, E. M., KAZIBWE, F., TWA-TWA, J. M., BARENZI, J. F. Z., ZARAMBA, S., STOTHARD, J. R., FENWICK, A. & BROOKER, S. 2005. Soil-transmitted helminthiasis in Uganda: epidemiology and cost of control. *Tropical Medicine & International Health,* 10, 1187-1189.

KABATEREINE, N. B., FLEMING, F. M., NYANDINDI, U., MWANZA, J. C. L. & BLAIR, L. 2006a. The control of schistosomiasis and soil-transmitted helminths in East Africa. *Trends in Parasitology,* 22, 332-339.

KABATEREINE, N. B., TUKAHEBWA, E., KAZIBWE, F., NAMWANGYE, H., ZARAMBA, S., BROOKER, S., STOTHARD, J. R., KAMENKA, C., WHAWELL, S., WEBSTER, J. P. & FENWICK, A. 2006b. Progress towards countrywide control of schistosomiasis and soil-transmitted helminthiasis in Uganda. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 100, 208-215.

KABATEREINE, N. B., BROOKER, S., KOUKOUNARI, A., KAZIBWE, F., TUKAHEBWA, E. M., FLEMING, F. M., ZHANG, Y. B., WEBSTER, J. P., STOTHARD, J. R. & FENWICK, A. 2007. Impact of a national helminth control programme on infection and morbidity in Ugandan schoolchildren. *Bulletin of the World Health Organization,* 85, 91-99.

KALBE, M., HABERL, B & HAAS, W. 1996. *Schistosoma mansoni* host-finding: species specificity of an Egyptian strain. *Parasitology Research,* 82, 8-13.

KALLESTRUP, P., ZINYAMA, R., GOMO, E., BUTTERWORTH, A. E., VAN DAM, G. J., ERIKSTRUP, C. & ULLUM, H. 2005. Schistosomiasis and HIV-1 infection in rural Zimbabwe: implications of coinfection for excretion of eggs. *Journal of Infectious Diseases,* 191, 1311-1320

KALTZ, O. & SHYKOFF, J. A. 1998. Local adaptation in host-parasite systems. *Heredity,* 81, 361-370.

KANE, R. A. & ROLLINSON, D. 1994. Repetitive sequences in the ribosomal DNA internal transcribed spacer of *Schistosoma haematobium, Schistosoma intercalatum* and *Schistosoma mattheei. Molecular and Biochemical Parasitology,* 63, 153-156.

KAPITO-TEMBO, A. P., MWAPASA, V., MASHNICK, S. R., SAMANYIKA, Y., BANDA, D., BOWIE, C. & RADKE, S. 2009. Prevalence, distribution and risk factors for *Schistosoma haematobium* infection among school children in Blantyre, Malawi. *PLoS Neglected Tropical Diseases,* 3, e361.

KARANJA, D. M. S., COLLEY, D. G., NAHLEN, B. L. OUMA, J. H. & SECOR, W. E. 1997. Studies on schistosomiasis in western Kenya: I. Evidence for immune-facilitated excretion of

schistosome eggs from patients with *Schistosoma mansoni* and Human Immunodeficieny Virus co-infections. *American Journal of Tropical Medicine and Hygiene,* 56, 515-521.

KATZ, N., CHAVES, A. & PELLEGRINO, J. 1972. A simple device for quantitative stool thick-smear technique in schistosomiasis mansoni. *Revista Instituto de Medicina Tropical de Sao Paulo,* 14, 397-400.

KATZ, N., ROCHA, R. S. & CHAVES, A. 1979. Preliminary trials with praziquantel in human infections due to *Schistosoma mansoni. Bulletin of the World Health Organization,* 57, 781-785.

KAZIBWE, F., MAKANGA, B., RUBAIRE-AKIIKI, C., OUMA, J., KARIUKI, C., KABATEREINE, N. B., BOOTH, M., VENNERVALD, B. J., STURROCK, R. F. & STOTHARD, J. R. 2006. Ecology of *Biomphalaria* (Gastropoda : Planorbidae) in Lake Albert, Western Uganda: snail distributions, infection with schistosomes and temporal associations with environmental dynamics. *Hydrobiologia,* 568, 433-444.

KAZIBWE, F., MAKANGA, B., RUBAIRE-AKIIKI, C., OUMA, J., KARIUKI, C., KABATEREINE, N. B., VENNERVALD, B. J., ROLLINSON, D. & STOTHARD, J. R. 2010. Transmission studies of intestinal schistosomiasis in Lake Albert, Uganda, and experimental compatibiity of local *Biomphalaria* spp. *Parasitology International,* 59, 49-53.

KEENEY, D. .B, KING, T. M., ROWE, D. L. & POULIN, R. 2009. Contrasting mtDNA diversity and population structure in a direct-developing marine gastropod and its trematode parasites. *Molecular Ecology,* 18, 4591-4603.

KEESING, F., HOLT, R. D. & OSTFELD, R. S. 2006. Effects of species diversity on disease risk. *Ecology Letters,* 9, 485-498.

KISHINO, H. & HASEGAWA, M. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in hominoidea. *Journal of Molecular Evolution,* 29, 170-179.

KITCHELL, J. F., SCHINDLER, D. E., OGUTU-OHWAYO, R. & REINTHAL, P. N. 1997. The Nile Perch in Lake Victoria: Interactions between predation and fisheries. *Ecological Applications,* 7, 653-664.

KIWANGO, Y. & WOLANSKI, E. 2008. Papyrus wetlands, nutrients balance, fisheries collapse, food security, and Lake Victoria level decline in 2000–2006. *Wetlands Ecology and Management,* 16, 89-96.

KLOOS, H. 1995. Human behavior, health education and schistosomiasis control: A review. *Social Science and Medicine,* 40, 1497-1511.

KNOPP, S., RINALDI, L., KHAMIS, S. I., STOTHARD, J. R., ROLLINSON, D., MAURELLI, M. P., STEINMANN, P., MARTI, H., CRINGOLI, G. & UTZINGER, J. 2009. A single FLOTAC is more sensitive than triplicate Kato-Katz for the diagnosis of low-intensity soil-transmitted helminth infections. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 103, 347-354.

KONGS, A., MARKS, G., VERLE, P. & VAN DER STUYFT, P. 2001. The unreliability of the Kato-Katz technique limits its usefulness for evaluating *S. mansoni* infections. *Tropical Medicine & International Health,* 6, 163-169.

KOSKELLA, B. & LIVELY, C. M. 2009. Evidence for negative frequency-dependent selection during experimental coevolution of a freshwater snail and a sterilizing trematode. *Evolution,* 63, 2213-2221.

KRISTENSEN, T. K., YOUSIF, F. & RAAHAUGE, P. 1999. Molecular characterisation of *Biomphalaria* spp. in Egypt. *Journal of Molluscan Studies,* 65, 133-136.

KRISTENSEN, T. K., MALONE, J. B. & MCCARROLL, J. C. 2001. Use of satellite remote sensing and geographic information systems to model the distribution and abundance of snail intermediate hosts in Africa: a preliminary model for *Biomphalaria pfeifferi* in Ethiopia. *Acta Tropica,* 79, 73-78.

KUMAR, P. N., PATRA, K. P., HOTI, S. L. & DAS, P. K. 2002. Genetic variability of the human filarial parasite, *Wuchereria bancrofti* in South India. *Acta Tropica,* 82, 67-76.

KUSEL, J. R. 1970. Studies on the structure and hatching of the eggs of *Schistosoma mansoni. Parasitology,* 60, 79-88.

LAKWO, T. L. & ODONGOAGINYA, E. I. 1990. Schistosomiasis on the Entebbe peninsula. *East African Medical Journal,* 67, 43-47.

LAM, P. K. S. & CALOW, P. 1988. Differences in the shell shape of *Lymnaea peregra* (Müller) (Gastropoda: Pulmonata) from lotic and lentic habitats; environmental or genetic variance? *Journal of Molluscan Studies,* 54, 197-207.

LAMBERTUCCI, J. R., SERUFO, J. C., GERSPACHER-LARA, R., RAYES, A. A. M., TEIXIERA, R., NOBRE, V. & ANTUNES, C. M. F. 2000. *Schistosoma mansoni*: assessment of morbidity before and after control. *Acta Tropica,* 77, 101-109.

LANGE, C. N. 2005. *Diversity patterns of freshwater gastropods in lake habitats with and without anthropogenic disturbances in Lake Victoria, Kenya.* Ph.D., University of Copenhagen.

LEAKEY, L. S. B., TOBIAS, P. V. & NAPIER, J. R. 1964. A new species of the genus *Homo* from Olduvai Gorge. *Nature,* 202, 7-9.

LEGESSE, M. & ERKO, B. 2007. Field-based evalution of a reagent strip test for diagnosis of *Schistosoma mansoni* by detecting circulating cathodic antigen iun urine before and after chemotherapy. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 101, 668-673.

LENZI, H., ROMANHA, W. D. S., SANTOS, R. Z. D., ROSAS, A., MOTA, E., MANSO, P., CAPUTO, L. & PELAJO-MACHADO, M. 2006. Four whole-istic aspects of schistosome granuloma biology: fractal arrangement, internal regulation, autopoietic component and closure. *Memorias do Instituto Oswaldo Cruz,* 101, 219-231.

LIANG, S., YANG, C., ZHONG, B. & QIU, D. 2006. Re-emerging schistosomiasis in hilly and mountainous areas of Sichuan, China. *Bulletin of the World Health Organization,* 84, 139-144.

LITTLEWOOD, D. T. & BRAY, R. S. 2000. *Interrelationships of the Platyhelminthes,* CRC Press.

LITTLEWOOD, D. T., TELFORD, M. J. & BRAY, R. A. 2004. Protostomes and Platyhelminthes. *In:* CRACRAFT, J. & DONOGHUE, M. J. (eds.) *Assembling the Tree of Life.* New York: Oxford University Press USA.

LITTLEWOOD, D. T. J., LOCKYER, A. E., WEBSTER, B. L., JOHNSTON, D. A. & LE, T. H. 2006. The complete mitochondrial genomes of *Schistosoma haematobium* and *Schistosoma spindale* and the evolutionary history of mitochondrial genome changes among parasitic flatworms. *Molecular Phylogenetics and Evolution,* 39, 452-467.

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.

LOFTY, W. M., DEJONG, R. J., ABDEL-KADER, A. & LOKER, E. S. 2005. A molecular survey of *Biomphalaria* in Egypt: is *B. glabrata* present? *American Journal of Tropical Medicine and Hygiene,* 73, 131-139.

LOHMANN, G. P. & SCHWEITZER, P. N. Year. On eigenshape analysis. *In:* ROHLF, F. J. & BOOKSTEIN, F. L., eds. Proceedings of the Michigan Morphometrics Workshop, 1990 University of Michigan Museum of Zoology, Ann Arbor. 145–166.

LOREAU, M. & BALUKU, B. 1987. Population dynamics of the freshwater snail *Biomphalaria pfeifferi* in Eastern Zaire. *Journal of Molluscan Studies,* 53, 249-265.

LOVERDE, P. T., DEWALD, J., MINCHELLA, D. J., BOSSHARDT, S. C. & DAMIAN, R. T. 1985. Evidence for host-induced selection in *Schistosoma mansoni. The Journal of Parasitology,* 71, 297-301.

LUNN, D. J., THOMAS, A., BEST, N. & SPIEGELHALTER, D. 2000. WinBUGS - a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing,* 10, 325-337.

LURIE, M. N., WILLIAMS, B. G., ZUMA, K., MKAYA-MWAMBURI, D., GARNETT, G. P., STURM, A. W., SWEAT, M. D., GITTELSOHN, J. & KARIM, S. S. A. 2003. The impact of migration on HIV-1 transmission in South Africa - A study of migrant and nonmigrant men and their partners. *Sexually Transmitted Diseases,* 30, 149-156.

MACEDO, A. M. & PENA, S. D. J. 1998. Genetic variability of *Trypanosoma cruzi*: Implications for the pathogenesis of Chagas Disease. *Parasitology Today,* 14, 119-124.

MACLEOD, N. 1999. Generalizing and extending the eigenshape method of shape space visualization and analysis. *Paleobiology,* 25, 107-138.

MADDOX, G., GIBLIN, J. L. & KIMAMBO, I. N. 1996. *Custodians of the land: ecology & culture in the history of Tanzania,* Athens, Ohio University Press.

MAGENDANTZ, M. 1972. The biology of *Biomphalaria choanomphala* and *B. sudanica* in relation to their role in the transmission of *Schistosoma mansoni* in Lake Victoria at Mwanza, Tanzania. *Bulletin of the World Health Organization,* 47, 331-341.

MAGNUSSEN, P. 2003. Treatment and re-treatment strategies for schistosomiasis control in different epidemiological settings: a review of 10 years' experiences. *Acta Tropica,* 86, 243-254.

MAJOROS, G., FEHER, Z., DELI, T. & FOLDVARI, G. 2008. Establishment of *Biomphalaria tenagophila* snails in Europe. *Emerging Infectious Diseases,* 14, 1812-1814.

MALEK, E. A. 1976. Effect of the Aswan Dam on prevalence of schistosomiasis in Egypt. *Tropical and Geographical Medicine,* 27, 359-364.

MALEK, E. A. 1985. *Snail hosts of schistosomiasis and other snail-transmitted tropical diseases in tropical America,* Washington, PAHO & WHO.

MANDAHL-BARTH, G. 1957. Intermediate hosts of *Schistosoma.* African *Biomphalaria* and *Bulinus*: 1. *Biomphalaria. Bulletin of the World Health Organization,* 15, 1103–1163.

MANDAHL-BARTH, G. 1958. Intermediate hosts of *Schistosoma*: African *Biomphalaria* and *Bulinus. World Health Organisation Monograph Series.* Geneva: World Health Organisation.

MANDAHL-BARTH, G. 1962. Key to the identification of East and Central African freshwater snails of medical and veterinary importance. *Bulletin of the World Health Organization,* 27, 135-150.

MANTEL, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research,* 27, 209-220.

MARTIN, L. K. & BEAVER, P. C. 1968. Evaluation of Kato thick-smear technique for quantitative diagnosis of helminth infections. *American Journal of Tropical Medicine and Hygiene,* 17, 382-391.

MARTINS, A. V. 1958. Non-human vertebrate hosts of *Schistosoma haematobium* and *Schistosoma mansoni*. *Bulletin of the World Health Organization*, 18, 931-944.

MASIFWA, W. F., TWONGO, T. & DENNY, P. 2001. The impact of water hyacinth, *Eichhornia crassipes* (Mart) Solms on the abundance and diversity of aquatic macroinvertebrates along the shores of northern Lake Victoria, Uganda. *Hydrobiologia*, 452, 79-88.

MAVÁREZ, J., AMARISTA, M., POINTIER, J.-P. & JARNE, P. Microsatellite variation in the freshwater schistosome-transmitting snail *Biomphalaria glabrata*. *Molecular Ecology*, 9, 1009-1011.

MAVÁREZ, J., POINTIER, J.-P., DAVID, P., DELAY, B. & JARNE, P. 2002. Genetic differentiation, dispersal andmating system in the schistosome-transmitting freshwater snail *Biomphalaria glabrata*. *Heredity*, 89, 258-265.

MCBRATNEY, A. B. & WEBSTER, R. 1986. Choosing functions for semi-variograms of soil properties and fitting them to sampling estimates. *European Journal of Soil Science*, 37, 617-639.

MCCLELLAND, W. F. J. & JORDAN, P. 1962. Schistosomiasis at Bukoba, Tanganyika, on Lake Victoria. *Annals of Tropical Medicine and Parasitology*, 56, 396.

MCCULLOUGH, F. S. 1972. Distribution of *Schistosoma mansoni* and *S. haematobium* in East Africa. *Tropical and Geographical Medicine*, 24, 199-207.

MCKERROW, J. H. & SALTER, J. 2002. Invasion of skin by *Schistosoma* cercariae. *Trends in Parasitology*, 18, 193-195.

MELLO, D. A. 1972. The comparative morphology of the genital system of some African species of *Biomphalaria* (Mollusca, Planorbidae). *Revista Brasileira de Biologia*, 32, 443-450.

MILLER, J. H. 1960. *Papio doguera*, (dog-face baboon), a primate reservoir of *Schistosoma mansoni* in East Africa. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 54, 44-46.

MILLIEN, V., KATHLEEN LYONS, S., OLSON, L., SMITH, F. A., WILSON, A. B. & YOM-TOV, Y. 2006. Ecotypic variation in the context of global climate change: revisiting the rules. *Ecology Letters*, 9, 853-869.

MINCHELLA, D. J. & LOVERDE, P. T. 1981. A cost of increased early reproductive effort in the snail *Biomphalaria glabrata*. *The American Naturalist*, 118, 876-881.

MINCHELLA, D. J., SOLLENBERGER, K. M. & PEREIRA DE SOUZA, C. 1995. Distribution of schistosome genetic diversity within molluscan intermediate hosts. *Parasitology*, 111, 217-220.

MKOJI, G. M., MUNGAI, B. N., KOECH, D. K. & LOKER, E. S. 1998. Experimental control of the schistosome-transmitting snail *Biomphalaria pfeifferi* by the ampullariid snail *Pila ovata*. *Annals of Tropical Medicine and Parasitology*, 92, 65-72.

MODENA, C. M., LIMA, W. D. S. & COELHO, P. M. Z. 2008. Wild and domesticated animals as reservoirs of *Schistosomiasis mansoni* in Brazil. *Acta Tropica*, 108, 242-244.

MONTRESOR, A., CROMPTON, D. W. T., BUNDY, D. A. P., HALL, A. & SAVIOLI, L. 1998. *Guidelines for the Evaluation of Soil-Transmitted Helminthiasis and Schistosomiasis at the Community Level*. World Health Organization, Geneva. WHO/CTD/SIP/98.1

MORGAN, J. A. T., DEJONG, R. J., SNYDER, S. D., MKOJI, G. M. & LOKER, E. S. 2001. *Schistosoma mansoni* and *Biomphalaria*: past history and future trends. *Parasitology*, 123, S211-S228.

MORGAN, J. A. T., DEJONG, R. J., KAZIBWE, F., MKOJI, G. M. & LOKER, E. S. 2003. A newly-identified lineage of *Schistosoma*. *International Journal of Parasitology*, 33, 977-985.

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*. *The Journal of Parasitology*, 89, 416-418.

MORGAN, J. A. T., DEJONG, R. J., ADEOYE, G. O., ANSA, E. D. O., BARBOSA, C. S., BREMOND, P., CESARI, I. M., CHARBONNEL, N., CORREA, L. R., COULIBALY, G., D'ANDREA, P. S., DE SOUZA, C. P., DOENHOFF, M. J., FILE, S., IDRIS, M. A., INCANI, R. N., JARNE, P., KARANJA, D. M. S., KAZIBWE, F., KPIKPI, J., LWAMBO, N. J. S., MABAYE, A., MAGALHAES, L. A., MAKUNDI, A., MONE, H., MOUAHID, G., MUCHEMI, G. M., MUNGAI, B. N., SENE, M., SOUTHGATE, V., TCHUENTE, L. A. T., THÉRON, A., YOUSIF, F., MAGALHAES, E. M. Z., MKOJI, G. M. & LOKER, E. S. 2005. Origin and diversification of the human parasite *Schistosoma mansoni*. *Molecular Ecology*, 14, 3889-3902.

MORRISON, L. J., MCLELLAN, S., SWEENEY, L., CHAN, C. N., MACLEOD, A., TAIT, A. & TURNER, C. M. R. 2010. Role for parasite genetic diversity in differential host responses to *Trypanosoma brucei* infection. *Infection and Immunity*, 78, 1096-1108.

MUGISHA, L., KUCHERER, C., ELLERBROK, H., JUNGLEN, S., OPUDA-ASIBO, J., JOSEPH, O. O., PAULI, G. & LEENDERTZ, F. H. 2010. Retroviruses in wild-born semi-captive East Africa sanctuary chimpanzees (*Pan trogodytes schweinfurthii*). *Open Veterinary Science Journal*, 4, 6-10.

MUTAPI, F., GRYSEELS, B. & RODDAM, A. 2003. On the calculation of intestinal schistosome infection intensity. *Acta Tropica*, 87, 225-233.

MURRAY, S., STEM, C., BOUDREAU, B. & GOODALL, J. 2000. Intestinal Parasites of Baboons (*Papio cynocephalus anubis*) and Chimpanzees (*Pan troglodytes*) in Gombe National Park. *Journal of Zoo and Wildlife Medicine*, 31, 176-178.

NELSON, G. S. 1958. *Schistosoma mansoni* infection in the West Nile District of Uganda. Part 1.The incidence of *S. mansoni* infection. *East African Medical Journal*, 35, 312-319.

NELSON, G. S. 1960. Schistosome infections as zoonoses in Africa. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 54, 301-316.

NELSON, G. S. & TEESDALE, C. 1965. *Schistosoma rodhaini* "taken for a ride" by *S. mansoni* in baboons (Demonstration). *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 59 (1), 7.

NIESTERS, H. G. 2002. Clinical virology in real time. *Journal of Clinical Virology*, 25 (Supp), S3-S12.

NOKES, C., MCGARVEY, S. T., SHIUE, L., WU, G., WU, H., BUNDY, D. A. & OLDS, G. R. 1999. Evidence for an improvement in cognitive function following treatment of *Schistosoma japonicum* infection in Chinese primary school children. *American Journal of Tropical Medicine and Hygiene*, 60, 556-565.

NORTON, A. J., WEBSTER, J. P., KANE, R. A. & ROLLINSON, D. 2008. Inter-specific parasite competition: mixed infections of *Schistosoma mansoni* and *S. rodhaini* in the definitive host. *Parasitology*, 135, 473-484.

NUTTER, F. 1993. A comparison of gastrointestinal parasites in two communities of chimpanzees at Gombe National Park, Tanzania. PhD thesis, Tufts University of Veterinary Medicine.

NYBAKKEN, J. & PERRON, F. 1988. Ontogenetic change in the radula of *Conus magus* (Gastropoda). *Marine Biology*, 98, 239-242.

OBENG, B. B., ARYEETEY, Y. A., DE DOOD, C. J., AMOAH, A. S., LARBI, I. A., DEELDER, A. M. TAZDANBAKHSH, M. HARTGERS, F. C., BOAKYE, D. A., VERWEIJ, J. J., VAN DAM, G. J. & VAN LIESHOUT, L. 2008. Application of a circulating-cathodic-antigen (CCA) strip test and real-

time PCR, in comparison with microscopy, for the detection of *Schistosoma haematobium* in urine samples from Ghana. *Annals of Tropical Medicine and Parasitology,* 102, 625-633.

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 (<3 years) along the Ugandan shoreline of Lake Victoria. *Annals of Tropical Medicine and Parasitology,* 100 315-326.

ODONGO-AGINYA, E. I. & MUGISA, C. 1987. The prevalence of *Schistosoma mansoni* in migrants from endemic areas living in the Entebbe Peninsula, Uganda. *East African Medical Journal,* 64, 571-579.

ODONGO-AGINYA, E. I. & EKKEHARD, D. 2008. Epidemiology, of Bilharzias (schistosomiasis) in Uganda from 1902 until 2005. *African Health Sciences,* 8, 239-243.

OGOT, B. A. 1967. *History of the southern Luo,* Nairobi, East African Pub. House.

OGUTU-OHWAYO, R. 1990. The decline of the native fishes of lakes Victoria and Kyoga (East Africa) and the impact of introduced species, especially the Nile perch, *Lates niloticus,* and the Nile tilapia, *Oreochromis niloticus. Environmental Biology of Fishes,* 27, 81-96.

OGUTU-OHWAYO, R. 2001. The fisheries of Lake Victoria: Harvesting biomass at the expense of biodiversity. *Thematic Workshop Blue Millennium: Managing Global Fisheries for Biodiversity.* Victoria, BC, Canada: UNEP / IDRC.

OLIVER, M. A. & WEBSTER, R. 1990. Kriging: a method of interpolation for geographical information system. *International Journal of Geographical Information Systems,* 4, 313-332.

OMONDI, R. & MWENDI, K. 2006. Macrophytes of Lake Victoria and succession after invasion of water hyacinth. *In:* 11th World Lakes Conference, 2006 Nairobi, Kenya. 600-602.

OPANDE, G. O., ONYANGO, J. C. & WAGAI, S. O. 2004. Lake Victoria: The water hyacinth (*Eichhornia crassipes* [MART.] SOLMS), its socio-economic effects, control measures and resurgence in the Winam gulf. *Limnologica,* 34, 105-109.

OSTFELD, R. S. & KEESING, F. The function of biodiversity in the ecology of vector-borne zoonotic diseases. *Canadian Journal of Zoology,* 78, 2061-2078.

OSTFELD, R. S. & HOLT, R. D. 2004. Are predators good for your health? Evaluating evidence for top-down regulation of zoonotic disease reservoirs. *Frontiers in Ecology and the Environment,* 2, 13-20.

PARAENSE, W. L. 1981. *Biomphalaria occidentalis* sp.n. from South America (Mollusca Basommatophora Pulmonata). *Memorias Do Instituto Oswaldo Cruz,* 76, 199-211.

PARAENSE, W. L. 2001. The schistosome vectors in the Americas. *Memórias do Instituto Oswaldo Cruz,* 96, 7-16.

PEARMAN, P. B. & GARNER, T. W. J. 2005. Susceptibility of Italian agile frog populations to an emerging strain of *Ranavirus* parallels population genetic diversity. *Ecology Letters,* 8, 401-408.

PETERS, W. 1978. Degradation of the radula in the snails *Biomphalaria glabrata* Say and *Limnaea stagnalis* L. (Gastropoda, Pulmonata). *Cell and Tissue Research,* 193, 283-295.

PETRŽELKOVÁ, K. J., HASEGAWA, H., APPLETON, C. C., HUFFMAN, M. A., ARCHER, C. E., MOSCOVICE, L. R., MAPUA, M. I., SINGH, J. & KAUR, T. 2010. Gastrointestinal parasites of the chimpanzee population introduced onto Rubondo Island National Park, Tanzania. *American Journal of Primatology,* 72, 307-316.

PICOZZI, K., TILLEY, A., FEVRE, E. M., COLEMAN, P. G., MAGONA, J. W., ODIIT, M., EISLER, M. C. & WELBURN, S. C. 2002. The diagnosis of trypanosome infections: applications of novel technology for reducing disease risk. *African Journal of Biotechnology,* 1, 39-45.

PIMENTEL, D. 1957. Life history of *Australorbis glabratus,* the intermediate snail host of *Schistosoma mansoni* in Puerto Rico. *Ecology,* 38, 576-580.

PITCHFORD, R. J. 1959. Cattle schistosomiasis in man in the Eastern Transvaal. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 53, 285-290.

PLAM, M., JØRGENSEN, A., KRISTENSEN, T. K. & MADSEN, H. 2008. Sympatric *Biomphalaria* species (Gastropoda: Planorbidae) in Lake Albert, Uganda, show homoplasies in shell morphology. *African Zoology,* 43, 34-44.

PLUMMER, M. 2005. Impact of invasive water hyacinth (*Eichhornia crassipes*) on snail hosts of schistosomiasis in Lake Victoria, East Africa. *Ecohealth,* 2, 81-86.

POINTIER, J.-P., DEJONG, R. J., TCHUEM TCHUENTÉ, L. A., KRISTENSEN, T. K. & LOKER, E. S. 2005. A neotropical snail host of *Schistosoma mansoni* introduced into Africa and consequences for the schistosomiasis transmission: *Biomphalaria tenagophila* in Kinshasa (Democratic Republic of Congo). *Acta Tropica,* 93, 191-199.

PONTES, L. A., OLIVIERA, M. C., KATZ, N., DIAS-NETO, E. & RABELLO, A. 2003. Comparison of a polymerase chain reaction and the Kato-Katz technique for diagnosing infection with *Schistosoma mansoni. American Journal of Tropical Medicine and Hygiene,* 68, 652-656.

PRENTICE, M. A. 1972. Distribution, prevalence and transmission of schistosomiasis in Uganda. *Uganda Medical Journal,* 1, 136-140.

PRESTON, H. B. 1910. Addition to the non-marine molluscan fauna of British and German East Africa and Lake Albert Edward. *The Annals and Magazine of Natural History including zoology, botany, and geology,* 6, 526-535.

PRUGNOLLE, F., LIU, H., MEEUS, T. & BALLOUX, F. 2005. Population genetics of complex life-cycle parasites: an illustration with trematodes. *International Journal for Parasitology,* 35, 255-267.

RAMASWAMY, K., HE, Y.-X., SALAFSKY, B. & SHIBUYA, T. 2003. Topical application of DEET for schistosomiasis. *Trends in Parasitology,* 19, 551-555.

RASO, G., MATTHYS, B., N'GORAN, E. K., TANNER, M., VOUNATSOU, P. & UTZINGER, J. 2005. Spatial risk prediction and mapping of *Schistosoma mansoni* infections among school children living in Western Côte d'Ivoire. *Parasitology,* 131, 97-108.

RATNASINGHAM, S. & HEBERT, P. D. N. 2007. The Barcode of Life Data System (http://www.barcodinglife.org). *Molecular Ecology Notes,* 7, 355-364.

RAYMOND, M. & ROUSSET, F. 1995. Genepop (Version-1.2) - Population Genetics Software for Exact Tests and Ecumenicism. *Journal of Heredity,* 86, 248-249.

RESH, V. H., BROWN, A. V., COVICH, A. P., GURTZ, M. E., LI, H. W., MINSHALL, G. W., REICE, S. R., SHELDON, A. L., WALLACE, J. B. & WISSMAR, R. C. 1988. The role of disturbance in stream ecology. *Journal of the North American Benthological Society,* 7, 433-455

REYNOLDS, J. E. & KITAKULE, J. A. 1991. Field Document No. 2: Socioeconomic aspectes of Lake Victoria and Albert fisheries: The 1991 FISHIN' community survey. *FISHIN Notes and Records.* FAO/UNDP.

ROBERTS, A. 2007. *Uganda's Great Rift Valley,* Kampala, Graphic Systems.

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. *International Journal for Parasitology,* 32, 843-851.

ROGERS, A. R. & HARPENDING, H. 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution,* 9, 552-569.

ROHLF, F. J. 2007a. *TpsDIG2.* Department of Ecology and Evolution, State University of New York at Stony Brook

ROHLF, F. J. 2007b. *TpsUTIL.* Department of Ecology and Evolution, State University of New York at Stony Brook

ROLLINSON, D. & SOUTHGATE, V. R. 1987. The genus *Schistosoma*: a taxonomic appraisal. *In:* ROLLINSON, D. & SIMPSON, A. J. (eds.) *The biology of schistosomes. From genes to latrines.* London: Academic Press.

ROLLINSON, D., STOTHARD, J. R. & SOUTHGATE, V. R. 2001. Interactions between intermediate host snails of the genus *Bulinus* and schistosomes of the *Schistosoma haematobium* group. *Parasitology,* 123, S245-S260.

ROLLINSON, D., WEBSTER, J. P., WEBSTER, B., NYAKAANA, S., JØRGENSEN, A. & STOTHARD, J. R. 2009. Genetic diversity of schistosomes and snails: implications for control. *Parasitology,* 136, 1801-1811.

RONQUIST, F. & HUELSENBECK, J. P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics,* 19, 1572-1574.

ROUSSET, F. 2008. Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Molecular Ecology Resources,* 8, 103-106.

RUBAIHAYO, J., MOGHUSU, E., CLOUDS, P. & ABAASA, A. 2008. Schistosomiasis transmission at high altitude crater lakes in Western Uganda. *BMC Infectious Diseases,* 8, e110.

RUDGE, J. W., STOTHARD, J. R., BASANEZ, M. G., MGENI, A. F., KHAMIS, I. S., KHAMIS, A. N. & ROLLINSON, D. 2008. Micro-epidemiology of urinary schistosomiasis in Zanzibar: Local risk factors associated with distribution of infections among schoolchildren and relevance for control. *Acta Tropica,* 105, 45-54.

SADUN, E. H., VON LICHTENBERG, F. & BRUCE, J. I. 1966. Susceptibility and comparative pathology of ten species of primates exposed to infection with *Schistosoma mansoni. American Journal of Tropical Medicine and Hygiene,* 15, 705-718.

SADUN, E. H. & GORE, R. W. 1970. *Schistosoma mansoni* and *S. haematobium*: Homocytotropic reagin-like antibodies in infections of man and experimental animals. *Experimental Parasitology,* 28, 435-449.

SAMUELSON, J. C., QUINN, J. J. & CAULFIELD, J. P. 1984. Hatching, chemokinesis, and transformation of miracidia of *Schistosoma mansoni. The Journal of Parasitology,* 70, 321-331.

SANDBACH, F. R. 1976. The history of schistosomiasis research and policy for its contro. *Medical History,* 20, 259–275.

SANDLAND, G. J., WETHINGTON, A. R., FOSTER, A.V. & MINCHELLA, D. A. 2009. Effects of host outcrossing on the interaction between an aquatic snail and its locally adapted parasite. *Parasitology Research,* 105, 555-561.

SAVIOLI, L., RENGANATHAN, E., MONTRESOR, A., DAVIS, A. & BEHBEHANI, K. 1997. Control of schistosomiasis -- A global picture. *Parasitology Today,* 13, 444-448.

SCHUTTE, C. H. J. & FRANK, G. H. 1964. Observations on the distribution of freshwater mollusca and chemistry of the natural waters in the south-eastern Transvaal and adjacent northern Swaziland. *Bulletin of the World Health Organization,* 30, 389-400.

SCHWETZ, J. 1954. On two schistosomes of wild rodents of the Belgian Congo: *Schistosoma rodhaini* Brumpt, 1931; and *Schistosoma mansoni* var. *rodentorum* ; and their relationship to *S. mansoni* of man. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 48, 89-100.

SCOTT, J. T., VEREECKEN, K, DIAKHATÉ, M. M., VANAGT, T., SAMBOU, V., DIOP, M., LY, A., VERCRUYSSE, J., GRYSEELS, B. & KESTENS, L. 2004. Are the cellular responses of children and adults with *Schistosoma mansoni* infection intrinsically different? Cytokines produced *ex vivo* in response to antigens and mirogens. *Parasitological Immunology,* 26, 29-36.

SHRIVASTAVA, J., QIAN, B. & WEBSTER, J. P. 2005. An insight into the genetic variation of *Schistosoma japonicum* in mainland China using DNA microsatellite markers. *Molecular Ecology,* 14, 839-849.

SHIMODAIRA, H. & HASEGAWA, M. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular Biology and Evolution,* 16., 1114.

SIMOONGA, C., UTZINGER, J., BROOKER, S., VOUNATSOU, P., APPLETON, C. C., STENSGAARD, A. S., OLSEN, A. & KRISTENSEN, T. K. 2009. Remote sensing, geographical information system and spatial analysis for schistosomiasis epidemiology and ecology in Africa. *Parasitology,* 136, 1683-1693.

SIRE, C., ROGNON, A. & THÉRON, A. 1998. Failure of *Schistosoma mansoni* to reinfect *Biomphalaria glabrata* snails: acquired humoral resistance or intra-specific larval antagonism? *Parasitology,* 117, 117-122.

SIRE, C., LANGARD , J., BARRAL, V. & THÉRON, A. 2001. Paraaite (*Schistosoma mansoni*) and host *(Biomphalaria glabrata*) genetic diversity: population structure in a fragmented landscape. *Parasitology,* 122, 545-554.

SLOOTWEG, R. 1987. Prey selection by molluscivorous cichlids foraging on a schistosomiasis vector snail, *Biomphalaria glabrata. Oecologia,* 74, 193-202.

SMITHERS, S. R. & TERRY, R. J. 1965. The infection of laboratory hosts with cercariae of *Schistosoma mansoni* and the recovery of the adult worms. *Parasitology,* 55, 695-700.

SMOLDERS, A., LUCASSEN, E., BOBBINK, R., ROELOFS, J. & LAMERS, L. 2010. How nitrate leaching from agricultural lands provokes phosphate eutrophication in groundwater fed wetlands: the sulphur bridge. *Biogeochemistry,* 98, 1-7.

SNYDER, S. D. & LOKER, E. S. 2000. Evolutionary relationships among the Schistosomatidae (Platyhelminthes: Digenea) and an Asian orign for *Schistosoma. The Journal of Parasitology,* 86, 283-288.

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.

SOUSA-FIGUEIREDO, J. C., PLEASANT, J., DAY, M., BETSON, M., ROLLINSON, D., MONTRESOR, A., KIZIBWE, F., KABATEREINE, N. B. & STOTHARD, J. R. 2010. Treatment of intestinal schistosomaisis in Ugandan preschool children: best diagnosis, treatment efficacy and side-effects, and an extended praziquantel dosing pole. *International Health,* 2, 103-113.

SOUTHGATE, V. R., TCHUEM TCHUENTÉ, L. A., THÉRON, A., JOURDANE, J., LY, A., MONCRIEFF, C. B. & GRYSEELS, B. 2000. Compatibility of *Schistosoma mansoni* Cameroon and *Biomphalaria pfeifferi* Senegal. *Parasitology,* 121, 501-505.

SOUTHGATE, V. R., TCHUEM TCHUENTÉ, L. A., SÈNE, M., DE CLERCQ, D., THÉRON, A., JOURDANE, J., WEBSTER, B. L., ROLLINSON, D., GRYSEEELS, G. & VERCRUYSSE, J. 2001. Studies on the biology of schistosomiasis with emphasis on the Senegal river basin. *Memorias do Instituto Oswaldo Cruz*, 96 (Supp), 75-78.

SOW, S., POLMAN, K., VEREECKEN, K., VERCRUYSSE, J., GRYSEELS, B. & DE VLAS, S. J. 2008. The role of hygienic bathing after defecation in the transmission of *Schistosoma mansoni*. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 102, 542-547.

SPATZ, L., VIDIGAL, T. H. D. A., SILVA, M. C. A., CAPPA, S. M. G. & CARVALHO, O. S. 2000. Characterization of *Biomphalaria orbignyi*, *Biomphalaria peregrina* and *Biomphalaria oligoza* by polymerase chain reaction and restriction enzyme digestion of the internal transcribed spacer resgion of the RNA ribsomal gene. *Memorias do Instituto Oswaldo Cruz*, 95, 807-814.

SPIELMAN, D., BROOK, B. W., BRISCOE, D. A. & FRANKHAM, R. 2004. Does inbreeding and loss of genetic diversity decrease disease resistance? *Conservation Genetics*, 5, 439-448.

SSEGAWA, P. & NKUUTU, D. N. 2006. Diversity of vascular plants on Ssese islands in Lake Victoria, central Uganda. *African Journal of Ecology*, 44, 22-29.

STADLER, T. & JARNE, P. 1997. Population biology, genetic structure, and mating system parameters in freshwater snails. *In:* STREIT, B., STADLER, T. & LIVELY, C. M. (eds.) *Evolutionary Ecology of Freshwater Animals*. Basel: Birkhauser Verlag.

STAGER, J. & JOHNSON, T. 2008. The late Pleistocene desiccation of Lake Victoria and the origin of its endemic biota. *Hydrobiologia*, 596, 5-16.

STANDLEY, C. J., KABATEREINE, N. B., LANGE, C. N., LWAMBO, N. J. S. & STOTHARD, J. R. 2010. Molecular epidemiology and phylogeography of *Schistosoma mansoni* around Lake Victoria *Parasitology*, 137 (13), 1937-1949.

STANDLEY, C. J., ADRIKO, M., ALINAITWE, M., KAZIBWE, F., KABATEREINE, N. B. & STOTHARD, J. R. 2009. Intestinal schistosomiasis and soil-transmitted helminthiasis in Ugandan schoolchildren: a rapid mapping assessment. *Geospatial Health*, 4, 39-53.

STANDLEY, C. J. & STOTHARD, J. R. 2010. Towards defining appropriate strategies for targeted NTD control. *Tropical Medicine & International Health*, 15, 772-773.

STANDLEY, C. J., POINTIER, J. P., ISSIA, L., WISNIVESKY-COLLI, C. & STOTHARD, J. R. 2011. The identification and characterisation of *Biomphalaria peregrina* (Orbignyi, 1835) from Agua Escondida in northern Patagonia, Argentina. *Journal of Natural History*, 45, 347-356.

STANLEY, H. M. 1876. Letters of Mr. H. M. Stanley on His Journey to Victoria Nyanza, and Circumnavigation of the Lake. *Proceedings of the Royal Geographical Society of London*, 20, 134-159.

STATACORP. 2009. *Stata Statistical Software: Release 11*. StataCorp LP

STAUFFER, J., MADSEN, H., MCKAYE, K., KONINGS, A., BLOCH, P., FERRERI, C., LIKONGWE, J. & MAKAULA, P. 2006. Schistosomiasis in Lake Malawi: Relationship of fish and intermediate host density to prevalence of human infection. *Ecohealth*, 3, 22-27.

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. 2008. Introgressive hybridization of human and rodent schistosome parasites in western Kenya. *Molecular Ecology*, 17, 5062-5074.

STEINAUER, M. L., HANELT, H., 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. *International Journal for Parasitology*, 39, 1353-1362.

STEINAUER, M. L., BLOUIN, M. S. & CRISCIONE, C. D. 2010. Applying evolutionary genetics to schistosome epidemiology. *Infection, Genetics and Evolution,* 10, 433-443.

STENSGAARD, A. S., JØRGENSEN, A., KABATEREINE, N., MALONE, J. B. & KRISTENSEN, T. K. 2005. Modeling the distribution of *Schistosoma mansoni* and host snails in Uganda using satellite sensor data and Geographical Information System. *Parassitologia,* 47, 115-125.

STENSGAARD, A. S., JØRGENSEN, A., KABATEREINE, N. B., RAHBEK, C. & KRISTENSEN, T. K. 2006. Modeling freshwater snail habitat suitability and areas of potential snail-borne disease transmission in Uganda. *Geospatial Health,* 1, 93-104.

STIREWALT, M. A. 1954. Effect of snail maintenance temperatures on development of *Schistosoma mansoni. Experimental Parasitology,* 3, 504-516.

STIREWALT, M. A. 1956. Penetration of host skin by cercariae of *Schistosoma mansoni.* I. Observed entry into skin of mouse, hamster, rat, monkey and man. *The Journal of Parasitology,* 42, 565-580.

STOTHARD, J. R., HUGHES, S. & ROLLINSON, D. 1996. Variation within the Internal Transcribed Spacer (ITS) of ribosomal DNA genes of intermediate snail hosts within the genus *Bulinus* (Gastropoda: Planorbidae). *Acta Tropica,* 61, 19-29.

STOTHARD, J. R. & ROLLINSON, D. 1996. An evaluation of random amplified polymorphic DNA (RAPD) for the identification and phylogeny of freshwater snails of the genus *Bulinus* (Gastropoda: Planorbidae). *Journal of Molluscan Studies,* 62, 165-176.

STOTHARD, J. R. & ROLLINSON, D. 1997. Partial DNA sequences from the mitochondrial cytochrome oxidase subunit I (COI) gene can differentiate the intermediate snail hosts *Bulinus globosus* and *B. nasutus* (Gastropoda: Planorbidae). *Journal of Natural History,* 31, 727-737.

STOTHARD, J. R., MGENI, A. F., KHAMIS, S., SETO, E., RAMSAN, M., HUBBARD, S. J., KRISTENSEN, T. K. & ROLLINSON, D. 2002. New insights into the transmission biology of urinary schistosomiasis in Zanzibar. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 96, 470-475.

STOTHARD, J. R., KABATEREINE, N. B., TUKAHEBWA, E. M., KAZIBWE, F., MATHIESON, W., WEBSTER, J. P. & FENWICK, A. 2005. Field evaluation of the Meade Readiview handheld microscope for diagnosis of intestinal schistosomiasis in Ugandan school children. *American Journal of Tropical Medicine and Hygiene,* 73, 949-955.

STOTHARD, J. R., KABATEREINE, N. B., TUKAHEBWA, E. M., KAZIBWE, F., ROLLINSON, D., MATHIESON, W., WEBSTER, J. P. & FENWICK, A.. 2006a. Use of circulating cathodic antigen (CCA) dipsticks for detection of intestinal and urinary schistosomiasis. *Acta Tropica,* 97, 219-228.

STOTHARD, J. R., MOOK, P., MGENI, A. F., KHAMIS, I. S., KHAMIS, A. N. & ROLLINSON, D. 2006b. Control of urinary schistosomiasis on Zanzibar (Unguja Island): a pilot evaluation of the educational impact of the Juma na Kichocho health booklet within primary schools. *Memorias Do Instituto Oswaldo Cruz,* 101, 119-124.

STOTHARD, J. R. 2009. Improving control of African schistosomiasis: towards effective use of rapid diagnostic tests within an appropriate disease surveillance model. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 103, 325-332.

STOTHARD, J. R., CHITSULO, L., KRISTENSEN, T. K. & UTZINGER, J. 2009a. Control of schistosomiasis in sub-Saharan Africa: progress made, new opportunities and remaining challenges. *Parasitology,* 136, 1665-1675.

STOTHARD, J. R., SOUSA-FIGUEIREDO, J. C., STANDLEY, C., VAN DAM, G. J., KNOPP, S., UTZINGER, J., AMERI, H., KHAMIS, A. N., KHAMIS, I. S., DEELDER, A. M., MOHAMMED, K. A. &

ROLLINSON, D. 2009b. An evaluation of urine-CCA strip test and fingerprick blood SEA-ELISA for detection of urinary schistosomiasis in schoolchildren in Zanzibar. *Acta Tropica,* 111, 64-70.

STOTHARD, J. R., WEBSTER, B. L., WEBER, T., NYAKAANA, S., WEBSTER, J. P., KAZIBWE, F., KABATEREINE, N. B. & ROLLINSON, D. 2009c. Molecular epidemiology of *Schistosoma mansoni* in Uganda: DNA barcoding reveals substantive genetic diversity within Lake Albert and Lake Victoria populations. *Parasitology,* 136, 1-12.

STURROCK, R. F. & BARNISH, G. 1973. The aerial application of molluscicides with special reference to schistosomiasis control. *Bulletin of the World Health Organization,* 49, 283-285.

STURROCK, R. F. 1974. Ecological notes on habitats of the freshwater snail *Biomphalaria glabrata,* intermediate host of *Schistosoma mansoni,* on St. Lucia, West Indies. *Caribbean Journal of Science,* 14, 149-162.

STURROCK, R. F. 2001. Schistosomiasis epidemiology and control: how did we get here and where should we go? *Memorias do Instituto Oswaldo Cruz,* 96 Suppl, 17-27.

SUKWA, T. Y., BULSARA, M. K. & WURAPA, F. K. 1986. The relationship between morbidity and intensity of *Schistosoma mansoni* infection in a rural Zambian community. *International Journal of Epidemiology,* 15, 248-251.

SWOFFORD, D. L. 2003. *PAUP*: phylogenetic analysis using parsimony, version 4.0b10.* Sinauer Associates

TAJIMA, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics,* 123, 585-595.

TAMURA, K., DUDLEY, J., NEI, M. & KUMAR, S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution,* 24, 1596-1599.

TEMPLETON, A. R., CRANDALL, K. A. & SINF, 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.

TEN HOVE, R. J., VERWEIJ, J. J., VEREECKEN, K., POLMAR, K., DIEYE, L. & VAN LIESHOUT, L. 2008. Multiplex real-time PCRfor the detection and quantification of *Schistosoma mansoni* and *S. haematobium* infections in stool samples collected in northern Senegal. *Transactions of the Royal Society for Tropical Medicine and Hygiene.* 102, 179-185.

TESKE, P. R., BARKER, N. P. & MCQUAID, C. D. 2007. Lack of genetic differentiation among four sympatric southeast African intertidal limpets (Siphonariidae): phenotypic plasticity in a single species? *Journal of Molluscan Studies,* 73, 223-228.

THAEWNON-NGIW, B., KLINGBUNGA, S., PHANWICHIEN, K., SANGDUEN, N., LAUHACHINDA, N. & MENASVETA, P. 2003. Genetic diversity of introduced (*Pomacea canaliculata*) and native (*Pila*) apple snails in Thailand revealed by randomly amplified ploymorphic DNA (RAPD) analysis. *ASEAN Journal on Science and Technology for Development,* 20, 289-306.

THÉRON, A. & JOURDANE, J. 1979. Séquence de reconversion des sporocystes de Schistosoma mansoniproducteurs de cercaires en vue de la production de nouvelles générations de sporocystes. *Zeitschrift für Parasitenkunde,* 60, 63-71.

THÉRON, A. 1984. Early and late shedding patterns of *Schistosoma mansoni* cercariae: ecological significance in transmission to human and murine hosts. *The Journal of Parasitology,* 70, 652-655.

THÉRON, A. & TOUASSEM, R. 1989. *Schistosoma rodhaini*: Intramolluscan larval development, migration and replication processes of daughter sporocysts. *Acta Tropica*, 46, 39-45.

THÉRON, 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, G. & MINCHELLA, D. 2007. Genetic diversity and population structuring of *Schistosoma mansoni* in a Brazilian village. *International Journal of Parasitology*, 38, 389-399.

THOMAS, A. 2000. Poverty and the end of development. *In:* ALLEN, T. & THOMAS, A. (eds.) *Poverty and Development.* Oxford: Oxford University Press.

TIAN-BI, Y.-M. T., N'GORAN, E. K., N'GUETTA, S. P., MATTHYS, B., SANGARE, A. & JARNE, P. 2008. Prior selfing and the selfing syndrome in animals: an experimental approach in the freshwater snail *Biomphalaria glabrata*. *Genetics Research*, 90, 61-72.

TURNER, M. G. 1989. Landscape ecology: The effect of pattern on process. *Annual Review of Ecology and Systematics*, 20, 171-197.

UTZINGER, J., MAYOMBANA, C., MEZ, K. & TANNER, M. 1997. Evaluation of chemical and physical-morphological factors as potential determinants of *Biomphalaria pfeifferi* (Krauss, 1848) distribution. *Memorias do Instituto Oswaldo Cruz*, 92, 323-328.

UTZINGER, J., RASO, G., BROOKER, S., DE SAVIGNY, D., TANNER, M., ORNBJERG, N., SINGER, B. H. & N'GORAN E, K. 2009. Schistosomiasis and neglected tropical diseases: towards integrated and sustainable control and a word of caution. *Parasitology*, 136, 1859-1874.

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 Tropica*, 86, 125-139.

VAN GOOL, T., VETTER, H., VERVOORT, T., DOENHOFF, M. J., WETSTEYN, J. & OVERBOSCH, D. 2002. Serodiagnosis of imported schistosomiasis by a combination of a commercial indirect hemagglutination test with *Schistosoma mansoni* adult worm antigens and an enzyme-linked immunosorbent assay with *S. mansoni* egg antigens. *Journal of Clinical Microbiology*, 40, 3432-3437.

VAN SCHAYCK, C. P. 1985. Laboratory studies on the relation between aquatic vegetation and the presence of two bilharzia-bearing snail species. *Journal of Aquatic Plant Management*, 23, 87-91.

VAN VALEN, L. 1973. A new evolutionary law. *Evolutionary Theory*, 1, 1-30.

VAZQUEZ PERERA, A. A., SANCHEZ NODA, J. & HEVIA JIMENEZ, Y. 2010. Distribution and habitat preferences of the genus *Biomphalaria* (Gastropoda: Planorbidae) in Cuba. *Memorias do Instituto Oswaldo Cruz*, 105, 41-44.

VEGA, R., PIÒERO, D., RAMANANKANDRASANA, B., DUMAS, M., BOUTEILLE, B., FLEURY, A., SCIUTTO, E., LARRALDE, C. & FRAGOSO, G. 2003. Population genetic structure of *Taenia solium* from Madagascar and Mexico: implications for clinical profile diversity and immunological technology. *International Journal for Parasitology*, 33, 1479-1485.

VERNON, J. G. 1995. Low reproductive output of isolated, self-fertilizing snails: Inbreeding depression or absence of social facilitation? *Proceedings of the Royal Society B: Biological Sciences*, 259, 131-136.

VERNON, J. G., JONES, C. S. & NOBLE, L. R. 1995. Random amplified polymorphic DNA (RAPD) markers reveal cross-fertilisation in *Biomphalaria glabrata* (Pulmonata: Basommatophora). *Journal of Molluscan Studies*, 61, 455-465.

VERSCHUREN, D., JOHNSON, T. C., KLING, H. J., EDGINGTON, D. N., LEAVITT, P. R., BROWN, E. T., TALBOT, M. R. & HECKY, R. E. 2002. History and timing of human impact on Lake Victoria, East Africa. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 269, 289-294.

VERWEIJ, J. J., PIT, D. S., VAN LIESHOUT, L., BAETA, S. M., DERY, G. D., GASSER, R. B. & POLDERMAN, A. M. 2001. Determining the prevalence of *Oesophagostomum bifurcum* and *Necator americanus* infections using specific PCR amplification of DNA from faecal samples. *Tropical Medicine & International Health*, 6, 726-731.

VERWEIJ, J. J., BRIENEN, E. A. T., ZIEM, J., YELIFERI, L., POLDERMAN, A. M. & VAN LIESHOUT, L. 2007. Simultaneous detection and quantification of *Anclyostoma duodenale, Necator americanus* and *Oesophagostomum bifurcum* in fecal samples using multiplex real-time PCR. *American Journal of Tropical Medicine and Hygiene*, 77, 685-690.

VIDIGAL, T. H., SPATZ, L., KISSINGER, J. C., REDONDO, R. A., PIRES, E. C., SIMPSON, A. J. & CARVALHO, O. S. 2004. Analysis of the first and second internal transcribed spacer sequences of the ribosomal DNA in *Biomphalaria tenagophila* complex (Mollusca: Planorbidae). *Memorias do Instituto Oswaldo Cruz*, 99, 153-158.

VRIJENHOEK, R. & GRAVEN, M. 1992. Population genetics of Egyptian *Biomphalaria alexandrina* (Gastrodpoda: Planorbidae). *Journal of Heredity*, 83, 255-261.

WAGNER, E. D. & MOORE, B. 1959. The development of *Schistosoma mansoni* in snails kept at certain constant temperatures. *Transactions of the American Microscopical Society*, 78, 424-428.

WEBBE, G. & JORDAN, P. 1966. Recent advances in knowledge of schistosomiasis in East Africa. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 60, 279-312.

WEBSTER, B. L., SOUTHGATE, V. R. & LITTLEWOOD, D. T. J. 2006. A revision of the interrelationships of *Schistosoma* including the recently described *Schistosoma guineensis*. *International Journal for Parasitology*, 36, 947-955.

WEBSTER, J. P. & DAVIES, C. M. 2001. Coevolution and compatibility in the snail-schistosome system. *Parasitology*, 123 Supp, S41-S56.

WEBSTER, J. P., DAVIES, C. M., NDAMBA, J., NOBLE, L. R., JONES, C. S. & WOOLHOUSE, M. E. J. 2001a. Spatio-temporal genetic variability in the schistosome intermediate host *Biomphalaria pfeifferi*. *Annals of Tropical Medicine and Parasitology*, 95, 515-527.

WEBSTER, J. P., DAVIES, C. M., HOFFMAN, J. I., NDAMBA, J., NOBEL, L. R. & WOOLHOUSE, M. E. 2001b. Population genetics of the schistosome intermediate host *Biomphalaria pfeifferi* in the Zimbabwean highveld: implications for co-evolutionary theory. *Annals of Tropical Medicine and Parasitology*, 95, 203-214.

WEIR, B. S. & COCKERHAM, C. C. 1984. Estimating F-statistics for the analysis of population structure. *Evolution*, 38, 1358-1370.

WEISS, R. A. & MCMICHAEL, A. J. 2004. Social and environmental risk factors in the emergence of infectious diseases. *Nature Medicine*, 10, S70-S76.

WELLINGTON, G. M. & KURIS, A. M. 1983. Growth and shell variation in the tropical eastern Pacific intertidal gastropod genus *Purpura*: Ecological and evolutionary implications. *The Biological Bulletin*, 164, 518-535.

WETHINGTON, A. R., ZAVODNA, M., SMITH, M. K., OLIVEIRA, G., LEWIS, F. & MINCHELLA, D. J. 2007. Population genetic structure of *Biomphalaria glabrata* in a schistosomiasis-endemic region in Brazil. *Journal of Molluscan Studies,* 73, 45-52.

WHO 2002. Prevention and control of schistosomiasis and soil-transmitted helminthiasis - report of a WHO expert committee. *In:* REPORT, W. T. S. (ed.) *WHO Technical Report Series* Geneva: WHO.

WITTE, F., GOLDSCHMIDT, T., GOUDSWAARD, P. C., LIGTVOET, W., VAN OIJEN, M. J. P. & WANINK, J. H. 1991. Species extinction and concomitant ecological changes in Lake Victoria. *Netherlands Journal of Zoology,* 42, 214-232.

WITTE, F., MSUKU, B. S., WANINK, J. H., SEEHAUSEN, O., KATUNZI, E. F. B., GOUDSWAARD, P. C. & GOLDSCHMIDT, T. 2000. Recovery of cichlid species in Lake Victoria: an examination of factors leading to differential extinction. *Reviews in Fish Biology and Fisheries,* 10, 233-241.

WOLLSCHEID-LENGELING, E., BOORE, J., BROWN, W. & W‰GELE, H. 2001. The phylogeny of Nudibranchia (Opisthobranchia, Gastropoda, Mollusca) reconstructed by three molecular markers. *Organisms Diversity & Evolution,* 1, 241-256.

WOODRUFF, D. S., MULVEY, M. & YIPP, M. W. 1985a. The continued introduction of intermediate host snails of *Schistosoma mansoni* into Hong Kong. *Bulletin of the World Health Organization,* 63, 621-622.

WOODRUFF, D. S., MULVEY, M. & YIPP, M. W. 1985b. Population genetics of *Biomphalaria straminea* in Hong Kong - a neotropical schistosome-transmitting snail recently introduced into China. *Journal of Heredity,* 76, 355-360.

WOODRUFF, D. S. & MULVEY, M. 1997. Neotropical schistosomiasis: African affinities of the host snail *Biomphalaria glabrata* (Gastropoda: Planorbidae). *Biological Journal of the Linnean Society,* 60, 505-516.

WOOLHOUSE, M. E. J. & CHANDIWANA, S. K. 1990. Population dynamics model for *Bulinus globosus,* intermediate host for *Schistosoma haematobium,* in river habitats. *Acta Tropica,* 47, 151-160.

WRIGHT, C. A., ROLLINSON, D. & GOLL, P. H. 1979. Parasites in *Bulinus senegalensis* (Mollusca: Planorbidae) and their detection. *Parasitology,* 79, 95-105.

XU, Y. & DRESDEN, M. H. 1986. Leucine aminopeptidase and hatching of *Schistosoma mansoni* eggs. *The Journal of Parasitology,* 72, 507-511.

XU, Y. & DRESDEN, M. H. 1990. The hatching of schistosome eggs. *Experimental Parasitology,* 70, 236-240.

XU, B., GONG, P., BIGING, G., LIANG, S., SETO, E. & SPEAR, R. 2004. Snail density prediction for schistosomiasis control using Ikonos and ASTER images. *Photogrammetric Engineering & Remote Sensing.* 70, 1285-1294.

YU, J. M., DE VLAS, S. J., JIANG, Q. W. & GRYSSELS, B. 2007. Comparison of the Kato-Katz technique, hatching test and indirect hemagglutination assay (IHA) for the diagnosis of *Schistosoma japonicum* infection in China. *Parasitology International,* 56, 45-49.

ZAHOOR, Z., DAVIES, A., KIRK, R., ROLLINSON, D. & WALKER, A. 2009. Nitric oxide production by *Biomphalaria glabrata* haemocytes: effects of *Schistosoma mansoni* ESPs and regulation through the extracellular signal-regulated kinase pathway. *Parasites & Vectors,* 2, 18.

ZHANG, G., VERNEAU, O., QIU, C., JORDANE, J. & XIA, M. 2001. Africa or Asia, which is the evolutionary origin of human schistosomes. *Comptes rendus de l'Académie des sciences. Série III,* 324, 1001-1010.

ZHANG, Y., KOUKOUNARI, A., KABATEREINE, N. B., FLEMING, F. M., KAZIBWE, F., TUKAHEBWA, E. M., STOTHARD, J. R., WEBSTER, J. P.. & FENWICK, A. 2007. Parasitological impact of 2-year preventative chemotherapy on schistosomiasis and soil-transmitted helminthiasis in Uganda. *BMC Medicine*, 5, e27.

ZHAO, Q. P., JIANG, M. S., LITTLEWOOD, D. T. J. & NIE, P. 2010. Distinct genetic diversity of *Oncomelania hupensis*, intermediate host of *Schistosoma japonicum* in mainland China as revealed by ITS sequences. *PLoS Neglected Tropical Diseases*, 4, e611.

13 Appendix

13.1 TE buffer recipe

- 10 mM tris(hydroxymethyl)aminomethane (Tris)

- Solution brought to desired pH with hydrochloric acid (HCl)

- 1 mM ethylenediaminetetraacetic acid (EDTA)

13.2 Lake Victoria water chemistry and environmental variables

Turbulence	Abund	Depth
Low = 1	no snails = 1	Shallow = 1
Mod = 2	<10 = 2	Shallow/Med = 2
High = 3	10-30 = 3	Medium = 3
	> 30 = 4	Medium/Deep = 4
		Deep = 5

Habitat	Substrate	Other spp
Lake = 1	Weeds/reeds = 1	Number of other species found
Marsh = 2	Organic debris/mud = 2	
Paddy = 3	Sand = 3	
Pond = 4	Rock = 4	
Ditch/Canal = 5	Combination incl. vegetation = 5	
Lake/marsh = 6	Combination excl. vegetation = 6	
Other combo = 7		
Other = 8		

Table 13.1 – Table displaying the codes used for classifying environmental data

All of the above variables were therefore analysed as categorical variables, apart from 'Other spp.' (number of other species) which was a continuous variable.

Variable	Minimum	Median	Maximum	Mean	Missing values
pH	6.5	8.2	10.1	8.187	6
Flouride	0.039	0.300	2.260	0.393	30
Chloride	1.820	7.730	194.000	13.320	30
Nitrate	0.000	0.000	34.400	1.461	30
Phosphate	0.000	0.122	6.860	0.465	30
Sulphate	0.000	0.960	121.500	4.302	30
Sodium	2.740	11.360	280.000	20.170	31
Potassium	1.590	7.220	60.850	10.999	31
Magnesium	0.604	2.790	42.900	3.972	31
Calcium	2.340	7.095	53.900	12.867	31
Bromide	0.000	0.000	0.460	0.020	77
Nitrite	0.000	0.000	38.300	0.517	77
Conductivity	24.80	120.2	2054	223.000	6
TDS	15.0	72.1	689	123.100	8
Salinity	0.3	0.4	0.7	0.413	28
Temperature	21.7	27.2	35.5	27.170	8
Other species	0	2	6	2.512	14

Table 13.2 – Summary table of all water chemistry and environmental measurements for Lake Victoria

'TDS' stands for total dissolved solids. All values are given to the number of decimal places to which they were measured originally.

Factor	Category	Number of sites
Turbulence	Low	114
	Medium	29
	High	30
	Missing	50
Habitat	Lake	128
	Marsh	54
	Other*	40
	Missing	1
Depth	Shallow	124
	Moderate	33
	Deep	49
	Missing	17
Substrate	Weeds/reeds	17
	Mud/organic debris	52
	Sand	23
	Rock	19
	Combination incl. vegetation	57
	Combination excl. vegetation	10
	Missing	45

Table 13.3 – Summary table for categorical environmental measurements for Lake Victoria

* Breakdown as follows: rice paddy (4), pond (9), ditch/canal (1), lake/marsh combo (16), other combo (8), other/unsure (2)

Site	Longitude°	Latitude°	pH	F	Cl	N03	P04	S04	Na	NH4	K	Mg	Ca	Bromide
T001	32.89392	-2.71347	8.8	0.54	10.2	0.10	0.00	0.92	NA	NA	NA	NA	NA	0.00
T002A	32.86958	-2.72572	7.4	0.58	9.69	1.17	0.73	1.50	19.2	NA	8.42	3.58	11.5	0.00
T002B	32.86958	-2.72572	9.0	0.58	6.75	0.43	0.18	0.90	16.4	NA	4.30	2.96	5.07	0.00
T003	32.95962	-2.64253	7.9	1.17	49.4	0.35	1.28	5.02	134	NA	13.2	5.43	9.12	0.22
T004	32.90055	-2.53082	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
T005	32.94063	-2.41402	8.0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
T006a	32.94998	-2.40463	7.3	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
T006b	32.94998	-2.40463	7.7	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
T007	32.88012	-2.50150	9.4	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
T008	33.38995	-2.58458	9.2	0.93	26.3	0.33	2.66	5.04	41.1	NA	39.4	3.92	27.5	0.13
T009A	33.39530	-2.52582	7.4	0.35	14.5	0.00	1.38	2.13	40.4	NA	13.5	8.93	45.7	0.00
T009B	33.39530	-2.52582	8.3	0.43	6.99	0.00	0.48	1.88	11.1	NA	5.06	2.45	5.95	0.00
T010	33.41522	-2.58758	7.6	1.15	16.2	0.00	0.00	0.55	49.7	NA	4.85	3.65	27.5	0.12
T011	33.51693	-2.45220	7.7	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
T012	33.80880	-2.26215	7.2	0.30	12.0	0.17	0.54	2.17	62.9	NA	9.47	6.99	30.0	0.00
T013	33.55147	-2.1676	8.9	0.25	4.00	0.00	0.00	0.72	8.78	NA	3.79	1.53	4.21	0.00
T014A	33.21043	-2.11240	7.0	0.27	18.1	1.70	1.60	5.07	19.4	NA	17.4	4.57	30.6	0.00
T014B	33.21043	-2.11240	7.5	0.42	9.79	0.00	1.13	2.64	12.8	NA	6.77	2.64	7.19	0.00
T015	33.04788	-2.12818	7.2	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
T016	33.07085	-2.11720	7.3	0.32	17.6	0.00	0.71	4.36	23.3	NA	4.94	4.19	45.3	0.00
T017	33.06497	-2.12400	7.9	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
T018	33.10333	-2.01845	7.1	0.29	8.24	0.00	0.33	0.19	13.5	NA	2.40	3.28	30.5	0.00
T019	33.11142	-1.99552	7.2	0.42	23.9	0.28	6.86	15.437	21.3	NA	21.7	5.76	22.7	0.00
T020	33.01798	-1.98392	7.0	0.11	10.6	0.00	0.34	3.95	8.47	NA	6.73	1.02	12.9	0.00
T021	32.86108	-1.94297	6.5	0.09	2.26	0.14	0.00	0.17	2.74	NA	2.16	0.604	7.04	0.00
T022	33.31210	-2.04797	7.3	0.37	5.19	0.00	0.17	1.30	10.1	NA	3.20	2.51	6.01	0.00
T023	33.32848	-2.13092	6.8	0.19	1.82	0.00	0.96	0.15	10.9	NA	1.59	2.21	9.96	0.00
T024	33.47742	-2.15658	7.1	0.45	96.1	0.00	2.52	80.7	92.9	NA	6.67	7.19	28.7	0.23
T025	33.74203	-2.07973	7.1	0.46	21.6	0.57	0.78	5.80	22.2	NA	12.6	9.12	42.0	0.00
T026A	32.75490	-2.53505	7.3	0.40	17.8	0.00	1.98	3.04	33.1	NA	18.8	5.94	36.3	0.00
T026B	32.75490	-2.53505	7.7	0.33	5.50	0.00	0.22	1.15	10.5	NA	3.76	2.49	6.33	0.00
T027A	32.54162	-2.54550	7.1	0.35	6.28	0.16	0.00	1.07	15.4	NA	4.00	3.01	11.4	0.00

Site	Longitude°	Latitude°	pH	F	Cl	N03	P04	S04	Na	NH4	K	Mg	Ca	Bromide
T027B	32.54162	-2.54550	8.7	0.41	7.82	0.00	0.23	1.48	10.9	NA	4.25	2.61	4.80	0.00
T028	32.23285	-2.53780	7.6	0.25	7.00	0.21	0.32	0.55	9.22	NA	5.57	0.956	4.24	0.00
T029	32.01462	-2.50673	7.4	0.28	6.32	0.00	0.28	1.45	11.2	NA	5.32	2.54	5.80	0.00
T030	31.98610	-2.49390	7.7	0.35	6.06	0.16	0.00	1.32	10.8	NA	3.81	2.54	5.82	0.00
T031A	31.98360	-2.46088	6.5	0.20	7.61	0.00	2.00	0.18	14.4	NA	12.7	6.20	27.7	0.00
T031B	31.98360	-2.46088	8.7	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
T032	32.00978	-2.43967	8.7	0.38	4.92	0.00	0.63	1.02	9.77	NA	3.31	2.42	4.62	0.00
T033A	32.05935	-2.40525	6.5	0.20	23.6	0.63	4.65	8.67	21.0	NA	22.9	5.21	20.4	0.00
T033B	32.05935	-2.40525	6.9	0.28	7.29	0.09	0.50	0.95	10.8	NA	4.58	2.40	5.71	0.00
T034A	32.04040	-2.34837	6.7	0.14	12.0	0.00	0.17	0.34	17.7	NA	10.2	4.38	39.3	0.00
T034B	32.04040	-2.34837	7.2	0.27	5.82	0.00	1.19	1.17	10.2	NA	7.54	2.29	4.73	0.00
T035	31.96692	-2.38317	8.5	0.24	6.43	0.00	0.25	0.49	10.7	NA	5.16	2.19	5.69	0.00
T036A	31.94512	-2.40662	7.2	0.15	11.5	0.35	0.22	9.28	19.4	NA	6.46	8.89	42.7	0.00
T036B	31.94512	-2.40662	7.3	0.37	7.61	0.39	0.25	1.85	10.7	NA	4.85	2.41	6.24	0.00
T037A	31.92260	-2.41502	6.8	0.28	5.90	0.10	0.19	1.39	10.7	NA	4.26	1.92	8.52	0.00
T037B	31.92260	-2.41502	8.1	0.29	3.70	0.00	0.17	0.43	9.69	NA	3.52	2.50	4.96	0.00
T038a	32.41142	-2.43602	7.3	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
T038b	32.41142	-2.43602	9.3	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
T039	32.84040	-2.53903	8.4	0.45	7.51	0.27	0.00	1.34	12.3	NA	4.62	2.69	6.82	0.00
T040	32.89540	-2.52758	7.3	0.59	29.5	0.25	0.61	10.4	27.8	NA	14.6	7.43	27.4	0.00
T041	33.52967	-1.96210	7.7	0.73	19.5	0.000	1.35	4.15	30.6	NA	32.2	7.88	19.5	0.000
T042a	33.46734	-1.95587	7.9	0.41	9.95	19.7	0.00	2.00	16.6	NA	10.3	2.98	11.9	0.000
T042b	33.46734	-1.95587	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
T043a	33.38072	-2.04744	7.7	0.34	15.1	17.7	0.20	4.26	15.5	NA	44.9	6.64	23.1	0.000
T042b	33.38072	-2.04744	8.2	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
T043c	33.38072	-2.04744	7.6	0.41	14.6	0.000	0.00	0.06	17.4	NA	17.7	7.89	51.4	0.000
T044a	33.38573	-2.01536	8.2	0.42	11.2	0.000	0.00	0.000	13.7	NA	12.7	3.09	51.5	0.000
T044b	33.38573	-2.01536	8.8	0.30	5.01	0.000	0.00	0.74	11.2	NA	5.20	2.61	7.04	0.000
T045	33.43285	-1.98415	8.1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
T046	33.47432	-2.01133	7.3	0.47	15.5	0.000	0.00	0.84	17.2	NA	21.9	2.95	7.17	0.000
T047a	33.41163	-1.80884	7.9	0.15	27.6	4.37	5.80	8.55	20.5	NA	46.9	5.22	16.4	0.000
T047b	33.41163	-1.80884	8.2	0.28	4.87	0.000	0.00	0.89	11.2	NA	5.70	2.53	5.30	0.000

Site	Longitude°	Latitude°	pH	F	Cl	N03	P04	S04	Na	NH4	K	Mg	Ca	Bromide
T048a	33.39669	-1.90757	7.7	0.42	9.32	6.31	0.18	50.5	55.1	NA	9.10	4.00	14.7	0.000
T048b	33.39669	-1.90757	9.0	0.38	7.73	0.000	0.12	0.15	11.6	NA	7.56	2.57	5.63	0.000
T049	33.46578	-1.84695	8.6	0.29	6.38	0.000	0.00	0.92	10.7	NA	6.61	2.48	5.11	0.000
T050a	33.62188	-1.78221	7.7	0.52	10.3	2.50	0.00	0.000	14.6	NA	14.1	6.48	42.6	0.000
T050b	33.62188	-1.78221	8.9	0.30	3.95	0.000	0.00	0.09	10.2	NA	4.11	2.43	4.61	0.000
T051	33.54084	-1.67999	9.1	0.28	5.95	0.000	0.00	0.18	10.1	NA	6.35	2.49	5.08	0.000
T052	33.61805	-1.67688	9.2	0.28	6.99	0.000	0.00	0.76	10.1	NA	6.84	2.50	2.42	0.000
T053a	33.68742	-1.68303	7.5	0.40	10.0	0.000	0.00	2.20	14.9	NA	8.28	2.37	14.0	0.000
T053b	33.68742	-1.68303	9.1	0.30	3.93	0.000	0.00	0.77	10.3	NA	3.79	2.43	5.45	0.000
T054	33.69483	-1.60456	8.7	0.32	6.91	0.28	0.00	0.83	11.1	NA	7.04	2.39	6.59	0.000
T055	33.68313	-1.54812	9.1	0.30	5.83	0.000	0.00	0.84	10.2	NA	5.98	2.45	5.34	0.000
T056	33.73884	-1.49598	9.1	0.30	5.97	0.000	0.00	0.22	10.4	NA	6.16	2.48	5.38	0.000
T057	33.8129	-1.32995	8.1	0.30	6.88	1.04	1.58	1.58	11.0	NA	5.02	3.24	4.58	0.000
T058	33.83332	-1.39163	8.8	0.26	9.32	0.20	0.00	0.67	11.3	NA	10.0	2.59	5.29	0.000
T059	33.85619	-1.45393	9.1	0.39	6.35	0.09	0.00	1.00	13.0	NA	6.02	2.52	6.65	0.000
T060a	33.89542	-1.49796	7.8	0.57	8.84	1.02	0.00	1.35	16.2	NA	8.97	2.22	10.6	0.000
T060b	33.89542	-1.49796	8.4	0.44	6.10	0.000	0.00	0.96	13.3	NA	5.11	2.41	6.19	0.000
T061	33.91271	-1.59873	8.5	0.95	28.0	3.17	0.00	6.18	33.5	NA	11.7	6.11	19.6	0.16
T062	33.82102	-1.51636	7.3	0.34	11.8	19.3	0.12	7.12	18.1	NA	8.65	2.34	2.59	0.000
T063	33.83153	-1.52643	7.5	0.79	23.4	34.4	2.93	34.2	40.8	NA	12.5	5.77	21.7	0.08
T064a	33.97015	-1.34673	7.6	0.50	25.4	6.46	0.71	7.85	21.2	NA	23.7	5.58	42.6	0.000
T064b	33.97015	-1.34673	8.8	0.33	4.11	25.0	0.70	0.96	10.0	NA	4.69	2.62	5.81	0.000
T065	34.13436	-1.40119	8.3	1.21	7.55	3.04	0.11	4.07	23.0	NA	10.1	9.35	39.2	0.000
T066a	33.95499	-1.30631	7.7	0.65	21.8	0.000	1.80	7.50	36.5	NA	24.9	5.48	51.2	0.000
T066b	33.95499	-1.30631	8.8	0.31	5.34	0.000	0.12	0.80	10.2	NA	5.39	2.48	5.71	0.000
T067a	33.99894	-1.12465	7.9	0.67	43.4	4.24	0.00	121.5	77.6	NA	21.9	13.8	37.4	0.12
T067b	33.99894	-1.12465	9.9	0.33	4.34	3.83	2.00	0.96	10.3	NA	6.23	2.76	6.10	0.000
T068a	34.0849	-1.03787	7.8	1.36	18.7	0.000	0.12	0.000	72.1	NA	7.22	8.19	47.6	0.000
T068b	34.0849	-1.03787	9.9	0.28	5.42	9.43	0.00	0.91	9.55	NA	5.17	2.56	2.51	0.000
T069	33.94331	-1.19345	9.2	0.28	4.75	0.000	0.00	0.15	10.2	NA	4.97	2.44	5.25	0.000
T070a	33.86792	-1.2545	7.8	0.48	5.94	0.08	0.00	2.14	11.4	NA	4.91	3.10	39.1	0.000
T070b	33.86792	-1.2545	8.8	0.28	3.65	0.000	0.00	0.82	10.1	NA	3.63	2.49	2.34	0.000

Site	Longitude°	Latitude°	pH	F	Cl	N03	P04	S04	Na	NH4	K	Mg	Ca	Bromide
U001	32.44847	0.078367	9.7	0.23	2.39	0.17	0.05	0.13	10.39	0.00	2.76	2.66	6.56	NA
U002	32.57508	-0.2338	9.3	0.25	2.60	0.00	0.09	0.41	10.47	0.00	2.67	2.88	7.12	NA
U003	32.57153	-0.35158	9.3	0.24	2.55	0.00	0.17	0.39	10.80	0.00	2.84	2.84	6.68	NA
U004	32.5755	-0.32025	9.4	0.23	3.71	0.00	0.10	0.38	10.63	0.00	4.08	2.81	7.17	NA
U005	32.29488	-0.36353	9.3	0.71	8.00	0.00	0.00	0.40	10.99	0.00	8.50	2.75	7.18	NA
U006	32.33168	-0.33442	9.4	0.22	2.56	0.00	0.00	0.37	10.39	0.00	2.70	2.87	6.82	NA
U007	32.30893	-0.32462	9.2	0.23	2.55	0.00	0.00	0.36	10.16	0.00	2.59	2.72	7.18	NA
U008	32.15842	-0.51085	7.5	0.20	13.18	0.00	0.00	0.48	9.17	0.00	14.98	2.42	5.70	NA
U009	32.29235	-0.31022	9.4	0.20	11.20	0.00	0.00	0.41	11.53	0.00	13.07	2.95	6.96	NA
U010	32.19448	-0.32447	8.4	0.29	21.02	0.00	0.00	0.59	11.97	0.00	24.27	3.04	7.36	NA
U011	32.06785	-0.24753	8.3	0.24	14.19	0.00	0.00	0.47	11.35	0.00	17.76	2.84	7.32	NA
U012	32.02652	-0.27327	6.7	0.22	19.79	0.00	0.00	0.14	11.15	0.00	25.65	2.81	7.63	NA
U013	32.03525	-0.30077	8.3	0.24	7.38	0.00	0.00	0.52	10.95	0.00	8.69	2.93	6.85	NA
U014	32.43157	0.004167	10.1	0.27	9.84	0.00	0.00	0.27	10.95	0.00	11.60	2.94	7.59	NA
U015	32.38817	0.015467	9.5	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
U016	32.38142	0.016383	9.3	0.21	22.08	0.00	0.00	0.20	11.49	0.00	35.50	3.00	6.96	NA
U017	32.43163	-0.00908	9.5	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
U018	32.4863	0.07265	7.7	0.23	15.03	0.02	0.00	0.38	11.55	0.00	17.85	2.90	8.28	NA
U019	31.7673	-0.91535	8.3	0.21	15.06	0.04	0.00	1.05	11.86	0.55	17.64	3.27	10.35	NA
U020	31.76323	-0.9391	6.5	0.09	21.27	0.19	0.00	1.73	5.69	0.00	22.29	3.54	6.66	NA
U021	31.79747	-0.65482	6.5	0.04	36.19	0.00	1.02	2.29	10.54	9.54	60.85	2.16	10.96	NA
U022	31.87958	-0.34838	7.7	0.05	9.05	0.29	0.00	0.16	3.10	0.18	10.62	0.87	2.41	NA
U023a	32.76717	0.014767	9.8	0.40	11.56	0.00	0.00	0.32	11.19	0.00	13.78	2.82	7.27	NA
U023b	32.76717	0.014767	7.5	0.24	8.55	0.00	0.32	0.65	12.99	1.34	11.56	3.06	8.84	NA
U024a	32.76352	-0.04157	7.2	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
U024b	32.76352	-0.04157	7.0	0.23	5.88	0.00	0.42	0.49	10.92	1.26	7.17	3.51	10.09	NA
U025a	32.90127	0.002017	10.0	0.23	6.31	0.00	1.32	0.79	13.17	3.14	15.24	3.17	7.07	NA
U025b	32.90127	0.002017	7.9	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
U026	32.764	-0.10997	8.4	0.26	5.83	0.03	0.16	0.62	10.76	0.46	6.35	2.70	6.74	NA
U027	32.65303	-0.09987	8.8	0.25	7.19	0.06	0.14	0.47	10.19	0.77	9.19	2.63	6.80	NA
U028	32.6522	-0.08638	7.7	0.22	7.15	0.00	0.29	0.53	11.36	1.33	9.31	3.17	7.82	NA
U029	33.60233	0.140983	8.5	0.32	37.42	0.00	0.22	0.38	10.66	0.00	47.31	2.90	6.96	NA

Site	Longitude°	Latitude°	pH	F	Cl	N03	P04	S04	Na	NH4	K	Mg	Ca	Bromide
U030b	33.60215	0.111767	9.0	0.23	10.88	0.00	0.20	0.39	10.70	0.00	46.86	2.77	6.85	NA
U031	33.60583	0.05235	8.0	0.24	19.26	0.33	0.08	1.59	10.95	0.88	22.61	3.44	9.58	NA
U032	33.64465	0.031083	7.2	0.26	10.81	0.00	1.76	1.76	12.10	4.62	13.06	5.39	18.52	NA
U033	33.65085	0.03285	7.1	0.26	17.29	0.03	1.37	1.87	11.46	4.20	18.68	4.68	16.04	NA
U034	33.6586	0.003033	6.9	0.20	9.48	0.08	0.58	0.54	11.04	3.84	18.04	2.88	6.65	NA
U035b	33.56602	0.156183	7.6	0.24	11.37	0.00	0.43	0.83	11.27	1.65	14.93	3.13	7.61	NA
U036	33.56192	0.172867	8.5	0.24	5.74	1.14	0.21	0.48	10.62	1.22	7.65	3.00	8.93	NA
U037	33.62703	0.318117	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
U038a	33.98475	0.263167	7.8	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
U038b	33.98475	0.263167	7.3	0.17	3.47	0.01	0.00	1.21	18.54	1.66	4.91	5.04	36.21	NA
U039	33.98893	0.253067	7.4	0.24	12.59	0.07	0.14	2.07	14.74	1.85	20.12	4.62	18.56	NA
U040	33.99178	0.241167	8.9	0.24	16.20	0.01	0.00	0.41	12.01	0.72	21.84	3.26	9.01	NA
U041	33.8912	0.5354	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
U042	33.671	0.6104	7.3	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
U043	33.28143	0.47555	7.5	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
U044	33.24102	0.437517	6.9	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
U045	32.6549	0.289967	8.7	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
U046	33.18397	0.1732	7.2	0.20	22.07	0.11	0.00	0.69	12.33	0.00	39.15	3.81	8.61	NA
U047	33.21522	0.185583	7.5	0.26	10.83	0.03	0.66	1.49	15.16	8.33	14.05	4.69	20.43	NA
U048b	33.26475	0.197683	7.1	0.24	8.98	0.13	0.52	0.90	11.72	3.74	9.86	3.32	7.81	NA
U049	33.2429	0.234283	7.2	0.25	9.92	0.20	0.16	0.88	10.98	2.20	11.76	3.06	7.81	NA
U050	33.21898	0.246767	7.1	0.23	5.59	0.20	0.00	0.94	11.41	2.20	7.97	2.70	7.22	NA
U051	33.20588	0.270217	8.8	0.24	5.25	0.12	0.00	0.48	10.34	0.52	6.22	2.81	6.77	NA
U052	33.15325	0.271317	7.5	0.23	4.61	0.16	0.61	0.92	10.49	2.83	4.63	3.35	7.68	NA
U053	33.13708	0.2403	7.0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
U054	NA	NA	7.4	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
U055	32.68385	-0.092	8.9	0.29	6.63	0.00	0.21	0.75	9.63	NA	6.54	2.18	4.55	0.00
UW003	31.79748	-0.65477	9.6	0.21	3.90	0.00	0.09	0.50	9.79	0.49	5.58	2.75	6.58	NA
UW001	32.41853	0.01555	9.7	0.22	12.88	0.00	0.00	0.28	10.57	0.00	18.24	2.94	6.75	NA
K001A	34.05849	-0.07313	7.9	0.63	178	10.6	0.34	13.7	183	NA	51.5	31.9	23.3	0.38
K001B	34.05849	-0.07313	9.8	0.30	9.59	0.00	0.00	0.98	10.2	NA	8.52	2.34	5.06	0.00
K002A	34.06501	-0.10968	8.5	0.67	194	0.00	3.77	11.9	280	NA	40.3	42.9	53.9	0.46

Site	Longitude°	Latitude°	pH	F	Cl	N03	P04	S04	Na	NH4	K	Mg	Ca	Bromide
K002B	34.06501	-0.10968	9.3	0.29	5.88	0.00	0.00	0.88	10.2	NA	4.83	2.33	6.02	0.00
K003A	34.38747	-0.18737	NA	0.57	10.0	13.4	0.61	4.71	18.8	NA	7.85	3.57	8.06	0.00
K003B	34.38747	-0.18737	8.3	0.52	8.93	1.22	0.27	2.95	16.0	NA	7.85	2.85	9.85	0.00
K004	34.59429	-0.14123	8.2	0.51	11.8	0.56	0.30	2.69	15.3	NA	11.0	2.39	5.91	0.00
K005	34.20749	-0.42048	8.1	0.41	24.3	0.39	3.41	3.93	18.9	NA	16.8	3.83	6.61	0.00
K006A	34.17073	-0.43388	7.6	0.42	15.3	2.69	0.15	1.48	21.5	NA	11.6	3.89	17.0	0.00
K006B	34.17073	-0.43388	8.4	0.31	5.90	0.11	0.00	1.02	10.0	NA	5.40	2.45	6.98	0.00
K007A	34.12877	-0.43074	8.7	0.86	21.4	0.10	0.79	7.20	73.3	NA	8.57	3.77	8.93	0.00
K007B	34.12877	-0.43074	8.8	0.26	9.50	0.09	0.30	1.17	14.9	NA	6.17	3.07	6.10	0.00
K008A	34.15962	-0.39753	8.6	1.52	48.7	0.00	0.67	9.39	158	NA	18.6	23.4	24.2	0.13
K008B	34.15962	-0.39753	9.2	0.37	6.43	0.00	0.00	1.49	15.6	NA	4.81	2.93	6.36	0.00
K009	34.13857	-0.42781	8.5	0.26	3.32	0.00	0.00	0.93	9.93	NA	3.37	2.35	6.41	0.00
K010	34.21333	-0.38058	8.8	0.31	6.77	0.00	0.00	1.13	11.3	NA	5.64	2.40	6.23	0.00
K011	34.32315	-0.45251	8.0	0.42	9.17	0.52	0.57	4.37	14.8	NA	7.26	2.42	5.85	0.00
K012	34.28831	-0.47401	7.8	0.37	4.87	0.00	0.00	1.68	12.3	NA	4.18	2.45	7.52	0.00
K013A	34.12972	-1.0097	8.1	0.29	3.25	0.00	0.41	0.66	6.83	NA	2.50	1.83	6.04	0.00
K013B	34.12972	-1.0097	8.8	0.31	7.49	0.00	0.75	0.95	9.95	NA	5.60	2.71	5.08	0.00
K014	34.09884	-1.0013	9.7	0.35	9.42	0.11	0.00	1.78	13.2	NA	7.41	2.51	6.74	0.00
K015A	34.18728	-0.85562	7.7	0.34	44.8	0.08	0.17	3.69	28.8	NA	11.8	8.02	29.2	0.23
K015B	34.18728	-0.85562	8.6	0.31	7.49	1.04	0.00	1.21	10.7	NA	6.06	2.34	4.84	0.00
K016B	34.11915	-0.81734	8.4	0.30	4.05	0.14	0.14	0.78	9.03	NA	3.21	2.13	5.07	0.00
K017	34.05849	-0.72602	8.4	0.32	27.9	0.16	1.13	2.79	34.2	NA	7.22	3.32	6.07	0.11
K018B	34.16403	-0.53805	8.8	0.37	8.16	0.00	0.00	2.71	14.4	NA	4.34	2.51	8.27	0.00
K019A	34.01513	-0.43738	8.0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
K019B	34.01513	-0.43738	8.5	0.33	5.71	0.06	0.18	1.39	11.6	NA	4.35	2.54	3.48	0.00
K020A	34.45482	-0.52292	8.1	0.70	21.9	0.37	0.34	19.2	43.1	NA	8.14	6.53	19.1	0.00
K020B	34.45482	-0.52292	8.2	0.71	35.5	23.9	1.31	81.3	75.4	NA	17.2	9.23	28.3	0.12
K021A	34.66333	-0.35414	8.3	0.60	9.29	13.6	0.29	6.94	19.2	NA	6.21	2.59	9.41	0.00
K021B	34.66333	-0.35414	8.7	0.50	6.56	0.00	0.00	2.70	16.3	NA	5.39	2.41	7.49	0.00
K022	34.84837	-0.31209	8.0	2.26	25.6	0.95	0.58	44.3	73.4	NA	15.3	7.79	40.7	0.27
K023	34.96699	-0.22569	7.0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
K024	34.93551	-0.1751	7.5	1.63	9.36	0.13	0.18	2.74	13.2	NA	14.0	2.90	16.9	0.00

Site	Longitude°	Latitude°	pH	F	Cl	N03	P04	S04	Na	NH4	K	Mg	Ca	Bromide
K025	34.90741	-0.17148	6.9	0.71	4.34	0.38	0.14	6.23	21.8	NA	7.49	6.93	21.4	0.12
K026	34.84986	-0.15562	6.6	1.47	7.53	0.44	0.00	4.90	18.0	NA	29.8	7.80	31.6	0.00
K027	34.68494	-0.08456	7.9	0.67	9.96	0.00	1.23	6.57	14.3	NA	13.8	2.82	9.61	0.00
K028	34.71829	-0.10519	7.9	0.56	6.23	0.29	0.12	3.74	16.0	NA	7.34	2.39	8.20	0.00
K029	34.74907	-0.09589	7.5	0.51	33.3	0.13	0.14	31.5	31.2	NA	14.6	6.14	34.3	0.00
K030A	34.26699	-0.32841	8.2	0.60	26.9	3.64	0.78	6.38	35.3	NA	17.5	13.8	30.7	0.00
K030B	34.26699	-0.32841	9.3	0.35	6.87	0.16	3.28	3.43	12.5	NA	5.08	3.50	6.43	0.00
K031	34.20963	-0.2394	9.2	0.30	5.55	0.21	0.00	1.50	11.3	NA	4.44	2.28	5.81	0.00
K032	34.26283	-0.18135	8.1	1.42	4.85	12.8	0.00	2.90	17.2	NA	17.7	2.72	11.9	0.00
K033A	33.96756	0.09862	8.5	0.32	4.97	2.73	0.23	3.08	11.6	NA	4.97	2.65	17.8	0.00
K033B	33.96756	0.09862	8.7	0.26	5.78	0.64	0.00	1.71	11.6	NA	4.58	2.60	7.15	0.00
U023.08	32.76717	0.014767	8.8	0.30	8.71	5.12	0.00	1.39	9.95	NA	7.21	2.15	4.91	0.00
U024B.08	32.76352	-0.04157	8.4	0.29	7.21	0.00	0.00	0.77	9.26	NA	6.85	2.18	4.55	0.00
U027.08	32.65303	-0.09987	8.7	0.28	5.04	0.00	0.00	0.74	9.44	NA	4.54	2.16	4.39	0.00
U028.08	32.6522	-0.08638	8.8	0.24	5.13	0.11	0.00	0.88	9.83	NA	4.70	2.30	4.44	0.00
Water 1 08	32.64527	0.009167	9	0.27	15.0	0.00	0.00	0.78	9.41	NA	13.9	2.18	4.48	0.00
S001	32.29242	-0.36340	9.0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
S002	32.22253	-0.56909	8.8	0.268	3.56	0.00	0.000	0.358	8.72	NA	3.01	2.25	4.45	0.00
S003	32.18170	-0.68952	8.4	0.275	3.83	0.00	0.311	0.375	8.40	NA	2.97	2.16	4.77	0.00
S004	32.26802	-0.73168	8.7	0.259	3.45	0.00	0.000	0.355	8.02	NA	2.80	2.09	4.14	0.00
S005	32.30853	-0.71890	NA	0.259	3.73	0.00	0.000	0.392	8.38	NA	3.09	2.22	4.44	0.00
S006	32.32633	-0.67167	8.7	0.268	3.74	0.00	0.174	0.395	8.86	NA	3.25	2.31	4.59	0.00
S007	32.35232	-0.64666	8.6	0.255	3.54	0.00	0.122	0.377	8.59	NA	3.18	2.24	4.30	0.00
S008	32.31278	-0.63200	8.9	0.274	3.62	0.00	0.000	0.368	8.79	NA	3.06	2.27	4.36	0.00
S009a	32.29731	-0.52895	7.8	0.039	9.78	11.56	0.684	2.24	6.80	NA	12.3	1.57	6.43	0.00
S009b	32.29731	-0.52895	9.5	0.254	4.05	0.07	0.000	0.433	8.67	NA	3.69	2.12	4.72	0.00
S010	32.44622	-0.48870	9.7	0.272	3.62	0.00	0.000	0.361	9.05	NA	3.20	2.31	4.76	0.00
S011	32.51683	-0.37463	9.2	0.256	3.54	0.00	0.000	0.338	8.90	NA	3.14	2.31	4.65	0.00
S012	32.57128	-0.35153	9.9	0.252	3.57	0.00	0.118	0.364	8.97	NA	3.10	2.22	3.76	0.00
S013	32.51855	-0.33335	9.9	0.226	2.34	0.00	0.000	0.221	6.28	NA	2.14	1.61	3.29	0.00
S014	32.55737	-0.22757	9.5	0.266	3.80	0.00	0.000	0.335	8.99	NA	3.26	2.31	4.57	0.00

Table 13.4 – Table showing raw environmental and water chemistry data for all Lake Victoria – Part I

Sample	Nitrite	uS	TDS	Salin	Turb	WaterTemp	ChoAbund	SudAbund	TotAbund	Lym	Bul	OtherSpp	Depth	Habitat	Substrate
T001	0.00	127.2	72.2	0.4	1	26.0	3	4	4	1	1	6	2	6	5
T002A	0.10	194.0	114.2	0.4	1	28.4	1	4	4	1	1	5	1	2	2
T002B	0.00	127.9	78.2	0.4	NA	27.6	2	1	2	1	1	5	5	1	5
T003	2.51	708.0	394.0	0.5	1	26.5	1	4	4	0	0	1	1	4	2
T004	NA	NA	NA	NA	1	NA	4	4	4	1	1	6	NA	6	5
T005	NA	240.5	150.2	0.4	1	26.9	1	1	1	0	1	2	1	3	NA
T006a	NA	1053.0	640.0	0.7	1	26.8	1	4	4	1	0	2	1	3	2
T006b	NA	114.8	70.5	0.4	NA	26.6	NA	1	1	1	0	2	5	1	6
T007	NA	99.9	62.2	0.4	NA	26.4	1	1	1	0	1	3	3	1	NA
T008	1.94	480.0	290.0	0.5	1	23.2	1	1	1	0	1	1	3	4	3
T009A	0.09	528.5	325.0	0.5	1	23.9	1	2	2	1	1	2	3	2	2
T009B	0.00	115.4	69.8	0.4	2	23.0	1	1	1	0	1	6	5	1	5
T010	0.08	405.0	244.0	0.5	1	22.6	1	1	1	1	1	3	1	4	NA
T011	NA	243.5	147.6	0.4	NA	26.2	4	1	4	0	1	3	1	1	5
T012	0.13	653.5	397.0	0.5	NA	27.6	1	1	1	1	1	3	1	2	2
T013	0.00	111.3	68.4	0.4	3	26.6	1	1	1	0	0	1	2	1	NA
T014A	2.64	349.0	211.0	0.4	1	24.3	1	1	1	1	0	3	1	2	NA
T014B	0.00	113.8	69.9	0.4	NA	27.0	2	1	2	1	1	3	1	1	4
T015	NA	250.5	151.4	0.4	1	24.7	1	4	4	1	1	3	1	2	2
T016	0.00	383.0	234.0	0.4	1	22.4	1	4	4	1	0	1	1	2	NA
T017	NA	125.2	73.0	0.4	NA	23.0	2	1	2	1	1	5	2	1	NA
T018	0.00	274.0	165.5	0.4	1	23.8	2	1	2	1	0	1	1	2	NA
T019	0.44	320.5	192.7	0.4	1	26.2	1	2	2	1	1	4	1	6	NA
T020	0.00	147.7	83.4	0.4	1	26.6	1	2	2	1	1	4	1	2	NA
T021	0.00	40.4	24.6	0.4	1	27.7	1	3	3	1	1	3	NA	2	2
T022	0.00	96.5	59.5	0.4	1	28.7	1	2	2	0	0	5	NA	2	NA
T023	0.00	81.7	48.8	0.3	1	25.3	1	NA	NA	0	0	2	NA	2	NA
T024	0.00	766.0	465.0	0.6	1	27.2	1	1	1	0	1	3	NA	2	NA
T025	0.17	395.5	312.0	0.5	1	31.4	1	NA	NA	0	1	2	NA	5	NA

Sample	Nitrite	uS	TDS	Salin	Turb	WaterTemp	ChoAbund	SudAbund	TotAbund	Lym	Bul	OtherSpp	Depth	Habitat	Substrate
T026A	0.00	435.0	261.0	0.5	1	27.6	1	4	4	1	1	3	1	7	2
T026B	0.00	112.7	65.5	0.4	NA	27.3	2	1	2	0	1	3	NA	1	1
T027A	0.00	143.9	86.6	0.4	1	28.8	1	4	4	NA	NA	NA	1	2	NA
T027B	0.00	96.3	59.3	0.4	NA	26.9	4	1	4	0	1	2	4	1	NA
T028	0.00	82.7	49.4	0.3	1	30.1	1	1	1	1	1	3	1	7	NA
T029	0.00	112.4	67.2	0.4	1	28.0	2	1	2	1	1	4	3	1	1
T030	0.00	154.6	75.2	NA	NA	28.2	4	1	4	1	1	3	3	1	NA
T031A	0.00	397.0	226.0	0.5	1	26.8	1	4	4	0	1	1	1	2	NA
T031B	NA	129.0	68.8	0.4	1	28.3	1	4	4	1	1	4	NA	1	NA
T032	0.00	95.7	58.1	0.4	NA	26.8	1	1	1	0	1	3	2	1	4
T033A	0.90	301.5	174.0	0.4	1	27.4	1	4	4	0	0	2	1	2	NA
T033B	0.00	401.0	244.0	0.5	2	26.4	4	1	4	1	1	3	5	1	NA
T034A	0.00	929.0	370.0	0.6	1	26.1	1	4	4	0	1	3	1	2	2
T034B	0.00	145.0	85.0	0.4	3	26.8	2	1	2	1	1	2	5	1	1
T035	0.00	170.5	102.2	0.4	NA	26.2	1	1	1	0	1	1	1	1	5
T036A	0.50	677.0	413.0	0.6	1	26.3	1	4	4	1	1	2	1	2	NA
T036B	0.00	92.8	57.5	0.4	NA	26.3	3	1	3	1	1	2	1	1	5
T037A	0.00	120.0	70.6	0.4	1	27.5	1	4	4	1	1	3	1	2	2
T037B	0.00	109.0	68.5	0.4	NA	26.7	2	1	2	0	0	0	5	1	6
T038a	NA	190.6	116.4	0.4	1	22.6	1	2	2	0	0	1	1	2	2
T038b	NA	101.3	60.6	0.4	NA	22.9	1	1	1	1	1	2	1	1	4
T039	0.00	107.8	64.1	0.4	NA	27.5	1	1	1	0	0	0	1	6	2
T040	0.19	329.5	176.9	0.4	1	22.7	1	4	4	1	0	2	1	6	NA
T041	0.000	559.5	359.0	0.5	1	26.7	1	1	1	0	1	3	1	2	2
T042a	0.05	209.0	131.4	0.4	1	26.4	1	2	2	0	0	0	3	2	2
T042b	NA	NA	NA	NA	2	NA	1	1	1	0	0	1	3	1	3
T043a	0.000	385.0	234.0	0.4	1	27.2	1	2	2	0	0	1	1	2	2
T042b	NA	119.0	78.0	0.4	1	27.2	1	1	1	0	0	0	3	1	3
T043c	0.44	412.0	265.0	0.5	1	29.0	1	1	1	1	0	1	1	2	2

Sample	Nitrite	uS	TDS	Salin	Turb	WaterTemp	ChoAbund	SudAbund	TotAbund	Lym	Bul	OtherSpp	Depth	Habitat	Substrate
T044a	0.000	316.5	203.0	0.4	1	29.3	1	1	1	0	1	1	1	7	2
T044b	0.000	99.8	59.7	0.4	2	25.8	2	1	2	0	0	0	NA	1	3
T045	NA	141.0	96.7	0.4	NA	29.3	1	2	2	1	1	3	3	1	NA
T046	0.000	198.0	131.5	0.4	NA	26.0	2	1	2	1	1	6	1	6	2
T047a	0.000	353.5	216.0	0.4	1	23.1	1	NA	NA	1	0	3	1	2	2
T047b	0.13	101.0	59.2	0.4	3	NA	1	1	1	0	0	0	3	1	3
T048a	0.000	346.5	214.0	0.4	1	23.5	1	2	2	1	1	2	1	4	NA
T048b	0.000	118.0	76.9	0.4	NA	24.8	1	1	1	0	0	3	1	1	NA
T049	0.000	158.7	65.4	0.4	2	25.2	1	1	1	0	0	0	5	1	3
T050a	0.000	360.5	217.0	0.4	1	25.2	1	1	1	1	0	1	1	2	2
T050b	0.09	106.1	61.3	0.4	2	25.0	1	1	1	0	1	1	4	1	5
T051	0.000	109.3	65.5	0.4	2	27.3	1	1	1	0	0	0	3	1	3
T052	0.000	114.0	69.1	0.4	1	27.0	1	1	1	1	0	1	1	1	3
T053a	0.000	176.5	110.1	0.4	1	28.6	1	1	1	1	0	1	1	2	2
T053b	0.000	101.6	63.5	0.4	NA	28.3	1	1	1	0	0	1	1	1	3
T054	0.000	114.8	65.9	0.4	3	27.2	1	1	1	0	1	1	5	1	5
T055	0.000	104.8	66.3	0.4	3	27.9	2	1	2	1	1	6	4	1	1
T056	0.000	108.7	66.2	0.4	3	27.3	1	1	1	0	0	0	1	1	3
T057	7.49	112.0	71.8	0.4	1	29.1	3	1	3	1	1	4	3	1	1
T058	0.000	127.0	76.4	0.4	1	28.8	2	1	2	0	1	6	1	8	6
T059	0.000	122.2	75.9	0.4	3	28.6	1	1	1	0	0	6	1	1	2
T060a	0.000	163.2	96.4	0.4	1	29.8	1	1	1	0	0	0	1	2	2
T060b	0.000	125.5	76.1	0.4	3	28.3	1	1	1	0	0	0	1	1	6
T061	0.000	325.0	198.1	0.4	1	28.8	1	1	3	1	1	2	1	4	2
T062	0.000	126.8	77.6	0.4	1	24.3	1	1	1	1	0	1	1	7	2
T063	0.000	329.5	201.0	0.4	1	25.5	1	3	3	1	1	4	1	7	2
T064a	0.000	438.5	265.0	0.5	1	29.5	1	4	4	0	0	2	1	2	2
T064b	0.000	108.3	62.9	0.4	1	28.3	1	1	1	0	0	2	1	1	3
T065	0.000	399.5	242.0	0.5	3	27.6	1	1	1	0	0	0	1	4	2

Sample	Nitrite	uS	TDS	Salin	Turb	WaterTemp	ChoAbund	SudAbund	TotAbund	Lym	Bul	OtherSpp	Depth	Habitat	Substrate
T066a	0.000	483.0	182.4	0.4	1	21.7	1	4	4	0	0	0	1	2	2
T066b	0.000	119.3	64.4	0.4	3	22.1	2	1	2	1	1	3	5	1	5
T067a	0.000	773.0	460.0	0.6	1	26.6	1	3	3	0	0	3	1	2	2
T067b	38.3	120.0	70.3	0.4	NA	28.2	1	1	1	0	0	2	1	1	3
T068a	0.000	678.5	403.0	0.5	1	26.4	1	1	1	1	1	2	1	2	NA
T068b	0.000	111.5	69.8	0.4	NA	28.6	2	1	2	NA	NA	NA	2	1	5
T069	0.000	105.1	61.9	0.4	3	24.8	1	1	1	0	0	0	5	1	3
T070a	0.000	299.5	182.3	0.4	1	24.8	1	2	2	1	0	1	1	2	5
T070b	0.000	102.9	60.9	0.4	2	25.4	1	1	1	1	1	2	1	1	5
U001	NA	107.5	70.0	0.4	2	28.5	3	1	3	1	1	6	1	1	1
U002	NA	92.2	57.2	0.3	3	26.9	1	1	1	0	1	4	1	1	4
U003	NA	90.5	52.6	0.3	NA	31.4	1	1	1	0	0	2	1	1	6
U004	NA	92.0	58.3	0.4	1	29.0	1	1	1	0	0	2	1	7	5
U005	NA	99.0	63.2	0.4	1	25.1	3	1	3	1	1	5	1	1	5
U006	NA	90.8	56.6	0.3	1	25.6	2	1	2	1	1	5	3	2	6
U007	NA	89.3	54.0	0.3	1	27.6	1	1	1	1	0	3	1	6	6
U008	NA	77.8	48.1	0.3	1	30.2	2	1	2	1	1	5	1	1	5
U009	NA	95.5	59.8	0.4	NA	25.4	2	1	2	1	1	5	1	1	5
U010	NA	99.5	62.2	0.4	2	25.5	2	1	2	1	1	5	1	1	3
U011	NA	95.6	58.5	0.4	NA	27.6	2	1	2	1	1	5	5	1	5
U012	NA	108.0	65.0	0.4	NA	27.6	NA	NA	4	0	0	2	4	2	5
U013	NA	90.9	56.2	0.3	2	25.5	2	1	2	1	1	4	5	1	5
U014	NA	106.5	66.0	0.4	NA	26.2	2	1	2	0	0	0	3	1	5
U015	NA	91.7	57.2	0.3	NA	26.1	2	1	2	0	0	1	5	1	2
U016	NA	89.5	54.0	0.3	NA	27.7	4	1	4	1	1	6	1	1	4
U017	NA	92.9	58.1	0.4	NA	27.3	1	1	1	0	0	0	NA	1	NA
U018	NA	99.0	61.0	0.4	1	NA	2	1	2	0	1	3	5	1	NA
U019	NA	121.0	75.2	0.4	NA	25.8	1	1	1	0	1	1	3	1	3
U020	NA	85.0	50.5	0.3	1	25.0	NA	NA	4	1	0	2	3	2	2

Sample	Nitrite	uS	TDS	Salin	Turb	WaterTemp	ChoAbund	SudAbund	TotAbund	Lym	Bul	OtherSpp	Depth	Habitat	Substrate
U021	NA	176.5	108.0	0.4	3	29.0	3	3	4	0	0	1	1	2	2
U022	NA	24.8	15.0	0.3	3	25.7	1	1	1	0	0	0	4	1	3
U023a	NA	101.8	64.2	0.4	2	25.0	4	1	4	1	1	3	5	1	1
U023b	NA	126.6	76.6	0.4	1	28.9	4	1	4	0	0	4	1	2	5
U024a	NA	103.0	63.4	0.4	3	26.9	NA	1	4	0	1	5	5	1	5
U024b	NA	113.6	74.0	0.4	3	27.0	NA	1	4	0	1	5	3	1	5
U025a	NA	102.0	64.0	0.4	NA	27.3	NA	1	3	NA	NA	NA	5	1	5
U025b	NA	95.8	58.4	0.4	NA	26.0	NA	1	3	NA	NA	NA	1	1	3
U026	NA	96.5	60.0	0.4	NA	27.9	2	1	2	0	1	4	4	1	5
U027	NA	94.5	58.8	0.4	2	25.7	NA	1	3	1	1	6	1	1	5
U028	NA	113.5	71.2	0.4	2	25.4	NA	1	4	1	1	2	1	1	5
U029	NA	91.6	56.5	0.3	NA	26.7	3	1	3	1	1	5	1	1	5
U030b	NA	96.5	59.7	0.4	NA	28.4	4	1	4	1	1	6	3	6	5
U031	NA	118.4	72.1	0.4	3	30.9	NA	1	4	1	1	4	1	1	1
U032	NA	186.0	115.4	0.4	3	27.2	NA	1	3	0	1	1	5	1	5
U033	NA	169.1	104.1	0.4	1	26.4	4	1	4	1	1	6	1	1	5
U034	NA	118.8	71.3	0.4	2	26.8	2	1	2	1	1	2	5	1	1
U035b	NA	113.5	67.5	0.4	NA	35.5	NA	1	4	0	0	1	3	1	5
U036	NA	108.5	67.4	0.4	3	29.7	2	1	2	0	0	0	4	1	NA
U037	NA	NA	NA	NA	2	NA	4	3	4	0	1	5	2	6	5
U038a	NA	109.4	67.4	0.4	1	30.3	2	1	1	1	0	4	1	1	6
U038b	NA	376.5	231.0	0.4	1	28.0	1	NA	4	1	0	4	3	2	2
U039	NA	202.0	119.1	0.4	1	30.9	NA	NA	4	1	1	3	2	2	2
U040	NA	109.5	66.6	0.4	3	29.4	3	1	3	0	1	5	1	1	5
U041	NA	NA	NA	NA	1	NA	1	1	1	1	0	2	2	2	2
U042	NA	445.5	276.0	0.5	1	31.5	NA	1	3	0	1	1	2	2	2
U043	NA	112.2	69.2	0.4	3	30.1	1	1	1	0	1	5	1	6	6
U044	NA	102.1	62.6	0.4	1	30.6	NA	NA	3	1	0	1	2	1	5
U045	NA	139.9	85.6	0.4	1	30.1	1	1	1	0	1	2	4	1	5

Sample	Nitrite	uS	TDS	Salin	Turb	WaterTemp	ChoAbund	SudAbund	TotAbund	Lym	Bul	OtherSpp	Depth	Habitat	Substrate
U046	NA	127.8	73.9	0.4	NA	26.8	NA	1	4	1	1	6	4	1	5
U047	NA	200.5	122.6	0.4	3	27.9	4	1	4	1	1	4	1	1	5
U048b	NA	120.2	72.2	0.4	NA	27.7	NA	1	4	0	1	3	1	1	5
U049	NA	109.0	66.4	0.4	2	27.7	NA	1	3	0	1	2	3	1	5
U050	NA	93.2	56.6	0.3	3	27.3	NA	NA	3	0	1	1	4	6	5
U051	NA	95.0	58.7	0.4	NA	28.8	1	1	1	0	0	1	1	1	2
U052	NA	166.6	102.7	0.4	NA	28.6	NA	1	4	1	1	5	2	1	4
U053	NA	116.3	70.4	0.4	NA	28.7	NA	NA	4	1	1	3	1	1	5
U054	NA	145.0	88.1	0.4	1	28.6	NA	NA	2	1	1	3	1	6	2
U055	0.00	119.7	70.2	0.4	NA	26.1	1	4	4	NA	NA	NA	NA	1	1
UW003	NA	92.0	57	0.4	2	27.8	NA	NA	NA	NA	NA	NA	5	1	NA
UW001	NA	93.9	58	0.4	2	25.7	NA	NA	NA	NA	NA	NA	5	1	NA
K001A	0.11	1588.0	NA	NA	1	28.0	1	3	3	0	0	1	1	2	2
K001B	0.07	123.4	81.7	0.4	1	28.5	1	1	1	0	0	0	5	1	3
K002A	0.00	2054.0	NA	NA	1	31.1	1	4	4	0	1	5	1	2	2
K002B	0.00	121.5	69.0	0.4	1	30.5	2	1	2	1	1	3	5	1	5
K003A	0.17	NA	NA	NA	1	NA	1	2	2	0	0	0	1	2	2
K003B	2.11	199.8	115.7	0.4	1	29.2	1	1	1	0	1	1	5	1	2
K004	1.63	182.5	107.7	0.4	1	25.7	1	3	3	0	1	4	3	1	5
K005	0.27	173.1	103.1	0.4	1	25.8	1	3	3	1	1	3	3	6	2
K006A	0.21	260.0	156.1	0.4	1	24.5	1	2	2	0	0	1	1	2	5
K006B	0.18	109.3	64.7	0.4	1	26.5	3	1	3	0	1	5	4	1	5
K007A	0.00	332.0	201.0	0.4	1	30.2	1	2	2	0	0	0	1	2	5
K007B	0.08	108.7	65.5	0.4	2	28.0	1	1	1	0	1	1	3	1	3
K008A	0.00	1060.5	625.0	0.7	1	33.0	1	3	3	0	1	1	1	2	5
K008B	0.00	149.4	83.9	0.4	2	30.2	2	1	2	0	1	3	3	1	5
K009	0.00	93.0	59.0	NA	1	29.2	2	1	2	0	1	1	1	6	NA
K010	0.05	62.2	67.1	0.4	3	30.3	1	1	1	0	0	0	5	1	3
K011	1.60	144.9	85.8	0.4	1	27.4	1	3	3	1	0	4	2	6	NA

Sample	Nitrite	uS	TDS	Salin	Turb	WaterTemp	ChoAbund	SudAbund	TotAbund	Lym	Bul	OtherSpp	Depth	Habitat	Substrate
K012	0.00	132.5	80.2	0.4	1	27.6	1	2	2	0	0	2	5	1	5
K013A	0.00	124.0	74.2	0.4	1	29.2	1	1	1	1	0	2	1	1	5
K013B	0.05	106.4	63.6	0.4	1	29.4	4	1	4	0	1	5	3	1	1
K014	0.00	132.8	80.9	0.4	3	28.4	1	1	1	0	0	0	5	1	3
K015A	0.07	472.5	282.0	0.5	1	29.3	1	2	2	0	0	1	1	2	1
K015B	1.26	132.2	76.4	0.4	3	29.6	2	1	2	0	1	4	5	1	5
K016B	0.00	101.8	62.0	0.4	3	28.4	1	1	1	0	1	2	5	1	5
K017	0.00	280.5	170.4	0.4	1	28.0	1	2	2	0	1	2	1	2	2
K018B	0.00	130.9	79.0	0.4	3	26.4	1	1	1	0	0	0	3	6	6
K019A	NA	1141.0	689.0	0.7	1	25.0	1	3	3	NA	NA	NA	1	2	1
K019B	0.05	107.7	65.8	0.4	1	23.9	2	1	2	NA	NA	NA	3	1	1
K020A	0.09	518.5	314.0	0.5	1	27.2	1	3	3	0	1	5	1	7	2
K020B	0.85	681.0	419.0	0.6	1	27.2	1	3	3	0	1	5	1	1	5
K021A	0.04	162.0	97.9	0.4	1	30.2	1	3	3	0	0	0	2	2	1
K021B	0.00	154.9	92.0	0.4	2	29.3	1	1	1	0	0	0	5	1	4
K022	3.49	548.5	330.0	0.5	1	28.4	1	2	2	0	1	2	3	7	NA
K023	NA	371.0	185.0	NA	1	27.2	1	1	1	0	1	3	NA	3	NA
K024	0.11	328.0	165.0	NA	1	27.0	1	1	1	1	1	3	1	4	5
K025	3.84	378.0	189.0	NA	1	23.2	1	1	1	0	1	2	4	3	NA
K026	0.47	503.0	191.0	NA	1	22.4	1	1	1	0	1	1	3	4	2
K027	0.08	186.1	112.1	0.4	1	32.4	1	2	2	0	1	4	2	8	2
K028	0.00	149.1	90.8	0.4	1	32.0	1	3	3	0	1	3	4	1	1
K029	0.05	514.0	309.0	0.5	1	29.5	1	4	4	1	1	5	1	1	1
K030A	0.00	539.0	333.0	0.5	1	32.8	1	2	2	1	0	2	1	2	2
K030B	0.00	114.1	67.4	0.4	NA	28.9	2	1	2	0	1	1	2	1	4
K031	0.00	115.8	69.4	0.4	2	28.8	2	1	2	0	1	4	1	1	4
K032	0.00	230.5	139.3	0.4	1	27.8	1	1	1	0	1	1	1	4	2
K033A	2.42	125.6	75.3	0.4	1	27.4	1	2	2	1	1	6	1	2	2
K033B	0.00	117.4	70.6	0.4	2	26.5	3	1	3	0	1	3	1	1	4

Sample	Nitrite	uS	TDS	Salin	Turb	WaterTemp	ChoAbund	SudAbund	TotAbund	Lym	Bul	OtherSpp	Depth	Habitat	Substrate
U023.08	0.00	113.8	71.9	0.4	NA	25.7	NA	1	3	NA	NA	NA	3	1	NA
U024B.08	0.00	107.8	66.2	0.4	NA	25.7	4	1	4	NA	NA	NA	NA	1	NA
U027.08	0.00	96.0	61.4	0.4	NA	27.5	NA	NA	NA	NA	NA	NA	NA	1	NA
U028.08	0.00	100.7	62	0.4	NA	26.4	NA	NA	NA	NA	NA	NA	NA	1	NA
Water 1 08	0.00	146.0	87.8	0.4	NA	25.2	NA	NA	NA	NA	NA	NA	NA	1	NA
S001	NA	92.0	47	NA	L0w	28.0	0	0	1	0	1	2	1	1	3
S002	0	94.0	45	NA	L0w	26.1	0	0	1	0	0	1	3	1	4
S003	0	93.0	46	NA	High	23.9	0	0	1	1	1	2	5	1	4
S004	0	93.0	45	NA	M0d	24.8	0	0	1	0	0	0	2	1	5
S005	0	NA	NA	NA	NA	NA	0	0	1	0	1	2	NA	NA	NA
S006	0	92.0	45	NA	M0d	23.6	0	0	1	0	0	0	2	1	NA
S007	0	93.0	47	NA	M0d	24.2	0	0	1	1	1	3	5	1	4
S008	0	95.0	46	NA	L0w	24.0	2	0	2	0	1	1	1	1	3
S009a	0	152.0	77	NA	L0w	27.2	0	4	4	0	0	0	1	2	2
S009b	0	92.0	46	NA	L0w	27.6	0	0	1	0	1	3	1	1	4
S010	0	96.0	47	NA	M0d	28.5	4	0	4	1	1	2	5	1	4
S011	0	96.0	45	NA	L0w	26.4	2	0	2	1	1	4	4	1	4
S012	0	90.0	45	NA	M0d	23.5	0	0	1	0	1	4	5	1	4
S013	0	91.0	45	NA	High	27.1	0	0	1	0	1	3	5	1	4
S014	0	90.0	44	NA	L0w	26.4	0	0	1	0	0	1	3	1	4

Table 13.5 - Table showing raw environmental and water chemistry data for all Lake Victoria – Part II

Site *GPS coordinates*	pH	F	Cl	NO3	PO4	SO4	Na	K	Mg	Ca	Bromide	Nitrite	uS	TDS	Water Temp	Shannon	Simpson	Fisher
KM01 *S0.08832; E32.65279*	10.9	0.278	3.387	0	0.121	0.347	9.067	2.985	2.181	4.355	0	0	91	45	28.8	2.024	0.864	1.857
KM02 *S0.08871; E32.65261*	10.9	0.28	3.388	0	0.108	0.32	9.082	2.962	2.149	4.142	0	0	92	46	28.8	1.881	0.83	1.607
KM03 *S0.08919; E32.65228*	10.8	0.282	3.357	0	0.16	0.327	9.048	2.925	2.129	3.876	0	0	93	46	28.5	1.696	0.754	1.95
KM04 *S0.08903; E32.65201*	10.8	0.266	3.341	0	0.216	0.317	8.976	2.962	2.062	3.846	0	0	95	47	27.5	1.665	0.772	1.972
KM05 *S0.08901; E32.65071*	11.2	0.286	3.354	0	0.333	0.364	9.122	2.869	2.079	3.974	0	0	92	46	28.6	1.547	0.743	1.957
KM06 *S0.08893; E32.65027*	11	0.308	3.019	0	0.298	0.376	9.365	2.598	2.191	4.666	0	0	92	46	28.8	1.242	0.656	1.727
KM07 *S0.08883; E32.64846*	11.2	0.303	2.935	0	0.563	0.465	8.289	2.521	1.882	3.642	0	0	128	64	28.3	1.335	0.694	1.312
KM08 *S0.08847; E32.64725*	11.1	0.264	3.088	0	0.335	0.283	9.281	2.609	2.223	4.599	0	0	92	46	28.3	1.306	0.621	1.462
KM09A *S0.08838; E32.64689*	10.5	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	110	53	29	0.974	0.593	0.617
KM09B *S0.08838; E32.64689*	11.1	0.265	3.242	0	0.257	0.316	9.287	2.812	2.208	4.304	0	0	93	47	29.3	1.847	0.827	1.642
KM10 *S0.08776; E32.64590*	10.9	0.312	3.24	0	0.129	0.298	8.966	2.77	2.137	4.186	0	0	91	45	27	1.538	0.74	1.286
KM11 *S0.08680; E32.64353*	10.9	0.249	3.315	0	0	0.322	9.099	2.889	2.291	4.379	0	0	90	44	27.3	1.419	0.735	1.033
KM12 *S0.08585; E32.64416*	10.8	0.269	3.36	0	0.205	0.266	9.286	3.032	2.283	4.54	0	0	100	50	27.2	1.844	0.827	1.816
KM13 *S0.08536; E32.64500*	11.1	0.313	3.161	0	0.185	0.19	8.876	2.806	2.028	3.729	0	0	102	51	27	1.551	0.743	1.78
KM14 *S0.08492; E32.64629*	11.1	0.092	0.193	0.075	0	0	0.184	0.089	0.18	0.69	0	0	93	46	28.3	0.326	0.148	0.764
KM15 *S0.08510; E32.64790*	11	0.309	3.309	0	0.228	0.341	9.063	2.936	2.102	3.839	0	0	94	47	29	1.657	0.778	1.252
KM16 *S0.08526; E32.64833*	11	0.277	3.323	0	0.194	0.326	9.01	2.915	2.082	3.75	0	0	96	47	29.2	1.704	0.786	1.464
KM17 *S0.08535; E32.64955*	11.4	0.274	3.28	0	0.166	0.319	8.934	2.903	2.11	3.96	0	0	93	46	29.5	1.592	0.759	1.495
KM18 *S0.08541; E32.65033*	11.4	0.272	3.596	0	0.238	0.401	9.219	3.203	2.181	4.314	0	0	97	48	28.3	1.687	0.779	1.805
KM19 *S0.08567; E32.25073*	11	0.304	3.631	0	0.135	0.44	9.251	3.176	2.218	4.379	0	0	92	46	28.1	1.738	0.787	1.792
KM20 *S0.08662; E32.65265*	11	0.271	3.402	0	0.225	0.348	9.112	2.953	2.215	4.182	0	0	92	45	29.2	1.762	0.799	1.721

Table 13.6 – Summary table of water chemistry and environmental measurements for Kimi Island

Site *GPS coordinates*	pH	F	Cl	NO3	PO4	SO4	Na	K	Mg	Ca	Bromide	Nitrite	uS	TDS	Water Temp	Shannon	Simpson	Fisher
NG01 *S0.09988; E32.65304*	10.2	0.259	3.334	0	0	0.344	8.991	2.908	2.231	4.353	0	0	93	46	26.9	1.595	0.756	1.497
NG02 *S0.10020; E32.65392*	10.1	0.264	3.367	0	0	0.353	8.939	2.866	2.235	4.357	0	0	92	46	27.1	1.605	0.751	1.737
NG03 *S0.10095; E32.65538*	10.2	0.259	3.269	0	0	0.338	8.567	2.774	2.137	4.22	0	0	90	45	28.3	1.543	0.701	1.654
NG04 *S0.10227; E32.65812*	10.3	0.259	3.211	0	0	0.335	8.462	2.709	2.135	4.301	0	0	92	45	27.8	1.23	0.586	1.48
NG05 *S0.10281; E32.65736*	10.5	0.249	3.219	0	0.065	0.336	8.541	2.728	2.095	4.39	0	0	98	49	28.9	1.458	0.694	1.63
NG06 *S0.10328; E32.65615*	10.1	0.26	3.362	0	0	0.332	9.004	2.86	2.274	4.771	0	0	96	48	28.6	0.805	0.528	0.498
NG07 *S0.10404; E32.65680*	9.4	0.259	3.381	0	0	0.313	8.962	2.854	2.229	4.476	0	0	89	45	29.2	1.709	0.78	1.656
NG08 *S0.10507; E32.65619*	10.5	0.268	3.315	0	0.118	0.331	8.839	2.782	2.244	4.484	0	0	97	49	29	1.035	0.608	0.925
NG09 *S0.10630; E32.65387*	10.6	0.258	3.297	0	0	0.31	8.973	2.894	2.251	4.985	0	0	101	50	28.2	1.311	0.702	1.049
NG10 *S0.10671; E32.65243*	10.6	0.229	3.317	0	0	0.324	9.082	2.925	2.173	4.525	0	0	101	50	28.3	1.436	0.733	1.196
NG11 *S0.10684; E32.65204*	10.2	0.265	3.367	0	0	0.353	8.976	2.877	2.239	4.693	0	0	99	49	28.8	1.308	0.651	1.936
NG12 *S0.10719; E32.65134*	10	0.264	3.222	0	0	0.341	8.687	2.77	2.24	4.644	0	0	95	47	28.3	1.556	0.78	1.363
NG13 *S0.10748; E32.65077*	10.3	0.268	3.377	0	0	0.361	9.033	2.919	2.291	4.82	0	0	96	48	28.2	1.493	0.739	1.246
NG14 *S0.10721; E32.64796*	10.1	0.23	3.17	0.074	0.133	0.328	8.332	2.684	2.073	4.142	0	0	93	46	27.1	1.491	0.713	1.468
NG15 *S0.10660; E32.64780*	10.2	0.222	2.651	0	0	0.279	7.288	2.321	2.015	4.503	0	0	96	48	27.3	1.424	0.702	1.316
NG16 *S0.10528; E32.64871*	10.4	0.255	3.356	0	0	0.352	9.115	2.939	2.33	4.794	0	0	93	46	28.2	1.466	0.749	1.406
NG17 *S0.10404; E32.64935*	10.1	0.254	3.091	2.509	0.254	0.332	8.106	2.591	2.042	3.701	0	0	91	45	28.1	1.055	0.562	0.973
NG18 *S0.10229; E32.65055*	10.2	0.271	3.429	0	0	0.351	9.07	2.906	2.267	4.768	0	0	93	46	27.3	1.642	0.733	1.632
NG19 *S0.10143; E32.65174*	9.9	0.266	3.344	0	0	0.35	8.963	2.855	2.241	4.344	0	0	92	45	28.1	1.724	0.8	1.381
NG20 *S0.10015; E32.65185*	10.2	0.254	2.851	0	0	0.302	7.595	2.382	1.927	4.109	0	0	91	46	28.2	1.887	0.829	1.557

Table 13.7 – Summary table of water chemistry and environmental measurements for Ngamba Island

Site	Wave action	Depth	Habitat	Substrate
NG01	Low	Shallow / Moderate	Lake	Combination incl. aquatic veg.
NG02	Low	Shallow / Moderate	Lake	Aquatic vegetation
NG03	Low	Shallow / Moderate	Lake	Combination incl. aquatic veg.
NG04	Low	Moderate / Deep	Lake	Aquatic vegetation
NG05	Low	Moderate / Deep	Lake	Aquatic vegetation
NG06	Low	Moderate	Lake	Organic debris/mud
NG07	Low	Moderate	Lake	Combination excl. aquatic veg.
NG08	High	Moderate	Lake	Rock
NG09	Moderate	Shallow / Moderate	Lake	Rock
NG10	Moderate	Shallow / Moderate	Lake	Rock
NG11	Moderate	Shallow / Moderate	Lake	Rock
NG12	Moderate	Shallow	Lake	Rock
NG13	High	Shallow	Lake	Rock
NG14	High	Deep	Lake	Rock
NG15	Moderate	Moderate	Lake	Rock
NG16	Low	Shallow	Lake	Rock
NG17	Low	Moderate / Deep	Lake	Rock
NG18	Low	Shallow / Moderate	Lake	Rock
NG19	Low	Deep	Lake	Combination incl. aquatic veg.
NG20	Low	Moderate / Deep	Lake	Aquatic vegetation
KM01	High	Moderate / Deep	Lake	Aquatic vegetation
KM02	Moderate	Moderate	Lake	Combination incl. aquatic veg.
KM03	Moderate	Shallow / Moderate	Lake	Rock
KM04	Low	Moderate	Lake	Aquatic vegetation
KM05	Low	Shallow	Lake	Combination incl. aquatic veg.
KM06	Low	Shallow / Moderate	Lake	Organic debris/mud
KM07	Low	Shallow / Moderate	Lake	Combination incl. aquatic veg.
KM08	Low	Moderate	Lake	Combination incl. aquatic veg.
KM09A	Low	Shallow	Marsh	Organic debris/mud
KM09B	Low	Shallow	Lake	Combination incl. aquatic veg.
KM10	Low	Deep	Lake	Combination incl. aquatic veg.
KM11	Low	Moderate / Deep	Lake	Combination incl. aquatic veg.
KM12	Low	Shallow / Moderate	Lake	Aquatic vegetation
KM13	Low	Shallow / Moderate	Lake	Combination incl. aquatic veg.
KM14	Moderate	Shallow / Moderate	Lake	Combination incl. aquatic veg.
KM15	Low	Moderate	Lake	Aquatic vegetation
KM16	Low	Shallow	Lake	Aquatic vegetation
KM17	Moderate	Shallow / Moderate	Lake	Aquatic vegetation
KM18	Moderate	Moderate	Lake	Combination incl. aquatic veg.
KM19	High	Moderate	Lake	Combination incl. aquatic veg.
KM20	High	Moderate / Deep	Lake	Aquatic vegetation

Table 13.8 – Summary table of categorical environmental variables for Kimi and Ngamba Islands

13.3 Chapter 7 supplements: *Biomphalaria* taxonomy

Morphogroup	*B. choanomphala*-like	*B. choanomphala*-intermediate	*B. sudanica*-like	*B. sudanica*-intermediate
B. choanomphala-like	0.016			
B. choanomphala-intermediate	0.019	0.017		
B. sudanica-like	0.018	0.020	0.015	
B. sudanica-intermediate	0.015	0.017	0.014	0.012

Table 13.9 – Table of uncorrected COI genetic distance values within and between *Biomphalaria* morphogroup

The diagonal (in italics) shows the within-morphogroup mean genetic distance.

Site	K029	T001	T033a	T033b	U023a	U023b	U046
K029	*0.004*						
T001	0.008	*0.008*					
T033a	0.012	0.013	*0.010*				
T033b	0.014	0.016	0.015	*0.014*			
U023a	0.020	0.022	0.015	0.021	*0.017*		
U023b	0.021	0.022	0.015	0.022	0.016	*0.016*	
U046	0.019	0.020	0.013	0.019	0.014	0.014	*0.010*

Table 13.10 – Table of uncorrected COI genetic distance values within and between sites used for the *Biomphalaria* taxonomic analysis

The diagonal (in italics) shows the within-site mean genetic distance.

Morphogroup	*B. choanomphala*-like	*B. choanomphala*-intermediate	*B. sudanica*-like	*B. sudanica*-intermediate
B. choanomphala-like	*0.007*			
B. choanomphala-intermediate	0.008	*0.007*		
B. sudanica-like	0.008	0.008	*0.008*	
B. sudanica-intermediate	0.006	0.006	0.007	*0.005*

Table 13.11 – Table of uncorrected 16S genetic distance values within and between *Biomphalaria* morphogroup

The diagonal (in italics) shows the within-morphogroup mean genetic distance.

Site	K029	T001	T033a	T033b	U023a	U023b	U046
K029	*0.003*						
T001	0.008	*0.004*					
T033a	0.007	0.005	*0.004*				
T033b	0.008	0.004	0.005	*0.004*			
U023a	0.010	0.008	0.007	0.008	*0.008*		
U023b	0.009	0.007	0.006	0.007	0.007	*0.007*	
U046	0.010	0.009	0.008	0.008	0.009	0.008	*0.008*

Table 13.12 – Table of uncorrected 16S genetic distance values within and between sites used for the *Biomphalaria* taxonomic analysis

The diagonal (in italics) shows the within-site mean genetic distance.

Morpho-group		Measurements				
	Descriptors	Height	Width	Depth 1	Depth 2	Whorls
B. choanomphala	Mean	5.31	6.43	2.21	2.94	2.8
	Standard dev.	0.74	0.90	0.30	0.37	0.27
	Variance	0.54	0.82	0.09	0.14	0.07
B. sudanica	Mean	9.54	10.78	2.93	3.41	4.7
	Standard dev.	1.90	2.17	0.41	0.53	0.41
	Variance	3.60	4.72	0.17	0.28	0.17
B. choanomphala intermediate	Mean	7.56	8.41	2.35	2.84	4.3
	Standard dev.	2.53	2.80	0.29	0.43	0.52
	Variance	6.39	7.86	0.08	0.18	0.27
B. sudanica intermediate	Mean	8.55	9.76	2.82	3.26	4.3
	Standard dev.	1.48	1.75	0.32	0.32	0.44
	Variance	2.20	3.07	0.10	0.11	0.20
B. pfeifferi	Mean	5.35	6.46	2.43	2.87	3.1
	Standard dev.	2.37	2.65	1.03	1.06	0.25
	Variance	5.61	7.03	1.06	1.13	0.06
B. choanomphala 'Types' - Berlin	Mean	6.73	7.97	2.68	3.47	3.6
	Standard dev.	1.34	1.80	0.71	0.92	0.29
	Variance	1.79	3.24	0.50	0.84	0.09
B. sudanica 'Types' - Berlin	Mean	9.28	10.76	2.95	3.38	4.3
	Standard dev.	3.18	3.60	0.62	0.77	0.64
	Variance	10.11	12.99	0.39	0.59	0.41

Table 13.13 – Mean, standard deviation and variance of shell measurements, per morpho-group

All measurements are in mm apart from whorls, which were counted; fractions were calculated as a proportion of the complete whorl.

Morphotype	Measurement			
	Descriptor	# Diverticula	Preputium	Penis sheath
B. choanomphala	Mean	17.4	2.34	1.82
	Standard dev.	3.08	0.42	0.35
	Variance	9.46	0.17	0.12
B. sudanica	Mean	23.0	2.95	2.13
	Standard dev.	3.29	0.64	0.58
	Variance	10.84	0.41	0.33
B. choanomphala intermediate	Mean	25.5	3.21	2.15
	Standard dev.	0.71	0.46	0.21
	Variance	0.50	0.21	0.04
B. sudanica intermediate	Mean	20.8	2.97	2.17
	Standard dev.	5.23	0.54	0.43
	Variance	27.31	0.29	0.19

Table 13.14 – Mean, standard deviation and variance of internal anatomy measurements, per morphotype

'Standard dev' refers to the standard deviation of the mean of the measurement. Preputium and penis sheath length were measured in mm.

13.4 Chapter 8 supplements: *Biomphalaria* population genetics

Site	Haplotypes	Frequency
K001a	4	4
	5	4
	6	1
	104	1
K002a	5	5
	7	1
	8	1
	10	1
	13	1
	105	1
K006a	9	3
	20	2
	113	4
	114	1
K006b	11	3
	14	1
	16	2
	106	1
	107	1
	108	1
	109	1
K013b	1	2
	5	1
	12	2
	15	2
	17	1
	110	2
	111	1
K020b	5	4
	8	3
	18	1
	19	1
	61	1
	112	1
K029	5	3
	7	2
	8	6
	103	1
T001	1	7
	7	1
	24	1
	82	1
	83	1
T011	21	2
	22	1
	23	2
	84	4
	85	1
	86	1
T016	1	4
	23	3
	24	4
T026a	1	1
	24	8
	25	1
	87	1
T027a	1	3
	24	3
	26	3
	27	1
	28	1
	29	1
	30	2
	31	1
T027b	88	3
	89	2
	90	1
	91	1
T033a	32	5
	33	3
	92	1
T033b	1	1
	2	2
	34	2
	35	1
	36	1
	37	3
	93	1

Table 13.15 – COI haplotype number and frequency per site – K001a to T033b

Site	Haplotypes	Frequency	Site	Haplotypes	Frequency
T036a	5	2	U023b	26	1
	29	5		28	1
	94	1		48	1
	95	1		49	1
	96	1		51	2
T040	1	3		52	1
	38	1		53	1
	97	1		83	1
	98	1		117	1
	99	1		118	1
	100	1	U028	54	1
	101	1		55	1
	123	1		56	3
T064a	1	1		73	1
	3	2		74	1
	9	4		75	1
	98	1		76	2
	102	1		77	1
	124	1	U030b	32	1
U005	1	1		57	2
	32	1		58	1
	39	1		59	2
	40	5		119	1
	41	1		120	1
	68	1		121	1
	69	1		122	1
U012	1	1		127	1
	42	6	U030c	8	2
	43	1		27	8
	70	2		79	1
U020	44	1	U037	8	1
	45	9		27	1
	71	1		60	1
U021	46	9		61	1
	72	1		62	2
U023a	47	1		63	1
	48	2		64	2
	49	1		65	1
	50	1		78	1
	52	1	U046	32	2
	115	1		66	1
	116	1		67	5
	125	1		80	2
	126	1		81	1

Table 13.16 – COI haplotype number and frequency per site – T036a to U046

Site	Haplotype	Frequency
K002a	69	5
	70	3
	71	2
	147	1
K006a	68	2
	70	1
	72	1
	73	1
	75	1
	85	1
	148	1
	149	1
	150	1
K006b	79	6
	80	3
		1
	74	
	76	1
	77	1
	151	1
	152	1
	153	1
	154	1
	155	1
	156	1
K013b	51	2
	78	4
	79	1
	157	2
	175	1
	176	1
	177	1
K020b	69	5
	70	1
	79	1
	158	1
	159	2
	160	1
K029	69	6
	79	1
	85	1
	146	1
T001	8	1
	9	4
	67	1
	89	1
	90	1
	91	1
	92	1
T011	9	1
	29	1
	30	1
	51	1
	93	3
	94	1
	95	1
	96	1
T016	9	3
	10	4
	11	4
T026a	9	2
	11	3
	12	1
	19	2
	65	1
	97	2
T027a	19	3
	28	1
	57	2
	59	1
	62	2
	81	2
	82	1
	98	1
T027b	24	1
	26	1
	27	1
	61	1
	99	1
	100	2
	101	1
	102	2
T033a	25	1
	28	1
	62	5
	84	3
T033b	13	1
	14	1
	22	1
	49	1
	50	1
	52	1
	63	1
	92	1
	131	1
	132	1
	133	1
	178	1
T036a	31	5
	134	1
	135	1
	136	1
	137	1
	138	1
	139	1
T040	9	2
	23	1
	57	1
	140	1
	141	1
	142	1
	143	1
	144	1
	174	1
T064a	51	1
	65	2
	66	3
	67	2
	68	1
	145	1

Table 13.17 – 16S haplotype number and frequency per site – K001a to T064a

Site	Haplotype	Frequency	Site	Haplotype	Frequency
U005	4	2	U028	40	1
	5	1		41	1
	33	1		53	1
	103	1		114	1
	104	1		115	1
	105	1		116	1
	106	1		117	1
	107	1		118	1
	108	1		119	1
	109	1		120	1
U012	6	1	U030b	42	1
	7	1		43	1
	9	1		44	1
	34	1		46	2
	35	1		169	1
	110	2		170	1
	111	3		171	1
U020	21	1		172	1
	36	1		173	1
	112	6		180	1
	113	1	U037	2	1
U021	19	8		47	1
	20	1		48	1
	83	1		54	1
U023a	17	2		55	1
	18	2		56	1
	37	1		58	1
	38	1		60	2
	161	1		121	1
	162	1	U046	3	1
	163	1		64	1
	164	1		86	1
	179	1		87	1
U023b	1	1		88	1
	15	1		124	1
	16	1		128	1
	17	1		129	1
	39	1		130	1
	165	1			
	166	1			
	167	1			
	168	1			
	181	1			
	182	1			

Table 13.18 – 16S haplotype number and frequency per site – U005 to U046

Site	K001 a	K002 a	K006 a	K006 b	K013 b	K020 b	K029	T001	T011	T016	T026 a	T027 a	T027 b	T033 a
K001a	*													
K002a	*0.072*	*												
K006a	0.247	0.217	*											
K006b	0.348	0.353	0.367	*										
K013b	0.290	0.262	0.285	0.266	*									
K020b	*0.009*	*-0.006*	0.126	0.347	0.245	*								
K029	*0.084*	*-0.020*	0.281	0.387	0.304	*-0.008*	*							
T001	0.309	0.288	0.270	0.303	0.046	0.269	0.307	*						
T011	0.317	0.302	0.301	0.253	0.139	0.297	0.330	*0.111*	*					
T016	0.497	0.490	0.433	0.383	0.170	0.414	0.520	*0.023*	0.184	*				
T026a	0.614	0.634	0.561	0.449	0.412	0.527	0.670	0.297	0.333	0.253	*			
T027a	0.283	0.262	0.285	0.225	0.150	0.261	0.294	0.126	*0.105*	0.178	0.265	*		
T027b	0.357	0.352	0.355	0.097	0.262	0.348	0.387	0.269	0.209	0.345	0.414	0.183	*	
T033a	0.493	0.481	0.442	0.206	0.318	0.452	0.524	0.290	0.215	0.399	0.473	*0.011*	0.153	*
T033b	0.414	0.392	0.353	0.225	0.289	0.374	0.424	0.300	0.271	0.405	0.489	0.197	0.191	*0.173*
T036a	0.304	0.255	0.278	0.321	0.263	0.242	0.325	0.282	0.288	0.444	0.573	0.215	0.309	0.399
T040	0.481	0.459	0.403	0.306	0.075	0.399	0.487	*0.007*	0.161	0.178	0.510	0.192	0.273	0.369
T064a	0.305	0.300	0.262	0.189	0.145	0.282	0.331	*0.127*	0.172	0.246	0.385	0.145	0.197	0.212
U005	0.380	0.366	0.380	0.197	0.267	0.361	0.404	0.257	0.217	0.362	0.435	*0.058*	0.159	*0.022*
U012	0.822	0.828	0.786	0.406	0.674	0.781	0.854	0.648	0.513	0.776	0.816	0.414	0.331	0.312
U020	0.800	0.802	0.772	0.408	0.677	0.758	0.830	0.665	0.549	0.784	0.818	0.453	0.341	0.399
U021	0.867	0.883	0.809	0.564	0.698	0.786	0.908	0.649	0.571	0.788	0.792	0.565	0.538	0.680
U023a	0.547	0.546	0.540	0.214	0.453	0.532	0.583	0.453	0.337	0.543	0.584	0.227	0.206	0.119
U023b	0.554	0.553	0.550	0.237	0.467	0.540	0.588	0.459	0.346	0.552	0.588	0.230	0.215	0.132
U028	0.675	0.676	0.663	0.298	0.571	0.648	0.709	0.569	0.458	0.668	0.702	0.321	0.247	*0.191*
U030b	0.306	0.332	0.330	0.146	0.185	0.315	0.361	0.195	0.207	0.308	0.404	0.141	0.144	0.159
U030c	0.497	0.426	0.409	0.397	0.357	0.358	0.521	0.382	0.370	0.614	0.752	0.319	0.386	0.533
U037	0.212	0.220	0.275	0.215	0.127	0.205	0.251	0.174	0.192	0.311	0.459	*0.082*	0.214	0.197
U046	0.670	0.669	0.656	0.332	0.570	0.643	0.701	0.559	0.454	0.664	0.699	0.321	0.280	0.202

Site	T033 b	T036 a	T040	T064 a	U005	U012	U020	U021	U023 a	U023 b	U028	U030 b	U030 c	U037
T033b	*													
T036a	0.345	*												
T040	0.327	0.366	*											
T064a	0.223	0.247	*0.154*	*										
U005	0.163	0.308	0.321	0.184	*									
U012	0.498	0.734	0.741	0.497	0.315	*								
U020	0.487	0.710	0.749	0.543	0.352	0.590	*							
U021	0.644	0.781	0.806	0.602	0.606	0.899	0.892	*						
U023a	0.277	0.484	0.498	0.342	0.104	0.247	0.327	0.683	*					
U023b	0.309	0.496	0.515	0.366	0.127	0.229	0.296	0.683	*-0.012*	*				
U028	0.357	0.605	0.632	0.435	0.127	0.295	0.378	0.789	0.076	0.072	*			
U030b	0.164	0.290	0.223	0.165	0.141	0.437	0.434	0.563	0.247	0.261	0.342	*		
U030c	0.435	*0.102*	0.529	0.331	0.409	0.851	0.827	0.921	0.581	0.589	0.709	0.386	*	
U037	0.223	0.186	0.237	0.158	0.146	0.591	0.573	0.687	0.338	0.356	0.450	0.123	0.252	*
U046	0.374	0.598	0.626	0.446	0.123	0.331	0.403	0.785	0.088	0.066	*0.059*	0.342	0.700	0.440

Table 13.19 - F_{ST} pairwise matrix of COI sequence data for 29 populations of *Biomphalaria*

Non significant values are in italics

Site	K001a	K002a	K006a	K006b	K013b	K020b	K029	T001	T011	T016	T026a	T027a	T027b	T033a	T033b
K001a	*														
K002a	*-0.018*	*													
K006a	0.220	0.143	*												
K006b	0.472	0.404	0.418	*											
K013b	0.419	0.315	0.273	0.339	*										
K020b	*0.007*	*0.004*	*0.165*	0.441	0.329	*									
K029	*0.052*	*0.031*	0.360	0.448	0.425	*-0.022*	*								
T001	0.375	0.267	0.227	0.244	*0.028*	0.287	0.372	*							
T011	0.491	0.389	0.401	0.260	0.199	0.428	0.483	0.133	*						
T016	0.535	0.402	0.382	0.372	0.153	0.453	0.557	*0.052*	0.291	*					
T026a	0.523	0.359	0.316	0.404	0.186	0.446	0.583	0.110	0.347	*0.173*	*				
T027a	0.349	0.240	0.259	0.206	0.175	0.298	0.352	0.100	0.209	0.218	0.191	*			
T027b	0.535	0.465	0.491	0.118	0.427	0.506	0.512	0.329	0.346	0.462	0.495	0.268	*		
T033a	0.555	0.455	0.533	0.234	0.429	0.523	0.548	0.337	0.389	0.502	0.559	0.151	0.223	*	
T033b	0.420	0.317	0.318	0.196	0.153	0.351	0.402	*0.051*	0.164	0.197	0.232	*0.054*	0.264	0.218	*
T036a	0.290	0.185	0.180	0.332	0.215	0.277	0.347	0.168	0.288	0.243	0.197	0.130	0.404	0.365	0.246
T040	0.439	0.319	0.281	0.307	*-0.002*	0.348	0.455	*-0.009*	0.177	0.126	0.151	0.153	0.397	0.425	0.134
T064a	0.380	0.282	0.274	0.189	0.150	0.309	0.355	*0.088*	0.201	0.220	0.221	*0.080*	0.280	0.291	0.089
U005	0.476	0.398	0.433	0.158	0.342	0.436	0.450	0.230	0.322	0.362	0.428	0.169	0.212	0.127	0.127
U012	0.656	0.585	0.632	0.291	0.565	0.631	0.640	0.469	0.525	0.561	0.637	0.376	0.313	0.331	0.378
U020	0.814	0.728	0.807	0.397	0.719	0.799	0.822	0.639	0.679	0.757	0.786	0.501	0.428	0.558	0.522
U021	0.716	0.520	0.632	0.448	0.399	0.667	0.770	0.309	0.472	0.410	0.252	0.282	0.554	0.668	0.347
U023a	0.503	0.436	0.467	0.113	0.406	0.475	0.477	0.315	0.338	0.426	0.468	0.207	0.145	0.145	0.212
U023b	0.535	0.467	0.513	0.179	0.448	0.511	0.510	0.370	0.402	0.488	0.529	0.233	0.214	0.114	0.263
U028	0.567	0.485	0.536	0.125	0.443	0.534	0.542	0.357	0.379	0.497	0.533	0.216	0.132	0.132	0.223
U030b	0.416	0.357	0.360	0.189	0.236	0.387	0.410	0.203	0.250	0.345	0.374	0.230	0.215	0.266	0.216
U030c	0.381	0.260	0.279	0.313	0.159	0.290	0.367	0.101	0.230	0.232	0.261	0.138	0.393	0.387	0.183
U037	0.387	0.307	0.318	0.191	0.163	0.331	0.384	0.110	0.163	0.277	0.322	0.114	0.242	0.213	0.137
U046	0.526	0.449	0.499	0.148	0.439	0.505	0.507	0.351	0.391	0.480	0.509	0.229	0.164	0.169	0.262

Site	T036a	T040	T064a	U005	U012	U020	U021	U023a	U023b	U028	U030b	U030c	U037	U046
T036a	*													
T040	0.187	*												
T064a	0.220	*0.138*	*											
U005	0.336	0.330	0.211	*										
U012	0.511	0.555	0.409	0.191	*									
U020	0.636	0.727	0.541	0.396	0.498	*								
U021	0.354	0.395	0.315	0.517	0.698	0.845	*							
U023a	0.371	0.388	0.245	0.096	0.188	0.317	0.525	*						
U023b	0.410	0.439	0.291	*0.058*	0.212	0.390	0.587	*0.039*	*					
U028	0.426	0.435	0.267	0.064	0.248	0.360	0.602	*0.040*	*0.019*	*				
U030b	0.277	0.195	0.213	0.227	0.402	0.545	0.481	0.264	0.282	0.269	*			
U030c	0.083	0.146	0.173	0.318	0.530	0.682	0.459	0.358	0.409	0.412	0.281	*		
U037	0.205	0.106	0.165	0.184	0.411	0.576	0.488	0.229	0.258	0.230	*0.080*	0.159	*	
U046	0.384	0.421	0.264	0.071	0.217	0.269	0.570	*0.050*	*0.036*	*0.023*	0.247	0.404	0.227	*

Table 13.20 - F_{ST} pairwise matrix of 16S sequence data for 29 populations of *Biomphalaria*

Non significant values are in italics

Site	K001 a	K002 a	K006 a	K006 b	K013 b	K020 b	T001	T011	T016	T026 a	T027 a	T027 b
K001a	*											
K002a	*0.028*	*										
K006a	*0.017*	*-0.017*	*									
K006b	0.127	*0.045*	*0.109*	*								
K013b	0.124	0.100	*0.090*	0.174	*							
K020b	0.060	0.042	*0.010*	0.116	0.160	*						
T001	*0.073*	*0.076*	*0.078*	0.183	*0.175*	0.138	*					
T011	0.073	0.104	*0.162*	0.203	0.194	0.126	*0.085*	*				
T016	0.088	0.081	*0.089*	0.197	0.134	0.148	*0.041*	0.132	*			
T026a	0.092	0.131	*0.226*	0.223	0.232	0.164	*0.135*	*0.170*	0.120	*		
T027a	0.059	0.050	*0.025*	0.090	0.107	0.099	0.102	0.179	0.091	0.135	*	
T027b	0.125	0.132	*0.143*	0.095	0.188	0.164	0.196	0.203	0.204	0.255	0.108	*
T033a	0.078	0.051	*0.086*	0.085	0.156	0.076	*0.142*	0.210	0.155	0.192	0.075	0.116
T033b	0.124	0.137	*0.113*	*0.079*	0.189	0.150	*0.145*	0.147	0.182	0.236	0.118	*0.021*
T036a	0.109	0.084	*0.138*	0.155	0.119	0.131	0.173	0.124	0.155	0.183	0.135	0.106
T040	0.103	0.057	*-0.009*	*0.089*	0.115	0.128	0.113	0.204	0.091	0.159	*0.010*	0.108
T064a	0.059	0.038	*0.098*	0.129	0.137	0.070	*0.099*	*0.040*	0.110	0.177	0.132	0.153
U005	0.047	0.043	*0.057*	0.152	0.141	0.095	*0.088*	0.054	0.084	0.080	0.106	0.171
U020	*0.044*	0.108	*0.184*	0.153	0.165	0.139	0.139	0.149	0.157	0.120	0.095	0.194
U021	*0.037*	0.112	*0.127*	0.163	0.162	0.137	0.143	0.148	0.182	0.187	0.124	0.169
U023a	0.087	0.079	*0.013*	*0.069*	0.079	0.125	0.147	0.190	0.166	0.217	0.078	0.078
U023b	*0.040*	0.061	*-0.022*	*0.060*	0.114	0.093	*0.118*	0.155	0.128	0.152	0.050	0.043
U028	0.044	0.033	*0.019*	0.068	0.096	0.075	*0.084*	0.110	0.082	0.141	0.050	0.080
U030b	*0.092*	0.059	*-0.018*	*0.072*	*0.189*	0.080	*0.178*	*0.182*	0.157	0.245	0.106	0.112
U037	0.043	0.019	*0.049*	0.114	0.121	0.045	*0.069*	0.091	0.114	0.162	0.096	0.153
U046	0.038	0.018	*0.059*	0.123	0.132	0.082	0.106	0.072	0.099	0.131	0.100	0.136

Site	T033 a	T033 b	T036 a	T040	T064 a	U005	U020	U021	U023 a	U023 b	U028	U030 b	U037	U046
T033a	*													
T033b	0.144	*												
T036a	0.133	0.156	*											
T040	0.076	0.117	0.133	*										
T064a	0.104	0.117	0.096	0.112	*									
U005	0.084	0.143	0.092	0.096	0.047	*								
U020	0.149	0.185	0.167	0.148	0.111	0.136	*							
U021	0.160	0.155	0.177	0.156	0.122	0.131	0.092	*						
U023a	0.101	*0.026*	0.145	0.070	0.083	0.123	0.144	0.127	*					
U023b	*0.060*	*0.034*	0.101	*0.047*	0.100	0.086	0.123	0.099	*-0.004*	*				
U028	*0.039*	0.058	0.077	0.056	0.059	0.041	0.137	0.125	*0.036*	*0.017*	*			
U030b	*0.093*	*0.073*	0.162	0.109	*0.053*	*0.090*	*0.162*	0.143	*0.023*	*-0.000*	0.071	*		
U037	0.057	0.122	0.109	0.111	*0.019*	0.038	0.143	0.119	0.074	0.063	*0.022*	*0.055*	*	
U046	0.083	0.121	0.099	0.103	*0.014*	*0.033*	0.134	0.125	0.065	0.061	0.049	*0.062*	*0.015*	*

Table 13.21 - F_{ST} pairwise matrix of microsatellite genotype data for 28 populations of *Biomphalaria*

Non significant values are in italics

Site	General information					Standard and molecular diversity indices							Mismatch analyses		Neutrality
	Shedding	*Susceptible*	*# Seqs*	*# Haps*	*# Polymorphic sites*	*Gene div.*	*Error*	*Nucleotide div.*	*Error*	*Theta k*	*Lower CI*	*Upper CI*	*Sum of squares p*	*Raggedness p*	*Tajima's p*
K001a	N	N	10	4	5	0.7333	0.1005	0.00258	0.00186	1.956	0.5982	6.1275	0.120	0.200	0.430
K002a	N	N	10	6	4	0.7778	0.1374	0.00224	0.00167	5.408	1.7798	16.628	0.250	0.410	0.630
K006a	N	N	10	4	7	0.7778	0.0907	0.00441	0.00286	1.956	0.5982	6.1275	0.120	0.330	0.778
K006b	N	N	10	7	55	0.9111	0.0773	0.02565	0.01413	9.023	2.9077	29.191	0.040	0.070	0.269
K013b	N	N	11	7	16	0.9273	0.0543	0.00744	0.00443	7.237	2.4762	21.639	0.070	0.010	0.341
K020b	N	N	11	6	13	0.8364	0.0887	0.00461	0.00293	4.627	1.5813	13.507	0.840	0.960	0.077
K029	N	N	12	4	3	0.7121	0.1053	0.00146	0.0012	1.672	0.5258	5.0177	0.320	0.300	0.435
T001	Y	Y	11	5	28	0.6182	0.1643	0.00799	0.00472	2.944	0.9756	8.6492	0.000	1.000	0.000*
T011	N	N	11	6	26	0.8545	0.0852	0.01377	0.00775	4.627	1.5813	13.507	0.590	0.580	0.564
T016	N	N	11	3	4	0.7273	0.0679	0.00300	0.00207	0.986	0.2732	3.3049	0.160	0.290	0.957
T026a	N	Y	11	4	5	0.4909	0.1754	0.00205	0.00155	1.795	0.5576	5.4902	0.000	1.000	0.233
T027a	N	Y	12	6	17	0.8636	0.0639	0.01087	0.00618	4.101	1.4398	11.521	0.060	0.160	0.915
T027b	N	N	10	6	44	0.8889	0.0754	0.02650	0.01457	5.408	1.7798	16.628	0.000	0.010	0.758
T033a	N	Y	9	3	13	0.6389	0.1258	0.01043	0.00615	1.138	0.3083	3.9476	0.010	0.000	0.991
T033b	Y	Y	11	7	24	0.9091	0.0656	0.01432	0.00804	7.237	2.4762	21.639	0.090	0.470	0.760
T036a	N	N	10	5	14	0.7556	0.1295	0.00618	0.00381	3.302	1.0677	10.048	0.460	0.400	0.205
T040	N	N	10	8	10	0.9333	0.0773	0.00421	0.00275	16.397	4.8904	59.01	0.930	0.760	0.186
T064a	N	N	10	6	22	0.8444	0.1029	0.01466	0.00831	5.408	1.7798	16.628	0.270	0.460	0.902
U005	N	N	11	7	34	0.8182	0.1191	0.01482	0.00831	7.237	2.4762	21.639	0.130	0.220	0.229
U012	N	N	10	4	12	0.6444	0.1518	0.00390	0.00258	1.956	0.5982	6.1275	0.140	0.650	0.016
U020	N	N	11	3	13	0.3455	0.1722	0.00428	0.00276	0.986	0.2732	3.3049	0.070	0.520	0.036
U021	N	Y	10	2	1	0.2000	0.1541	0.00030	0.00046	0.430	0.0976	1.7949	0.160	0.290	0.196
U023a	N	N	10	9	34	0.9778	0.0540	0.01652	0.00930	38.778	9.4026	169.08	0.360	0.320	0.329
U023b	Y	Y	11	10	43	0.9818	0.0463	0.01607	0.00896	48.112	11.7836	208.66	0.080	0.030	0.081
U028	Y	Y	11	8	27	0.9273	0.0665	0.00943	0.00548	11.688	3.8705	37.12	0.350	0.620	0.045
U030b	N	N	11	9	43	0.9636	0.0510	0.01641	0.00914	20.730	6.2965	73.767	0.620	0.580	0.119
U030c	N	N	11	3	3	0.4727	0.1617	0.00167	0.00133	0.986	0.2732	3.3049	0.000	0.950	0.647
U037	Y	Y	11	9	21	0.9636	0.0510	0.00902	0.00526	20.730	6.2965	73.767	0.480	0.800	0.229
U046	N	N	11	5	30	0.7818	0.1073	0.00985	0.00570	2.944	0.9756	8.6492	0.270	0.380	0.031

Table 13.22 – Summary table of molecular diversity indices, mismatch analysis and neutrality testing for COI sequence data

Site	General information					Standard and molecular diversity indices							Mismatch analyses		Neutrality
	Shedding	*Susceptible*	*# Seqs*	*# Haps*	*# Polymorphic sites*	*Gene div.*	*Error*	*Nucleotide div.*	*Error*	*Theta k*	*Lower CI*	*Upper CI*	*Sum of squares p*	*Raggedness p*	*Tajima's p*
K001a	N	N	11	4	5	0.7455	0.0978	0.00500	0.00340	1.795	0.5576	5.4903	0.020	0.020	0.979
K002a	N	N	10	9	16	0.9778	0.0540	0.01075	0.00652	38.778	9.4026	169.0841	0.350	0.790	0.450
K006a	N	N	9	2	4	0.5000	0.1283	0.00474	0.00334	0.455	0.1028	1.9183	0.000	0.900	0.892
K006b	N	N	9	9	55	1.0000	0.0524	0.04272	0.02376	NA	NA	NA	0.860	0.930	0.408
K013b	N	N	12	7	19	0.8788	0.0751	0.01186	0.00697	6.155	2.1926	17.4179	0.260	0.340	0.034
K020b	N	N	11	6	8	0.8000	0.1138	0.00576	0.00381	4.627	1.5813	13.5073	0.130	0.090	0.647
K029	N	N	9	4	6	0.5833	0.1833	0.00447	0.00318	2.177	0.6518	7.0364	0.340	0.320	0.699
T001	Y	Y	10	7	32	0.8667	0.1072	0.01745	0.01008	9.023	2.9077	29.1905	0.010	0.910	0.002
T011	N	N	10	8	23	0.9333	0.0773	0.01630	0.00948	16.397	4.8904	59.0103	0.620	0.700	0.265
T016	N	N	11	3	7	0.7273	0.0679	0.00837	0.00519	0.986	0.2732	3.3049	0.050	0.020	1.000
T026a	N	Y	11	6	7	0.8909	0.0633	0.00521	0.00351	4.627	1.5813	13.5073	0.240	0.800	1.000
T027a	N	Y	13	8	20	0.9231	0.0500	0.01815	0.01017	7.883	2.9015	21.7809	0.140	0.170	0.604
T027b	N	N	10	8	35	0.9556	0.0594	0.03315	0.01839	16.397	4.8904	59.0103	0.380	0.460	0.716
T033a	N	Y	10	4	17	0.7111	0.1175	0.01436	0.00844	1.956	0.5982	6.1275	0.080	0.090	0.675
T033b	Y	Y	12	12	30	1.0000	0.0340	0.02086	0.01164	NA	NA	NA	0.490	0.360	0.362
T036a	N	N	11	7	31	0.8182	0.1191	0.01633	0.00937	7.237	2.4762	21.6389	0.420	0.410	0.034
T040	N	N	10	9	15	0.9778	0.0540	0.01077	0.00653	38.778	9.4026	169.0841	0.670	0.650	0.305
T064a	N	N	10	6	21	0.8889	0.0754	0.02191	0.01245	5.408	1.7798	16.6277	0.330	0.790	0.815
U005	N	N	11	10	33	0.9818	0.0463	0.02688	0.01489	48.112	11.7836	208.6648	0.620	0.540	0.297
U012	N	N	10	7	21	0.9111	0.0773	0.01964	0.01124	9.023	2.9077	29.1905	0.310	0.120	0.438
U020	N	N	9	4	13	0.5833	0.1833	0.00721	0.00469	2.177	0.6518	7.0364	0.340	0.700	0.116
U021	N	Y	10	3	4	0.3778	0.1813	0.00189	0.00168	1.052	0.2887	3.5803	0.400	0.620	0.087
U023a	N	N	10	9	45	0.9778	0.0540	0.03338	0.01849	42.175	42.1751	42.1751	0.350	0.230	0.659
U023b	Y	Y	11	11	49	1.0000	0.0388	0.02810	0.01552	NA	NA	NA	0.810	0.770	0.034
U028	Y	Y	10	10	35	1.0000	0.0447	0.02479	0.01396	NA	NA	NA	0.610	0.520	0.152
U030b	N	N	11	10	40	0.9818	0.0463	0.02822	0.01560	48.112	11.7836	208.6648	0.720	0.960	0.153
U030c	N	N	10	6	16	0.8444	0.1029	0.01427	0.00840	5.408	1.7798	16.6277	0.160	0.300	0.471
U037	Y	Y	10	9	24	0.9778	0.0540	0.01894	0.01087	38.778	9.4026	169.0841	0.630	0.630	0.376
U046	N	N	9	9	41	1.0000	0.0524	0.03002	0.01696	NA	NA	NA	0.770	0.780	0.140

Table 13.23 – Summary table of molecular diversity indices, mismatch analysis and neutrality testing for 16S sequence data

Site	General information					Standard and molecular diversity indices							Mismatch analyses		Neutrality
	Shedding	*Susceptible*	*# Seqs*	*# Haps*	*# Polymorphic sites*	*Gene div.*	*Error*	*Nucleotide div.*	*Error*	*Theta k*	*Lower CI*	*Upper CI*	*Sum of squares p*	*Raggedness p*	*Tajima's p*
K001a	N	N	11	4	5	0.7455	0.0978	0.00500	0.00340	1.795	0.5576	5.4903	0.020	0.020	0.979
K002a	N	N	10	9	16	0.9778	0.0540	0.01075	0.00652	38.778	9.4026	169.0841	0.350	0.790	0.450
K006a	N	N	9	2	4	0.5000	0.1283	0.00474	0.00334	0.455	0.1028	1.9183	0.000	0.900	0.892
K006b	N	N	9	9	55	1.0000	0.0524	0.04272	0.02376	NA	NA	NA	0.860	0.930	0.408
K013b	N	N	12	7	19	0.8788	0.0751	0.01186	0.00697	6.155	2.1926	17.4179	0.260	0.340	0.034
K020b	N	N	11	6	8	0.8000	0.1138	0.00576	0.00381	4.627	1.5813	13.5073	0.130	0.090	0.647
K029	N	N	9	4	6	0.5833	0.1833	0.00447	0.00318	2.177	0.6518	7.0364	0.340	0.320	0.699
T001	Y	Y	10	7	32	0.8667	0.1072	0.01745	0.01008	9.023	2.9077	29.1905	0.010	0.910	0.002
T011	N	N	10	8	23	0.9333	0.0773	0.01630	0.00948	16.397	4.8904	59.0103	0.620	0.700	0.265
T016	N	N	11	3	7	0.7273	0.0679	0.00837	0.00519	0.986	0.2732	3.3049	0.050	0.020	1.000
T026a	N	Y	11	6	7	0.8909	0.0633	0.00521	0.00351	4.627	1.5813	13.5073	0.240	0.800	1.000
T027a	N	Y	13	8	20	0.9231	0.0500	0.01815	0.01017	7.883	2.9015	21.7809	0.140	0.170	0.604
T027b	N	N	10	8	35	0.9556	0.0594	0.03315	0.01839	16.397	4.8904	59.0103	0.380	0.460	0.716
T033a	N	Y	10	4	17	0.7111	0.1175	0.01436	0.00844	1.956	0.5982	6.1275	0.080	0.090	0.675
T033b	Y	Y	12	12	30	1.0000	0.0340	0.02086	0.01164	NA	NA	NA	0.490	0.360	0.362
T036a	N	N	11	7	31	0.8182	0.1191	0.01633	0.00937	7.237	2.4762	21.6389	0.420	0.410	0.034
T040	N	N	10	9	15	0.9778	0.0540	0.01077	0.00653	38.778	9.4026	169.0841	0.670	0.650	0.305
T064a	N	N	10	6	21	0.8889	0.0754	0.02191	0.01245	5.408	1.7798	16.6277	0.330	0.790	0.815
U005	N	N	11	10	33	0.9818	0.0463	0.02688	0.01489	48.112	11.7836	208.6648	0.620	0.540	0.297
U012	N	N	10	7	21	0.9111	0.0773	0.01964	0.01124	9.023	2.9077	29.1905	0.310	0.120	0.438
U020	N	N	9	4	13	0.5833	0.1833	0.00721	0.00469	2.177	0.6518	7.0364	0.340	0.700	0.116
U021	N	Y	10	3	4	0.3778	0.1813	0.00189	0.00168	1.052	0.2887	3.5803	0.400	0.620	0.087
U023a	N	N	10	9	45	0.9778	0.0540	0.03338	0.01849	42.175	42.1751	42.1751	0.350	0.230	0.659
U023b	Y	Y	11	11	49	1.0000	0.0388	0.02810	0.01552	NA	NA	NA	0.810	0.770	0.034
U028	Y	Y	10	10	35	1.0000	0.0447	0.02479	0.01396	NA	NA	NA	0.610	0.520	0.152
U030b	N	N	11	10	40	0.9818	0.0463	0.02822	0.01560	48.112	11.7836	208.6648	0.720	0.960	0.153
U030c	N	N	10	6	16	0.8444	0.1029	0.01427	0.00840	5.408	1.7798	16.6277	0.160	0.300	0.471
U037	Y	Y	10	9	24	0.9778	0.0540	0.01894	0.01087	38.778	9.4026	169.0841	0.630	0.630	0.376
U046	N	N	9	9	41	1.0000	0.0524	0.03002	0.01696	NA	NA	NA	0.770	0.780	0.140

Table 13.24 – Summary table of molecular diversity indices, mismatch analysis and neutrality testing for microsatellite allele data