University of California

Los Angeles

Studies on the Ecology and Adaptation of *Anopheles gambiae* in Mali and Their Impacts on Malaria Transmission and Control

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Biology

by

Nicholas Chirivas Manoukis

2006

The dissertation of Nicholas Chirivas Manoukis is approved.

Victoria L. Sork

Henry A. Hespenheide

Glen M. MacDonald

Yongkang Xue

Charles E. Taylor, Committee Chair

University of California, Los Angeles

2006

To my Valeria

Table of Contents

List of Figures

List of Tables

Acknowledgments

This work would not have been possible without the intellectual investment and unflinching support of Charles Taylor. I owe its successes to his generosity and vision.

I must also thank my many colleagues at UCLA and at the Malaria Research Center (MRTC) in Bamako, Mali. I thank Mahamoudou Touré for his friendship and the many hours of hard work we put in together. Thanks to my lab mates Maria Diuk-Wasser, Travis Collier, Yoosook Lee, Sigrid Rian, Saul Lozano-Fuentes and John Marshall among others gave me the warmth of friendship during these years. I also thank David Jacobs for the first two years I spent at UCLA.

Others associated with the MRTC and who generously aided me through five visits since 2002 are Sekou Traoré, Seydou Doumbia, Richard Sakai and Robert Gwadz. Thank you also to Ibrahim Sissoko and Adama Dao for a great collaboration and Sibiry Samake and Boubacar Guindo for assistance in the field.

The people of Niono and Banambani were generous and welcoming of me and my research. I thank the many field workers and especially Kefa Diarra for many hours in the hot sun collecting larvae and adult mosquitoes. I hope that the work we did will indeed lead to better control of malaria.

Thank you to my committee, Victoria Sork, Henry Hespenheide, Glen MacDonald and Yongkang Xue for the time they spent with me discussing strategy and the projects in this dissertation. I also thank Sally Blower for the many conversations with her and laboratory members, furthering my understanding of epidemiology and modeling.

Most of the work in this dissertation has already been published and is the re-

sult of collaborations with others. Chapter 2 is a version of Maria A. Diuk-Wasser, Mahamoudou B. Touré, Guimogo Dolo, Magaran Bagayoko, Nafomon Sogoba, Sekou F. Traoré, Nicholas C. Manoukis and Charles E. Taylor. 2005. Vector abundance and malaria transmission in rice-growing villages in Mali. *American Journal of Tropical Medicine and Hygiene* 16:725 – 731. The field work was conducted by the MRTC team with the analysis and writing conducted by those at UCLA including myself.

Chapter 3 has been published in slightly different form in: Nicholas C. Manoukis, Mahamoudou B. Touré, Ibrahim Sissoko, Seydou Doumbia, Sekou F. Traoré, Maria A. Diuk-Wasser, and Charles E. Taylor. 2006. Is Vector Body Size the Key to Reduced Malaria Transmission in the Irrigated Region of Niono, Mali? *Journal of Medical Entomology*, 43:820 – 827. Dr. Touré and I led the field work and I was primarily responsible for the analysis and writing.

Chapter 4 is a revision and extension of a chapter I wrote with Charles Taylor: Charles E. Taylor and Nicholas C. Manoukis. 2003. Effective Population Size in Relation to Genetic Modification of *Anopheles gambiae* s.s. *Ecological Aspects for Application of Genetically Modified Mosquitoes*, Chapter 10. Tom W. Scott and Willem Takken, Editors. Frontis, Wageningen, Netherlands. Dr. Taylor and I worked together on the writing and I produced the visualization for his simulation results.

The results in Chapter 5 are in preparation. It is the product of a collaboration between an MRTC field work team led by Mahamoudou Touré and members of Gregory Lanzaro's laboratory at UC Davis. I thank Allan Rae and Tara Thiemann for their work on the molecular data and Greg Lanzaro for supporting it. I am responsible for the analysis and writing of the material in that chapter.

The work in Chapter 6 has been submitted for publication to *Journal of*

Evolutionary Biology, and is the result of a collaboration between myself, Travis Collier and Charles Taylor. I wrote most of the computer code, conducted the analysis and been primarily responsible for the writing of the material in that chapter. I gratefully acknowledge the help of Mario Coluzzi for guidance and insight based on years of careful observation and study. I also thank Jesus K. Estrada for work on the parallelization program and Barbara Coluzzi for useful discussions on computational issues.

I acknowledge the collaboration of the Niono Health Center, the Office du Niger, the Institute d'Economie Rurale in Niono. Much of the work in this thesis was supported by NIH grant 5RO1AI51633

Finally I thank the members of my family for giving me the strength to complete this project. To my mother, Chrysanthi, for love and pushing; to my brother Dimitrios, for all the hours of conversation. Thanks to Tim and Nancy Grab for all the warmth, meals and help. To my son Demos: you fill me with joy. Most of all thank you to my wife Valerie for allowing me to chase my dreams, even when they consisted of cutting wings off of mosquitoes. You give me measure of all things.

Vita

1974	Born, New York NY
1990–1992	International Baccalaureate Diploma, The British School, Rio de Janeiro, Brazil
1995	Recipient of Reed College Independent Fieldwork Grant
1995–1997	Head lab assistant, Kaplan Lab (Reed College, Portland OR)
1993–1997	B.A. Biology, Reed College, Portland OR
1997–1998	Lab Assistant for Life Science 1 course, (University of Southern California, Los Angeles CA)
1998–2000	Fifth grade teacher, Garfield Elementary School (Long Beach Unified School District, Long Beach CA)
2000–2001	Recipient of California Genetic Resources Conservation Program Grant
2000-2005	Teaching Assistant, Department of Ecology and Evolutionary Biology, UCLA. Courses: Life Sciences 1 (2000), Evolution (2001), Computational Biology (2005)
2003–2004	Recipient of Systems and Integrative Biology Training Grant (NIH)
2001–2006	Graduate Student Researcher for Dr. Charles Taylor

Publications

Maria A. Diuk-Wasser and Guimogo Dolo and Magaran Bagayoko and Nafomon Sogoba and Mahamoudou B. Touré and M. Moghaddam and Nicholas C. Manoukis and Sigrid Rian and Sekou F. Traoré and Charles E. Taylor. 2006. Patterns of irrigated rice growth and malaria vector breeding in Mali using multi-temporal ERS-2 synthetic aperture radar. *International Journal of Remote Sensing* 27:535 – 548.

Maria A. Diuk-Wasser, Mahamoudou B. Touré, Guimogo Dolo, Magaran Bagayoko, Nafomon Sogoba, Sekou F. Traoré, Nicholas C. Manoukis and Charles E. Taylor. 2005. Vector abundance and malaria transmission in rice-growing villages in Mali. *American Journal of Tropical Medicine and Hygiene* 16:725 – 731.

Nicholas C. Manoukis and David K. Jacobs. 2001. Conservation of the California Tree Frog, *Hyla cadaverina*, from desert oasis areas in Joshua Tree National Park. *Technical Report to the Genetic Resources Conservation Program, U.C. Davis.*

Nicholas C. Manoukis, Yoosook Lee, Edgar Vallejo, and Charles E. Taylor. 2004. Detecting recurrent extinction in a meta-populations of Anopheles gambiae: preliminary results using simulation. *Proceedings of the 6th WSEAS International Conference on Algorithms, Scientific Computing, Modeling and Simulation.*

Nicholas C. Manoukis, Mahamoudou B. Touré, Ibrahim Sissoko, Seydou Doumbia, Sekou F. Traoré, Maria A. Diuk-Wasser, and Charles E. Taylor. 2006. Is Vector Body Size the Key to Reduced Malaria Transmission in the Irrigated Region of Niono, Mali? *Journal of Medical Entomology*, 43:820 – 827.

Charles E. Taylor and Nicholas C. Manoukis. 2003. Effective Population Size in Relation to Genetic Modification of *Anopheles gambiae* s.s. *Ecological Aspects for Application of Genetically Modified Mosquitoes*, Chapter 10. Tom W. Scott and Willem Takken, Editors. Frontis, Wageningen, Netherlands.

Abstract of the Dissertation

Studies on the Ecology and Adaptation of *Anopheles gambiae* in Mali and Their Impacts on Malaria Transmission and Control

by

Nicholas Chirivas Manoukis

Doctor of Philosophy in Biology

University of California, Los Angeles, 2006

Professor Charles E. Taylor, Chair

Anopheles gambiae ecology leads to particular patterns of population density and structure in Mali which have significant impacts on species evolution and prospects of malaria control through genetically modified vectors. How density affects transmission was studied in Niono, Mali, where villages in irrigated areas usually have more anopheline vectors than adjacent non-irrigated villages, but overall malaria prevalence is lower. An analysis of correlations between adult mosquito density and survival, anthropophily and vectorial capacity using field collected adults within the irrigated area showed a decrease in anthropophily and survival with adult density, supporting the "competition hypothesis". This hypothesis holds that high larval densities lead to smaller mosquitoes, which suffer elevated mortality, leading them to be less efficient vectors. Further field experiments showed a modest positive relationship between densities of immatures and larval size and a strong relationship between larval and adult size. Adult survivorship was found to be higher in non-irrigated areas; However, there was no effect of size on survivorship between comparable samples, rejecting the

competition hypothesis. Studies in Banambani, Mali, focused on the population structure of the *An. gambiae* complex. An analysis of bi-weekly adult samples through molecular and demographic techniques revealed a succession of species with changes in density due to climate, and a severe reduction in population size for some types. The bi-weekly samples point to an important role of annual extinction and metapopulation dynamics to malaria control. Simulation models for studying this question indicated that structure can qualitatively affect the outcome of malaria control schemes using a transposable element to drive a gene for refractoriness into the vector population. Finally, a mechanism of adaptation and speciation dependent on structure, called "ecotypification", was examined through simulations. Results indicated that ecotypification is a viable explanation for chromosomal inversion polymorphism, a precursor to speciation in this taxon. Inversion polymorphism was found to be dependent on population size, structure and niche differentiation. Overall, simulated and empirical data presented strongly suggest that the density and population structure of *Anopheles gambiae* are of central importance to its evolution and to malaria control.

CHAPTER 1

Vector biology and malaria control

1.1 Introduction

This dissertation is presented in the form of five research projects I have worked on during my time at UCLA. All five extend our understanding of the biology of *Anopheles gambiae* s.s. relating to the transmission and control of malaria. This work illustrates that knowledge of the ecology and evolution of this disease vector has been and is important to attacking the disease it spreads. It will be of increasing importance as more technologically advanced interventions are contemplated.

Malaria is one of the world's greatest health challenges, with 250 – 450 million clinical cases and around one million deaths per year. 90% of the cases occur in sub Saharan Africa (Greenwood and Mutabingwa, 2002). It is difficult to overstate the burden of a lethal disease that threatens over 1/3 of the world's population (Trigg and Kondrachine, 1998) or the economic, political and social toll it takes on its poorest countries. The disease disproportionately affects pregnant women and children, who are most susceptible to its fatal effects (Roll Back Malaria Project, 2003). It is currently responsible for 23 – 37% of the total child mortality in Africa (Korenromp et al., 2003).

Scientific research targeted at reducing malaria may be divided into two classes of efforts: those focused on the parasite itself (usually while in the human part

of its life cycle) and those focused on the parasite's insect vectors. In the first is development of a vaccine to either prevent transmission or immunize humans to the *Plasmodium* parasite and the development of new chemotherapeutic and prophylactic medications. Two topics within the second area, interfering with the transmission of the parasite through mosquitoes of the genus *Anopheles*, are the focus of the work in this dissertation.

Since the time that mosquitoes were shown to transmit malaria, around 1900, humans have tried to control or eradicate the host. Some of that history is discussed in this chapter. The current understanding of the vector's ecology and evolution is provided, together with a brief description of the parasite's life cycle. The rest of the dissertation considers current and future challenges to malaria control based on the study of vectors. The last sections in this chapter describe the areas where research was conducted.

With the growing population pressures on Africa, the continent has seen a large increase in the amount of irrigated crop land to feed its human population. Estimates of irrigated land area in Africa by 2020 are of about 15.9 million ha, a 30% increase from 1990 (Rosengrant and Perez, 2000). This increase in irrigation may increase the frequency of vector-borne diseases such as malaria, making it an important public health question in sub-Saharan Africa.

Among the many environmental changes driven by irrigation, often an increase in mosquito densities is most noticeable. In our study region of Niono, Mali biting densities of mosquitoes have been measured to be ten times greater (up to 550 bites per person per night) in irrigated villages than in adjoining non-irrigated ones (Dolo et al., 2004). This increase in vectors does not always bring about increased disease in the case of malaria, however.

In fact malaria cases and transmission can increase, decrease or show no

change with irrigation (Chapter 2). The likely cause or causes of this discrepancy lie in the details of vector biology, human behavior and the parasite's life cycle. In Niono increased density brings decreases in some measures of transmission efficiency according to results presented in Chapter 2.

The exact mechanism of reduced transmission in Niono is not clear. Among several possibilities is the "competition hypothesis", a test of which is the topic of Chapter 3. The hypothesis holds that increased larval density causes resource competition in irrigated rice fields, resulting in smaller larvae. These smaller larvae become smaller adult mosquitoes, which suffer elevated mortality and so are less efficient vectors of malaria. The results in Chapter 3 test each part of this link between larval competition and adult survivorship.

Anopheles gambiae biology is also important in developing strategies to control transmission by genetically modifying the vector. One of these strategies involves using transposable element with an attached effector gene to drive refractoriness into populations. This approach is considered promising since insecticide resistance in the mosquito has threatened vector control approaches in the past, a problem that some believe can be avoided with a genetic approach (Hemingway, 2004).

Regarding genetic modification most work is concentrated on development of the genes and molecules needed to create a refractory (uninfectable) phenotype in the mosquito, unarguably a difficult problem. Much less attention is currently devoted to exactly how such genes and mobile genetic elements may spread (or not) if they are introduced to natural populations. This basic question of transposable element spread through a population of *Anopheles gambiae*, the putative target of intervention, depends on the species' ecology and its adaptation to the environment. These topics occupy the last three chapters of this dissertation.

The population structure of *An. gambiae* will have a large and important effect on the spread of a transposable element, as illustrated in Chapter 4 with simple simulations. That chapter also discusses the current understanding of structure in this species and the areas where our understanding needs to be extended.

The structure of the population is in part determined by the adaptation of the species to its environment, the topic of Chapter 5. In it evidence is presented for significant ecological differences between cryptic and incipient species in a village in Mali by examining the frequency of each type over one year. The question of population persistence during the dry season, a critical issue raised in Chapter 3, is also addressed.

The implications of adaptation extend beyond events in an ecological timescale to differentiation and speciation within the *Anopheles* complex. Simulations seeking to clarify a long standing hypothesis of speciation by chromosomal inversion based on local adaptation are presented in Chapter 6. Known as "ecotypification," this process contributes to the great environmental flexibility in *An. gambiae*. The evolutionarily short time involved serves as an additional dimension of complexity for malaria control efforts through genetic modification of the vector.

Most of the chapters in this dissertation were written for publication or have been published, so they generally contain enough information to be read on their own. As such, some of the material presented in this chapter may be repeated elsewhere.

1.1.1 Objectives and hypotheses

General hypothesis

> *Anopheles gambiae* ecology leads to particular patterns of population density and structure in Mali which have significant impacts on the species evolution and the prospects of malaria control through genetically modified vectors.

Chapter 2

> Vectorial efficiency decreases with increasing adult *Anopheles* density.

Chapter 3

- Increased *Anopheles* larval density leads to smaller larvae in the rice fields of Niono.
- Smaller larvae become smaller adult females.
- Smaller adult females survive less well than larger ones.

Chapter 4

> Population structure of *Anopheles gambiae* s.s. around Banambani has significant effects on the spread of a transposable element as a method of malaria transmission reduction.

Chapter 5

- The two molecular forms of *An. gambiae* s.s. and *An. arabiensis* in Banambani have differing ecological niches; the predominant form varies over a year.
- The main malaria vector present in the area changes over the year as a consequence of the above.
- Some forms/species are more likely to undergo annual extirpation than others.

Chapter 6

- Inversion polymorphism can result in *Anopheles* from a process of ecological specialization and chromosomal inversion ("ecotypification").
- Some demographic, environmental or genetic factors are more important to the maintenance of inversion polymorphism than others.

1.2 Vector biology and malaria: A brief history

1.2.1 Discovery

The malaria parasite has probably afflicted humans since prehistoric times. This idea is suggested by fossilized mosquitoes displaying fossilized parasites dating to the Paleozoic period (Russell, 1952). Enlarged spleens from 3,000 year old Egyptian mummies (Sherman, 1998) also indicate the presence of malaria. Until about 1700 CE no treatment was known to Europeans, at least none that differed from the treatments for other maladies. With the importation of chiconia bark from the Andean rain forests of Peru in the 1640s (see Honigsbaum 2001 for an entertaining and complete account), European doctors had their first chemotherapy for malaria. The availability of what we now know as quinine enabled some, such as Torti in Italy, to differentiate differentiated fevers alleviated by chiconia from others that did not respond.

Differentiating malaria fevers from others enabled it to be named. In English, it became known as malaria after the Italian *mal'aria* ("bad air") when Horace Walpole imported the name in 1740. He also referred to it as the "fever that comes to Rome every summer" (Russell, 1952), so that it was sometimes called the "Roman fever". In any event, the name indicates that early medicine had little idea of what caused malaria or how it spread.

An important date for research on malaria vectors to control the disease is 20 August 1897. On that day Ronald Ross, who had been trying to find evidence of the parasite in mosquitoes mostly in Secunderabad, India, found pigmented cells on the outside of the stomach of a dissected specimen (probably *An. stephensii*). As he wrote to his mentor Patrick Manson in London on 22 August 1897:

> I am so familiar with the mosquito's stomach that these bodies struck

> me at once; & you may imagine how much more struck I was when, on focusing carefully I found they contained *pigment indistinguishable in colour, shape etc from that of the haemamoeba!*" (Bynum and Overy, 1998).

Ross was not the first to hypothesize that mosquitoes could transmit disease. Raimbert had showed that anthrax could be dispersed by flies in 1869. Others had suspected malaria specifically was transmitted by mosquitoes. In 1854 Louis Baruperthuy, working in Caracas, Venezuela, had suggested that malaria was caused by living germs in mosquitoes. Alphonse Laveran identified and described *Plasmodium malariae* from human blood in 1880 while working in Algeria, and in 1884 Laveran also suggested that mosquitoes could actually transmit malaria (Harrison, 1978). They were among several others, but Manson especially mentored and encouraged Ross to pursue the question, competing with Italian, German and American researchers (Bynum and Overy, 1998).

From Ross and Manson's correspondence it is clear they knew little about the mosquito. They did not systematically distinguish the species being dissected, and so spent additional time surveying species that we now know can not transmit malaria. In fact, Ross referred to "a grey one" (probably *Culex pipiens*) and a "brindled one" (probably *Aedes aegypti*). The one that eventually yielded the parasite was "a big brown fellow". This illustrates how little was known and how little attention was given to the insect by those who were medical doctors.

The early focus by these men was entirely on the parasite. Occasionally this led to errors, such as those committed by Manson himself some 20 years earlier. Manson had shown in 1877 that mosquitoes contained filariasis-causing worms from his work in China the previous decade or so. His lack of knowledge about the mosquito prevented him from understanding how transmission actually occurs. He did not know that the species he was studying could live more than very few

days (they did not survive for long in his laboratory) or that they could have more than one feeding in their lifetimes. Manson was thus unable to follow the worm's full development to the mosquito's proboscis, so the actual mode of transmission of filariasis was not correctly elucidated in his work (Harrison, 1978).

The emphasis of Manson and Ross' approach does not diminish the importance of their work, but it does clearly illustrate that the biology of the mosquito is of much consequence to comprehending the dynamics of malaria. It should also be noted that not all researchers were inclined as Ross and Manson. Giovanni Grassi was the preeminent medical zoologist of the time and considered their lack of attention to the mosquito to mark them as amateurs (Spielman and D'Antonio, 2001). The deep animosity among Ross, Grassi and others foreshadows the current divisions between medical and entomological researchers and, to a lesser degree today, the nationalistic and competitive nature of the research.

With Ross's results, work on the transmission cycle gained speed and urgency. In 1898 the malaria cycle between humans and mosquitoes was described by Grassi, Bignami and Bastianelli in Italy. About 50 years later the achievements from the end of the 19^{th} century would spur researchers and governments to work on a vision of malaria eradication.

1.2.2 Eradication

The first efforts to reduce or eliminate malaria were those of the Ancient Greeks and Romans, who "improved" swamps and water drainage; they knew these to be associated with disease (Russell, 1952). After the discovery of the parasite and its mode of transmission, efforts focused on eliminating mosquito larvae to eliminate transmission. In 1899 Ross led anti-larval measures in Sierra Leone. Another effort that year in Cuba by William C. Gorgas, a U.S. army major, was

highly successful. It was to become a template for later efforts, such as subsequent programs in Malaya led by Watson and in Egypt by Ross (1901 – 1904) (Gilles and Warrell, 1993).

The most famous project to follow the Cuban mold was the eradication effort in the Panama Canal Zone. Gorgas himself led that effort while the United States government built the canal. Without this, construction of the canal would likely not have been completed at that time; Indeed, an earlier French effort was abandoned in 1893 after thirteen years of work due to the high number of fatalities (suggested to be around 22,000), many of them from malaria. Despite controversy and political opposition Gorgas strictly enforced a costly vector elimination program resulting in elimination of Yellow Fever and massive reduction in malaria within the canal zone. During the decade of its construction it has been estimated that about 2% of the US workforce was hospitalized at any given time. The figure for the earlier French construction was about 30% (Spielman and D'Antonio, 2001).

The next two and a half decades brought possibly the most impressive instance of vector elimination. The 1939 – 1940 campaign, which eliminated an invasive population of *Anopheles gambiae* from North Eastern Brazil, led by Fred Soper (Killeen et al., 2002; Coura et al., 2006) marks a turning point in how governments confronted malaria and its vectors. The United States government and the Rockefeller Foundation for which Soper worked were convinced that *An. gambiae* could spread north and eventually threaten the U.S. with renewed malaria transmission.

Soper advocated complete vector extirpation to eliminate malaria, and convinced the Brazilian government of President Getulio Vargas that the region could be made malaria free, as he argued it was before, by getting rid of the

foreign mosquito (Soper and Wilson, 1943). Others, such as Evandro Chagas, argued that improving the living conditions of the people in the region and treating them with quinine was a better course of action in battling the parasite. (Valentine, 2005).

Soper won the day and permission to conduct a military-style operation treating breeding sites with Paris Green, a larvicide, and attacking adults with pyrethrum sprays (Harrison, 1978). In 1940 the area was declared free of *An. gambiae*, but interestingly it was not free of malaria. The area had not been completely malaria free before the introduction of *Anopheles gambiae*, so the outbreak of the disease just before 1939 may have been worsened by that species, but the parasite was able to sustain itself through other, native, species for years before and afterwords.

The heyday of malaria eradication programs probably started in 1943. At that time serious attempts were undertaken to eliminate malaria from the United States (Oaks and Mitchell, 1991), and there was a blossoming of research on malaria vectors around the world (e.g., MacDonald 1946). This was spurred by the serious problem malaria presented to troops fighting in World War II, and continuing for U.S. armed forces in Korea and Vietnam. Thus the U.S. military funded significant numbers of entomological studies from that point to the present day.

In general the resources and effort in these programs was directed to massive spraying rather than in-depth medical entomology. However even this focus of vector eradication programs was based on some new understanding of the vector. The shift to adult mosquitoes from larvae was precipitated by MacDonald's classic model that showed that killing adult mosquitoes was a more efficient way to reduce transmission potential (MacDonald, 1957).

During the Eight World Health Assembly in 1955 (WHO, 1955) a plan for world-wide malaria eradication was presented, largely relying on residual spraying of dichloro-diphenyl-trichloroethane (DDT). This compound's insecticidal properties were excellent, described in 1939 by a Swiss chemist named Paul Müller. However, the problem of insect resistance was known from the start (WHO, 1955), based on the results of an eradication program conduced in Greece after World War II. In those years *An. sacharovi* became resistant to DDT and by 1956 resistance was found in both adults and larvae to chloroquine and dieldrin, completely different insecticidal compounds (Harrison, 1978). So the time frame for eradication was purposely made as short as possible (under a decade). Sub-Saharan Africa was excluded from the WHO effort of 1955, though many individual projects were launched there over the next decade.

While the eradication campaign generated successes in North America and Europe, it achieved only temporary or little reduction of malaria in the rest of the world for technical, social, economic and political reasons (Oaks and Mitchell, 1991). By 1969 the goal of eradication for the remaining malarious regions of the world was replaced with control (WHO, 1969).

1.2.3 Control

The mainstream eradication efforts of the 1950's and 1960's did not include sub-Saharan Africa for two reasons: First, the problem there was thought to be too severe and local capacity was judged to be lacking for implementation of an eradication program. Second, there was the view that young African children in particular should not be treated for their first infection of malaria to allow immunity to develop, advanced by some malariologists in the mid 1930s. Indeed a group of experts at the 1950 WHO meeting argued for drug therapy as a main

weapon rather than vector control in Africa for fear that weakening adult immunity could lead to devastating epidemics (Harrison, 1978). Areas of holoendemic (stable) malaria transmission were thought best left alone. For these regions, and soon for all remaining areas with endemic malaria, control was judged to be a more reasonable objective.

Control is defined as "Reduction of disease incidence, prevalence, morbidity or mortality to a locally acceptable level as a result of deliberate efforts" (Molyneux et al., 2004). It is worth noting that this is a significant philosophical shift from the optimistic idea that the transmission cycle can be interrupted and the disease eliminated completely from an area (Najera, 1989), epitomized by such book titles as "Man's mastery of malaria" (Russell, 1955).

Despite Africa's being treated separately from the major eradication efforts of the mid 20^{th} century, there were about 20 pilot projects carried out there during the eradication efforts by national governments and the WHO. These often ended with poor results, partially because they did not include careful measurement of epidemiological parameters or of the entomological details within the areas where programs were being carried out (Molineaux and Gramiccia, 1980).

Due to these failures and the rise of resistance to insecticides in the mosquitoes and drug-resistant *Plasmodium*, much of the optimism of the previous years was lost after 1970 (Trigg and Kondrachine, 1998) and until quite recently as discussed below. There were some bright spots, however. One of the most important was the so-called "Garki Project" a joint WHO/Government of Nigeria initiative which ran from 1969 – 1976 (Molineaux and Gramiccia, 1980).

Though it ran at a significant cost at the time, the Garki project addressed all aspects of the malaria cycle in one location for an extended period of time. Workers collected extensive baseline data pre-intervention, tested an intervention

based on residual spraying and mass drug administration, and created a mathematical model of malaria transmission to be parametrized by and inform their field data. The Garki project was so carefully directed and thoroughly funded because plans to control malaria tend to be more complex than those to eliminate it. This work gave enhanced prominence to modeling efforts, which continues to this day.

In the Garki project there was a heavy focus on entomological work and measurement, including parameters such as vectorial capacity (C) used to inform the model. It was not the first attempt to use entomological knowledge to understand the effect of interventions with mathematical models. Ross pioneered this approach in 1911 (Ross 1911; see Fine 1975). Most of the models developed since then are systems of differential equations, typically based on "compartments," where individuals in a population are divided into classes ("susceptible" or "infected," for example). A compartment approach is the basis for the classic MacDonald (1957) extension of the Ross archetype, which showed that reducing the life span of the adult mosquito vector was an efficient way to drop transmission below the minimum threshold that permits the parasite to remain endemic.

The model created as a result of the Garki project also followed a compartment approach, but was very detailed: it included seven rather than the usual three human categories. The high degree of detail enabled the study of immunity and superinfection, outstanding questions for many years (Macdonald, 1965). The Garki project was one of several advances in modeling of the era, centered around a dynamical approach, allowing a wider view of the population biology of disease agents (Anderson and May, 1979).

Despite the dispiriting collapse of eradication efforts, optimism seems to be returning to the malaria and *Anopheles* research communities. A major cause

for enthusiasm are recent efforts aimed at vaccine development. Another idea that has done much to revive interest in *Anopheles* entomology particularly is the possibility of controlling the disease through the introduction of a genetically modified vector, first proposed in the 1980's (Aultman et al., 2001).

Today several labs are working intensely on developing the genes and approaches that might be used under such a control program, buoyed by the recent investment from the public and private sectors. The work in this dissertation shows that, just as in the past, details of vector biology are of critical importance to control.

1.3 *Anopheles gambiae*

Anopheles gambiae is the most efficient malaria vector in the world, due primarily to its anthropophilic habit and its receptivity to *Plasmodium*. It has been heavily studied, with a massive literature devoted to its ecology, evolution and vectorial properties. The concentration of effort on this species is not just due to its status as a malaria vector, however. It is also because it is a biologically complex species, particularly with respect to its taxonomy and mode of adaptation to the environment.

1.3.1 Taxonomy and evolution

The *Anopheles gambiae* species complex (*An. gambiae* sensu lato) is composed of seven closely related, morphologically cryptic species (Davidson, 1964; Hunt et al., 1998): *Anopheles arabiensis* (Patton, 1905) , *An. merus* (Donitz, 1902), *An. melas* (Theobald, 1903), *An. quadriannulatus* (Theobald, 1911) species A and B, *An. bwambae* (White, 1985) and *An. gambiae sensu stricto* (s.s.)

(Giles, 1902). These species differ in their environmental niches and efficiency of malaria transmission, complicating control programs in Africa. For example, *An. quadriannulatus* is found only in East Africa and is not known to transmit malaria. *An. gambiae* s.s. and *An. arabiensis* are pan-continental south of the Sahara and major vectors of malaria.

The origins of *An. gambiae* are complex and explanations are heavily contingent on the method used to study them. In terms of *An. gambiae* s.s., the problem is compounded because groups are morphologically equivalent and because the split is considered to have been recent and to be incomplete. Most approaches have been genetic because the fossil record is limited.

Anopheles gambiae have a genome made up of two pairs of autosomes and a pair of sex chromosomes. The polytene complement consists of five chromosomal arms with characteristic banding patterns which have been compared among species. There are abundant chromosomal inversions seen in these banding patterns: ten are fixed between species, but more than 120 have been found to be polymorphic between them in the field (Coluzzi et al., 2002).

The earliest work elucidating origins thus utilized cytogenetic techniques, which led to designating the above-mentioned chromosomal forms. Based on this approach *An. gambiae* s.s. was argued to be most closely related with *An. merus* because they share the sex-linked Xag inversion (Coluzzi et al., 1979). However, this designation conflicts with the evidence from shared chromosome 2 inversions between *An. gambiae* and *An. arabiensis*. Those similarities were hypothesized to result from introgression, which is possible for chromosome 2 but not for heterochromatic X chromosomes. Subsequently this interpretation has been supported by other studies (see Krzywinski and Besansky 2003).

The origin of *An. gambiae* is thought to have occurred with the introduction

of agriculture and human population expansion in Africa (Coluzzi et al., 1985), occurring between 10,000 and 4,000 ybp (Cavalli-Sforza et al., 1993). This argument is based on the adaptation of *An. gambiae* to humans and its strong anthropophilic feeding habit. In addition, molecular studies show that the split is recent and differentiation low, concordant with the timing of the human population expansion (Besansky et al., 1994; Donnelly et al., 2001; Coluzzi et al., 2002), although this interpretation is still being tested.

As mentioned above, a number of paracentric inversions within *An. gambiae* s.s. have been described on the right arm of the second chromosome (Coluzzi and Sabatini, 1967); these were called *j,b,c,u* and *d*. The inversions appear not to be randomly associated in nature, but rather appear in a dozen common combinations (Black and Lanzaro, 2001). Based on these associations between inverted and wild-type karyotypes five "chromosomal forms" were described and given the non-Linnean names "Mopti", "Savanna", "Bamako", "Forest" and "Bissau"(Coluzzi et al., 1979; Touré et al., 1998a). Each comprises several inversion karyotypes. Of particular interest are Mopti, Savana and Bamako forms which occur sympatrically and are stable in Mali.

The inversion karyotypes can be denoted with a "+" for wild type and the inversion letter (j, b, c, u or d, in this order) if inverted. Forest is thought to be the ancestral form of *An. gambiae* s.s., mainly displaying +++++ karyotypes with occasional b, c, u or d inversions. Savana is the most broadly distributed 2R form and has +bc++, ++cu+, +bc+d, +bcu+, +b+u+ or +b++d patterns. The Mopti form is associated with rice fields and floodplains in West Africa and has the following inversion patterns: +++++, +bc++ and +++u+. Bamako is another form only found in West Africa, associated with the Niger river, and is the only one to display the *j* inversion: j+cu+ and jbcu+. Finally, the Bissau

form is found only in The Gambia and has +++++ and ++++d karyotypes (Black and Lanzaro, 2001).

The five inversions on the right arm of the second chromosome are not the only ones that appear to differ among these genetically discontinuous forms. The $2La$ inversion is nearly fixed in the Bamako and Mopti forms but occurs at significantly lower frequency in sympatric Savana populations (Coluzzi et al., 2002). $2La$ heterokaryotypes do not appear to suffer any fitness loss in the laboratory at least. Such observations are suggestive of assortative mating.

Hybridization between the Malian chromosomal forms varies from very little (Mopti and Savanna; Bamako and Savana) to zero (Bamako and Mopti) (Touré et al., 1983; Coluzzi et al., 2002). Estimates based on inversion karyotypes are complicated by shared inversions, however, and laboratory experiments did not reveal postmating isolation (Di Deco et al., 1980). Currently it is difficult to conclude that chromosomal forms are completely isolated from each other and "good" biological species, but they may be in the early stages of speciation (recently argued by della Torre et al. 2002, 2005; Fanello et al. 2003).

The delineation of forms themselves within *An. gambiae* s.s. is still to be fully worked out, especially on a continent-wide scale. Molecular differences between the forms have been sought for about the last decade (Favia et al., 1997). Based on differences in ribosomal DNA (rDNA) markers, two "molecular forms" have been defined, named "M" and "S" forms (della Torre et al., 2001). These correspond well with Mopti and Savanna/Bamako chromosomal forms in Mali and Burkina Faso, where the three coexist (Gentile et al., 2002), but do not align in other parts of the continent (della Torre et al., 2001; Wondji et al., 2005). Hybrids between the molecular forms are appear much less frequent than among the groups based on chromosomal arrangements (Tripet et al., 2001; Taylor et al., 2001; della Torre

et al., 2005).

1.3.2 Life history

Like other Culicidae, *An. gambiae* has a complex life cycle: an aquatic larval stage and a terrestrial adult phase. Only female adults feed on blood, necessary to produce a batch of eggs; both sexes feed on plant nectar and juices. The adult life span is usually a few weeks, though this depends on environmental conditions (see Chapter 3). During her life span a female may feed and reproduce several times.

Eggs are laid singly in water, and measure about 1 mm in diameter. Larvae hatch from the eggs, progressing through four larval instar stages (Gilles and De Meillon, 1968). The larvae have an elongated appearance, thicker on the head end. Anopheline larvae float horizontally against the water surface keeping their breathing tube, located near its tail, open to the air above. In contrast, larvae of *Culex* spp. rest at an angle to the water surface, their heads pointing toward the bottom at an angle of about 45 degrees. While the larvae spend most of their time at the surface feeding on microscopic organisms and particles, they may wiggle and dive if disturbed by vibrations or a shadow. Growth occurs between stages with molting.

The last aquatic stage occurs when the 4^{th} larva molts into a pupa. The pupae remain in the water, but don't feed and are generally unresponsive, a state "well described as an animated coma" (MacDonald, 1946). After about 24 h the adult mosquito emerges through an opening at the head to the water surface, where it can dry its wings. It then flies to a nearby resting place, leaving behind an abandoned shell. The length of the larval phase and pupation are dependent on temperature and other environmental factors, usually lasting a several days.

After a few days for maturation, mating occurs in crepuscular swarms near the emergence site, usually near a marker of horizontal contrast (Charlwood et al., 2002). Recognition of females entering the swarm is generally thought to be by wing beat frequency, though in *An. gambiae* chemical recognition is also known to be involved (Yuval, 2006). There is a 10 – 30 second copulation on the wing with no known courtship, resulting in a spermatheca being deposited in the female. There are low rates of polyandry in *An. gambiae* (Tripet et al., 2003), completely absent in other Culicidae (Yuval, 2006).

After mating the female will need to obtain a blood meal in order to lay her eggs. She will only seek one if inseminated. In the case of *An. gambiae*, which feeds almost exclusively on humans, she is attracted to nearby homes by a scent plume. At close range she homes in on the carbon dioxide put out by the usually-sleeping human. She will land softly and probe for small blood vessels with her proboscis. Guided by heat, this probing finally results in her nicking an arteriole or venule. She delivers saliva containing an anti-coagulant throughout the feeding. In about 1-1/2 minutes she will have ingested about two to three times her body weight in blood. (Spielman and D'Antonio, 2001).

The female must then struggle to reach a nearby resting place, usually a vertical surface such as a wall. She is an easy target at this time, and the danger of undertaking a blood meal is at its highest when she slowly flies away from the human. For about one hour she will expel water from the blood she has taken, keeping the solids for digestion and developing her eggs.

The adult male *Anopheles* is smaller and thought to be shorter-lived, subsisting on nectar, usually away from humans and other animals. During the day they may rest in a shady area, particularly when conditions are dry or hot.

It is very difficult to assess whether infection with *Plasmodium* has a negative

impact on the fitness of the mosquito. Most studies support some fitness cost to refractoriness (Lyimo and Koella, 1992; Hogg and Hurd, 1997), but this cost is probably equivalent to that of supporting infection (Hurd et al., 2005).

Effects on mosquito behavior from infection have been documented, such as increased rates of feedings on multiple people in the same night in infected mosquitoes (Koella et al., 1998). Other effects, such as modulating vector reproductive output, are more arguably behavioral manipulation by the parasite, since it is not clear how they improve parasite fitness (Hurd, 2003).

1.3.3 Larval and adult ecology

Anopheles gambiae s.s. has been successful at colonizing diverse environments associated with human activity. Larvae may be found in pit, drains, tire ruts, puddles, rice fields and swamps (Gilles and De Meillon, 1968). In general the preferred environments have low or no vegetation, some exposure to sun light and clear to slightly cloudy water (MacDonald, 1946). Related species are often found to have very different larval habitat preferences; *An. merus*, for example, breeds in salt water. *An. funestus*, which often co-occurs with *An. gambiae* tends to be found in shaded pools with clear water and abundant vegetation, as another example.

The larval habitat remains poorly understood. Adjacent rice fields, which share the same water may have completely different abundances of larvae (see Chapter 3). This may be due to predation differences (e.g., Service 1977), water quality issues (Mutero et al., 2004a) or even more complex dynamics involving food limitation and cannibalism (Koenraadt et al., 2004).

The adult habitat of *An. gambiae* tends to be around human settlements for feeding and around water for breeding and laying eggs. Different chromosomal

forms are adapted to differing levels of aridity, a factor that has a significant effect on mosquito survival, explored in detail in Chapter 5. A survey of data from across the Afro-tropical region over about 50 years shows *An. arabiensis* more often in areas of lower rainfall than *An. gambiae* s.s. Within the latter the Mopti chromosomal form is more often associated with lower rainfall than other forms (Bayoh et al., 2001).

The same pattern of Mopti adaptation to drier habitats than the other chromosomal forms was found in an extensive study by Touré et al. (1998a) in Mali. They also found a strong seasonal pattern to chromosomal form composition, with Savana and Bamako forms becoming more abundant in the wetter part of the year. Furthermore, Bamako is found in particular parts of the Niger river basin, described in more detail in Chapter 6.

Ecological differences between molecular forms are much less clear. There is some evidence for differences in molecular form composition between larval habitat types (Edillo et al., 2002, 2006). These studies were conducted in an area where the molecular forms correspond to particular chromosomal forms, however, so they don't show that the differences observed are associated with the molecular form rather than the chromosomal form of the larvae. Other evidence has been cited (della Torre et al., 2002), but is also inconclusive. There is no evidence yet of differences between molecular forms in their adult habitat or feeding preferences (Onyabe et al., 2003; Abdoulaye et al., 2003). However, not much time has elapsed since the definition of the molecular forms in 1997 and the issue has not been exhaustively studied, so results supporting ecological differences between adults of differing molecular forms may yet appear.

1.4 The malaria cycle

Malaria in humans can be caused by any one of four unicellular protozoans of the genus *Plasmodium* (Coccidiida: Plasmodiidae). They are *P. malariae* (Laveran, 1881), *P. vivax* (Grassi and Feletti, 1890), *P. ovale* (Stephens 1922) and *P. falciparum* (Welch, 1897). The most deadly and complex disease is associated with the last of these, which mainly infects humans in sub Saharan Africa, though it is also found in other tropical or subtropical regions. Due to its importance in Africa it is the focus of discussion here, though the life cycle of all four species is essentially the same.

The definitive hosts of the *Plasmodium* parasite are mosquitoes of the genus *Anopheles*, wherein it undergoes the sexual phase of its life cycle (sporogony). Upon transmission to a vertebrate host the parasite multiplies asexually (schizogony) to prepare for infection of the next Anopheline. The genus *Plasmodium* has been defined by asexual multiplication in cells other than erythrocytes of the vertebrate host (*exo-erythrocytic schizogony*) and by Dipteran invertebrate hosts (Gilles and Warrell, 1993).

Traditionally, the life cycle of the malaria parasite is described beginning with its sexual phase in the mosquito (Russell, 1952; Bruce-Chwatt, 1985). A female *Anopheles* becomes infected upon feeding on a human harboring the parasite. She may take up infected erythrocytes and free-floating sexual stage gametocytes. The male gametocytes form a thread-like structure in a process known as exflagellation, which produces 8 – 10 microgametes (also known as gamonts). The microgametes seek out the female macrogametes, which are then fertilized by fusing with the microgametes, producing a zygote. The zygote matures to an elongated form including a cytoskelleton and is now called an oökinate.

The oökinates migrate through the insect stomach wall, ending in round structures on the outer surface of the organ, known as oöcysts. These may number in the thousands or be few. Oöcysts increase in size until they burst, releasing what are known as sporozoites into the body cavity. These eventually reach the female's salivary glands, penetrate them, and are injected into the human host with saliva upon the next feeding. The time between infection of the mosquito and readiness to infect another human from the salivary gland is a few days.

The asexual development of the parasite begins in the human host. Sporozoites disappear from the blood stream in about 1/2 an hour. Some are phagocytosed, but others enter parenchymal cells of the liver (hepatocytes). In relapsing malaria (*P. vivax* and probably *P. ovale*), some sporozoites may further differentiate into hypnozoites or developing tissue schizonts. These may remain for a long, predetermined time before entering exoerythrocytic schizogony, described below.

Exoerythrocytic schizogony occurs when sporozoites form into schizonts, which grow in hepatocytes for 6 – 16 days post infection. After this growth period they appear in the blood in the form of thousands of tiny merozoites. In the blood stream the parasite enters the erythrocytic development phase. Merozoites invade erythrocytes, losing two of three membranes and assuming a rounded shape within the red blood cell. There the merozoite initiates a development process leading to trophozoites, which undergo asexual division (erythrocytic schizogony) ending in a population of schizonts. When fully developed these produce more meozoites, which burst the host erythrocyte and enter the blood stream, where they invade further erythrocytes.

This cycle is repeated, increasing parasetemia levels. It ends up equilibrating into a characteristic three or four day peak activity pattern, which led malaria

to also be known as the tertan or quartan fever. After several such cycles sexual gametocytes may be formed, bringing us back to the beginning of the *Plasmodium* life cycle.

1.5 Study areas

Research was conducted in two parts of Mali. The first is the village of Banambani (latitude 8°03' W, longitude 12°48' N), which has been an important entomological research site since 1979 (Touré, 1979). The large number of research projects conducted on Anophelines in particular has led to a highly knowledgeable local population, many of whom are proficient in collection methods and species identification. Beyond its historical importance for vector research Banambani is useful as a typical Sudan savana setting, with small scale agriculture.

The second area is Niono, home to the largest irrigated risiculture project in Mali. This region has been the focus of much recent research into the role of irrigation on disease transmission. With the increasing population pressures and the attendant need for greater food production areas such as Niono are likely to become more common. The environmental effects of this kind of massive agriculture project are therefore important to the future of an ever increasing number of West Africans.

1.5.1 Banambani

Banambani is located about 25 km north east of Bamako in the northern Sudan Savanna zone abutting the Malian Sahel. The nearest large town is Kati, a settlement of several thousand inhabitants. A temporary stream runs near Banambani, which is surrounded by agricultural fields and beyond those, hills (Figure 1.1).

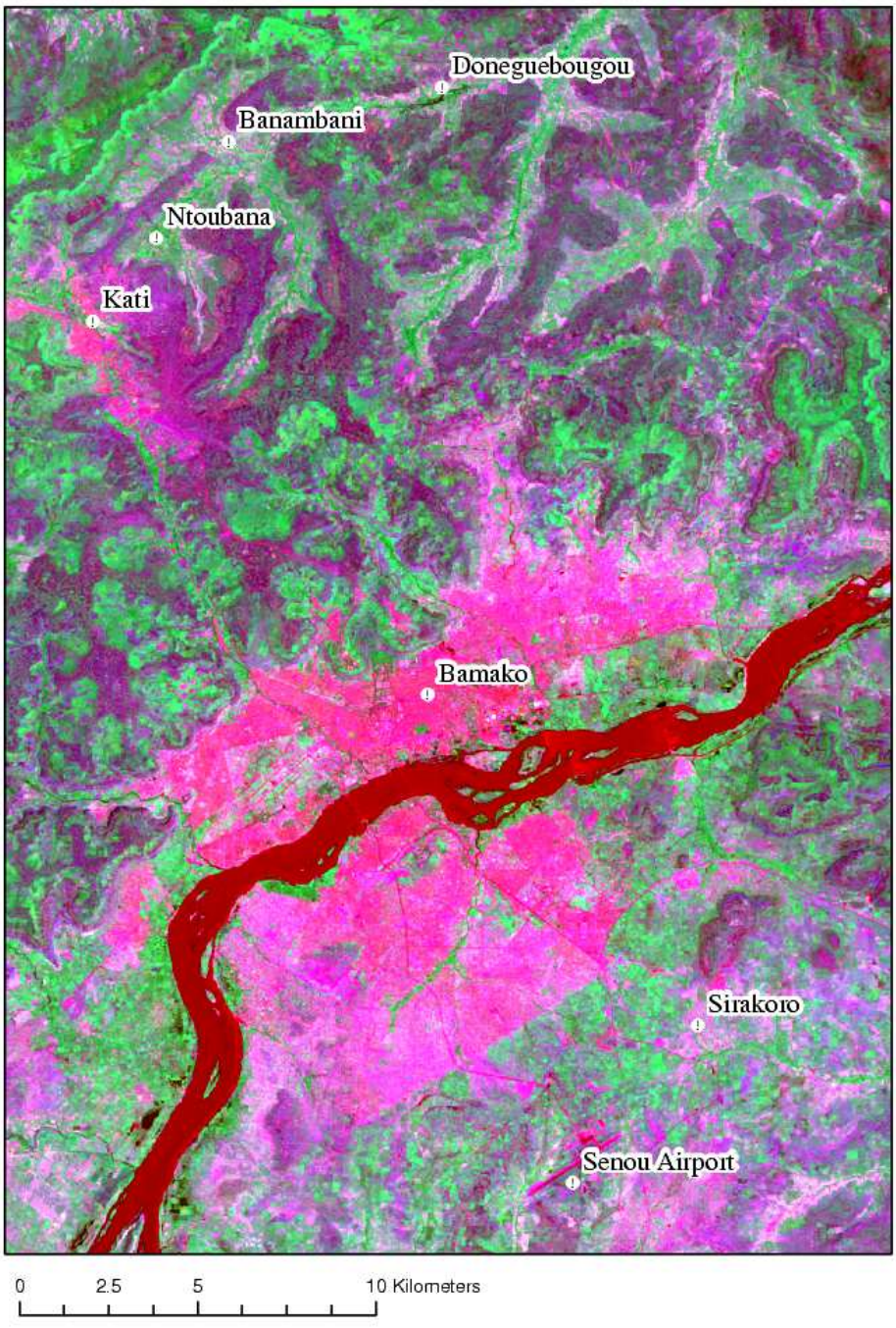

Figure 1.1: Landsat 7 ETM+ false color satellite image of the area around Bamako. The villages of Banambani, N'Toubana, Doneguebougou and Sirakoro are those included in the simulations presented in Chapter 4

Annual rainfall is between 500 – 1000 mm in a highly seasonal pattern. Almost all the precipitation occurs during the wet season from May to October. There is an alternating dry season from November to April (Touré, 1979).

The human population of about 700 are mostly farmers of the Bambara ethnic group. Most of the residents are animists, though there are some Muslim and Christian families also (Touré, 1979). They inhabit about 250 houses organized in around 70 compounds arranged in a low density pattern. Most of the structures are around the village center, though there are a few compounds in an area about 500m from it and separated by some crops (Figure 1.2).

The crops grown are primarily millet, sorghum, maize, peanuts, beans, potatoes and, seasonally, rice. There are some livestock, goats, sheep and cows, living around the houses. The houses are almost all made of mud brick with one of three roof types: a conical thatch roof (on round cooking buildings), a flat mud roof or a flat sheet metal roof.

The area has been extensively surveyed for mosquito larvae. *Anopheles* larvae are found in three major habitat types: rock pools by the river, swamps and temporary road puddles (Edillo et al., 2002). There is also some mosquito production from rice fields, though this is probably not a major contributor to mosquito density.

Malaria transmission in Banambani is seasonal, coincident with the rains and higher vector densities (between June and November). The transmission intensity during the peak of the season has recently been estimated at 19.2 – 21.1 infectious bites per person when measured by spray catch in the village of Doneguebougou (Dicko et al., 2004), which is about 2km away from Banambani (Figure 1.1).

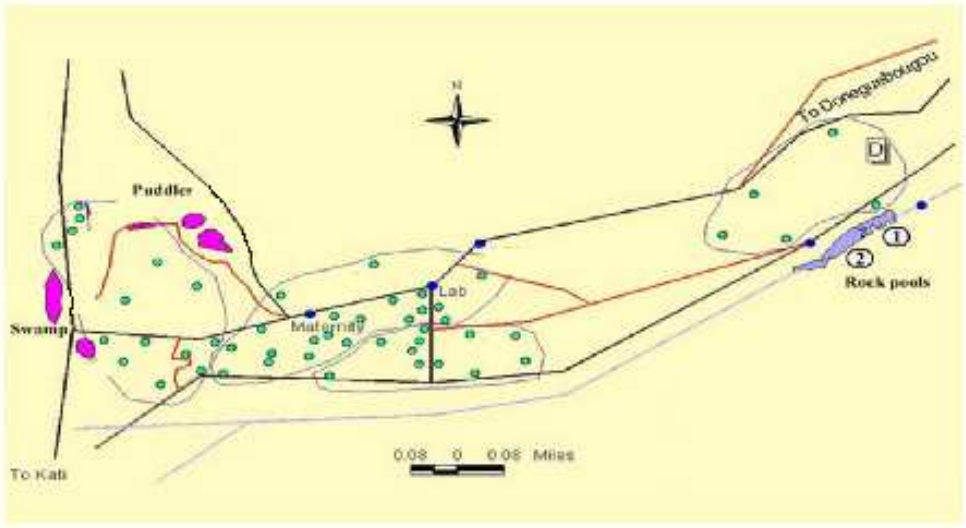

Figure 1.2: Map of Banambani, Mali showing compound and larval habitat locations. Larval habitats are classified as puddles, rock pools or swamp. Figure adapted From Edillo et al. (2006)

1.5.2 Niono

The district of Niono is located in the Sahelian area of Mali, 350 km north of Bamako near the town of Segou. It is accessible year round via a newly paved road from Markala. Three distinct climatic seasons can be distinguished: a wet season of about three months (July to September), a cold dry season (October to February), and a hot dry season (March to June). Overall there are about 400 mm of annual precipitation. Various villages were included in the studies reported in Chapters 2 and 3 (exact locations are given in Table 1.1).

The Niono irrigation project was established in 1932. Since that time the population of the irrigated area has increased to 150,000 people mostly of the Miniaka, Bambara, Peulh, Bela and Bozo ethnic groups. A large majority in Niono are Muslim. The main irrigated crop grown in the region is rice, though there is also some vegetable cultivation, livestock such as cattle and goats, and

Table 1.1: Location of villages studied in the Niono region, Mali

Name	Type	Latitude (^oW)	Longitude (^oN)
Hamdalaye	irrigated	6.001604	14.39137
Coccodi	irrigated	6.022042	14.33674
NionoKoroni	irrigated	6.034000	14.28900
Molodo	irrigated	6.047055	14.24913
Sokourani	irrigated	6.050783	14.22004
Mourdian	irrigated	6.024810	14.18114
KolodougouCoura	irrigated	5.969906	14.20069
KoyanCoura	irrigated	5.931101	14.26810
Nango	irrigated	5.963000	14.27600
Siengo	irrigated	5.978944	14.39583
Nara	irrigated	5.990388	14.47661
Toukoun Courá	irrigated	5.948208	14.49762
SounkaloKan	irrigated	5.899817	14.46451
Sarango	irrigated	5.920917	14.41061
Ténégué	irrigated	5.945522	14.34334
Tissana	irrigated	5.912837	14.36103
Niessoumana	irrigated	5.965918	14.32681
Tigabougou	irrigated	5.954973	14.31098
Toumakoro	nonirrigated	6.182000	14.06000
Dokoboukou	nonirrigated	6.130000	14.16000
Kalanampala	nonirrigated	5.862000	14.14700

fishing.

The area outside the irrigation project is less densely populated with Bambara, Fulani (Peulh), Maure and Sarakolé peoples. There is millet and sorghum cultivation and come cattle, mostly raised by the Fulani.

The gravity-irrigation system is fed by the Niger river at the Markala Dam and managed by the *Office du Niger*. Originally the project was intended to produce rice and cotton for the whole of West Africa, but cotton cultivation is no longer practiced due to pest problems and rising water table levels (Dolo et al., 2004). Rehabilitation and improvement projects were initiated in the 1980s and

are ongoing. Since a 1990 deregulation, double cropping and the introduction of privately owned husking machines has increased productivity per hectare.

The pattern of irrigation and rice cultivation is in step with the seasons. In general, flooding for irrigation begins in June/July and the rice is harvested in October/November. There is variation in the precise rice cultivation schedule between fields because cropping cycles are constrained by the water distribution scheme (Klinkenberg et al., 2003). Most fields are cultivated one time per year but some farmers are able to grow a second crop (double cropping) between January and May.

The large irrigated risiculture area in Niono (Figure 1.3) means that breeding sites are available for anophelines year round (Klinkenberg et al., 2003; Dolo et al., 2004). However, not all stages of rice growth produce mosquitoes equally. The peak production of anophelines occurs in August/September, when the rice has been transplanted, but is not yet dense enough to shade the water. The peak in malaria transmission typically occurs shortly after the peak mosquito production, and is low when the fields are fallow during January to June. In some of the irrigated areas, and in some years, a second rice crop is also grown, beginning in March, and harvested in June. In such areas there are anophelines and malaria transmission through much of the year, with less marked seasonality (Dolo et al., 2004).

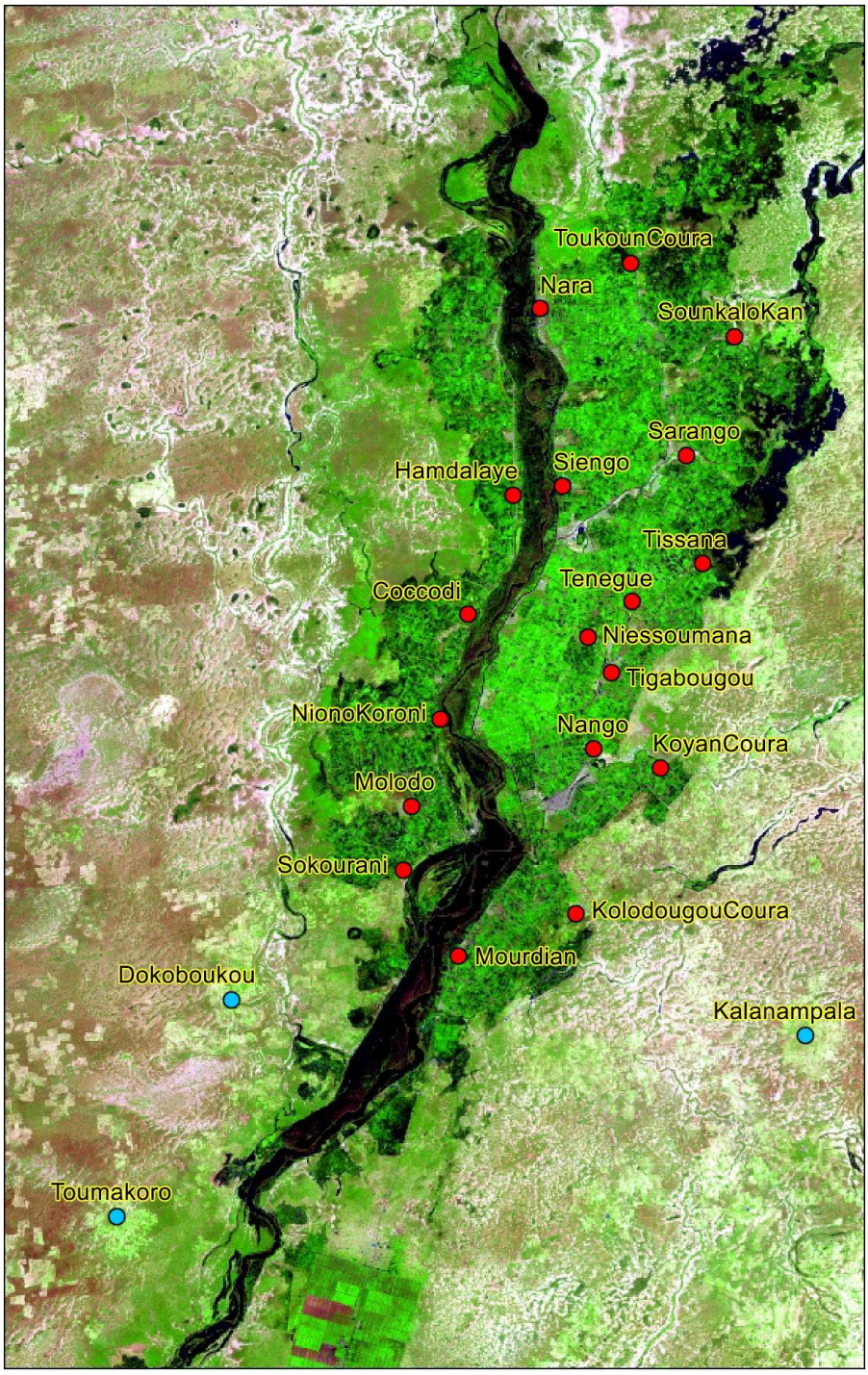

Figure 1.3: Landsat 7 ETM+ false color satellite image of the area around Niono, Mali. Villages marked in red are in the irrigated zone. Those in blue are in the adjoining non-irrigated region. Figure from Diuk-Wasser et al. (2005a)

1.6 Conclusion

This chapter provides the background for understanding the relationship between vector biology and malaria in this chapter. I have summarized the historical development of the field and ideas on combating malaria, a description of the current understanding of vector ecology and evolution together with a brief description of the malaria parasite's life cycle. This is presented with a description of my research objectives and of the areas in which I conducted research. The overarching theme, that *Anopheles gambiae* ecology leads to particular patterns of population density and structure in Mali which have significant impacts on its evolution and on prospects for malaria control, are re-visited in the final chapter.

CHAPTER 2

Anopheles gambiae density and transmission efficiency in Niono, Mali

2.1 Introduction

Dams and projects for irrigation of rice can have dramatic consequences for the health of people living in the areas around them. Along with their economic benefits, such projects may also bring sharply increased incidences of disease, including malaria, schistosomiasis, and filariasis. Increased malaria may result from higher numbers of mosquito vectors, but these may not always transmit more of the disease (Ijumba and Lindsay, 2001). Surveys of published reports indicate that no-change or even decreases in malaria transmission occur as often as increases (Table 2.1).

Several hypotheses have been suggested to explain why irrigation might lead to lower rates of malaria transmission. These include: (1) Irrigation projects lead to greater wealth, which in turn leads to better health care and an increase in personal protection against mosquitoes; (2) The large number of mosquitoes from irrigated fields cause so much nuisance that it induces people to use their bed

This chapter is a version of Diuk-Wasser et al. 2005a

Table 2.1: Studies comparing vectorial capacity (C), entomological innoculation rate (EIR) or human prevalence in irrigated and non-irrigated environments. These studies include either separate irrigated and non-irrigated areas or the same area before and after irrigation.

Country	Effect	Source*	Parameters Used
Burkina Faso		1	EIR
Cameroon	No Difference	2,3	prevalence
Ivory Coast		4	EIR
Senegal		5	EIR, prevalence
Mali		6,7	EIR, prevalence
Tanzania	Less Transmission	8	EIR (but incr. C)
Burkina Faso		9,10	prevalence
The Gambia		11	prevalence
Burkina Faso		12	EIR
Burundi		13,14	C, prevalence
Cameroon		15	prevalence
Guinea-Bissau	More Transmission	16	prevalence
Kenya		17-19	EIR, prevalence
Madagascar		20,21	EIR, prevalence
Sierra Leone		22	prevalence

* [1] Robert et al. 1985; [2] Couprie et al. 1985; [3] Audibert et al. 1990; [4] Dossou-Yovo et al. 1998; [5] Faye et al. 1993; [6] Dolo et al. 2004; [7] Sissoko et al. 2004; [8] Ijumba et al. 2002; [9] Carnevale and Robert 1987; [10] Boudin et al. 1992; [11] Thompson et al. 1994; [12] Baldet et al. 2003; [13] Coosemans et al. 1984; [14] Audibert et al. 1985; [15] Robert et al. 1992; [16] Gonçalves et al. 1996; [17] Githeko et al. 1993; [18] Hunter et al. 1993; [19] Mutero et al. 2004b; [20] Laventure et al. 1996; [21] Marrama et al. 2004; [22] Gbakima 1994.

nets more than they would if numbers were smaller; (3) Irrigated fields may favor certain species/forms of vectors that are less effective; and (4) High *Anopheles gambiae* densities in the larval stages lead to adults that are less efficient at malaria transmission. The last of these is termed the "competition hypothesis" here. None of these explanations are mutually exclusive, and all might contribute.

Intensive studies of the relation between irrigation and disease have been conducted by the Malaria Research and Training Center (MRTC) at the University of Bamako in Mali. One of these included three irrigated and three non-irrigated

villages (Dolo et al., 1999, 2004; Sissoko et al., 2004). Researchers observed that malaria was quite seasonal in the non-irrigated villages, concentrated in the wet season. In irrigated villages, especially those where a second crop of rice was grown, mosquitoes and malaria were observed through more of the year. Nonetheless, the total incidence of malaria transmission in the irrigated villages was much smaller than in non-irrigated ones. For example, the average Entomological Infection Rate (EIR; see Appendix A), was 8.7 infective bites per month for the non-irrigated villages, but only 2.4 infective bites/month for the irrigated villages. This paralleled the numbers of clinically diagnosed cases of malaria in children at the same times (Sissoko et al., 2004).

In an effort to explain this surprising relation, Dolo and colleagues favored the competition hypothesis. They could rule out the species/form composition hypothesis (no. 3), because the composition was substantially the same in both regions. Bednets were used almost universally in both regions, based on personal observations, ruling out explanation no. 2 for them. Generally the same medical care was available throughout the region at the time of their study, suggesting that differences in health care, no. 1, was unlikely to account for the differences in malaria transmission (though actual use of services could be affected by social conventions and distance to the health care facility). Finally, the competition hypothesis was supported by some results of their study. Specifically, they observed that during the season of peak density adult survivorship was at its lowest. Also, sporozoite rates and vectorial capacity were reduced when densities were high.

The goal of the work in this chapter is to better understand the relation between rice irrigation and malaria in a focal irrigation project of Mali, known as the *Office du Niger* and located in Niono; More specifically, why it is that malaria transmission is (generally) lower there than in nearby non-irrigated villages. If

competition is responsible for lower malaria transmission in the irrigated area then one should find that measures of malaria transmissability decrease with increasing vector densities. In other words, Anophelines should become less efficient vectors at higher densities.

To estimate vector efficiency we concentrated on two parameters affecting transmission: anthropophily (proportion of bites on humans) and survivorship. We also estimated one overall measure of insect population receptivity to the parasite, vectorial capacity (see Appendix A). We utilized a larger dataset than previous studies, resulting in the best estimate of the effect of density on transmission in Niono to date.

2.2 Methods

2.2.1 Collections

In their original study of malaria transmission in Niono, Dolo et al. (2004) included three non-irrigated villages (Toumakoro, Dokoboukou and Kalanampala) and three irrigated villages (Niessoumana, Ténégué and Tissana) (Figure 1.3). This study does not include any non-irrigated villages, but does include the three irrigated villages mentioned above plus 15 additional ones. The eighteen villages studied are distributed within the three administrative sub-zones of the irrigated area: 8 in the Niono sub-zone (Nango, Tigabougou, Niessoumana, Tissana, Ténégué, Koyan Coura, Kolodougou Coura, and Mourdian); 5 in the N'Débougou sub-zone (Siengo, Nara, Toukoun Courá, Sounkalokan and Sarango); and 5 in the Molodo subzone (Sokourani, Molodo, Niono Koroni, Coccodi and Hamdalaye).

The selection criteria for locations were: 1) the villages be at least 2 km apart (to decrease the likelihood of capturing mosquitoes coming from the rice fields of

a neighboring village) 2) accessibility and 3) willingness of villagers to co-operate. The fact that villages were located in different sub-zones (Niono, N'Débougou and Molodo) increased the variability in the numbers of mosquitoes captured, since each sub-zone is managed separately and had been subject to different levels of rehabilitation.

2.2.2 Malaria transmission

Malaria in this region is transmitted predominantly by *An. gambiae* s.l. and the *An. funestus* group. Dolo and others found that 99.6% of all *An. gambiae* s.l. were *An. gambiae* s.s. and of those, 98.6% were of the Mopti chromosomal form/M molecular form (In Mali, there is nearly a 1:1 association between the two, see Chapter 1). Irrigation typically provides breeding sites for these species, with attendant increases in mosquito densities. The simple number of mosquitoes is not the same as their ability to transmit malaria, however.

A useful and widely adopted estimate of a vector population's ability to transmit malaria is the vectorial capacity, C. It may be described as follows (Molineaux et al., 1988): Let the number of vectors per human be m and the number of bites per mosquito per night on humans is a, then a human is bit ma times per day, on average. Assuming an exponential survival rate with daily survival p, then proportion p^n of these vectors survives the incubation period (sporozoite cycle) of the parasite (n), so that they could then transmit the pathogen. The vectors are then expected to survive another $1/log(p)$ days, and bite other persons a times per day, on average.

Combining this vectorial capacity, a measure of transmission, is defined to be (Molineaux et al., 1988):

$$C = \frac{ma^2p^n}{-log(p)} \tag{2.1}$$

These terms and how they were measured are summarized in Table 2.2.

2.2.3 Entomological studies

A team from MRTC conducted eight entomological surveys were between April 1999 and January 2001, during the middle and end of the rainy season (August and October, respectively), the off-season crop (April) and the harvesting period (January). Each survey consisted of two-day visits to each of the study villages. Included in the survey were two types of entomological surveys: day collections and night captures. For the day collections, a team of three people estimated indoor resting density (N_t) using the pyrethrum spray catch method in 30 randomly chosen houses between 15:00 and 18:00 h. This consisted of covering all exposed surfaces with white sheets, spraying the rooms and collecting all fallen specimens (Service, 1993). Anopheline mosquitoes of interest were identified to species (*An. gambiae* s.l. or members of the *An. funestus* group).

Night captures were conducted at two houses in each village, at least 200 m apart, between 18:00 and 06:00 h with a personnel change at midnight. At each house, a collector was posted indoors and another outdoors with a flashlight and a mouth aspirator, collecting females as they landed (Service, 1993).

Collected females were classified by abdominal status (unfed, fed, semi-gravid and gravid) in the field when possible. At high density specimens were conserved in Carnoy's fixer (3 parts ethanol: 1 part glacial acetic acid) and classification was conducted in the laboratory. The number of human occupants during the previous night was recorded for each surveyed house. To estimate the proportion

Table 2.2: Variables used in the calculation of vectorial capacity. The values of n (length of sporozoite cycle) and g (length of gonotrophic cycle) are taken from the literature. Their values are 12 and 2 days, respectively. PSC, pyrethrum spray catch.

Variable	Definition	Estimation Method
N_t	Indoor resting density	Sum of all mosquitoes captured by PSC per house
N_f	Number of recently fed female mosquitoes	Sum of fed and semi-gravid mosquitoes captured by PSC per house
N_s	Number of human sleepers in rooms where PSC was carried out	Number of humans in rooms surveyed
m	Relative density of female mosquitoes to humans	N_t/N_s
ma	Human biting rate (bites/ human/ day)	N_f/N_s
P	Proportion of parous females (parity rate)	Detinova 1962
p	Probability of daily survival	g^{th} root of parity rate
A	Proportion of bites on humans (anthropophilic rate)	From ELISA
a	Number of bites per mosquito per night ("Man Biting Habit")	A/g
C	Vecotrial Capacity	$(ma \cdot a \cdot p^n)/-ln(p)$

of the blood-fed and semi-gravid *An. gambiae* s.l. and *An. funestus* that had fed on humans (anthropophilic rate), a blood aliquot was extracted from fed mosquitoes, conserved in Carnoy's and analyzed with ELISA (Burkot et al., 1984) by a separate MRTC team in Bamako. Parity rates were estimated from the night catches in the field using the method of Detinova (1962) the day after capture.

2.2.4 Statistical methods

Regression analysis was used to test density as a predictor of anthropophily, survivorship and vectorial capacity. Regression was chosen because village samples were comparable (all within the irrigated zone) and on a density continuum. Since samples per village were separated by three months we assumed they were approximately independent.

Density per village was estimated by the mean number of *An. gambiae* per sleeper ($N_t/N_s = m$). All the variables were tested for normality using Kolmogorov-Smirnov one sample tests and transformed as necessary.

2.3 Results

2.3.1 Spatial and seasonal patterns of anopheline indoor resting density

The geometric mean of the indoor resting density for all surveys, N_t, was 19.1 *An. gambiae* per house per night, and 3.6 *An. funestus* per house per night. *An. gambiae* was more abundant than *An. funestus* in all surveys but one, October 2000 (results not shown). This difference is consistent with the observation that *An. gambiae* breeds principally in irrigated fields, abundant during the rainy season, whereas *An. funestus* typically breeds in more permanent water sources.

Hence *An. funestus* predominated when irrigation was absent and densities lower.

N_t varied significantly over seasons and years, as did the mean values of the malaria transmissability parameters obtained in each of the surveys (Table 2.3). The highest *An. gambiae* N_t occurred in the middle of the main cropping season (August) in both years, followed by the off- season crop (April). In April 2000, N_t was significantly lower than in April 1999, coincident with a shortage of irrigation due to work on channel maintenance. Given these differences between the 2 years of the study, we will hereafter refer to surveys ($N = 8$) instead of years/seasons.

The seasonal change described above occurred simultaneously in the villages we studied (Figure 2.1). However, large variability was found within single village surveys (e.g. in August 1999, the numbers captured in one house in Toukoun Courá was 0 while in another house it was 2,487). Even with this large within-village variability, the differences in mosquito numbers among surveys were statistically significant, for both *An. gambiae* ($\chi^2 = 155.40$, $df = 17$, $p < 0.0001$) and *An. funestus* ($\chi^2 = 409.32$, $df = 17$, $p < 0.00001$).

2.3.2 Adult density and vectorial efficiency

Regression analyses were used to separately test the predictive power of adult mosquito density (m) on anthropophily (A: Figure 2.2a and Table 2.4a), survivorship (p: Figure 2.2b and Table 2.4b) and the calculated values of vectorial capacity (C: Figure 2.2c and Table 2.4c) per survey. Increasing density was strongly related to a decrease in anthropophily. A significant negative relationship with survivorship was also detected, but this was a weaker effect. Density was significantly positively associated with vectorial capacity, though the slope was shallow.

Table 2.3: Survey summaries of the variables used in the calculation of vectorial capacity of *An. gambiae*

	1999			2000				2001
Variable	Apr	Aug	Oct	Jan	Apr	Aug	Oct	Jan
N_t	70.39	337.72	7.01	1.60	17.34	356.78	5.12	6.13
N_f	40.91	75.52	4.59	1.72	10.36	129.04	3.48	3.55
N_s	3.65	3.79	3.84	3.90	3.65	3.54	3.67	3.95
ma	11.21	19.92	1.20	0.44	2.84	36.43	0.95	0.90
A	0.30	0.29	0.43	0.64	0.36	0.16	0.36	0.55
a	0.15	0.14	0.22	0.32	0.18	0.08	0.18	0.28
P	0.62	0.52	0.77	0.93	0.87	0.58	0.55	0.80
C	0.39	0.16	0.45	2.53	3.08	0.41	0.02	0.60

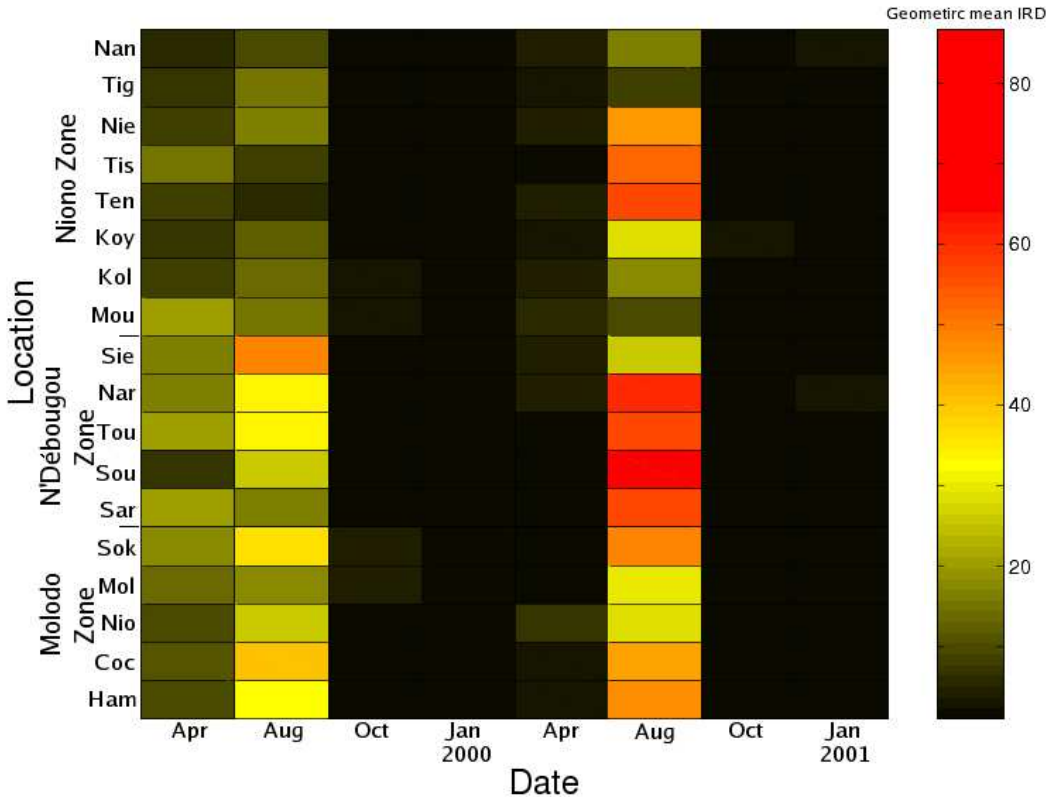

Figure 2.1: Geometric mean of indoor resting density (N_t) of *An. gambiae* in irrigated villages samples from 1999 to 2000

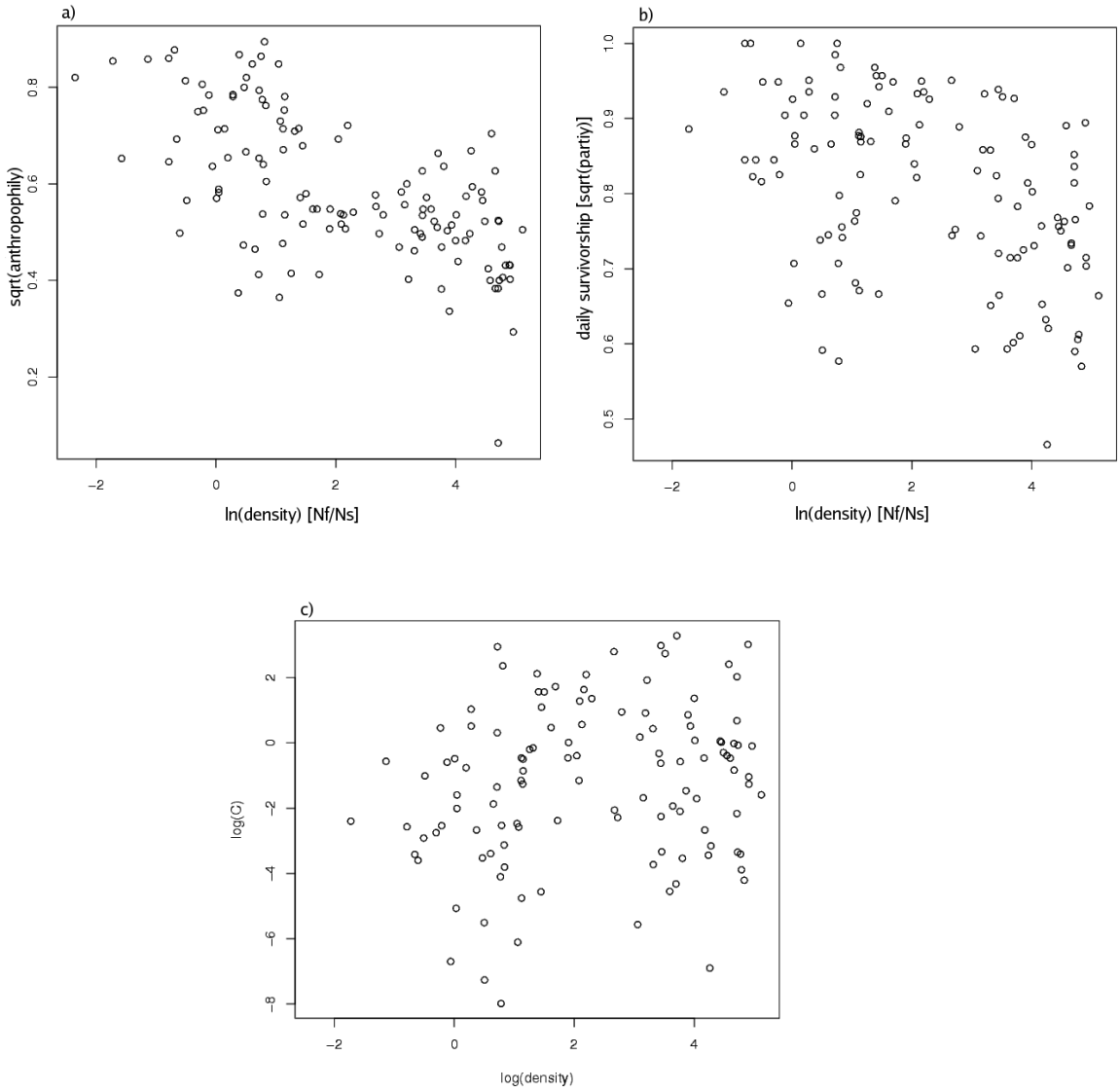

Figure 2.2: Regression plots of density (m) versus (a) percent feeding on humans, (b) survivorship and (c) vectorial capacity

Table 2.4: Density (m) as a predictor of (a) anthropophily, (b) survivorship and (c) vectorial capacity.

a) Anthropohily: $\sqrt{A} \sim ln(m)$				
	Estimate	Std. Error	t value	P
Intercept	0.694	0.016	44.53	< 0.0001
$ln(m)$	-0.050	0.005	-9.13	< 0.0001
Residual SE	0.116			
F	83.35 on 1 and 124 degrees of freedom			
Adjusted r^2	0.402			

b) Survivorship: $p \sim ln(m)$				
	Estimate	Std. Error	t value	P
Intercept	0.871	0.015	56.94	< 0.0001
$ln(m)$	-0.028	0.005	-5.33	< 0.0001
Residual SE	0.107			
F	28.42 on 1 and 120 degrees of freedom			
Adjusted r^2	0.192			

c) Vectorial capacity: $ln(C) \sim ln(m)$				
	Estimate	Std. Error	t value	P
Intercept	-1.820	0.355	-5.13	< 0.0001
$ln(m)$	0.249	0.122	2.05	< 0.05
Residual SE	2.375			
F	4.216 on 1 and 116 degrees of freedom			
Adjusted r^2	0.035			

2.4 Discussion

This study found a significant negative effect of adult anopheline density on anthropophily and survivorship, both determinants of transmission efficiency. The relationship is particularly strong in the case of anthropophily, an observation which is further discussed below.

In contrast, vectorial capacity, a composite measure of transmission ability, was positively related with density. This apparently contradictory result is miti-

gated by two factors: First, the relationship between density and vectorial capacity displays a shallow slope, indicating that while it was found to be significant it is not important. Second, there is great deal of variation in vectorial capacity that is not explained by density (Figure 2.2c and note the low r^2 value in Table 2.4c). Overall, density is likely not a sufficient explanation for the wide differences in estimated vectorial capacity.

The survivorship result bears directly on the competition hypothesis. It indicates that there may be some intra specific competition effect, though this study does not address whether it might be occurring at the larval or adult stages. Recall that the competition hypothesis holds that body size decreases with larval density and that smaller adult mosquitoes survive less well. There is ample reason to believe that competition in *An. gambiae* occurs in the immature stages (Service, 1993; Schneider et al., 2000; Koenraadt et al., 2004) and that smaller adults survive less well (Ameneshewa and Service, 1996; Takken et al., 1998; Gimnig et al., 2002). Experiments designed to explicitly test these issues are presented in Chapter 3.

In this study we used adult mosquitoes from indoor PSC catches as a measure of density. Adult density may or may not reflect larval density and competition in adjoining rice fields. In either event adult or larval density dependence will depend on numbers of mosquitoes relative to available resources. There is no good way to accurately assess how many resources are available, or to measure how good they are for *An. gambiae*. One house in a village may have no mosquitoes while another is found to contain thousands of females on the same night.

Larval habitat suitability is no easier to assess (Chapter 3). Density-dependent survival has been shown in the laboratory and under semi-natural circumstances, but more work needs to be done to demonstrate the same in field. Larval densi-

ties in previous studies (Schneider et al., 2000; Gimnig et al., 2002) are typically much greater that those commonly observed in natural water. Service (1993) has reviewed much of this work, citing a variety of ways the question has been addressed and referencing several dozen studies that found evidence pro- or con-density dependence. He could conclude only that "...it may be difficult to prove the existence, or otherwise, of density-dependent population regulation." (p. 757).

This study suggests that the strongest relation between density and transmission efficiency is with the anthropophilic rate (Figure 2.2a). The most obvious reason why this might occur is that when densities are high, people protect themselves and the mosquitoes are forced to alternate hosts. Human night catches have been made in this area, and during the rainy season the number of bites per person may exceed 550 bites/night by *An. gambiae* alone (Dolo et al., 1999, 2004). While Dolo and colleagues report near-universal use of bed-nets, based on personal observation, a more detailed study is warranted.

CHAPTER 3

Body size of *Anopheles gambiae* and malaria transmission in Niono, Mali: A test of the competition hypothesis

3.1 Introduction

Irrigation in Sub-Saharan Africa can be a mixed blessing. On the one hand, it contributes to greater food production and income; on the other hand, it may increase the incidence of diseases like schistosomiasis and malaria (Service, 1989a,b; Dzodzomenyo et al., 1999; van der Hoek, 2004). It is surprising, then, that several investigators have reported malaria transmission to be the same, or even less, in irrigated areas with high vector densities than in nearby non-irrigated ones with lower numbers of mosquitoes (Table 2.1). For example, Dolo et al. (2004) found densities of *Anopheles gambiae* Giles in the irrigated region of Niono, Mali in excess of 550 bites per person per night, compared to "only" 30 – 50 bites per person per night in nearby non-irrigated villages. However, malaria prevalence measured by longitudinal surveys was lower in the irrigated areas (Sissoko et al., 2004). Reasons for this departure from expectation might include: (a) the nui-

This chapter has been published in *J. Med. Ent* (Manoukis et al., 2006)

sance of so many mosquitoes might compel more people to use bednets; (b) the greater prosperity of irrigated areas might permit access to better health care and/or protection; or, (c) the mosquitoes that emerge from the high density irrigated areas during the rainy season might be less efficient at transmitting malaria, e.g. if they do not survive as well (Diuk-Wasser et al., 2005a).

This study tests the hypothesis that body size is the cause of reduced vector survivorship at high density, leading to lower malaria transmissability (Diuk-Wasser et al., 2005a) and prevalence (Sissoko et al., 2004) in the irrigated areas of Niono. Several aspects of this hypothesis are supported by previous work. In the irrigated zone, high density during the larval stage may lead to smaller larvae through competition for resources (Schneider et al., 2000; Gimnig et al., 2002). Since larval and adult sizes are correlated (Lyimo et al., 1992), competition results in smaller adults. Mosquitoes of smaller body size have been reported to survive less well than larger ones in several species (Nasci, 1986b,a, 1987; Kitthawee et al., 1990, 1992; Lounibos and Conn, 1991; Ameneshewa and Service, 1996; Takken et al., 1998). A combination of these factors could explain a decrease in malaria transmission (Kitthawee and Edman, 1995) within irrigated areas. If higher densities reduce the vectorial effectiveness of *An. gambiae* by decreasing body size and consequently reducing survivorship one expects to find that larval and/or adult densities have a negative correlation with adult body size and that adult body size in turn is positively correlated with adult survivorship.

3.2 Methods

3.2.1 Larval collections

Larval collections were conduced in Tissana, Ténégué and Niessoumana using 350 ml dippers (BioQuip Products; Rancho Dominguez, CA, USA) along the edge of rice fields. Two groups of three adjacent rice fields per village were sampled two times each in August and September 2003 (total: 36 samples or field/visits). There was at least a two week period between successive samplings of any particular field to render the samples roughly independent. The same villages were visited in September 2004 when 12 fields in each were sampled, including the six visited in 2003 (total: 36 fields). Larval collections were not conducted in Kalanampala because we did not locate habitats with more than a few *An. gambiae* larvae in or immediately around that village.

Two workers collected larvae at intervals of 1-2 m along the edge of each field following standard technique (Service, 1993). Density was estimated as the number of dips required to collect 25 larvae up to a maximum of 45 dips by the same worker. A fixed number of larvae was collected from each field in an attempt to standardize the variance of size estimates between fields. The sample collected by the first worker will be referred to as sample A; this was the sample used to estimate larval density and the instar composition in the field. The second worker collected sample B, which also numbered about 25 larvae. The number of fourth instar larvae in sample B was enriched through an extra collecting effort after the initial 25 larvae were captured with the aim of capturing at least ten individuals of this stage.

Samples A and B were taken back to the laboratory in separate plastic bags filled with water from the fields. Larvae in sample A were sorted by instar and

fixed in Carnoy's solution for later measurement. All larvae from sample B were placed in one growth chamber per field (BioQuip Products; Rancho Dominguez, CA, USA) filled with tap water. Pupation of fourth instar larvae generally occurred within 24 h. Since pupae do not feed, any effects of habitat on adult size should be reflective of conditions in the rice field and not the artificial growth chamber. Early each morning, for up to three days after the collection, any emerged adults in the growth chambers were fixed in Carnoys solution.

3.2.2 Density estimation in rice fields

In order to estimate the reliability of our measures of larval density, transects were conducted along the middle of the length (about 20 m) and width (about 8 m) of three rice fields, two in Ténégué and one in Niessoumana during September 2003. This resulted in a total of six transects. For each of these, a single dip sample was taken roughly every meter. These data were collected to assess the representativeness of edge dips to counts from elsewhere in the field.

An existing dataset (Diuk-Wasser et al., 2005b) of 1,448 rice fields that had each been sampled with 20 dips from April 2000 to January 2001 was also used. For that dataset five dips were taken at regular intervals along each side of the fields and the number of larvae in each dip recorded.

3.2.3 Adult collections

Collections of adults were made by night landing catches from Tissana and Ténégué in the irrigated zone and from Kalanampala during September 2004. A dry season collection was also made in March 2004 from only Ténégué. These villages all show the pattern of adult mosquito abundance typical to this region (Table 3.1). In non-irrigated villages *An. gambiae* is common during the rainy

Table 3.1: Number of adult *An. gambiae* captured per collector by landing catch

Sample	*N*	No. Nights	Int, Mean (SE)	Ext, Mean (SE)
Ténégué (dry)	8	2	33(10.0)	12 (2.5)
Ténégué (rainy)	4	1	41 (7.8)	40 (2.8)
Tissana (rainy)	4	1	57 (4.2)	37 (0.7)
Kalanampala* (rainy)	10	2	23 (7.6)	7 (2.2)

* Non irrigated village; *N*, Number of human-nights (one interior and one exterior per station per night)

season but apparently absent during the dry season; in the irrigated villages *An. gambiae* is very abundant during the rainy season and moderately abundant during the dry part of the year, when some farmers irrigate for a second rice crop in March – April.

Densities were high during the September collections, as expected, so sufficiently large sample sizes were obtained with one or two night catches. During the dry season there were fewer mosquitoes, so it was necessary to supplement night catches with daytime collections of unfed females using mouth aspirators from inside homes and other structures. Night landing catches were conducted from 18:00 to 06:00 h at two collecting stations over 300 m apart in the village, each with one collector indoors and one outdoors. Standard methods were employed for these collections (Service, 1993).

3.2.4 Estimation of survivorship and size

Adult females were dissected between 09:00 and 12:00 h on the morning following night catches. Parity rates were estimated by examination of the ovaries following the method of Detinova (1962). Under the assumption of no net change in population size (i.e. constant recruitment and death rate), survivorship can be estimated as the g^{th} root of the parity rate (Davidson and Draper, 1953; Davidson,

1954), where g is the number of days between a blood meal and oviposition (the gonotrophic cycle, WHO 1975). Numerous laboratory experiments by Y.T. Touré and colleagues (unpublished data) have found that g of *An. gambiae* s.l. in Mali is very close to 2.0, identical to other estimates from West Africa (Thompson, 1948).

As previous research has shown that wing length is highly correlated to the first two principal components of adult size (Petrarca et al., 1998), this measure was used to estimate adult body size. Samples were dried in individual tubes for 24 – 48 h before dissection of a single wing from each individual, which was mounted on a slide with a small drop of ethanol. Cover slips were attached with nail varnish and wing length measured from the alular notch to the wing tip excluding fringe using a filar micrometer and dissecting microscope. This procedure was repeated to estimate the body size of adult *An. gambiae* from the growth chambers.

Larvae were also measured using a filar micrometer and dissecting microscope. The larvae were measured for total length, from the distal point of the head to the end of the anal segment excluding antennae, feeding brush and caudal hair. We distinguished instars of *An. gambiae* and separated these from those of *Anopheles funestus* Giles (Gilles and De Meillon, 1968; Gilles and Coetzee, 1987) with the aid of a dissecting microscope.

3.2.5 Statistical testing

Using the transect data, the number of larvae per dip at field edges were compared to counts from field interiors with a Kruskal-Wallis tests to ascertain if sampling only from the edge of fields would bias density estimates. In addition, the general reliability of density estimation by dipping was tested using the Diuk-Wasser

et al. (2005b) dataset with a one-way intra-class correlation on average measure (ICC(k) in the terminology of McGraw and Wong 1996). This test produces a measure of the variance that can be attributed to the rice field itself (the object whose density we wish to estimate) as a proportion of the overall variance in the dip data. In other words, if each dip exactly measured the actual density in the field (agreed with all other dips), then ICC(k) = 1.

The effect of larval density on larval body size was tested using the mean values for length measurements per field. Mean length was calculated for each instar and species when more than one was present. Density (necessarily a single value per field) was estimated by the number of dips required to collect 25 larvae, as described above.

A relationship between adult size and survivorship was sought by comparing the body size distribution of parous vs. nulliparous female *An. gambiae* per village sample with a Kolmogorov-Smirnov two-sample test (Sokal and Rohlf, 1969). This approach allows the detection of significant differences in any size class and thus is more sensitive than testing the means alone. It also controls for other factors affecting survivorship between villages that may be confounded with adult size.

3.3 Results

3.3.1 Reliability of Larval Density Estimates

The number of larvae captured per dip in transects varied between zero and six. A statistically significant difference in number of larvae was not found between edge and interior dips within fields (Kruskal-Wallis test, Field 1: $N_{edge} = 8$, $N_{interior} = 17$; $\chi^2_1 = 1$, $P = 0.16$; Field 2: $N_{edge} = 8$, $N_{interior} = 19$; $\chi^2_1 = 0.085$,

$P = 0.77$; Field 3: $N_{edge} = 8$, $N_{interior} = 17$; $\chi^2_1 = 0.070$, $P = 0.79$).

The intra-class correlation showed that 86% of the variance in the data was between fields, indicating a high level of reliability ($N = 1448$ fields, 20 observations per field; ICC $F_{1447,27512} = 7.38$, $P <0.001$).

3.3.2 Density and larval and adult body size

Differences in instar composition and length per instar between years were tested to see if pooling was justified. Instar composition varied significantly between 2003 and 2004 (Fishers exact test $P < 0.035$) as did length per instar based on ANOVA (R Development Core Team, 2004): Instar: $F_{2,673} = 671.4$, $P < 0.001$; Year: $F_{1,673} = 4.2$, $P = 0.04$; Species: $F_{1,673} = 3.2$, $P = 0.08$. This suggests a need to include year as a blocking variable when examining the effect of density on larval size in this dataset.

An ANCOVA model (StataCorp, 2003) of these data found a significant effect of density on mean larval size (Table 3.2), but this effect disappears when one or more of the other terms are removed (results not shown), significant or not (such as year or species). This indicates a significant (and positive: see Figure 3.1) but slight effect of density on mean larval size, which can easily be missed if other factors are not controlled. The mean body size of female adult *An. gambiae* emerging from the fourth instar larvae of sample B had a strong relationship to the mean size of larvae of the same instar and species from sample A paired by field (Figure 3.2). The correlation between the two measurements is significant only for the comparison of 4^{th} instar larvae, and not significant between sample B adults and earlier instars of sample A. This test was only possible for the 2003 samples because no adult *An. gambiae* emerged from the growth chambers in 2004, possibly due to contaminated water.

Table 3.2: Analysis of covariance of mean larval size per field as determined by density. Larvae per dip is a continuous variable. Larval stage, species and collection year were discrete variables. (SS) are partial because the factors in the model are not perfectly orthogonal.

Source	Partial SS	d.f.	MS	F	P
Model	314.26	7	44.89	88.74	< 0.01
Density	2.56	1	2.56	5.06	0.03
Instar	126.40	2	63.20	124.92	< 0.01
Species	0.07	1	0.07	0.14	0.71
Year	1.33	1	1.33	2.62	0.11
Density*Instar	1.36	2	0.68	1.35	0.26
Residual	123.95	245	0.51		
Total	438.22	252	1.74		

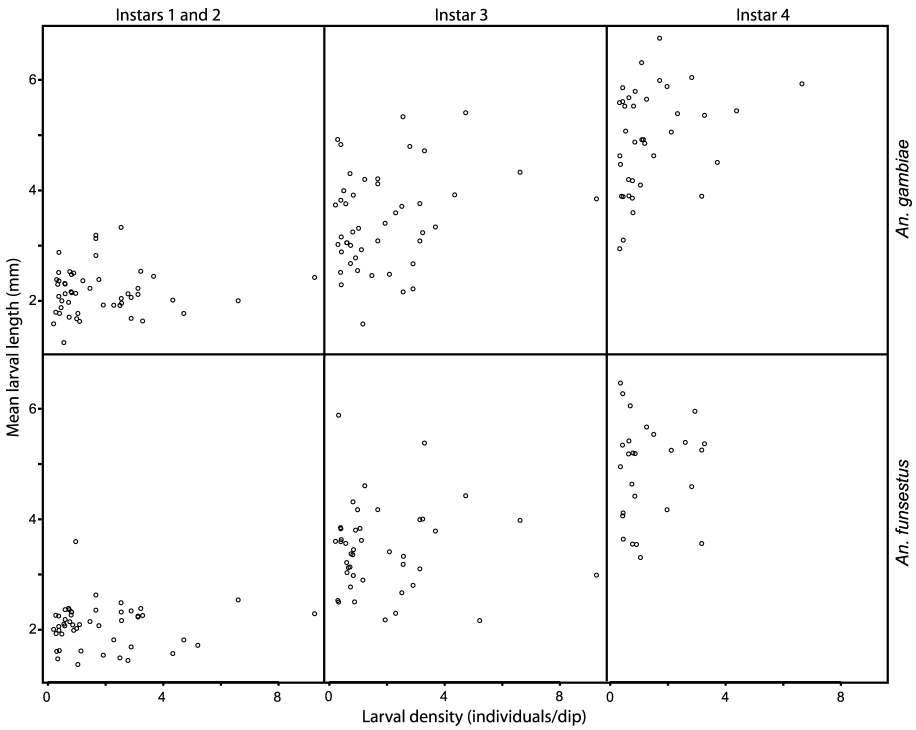

Figure 3.1: Estimated larval density versus mean length of larvae per growth stage and species. Each point represents the mean for a separate field from the 2003 and 2004 collections.

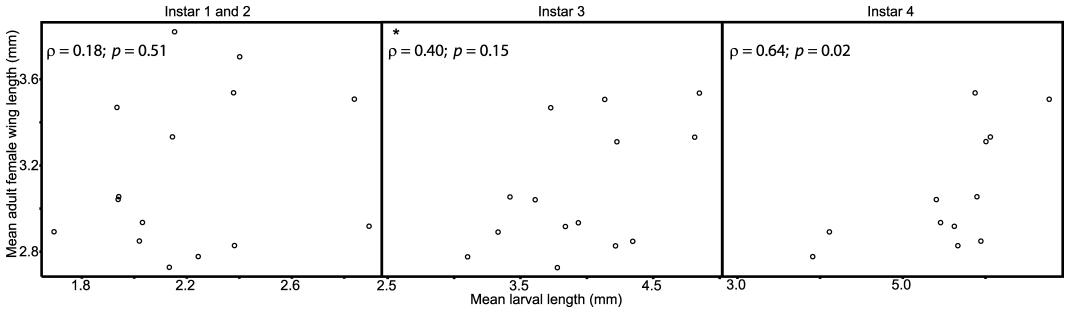

Figure 3.2: Mean larval length for *An. gambiae* versus mean adult *An. gambiae* female wing length from growth chambers collected from the same field. Larval means from sample A; N per mean 1 – 12 individuals, mean number of observations 3.13. Adult wing length means from sample B; N per mean 1 – 3 individuals, mean number of observations 1.44. Spearman rank correlations were conducted for each of the three stage classes, with results shown in the figure panels. The point marked with an asterisk (*) was considered an outlier and was not included in the analysis.

The wing lengths of all the emerged adult females of sample B ($N = 27$ individuals; mean = 3.14, SE = 0.07) were compared to those of the nulliparous females collected in the irrigated villages during the wet season ($N = 186$ individuals; mean = 3.07, SE = 0.01). There was not a statistically significant difference in size between the two samples, which supports the assumption that nulliparous females collected from the village are a representative sample of the size of adults produced in the fields ($t_{27.5} = 2.04$; $P > 0.05$).

3.3.3 Adult body size and survival

Differences between samples in parity rates and wing lengths were tested with Bonferroni corrected pairwise chi square tests (R Development Core Team, 2004) and a Student-Newman-Keuls multiple range grouping test (StataCorp, 2003), respectively. The two rainy season irrigated village samples (Ténégué, rainy season and Tissana) were the only ones which did not differ significantly in either

of these tests, and so were combined for a joint irrigated rainy season estimate (Table 3.3).

Estimated daily survival rates ranged from 0.53 in Tissana (rainy) to 0.67 in Ténégué (dry). Importantly, when the mosquito densities are lower (irrigated villages during the dry season and non-irrigated villages) then daily survival is higher, and when mosquito densities are high (irrigated villages during the rainy season) then daily survival is lower. Both this result and the estimates of survivorship obtained here are similar to those found by others (Dolo et al., 2004). Finally, Table 3.3 also shows that the distribution of wing lengths between parous and nulliparous females were in all cases not significantly different by Kolmogorov-Smirnov two sample tests.

Table 3.3: Parity rates, estimated daily survivorship and wing length measurements for all adult female *An. gambiae* collected in 2003 and 2004. Significant differences in parity rates between individual samples were tested with Boneferroni corrected chi squared tests. Comparisons were only significant between Ténégué (dry season) and Tissana, $\chi^2 = 9.77$, $p = 0.02$. Differences in wing length of all females from each sample were tested with a Student-Newman-Keuls (SNK) multiple range grouping test (subsets for $\alpha = 0.05$) on a one way ANOVA (wing length $\sim$ sample: $F_{3,565} = 94.47$, $p < 0.001$). For this test groupings are represented by the letters A, B and C; samples with the same grouping do not differ significantly in wing length. Differences in wing length between parous and nulliparous females within each sample were tested with two-sample Kolmogorov-Smirnoff (K-S) tests

				—Wing length:Mean (SE)—			—K-S Test—	
Sample Name	N	PR	ES	Parous	Nuliparous	All females/SNK	D	P
Ténégué (dry season)	157	0.45†	0.672	3.24 (0.25)	3.25 (0.26)	3.24 (0.26)/A	0.143	0.41
Ténégué (rainy season)	117	0.35	0.592	3.01 (0.17)	3.03 (0.17)	3.02 (0.17)/B	0.182	0.34
Tissana (rainy season)	153	0.28†	0.530	3.00 (0.16)	2.99 (0.15)	2.99 (0.15)/B	0.102	0.90
Kalanampala* (rainy season)	142	0.43	0.655	2.82 (0.23)	2.89 (0.19)	2.86 (0.21)/C	0.139	0.52
Irrigated zone (rainy season)	270	0.31	0.558	3.00 (0.16)	3.01 (0.16)	-	0.140	0.21

* Non irrigated village; N, sample size; PR, Parity Rate; ES, Estimated daily survivorship; † Significantly different parity rates at $\alpha = 0.05$

3.4 Discussion

While some portions of the size based explanation of the link between density and transmission are substantiated by these results, the entire chain of events is not likely in Niono. The results of this study support alternate causes for the previously observed decrease of malaria prevalence in villages with high vector density (Dolo et al., 2004; Sissoko et al., 2004). The potential role of bednet usage is noteworthy as discussed below.

An association between larval density and larval body size was seen, but it was weak and sensitive to other variables. This may be the result of small sample sizes per field. Surprisingly, the relationship observed was positive in the case of *An. gambiae*, the opposite of what is expected under resource competition (Gimnig et al., 2002). The positive relationship suggests that some fields are simply superior larval habitats to others, and that these produce both more and slightly larger larvae. Density independent factors that determine field quality appear to be important: adjacent fields of the same rice growth stage that shared a water source were often seen to have very different larval densities. Agricultural practices such as fertilizer usage have been shown to affect larval densities (Victor and Reuben, 2000; Mutero et al., 2004a), and have been proposed to be a factor in Niono (Diuk-Wasser et al., 2005b). In addition, predation may affect larval composition and numbers (Service, 1977).

Density dependence and competition among larval anophelines has been shown under controlled conditions without predation (Gimnig et al., 2002) and under strict laboratory conditions (Schneider et al., 2000). The density of larvae in such studies was in the hundreds per liter, however, while the highest estimated density observed in this study was under 20 per liter. This agrees with densities reported from other, much more extensive, sampling efforts in this area (Diuk-

Wasser et al., 2005b). We have found the accuracy of density estimation in these rice fields to be better than generally reported (WHO, 1975; Service, 1993), and that edge samples are reflective of density in other parts of the field; it is therefore unlikely that densities ever get high enough for there to be resource limitation.

A significant relationship between larval and adult body size was found for fourth instar larvae. This was absent from earlier instars collected from the same fields, as expected, because other cohorts are the product of a different environmental and parental composition than the one from which adults were produced for comparison. Since body size may affect multiple aspects of the mosquito vector competence (Lounibos and Conn, 1991; Takken et al., 1998; Suwanchaichinda and Paskewitz, 1998; Nghabi et al., 2005) apart from survival, body size will remain of interest to malaria researchers (Mwangangi et al., 2004). For this study, however, the importance of body size to transmission rests on it having a significant effect on adult survivorship.

The two samples with the highest estimated survivorship had very large or very small size compared to the other samples. This indicates that there is a low linkage between adult body size and survivorship. The survivorship estimates are only rough, however, so this result is not conclusive. Independent from estimated survivorship, a better test is to compare adult body sizes of older and younger females within samples (villages). Results presented here suggest that nulliparous females are representative in size of females being produced in the rice fields. They are predominantly those that have not yet oviposited but will do so, and have been successful only in flying to the village and landing on a human to feed. Parous females are necessarily older and have been successful in feeding, laying eggs and returning to feed again at least once. If adult size is associated with survivorship and larger females have an advantage, then the size of parous (older) females

should be larger than that of the nulliparous (younger) females. Comparison of the sizes of these females showed no difference between older (parous) females and younger (nulliparous) ones. Overall, adult body size in these samples was not closely related to female daily survival – by either test – and so is unlikely to underlie the reduction in malaria transmission that occurs in irrigated villages during the rainy season (Sissoko et al., 2004).

The densities of adult mosquitoes observed from night catches are substantially less than those reported by others (Dolo et al., 1999, 2004). The largest number of bites per person per night recorded here was 57 compared to > 500 in the earlier study. While the rank order of density was similar in both studies, the reason for this difference in maximum density is not clear. Possibilities include the choice of collecting locations within villages, collectors with varying levels of experience or simple year-to-year differences in mosquito abundance.

Despite the low relationship between density and larval size, significant variation in adult body size between dry versus rainy season and irrigated versus non-irrigated zones was seen. Based on our results, this reflects varying larval sizes, but causes at the larval stage remain unclear. Others have observed a difference in size of adult *Anopheles dirus* between rainy and dry seasons, perhaps due to nutritional stress on larvae from the dilution of aquatic habitats or physiological changes with temperature (Kitthawee et al., 1992).

Species other than *An. gambiae* were excluded from our samples and so did not interfere with adult size measurements. *An. funestus* were not abundant and are easily distinguishable. Dolo et al. (1999) further reported that *Anopheles arabiensis* Patton, while present, is not common in this area. Within *An. gambiae* s.s., almost 100% have been reported as belonging to the Mopti chromosomal form over the year (Dolo et al., 1999). Consequently, our samples are very likely to be

almost all *An. gambiae* s.s., Mopti chromosomal form, with some *An. arabiensis* and other chromosomal forms present in very low numbers (Dolo et al., 2004).

Size does not appear to be the causal factor linking higher vector density to lower malaria prevalence through survivorship. Survivorship is generally an important determinant of vector competence, however, and may be affected by other factors in Niono. Of the hypotheses enumerated in the introduction and elsewhere (Dolo et al., 2004; Diuk-Wasser et al., 2005a), increased bednet usage with high mosquito density is a likely causative factor because anthropophily very clearly decreases with increasing density (Diuk-Wasser et al., 2005a). This factor alone may be sufficient to explain the reduction in malaria prevalence, or it may be a product of lower survivorship combined with increased zoophily in these anophelines.

On balance, the results presented here agree qualitatively with previous entomological studies in the Niono region (Dolo et al., 2004). They do not support the hypothesis that density effects on size are responsible for reduced malaria transmission in Niono through lower vector survivorship. A modest and unexpectedly positive effect of density on larval size is suggested by the data and a relationship between larval and adult size in *An. gambiae* were detected, but there was no evidence of reduced survivorship for smaller females in any of the samples. Previous findings about estimated survivorship were confirmed: during periods of high mosquito abundance daily survival is higher in the non-irrigated villages than in the irrigated area. It is also apparent from these results that the adult size of *An. gambiae* varies significantly between zones and between seasons.

Previous research indicated that female *An. gambiae* from non-irrigated villages are more effective vectors of malaria than those from the rainy season irrigated villages, when the number of mosquito bites per night per person can be

extremely high (Dolo et al., 2004). This study indicates this difference is not mediated by differences in body size produced by larval crowding.

CHAPTER 4

N_e and metapopulation structure of *Anopheles gambiae* s.s. in Banambani and its effect on transposable element dynamics

4.1 Introduction

The abstract ideal for most theory in population genetics and ecology is a panmictic population – a single, randomly mating group of individuals in which the population size stays constant, all offspring have an equal chance to reproduce and there is no geographic variation in gene or genotype frequencies. Like frictionless surfaces or perfect vacuums in theoretical physics, the ideal panmictic population plays an important role in theoretical population genetics, though no such ideal population actually exists. Rather, all real populations are in some way(s) structured. That is to say, they consist of subpopulations that are finite in size, may or may not mate randomly, may or may not share migrants among one another, and probably change size over time. Making inferences or predictions about these populations requires that their structure be understood and described in a manner commensurate with existing theory. The notion of

This chapter is a revision and extension of Taylor and Manoukis 2003

effective population size, N_e, plays an especially central role for describing this population structure. N_e may be defined as the size of an ideal population that exhibits the same rate of drift as the actual population it characterizes. As Futuyma (Futuyma, 1998) puts it: if we count 10,000 adults in a population but only 1,000 of them successfully breed, genetic drift proceeds at the same rate as if the population size were 1,000, and this is the effective size. Alternatively, as discussed later in this chapter, there are huge seasonal differences in the population size of *An. gambiae*. The harmonic mean of these sizes is N_e. N_e was introduced by Sewall Wright, with important subsequent contributions by G. Malecot, J. F. Crow, M. Kimura and others (Malecot, 1969; Wright, 1969; Crow and Kimura, 1970; Kimura and Ohta, 1971)

This chapter is concerned with the metapopulation structure of *An. gambiae* s.s. in a typical setting in Mali and how N_e can be used to make inferences about it with an emphasis on informing efforts to introduce a genetically modified vector to control malaria. After a few introductory remarks about taxonomic questions methods of inferring population structure and N_e are discussed. With these components to anchor the discussion the metapopulation structure of *An. gambiae* around Banambani, Mali is described together with simulations of transposable element spread through it.

The population structure of *An. gambiae* is quite complex. The highest taxonomic level in the system is *An. gambiae* sensu lato (s.l.) that comprises at least 7 species, one of which is *An. gambiae* sensu stricto (s.s.). *An. gambiae* s.s. in turn has as many as 5 different "chromosomal forms". In some locations (e.g. Mali), distributions of chromosomal forms coincide with molecular forms which can be distinguished with polymerase chain reaction (PCR) assays (Favia et al., 1997, 2001). This is not, however, the case in all locations (della Torre et al.,

2001). Gene flow across the forms is limited, so this is an important part of the population structure.

The principal focus in this chapter is on *Anopheles gambiae* s.s. in the village of Banambani, Mali, and its surrounding area, where three chromosomal forms are present - termed Bamako, Savanna and Mopti. There are at least three types of structure that relate to N_e and result in what may be best considered a metapopulation in this area. The first is the extent of population-size changes over the year and between years. The second is the structure imposed by chromosomal forms. Finally there is geographic structure because the species inhabits discrete patches (villages), among which gene flow is limited.

Metapopulation analysis holds much promise for furthering our understanding of structure in this species (Slatkin, 1977; Whitlock, 1999; Pannell and Charlesworth, 2000; Gavrilets et al., 2000). A metapopulation is a set of local populations, many of which may be unable to sustain themselves, where local extinction may be frequent, but where migration and recolonization can retain a dynamic equilibrium.

Several types of metapopulations can be distinguished (Hanski and Gilpin, 1997). Three of these have particular significance for the introduction of genetically modified *An. gambiae*: (1) The "classical" metapopulation of Levins (1969), a network of equivalent local populations that inhabit discrete patches. The probability of extinction is equal for all patches, as is the probability of recolonization and origin of migrants. (2) A mainland-island metapopulation, a system of habitat patches which are within dispersal distance of a very large (mainland) patch which never suffers extinction. (3) A source-sink metapopulation, where some patches have a negative growth rate (sinks) and are maintained by migration from patches with a positive growth rate (sources). Which patches are sources or sinks may vary seasonally or otherwise.

Describing the metapopulation of *An. gambiae* around Banambani as one of these types or as an intermediate will require careful application of direct and indirect methods of assessing structure. Genetic data can be used to test particular hypotheses about structure and type as an extensive body of theory now exists that describes how the patterns of gene frequencies and molecular evolution are determined by metapopulation structure. A better description of the importance of drift for the population can be attained by examining this type of question.

A significant understanding of the complex structure in the study area is critical for the successful introduction of genetically modified *An. gambiae*. Introduction of a gene for refractoriness to malaria (inability to transmit the disease) through genetic modification has been proposed as a control method (Aultman et al., 2001). A leading approach, examined in detail in this chapter, involves attaching the gene that confers a refractory phenotype ("effector gene") to a mobile genetic element known as a transposable element (James and Handler, 2000). This is argued to help the effector gene to attain a high frequency in the population despite a possible fitness detriment through the "meiotic dive", or non-Mendelian segregation, of the transposable element. Meiotic drive leads the transposable element to be found in more than half of the gametes of an individual that is heterozygous for the transposable element though its ability to replicate itself within the genome.

For such a scheme to be successful, however, many questions must be addressed before a release. It is important to know, for example, how many individuals will need to be released in order to alter the population, how fast the introduced genetic element will spread from the site of release and if it will move from one form or species into another. Even simple models for known features of

the metapopulation of this species lead to quite complex patterns of gene flow, as shown below.

4.1.1 Measuring population structure

Several methods have been used to measure the principal features of population structure, population sizes and migration rates. These have been classed as "direct" or "indirect" by Slatkin (Slatkin, 1985).

Direct measures of population structure consist of those measures of population properties themselves. For example, several groups have conducted MRR studies of adult female *An. gambiae* (Costantini et al., 1996; Touré et al., 1998b). Relative density can be estimated from how many bites a sedentary person receives during fixed periods (Service, 1993), and larval dipping have been used to estimate numbers of immatures (Service, 1993). Movement and migration are inferred by recaptures from distances or villages away from a release site and interbreeding between forms by the numbers of hybrids that can be detected. Service (1993) has an extended discussion about application of these methods for mosquitoes and several other authors have more general discussion on these topics (Blower et al., 1981; Johnson, 1969; Turchin, 1998).

Indirect measures of population structure are the methods of primary interest here. A particular history of size change and gene flow should give rise to corresponding patterns of allele frequencies in space and/or time. If those patterns can be appropriately measured, then it is frequently possible to work backwards and infer the population structure that gave rise to the pattern. A common approach is to estimate N_e as a starting point to understanding population structure and history (Table 4.1).

Briefly, in Table 4.1, n_e refers to effective number of alleles and θ is a measure

Table 4.1: Some indirect methods used to infer the structure and history of populations through patterns of genetic variation

Measure	Formulae	Reference*
1. Amount of polymorphism	$n_e = 4N_e\mu + 1$	1
	$\theta \approx 4N_e\mu$	2,3
2. Differentiation of populations		4,5
Island model	$F_{ST} = 1/(4N_em + 1)$	5
Islation-by-distance model	$\phi(x) \propto \frac{e^{-x(2\mu)^{1/2}}}{\sqrt{\sigma}}$	5
Stepping-stone model	$r(\rho) \propto \frac{e^{-\rho(4m_\infty/m_1)^{1/2}}}{\sqrt{\rho}}$	5
3. Temporal changes in frequency	$\sigma^s(p) \approx p(1-p)[1-(1-1/2N_e)]$	6,7
4. Coalescence	$tMRCA \approx N_em$	2,3

* [1] Hartl and Clark 1997; [2] Hudson 1990; [3] Fu and Li 1999; [4] Crow and Kimura 1970; [5] Kimura and Ohta 1971; [6] Anderson et al. 2000; [7] Waples 1989

of genetic diversity derived from allele frequencies by Watterson (1975), often used in coalescent theory. These both depend on the product of effective population size and neutral mutation rate, μ. There are several models of gene flow that have received theoretical attention; these differ about the importance of long-distance (m or m_∞), or short-distance (m_1) migration and whether there are discrete populations of size N_e or if they are distributed with a constant density. Distance, whether continuous (x) or discrete population steps (ρ), affects how correlated are populations to one another. F_{ST} is a correlation of allele frequencies across subpopulations, closely related to $r(\rho)$ and $\phi(x)$. Because large populations are expected to drift less than small populations, change in allele frequency (p) from generation to generation, measured by $\sigma^2(p)$, will depend on effective population size. And finally, when gene trees are constructed, the time, t, to the most recent common ancestor (MRCA) will depend on effective population size and mutation rate.

Patterns of detectable genetic variation shown in Table 4.1 are all sensitive to

the structure and history of the population being studied. Taken together with estimates of N_e, these provide insight into how populations of *An. gambiae* are structured, though there may be more than one interpretation for any particular pattern.

4.1.2 Estimating N_e

In the equations in Table 4.1 the estimates of effective population size can be taken as similar to the actual population size, N, but because of the extensive structure already discussed, N is likely to be a poor approximation of the effective population size. Recall from above that N_e is defined to be the size of an ideal population that exhibits the same rate of drift as the actual population (Crow and Kimura, 1970). Such drift might affect inbreeding, variance in gene frequencies, or rate of extinction of alleles. Consequently, one can distinguish inbreeding-effective population size, variance-effective population size, and eigenvalue-effective population size or extinction-effective population size (Crow, 1956). In most cases these will be similar to one another and no distinction need be made, but in some instances they can differ substantially (Kimura and Ohta, 1971).

An ideal population may drift at the same rate as the average drift of an actual population. For instance, When the population size is low, as at the beginning and end, then the allele frequencies will drift rapidly, but when the population size is large, then the drift will be slower. Between these extremes there must be an average rate that can be calculated by noting that the rate of drift at time t is proportional to $1/N_t$ (Figure 4.1). The average rate of drift, described by $1/N_e$, is seen to be:

$$\frac{1}{N_e} = \frac{1}{k}\left(\frac{1}{N_1} + \frac{1}{N_2} + \ldots + \frac{1}{N_k}\right) \tag{4.1}$$

Other adjustments can be made for other departures from the idealized, panmictic, population. For example when there are N_m males and N_f females then asymmetrical contributions to the next generation can be adjusted by:

$$N_e = \frac{4N_mN_f}{N_m + N_f} \tag{4.2}$$

How important unequal sex ratios are for *Anopheles* is, at present, unknown. We do not know how many males actually mate and this number might be highly skewed.

A third type of adjustment, for non-Poisson survival of offspring, is

$$N_e = \frac{4N - 2}{V_k + 2} \tag{4.3}$$

where V_k is the variance in numbers of offspring per parental pair. When survival of each egg has a Poisson distribution with mean 2, then $V_k = 2$. The true variance is probably much greater. Consider a common larval site for *Anopheles*, puddles. Survival of all the eggs laid there is hit or miss, in large part dependent on whether the puddle dries up or is washed away. Assuming that each female lays several eggs when she oviposits, then some few females who lay in fortunate sites will produce more offspring than others, thereby increasing V_k and decreasing N_e.

Note that these equations all refer to discrete generations, an idealization that is not really appropriate for *An. gambiae*. It is possible to make adjustments when a stable age distribution can be inferred, but a rough approximation – that N_e is roughly the number of mosquitoes born during some time interval that make

it to the age of reproduction times the number of time intervals during which they reproduce – is more practical. In general multiple considerations to correct N so it approaches N_e can be combined with one another, simply by taking a composite function (Kimura and Ohta, 1971).

4.2 Structure of *Anopheles gambiae* s.s. around Banambani

Mark-release-recapture (MRR) studies conducted at peak seasonal abundance in Banambani during a six-year period indicate that there is much variation in the size of local populations of *An. gambiae* s.s. Annual abundance estimates varied at least 4-fold, as shown in Table 4.2, from Taylor et al. (2001). Compounding the variability of population size between years, the density of *An. gambiae* s.s. changes dramatically within a single year. For example, a survey in non-irrigated villages near the Niono irrigation projects (Dolo et al., 1999) found peak numbers of bites in villages during the wet season to be more than 1000 times the number recorded during the dry season.

Table 4.2: Population sizes from 1993 to 1998 estimated from mark-release-recapture experiments at the village of Banambani. All measurements were made at the peak of seasonal abundance (August-September). Estimates from Taylor et al. 2001

Year	Replicates	Released	Recaptured	Daily Survival	Estimated N
1993	3	938	83	0.80	20,178
1994	4	1,913	57	0.80	64,002
1996	4	1,421	44	0.92	63,006
1997	2	1,002	24	0.97	53,400
1998	4	1,205	21	0.97	79,280

Seasonality seems to vary from place to place (see Chapter 5); near our study site in Banambani we have estimated dry-season population sizes to be 5 to 10% of the wet-season peak (Taylor et al. 2001, Touré, personal communication). Assuming a more or less exponential growth and decline between the seasonal low and highs, a typical annual cycle in the *An. gambiae* s.s. population at Banambani is depicted in Figure 4.1.

Incidence of the chromosomal forms of *An. gambiae* s.s. also vary during the year. All three forms, Bamako, Savanna and Mopti are about equally numerous during the wet-season peak, in late August to early September, but during the drier part of the year the Savanna and Bamako forms decrease in numbers, so Mopti individuals comprise nearly all of the *An. gambiae* s.s. that can be collected then. Based on the data of Touré et al. (1998a) the numbers of each form are approximately as shown in Figure 4.2.

This estimate is independent of the molecular form composition and population size figures given in Chapter 5, but is in good agreement with it. Recall that the S molecular form encompasses both the Savana and Bamako chromosomal forms when comparing them.

The three forms found at each location are thought to exchange genes among themselves, though how much and the significance of this is still unclear. Nonetheless, some exchange certainly occurs. Based on the number of hybrids observed, the number of migrants from one form to another is estimated to be in the order of 0 - 11% depending on the forms exchanging genes and the manner in which gene flow was estimated (Touré et al., 1998a; Tripet et al., 2001). There is currently insufficient information to determine whether gene flow is symmetrical or not.

Finally, there appears to be some movement of individuals among villages.

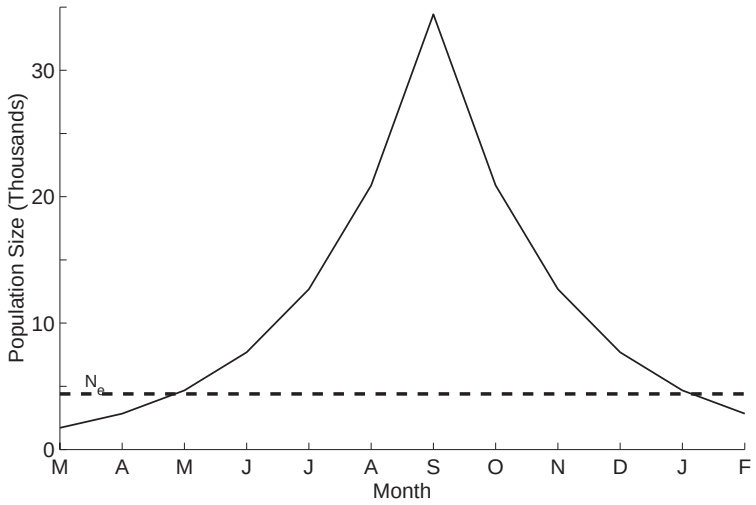

Figure 4.1: Estimated seasonal change in size of the adult population of *Anopheles gambiae* s.s. in Banambani, Mali. The broken line represents the value of N_e, calculated as per equation 4.1. Note that there are times when the actual and effective population sizes are equal.

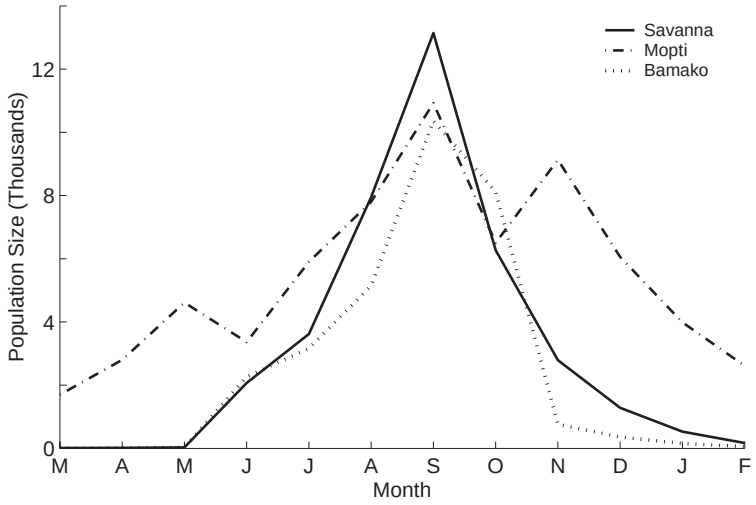

Figure 4.2: Estimated size of the adult populations of each of three chromosomal forms of *Anopheles gambiae* s.s. collected in and around Banambani, Mali.

Touré et al. (Touré et al., 1998b) and Dolo (Dolo, 2000) report on MRR studies where released females were captured in neighboring villages, but the numbers were low and the confidence limits large. As would be expected, there appears to be less gene exchange among villages farther apart than among those closely together (Johnson, 1969). Carnahan et al. (Carnahan et al., 2002) collated all of the available microsatellite DNA studies and found a linear relation between distance and log gene flow (i.e. $log(N_e m)$), as expected from theoretical considerations (see Table 4.1, and (Kimura and Ohta, 1971)). In this case m is the number of migrants entering a subpopulation/the size of that subpopulation. For distances a few kilometers apart, the typical inter-village distance around Banamabani, m is about 0.008-0.039.

4.2.1 Population structure visualization

Frames from a movie for visualizing the metapopulation of *An. gambiae* around Banambani can be used to illustrate its complexity (Figure 4.3). The series of images used to produce the movie include the best information available.

The series begins in the dry season (March) and proceeds through the following dry season (February). The population sizes of the three chromosomal forms are shown as discs, with the area proportional to the size of the population. The discs at each location represent (from bottom to top): Mopti, Savanna and Bamako forms. Locations in the map shown hosting populations are those sampled by Touré et al. (Touré et al., 1998b). Villages between them are not shown. All connecting pipes represent gene flow; white is between chromosomal forms and black is between villages. The diameter of the pipe is proportional to the magnitude of migration. The time bar in the upper right corner of each frame represents one year from March to March.

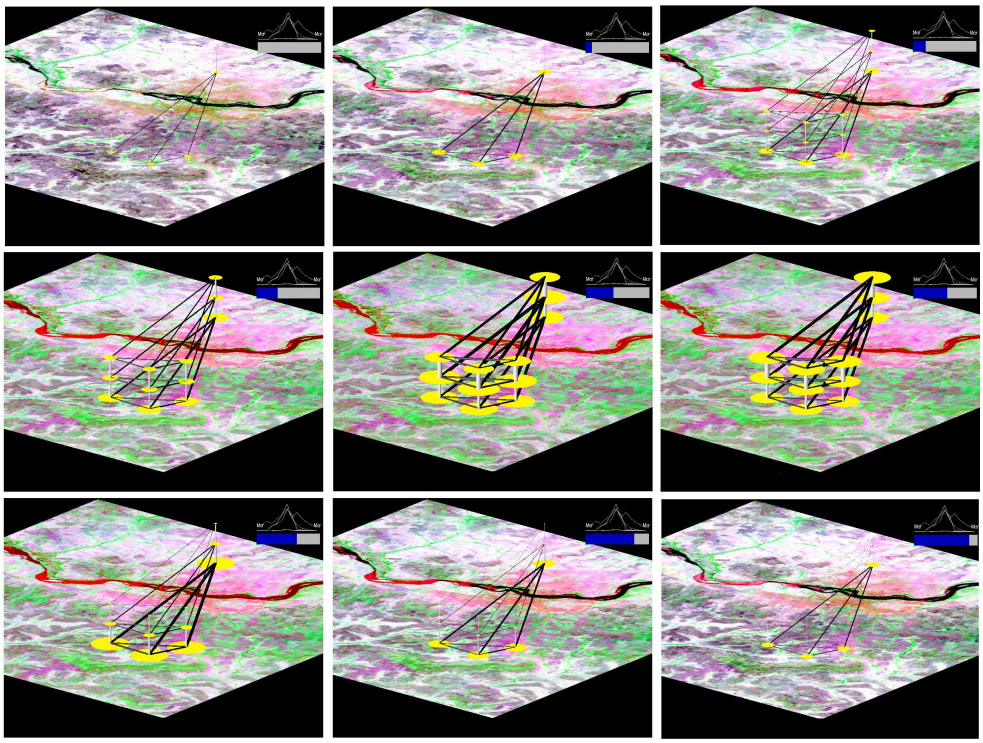

Figure 4.3: Summary of the population structure of the three chromosomal forms around Banambani, Mali. For details please refer to the text. The full movie from which these frames were excerpted is available online at http://taylor0.biology.ucla.edu/∼manoukis/Structure.

4.3 Metapopulation simulations

In this section the metapopulation of *An. gambiae* s.s. around Banambani is presented in two alternative forms, and the spread of a transposable element combined with an effector gene for refractoriness to malaria is simulated in the simplest terms. The conditions are oversimplified (Boete and Koella, 2002), but the purpose here is to explore the effect of structure on the movement of an ideal transposable element and gene for malaria refractoriness. The incorporation of more realistic transposable element dynamics to this simulation is a future goal.

4.3.1 *Anopheles gambiae* metapopulation parametrization

Combining the information in the preceding sections we get the following picture of genetic structure on *Anopheles gambiae* s.s. at our focal research site, Banambani village, Mali (Table 4.3). (Taylor et al., 1993; Lanzaro et al., 1998). The total effective population size is slightly larger than the sum of those for the forms it comprises. This is because of rounding error and because there are some unassigned or hybrid individuals.

Table 4.3: Estimates of effective population size presented in relation to some of the structure in *An. gambiae* around Banambani, Mali

Parameter	Notation	Value (Banambani)
Effective population size	$N_e tot$	4,400
	N_e, Bam	900
	N_e, Sav	1,500
	N_e, Mop	1,900
Migration to adjacent population	m_{ss}	0.008
	$N_e mBS$	16
Gene flow among forms	$N_e mSM$	12
	$N_e mBM$	2

The values in Table 4.3 vary in their reliability, particularly the numbers and migration rates when population sizes are small during the dry season. Frankham et al. (2002) report that an average ratio of N_e/N across many species is approximately 0.1, not far off from the ratio estimated from Tables 4.2 and 4.3. There is little allowance here for year-to-year variation, which we know to be substantial in our study area (Table 4.2). In addition to temporal structure, the population of this species around Banambani is structured spatially (between villages) and non- dimensionally (non-random mating due to chromosomal forms), in the terminology of Taylor and Powell (1983). This degree of complexity is hard to capture and quantify even with multiple patterns of genetic variation.

4.3.2 Levins metapopulation simulation

We have used computer simulation to explore how a transposable element might move through a classical metapopulation under very simple conditions. The outcome of one such simulation (termed Model 1) is shown in Figure 4.4, for a transposable element released at a frequency of 0.1 in the Mopti population of Banambani in March. This hypothetical transposable element is associated with a single gene which induces 100% refractoriness to *Plasmodium*. We assume that there is no dissociation of the transposable element and the gene of interest. In addition, the transposable element has a fixed level of meiotic drive so that heterozygotes would contribute $2p(1-p)(1+i)$ of the modified gene rather than the $2p(1-p)$ of Mendelian segregation. We took $i = 0.75$ for these simulations. We assumed that the transposable element and effector gene had no negative fitness effect on the carrier.

The status of the population at two-month intervals for two years is illustrated. The degree of red coloration in the figure shows what proportion of the population is carrying the transposable element. It is evident that incorporation of the transposable element is far from simple and depends critically on the population structure.

The simulation shows transposable element frequency increasing first in the Mopti chromosomal for in the release village then spreading to Savana in the same location (Figure 4.4). With the decrease in population size at the end of the first year the transposable element spreads to neighboring villages in both forms before increasing in Bamako at all four locations. The structure of the metapopulation had clear effects on the movement of the gene even under the condition of a highly idealized and unrealistic transposable element.

If i is decreased to 0.25 and 1000 individuals used for the release in June

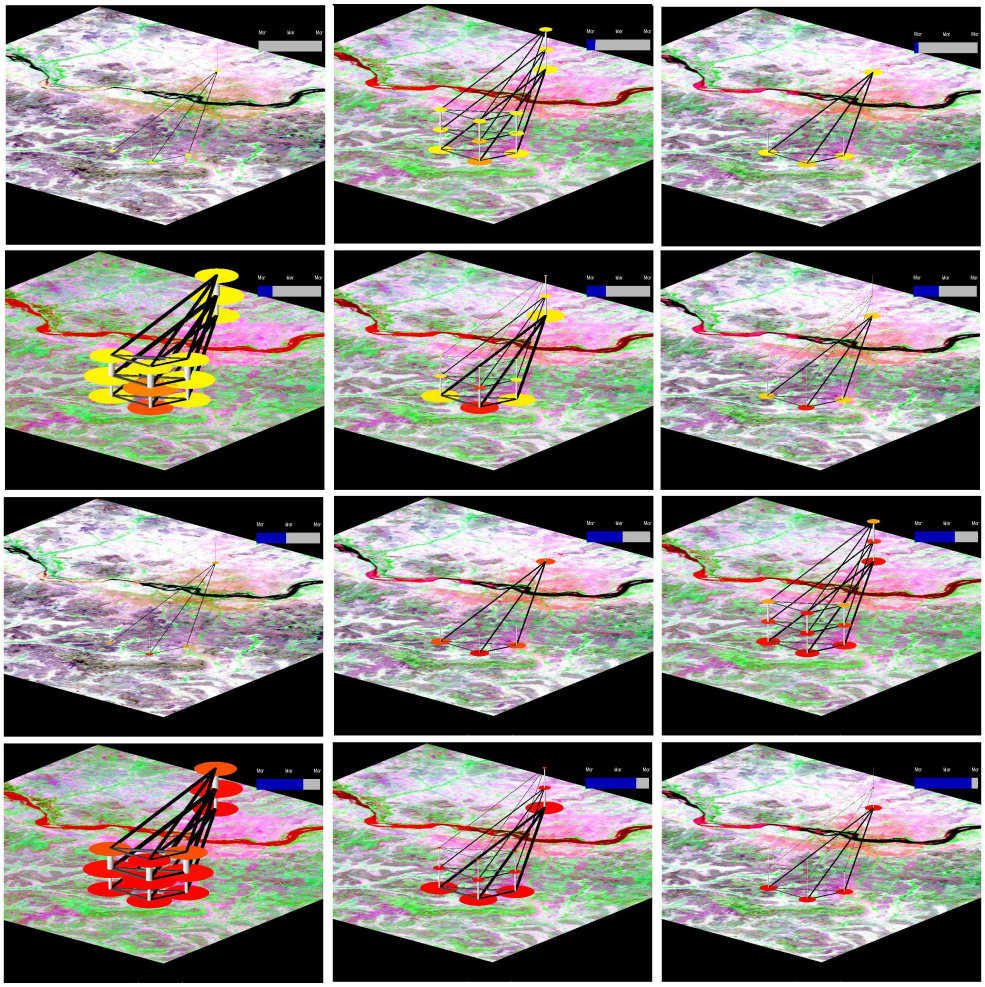

Figure 4.4: Simulated movement and frequency change of a transposable element through a Levins metapopulation of *Anopheles gambiae* s.s. around Banambani, Mali (Model 1). Symbolic representations of the metapopulation are as in Figure 4.3, with the addition of red coloration to depict the proportion of each subpopulation possessing the transposable element. The full movie from which these frames were excerpted is available at: http://taylor0.biology.ucla.edu/∼manoukis/Structure.

instead of a fixed proportion (these values are used for all subsequent simulations) the spread of the transposable element is slower, but follows the same course. This is shown in the results for Model 2 (Table 4.4). Changing the chromosomal form into which it is released has a qualitative effect, however. A Savana release (Model 3) will have the frequency of the transposable element increase in that form only for the first year, but after the bottleneck it actually decreases in Savana and becomes dominant in the Bamako form (Figure 4.5a). Even under the highly simplified conditions utilized here the unexpected outcome is that the transposable element "hops" from Savana to Bamako.

4.3.3 Source-sink metapopulation simulation

Further complications occur when the metapopulation is modeled with a source-sink dynamic. One of the four populations, the village of Sirakoro near the Niger river, was fixed at the wet season maximum population size. The form composition was varied according to the same proportions as the other three populations.

Two source-sink scenarios were simulated. The first was a release of 1000 Savana mosquitoes carrying the transposable element into a sink population (Model 4 – Figure 4.5b). In this simulation the frequency of the transposable element remained very low, and in a stochastic simulation would almost certainly have been lost. In the second source-sink simulation the same number of individuals were released into the source population's Savana group, with very different effect (Model 5 – Figure 4.5c). In this simulation the transposable element "jumped" to the Bamako form and spread to all four populations.

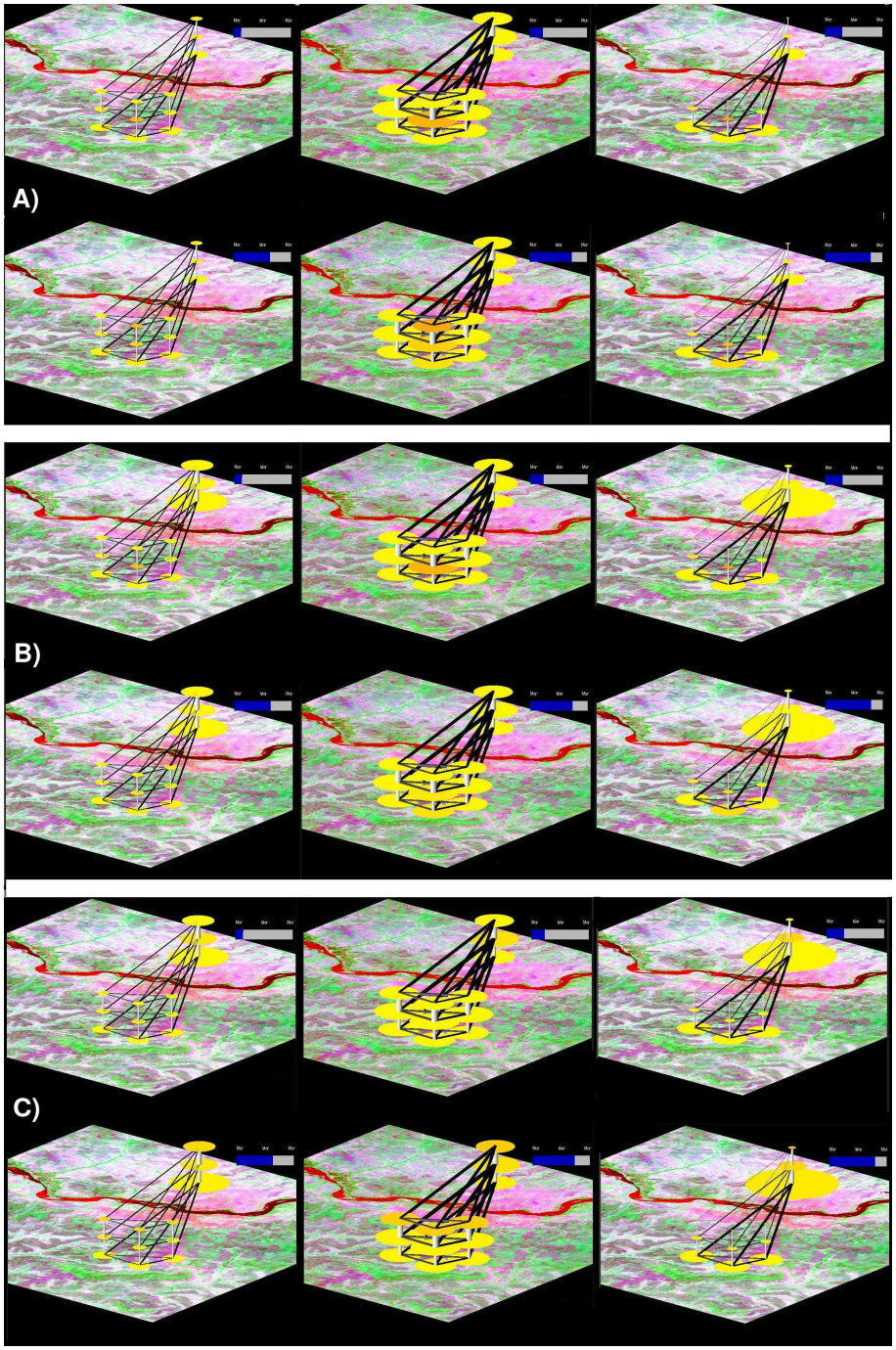

Figure 4.5: Simulated movement and frequency change of a transposable element through a Levins metapopulation of *Anopheles gambiae* s.s. around Banambani, Mali (Models 3 - 5). For a more detailed description of the models and results refer to the text and Table 4.4. The full movies from which these frames were extracted are available at: http://taylor0.biology.ucla.edu/~manoukis/Structure/Transposon_Movement.html

Table 4.4: Proportion chromosomal forms in simulated metapopulations carrying an idealized transposable element over time under five models. Columns show proportion of Mopti (M), Savana (S) and Bamako (B) form carrying the introduced transposable element across all populations. The "Time" column shows sequential months of the two years simulated.
Model 1: Levins dynamic, $i = 0.75$, release $p = 0.1$ of Banambani Mopti population size
Model 2: Levins dynamic, $i = 0.25$, release 1000 individuals into Banambani Mopti population
Model 3: Levins dynamic, $i = 0.25$, release 1000 individuals into Banambani Savana population
Model 4: Source-sink dynamic, $i = 0.25$, release 1000 individuals into Banambani Savana population (sink)
Model 5: Source-sink dynamic, $i = 0.25$, release 1000 individuals into Sirakoro Savana population (source)

Time	Model 1			Model 2			Model 3			Model 4			Model 5		
	M	S	B	M	S	B	M	S	B	M	S	B	M	S	B
June	0.089	0.049	0.000	0.057	0.000	0.000	0.000	0.081	0.000	0.000	0.044	0.000	0.000	0.059	0.000
Sept	0.208	0.146	0.006	0.094	0.002	0.000	0.002	0.120	0.005	0.002	0.118	0.005	0.001	0.045	0.002
Dec	0.304	0.258	0.034	0.140	0.009	0.000	0.006	0.161	0.021	0.004	0.088	0.011	0.005	0.166	0.018
Mar	0.472	0.376	0.139	0.194	0.046	0.003	0.013	0.180	0.059	0.003	0.012	0.003	0.018	0.362	0.084
June	0.804	0.657	0.309	0.250	0.215	0.011	0.025	0.047	0.102	0.012	0.011	0.005	0.023	0.058	0.177
Sept	0.984	0.943	0.570	0.313	0.272	0.033	0.046	0.084	0.156	0.039	0.032	0.010	0.027	0.070	0.289
Dec	0.999	0.994	0.904	0.398	0.344	0.081	0.081	0.139	0.218	0.040	0.037	0.021	0.086	0.200	0.445
Mar	1.000	0.999	0.995	0.517	0.453	0.174	0.135	0.198	0.287	0.029	0.024	0.036	0.222	0.402	0.607

4.4 Discussion

The importance of N_e lies in its utility to probe the structure of the population by allowing us to relate observed patterns of allele frequencies to particular aspects of population structure. This is critical for determining the pattern of transformation following the introduction of a genetically modified vector to reduce malaria transmission.

Both direct and indirect measures of gene flow will be necessary for estimates of population structure. They each give somewhat different information because they are based on different assumptions (Slatkin, 1987). Direct methods are necessarily bound by space and time and have low sensitivity, which is to say they are not well suited to detecting rare events. On the other hand, recorded events are real and may give insight into yearly cycles or median conditions, given enough observations. Indirect methods typically assume complete neutrality and equilibrium conditions, and typically give results that are open to multiple interpretations.

The use of both approaches helps to fill in the gaps left by the assumptions of one method or another, making estimates more robust. Especially in the case of rare events, which are of enormous importance to the prospects of genetic modification of *An. gambiae*, multiple sorts of evidence increase our confidence in the estimates.

With the prospect of introducing a transposable element into the population, simulation results presented above indicate that it is critical to understand more about the equivalency, permanence and demographic synchrony of patches. The results attained this far go part of the way to describing important characteristics of the metapopulation. The mathematical analyses and computer simulations

made to date are still only rough. Even in this basic form they demonstrate the qualitatively important effects of metapopulation structure on transposable element spread. Much more detail, especially with regard to population properties during the dry season, contributions from different locations where immatures develop and more information on mating behavior are needed to attain a level of resolution that will make modeling of the effect of such a transposable element effective and any introduction successful.

CHAPTER 5

Density and species composition of *Anopheles gambiae* s.l. in Banambani, Mali

5.1 Introduction

This chapter addresses issues of population size and species composition, which have important consequences for malaria transmission and control, by fulfilling two objectives. First, the size of the population of *An. gambiae* s.l. in Banambani is estimated over one year with special attention to the minimum size. Second, the species composition in the sample is examined with reference to climatic conditions, several variables of which are known to affect the species that make up the complex to differing degrees.

Villages in the Sudan savana region of Mali, such as Banambani, are at the northern edge of *Anopheles gambiae*'s range. These localities can experience mosquito vector densities more than 100 fold greater during the rainy season as compared to the dry season (Taylor and Manoukis, 2003). Such a massive increase and crash in population sizes can have important evolutionary consequences (Carson, 1970) and in the case of *An. gambiae* may drive speciation within the complex, enabling adaptation to new habitats (Ayala and Coluzzi 2005 and Chapter 6). Increased density may have complex effects on malaria transmissability and prevalence based on studies conducted in irrigated areas

(Chapters 2 and 3).

In addition to sheer vector numbers, population permanence during the dry season is of particular importance to malaria control schemes in at least three ways. 1) Extinction during the dry season is a core consideration for plans to control malaria transmission through the introduction of a genetically modified mosquito (Taylor and Manoukis, 2003). Introduction to a population sink could slow or foil such a plan (Chapter 3). 2) Detailed simulations of population dynamics for testing any intervention through simulations (e.g., Depinay et al. 2004) will be disproportionately affected by exactly what occurs during the dry season (Chapter 4). 3) The minimum size of populations may be a critical factor in speciation and adaptation within the *An.gambiae* complex (Coluzzi et al., 1985; Manoukis et al., 2007), which may have effects on patterns of resistance.

Malaria transmission is also affected by a characteristic change in the members of the *An. gambiae* s.l. (*sensu lato*) species complex that are present over one year (Touré et al., 1998a; Taylor et al., 1993). The vectorial capacity of these species is variable (White, 1978), so knowing which species are present at a particular time can in part explain transmission intensity.

Banambani is assumed to be a typical village in South-central Mali, although it has been more intensively studied than others in the area (Chapter 1). Dry season population sizes have been estimated to be $\sim 10\%$ those at the peak of the wet season in Banambani (Taylor et al., 2001). Persistence may occur through aestivation, gonotropic dissociation or reproductive quiescence (diapause) (Taylor et al., 1993) of adult mosquitoes. Alternatively, there may just be individuals surviving throughout the year, perhaps with a shortened life span during the times of harshest conditions. None of these possibilities or recurrent extirpation have been strongly supported despite many years of intensive study on the vector

populations in this area (Taylor and Manoukis, 2003).

The continued doubt surrounding the question of extirpation (Warburg and Touré, 2002) may be partially attributed to the difficulty in proving absence by direct observation for any species. In addition, there has been mixed evidence from other parts of Africa for aestivation in *An. gambiae*, with some evidence supporting the possibility (Minakawa et al., 2001) and other studies rejecting it (Charlwood et al., 2000) in the same part of the continent. Finally, there are some important biological properties of the species complex that make such an undertaking difficult. The complexity stems from the taxonomy of the species, which is not fully resolved.

Anopheles gambiae s.s. (*sensu stricto*) is one of the seven species of the *An. gambiae* s.l. species complex. It is made up of five "chromosomal forms" of uncertain taxonomic status (Coluzzi et al., 2002). These forms are distinguished by combinations of four paracentric inversions on the right arm of the second chromosome, which is thought to contain loci under selection (Turner et al., 2005). These forms differ ecologically (Coluzzi et al., 1979; Touré et al., 1994; Carnahan et al., 2002; Brooke et al., 2002) and hybridize only rarely (Tripet et al., 2001; della Torre et al., 2005), leading to speculation that they may be in the process of undergoing speciation (della Torre et al., 2002; Fanello et al., 2003; Manoukis et al., 2007).

Three of these chromosomal forms are present in Banambani, called "Mopti", "Savanna" and "Bamako" as is the sibling species *Anopheles arabiensis*, also a member of the sensu lato group. Mopti are known to be better adapted to arid environments and have previously been found to be more abundant during the dry season (Touré et al., 1998a). The other two forms may outcompete it during other times of the year or in different habitats. *An. arabiensis* also is adapted to

arid environments and is an important but secondary malaria vector compared to *An. gambiae* s.s. (Girod et al., 1999).

The Mopti chromosomal form can be distinguished from the other two on the basis of markers on their X-linked ribosomal DNA (Favia et al., 1997). From this "molecular" forms, named "M" and "S" were described (della Torre et al., 2001; Krzywinski and Besansky, 2003). This association breaks down in many other parts of Africa (della Torre et al., 2005) but in Mali it is possible to equate the M molecular forms with Mopti chromosomal arrangements (della Torre et al., 2001; Lozano-Fuentes et al., 2007).

5.2 Methods

5.2.1 Sampling

Male and female mosquitoes were collected through pyrethrum spray catch in Banambani by members of the Malaria Research and Training Center (MRTC) in Bamako, Mali following standard protocol (Service, 1993). These collections were conducted biweekly from 5 January to 29 November 2005. Females were scored for gonotrophic state and all mosquitoes were stored in tubes with silica gel desiccant.

A constant set of 20 houses was selected randomly for sampling. Members of the collecting team worked in a single group visiting these houses in a random order. When densities became too large for every specimen to be scored and preserved (23 June 2005) sampling was capped at 100 mosquitoes per collection trip. This was accomplished by sampling no more than 10 *An. gambiae* from any one house. If 100 total individuals were preserved the collection was halted and the remaining homes were not sampled. If fewer than 100 total *An. gambiae*

were collected then all 20 houses were sampled.

5.2.2 Genetic screening

Samples collected after 23 June 2005 were preserved in one tube per mosquito. Prior to this numbers were considered low enough for all the individuals from a single house to be saved in a single tube with silica gel. It was important to ensure that material used for PCR was only from a single individual so DNA was extracted from only the head and thorax of each mosquito, with care taken to avoid including any legs or parts thereof.

Molecular work was conducted by members of the laboratory of Greg Lanzaro at U.C. Davis. The DNAzol kit from Invitrogen Inc. (Carlsbad, CA) was used for all extractions. Once the tDNAs were available two diagnostic PCRs were carried out: The first was the Scott et al. (1993) method to distinguish *An. arabiensis* from *An. gambiae*; primers UN, GA and AR were used for this test. The second diagnostic followed the method of Favia et al. (2001) to distinguish the M or S molecular forms of *An. gambiae*.

5.2.3 Population size estimates

The change in sampling procedure after 23 June 2005 (see above) resulted in collections after that date not being strictly comparable to the ones before. To mitigate this problem an estimated sample size was produced based on the expected number of mosquitoes that would have been collected had the initial procedure been employed throughout. Two procedures were employed to obtain estimated sample size. First the houses that had numbers capped at ten individuals had those values replaced with the mean number of mosquitoes captured in houses with over ten individuals prior to 23 June 2005. Second, for the one week when

not all the houses were visited (3 August 2005) houses were divided into two groups: sampled and non-sampled. By multiplying the number of mosquitoes in the sampled houses by the ratio of non-sampled to sampled from all the other weeks the expected total number could be estimated. The estimated sample size is still expected to be an underestimate of the number of mosquitoes during the peak of the wet season, but represents an improvement over the uncorrected data.

Past research in Banambani has provided a number of estimates of maximum population size. The average of the mark-release-recapture (MRR) estimates listed in Taylor et al. (2001) indicates an average peak size of about 56,000 *Anopheles gambiae* s.l. in the village during the wettest part of the season. If we take the largest estimated collection number to equal a population peak of 56,000 total individuals the minimum (non-zero) collection size can be used to estimate the minimum total population size as a ratio of the maximum.

A moving average method was used to estimate the population size of the M and S molecular forms in Banambani. This average is based on estimates the total population size of *An. gambiae* s.l. for each week (N_{sl}) obtained as described above, utilizing a moving average of both the number of *An. gambiae* s.s. (N_{ss}) and the proportion of M and S molecular forms. The weekly estimate of (N_{ss}) was calculated as follows

$$N_{ss} = N_{sl} - \left(N_{sl} \frac{R}{R+G} \right) \tag{5.1}$$

where R is the number of *An. arabiensis* typed and G is the number of *An. gambiae* s.s. typed for the same week.

This estimated population size of *An. gambiae* s.s. was then multiplied by the proportion of M form (m) and S form individuals in the G sample to obtain

the sizes of populations of each form:

$$M = \bar{m}\bar{N}_{ss} \tag{5.2}$$

$$S = 1 - M \tag{5.3}$$

M and S are the estimated number of M and S forms, respectively. Symbols marked with a bar indicate that a moving average was used This means that for week i the moving average was the mean value x_i, x_{i-1} and x_{i+1} where x may be m or N_{ss}.

5.2.4 Climate data

Climate data from the Bamako - Senou meteorological station were used to compare with the sample sizes and species composition data. This station is 30 km from the village of Banambani and 16 km from Bamako (see Chapter 1). Comparison of data from this station with values from a weather station installed in Banambani after June 2005 revealed a close correlation (S. Lozano-Fuentes, pers. comm.). Averages for 14 days prior to the collection date were used for mean daily temperature and relative humidity. For precipitation the sum of the values from the same period prior to the collection were used.

5.3 Results

5.3.1 Population size

Population size peaked from about the end of June to November (Figure 5.1). The numbers reported are necessarily underestimates for the period between 23

June and 19 October 2005 for the reasons discussed above. Nine collections of 25 had to be adjusted to account for capping per house (maximum = 10) and one of these had to be additionally adjusted for incomplete sampling of houses. These procedures gave the estimated sample sizes presented in Figure 5.1.

If we take the maximum estimated collection size in these samples to be comparable to historically estimated maximum population sizes as described above, the minimum population size was of 1150 *An. gambiae* s.l. during the smallest non-zero collection.

5.3.2 Species and molecular form composition

An. arabiensis was present through most 2005 and outnumbered *An. gambiae* s.s. during January and November. The M molecular form also was present most often, but was less numerous than the S form when population sizes increased during the wet season (Figure 5.2).

The moving average of each form in Banambani (Figure 5.3A) closely tracks the averaged climate data (Figure 5.3B-D): the M molecular form was the only *An. gambiae* present when precipitation and humidity were low, and temperatures were high. The S form appeared and outnumbered the M form under the opposite conditions.

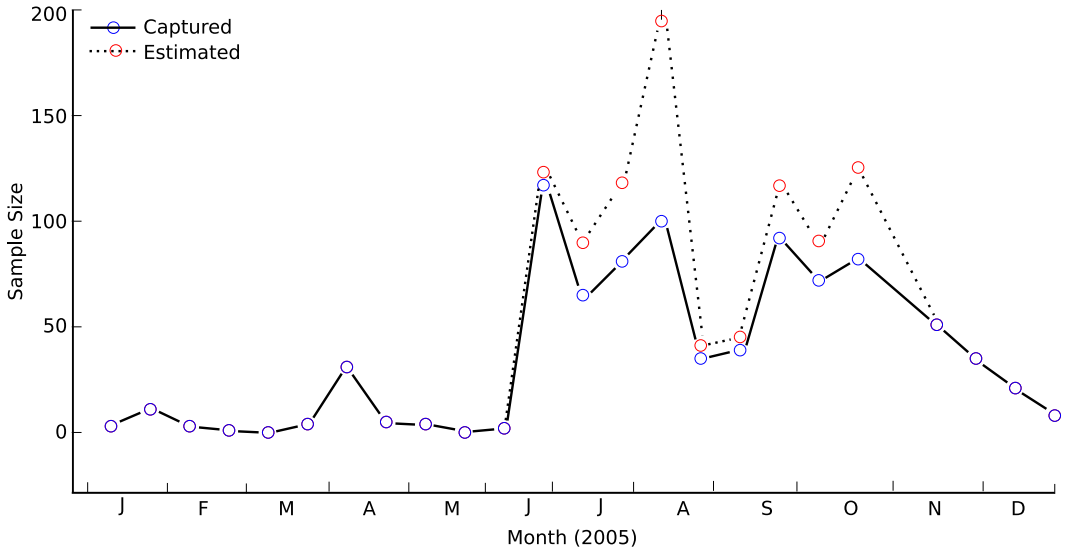

Figure 5.1: Numbers of *An. gambiae* s.l. captured and estimated sample sizes from Banambani during 2005. The estimated numbers presented are still likely to be an underestimate, see text for more details.

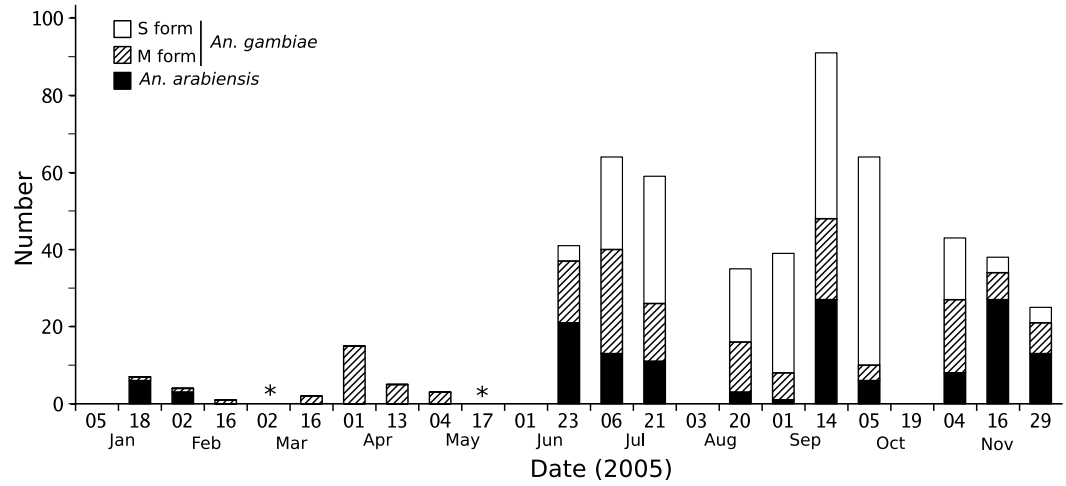

Figure 5.2: Numbers of *An. arabiensis* and *An. gambiae* M and S typed from 2005 collections. * collections that resulted in zero *An. gambiae* s.l. being captured; other collections may show no individuals because those collected could not be typed sucessfully.

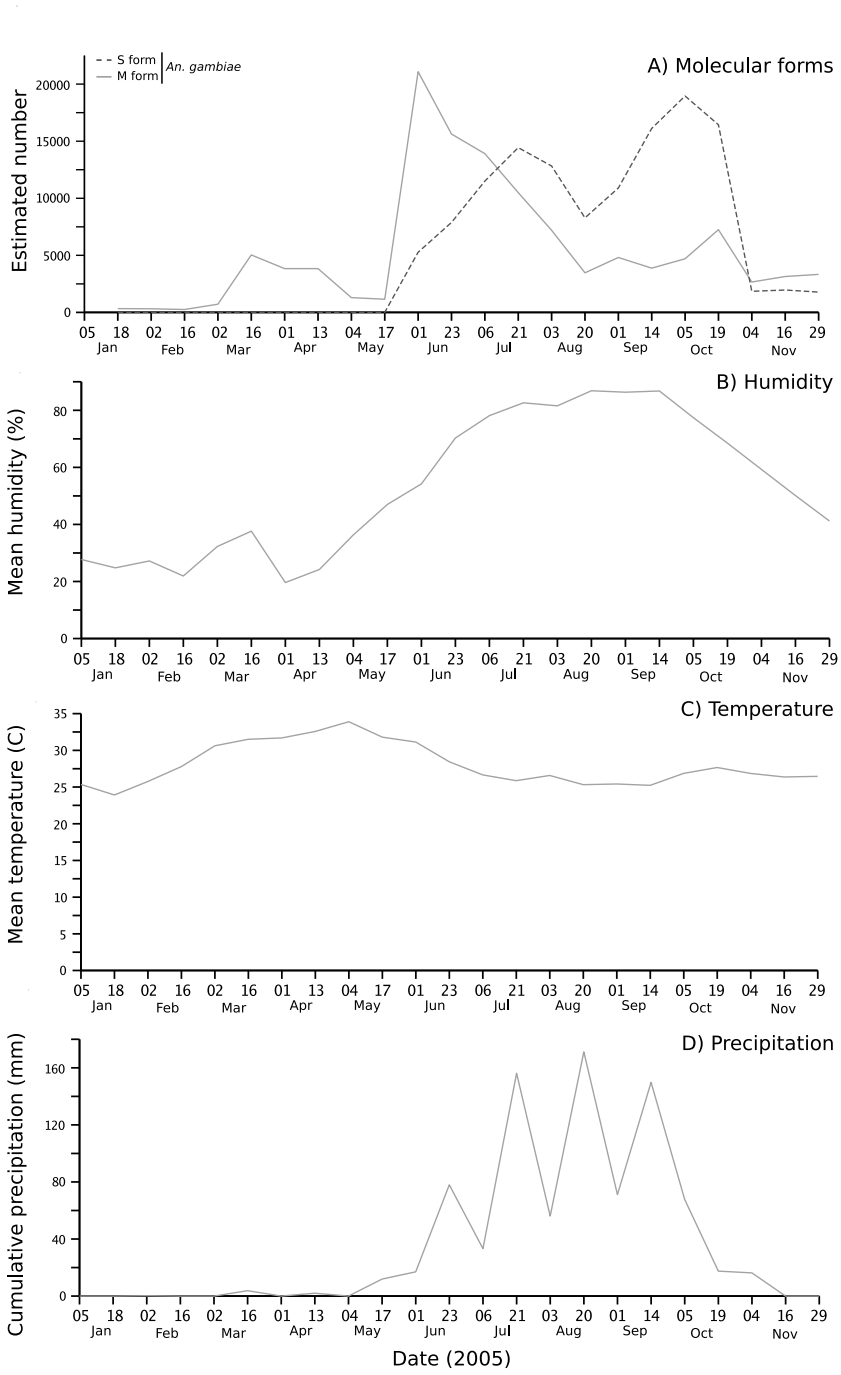

Figure 5.3: Estimated changes in population size of *An. gambiae* s.s. in Banamabani during 2005 with climate data. A) Population sizes and proportions of each form are moving average of each samples estimate and the previous and following sample (see text for further details. For B), D) values per survey date are averages of the previous 14 days, while for C) it is a sum for the same period. Climate data from the Bamako-Senou meteorological station obtained from http://tutiempo.net.

5.4 Discussion

A succession of species and forms was observed in Banambani. At the beginning of the year most of the individuals captured were *An. arabiensis* with few M form *An. gambiae*. As humidity begins to increase the M form is the only type found in the sample, around March. With increasing wetness in May the S form of *An. gambiae* makes its appearance, increasing in frequency to be more numerous than the M form and *An. arabiensis* during the wettest part of the year. This succession reflects the known adaptation of *An. arabiensis* and the Mopti chromosomal form to aridity compared to Savana and Bamako (Powell et al., 1999; Touré et al., 1998a). This difference is usually considered a consequence of adult tolerances to dessication (Coluzzi et al., 1985), likely due to higher body water content (Gray and Bradley, 2005). Other factors may also play a role, however, such as differences in breeding site specialization (Taylor et al., 1993).

The data suggest that *An. arabiensis* is present during the very driest times, though population sizes are small. This confirms indirect estimates of its persistence (Taylor et al., 1993). However, *An. arabiensis* may be competitively displaced by the M form (Touré et al., 1998a) once humidity increases just above the minimum. Similar collections in the village of Donegubougou found a co-occurrence of *An. arabiensis* and the Mopti chromosomal form between June 1998 and August 1999 (Warburg and Touré, 2002) so that competitive exclusion, if it is occurring, may not be a universal phenomenon in this area.

The results presented here (Figure 5.3) may be directly compared to an earlier attempt to estimate numbers of each chromosomal form over one year given in Chapter 3. Two distinct peaks of S population size are seen in this chapter, despite the extensive averaging. This may represent a transition between the Savana and Bamako form halfway through the wet season in agreement with the

results in Chapter 3. These chromosomal forms are both subsumed by the S molecular form, so this interpretation is speculative but consistent with previous observations.

Population sizes and form changes were similar to those estimated previously (Chapter 3). The minimum non-zero estimated population size was 1150 individuals, but this number represents only one individual captured from the 20 houses, and so is the limit of our resolution. The actual minimum population size of *An. gambiae* s.l., if it is non-zero, is likely to be smaller than that. It is impossible to ascertain with more certainty from these data what the minimum number may be.

The smallest population sizes closely track the dips in relative humidity (Figure 5.3B). The average number of mosquitoes collected during the driest months (January – April) was 7.25 for all 20 houses, while the number for the wettest months (May – November) was 57.42.

The results do not conclusively demonstrate that there is or is not annual extirpation of the *Anopheles gambiae* s.s. population in Banambani, but they do yield some insight. First, if there is extirpation and recolonization of the M form, it is unlikely to be a single event per season, based on the time between the two zero collections (2 March and 17 May 2005). Extirpation of the S form is more likely because they are absent from the collections from January to May.

Evidence for large scale estivation is lacking for *An. gambiae* in Mali. An exhaustive survey of the nearby ($\sim$ 2km) village of Donegubougou by Warburg and Touré (2002) revealed only 10 *An. gambiae* s.l. in about 400 nearby possible estivation sites, which did not support the hypothesis that there were large numbers of estivating *An. gambiae* around the village. Thus estivation seems unlikely in this area; rather, there is probably movement between villages at least for the

S form.

N'Gabakoro Doit, another local village, was sampled concurrently with Banambani, and the profile of species composition was completely different. *An. arabiensis* and the S form of *An. gambiae* were virtually absent year-round, and population sizes appear to have been much larger (unpublished data). These observations may be due to N'Gabakoro Droit's proximity to the Niger river.

The different species appear thus to be very sensitive to climate and geography and these can make for a completely different composition and extinction dynamics. Based on these observations it seems likely that the question of extinction and species composition is variable between villages.

Collections are ongoing in both Banambani and N'gabakoro Droit, which will allow for inter-year comparisons. From the comparison of these two villages plus Donegubougou (Warburg and Touré, 2002) it seems likely that more villages will have to be sampled before the question of local extinction can be fully addressed. Experimental techniques beyond passive collections will likely also have to be employed, such as multi village Mark-Release-Recapture (MRR) studies to assess the degree of movement and local population sizes.

The seasonal malaria transmission in Banambani, which peaks during the wet season, is consistent with the large population size and predominance of *An. gambiae* s.s. during that time. Further, consideration of species composition and density as indicated by this study will be important to intervention efforts using genetically modified mosquitoes or other methods.

CHAPTER 6

Speciation by ecotypification in *Anopheles gambiae*: A simulation study

6.1 Introduction

6.1.1 Inversion polymorphism and speciation

Chromosomal inversions have been considered precursors of speciation in at least two ways. The first is through underdominance of the heterokaryotype (e.g., in stasipatric speciation, see White 1978). A major difficulty of that type of model, remarked by several authors (e.g. Bengtsson 1985; Coyne and Orr 2004; Ayala and Coluzzi 2005), is that if there is enough underdominance for reinforcement to evolve, an inversion is unlikely to ever establish itself, especially since it must first occur in a heterokaryotypic individual. This and other problems of underdominance based mechanisms (Navarro and Barton, 2003a) have led to their limited acceptance among evolutionary biologists.

The second reason chromosomal inversion polymorphism may be found between species is that recombination suppression due to the inversion may cause speciation (Dobzhansky, 1947; Trickett and Butlin, 1993; Navarro and Barton,

The work in this chapter has been submitted for publication (Manoukis et al., 2007)

2003a). Speciation ultimately results because the region of suppressed recombination creates a genetic barrier allowing divergence (Ortiz-Barrientos et al., 2002). There is good empirical evidence for this mechanism in sunflowers (Livingstone and Rieseberg, 2004), *Drosophila* (Noor et al., 2001) and humans (Navarro and Barton, 2003b; Zhang et al., 2004). Unlike underdominance based models, suppressed recombination based models of reproductive isolation impose no specific difficulty to the establishment of an inversion. Thus mechanisms leading to inversion polymorphism can be usefully investigated separately from the actual evolution of reproductive isolation.

The question of how chromosomal inversions are established and maintained has received limited theoretical attention until recently (Kirkpatrick and Barton, 2006), and has not been quantitatively tested in a realistic evolutionary context. A taxon where the evolution of chromosomal inversion polymorphism is of special significance is the *Anopheles gambiae* species complex (*Anopheles gambiae* sensu lato, hereafter *An. gambiae* s.l.) because of the many inversions observed, described below. In addition, *Anopheles gambiae* is a candidate for studying the question of chromosomal inversion establishment quantitatively owing to the many years of research on this species.

6.1.2 "Ecotypification " in *Anopheles gambiae*

Anopheles gambiae is the principal vector of malaria in Subsaharan Africa. *An. gambiae* s.l. is considered to consist of seven species: *An. arabiensis*, *An. merus*, *An. melas*, *An. quadriannulatus* A, *An. quadriannulatus* B, *An. bwambae* and *An. gambiae sensu stricto* (hereafter, *An. gambiae* s.s.). There appears to be strong reproductive isolation among each of these "good" species (Black and Lanzaro, 2001). Within *An. gambiae* s.s., however, there is a complex pattern of

reproductive isolation that is poorly understood.

Five "chromosomal forms" of *An. gambiae* s.s. have been distinguished based primarily on non-random associations of inversions of the 2R chromosome. These forms have been named 'Mopti", "Savanna", "Bamako", "Forest" and "Bissau"(Coluzzi et al., 1985), each comprising several inversion karyotypes. They show somewhat different patterns of geographic distribution and varying ecological preferences (Coluzzi et al., 1979; Carnahan et al., 2002; Brooke et al., 2002), though in some places as many as three are sympatric with limited gene flow among them (Touré et al., 1983).

Using a different criterion, ribosomal DNA genes on the X chromosome, it is possible to distinguish "molecular forms" of *An. gambiae* s.s. (Favia et al., 1997; della Torre et al., 2001) that correspond well with the chromosomal forms – "M" (corresponds to Mopti) and "S" (corresponds to Bamako and Savanna) (Gentile et al., 2002) – in some of the geographic range of the species, though not all of it (della Torre et al., 2005).

While one might be tempted to infer that the M and S molecular forms are two incompletely isolated species distinguished by their X chromosome, or that the chromosomal forms are five incompletely isolated species distinguished by their second chromosome, the lack of consistent associations over the range argues against this. Instead, it suggests that the gene pool of *An. gambiae* is in flux, with metaphorical eddies, currents, counter currents and backwaters (Coluzzi et al., 1985; Tripet et al., 2001; della Torre et al., 2002; Fanello et al., 2003; della Torre et al., 2005).

Coluzzi (1982) has suggested that this complexity might reflect ongoing, incomplete speciation associated with the widespread chromosomal inversions in this species group through a suppressed recombination mechanism. His argu-

ments have been verbal and based on observation. They may explain the documented ecological adaptability of this species.

In an effort to understand how such complexities arise we have focused specifically on the Bamako chromosomal form by hypothesizing on how its characteristic inversion may have arisen in West Africa and started it on the path to becoming a separate species. As a shorthand, we refer to the process as "ecotypification" (M. Coluzzi, pers. comm. 2005), associated with specialization to particular niches near the edge of the species distribution.

Ecotypification in *Anopheles* may be described as follows: When conditions are favorable, mosquitoes are able to increase their range and expand into habitats in the periphery of their range. New alleles may arise in this peripheral area that allow better utilization of that environment. These may rise in prevalence as conditions become harsher. If those alleles are captured by a region of chromosomal inversion then they will be protected from genetic homogenization by suppression of recombination with maladapted alleles. Other adaptations that favor the mosquito in the peripheral environment may then accumulate in or near the region of suppressed recombination and this accumulation may eventually lead to the evolution of reproductive isolation.

This model of inversion polymorphism leading to reproductive isolation is similar to others (Kirkpatrick and Barton, 2006). It differs significantly in that there is a large change in population size during the year. The explosive increase in numbers of *An. gambiae* during the wet season seen in many parts of Africa (e.g. Taylor and Manoukis 2003) is important because it allows peripheral populations to be effectively colonized then become isolated. It also will affect the degree of random genetic drift experienced.

As mentioned above, inversion polymorphism is proposed to arise through an

environmental difference between the core and peripheral habitats. The peripheral habitats in the case of the Bamako form may be laterite pools in the rocks of the Niger river in Mali and Guinea. Coluzzi proposes that a specific adaptation of the Bamako form is a particular "hooking" behavior by the larvae which prevents them from being washed away from these rock pools during periods of high water flow (Coluzzi, pers. comm. 2005).

Inversion polymorphism is not the same as speciation (Maynard Smith, 1966). We have purposely avoided including genes for reproductive isolation in this work because to do so would be speculative in the case of *An. gambiae*, despite recent advances (Turner et al., 2005; Stump et al., 2005). We note, however, that these recent results are fully consistent with the process of ecotypification as described, as are the well documented differences in ecological adaptations between chromosomal forms (Coluzzi et al., 1977; Touré et al., 1998a). A condition of inversion polymorphism is an essential precursor to the differentiation and isolation between forms through suppressed recombination that is now being elucidated at the level of DNA sequences.

The objectives of the work presented here are 1) to use numerical, stochastic simulations to examine if ecotypification is a sufficient explanation for how inversion polymorphism arises in *An. gambiae*, and 2) if so, use the simulation model to determine which factors most affect the maintenance and final frequency of chromosomal inversions. We find that inversion polymorphism is not a rare outcome under parameter ranges realistic for *An. gambiae* and that several factors can contribute to its establishment. In order of importance they were found to be random genetic drift and selection followed by migration rate and recombination rate. Once an inversion is established its final frequency was affected only by migration rate and selection.

6.1.3 Model outline

We have created a "minimal model" (Roughgarden et al., 1996) that includes the essential elements of the ecotypification hypothesis and their relation to inversion polymorphisms of *An. gambiae* using parameters based on field and laboratory studies. We simulate two populations of mosquitoes connected by migration. The first population represents the core portion of the range and supports a large number of individuals. The second population represents a marginal habitat supporting fewer individuals. Each mosquito has a diploid genome of 15 equally spaced bi-allelic loci under selection. Over one year in the simulation (10 generational time steps) population size changes from a dry season minimum (S_0 for core and S_1 for peripheral) to a wet season maximum ($100 \cdot S_0$ and $100 \cdot S_1$) then back down to near minimum size. The evolutionary forces represented in the simulation are selection (s), recombination (r) and migration (m) between populations. For further details and symbols refer to Table 6.1 and Appendix B.

Each simulation starts with randomly chosen alleles for all individuals then is run for 289 generations to attain migration-selection balance. During the year after this time (generations 290 to 299) an inversion is introduced at a single time step (i.e. population size) t_i. This inversion always is introduced to the peripheral population and captures a set number of adjacent loci n_i. All the loci captured have alleles adapted to the peripheral habitat. The simulation is then extended to generation 600 and the frequency of the inversion in both populations is recorded. The final outcome is either inversion polymorphism (frequency of inversion > 0) or inversion loss (frequency of inversion $= 0$).

To experiment with the system, we varied the input parameters over ranges taken from previous studies of *An. gambiae* (Table 6.1). Each parameter was uniformly sampled and simulations were algorithmically evenly distributed over

the parameter space. Parameter values and inversion frequency outputs were recorded for thousands of simulations providing data for statistical analysis of the relationship between them.

Table 6.1: Model parameters. Parameters required for each simulation were chosen from the ranges indicated below. S_0 and S_1 are the minimum population sizes with the maximum always 100 times that value yielding effective population sizes (N_e) noted. m_x is the probability of selecting a mate at random from the other population. s is the fitness of a uniformly maladapted genome relative to an adapted one. t_i covers all the population sizes just before generation 300. n are the number of loci captured in an inversion. r is the probability that a break point occurs between adjacent loci.

Parameter	Abr.	Range	Sources*
Core population (P_0) min. size	S_0	200 – 600 ($N_e = 1648 - 4945$)	1 – 5
Peripheral population (P_1) min. size	S_1	5 – 200 ($N_e = 41 - 1648$)	
Migration rate (P_1 to P_0)	m_0	$0.001 - 0.01$	3, 6 – 11
Migration rate (P_0 to P_1)	m_1	$0.001 - 0.5$	
Relative fitness of maladapted allele	s	0.1 - 1.0	1, 12 – 14†
Time/pop. size when inversion occurs	t_i	$290 - 299$	N/A
Number of loci captured by inversion	n_i	1 - 15	15
Recombination rate	r	0.006 - 0.06	15, 16

* [1]Touré et al. 1998a; [2] Pinto et al. 2003; [3] Taylor et al. 2001; [4] Charlwood et al. 2000; [5] Lehmann et al. 1998; [6] Costantini et al. 1996; [7] Besansky et al. 1997; [8] Carnahan et al. 2002; [9] Onyabe et al. 2003; [10] Tripet et al. 2005; [11] Kayondo et al. 2005; [12] Touré et al. 1994; [13] Onyabe and Conn 2001; [14] Bayoh et al. 2001; [15] Turner et al. 2005; [16] Coluzzi et al. 2002. †, these sources give qualitative comparisons, so we adopted a large range.

6.2 Results

6.2.1 Model validation

We tested the model's response to several single parameters and compared simulation results against expectations from Kirkpatrick and Barton (2006) before varying them in tandem. We individually tested migration rate, recombination rate and number of loci captured by the inversion. In all cases the model's behavior conformed to expectations, supporting the assumption that it captures the essential elements of the ecotypification hypothesis.

As an example, Figure 6.1 shows the results of some tests on migration rate. We used a peripheral population size of 500 individuals with two loci under selection. These individuals may breed with migrants, which are maladapted to the peripheral environment (fixed for the disadvantageous allele at all loci). An inversion is introduced encompassing both loci. We found that higher migration rates lead to a faster increase in inversion frequency as expected from Kirkpatrick and Barton's work: the selective advantage of not having recombination with maladapted (migrant) individuals is higher when there are more migrants (Figure 6.1). Inversion frequencies stabilize at varying levels depending on the degree of influx of wild type individuals.

6.2.2 Factors leading to inversion polymorphism

We wished to determine whether inversion polymorphism was sustainable in *An. gambiae* over what are thought to be realistic parameter ranges. Since stochastic factors are expected to be important, we ran experiments with 1000, 2500 and 5000 simulations under the same parameter ranges. These had 7.20, 7.44 and 7.7% of their runs end with inversion frequencies that were greater than zero,

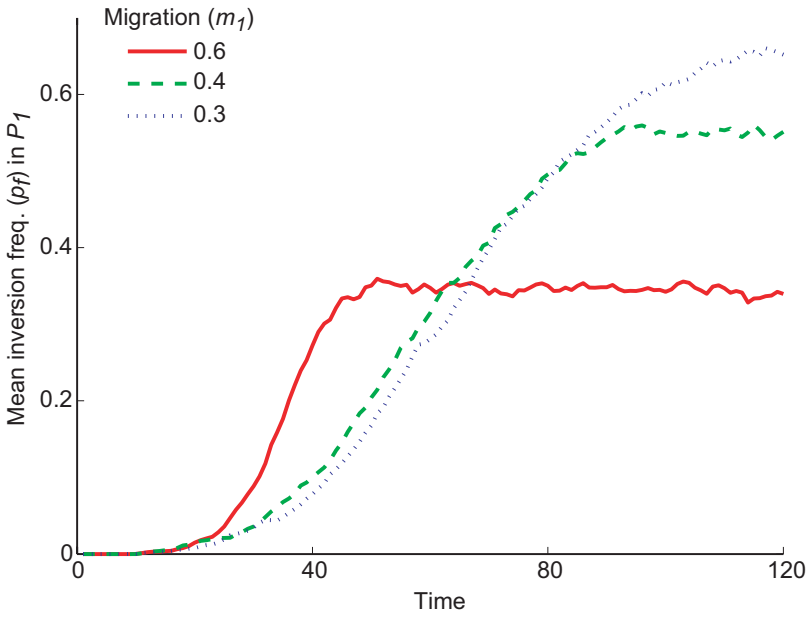

Figure 6.1: Higher migration leads to higher rate of increase in inversion frequency. Each line is the mean frequency of inverted chromosomes in the peripheral population from ten runs where the inversion is not lost. To be comparable with the results of Kirkpatrick and Barton (2006) all migrants are maladapted to the peripheral environment (so S_0 is irrelevant) and fitness is calculated multiplicatively. Other parameters: $r = 0.5$, $t_i = 10$, $n_i = 2$, $m_0 = 0$, $S_1 = 500$ and $s = 1$.

respectively. The inverted karyotype never reached fixation in either population for any simulation, so all simulations with inversion frequencies (p_f) > 0 can be considered to have had a polymorphic outcome. For all subsequent analysis we have used the dataset with 5000 simulations to maximize statistical power, though the results were qualitatively the same as for the 2500 simulation experiment.

Several factors affected whether inversions could persist. To determine what they were and their relative importance, we coded runs with $p_f > 0$ as 1 (polymorphism: $p = 1$) and those where the inversion was lost as 0 (loss of inversion: $p = 0$. The results of a binomial regression of parameters varied on p showed that all but S_0, m_0 and s had a significant effect on p (Table 6.2). This out-

come was confirmed with a classification analysis by classification and regression trees (CART: Breiman et al. 1984) (results not shown). The CART analysis did not detect any higher level interactions, so the regression model appropriately describes the data.

Table 6.2: Factors that lead to inversion polymorphism. Binomial regression† and analysis of deviance of model parameters. For analysis of deviance terms were added sequentially in the order presented below.

	Binomial regression				Analysis of Deviance	
Coef.	Estimate	Std. Error	z	P	Resid. dev	% Deviance
Int.	192.2	10.8	17.88	< 0.0001	2713.80	
S_0	-5.4×10^{-4}	5.5×10^{-4}	-0.98	0.33	2713.05	0.03
S_1	-2.7×10^{-3}	1.1×10^{-3}	-2.39	0.02	2710.13	0.11
m_0	-361.0	243.2	-1.48	0.14	2709.60	0.02
m_1	4.4	0.5	9.54	< 0.0001	2630.48	2.82*
r	38.4	3.9	9.82	< 0.0001	2536.10	0.04*
s	7.0×10^{-4}	2.4×10^{-2}	-0.29	0.77	2535.00	0.04
t_i	-0.6	3.7×10^{-2}	-18.34	< 0.0001	1945.47	21.72*
n_i	0.2	1.7×10^{-2}	13.67	< 0.0001	1714.76	8.50*

* terms with $P(\chi^2) < 0.05$; †Null deviance = 2714 on 4999 d.f.; residual deviance = 1715 on 4991 d.f.

An analysis of deviance on the binomial regression model was used to rank the importance of the variables. It revealed two classes of significant predictors (Table 6.2). t_i and n_i explained much more of the deviance, while r and m_1 had relatively minor roles. All the other factors were not significant at $\alpha = 0.05$ in the analysis of deviance.

Examination of t_i and n_i individually shows that inversions persisted most often when introduced to a small population (Figure 6.2A) and when the inversion encompassed more loci (Figure 6.2B).

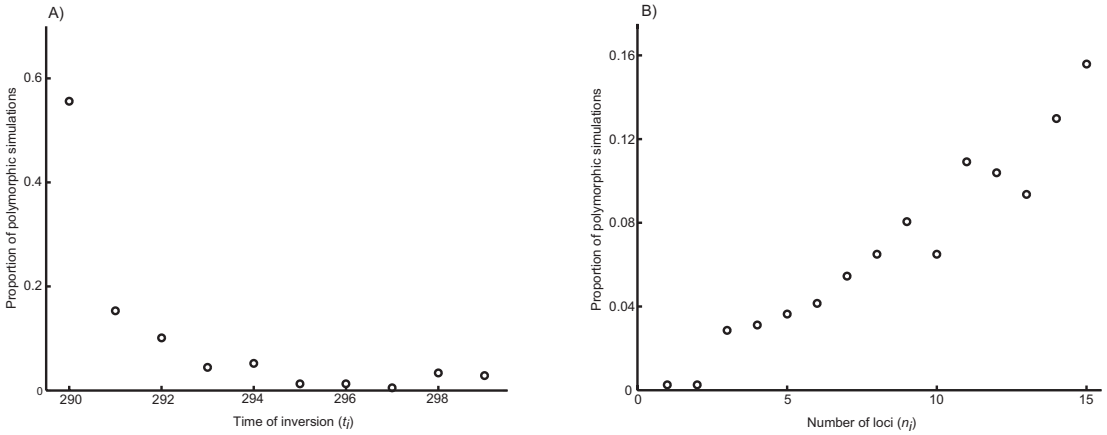

Figure 6.2: Main factors affecting inversion polymorphism. Proportion of simulations ending in inversion polymorphism at each of the levels of A) time of inversion introduction (t_i) and B) number of loci captured in the inversion (n_i). Note that for A) S_1 is at its minimum at time step 290, increases to a maximum of 100 x that by step 295 then decreases again through step 299. The total number of simulations where $p = 1$ was 385.

6.2.3 Factors affecting final inversion frequency

We performed a second analysis restricted to those simulations which ended in inversion polymorphism. With this test we created a conditional model that has a simple interpretation (Welsh et al., 1996). Those runs with $p = 1$ yielded a dataset of 385 simulations with inversion frequencies in the peripheral population that were approximately normally distributed ($\bar{p_f} = 0.57$, SD(p_f) = 0.17; one sample Kolmogorov-Smirnov test, $D = 0.05$, $p = 0.25$). The inversion frequencies in the core population were significantly lower ($\bar{p_f} = 0.0006$, SD(p_f) = 0.001; Mann-Whitney U test, $W \sim 0$, $p < 0.001$) and not normally distributed (one sample Kolmogorov-Smirnov test, $D \sim 1$, $p < 0.001$) compared to those in the peripheral population.

Linear regression of the predictors on p_f for the peripheral population showed m_1, s and n_i to be significant determinants of final inversion frequency (Table

6.3).

Table 6.3: Factors affecting final inversion frequency in the peripheral population using runs ending in polymorphism.*

Coefficient	Estimate	Std. Error	t	P
Intercept	-8.5×10^{-2}	0.7	-0.12	0.90
S_0	1.7×10^{-5}	4.5×10^{-5}	0.38	0.71
S_1	8.2×10^{-5}	9.2×10^{-5}	0.89	0.38
m_0	12.0	20.4	0.59	0.56
m_1	-0.9	4.2×10^{-2}	-22.22	< 0.0001
r	0.5	0.3	1.44	0.15
s	2.8×10^{-2}	2.4×10^{-3}	11.78	< 0.0001
t_i	2.1×10^{-3}	2.3×10^{-3}	0.91	0.36
n_i	1.5×10^{-2}	1.5×10^{-3}	9.86	< 0.0001

* Residual standard error = 0.104 on 376 d.f.; multiple $r^2 = 0.64$; $F = 82.33$ on 8 and 376 d.f., $P < 0.0001$

6.3 Discussion

Using our simulation as an experimental system (Peck, 2004) we have found that inversion polymorphism is possible under conditions that exist for *An. gambiae*. Furthermore, we were able to determine the relative importance of factors that affect whether polymorphism is maintained and speciation can occur. These results rely on the simulation capturing essential elements of the real0-world situation. Based on the description of ecotypification (Coluzzi, 1982) and theoretically expected behavior (Kirkpatrick and Barton, 2006) the model seems to conform to that assumption.

Population size was the single most important factor affecting eventual polymorphism. This is reflected in the statistical significance of both the absolute size of the peripheral population and the time at which an inversion occurs there.

Smaller populations had inversion polymorphism persist more often as did populations that had an inversion introduced when they were at their minimum size. If an inversion is introduced to a large and/or shrinking population it has a greater chance of being lost despite selection. This indicates that random genetic drift, intensified by population size fluctuations, is the most important factor contributing to the establishment of inversion polymorphism in *An. gambiae*. This will be the case in other species with a small effective population size in the peripheral habitat.

In our simulations stochasticity due to drift is much more important than parameters that have been theoretically shown to be significant (Kirkpatrick and Barton, 2006) to inversion maintenance: migration rate into the peripheral population and recombination rate. However, once an inversion rises in frequency enough to overcome loss by chance, the simulation shows the expected factors affect final frequency: rate of in migration of non-inverted individuals, selective difference between habitats, and the combined selective effect of loci within the inversion.

The second most important parameter to inversion polymorphism in our simulations was the number of loci captured by the inversion. More generally, this is the combined selective advantage of the loci within the inversion. This result is expected, highlighting that genomic regions that are candidates for inversion polymorphism and eventual speciation are likely to have multiple loci under directional selection in close proximity (Turner et al., 2005).

When the introduced inversion persisted it was at a significantly higher frequency in the peripheral population than in the core one. A comparison of mean F_{st} between simulations with inversion polymorphism and others with the same parameter set, but no inversion, revealed higher differentiation when an inversion

was present (results not shown). This differentiation indicates a barrier to gene flow, setting the stage for reproductive isolation by various mechanisms (Coyne and Orr, 2004). In the case of the Bamako form of *An. gambiae* s.s. isolation is likely to depend on specifics of its larval ecology.

Larvae of the Bamako form often display behavior different from larvae of the other forms, most notably a hook-like posture that appears to anchor them better in laterite pools (M. Coluzzi, pers. comm. 2005). We presume that such adaptations can arise in the ancestor Savanna population, but would often find themselves in combination with other adaptations to more typical breeding sites, such as puddles, due to random assortment. However, if a new inversion were to arise which by chance included two or more of the genes for laterite adaptations, then these associations would not be broken up each generation. The *j* inversion, unique to Bamako, may play such a role. There is new evidence that the divergence between the Bamako and Savanna is very recent (Slotman et al., 2006) and that loci within the *j* inversion are more differentiated than those outside of inversion regions (Tripet et al., 2005), consistent with this possibility. The evolution of Bamako from Savanna may thus be enabled by a condition of inversion polymorphism, which according to our results is less than rare.

An important simplification is that we only simulated a single inversion, while *An. gambiae* and related species have several. In nature inversions are also not randomly associated. Unless inversions occur simultaneously so that haplotypes with multiple unrelated inversions are competing, multiple non-overlapping inversions should not alter the outcome described here – single inversion frequencies stabilize quite quickly.

Chromosomal inversion may not be the sole cause of recombination suppression leading to reproductive isolation. If the loci are in the heterochromatic region

or near the centromere, for example, the outcome would be the same. There is recent evidence of the latter (Stump et al., 2005).

Finally, suppressed recombination is not the exclusive mechanism likely to be driving speciation in *An. gambiae* s.s. Given the high differentiation of loci on the X chromosome (Turner et al., 2005; Stump et al., 2005) sex linked reinforcement genes may lead to hybrid incompatibilities and speciation (Lemmon and Kirkpatrick, 2006). The variety of non-exclusive speciation forces acting on *An. gambiae* probably contributes to the high degree of complexity revealed by genetic analysis.

Our results suggest that field research on inversion polymorphism should focus on simple demography, particularly the size of populations in peripheral habitats. The long standing question of local extinction and recolonization versus severe bottlenecking in *Anopheles* (Charlwood et al., 2000; Minakawa et al., 2001; Taylor and Manoukis, 2003) during the dry season is of great consequence according to our analysis. Factors such as migration rate, which are hard to estimate in *An. gambiae*, should not be the first object of study. We expect that geographic regions that have high ecological variability and seasonality with many small, niche populations are more likely to be the source of inverted karyotypes than areas that support few large populations.

We conclude that inversion polymorphism can result as outlined in the ecotypification hypothesis under parameters reasonable for *An. gambiae*. This crucial first step sets the stage for actual differentiation, isolation and speciation. The simplest of demographic and genetic factors play the largest role in a polymorphic outcome: population size and loci captured in the inversion. We suggest that these should be the focus of further work on chromosomal speciation in *Anopheles gambiae*.

CHAPTER 7

Summary and conclusions

> Mosquitoes in general, and the malaria carriers in particular, have been the subject of an enormous amount of study, whose results have been reported in a voluminous literature. Much of this literature is an uncritical accumulation of facts that were easy to record, or of facts that were related to some momentarily fashionable subject of study, or of facts that were needed for some immediately practical objective. As a mere accumulation of facts the literature represents ... an encumbrance: but it is an encumbrance waiting to be converted into an orderly and useful structure of knowledge. (Bates, 1949)

7.1 Overview of results

The work presented in this dissertation furthers our understanding of two topics in *Anopheles gambiae* biology that are important to malaria transmission and control in sub-Saharan Africa. The first is how irrigation and increased density can affect vectorial efficiency. The second topic is the influence of *Anopheles* population structure and adaptation on the prospects of malaria control by novel methods, particularly through genetic modification of the insect vector.

The results of the Niono studies should be immediately useful for framing future research and directing control efforts. In Chapter 2 increased density was found to be correlated with reduction in some variables known to increase vectorial efficiency. Decreased anthropophily with density increases were notably sharp, though survivorship was also lower at high density. Other studies of

malaria prevalence in the same area show decreased malaria in irrigated areas compared to non-irrigated villages, which have lower vector densities (Sissoko et al., 2004).

In Chapter 3 a leading hypothesis for how survivorship could be decreased in irrigated regions was rejected. The competition hypothesis held that high larval densities would result in adult *Anopheles* with smaller body size and that these would have a shorter life span, explaining both the survivorship results of Chapter 2 and the documented reduction in malaria. This idea was based on several carefully executed studies from across the continent, but a systematic test of each component of the hypothesis revealed that it is not a likely reason for reduced transmission. In particular smaller adult females were not found to suffer increased mortality compared to larger ones, thus rejecting the crucial last step in the process.

The other chapters reveal challenges facing new proposed methods of malaria control, particularly through genetic modification of the vector for refractoriness. In Chapter 5 details of *Anopheles* population dynamics in Banambani were investigated. Large changes in population size with the annual climatic cycle were observed. Also evident was the tight linkage of *An. gambiae* form/species composition with environmental conditions. *An. arabiensis* was predominant during the driest parts of the year, replaced by the M molecular form of *An. gambiae* with the increased humidity. The M form represents the Mopti chromosomal form, thought to be adapted to aridity (Touré et al., 1998a). During the wettest part of 2005 the S molecular form, representing both the Savana and Bamako chromosomal forms, made its appearance and outnumbered the others as population size reached its peak.

The results suggest that the S form is the most likely to be undergoing seasonal

extinction. These factors are of importance to malaria control schemes not only because different species and densities may result in varying vectorial capacities but also because population structure is argued to be important to the spread of genes, as discussed below.

The ecological flexibility just described has been observed for some time, reflected by the large range of *An. gambiae*. In Chapter 6 a mechanism by which *An. gambiae* may adapt to new environments and eventually speciate was investigated through simulation. This work is the first quantitative test of the ecotypification hypothesis, explaining the abundance of morphologically indistinguishable forms of *An. gambiae* (Coluzzi, 1982). The model shows that specialization and speciation by this method is at least possible insofar as inversion polymorphism is not a rare outcome of the particular demography of this species. The results in Chapter 6 suggest how *An. gambiae* may be able to respond to selection pressures, such as those that may be introduced by insecticide spraying programs. It also indicates geographic and demographic conditions that might be conducive to *Anopheles* adaptation.

The ecology and population structure of *An. gambiae* s.s. combine to significantly complicate the spread of a hypothetical transposable element even under the simplest assumptions, as shown in Chapter 4. This is of significance because transposable elements have been proposed as vehicles for driving malaria refractoriness genes into Anophelines through their meiotic drive. Chapter 4 argues that even under optimistic conditions the details of population structure could result in such an effort having qualitatively different outcomes.

7.1.1 Summary of hypotheses

Chapter 2

Vectorial efficiency decreases with increasing adult *Anopheles* density.

Anthropophily and survivorship both decreased with increasing density in Niono, Mali. Vectorial capacity, however, did not.

Chapter 3

- Increased *Anopheles* larval density leads to smaller larvae in the rice fields of Niono.

  Evidence supporting decreased larval size with larval density increases was found for *An. gambiae*.

- Smaller larvae become smaller adult females.

  This hypothesis was supported for *An. gambiae* females.

- Smaller adult females survive less well than larger ones.

  No evidence of lower survivorship by smaller females was found, rejecting this hypothesis.

Chapter 4

Population structure of *Anopheles gambiae* s.s. around Banambani has significant effects on the spread of a transposable element as a method of malaria transmission reduction.

Simulation models support this hypothesis. Factors highlighted were the metapopulation dynamic (Source-sink or Levins) and yearly extinction.

Chapter 5

- The two molecular forms of *An. gambiae* s.s. and *An. arabiensis* in Banambani have differing ecological niches; the predominant form varies over a year.

  This hypothesis was supported by results showing composition changes with climate.

- The main malaria vector present in the area changes over the year as a consequence of the above.

  An. arabiensis and the M form of *An. gambiae* predominated during the drier parts of the year; they were outnumbered by the S form of *An. gambiae* during the wet season.

- Some forms/species are more likely to undergo annual extirpation than others.

  No conclusive evidence was found regarding this hypothesis.

Chapter 6

- Inversion polymorphism can result in *Anopheles* from a process of ecological specialization and chromosomal inversion ("ecotypification").

  The simulation model studied indicates that ecotypification could account for chromosomal inversion polymorphism in *An. gambiae*.

- Some demographic, environmental or genetic factors are more important to the maintenance of inversion polymorphism than others.

  Introduction to a small population size and capturing a larger number of selected loci were found to be the most significant factors leading to inversion polymorphism. Migration rate, recombination rate and maximum population size were found to be of secondary importance.

7.2 Recomendations

7.2.1 Irrigated areas

In Niono the next question addressed should be the relationship between vector density and bed net usage. The anthropophily results in Chapter 2 suggest there is a relationship between them, but this must be shown directly. If bed net usage is indeed the driver of reduced vectorial efficiency at high density then this suggests that renewed attention might be given to bed net programs in Niono at

least. A medical-social science approach to examining the effect of irrigation on disease and health will probably be most useful. Plaen et al. (2004) exemplify this type of effort, finding that social changes induced by irrigation can be as significant to the levels of disease as can the entomological changes.

The epidemiology of malaria in irrigated areas will have to be examined in multiple instances before generalities can be found, if they do exist. The high degree of variability in estimates of transmission intensity across irrigation projects suggests that different factors may be important in different areas. These will probably all have to be investigated separately.

One line of inquiry that should have benefits for many areas is relating agricultural practices to larval productivity. From informal observations during the field work for Chapter 3 I often observed large discrepancies in larval density between adjoining rice fields. These sometimes seemed identical to the eye, even sharing a water source.

The issue of water quality (Edillo et al., 2006) should also be examined in more detail. This would add to our current understanding of which factors determine larval productivity of particular rice fields (Diuk-Wasser et al., 2005b). This should lead to an enhanced ability to target interventions.

7.2.2 Population structure

Many of the research questions that are important to investigating speciation by ecotypification are also important to future models of the spread of genetically modified elements. The simulation results in Chapter 6 show that simple demography plays an important role in chromosomal inversion polymorphism. Direct ecological estimation of population sizes and movement between ecological niches are called for. These would also aid in characterizing the population structure of

An. gambiae, a critical issue in the movement of transposable elements.

In this regard, one of the most urgent needs in is to establish the metapopulation structure of *An. gambiae* by directly measuring and indirectly inferring parameters specific to metapopulation characteristics, such as synchrony, extirpation regime and patch equivalency. One important area here is to examine the extinction regime of the populations in the villages, which may be undertaken by both direct and indirect methods. The work in Chapter 5 is a step in this direction, but other approaches will be needed. Methods based on marking individual mosquitoes during the dry season and recapturing may be best suited for resolving the question of extirpation or permanence.

This sort of investigation would be most useful if extended to other sites. In addition to the power of comparative studies that would be gained, some sense of the limits of variability in system dynamics could also be explored. Such an understanding will ultimately prove critical in predicting how a given system might change when faced with attempts at genetic manipulation or externally varying conditions, like changes in climate or increasing human development, for example.

Once we have a more complete quantification of the factors we already know are important to the population structure of the species, realistic modeling of the effects of genetic modification of *An. gambiae* will be possible. The more complete our understanding of the system is as a metapopulation the more useful and accurate the modeling efforts can be.

Specifically for the question of ecotypification, attention should be given to larval and adult habitat differences. The selection gradient between the periphery and core portions of the range should be quite steep for ecotypification to occur. Thus areas that are highly variable should be the starting focus of investigation.

Furthermore, regions that have novel inversion haplotypes are also likely to be of interest, such as the southwestern Mali. For the j inversion in particular M. Coluzzi has already suggested the laterite pools on the border with Guinea. These are already being studied by members of our research group.

Finally, adaptive differences between chromosomal forms should continue to be sought at the DNA level. The development of a microarray "chip" for *An. gambiae* will be a great aid in this regard.

7.3 On the future of vector biology research

We are in the midst of a resurgence of interest in the malaria problem. Many resources are now being directed to methods of control based on the mosquito vector, exemplified by the Gates Foundation's investment in genetic modification. At the same time there is a serious risk that these efforts will repeat the shortcomings of those from the passed, though for differing reasons.

One important question is how research efforts should be directed. Technologically advanced means for reducing the malaria burden such as introducing genetically modified mosquitoes are in vogue, but they have hazy prospects for success. Clearly the strategies that worked in the mesoendemic edges of the parasite's range in the 20^{th} century are unlikely to work in sub-Saharan Africa for many reasons (Amorosa et al., 2005). Still, some have recently re argued that the simplest measures, such as mosquito-proofing houses (Lindsay et al., 2002), or raising them off the ground (Charlwood et al., 2003) can significantly reduce human contact with the vector and thus reduce malaria. There is a difficult balance between "silver bullet" type technical efforts and simple, inexpensive and only partially effective methods. It is probably the case that the largest improvement

in the malaria crisis in Africa will result from a combination of efforts, from the economic to the medical and the vector-targeted.

Despite this difficult balance increasingly clear that a lot of work on *Anopheles* ecology is required for a clearer assessment of the probability of success for any, especially the technically complex, control program. It would be ideal if the results of such ecological studies were general enough to be applied to study of other possible control methods and toward comprehension of the evolution of the vector.

With increasing human population size in Africa future vector research will by necessity be focused on areas of heavy anthropogenic environmental impacts, such as the Niono irrigation project. These impacts clearly hold the potential for modifying biological relationships and with them transmission patterns. Study of these areas should suggest new ways to interrupt or reduce the disease cycle.

An informal survey of the literature gives the impression that currently most of the efforts of the vector research community are directed toward molecular and population genetic studies. While these results are important and interesting, malaria control efforts may be better aided by ecological studies at this point. These are more time-consuming and expensive, but are the only way to apply the understanding of molecular markers (such as the M and S distinction) to the situation in nature. We should heed the advice of M. Bates at the beginning of this chapter and work towards a useful structure of knowledge. Increasing the amount of ecological and field work to catch up to the mass of genetic work already in hand would be an important step in this regard.

Looking to the past, the greatest successes in the war against malaria came from massive public investment and global political will. The early efforts in Cuba, the Brazilian campaign against *An. gambiae*, the DDT residual spraying

campaigns of the WHO and the modeling results of the Garki project were driven by governmental enthusiasm and investment. In such an environment research can thrive and play an important role in directing efforts, enabling success or failure of control plans. We are fortunate to be in a period today of increasing public investment, and it is my hope that we rise to the challenge and make the crucial difference to reducing the malaria burden on humanity.

APPENDIX A

Measuring malaria transmission

A.1 Entomological Inoculation Rate (EIR)

Malaria transmission may be measured by the Entomological Inoculation Rate (EIR) when it is being estimated by study of the vector population rather than the humans. EIR is defined as the number of positive bites one person receives in one night. It is calculated as:

$$EIR = ma \cdot s \tag{A.1}$$

where ma is the "man biting rate". ma is the product of: m, the anopheline density relative to sleepers (female mosquitoes per sleeper) and a is the "man biting habit of the species" (the proportion of bites on humans divided by the length of the gonotrophic cycle). s is the sporozoite rate in the biting population (the proportion of female mosquitoes that are infectious: bearing sporozoites on their salivary glands).

A.2 Vectorial capacity (C)

The sporozoite rate is usually quite low, particularly in Niono (Dolo et al., 2004), meaning that thousands of mosquitoes may have to be dissected for a reliable es-

timate of EIR. In this case or when the objective is to estimate the transmission efficiency of a vector population in the absence of any measure of actual transmission, Vectorial Capacity (C) may be used. This index essentially measures the receptivity to malaria of a vector population.

C is also a useful measure for assessing the influence of entomological parameters on transmission apart from other factors (parasitological or immunological, for example).

Vectorial capacity is calculated according to:

$$C = \frac{ma^2 p^n}{-log(p)} \tag{A.2}$$

Where ma is calculated as above, then multiplied by a estimated as the number of blood meals per vector per day, yielding ma^2. p is the proportion of vectors surviving one day. n is the extrinsic period of development of the parasite, typically 12 days at 25^oC.

Vectorial capacity may be verbally described as follows (Molineaux et al., 1988): Let the number of vectors per human be m and the number of bites per mosquito per night on humans is a, then a human is bit ma times per day, on average. Assuming an exponential survival rate with daily survival p, then proportion p^n of these vectors survives the incubation period (sporozoite cycle) of the parasite (n), so that they could then transmit the pathogen. The vectors are then expected to survive another $1/log(p)$ days, and bite other persons a times per day, on average.

APPENDIX B

Ecotypification simulation details

The simulation is an object oriented program consisting of two populations (P_0 and P_1). Each population contains a number of individuals, each of which has a genome made up of two chromosomes. These are represented as two binary arrays of length equal to the number of adaptive loci being simulated (N). The state of each position represents one of two alleles: one adapted to each population's habitat. Each locus contributes additively a set amount to the fitness (W) of an individual. Mosquitoes are selected for breeding based on their fitness relative to others in the population. The two populations are connected by migration (m_0 and m_1), described in more detail below.

Each time step of the simulation represents a generation and involves creating a new group of individuals from the current one. The size of the population after reproduction is defined by a linear function which increases from a minimum size (S_0 for P_0 and S_1 for P_1) to a maximum size 100 times that value after five time steps. The population size is then decreased back to the minimum size after the tenth step.

The major steps for a generation are: 1) calculate the fitness of all individuals 2) perform a "breeding" step the number of times necessary to create the offspring population.

B.1 Calculating fitness

Individuals are selected for breeding with a probability proportional to their fitness, so at the beginning of each time step the fitness of all individuals is calculated. For each locus, if the allele state is homozygous maladapted to the environment in which the individual is located then that locus contributes 1 to the fitness of the organism. The fitness of the whole organism is calculated in an additive manner, so if the individual is homozygous maladapted in all loci to the environment then its fitness is N.

If the individual is homozygous adapted to the environment then that locus contributes $1 + s$ to the individual's fitness, where s is the adaptive value of the locus (this value is the same for all loci but is varied between simulations). Thus the maximum fitness attainable (an individual that is homozygous adapted in all its loci) is $N(1 + s)$.

If an individual is a heterozygote at a particular locus then the fitness contribution of that locus to the summation is $1 + (s/2)$. The fitness calculation for a single individual may be written as follows

$$W = H(1 + s) + h(1 + \frac{s}{2}) + (N - [H + h]) \tag{B.1}$$

where W is the individual's fitness, H is the number of homozygous adapted loci, h is the number of heterozygous loci.

B.2 Breeding

The breeding function creates a new individual from two parents. The first parent is selected with probability proportional to its fitness from the population which

is being stepped into the next generation. A second individual is chosen from the other population with probability m, where m is the migration rate. If a migrant is selected the second parent is chosen randomly from the other population. With probability $1 - m$ the second parent is from the same population and is selected with probability proportional to its fitness. In either case the parents must be two distinct individuals. An individual may be selected for breeding multiple times or not at all during a particular breeding cycle.

One chromosome is chosen at random from each parent and combined in a new individual. At this point recombination may occur. Each gap between loci has the same probability (r) of being the site of a chromosomal break-point. A draw is made for each gap, in order from locus 0 to locus i_N. If a break occurs at locus i_j then all alleles from i_j through i_N are swapped from one chromosome to another. Thus multiple break-points and compound recombination patterns are possible.

If a break point is chosen in an interval between loci that are within an inversion in that particular individual then the recombination event is not carried out. This is the case if there is only one or if both chromosomes carry the inversion.

B.3 Chromosomal inversion

At a time t_i an inversion is introduced to one chromosome of a single individual in the peripheral population. The number of consecutive loci marked as being in an inversion is a simulation parameter, n_i. All the loci within the inversion are adapted to the peripheral environment and there can be no crossing over in the gaps between them as noted above.

If the frequency of the inversion becomes zero after t_i the simulation is stopped since inversions are not introduced at any other time and so the frequency will not change from zero.

B.4 Experimental parameter sets and parallelization

Individual simulations were run on 11 worker cpus, tasked by a master program which sampled from the parameter space. The master program used Latin Hypercube Sampling (LHS) (Blower and Dowlatabadi, 1994) to select parameter sets. To do this each parameter range is divided into uniform subranges ("slices") numbering the overall number of simulations. To obtain a parameter set we randomly chose one of the slices for each parameter without replacement. The actual value of each parameter sent to the simulation was the mid-point of the slice. This process is repeated until no slices remain and the requisite number of simulations are complete.

References

Abdoulaye, D., Thierry, B., Chandre, C., Roch, D., Pierre, K., Robert, G., Frederic, S., Pierre, G., Janet, H., and Marc, H. (2003). KDR mutation, a genetic marker to assess events of introgression between the molecular m and s forms of *Anopheles gambiae* (diptera : Culicidae) in the tropical savannah area of West Africa. *Journal of Medical Entomology*, 40:195 – 198.

Ameneshewa, B. and Service, M. (1996). The relationship between female body size and survival rate of the malaria vector *Anopheles arabiensis* in Ethiopia. *Medical and Veterinary Entomology*, 10(2):170–172.

Amorosa, L., Corbellini, G., and Coluzzi, M. (2005). Lessons from malaria: Italy's past and sub-Sahara's future. *Health and Place*, 11:67 – 73.

Anderson, E., Williamson, E., and Thompson, E. (2000). Monte Carlo evaluation of the likelihood for N_e from temporally spaced samples. *Genetics*, 156:2109–2118.

Anderson, R. and May, R. (1979). Population biology of infectious diseases: Part I. *Nature*, 280:361 – 367.

Audibert, M., Josseran, R., Josse, R., and Adjidji, A. (1985). Comparison of malarial endemicity in a rice-growing area and a cotton-growing area of the Rusizi Plain, Burundi. *Annales de la Societe Belge de Medecine Tropicale*, 65 Supplement 2:187 – 200.

Audibert, M., Josseran, R., Josse, R., and Adjidji, A. (1990). Irrigation, schistosomiasis, and malaria in the Logone Valley, Cameroon. *American Journal of Tropical Medicine and Hygiene*, 42:550 – 560.

Aultman, K., Beaty, B., and Walker, E. (2001). Genetically manipulated vectors of human disease: a practical overview. *Trends in Parasitology*, 17:507 – 510.

Ayala, F. and Coluzzi, M. (2005). Chromosome speciation: Humans, *Drosophila* and mosquitoes. *Proceedings of the National Academy of Sciences, USA*, 102:6535 – 6542.

Baldet, T., Diabat, A., and Guiguemd, T. R. (2003). Malaria transmission in 1999 in the rice field area of the Kou Valley (Bama), (Burkina Faso). *Sante*, 13(1):55–60.

Bates, M. (1949). *The natural history of mosquitoes*. The Macmillan Co.

Bayoh, M. N., Thomas, C. J., and Lindsay, S. W. (2001). Mapping distributions of chromosomal forms of *Anopheles gambiae* in West Africa using climate data. *Medical and Veterinary Entomology*, 15(3):267–274.

Bengtsson, B. (1985). *Evolution: Essays in honour of John Maynard Smith*, chapter The flow of genes through a genetic barrier, pages 31 –42. Cambridge University Press.

Besansky, N. J., Lehmann, T., Fahey, G. T., Fontenille, D., Braack, L. E. O., Hawley, W. A., and Collins, F. H. (1997). Patterns of mitochondrial variation within and between African malaria vectors, *Anopheles gambiae* and *An-arabiensis*, suggest extensive gene flow. *Genetics*, 147(4):1817–1828.

Besansky, N. J., Powell, J. R., Caccone, A., Hamm, D. M., Scott, J. A., and Collins, F. H. (1994). Molecular phylogeny of the *Anopheles gambiae* complex suggests genetic introgression between principal malaria vectors. *Proceedings of the National Academy of Sciences, USA*, 91(15):6885–6888.

Black, W. C. and Lanzaro, G. C. (2001). Distribution of genetic variation among chromosomal forms of *Anopheles gambiae* s.s.: introgressive hybridization, adaptive inversions, or recent reproductive isolation? *Insect Molecular Biology*, 10(1):3–7.

Blower, J., Cook, L., and Bishop, J. (1981). *Estimating the size of animal populations*. Allen & Unwin, London.

Blower, S. and Dowlatabadi, H. (1994). Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model, as an example. *International Statistical Review*, 2:229 – 243.

Boete, C. and Koella, J. (2002). A theoretical approach to predicting the success of genetic manipulation of malaria mosquitoes in malaria control. *Malaria Journal*, 1:1 – 7.

Boudin, C., Robert, V., Carnevale, P., and Thomas, P. (1992). Epidemiology of *Plasmodium falciparum* in a rice field and a savannah area iin Burkina Faso. *Acta Tropica*, 51:103–111.

Breiman, L., Freidman, J., Olshen, R., and Stone, C. (1984). *Classification and regression trees*. Wadsworth, Monterey, CA.

Brooke, B. D., Hunt, R. H., Chandre, F., Carnevale, P., and Coetzee, M. (2002). Stable chromosomal inversion polymorphisms and insecticide resistance in the malaria vector mosquito *Anopheles gambiae* (Diptera : Culicidae). *Journal of Medical Entomology*, 39(4):568–573.

Bruce-Chwatt, L. (1985). *Essential Malariology*. John Wiley and Sons, New York, 2nd edition.

Burkot, T., Zavala, F., Gwadz, R., Collins, F., Nussenzweig, R., and Roberts, D. (1984). Identification of malaria-infected mosquitos by a 2-site enzyme-linked immuno-absoprbent assay. *American Journal of Tropical Medicine and Hygiene*, 33:227 – 231.

Bynum, W. and Overy, C., editors (1998). *The beast in the mosquito: The correspondence of Ronald Ross and Patrick Manson*. Rodopi, Amsterdam.

Carnahan, J. M., Zheng, L., Taylor, C. E., Touré, Y. T., Norris, D. E., Dolo, G., Diuk-Wasser, M. A., and Lanzaro, G. C. (2002). Genetic differentiation of *Anopheles gambiae* s.s. populations in Mali, West Africa, using microsatellite loci. *Journal of Heredity*, 93:249–253.

Carnevale, P. and Robert, R., editors (1987). *Introduction of irrigation in Burkina Faso and its effect on malaria transmission*. Edited vers. of working papers presented at the 7^th annual meeting of joint WHO/FAO/UNEP panel of experts on environmental management for vector control, Sept. 1987.

Carson, H. (1970). Chromosome tracers of the origin of species. *Science*, 168:1414 – 1418.

Cavalli-Sforza, L., Menozzi, P., and Piazza, A. (1993). Demic expansions and human evolution. *Science*, 259:639 – 646.

Charlwood, J., Pinto, J., Ferrara, P., Sousa, C., Ferreira, C., Gil, V., and Rosario, V. (2003). Raised houses reduce mosquito bites. *Malaria Journal*, 2:45.

Charlwood, J., Pinto, J., Sousa, C., Madsen, H., Ferreira, C., and do Rosario, V. (2002). The swarming and mating behaviour of *Anopheles gambiae* s.s. (Diptera : Culicidae) from Sao Tome island. *Journal of Vector Ecology*, 27:178 – 183.

Charlwood, J. D., Vij, R., and Billingsley, P. F. (2000). Dry season refugia of malaria-transmitting mosquitoes in a dry savannah zone of East Africa. *American Journal of Tropical Medicine and Hygiene*, 62(6):726–732.

Coluzzi, M. (1982). *Spatial distribution of chromosomal inversions and speciation in Anopheline mosquitoes*, pages 143 – 153. Alan R. Liss, New York.

Coluzzi, M., Petrarca, V., and Dideco, M. A. (1985). Chromosomal inversion intergradation and incipient speciation in *Anopheles gambiae*. *Bollettino di Zoologia*, 52(1-2):45–63.

Coluzzi, M. and Sabatini, A. (1967). Cytogenetic observations on species A and B of the *Anopheles gambiae* complex. *Parassitologia*, 9:71 – 88.

Coluzzi, M., Sabatini, A., della Torre, A., Di Deco, M., and Petrarca, V. (2002). A polytene chromosome analysis of the *Anopheles gambiae* complex. *Science*, 298:1415 – 1418.

Coluzzi, M., Sabatini, A., Petrarca, V., and Di Deco, M. (1977). Behavioural divergences between mosquitoes with different inversion karyotypes in polymorphic populations of the *Anopheles gambiae* complex. *Nature*, 266:832 – 833.

Coluzzi, M., Sabatini, A., Petrarca, V., and Di Deco, M. A. (1979). Chromosomal differentiation and adaptation to human environments in the *Anopheles gambiae* complex. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 73:483 – 497.

Coosemans, M., Wery, M., Storme, B., Hendrix, L., and Mfisi, B. (1984). Epidemiology of malaria in the Ruzizi valley, Burundi. *Annales de la Societe Belge de Medecine Tropicale*, 64:135–158.

Costantini, C., Li, S. G., DellaTorre, A., Sagnon, N., Coluzzi, M., and Taylor, C. E. (1996). Density, survival and dispersal of *Anopheles gambiae* complex mosquitoes in a West African Sudan savanna village. *Medical and Veterinary Entomology*, 10(3):203–219.

Couprie, B., Claudot, Y., Sameekobo, A., Issoufa, H., Legerdebruyne, M., Tribouley, J., and Ripert, C. (1985). An epidemiological-study of the malaria in the rice-growing regions of Yagoua and Maga (North-Cameroun). *Bulletin de la Societe de Pathologie Exotique (Paris)*, 78:191 – 204.

Coura, J., Suarez-Mutis, M., and Ladeia-Andrade, S. (2006). A new challenge for malaria control in Brazil: asymptomatic *Plasmodium* infection - a review. *Memorias do Instituto Oswaldo Cruz*, 101:229 – 237.

Coyne, J. and Orr, H. (2004). *Speciation*. Sinauer, Sunderland, MA.

Crow, J. (1956). *Statistics and mathematics in biology*, chapter Breeding structure of populations. II. Effective population number. University of Iowa Press, Ames, Iowa.

Crow, J. and Kimura, M. (1970). *Introduction to population genetics theory*. Harper and Row, New York.

Davidson, G. (1954). Estimation of the survival rate of Anopheline mosquitoes in nature. *Nature*, 174:792–793.

Davidson, G. (1964). *Anopheles gambiae*, a complex of species. *Bulletin of the World Health Organization*, 31:371 – 374.

Davidson, G. and Draper, C. (1953). Field studies of some of the basic factors concerned in the transmission of malaria. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 49:522–535.

della Torre, A., Costantini, C., Besansky, N. J., Caccone, A., Petrarca, V., Powell, J. R., and Coluzzi, M. (2002). Speciation within *Anopheles gambiae*: the glass is half full. *Science*, 298(5591):115–117.

della Torre, A., Fanello, C., Akogbeto, M., Dossou-yovo, J., Favia, G., Petrarca, V., and Coluzzi, M. (2001). Molecular evidence of incipient speciation within *Anopheles gambiae* s.s. in West Africa. *Insect Molecular Biology*, 10(1):9–18.

della Torre, A., Tu, Z. J., and Petrarca, V. (2005). On the distribution and genetic differentiation of *Anopheles gambiae* s.s. molecular forms. *Insect Biochemistry And Molecular Biology*, 35(7):755–769.

Depinay, J., Mbogo, C., Killeen, G., Knols, B., Beier, J., Carlson, J., Dushoff, J., Billingsley, P., Mwambi, H., Githure, J., Touré, A., and McKenzie, E. (2004). A simulation model of African *Anopheles* ecology and population dynamics for the analysis of malaria transmission. *Malaria Journal*, 3:29.

Detinova, T. (1962). *Age-grouping methods in Diptera of medical importance with special reference to some vectors of malaria*. World Health Organization, Geneva.

Di Deco, M., Petrarca, V., Villani, F., and Coluzzi, M. (1980). Polimorfismo cromosomico da inversioni paracetriche ed eccesso degli eterocariotipi in ceppi de *Anopheles* allevati in laboratorio. *Parassitologia*, 22:304 – 306.

Dicko, A., Klion, A., Thera, M., Sangara, I., Yalcouye, D., Niambele, M., Sogoba, M., Dolo, G., Dao, A., Diallo, D., Doumbo, O., and Miller, L. (2004). The etiology of severe anemia in a village and a periurban area in Mali. *Blood*, 104:1198 – 1200.

Diuk-Wasser, M., Touré, M., Dolo, G., Bagayoko, M., Sogoba, N., Traoré, S., Manoukis, N., and Taylor, C. (2005a). Vector abundance and malaria transmission in rice-growing villages in Mali. *Journal of the American Society of Tropical Medicine and Hygiene*, 72:725 – 731.

Diuk-Wasser, M. A., Dolo, G., Bagayoko, M., Sogoba, N., Touré, M. B., Moghaddam, M., Manoukis, N. C., Rian, S., Traoré, S. F., and Taylor, C. E. (2005b). Patterns of irrigated rice growth and malaria vector breeding in Mali using multi-temporal ERS-2 synthetic aperture radar. *International Journal of Remote Sensing*, 27:535–548.

Dobzhansky, T. (1947). Genetics of natural populations. A response of certain gene arrangements in the third chromosome of *Drosophila pseudoobscura* to natural selection. *Genetics*, 32:142 – 160.

Dolo, G. (2000). *Etude des populations d'*An. gambiae *s.l. par marque, lacher et recapture a Banambani en 1993 et 1994.* PhD thesis, Universite du Mali, Bamako, Mali.

Dolo, G., Briet, O., Dao, A., Traoré, S., Bouaré, M., Sogoba, N., Niaré, O., Bagayogo, M., Sangaré, D., Teuscher, T., and Touré, Y. (2004). Malaria transmission in relation to rice cultivation in the irrigated Sahel of Mali. *Acta Tropica*, 89(2):147–159.

Dolo, G., Dao, A., Traoré, S., Bouaré, M., Sogoba, N., Niaré, O., Bagayogo, M., Sangaré, D., and Tourè, Y. (1999). Rapport de l'étude entomologique sur la transmission du paludisme dans six villages (aout 1995-fevrier 1998). Technical report, West African Rice Development Association.

Donnelly, M., Licht, M., and Lehmann, T. (2001). Evidence for recent population

expansion in the evolutionary history of the malaria vectors *Anopheles arabiensis* and *Anopheles gambiae*. *Molecular Biology and Evolution*, pages 1353 – 1364.

Dossou-Yovo, J., Doannio, J., Diarrassouba, S., and Chauvancy, G. (1998). The impact of rice fields on malaria transmission in the city of Bouaké, Côte d'Ivoire. *Bulletin de la Societe de Pathologie Exotique*, 91:327 – 323.

Dzodzomenyo, M., Dunyo, S., Ahorlu, C., Coker, W., Appawu, M., Pedersen, E., and Simonsen, P. (1999). Bancroftian filariasis in an irrigation project community in southern Ghana. *Tropical Medicine & International Health*, 4(1):13–18.

Edillo, F., Tripet, F., Touré, Y., Lanzaro, G., Dolo, G., and Taylor, C. (2006). Water quality and immatures of the M and S forms of *Anopheles gambiae* s.s. and *An. arabiensis* in a Malian village. *Malaria Journal*, 5:35.

Edillo, F. E., Touré, Y. T., Lanzaro, G. C., Dolo, G., and Taylor, C. E. (2002). Spatial and habitat distribution of *Anopheles gambiae* and *Anopheles arabiensis* (Diptera : Culicidae) in Banambani Village, Mali. *Journal of Medical Entomology*, 39(1):70–77.

Fanello, C., Petrarca, V., della Torre, A., Santolamazza, F., Dolo, G., Coulibaly, M., Alloueche, A., Curtis, C. F., Touré, Y. T., and Coluzzi, M. (2003). The pyrethroid knock-down resistance gene in the *Anopheles gambiae* complex in Mali and further indication of incipient speciation within *An. gambiae* s.s. *Insect Molecular Biology*, 12(3):241–245.

Favia, G., della Torre, A., Bagayoko, M., Lanfrancotti, A., Sagnon, N., Touré, Y. T., and Coluzzi, M. (1997). Molecular identification of sympatric chromo-

somal forms of *Anopheles gambiae* and further evidence of their reproductive isolation. *Insect Molecular Biology*, 6(4):377–383.

Favia, G., Lanfrancotti, A., Spanos, L., Siden-Kiamos, I., and Louis, C. (2001). Molecular characterization of ribosome DNA polymorphisms discriminating among chromosomal forms of *Anopheles gambiae* s.s. *Insect Molecular Biology*, 10:19 – 23.

Faye, O., Gaye, O., Herv, J., Diack, P. A., and Diallo, S. (1993). Malaria in the Saharan region of Senegal. 2. Parasitological indices. *Annales de la Societe Belge de Medecine Tropicale*, 73:31 – 36.

Fine, P. (1975). Ross' *a priori* pathometry – a perspective. *Proceedings of the Royal Society of Medicine*, 68:547 – 551.

Frankham, R., Ballou, J., and Briscoe, D. (2002). *Introduction to conservation biology*. University Press, Cambridge.

Fu, Y. and Li, W. (1999). Coalescing into the 21^{st} century: An overview and prospects for coalescent theory. *Theoretical Population Biology*, 56:1 – 10.

Futuyma, D. (1998). *Evolutionary Biology*. Sinauer, Sunderland.

Gavrilets, S., Acton, R., and Gravner, J. (2000). Dynamics of speciation and diversification in a metapopulation. *Evolution*, 54:1493 – 1501.

Gbakima, A. A. (1994). Inland valley swamp rice development: malaria, schistosomiasis, onchocerciasis in south central Sierra Leone. *Public Health*, 108(2):149–157.

Gentile, G., della Torre, A., Maegga, B., Powell, J. R., and Caccone, A. (2002).

Genetic differentiation in the African malaria vector, *Anopheles gambiae* ss, and the problem of taxonomic status. *Genetics*, 161(4):1561–1578.

Gilles, H. and Warrell, D. (1993). *Bruce-Chwatt's Essential Malariology*. Oxford University Press, New York, 3rd edition.

Gilles, M. and Coetzee, M. (1987). *A supplement to the Anophelinae of Africa south of Sahara*. Publications of the South African Institute of Medical Research No. 55, Johannesburg, Johannesburg, 1st edition.

Gilles, M. and De Meillon, B. (1968). *The Anophelinae of Africa, South of sahara (Ethiopian Zoogrographical Region)*. Publications of the South African Institute of Medical Research No. 54, Johannesburg, Johannesburg, 2nd edition.

Gimnig, J., Ombok, M., Otieno, S., Kaufman, M., Vulule, J., and Walker, E. (2002). Density-dependent development of *Anopheles gambiae* (Diptera : Culicidae) larvae in artificial habitats. *Journal of Medical Entomology*, 39(1):162–172.

Girod, R., Salvan, M., Simard, F., Andrianaivolambo, L., Fontenille, D., and Laventure, S. (1999). Evaluation of the vectorial capacity of *Anopheles arabiensis* (Diptera:Culicidae) on the island of Reunion: an approach to the health risk of malaria importation in an area of eradication. *Bulletin de la Societe de Pathologie Exotic (Paris)*, 92(3):203–209.

Githeko, A. K., Service, M. W., Mbogo, C. M., Atieli, F. K., and Juma, F. O. (1993). *Plasmodium falciparum* sporozoite and entomological inoculation rates at the Ahero rice irrigation scheme and the Miwani sugar-belt in western Kenya. *Annals of Tropical Medicine and Parasitology*, 87:379 – 91.

Gonçalves, A., Ferrinho, P., and Dias, F. (1996). The epidemiology of malaria in Probis, Guinea-Bissau. *Memorias do Instituto Oswaldo Cruz*, 91:11 – 17.

Gray, E. and Bradley, T. (2005). Physiology of desiccation resistance in *Anopheles gambiae* and *Anopheles arabiensis*. *American Journal of Tropical Medicine And Hygiene*, 73:553 – 559.

Greenwood, B. and Mutabingwa, T. (2002). Malaria in 2002. *Nature*, 415:670 – 672.

Hanski, I. and Gilpin, M., editors (1997). *Metapopulation biology : ecology, genetics, and evolution*. Academic Press, San Diego.

Harrison, G. (1978). *Mosquitos, malaria and man: A history of the hostilities since 1880*. E.P. Dutton, New York.

Hartl, D. L. and Clark, A. G. (1997). *Principles of Population Genetics*. Sinauer, Sunderland, MA, 3rd edition.

Hemingway, J. (2004). Taking aim at mosquitoes. *Nature*, 430:936.

Hogg, J. C. and Hurd, H. (1997). The effects of natural *Plasmodium falciparum* infection on the fecundity and mortality of *Anopheles gambiae* s. l. in north east Tanzania. *Parasitology*, 114:325–331.

Honigsbaum, M. (2001). *The fever trail: The hunt for the cure for malaria*. MacMillan.

Hudson, H. (1990). *Oxford survey of evolutionary biology*, chapter Gene genealogies and the coalescent process, pages 1 – 44. Oxford University Press, Oxford.

Hunt, R., Coetzee, M., and Fettene, M. (1998). The *Anopheles gambiae comlex*: a new species from Ethiopia. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 92:231 – 235.

Hunter, J., Rey, L., Chu, K., Adekolu, John, E., and Mott, K. (1993). *Parasitic Diseases in Water Resources Development*. WHO, Geneva.

Hurd, H. (2003). Manipulation of medically important insect vectors by their parasites. *Annual Review of Entomology*, 48:141–161.

Hurd, H., Taylor, P. J., Adams, D., Underhill, A., and Eggleston, P. (2005). Evaluating the costs of mosquito resistance to malaria parasites. *Evolution*, 59(12):2560–2572.

Ijumba, J. and Lindsay, S. (2001). Impact of irrigation on malaria in africa: paddies paradox. *Medical and Veterinary Entomology*, 15(1):1–11.

Ijumba, J. N., Mosha, F. W., and Lindsay, S. W. (2002). Malaria transmission risk variations derived from different agricultural practices in an irrigated area of northern Tanzania. *Medical and Veterinary Entomology*, 16(1):28–38.

James, A. and Handler, A., editors (2000). *Insect Transgenesis: Methods and Applications*. CRC Press, Boca Raton, FL.

Johnson, C. (1969). *Migration and dispersal of insects by flight*. Methuen, London.

Kayondo, J. K., Mukwaya, L. G., Stump, A., Michel, A. P., Coulibaly, M. B., Besansky, N. J., and Collins, F. H. (2005). Genetic structure of *Anopheles gambiae* populations on islands in northwestern Lake Victoria, Uganda. *Malaria Journal*, 4:59.

Killeen, G., Fillinger, U., Kiche, I., Gouagna, L., and Knols, B. (2002). Eradication of *Anopheles gambiae* from Brazil: lessons for malaria control in Africa? *The Lancet Infectious Diseases*, 2:618 – 627.

Kimura, M. and Ohta, T. (1971). *Theoretical aspects of population genetics*. Princeton University Press, Princeton.

Kirkpatrick, M. and Barton, N. (2006). Chromosome inversions, local adaptation and speciation. *Genetics*, 173:419 – 434.

Kitthawee, S., Edman, D., and Sattabongkot, J. (1990). Evaluation of survival potential and malaria suseptibility among different size classes of laboratory-reared *Anopheles dirus*. *American Journal of Tropical Medicine and Hygiene*, 43(4):328–332.

Kitthawee, S. and Edman, J. (1995). Adult body size and biting activity of field populations of *Anopheles dirus* (Diptera : Culicidae). *Southeast Asian Journal of Tropical Medicine and Public Health*, 26(3):582–585.

Kitthawee, S., Edman, J., and Upatham, E. (1992). Relationship between female *Anopheles dirus* (Diptera : Culicidae) body size and parity in a biting population. *Journal of Medical Entomology*, 29(6):921–926.

Klinkenberg, E., Takken, W., Huibers, F., and Touré, Y. (2003). The phenology of malaria mosquitoes in irrigated rice fields in Mali. *Acta Tropica*, 85(1):71–82.

Koella, J. C., Sorensen, F. L., and Anderson, R. A. (1998). The malaria parasite, *Plasmodium falciparum*, increases the frequency of multiple feeding of its mosquito vector, *Anopheles gambiae*. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 265(1398):763–768.

Koenraadt, C., Majambere, S., Hemerik, L., and Takken, W. (2004). The effects of food and space on the occurrence of cannibalism and predation among larvae of *Anopheles gambiae* sl. *Entomologia Experimentalis et Applicata*, 112:125 – 134.

Korenromp, E. L., Williams, B. G., Gouws, E., Dye, C., and Snow, R. W. (2003). Measurement of trends in childhood malaria mortality in Africa: an assessment of progress toward targets based on verbal autopsy. *Lancet Infectious Diseases*, 3(6):349–358.

Krzywinski, J. and Besansky, N. (2003). Molecular systematics of *Anopheles*: from subgenera to subpopulations. *Annual Review of Entomology*, 48:111 – 139.

Lanzaro, G. C., Touré, Y. T., Carnahan, J., Zheng, L., Dolo, G., Troaré, S., Petrarca, V., Vernick, K. D., and Taylor, C. E. (1998). Complexities in the genetic structure of *Anopheles gambiae* populations in West Africa as revealed by microsatellite dna analysis. *Proceedings of the National Academy of Sciences, USA*, 95:14260–14265.

Laventure, S., Mouchet, J., Blanchy, S., Marrama, L., Rabarison, P., Andrianaivolambo, L., Rajaonarivelo, E., Rakotoarivony, I., and Roux, J. (1996). Rice: source of life and death on the plateaux of Madagascar. *Sante*, 6(2):79–86.

Lehmann, T., Hawley, W. A., Grebert, H., and Collins, F. H. (1998). The effective population size of *Anopheles gambiae* in Kenya: Implications for population structure. *Molecular Biology and Evolution*, 15(3):264–276.

Lemmon, A. and Kirkpatrick, M. (2006). Reinforcement and the genetics of hybrid incompatabilities. *Genetics*, 173:1145 – 1155.

Levins, R. (1969). Some demographic and genetic consequences of environmental hetrogeneity for biological control. *Bulletin of the Entomological Society of America*, 15:237 – 240.

Lindsay, S., Emerson, P., and Charlwood, J. (2002). Reducing malaria by mosquito-proofing houses. *Trends in Parasitology*, 18:510 – 514.

Livingstone, K. and Rieseberg, L. (2004). Chromosomal evolution and speciation: a recombination-based approach. *New Phytologist*, 161(1):107–112.

Lounibos, L. and Conn, J. (1991). Fecundity, parity and adult feeding relationships among nyssorhynchus malaria vectors from Venezuela. *Memorias do Instituto Oswaldo Cruz*, 86(1):57–66.

Lozano-Fuentes, S., Lee, Y., Touré, M., Doumbia, S., S.F., T., Lanzaro, G., and Taylor, C. (2007). *Anopheles gambiae* s.s. in mali: Towards releasing a GM mosquito. *Nature*, submitted:–.

Lyimo, E., Takken, W., and Koella, J. C. (1992). Effect of rearing temperature and larval density on larval survival, age at pupation and adult size of *Anopheles gambiae*. *Entomologia Experimentalis et Applicata*, 63(3):265–271.

Lyimo, E. O. and Koella, J. C. (1992). Relationship between body size of adult *Anopheles gambiae* s.l. and infection with the malaria parasite *Plasmodium falciparum*. *Parasitology*, 104 (Pt 2):233–237.

MacDonald, G. (1946). *Malaria and its control*. London School of Hygiene and Tropical Medicine, London.

MacDonald, G. (1957). *The epidemiology and control of malaria*. Oxford University Press, London.

Macdonald, G. (1965). Eradication of malaria. *Public Health Reports*, 80:870–880.

Malecot, G. (1969). *The mathematics of heredity*. W.H. Freeman, San Fransisco.

Manoukis, N. C., Collier, T. C., and Taylor, C. E. (2007). Inversion polymorphism and speciation in *Anopheles gambiae* s.s.: a simulation study. *Journal of Evolutionary Biology*, Submitted:–.

Manoukis, N. C., Touré, M. B., Sissoko, I., Doumbia, S., Traoré, S. F., Diuk-Wasser, M. A., and Taylor, C. E. (2006). Is vector body size the key to reduced malaria transmission in the irrigated region of Niono, Mali. *Journal of Medical Entomology*, 43:820 – 827.

Marrama, L., Jambou, R., Rakotoarivony, I., Tsi, J. M. L. P., Duchemin, J. B., Laventure, S., Mouchet, J., and Roux, J. (2004). Malaria transmission in Southern Madagascar: influence of the environment and hydro-agricultural works in sub-arid and humid regions. Part 1. Entomological investigations. *Acta Tropica*, 89(2):193–203.

Maynard Smith, J. (1966). Sympatric speciation. *American Naturalist*, 100:637 – 650.

McGraw, K. and Wong, S. (1996). Forming inferences about some intraclass correlation coefficients. *Pyschological Methods*, 1(1):30 – 46.

Minakawa, N., Githure, J. I., Beier, J. C., and Yan, G. Y. (2001). Anopheline mosquito survival strategies during the dry period in western Kenya. *Journal of Medical Entomology*, 38(3):388–392.

Molineaux, L. and Gramiccia, G. (1980). *The Garki Project: Research on the Epidemiology and Control of Malaria in the Sudan Savanna of West Africa*. World Health Organization, Geneva.

Molineaux, L., Muir, D. A., Spencer, H. C., and Wernsdorfer, W. H. (1988). The epidemiology of malaria and its measurement. In Wernsdorfer, W. H. and McGregor, I., editors, *Malaria : principles and practice of malariology*, pages 999 – 1089. Churchill Livingston, New York.

Molyneux, D., Hopkins, D., and Zagaria, N. (2004). Disease eradication, elimination and control: the need for accurate and consistent usage. *Trends in Parasitology*, 20:347 – 351.

Mutero, C., Ngánga, P., Wekoyela, P., Githure, J., and Konradsen, F. (2004a). Ammonium sulphate fertilizer increases larval populations of *Anopheles arabiensis* and culicine mosquitoes in rice fields. *Acta Tropica*, 89:187 – 192.

Mutero, C. M., Kabutha, C., Kimani, V., Kabuage, L., Gitau, G., Ssennyonga, J., Githure, J., Muthami, L., Kaida, A., Musyoka, L., Kiarie, E., and Oganda, M. (2004b). A transdisciplinary perspective on the links between malaria and agroecosystems in Kenya. *Acta Tropica*, 89(2):171–186.

Mwangangi, J., Mbogo, C., Nzovu, J., Kabiru, E., Mwanbi, H., Githure, J., and Beier, J. (2004). Relationships between body size of *Anopheles* mosquitoes and *Plasmodium falciparum* sporozoite rates along the Kenya coast. *Journal of the American Mosquito Control Association*, 20(4):390–394.

Najera, J. (1989). Malaria and the work of WHO. *Bulletin of the World Health Organization*, 67:229 – 243.

Nasci, R. (1986a). The relationship between adult mosquito body size and parity in field populations. *Environmental Entomology*, 15(4):874–876.

Nasci, R. (1986b). The size of emerging and host-seeking *Aedes aegypti* and the

relation of size to blood feeding success in the field. *Journal of the American Mosquito Control Association*, 2(1):61–62.

Nasci, R. (1987). Adult body size and parity in field populations of the mosquitoes *Anopheles crucians, Aedes taeniorhychus* and *Aedes sollicitans*. *Journal of the American Mosquito Control Association*, 3(4):636–637.

Navarro, A. and Barton, N. H. (2003a). Accumulating postzygotic isolation genes in parapatry: A new twist on chromosomal speciation. *Evolution*, 57(3):447–459.

Navarro, A. and Barton, N. H. (2003b). Chromosomal speciation and molecular divergence: Accelerated evolution in rearranged chromosomes. *Science*, 300(5617):321–324.

Nghabi, K., John, B., Nkwengulila, B., Knols, G., Killeen, G., and Ferguson, H. (2005). Effect of larval crowding on mating competitiveness of *Anopheles gambiae* mosquitoes. *Malaria Journal*, 4:49.

Noor, M. A. F., Grams, K. L., Bertucci, L. A., Almendarez, Y., Reiland, J., and Smith, K. R. (2001). The genetics of reproductive isolation and the potential for gene exchange between *Drosophila pseudoobscura* and *D. persimilis* via backcross hybrid males. *Evolution*, 55:512–521.

Oaks, S. and Mitchell, V. (1991). *Malaria: Obstacles and opportunities*. National Academy Press, Washington, DC.

Onyabe, D. and Conn, J. (2001). Genetic differentiation of the malaria vector *Anopheles gambiae* across Nigeria suggests that selection limits gene fow. *Heredity*, 87:647–658.

Onyabe, D. Y., Vajime, C. G., Nock, I. H., Ndams, I. S., Akpa, A. U., Alaribe, A. A., and Conn, J. E. (2003). The distribution of M and S molecular forms of *Anopheles gambiae* in Nigeria. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 97(5):605–608.

Ortiz-Barrientos, D., Reiland, J., Hey, J., and Noor, M. A. F. (2002). Recombination and the divergence of hybridizing species. *Genetica*, 116(2-3):167–178.

Pannell, J. and Charlesworth, B. (2000). Effects of metapopulation processes on measures of genetic diversity. *Philosophical Transactions of the Royal Society of London Series B*, 355:1851 – 1864.

Peck, S. L. (2004). Simulation as experiment: a philosophical reassessment for biological modeling. *Trends In Ecology & Evolution*, 19(10):530–534.

Petrarca, V., Sabatinelli, G., Touré, Y., and Di Deco, M. (1998). Morphometric multivariate analysis of field samples of adult *Anopheles arabiensis* and *Anopheles gambiae s.s* (diptera: Culicidae). *Journal of Medical Entomology*, 35(1):16–25.

Pinto, J., Donnelly, M. J., Sousa, C. A., Malta-Vacas, J., Gil, V., Ferreira, C., Petrarca, V., do Rosario, V. E., and Charlwood, J. D. (2003). An island within an island: genetic differentiation of *Anopheles gambiae* in Sao Tome, West Africa, and its relevance to malaria vector control. *Heredity*, 91(4):407–414.

Plaen, R. D., Seka, M.-L., and Koutoua, A. (2004). The paddy, the vector and the caregiver: lessons from an ecosystem approach to irrigation and malaria in Northern Cote d'Ivoire. *Acta Tropica*, 89(2):135–146.

Powell, J. R., Petrarca, V., della Torre, A., Caccone, A., and Coluzzi, M. (1999).

Population structure, speciation, and introgression in the *Anopheles gambiae* complex. *Parassitologia*, 41(1-3):101–113.

R Development Core Team, R. (2004). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. 3-900051-07-0.

Robert, V., Gazin, P., Boudin, C., Molez, J., Ouedraogo, V., and Carnevale, P. (1985). The transmission of malaria in a wooded savannah area and a rice-growing area around Bobo Dioulasso (Burkina Faso). *Annales de la Societe Belge de Medecine Tropicale*, 65 Supplement 2:201 – 214.

Robert, V., van den Broek, A., Stevens, P., Slootweg, R., Petrarca, V., Coluzzi, M., Le Goff, G., Di Deco, M. A., and Carnevale, P. (1992). Mosquitoes and malaria transmission in irrigated rice-fields in the Benouevalley of northern Cameroon. *Acta Tropica*, 52:201 – 204.

Roll Back Malaria Project, W. (2003). Africa malaria report 2003. Technical report, WHO/UNICEF.

Rosengrant, M. and Perez, N. (2000). Water resources development for Africa: a review and synthesis of issues. Technical report, World Bank, Washington DC.

Ross, R. (1911). *The Prevention of Malaria*. Murray, London.

Roughgarden, J., Bergman, A., Shafir, S., and Taylor, C. (1996). *Adaptive individuals in evolving populations*, chapter Adaptive communication in Ecology and Evolution: A guide for future research, pages 25 – 30. Addison-Wesley, Reading, MA.

Russell, P. (1952). *Malaria: Basic principles briefly stated*. Blackwell Scientific Publications, Oxford.

Russell, P. (1955). *Man's mastry of malaria*. Oxford University Press, London.

Schneider, P., Takken, W., and McCall, P. (2000). Interspecific competition between sibling species larvae of *Anopheles arabiensis* and *An. gambiae*. *Medical and Veterinary Entomology*, 14(2):165–170.

Scott, J., Brogdon, W., and Collins, F. (1993). Identication of single specimens of the *Anopheles gambiae* complex by polymerase chain reaction. *American Journal Of Tropical Medicine And Hygiene*, 49:520–529.

Service, M. (1977). Mortalities of immature stages of species-b of *Anopheles gambiae* complex in Kenya: comparison between rice fields and temporary pools, identification of predators, and effects of insecticidal spraying. *Journal of Medical Entomology*, 13:535 – 545.

Service, M. (1989a). Irrigation: Boon or bane. In Service, M. W., editor, *Demography and Vector-Borne Diseases*, pages 238 – 254. CRC Press, Boca Raton, Florida.

Service, M. (1989b). Rice, a challenge to health. *Parasitology Today*, 5(5):162–165.

Service, M. (1993). *Mosquito Ecology: Field Sampling Methods*. Elsevier Applied Science Publishers, London, 2nd edition.

Sherman, I. (1998). *Malaria: Parasite biology, pathogenesis and protection*, chapter A brief history of malaria and discovery of the parasite's life cycle, pages 3 – 10. ASM Press, Washington DC.

Sissoko, M., Dicko, A., Briet, O., Sissoko, M., Sagara, I., Keita, H., Sogoba, M., Rogier, C., Touré, Y., and Doumbo, O. (2004). Malaria incidence in relation to rice cultivation in the irrigated Sahel of Mali. *Acta Tropica*, 89(2):161–170.

Slatkin, M. (1977). Gene flow and genetic drift in a species subject to frequent local extinctions. *Theoretical Population Biology*, 12:253 – 262.

Slatkin, M. (1985). Gene flow in natural populations. *Annual Review of Ecology and Systematics*, 16:393 – 430.

Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. *Science*, 236:787 – 792.

Slotman, M. A., Mendez, M. M., della Torre, A., Dolo, G., Touré, Y., and Caccone, A. (2006). Genetic differentiation between the Bamako and Savanna chromosomal forms of *Anopheles gambiae* as indicated by amplified restriction fragment length polymorphism analysis. *American Journal of Tropical Medicine and Hygiene*, 74:641 – 648.

Sokal, R. and Rohlf, F. (1969). *Biometry*. W.H. Freeman and Company, San Fransisco, 1st edition.

Soper, F. and Wilson, D. (1943). Anopheles gambiae *in Brazil: 1930 to 1940*. Rockafeller Foundation, New York.

Spielman, A. and D'Antonio, M. (2001). *Mosquito: The story of Man's deadliest foe*. Hyperion, New York.

StataCorp (2003). *Stata Statistical Software: Release 8*. StataCorp LP, College Station, TX:.

Stump, A., Fitzpatrick, M., Lobo, N., Traoré, S., Sagnon, N., Costantini, C., Collins, F., and Besansky, N. (2005). Centromere-proximal differentiation and speciation in *Anopheles gambiae*. *Proceedings of the National Academy of Sciences, USA*, In Press:–.

Suwanchaichinda, C. and Paskewitz, S. (1998). Effects of larval nutrition, adult body size, and adult temperature on the ability of *Anopheles gambiae* (Diptera : Culicidae) to melanize Sephadex beads. *Journal of Medical Entomology*, 35(2):157–161.

Takken, W., Klowden, M., and Chambers, G. (1998). Effect of body size on host seeking and blood meal utilization in *Anopheles gambiae* sensu stricto (Diptera : Culicidae): The disadvantage of being small. *Journal of Medical Entomology*, 35(5):639–645.

Taylor, C. and Powell, J. (1983). *The genetics and biology of Drosophila Vol. 3D*, chapter Population structure of Drosophila: genetics and ecology, pages 29 – 60. Academic Press, London.

Taylor, C., Touré, Y., Coluzzi, M., and Petrarca, V. (1993). Effective population size and persistence of *Anopheles arabiensis* during the dry season of West Africa. *Medical and Veterinary Entomology*, 7:351 – 357.

Taylor, C. E. and Manoukis, N. C. (2003). Effective population size in relation to genetic modification of *Anopheles gambiae* sensu stricto. In *Ecological Aspects for Application of Genetically Modified Mosquitos*. Proc. of the Frontis workshop on ecological challenges.

Taylor, C. E., Touré, Y. T., Carnahan, J. M., Norris, D. E., Dolo, G., Traoré, S. F., Edillo, F. E., and Lanzaro, G. C. (2001). Gene flow among populations

of the malaria vector *Anopheles gambiae*, in Mali, West Africa. *Genetics*, 157(2):743–750.

Thompson, M., D'Alessandro, U., Bennetts, S., Connor, S., Langerock, R., Jawara, M., Todd, J., and Greenwood, B. (1994). Malaria prevalence is inversely related to vector density in The Gambia, West Africa. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 94:159–163.

Thompson, R. (1948). Studies on *Anopheles gambiae* and *A. melas* in and around Lagos. *Bulletin of Entomological Research*, 36:395 – 417.

Touré, Y., Petrarca, V., and Coluzzi, M. (1983). Il complesso *Anopheles gambiae* in mali. *Parassitologia*, 25:367 – 370.

Touré, Y., Petrarca, V., Traoré, S., Coulibaly, A., Maiga, H., Sankare, O., Sow, M., Di Deco, M., and Coluzzi, M. (1998a). The distribution and inversion polymorphism of chromosomally recognized taxa of the *Anopheles gambiae* complex in Mali, West Africa. *Parassitologia*, 40:477 – 511.

Touré, Y. T. (1979). *Bio-ecologie des* Anopheles *(Diptera:Culicidae) dans une zone rurale de savane soudanienne au Mali village de Banambani – Arrondissement de Kati: Incidence sur la transmission du Paludisme et de la Filariose de Bancroft*. PhD thesis, Centre Pedagogie Superieur. Bamako, Mali.

Touré, Y. T., Dolo, G., Petrarca, V., Traoré, S. F., Bouaré, M., Dao, A., Carnahan, J., and Taylor, C. E. (1998b). Mark-release-recapture experiments with *Anopheles gambiae* sl in Banambani Village, Mali, to determine population size and structure. *Medical and Veterinary Entomology*, 12(1):74–83.

Touré, Y. T., Petrarca, V., Traoré, S. F., Coulibaly, A., Maiga, H. M., Sankaré, O., Sow, M., Dideco, M. A., and Coluzzi, M. (1994). Ecological genetic-studies

in the chromosomal form Mopti of *Anopheles gambiae* s-str in Mali, West Africa. *Genetica*, 94(2-3):213–223.

Trickett, A. J. and Butlin, R. (1993). Recombination suppressors and the evolution of new species. *Heredity*, 73:339 – 345.

Trigg, P. and Kondrachine, A. (1998). *Malaria: Parasite biology, pathogenesis and protection*, chapter The current global malaria situation, pages 11 – 24. ASM Press, Washington DC.

Tripet, F., Dolo, G., and Lanzaro, G. C. (2005). Multilevel analyses of genetic differentiation in *Anopheles gambiae* ss reveal patterns of gene flow important for malaria-fighting mosquito projects. *Genetics*, 169(1):313–324.

Tripet, F., Touré, Y., Dolo, G., and Lanzaro, G. (2003). Frequency of multiple inseminations in field collected *Anopheles gambiae* revealed by DNA analysis of transferred sperm. *American Journal of Tropical Medicine and Hygiene*, 68:1 – 5.

Tripet, F., Touré, Y. T., Taylor, C. E., Norris, D. E., Dolo, G., and Lanzaro, G. C. (2001). DNA analysis of transferred sperm reveals significant levels of gene flow between molecular forms of *Anopheles gambiae*. *Journal of Molecular Evolution*, 10:1725–1732.

Turchin, P. (1998). *Quantitative analysis of movement : measuring and modeling population redistribution in animals and plants*. Sinauer, Sunderland, MA.

Turner, T. L., Hahn, M. W., and Nuzhdin, S. V. (2005). Genomic islands of speciation in *Anopheles gambiae*. *PLoS Biology*, 3(9):e285.

Valentine, V. (2005). Fred L. Soper and Brazil's alien invasion. Internet resource.

van der Hoek, W. (2004). How can better farming methods reduce malaria? *Acta Tropica*, 89:95 – 97.

Victor, T. J. and Reuben, R. (2000). Effects of organic and inorganic fertilizers on mosquito populations in rice fields of Southern India. *Medical and Veterinary Entomology*, 14:361 – 368.

Waples, R. (1989). A generalized approach for estimating effective population size from temporal changes in allele frequency. *Genetics*, 121:379 – 391.

Warburg, A. and Touré, Y. (2002). Estivation of *Anopheles gambiae*: Potential habitats and physiology (report pn-acr-402). Technical report, US Agency for International Development.

Watterson, G. (1975). On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, 7:256 – 276.

Welsh, A., Cunningham, R., Donnelly, C., and Lindenmayer, D. (1996). Modelling the abundance of rare species: statistical models for counts with many zeros. *Ecological Modelling*, 88:297 – 308.

White, M. (1978). *Modes of speciation*. W.H. Freeman and Co., San Fransisco.

Whitlock, M. (1999). Neutral additive genetic variance in a metapopulation. *Genetical Research*, 74:215 – 221.

WHO (1955). *WHA8.30 Malaria Eradication, from the ninth plenary meeting, May 26, 1955; WHO official records no. 63*. World Health Organization, Geneva.

WHO (1969). *Re-examination of the global strategy of malaria eradication. A*

report by the director-general of the 22nd world health assembly, May 30, 1969. World Health Organization, Geneva.

WHO (1975). *Manual on practical entomology in malaria, part II of WHO offset publication No. 13*, volume Part II of *WHO Offset Publication No. 13*. WHO Division of Malaria and Other Parasitic Diseases, Geneva.

Wondji, C., Simard, F., Petrarca, V., Etang, J., Santolamazza, F., Della Torre, A., and Fontenille, D. (2005). Species and populations of the *Anopheles gambiae* complex in Cameroon with special emphasis on chromosomal and molecular forms of *Anopheles gambiae* s.s. *Journal of Medical Entomology*, 42(6):998–1005.

Wright, S. (1969). *Evolution and the genetics of populations. Vol. II. The theory of gene frequencies*. University of Chicago Press, Chicago, IL.

Yuval, B. (2006). Mating systems of blood-feeding flies. *Annual Review of Entomology*, 51:413 – 440.

Zhang, J., Wang, X., and Podlaha, O. (2004). Testing the chromosomal speciation hypothesis for humans and chimpanzees. *Genome Research*, 14:845–851.