Keele
University

An investigation into the fitness and mating competitiveness of laboratory and transgenic strains of *Anopheles gambiae* sensu stricto

Douglas Graeme Paton

Doctor of Philosophy

June 2013

Keele University

Abstract

Malaria is a deadly parasitic disease of humans spread through the bite of *Anopheles* spp. mosquitoes. Current control methods for the disease are broadly effective, but the spread of insecticide resistance in the principle Anopheline vectors of the disease raises the possibility of an increase in disease burden in the future. Transgenesis offers a novel alternative approach to vector control but requires the mass-release of virile, competitive, genetically altered male mosquitoes, thus the success of future transgenic release programmes depends greatly on how capable the transgenic strain is of surviving in field conditions and successfully introgressing with wild mosquito populations.

Using a combination of laboratory and field-based ecological experiments, along with molecular biological and genomic approaches, we assessed both the fitness of two transgenic strains of *Anopheles gambiae* s.s. and the genetic and environmental factors determining survival, mating success and assortative mating behaviour in lab and field derived samples of non-transgenic *Anopheles gambiae* s.s.

We found that imposed a fitness cost in both a transgenic strain carrying a phenotypic marker, and a second strain carrying a putative anti-malarial peptide sequence. Overt fitness costs were confined to larval development in both strains, although there was some evidence of a difference in egg production and morphology between strains. The anti-malarial peptide-carrying strain was significantly less fit, and suffered a fitness burden in hemizygote individuals as well as homozygotes. The possible sources of fitness differences are discussed. In semi-field-based crosses, we observed a significant interaction between the genetic and environmental background in the survival and mating success of non-transgenic strains; furthermore, the combination of laboratory rearing and a laboratory genetic background was sufficient to abolish the assortative mating behaviour normally observed between M- and S-molecular form *An. gambiae* populations.

SUBMISSION OF THESIS FOR A RESEARCH DEGREE

Part I. DECLARATION by the candidate for a research degree. To be bound in the thesis

Degree for which thesis being submitted

Title of thesis

This thesis contains confidential information and is subject to the protocol set down for the submission and examination of such a thesis.

YES/NO [please delete as appropriate; if YES the box in Part II should be completed]

Date of submission   Original registration date
(Date of submission must comply with Regulation 2D)

Name of candidate

Research Institute   Name of Lead Supervisor

I certify that:

(a) The thesis being submitted for examination is my own account of my own research
(b) My research has been conducted ethically. Where relevant a letter from the approving body confirming that ethical approval has been given has been bound in the thesis as an Annex
(c) The data and results presented are the genuine data and results actually obtained by me during the conduct of the research
(d) Where I have drawn on the work, ideas and results of others this has been appropriately acknowledged in the thesis
(e) Where any collaboration has taken place with one or more other researchers, I have included within an 'Acknowledgments' section in the thesis a clear statement of their contributions, in line with the relevant statement in the Code of Practice (see Note overleaf).
(f) The greater portion of the work described in the thesis has been undertaken subsequent to my registration for the higher degree for which I am submitting for examination
(g) Where part of the work described in the thesis has previously been incorporated in another thesis submitted by me for a higher degree (if any), this has been identified and acknowledged in the thesis
(h) The thesis submitted is within the required word limit as specified in the Regulations

Total words in submitted thesis (including text and footnotes, but excluding references and appendices)   **46272**

Signature of candidate   …………………………………Date   …………………

Note

*Extract from Code of Practice: **If the research degree is set within a broader programme of work involving a group of investigators – particularly if this programme of work predates the candidate's registration – the candidate should provide an explicit statement (in an 'Acknowledgments' section) of the respective roles of the candidate and these other individuals in relevant aspects of the work reported in the thesis. For example, it should***

*make clear, where relevant, the candidate's role in designing the study, developing data collection instruments, collecting primary data, analysing such data, and formulating conclusions from the analysis. Others involved in these aspects of the research should be named, and their contributions relative to that of the candidate should be specified (***this does not apply to the ordinary supervision, only if the supervisor or supervisory team has had greater than usual involvement***).*

Contents

Acknowledgements

It is a curious thing that this section of the thesis is the one I have looked forward to writing the most but, now that I come to it, is proving to be the most taxing of all. In a fashion that anyone who knows me even in passing will be familiar with, the original version of the acknowledgements was a rush job, completed as it was on the morning of submission, this revised and final version of my thesis allows me the opportunity to address not only a number of mistakes and corrections to the core text, but also to expand this section to a more appropriate size given the number of people that have, in ways both major and minor, contributed to my 'ascension' to PhD.

It's evolution baby!
Eddie Vedder

1. General Introduction

1.1 History: Man and Malaria

The history of man and malaria is a long and complex one; stretching back 6 million years into the evolutionary past to when the ancestors of modern-day chimpanzees and those of what would become modern humans went their separate ways. Even then, at the dawn of man, there was malaria (Escalante *et al* 1995), a wholly unwelcome witness to the span of human history.

Oblique reference to autumnal (aestival) agues, killing fevers and the relative risks of living in close proximity to marshlands can be found throughout man's early writings - Homer alludes to malaria in the Iliad[1] - but it is Hippocrates (c460 BCE – c370 BCE) who is credited with the first unequivocal description of malaria, writing as he did of the 48-hour cycle of fever and chills characteristic of the disease in his *Aphorisms,* and implicating living in proximity of marshy ground with risk of recurring fever. It would however, take almost two-and-a-half millennia before humanity would even begin to understand the true nature of malaria.

Following on from the near epoch-defining development of the germ theory of disease by Pasteur and Koch, pioneering work by Laveran in the late 1870's and early 1880's led him to correctly identify a protozoan parasite as the causative agent for malaria (Laveran, 1884), which was created *Plasmodium* spp. in 1888 by Marchiafava and Celli. The contemporary elucidation of the mosquito-dependent life-cycle of filarial worms by Manson in 1878, along with an extensive but ultimately unsuccessful search for a free-living stage of the malaria parasite, began to shift the paradigm of the day towards a

---

[1] *[Achilles], the old man Priam was first to behold with his eyes, as he sped all agleam across the plain, like to the star that cometh forth at harvest-time, and brightly do his rays shine amid the host of stars in the darkness of night, the star that men call by name the Dog of Orion. Brightest of all is he, yet withal is he a sign of evil, and bringeth much fever upon wretched mortals* **The Iliad,** **Book XXII, lines 25 - 32**

mosquito-centric model of disease transmission. This model was confirmed in 1898 by the much-celebrated work of (the soon to be Sir) Ronald Ross, an event which, along with Manson's filarial work, effectively precipitated the birth of the modern fields of tropical medicine and medical entomology. Ross's work was quickly followed by further important advancements such as Grassi's implication of Anopheline mosquitoes in the transmission of human malaria (Grassi, 1900).

The colonial powers of the day suddenly found themselves armed with all information needed to mount a directed assault on malaria, and took it upon themselves to instigate a massive and ostensibly successful programme of control, based primarily on the destruction of breeding sites. Therefore it is perhaps understandable, that over the next decades the scientific community was in ebullient mood; Ross even went as far as to call his 1923 autobiography: *Memoirs - with a full Account of the Great Malaria Problem and its Solution*. The subsequent development of synthetic anti-malarial drugs, during both World Wars, particularly that of chloroquine during World War II (Coatney, 1963), along with the Nobel Prize-winning discovery of DDT's application as an insecticide by Müller in 1939 (Fischer, 1948) added potent tools and renewed momentum to the push to eradicate malaria. So much so, that by the time the World Health Organisation's inaugural constitution came into force in 1948 (WHO, 2007) the disease was on the decline across much of Europe and North America and was permanently eradicated in the US by the late 1950's and in Europe by 1976 (WHO, 1978). This pattern of gradual retreat was reflected globally, to the extent that at the peak of their success, WHO mediated control programmes meant that 83% of the world's population were living in areas, in many cases newly, free from endemic transmission of malaria (Brown *et al* 1976).

Unfortunately success, in this and other cases of vector and disease control, was short-lived: many of the control programs lapsed in the absence of a perceived threat,

fuelled by the atrophy of both the political will and financial backing required for their maintenance (Killeen *et al* 2002). The situation was compounded by the socio-economic and political turmoil of the 1960's and 70's and the emergence of resistance in *Plamodium* to the front-line drug chloroquine (Young and Moore 1961). The end result was the eventual return of the endemic transmission of malaria to the majority of its former range within the tropics, a situation that did not significantly improve over the course of the 80's and early 90's (WHO, 2000).

The turn of the century has seen the sequencing of the genome of *Plasmodium falciparum* (Gardner *et al* 1998, Gardner *et al* 2002) and the principal Afro-tropical mosquito vector *Anopheles gambiae* sensu stricto (Holt *et al* 2002). This, in conjunction with the development and implementation of several new technologies such as the potent anti-malarial artesunate and its derivatives, insecticide treated materials (ITM) and recent progress in the genetic modification of vector species, will go some way to counteracting the recrudescence of malaria. In addition, there has been a general re-prioritisation of the problem of globally resurgent endemic malaria led primarily by the WHO-mediated Special Programme for Research and Training in Tropical Diseases (TDR) and the Roll Back Malaria (RBM) Programme (TDR 2005, RBM 2005) further facilitated by extensive charitable and philanthropic donations from organisations such as the Wellcome Trust and the Bill and Melinda Gates Foundation.

Nevertheless today, whilst global malaria mortality has reduced by 13% since 2000, there has been little change in the number of reported cases. Thus a large proportion of the world's population live at risk from malaria infection and with over 215 million reported cases and 655,000 deaths in 2010 (WHO 2011) malaria is still a massive problem. Even with renewed commitment to control this disease, it is clear from the scale of the problem, and from the lessons of the past, that there can be no singular solution to the

malaria issue. Thus, there is a continual and ever-growing need for research into the basic biology of both parasite and vector; the mechanisms of disease and transmission; and, leading from these, the development of new treatments, therapies and control methods with the goal of achieving a manifold approach to tackling this most deadly of parasitic diseases.

1.2 Anopheles spp: Global Distribution, Tropical Vector

All mosquitoes exist within the family Culicidae, which consists of three subfamilies: Toxorhynchitinae, Culicinae and Anophelinae. Members of the subfamily Toxorhynchitinae do not require a blood meal for egg production (i.e. are autogenous) and are therefore not medically important as a vector - although they have been proposed as a form of biological control as their larval stages are natural predators of other mosquito larvae that *are* important vectors. Members of both the Culicinae and Anophelinae are anautogenous (i.e. they require a blood-meal for egg production) and both subfamilies contain important vectors of human diseases. The Culicinae comprises, among others, the genera *Aedes* spp., *Culex* spp. and *Mansonia* spp. which are collectively responsible for the transmission of dengue and yellow fever, Japanese encephalitis and several other arboviruses, as well as *Wuschereria bancrofti, Brugei malayi and B. timori*, the filarial worms responsible for lymphatic filariasis. The Anophelinae are also vectors of lymphatic filariasis and arboviruses such as o'nyong'nyong, but are of primary importance as the vectors of human malaria (*Plasmodium* spp.) (Service 2004).

Mosquitoes of the genus *Anopheles* are effectively ubiquitous. There are approximately 460 recognised species within the genus (Krzywinski and Besansky 2003), occurring on every continent, barring Antarctica, at altitudes up to 3000m and across a diverse range of local, regional and continental climactic conditions (Reiter 2001). Despite

this wide range and the consequent variation in local environmental conditions, the essential details of the life cycle of *Anopheles* mosquitoes are constant: eggs are laid individually by gravid females (contrast with *Culex* spp./*Mansonia* spp. which lay eggs in rafts or in batches on fresh-water plants respectively). Broods consist typically of between 50-200 eggs; but this can vary dependent on the species, the size of the blood meal and the interaction of genetic and environmental factors (Hogg and Hurd 1995, Takken *et al* 1998). Brood size (fecundity) is a strong measure of fitness (Moreira *et al* 2004, Moreira *et al* 2007). Eggs are dark brown or black in colour and between 1-2 mm in length, have characteristic air-filled sacs - floats - positioned laterally, and are susceptible to desiccation. Oviposition sites vary between species but are typically free of organic pollutants (in contrast to *Culex quinquifasciatus*) and can range from the temporary, such as rain-filled hoof prints, to permanent bodies of fresh and in some cases brackish water (e.g. *Anopheles meras*). Eggs of tropical species will hatch releasing first instar larva within 2 days of the time of oviposition. This takes significantly longer in temperate climes (Service 2004).

Anopheles spp. larvae characteristically position themselves parallel to the water surface, breathing through spiracles located on the 8th abdominal segment and utilise brushes on their mouth-parts to filter feed on small food particles and micro-organisms. Larvae progress through four larval instars, separated by three moults, which in optimal conditions takes 10 to 14 days, although this can again take significantly longer in temperate species. Fourth instar larvae metamorphose into pupae, which are motile but do not feed. Pupae are 'comma' shaped and breathe through two conical respiratory trumpets situated dorsally on the thorax. The process of metamorphosis into the imago (adult) life stage is, again, dependent on the prevalent environmental conditions, with ecdysis

occurring within 2-3 days in the tropics, the pupal exoskeleton splits centrally along the dorsal surface allowing the adult mosquito to emerge (Service 2004).

Immediately post-emergence the adult cuticle hardens and darkens and the wings dry. Adult mosquitoes are not fully mature upon ecdysis from the pupal stage; both males and females require a further 2-3 days to reach sexual maturity as the testes and ovaries develop respectively, in addition the external male genitalia, which face dorsally upon emergence, rotate through 180º to face ventrally during this time, this facilitates copulation. Similarly to larvae, adults can be easily distinguished from Culicine mosquitoes. Adult *Anopheles* are typically dark grey or dark brown and have a characteristic banded pattern of dark and pale scales along the anterior margin of the wing (this feature can also facilitate identification to the species level). Resting adult *Anopheles* mosquitoes also hold their body at a characteristic angle to the surface upon which they are resting which is in contrast to the body position of Culicine mosquitoes. Male Anophelines have palps that are the same length as the proboscis and are slightly swollen or clubbed at the end, females also have long palps but lack the clubbed end. Discrimination of sex is also achieved by considering the antennae, which in males are plumose but non-plumose in females. In addition the last abdominal segment in males has characteristic claspers, which are utilised to hold the female in position during copulation. Specific determination requires the observation of microscopic morphological features such as setae on the gonocoxites of the male genitalia or observation of characteristic chromosomal inversion patterns of stained polytene chromosomes from adult female ovarian nurse cells and 4th instar larval salivary glands (Service 2004, Coluzzi *et al* 2002).

The behaviour of adult mosquitoes in the field, particularly that of the female which requires a blood meal from a vertebrate host for egg development, can have a profound effect on the efficiency with which it will transmit malaria from human to human. Host

seeking and choice, feeding and post-blood meal behaviour as well as mosquito longevity all interact to determine the vectorial capacity of a given species. Mosquitoes which feed primarily on humans are termed anthropophagic (or anthropophilic), whilst those that feed primarily on other animals are termed zoophagic. This is not an absolute distinction: different species of *Anopheles* spp. exhibit varying degrees of anthropophagy and the degree of anthropophagy can vary within a species depending on environmental factors (Pates *et al* 2001). Clearly a high degree of anthropophagy increases the likelihood that a given mosquito will take a blood meal from an infected human host and subsequently transfer the infection. Where and when the mosquito feeds can also effect it's efficiency as a malaria vector. Mosquitoes can be classed as either exophagic; feeding primarily outside or endophagic; feeding indoors, again this is not absolute. The majority of *Anopheles* spp. are crepuscular or nocturnal feeders, thus, those that preferentially feed indoors have a potentially higher capacity for transmission, although this is also dependent on the behaviour of the human host (Service 2004). Finally the post-blood meal behaviour of the mosquito: either resting indoors (endophilic) or outdoors (exophilic) can effect vectorial capacity. Overarching these behaviours is the intrinsic longevity of the mosquito: in the tropics, *Plasmodium* spp. requires approximately 2 weeks to develop to an infectious state within a host mosquito, thus factors which effect the longevity of the mosquito effect its capacity to transmit malaria.

1.3 The Anopheles gambiae *sensu lato species complex*

In the Afrotropical region the principal vectors of malaria are members of the *Anopheles gambiae* species complex and *An. funestus* group. The *An. gambiae* species complex comprises seven morphologically indistinguishable but behaviourally and, to a certain extent, genetically distinct species that mate assortively but are incompletely

reproductively isolated such that female hybrids are fertile but males are not (Coluzzi *et al* 2002). These are: *Anopheles gambiae* sensu stricto, *An. arabiensis, An. quadriannulatus* A and B, *An. merus, An. melas* and *An. bwambae. An. merus* and *An. melas* breed primarily in brackish habitats on the Eastern and Western coastal areas of Africa respectively and, whilst they are locally important vectors of malaria, their limited range reduces their impact as a public health issue. Similarly *An. bwambae* whilst a vector, is confined to mineral springs in the East African Rift Valley area. The remaining members; *An. gambiae* s.s., *An. arabiensis* and *An. quadriannulatus* have a continental distribution but vary wildly in their vectorial capacity: *An. quadriannulatus* is a primarily zoophagic species and not considered an important vector. In contrast *An. arabiensis* and, in particular, *An. gambiae* s.s. are both highly anthropophagic and are thus extremely efficient vectors of malaria (Coluzzi *et al* 1979, Hunt *et al* 1998, della Torre 2005). Identification of these sibling species is dependent on observation of fixed chromosomal inversion polymorphisms that exist between members, although PCR-based technologies have been developed that allow differentiation between the various sibling species (Scott *et al* 1993).

1.4 Population Structure of Anopheles gambiae *sensu stricto*

Anopheles gambiae s.s. exhibits a particularly high degree of synanthropy and is adapted to take advantage of both temporary and man-made breeding sites as well as large range of climactic conditions. These factors combined with a propensity for endophily, endophagy, almost exclusive anthropophagy and relatively long-lived adult females mean that *An. gambiae* s.s. (from here onwards *An. gambiae*) is an extremely efficient vector of malaria (Coluzzi *et al* 1979). For these reasons *An. gambiae* is one of the most studied and medically important single vector species.

Investigations during the late '70s and 1980s into the feeding behaviour and environmental adaptation of *An. gambiae* to the diverse conditions of relative aridity, ranging from dry Sahel savannah to humid tropical forest areas that are found throughout the range of this mosquito, led to the discovery of a number of paracentric chromosome inversion polymorphisms which showed variation in frequency along aridity clines. Inversions that are present at or near fixation in arid environments - such as inversion 2La, are present at very low frequencies in more humid areas and, proceeding along a gradient of relative aridity from humid to arid, the frequency of such inversions alters in line with changing environmental conditions (Coluzzi *et al* 1985). These inversions all occur on chromosome 2 (*Anopheles* spp. have three chromosome pairs: the autosomal chromosome pairs 2 and 3 and the x-body), with one inversion (2La) on the left arm of the chromosome, and the remainder (2Rb, c, d, j and u) on the right arm (Coluzzi *et al* 2002, della Torre *et al* 2005). The discovery of deviations from Hardy-Weinberg equilibrium (HWE) and an absence of heterokaryotypes for specific combinations of predicted hybrid chromosomal inversion patterns by a number of investigators resulted in the designation of five 'chromosomal' forms within *An. gambiae.* These chromosomal forms were believed to be at least partially, and in some cases almost totally reproductively isolated in the field. They are designated Forest, Savannah, Mopti, Bamako and Bissau and were believed to represent populations undergoing incipient speciation (Coluzzi 1992, Touré *et al* 1998).

The Forest form is believed to be ancestral and typically displays the standard arrangement on chromosome 2, with 2Ru, 2Rb and 2Rd present at low frequencies. This form is associated with humid tropical forest environments and the transitional environments between forest and humid savannah areas. The Savannah form is characterised by presence of the 2Rb and 2La inversions. In Southern Mali, Savannah forms also exhibit inversions 2Rcu and 2Rbcu in HWE with the standard arrangement. In

Nigeria the balanced polymorphism 2Rcu/bcd also exists within the 2Rb$^+$, 2La$^+$ typical of the Savannah form. This form is widespread across Africa though its frequency is highly dependent on season, being primarily associated with the wet-season and the concomitant increase in natural and temporary breeding sites that increased rainfall and humidity brings. The Bamako chromosomal form is characterised by the presence of 2Rj and is found primarily in Southern Mali. The Mopti chromosomal form exhibits the balanced polymorphism 2Rbc/u and is fixed for 2La, this form is typically associated with arid/sahel savannah such as areas of central Mali and also with artificial/man-made permanent breeding sites and thus can maintain a relatively high population even over the tropical dry season. Finally the Bissau form, which is found in coastal areas of West Africa and frequently associated with agriculture, is characterised by high frequencies of the 2Rd inversion (della Torre 2002, della Torre *et al* 2005). Low level introgression, evidenced by the observation of hybrid heterokaryotypes is believed to occur in the field between some chromosomal forms - for example between sympatric Mopti and Savannah forms in Mali (Touré *et al* 1998). However it has never been adequately demonstrated that putative introgressed or hybrid heterokaryotypes observed in the field do not represent rare 'non-form' inversions that exist within a form at a low level. The true degree of assortative mating occurring in the field between chromosomal forms remains a controversial and debated topic (Yawson *et al* 2007, Lee *et al* 2009) and further discoveries have brought the taxonomic status of chromosomal forms into question.

Investigating paracentric inversion karyotype frequencies by observation of polytene chromosomes from the ovarian nurse-cells of semi-gravid females is labour intensive and reduces the testable population by limiting the number of individuals that can be studied as it excludes all males and a sub-set of females from any study. Understandably, there has been an effort to discover a reliable molecular method for

distinguishing sub-populations of *An. gambiae* s.s. In 1994, Favia and colleagues reported that they were able to distinguish between forms using RAPD (Random Amplification of Polymorphic DNA). Unfortunately this method has since proven unreliable. However the same group later reported that restriction digest of a 1.3 kilobase (kb) PCR amplicon comprising the 3' end of the X-linked 28s ribosomal DNA and the downstream intergenic spacer (IGS) region produced a banding pattern that was able to differentiate Mopti, and Savannah/Bamako chromosomal form mosquitoes sampled from Burkina Faso and Mali (Favia *et al* 1997, Fanello *et al* 2002). Subsequent studies, which sequenced these regions, identified a single nucleotide polymorphism within the IGS region of the rDNA as the cause of the differing restriction digest banding pattern. These 'molecular forms' were designated M-form - corresponding to the Mopti chromosomal form in Mali - and S-form - corresponding to the Savannah and Bamako chromosomal form in Mali (Gentile *et al* 2001, Gentile *et al* 2002). A site within the internal transcribed sequence (ITS) of the same rDNA sequence containing three fixed nucleotide differences in complete linkage disequilibrium with the M- and S-form was also identified by the same investigators. These were designated Type I (in linkage with S-Form) and Type II (in linkage with M-form). A third ITS type was discovered in the isolated mosquito populations from the island of São Tomé off the West coast of Africa (Gentile *et al* 2002).

The availability of an unambiguous and co-dominant molecular marker has allowed investigators to study the complexities of the population structuring within in *An. gambiae* s.s. on a much larger scale than was possible with karyotypic analysis. Studies had until recently uniformly found a high degree of assortive mating between M- and S-form populations, though both hybrids and hybrid mating events have been rarely ($\ll$1% della Torre *et al* 2005) observed in adults (Tripet *et al* 2001) and in larvae (Edillo *et al* 2002). M-form individuals are, with the exception of one population from Kanyemba, Zimbabwe,

concentrated to Western areas of Africa, while the S-form has a continental distribution (della Torre *et al* 2005). Interestingly, the tight association between chromosomal and molecular form observed in Burkina and Mali breaks down outside of these areas. The Savannah form, which is exclusively S-form in Mali, is also observed throughout the Gambia, but in the populations sampled there, is almost completely Type II/M-form (della Torre *et al* 2005). The Savannah-M combination of karyotype and molecular form is also observed in populations from Angola. Similarly Forest form (i.e. the standard chromosome arrangement) individuals in Cameroon exhibit both M- and S-form IGS-types in sympatry and, based on the degree of genetic differentiation exhibited at a sub-set of microsatellites, appear to be completely reproductively isolated (Wondji *et al* 2002).

1.5 Current Control

The mass-distribution of insecticide treated nets (ITNs) and large-scale indoor residual spraying (IRS) of insecticides have been extremely effective in reducing the incidence of malaria in endemic countries (Lengeler 2004). Despite these successes, there is a real danger that the great strides in malaria control made in the last 20 or so years could be undone; resistance to insecticides in the key sub-Saharan malaria vectors *Anopheles gambiae sensu stricto, An. arabiensis* and *An. funestus*, either through target-site insensitivity or enzyme-mediated metabolic detoxification has been described in a number of sub-populations and is spreading. The situation is compounded by the fact that the four available classes of insecticide act on only 2 targets within the mosquito, not only vastly increasing the selective pressure on the respective target site but causing cross-reactive target site resistance (reviewed in Ranson *et al* 2011). Further, metabolic cross-resistance was recently described for the first time in *An. gambiae s.s*. (Mitchell *et al* 2012).

Clearly there is a demonstrable need for development of not only new insecticides, but also novel and alternative approaches to vector control. In response to this, since the late nineties, there has been a renewed interest in - and rapid expansion of - research into alternative methods of vector control based on the genetic manipulation of the vector; either through radio- or chemosterilisation (sterile insect technique, Benedict and Robinson 2003) or, genetic manipulation of the vector (transgenesis, Alphey *et al* 2002). Underpinning these approaches is the requirement to consistently raise, sort and release potentially millions of healthy and sexually competitive male mosquitoes into the environment. This presents a number of challenges which need to be overcome: ranging from the environmental, behavioural and genetic (Jones and Gubbins 1978, Reisen *et al* 1980) to the political and regulatory issues surrounding the release of genetically modified organisms (Coleman and Alphey 2004, C. Curtis Pers. Comm.).

Engineering a Solution by Engineering the Vector

In 1982, scientists successfully transformed *Drosophila melanogaster* using an endogenous transposable element (TE) known as the *P*-element (Rubin and Spradling 1982). TEs are naturally occurring sequences that can separate and integrate into a genome, either by transcription (Class 1), which results in multiple copies of the sequence, or by excision and integration without an RNA mediated step (Class 2) (Brock *et al* 2000). Rubin and Spradling were able to utilise these properties to insert a dominant eye-colour gene into a white-eyed population of *Drosophila* by cloning the *cinnabar* gene into a *P*-element and subsequently micro-injecting this into the developing embryo.

This was the first time a multi-cellular organism had been transformed and understandably the results were announced with some fanfare. However, it would not be until 1998, fully 16 years later, that the first transgenic mosquitos, *Aedes aegypti*

transiently expressing a co-dominant *Drosophila* eye colour gene, would be created (Jasinskiene *et al* 1998). This delay was primarily caused by the assumption that the *P*-element that has proven so successful in *Drosophila* would function in the Culicidae - an assumption that would ultimately prove erroneous - and a dearth of robust phenotypic markers (Jacobs-Lorena 2003). The initial transformation of *Aedes* utilised a transposable element known as *Hermes* that was originally isolated from the house fly *Musca domestica* (Jasinskiene *et al* 1998) and there are now several other robust TE systems available that have proven efficacy in mosquito transgenics such as *piggyBac*, *Minos* and *mariner* (O'Brochta and Handler 2008). In addition the development of fluorophores such as the green and cyan fluorescent proteins (GFP, CFP) and *Ds*Red as dominant markers (Pinkerton *et al* 2000) has greatly enhanced the process of creating transgenic mosquitoes and ameliorated the need to characterise and clone endogenous phenotypic markers (Jacobs-Lorena 2003).

Scientists now have several robust transformation mechanisms and unambiguous markers available to them through which they can create transgenic mosquitoes, and in the 11 years since the first transgenic *Aedes aegypti,* they have successfully transformed both Culicine and subsequently, Anopheline mosquitoes. *Anopheles gambiae* has proven particularly difficult to transform, although this has now also been achieved. With the technology and expertise in place, focus has now shifted to the creation of transgenic mosquitoes with a practical application. There are two major paradigms being pursued: Release of insects with a dominant lethal mutation (RIDL, Alphey 2002) - effectively a genetically enhanced sterile insect technique (SIT), and population replacement - creation and release of malaria-refractory mosquitoes with the goal of permanently altering the genetic make-up of wild populations (O'Brochta 2003).

Sterile insect technique (SIT) involves the captive rearing of many millions of insects, separating the males and sterilising them by exposing them to ionising (usually x- or γ-) radiation and releasing them into infested areas. Sterile males compete for females and as a result reduce mating success and the eventually achieves eradication of the target organism. Advantages of this form of control are that, in contrast to insecticide-based control strategies, the efficacy of SIT actually increases as the population density is reduced by successive rounds of sterile male release, as the ratio of sterile to fertile males increases, and it is ecologically benign (Feldman and Hendrichs 2001).

Inducing changes in vector or pest insects in captivity and releasing them into wild populations as a method of control is not a new idea. It was initially conceived in the 1930's, successfully tested in the late 40's and eventually used to successfully eradicate the New World Screwworm, *Cochliomyia hominovorax,* from North and Central America in the latter half of the 20th century, and laterally to eradicate a population accidentally introduced into North Africa (Klassen and Curtis 2005). Attempts at applying these techniques to mosquito control have been broadly unsuccessful, primarily due to the reduced competitiveness of radiation sterilised males and opposition from national governments (Benedict and Robinson 2003) and unexpected population structuring (Reisen *et al* 1982). However, the concept has undergone something of a renaissance since the turn of the century largely due to the advent of robust transgenic technologies (Alphey 2002, Benedict and Robinson 2003). The basic premise of RIDL is effectively identical to SIT, in that it involves the inundatory release of male mosquitoes that, rather than having been radiation-sterilised, carry a transgene which induces mortality in their offspring (Alphey 2002).

The population replacement strategy relies on the creation of transgenic *Anopheles* expressing one or several anti-malarial peptides that attack the parasite as it develops

within the mosquito. This in turn requires the identification of anti-*Plasmodium* effecter genes that are not toxic to the mosquito, and the characterisation of endogenous genetic regulatory sequences to control the expression of this effecter gene. Once a refractory mosquito has been produced it would then be released into, and introgress with, wild populations thus spreading malaria refractoriness and, theoretically, interrupting transmission of the disease.

There has been an extensive search for candidate anti-malarial peptides not only as potential transgenes, but also as therapeutics. Mosquitoes naturally express three classes of AMPs that are active against *Plasmodium*: Cercropins, maganins and defensins (Lehmann *et al* 2009). Exogenous AMPs from other arthropods such as PLA2 (*Bombus* spp. Moreira *et al* 2002) scorpine (*Pandinus imperator,* Zhu and Tytgas 2004) and gomesin (*Acanthoscurria gomesiana,* Moreira *et al* 2007) also show promising anti-malarial activity. Finally, synthetic AMPs such as SM1 (Ghosh *et al* 2001) or VIDA3 (Arrighi *et al* 2002) have been created based on phage-display assays and structural homology with natural peptides respectively.

The activity of these AMPs has been primarily assayed *in vitro* or using non-natural model systems such as the *Plasmodium berghei*/*Anopheles gambiae* combination which has questionable worth for assessing AMP efficacy (Tripet 2009). Very few have been demonstrated to have anti-*P. falciparum* activity in anything approaching a natural system - *P. falciparum* is notoriously difficult to culture through a full sexual/infectious cycle *ex vivo* (Kaushal *et al* 1980, Fivelman *et al* 2007)) and fewer still have actually been expressed in a transgenic mosquito - PLA2, SM1 and VIDA3 are notable exceptions. However new technologies such as site-specific transgenic integration and *Plasmodium falciparum*/*Anopheles gambiae* infection models should make assessing the efficacy of these AMPs both easier and more demonstrative.

Tailoring the expression of a putative AMP in a transgenic is arguably as important as the efficacy of the transgene itself. Expression should, ideally, be limited to specific tissues and/or developmental stages to limit any potentially deleterious effects of foreign gene expression on the transgenic mosquito, but must also ensure that peak levels of the transgene coincide with the presence of the parasite. Marker genes are required to be expressed constitutively (i.e. always on), in transformants, this is typically achieved using the synthetic P3 promoter, which is constitutively expressed in the mosquito eye and, in some cases, the larval nerve ganglia and anal papillae (Horn and Wimmer 2000).

Plasmodium is susceptible to attack by AMPs in two tissues in the mosquito: the mid-gut and the haemocoel and promoters have been characterised for both of these compartments. The *Anopheles gambiae carboxypeptidase* (*AgCP*) gene is up-regulated in response to a blood meal and the enzymatic product is excreted into the mid-gut lumen, reaching peak production 24 hrs post-blood meal (Edwards *et al* 1997). This pattern is also observed when the *AgCP* promoter and 3' UTR are used to drive transgene expression and could be used to drive the expression on an AMP that is active against the *Plasmodium* ookinete (Moreira *et al* 2000, Moreira *et al* 2002). It has the additional advantage of being both sex- and developmental stage-specific as it is only expressed by adult female mosquitoes. Similarly the *Anopheles gambiae vitellogenin* (*AgVG*) promoter is up-regulated in adult females following a blood meal but is active in the fat-body of the mosquito (Nirmala *et al* 2006). Transgenic mosquitoes using *AgVG* to drive expression of a reporter gene demonstrated that gene products freely diffused into the haemocoel. Transgenic mosquitoes expressing AMPs have been created which utilise both the *AgVG* and *AgCP* promoter sequences (Moreira *et al* 2002, Li *et al* 2009).

1.6 Aims and Objectives

The primary goal of population replacement is to drive a foreign, novel anti-malarial gene into wild populations. Thus transgenic mosquitoes must be highly competitive relative to non-transgenic wild populations. This goal is complicated by a number of factors namely. 1) The potential for lack of fitness or competitiveness in transgenic strains induced by the process of transformation; either through an unintended directly deleterious effect of carrying the transgene, or through the fixation of deleterious alleles linked to the transformation site. 2) A lack of fitness or competitiveness induced by long-term colonisation and rearing of transgenic strains and their non-transgenic primogenitors. 3) Cryptic population structuring within a target population which may erect significant barriers to gene flow and thus supress the spread of a transgene into a target population.

In this thesis, we present data from a series of experiments investigating the fitness and mating competitiveness of laboratory and transgenic strains of *Anopheles gambaie* sensu stricto. We investigate the effects of transgenesis and inbreeding on the performance of two transgenic strains in the lab and attempted to elucidate some of the environmental and genetic factors determining mating behaviour and survival in colonised and field-derived *An. gambiae s.s*

Our main objectives were as follows:

1. Investigate transgenesis in the context of the key ecological factors of egg production and larval development in two transgenic strains - EE and EVida3 (described previously) - notable due to their shared transgenic loci but differing transgenic cassettes.

2. Using cage invasion experiments, assess the long-term stability of these transgenic cassettes in mixed populations and partition the sources (if any) of reduced fitness in transgenic individuals

3. Quantify the main and interactive effects of genetic and environmental factors typically associated with colonisation and laboratory rearing, using a small-scale semi-field system.

4. Using high throughput ultra-deep pyrosequencing, investigate whether genetic variability at potential assortative mating or speciation genes is lost during the colonization and lab rearing process. We will also attempt to determine if selection for lab-mating might have selected for the fixation of mutant alleles or rare polymorphisms.

2 General Methods

2.1 Insectary Conditions

All strains in the UK were maintained in dedicated insectary facilities at Keele University. Temperature is maintained at a constant 27±2ºC, and relative humidity at 70±5. A 12h light/dark cycle was maintained using a digital timer and light control. Larvae were grown at a density of 200 larvae/l and initially fed with a suspension of yeast cells (Liquifry, Tetramin), followed by an optimized regimen of ground fish food (Tetramin). Upon pupation, pupae were transferred by aspiration to a standard rearing cage made of a 5l white polypropylene bucket (~20.5 cm height×20 cm diameter) with a sleeved side opening for introducing and removing mosquitoes and accessories, and with the top covered with a mosquito netting cylindrical enclosure. Adults were provided with water and a 5% glucose solution. To induce egg production, adult flies were fed a blood meal consisting of defibrinillated human blood warmed to 37°C delivered through a membrane feeder system (Hematek). After 48h gravid females were provided with a polystyrene oviposition cup lined with filter paper (grade 1, Whatmann) and filled to ~2cm depth with ddH$_2$O. After a further 72h, the ovipostion cup was retrieved and the eggs distributed into fresh trays as described above. Stock population size was typically maintained at 800-1200 larvae per generation.

2.2 Insect Strains

In our experiments investigating the effects of transgenesis on fitness we utilised two transgenic strains previously described in Meredith *et al* 2011. The Phase 1 EE strain carries a transgene cassette consisting of the phenotypic marker ECFP under the control of the 3xP3 promoter driving its expression in the eyes and other nerve tissues, and the phiC31 integrase recognition sequence *attP* (Thyagarajan *et al* 2001)(figure 2.1,2.3). The
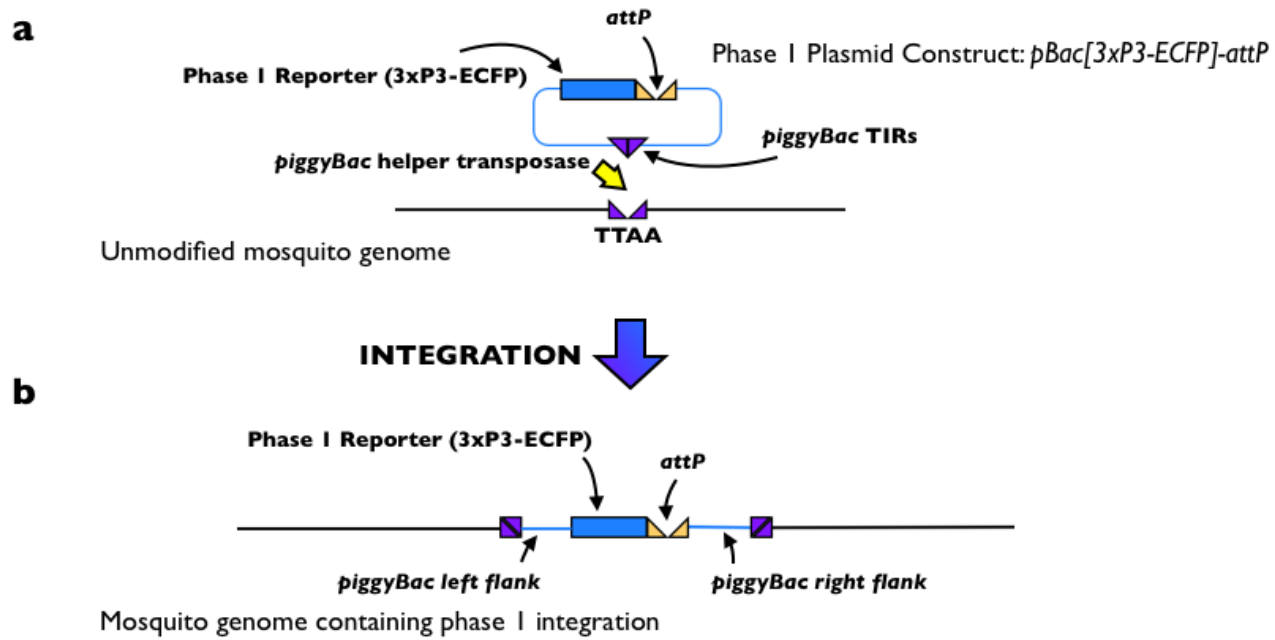
Figure 2.1: Schematic representation of phase 1 of the two phase piggBac/phiC31 mediated site specific transgene integration system. A) Pre-insertion at a random TTAA site. B) Post-insertion at a random TTAA site.
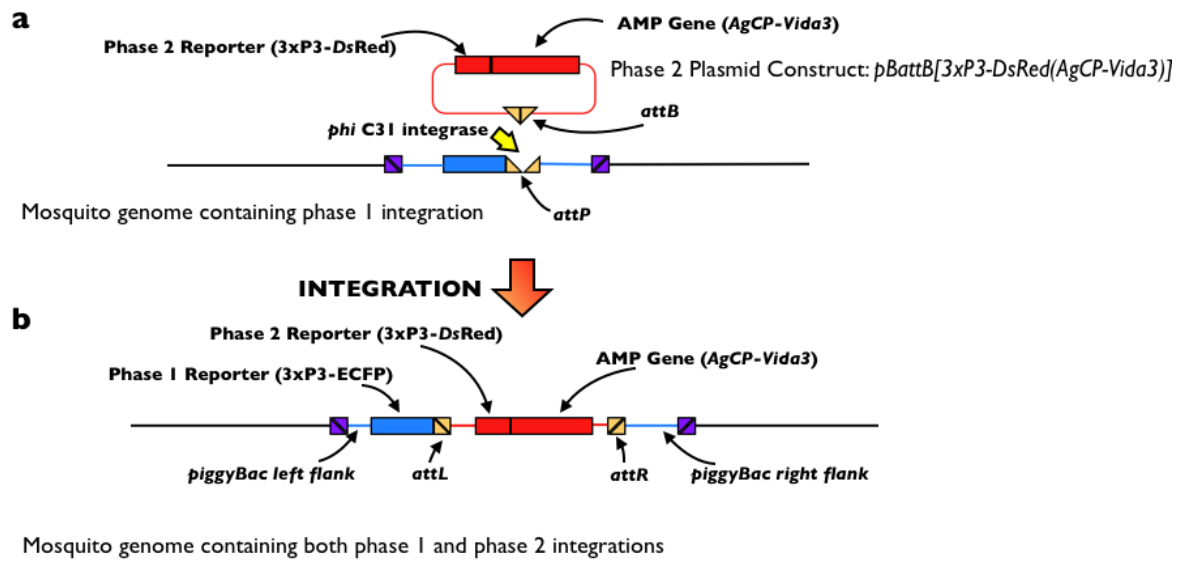
Figure 2.2: Schematic representation of phase 2 of the two phase piggBac/phiC31 mediated site specific transgene integration system. A) Pre-insertion at the phase 1 specific *attP* site. B) Post-insertion at at the phase 1 specific *attP* site.

Phase 2 EVida3 strain is derived from the EE strain in a second transformation step and carries a cassette consisting of 3xP3 ECFP, an additional marker 3xP3 DsRed and the putative AMP Vida3 sequence with the An. *gambiae carboxypeptidase* promotor, signal peptide and UTRs (Meredith *et al* 2011)(figure 2.2,2.3). The Phase 1-2 integration site has been identified on chromosome 3R (position 15801959 - band 31B) and is therefore located away from any of the inversion polymorphisms commonly found in An. gambiae s.s. (Coluzzi *et al* 1985, Toure *et al* 1998). The two transgenic lines were derived from the wild-type strain KIL originally colonized from Tanzania in the 1970's. Both transgenic strains are of the M molecular form (della Torre *et al* 2001). The wild-type strain used in all experiments is a Mopti, M-form population originally colonized from the village of N'Gabakoro Droit, Koulikorou, Mali (12°39'46"N, 7°50'34"W) 2003. Since it has been in our laboratory, the Mopti strain has refreshed yearly by outcrossing to the F1 of field caught individuals from the same site.

2.3 Whole Body gDNA Extraction

DNA was extracted from individuals using a modified DNAzol (Invitrogen) protocol: mosquito carcasses were homogenised in 100 µl DNAzol reagent, the soluble fraction was isolated by centrifugation at 10000 *g* for 10 minutes and the supernatant retained in fresh 1.5 ml microcentrifuge tube. DNA was precipitated by the addition of 50 µl 100% ethanol, and peletted by centrifugation at 10000 *g* for 10 minutes. The isolated DNA was then washed twice by resuspending in 750 µl 70% ethanol and centrifugation at 10000 *g* for 10 minutes. Following washing, the ethanol wash was discarded and the remaining pellet of DNA allowed to dry in a 37 °C incubator for 3-24 hours. Once dry, the DNA was resuspended in 200 µl ddH$_2$O. Samples were stored at -20°C for up to a month, and at -80°C for longer periods.
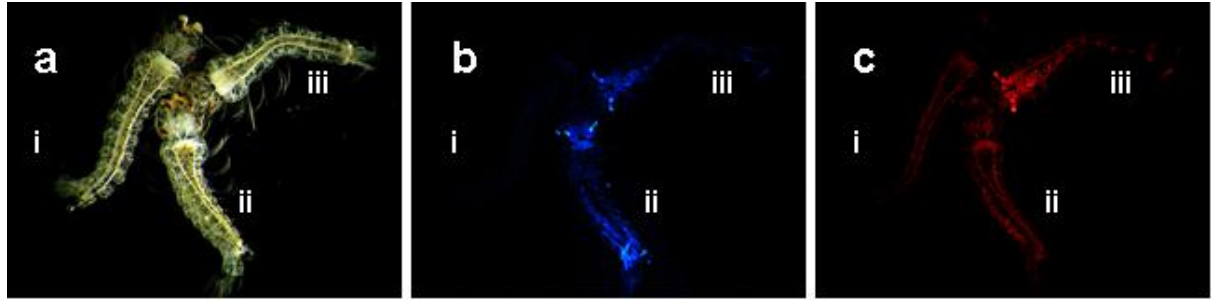
Figure 2.3: KIL non-transgenic (i) EE, Phase 1 transgenic (ii), and EVida3, Phase 2 transgenic (iii) under differing lighting conditions. (a) Natural light. (b) 439nm (violet) excitation/476nm (cyan) emission. (c) 563nm (green) excitation/582nm (orange/red) emission. Images courtesy of Dr S. Basu.

2.4 Co-dominant PCR assay to determine transgenic status.

A prerequisite for many of our planned experiments was an unambiguous, reliable method for determining the genotype of individual mosquitoes sampled from mixed transgenic and non-transgenic populations. To this end, we set out to develop a co-dominant PCR assay specific to the transgenic lines we were working with.

Previous work using inverse PCR had successfully localised the positions of the transgenic insert in lines EE and E-Vida to chromosome 3R, position 15801959 (band 31B, S. Basu and J. Meridith personal communication). Due to a peculiarity in the 2-stage transformation process used to create these strains, the sequences within the transgenic cassette immediately adjacent to the insertion point were identical in both EE and E-Vida. Taking advantage of this, we designed primers flanking the insertion point in the unmodified genome, as well as a number of primers specific to sequences within the insert itself. We envisioned a system by where a primer targeting a region upstream and adjacent to the insertion point would act as a 'universal' forward primer with primers both specific to the region downstream of the insertion point and also within the left flank of insertion; acting as wild-type and transgenic-specific reverse primers respectively. Multiple sequences for these three target region were prototyped using Primer3 (http://primer3.sourceforge.net) and synthesized (MWG Operon). Triplet combinations of potential primers were tested for target sensitivity and specificity as well as for undesirable cross-reactivity or self-annealing using multiple transgenic and non-transgenic DNA templates and template-negative controls. Candidate PCR combinations were blind-tested against a set of 20 DNA templates homozygous EE, EVida3 and Mopti individuals, as well as hemizygous EE/Mopti and EVida3/Mopti samples. Following testing, we arrived at a system as follows: A universal forward primer (UnFwd 5' - CCA TCC CCA AAA AAA TGA ACT GAA A -3') targeting a region just upstream of the transgene insertion point in

EE and EVIDA3, combined with two reverse primers, one - transgenic specific - targeting a region on the transgene's left flank (TGRev 5'- GCA GAC TAT CTT TCT AGG GTT AAA CTG -3'); the other, wild-type specific primer binds downstream of the insertion point (5'- TCC CTC TTA TAA GTA AGG GTT GC -3'). These primers produce an amplicon of 172bp in size in the presence of a homozygous wild type individual and an amplicon of 166bp in size in the presence of a homozygous EE-derived transgenic strain (EE or EVIda3). Hemizygotes produce two amplicons. This system proved capable of correctly identifying all 20 'unknown' template sources and was used in subsequent experiments as required (see Chapters 3 and 4).

3. Egg production and larval development in transgenic *Anopheles gambiae* s.s

Abstract

Potentially one of the most powerful new approaches to malaria control, vector transgenesis, has been the focus of sustained research over the last decade which is now beginning to generate laboratory strains of transformed mosquitoes that carry exogenous genes that block parasitic development within the definitive vector host. Laterally, focus has moved from developing the methodology for efficiently creating new transformants to assessing the efficacy of anti-parasitic transgenes *in vivo* and the ability of transgenic mosquito strains to outcompete wild-type strains and persist in a field context.

In an effort to determine the physiological manifestation of potential transformation-mediated fitness effects, we conducted studies of two *Anopheles gambiae s.s*. transgenic lines recently developed using a two-phase targeted genetic transformation system. Oviposition success, egg batch size and egg morphology from primagravidae females of the EE Phase-1 docking strain and EVida3 Phase-2 strain were assessed relative to a non-transgenic laboratory strain. Furthermore the effect of growth density and growth medium quality on the development time and survival to ecdysis of larvae from these strains was also investigated.

Our results suggest there is a significant fitness load suffered by the phase 2 transgenic during egg development and/or oviposition relative to the phase 1. Furthermore both transgenic strains were more sensitive to increased larval growth density than the wild-type. This may be indicative of fitness costs within these strains and has important implications for future mass rearing and release strategies.

3.1 Introduction

In any future genetic malaria control strategy, the fitness and competitiveness - both in terms of adult mating success and the development of immature life stages - of a candidate transgenic strain is of key importance, particularly in strategies that are based around the spread of transgenic refractoriness through a target wild population (O'Brochta 2003). Released homozygous transgenic or sterile males must be able to successfully survive in the wild and compete for mates with their wild counterparts, furthermore, with a population replacement strategy - where a long term/permanent change in a population is affected through the introgression of transgenic malaria refractoriness - transgenic larvae must also be capable of surviving in the habitats typical of wild mosquitoes. Thus these larvae must be able not only to survive in a field context, but also to compete for potential limited and limiting resources with unmodified wild larvae. A discussion of the broad mechanisms and genetic basis of the impact of transgenesis on the long term fitness of transformed strains can be found in Chapter 4. Here we consider transgenesis in the context of the key ecological factors of egg production and larval development in two transgenic strains - EE and EVida3 (described previously) - notable due to their shared transgenic loci but differing transgenic cassettes.

Egg production and oviposition in Anophelines is tightly correlated with the metabolic reserves of an individual. These reserves are built up during the trophic larval stages and are themselves tightly positively correlated with adult body size. It has been demonstrated that smaller adult female *An. gambiae* require a 'pregravid' blood meal - which they do not produce an egg batch with - to build up their metabolic reserves prior to producing eggs with a second blood meal (Takken *et al* 1998). Larger adult female Anophelines have been shown to have larger metabolic reserves, take larger blood meals and are more likely to produce an egg batch following their first blood meal. Furthermore,

several studies have demonstrated a positive correlation between female body size and egg batch size.

A number of transgenic mosquito strains drive AMP transcription using regulatory sequences from genes naturally upregulated following a blood meal, for example *Anopheles gambiae carboxypeptidase* (*Agcp,* Edwards *et al* 1997) or *vitellogenin* (*vtg* Nirmala *et al* 2006). This kind of conditional expression means that AMPs are upregulated to coincide with the ingestion of *Plasmodium* gametocytes and can target the parasite at the developmental bottleneck before it undergoes major multiplicative replication in the oocyst. Furthermore conditional expression means that the AMP is only expressed in a limited number of tissues and in a narrow timeframe, limiting the potential fitness impact on the transgenic mosquito of AMP expression. Nevertheless, timing expression to coincide with a blood meal or oogenesis represents a potential source of stress - through an AMP-stimulated endogenous immune response - or reduced fitness - through AMP toxicity - which may affect female blood meal utilisation, egg production and egg morphology. For this reason, assessing these factors in transgenic mosquito populations is a logical first step in the determining the fitness impact of transgenesis.

Moreira and colleagues (Moreira *et al* 2004) assessed the blood meal ingestion and egg production of two strains of transgenic *Anopheles stephensi*, one expressing the AMP SM1, and another expressing bee-venom phospholipase 2 (PLA2). Both transgenes were under the control of the *Agcp* regulatory sequence which drives expression of a product into the midgut of a female shortly after the ingestion of a blood meal. In this study they found that, whilst the SM1-expressing transgenic strain had comparable egg production and blood meal ingestion with their non-transgenic control, the PLA2-expressing strain ingested significantly less blood and laid significantly fewer eggs than the control. In addition, even with the observed reduction in ingested blood this strain developed fewer

eggs than expected, indicating a reduction in blood meal utilisation. It was later found that the expressed PLA2 was permeabilising the midgut lumen through enzymatic activity (Abraham *et al* 2005). A more recent study by Li and colleagues (Li *et al* 2009) assessed egg-production and blood meal utilisation in transgenic *Anopheles stephensi* again expressing the AMP SM1. In this case the transgene was under the control of the *vitellogenin* regulatory sequences which drives expression into the haemocoel following a blood meal. In this study a reduction in egg production and fecundity was observed in two transgenic strains expressing this cassette despite comparable body size and blood meal ingestion levels indicating that the act of transgenesis or expression of the transgene was having a negative effect on blood meal utilisation and/or egg development.

The conditions in which larvae develop affect a number of adult traits key to the survival and persistence of a mosquito; these include: metabolic reserves and desiccation resistance (Aboagye-Antwi and Tripet 2010), female blood meal utilisation and brood size (Takken *et al* 1998) and male mating success (Ng'habi *et al* 2005). In *An. gambiae*, increasing larval density has been shown to increase development time and reduce survival to eclosure and adult body size (Schneider *et al* 2000, Gimnig *et al* 2002). In the first study, Schneider and colleagues raised *An. gambiae* and *An. arabiensis* larvae both singly and in a 1:1 mix at three larval growth densities: 100, 200 and 400 larvae/litre under insectary conditions with a set amount of food per larva. They observed a decrease in larval survival of ~30% between the lowest and highest densities in *An. gambiae* grown singly, but no change in the mean time to pupation between densities. Intriguingly, when *An. gambiae* were mixed with *An. arabiensis*, there was no reduction in *An. gambiae* larval survival between densities, but *An. arabiensis* survival was significantly reduced in mixed conditions compared to single strain controls, indicating they were at a significant competitive disadvantage to *An. gambiae* larvae despite their slower mean development

time. In another, separate study, Gimnig and colleagues performed a similar experiment investigating the effects of larval growth density on *An. gambiae* in artificial habitats in a field context, with static food availability between densities (Gimnig *et al* 2002). Here, increasing larval density reduced larval development time and adult body size, but had no effect on net survival.

In transgenic mosquitoes, transgenesis has been observed to negatively affect larval survival and increase larval development time in studies of transgenic fitness, such as Irvin *et al* (2004) who observed a marked reduction in larval survival in homozygous transgenic *Aedes aegypti* expressing EGFP and transposase genes facultatively. Although a number of groups have reported no impact on larval development from transgenesis, for example the Li study cited previously (Li *et al* 2009). Koenraadt and colleagues demonstrated that increasing intraspecific competition in larvae by reducing food availability had an extremely negative effect on transgenic *Aedes aegypti* - expressing EGFP facultatively - in terms of larval survival and development time that was not suffered by wild-type non-transgenic larvae either when raised singly or when transgenic and non-transgenic larvae were grown mixed in direct competition (Koenraadt *et al* 2010). In a similar experiment - although lacking the mixed larval trays of Koenraadt *et al* - Bargielowski and colleagues demonstrated that increasing larval growth density had a negative effect on larval development time, survival and adult body size in transgenic *Ae. aegypti* carrying a tetracycline repressible dominant lethal cassette. Although the negative effect of increasing density on larval development and survival was similar between the transgenic and non-transgenic strains, the reduction in adult body size in transgenic individuals induced by increased larval growth density was greater in the transgenic strain (Bargielowski *et al* 2011).

We set out to investigate egg production and larval development in the transgenic strains EE and EVida3 in relation to the non-transgenic wild-type Mopti strain (described previously). As the activity of the *Agcp* promoter driving AMP expression in EVida3 is upregulated immediately following a blood meal, and remains so for up to 48 hours PBM (Meredith *et al* 2011), we investigated whether AMP expression had an effect on female brood-size and the gross morphology of brood eggs. In addition we assessed the mean survival, median time to eclosure and adult body size in these strains at 3 larval growth densities in both single strain and mixed competition comparisons.

3.2 Materials and Methods

3.2.1 Strains and Rearing

In these experiments we investigated the effects of transgenesis and competition on larval development, adult egg production and egg morphology, and adult body size using the EE and EVida3 transgenic strains described previously. The wild-type strain used in this experiment was our Mopti, M-form population originally colonized from Mali in 2003. At the time of the experiment, this strain had been recently refreshed by outcrossing to the F1 progeny of field-caught gravid females. Both transgenic stocks are maintained as true-breeding homozygotes and, along with wild-type strains, are maintained in dedicated insectaries as described previously (see Chapter 2).

3.2.2 Experimental Design

With a range of experiments we investigated the effect of transgenesis on key developmental stages in our three strains of *An. gambiae*. Comparisons between Mopti and EE/EVida3 allowed us to determine the effects of transgenesis, and a comparison of the
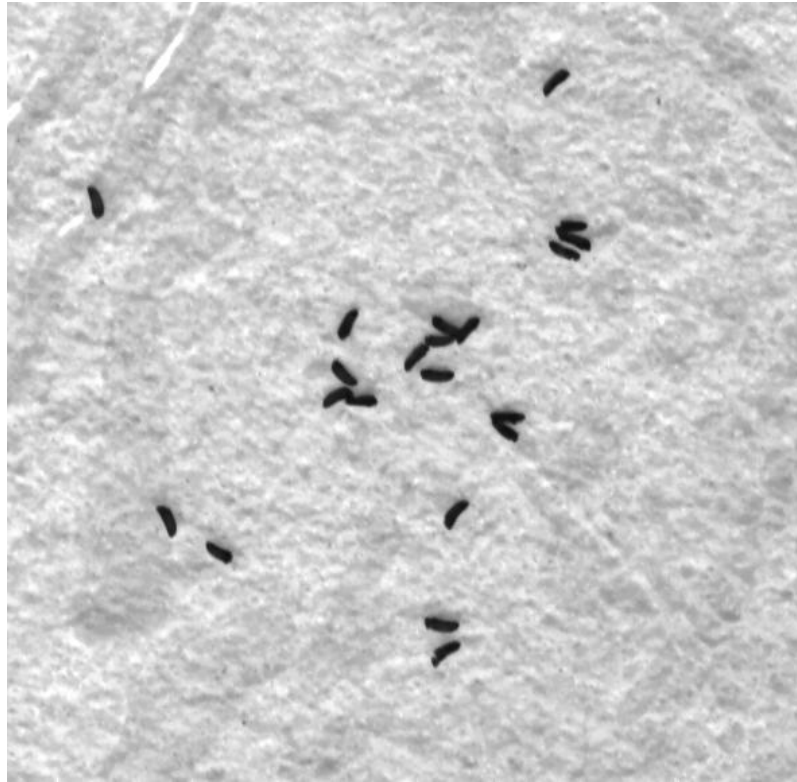
Figure 3.2.1: Section of a scanned filter paper with immobilised *An. gambiae* eggs captured at 1200ppi.

relative performance of EE and EVida3 isolated the effects of carrying and expressing the Vida3 tetramer from any broader impact of transgenesis.

3.2.3 Egg morphology and brood size

48 hours after a blood meal from a human volunteer (DP), 120 gravid females from each mosquito population were selected at random and transferred by aspiration to single female oviposition tubes as follows: females from each strain were transferred by aspiration to 8 cm x 2 cm cylindrical glass tubes, sealed with a small square of mosquito netting. After approximately 2 hours, 5 ml ddH_2O was carefully transferred into the tube by pipetting the fluid slowly down the side of the tube so as to minimise the stress to the mosquito and chance of inundation. These individual oviposition tubes were transferred to a rack and left to allow the mosquitoes to oviposit. After 48 hours and every 24 hours thereafter, tubes were checked for eggs or the death of the individual until all tubes were accounted for. After 72 hours, a random sub-sample of 32 oviposition tubes from each strain was taken. Females were recaptured by aspiration, and knocked down by transfer to a 1.5 ml microcentrifuge tube containing 500 µl of 70% ethanol and stored at -20 °C. Egg broods were isolated by gently vacuum filtering the oviposition tube contents through filter paper (5 cm diameter, type I, Whatmann), taking care to thoroughly rinse the tube to ensure all eggs were transferred. Once the filtration was complete, egg papers was sealed between two layers of cellophane and scanned using an HP ScanJet 5370. Images were captured in 8-bit greyscale at a resolution 1200 pixels per inch. Adult females were scored for oviposition status and wing length. Broods were scored for egg number and the presence or absence of hatching. Egg morphology was investigated by digital image analysis of a subset of eggs from each brood scan using the ImageJ digital analysis suite.

3.2.4 Density-dependent effects on larval development time and survival

In this experiment, we investigated the effects of larval population density on development time (time to eclosure) and survival. In two independent experiments, we compared Mopti to homozygous EE transgenics (exp. 1); and Mopti to homozygous EVida3 transgenics (exp.2) as follows: Newly hatched L1 larvae were separated into 3 larval growth density cohorts - 100 larvae/l, 200 larvae/l and 400 larvae/l - and two groups - a control group consisting of a single strain, and a competition group consisting of a 1:1 mix of wild-type (Mopti) and transgenic (exp.1: EE, exp.2: EVida3) larvae. Larvae were raised in standard growth conditions as described previously. Growth tray position was rotated over the course of larval development to control for any environmental variation within the insectary space. Pupae were transferred by aspiration to small polystyrene eclosure pots on a daily basis. Newly emerged adults were captured by aspiration and knocked down with an aerosol of 70% ethanol and stored at -20 °C. Capture of newly emerged adults was performed at the same time each day to ensure that each sample was representative of exactly one 24 hour period. Both the control and competition group adults were scored for emergence day, gender and wing length. In addition, all individuals surviving to eclosure from the competition group had their genotype determined by performing our WT/EE-based transgenic discriminative PCR (described previously) on extracted DNA.

3.2.5 Dose-dependent effects of used growth medium on larval development

We investigated whether there was any evidence of the accumulation of toxic substances in the larval growth media of transgenic strains. We hypothesised that media previously used to raise larvae would have an impact on survival and development time which was independent of the genotype of larvae being raised in it. To test this hypothesis, we devised an experiment in which larvae were raised in media harvested from mature larval growth

trays. Combining larvae from a given strain with media from other strains allowed us to determine the relative toxicity of these media.

Prior to the launch of this experiment, 9-12 larval trays per strain containing 400 larvae in 1l of ddH_2O were set up for each of our three strains (Mopti, EE, EVida3). With the exception of the larval density, trays were grown in standard conditions for 12 days. After this period, the contents of the trays were run through a course, 2 mm, nylon mesh filter. Large food particles, larvae, pupae and course detritus were discarded. The 'used' growth medium was retained and stored at 4 °C for later use.

For the experiment proper, larvae from each strain were segregated into one of three treatment groups: a control group grown in ddH_2O (control), a group grown in a 50% v/v mix of used medium and ddH_2O (Test 50%) and a group grown in 100% used medium (Test 100%). Within the test groups, larvae were grown in their 'self' medium (ie Mopti larvae raised in Mopti medium, EE larvae in EE medium etc). In addition both transgenic strains were raised in Mopti medium, and Mopti larvae were raised in both the EE and EVida3 medium. Growth density in all trays was 400 larvae/l but standard rearing protocols were otherwise used. As in the previous experiment, pupae were transferred by aspiration daily to a small eclosure pot and newly emerged adults were captured and stored in 70% ethanol at -20 °C. All individuals surviving to eclosure were scored for emergence day, gender and wing length.

3.2.6 Data Analysis

Egg morphology and brood size

Initially, gross variation in total brood size between genotypes was assessed by ANOVA with a Tukey post-test. We investigated whether there was a correlation between brood size and wing length by constructing an X/Y scatter and regression analysis of these

variables. Comparative oviposition success was assessed by pairwise Chi-square. Finally, multivariate models were constructed testing the effects of the independent variables of genotype and wing length, and their interaction, on the dependent variables of oviposition success (Logistic Regression [LR]) and brood size (General Linear Model [GLM]).

To examine egg morphology in detail, up to 23 egg silhouettes were sampled from each egg brood (example image figure 3.2.1) and analysed using the Analyse Particle function of the ImageJ image analysis software. This algorithm is capable of returning up to 20 variables per egg. In our first pass to eliminate non-discriminative characteristics, variables based on grey value (i.e. colour), angles relative to the image's x-axis, and the area of a selection of various bounding polygons were eliminated as not being of physiological importance. This left us with four candidate variables which were as follows: Area - defined as the total number of pixels within the identified egg silhouette; Perimeter - the length of the outside boundary of the silhouette; Feret's Diameter - the longest distance between any two points within the designated particle - effectively equivalent to the length at the long axis of the egg; and Circularity: $4\pi(\text{area/perimeter}^2)$, this variable describes the degree of similarity of the analysed shape to a circle - a perfectly circular egg would have a Circularity of 1, and would tend towards 0 in progressively more elongated shapes.

GLM models were constructed for each variable to determine the relationship between, genotypes and individuals within genotypes for each of our 4 candidate variables. As we made no *a priori* assumptions about the relative discriminative ability, importance of, and relationship between individual variables we constructed a covariance matrix to assess the relationship between variables. Furthermore, to avoid a complex higher dimensional statistical analysis, egg parameters were fed into a principal component analysis. This allowed us to identify redundant co-variants and collapse multiple variables into a smaller number of orthogonal (i.e. uncorrelated) principal components. Once these

had been determined, GLM models were constructed to assess the relative impact of variation within and between genotypes on each identified principal component. Finally, the ability of the variables, including the principal components to successfully discriminate between individuals of different genotypes was assessed using a multivariate discriminative analysis.

Density-dependent effects on larval development time and survival

Differences in mean body size were assessed by taking a random sub-sample of 30 wing lengths from each growth density and genotype cohort. Variance in mean wing length between groups was analysed by ANOVA and *post hoc* Tukey tests.

The factors affecting survival (survival to eclosure) and development time (time to eclosure) were assessed by constructing a linear regression model and proportional hazards survival model respectively. Model effects were constructed from the following independent variables: larval growth density, genotype, replicate and treatment group. These analyses were also broken down by larval growth density. Differences in survival between genotypes, larval growth densities and the control and competition groups were tested by pairwise Chi-square. Differences in development time between genotypes, growth densities and the control and competition groups were assessed using pairwise survival analysis.

Dose-dependent effects of used growth medium on larval development time and survival

The analysis of the factors affecting survival and development time was carried out in the same manner as in our density dependent experiments (above). Multivariate linear regression models were constructed factoring the effects of genotype, replicate and media source and concentration on survival. Analyses nested by comparison (i.e. Mopti x EE,

Mopti x EVida3) were also carried out. Factors affecting time to eclosure were assessed in the same manner using a proportional hazards survival model. Again, the direction and nature of differences in development time and survival between media and genotypes were determined by pairwise survival and Chi-square analyses respectively.

3.3 Results

3.3.1 Egg morphology and brood size

Brood size was normally distributed within all genotypes. Overall, homozygous EE females successfully oviposited most frequently (25 broods from 32 oviposition tubes sampled) and had the largest average broods (mean 97.64 eggs per brood 95% CI ±12.9). Homozygous E-Vida3 and Mopti females both laid successfully in 14 and 15 of 32 oviposition tubes and had mean brood sizes of 64.33±16.46 and 59.43±19.17 respectively (Figure 3.3.1). The difference in both oviposition success and brood size between EE and the remaining genotypes was significant (Brood Size: ANOVA n=53, df=2, F=8.7, $P<0.001$, Tukey, $P=0.002$, Oviposition Success: Chi-square, n=96, df=2, $\chi^2=10.7$, $P=0.005$). Wing length within genotypes was not normally distributed. EE females had the highest median wing length (3.05mm, interquartile range 2.96-3.08mm) followed by Mopti females (2.84mm, interquartile range 2.74-3.02mm) and EVida3 (2.77m, interquartile range 2.74-2.94mm) females. The difference in median wing length between EE and the remaining genotypes was significant (Wilcoxon, n=80, df=2, Z<-3.87, $P<0.001$), the difference in median wing length between Mopti and Evida3 was not significant. There was no correlation between wing length and brood size within any genotype cohort (Linear Regression: EE, n=24, F=0.037, $P=0.850$, Mopti, n=13, $F=0.031$, $P=0.864$, EVida3, n=10, F=2.053, $P=0.189$). Finally, the factors affecting the dependent variables of brood size and oviposition success were investigated by constructing linear regression (LR) and general
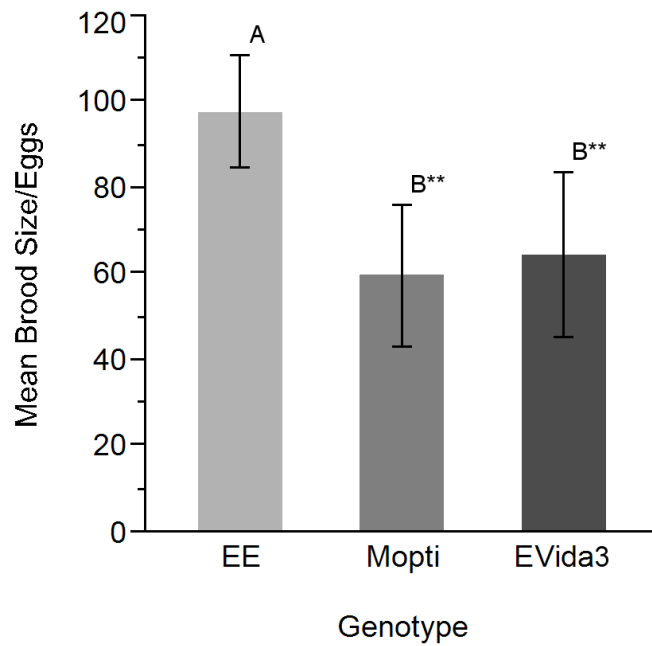
Figure 3.3.1: Mean egg brood size between homozygous female non transgenic (Mopti) and transgenic (EE, EVida3) females. Error bars represent 95% confidence intervals, the results of a Tukey *post hoc* test are indicated - ns, not significant, *: P> 0.05, ** P> 0.01, *** P> 0.001.

linear (GLM) models respectively. Model effects consisted of wing length, genotype and individual mosquito nested within genotype (accounting for variation between individuals within genotypes). Genotype was the only significant determinant of both brood size (GLM: n=47, df=2, F=7.5, P=0.002) and oviposition success (LR: n=80, df=2, χ^2=6.907, P=0.032).

Having identified four candidate variables for describing comparative egg morphology between genotypes (Area, Perimeter, Feret's Diameter and Circularity, see methods), we investigated the relationship between each variable and our three genotypes using a GLM. Model effects were genotype and individual mosquito nested within genotype (accounting for variation between individuals within genotypes) (Figure 3.3.2). There was a significant difference in mean egg Area between genotypes (GLM: n=1213, df=2, F=196.2, P<0.001). A Tukey post-test revealed that all genotypes were significantly different from each other. In terms of Perimeter, there was again a significant difference between genotypes (GLM: n=1213, df=2, F=66.41, P<0.001). In this case the EE genotype had significantly lower egg perimeter than either Mopti or EVida3, but there was no significant difference between Mopti and EVida3 eggs. Feret's Diameter followed exactly the same pattern as Perimeter, with EE eggs having a significantly lower Diameter than either EVida3 or Mopti, and no difference between these groups (GLM: n=1213, df=2, F=41.85, P<0.001). Finally, when we considered Circularity, we found that there was again a significant difference between genotypes (GLM: n=1213, df=2, F=128.03, P<0.001). In the case EVida3 eggs were, on average, significantly more round/closer to circular than either EE or Mopti eggs. There was no significant difference in mean circularity between Mopti and EE eggs. There was significant variation between individuals within each genotype in all analyses (GLM: n=1213, df=55, F>6.409, P<0.001).
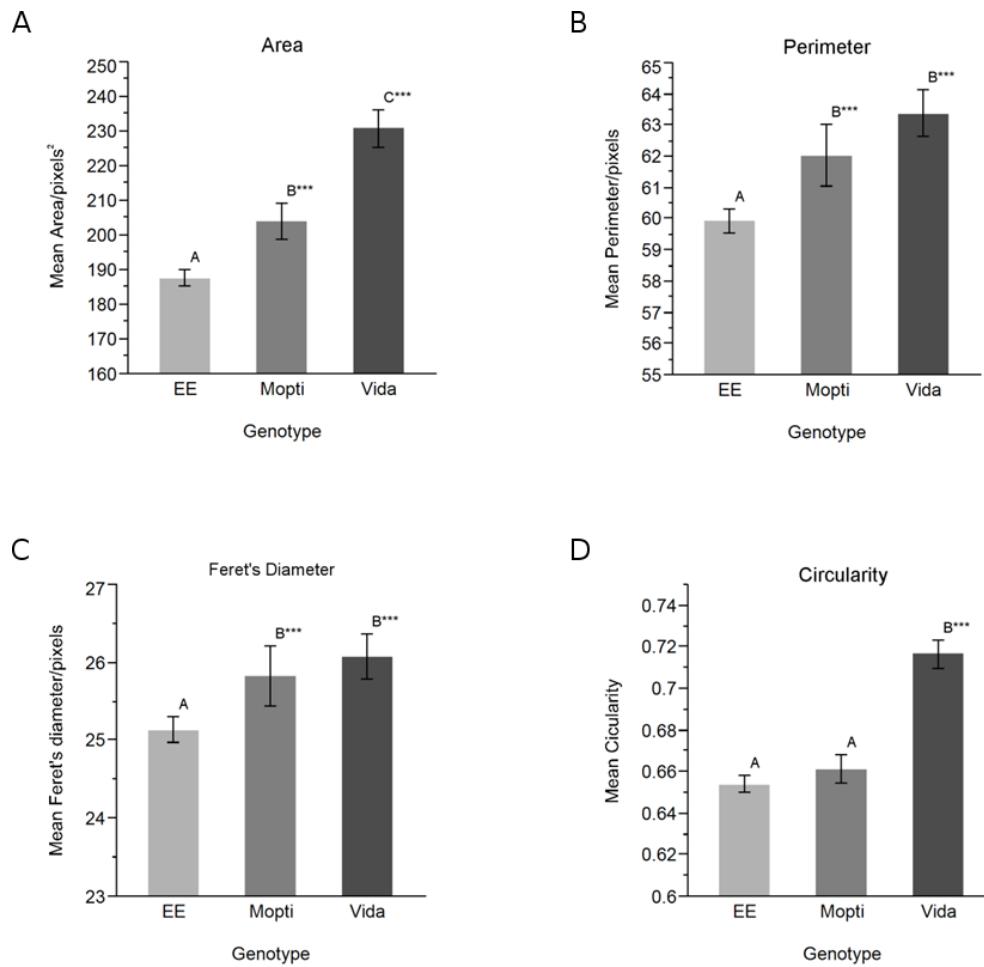
Figure 3.3.2: Mean values for four potentially discriminative morphological characteristics between eggs from three strains. A) Area in pixels^2. B) Perimeter in pixels. C) Feret's Diameter in pixels. D) Circularity (no units). The results of a GLM analysis for the effect 'Genotype' are indicated - ns, not significant, *: P> 0.05, ** P> 0.01, *** P> 0.001.

Covariance and correlation between our four candidate variables was investigated by constructing a correlation matrix (table 3.3.1) and scatterplot matrix (figure 3.3.3). Area, Perimeter and Feret's Diameter were all strongly positively correlated with one-another (co-efficient of correlation 0.7673 - 0.9476). Circularity correlated positively but weakly with Area (CoC: 0.286) and negatively and again weakly with Perimeter and Feret's Diameter (CoC: -0.1878 - -0.3055). When these co-varying morphological characteristics were included in a multivariate Principal Component Analysis, the four variables collapsed into two new orthogonal principal components (Table 3.3.2): Principal Component 1 (PC1, Eigenvalue: 2.74) comprised of the variables Area, Perimeter and Feret's Diameter and accounted for 68.49% of all variation within genotypes. Principal Component 2 (PC2, Eigenvalue: 1.20) comprised of the variables Area and Circularity and accounted for 29.81% of variation within genotypes. In total PC1 and PC2 accounted for 98.30% of all variation in our sample and were, by definition, completely uncorrelated. We constructed two GLM analyses using our principal components as the dependent variables. Effects comprised of, as before genotype, and individual mosquito nested within genotype. Once variation between individuals within genotypes had been taken into account, we found there was a significant difference between genotypes in terms of PC1 (GLM: n=1213, df=2, F=10.51, P<0.001) but not in terms of PC2 (GLM: n=1213, df=2, F=0.26, P=0.771). A Tukey post test revealed that EE had a significantly lower PC1 score than either the EVida3 or Mopti cohort, and there was no significant difference between the EVida3 and Mopti.

Finally, we built the four morphological characteristics and our principal components into a multivariate discriminative analysis. The highest proportion of correctly

	Area	Perimeter	Circularity	Feret's Diameter
Area	1	**0.8766**	0.286	**0.7673**
Perimeter	**0.8766**	1	-0.1878	**0.9476**
Circularity	0.286	-0.1878	1	-0.3055
Feret's Diameter	**0.7673**	**0.9476**	-0.3055	1

Table 3.3.1: Correlation matrix between four potentially discriminative morphological characteristics. Co-efficients of correlation greater than 0.750 are highlighted in **bold**.
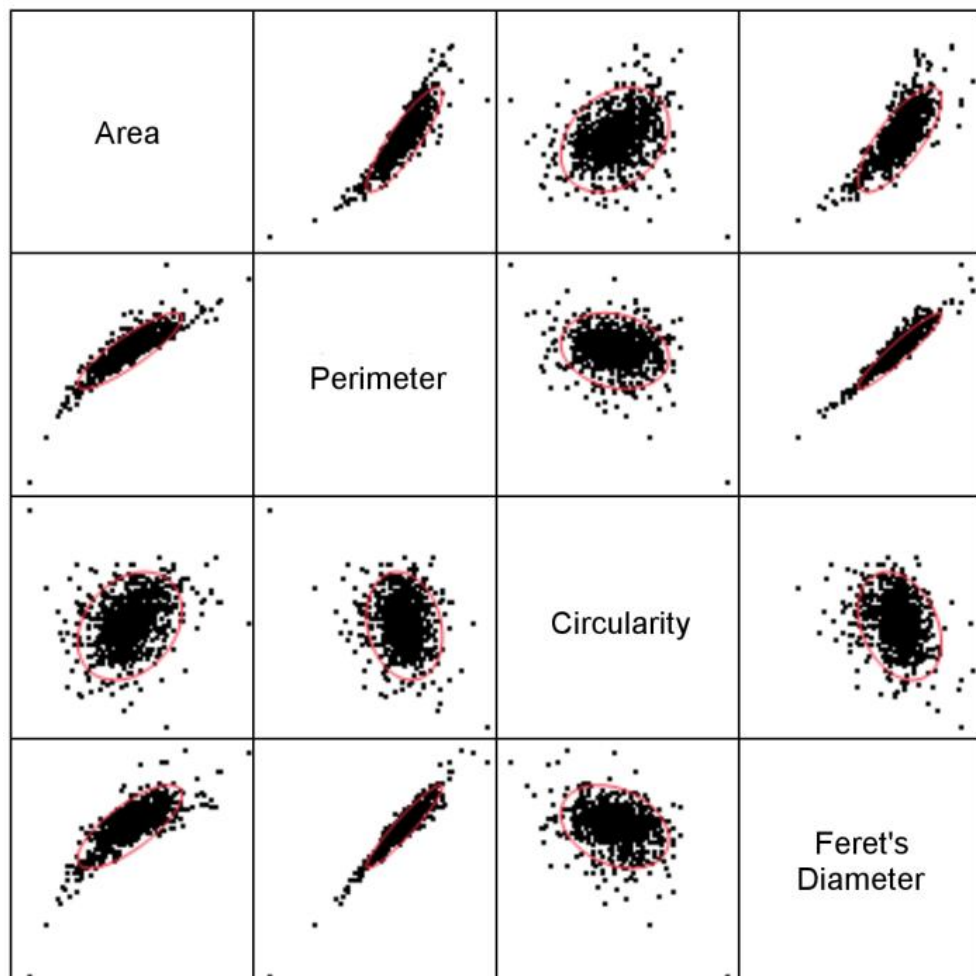
Figure 3.3.3 Covariance scatterplot matrix of four potentially discriminative morphological characteristics.

identified egg genotypes we were able to achieve with any combination of covariates was 40%, thus these features were not discriminative.

3.3.2 Density-dependent effects on larval development time and survival

In our first larval competition experiment we compared the development time - defined as the number of days between egg hatching and eclosure – and survival – defined as the proportion of larvae surviving to adulthood – of wild-type Mopti and transgenic EE individuals in both single-strain and mixed growth conditions at varying larval growth densities.

Median time to eclosure in experiment 1 was as follows: in the Control (single strain) Mopti group (figure 3.3.4), larvae grown at a density of 100 larvae/litre had a median time to eclosure of 9 days, (interquartile range [IQR] 8-9 days), at the 200 larvae/litre density median time to eclosure was also 9 days (IQR 8-10 days). Median time to eclosure in Mopti larvae at 400 larvae/litre was 11 days (IQR 11-12 days). There was a significant difference in time to eclosure between all growth densities (Chi-square: df=2, χ^2=553.95, *P*<0.001), pairwise comparisons of time to eclosure between densities in the Mopti control were also all significant (Chi-square: df=1, χ^2>7.23, *P*<0.007). In the EE control (figure 3.3.4), median time to eclosure was 8 days at 100 larvae/litre (IQR 8-9 days), 9 days at 200 larvae/litre (IQR 8-9 days) and 11 days at 400 larvae/litre (IQR 9-12 days). Overall there was a significant difference in time to eclosure between densities (Chi-square: df=2, χ^2=399.53, *P*<0.001), although pairwise comparison revealed that the difference in time to eclosure between EE larvae grown at 100 and 200 larvae/litre was not significant (Chi-square: df=1, χ^2=3.29, *P*=0.070). When we compared time to eclosure between the EE and Mopti genotypes at each larval growth density (figure 3.3.5), we found

A

PC No.	Eigenvalue	% Variation	χ^2	df	*P*
1	2.7396	68.491	7704.64	4.808	<.0001
2	1.1923	29.808	5771.61	5.061	<.0001
3	0.0591	1.478	949.645	2.068	<.0001
4	0.0089	0.223	0 -		-

B

Variable	PC1	PC2	PC3	PC4
Area	**0.90291**	**0.41453**	-0.0967	-0.05977
Perimeter	**0.99221**	-0.0573	-0.08777	0.06732
Circularity	-0.12749	**0.98871**	0.07398	0.02699
Feret's diameter	**0.96108**	-0.19914	0.19127	-0.00977

Table 3.3.2: Results of a principal component analysis of four potentially discriminative morphological characteristics. A) Eigenvalues. B) Loading Matrix.

that EE larvae had a significantly lower median time to eclosure than Mopti larvae at all densities (Chi-square: df=1, χ^2>6.67, P<0.001).

In the competition comparison (figure 3.3.4) median time to eclosure in both Mopti and EE larvae was 9 days at 100 larvae/litre (IQ range 8-9 days), 9 days at 200 larvae/litre (IQ range 8-9 days) and 11 days at 400 larvae/litre (IQ range 9-12 days). There was no significant difference in time to eclosure between genotypes (Chi-square: df=1, χ^2<0.2843, P>0.395) (figure 3.3.5). In both Mopti and EE there was no significant difference in time to eclosure between the 100 and 200 larvae/litre densities (Chi-square: df=1, χ^2<2.39, P>0.122), all other pairwise density comparisons were significant (Chi-square: df=1, χ^2>138.25, P<0.001).

In terms of survival, in the Mopti control (figure 3.3.6), the proportion of larvae surviving to eclosure (Mean ±95% CIs) was 67±5.4% at 100 larvae/litre, 71.8±3.6% at 200 larvae/litre and 67.3±2.7% at 400 larvae/litre. There was no significant difference in survival between densities (Chi-square: df=2, χ^2=4.32, P=0.115). In the EE control, mean survival was 74.3±5% at 100 larvae/litre, 57±4% at 200 larvae/litre and 48.7±2.9% at 400 larvae/litre. In this case the difference in survival between growth densities was significant (Chi-square: df=2, χ^2=65.58, P<0.001). All pairwise comparisons were likewise significant. When we compared survival between genotypes at each growth density (figure 3.3.7), we found that in the controls, significantly more Mopti larvae survived to eclosure than EE at all densities (Chi-square: df=1, χ^2>3.89, P<0.048).

In the competition comparison, survival in Mopti larvae was 65.3±7.7 at 100 larvae/litre, 68±5.3% at 200 larvae/litre and 63.5±3.8% at 400 larvae/litre. EE survival was 73.3±7.2% at 100 larvae/litre, 50.7±5.7% at 200 larvae/litre and 45.7±4% at 400 larvae/litre. Between densities, survival followed a similar pattern as the controls, with no difference in survival between densities in the Mopti (Chi-square: df=2, χ^2=1.83, P=0.409)
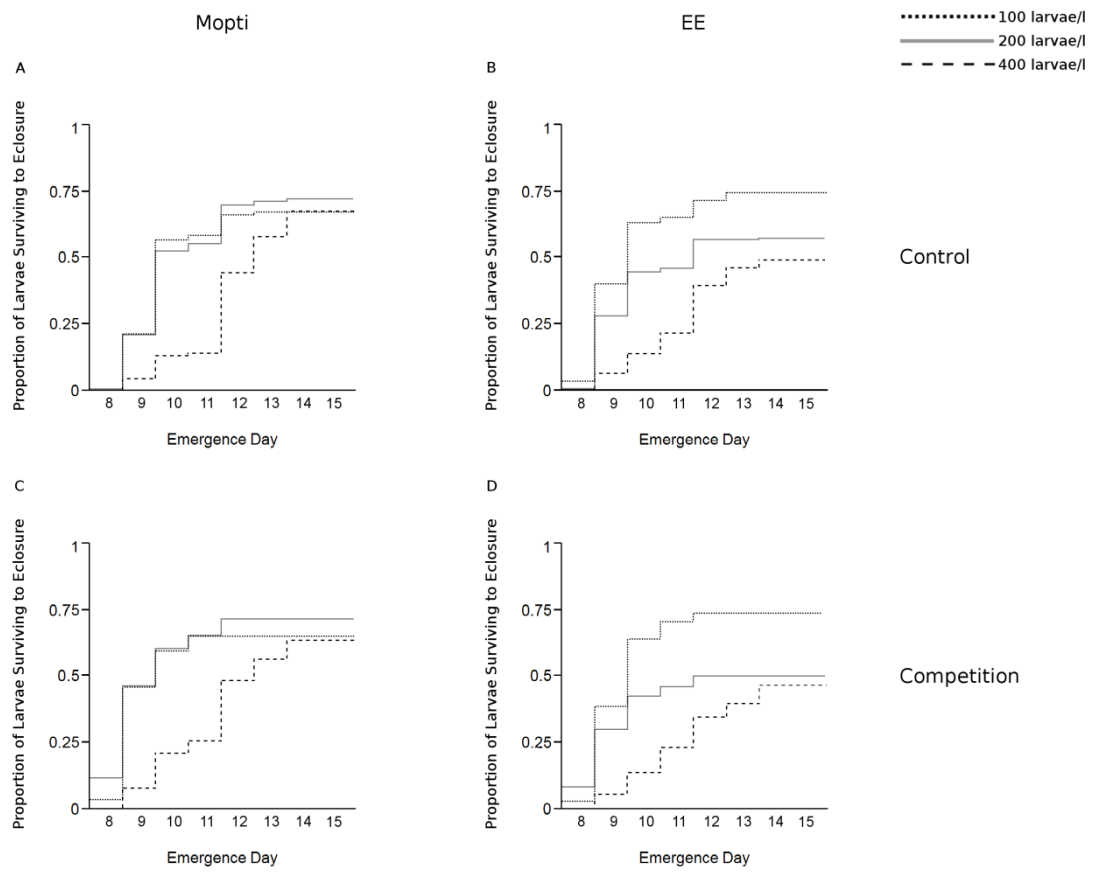
Figure 3.3.4: Cumulative frequency of larvae surviving to eclosure by genotype and comparison in Experiment 1 (Mopti x EE). Plots for the Mopti (A and C) and EE (B and D) strains for both single strain (A and B) and competition comparisons (C and D). Cumulative frequency plots for the larval growth densities of 100, 200 and 400 larvae/litre are indicated. Endpoints represent total proportion of larvae surviving to eclosure in each comparison.

Figure 3.3.5: Cumulative frequency of larvae surviving to eclosure by growth density and comparison in Experiment 1 (Mopti x EE). Plots for the single strain control (A, C and E) and competition comparisons (B, D and F) at 100 larvae/litre (A and B), 200 larvae/litre (C and D) and 400 larvae/litre (E and F). Separate cumulative frequency plots for each genotype (Mopti, EE) are indicated. Endpoints represent total proportion of larvae surviving to eclosure in each comparison.

Overall

Model Effect	df	χ^2	P
Replicate	2	671.38	<0.001
Genotype	1	23.84	<0.001
Comparison	1	14.58	<0.001
Growth Density	2	1688.56	<0.001
Genotype*Comparison	1	23.08	<0.001
			n=3807

100 larvae/litre

Model Effect	df	χ^2	P
Replicate	2	418.599	<0.001
Genotype	1	4.87	0.027
Comparison	1	6.48	0.011
Genotype*Comparison	1	15.31	<0.001
			n=632

200 larvae/litre

Model Effect	df	χ^2	P
Replicate	2	510.72	<0.001
Genotype	1	24.65	<0.001
Comparison	1	4.09	0.045
Genotype*Comparison	1	15.14	<0.001
			n=1129

400 larvae/litre

Model Effect	df	χ^2	P
Replicate	2	150.01	<0.001
Genotype	1	6.05	0.014
Comparison	1	7.46	0.006
Genotype*Comparison	1	5.40	0.020
			n=2046

Table 3.3.3: Proportional Hazards Survival model effects for factors affecting time to eclosure in Experiment 1 (Mopti x EE).

Overall

Model Effect	df	χ^2	P
Replicate	2	361.22	<0.001
Genotype	1	33.83	<0.001
Comparison	1	8.21	<0.001
Growth Density	1	63.04	<0.001
Growth Density*Genotype	2	52.45	<0.001
			n=6300

100 larvae/litre

Model Effect	df	χ^2	P
Replicate	2	59.71	<0.001
Genotype	1	6.56	0.011
Comparison	1	0.18	0.670
			n=900

200 larvae/litre

Model Effect	df	χ^2	P
Replicate	2	103.2	<0.001
Genotype	1	50.54	<0.001
Comparison	1	4.79	0.025
			n=1800

400 larvae/litre

Model Effect	df	χ^2	P
Replicate	2	245.73	<0.001
Genotype	1	133.21	<0.001
Comparison	1	4.11	0.042
			n=3600

Table 3.3.4: Logistic Regression model effects for factors affecting survival to eclosure in Experiment 1 (Mopti x EE).
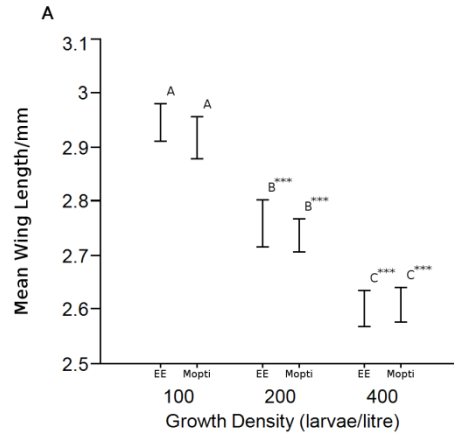
Figure 3.3.6: Mean wing length between genotypes and growth densities in a random sample of 30 individuals per genotype per density in Experiment 1 (Mopti x EE).. Error bars represent 95% confidence intervals. The results of an ANOVA and Tukey *post hoc* test are indicated as follows - ns, not significant, *: P> 0.05, ** P> 0.01, *** P> 0.001.

whilst in the EE there was an overall significant difference in survival between densities (Chi-square: df=2, χ^2=38.07, P<0.001), although the difference in EE survival between the 200 larvae/l and 400 larvae/l groups was not significant (Chi-square: df=1, χ^2=2.00, P=0.57). Mopti larvae survived significantly better than EE larvae at both the 200 larvae/litre and 400 larvae/litre growth densities (Chi-square: df=1, χ^2>18.79, P<0.001). There was no significant difference between Mopti and EE survival in the competition groups at the 100 larvae/litre growth density (Chi-square: df=1, χ^2=2.26, P=0.132).

Finally we compared time to eclosure and survival between the competition (mixed genotype) and control (single genotype) groups. Here we found there was no significant difference in survival in either Mopti or EE at any density between the control and competition groups (Chi-square: df=1, χ^2<3.23, P>0.07), however Mopti individuals at all densities had a significantly shorter time to eclosure when mixed with EE larvae compared to Mopti in the unmixed control group (Chi-square: df=1, χ^2>7.59, P<0.006). Time to eclosure in EE was unaffected by competition (Chi-square: df=1, χ^2<0.53, P>0.474).

The factors effecting both development time and survival were assessed by constructing a proportional hazards survival (PHS) test and logistic regression (LR) respectively. In both cases model effects were constructed from the independent variables of Replicate, Growth Density (100, 200 or 400 larvae/litre), Genotype (Mopti or EE) and Comparison (single or mixed strains), as well as interactions between these factors. In terms of development time (table 3.3.3), once the (significant) effect of Replicate had been partitioned, Growth Density, Genotype and Comparison were all identified as significant factors. There was also a significant interaction between Genotype and Comparison. As Growth Density had such a strong effect (PHS: n=3807, df=2, χ^2=1688.56, P<0.001), we also broke down the analysis by density, although Genotype, Comparison and the interaction of the two were still identified as significant factors in time to eclosure at all

densities (table 3.3.3). In terms of survival, all model effects were again significant (3.3.4). There was a significant interaction between Genotype and Growth Density. Again, the analysis was broken down by density and all factors were preserved as significant with the exception of Comparison at 100 larvae/litre, there were no significant interaction detected (table 3.3.4).

Random samples of the wing lengths of 30 individuals from each growth density, genotype and both the control and competition groups were compared (figure 3.3.8). There was no significant difference in mean wing length between genotypes or between the control and competition groups. There was however a significant difference in mean wing length between growth densities, with adults raised at a density of 100 larvae/litre having a significantly larger mean wing length than adults raised at 200 larvae/litre. Similarly, adults raised at 200 larvae/litre had in turn a significantly larger mean wing length than those raised at 400 larvae/litre.

In our second experiment, we compared development time and survival between wild-type Mopti and transgenic EVida3. As with the first comparison, larvae were grown in both single strain control groups and in a mixed competition group at three larval growth densities.

Median time to eclosure in the experiment 2 (figure 3.3.9) Mopti control was 9 days (IQR, 8-10) at 100 larvae/litre, 10 days (IQR, 8-11 days) at 200 larvae/litre and 12 days (11-12 days) at 400 larvae/litre. The difference in time to eclosure between all growth densities was significant (Chi-square: df=2, χ^2=461.84, P<0.001), pairwise comparisons of time to eclosure between densities in the Mopti control were also all significant (Chi-square: df=1, χ^2>95.75, P<0.001). In the EVida3 single strain control, median time to eclosure was 8 days (IQR 7-9 days), 10 days (IQR 8-11 days) and 11 days (IQR, 9-12 days) at the 100, 200 and 400 larvae/litre growth densities respectively. Overall variation
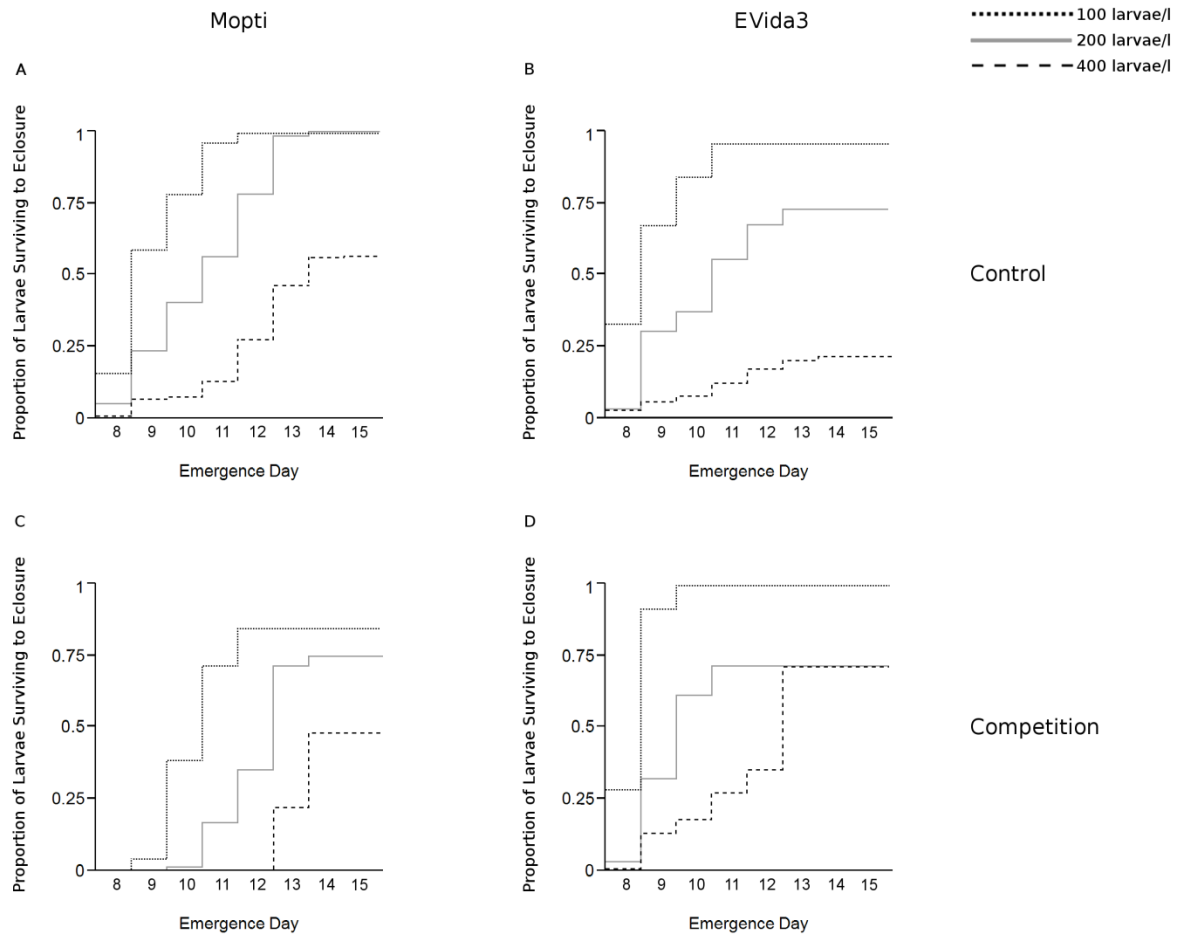
Figure 3.3.7: Cumulative frequency of larvae surviving to eclosure by genotype and comparison in Experiment 2 (Mopti x EVida3). Plots for the Mopti (A and C) and EVida3 (B and D) strains for both single strain (A and B) and competition comparisons (C and D). Cumulative frequency plots for the larval growth densities of 100, 200 and 400 larvae/litre are indicated. Endpoints represent total proportion of larvae surviving to eclosure in each comparison.
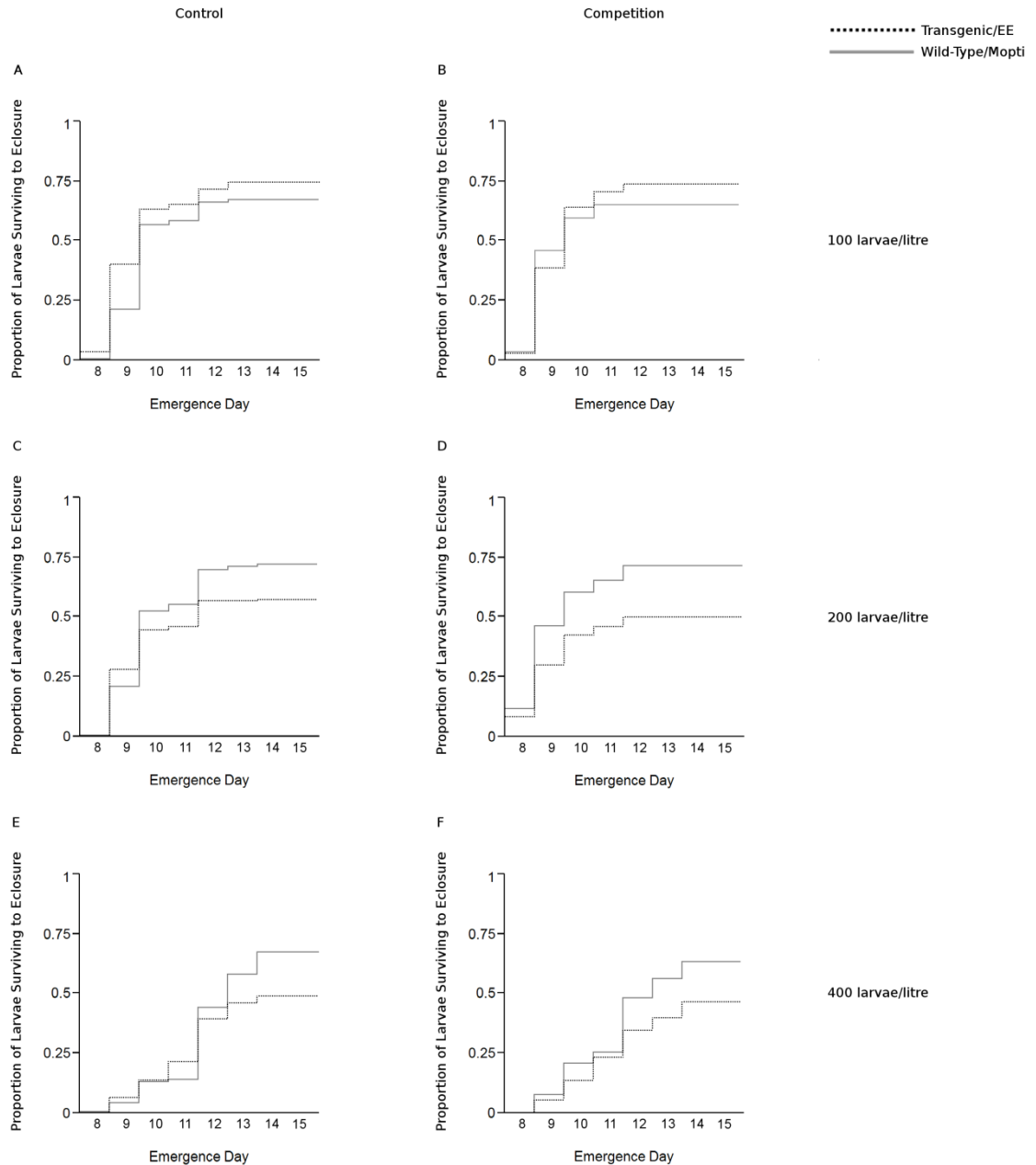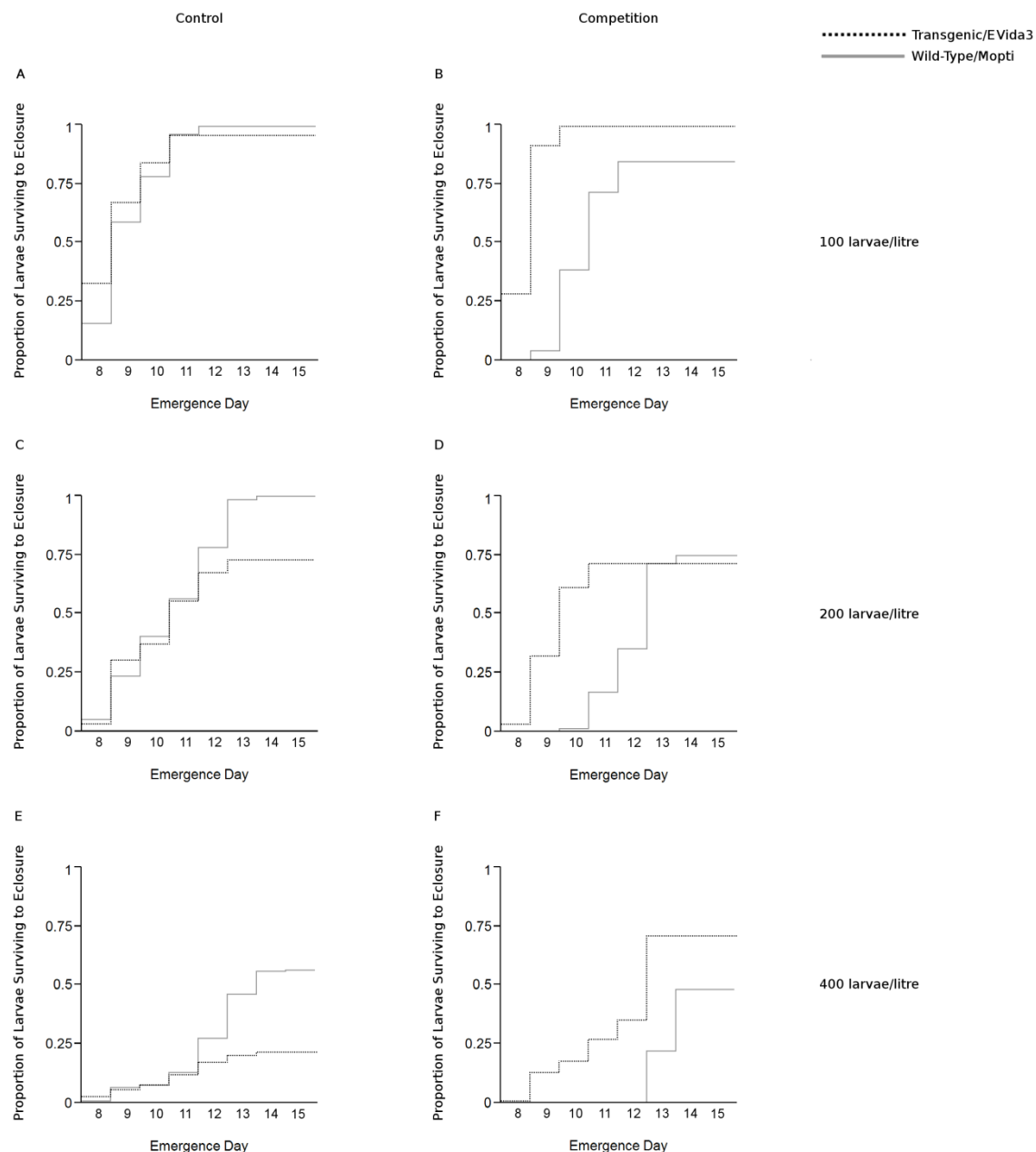
Figure 3.3.8: Cumulative frequency of larvae surviving to eclosure by growth density and comparison in Experiment 2 (Mopti x EVida3). Plots for the single strain control (A, C and E) and competition comparisons (B, D and F) at 100 larvae/litre (A and B), 200 larvae/litre (C and D) and 400 larvae/litre (E and F). Separate cumulative frequency plots for each genotype (Mopti, EVida3) are indicated. Endpoints represent total proportion of larvae surviving to eclosure in each comparison.

Overall

Model Effect	df	χ^2	P
Replicate	1	10.66	0.001
Genotype	1	596.09	<0.001
Comparison	1	43.08	<0.001
Growth Density	2	1245.11	<0.001
Genotype*Comparison	1	144.61	<0.001
Density*Comparison	2	12.82	0.002
			n=2713

100 larvae/litre

Model Effect	df	χ^2	P
Replicate	1	0.012	0.912
Genotype	1	158.11	<0.001
Comparison	1	7.31	0.007
Genotype*Comparison	1	20.33	<0.001
			n=629

200 larvae/litre

Model Effect	df	χ^2	P
Replicate	1	7.84	0.005
Genotype	1	225.22	<0.001
Comparison	1	9.64	0.002
Genotype*Comparison	1	96.67	<0.001
			n=990

400 larvae/litre

Model Effect	df	χ^2	P
Replicate	1	6.44	0.011
Genotype	1	207.61	<0.001
Comparison	1	45.74	<0.001
Genotype*Comparison	1	44.34	<0.001
			n=1094

Table 3.3.5: Proportional Hazards Survival model effects for factors affecting time to eclosure in Experiment 2 (Mopti x EVida3).

Overall

Model Effect	df	χ^2	P
Replicate	1	43.75	<0.001
Genotype	1	14.06	<0.001
Comparison	1	0.374	0.54
Growth Density	2	719.09	<0.001
Genotype*Comparison	1	238.14	<0.001
Genotype*Density	2	15.67	<0.001
Density*Comparison	2	111.67	<0.001
			n=4200

100 larvae/litre

Model Effect	df	χ^2	P
Replicate	1	4.52	0.035
Genotype	1	0	1
Comparison	1	0	1
Genotype*Comparison	1	33.16	<0.001
			n=600

200 larvae/litre

Model Effect	df	χ^2	P
Replicate	1	32.10	<0.001
Genotype	1	42.59	<0.001
Comparison	1	33.12	<0.001
Genotype*Comparison	1	28.46	<0.001
			n=1200

400 larvae/litre

Model Effect	df	χ^2	P
Replicate	1	16.51	0.001
Genotype	1	10.14	<0.001
Comparison	1	104.50	<0.001
Genotype*Comparison	1	193.48	<0.001
			n=2400

Table 3.3.6: Logistic Regression model effects for factors affecting survival to eclosure in Experiment 2 (Mopti x EVida3).

B
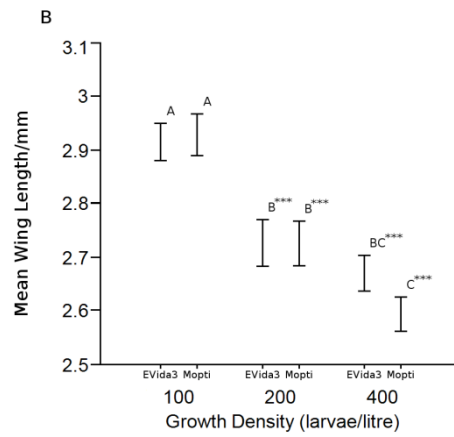
Figure 3.3.9: Mean wing length between genotypes and growth densities in a random sample of 30 individuals per genotype per density in Experiment 2 (Mopti x EVida3). Error bars represent 95% confidence intervals. The results of an ANOVA and Tukey *post hoc* test are indicated as follows - ns, not significant, *: $P > 0.05$, ** $P > 0.01$, *** $P > 0.001$.

in median time to eclosure between growth densities was significant (Chi-square: df=2, χ^2=260.50, P<0.001) as were all pairwise comparisons (Chi-square: df=1, χ^2>52.04, P<0.001). Comparing time to eclosure between the EVida3 and Mopti genotypes in the single genotype control at each larval growth density (figure 3.3.8), we found that EVida3 larvae had a significantly lower median time to eclosure than Mopti larvae at all densities (Chi-square: df=1, χ^2>8.16, P<0.001).

In the competition cohort (figure 8), median time to eclosure for Mopti individuals was 10 days (IQR, 9-10) at 100 larvae/litre, 12 days (IQR, 11-12 days) at 200 larvae/litre and 13 days (IQR, 12-13 days) at 400 larvae litre. The differences in time to eclosure between densities again proved to be significant both overall (Chi-square: df=2, χ^2=461.84, P<0.001)(figure 3.3.7) and in pairwise comparisons (Chi-square: df=1, χ^2>95.75, P<0.001)(figure 3.3.8). In EVida3 individuals in the competition cohort, median time to eclosure was 8 days (IQR, 7-8 days), 9 days (IQR, 8-9 days) and 12 days (IQR, 10-12 days) at the 100, 200 and 400 larvae/litre growth densities. Time to eclosure between densities again proved to be significantly different overall (Chi-square: df=2, χ^2=371.98, P<0.001)(figure 3.3.7) and in pairwise comparisons (Chi-square: df=1, χ^2>69.39, P<0.001). When we looked at time to eclosure between mixed Mopti and EVida3 larvae at each growth density (figure 3.3.8), we again found that the transgenic EVida3 had a significantly shorter time to eclosure than the Mopti at all densities Chi-square: df=1, χ^2>143.95, P<0.001).

In terms of survival in the single strain controls in experiment 2 (Mopti x EVida3, figure 3.3.7), the single strain control Mopti survival was 100±0% at 100 larvae/litre, 94.68±2.12% at 200 larvae/litre and 56.13±3.46% at 400 larvae/litre. Differences in survival were significant between all pairwise growth density comparisons (Chi-square: df=1, χ^2>20.84, P<0.001). In the EVida3 single strain control survival was 92.31±3.65% at

A

Comparison	Growth Density (larvae/litre)	Strain	Median Time to Eclosure (days)	IQR[1]	Mean Survival (%)	95% CI[2]
Single Strain	100	Mopti	9	8-9	67	5.4
		EE	8	8-9	74.3	5
	200	Mopti	9	8-10	71.8	3.6
		EE	9	8-9	57.4	4
	400	Mopti	11	11-12	67.3	2.7
		EE	11	9-12	48.7	2.9
Mixed Strains	100	Mopti	9	8-9	65.3	7.7
		EE	9	8-9	73.3	7.2
	200	Mopti	9	8-9	68	5.3
		EE	9	8-9	50.7	5.7
	400	Mopti	11	9-12	63	3.8
		EE	11	9-12	45.7	4

B

Comparison	Growth Density (larvae/litre)	Strain	Median Time to Eclosure (days)	IQR	Mean Survival (%)	95% CI
Single Strain	100	Mopti	9	8-10	100	0
		EVida3	8	7-9	92.31	3.65
	200	Mopti	10	8-11	94.68	2.12
		EVida3	10	8-11	72.5	4.39
	400	Mopti	12	11-12	56.13	3.46
		EVida3	11	9-12	21.25	2.84
Mixed Strains	100	Mopti	10	9-10	92.31	5.58
		EVida3	8	7-8	100	0
	200	Mopti	12	11-12	74.45	6.09
		EVida3	9	8-9	71	6.34
	400	Mopti	13	12-13	48	4.92
		EVida3	12	10-12	70.75	4.48

[1] Interquartile range

[2] Confidence interval

Table 3.3.7: Summary data for density-dependent effects on larval development time and survival all comparisons. A) Experiment 1 (EE vs Mopti) **B)** Experiment 2 (EVida3 vs Mopti). Strains were grown in varying densities and either as a single strain (single comparisons) or in a 1:1 mix of strains (mixed strains). See sectiosn 3.3.2 and onwards.

100 larave/litre, 72.5±4.39% at 200 larvae/litre and 21.25±2.84 at 400 larvae/litre. Differences in survival were again significant across all pairwise density comparisons (Chi-square: df=1, χ^2>37.142, P<0.001). When we compared survival in the control cohort Mopti and EVida3 at each growth density (figure 3.3.8), we found that Mopti individuals had a greater proportionate survival relative to EVida3 at all densities (Chi-square: df=1, χ^2>25.31, P<0.001).

In the competition comparison, Mopti survival was 92.31±5.58%, 74.45±6.09% and 48±4.92% in the 100, 200 and 400 larvae/litre comparisons. Observed differences in survival were again significant between all growth densities (Chi-square: df=1, $\chi^2$14.29, P<0.001). The EVida3 individuals in the competition cohort had the following rate of survival to eclosure: 100±0% at 100 larvae/litre, 71±6.34% at 200 larvae/litre and 70.75±4.48% at 400 larvae/litre. Survival was significantly different between EVida3 individuals grown at 100 larvae/litre and individuals of the same genotype grown at 200 and 400 larvae/litre (Chi-square: df=1, χ^2>11.34, P<0.001). There was no significant difference in survival between EVida3 individuals in the 200, and 400 larvae/litre competition group comparisons (Chi-square: df=1, χ^2=0.004, P=0.949). In contrast to the single-strain controls, EVida3 individuals survived significantly better than Mopti individuals at 100 larvae/litre (Chi-square: df=1, χ^2=11.34, P<0.001) and 400 larvae/litre (Chi-square: df=1, χ^2=42.37, P<0.001) when the two were mixed in equal proportion. There was no difference in survival between genotypes at 200 larvae/litre (Chi-square: df=1, χ^2=0.618, P=0.432).

Finally, comparing the relative performance of each genotype between controls and competition at each density we found that EVida3 individuals both survived better and reached eclosure faster in the competition trays compared to the single strain controls. This increase in mean survival was significant at the 100 and 400 larvae/litre densities (Chi-

square: df=1, χ^2>14.24, *P*<0.001) and the decrease in median time to eclosure was significant at the 100 and 200 larvae/litre densities (Chi-square: df=1, χ^2>7.54, *P*<0.006). In contrast, wild type Mopti individuals survived both significantly worse on average (Chi-square: df=1, χ^2>7.07, *P*<0.008) and reached eclosure in significantly more time (Chi-square: df=1, χ^2>8.95, *P*<0.003) in the competition comparison compared to Mopti individuals raised in the single-genotype controls at all densities.

The factors effecting both development time and survival were assessed by constructing a proportional hazards survival (PHS) test and logistic regression (LR) respectively. In both cases model effects were constructed from the independent variables of Replicate, Growth Density (100, 200 or 400 larvae/litre), Genotype (Mopti or EVida3) and Comparison (single or mixed strains), as well as interactions between these factors. In terms of development time (table 3.3.5), once the effect of Replicate had been partitioned, Growth Density, Genotype and Comparison were all identified as significant factors. There were also significant interactions between Genotype and Comparison, and Growth Density and Comparison. As Growth Density had such a strong effect (PHS: n=4284, df=2, χ^2=1245.11, *P*<0.001), we also broke down the analysis by density: Genotype and Comparison as well as the interaction between the two were significant at all densities (table 3.3.5). In terms of survival, all model effects were again significant with the exception of Comparison (table 3.3.6). There was a significant interaction between all combinations of Comparison, Genotype and Growth Density. Again, the analysis was broken down by density and all factors were preserved as significant with the exception of Comparison at 100 larvae/litre, there were a significant interaction between Genotype and Comparison (table 3.3.6).

Random samples of the wing lengths of 30 individuals from each growth density, genotype and both the control and competition groups were compared (figure 3.3.9). There
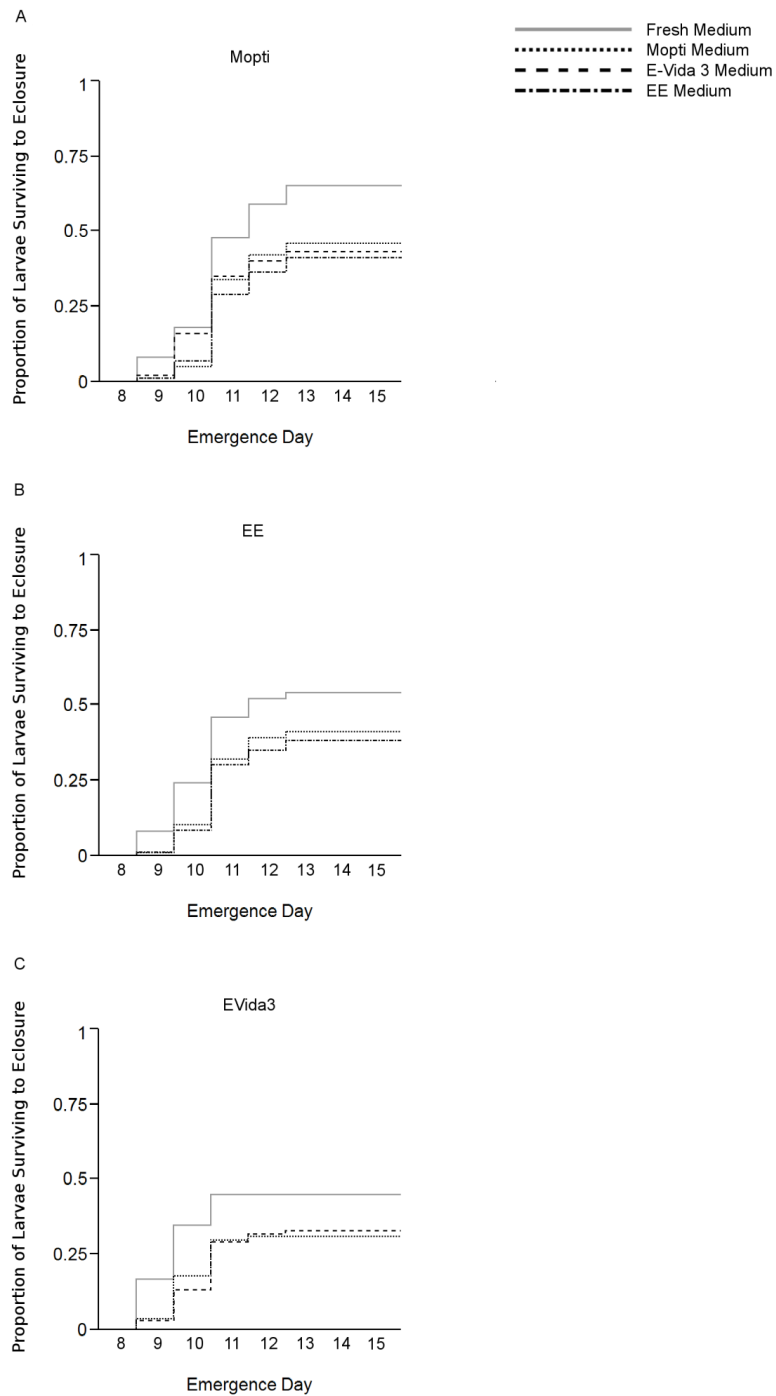
Figure 3.3.10: Cumulative frequency of larvae surviving to eclosure by genotype in the dose dependent 'medium swap' experiment. A) Mopti. B) EE. C) EVida3. Separate cumulative frequency plots for each media source (Fresh, 50% Mopti, 50% EE, 50% EVida3) are indicated. Endpoints represent total proportion of larvae surviving to eclosure in each comparison.

was no significant difference in mean wing length between genotypes or between the control and competition groups. There was however a significant difference in mean wing length between growth densities, with adults raised at a density of 100 larvae/litre having a significantly larger mean wing length than adults raised at 200 larvae/litre.

3.3.3 Dose-dependent effects of used growth medium on larval development time

After observing such a marked difference in survival and development time between Mopti and EVida3 individuals, and the change in these variables induced by mixing the two strains (see preceding section), we hypothesised that, among other potential explanations (see discussion) that patterns of survival observed previously could be explained if the EVida3 were toxifying their larval growth environment through the leaky expression and subsequent excretion of the Vida3 AMP. This process could potentially explain why the EVida3 survived better whilst mixed at a given growth density compared to the unmixed control and equally why the Mopti survived poorly relative to the unmixed control.

To test this hypothesis we designed an experiment where, along with fresh water controls, our three strains (Mopti, EE and EVida3) were grown in 'used' medium from each strain, both in a 50% v/v mix with ddH_2O and in 100% used medium.

Mopti larvae grown in fresh medium had a median time to eclosure of 11 days (IQR, 10-12 days) and mean survival of 65.08±2.7%. In 50% v/v 'Mopti' used medium, median time to eclosure was 11 days (IQR, 11-12 days) and mean survival was 45.75±2.82%. In 50% v/v 'EE' used medium eclosure time was 11 days (IQR, 11-12 days) and survival 40.83±2.58%. Finally eclosure time and survival of Mopti larvae in 50% v/v 'EVida3' medium was 11 days (IQR, 10-11 days) and 42.83±2.8% respectively. EE larvae grown in fresh medium had a median eclosure time of 11 days (IQR, 10-11 days) and mean survival of 54.42±2.81%. In 50% v/v 'EE' medium, median eclosure time was 11

days (IQR, 11-11 days) and survival was 38.08±2% and EE larvae grown in 50% v/v 'Mopti' medium had a median time to eclosure of 11 days (IQR, 10-11 days) and mean survival of 41.08±2.79%. EVida3 larvae grown in fresh water had a median time to eclosure of 10 days (IQR, 9-10 days) and a mean survival of 45.25±2.82%. In 50% v/v 'EVida3' medium, median time to eclosure was 10 days (IQR, 10-11 days) and mean survival 31.42±2.62%. Finally EVida3 larvae grown in 50% v/v 'Mopti' medium had a median time to eclosure of 11 days (IQR, 10-11 days) and a mean survival of 33.42±2.66% (figure 3.3.10). There were no survivors to eclosure of any strain in any of the 100% media comparisons.

We constructed logistic regression and proportional hazard survival models to partition the effects of experimental variables on survival and time to eclosure respectively. To ensure a balanced analysis, models were nested by the comparisons Mopti x EE and Mopti x EVida3. In our initial analysis, model effects were constructed from the variables Replicate, Genotype (wild-type, transgenic) and Media Source (fresh, used (50%)) as well as the interaction of Genotype and Media Source. In the Mopti x EE comparison both Genotype and Media Source had a significant effect on survival (LR: n=7200, df=1, χ^2>34.39, P<0.001). The interaction between Genotype and Media Source was also significant (LR: n=7200, df=1, χ^2=8.20, P=0.004). The same factors were also significant in terms of time to eclosure in the Mopti x EE comparison. In the Mopti x EVida3 comparison Genotype, Media Source and their interaction were again significant factors in both survival and time to eclosure.

Our second analysis excluded the cohort of larvae grown in fresh water: model effects were constructed from Replicate, Genotype and Media Source (used: self, or used: other) as well as the interaction of Genotype and Media Source. With this analysis, in the Mopti x EE comparison, Genotype was the only significant factor in terms of time to

eclosure (PHS: n=7200, df=1, χ^2=8.19, P=0.004). Both Genotype (LR: n=7200, df=1, χ^2=6.77, P=0.009) and the interaction of Genotype and Media Source (LR: n=7200, df=1, χ^2=7.72, P=0.006) were significant in terms of survival. In the Mopti x EVida3 comparison Genotype was the only significant factor in terms of survival (LR: n=7200, df=1, χ^2=71.87, P<0.001). In terms of time to eclosure, both Genotype (PHS: n=7200, df=1, χ^2=65.19, P<0.001) and the interaction of Genotype and Media Source (PHS: n=7200, df=1, χ^2=34.78, P<34.78) were significant factors.

3.4 Discussion

We assessed egg production and larval development in our two transgenic strains relative to the non-trangenic Mopti. Egg production and larval survival and development are important sentinel traits for detecting gross fitness effects in transgenic strains and provide an insight into how a transgene might perform in a release scenario. This is one of only a few previous studies assessing larval stress-factors and competition in transgenic mosquitoes, and the first to consider these factors in the key Afro-tropical malaria vector *Anopheles gambiae* s.s.

When we looked at egg production, we found that EE females laid significantly more eggs and successfully oviposited more frequently than either the phase 2 transgenic EVida 3 or the non-transgenic Mopti. Despite the EE also having a significantly larger median wing length we could detect no correlation between wing length and brood size in our experimental sample. This is in contrast to several studies in the literature that have found a strong correlation between these factors (Hogg *et al* 1996, Takken *et al* 1998). The fact that relatively few Mopti and EVida3 individuals successfully oviposited (14 and 15 of 32 respectively) compared to EE, combined with the lower egg-count and lower median wing length may suggest that a proportion of the Mopti and EVida3 females selected were

in a pre-gravid state compared to EE. However, if this were the case we would expect there to be a significant correlation between wing-length and oviposition success. When we modelled these effects, genotype was the only significant factor. This may indicate a fitness load in the EVida3 not encountered by the EE causing a reduction in relative oviposition success and brood size. The difference in performance between EE and Mopti is harder to account for, but may be explained by the difference in genetic background between the two strains. Both EE and EVida3 are derived from the long-established KIL strain and thus may be better adapted to the higher stress conditions of the single oviposition tubes used in this experiment compared to the relatively recently colonised (2003) and refreshed (2008) Mopti strain. Whilst the rationale for using a relatively genetically diverse strain was a strong one in our larval development (this chapter) and transgene stability (chapter 4) experiments, here, without also including the KIL strain, we may have introduced a confounding effect of genetic background into comparisons between Mopti and the two transgenic strains. Nevertheless, this would not affect the relative performance of EE and EVida3, where we did observe a significant difference in oviposition successs and brood size independent of wing length.

In terms of egg morphology, we were able to detect difference between each strain in all of the morphological traits we investigated. However none of these traits proved to be capable of accurately discriminating between genotypes. The four morphological features (Area, Perimeter, Feret's Diameter and Circularity) collapsed in to two principal components broadly describing egg size (PC1) and egg shape (PC2). PC2 did not vary between genotypes, suggesting that eggs were of a similar shape in all strains. EE individuals had a significantly smaller score for PC1 suggesting that their eggs were smaller overall - this is corroborated by the individual morphological features. Despite this, we could not detect a correlation between PC1 and brood size.

In addition to the egg production and morphology experiments, we also considered the effects of transgenesis, larval growth density and competition on body size, time to- and survival to eclosure. In the single strain comparisons we found that the non-transgenic Mopti survived significantly better but developed significantly slower at all densities than both transgenic strains. Increasing larval growth density reduced wing length and increased development time across all comparisons. In terms of survival, there was a clear negative effect on survival with increasing larval growth density in both transgenic strains, with EVida3 in particular surviving extremely poorly at high larval growth densities - again suggestive of an additional fitness load present in the EVida3 compared to EE. The effects of density on survival to eclosure in the Mopti was less consistent: we detected no difference in survival at increasing larval growth densities in singly grown Mopti in our first experiment (Mopti x EE) but found a strong negative effect of density in singly grown Mopti in the second experiment (Mopti x EVida3). Although the survival of Mopti did vary temporally. Reconciling these differing results in the single strain Mopti comparisons is difficult. Although both neutral (Solomon *et al* 2000) and negative (Gimnig *et al* 2002) effects of increasing larval density on survival have been independently reported in *An. gambiae* larvae previously, the experimental conditions of the two experiments were wildly different. The study by Solomon *et al* in 2000 matched our experimental design closely and this group reported a strong negative effect of density on survival but not on development time. The Gimnig study was a semi-field experiment only tangentially similar to our own, but reported a significant increase in time to eclosure but no change in survival at increasing densities. As our two experiments were separated by several months, it is possible that there some variability in the extrinsic environmental conditions - although both experiments were carried out in an environmentally controlled insectary - or in the quality of the larval food source - but again this is closely monitored and maintained.

When we mixed Mopti with each of our transgenic strains we observed some surprising results, particularly in the Mopti X EVida3 comparison. In the Mopti x EE competition comparison we found no change in overall survival in either genotype relative to the single strain controls, indicating no competitive advantage/disadvantage in either strain. We did however observe that Mopti larvae reached eclosure faster in the competition comparison compared to the single strain control.

In the Mopti x EVida3 competition comparison we observed and almost complete reversal in the pattern of survival between Mopti and EVida3 compared to the results of the single strain controls. EVida3 survived better and developed faster than both single raised EVida3 and the individuals with which they were mixed over all densities. In contrast, Mopti development time and survival were both strongly negatively affected by competition with the EVida3, indicating they were at a significant competitive disadvantage. We hypothesised that EVida3 individuals may be intoxicating their growth medium through 'leaky' expression and excretion of their AMP transgene. This hypothesis could explain both why singly raised EVida3 at high density survived comparatively worse than either EE or Mopti, and also why Mopti larvae were at such a marked competitive disadvantage when mixed with EVida3. We tested this hypothesis by raising Mopti, EE and EVida3 larvae at high density and both fresh water and either a 50% or 1005 solution of used medium from each strain. Whilst we did detect a negative effect on survival and development time in larvae of all strains raised in 50% used medium compared to those grown in fresh water, there was no difference in survival between individuals grown in their 'self' used medium compared to those grown in medium from another strain. This indicates that whilst used medium was indeed more toxic than fresh water, there was no difference in toxicity between media from each strain. An alternative hypothesis is that, due to the significantly faster development time observed in the EVida3 compared to

Mopti, there was a sufficient difference in the timing to each instar molt that EVida3 effectively avoided direct competition with the slower developing Mopti larvae. This would reduced the size of the effective population of EVida3, increasing survival but also decreasing food availability for the competing Mopti, decreasing their survival.

These data present a complex picture of the effects of transgenesis on larval development and female egg production. Whilst we detected potential fitness costs in the EVida3 relative to EE in terms of both egg production and survival at high density during larval development. The EVida3 proved to be at a considerable competitive advantage over the non-transgenic Mopti in mixed larval development. How these effects interact will be the subject of the next experimental chapter, where we look at the stability of both transgene cassettes over multiple generations.

4. Long term performance of transgene cassettes in mixed populations.

Abstract

The deployment of transgenic mosquitoes carrying genes for refractoriness to malaria has long been seen as a futuristic scenario riddled with technical difficulties. The integration of anti-malarial effector genes and a gene-drive system into the mosquito genome without affecting mosquito fitness is recognized as critical to the success of this malaria control strategy.

Here we conducted detailed fitness studies of two *Anopheles gambiae s.s.* transgenic lines recently developed using a two-phase targeted genetic transformation system. In replicated cage-invasion experiments, males and females of the EE Phase-1 docking strain and EVida3 Phase-2 strain loaded with an antimicrobial peptide (AMP) expressed upon blood-feeding, were mixed with individuals of a recently-colonized strain of the Mopti chromosomal form. The experimental design enabled us to detect initial strain reproductive success differences, assortative mating and hybrid vigor that may characterize mosquito release situations. In addition, the potential fitness costs of the unloaded Phase-1 and loaded Phase-2 genetic constructs, independent of the strains' original genetic backgrounds, were estimated between the 1st instar larvae, pupae and adult stages over 10 generations.

The Phase-1 unloaded docking cassette was found to have significantly lower allelic fitness relative to the wild type allele during larval development. However, overall genotypic fitness was comparable to the wild type allele across all stages leading to stable equilibrium in all replicates. In contrast, the Phase-2 construct expressing EVida3 disappeared from all replicates within 10 generations due to lower fitness of hemi- and homozygous larvae, suggesting costly background AMP expression and/or of the DsRed2

marker. This is the first study to effectively partition independent fitness stage-specific determinants in unloaded and loaded transgenic strains of a Phase-1-2 transformation system. Critically, the high fitness of the Phase-1 docking strain makes it the ideal model system for measuring the genetic load of novel candidate anti-malarial molecules *in vivo*.

4.1 Introduction

As the technological and methodological hurdles of achieving efficient transgenesis and developing gene-drive systems capable of spreading effector genes into target populations look to be overcome in the very near future, there has been a growing focus on the practical implementation of transgenic mosquito release as a means of disease control. The recent publication of work describing a functional, transgenic, homing endonuclease gene drive system in *Anopheles gambiae* (Windbachler *et al* 2011) is a massive step towards developing and deploying an effective population replacement strategy. This, coupled with a rapidly expanding repertoire of potential anti-malarial effector genes (Isaacs *et al* 2011), newly characterised expression systems (Nolan *et al* 2011) and increasingly efficient transformation protocols (Meredith *et al* 2011) mean that we are better placed than ever to develop a system to drive transgenic malaria refractoriness into wild mosquito populations.

All transgenic control strategies rely on transformed male mosquitoes being able to successfully compete with wild males for mates once released in the field. In the case of population replacement strategies, the F1 and subsequent generations carrying transgenic constructs post-release must also be vigorous, fecund and robust enough to ensure the continuing spread of these genes through the target population. Thus, assessing the fitness and mating competitiveness of transgenic lines, but most critically of the transgenic alleles once it spreads within the wild type population is a vital step in the development of functional transgenic mosquitoes for the control of malaria transmission.

There are a number of ways in which transformation could potentially affect fitness (reviewed in Marrelli, 2006). Firstly the strong expression of exogenous genes may reduce the competitiveness of a transgenic individual by having a deleterious behavioural or physiological effect as it accumulates in tissues (e.g. Moreira *et al* 2004), or simply by imposing an additional metabolic cost on the transgenic strain not suffered by a wild type competitor (e.g. Catteruccia *et al* 2003). Secondly, and independent of transgene expression, the site at which a transgenic construct integrates into the target genome can itself have a significant effect on fitness. For example, the transgene may integrate into the open reading frame or regulatory sequence of an endogenous gene, thus interrupting its function and leading to fitness costs or even recessive lethality (e.g. Irvin *et al* 2004). Thirdly, the process by which a transgenic lineage is created necessarily involves at least one - and in some cases two - severe genetic bottlenecks where a single mosquito is the progenitor of the entire subsequent population of transgenic insects, leading to inbreeding depression and fixation of deleterious recessive alleles by random genetic drift. This effect can be, theoretically, ameliorated by successive generations of outcrossing to more genetically diverse populations. Finally, and depending on the site of integration and the genetic background of the mosquito, deleterious recessive alleles at loci proximal to the site of the transgene integration can - in a process known as hitchhiking - be positively selected for through tight-linkage with the transgene insert and may impose a fitness cost in homozygous individuals (Marrelli *et al* 2006).

Evaluating the fitness of transgenic mosquito lines can be done in several ways. Direct comparisons of genetically-modified strains to their unmodified parental strain or a wild-type colony have been made in order to compare fitness components such as adult fecundity as well as developmental rates and survival at different life stages. In theory, such comparisons do not allow partitioning of the fitness costs linked to the transgenic

mosquito genetic background (e.g. inbreeding depression) from those linked to the genomic location of the transgenic construct or the expression of its effector molecules. However, since the properties of these molecules - e.g. antiparasitic - are often tested on homozygous transgenic lines, direct mosquito fitness comparisons may serve to objectively identify grossly unfit homozygous lines that may not be worth further characterization. Direct comparisons, have revealed strong fitness costs in terms of fertility and survivorship in transgenic lines of *Aedes aegypti* carrying an enhanced GFP gene or expressing transposase from the Hermes and MOS1 elements (Irvin *et al* 2004). They also showed reduced size, survival and longevity in the OX513A Line of *Aedes aegypti* that carries a tetracycline repressible, dominant lethal positive feedback system for sterile insect release (Bargielowski *et al* 2011). The confounding effects of genetic background inherent to the direct comparisons approach are typically decreased by repeatedly backcrossing transgenic lines into a wild-type line in order to increase their heterozygosity prior to experimentation. For example, comparisons of non-transgenic and transgenic lines have revealed differences in fertility and survival between *An. stephensi* transgenic lines expressing active bee venom and non-trangenic lines suggesting a negative effect on their midgut nutrient absorption (Moreira *et al* 2004). Further comparisons in *Aedes fluviatilis* expressing inactive bee venom revealed no apparent negative effects of the protein, no difference in fertility, and even increased survival in some transgenic lines compared to non-transgenic ones (Rodrigues *et al* 2008).

A second approach for evaluating the fitness of transgenic lines that resolves some of the limitations of direct strain comparisons has been to compare the fitness parameters of individuals hemizygous for the transgene, with those of sibling wild-type individuals (Amenya *et al* 2010, Isaacs *et al* 2012). Hemizygosity is achieved by first crossing homozygous transgenic with wild-type individuals and eliminates the confounding factors

of inbreeding depression and potential costs of recessive alleles hitch-hiking with the construct. Although this constitutes a vast improvement over direct homozygous strain comparisons, fitness costs that usually affect individuals homozygous for the transgene construct (i.e. recessive and co-dominant effects) cannot be measured. The lack of evaluation of transgene fitness costs at the homozygous state is made particularly obvious in studies that test the effects of antiparasitic effector molecules using homozygous individuals but transgene fitness costs on hemizygous ones (Amenya *et al* 2010, Isaacs *et al* 2012).

Finally, the fitness of the transgenic construct independent of the transgenic line's genetic background can be followed using cage-invasion experiments in which the transgenic allele is introduced into a wild-type population and its frequency monitored over time (i.e. Catteruccia *et al* 2003, Moreira *et al* 2004, Marrelli *et al* 2007, Li *et al* 2008). These experiments best simulate real release-like situations but require carefully planned and comparatively complicated design. The main advantages of such approach are that: (1) they allow direct competition between transgenic and wild-type alleles; (2) they enable the independent assessment of the fitness of individuals hemi- and homozygous for the transgene (i.e. recessive, co-dominant, dominant effects); (3) Several generations-worth of recombination breaks down the linkage between the construct and all but the closest recessive deleterious genes that may be hitch-hiking with it. Depending on the design of the experiment, one can also assess the initial fitness of homozygous transgenic and wild-type individuals, potential problems associated with assortative mating amongst released homozygous transgenic individuals, and the importance of hybrid vigor in first generation hemizygous individuals. All of these aspects contribute to making cage-invasion experiments not only the most rigorous for assessing the fitness of transgenic strains but

also the most useful in terms of generating the fitness parameters required for population dynamic models of transgenic spread in target populations.

Only a handful of studies explicitly investigating transgenic mosquito fitness have described a fitness-neutral transformation that is stable in mixed populations over multiple generations. Cage-invasion experiments complementing direct strain comparisons between wild-type and a transgenic *An. stephensi* line expressing SM1 demonstrated that the transgenic construct subsisted in test populations for 5 generations (Moreira *et al* 2004). Using the same approach *Aedes fluviatilis* lines expressing inactive been venom enzyme PLA$_2$ were shown to bear no apparent fitness costs (Rodrigues *et al* 2008). However, most other studies investigating the stability of a given transgenic constructs over multiple generations have observed a rapid decrease in transgene frequency, and in some cases total extinction of the transgenic allele (Catteruccia *et al* 2003, Moreira *et al* 2004, Irvin *et al* 2004, Li *et al* 2008).

Recent progress in the development of site-specific transgene integration systems in *Ae. aegypti* (Nimmo *et al* 2006) and *An. gambiae* (Meredith *et al* 2011) can potentially provide the scientific community with the means to thoroughly evaluate the potential fitness of a whole suite of effector transgenes. Site-specific transgene integration relies on two steps of genetic transformation: Phase 1 uses transposon-like integration to create a so-called docking strain carrying a phenotypic marker and site-specific *phiC31* integrase recognition *attP* sequences; Phase 2 uses the *attP* sequences and endo- or exogenous transposase in order to integrate a second phenotypic marker and an effector gene within the docking cassette. The power of this approach lays in the possibility to efficiently produce and compare different Phase 2 loaded transgenic lines produced from on a single well-characterized Phase 1 docking strain. Having different effector genes and their promoter sequences located precisely in same location in the mosquito genome, effectively

controls for variation in potential fitness costs caused by gene-hitchhiking, positional expression effects and the site of integration.

As a proof of principal, we set out to assess and compare the fitness of the unloaded Phase-1 EE docking strain and loaded Phase 2 EVida3 transgenic lines recently developed using the two-phase targeted genetic transformation system in *An. gambiae* s.s. (Meredith *et al* 2011). Preliminary studies of the EE docking strain and the EVida3 strain which expresses a tetramer of the putative Vida AMP (Arrighi *et al* 2002) under the control of the *An. gambiae carboxypeptidase* promoter suggested that the two strains bred and survived well under standard laboratory conditions. Here we performed replicated cage-invasion experiments to assess the long-term stability of the Phase 1 and 2 genetic constructs independent of their genetic background when competing against wild-type alleles. In addition, the design of the experiment allowed us to detect initial differences in fitness and assortative mating in the transgenic strains, as well as to evaluate the importance of heterosis in their F1 progeny. The results highlight the power of cage-experiments for partitioning the different sources of fitness costs potentially affecting genetically-modified alleles in a mosquito release context. The EE docking line provides researchers with the ideal system to test the potential genetic load of candidate transgenic constructs carrying effector genes targeting the malaria parasite or other mosquito traits affecting malaria transmission.

4.2 Materials and Methods

4.2.1 Mosquito strains and insectary conditions

The EE and EVida3 transgenic strains of *An. gambiae.* developed by Meredith *et al.* (Meredith *et al* 2011) were used to assess the different sources of fitness costs potentially affecting transgenic lines. The Phase 1 EE strain carries a transgene cassette consisting of

the phenotypic marker ECFP under the control of the 3xP3 promoter driving its expression in the eyes and other nerve tissues, and the *phiC31* integrase recognition sequence *attP* (Thyagarajan *et al* 2001). The Phase 2 EVida3 strain derived from the EE strain in a second transformation step carries a cassette consisting of 3xP3 ECFP, an additional marker 3xP3 *Ds*Red and the putative AMP Vida3 sequence with the *An. gambiae carboxypeptidase* promotor, signal peptide and UTRs (Meredith *et al* 2011). The Phase 1-2 integration site is located on chromosome 3R (position 15801959 - band 31B) and is some distance from any of the inversion polymorphisms commonly found in *An. gambiae* s.s. (Coluzzi *et al* 1985, Toure *et al* 1998). The two transgenic lines were derived from the wild-type strain KIL which was originally colonized from Tanzania in the 1970's. Both transgenic strains are of the M molecular form (della Torre *et al* 2001). The wild-type strain used in this experiment is a Mopti, M-form population originally colonized from the village of N'Gabakoro Droit, Mali in 2003. Since it has been in our laboratory, the Mopti strain has refreshed yearly by outcrossing to the F1 of field caught individuals from the same site. Both transgenic stocks are maintained as true-breeding homozygotes and, along with wild-type strains, are maintained in dedicated insectaries as described previously (see Chapter 2)

4.2.2 Cage invasion experiments

Cage invasion experiments were initiated by mixing 100 male and 100 female homozygous wild-type mosquitoes (WT) with 100 male and 100 female homozygous transgenic mosquitoes (TT). All individuals were 3-5-day old and unmated prior to mixing. After allowing 2 dark cycles for mating, mosquitoes were blood fed to produce eggs and, after a further 2d, provided with a ~10cm diameter pot lined with wet filter paper (grade 1, Whatman) for oviposition. Eggs were hatched in 1l of ddH$_2$O and L1 larvae separated into

growth trays at a density of 200 larvae/l with a total of 6 trays/1200 larvae per generation per experimental replicate. Larvae were maintained in the same conditions as the stock populations (see above). Once pupated, individuals were transferred to a standard 5l adult enclosure to emerge. Adult were maintained in the same conditions as the stock populations (see above) and left to mature and mate. 4 days after adding the last pupae to the cage, adult females were blood fed to produce the next generation. Mixed populations were maintained in this way for 10 generations.

4.2.3 Sampling

The frequency of the transgene was determined at three key life stages: L1 larvae (sample 1), pupae (sample 2) and 2 day post-emergence adults (sample 3). At each life stage, 48 individuals were selected at random from each population and genomic DNA was extracted using a modified DNAzol gDNA extraction protocol (Invitrogen). Transgenic status was then determined by carrying out a PCR on the extracted DNA using primers designed to produce characteristic gel bands for homozygous transgenic (TT), homozygous wild type (WW) or a hemizygous hybrid (TW). Hence the precise genotypic and allelic frequencies could be calculated for each life stages.

4.2.4 Data analyses

Mating and fitness in the initial generation (F0-F1)

Assortative mating/hybrid deficiencies in both experiments were tested by comparing the observed frequency of hybrids and homozygotes genotypes in the L1 larvae sample of the F1 progeny to the 50:50 ratio predicted given the equal numbers of WW and TT males and females used to initiate each experiment using Chi-square Goodness of Fit tests. Similarly, the overall fitness of transgenic and non-transgenic lines was assessed prior to any

recombination event by comparing the frequencies of transgenic and wild-type alleles in the F1 progeny (L1 Larvae in both experiments) using Chi-square Goodness of Fit tests. Finally, the effects of heterozygosity and/or hybrid vigor on survival from the larval to the pupae and from pupae to the adult stage were tested by comparing the absolute genotypic fitness of hemizygotes and homozygotes between the F1 L1 larvae and F1 adult stages using Chi-square Goodness of Fit tests.

Absolute genotypic fitness ($W_{(abs)}$) was estimated as the change in frequency (f) of a given genotype over time, either between generations or between samples, (where AA is the genotype considered):

$$W_{(abs)} = f_{(AA)\ Fn} / f_{(AA)\ Fn-1}\ \textbf{(Eq. 1).}$$

Similarly, absolute allelic fitness was calculated as (where A is the allele considered):

$$W_{(abs)} = f_{(A)\ Fn} / f_{(A)\ Fn-1}\ \textbf{(Eq. 2)}$$

Transgenic vs wild-type fitness comparisons (F1-10)

Genotypic and allelic fitness was assessed by monitoring allele and genotype frequency over time, both between generations and between samples within each generation. Based on the starting conditions (100 males and 100 female homozygous wild-type WW and 100 males and 100 female homozygous TT of either phase 1 EE or phase 2 EVida3) and assuming random mating and no fitness costs on the transgenic strains and transgenic allele, the expected frequencies are 0.25 for homozygote WW and TT and 0.50 for hemizygous WT individuals.

EE vs Evida relative fitness comparisons (F1-10)

Allelic and genotypic fitness relative to the wild-type strain $W_{(rel)}$ was calculated and plotted in graphs as the absolute fitness $W_{(abs)}$ normalized by dividing it by the absolute fitness of the wild-type strain $W_{(abs\ WT)}$:

$$W_{(rel)} = W_{(abs)} / W_{(abs\ WT)} \textbf{ (Eq. 3)}$$

The overall and generation by generation frequencies of EE and EVida3 transgenic alleles relative to those of wild-type Mopti allele were formally compared using Logistic Regression (LR) on the combined data from the 3 replicates.

Life stages-specific fitness costs (F2-F5)

The relative genotypic and allelic fitness $W_{(rel)}$ between generations and between samples within each generation was calculated from the differences in genotypic frequencies observed in generations F2-F5 following (Eq. 1, 2 and 3), but using between-stage changes rather than between generations ones. For simplicity and given the adequate sample sizes, the actual values of relative fitness $W_{(rel)}$ were directly used in non-parametric comparisons rather than working with the raw genotypic and allelic frequencies.

All statistical analysis and graphing were carried out using JMP (SAS Institute inc.). Significant differences between replicates were checked in every analysis and reported whenever appropriate.

4.3 Results

4.3.1 Assortative mating amongst strains

Evidence of assortative mating/hybrid deficiency in both experiments was tested for by
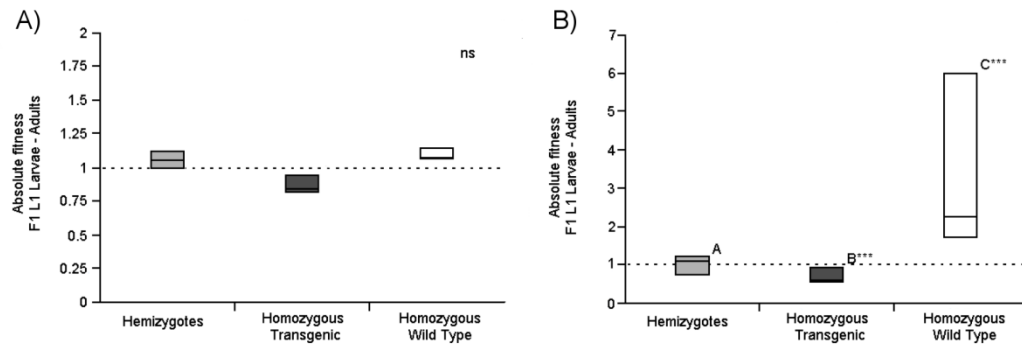
Figure 4.3.1- Genotypic fitness for hemizygotes, homozygous transgenic and homozygous wild type individuals from F1 larvae to adults - In (a) The fitness of the Mopti wildtype allele was compared to the Phase 1 EE allele (Experiment 1); in (b) to the Phase 2 Evida3 allele (experiment 2). Boxplots were median, quartiles and min-maximum values. The significance levels of a Chi-square test are indicated - ns, not significant, *: $P>$ 0.05, ** $P>$ 0.01, *** $P>$ 0.001.

comparing the observed frequency of hybrids and homozygote genotypes in the L1 larvae sample of the F1 progeny to those predicted given the equal numbers of homozygous males and females used at the start of each experiment (50:50 ratio). Significant assortative mating was observed in experiment 1 (Mopti vs EE) over all replicates (Chi-square Goodness of Fit: $n= 144$, df$= 1$, $\chi^2= 16.3$, $P< 0.001$) and within each replicate ($P< 0.05$ in all cases). In contrast, significant assortative mating in experiment 2 was only detected in replicate 2 (Chi-square: $n= 48$, df$= 1$, $\chi^2= 5.4$, $P= 0.018$) but was not significant over all replicates ($P< 0.05$).

4.3.2 Fitness differences between strains (F0-F1)

The overall fitness of transgenic and non-transgenic lines including the combined effects of male mating success, adult survival and female fertility, was assessed prior to any recombination event by comparing the frequencies of transgenic and wild-type alleles in the F1 progeny (L1 Larvae in both experiments). In comparisons of Mopti vs. EE (experiment 1) no overall significant difference was found between the fitness of the two strains (Chi-square Goodness of Fit: $n= 288$, df$= 1$, $\chi^2= 2.7$, $P= 0.099$) nor within any of the replicates ($P> 0.152$ in all cases). In contrast in the 2nd experiment, the EVIDA3 strain had higher initial fitness than the Mopti strain in the first and second replicate, leading to an overall significant difference across replicates (Chi-square: $n= 288$, df$= 1$, $\chi^2= 29.9$, $P< 0.001$)

4.3.3 Strain fitness and hybrid vigor (F1)

Evidence of any heterozygosity or hybrid vigor in the form of increased survival from the larval to the pupae and from pupae to the adult stage was specifically tested by comparing the absolute genotypic fitness of hemizygotes and homozygotes between the F1, L1 larvae

and F1 adults stages (Figure 4.3.1). In the first experiment 1 (Mopti vs EE, figure 4.3.1), there was no overall significant difference in absolute fitness between hemizygote TW and homozygotes TT and WW individuals from larval to adult stage over the 3 replicates (Chisquare Goodness of Fit: $n= 18$, df = 2, $\chi^2= 0.957$, $P= 0.620$).

In the second experiment (Mopti vs EVida3, figure 4.3.1), hemizygous and homozygous individuals for the transgene had significantly lower overall fitness than wild-type individuals indicating a strong negative effect of the transgene (Chi-square: $n= 72$, df = 2, $\chi^2= 42.0$, $P< 0.001$) across all 3 replicates. This was also observed within each replicate ($P<0.001$ in all cases). These negative fitness effects also affected the overall survival of F1 hemi and homozygous progeny from L1 larvae to pupae stages (Chi-square: $n= 32$, df = 2, $\chi^2= 18.3$, $P< 0.001$). Fitness costs also affected the overall emergence of homozygous but not hemizygous pupae to the adult stage (Chi^2 Goodness of Fit: $n= 65$, df = 2, $\chi^2= 18.3$, $P< 0.001$) although this effect was the strongest and significant only for replicate 2 (Chi^2 Goodness of Fit: $n= 31$, df = 2, $\chi^2= 27.7$, $P< 0.001$).

4.3.4 Transgenic vs wildtype fitness comparisons

Following mixing and recombination between the transgenic lines and the wild-type strain (Mopti) over 10 generations, the two transgenic elements exhibited strikingly different trajectories over time (figure 4.3.2). After 10 generations, the phase 1 (EE) transgene (experiment 1, figure 4.3.2) was present in all 3 replicates. Despite some fluctuations between F2 and F5, by generation F10 the observed genotypic frequencies of WT, TT and WW did not deviate significantly from Hardy Weinberg equilibrium nor from the 50:25:25 ratio predicted from starting conditions (Chi-square Goodness of Fit, $P > 0.05$ in all cases, figure 4.3.2). In contrast, the frequency of the phase 2 EVida3 transgenic construct (figure 4.3.2) decreased rapidly and was no longer detectable after 5 generations in two replicates,
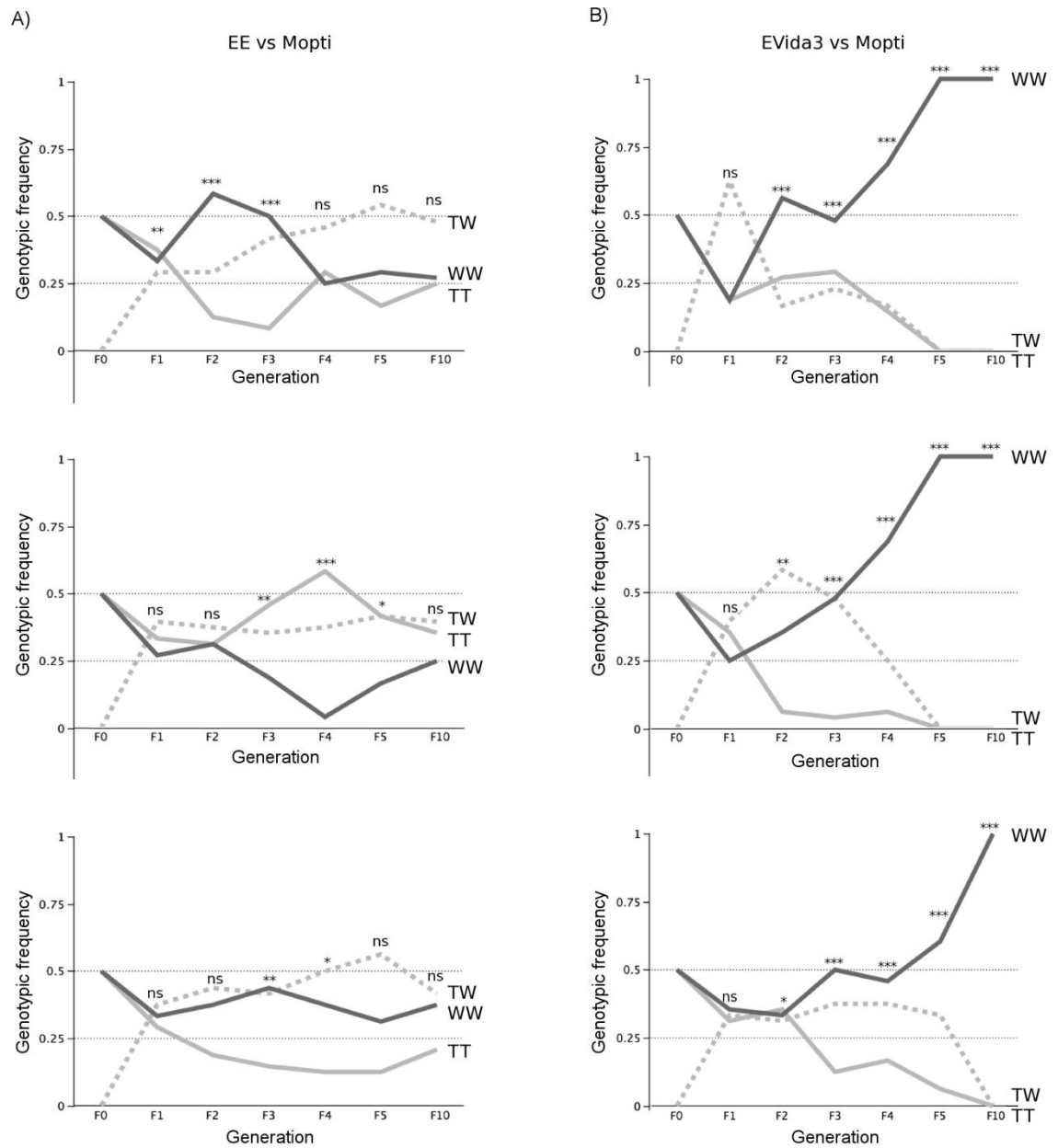
Figure 4.3.2- Frequency of hemizygote, homozygous transgenic and homozygous wildtype genotypes over 10 generations, (a) The frequency of homozygous Phase 1 EE (TT), homozygous wildtype (WW) and hemizygote (TW) genotypes (Experiment 1); (b) The frequency of homozygous Phase 2 Evida3 (TT), homozygous wildtype (WW) and hemizygote (TW) genotypes (Experiment 2). Each replicate was considered separately. The significance levels of a Chi-square test based on HWE are indicated - ns, not significant, *: $P > 0.05$, ** $P > 0.01$, *** $P > 0.001$.
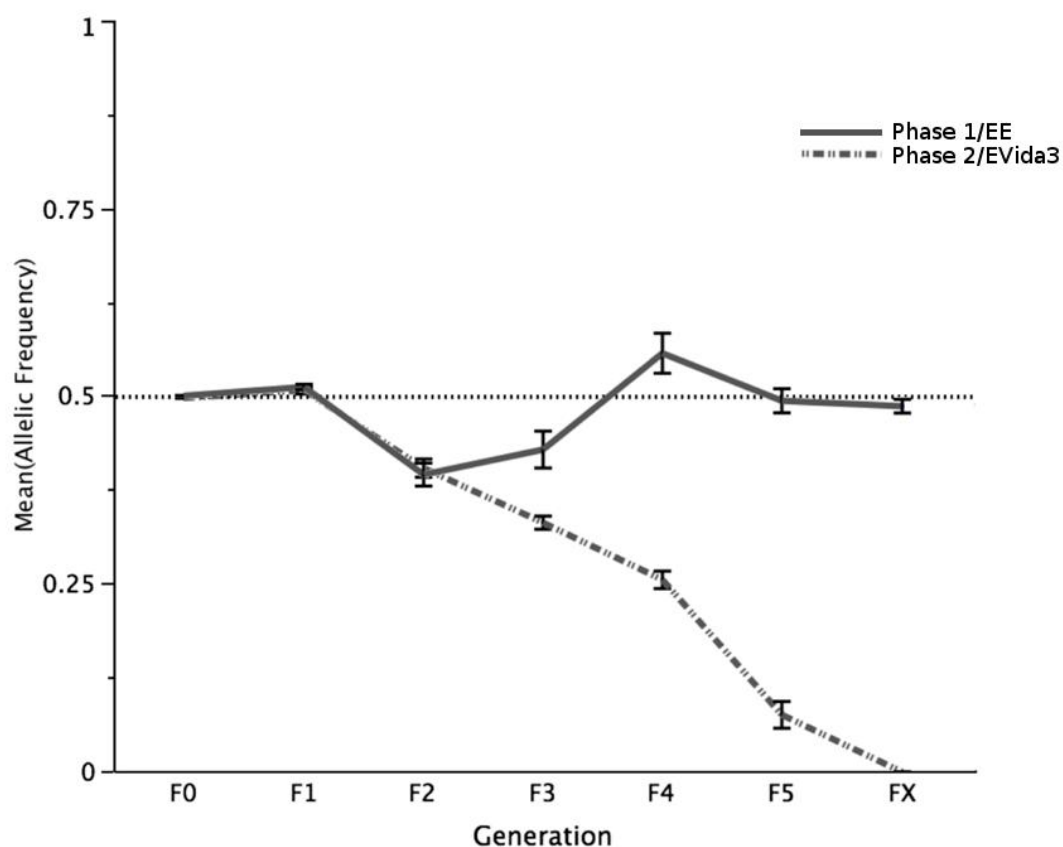
Figure 4.3.3 - Overlay of mean frequency of transgenic alleles over 10 generations.
The solid line represents the change in mean allele frequency of the EE/Phase 1 transgenic
allele. The hashed line represents the change in mean allele frequency of the EVida3/Phase
2 transgenic allele. Points are the mean allele frequency from three replicates, error bars
represent 95% confidence intervals.

and by generation 10 in the third. Deviations from HWE frequencies and from the ratio predicted from starting conditions were highly significant from the F2 onwards in all replicates. Allelic frequencies followed a similar pattern in both experiments (results not shown).

4.3.4 EE vs Evida relative fitness comparisons

The frequencies of EE and EVida3 transgenic alleles relative to the wild-type Mopti allele were formally compared using Logistic regression on the combined frequency data of the 3 replicates. As expected, transgenic allele frequencies were significantly higher in experiment 1 than in experiment 2 (Logistic regression LR: $n= 2880$, df= 1, $\chi^2= 77.6$, $P< 0.001$) and varied significantly between generations (Logistic regression: $n= 2880$, df= 4, $\chi^2= 65.5$, $P< 0.001$, figure 4.3.3). Breaking down the analysis by generation showed that there was no significant difference in transgenic allele frequencies between the two experiments in generations F1 (Logistic regression: $n= 576$, df= 1, $\chi^2= 0.0$, $P= 1.000$) and F2 ($\chi^2= 0.12$, $P= 0.734$, figure 4.3.3). However, from generation F3 ($\chi^2= 5.4$, $P= 0.020$), the frequency of the EE docking construct was significantly higher than that of the EVida3 cassette ($P< 0.001$ in both F4 and F5 generations, figure 4.4.3).

4.3.5 Life stages-specific fitness costs (F2-F5)

Analyses of stage-specific fitness of F2-F5 for the 3 replicates combined showed that was no significant reduction in fitness of the EE and EVida3 alleles and relative to the wild type from adults to the next generation's L1 larvae (Kruskal-Wallis: n=12, df= 2, $P= 0.411$) and during development from pupae to adults (Kruskal-Wallis: n=12, df= 2, $P= 0.053$, Figure 4.3.4). However, allelic fitness relative to the wild type was significantly reduced in both the phase 1 EE and phase 2 EVida3 strains during larval development (Kruskal-Wallis:
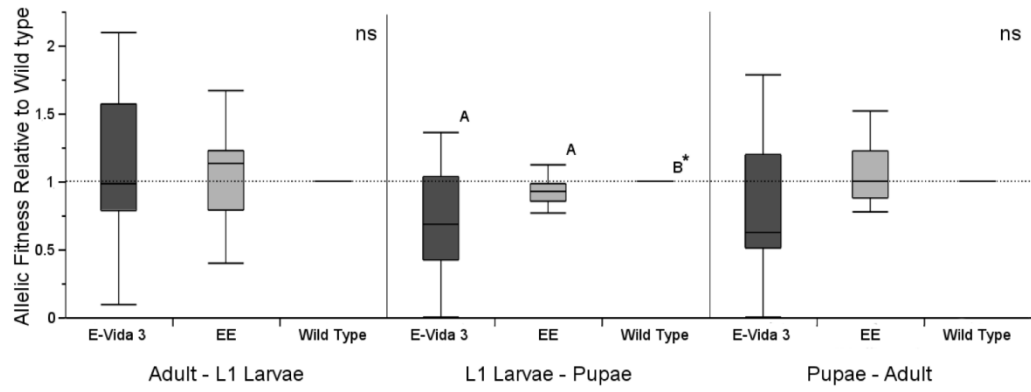
Figure 4.3.4 - Allelic fitness for transgene alleles relative to wildtype from three developmental periods. Boxplots were median, quartiles and min-maximum values. The significance levels of a pairwise Wilcoxon test are indicated - ns, not significant, *: $P > 0.05$, ** $P > 0.01$, *** $P > 0.001$.
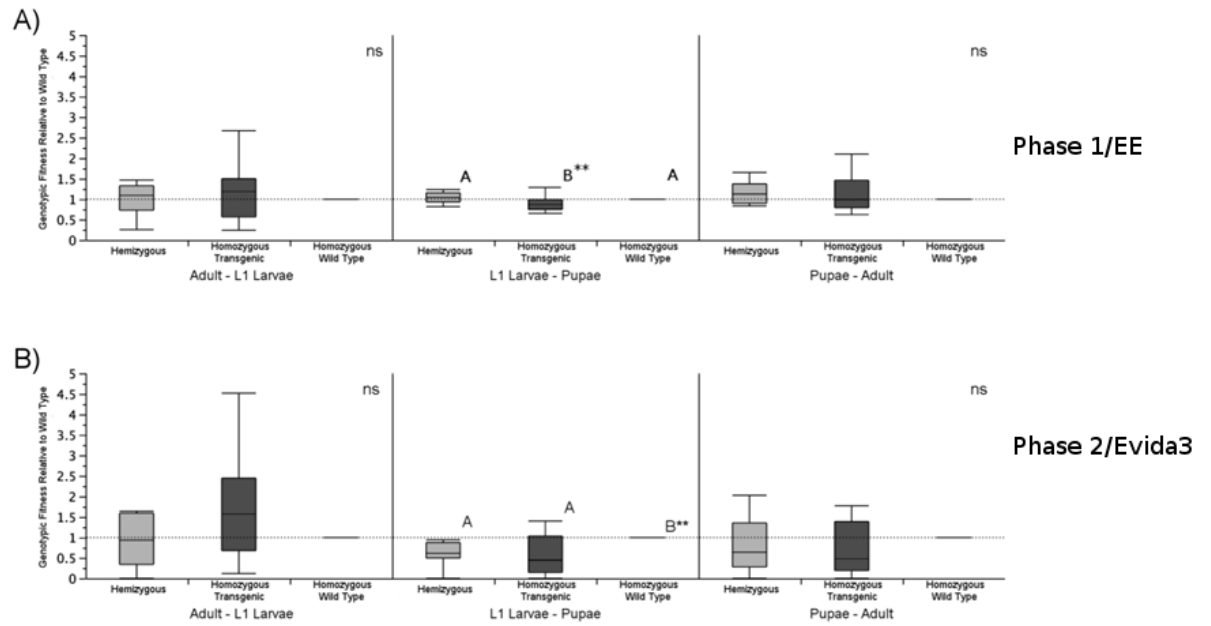
Figure 4.3.5 - Genotypic fitness for hemizygotes and transgenic homozygotes relative to homozygous wild type over three developmental periods. In (a) The fitness of the homozygous Phase 1 EE and hemizygous genotype were compared to wild type homzygotes (Experiment 1); in (b) The fitness of the homozygous Phase 2 Evida 3 and hemizygous genotype were compared to wild type homzygotes (Experiment 2); Boxplots were median, quartiles and min-maximum values. The significance levels of a pairwise Wilcoxon test are indicated - ns, not significant, *: $P > 0.05$, ** $P > 0.01$, *** $P > 0.001$.

n=12, df= 2, *P*= 0.024, figure 4.3.4).

In experiment 1 (EE vs Mopti) no significant differences in genotypic fitness relative to the homozygous wild type were found in hemi- or homozygous transgenic genotypes from the adult to L1 larvae stages (Kruskal-Wallis: n=12, df= 2, *P*= 0.463) or pupae to adults (Kruskal-Wallis: n=12, df= 2, *P*= 0.432, figure 4.3.5). The fitness of individuals during development from L1 larvae to pupae was not significantly reduced in hemizygotes relative to homozygous wild-types (Wilcoxon: n=12, df= 1, *P*= 0.257), but homozygous phase 1 EE transgenic fitness was significantly lower than that of the wild-type (Wilcoxon: n=12, df= 1, *P*= 0.008) and hemizygous genotypes (Wilcoxon: n=12, df= 1, *P*= 0.040, fig4.3.5).

In experiment 2 (EVida3 vs Mopti) there were again no significant differences in genotypic fitness relative to the wild type from adult to L1 larvae (Kruskal-Wallis: n=12, df= 2, *P*= 0.565) and pupae to adult stages (Kruskal-Wallis, n=12, df= 2, *P*= 0.398) developmental periods (Figure 4.3.5). However during development from L1 larvae to pupae the relative fitness of both the hemizygous (Wilcoxon: n=12, df= 1, *P*< 0.001) and homozygous EVida3 transgenics (Wilcoxon: n=12, df= 1, *P*= 0.008) were significantly reduced (Figure 4.3.5).

4.4 Discussion

We assessed the fitness of two *Anopheles gambiae s.s.* transgenic lines recently developed using a two-phase targeted genetic transformation system. The experimental design enabled us to detect initial strain fitness differences, assortative mating and hybrid vigor - all factors important in a future field-release scenario. In addition, the potential fitness costs of the unloaded Phase 1 and loaded Phase 2 genetic constructs, independent of the

strains' original genetic backgrounds, were estimated at the 1st instar larvae, pupae and adult stages over the next 5 generations.

This is only the second study to consider both allelic and genotypic frequency over time for an AMP-carrying strain of *Anopheles* (Li *et al* 2008), and the first to consider these in an *An. gambiae* strain. Furthermore, whilst a number of studies have considered the relative performance of the same transgenic cassette in different genomic loci (e.g. Amenya *et al* 2008) this is the first study to consider the long term stability and relative fitness effects of two different transgenes at the same genomic loci; this allowed us to control for a number of potential confounding effects such as inbreeding depression, gene hitchhiking and general genetic background effects by comparison of the relative performance of these two transgenic strains to a baseline wild-type strain. Finally,sampling at multiple, key developmental stages within each generation over the course of this experiment has allowed us to determine not only *if* there are fitness effects suffered by our transgenic strains, but also *when*; an important consideration when optimizing a potential transgene.

When we considered the performance of the unloaded, phase 1 transgenic cassette (EE, experiment 1), we found that, over 10 generations, it was stably integrated into a mixed population and achieved HWE in all replicates. Whilst we observed no evidence for the potential confounding effects of differential fitness - in terms of male mating success, adult survival and female fertility - between Mopti and EE prior to recombination (F0-F1), and of hybrid vigour in the F1, we did observe a deficiency in hemizygotes in the F1 indicating some assortative mating. However this phenomenon had no effect on the outcome of the experiment, as in successive generations from F2 onwards the frequency of hemizygotes normalised to be consistent with the predictions of Hardy-Weinberg - which assumes random mating in a population. Despite the overall stability of the transgene,

when we looked in detail at the relative genotypic fitness of homozygous EE, hemizygotes and homozygous Mopti at three key developmental stages, we found that during larval development there was a significant reduction in fitness in homozygous transgenic individuals. As this fitness cost was only observed in homozygotes and only at this developmental stage, it is unlikely to be due to expression of the ECFP phenotypic marker which is expressed co-dominantly and throughout all life stages - although an overdominance effect cannot be completely ruled out (Liu *et al* 1999). Thus, the most likely explanation is that this effect is a result of a recessive, deleterious gene linked to the transgene insertion position and selected for during initial screening. Nevertheless this fitness cost, whilst observed consistently in all replicates, did not affect the eventual outcome of the experiment over time, as the effects were ameliorated by higher (but not statistically significantly) fitness relative to wild-type at other stages. Previous studies, albeit only considering adults each generation, have also found similar, recessive fitness effects in otherwise stable transgenic strains. For example, despite reporting a transgenic strain of *An. stephensi* expressing the AMP SM1 being stable in mixed transgenic and non-transgenic cage invasion experiments (Moreira *et al* 2004), a later study investigating transgenics from the same strain, detected a homozygous fitness load (Marelli *et al* 2007),

In the second experiment we investigated the performance of the phase 2 AMP-loaded transgene cassette (EVida3) and found that within 10 generations the transgene could not be detected either visually or through PCR analysis, in any of the 3 replicates, this despite observing that pre-recombination fitness parameters (the combination of F0 male mating success, adult survival and female fertility) were significantly higher than the Mopti. Although we did not observe the potential confounding effect of assortative mating (no hemizygote deficiency in the F1) we did observe a significant, immediate decrease in absolute fitness in both F1 homo- and hemi-zygotes. Due to the drop in fitness of both

homo- and hemizygous groups it was impossible to determine the effects (if any) of heterozygosity. Despite the rapid reduction in the frequency of the EVida3 transgenic cassette between the ensuing 10 generations, when we considered relative genotypic fitness within each generation we found that significant fitness costs were confined to the larval development sample. In contrast to experiment 1, however, significantly reduced fitness was observed in both homozygous and hemizygous individuals. The fact that the fitness cost was observed in hemizygote E-Vida3 contrasts with other studies such as Li *et al* (2008). This study considered long term transgene stability on *An. stephensi* expressing the AMP SM1 under the control of the *Agvg* promoter and found that whilst hemizygotes persisted at high frequency (~0.4) in the cage invasion populations, homozygote transgenics were found at very low frequency (<0.1) suggesting a recessive fitness load..

Clearly the fitness costs imposed by the EVida3 construct at the larval stage, cannot be wholly explained by the fitness costs observed in the first experiment in homozygous larval EE - although these were likely to contribute to the much lower fitness of homozygous larvae. One possible explanation for this fitness cost is leaky expression of the vida3 AMP during larval development. Despite being controlled by the *Anophles gambiae carboxypeptidase* promoter, which drives expression in midgut tissues post blood meal in adult females, there is evidence that the position of transgene insertion can alter the timing, intensity and tissue specificity of the expression of transgenes linked to promoters with otherwise predictable expression profiles due to epigenetic effects (Amenya *et al* 2010). Alternatively, confinement of fitness costs to larval development may indicate that they are due to reduced maternal egg-investment caused by increased post-blood meal stress due to expression of the transgene. This could affect blood-meal utilization and egg development leading to reduced fitness in larvae being independent of the larval genotype but instead dependent on the maternal genotype. Additionally, as with our EE/Mopti

comparison, we cannot rule out dose-dependent toxicity of phenotypic markers: Liu *et al* (1999) observed increased apoptosis in *in vitro* cell lines carrying GFP and EGFP plasmids, and subsequent studies have indicated that prolonged excitation of fluorophores can increase the incidence of active oxygen species in neurones *in vitro* (Dixit and Cyr 2003) and interrupt post-translational polyubiquitination in mice *in vivo* (Baens *et al* 2006). Finally it is possible that there is an independent deleterious effect caused by transgenic insert size (EE ~4kb, E-Vida ~11kb). There is some evidence from studies in *Drosophila* that fitness was reduced in individuals carrying larger (non-coding) transgenic inserts (Kaiser *et al* 1997) relative to those carrying a smaller non-coding insert. Furthermore, transformation efficiency is widely reported to be inversely proportional to insert size in Drosophila (Venken et al 2006, Meredith et al 2011) which may indicate that larger transgenic constructs induce dominant deleterious effects through their size alone.

Whilst it is disappointing that the EVida3 is uncompetitive and thus unlikely to be a strong candidate for a future transgenic release - despite its demonstrated, transgenically-induced refractoriness to *Plasmodium* infection (Meredith *et al* 2011). It is an important proof of concept of the power of the site-specific two-stage transformation process. Furthermore, in the EE line, we have identified a fit, competitive base on which to build, test and evaluate future phase 2 transgenic lines and a powerful tool for investigating the mechanics of transgenesis and how it affects the physiology of transformed mosquitoes.

5. The effects of genetic background and larval rearing conditions on the survival, mating success and assortative mating behaviour under semi-field conditions

Abstract

Anopheles gambiae, the main vector of malaria in Africa, is characterized by a vast geographical range and complex population structure. The co-occurrence of reproductively isolated cryptic forms maintained by strong assortative mating in many populations poses unique conceptual and logistical challenges for programs aiming to release sterile or genetically-modified male mosquitoes in order to decrease malaria incidence. Whether mass-reared *Anopheles gambiae* males of a given cryptic taxa can successfully compete against their wild counterparts to inseminate females and whether they would mate assortatively is still unknown and yet crucial to such approaches.

Here, the independent effects of genetic and environmental factors associated with laboratory rearing on male and female survival, mating success and assortative mating were evaluated in semi-field enclosures in the Mopti form of *Anopheles gambiae*. Males and females from a laboratory strain as well as the progeny of field-collected individuals reared at the larval stage in the laboratory exhibited significantly lower survival and mating success than field progeny reared outdoors. However, rearing laboratory progeny outdoors did not result in improved survival or mating success. Importantly, laboratory individuals reared indoors were unable to mate assortatively, whilst field progeny reared either outdoors or in the laboratory, as well as laboratory progeny reared outdoors all mated significantly assortatively.

These results highlight the importance of genetic * environment interactions for the development of *Anopheles gambiae's* full mating behavioral repertoire and underlines the challenges this creates for mosquito-release vector control strategies.

5.1 Introduction

The direct fitness effects of sterilisation or transgenesis notwithstanding (Grover *et al* 1976, Marrelli *et al* 2006, Chapters 3, and 4), one of the primary factors affecting the competitiveness and fitness of a release-candidate strain is that of the effect of colonisation. During the process of establishing a new laboratory colony, the mosquito population undergoes at least one, and possibly several genetic bottlenecks due to the strong selection pressures imposed on the newly colonised strain as it adapts to the conditions of the insectary. This selective pressure is primarily faced by females due to major differences in bloodmeal delivery - usually through an artificial membrane system and oviposition site. A small starting population can further compound this problem by increasing the incidence of consanguineous mating. Norris and colleagues (Norris *et al* 2001) reported an 8-fold decrease in allelic richness, and a 3.5 fold decrease in heterozygosities in laboratory populations of *An. gambiae s.s.* when compared to field samples. This loss of genetic diversity can lead to a consequent loss of relative competitiveness caused by inbreeding depression, an effect which can be magnified outside of the uniform insectary environment; although the opposite can also be true (Armbruster *et al* 2000). As the colonised strain adapts to its new environment there is also a strong possibility that it will develop 'aberrant' swarming and mating behaviour in response to new environmental conditions (Jones and Gubbins 1978, Marchand 1985): the 12hr:12hr light:dark cycle and lack of a crepuscular transition typically employed in insectaries can cause significant changes in the time of both male swarming behaviour and peak female mate-seeking behaviour - both important determinants of mating success in members of the *An. gambiae* species complex (Charlwood and Jones 1980). In addition the lack of natural horizon or swarm markers in laboratory enclosures can cause further

divergence in mating behaviour between a colonised strain and the wild population it is derived from (Marchand 1985).

These effects can erect significant pre-zygotic barriers to mating between lab and field populations - particularly outside of the laboratory. This phenomenon was observed during a large-scale field release of sterile male *Culex tritaeneoryhchus* in India in 1977 (Baker *et al* 1979). In this trial, Baker and colleagues released large number of *C. tritaeneorynchus* males carrying a complex, sterility inducing chromosomal aberration alongside marked females from a lab strain into a village near Lahore, India. Subsequent population sampling from this area revealed that the released sterile males were highly competitive in terms of mating with the laboratory-reared females, but non-competitive with females from the local wild population. The sterile male strain's genetic background was primarily derived from a 6 year+ colonised wild-type *C. tritaeneoryhnchus* strain. This suggests that in this case, laboratory colonisation has induced significant assortative mating behaviour. In an attempt to ameliorate this, a subsequent study outcrossed the sterile male strain to the F1 progeny of field-captured females (Reisen *et al* 1980), however, despite the introgression of genetic material from the field sample, significant assortative mating was again observed. This suggests that the traits determining the observed assortative mating behaviour are either selected for very rapidly (i.e. within one generation) or there are both genetic (in terms of selection for lab conditions) and environmental (in terms of the difference in field and laboratory rearing conditions) elements determining mating behaviour. This latter hypothesis is supported by similar studies carried out between 1977 and 1981 assessing the competitiveness of radiosterilised *C. tarsalis* in California. An initial study had determined that male *C. tarsalis* collected as pupae from the field and radiosterilised as adults were competitive for wild females and able to induce sterility in a target population (Reisen *et al* 1981). However, field collection of pupae was not able to

provide sufficient mosquitoes for a full scale trial. For this reason, in a subsequent study, Reisen and colleagues established a laboratory colony from ~3000 field caught females which were maintained under insectary conditions for 9-16 generations (Reisen *et al* 1982). Upon the release of radiosterilised males from this population it was found that significant assortative mating behaviour had developed during the short time the population had been colonised. In large-cage mating competition experiments lab colonised sterile males were found to be super-competitive for lab colony females (68% of mated females mated by sterile lab males) and uncompetitive for field females captured from the field as pupae (21% mated by sterile males). Interestingly in large cages where field and lab males competed for lab-reared female F1 progeny of field capture females (i.e. with 'field-type' genetic background, but a 'lab-type' environmental background), mating competitiveness was found to be roughly intermediate, reinforcing the hypothesis that both genetic and environmental factors affect mating behaviour. The potential for interaction between the genetic background of a strain and the larval rearing environment was highlighted in a recent study by Dao *et al* 2008. Here the investigators were assessing indoor mating behaviour in Malian populations of *An. gambiae* s.s. As part of this study they examined assortative mating behaviour between M- and S-molecular form groups using mark, release, recapture (MRR) experiments. These experiments were carried out within typical rural single room houses using the lab-reared F1 progeny of field caught females of both molecular forms. In this setting, Dao and colleagues observed a breakdown in assortative mating behaviour between forms that was inconsistent with the previously observed low rate of between-form hybridisation reported in Mali (Tripet *et al* 2001). The authors concluded that non-specific indoor mating could be an important source of gene flow between these otherwise reproductively isolated forms. However they acknowledge that, as their experimental mosquitoes were raised under laboratory conditions, they could

not rule out the interaction of environmental confounding factors on the mating behaviour of these strains. For this reason elucidating the potential impact of larval rearing environment on assortative mating behaviour was a key objective of our study.

The baseline performance of a strain of interest can be assessed most easily by using smaller scale, laboratory-based studies. These preliminary experiments provide an answer to the question of whether a given approach functions as expected, allowing us to determine, for example, the effectiveness of sterilisation protocol (Helsinki and Knols 2009), or the penetrance of a transgenic lethality (Harris *et al* 2011) or resistance phenotype (Meredith *et al* 2011). Secondly, we can investigate the compatibility of a manipulated lab strain with any given target field strain. As a precursor to a SIT intervention, Munhenga and colleagues recently demonstrated that long-term colonised populations of *An. arabiensis* mated readily with the F1 progeny of field captured individuals from a target site under laboratory conditions (Munhenga *et al* 2011). Although it is worth noting that the propensity for two strains to mate in a laboratory setting does not guarantee they will mate in the field, as evidenced by the readiness which *An. arabiensis* and An. gambiae mate under insectary conditions, but not in the field (Davidson 1969), despite occurring in sympatry (Marchand 1984). Furthermore, lab-based studies can be a useful for assessing the relative competitiveness of two strains created using different sterilising or transgenic protocols, or expressing different transgenes (Helsinki and Knols 2009, Chapters 3 and 4). This approach can highlight gross differences in competitiveness induced by differing approaches, and allows early elimination of underperforming strains. Finally, lab-based studies can provide a broad estimate of the potential field performance of a manipulated mosquito strain. However, the accuracy of lab-derived estimates of field performance is notoriously problematic: several trials of sterile mosquitoes in the 1970's were partially undermined by large differences in predicted and actual competitiveness of

released males. In one notable example, despite being 100% competitive in the laboratory, under field conditions both chemosterilised and cytoplasmically incompatible male *C. pipiens* Fatigans exhibited a level of competitiveness only between 25 and 50% relative to their performance under laboratory conditions (Grover *et al* 1976). In a separate study, a North American trial of heterozygous translocated *C. tarsalis* observed a drop in field competitiveness from 75% - 25% after 2 years under laboratory conditions (Milby 1980). This tendency for lab studies to overestimate field competitiveness has caused some researchers to question their validity in this context (Ferguson *et al* 2008, Lee *et al* 2011), and has driven the expansion of purpose built semi-field facilities throughout the tropics (Knols *et al* 2001, Ferguson *et al* 2008). A large (the largest) semi-field facility was recently completed at the Ifakara Health Institute, Kilombero, Tanzania, totalling over 625 m^2 in size and comprising 4 experimental and insectary chambers (Ferguson *et al* 2008), researchers have since reported the establishment of a self propagating (>20 generation) population of a local *An. arabiensis* strain (Ng'habi *et al* 2010).). Meanwhile, in Sudan, Hassan and colleagues, working in a smaller permanent semi-field facility (144 m^2), have assessed the mating competitiveness of radio-sterilised *An. arabiensis* as part of a wider study evaluating rearing, transport and distribution methodologies for a future sterile insect release. In this setting the sterile strain (taken from a 68th generation insectary colony) was able to compete with the non-sterile males - field-caught as 3-4th instar larvae and pupae - for field caught virgin females. The researchers estimated the mean competitiveness index (CI) of the sterile strain as 0.71 i.e. sterile males inseminated 29% less females than their non-sterile counterparts (Hassan *et al* 2010), although they did not confirm this result by investigating whether recaptured females produced viable offspring following a blood meal. More recently, in Malaysia, Lee and colleagues reported the first semi-field experiments using a transgenic mosquito population. In this case a strain of *Aedes aegypti*,

expressing a repressible dominant lethality phenotype. Their results suggest that the Mexican-derived transgenic strain was compatible with a local lab-strain of *Ae. aegypti*. In terms of competition they found no deviation from the result expected using a hypothetical, 100% competitive strain (Lee *et al* 2011).

In this chapter we describe experiments designed to quantify the main and interactive effects of genetic and environmental factors typically associated with colonisation and laboratory rearing. Using a small-scale semi-field system we investigated the effects of genetic background and larval rearing conditions on the survival and mating success of M-form *Anopheles gambiae* s.s. under semi-field conditions. We also investigated how the interaction of genetic background and environmental conditions affected assortative mating behaviour between M- and S- form *An. gambiae* and discuss this in the context of findings elsewhere. These data offer insights into both the effect of long term colonisation and larval rearing conditions on the performance of *An. gambiae* in a field-like setting. Both factors that should be taken into account in future genetic control strategies and ecological experiments.

5.2 Materials and Methods

5.2.1 Facilities: Laboratory/Insectary

A key brief of the Wellcome Trust grant which funded this study was to establish facilities purpose built for state of the art transgenic and ecological entomological investigations, coupled with parasite culture facilities. To this end laboratory space at the Malaria Research and Training Centre, Universite de Bamako, Bamako, Mali was refitted into a modern, category 3 biosecurity molecular entomology and parasitology lab; fully equipped for insect transgenesis and parasite culture. A large, adjoining insectary was constructed with space for ~100 larval rearing trays in the main rearing area. One wall of the insectary

was partially composed of glass bricks providing a natural day-dusk-night-dawn light cycle. Air temperature was maintained at a constant 27±2°C and relative humidity was kept at 70±5%. The water temperature in larval growth trays was 22.5±0.5°C - although this was not actively regulated. Within the insectary, a screened-off area provided a dedicated, secure space for *Anopheles*/*Plasmodium* infection studies; and a dark room equipped with a digital, programmable light cycler allowed us to manipulate the larval rearing and adult photoperiod. The laboratory and insectary facility was sealed off from the rest of the MRTC facility by a magnetically interlocking door/antechamber system with a mechanical code lock. Entry into the insectary areas from within the laboratory was controlled with a second pair of magnetically interlocking doors.

5.2.2 Facilities: Field Site

We established a small field site on the outskirts of the village of N'Gabakoro Droit, consisting of four, 4×4×2 m and one 2×2×2m custom-made plastic netting (1mm weave mesh) enclosures (HowieNet) supported within a timber frame and covered with a tarpaulin roof (GalaTent). Each enclosure possessed an antechamber to prevent unwanted escapees. Three of the 4×4×2m enclosures were used exclusively for mating and survival studies, with the fourth acting as either an insectary or an additional experimental enclosure depending on circumstance. The smaller 2×2×2m enclosure was used as an insectary and field lab. Each mating enclosure was provided with the following: a ~3cm depth floor-covering of coarse fluvial gravel, kept moist to enhance humidity; 3 large (30×80cm) cylindrical clay pots with a ~10cm deep layer of wet gravel, providing a shaded and humid microclimate; and two large leafy plants to provide shade and additional humidity through transpiration. Additionally, each plant had several cotton wool pads

attached. These were soaked with 10% sucrose solution to provide an energy source during each experimental period.

The three dedicated mating enclosures were arranged in a straight line running approximately on a North-South axis, with the fourth, and the smaller enclosure, offset slightly to the West. This minimised variation in the level of sunshine and shade each enclosure received. The enclosures were bordered on the North by a breezeblock wall, to the East and South with marshy scrubland and to the West by a dirt road. The whole site was removed ~100m from the village buildings.

Air temperature at the field site ranged from daytime highs of 34-42°C to lows at night of between 24 and 28°C. Water temperature in the larval growth trays ranged from 24-32°C over the course of a typical 24 hour period, with an observed high of 36°C. Relative humidity within the field enclosures was between 40 and 80% and between 60 and 80% inside the emergence cages. The temperature within the clay pot refuges in each mating enclosure was a consistent 4-5°C below the daytime ambient air temperature. RH was between 60 and 80% within the pots.

5.2.3 Mosquito Colonisation and Field Captures

An initial colony of Mopti, M-form *Anopheles gambiae* s.s. was established from the F1 progeny of field-captured gravid females. Collections were carried out in 2008 from the village of N'Gabakoro Droit, Koulikorou, Mali (12°39'46"N, 7°50'34"W) and maintained in the insectary facilities at the MRTC. At the start of the experiment proper, this lab colony had reached generation F42, and was well-adapted to laboratory culture. To provide field F1 progeny during each experimental procedure, field capture of gravid females was carried out both from N'Gabakoro Droit and the village of Bankoumana, West of Bamako, Kati District, Mali (12°12'2"N, 8°15'54"W). Thus experimental mosquitoes were either

drawn from a well-established laboratory population, or were the F1 progeny of 'fresh' field captured individuals.

Gravid females were captured between 6 and 8 AM, by mouth aspiration in the homes and outbuildings of N'Gabakoro and Bankoumana. Captured mosquitoes were then transferred to a 5l adult enclosure and provided with a 10% sucrose solution and H_2O *ad libitum* and transported by car to our insectary at the MRTC.

5.2.4 Genotyping of Field F1 Broods

48 hours post field-capture; individual females were transferred by aspiration to 8cm×2cm cylindrical glass tubes sealed with a small square of mosquito netting. After approximately 2h, 5ml H_2O was carefully transferred into the tube by slow pipetting so as to minimise the stress to the mosquito and chance of inundation. These individual oviposition tubes were transferred to a rack and left to allow the mosquitoes to oviposit. Tubes were checked for eggs or the death of the individual mosquito twice daily, until all were accounted for. Each individual egg batch was transferred to 15ml H_2O in a 25ml plastic weigh boat and provided with a suspension of yeast cells (Liquifry, Tetra) until genotyped. Females which laid eggs were captured with forceps and transferred to a 1.5ml microcentrifuge tube for DNA extraction and PCR genotyping. DNA extraction was carried out using a modified DNAzol (Invitrogen) kit protocol as described previously. Genotyping to species and sub-species level was performed by PCR/RFLP using the DreamTaq (Fermentas) PCR kit and *HhaI* restriction endonuclease (Promega), as described by Fanello *et al* 2002 (see chapter 2). Once successfully genotyped, M-form *Anopheles gambiae* s.s. broods were pooled and either prepared for transport to the field site or separated into growth trays as per the lab insectary conditions described below. S-form broods were further characterised to determine if they were Savannah or Bamako chromosomal form using the diagnostic PCR

based on *J* inversion polymorphisms developed by Coulibaly *et al* (2007). S-form, Savannah broods were pooled and prepared for transport to the field site. Any S-form Bamako or *An. arabiensis* broods were killed by freezing and securely discarded.

5.2.5 Larval Rearing

Eggs were obtained from the 'lab' genetic background cohort by blood feeding from a human volunteer (DP) and, after allowing 48hrs for egg development, providing an oviposition pot consisting of a polystyrene cup (8cm diameter, 3cm depth) containing moistened filter paper. After a further 48h the ovipostion cup was removed and the newly laid eggs suspended in 1l H_2O to hatch. Once genotyped (above) the field F1 broods obtained by field capture were pooled by molecular form in 1l H_2O. 'Lab' and 'Field' genetic background larvae to be reared in the semi-field system were transported by car to the field site from the MRTC suspended in 1l of H_2O in a glass Duran bottle.

In both the lab and the SFS, L1/early L2 larvae were separated into 30×15×4cm plastic trays and suspended in 1l of H_2O (~2.22cm depth) at a larval growth density of 200 larvae/litre. The number of trays set up for each group varied depending on the number of adults required for each experiment. In the SFS insectary, larval trays were stored at or near ground level in an effort to provide a natural horizon for developing larvae. During development through the L1-L4 larval instars, larvae were supplied with, initially, a yeast cell suspension (Liquifry, Tetra) followed by an optimised regimen of ground fish food (Tetramin, Tetra). Upon pupation, pupae were segregated by sex using a binocular dissecting microscope (Leica) and transferred by aspiration to small polystyrene cups. In the lab, pupae were left to eclose in standard rearing cage made of a 5l cylindrical polypropylene bucket (~20.5cm height×20cm diameter) with a sleeved side opening and a mosquito netting top. In the SFS pupae were transferred to a large 50×50×100cm steel-

framed netting enclosure with a sleeved side-opening. Adult flies in both the lab and SFS insectary were supplied with a 10% sucrose solution and H_2O *ad libitum*. In the SFS, additional shade and humidity were provided by covering the top of the emergence enclosures with a layer of wet cotton wool beneath wet, rough-spun cotton towels. To maintain a steady supply of experimental mosquitoes, lab colonies were blood fed and field collections were carried out every 2 days.

5.2.6 Experimental Design

In our first experiment, we set out to test the effects of colonisation (genetic background) and the larval rearing and adult maintenance conditions (environmental background) on survival and mating success in our Mopti M-form populations from N'Gabokoro. Mosquitoes were segregated into two groups in relation to their genetic background: generation F42+ lab colonised (lab) and F1 field captures (field). Additionally, our lab and field facilities allowed us to raise mosquitoes in one of two different larval rearing and adult maintenance conditions: lab-reared (lab) and semi-field enclosure-reared (field). These genetic and environmental combinations resulted in a total of four experimental groups (genetic/environmental background respectively): Field/Field, Field/Lab, Lab/Field and Lab/Lab. In all experiments males and females from each of these 4 treatment groups (treatment males/females) were mated with females and males from the Field/Field treatment group (field females/males).

A second experiment combined expanded on this with the additional factor of assortative mating behaviour between M- and S- molecular forms observed in Malian *Anopheles gambiae* s.s. populations. In this experiment, our four M-form Field/Field, Field/Lab, Lab/Field and Lab/Lab males and female treatment group were mated with

'field' females and male samples made up from a 1:1 mix of M-form Field/Field and S-form Field/Field individuals.

5.2.7 Experimental Procedure

Effect of colonization and rearing conditions on survival and mating success

This experiment was split into two sections and carried out over both the 2010 and 2011 field seasons. In the 2010 season we carried out mating experiments crossing Field mosquitoes with Treatment mosquitoes from the Field/Field, Field/Lab and Lab/Lab cohorts. In 2011 we investigated the Field/Field, Lab/Field and Lab/Lab cohorts. This design allowed us to optimize the use of the 3 enclosures available in 2010 for the project whilst keeping a balanced design for statistical analyses (see below). The experimental procedure was identical between each field season as was a follows: A sample of 50, 3-5 day post-eclosure, virgin adults were collected at random by aspiration from each of the three treatment groups investigated (see above; 2010: Field/Field, Field/Lab and Lab/Lab; 2011: Field/Field, Lab/Field and Lab/Lab) and placed in standard 5l adult enclosures. Lab-reared samples (Lab/Field, Lab/Lab) were transported to the SFS site by car. Three samples of 50 Field/Field individuals of the opposite gender were similarly prepared at the SFS. All enclosures were provided with water and sugar solution and left to acclimatise in the SFS insectary enclosure for 2-3 hours. At approximately 1700hrs local time each treatment cage was paired up with a random field cage and the mosquitoes released into the large experimental mating enclosures (above). After approximately 40 hours (2 nights) surviving individuals were recaptured from within their enclosures using a large backpack aspirator (JW Hock & co.). A total of three sweeps over ~2 hours were carried out to maximise the number of individuals recaptured, and the experimental enclosures reset for the next cross. Each cross and the reciprocal was replicated twice for a total of 3 replicates

and 18 experimental crosses in each field season (36 comparisons in total). The experimental enclosure used for each cross was rotated between replicates to ameliorate the effects of any environmental variation between enclosures.

Recaptured individuals were transported back to our lab at the MRTC. The body size of both male and female individuals was estimated by measuring the size of the wing from the posterior anal cell margin to the tip of radial vein 3 at 20x magnification. Females were stored at -20°C in 70% ethanol for at least 24h and then dissected to ascertain their mating status based on the presence of absence of a sperm bundle within the spermathecae.

Effects of colonization and rearing conditions on assortative mating

This experiment was carried out in its entirety in the 2011 field season and a fourth enclosure allowed us to run all four treatment groups (Field/Field, Field/Lab, Lab/Field and Lab/Lab) in parallel. In this experiment, samples of 50, 3-5 day post-eclosure, virgin adults were collected at random by aspiration from each of the four treatment groups. Four samples of Field/Field mosquitoes of the opposite sex, composed of 25 M-form and 25 S-form Field/Field individuals, were also prepared. As previously, lab-reared samples were transported to the SFS by car and allowed 2-3 hours to acclimatise before each cross was launched. As before, crosses were released at around 1700hrs local and left within each experimental enclosure for approximately 40 hours before being recaptured by aspiration. Each cross and its reciprocal was replicated twice for a total of 3 replicates and 24 experimental crosses. The experimental enclosure used for each cross was rotated between replicates to ameliorate the effects of any environmental variation between enclosures.

The wing length of all recaptured individuals was measured. The genotype of all survivors was determined by PCR/RFLP (described previously). In addition, recaptured females were stored for at least 24 hours in 70% ethanol at -20 °C, and then dissected to

determine their mating status based on the presence of absence of a sperm bundle within the spermathecae. Further, for treatment females crossed with a mixture of M- and S-molecular form Field/Field males the genotype of the successful male was determined by PCR analysis of transferred sperm. Once isolated by dissection, the sperm bundle was washed and DNA was isolated from it using a modified magnetic nano-particle DNA extraction kit protocol (Invitrogen, ChargeSwitch DNA extraction kit), and the male genotype in terms of molecular form determined by PCR/RFLP.

5.2.8 Data Analysis

Body size

Due to the unequal numbers of recovered mosquitoes in each treatment groups, differences in body size between experimental treatments were tested by randomly sub-sampling 30 wing length measurements from each gender, treatment group and from both the 2010 and 2011 field seasons. Comparison was by T-Test, ANOVA and pairwise Tukey-Kramer Honestly Significant Difference (HSD) tests as appropriate.

Survival and mating success

For the purposes of our analysis, survival was estimated based on the rate of recapture of mosquitoes in each experimental group after 2 nights within an experimental enclosure and mating success was defined as the proportion of recaptured females with identifiable spermatozoa present within their spermathecae. Separate multivariate logistic regression models were constructed to assess the effects of the independent variables of genetic and environmental background, sex, replicate/enclosure and (where applicable) the experimental year on the dependent variables of survival and mating success. In models were all four experimental groups and both years were considered simultaneously,

experimental groups were nested within year. For simplicity, the data from both years was combined to produce bar-plots and other summary statistics.

The presence or absence of significant assortative mating behaviour in our second experiment was assessed by comparison of the observed frequency of 'assortative' mating events (i.e. M-form with M-form) to non-assortative mating events (M-form with S-form) to the frequency of these expected in the absence of significant assortative mating - a 1:1 ratio of assortative:non-assortative mating events. This was assessed using a Chi-square Goodness of Fit test.

All statistical analyses and graphs were constructed using JMP 9 (SAS Institute inc.). Significant differences between replicates were checked in all analyses and reported whenever appropriate.

5.3 Results

5.3.1 Effects of colonization and rearing conditions on body size, survival and mating success

Body size

In each field season, a total of 18 experimental crosses were attempted, utilizing 1800 mosquitoes from six rearing batches each season; each batch corresponding to one reciprocal cross within each replicate. When we considered the mean wing length of each gender within the three treatment groups in the 2010 field season (figure 5.3.1), there was no significant difference in wing lengths between treatments in Field/Field males (2.94±0.06mm) Field/Lab males (2.89±0.07mm) and Lab/Lab males (2.90±0.05mm) (ANOVA, $F_{2,85}$=0.55, P=0.580) and Field/Field females (2.86±0.07mm) Field/Lab females (2.96±0.06mm) and Lab/Lab females (2.90±0.06mm)(ANOVA $F_{2,74}$=2.37 P=0.102).
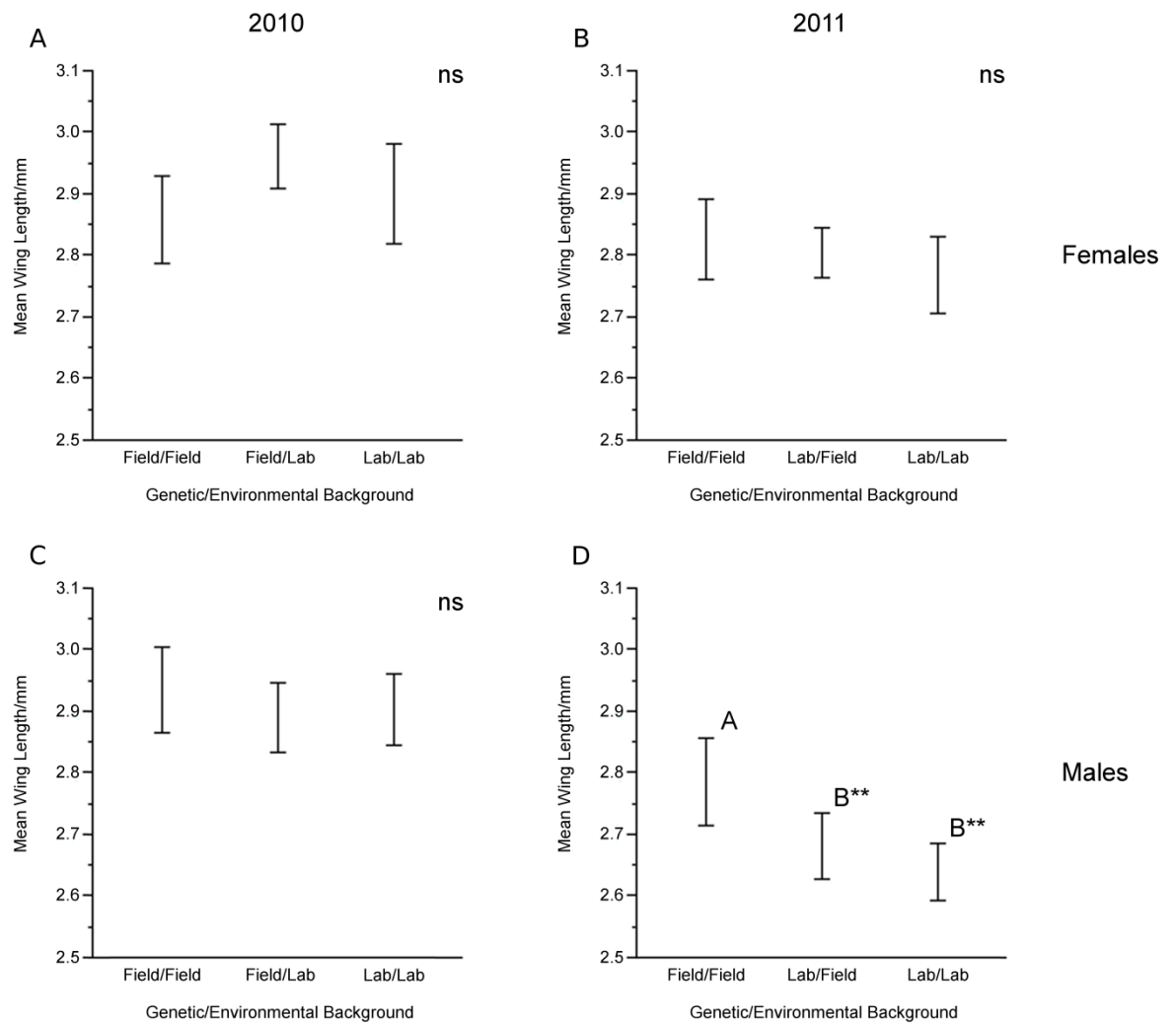
Figure 5.3.1: Comparison of mean wing length between random samples of 30 individuals from genetic/environmental background cohorts, segregated by experimental year and gender. (a) Treatment females, 2010. (b) Treatment females 2011. (c) Treatment males, 2010. (d) Treatment males, 2010. The results of a one-way ANOVA and Tukey *post hoc* test are indicated - ns, not significant, *: P< 0.05, ** P< 0.01, *** P< 0.001. Error bars represent 95% confidence intervals.

When we compared wing lengths in sampled populations from the 2011 season, we found no overall significant difference in wing length between Field/Field females (2.83±0.06mm) Lab/Field females (2.81±0.05mm) and Lab/Lab females (2.77±0.06mm) from each treatment group (ANOVA, $F_{2,87}$=1.13 P = 0.33. In males from the 2011 treatment groups we found that there was an overall significant difference in mean wing length (ANOVA, $F_{2,85}$=7.12, P=0.001) such that both Lab/Field males (2.68±0.06mm, Tukey P=0.032) and Lab/Lab males (2.64±0.05mm, Tukey, P=0.001) were significantly smaller than Field/Field males (2.79±0.06mm).

Survival

Over two wet seasons in Mali, we tested the relative survival of adult male and female *Anopheles gambiae* s.s. in our custom-built, SFS. Mosquitoes were drawn from one of two genetic backgrounds (a well-established (generation F42+) laboratory colony or the F1 progeny of gravid females captured by resting catch from the field) and one of two larval rearing conditions (standardised laboratory conditions or from a population raised in a semi-field insectary within our SFS facility). Males and females from each of the resultant four treatment groups (genetic/environmental background respectively: Field/Field, Field/Lab, Lab/Field and Lab/Lab) were mixed with an equal number of Field/Field mosquitoes of the opposite gender and left to mate in a large (4m×4m×2m) semi-field mating enclosure. Surviving individuals were recaptured after 2 nights.

Survival across all replicates after 2 nights was highest in the Field/Field cohort (55.42±5.81% survived), followed by Lab/Lab (51.67±6.58%) and Field/Lab (48.67±8.09%) with Lab/Field (44.67±8.06%) having the lowest survival overall (Figure 5.3.2). In males, the cumulative survival across all experiments after 2 nights was also

A

Gender	Genetic Background	Environmental Background	Survival (%)	95% CI	Mating Success (%)	95% CI
Male[1]	Field M	Field	42.25	2.8	35.9	6.8
	Field M	Lab	41.33	7.97	25	9.01
	Lab M	Field	24.67	6.99	27.08	12.04
	Lab M	Lab	33	5.02	27.94	7.64
Female[2]	Field M	Field	55.42	5.81	31.03	6.94
	Field M	Lab	48.67	8.09	16.9	8.93
	Lab M	Field	44.67	8.06	16.42	9.11
	Lab M	Lab	51.67	5.58	19.5	6.22

[1] treatment males were crossed with Field M/Field females

[2] treatment females were crossed with Field M/Field males

B

Gender	Genetic Background	Environmental Background	Survival (%)	95% CI	Mating Success (%)	95% CI	Prop Mated Assortatively[3]
Male[1]	Field M	Field	35.33	4.95	36.25	10.75	0.759
	Field M	Lab	9.09	4.03	22.5	13.52	0.667
	Lab M	Field	17.33	6.12	25	15.86	0.625
	Lab M	Lab	32	7.55	35.9	16.75	0.429
Female[2]	Field M	Field	65.33	7.7	16	7.32	0.875
	Field M	Lab	32	7.55	22.92	12.33	1
	Lab M	Field	28.67	7.31	20.45	12.41	1
	Lab M	Lab	61.33	7.91	18.6	8.39	0.563

[1] treatment males were crossed with a 1:1 mix of Field M/Field and Field S/Field females

[2] treatment females were crossed with a 1:1 mix of Field M/Field and Field S/Field males

[3] assortative mating defined as mating with an individual of the same molecular form

Table 5.3.0: Summary data for the effects of genetic and environmental background on survival, mating success and assortative mating behaviour. A) Experiment 1: Survival and mating success in M-form *An. gambiae* **B)** Experiment 2: Survival, mating success and assortative mating behaviour in M-form *An. gambiae* given a choice of M- or S-form mates.
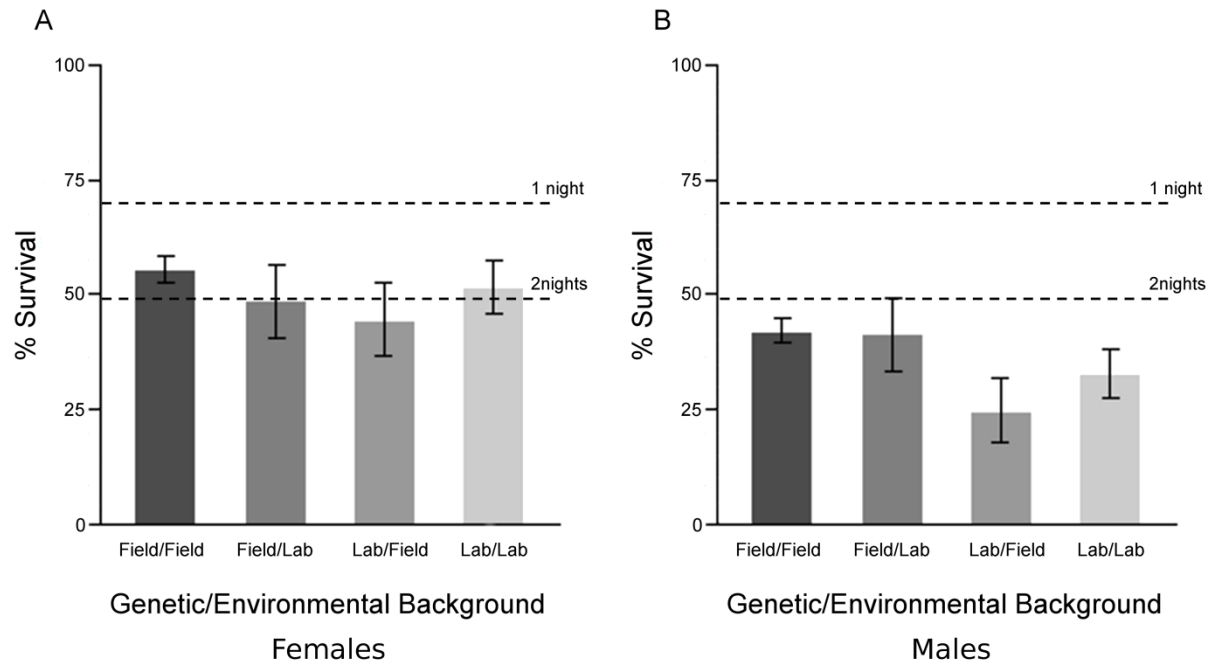
Figure 5.3.2 - Proportion of treatment cohort *An. gambiae* surviving after 2 nights under semi-field conditions, (a) females, (b) males. Error bars represent 95% confidence intervals. Theoretical survival corresponding to a daily mortality of 30% are indicated.

Reciprocal Cross Type:

Treatment Males x Field/Field Females			**Females**
Source	**df**	**LR Chi-square**	***P***
Experiment Year	1	34.93	<0.001*
Gen/Env [Experimental Year]	4	55.84	<0.001*
Enclosure	2	8.07	0.018*
Enclosure*Experimental Year	2	52.45	<0.001*
Enclosure*Gen/Env[Experimental Year]	8	88.09	<0.001*

Reciprocal Cross Type:

Treatment Males x Field/Field females			**Males**
Source	**df**	**LR Chi-square**	***P***
Experiment Year	1	8.04	0.005*
Gen/Env [Experimental Year]	4	19.96	<0.001*
Enclosure	2	25.08	<0.001*
Enclosure*Experimental Year	2	23.28	<0.001*
Enclosure*Gen/Env[Experimental Year]	8	66.93	<0.001*

Reciprocal Cross Type:

Field/Field Males x Treatment Females			**Females**
Source	**df**	**LR Chi-square**	***P***
Experiment Year	1	0.64	0.4221
Gen/Env [Experimental Year]	4	18.8	0.002*
Enclosure	2	51.82	<0.001*
Enclosure*Experimental Year	2	6.81	0.033*
Enclosure*Gen/Env[Experimental Year]	8	56.03	<0.001*

Reciprocal Cross Type:

Field/Field Males x Treatment Females			**Males**
Source	**df**	**LR Chi-square**	***P***
Experiment Year	1	24.07	<0.001*
Gen/Env [Experimental Year]	4	22.13	<0.001*
Enclosure	2	9.19	0.010*
Enclosure*Experimental Year	2	53.1	<0.001*
Enclosure*Gen/Env[Experimental Year]	8	65	<0.001*

Table 5.3.1: Nominal logistic regression (Likelihood-ratio, n = 900) of the effects of Experimental Year, Genetic and Environmental Background and Mating Enclosure on the survival of *An. gambiae* s.s. mosquitoes in a semi-field mating enclosure.

highest in the Field/Field cohort (42.25±2.80% survived), followed by Field/Lab (41.33±7.97%) and Lab/Lab (33.00±5.02%), with Lab/Field (24.67±6.99%) having the lowest survival overall (Figure 5.3.2).

The effect of the different sources of variation inherent to the experimental design and potentially affecting survival (dependent variable) were formally partitioned and statistically tested by constructing a logistic regression model. Model effects were nested by experimental year to account for the two field seasons the experiment was carried out over (experimental year). The effects of the independent variables of genetic background and larval rearing conditions (gen/env) and mating enclosure (enclosure, synonymous with replicate) were modelled along with any interactions between variables. In total, four analyses - corresponding to each gender in each of the two reciprocal crosses (treatment males × Field/Field females, Field/Field males ×treatment females) - were carried out.

After accounting for the effects of experimental year - highly significant in all but one comparison - and enclosure - highly significant in all comparisons - (Table 5.3.1) the effect of larval rearing treatment and genetic background on survival was significant both in treatment males (Logistic Regression LR: n=900 df=4, χ^2=19.96, $P<$ 0.001) and treatment females (LR: n=900 df=4, χ^2=16.799, P=0.002, Table 5.3.1). Unexpectedly, despite accounting for other confounding effects, the genetic and environmental background of 'treatment' males was also a significant factor in the survival of the Field/Field females they were crossed with (LR: n=900, df=4, χ^2=55.84, $P<$0.001), and the same was also true for Field/Field males crossed with females from the four treatment cohorts (LR: n=900 df=4, χ^2=22.130, $P<$0.001)(Table 5.3.1).
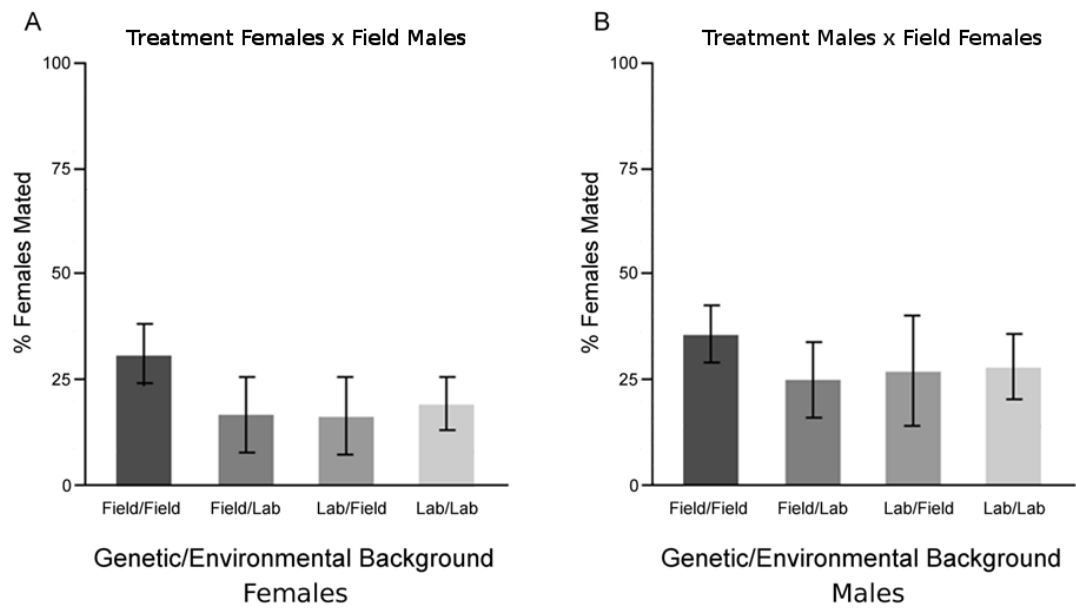
Figure 5.3.3: **Proportion of female *An. gambiae* mated after 2 nights in semi-field conditions.** (a) Treatment cohort females x Field cohort males. (b) Treatment cohort males x Field cohort females. Error bars represent 95% confidence intervals.

Source	df	LR Chi-square	P
Experiment Year	1	0.13	0.72
Reciprocal Cross Type	1	5.64	0.018*
Gen/Env [Experimental Year]	4	14.96	0.005*
Enclosure	2	3.45	0.18

Table 5.3.2: Nominal logistic regression (Likelihood-ratio, n = 942) of the effects of Experimental Year, Reciprocal Cross Type and Genetic and Environmental Background on mating success of *An. gambiae* s.s. mosquitoes in a semi-field mating enclosure. Square brackets indicate effect nesting.

Mating Success

Of 959 surviving females recaptured, 252 were found to have identifiable spermatozoa within their spermathecae after dissection (26.28%). In the 'treatment' male cohort (treatment males × Field/Field females), Field/Field males were the most successful (35.90±6.80% females mated). Followed by Lab/Lab males (27.94±7.64%), Lab/Field males (27.08±12.04%) and Field/Lab males (25±9.01% females mated) (Figure 5.3.3). In the reciprocal experiment (treatment females × Field/Field Males), Field/Field females had been mated most frequently (31.03±6.94%), followed by Lab/Lab females (19.50±6.22%), Field/Lab females (16.90±8.93%) and finally Lab/Field females (16.42±9.11%)(Figure 5.3.3).

We constructed a logistic regression model to assess the factors determining mating success (dependent variable). Model effects were experiment year, genetic and environmental background (nested within experiment year), reciprocal cross type (treatment males × Field/Field females, treatment females × Field/Field males) and enclosure.

In this case, neither experiment year (LR: n=952 df=1, χ^2=0.13, *P*=0.720) nor enclosure were significant factors in determining the proportion of mated females (LR: n=952, df=2, χ^2=0.3.45, *P*=0.18)(Table 5.3.2). Genetic and environmental background had a significant effect (LR: n=952 df=4, χ^2=14.96, *P*=0.005) on mating success, as did reciprocal cross type (LR: n=952 df=1, χ^2=5.64, *P*=0.018)(Table 5.3.2).
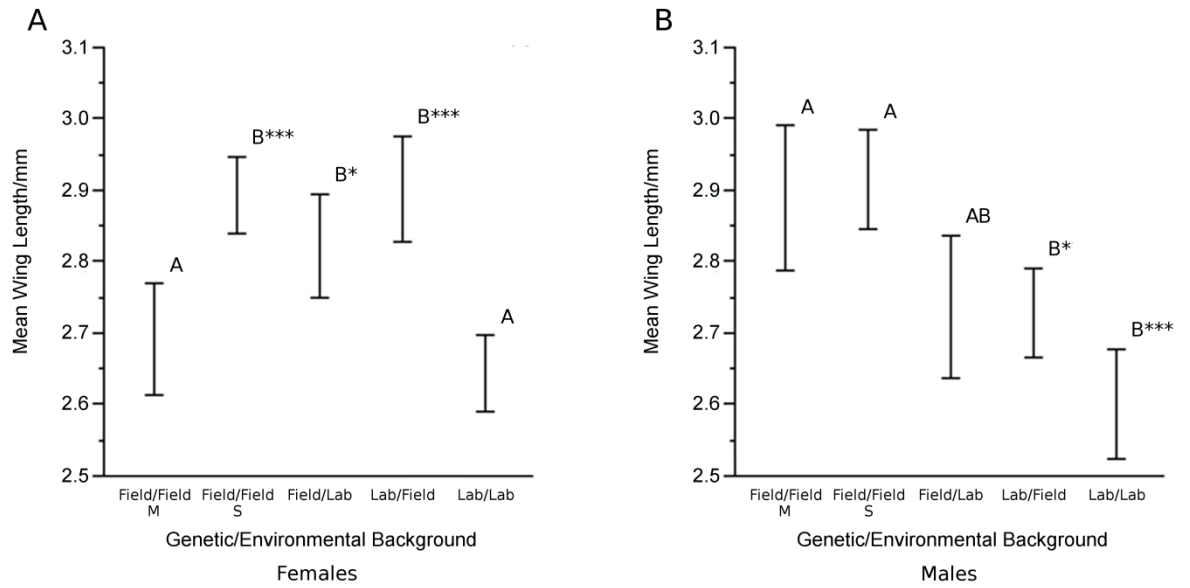
Figure 5.3.4: Comparison of mean wing length between random samples of 30 individuals from genetic/environmental background cohorts, segregated by gender. (a) Females. (b) Males 2011. The results of a one-way ANOVA and Tukey *post hoc* test are indicated - ns, not significant, *: P< 0.05, ** P< 0.01, *** P< 0.001. Error bars represent 95% confidence intervals.

5.3.2 Effects of colonization and rearing conditions on survival, mating success and assortative mating

Body size

In this second experiment, carried out entirely within the 2011 field season, a total of 2400 mosquitoes from 6 rearing batches were reared and released into our SFS experimental enclosures over 2 weeks. We took a random sample across all replicates of 30 wing length measurements from each gender and treatment group (Figure 5.3.4). We found significant variation in mean wing length between treatment groups in both females (ANOVA: $F_{4,145}$=12.61, P<0.001) and males (ANOVA, $F_{4,116,}$=12.08, P< 0.001). In females, the Field/Field S (2.90±0.05mm), Field/Lab (2.82±0.08mm) and Lab/Field (2.90±0.07mm) treatment cohorts formed a cluster of individual with significantly larger wing lengths than both the Field/Field M (2.69±0.08mm, Tukey: P< 0.049) and Lab/Lab (2.65±0.05mm, Tukey, P < 0.011) cohorts (Figure 5.3.4). In males, both the Field/Field M (2.89±0.10mm) and Field/Field S (2.91±0.07mm) cohorts were significantly larger than the Lab/Field (2.73±0.06mm) and Lab/Lab (2.60±0.08mm) cohorts (Tukey, P < 0.05). Mean wing length in the Field/Lab (2.74±0.10mm) cohort was intermediate between these clusters, and was not significantly different from any other male treatment cohort (Figure 5.3.4).

Survival

Males and females from each of our four treatment groups (genetic/environmental background respectively: Field/Field, Field/Lab, Lab/Field and Lab/Lab) were mixed with an equal number of a 1:1 mix of M- and S-molecular form Field/Field mosquitoes of the opposite gender and left to mate in a large (4m×4m×2m) semi-field mating enclosure. Surviving individuals were recaptured after 2 nights.
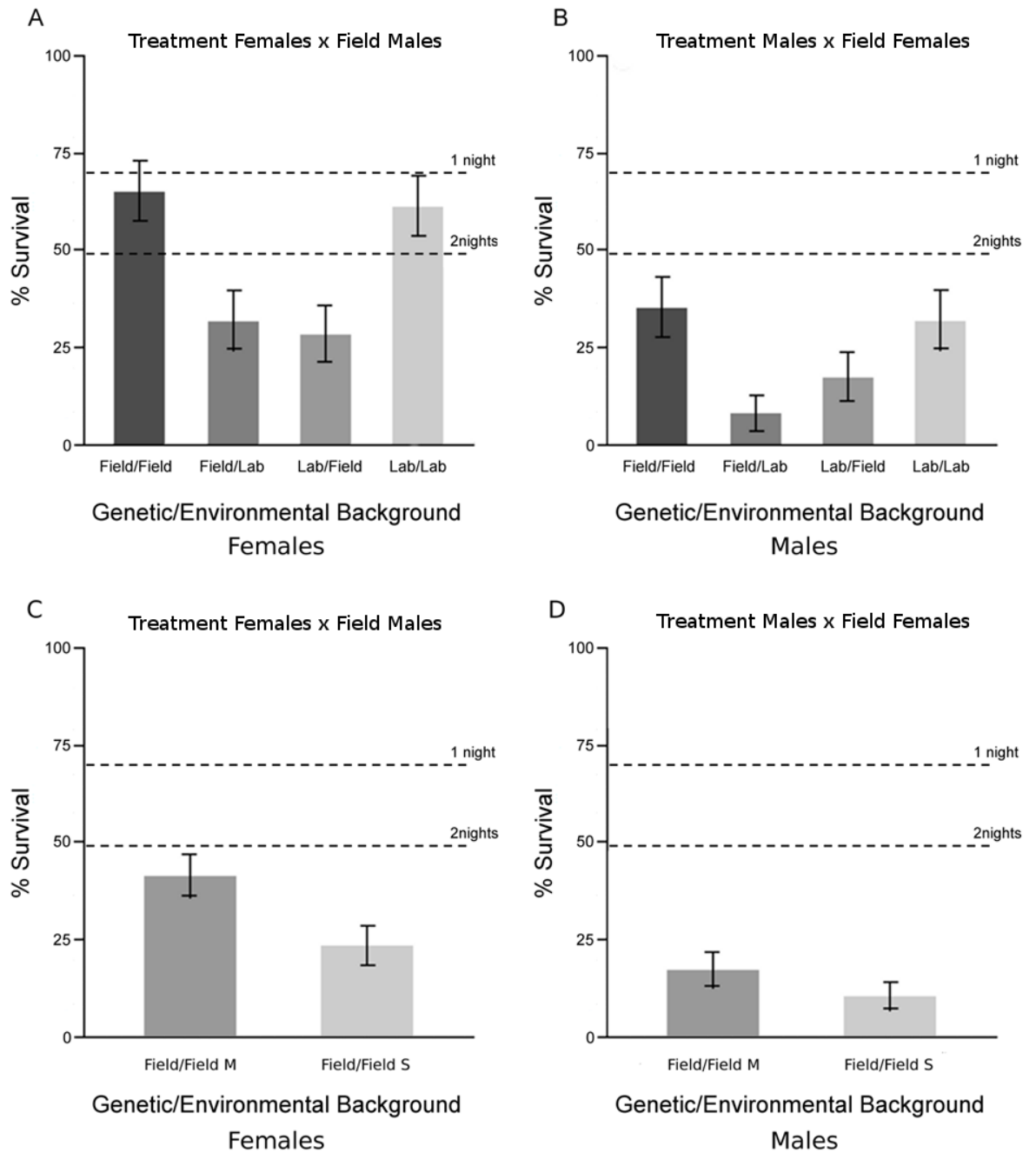
Figure 5.3.5 - Proportion of treatment and field cohort *An. gambiae* surviving after 2 nights under semi-field conditions, (a) Treatment cohort females, (b) Treatment cohort males. (c) Field cohort females. (d) Field cohort males. Error bars represent 95% confidence intervals. Theoretical survival corresponding to a daily mortality of 30% are indicated.

Reciprocal Cross Type:

Treatment M Males x M/S Field/Field Females   **Females**

Source	df	LR Chi-square	P
Molecular Form	1	27.93	<0.001*
Gen/Env	3	44.75	<0.001*
Enclosure	2	2.31	0.32
Gen/Env*Enclosure	6	12.85	0.045*

Reciprocal Cross Type:

Treatment M Males x M/S Field/Field Females   **Males**

Source	df	LR Chi-square	P
Molecular Form	0	N/A	N/A
Gen/Env	3	46.12	<0.001*
Enclosure	2	4.09	0.13
Gen/Env*Enclosure	6	6.33	0.39

Reciprocal Cross Type:

M/S Field/Field Males x Treatment M Females   **Females**

Source	df	LR Chi-square	P
Molecular Form	0	N/A	N/A
Gen/Env	3	68.73	<0.001*
Enclosure	2	6.03	0.049*
Gen/Env*Enclosure	6	3.07	0.80

Reciprocal Cross Type:

M/S Field/Field Males x Treatment M Females   **Males**

Source	df	LR Chi-square	P
Molecular Form	1	5.94	0.015*
Gen/Env	3	29.61	<0.001*
Enclosure	2	1.15	0.56
Gen/Env*Enclosure	6	7.95	0.24

Table 5.3.3: Nominal logistic regression (Likelihood-ratio, n = 600) of the effects of Genetic and Environmental Background and Mating Enclosure on the survival of *An. gambiae* s.s. mosquitoes in a semi-field mating enclosure

Of the 2400 mosquitoes released as part of this experiment, we were able to recapture 702 after 2 nights in our SFS mating enclosures. In the females treatment groups (figure 5.3.5), Field/Field M females had the highest proportional survival (65.33±7.70%), followed by Lab/Lab (61.33±7.91% survived), Field/Lab (32.00±7.55% survived) and Lab/Field (28.67±7.31% survived). In the females from the 'field' group, we observed that M-form Field/Field females survived better (41.59±5.47% survived) compared to S-form Field/Field females (23.51±4.95% survived, figure 5.3.4). In treatment males, Field/Field males again had the highest rate of survival (35.33±7.74% survived) followed by Lab/Lab males (32.00±7.55% survived), Lab/Field males (17.33±6.12% survived) and Field/Lab males (9.09±4.30% survived, figure 5.3.5). In the field cohort, M-form Field/Field males (17.33±4.28% survived) survived better than S-form Field/Field males (10.67±3.50% survived), although survival here was a great deal less than what we might have expected given the rate of survival of the Field/Field males from the 'treatment' cohort which were drawn from the same population (figure 5.3.5).

The factors determining the rate of survival (dependent variable) were modelled using logistic regression analysis (table 3). Independent variables factored into the analysis were as follows: Genetic and environmental background (Gen/Env: Field/Field, Field/Lab, Lab/Field, Lab/Lab), the mating enclosure used (enclosure, synonymous with replicate) and molecular form (M- or S-form, only applicable in models of survival in the field cohort). The interaction between genetic and environmental background and enclosure (Gen/Env×Enclosure) was the final factor. In females from the treatment cohort the effect of genetic and environmental background was extremely significant (LR: n=600 df=3, χ^2=68.73, P<0.001) and the same was also true for treatment males (LR: n=600 df=3, χ^2=46.12, P<0.001). As with the previous experiment, the genetic and environmental background of the treatment cohort also had an unexpected, significant effect on the
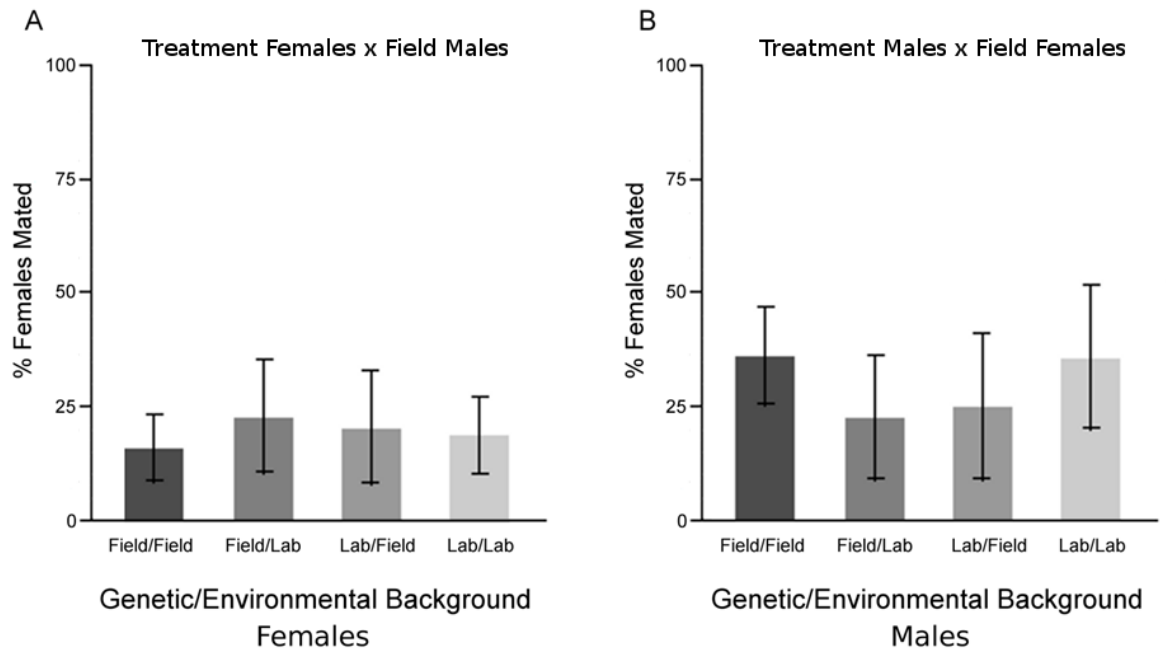
Figure 5.3.6: **Proportion of female *An. gambiae* mated after 2 nights in semi-field conditions.** (a) Treatment cohort females x M/S Field cohort males. (b) Treatment cohort males x M/S Field cohort females. Error bars represent 95% confidence intervals.

Source	df	LR Chi-square	P
Reciprocal Cross Type	1	10.46	0.001*
Gen/Env	3	0.42	0.94
Enclosure	2	1.42	0.49

Table 5.3.4: Nominal logistic regression (Likelihood-ratio, n = 472) of the effects of Reciprocal Cross Type, Genetic and Environmental Background and Mating Enclosure on mating success of *An. gambiae* s.s. mosquitoes in a semi-field mating enclosure.

survival of the field cohort in both males (LR: n=600 df=3, χ^2=29.61, P<0.001) and females (LR: n=600 df=3, χ^2=44.75, P<0.001). Molecular form was a significant factor in field cohort survival in both males (LR: n=600 df=1, χ^2=5.94, P = 0.015) and females (LR: n=600 df=1, χ^2=27.93, P<0.001, table 5.3.3).

Mating Success

In total, we recovered 104 intact sperm bundles from the spermathecae of 479 surviving females. Field/Field M- and S-form females had been successfully mated by M-form males from treatment cohorts as follows (figure 5.3.6): Field/Field males successfully mated most frequently (36.25±10.75% of M/S females mated), followed by Lab/Lab males (35.90±16.75% mated), Lab/Field (25.00±15.86% mated) and Field/Lab (22.50±13.52%). In the reciprocal cross (figure 5.3.6), M- and S-form Field/Field males mated most frequently with Field/Lab females (22.92±12.33% mated), followed by Lab/Field females (20.45±12.41% mated), Lab/Lab females (18.6±8.39% mated) and Field/Field females (16.00±7.32% mated).

A logistic regression model was constructed to assess the factors determining mating success (dependent variable) in this experiment. Model effects were constructed using the independent variable Enclosure, Genetic and Environmental Background and Reciprocal Cross Type. In this case (table 5.3.4), only Reciprocal Cross Type had a significant effect on mating success (LR: n=479 df=1, χ^2=10.46, P=0.001). All interactions were non-significant.

Assortative Mating Behaviour

We assessed the degree of assortative mating behaviour exhibited by each of our treatment groups by PCR/RFLP analysis of successfully mated females and (where appropriate)
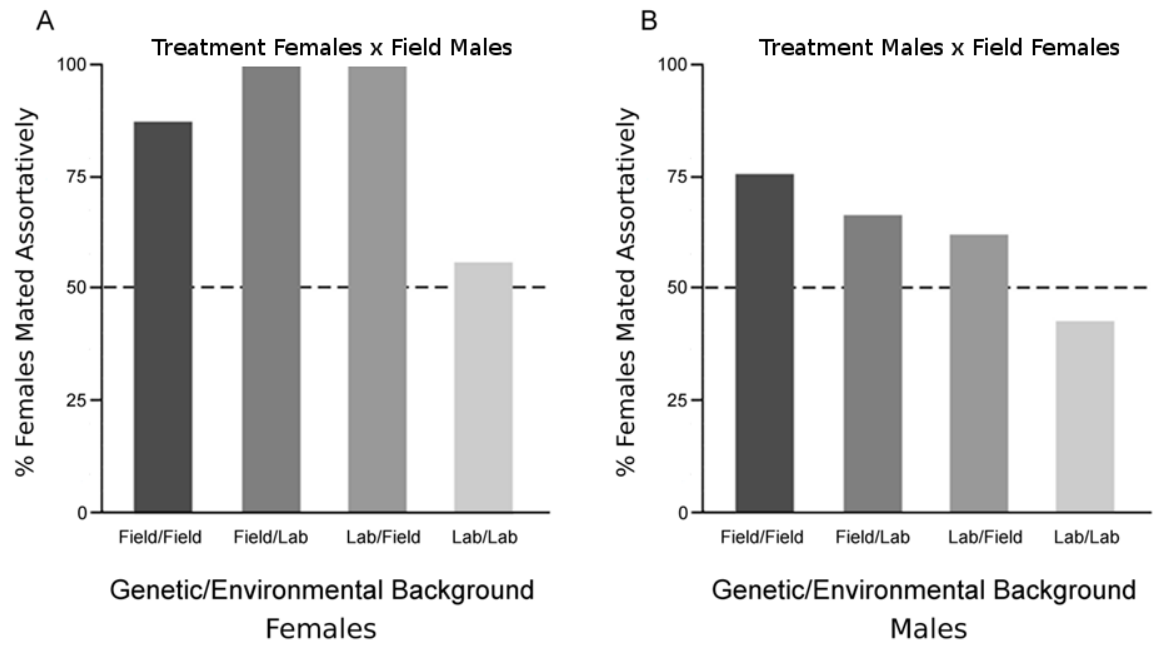
Figure 5.3.7: Proportion of females mated assortatively (ie within molecular form) after 2 nights under semi-field conditions (a) Treatment cohort females x M/S Field cohort males. (b) Treatment cohort males x M/S Field cohort females.

Source	df	LR Chi-square	*P*
Reciprocal Cross Type	1	4.81	0.028*
Gen/Env	3	11.13	0.011*
Enclosure	2	1.08	0.58

Table 5.3.5: Nominal logistic regression (Likelihood-ratio, n = 104) of the effects of Reciprocal Cross Type, Genetic and Environmental Background and Mating Enclosure on the assortative mating behaviour of *An. gambiae* s.s. mosquitoes in a semi-field mating enclosure.

transferred sperm. Where females from our 4 treatment groups were crossed with a 1:1 M:S mix of M- and S- molecular form males we determined the genotype of the successful male by PCR/RFLP of the transferred sperm bundle. From a total of 44 successfully mated females we observed that 14/16 Field/Field females (87.5%), 5/5 Field/Lab females (100%), 7/7 Lab/Field females (100%) and 9/16 (56.25%) Lab/Lab females had been inseminated by M-form males (figure 5.3.7). We did not observe any instances of multiple mating.

In the reciprocal, we crossed males from our four treatment groups with a 1:1 M:S mix of M- and S- molecular form females. In this case the molecular form of all successfully mated females was determined by PCR/RFLP on a whole-body DNA extract. A total of 60 successfully mated females were recovered. Field/Field males inseminated a total of 29 females, of which 22 (75.86%) were M-form. Field/Lab males inseminated 9 females, 6 of which were M-form (66.67%). Lab/Field males inseminated 8 females, 5 M-form (62.5%) and finally Lab/Lab males inseminated 14 females of which 6 were M-form (42.86%, figure 5.3.7).

Combining data by treatment cohort, we found that (independent of gender) the proportion of M-form individuals mating assortatively (i.e. with an M-form Field/Field individual) from the Field/Field (Chi-square, df=1, χ^2=17.34, P<0.001), Field/Lab (Chi-square, df=1, χ^2=4.86, P=0.028), and Lab/Field (Chi-square, df=1, χ^2=5.78, P=0.016), cohorts deviated significantly from the hypothesised 1:1 M:S mating ratio we would expect in the absence of assortative mating behaviour. There was no evidence of significant assortative mating in the Lab/Lab treatment cohort (Chi-square, df=1, χ^2=0, P=1).

Finally, we constructed a linear regression model to assess the factors determining variation in the proportion of individuals mating assortatively (dependent variable, table 5.3.5). Model effects were constructed from the independent variables Reciprocal Cross

Type, Genetic and Environmental Background, and Enclosure/Replicate. Interactions between these variables were also built into the model. In this instance we found that both reciprocal cross type (LR: n=104 df=1, χ^2=4.81, P=0.028) and genetic and environmental background (LR: n=479 df=3, χ^2=11.13, P=0.011) were significant. There was no significant effect from enclosure. Similarly, interactions were non-significant.

5.4 Discussion

The interaction of genetic background and the environmental conditions in which mosquito populations are raised and maintained has important consequences for the modelling, planning and implementation of genetic control strategies. This is demonstrated by past failures in population suppression caused by laboratory-induced population structuring (Reisen *et al* 1982). Understanding and quantifying these interactions will allow us to maximise the competitiveness of a released strain through optimised pre-release rearing conditions and outcrossing and provide additional context for ecological experiments involving the release of marked mosquitoes. There is a strong body of evidence detailing the genetic and behavioural consequences of long-term colonisation of mosquitoes in insectary conditions. Studies have shown a clear reduction in genetic differentiation following colonisation in terms of microsatellite allelic richness and heterozygosity (Norris *et al* 2001) as well as transcriptional changes in immune genes (Aguilar *et al* 2010). The changes in peak activity and assortative mating behaviour in *Anopheles gambiae* sensu lato following colonisation have been known for many years (Jones and Gubbins 1978). The role of short-term (within generation) environmental conditions, particularly during larval development, is less clear. Outcrossing to field-captured, lab-reared F1s failed to restore non-assortative mating behaviour in sterile *C. tritaeneorhyncus* in India during the late

1970s (Baker *et al* 1979, Reisen *et al* 1980) suggesting that genetic background was not the only determinant of population structuring in this study. Similarly a lab-colonised strain of *C. tarsalis* was found to be non-competitive with field-reared females, but competitive with the lab-reared F1 progeny of field captured females (Reisen *et al* 1982), suggesting that larval rearing conditions had affected adult mating behaviour. A recent study in Mali found that assortative mating behaviour between M- and S- molecular forms of *An. gambiae* s.s. breaks down between field-captured, lab-reared mosquitoes released inside houses (Dao *et al* 2008).

Over two wet season in Mali, West Africa, we investigated the genetic and environmental factors affecting survival, mating success and the M-/S-molecular form assortative mating phenotype in *An. gambiae* s.s under semi-field conditions. By raising lab-derived and field-derived *An. gambiae* under both standard lab conditions and field-like conditions we were able to partition the effects of genetic background and environmental conditions during rearing.

We observed a striking difference in assortative mating behaviour between our experimental cohorts. When crossed with a 1:1 mix of M- and S-form individuals Field-derived, field-raised (Field/Field) M-form individuals exhibited a significant preference for mating within molecular form, as did field-derived M-form individuals raised in the lab (Field/Lab), and lab-derived M-form individuals raised in the field (Lab/Field). In contrast lab-derived individuals raised in the lab (Lab/Lab) exhibited no assortative mating behaviour whatsoever. This result demonstrates both clear genetic and environmental factors in determining assortative mating behaviour in this setting. The fact that the Field/Lab cohort mated assortatively precludes this being a purely environmental effect - as lab-rearing was not able to induce non-assortative mating in the field-derived cohort. Similarly the fact that assortative mating behaviour was observed in the Lab/Field cohort

suggests that the loss of assortative mating in the Lab/Lab is not a purely genetic affect, as we were able to restore the assortative phenotype by raising lab-derived mosquitoes in our semi-field system. The exact mechanisms of recognition within cryptic taxa that ultimately lead to strong assortative mating are not fully known, but may involve several components, including spatial segregation of larval habitats and adult swarming (Diabate *et al* 2008, Diabate *et al* 2009) and the recognition of specific flight tones (Sanford *et al* 2011, Pennetier *et al* 2010). Olfaction and taste may also play a part as males and females establish contact before copulation. There were a number of differences between larval rearing and adult maintenance between the Lab/Field and Lab/Lab that may explain the observed differences in assortative mating. The key differences between the environmental conditions in our lab and field insecataries were as follows: 1) Intensity and change in intensity of incident light: Whilst we have no quantitative measure of light intensity, in terms of *lux*, from either the lab or field insecaries, intensity of incident light was likely to be much higher for both larvae and post-eclosure, pre-experimental adults in the field both due to increased refraction and as both rearing trays and adult enclosures were subject direct sunlight for 1-2h per day. Photoperiod was the same between lab and field (see method) but alterations in light intensity have been shown to induce changes in peak activity in *An. gambiae* s.l. under laboratory conditions (Jones *et al* 1972), and the reduction in light intensity at dusk in combination with the mosquito's natural circadian rhythm have been demonstrated as the key factors for initiating male swarming behaviour (reviewed in Howell and Knols 2009). 2) Variation in temperature over a 24h period: In the lab, water temperature in rearing trays was a constant $22.5\pm0.5°C$ whereas in the SFS larval rearing trays temperatures ranged from mid-day highs of 34-36°C to night-time lows of 24°C. Whilst there is no literature on the effects of temperature delta on circadian activity in mosquitoes, it would be an interesting topic for a further study attempting to

partition the varying environmental conditions between the lab and SFS. 3) Visible horizon was constant for field-raised individuals between rearing, adult maturation and experimental crosses, whereas lab-reared individuals would have had no experience of potential orienting features and swarm markers until the start of the experiment. *An. gambiae* are known to form swarms in the field based on both ground and horizon markers (Marchand 1984), and, whilst we did not directly observe swarm formation within the SFS, and it is not known precisely which factors determine aggregation sites for male mate-seeking behaviour, it is possible that the lack of pre-exposure to the SFS conditions and situation meant that lab-reared mosquitoes had a differential response to swarm triggers. Thus, it is possible that, the loss of the assortative mating phenotype in the Lab/Lab is as a result of a shift in activity time, position or both, that caused these individuals to encounter S-form mosquitoes more frequently than the other treatment cohorts, causing a consequent increase in mating between-types. Dao *et al* (2008) hypothesised that the loss of assortative mating they observed in lab-raised field-derived individuals may have been due to the artificially increased conspecificity in their enclosed experimental space and, whilst the number of mosquitoes in our experimental enclosures was much lower (100 vs 300 in Dao *et al* 2008)) - which may explain why we did not replicate their results in terms of assortative mating in Field/Lab individuals - a temporal rather than a spatial change/increase in conspecificity may explain the pattern of assortative mating we observed here.

Of course, as we only observed non-assortative mating in the Lab/Lab cohort these potential environmental determinants of assortative mating behaviour must also be acting in conjunction with the genetic background of the laboratory strain. The genetic basis of assortative mating is discussed in more detail in Chapter 6.

In terms of survival, the first experiment, genetic and environmental background had a significant effect on survival in all comparisons (table 5.3.1, figure 5.3.2). In males from the 'treatment' cohort, survival was clustered by genetic background, with both Field/Field and Field/Lab males surviving better than Lab/Field and Lab/Lab males. This result indicates that in this instance there is little interaction between disparate genetic and environmental factors, and that survival is primarily determined by the genetic background only. This is consistent with the findings of Huho and colleagues (Huho *et al* 2007) who found that the male F1 progeny of field-captured females maintained a larger tenereal lipid reserve - a factor positively associated with survival. In treatment females, despite a significant model effect from genetic and environmental background, no clear pattern of survival was apparent although the Field/Field cohort survived consistently better than other cohorts suggesting that there was an interaction between genetic background and the larval rearing environment. In the second experiment, carried out wholly within the 2011 season, genetic and environmental background was again a significant factor in the survival of both males and females from the treatment cohorts. In the case of both genders, Field/Field and Lab/Lab populations survived significantly better than the Field/Lab and Lab/Field cohorts (figure 5.3.5). This pattern suggests a significant interaction between genetic background and larval rearing environment.

The most unexpected result from our LR analysis was that the genetic and environmental background of the treatment groups affected the survival of the field/field group they were crossed with. This phenomenon was consistent between experimental years, rearing batches, replicates and enclosures. Indeed the effect was significant even after all potential environmental variables had been taken into account in our logistic regression analysis. The most obvious source of this relationship is from an external environmental effect. However, the expectations of variability caused by unavoidable

environmental variation were implicitly built into the experimental design from the earliest stages and were taken into account in the analysis. And the effect was still significant. One possible explanation for this phenomenon is that a lower density of mosquitoes (male or female) caused by significant mortality in the treatment cohort (dependent on treatment) reduces the per-mosquito chance of interaction and mating, this reduced mating chance has the net result of increasing mate-seeking behaviour (swarming, swarm seeking) in the field cohort, placing an additional energetic cost on the individual and leading to higher mortality. This effect could be further magnified by the relatively high-stress enclosure environment. This explanation, whilst plausible, needs to be tested to determine whether this is a bona fide ecological/behavioural effect or an environmental artefact.

Data is slim on short term semi-field survival but a recent, comparable semi-field study investigating survival and mating success in radiosterilised *An. arabiensis* have reported one-night survival/recapture percentages of between 70 and 90%. Assuming a linear progression of survival over time, this corresponds to a two-night survival proportion of between 49 and 81% (Hassan *et al* 2010). Whilst we did achieve this degree of survival (particularly in females, which consistently survived better than males) in some comparisons, survival was typically between 25 and 50 after two nights. A closer comparison may be to the indoor mating study of Dao *et al* (2008) as the domestic environment was of a similar size and configuration to our SFS. Here, one-night survival was between 25 and 70%, a range that more closely reflects our own.

These data present an intriguing picture of the interaction of genetic and environmental factors in determining the survival and mating behaviour of *An. gambiae* in a semi-field system and provide an insight into some of the factors that may have caused unexpected structuring in attempts at population suppression in mosquito populations in the past. They also highlight a number of factors that should be taken into account in the

planning of the mass rearing and release mosquitoes in a hypothetical future genetic control strategy.

6. Genetic variability at potential assortative mating or speciation gene loci.

Abstract

In organisms with known genomes, the use of amplicon pools based on many individuals per population combined with ultra-sequencing can unravel most population-level genetic variants at specific loci in a cost and time-efficient way. We applied this approach to a pilot study of patterns of genetic variation within genes and between populations at nine olfactory or gustatory receptors and cuticular protein loci located within the so-called *2R* and *2L* islands of speciation of *Anopheles gambiae*.

In this study, we investigated whether genetic variability at potential assortative mating or speciation loci is lost during the colonization and lab rearing process. We also attempted to determine if selection for lab-mating might have selected for the fixation of mutant alleles or rare polymorphisms. We characterised a number of regions of interest associated with the genomic islands of speciation described previously and generated PCR amplicons covering the ORFs and adjacent UTRs of nine olfactory or gustatory receptors and cuticular protein loci within these regions from 30 individuals and two samples of an Mopti, M-form *Anopheles gambiae* s.s. population from Mali: the first from a stable laboratory colony originally colonised in 2003 and the second from field samples collected in 2007

On average we observed a 2.7-fold reduction in genetic variatiability in the laboratory colony relative to the field isolate though this reduction was limited to loci on chromosome 2R. Variability was conserved between field and lab populations in loci sampled from with 2La inversion site. In addition we identified 750 SNPs - 693 which were previously undescribed - and 11 fixed differences between the two sampled isolates. This difference in the pattern of variation within the *Anopheles gambiae* genome supports

the theory that 2La loci are important for adaptation to colonisation in some way, although without further evidence that the reduction in variability observed in the 2R is observed throughout other regions of the genome - as seems likely - a conclusive interpretation of this result is difficult.

6.1 Introduction

Understanding the structure of the *Anopheles gambiae* population and how genes flow between sub-populations is important for - amongst other things - understanding the epidemiology of malaria, modelling the spread of insecticide resistance and planning and optimising genetic control release strategies. Given the observation that laboratory rearing and colonisation of M-form *An. gambiae* s.s. can abolish 'canonical' assortative mating behaviour between M- and S-form mosquitoes under semi-field conditions (Chapter 5), assessments of genetic differentiation and fixed differences between the genomes of field- and lab-derived samples may allow us to determine some of the genetic mechanisms that control (in part) assortative mating behaviour between molecular forms. Insights into these mechanisms will help in the development of new transgenic and sterile strains - as discussed previously, assortative mating can significantly affect the outcome of an indundatory release program - and are furthermore of interest in term of evolutionary biology and the opportunity to study an active, ongoing reproductive isolation and incipient speciation event within a broadly sympatric population.

As the genetic and molecular tools for investigating population structuring developed, studies in the late nineties confirmed and expanded on the picture of gene flow in *An. gambiae* s.s. suggested from analysis of paracentric inversion karyotypes (Coluzzi *et al* 1985). Pre-genomic, studies of genetic differentiation on a continental scale using isozyme and microsatellite data (Lehmann et al 1996) detected very low levels of genetic

differentiation between sampled populations from Kenya and Senegal. These results were in contrast to the cytological evidence from inversion polymorphisms and suggested contemporary gene flow across huge distances and significant geographical barriers. It was subsequently heavily criticised (Lanzaro and Tripet 2003). Indeed, Lehmann and colleagues have since performed a study of microsatellite data along a transect of 10 countries from West to East Africa, in which they identified two distinct populations and a third, bridging community, with the Great Rift Valley identified as a putative barrier to gene flow between the two groups (Lehmann *et al* 2003).

Until recently the pattern of strong assortative mating between molecular forms with limited genetic exchange appeared to be the case, to a greater or lesser extent, across the continental range of *An. gambiae* s.s. The integrity of the M- and S-form distinction had been confirmed by several investigators and from a diverse body of evidence. For example, the knockdownresistance gene, kdr, a single nucleotide polymorphism in the sodium gated channel peptide gene para , found proximal to the centromere on the left arm of chromosome 2 (band 20C, Ranson 2000) confers resistance to synthetic pyrethroids and DDT (Martinez-Torres *et al* 1998). This allele is widespread throughout S-form populations but, with some exceptions, absent from the M-form (N'Guessan *et al* 2003, Fannello *et al* 2003, N'Guessan *et al* 2007). Forest M-form populations in Benin possess the kdr allele, and it has been demonstrated that the kdr phenotype observed in Beninois M-form populations has arisen by introgression with sympatric S-form populations as opposed to a like-for-like mutation (Weill *et al* 2000). Kdr has also been observed in M-form populations in Burkina, Ghana and Mali but, perhaps tellingly, only in populations that exist in sympatry with the S-form, suggesting limited gene flow between molecular form sub-populations (Yawson *et al* 2004). Microsatellite data also strongly supports the distinction. Wondji and colleagues investigated genetic differentiation within Forest

chromosomal forms in Cameroon. Using 10 genome-wide microsatellite loci they confirmed that genetic differentiation was lowest between populations of the same molecular form, even over long distances. Interestingly differentiation was highest between sympatric M- and S-form populations (Wondji *et al* 2002). A similar study in Mali utilising 25 microsatellite loci, determined that whilst across the whole genome, the two molecular forms were largely undifferentiated, there is a high degree of differentiation at loci proximal to the centromere on the X chromosome (band 5d and 6), and by extension, proximal to the rDNA responsible for the original M and S designations (Wang *et al* 2001). Studies investigating differential transposable element insertion pattern polymorphism (TE display, Boulesteix *et al* 2007) and short interspersed elements (SINE, Barnes *et al* 2005, della Torre et al 2005) have further reinforced the distinction between M- and S-form populations. This has been confirmed on the continental scale, where TE display was able to resolve M- and S- populations from 10 countries from Benin to Madagascar (Esnault *et al* 2008).

However, recent work has made the current picture of gene flow and population dynamics in An. gambiae more complicated than simply assortively mating M- and S-form populations. Yawson and colleagues, following up on their work on the frequency of the kdr allele in populations from Ghana and Southern Burkina Faso (Yawson *et al* 2004), discovered population structuring that did not primarily follow the canonical molecular-form arrangement. Sampling *An. gambiae* s.s. populations from three different ecological zones - coastal mangrove, deciduous forest and Sahel Savannah - they discovered that genetic differentiation between M- and S-forms was relatively low across 7 microsatellite loci - although these had exhibited a high degree of differentiation in previous independent investigations (Wondji *et al* 2002). Differentiation was highest between mosquitoes from different ecological zones irrespective of molecular form. This result indicated that

ecological factors may be more important as barriers to gene flow than reproductive isolation in this setting. The research has subsequently come in for some criticism in the literature, primarily because the investigators failed to assess the paracentric inversion karyotypes of their captured specimens (Lee *et al* 2009).

Paracentric inversions notwithstanding, the study is still suggestive of further sources of genomic differentiation beyond the X-linked molecular form in at least some West African Anopheles gambiae s.s. populations. This view is given further credence by research published by Slotman and colleagues in 2007. This study investigated genetic differentiation across 12 microsatellite loci on chromosome 3 within and between mosquito populations in Mali and Cameroon, and considered both molecular and chromosomal form. Forest-S and Savannah-S populations exhibited the lowest degree of genetic differentiation, even over very large distances. As expected there was a high degree of genetic differentiation between both sympatric and spatially separated M- and S-form populations. However, there was an even greater degree of differentiation between Malian (Mopti-M) and Cameroonian (Forest-M) M-form populations: some 7.5 times the level of differentiation between comparable S-form populations, which the authors interpret as subdivision within the M-form in West Africa, with paracentric inversion playing a key role in maintaining differentiation (Slotman *et al* 2007). A subsequent study found that microsatellite, ecological and inversion data all support the theory that the Forest-M form is genetically distinct from other populations within *An. gambiae* s.s. (Lee et al 2009).

The apparent lack of genetic differentiation between molecular forms outside bands 5d and 6 on the X chromosome had been verified by numerous microsatellite studies (Wang *et al* 2001, Wondji *et al* 2002, Lehmann *et al* 2003, Stump *et al* 2005). However Turner and colleagues (2005) utilising a DNA-hybridisation microarray to map genetic differentiation across 142,000 loci in 7 M- and 7 S-form mosquitoes from Cameroon

(contrast to one of the most detailed microsatellite studies to date, which has mapped 42 loci (Wang-Sattler *et al* 2007)) not only described the predicted area of high differentiation proximal to the X chromosome centromere, but also identified a statistically robust 'island' of high genetic differentiation proximal to the centromere of chromosome 2L and a less well supported region on 2R within the 2Rb inversion. Sequenced loci in the 2L and 2R regions confirmed the presence of fixed SNPs in complete linkage disequilibrium with molecular form within these regions, but shared polymorphisms at loci adjacent to, but outside these areas (Turner *et al* 2005). A follow-up investigation confirmed the presence of increased differentiation in the centromeric region of 2L in 52 individuals from both Mali and Cameroon. However differentiation in the 2R region only remained significant for individuals from Cameroon, particularly at one gene locus: GPRor38 a putative gustatory receptor (Turner and Hahn 2007). As the authors state, these were significant findings, as they described fixed differences on chromosome 2L in complete linkage disequilibrium with the classic molecular form loci on the X-chromosome. Furthermore, a subsequent study has found pericentromeric regions on the left arm of chromosome 3 in complete linkage disequilibrium with both the X- and 2L loci (White *et al* 2010), an observation that can only be possible with complete reproductive isolation or powerful selection against hybridisation. However, no consistent bias in inheritance of the 3 segregating island genotypes was found in F2 progeny of recombinant crosses, precluding strong negative selection, and, in addition, it has recently been demonstrated that the X and 2L islands *sensu* Turner *et al* 2005 maintain a high degree of differentiation in population from the Gambia that exhibit high levels of molecular form hybridisation (Weetman *et al* 2011).

Reduced recombination in areas proximal to the centromere has been proposed as a plausible mechanism for maintaining genetic differentiation in the face of independent

assortment, selection and hybridisation and has been described in several species. It has since been confirmed that recombination is reduced 16-fold in An. gambiae in the centromeric region of the X-chromosome (Slotman *et al* 2006) and is likely to be severely limited in the 2L island (Stump *et al* 2006) although in the case of the latter a lack of informative microsatellite loci precluded a definitive analysis. Indeed the relatively poor coverage of pericentromeric regions in the published genome combined with reduced recombination makes a definitive conclusion on the importance of reduced recombination on the process of speciation difficult (Turner *et al* 2010). Additionally, reduced recombination (Stump *et al* 2006) and increased genetic differentiation (White *et al* 2007) have been described for the 2La inversion, with the latter being particularly marked in two regions within 2La, adjacent to the proximal and distal inversion breakpoints. The patchy distribution of genetic differentiation associated with reduced recombination suggests that gene flow can occur freely between assortatively mating populations at loci that are selectively neutral in hybrids, but that realised gene flow is reduced at loci that are selectively disadvantageous in hybrids, resulting in a mosaic genome architecture and areas of differential gene flow maintained by selection against hybrids at specific loci rather than in a genome-wide manner (Wang-Sattler *et al* 2007). It is likely, therefore, that the individual genes responsible for reproductive isolation/assortative mating would be found within these areas of reduced gene flow or 'islands of speciation' (Turner and Hahn 2007), making these regions strong candidates for detailed characterisation and investigation.

As mentioned previously, the physiological basis of assortative mating behaviour between the M- and S-molecular forms of An. gambiae in parts of their range is poorly understood. Diabate and colleagues have identified complete segregation in male swarming between molecular forms in Mali (Diabate *et al* 2009) and recent studies have demonstrated the use of differential wing beat harmonics for mate recognition in *An.*

gambiae (Pennetier *et al* 2010, Sanford *et al* 2011). Given the fact that contact plays a part in the initiation of mating, taste and/or olfaction may also play a role in mate choice. By identifying and characterising highly differentiated loci within the putative islands of speciation between samples exhibiting different degrees of assortative mating behaviour we may be able to elucidate the genetic mechanisms controlling mate choice. Previously, (Chapter 5) we described the breakdown in assortative mating behaviour under semi-field conditions in *An. gambiae* M-form mosquitoes sampled from a F42 laboratory colony and raised in standard laboratory conditions. Laboratory colonisation causes a loss of genetic diversity through random genetic drift, reduction in the size of the gene pool, increased consanguineous mating and strong selection for females capable of completing a gonotrophic cycle using the artificial feeding and oviposition equipment routinely employed in a typical insectary. Hard data regarding the loss of diversity in colonised *An. gambiae* is lacking, but an 8-fold decrease in allelic richness and 3.5-fold reduction in heterozygosity has been reported in microsatellites (Norris *et al* 2001) following colonisation.

In this study, we investigated whether genetic variability at potential assortative mating or speciation loci is lost during the colonization and lab rearing process. We also attempted to determine if selection for lab-mating might have selected for the fixation of mutant alleles or rare polymorphisms. We characterised a number of regions of interest associated with the genomic islands of speciation described previously and generated PCR amplicons covering the ORFs and adjacent UTRs of nine olfactory or gustatory receptors and cuticular protein loci within these regions from 30 individuals and two samples of an Mopti, M-form Anopheles gambiae s.s. population from Mali: the first from a stable laboratory colony originally colonised in 2003 and the second from field samples collected

in 2007. This study also acted as a pilot for the use of ultra-deep 454 pyrosequencing approach.

6.2 Material and Methods

6.2.1 Selection and Characterisation of SNP Target Regions

Previous studies investigating levels of gene flow between sub-populations of An. gambiae have identified several discrete regions where the recombination rate is significantly reduced. These 'genomic islands of speciation' (Turner et al 2005), are thought to contain genes driving the process of incipient speciation between An. gambiae sub-populations (Turner and Hahn 2007) and, in the case of areas associated with the 2La inversion, aridity tolerance (White et al 2007) and have recombination rates up to 16 times lower than the mean for the genome (Slotman et al 2006). Candidates for PCR amplification and sequencing were chosen from the X, 2R, 2L (distal), 2L (proximal), 2L (centromeric) and 3L islands.

The gene annotation for each candidate region was extracted in the BioMart format from the latest version of the Anopheles gambiae genome (PEST 3.4, build 51, ensembl.org/vectorase.org). Scale gene maps of each area were constructed from gene and intergenic space size data using the GNU Image Manipulation Program for Mac (www.gimp.org). These were then used to identify 20 kilobase areas of interest ('chunks') that would be the target of amplification. The selection criteria for chunks were primarily based on high gene density and annotation/putative function.

6.2.2 SNP PCR Primer Design, DNA Extraction and PCR Amplification

To maximise primer cross-specificity a Basic Local Alignment Search Tool (BLAST, www.vectorbase.org) search was carried out between a ~1 kb candidate primer position

Primer	20k Fwd	29 Fwd	30 Fwd	31 Fwd	32 Fwd
29 Rev	8, 647	**2, 859**	**506**	-	-
30 Rev	**10, 330***	**4, 542**	**2, 189**	**519**	-
31 Rev	12, 739	6, 951	**4, 598**	**2, 853**	1, 333
32 Rev	16, 314	**10, 526***	8, 173	6, 428	**4, 908**
20k Rev	**18, 856***	13, 166	**10, 713***	9, 043	7, 448

Table 6.2.1: Primer matrix for the GPRGR29-32 array within the distal 2La breakpoint island of speciation. Tested primer combinations are in bold. Primer pairs that produced a strong single band of the expected length are indicated with an asterisk(*).

region from the published genome against the whole genome shotgun (WGS) trace files for both the M- and S-molecular forms of An. gambiae. A sequence alignment of the top 5 hits from each WGS file against PEST was then constructed using the ClustalW algorithm. Alignment and sequence manipulation was carried out using the programs BioEdit for x86 windows (www.mbio.ncsu.edu/BioEdit/bioedit.html) and eBioX for Mac (www.ebioinformatics.org/ebiox/). Primers were designed using the program Primer3 (www.primer3.sourceforge.net) and were targeted to areas of 100% consensus. The target amplicon length was 20 kb. As this was an ambitious size for an fragment amplified from genomic DNA targets, additional primers were designed within each chunk that would generate amplicons of ~500, ~3000, ~5000 and ~10000 bp in length (figure 6.2.1, table 6.2.1). This modular design allowed the determination of the largest reliable amplicon.

Our Field sample consisted of 30 individuals from a Mopti M-form population captured by aspiration in the village of Bankoumana, Kati District, Mali (12°12'2"N, 8°15'54"W). These samples had been stored as whole-body gDNA extracts suspended in ddH$_2$O and stored at -80 ℃ for approximately 12 months. The Lab sample was composed of 30 individuals from a laboratory population of Mopti/M-form mosquitoes colonised from the village of N'Gabakoro Droit, Koulikorou, Mali (12°39'46"N, 7°50'34"W) and maintained in insectary conditions since 2003. DNA extraction was using a modified protocol of the DNAzol DNA extraction system (Invitrogen) described previously.

For target amplicons to 5 kb in size, PCR was carried out using the GoTaq Flexi PCR kit (Promega, UK). The reaction mix consisted of 5 µl 5x GoTaq Buffer, 1.5 µl 25 mM MgCl$_2$, 0.5 µl dNTP (10mM each), 0.5 µl Primer 'A', 0.5 µl Primer 'B', 0.125 µl Taq polymerase (0.5u), 2µl genomic DNA template (5-50 ng/µl) and 14.875 µl ddH$_2$O - total reaction volume 25 µl. The PCR was carried out using a PTC-200 'DNA-Engine' thermocycler (MJ research, now BioRad) with the following steps: Initial denaturation 3

minutes at 95 ℃, 35 cycles of: 15 s at 95 ℃, 30 s at 62 ℃ and 1 min/kb at 72 ℃ followed by 10 min at 72 ℃. PCR for 10 and 20 kb target amplicons was carried out using the Qiagen LongRange PCR kit (Qiagen) with a reaction mix as follows: 2.5 µl LongRange PCR Buffer w/MgCl$_2$, 1.25 µl dNTP (10mM each), 5 µl Q-solution, 4µl Primer 'A', 4 µl Primer 'B', 0.2 µl LongRange PCR enzyme mix (1.0u), 5 µl DNA template (5-50 ng/µl) and 2.05 µl ddH$_2$O - total reaction volume 25 µl. PCR reactions were carried out as above with the following thermocycler programs: 10 kb fragments, initial denaturation 3 min at 93 ℃ followed by 35 cycles of 15 s at 93 ℃, 30 s at 62 ℃ and 10 min at 68 ℃. 20 kb fragments: 3 min at 93 ℃, 10 cycles of 15 s at 93 ℃, 30 s at 62 ℃, 20 min at 68 ℃, followed by 28 cycles of 15 s at 93 ℃, 30 s at 62 ℃ and 20 min + 20 s per cycle at 68 ℃. PCR products were fractionated by electrophoresis using a 1% w/v agarose gel w/ 10 ppm ethidium bromide (10 µl EtBr in 1l 1% w/v agarose-TBE). Each PCR was judged to be a success if a band of the predicted size was visible on the gel. Failed PCRs where rerun once.

6.2.3 454 Pyrosequencing

5 µl from each successful PCR were pooled by strain and purified using the Qiagen MinElute PCR Cleanup kit (Qiagen) following the manufacturer's instructions. The concentration of amplified DNA was normalised prior to pooling. The DNA concentration of pooled and purified samples was quantified using a Nanodrop N-1000 spectrophotometer (Thermo Scientific). Samples were delivered to collaborators at the Centre for Genomic Research at the University of Liverpool. Pooled amplicons were fragmented at into random 400 bp stretches and labelled with population-specific Multiple Identifiers tags (MID), then combined and re-sequenced using the facility's 454 GS FLX Titanium Series pyrosequencer (454 sequencing Roche). Short reads were split by

population and mapped to reference sequences from the PEST genome using GS Reference Mapper in Newbler2.3.

6.2.4 Genomic Analysis

The consensus sequences for the five loci (see results) and two populations generated by the GS Reference Mapper program were aligned to the reference genome sequences using ClustalX2.0.12. A dataset of all variants from the reference sequences within each loci and populations, including their frequency and the depth of sequencing for each variant was created from the GS Reference Mapper output files. The consensus sequence alignments were used to indentify gaps in transcribed regions and areas of high complexity that were the results of mapping errors. The variants dataset was screened for fixed differences between populations and all such fixed differences double-checked manually from the raw assemblies data and consensus alignments. Exons containing fixed differences between populations were translated from the alignments using Jalview2.5.1. A subset of all SNPs was created and filtered to remove artefacts generated in the assembly steps. SNPs identified by GS Reference Mapper within or directly adjacent areas of high complexity - i.e. multiple substitutions, insertions or deletions - were usually errors and thus eliminated from subsequent analyses. The SNP dataset were compared to existing SNP data extracted from the Ensembl SNPdb using the Biomart search tools available at http://www.ensembl.org/biomart/martview/. The degree of variation within and between populations and the factors affecting variation such as coverage, sequencing depth and the number of sequenced haplotypes were investigated and modelled using a General Linear Model (GLM).

In addition to the identification of fixed differences between populations from the variant frequency data (see above), the program BayeScan2.0 was used to conduct

Bayesian scans to for detecting positive and stabilizing selection within the high-confidence SNP dataset.

6.3 Results

6.3.1 Selection and Characterisation of SNP Target Regions

In total, 6 candidate regions were assessed for candidate amplicon placements (figure 6.3.2). The X island is approximately 2.7 Mb long and contains ~40 open reading frames (ORFs) - 67.5 kb per gene on average - consistent with its position in the heterochromatin proximal to the X chromosome centromere. Only two ORFs have a functional annotation: Q8WQP8_ANOGA a member of the Cytochrome P450 protein super-family and Q5PTI3_ANOGA an HSP40 chaperone protein. The 2R island is significantly smaller than the X island - ~0.11Mb - but has a higher gene density - 11kb/gene - again consistent with its location in the more transcriptionally active euchromatin. The most notable annotated genes in this region are two olfactory receptors: GPROR38 and GPROR39 and a gustatory receptor: GPRGR13. The 3L island lies in a centromere-proximal region in the heterochromatin. There are 43 ORFs in this 2Mb region (46 kb/gene) but only two are functionally annotated: SOD3, a superoxide dismutase and RM23_ANOGA, a 39S ribosomal protein. The $2L_{distal}$ island has a comparable gene density to the 2R island: 10.4kb/gene - 174 genes in 2Mb again consistent with its euchromatic location. There are 52 genes with a robust functional annotation in this region, most notably an array of 37 cuticular proteins (CPR) and 4 gustatory receptors: GPRGR29-32. The $2L_{proximal}$ island, in contrast, has the lowest gene density of any of the euchromatic regions considered - 52 genes in 2Mb (38 kb/gene) - and only 3 functional annotations. These are: Pigment Dispersing Hormone (PDH), Q1WJM2_ANOGA - an Iduronate 2-sulfatase and
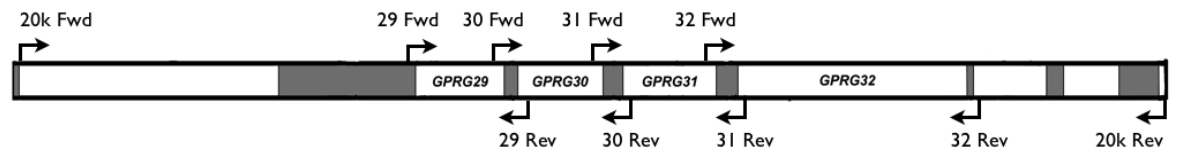
Figure 6.3.1: Scale representation of a 20 kilobase region of the 2L$_{distal}$ island encompassing an array of four putative gustatory receptors. Grey regions are intergenic DNA, white regions represent open reading frames. Primer target binding regions are indicated.

RL8_ANOGA - a 60S ribsomal protein. Finally, the $2L_{centromeric}$ island has the lowest gene density of any of the heterochromatic regions considered: 22 ORFs in 2 Mb or 90 kb/gene. In addition there is only one annotated gene: CPF3 a cuticular protein with a hydrocarbon binding motif.

6.3.2 SNP PCR Primer Design, DNA Extraction and PCR Amplification

All of the targets located in centromere-proximal speciation islands (X, $2L_{centromeric}$, 3L) proved completely refractory to amplification at any amplicon size. Furthermore, we were unable to amplify targets larger than approximately 3kb from the 2R and $2L_{distal}$ islands, the lone target amplicon in the $2L_{proximal}$ island (CPF3) also proved refractory to amplification (summarised Table 6.3.1). From the 2L chromosome, we were able to amplify the CPR64/34 locus comprising two cuticular protein-coding genes, CPR64 and CPR34 that are part of a larger array of cuticular protein genes in that genomic region (figure 6.3.2); and the GPRGR29-30-31-32 locus, consisting of 3 overlapping amplicons and covering the gustatory receptor genes GPRGR29, GPRGR30, GPRGR31 and the GPRGR32 5'UTR and first exon. On the 2R chromosome, we were able to amplify the GPRGR13 gustatory receptor sequence, and amplicons comprising the ORFs and up/downstream regions adjacent to GPROR38 and GPROR39 (Table 6.3.2).

6.3.3 454 Pyrosequencing and Genomic Analysis

An average of 44.23 ± 5.25 chromosomes were sequenced per locus, with a mean depth (ie number of short read sequences per variant) of 302.87 ± 47.05 resulting in a total coverage (number of reads per haplotype) of 6.42 ± 0.67. A total of 1016 unique variants were detected - 815 in the Field population and 611 in the Lab population - with 804 of these as SNPs, across the entirety of the 17413bp cumulative sequence being considered this

Chromosome	Position	Gene	Predicted Function	Amplicon Length (bp)
X	Centromeric, band 6	Q8WQP8_ANOGA	Cytochrome P450	2812
		Q5PTI3_ANOGA	HSP40 chaperone protein	2646
2R	2Rb inversion, band 12E	GPROR38	Olfcatory Receptor	2699*
		GPROR39	Olfcatory Receptor	2633*
		GPRGR13	Gustatory Receptor	2758*
2L	Distal 2La breakpoint, band 26C-D	GPRGR29	Gustatory Receptor	2758*
		GPRGR30	Gustatory Receptor	2699*
		GPRGR31	Gustatory Receptor	2633*
		CPR64	Cuticular Protein	2701*
		CPR34	Cuticular Protein	
	Proximal 2La breakpoint, band 23A-B	RL8_ANOGA	60S ribsomal protein	2544
	Centromeric, band 20A-C	CPF3	Cuticular Protein	2918
3L	Centromeric, band 38A-B	AGAP010317	N/A	2951

Table 6.3.1: Genes targeted for amplification and re-sequencing from within the 6 islands of speciation. Chromosomal location and putative function are indicated. Amplicons marked with an asterisk (*) were successfully amplified and re-sequenced.

Locus	Location	Start	Gene	Start Transcript	Length (bp)	Structure
CPR64/34	2L	39,220,924	CPR34	39,221,590	614	2 exons, 1 intron
			CPR64	39,222,587	930	2 exons, 1 intron
GPRGR29-30-31-32	2L	39,308,119	GPRGR29	39,308,977	1,536	4 exons, 3 intron
			GPRGR30	39,310,745	1,474	4 exons, 3 intron
			GPRGR31	39,312,570	1,618	4 exons, 3 intron
			GPRGR32	39,314,564	507 (incomplete)	1 exon
GPRGR13	2R	24,810,442	GPRGR13	24,811,173	1,441	4 exons, 3 intron
GPROR38	2R	24,857,029	GPROR38	24,857,474	1,622	6 exons, 5 intron
GPROR39	2R	24,849,688	GPROR39	24,850,239	1,608	6 exons, 5 intron

Table 6.3.2: Co-ordinates, position, length (bp) and genetic structure of the 5 loci and 9 genes successfully isolated for study by comparative ultradeep-sequencing.

Figure 6.3.2 (following pages): Scale representation of the six 'genomic islands of speciation' representing areas of reduced recombination on the Anopheles gambiae s.s. genome. Grey regions are intergenic DNA, white regions represent open reading frames. Coding regions with a functional annotation are identified. (A) The 2L 'proximal' island (band 23) (B) The 2L 'distal' island (band 26). (C) The 2L centromeric island (band 20). (D) The X island, proximal to the centromere (band 6) (E) The 2R island, note the larger scale. (F) The 3L island, proximal to the centromere (band 38).

A

Anopheles gambiae s.s. Proximal 2La Break Island   *(20.5 - 22.1 Mb, Band 23)*

☐ Gene/Coding Region
▦ Intergenic Space

0.2 Mb

PDH   Q1WJM2_ANOGA   Uncharacterised/Novel   RL8_ANOGA   Uncharacterised/Novel
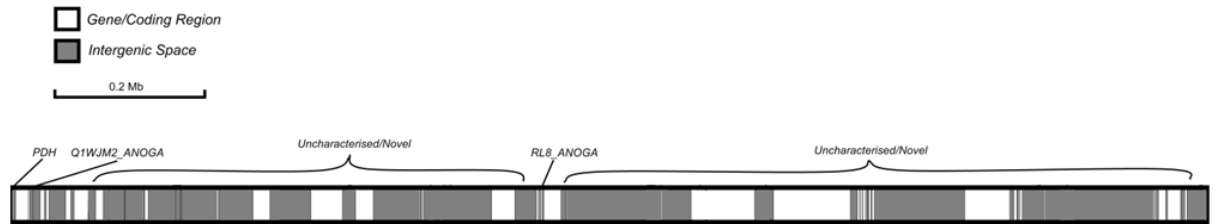
B

Anopheles gambiae s.s. Distal 2La Break Island   *(38.5 - 40.5 Mb, Band 26)*

☐ Gene/Coding Region
▦ Intergenic Space

0.2 Mb

CPR   GPRGR 29 30 31 32   A0A114_ANOGA   Q683N9_ANOGA   SRPN 1 2 3   GPRGR55   Q6VFH8_ANOGA CPR111   CLIPA10   HSP83_ANOGA   Toll 9

C
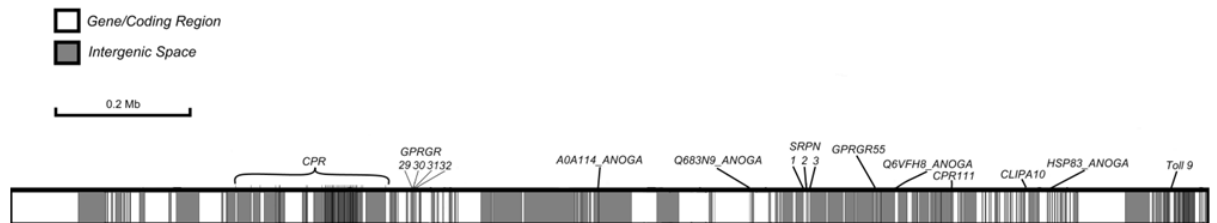
Anopheles gambiae s.s. 2L Centromeric Speciation Island   *(0 - 2 Mb, band 20)*

☐ Gene/Coding Region
▦ Intergenic Space

0.2 Mb

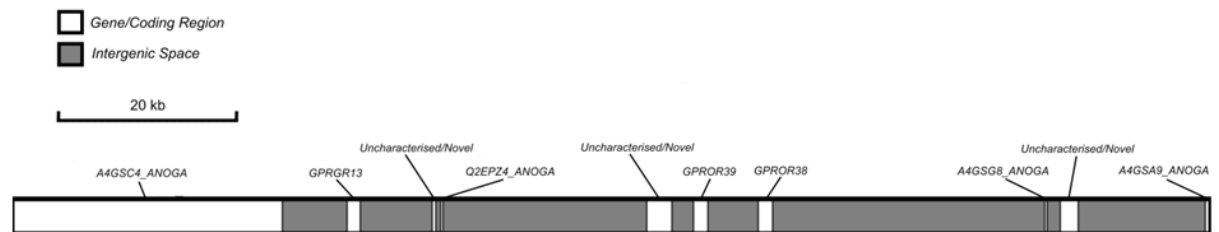Uncharacterised/Novel   CPF3   Uncharacterised/Novel

D
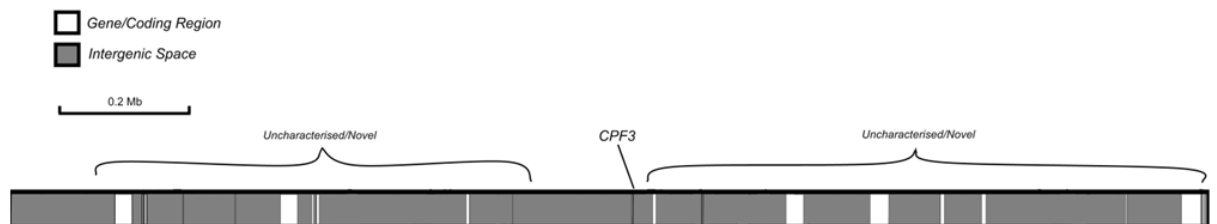
Anopheles gambiae s.s. X Speciation Island (21.6 - 24.3 Mb, Band 6)
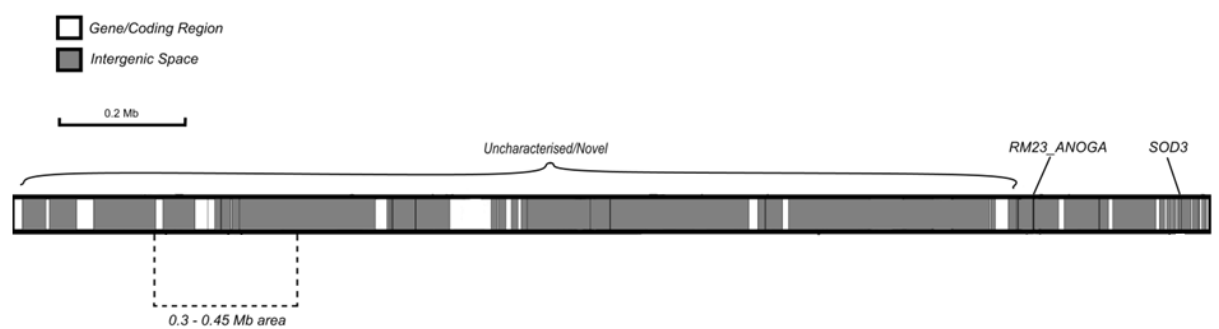

E
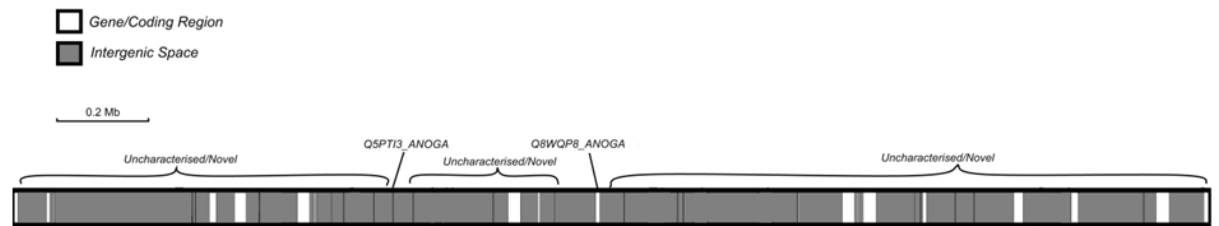
Anopheles gambiae s.s. 2R Speciation Island (24.8 - 24.9 Mb, Band 12e)

F

Anopheles gambiae s.s. 3L Island (0 - 2 Mb, Band 38)

equated to 4.61SNPs/100bp, or alternatively 22bp per SNP. Within populations the Lab cohort was less variable - 509 SNPs (2.92 SNPs/100bp) - than the Field cohort - 663 SNPs (3.87 SNPs/100bp). In terms of loci/amplicons, GPROR39 was the least variable (1.5 ±0.68 SNPs/100bp) with CPR64-34 proving to be the most variable (5.32 ±1.12 SNPs/100bp). Finally, introns (3.99 ±1.63) were on average more polymorphic than non-transcribed regions (3.54 ±1.50); and exons were the most conserved (2.41 ±1.23)(Table 6.3.3). To test the effects of Population (Lab or Field), Locus (amplicon), Sequence Type (exon, intron, non-transcribed) Coverage (depth/chromosome), and PCR success on variability (SNPs/100bp) we constructed a General Linear Model (GLM). With the exception of PCR success (GLM: df=1, F=0.652, P=0.429) all other factors were significant (Table 6.3.4). There was an inversely proportional relationship between coverage and variability - i.e. the number of SNPs identified reduced as coverage increased (Figure 6.3.3).

Due to the potential for the assembly and alignment process to introduce apocryphal SNPs, particularly adjacent to indels or complex variants, we undertook a manual 'enrichment' pass to generate a high confidence SNP sample. SNPs were excluded if directly adjacent to an indel or multi-base, complex indel/substitution. In this manner, 54 SNPs and 212 indels and complex variants were excluded leaving 750 high confidence SNPs. In contrast, the dbSNP database contains entries for 90 polymorphisms in the same area, of these 90 SNPs, we were able to cross reference 53 to our sample of 750 SNPs identified by deep sequencing. The Lab population remained less variable than the Field, with a total of 460 SNPs (2.64SNPs/100bp) compared to 628 SNPs (3.61SNP/100bp). A similar pattern to variation in terms of both locus and sequence type was observed, with CPR64-34 again the most variable locus (5.07 ± 1.33 SNP/100bp) and GPROR39 the least variable (1.42 ±0.58 SNP/100bp), and Exons being again the most conserved (2.32 ±
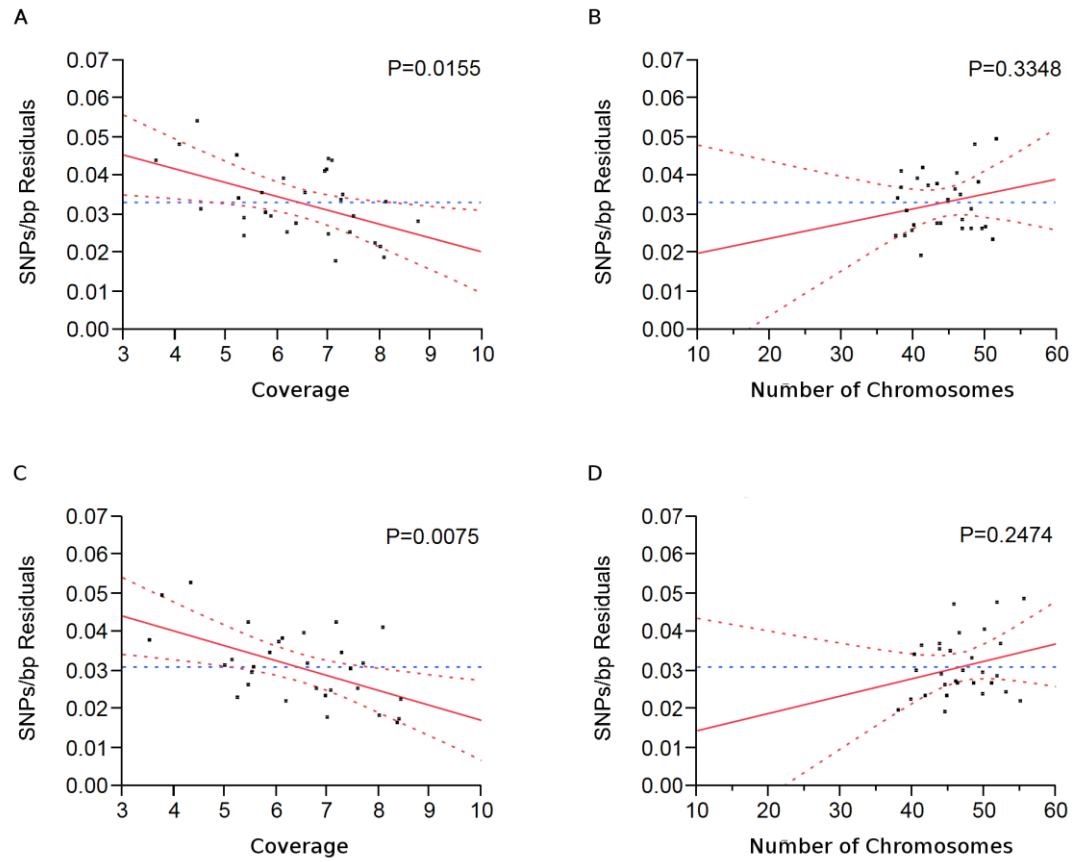
Figure 6.3.3: Effect of the coverage (A,C) and number of chromosomes (B,D) on the raw total number of variants (A,B) and a subset of high-confidence SNPs detected across all populations (C,D). Data was corrected for the effects of population, locus and region within locus (exon, intron or non-transcribed) through general linear modelling.

Locus	Region Seq Type	Lab SNPs/100bp	Field SNPs/100bp
2LCPR34-64	Exons	4.55	4.33
	Introns	6.96	6.33
	NT	3.75	4.52
2LGPRGR29-32	Exons	3.44	3.81
	Introns	5.35	4.53
	NT	5.17	5.45
2RGPRGR13	Exons	0.4	1.51
	Introns	1.1	4.97
	NT	1.9	5.46
2RGPROR38	Exons	0.41	2.78
	Introns	0.5	3.77
	NT	0.56	2.88
2RGPROR39	Exons	0.65	1.3
	Introns	1.84	2.1
	NT	0.93	1.67
Mean	Exons	1.89 ± 2.44	2.75 ± 1.68
	Introns	3.15 ± 3.18	4.34 ± 1.94
	NT	2.80 ± 2.43	4.00 ± 2.08
Grand mean	**All**	**2.50 ± 1.21**	**3.69 ± 0.87**

Table 6.3.3: Average variability in terms of SNPs/bp per sequence region (exon, intron and non-transcribed) per population in the assemblies of short-read sequences to 5 reference sequences located in speciation islands on chromosomes 2L and 2R.

1.17SNP/100bp) and introns the least (3.74 ±1.59SNP/100bp). A GLM model again identified all factors as significant in terms of their effect on variability, with the exception of PCR success (Table 6.3.4). There was again, a significant, inversely proportional effect of coverage on variability.

6.3.4 Genetic Differentiation, Selection and Fixed Differences

Overall genetic differentiation between the Lab and Field samples was extremely low (F_{ST} 0.004). The degree of genetic differentiation between the Lab and Field samples was similarly very low between coding (Exons, F_{ST} 0.004) and non-coding regions (Introns, F_{ST} 0.0003, non-transcribed regions, F_{ST} 0.008). A total of 10 fixed differences were identified between the populations. All differences identified were SNPs. Of the ten, 2 were located in non-coding, untranslated regions, 4 were located in introns and 4 were in exons. Five fixed differences were identified in the combined GPRGR29-31 locus, 4 in the GPROR38 locus and one in GPROR39 (Table 6.3.5). The SNP at position 1285 within then GPRGR29-31 array (specifically GPRGR29) was an A - G substitution in position 2 of codon 150, causing a change in the amino acid in the translated sequence from histidine (CAC) to argentine (CGC). The second fixed difference within an exon was also in the GPRGR29-31 array (specifically GPRGR31) is a silent mutation, CTC (leu) - CTT (leu), at residue 62 in the translated sequence. The third fixed difference causes a change from argenine (AGA) to lysine (AAA) at residue 39 in GPROR38. The fourth and final fixed difference we detected within an exon causes a switch from glutamic acid (GAA) to aspartic acid (GAT) at the carboxy terminal of GPROR39.

In addition to identifying fixed differences between populations and highly differentiated exons, we performed a Bayesian F_{ST} scan to detect individuals SNPs under significant positive directional selection or stabilizing purifying selection (False Discovery

Rate <0.05). However the scan was unable to identify any significant selection, although the ten previously identified fixed differences formed an easily distinguishable cluster under positive (but not significant) selection in the analysis (Figure 6.3.4).
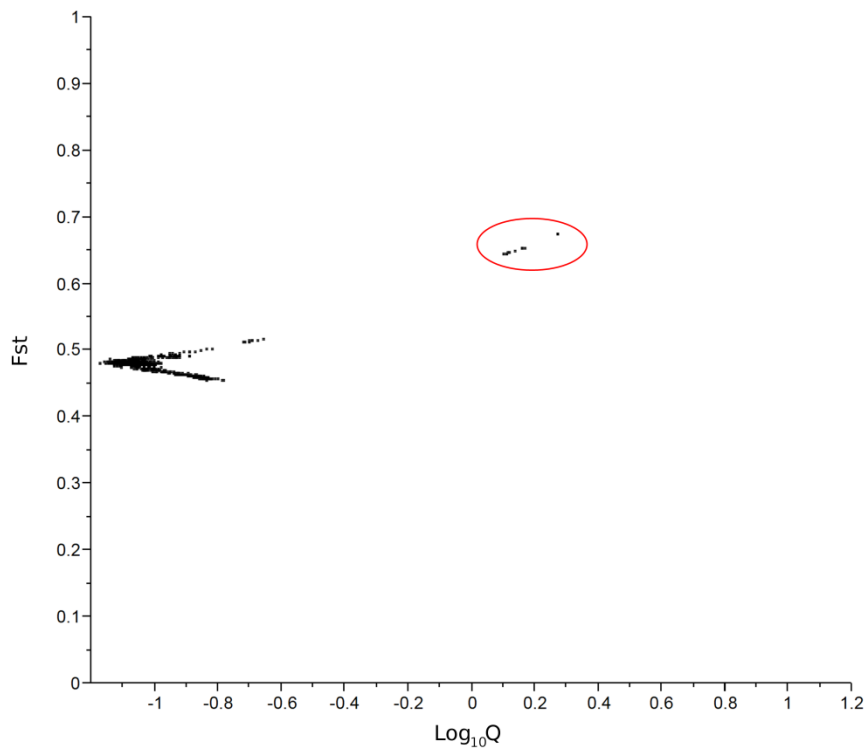
Figure 6.3.4: Bayesian F_{ST} scan to detect individuals SNPs under significant positive directional selection or stabilizing purifying selection. The cluster of results corresponding to the 10 identified fixed differences between lab and field population are highlighted in red. False Discovery Rate <0.05. No significant selection was detected for any SNP.

Raw Variant Data

Source	DF	F Ratio	P
Population	1	4.7587	0.0413*
Loci	4	17.48	<0.001*
Sequence Type	2	5.2442	0.0148*
Coverage	1	7.0074	0.0155*
No. of Chromosomes	1	0.9766	0.3348
			n=1016

High Confidence Subset

Source	DF	F Ratio	P
Population	1	7.0791	0.015*
Loci	4	14.7384	<.0001*
Sequence Type	2	4.3623	0.0268*
Coverage	1	8.8605	0.0075*
No. of Chromosomes	1	1.4199	0.2474
			n=750

Table 6.3.4: General Linear Model of the effects of Population (Lab, Field), Loci, Sequence Type (Exon, Intron, Non-coding), Coverage (mean number of short reads per chromosome) and Number of Chromosomes on the variability (SNPs/100bp) in the raw variant dataset (n=1016) and a subset of high confidence SNPs (n=750).

Chromosome	Loci	Sequence Type	Position	Canon (PEST)	Variant	Change
2L	*GPRGR29*	Non-Coding	39308286	T	A	
	GPRGR29	**Exon**	**39309403**	**A**	**G**	Histidine to Arginine
	GPRGR29	Intron	39310014	A	C	
	GPRGR31	**Exon**	**39312730**	**C**	**T**	Silent
	GPRGR31	Intron	39314109	C	A	
2R	*GPROR38*	Non-Coding	24857266	A	G	
	GPROR38	Intron	24857859	T	C	
	GPROR38	**Exon**	**24858476**	**G**	**A**	Arginine to Lysine
	GPROR38	**Exon**	**24858753**	**A**	**T**	Glutamic Acid to Aspartic Acid
	GPROR39	Intron	24851509	A	G	

Table 6.3.5: Fixed non-coding and coding genetic differences between populations identified by comparison of the population-specific assemblies at 5 loci located within islands of speciation of chromosomes 2L and 2R. Fixed differences located within exons are in bold.

6.4 Discussion

Elucidating the genetic basis of reproductive isolation, assortative mating and incipient speciation is of great importance for the control of malaria. Such population structuring can hinder the spread of insecticide resistance genes and, conversely, present a significant barrier to the introgression of transgenic resistance cassettes into target populations. The publication of the annotated *Anopheles gambiae* genome (Holt *et al* 2002) and the rapid expansion, and reduction in cost of next-generation sequencing technology in the latter half of the last decade (Ekblom *et al* 2011) means that broad ranging exploration of genetic differentiation between populations and subpopulations can be carried out on more samples and populations and by more research groups than ever before. Using post-genomic approaches, studies in the latter half of the last decade have indicated that genetic differentiation is maintained between the molecular forms of *Anopheles gambiae* despite contemporary gene flow (Turner *et al* 2005, Weetman *et al* 2011). So-called genomic islands of speciation found proximal to the centromeres of the X, 2L and 3L islands (Turner *et al* 2005, Turner and Hahn 2007, White *et al* 2010) are thought to harbour genes responsible for the reproductive isolation of the M- and S-molecular form and are a logical starting point for a 'reverse ecology' (Lawniczak *et al* 2010) approach to investigating the genetic basis of assortative mating behaviour.

Here, we carried out a small pilot study using ultra-deep 454 pyrosqeuencing to investigate whether genetic variability at potential assortative mating or speciation genes is lost during the colonization and lab rearing process. We were able to identify 750 high confidence SNPs across all amplicons, including 53 of the 90 SNPs present for the same regions in dbSNP.

We were able to quantify an overall reduction in genetic variability across all re-sequenced loci between field and lab samples (2.7-fold) that was broadly in agreement

with the findings of Norris *et al* (2001). This effect remained significant after the effects of variability of coverage were taken into account. Interestingly, the reduction in variability between the field and lab samples was not uniformly distributed across all loci: there was a significant reduction in variability in the lab sample for loci on chromosome 2R (3.95-fold reduction in variability), but no significant difference in variability between the field and lab samples for loci on chromosome 2L. Variability at all loci in the field sample (mean variability 3.65 SNP/bp) and at the 2L loci in the lab sample (4.87 SNP/100bp) was consistent with the rate of variability reported for the these regions in previous studies (2R: 3-4 SNP/100bp, Turner *et al* 2005, 2L: 4-5 SNP/bp White *et al* 2007) in field samples and with recent estimates of the rate of variability across a much larger area of the *An. gambiae* genome (~300kbp, 2.94 SNP/100bp, Wilding *et al* 2009). The reduction in variability in the 2R loci is consistent with the effects of a genetic bottlenecking and, given the questionable status of the 2R locus as a true speciation island (Turner and Hahn 2007) it seems likely that the observed reduction in variability is representative of the general loss in variability we would expect after laboratory colonisation. The maintenance of variability at field-sample levels in the lab sample 2L loci we successfully re-sequenced was intriguing. The loci considered here, *CPR34-64* and *GPRGR29-32* are found in the island of speciation identified by White *et al* (2007) associated with the chromosomal breakpoint distal to the centromere for the 2La inversion polymorphism and the 2La inversion is closely linked with aridity tolerance and is fixed in Mopti M-form populations in Mali (Coluzzi *et al* 1985). The fact that variability is maintained in this region despite a large reduction in variability in other parts of the genome may suggest that this loci are important for adaptation to colonisation in some way, although - without further evidence that the reduction in variability observed in the 2R is, as seems likely, observed elsewhere in the genome - interpretation of this result is difficult.

We were unable to amplify several regions of interest from within the previously identified centromeric islands of speciation. Given the poor coverage of the genome assembly in centromeric regions, combined with the high rate of variants in the *An. gambiae* genome reported here (~3.6 SNP/bp) and elsewhere (2.94 SNP/bp, Wilding *et al* 2009) it is likely that primer-site polymorphism played a large role in a failure to amplify any of these regions.

In addition to assessing the change in variability between Field and Lab populations, we were also able to identify 10 fixed SNPs between our two samples, this was low given the fixed differences observed between populations in these regions previously (Turner *et al* 2005, White *et al* 2007) and the length of time the lab sample had been colonised (~4 years) Four of the ten identified fixed SNPs were in coding regions, and the changes induced in the peptide sequence were investigated. Unfortunately, of our 4 coding fixed SNPs, one was a silent mutation and the remaining three were functionally synonymous and thus unlikely to induce a change in protein function. Furthermore a Bayesian scan for SNPs under selection failed to identify any loci under significant positive directional or stabilizing purifying selection. Although given the relatively small size of the sample - 17 kbp out of several Mbp of speciation islands, and the arbitrary nature of the selection process for target amplicons (see methods), discovering a fixed, functionally significant SNP was unlikely.

Despite the limitations imposed by our failure to amplify material from the centromeric regions implicated in reproductive isolation, using ultra-deep re-sequencing we were able to characterise 750 SNPs within the loci we studied of which 697 were previously unknown. Furthermore we were able to quantify the effects of colonisation on the genetic variability in two regions of the *An. gambiae* s.s. genome comprising the ORFs

and untranslated regions of 9 genes and identify a difference in the reduction in variability between two regions.

7. General Discussion

Despite the success of current malaria control strategies and the concomitant reduction in malaria deaths we have seen over the last decade, the spectre of the past recrudescence of malaria following a previous, successful, global control program casts a heavy shadow over any progress made in controlling the disease today. Given that the development of resistance to the current countermeasures for both parasite and vector is an evolutionary inevitability, and the development of new functional classes of insecticides and antimalarial drugs is a necessarily slow process, there is a demonstrable need to broaden the toolset for vector control beyond the classic paradigm of physical barriers and insecticidal substances. Although it could hardly be referred to as a new idea, genetic control - whether enhanced with modern transgenic techniques or relying on 'classical' radio- or chemosterilisation - represents a genuinely promising alternative approach to vector control but also presents a number of problems that must be overcome before it can be used as an effective tool. First and foremost amongst these is ensuring that a modified mosquito is capable of performing the task required of it: namely to persist and (out) compete with its wild counterparts for mates, repeatedly if possible. History tells us that this is not as straightforward as simply raising, sterilising and releasing hundreds of thousands of mosquitoes - although there were some notable successes with this approach - and for this reason a broad understanding of how chemical, radio or genetic modification affects its target, and what effects the conditions in which we maintain our mosquito stocks has on their ability to perform in the field are important aspects of the development of genetic control strategies.

In this thesis, we have presented data from a series of experiments investigating the fitness and mating competitiveness of laboratory and transgenic strains of *Anopheles gambaie* sensu stricto. Using a broad range of experimental techniques from classical lab-

based ecological comparisons to cutting-edge genomic techniques (via a tent in a field in Mali) we have investigated the effects of transgenesis and inbreeding on the performance of two transgenic strains in the lab and attempted to elucidate some of the environmental and genetic factors determining mating behaviour and survival in colonised and field-derived mosquitoes.

We detected a significant fitness burden in the transgenic 'phase 1' EE strain - expressing ECFP facultatively in the eyes and other nerve tissue - during larval development relative to the wild-type Mopti. This effect was apparent only in homozygotes and was observed in both the short-term, single generation larval development experiment and the multi-generation long term stability experiment. Despite this cost, the EE transgenic cassette was stable in mixed populations. Egg production data indicated that EE laid more eggs than either EVida3 or Mopti and it may be the case that reduced homozygous larval survival in the long term stability experiment was offset by increased egg production. EE laid smaller eggs on average compared to Mopti and EVida3, but this did not appear to either effect long term transgene stability (EVida3 were significantly less fit - see below - despite laying larger eggs) or indeed to be correlated with variables such as brood and body size. EVida3 was found to be broadly unfit during larval development (and only during larval development), both as a homozygote and as a hemizygote. The EVida3 transgenic cassette became extinct in all replicates of the long-term stability experiment. The partitioning of this strong, and apparently co-dominant fitness effect to larval development only is curious, as the only difference between the EE and EVida3 strains is the addition of ~7kb of transgenic material containing [3xP3,*Ds*Red] the *Agcp* regulatory sequences and the sequence encoding the Vida3 tetramer. Clearly there were some shared fitness effects observed in both strains: reduced larval survival relative to wild-type, which is likely due to a deleterious - and recessive - hitchhiker shared

between both transgenics, and other shared traits such as faster larval development relative to Mopti - almost certainly an adaptation to 30+ years of laboratory rearing. In chapter 4 we discussed a number of possible explanations for the clear difference in fitness between the EE and EVida3 strains. One of the possibilities discussed was the potential for 'leaky' expression of the AMP or expression outside of the expected profile of the promoter used. This is a plausible explanation, as the site of transgene integration has been shown to affect the spatial and temporal specificity of expression in transgenic *Anopheles stephensi* (Amenya *et al* 2010), however, neither the Vida3 tetramer mRNA transcript nor the peptide itself have been recovered from larval stage (J Meredith personal communication) and we detected no differential fitness costs when rearing Mopti larvae at high density in medium harvested from mature EVida3 larval rearing trays - although we acknowledge that this would also require the additional step of excretion of the peptide in addition to non-standard expression. A second possibility is that maternal expression of the Vida3 AMP has caused a reduction in blood meal utilisation and the maternal energetic investment in individual larvae, leading to a reduction in fitness in larvae independent of larval genotype and instead determined by the maternal genotype. This explanation has the advantage of being dependent on the normal expression profile of Vida3 and is supported tangentially by the fact that we did observe a reduction in brood size in the EVida3 in relation to the EE, which may be an indication that EVida3 females were utilising their blood meals less efficiently compared to the otherwise similar EE. However, despite the smaller brood size, EVida3 females did lay significantly larger eggs than EE, and - whilst there is no quantitative data about the relationship of within- or between-strain egg size variation and larval performance in mosquitoes in the literature, it is logical to assume that a larger egg results in a 'fitter' larva - although this would be an interesting avenue to investigate in further study. The maternal investment theory is lent some additional weight

by a study by Grech *et al* (2007) which demonstrated that the body condition of female *An. stephensi* affected the fecundity of their female offspring and is worthy of further investigation. Finally, there is a small body of literature that has demonstrated the toxic effects of fluorophore expression and/or excitation, this is again, an intriguing possibility, and if the case has potentially far-reaching consequences due to the extensive use of fluorophores as phenotypic markers in transformation. However, as almost all transgenic mosquitoes carry a fluorescent phenotypic marker, partitioning the effects of - for example - ECFP expression and excitation from deleterious allele linkage and more general inbreeding depression is difficult. Indeed, if ECFP expression does have a deleterious effect on the mosquito, it may account for the both the reduction in homozygous EE larval fitness and the increased magnitude of the fitness effect observed in EVida3, due to the increased number of phenotypic markers being expressed in that strain.

Our work investigating transgenesis presented in chapters 3 and 4 represent the first detailed study of the fitness of transgenic *Anopheles gambiae*: previous studies having considered - in terms of malaria vectors - almost exclusively the Asian malaria vector *An. stephensi* and - in the wider context of transgenic mosquitoes - *Aedes aegypti*. The development and study of transgenic *An. gambiae* is vital for the development of tools for malaria control within the key sub-Saharan Africa setting as this is where the greatest burden of mortality and morbidity is suffered. The strains described and assessed above represent an important milestone not only for transgenesis within *An. gambiae* - an organism that is notoriously difficult to transform (J Meredith, pers. comm.) - but also for the innovative and potentially very powerful *phi*C31 site-specific integration technique. The lack of any effective fitness costs associated with carrying the phase 1 docking site cassette makes the EE strain an ideal platform for testing the efficacy of putative anti-malarial peptide products, new regulatory sequences for controlling expression and

laterally, any fitness costs induced by these elements. Current comparisons of expression profiles and AMP efficacy between independently developed transgenic strains are confounded by positional effects caused by the random insertion of these cassettes into the mosquito genome. Using the 2-phase system described above removes these confounding factors and - theoretically - allows for the rapid development of, transformation with, and iteration on different effector gene/promoter combinations. The observed lack of fitness observed in the phase 2 transformant - EVida3 - whilst in and of itself unsurprising given the abundance of examples in the literature of fitness negative transformation of mosquitoes, was surprising in the manner in which those fitness cost were manifest. Specifically that the greatest fitness burden was observed during larval development; a time when the anti-malarial effector gene is not actively expressed. Ours was the first study to not only consider changes in transgene frequency over multiple generations but also over key developmental milestones within generations and the fact that the we detected the greatest fitness load during larval development validates this approach and may be a worthwhile approach in future studies of transgenic fitness. The implications of this are less clear but answers to the key questions raised by these data - namely what exactly is causing the fitness burden during development - will be answered in future studies investigating fitness of new transgenic strains created from the EE platform.

The second theme presented in this thesis was an exploration of the genetic and environmental factors associated with colonisation, and their effects on the genetic variation, survival, mating success and assortative mating behaviour of lab- and field-derived samples of a Mopti M-form population from Mali, West Africa. We observed significant interactions between genetic and environmental effects on the survival and mating success of mosquitoes within our semi-field system, primarily demonstrating that field-derived, field-reared individuals performed best in this setting - a result with

important consequences for future release strategies and field-based ecological experiments, particularly given how poor the performance of field derived, lab reared individuals was in some comparisons in this setting. When we investigated assortative mating behaviour, we observed a complete breakdown in assortative mating behaviour in lab-colonised Mopti M-form mosquitoes raised in the insectary, whilst rearing mosquitoes from the same strain in the more field-like conditions of the semi-field system restored the assortative mating phenotype. The principal differences in environmental conditions between rearing in the lab and the field consisted principally of changes in the intensity of incident light, large differences in the daily temperature cycle and differential exposure to swarm markers. We speculated (see chapter 6) that the differences in observed assortative mating behaviour between treatment cohorts may be due changes in peak mate-seeking activity periods bringing Lab/Lab mosquitoes into closer contact temporally and spatially with S-form individuals. This is a potentially important result for two reasons: 1) it demonstrates a clear interaction between genotype and larval rearing and adult maintenance conditions in determining assortative mating behaviour. 2) It presents the possibility that assessments of genetic differentiation and fixed differences between the genomes of field- and lab-derived samples may allow us to determine some of the genetic mechanisms that control (in part) assortative mating behaviour between molecular forms. The reduction in survival and mating success in this field-like context in the lab-derived and lab reared groups relative to field-derived field-reared mosquitoes has potentially important consequences for the design and implementation of mass rearing protocols for future transgenic control strategies. Mass-rearing mosquitoes for release in 'standard' laboratory conditions may not be sufficient to guarantee effective control even given an effective sterile or anti-malarial transgenic strain. Although this problem is theoretically ameliorated by simply increasing the number of released males keeping operating costs to

a minimum will be a large part of what determines the effective success of any future control programme, thus investigating which factors improve survival and competitiveness should feed back into maximising the effectiveness of a control strategy. Conversely, whilst the reduced survival and mating success in lab-derived and lab-reared mosquitoes could negatively affect release programs the fact that we observed the breakdown of assortative mating behaviour in lab/lab mosquitoes may have important, positive consequences for genetic control strategies. One of the potential weaknesses of sterile or transgenic insect control for *Anopheles* mosquitoes, particularly in the Afrotropical setting, is the extremely complex nature of population structuring and the pattern of gene flow within and between members of the *An. gambiae* s.l. species complex. This complex structuring means that interventions targeting a single member of the complex (e.g. M-form *An. gambiae* s.s.), are vulnerable to failure by, for example, the expansion of untargeted sibling-species into areas of SIT-surpression or through the failure of anti-malarial effector genes to spread through patchy populations of the targeted vector. Developing a behaviourally 'promiscuous' strain capable of mating with multiple groups within the complex could potentially sidestep this problem without the need to develop population specific interventions for each targeted region. Clearly doing so reliably requires an understanding of what precisely determines the choosiness of mosquitoes in this context. Having demonstrated that assortative mating does indeed break down under certain rearing and genetic conditions (chapter 5) the next objective is to determine what feature or features determine this behaviour. Our limited investigation into genetic differentiation and changes in variability between lab- and field-derived samples presented some potentially leading results in that direction given that we observed a reduction in variability between field and lab samples in re-sequenced loci on the 2R chromosome (in line with previous estimates of the loss of genetic variability following lab colonisation),

but did not observe a change in variability at the loci within the 2La inversion. This is precisely the kind of pattern of variability we would expect in areas containing genes that are important for survival and persistence within the lab including those governing dessication resistance and the behavioural reponses underpinning mate selection. Given the extremely limited area considered, definitive statements are difficult and clearly a wider investigation will be required to get a more complete picture of the pattern of variability-loss as mosquitoes are colonised. Over the past 4 years, sequencing technology has developed to the point where even relatively small research groups are now able to sequence and datamine multiple genomes, whilst this means that the approach used here is nearing obsolescence, further studies underway now may allow us to consider these questions in unprecedented detail.

In the 15 years since the creation of the first transgenic mosquito, WHO-backed interventions leveraging artemesenin derivative chemotherapy and widespread deployment of ITMs - primarily permethrin-treated bed nets - has seen an impressive reduction in malaria mortality and morbidity. With a number of countries - for example Sri Lanka - preparing to declare themselves 'malaria-free' for the first time, top-down policy discussions are starting to switch from discussing control to the prospect of elimination of malaria as a public health problem. Given the rise and spread of drug and insecticide resistance, achieving this lofty (and laudable) goal will become increasingly difficult without new tools to complement these (currently) potent interventions, particularly as the focus of elimination switches to the more challenging and complex geographical contexts of continental Africa, and the mainland Indian subcontinent and far East Asia. Transgenic approaches represent a potentially effective addition to current malaria control strategies. Although the full deployment of a population replacement-based transgenic intervention is still some way off - a truly effective transgenic product capable of completely blocking

malaria transmission has yet to be demonstrated *in vivo* - and there are some serious environmental concerns surrounding the concept of inducing a permanent change in a natural population which *must* be addressed before even a trial release of a transgenic strain carrying malaria-blocking transgenes, enormous progress has already been made towards developing an effective transgenic malaria-blocking strain. Exactly how effective an intervention this approach ultimately proves to be however, remains to be seen. What is increasingly clear is that the transgenic approach is unlikely to represent a complete panacea against malaria - despite the more hyperbolic claims of the popular press and some individuals associated with the field - and must operate within the context of an integrated, multi-focus approach to control. Nevertheless transgenesis has the potential to become a valuable tool in the fight against malaria and, with this thesis, it is our hope that we have contributed in some way towards developing an effective intervention.

8. References

Abraham, E G, M Donnelly-Doman, H Fujioka, A Ghosh, L Moreira, and M Jacobs-Lorena. 2005. "Driving Midgut-Specific Expression and Secretion of a Foreign Protein in Transgenic Mosquitoes with AgAper1 Regulatory Elements." Insect Molecular Biology 14(3): 271–79.

Allen, Margaret L, and Philip J Scholl. 2005. "Quality of Transgenic Laboratory Strains of Cochliomyia Hominivorax (Diptera: Calliphoridae)." Journal of Economic Entomology 98(6): 2301–6.

Alphey, Luke. 2002. "Re-Engineering the Sterile Insect Technique." Insect Biochemistry and Molecular Biology 32(10): 1243–47.

Alphey, Luke, C Ben Beard, Peter Billingsley, Maureen Coetzee, Andrea Crisanti, Chris Curtis, Paul Eggleston, Charles Godfray, Janet Hemingway, Marcelo Jacobs-Lorena, Anthony A James, Fotis C Kafatos, Louis G Mukwaya, Michael Paton, Jeffrey R Powell, William Schneider, Thomas W Scott, Barbara Sina, Robert Sinden, Steven Sinkins, Andrew Spielman, Yeya Touré, and Frank H Collins. 2002. "Malaria Control with Genetically Manipulated Insect Vectors." Science 298(5591): 119–21.

Amenya, D A, M Bonizzoni, A T Isaacs, N Jasinskiene, H Chen, O Marinotti, G Yan, and A A James. 2010a. "Comparative Fitness Assessment of Anopheles Stephensi Transgenic Lines Receptive to Site-Specific Integration." Insect Molecular Biology 19(2): 263–69.

Amenya, D A, M Bonizzoni, A T Isaacs, N Jasinskiene, H Chen, O Marinotti, G Yan, and A A James. 2010b. "Comparative Fitness Assessment of Anopheles Stephensi Transgenic Lines Receptive to Site-Specific Integration." Insect Molecular Biology 19(2): 263–69.

Andreasen, M H, and C F Curtis. 2005. "Optimal Life Stage for Radiation Sterilization of Anopheles Males and Their Fitness for Release." Medical and Veterinary Entomology 19(3): 238–44.

Armbruster, P, R A Hutchinson, and T Linvell. 2000. "Equivalent Inbreeding Depression Under Laboratory and Field Conditions in a Tree-Hole-Breeding Mosquito." Proceedings of the Royal Society of London. Series B: Biological Sciences 267(1456): 1939–45.

Arrighi, Romanico B G, Chikashi Nakamura, Jun Miyake, Hilary Hurd, and J Grant Burgess. 2002. "Design and Activity of Antimicrobial Peptides Against Sporogonic-Stage Parasites Causing Murine Malarias." Antimicrobial Agents and Chemotherapy 46(7): 2104–10.

Baens, Mathijs, Heidi Noels, Vicky Broeckx, Sofie Hagens, Sabine Fevery, An D Billiau, Hugo Vankelecom, and Peter Marynen. 2006. "The Dark Side of EGFP: Defective Polyubiquitination." PLoS ONE 1: e54.

Baker, R H, W K Reisen, R K Sakai, C G Hayes, M Aslamkhan, U T Saifuddin, F Mahmood, A VEEN, and S Javed. 1979. "Field Assessment of Mating Competitiveness of Male Culex Tritaeniorhynchus Carrying a Complex Chromosomal Aberration." Annals of the Entomological Society of America 72(6): 751–58.

Balestrino, Fabrizio, Jérémie R L Gilles, Sharon M Soliban, Anton Nirschl, Quentin E Benedict, and Mark Q Benedict. 2011. "Mosquito Mass Rearing Technology: a Cold-Water Vortex Device for Continuous Unattended Separation of Anopheles Arabiensis Pupae From Larvae." Journal of the American Mosquito Control Association 27(3): 227–35.

Baqar, S, C G Hayes, and T Ahmed. 1980. "The Effect of Larval Rearing Conditions and Adult Age on the Susceptibility of Culex Tritaenio-Rhynchus to Infection with West Nile Virus.." Journal of the American Mosquito Control Association 40(2): 165–71.

Bargielowski, Irka, Luke Alphey, and Jacob C Koella. 2011. "Cost of Mating and Insemination Capacity of a Genetically Modified Mosquito Aedes Aegypti OX513A Compared to Its Wild Type Counterpart." PLoS ONE 6(10): e26086.

Barnes, M J, N F Lobo, M B Coulibaly, N F Sagnon, C Costantini, and N J Besansky. 2005. "SINE Insertion Polymorphism on the X Chromosome Differentiates Anopheles Gambiae Molecular Forms." Insect Molecular Biology 14(4): 353–63.

Benedict, Mark Q, and Alan S Robinson. 2003. "The First Releases of Transgenic Mosquitoes: an Argument for the Sterile Insect Technique." Trends in Parasitology 19(8): 349–55.

Benedict, Mark Q, Bart G J Knols, Hervé C Bossin, Paul I Howell, Eric Mialhe, Carlos Caceres, and Alan S Robinson. 2009. "Colonisation and Mass Rearing: Learning From Others." Malaria Journal 8 Suppl 2: S4.

Billker, O, M K Shaw, G Margos, and R E Sinden. 1997. "The Roles of Temperature, pH and Mosquito Factors as Triggers of Male and Female Gametogenesis of Plasmodium Berghei in Vitro." Parasitology 115 (Pt 1): 1–7.

Billker, O, V Lindo, M Panico, A E Etienne, T Paxton, A Dell, M Rogers, R E Sinden, and H R Morris. 1998. "Identification of Xanthurenic Acid as the Putative Inducer of Malaria Development in the Mosquito." Nature 392(6673): 289–92.

Boulesteix, M, F Simard, C Antonio-Nkondjio, H P Awono-Ambene, D Fontenille, and C Biémont. 2007. "Insertion Polymorphism of Transposable Elements and Population

Structure of Anopheles Gambiae M and S Molecular Forms in Cameroon." Molecular Ecology 16(2): 441–52.

Brown, A W, J Haworth, and A R Zahar. 1976. "Malaria Eradication and Control From a Global Standpoint." Journal of Medical Entomology 13(1): 1–25.

Cassone, Bryan J, Karine Mouline, Matthew W Hahn, Bradley J White, Marco Pombi, Frederic Simard, Carlo Costantini, and Nora J Besansky. 2008. "Differential Gene Expression in Incipient Species of Anopheles Gambiae." Molecular Ecology 17(10): 2491–2504.

Catteruccia, Flaminia. 2007. "Malaria Vector Control in the Third Millennium: Progress and Perspectives of Molecular Approaches." Pest Management Science 63(7): 634–40.

Catteruccia, Flaminia, H Charles J Godfray, and Andrea Crisanti. 2003. "Impact of Genetic Manipulation on the Fitness of Anopheles Stephensi Mosquitoes." Science 299(5610): 1225–27.

Charlwood, J D, and MDR Jones. 1980. "Mating in the Mosquito, Anopheles Gambiae Sl." Physiological Entomology 5(4): 315–20.

Coleman, P G, and L Alphey. 2004. "Genetic Control of Vector Populations: an Imminent Prospect.." Tropical Medicine and International Health 9(4): 433.

Coluzzi, M. 1992. "Malaria Vector Analysis and Control." Parasitology Today (Personal ed) 8(4): 113–18.

Coluzzi, M. 2002. "A Polytene Chromosome Analysis of the Anopheles Gambiae Species Complex." Science 298(5597): 1415–18.

Coluzzi, M, A Sabatini, V Petrarca, and M A Di Deco. 1979. "Chromosomal Differentiation and Adaptation to Human Environments in the Anopheles Gambiae Complex." Transactions of the Royal Society of Tropical Medicine and Hygiene 73(5): 483–97.

Coluzzi, M, V Petrarca, and M A Di Deco. 1985. "Chromosomal Inversion Intergradation and Incipient Speciation in Anopheles Gambiae." Bollettino di Zoologia 52(1-2): 45–63.

Dame, David A, Christopher F Curtis, Mark Q Benedict, Alan S Robinson, and Bart G J Knols. 2009. "Historical Applications of Induced Sterilisation in Field Populations of Mosquitoes." Malaria Journal 8 Suppl 2: S2.

Dao, Adama, Abdoulaye Adamou, Alpha Seydou Yaro, Hamidou Moussa Maïga, Yaya Kassogue, Sékou Fantamady Traoré, and Tovi Lehmann. 2008. "Assessment of Alternative Mating Strategies in Anopheles Gambiae: Does Mating Occur Indoors?." Journal of Medical Entomology 45(4): 643–52.

Davidson, G. 1969. "The Potential Use of Sterile Hybrid Males for the Eradication of Member Species of the Anopheles Gambiae Complex." Bulletin of the World Health Organization.

Diabate, Abdoulaye, Adama Dao, Alpha S Yaro, Abdoulaye Adamou, Rodrigo Gonzalez, Nicholas C Manoukis, Sékou F Traoré, Robert W Gwadz, and Tovi Lehmann. 2009. "Spatial Swarm Segregation and Reproductive Isolation Between the Molecular Forms of Anopheles Gambiae." Proceedings Biological sciences / The Royal Society 276(1676): 4215–22.

Diabate, Abdoulaye, Roch K Dabiré, Kyle Heidenberger, Jacob Crawford, William O Lamp, Lauren E Culler, and Tovi Lehmann. 2008. "Evidence for Divergent Selection

Between the Molecular Forms of Anopheles Gambiae: Role of Predation." BMC Evolutionary Biology 8: 5.

Dixit, Ram, and Richard Cyr. 2003. "Cell Damage and Reactive Oxygen Species Production Induced by Fluorescence Microscopy: Effect on Mitosis and Guidelines for Non-Invasive Fluorescence Microscopy." The Plant Journal: for Cell and Molecular Biology 36(2): 280–90.

Edillo, Frances E, Yeya T Touré, Gregory C Lanzaro, Guimogo Dolo, and Charles E Taylor. 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.

Edwards, M J, F J Lemos, M Donnelly-Doman, and M Jacobs-Lorena. 1997. "Rapid Induction by a Blood Meal of a Carboxypeptidase Gene in the Gut of the Mosquito Anopheles Gambiae." Insect Biochemistry and Molecular Biology 27(12): 1063–72.

Ekblom, R, and J Galindo. 2011. "Applications of Next Generation Sequencing in Molecular Ecology of Non-Model Organisms." Heredity 107(1): 1–15.

Escalante, A A, E Barrio, and F J Ayala. 1995. "Evolutionary Origin of Human and Primate Malarias: Evidence From the Circumsporozoite Protein Gene." Molecular Biology and Evolution 12(4): 616–26.

Esnault, Caroline, Matthieu Boulesteix, Jean Bernard Duchemin, Alphonsine A Koffi, Fabrice Chandre, Roch Dabiré, Vincent Robert, Frederic Simard, Frédéric Tripet, Martin J Donnelly, Didier Fontenille, Christian Biémont, and Philip Awadalla. 2008. "High Genetic Differentiation Between the M and S Molecular Forms of Anopheles Gambiae in Africa." PLoS ONE 3(4): e1968.

Fanello, C, F Santolamazza, and A Della Torre. 2002. "Simultaneous Identification of Species and Molecular Forms of the Anopheles Gambiae Complex by PCR-RFLP." Medical and Veterinary Entomology 16(4): 461–64.

Fanello, C, V Petrarca, A Della Torre, F Santolamazza, G Dolo, M Coulibaly, A Alloueche, C F Curtis, Y T Touré, and M Coluzzi. 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): 241–45.

Favia, G, A Della Torre, M Bagayoko, A Lanfrancotti, N Sagnon, Y T Touré, and M Coluzzi. 1997. "Molecular Identification of Sympatric Chromosomal Forms of Anopheles Gambiae and Further Evidence of Their Reproductive Isolation." Insect Molecular Biology 6(4): 377–83.

Favia, G, G Dimopoulos, and C Louis. 1994. "Analysis of the Anopheles Gambiae Genome Using RAPD Markers." Insect Molecular Biology 3(3): 149–57.

Feldmann, U, and J Hendrichs. 2001. "Integrating the Sterile Insect Technique as a Key Component of Area-Wide Tsetse and Trypanosomiasis Invention." PAAT Technical and Scientific Series (FAO) 3.

Ferguson, Heather M, Kija R Ng'habi, Thomas Walder, Demetrius Kadungula, Sarah J Moore, Issa Lyimo, Tanya L Russell, Honorathy Urassa, Hassan Mshinda, Gerry F Killeen, and Bart Gj Knols. 2008. "Establishment of a Large Semi-Field System for Experimental Study of African Malaria Vector Ecology and Control in Tanzania." Malaria Journal 7: 158.

Fivelman, Quinton L, Louisa McRobert, Sarah Sharp, Cathy J Taylor, Maha Saeed, Claire A Swales, Colin J Sutherland, and David A Baker. 2007. "Improved Synchronous

Production of Plasmodium Falciparum Gametocytes in Vitro." Molecular & Biochemical Parasitology 154(1): 119–23.

Gentile, G, M Slotman, V Ketmaier, J R Powell, and A Caccone. 2001. "Attempts to Molecularly Distinguish Cryptic Taxa in Anopheles Gambiae S.S." Insect Molecular Biology 10(1): 25–32.

Gentile, Gabriele, Alessandra della Torre, Bertha Maegga, Jeffrey R Powell, and Adalgisa Caccone. 2002. "Genetic Differentiation in the African Malaria Vector, Anopheles Gambiae S.S., and the Problem of Taxonomic Status." Genetics 161(4): 1561–78.

Geurts, Aron M, Ying Yang, Karl J Clark, Geyi Liu, Zongbin Cui, Adam J Dupuy, Jason B Bell, David A Largaespada, and Perry B Hackett. 2003. "Gene Transfer Into Genomes of Human Cells by the Sleeping Beauty Transposon System." Molecular Therapy : the Journal of the American Society of Gene Therapy 8(1): 108–17.

Ghosh, A K, P E Ribolla, and M Jacobs-Lorena. 2001. "Targeting Plasmodium Ligands on Mosquito Salivary Glands and Midgut with a Phage Display Peptide Library." Proceedings Of The National Academy Of Sciences Of The United States Of America 98(23): 13278–81.

Ghosh, A, M J Edwards, and M Jacobs-Lorena. 2000. "The Journey of the Malaria Parasite in the Mosquito: Hopes for the New Century." Parasitology Today (Personal ed) 16(5): 196–201.

Gilles, J R L, R S Lees, S M Soliban, and M Q Benedict. 2011. "Density-Dependent Effects in Experimental Larval Populations of Anopheles Arabiensis (Diptera: Culicidae) Can Be Negative, Neutral, or Overcompensatory Depending on Density and Diet Levels." Journal of Medical Entomology 48(2): 296–304.

Gimnig, John E, Maurice Ombok, Samson Otieno, Michael G Kaufman, John M Vulule, and Edward D Walker. 2002. "Density-Dependent Development of Anopheles Gambiae (Diptera: Culicidae) Larvae in Artificial Habitats." Journal of Medical Entomology 39(1): 162–72.

Grassi, G. (1900). "Studi di uno Zoologo sulla Malaria " Mem Rend Accad Lincei 53: 299-502.

Grech, Katrina, Liam Aye Maung, and Andrew F Read. 2007. "The Effect of Parental Rearing Conditions on Offspring Life History in Anopheles Stephensi." Malaria Journal 6: 130.

Grover, K K, C F Curtis, V P Sharma, KRP Singh, K Dietz, H V Agarwal, R K Razdan, and V Vaidyanathan. 1976. "Competitiveness of Chemosterilized Males and Cytoplasmically Incompataible Translocated Males of Culex Pipiens Fatigans Widedemann (Diptera: Culicidae) in the Field." Bulletin of Entomological Research 66(03): 469–80.

Han, Yeon Soo, and Carolina Barillas-Mury. 2002. "Implications of Time Bomb Model of Ookinete Invasion of Midgut Cells." Insect Biochemistry and Molecular Biology 32(10): 1311–16.

Harris, Angela F, Derric Nimmo, Andrew R McKemey, Nick Kelly, Sarah Scaife, Christl A Donnelly, Camilla Beech, William D Petrie, and Luke Alphey. 2011. "Field Performance of Engineered Male Mosquitoes." Nature Biotechnology 29(11): 1034–37.

Hassan, M, W M El-Motasim, R T Ahmed, and B B El-Sayed. 2010. "Prolonged Colonisation, Irradiation, and Transportation Do Not Impede Mating Vigour and

Competitiveness of Male Anopheles Arabiensis Mosquitoes Under Semi-Field Conditions in Northern Sudan." MalriaWorld Journal 1(1): 1.

Hay, Simon I, Carlos A Guerra, Peter W Gething, Anand P Patil, Andrew J Tatem, Abdisalan M Noor, Caroline W Kabaria, Bui H Manh, Iqbal R F Elyazar, Simon Brooker, David L Smith, Rana A Moyeed, and Robert W Snow. 2009. "A World Malaria Map: Plasmodium Falciparum Endemicity in 2007." PLoS Medicine 6(3): e1000048.

Helinski, M E H, and B G J Knols. 2009. "The Influence of Late-Stage Pupal Irradiation and Increased Irradiated: Un-Irradiated Male Ratio on Mating Competitiveness of the Malaria Mosquito Anopheles Arabiensis Patton." Bulletin of Entomological Research 99(3): 317–22.

Hogg, J C, and H Hurd. 1995. "Malaria-Induced Reduction of Fecundity During the First Gonotrophic Cycle of Anopheles Stephensi Mosquitoes." Medical and Veterinary Entomology 9(2): 176–80.

Hogg, J C, M C Thomson, and H Hurd. 1996. "Comparative Fecundity and Associated Factors for Two Sibling Species of the Anopheles Gambiae Complex Occurring Sympatrically in the Gambia." Medical and Veterinary Entomology 10(4): 385–91.

Holt, Robert A, G Mani Subramanian, Aaron Halpern, Granger G Sutton, Rosane Charlab, Deborah R Nusskern, Patrick Wincker, Andrew G Clark, José M C Ribeiro, Ron Wides, Steven L Salzberg, Brendan Loftus, Mark Yandell, William H Majoros, Douglas B Rusch, Zhongwu Lai, Cheryl L Kraft, Josep F Abril, Veronique Anthouard, Peter Arensburger, Peter W Atkinson, Holly Baden, Veronique de Berardinis, Danita Baldwin, Vladimir Benes, Jim Biedler, Claudia Blass, Randall Bolanos, Didier Boscus, Mary Barnstead, Shuang Cai, Center, Angela, Kabir Chaturverdi, George K Christophides, Mathew A

Chrystal, Michele Clamp, Anibal Cravchik, Val Curwen, Ali Dana, Art Delcher, Ian Dew, Cheryl A Evans, Michael Flanigan, Anne Grundschober-Freimoser, Lisa Friedli, Zhiping Gu, Ping Guan, Roderic Guigo, Maureen E Hillenmeyer, Susanne L Hladun, James R Hogan, Young S Hong, Jeffrey Hoover, Olivier Jaillon, Zhaoxi Ke, Chinnappa Kodira, Elena Kokoza, Anastasios Koutsos, Ivica Letunic, Alex Levitsky, Yong Liang, Jhy-Jhu Lin, Neil F Lobo, John R Lopez, Joel A Malek, Tina C McIntosh, Stephan Meister, Jason Miller, Clark Mobarry, Emmanuel Mongin, Sean D Murphy, David A O'brochta, Cynthia Pfannkoch, Rong Qi, Megan A Regier, Karin Remington, Hongguang Shao, Maria V Sharakhova, Cynthia D Sitter, Jyoti Shetty, Thomas J Smith, Renee Strong, Jingtao Sun, Dana Thomasova, Lucas Q Ton, Pantelis Topalis, Zhijian Tu, Maria F Unger, Brian Walenz, Aihui Wang, Jian Wang, Mei Wang, Xuelan Wang, Kerry J Woodford, Jennifer R Wortman, Martin Wu, Alison Yao, Evgeny M Zdobnov, Hongyu Zhang, Qi Zhao, Shaying Zhao, Shiaoping C Zhu, Igor Zhimulev, Mario Coluzzi, Alessandra della Torre, Charles W Roth, Christos Louis, Francis Kalush, Richard J Mural, Eugene W Myers, Mark D Adams, Hamilton O Smith, Samuel Broder, Malcolm J Gardner, Claire M Fraser, Ewan Birney, Peer Bork, Paul T Brey, J Craig Venter, Jean Weissenbach, Fotis C Kafatos, Frank H Collins, and Stephen L Hoffman. 2002. "The Genome Sequence of the Malaria Mosquito Anopheles Gambiae." Science 298(5591): 129–49.

Horn, C, and E A Wimmer. 2000. "A Versatile Vector Set for Animal Transgenesis." Development Genes and Evolution 210(12): 630–37.

Howell, P I, and M Q Benedict. 2009. "Mating Competitiveness of Anopheles Arabiensis Males as a Function of Transgenic State and Genetic Similarity to Females." Journal of Insect Behaviour.

Huho, B J, K R Ng'habi, G F Killeen, G Nkwengulila, B G J Knols, and H M Ferguson. 2007. "Nature Beats Nurture: a Case Study of the Physiological Fitness of Free-Living and Laboratory-Reared Male Anopheles Gambiae S.L." The Journal of Experimental Biology 210(Pt 16): 2939–47.

Hunt, R H, M Coetzee, and M Fettene. 1998. "The Anopheles Gambiae Complex: a New Species From Ethiopia." Transactions of the Royal Society of Tropical Medicine and Hygiene 92(2): 231–35.

Hurd, H, V Carter, and A Nacer. 2005. "Interactions Between Malaria and Mosquitoes: the Role of Apoptosis in Parasite Establishment and Vector Response to Infection." Current Topics in Microbiology and Immunology 289: 185–217.

Irvin, Nic, Mark S Hoddle, David A O'brochta, Bryan Carey, and Peter W Atkinson. 2004. "Assessing Fitness Costs for Transgenic Aedes Aegypti Expressing the GFP Marker and Transposase Genes." Proceedings Of The National Academy Of Sciences Of The United States Of America 101(3): 891–96.

Isaacs, Alison T, Fengwu Li, Nijole Jasinskiene, Xiaoguang Chen, Xavier Nirmala, Osvaldo Marinotti, Joseph M Vinetz, and Anthony A James. 2011. "Engineered Resistance to Plasmodium Falciparum Development in Transgenic Anopheles Stephensi." PLoS Pathogens 7(4): e1002017.

Ito, J, A Ghosh, L A Moreira, E A Wimmer, and M Jacobs-Lorena. 2002. "Transgenic Anopheline Mosquitoes Impaired in Transmission of a Malaria Parasite." Nature 417(6887): 452–55.

Jacobs-Lorena, Marcelo. 2003. "Interrupting Malaria Transmission by Genetic Manipulation of Anopheline Mosquitoes." Journal of Vector Borne Diseases 40(3-4): 73–77.

Janse, C J, P F van der Klooster, H J van der Kaay, M van der Ploeg, and J P Overdulve. 1986. "DNA Synthesis in Plasmodium Berghei During Asexual and Sexual Development." Molecular & Biochemical Parasitology 20(2): 173–82.

Jasinskiene, N, C J Coates, M Q Benedict, A J Cornel, C S Rafferty, A A James, and F H Collins. 1998. "Stable Transformation of the Yellow Fever Mosquito, Aedes Aegypti, with the Hermes Element From the Housefly." Proceedings Of The National Academy Of Sciences Of The United States Of America 95(7): 3743–47.

Jones, MDR, and S J Gubbins. 1978. "Changes in the Circadian Flight Activity of the Mosquito Anopheles Gambiae in Relation to Insemination, Feeding and Oviposition." Physiological Entomology 3(3): 213–20.

Kaiser, M, M Gasser, R Ackermann, and S C Stearns. 1997. "P-Element Inserts in Transgenic Flies: a Cautionary Tale." Heredity 78 (Pt 1): 1–11.

Kappe, Stefan H I, Carlos A Buscaglia, Lawrence W Bergman, Isabelle Coppens, and Victor Nussenzweig. 2004. "Apicomplexan Gliding Motility and Host Cell Invasion: Overhauling the Motor Model." Trends in Parasitology 20(1): 13–16.

Kaushal, D C, R Carter, L H Miller, and G Krishna. 1980. "Gametocytogenesis by Malaria Parasites in Continuous Culture." Nature 286(5772): 490–92.

Killeen, Gerry F, Ulrike Fillinger, Ibrahim Kiche, Louis C Gouagna, and Bart G J Knols. 2002. "Eradication of Anopheles Gambiae From Brazil: Lessons for Malaria Control in Africa?." The Lancet Infectious Diseases 2(10): 618–27.

Klassen, W, and C Curtis. 2005. "History of the Sterile Insect Technique." Sterile insect Technique. Principles and Practice in Area-Wide Integrated Pest Management

Knols, B G J, B N Njiru, E M Mathenge, W R Mukabana, J C Beier, and G F Killeen. 2002. "MalariaSphere: a Greenhouse-Enclosed Simulation of a Natural Anopheles Gambiae (Diptera: Culicidae) Ecosystem in Western Kenya." Malaria Journal 1(1): 19.

Koenraadt, Constantianus Jm, Matthias Kormaksson, and Laura C Harrington. 2010. "Effects of Inbreeding and Genetic Modification on Aedes Aegypti Larval Competition and Adult Energy Reserves." Parasites and Vectors 3: 92.

Kristan, M, H Fleischmann, A Della Torre, A Stich, and C F Curtis. 2003. "Pyrethroid Resistance/Susceptibility and Differential Urban/Rural Distribution of Anopheles Arabiensis and an. Gambiae S.S. Malaria Vectors in Nigeria and Ghana." Medical and Veterinary Entomology 17(3): 326–32.

Krzywinski, Jaroslaw, and Nora J Besansky. 2003. "Molecular Systematics of Anopheles: From Subgenera to Subpopulations." Annual Review of Entomology 48: 111–39.

Lambrechts, L, J C Koella, and C Boëte. 2008. "Can Transgenic Mosquitoes Afford the Fitness Cost?." Trends in Parasitology 24(1): 4–7.

Lanzaro, G C, and F Tripet. 2003. "Gene Flow Among Populations of Anopheles Gambiae: a Critical Review." Ecological Aspects for Application of Genetically Modified Mosquitoes. Kluwer Academic Publishers, Dordrecht: 109–32.

Lawniczak, M K N, S J Emrich, A K Holloway, A P Regier, M Olson, B White, S Redmond, L Fulton, E Appelbaum, J Godfrey, C Farmer, A Chinwalla, S-P Yang, P Minx, J Nelson, K Kyung, B P Walenz, E Garcia-Hernandez, M Aguiar, L D Viswanathan, Y-H Rogers, R L Strausberg, C A Saski, D Lawson, F H Collins, F C Kafatos, G K

Christophides, S W Clifton, E F Kirkness, and N J Besansky. 2010. "Widespread Divergence Between Incipient Anopheles Gambiae Species Revealed by Whole Genome Sequences." Science 330(6003): 512–14.

Lawson, D, P Arensburger, P Atkinson, N J Besansky, R V Bruggner, R Butler, K S Campbell, G K Christophides, S Christley, E Dialynas, D Emmert, M Hammond, C A Hill, R C Kennedy, N F Lobo, M R Maccallum, G Madey, K Megy, S Redmond, S Russo, D W Severson, E O Stinson, P Topalis, E M Zdobnov, E Birney, W M Gelbart, F C Kafatos, C Louis, and F H Collins. 2007. "VectorBase: a Home for Invertebrate Vectors of Human Pathogens." Nucleic Acids Research 35: 503–5.

Lee, H L, Seshadri Vasan, Nazni Wasi Ahmad, Iswarti Idris, Norhaida Hanum, S Selvi, Luke Alphey, and Shahnaz Murad. 2012. "Mating Compatibility and Competitiveness of Transgenic and Wild Type Aedes Aegypti (L.) Under Contained Semi-Field Conditions." Transgenic Research.

Lee, Yoosook, Anthony J Cornel, Claudio R Meneses, Abdrahamane Fofana, Aurélie G Andrianarivo, Rory D McAbee, Etienne Fondjo, Sékou F Traoré, and Gregory C Lanzaro. 2009. "Ecological and Genetic Relationships of the Forest-M Form Among Chromosomal and Molecular Forms of the Malaria Vector Anopheles Gambiae Sensu Stricto." Malaria Journal 8: 75.

Lehmann, T, M Licht, N Elissa, BTA Maega, J M Chimumbwa, F T Watsenga, C S Wondji, F Simard, and W A Hawley. 2003. "Population Structure of Anopheles Gambiae in Africa." Heredity 94(2): 133.

Lehmann, T, W A Hawley, L Kamau, D Fontenille, F Simard, and F H Collins. 1996. "Genetic Differentiation of Anopheles Gambiae Populations From East and West Africa: Comparison of Microsatellite and Allozyme Loci." Heredity 77 (Pt 2): 192–200.

Lehmann, Tovi, Jen C C Hume, Monica Licht, Christopher S Burns, Kurt Wollenberg, Fred Simard, and José M C Ribeiro. 2009. "Molecular Evolution of Immune Genes in the Malaria Mosquito Anopheles Gambiae." PLoS ONE 4(2): e4549.

Li, Chaoyang, Mauro T Marrelli, Guiyun Yan, and Marcelo Jacobs-Lorena. 2008. "Fitness of Transgenic Anopheles Stephensi Mosquitoes Expressing the SM1 Peptide Under the Control of a Vitellogenin Promoter." Heredity 99(3): 275–82.

Liu, H S, M S Jan, C K Chou, P H Chen, and N J Ke. 1999. "Is Green Fluorescent Protein Toxic to the Living Cells?." Biochemical and Biophysical Research Communications 260(3): 712–17.

Mabaso, MLH, B Sharp, and C Lengeler. 2004. "Historical Review of Malarial Control in Southern African with Emphasis on the Use of Indoor Residual House-Spraying." Tropical Medicine and International Health 9(8): 846–56.

Manson, P. (1878). "On the development of Filaria sanguinis hominis, and on the mosquito considered as a nurse." Journal of the Linnean Society London, Zoology 14: 304-311.

Marchand, R P. 1983. "Field Observations on Swarming and Mating in Anopheles Gambiae Mosquitoes in Tanzania1." Netherlands Journal of Zoology 34(3): 367–87.

Marchand, R P. 1985. "A New Cage for Observing Mating Behavior of Wild Anopheles Gambiae in the Laboratory.." Journal of the American Mosquito Control Association 1(2): 234.

Marchiafava, E. a. C., A. (1888). "On the nature of malaria." The New Sydenam Society, Selected Monographs 121.

Marrelli, M T, C K Moreira, D Kelly, L Alphey, and M Jacobs-Lorena. 2006. "Mosquito Transgenesis: What Is the Fitness Cost?." Trends in Parasitology 22(5): 197–202.

Marrelli, M T, C Li, J L Rasgon, and M Jacobs-Lorena. 2007. "Transgenic Malaria-Resistant Mosquitoes Have a Fitness Advantage When Feeding on Plasmodium-Infected Blood." Proceedings of the National Academy of Sciences 104(13): 5580.

Martinez-Torres, D, F Chandre, M S Williamson, F Darriet, J B Bergé, A L Devonshire, P Guillet, N Pasteur, and D Pauron. 1998. "Molecular Characterization of Pyrethroid Knockdown Resistance (Kdr) in the Major Malaria Vector Anopheles Gambiae S.S." Insect Molecular Biology 7(2): 179–84.

Menge, David M, Tom Guda, Daibin Zhong, Aditi Pai, Goufa Zhou, John C Beier, Louis Gouagna, and Guiyun Yan. 2005. "Fitness Consequences of Anopheles Gambiae Population Hybridization." Malaria Journal 4: 44.

Meredith, Janet M, Sanjay Basu, Derric D Nimmo, Isabelle Larget-Thiery, Emma L Warr, Ann Underhill, Clare C McArthur, Victoria Carter, Hilary Hurd, Catherine Bourgouin, and Paul Eggleston. 2011. "Site-Specific Integration and Expression of an Anti-Malarial Gene in Transgenic Anopheles Gambiae Significantly Reduces Plasmodium Infections." PLoS ONE 6(1): e14587.

Milby, M M, R L Nelson, and P T McDonald. 1980. "Release of Heterozygous Translocated Adult Males for Genetic Control of Ciilcx Tarsalis at an Isolated Site.." Journal of the American Mosquito Control Association 40(1): 83–90.

Mitchell, Sara N, Bradley J Stevenson, Pie Müller, Craig S Wilding, Alexander Egyir-Yawson, Stuart G Field, Janet Hemingway, Mark J I Paine, Hilary Ranson, and Martin James Donnelly. 2012. "Identification and Validation of a Gene Causing Cross-Resistance Between Insecticide Classes in Anopheles Gambiae From Ghana." Proceedings Of The National Academy Of Sciences Of The United States Of America 109(16): 6147–52.

Moreira, C K, F G Rodrigues, A Ghosh, F P Varotti, A Miranda, S Daffre, M Jacobs-Lorena, and L A Moreira. 2007. "Effect of the Antimicrobial Peptide Gomesin Against Different Life Stages of Plasmodium Spp.." Experimental Parasitology 116(4): 346–53.

Moreira, L A. 2002. "Bee Venom Phospholipase Inhibits Malaria Parasite Development in Transgenic Mosquitoes." Journal of Biological Chemistry 277(43): 40839–43.

Moreira, L A, M J Edwards, F Adhami, N Jasinskiene, A A James, and M Jacobs-Lorena. 2000. "Robust Gut-Specific Gene Expression in Transgenic Aedes Aegypti Mosquitoes." Proceedings Of The National Academy Of Sciences Of The United States Of America 97(20): 10895–98.

Moreira, Luciano A, Jing Wang, Frank H Collins, and Marcelo Jacobs-Lorena. 2004. "Fitness of Anopheline Mosquitoes Expressing Transgenes That Inhibit Plasmodium Development." Genetics 166(3): 1337–41.

Moskalyk, L A, M M Oo, and M Jacobs-Lorena. 1996. "Peritrophic Matrix Proteins of Anopheles Gambiae and Aedes Aegypti." Insect Molecular Biology 5(4): 261–68.

Munhenga, Givemore, Basil D Brooke, Tobias F Chirwa, Richard H Hunt, Maureen Coetzee, Danny Govender, and Lizette L Koekemoer. 2011. "Evaluating the Potential of the Sterile Insect Technique for Malaria Control: Relative Fitness and Mating

Compatibility Between Laboratory Colonized and a Wild Population of Anopheles Arabiensis From the Kruger National Park, South Africa." Parasites and Vectors 4: 208.

Myung, Joon Mo, Patricia Marshall, and Photini Sinnis. 2004. "The Plasmodium Circumsporozoite Protein Is Involved in Mosquito Salivary Gland Invasion by Sporozoites." Molecular & Biochemical Parasitology 133(1): 53–59.

N'Guessan, R, F Darriet, P Guillet, P Carnevale, M Traore-Lamizana, V Corbel, A A Koffi, and F Chandre. 2003. "Resistance to Carbosulfan in Anopheles Gambiae From Ivory Coast, Based on Reduced Sensitivity of Acetylcholinesterase." Medical and Veterinary Entomology 17(1): 19–25.

N'Guessan, Raphael, Vincent Corbel, Martin Akogbéto, and Mark Rowland. 2007. "Reduced Efficacy of Insecticide-Treated Nets and Indoor Residual Spraying for Malaria Control in Pyrethroid Resistance Area, Benin." Emerging infectious diseases 13(2): 199–206.

Ng'habi, K R, B G J Knols, Y Lee, H M Ferguson, and G C Lanzaro. 2011. "Population Genetic Structure of Anopheles Arabiensis and Anopheles Gambiae in a Malaria Endemic Region of Southern Tanzania." Malaria Journal 10(1): 289.

Ng'habi, Kija R N, Dickson Mwasheshi, Bart G J Knols, and Heather M Ferguson. 2010. "Establishment of a Self-Propagating Population of the African Malaria Vector Anopheles Arabiensis Under Semi-Field Conditions." Malaria Journal 9: 356.

Ng'habi, Kija R, Bernadette John, Gamba Nkwengulila, Bart G J Knols, Gerry F Killeen, and Heather M Ferguson. 2005. "Effect of Larval Crowding on Mating Competitiveness of Anopheles Gambiae Mosquitoes." Malaria Journal 4: 49.

Nicastri, Emanuele, Nazario Bevilacqua, Monica Sane Schepisi, Maria G Paglia, Silvia Meschi, Shaali M Ame, Jape A Mohamed, Sabina Mangi, Robert Fumakule, Antonino Di Caro, Maria R Capobianchi, Andrew Kitua, Fabrizio Molteni, Vincenzo Racalbuto, and Giuseppe Ippolito. 2009. "Accuracy of Malaria Diagnosis by Microscopy, Rapid Diagnostic Test, and PCR Methods and Evidence of Antimalarial Overprescription in Non-Severe Febrile Patients in Two Tanzanian Hospitals." The American Journal of Tropical Medicine and Hygiene 80(5): 712–17.

Nimmo, D D, L Alphey, J M Meredith, and P Eggleston. 2006. "High Efficiency Site-Specific Genetic Engineering of the Mosquito Genome." Insect Molecular Biology 15(2): 129–36.

Nirmala, X, O Marinotti, J M Sandoval, S Phin, S Gakhar, N Jasinskiene, and A A James. 2006. "Functional Characterization of the Promoter of the Vitellogenin Gene, AsVg1, of the Malaria Vector, Anopheles Stephensi." Insect Biochemistry and Molecular Biology 36(9): 694–700.

Noedl, Harald, Youry Se, Kurt Schaecher, Bryan L Smith, Duong Socheat, Mark M Fukuda, and Artemisinin Resistance in Cambodia 1 ARC1 Study Consortium. 2008. "Evidence of Artemisinin-Resistant Malaria in Western Cambodia." The New England Journal of Medicine 359(24): 2619–20.

Nolan, Tony, Elisa Petris, Hans-Michael Müller, Ann Cronin, Flaminia Catteruccia, and Andrea Crisanti. 2011. "Analysis of Two Novel Midgut-Specific Promoters Driving Transgene Expression in Anopheles Stephensi Mosquitoes." PLoS ONE 6(2): e16471.

Norris, D E, A C Shurtleff, Y T Touré, and G C Lanzaro. 2001. "Microsatellite DNA Polymorphism and Heterozygosity Among Field and Laboratory Populations of Anopheles Gambiae Ss (Diptera: Culicidae)." Journal of Medical Entomology 38(2): 336–40.

O'brochta, D A. 2003a. "Gene Vector and Transposable Element Behavior in Mosquitoes." Journal of Experimental Biology 206(21): 3823–34.

O'brochta, D A. 2003b. "Transgenic Mosquitoes: the State of the Art." Ecological Aspects for Application of Genetically Modified Mosquitoes: 15–24.

Oliva, Clelia F, Mark Q Benedict, Guy Lempérière, and Jérémie Gilles. 2011. "Laboratory Selection for an Accelerated Mosquito Sexual Development Rate." Malaria Journal 10: 135.

Orozco-Dávila, D, R Hernández, S Meza, and J Domínguez. 2007. "Sexual Competitiveness and Compatibility Between Mass-Reared Sterile Flies and Wild Populations of Anastrepha Ludens (Diptera: Tephritidae) From Different Regions in Mexico." Florida Entomologist 90(1): 19–26.

O'brochta, David A, and Alfred M Handler. 2008. "Perspectives on the State of Insect Transgenics." Advances in Experimental Medicine and Biology 627: 1–18.

Pates, H V, W Takken, C F Curtis, P W Huisman, O Akinpelu, and G S Gill. 2001. "Unexpected Anthropophagic Behaviour in Anopheles Quadriannulatus." Medical and Veterinary Entomology 15(3): 293–98.

Paul, Richard E L, Paul T Brey, and Vincent Robert. 2002. "Plasmodium Sex Determination and Transmission to Mosquitoes." Trends in Parasitology 18(1): 32–38.

Pennetier, Cédric, Ben Warren, K Roch Dabiré, Ian J Russell, and Gabriella Gibson. 2010. "'Singing on the Wing' as a Mechanism for Species Recognition in the Malarial Mosquito Anopheles Gambiae." Current Biology : CB 20(2): 131–36.

Phuc, Hoang, Morten H Andreasen, Rosemary S Burton, Céline Vass, Matthew J Epton, Gavin Pape, Guoliang Fu, Kirsty C Condon, Sarah Scaife, Christl A Donnelly, Paul G Coleman, Helen White-Cooper, and Luke Alphey. 2007. "Late-Acting Dominant Lethal Genetic Systems and Mosquito Control." BMC Biology 5(1): 11.

Pimenta, P F, M TOURAY, and L MILLER. 1994. "The Journey of Malaria Sporozoites in the Mosquito Salivary Gland." Journal of Eukaryotic Microbiology 41(6): 608–24.

Pinkerton, A C, K Michel, D A O'brochta, and P W Atkinson. 2000. "Green Fluorescent Protein as a Genetic Marker in Transgenic Aedes Aegypti." Insect Molecular Biology 9(1): 1–10.

Pradel, G. 2007. "Proteins of the Malaria Parasite Sexual Stages: Expression, Function and Potential for Transmission Blocking Strategies." Parasitology 134(Pt.14): 1911–29.

Ranson, Hilary, Raphael N'Guessan, Jonathan Lines, Nicolas Moiroux, Zinga Nkuni, and Vincent Corbel. 2011. "Pyrethroid Resistance in African Anopheline Mosquitoes: What Are the Implications for Malaria Control?." Trends in Parasitology 27(2): 91–98.

Reisen, W K, R K Sakai, R H Baker, H R Rathor, K Raana, K Azra, and S Niaz. 1980. "Field Competitiveness of Culex Tritaeniorhynchus Giles Males Carrying a Complex Chromosomal Aberration: a Second Experiment." Annals of the Entomological Society of America 73(4): 479–84.

Reisen, W, M Milby, S Asman, M Bock, and R Meyer. 1982. "Attempted Suppression of a Semi-Isolated Culex Tarsalis Population by the Release of Irradiated Males: a Second

Experiment Using Males From a Recently Colonized ….” Journal of the American Mosquito Control Association.

Reiter, P. 2001. “Climate Change and Mosquito-Borne Disease.” Environmental Health Perspectives 109 Suppl 1: 141–61.

Rodrigues, F G, M N Santos, T X T de Carvalho, B C Rocha, M A Riehle, P F P Pimenta, E G Abraham, M Jacobs-Lorena, C F Alves de Brito, and L A Moreira. 2008. “Expression of a Mutated Phospholipase A2 in Transgenic Aedes Fluviatilis Mosquitoes Impacts Plasmodium Gallinaceum Development.” Insect Molecular Biology 17(2): 175–83.

Rubin, G M, and A C Spradling. 1982. “Genetic Transformation of Drosophila with Transposable Element Vectors.” Science 218(4570): 348–53.

Sanford, Michelle R, Berna Demirci, Clare D Marsden, Yoosook Lee, Anthony J Cornel, and Gregory C Lanzaro. 2011. “Morphological Differentiation May Mediate Mate-Choice Between Incipient Species of Anopheles Gambiae S.S.” PLoS ONE 6(11): e27920.

Schneider, P, W Takken, and P J McCall. 2000. “Interspecific Competition Between Sibling Species Larvae of Anopheles Arabiensis and an. Gambiae.” Medical and Veterinary Entomology 14(2): 165–70.

Scott, J A, W G Brogdon, and F H Collins. 1993. “Identification of Single Specimens of the Anopheles Gambiae Complex by the Polymerase Chain Reaction.” The American Journal of Tropical Medicine and Hygiene 49(4): 520–29.

Sidén-Kiamos, I, A Ecker, S Nybäck, and C Louis. 2006. “Plasmodium Berghei Calcium-Dependent Protein Kinase 3 Is Required for Ookinete Gliding Motility and ….” Molecular Microbiology.

Slotman, M A, F Tripet, A J Cornel, C R Meneses, Y Lee, L J Reimer, T C Thiemann, E Fondjo, A Fofana, S F Traoré, and G C Lanzaro. 2007. "Evidence for Subdivision Within the M Molecular Form of Anopheles Gambiae." Molecular Ecology 16(3): 639–49.

Slotman, M A, L J Reimer, T Thiemann, G Dolo, E Fondjo, and G C Lanzaro. 2006. "Reduced Recombination Rate and Genetic Differentiation Between the M and S Forms of Anopheles Gambiae S.S.." Genetics 174(4): 2081–93.

Slotman, Michel, Alessandra della Torre, and Jeffrey R Powell. 2005. "Female Sterility in Hybrids Between Anopheles Gambiae and a. Arabiensis, and the Causes of Haldane's Rule." Evolution 59(5): 1016–26.

Stump, A D, M Pombi, L Goeddel, J M C Ribeiro, J A Wilder, A Della Torre, and N J Besansky. 2007. "Genetic Exchange in 2La Inversion Heterokaryotypes of Anopheles Gambiae." Insect Molecular Biology 16(6): 703–9.

Takken, W, M J Klowden, and G M Chambers. 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): 639–45.

Talman, A M, O Domarle, F E McKenzie, F Ariey, and V Robert. 2004. "Gametocytogenesis: the Puberty of Plasmodium Falciparum." Malaria Journal 3(1): 24.

TDR. 2005. "30 Years of Research and Capacity Building in Tropical Diseases ."

Thathy, Vandana, Hisashi Fujioka, Soren Gantt, Ruth Nussenzweig, Victor Nussenzweig, and Robert Menard. 2002. "Levels of Circumsporozoite Protein in the Plasmodium Oocyst Determine Sporozoite Morphology." EMBO Journal 21(7): 1586–96.

Thyagarajan, B, E C Olivares, R P Hollis, D S Ginsburg, and M P Calos. 2001. "Site-Specific Genomic Integration in Mammalian Cells Mediated by Phage phiC31 Integrase." Molecular and Cellular Biology 21(12): 3926–34.

Torre, Della, A. 2002. "Speciation Within Anopheles Gambiae-- the Glass Is Half Full." Science 298(5591): 115–17.

Torre, della, Alessandra, Zhijian Tu, and Vincenzo Petrarca. 2005. "On the Distribution and Genetic Differentiation of Anopheles Gambiae S.S. Molecular Forms." Insect Biochemistry and Molecular Biology 35(7): 755–69.

Touré, Y T, V Petrarca, S F Traoré, A Coulibaly, H M Maiga, O Sankaré, M Sow, M A Di Deco, and M Coluzzi. 1998. "The Distribution and Inversion Polymorphism of Chromosomally Recognized Taxa of the Anopheles Gambiae Complex in Mali, West Africa." Parassitologia 40(4): 477–511.

Tripet, F. 2009. "Ecological Immunology of Mosquito-Malaria Interactions: of Non-Natural Versus Natural Model Systems and Their Inferences." Parasitology: 1–8.

Tripet, F, Y T Touré, C E Taylor, D E Norris, G Dolo, and G C Lanzaro. 2001. "DNA Analysis of Transferred Sperm Reveals Significant Levels of Gene Flow Between Molecular Forms of Anopheles Gambiae." Molecular Ecology 10(7): 1725–32.

Turner, T L, and M W Hahn. 2007. "Locus-and Population-Specific Selection and Differentiation Between Incipient Species of Anopheles Gambiae." Molecular Biology and Evolution 24(9): 2132.

Turner, Thomas L, and Matthew W Hahn. 2010. "Genomic Islands of Speciation or Genomic Islands and Speciation?." Molecular Ecology 19(5): 848–50.

Turner, Thomas L, Matthew W Hahn, and Sergey V Nuzhdin. 2005. "Genomic Islands of Speciation in Anopheles Gambiae." PLoS Biology 3(9): e285.

Venken, Koen J T, Yuchun He, Roger A Hoskins, and Hugo J Bellen. 2006. "P[Acman]: a BAC Transgenic Platform for Targeted Insertion of Large DNA Fragments in D. Melanogaster." Science 314(5806): 1747–51.

Vinetz, J M. 2005. "Plasmodium Ookinete Invasion of the Mosquito Midgut." Current topics in microbiology and immunology 295: 357–82.

Vlachou, D, T Zimmermann, R Cantera, and C Janse. 2004. "Real-Time, in Vivo Analysis of Malaria Ookinete Locomotion and Mosquito Midgut Invasion." Cellular Microbiology.

Wang, R, L Zheng, Y T Touré, T Dandekar, and F C Kafatos. 2001. "When Genetic Distance Matters: Measuring Genetic Differentiation at Microsatellite Loci in Whole-Genome Scans of Recent and Incipient Mosquito Species." Proceedings Of The National Academy Of Sciences Of The United States Of America 98(19): 10769–74.

Wang-Sattler, Rui, Stephanie Blandin, Ye Ning, Claudia Blass, Guimogo Dolo, Yeya T Touré, Alessandra delle Torre, Gregory C Lanzaro, Lars M Steinmetz, Fotis C Kafatos, and Liangbiao Zheng. 2007. "Mosaic Genome Architecture of the Anopheles Gambiae Species Complex." PLoS ONE 2(11): e1249.

Weetman, David, Craig S Wilding, Keith Steen, João Pinto, and Martin J Donnelly. 2012. "Gene Flow-Dependent Genomic Divergence Between Anopheles Gambiae M and S Forms." Molecular Biology and Evolution 29(1): 279–91.

Weill, M, F Chandre, C Brengues, S Manguin, M Akogbéto, N Pasteur, P Guillet, and M Raymond. 2000. "The Kdr Mutation Occurs in the Mopti Form of Anopheles Gambiae S.S. Through Introgression." Insect Molecular Biology 9(5): 451–55.

White, Bradley J, Matthew W Hahn, Marco Pombi, Bryan J Cassone, Neil F Lobo, Frederic Simard, and Nora J Besansky. 2007. "Localization of Candidate Regions Maintaining a Common Polymorphic Inversion (2La) in Anopheles Gambiae." PLoS Genetics 3(12): e217.

WHO (1978). "The malaria situation in 1976." WHO Chron. 32: 9-17.

WHO (2000). "Technical Report Series 892. WHO Expert Committee on Malaria: Twentieth Report."

WHO (2007). "Working for Health: An Introduciton to the World Health Organisation."

WHO (2011) "World Health Report 2011"

Wilding, Craig S, David Weetman, Keith Steen, and Martin J Donnelly. 2009. "High, Clustered, Nucleotide Diversity in the Genome of Anopheles Gambiae Revealed Through Pooled-Template Sequencing: Implications for High-Throughput Genotyping Protocols." BMC Genomics 10: 320.

Windbichler, Nikolai, Miriam Menichelli, Philippos Aris Papathanos, Summer B Thyme, Hui Li, Umut Y Ulge, Blake T Hovde, David Baker, Raymond J Monnat, Austin Burt, and Andrea Crisanti. 2011. "A Synthetic Homing Endonuclease-Based Gene Drive System in the Human Malaria Mosquito." Nature 473(7346): 212–15.

Wondji, C, F Simard, and D Fontenille. 2002. "Evidence for Genetic Differentiation Between the Molecular Forms M and S Within the Forest Chromosomal Form of Anopheles Gambiae in an Area of Sympatry." Insect Molecular Biology 11(1): 11–19.

Yakob, L, L Alphey, and M B Bonsall. 2008. "Aedes Aegypti Control: the Concomitant Role of Competition, Space and Transgenic Technologies." Journal of Applied Ecology 45(4): 1258–65.

Yawson, A E, P J McCall, M D Wilson, and M J Donnelly. 2004. "Species Abundance and Insecticide Resistance of Anopheles Gambiae in Selected Areas of Ghana and Burkina Faso." Medical and Veterinary Entomology 18(4): 372–77.

Yawson, Alexander E, David Weetman, Michael D Wilson, and Martin J Donnelly. 2007. "Ecological Zones Rather Than Molecular Forms Predict Genetic Differentiation in the Malaria Vector Anopheles Gambiae S.S. in Ghana." Genetics 175(2): 751–61.

Young, M, and D Moore. 1961. "Chloroquine Resistance in Plasmodium Falciparum." The American Journal of Tropical Medicine and Hygiene.

Zhu, S, and J Tytgat. 2004. "The Scorpine Family of Defensins: Gene Structure, Alternative Polyadenylation and Fold Recognition." Cellular and Molecular Life Sciences 61(14): 1–13.