Sex Determination and Sex Ratio Distortion in The Mosquito *Aedes aegypti*

Tze Min Teo

Department of Zoology

Green Templeton College

University of Oxford

August 2010

Abstract

Aedes aegypti is one of the most important mosquito vectors of human disease, transmitting dengue, Chikungunya and yellow fever viruses. New control methods are much needed for this species and an ability to skew sex ratios toward males would be very useful for a number of strategies. In this study, male-specific chromosomal regions were searched for in *Ae. aegypti* using novel approaches such as microdissection of chromosomes. The sex determination pathway of *Ae. aegypti* was also explored, using *Drosophila melanogaster* as a model, to find candidate genes that could be used to induce male biased sex ratios. The *transformer-2 (tra-2)* gene is necessary for sexual differentiation of females in *D. melanogaster*, but its role remains unknown in mosquitoes. A homolog of *tra-2* was identified in *Ae. aegypti*. Rapid Amplification of cDNA ends (RACE) experiments were conducted to characterize this gene in *Ae. aegypti*. The *Ae. aegypti tra-2* gene, as for *D. melanogaster*, was highly variable in transcription due to alternative splicing and alternative polyadenylation, with 9 different variants identified. RNA interference (RNAi) was then used to determine if knockdown of all variants of *Ae. aegypti tra-2* can be achieved and if it would cause gender switching in individuals, initially by means of direct injection of double stranded RNA, and then progressing to germline transformation. A construct designed to produce *tra-2* dsRNA was injected into *Ae. aegypti* embryos and integrated into the *Ae. aegypti* genome. The transgenic population showed up to 100% male bias in single pair crosses. The male bias effect could still be seen in the heterozygous population. The results suggest that knockdown of *tra-2* could provide a useful tool for sex ratio distortion as part of the development of novel control methods for *Ae. aegypti*.

Acknowledgements

I would like to thank my supervisor, Dr. Steven Sinkins, for his support and allowing me to have this opportunity to work on this interesting topic.

I would also like to extend my thanks to the Sinkins Lab, particularly the following people:

Dr. Kim Phuc Hoang for working with me on many of the projects, numerous technical teachings, and 'long-time' discussions; Dr. Zakaria Kambris for his support and advise, particularly in qRT-PCR and microarray techniques, and Fridays; Dr. Kirsty Stainton for countless proof-readings, Tesco runs, and other non-scientific activities; Dr. Lisa Klasson for bioinformatics help and Chiangmai; Dr. Tom Walker for lab advice and motivation; Dr. Peter Cook for microarray design and general mischief; Dr. Anthony Brown for Jenga and 'drugs'; Dr. Sofia Pinto for tips and technical advice; Limb Hapairai for muscles and laughter; Arun Raju for his support during my first year; Harriet Kemp for the extra pair of hands.

I am also very grateful to Diva Jain for chai, cake, and long drives; Jay and Emma for movie nights; Tom G for LaTex help; Prof. Ken Reid and Margery Reid, for their support and concerns.

I thank my siblings, who are always around in my time of need, especially Qin whom have been helping me in whatever way she can throughout the years.

Finally, my deepest and heartfelt gratitude to my parents for all they have done for me. Without them, this thesis would not have been made possible.

ABBREVIATIONS

Ae.	*Aedes* genus
An.	*Anopheles* genus
ATP	Adenosine 5'-triphosphate
bp	Base pairs (of DNA sequence)
BLAST	Basic Local Alignment Search Tool
BSA	Bovine serum albumin
BTB	Bric-a-brac, Tramtrack, Broad-complex
C.	*Ceratitis* genus
cDNA	Complementary deoxyribonucleic acid
CGH	Comparative genomic hybridization
CI	Cytoplasmic Incompatibility
CREB	Cyclic AMP Response Element Binding protein
CSPD	Disodium 3-(4-methoxyspiro {1,2-dioxetane-3,2-(5'-chloro)tricyclo [3.3.1.1^{3,7}]decan}-4-yl)phenyl phosphate
Cu.	*Culex* genus
D.	*Drosophila genus*
DENV	Dengue virus
DEPC	Diethylpyrocarbonate
DDT	Dichlorodiphenyltrichloroethane
DF	Dengue fever
DHF	Dengue hemorrhagic fever
DIG	Digoxigenin
DMDC	Dimethyl dicarbonate
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotide triphosphates (adenine, guanine, thymine, cytosine)
ds	Double stranded
dsRNA	Double stranded Ribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
EGFP	Enhanced green fluorescent protein
EivF	Early region IV F
EtBr	Ethidium Bromide
FISH	Fluorescence in situ hybridization
F$_n$	Progeny (n= number of generation)
g	standard gravity
GABA	Gamma-aminobutyric acid
GArC3	GG, AT-rich, CC 3
Gb	Gigabase
gDNA	Genomic DNA
G$_n$	Generation (n= number of generation)
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HSV IE	Herpes Simplex Virus Immediate Early
IPTG	Isopropyl β-D-1-thio-galactopyranoside

ITN	Insecticide treated nets
JCV	JC Virus (human papovavirus)
kb	Kilo-base pairs = 1000 base pairs (of DNA)
LB	Lysogeny Broth
M	Molar
M.	*Musca* genus
μg	Micrograms
min.	Minutes
ORF	Open reading frame
PCR	Polymerase chain reaction
rDNA	Ribosomal DNA
RIDL	Release of Insects carrying a Dominant Lethal
RNA	Ribonucleic acid
RNAi	RNA interference
RPM	Revolutions per minute
RRM	RNA Recognition Motif
SDS	Sodium dodecyl sulfate
SSC	Saline-sodium citrate
TAE	Tris-acetate-EDTA buffer
TFIID	Transciption Factor II D
TRA	Transformer
Tris	Tris(hydroxymethyl)aminomethane
ULS	Universal linkage system
UV	Ultra-violet
WHO	World Health Organisation
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
ZF	Zinc finger

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 General introduction to mosquitoes

Mosquitoes are members of the insect order Diptera, family Culicidae, and are traditionally classified into three subfamilies (Edwards, 1932): *Toxorhynchitinae*, *Anophelinae*, *Culicinae*. *Toxorhynchitinae* have only one known genus that is *Toxorhynchites*. There was some ambiguity in the subfamily *Toxorhynchitinae*, where it was lowered to tribal status Toxorhynchitini (Harbach and Kitching, 1998; Besansky and Fahey, 1997; Shepard et al., 2006; Harbach, 2007), therefore *Toxorynchites* can be considered a genus within the *Culicinae* family. Anophelinae are further subdivided into three genera: *Anopheles*, *Bironella*, *Chagasia*.

Culicinae is the largest subfamily of mosquitoes. There are over 100 genera, including *Culex*, *Aedes*, *Ochelarotatus*, *Culiseta*, *Mansonia*, *Malaya*, *Sabethes*, and *Wyeomyia*.

Vector borne diseases carried by mosquitoes have been a health problem throughout human history. *Anophelinae* and *Culicinae* subfamilies contain species that are known to be a serious health concern as they harbour microorganisms that can be transmitted to humans or animals when the female mosquito feeds on their blood. Examples of the diseases they can carry are malaria parasites, dengue and West Nile Virus.

1.2 *Aedes aegypti*

Aedes species belong to the *Culicinae* subfamily of mosquitoes. A well-known species is *Aedes (Stegomyia) aegypti* (Linnaeus), which is also commonly known as the Yellow Fever mosquito. It is a species that has been found to live within or close-by human environments, often biting indoors or in sheltered areas. This species prefers to seek blood by day in shaded areas, with peak activities at mid-morning and late afternoon (Scott et al., 1993). It also occasionally feeds early in the evening. *Aedes aegypti* is easily identified by the prominent silvery white thoracic and abdominal markings, white banded legs, and the silvery lyre-shaped marking on the scutum (see figure 1.1). *Aedes aegypti* tends to breed in all forms of receptacle in urban areas, especially after intermittent rainfall in tropical regions (Chou et al., 1998). The eggs are laid singly on wet surfaces and are resistant to desiccation. *Aedes aegypti* eggs are able to survive and lay dormant for several months when the environment become hostile. When conditions are favourable again, the eggs could then hatch and the life cycle begins once more. One female can lay more than a hundred eggs at a time. Thus, eradication of this species is difficult due to its capacity for rapid population recovery. A hardy species, *Ae. aegypti* is successful in colonising new habitats and is now found in many parts of the world away from its African origins.

The hardiness of *Ae. aegypti* also means that this particular species is favoured to study as it is easily reared in an artificial environment. The embryos are easy to handle and therefore *Ae. aegypti* is a good candidate for genetic manipulation.

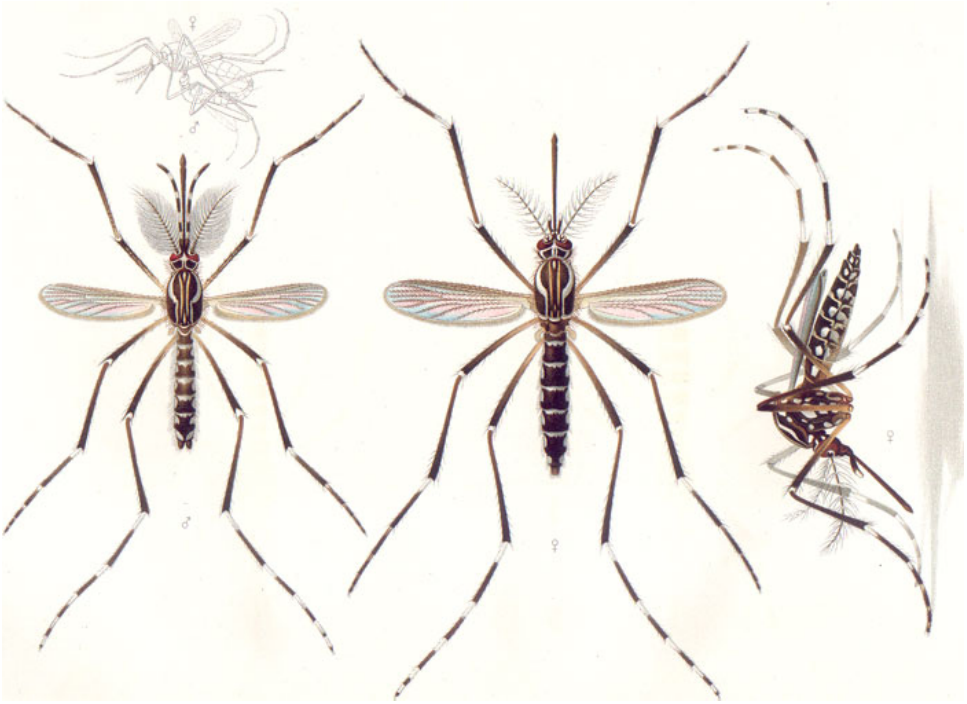

Figure 1.1: Illustrations of *Aedes aegypti* adults by E. M. Goeldi. At left, the male, in the middle and on the right, the female - ventral and side view. Above left, a flying pair in copula (Goeldi, 1905).

1.2.1 *Aedes aegypti* as vectors

Apart from being a biting pest, *Ae. aegypti* is a major disease vector in tropical and sub-tropical areas around the world where the mosquito thrives.

One of the diseases *Ae. aegypti* transmit is dengue. Dengue disease caused by the flavivirus dengue virus (DENV). Global cases of dengue fever (DF) and dengue hemorrhagic fever (DHF) have dramatically increased in recent years. Dengue is currently classified as an emerging or re-emerging infectious disease by the World Health Organisation (WHO).

In many tropical and subtropical regions, dengue is endemic and causes annual or periodic outbreaks (Chow et al., 1998). No dengue vaccine is available although candidate vaccine viruses have been developed. According to WHO, it is estimated that there are 2.5 billion people at risk from dengue and there

may be 50 million infections occuring worldwide per year. An estimated 500 000 people with DHF require hospitalisation each year and DHF case fatality rates can exceed 20% if left untreated (WHO, 2009b). In 2005, dengue was the most important mosquito-borne viral disease affecting humans; its global distribution is comparable to that of malaria (Centre for Disease Control and Prevention, 2006b). Dengue is now endemic in over 100 countries around the world (see figure 1.2), a large increase in comparison to pre-1970 when only nine countries had experienced DHF epidemics. Not only is the number of cases increasing as the disease is spreading to new areas, explosive outbreaks are also occurring as well (WHO, 2009b).

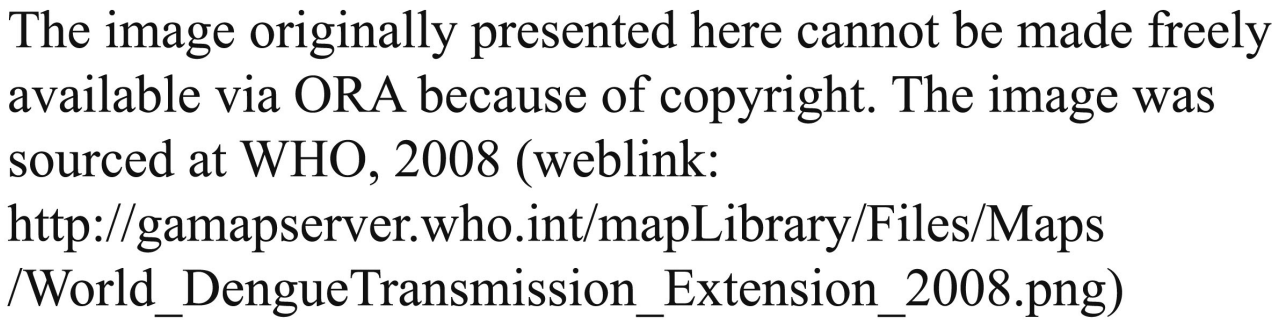

Figure 1.2: World distribution of Dengue surveyed in 2008. Source website: http://gamapserver.who.int/mapLibrary/Files/Maps /World_DengueTransmission_Extension_2008.png.

Aedes aegypti is also a major vector for yellow fever and Chikungunya. Yellow fever is an acute viral hemorrhagic disease. When epidemics occur in unvaccinated populations, case-fatality rates may exceed 50% (WHO, 2009a). The WHO estimates that yellow fever causes 200,000 illnesses and 30,000 deaths every year in unvaccinated populations; around 90% of the infections occur in Africa (WHO, 2009a).

Chikungunya is a viral disease that shows similar clinical signs to dengue. Like dengue, Chikungunya has no known cure. Chikungunya occurs in Africa, Asia and the Indian subcontinent (figure 1.3). Human infections in Africa have been at relatively low levels for a number of years, but in 1999-2000 there was a large outbreak in the Democratic Republic of the Congo, and in 2007 there was an outbreak in Gabon (WHO, 2008b).

According to WHO, a major outbreak of Chikungunya occurred in islands of the Indian Ocean in 2005. A large outbreak also occurred in India in 2006 and 2007 (Centre for Disease Control and Prevention, 2006a; WHO, 2008a). Several other countries in South-East Asia were also affected. In 2007, transmission was reported for the first time in Europe. A localised outbreak was reported in north-eastern Italy (Beltrame et al., 2007), although the mosquito responsible for transmission was *Ae. albopictus* and not *Ae. aegypti*.

Other viruses that *Ae. aegypti* can transmit are Ross River viruses (Gubler, 1981), Murray Valley encephalitis virus (Kay, 1980)and Rift Valley fever virus (Hoch et al., 1985). Apart from viruses, *Ae. aegypti* is also known to be a natural carrier of filariasis caused by parasite *Dirofilaria immitis*, also known as Dog heartworm (Hendrix et al., 1986).

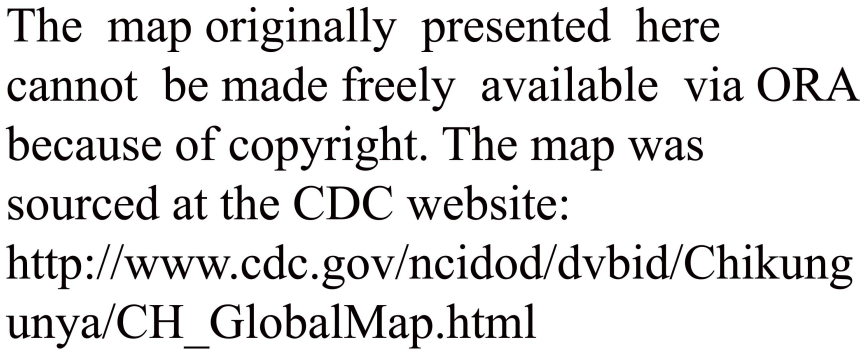

Figure 1.3: Estimated world distribution of Chikungunya in 2010. Areas highlighted in light blue indicates countries with reported cases of Chikungunya. Source: CDC: Chikungunya Distribution and Global Map (website: http://www.cdc.gov/ncidod/dvbid/Chikungunya/CH_GlobalMap.html).

1.2.2 Traditional strategies

Strategies used to curb vector transmitted diseases include vaccines, drug treatment of infected individuals, and population control of the vector. Traditional vector control strategies involve the use of insecticides (including larvicide use) and reducing potential breeding sites. Thermal fogging, originating from military smoke generators in the early 1940s (LaMer et al., 1947), and ultra-low-volume (ULV) cold spraying, originally used from modified agricultural sprayers in the 1950s (Lofgren, 1970), are often used to control outdoor adult mosquitoes in dengue epidemics (PAHO, 1982). Surface spraying was also used in mosquito control to give a longer lasting effect by leaving a concentrated residual insecticide (Payne et al., 1976). The introduction of indoor residual insecticide spraying had a huge impact on the malaria control. In many African countries, incidences of malaria were reduced

after the implementation of residual spraying (Mastbaum, 1954; Alves and Blair, 1955; Brink, 1958; Chayajabera et al., 1975; Sharp et al., 2002). Use of insecticide-impregnated bed nets (also known as insecticide treated nets, ITNs) reduced as much as half of the malarial cases in trials in African countries (Lengeler, 2004), including Burkina Faso (Habluetzel et al., 1997), Cameroon (Moyou-Somo et al., 1995), Gambia (Snow et al., 1988; Lindsay et al., 1989; D'Alessandro et al., 1995; Thomson et al., 1995), Kenya (Mbogo et al., 1996; Sexton et al., 1990) and Ghana (Binka et al., 1996). Since *Ae. aegypti* is a day biting mosquito, the use of insecticide-treated bed nets would not reduce the dengue vectors significantly. However, a similar strategy used insecticide treated window curtains and trials done in Mexico and Venezuela suggest it can reduce densities of dengue vector *Ae. aegypti* to low levels and affect dengue transmission (Kroeger et al., 2006). Potential breeding sites are reduced by removing unused water receptacles such as plastic containers, old tires; and by unclogging drains and gutters. Larger bodies of standing water that cannot be eliminated are treated with larvicide. Other less prevalent methods of control or protection from adult mosquitoes include traps, house screening, mosquito coils and personal repellents.

1.2.2.1 Limitations of traditional strategies

Traditional vector control strategies are becoming increasingly ineffective. One possible reason behind the declining effectiveness of these traditional methods is that they involve carrying out consistent and rigorous monitoring in large epidemic areas or places. This could be exacerbated by insufficient man power and funds for such control programmes. Even well established programmes still face many problems, as reported in Singapore where success in reducing *Aedes* population to a house index (calculated by number of infested houses

inspected divided by number of houses inspected, multiplied by 100) as low as 1% did not decrease incidence of dengue (Wang, 1994). Non-specific insecticide application is also a concern to public human health as increasingly large, repeated doses of insecticide are required and sprayed over a wide area most often inhabited by people. It is also a concern that insecticide usage will harm the environment. An example is dichlorodiphenyltrichloroethane (DDT), first used during 1940s to combat malaria, typhus and other vector-borne diseases in military areas (Curtis, 2002). DDT was commercially available until evidence of human, animal and environmental health problems linked to its usage eventually lead to banning it in most countries (Metcalf, 1973). Malathion, an insecticide of relatively low human toxicity, is still used widely in fogging programs. However, absorption or ingestion of malathion into the human body readily results in its metabolism to malaoxon, which is substantially more toxic (Edwards, 2006). Also, malathion has been known to be toxic to many species of aquatic life forms, including fish (Burdick, 1967; Bender, 1969); and beneficial insects such as honeybees (Bai and Reddy, 1977; Sharma and Abrol, 2005). Insecticides such as permethrin were often used to treat ITNs. It is a neurotoxin which is found to have more acute effects in cold-blooded animals than warm-blooded animals. Thus there is a concern for the potential impact to the environment from permethrin usage. An example is a study carried out on the effect of chemical fogging in natural field conditions in Malaysia using insecticide Aqua-Resigen which consisted of S-bioallethrin 0.14% w/w, permethrin 10.11% w/w, piperonyl butoxide 9.96% w/w, and inactive base 79.79% w/w (Chua et al., 2005). The study found that not only was chemical fogging ineffective in breaking the reproduction lifecycle of *Ae. aegypti*, but there were evidence of deaths to other insects and natural predators of mosquitoes such as spiders and frogs. Neurotoxin effects are also seen in humans when exposure is high (Baselt, 2008). There was also

a study which suggested a link of permethrin exposure to Parkinsons disease (Bloomquist et al., 2002).

Another problem in using insecticides is the development of resistance in the vectors. Most resistance machanisms can be divided into two groups: i) metabolic resistance, in which activities or levels of detoxification proteins such as cytochrome P450s, carboxylesterases, and glutathione transferases, are altered; ii) target site alteration, where mutations in the sodium channel, acetylcholinesterase and gamma-aminobutyric acid (GABA) receptor genes causes insecticides to be less effective (Hemingway et al., 2004).

Resistance to insecticides appeared in insects from every major genus (Brogdon and McAllister, 1998). As of 1992, the list of insecticide-resistant vector species included 56 anopheline and 39 culicine mosquitoes, body lice, bedbugs, triatomids and eight species of fleas (WHO report, 1992).

Resistant populations of *Ae. aegypti* have been detected in several countries throughout the geographical range of this species (Rodriguez et al., 2000; da Cunha et al., 2005; Cui et al., 2006; Jirakanjanakit et al., 2007) and, in some areas, insecticide resistance has been linked to failure of the dengue control programme (Cui et al., 2006; Coto et al., 2000; Focks et al., 2000). In Brazil, it has been reported that *Ae. aegypti* is resistant to organophosphate insecticides malathion, fenitrothion and temephos that were extensively used to control dengue outbreaks (Lima et al., 2003). Mosquito populations in Cuba were reported to be resistant to pyrethroid insecticide, deltamethrin, and DDT (Rodriguez et al., 2005). In South East Asia, all strains of *Ae. aegypti* larvae showed some degree of resistance in Malaysia (Chen et al., 2005), while in Thailand deltamethrin and permethrin resistance was reported in wild populations in Ratchaburi Province (Paeporn et al., 2004). In the Carribean, *Ae. aegypti* strains in some countries such as Antigua, St. Lucia, and Tortola

had consistently high resistance ratios to temephos (Rawlins, 1998).

1.2.3 Alternative control strategies

Alternative strategies are increasingly being explored to overcome the obstacles in traditional vector population control. Some are proving effective as alternative control methods, while others are still in field trials.

1.2.3.1 Biological control

Bacillus thuringiensis (Bt) spores and crystalline insecticidal proteins have been used to control insect pests since the 1920s (Lemaux, 2008). Due to their specificity, these pesticides are regarded as environmentally friendly, with little or no effect on humans, wildlife, pollinators, and most other beneficial insects. The Belgian company Plant Genetic Systems was the first company (in 1985) to develop genetically engineered (tobacco) plants with insect tolerance by expressing endotoxic genes from *B. thuringiensis* (Hofte et al., 1986). *Bacillus thurigiensis*-based insecticides are often applied as liquid sprays on crop plants (as permitted in organic farming), where the insecticide must be ingested to be effective (Broderick et al., 2006). Upon sporulation, *B. thuringiensis* forms crystals of proteinaceous insecticidal delta-endotoxins (called crystal proteins or Cry proteins), which are encoded by *cry* genes (Crickmore, 2010). In most strains of *B. thuringiensis* the *cry* genes are located on the plasmid (Stahly, 1984). Recent research has suggested that the midgut bacteria of susceptible larvae are required for *B. thuringiensis* insecticidal activity (Broderick et al., 2006). Findings from Broderick et al., 2006 showed that *B. thuringiensis* does not induce death when insect gut bacteria were absent.

Bacillus thuringiensis israelensis (Bti) is widely used as a larvicide against

mosquito larvae, and it is considered an environmentally friendly method of mosquito control. Trials on dengue vectors *Ae. aegypti* and *Ae. albopictus* using *B. thuringiensis israelensis* were conducted in countries with known widespread resistance to chemical larvicide, temephos, and has been found to be effective in eliminating vector larvae (Benjamin et al., 2005). There are concerns, however, that widespread *Bti* resistance may also develop in mosquitoes. Evidence of high-level *Bti* resistance has already been reported in *Cu. quinquefasciatus* populations in Thailand (Su and Mulla, 2004). Other biological control of mosquitoes include the use of fish such as *Gambusia affinis* (Meisch, 1985) and crusteans called cyclopoid copepods (Nam et al., 1997) that eat mosquito larvae. Larvae of the *Toxorhynchites* genus (eg. *Toxorhychites splendens*) which preys on other mosquito larvae (Chuah and Yap, 1984), and Microsporidia fungi which infects insects (Koella et al., 2009) were also used to control mosquito populations.

1.2.3.2 Sterile Insect Technique (SIT)

A powerful alternative to insecticides is the Sterile Insect Technique (SIT) first used successfully to control the crop-devastating screwworm fly (*Cochliomyia hominivorax*) in the Americas (Knipling, 1955). SIT is environmentally friendly and species-specific. Males are bred in large numbers (females that are produced are separated and largely discarded), sterilised at pupal stage by irradiation or chemically, and then released at adult stage into the wild population. The sterile males will mate with the females in the wild, which will result in failure to produce offspring by females. Large over-flooding ratios can be achieved by releasing sterile insects into seasonally or conventionally suppressed populations (Benedict and Robinson, 2003). This technique has been successful in controlling the aforementioned screwworm fly and the Medfly

(*Ceratitis capitata*) in Chile and areas of Southern Mexico and California (Hendrichs et al., 1995). The tsetse fly (*Glossina spp.*), which transmits trypanosomiasis, was also successfully controlled by SIT in African countries such as Zanzibar (Msangi et al., 2000), Bukina Faso (Politzar and Cuisance, 1984) and Nigeria (Oladunmade et al., 1990). SIT was used successfully to control *An. albimanus* in El Salvador, reducing the local population in Lake Apastepeque, El Salvador by 99% (Weidhass et al., 1974). On an island off Florida in the USA, *Cu. quinquefasciatus* was eliminated using chemosterilised (sterilised using chemicals such as caesium or thiotepa) males (Patterson et al., 1970).

However, SIT is used against only a modest range of pest species (Gong et al., 2005). Irradiation can negatively impact subsequent performance of the insects to be released (Barry et al., 2003; Lance et al., 2000; Shelly et al., 1994; Kraaijeveld and Chapman, 2004). Problems include the poor dispersal of released males and the reduced longevity. In the case of Medfly, these effects include reduced number of emergence, flight ability, and survivorship (Lance et al., 2000). The less fit insects are thus quickly lost in the natural environment. This is also seen in mosquitoes: irradiated *An. arabiensis* had difficulty competing with wild males for wild females (Helinski et al., 2006). Field trials of chemosterilised *Ae. aegypti* (Seawright et al., 1978) and *An. culicifacies* (Reisin, 1981) also demonstrated poor competitiveness despite finding of equal competitiveness under laboratory conditions. Keeping large numbers of pest insects in an SIT programme runs the risk of accidental release before they are sterilised. One such incident did occur in Mexico in 2003 where non-irradiated New World screwworms were released unintentionally (del Valle, 2003). Another limiting factor is that females must be removed from the population to be released, as they are still potential pests regardless

of sterility. Therefore, sorting large numbers of insects is very time consuming, costly and is prone to errors (Baker and Wolfner, 1988). Another technical problem is the immigration of target species into release areas. In India, large releases of *Cu. quinquefasciatus* were carried out, but only modest effects were observed (Pal, 1974). The cause of failure was due to unexpected migration of females that were already mated with wild type males over several kilometres into control areas.

1.2.3.3 Genetic transformation as a control strategy

Advances in genetic technology enable us to explore the possibility of altering vectors genetically to either reduce or replace the wild population with refractory strains. Transformations made in *Drosophila* were achieved some time ago using a transposable element called P element to deliver gene(s) into the genome (Rubin and Spradling, 1982). Transformation was achieved only recently in the mosquito with the discovery of new transposable elements *Hermes* and *Minos*. A notable use of this type of genetic modification in the mosquito was accomplished by a group in Imperial College London (Catteruccia et al., 2000). They have successfully developed transgenic sexing lines for the mosquito *An. stephensi*, using *Minos* as the delivery system (Catteruccia et al., 2000). Male mosquitoes were made to express enhanced green fluorescent protein (EGFP) under the control of the testis-specific β2-tubulin promoter, thus enabling identification by their fluorescent gonads in early larval stages, which facilitated their efficient separation from females (Catteruccia et al., 2005). This progress could significantly increase the effectiveness of programmes such as SIT.

1.2.3.4 Release of Insects carrying a Dominant Lethal (RIDL)

Thanks to advances in genetic manipulations mentioned previously, Release of Insects carrying a Dominant Lethal (RIDL) was made possible and could solve the problems encountered in SIT programs. RIDL is a new technique in controlling insect population by inserting a dominant lethal gene into the insects' genome. It was first used in a way that the gene was only expressed in females, where tetracycline-repressible transcriptional activator fusion protein (tTA) was expressed under the control of the ovarian fat body activator called yolk protein 3 (*Yp3*) (Thomas et al., 2000). *Yp3* promotes expression in female larvae and adults, but not in males. Expression of the cytotoxic tTA is thus expected to occur in the *Yp3* pattern, allowing only males to survive and to be released into the environment (Thomas et al., 2000). This creates a genetic system for hereditary sexing. Current RIDL technology is used in both sexes. The released males of the modified insects are not sterile, therefore are able to mate and produce offspring with wild females. Any offspring produced will then inherit the dominant lethal gene from the father and die, if it is a female, leading to a decline of females in the wild population. An example of this is the Mediterranean fruit fly that is engineered to carry tTA that causes lethality in early developmental stages of the heterozygous progeny but has minimal effect on the survival of parental transgenic tTA insects (Gong et al., 2005). The released flies are kept alive by giving a dietary additive (in this case, antibiotic tetracycline), but they and their progeny will die in the wild once that supplement becomes unavailable. This technique also would mean that accidentally released insects would not survive past one generation in the wild. RIDL technique was also replicated in *Ae. aegypti* (Phuc et al., 2007), where the progeny of transgenic *Ae. aegypti* strain LA513A will die during the late larval-pupal boundary of development. This mosquito strain is currently

undergoing release trials in Malaysia by Oxitec to test the effectiveness in population control (Cyranoski, 2008). Transgenic *Ae. aegypti* were further engineered to combine all of the genetic features to produce a strain that is highly penetrant, dominant, late-acting, and has female-specific lethality, by using a promoter derived from *Ae. aegypti* called Actin-4 (Fu et al., 2010). This construct gave a flightless phenotype in females due to the disruption in the development of female-specific flight muscle.

RIDL has certain advantages over SIT. RIDL insects produced are in theory more competitive in the wild. There is no risk associated with radiation or with the release of non-irradiated insects and thus it can be applied to insects that do not tolerate radiation. Also, RIDL is not associated with high financial costs. Moreover, the RIDL that induces late stage lethality is shown to be more effective for control as 'doomed' progeny as larvae will compete for the same resources as the wild-type. The most hopeful type of sites for eradication using RIDL technology in release programmes is urban areas with one species of vector, but where the surrounding rural area has another. This is because the target species would be an isolated population and thus no unexpected migration of females that mated with wild-type males from another region as reported by Pal, 1974. An example of such cases are *An. stephensi* in Indian cities (Rao, 1984) and *An. arabiensis* in Southern-Nigerian cities surrounded by *An. gambiae* in rural areas (Coluzzi et al., 1979; Kristan et al., 2003). Incidence of dengue surged in the 1990s in Singapore, despite successful vector control to keep mosquito population low (Goh, 1998). One of the explanations is that the human population in Singapore now has little immunity because of the impact of legally enforced larval control (Ooi et al., 2001). This form of larvae control has not been able to eradicate *Ae. aegypti* completely. An attempt to do so with RIDL males would seem

to be appropriate. Mathematical models for mosquito engineered carrying technology like RIDL, based on data for dengue fever, suggest that eradication is feasible for affected human populations (Atkinson et al., 2007).

1.3 Ethical considerations regarding release of transgenic mosquitoes

Ethical concerns about introducing genetically modified mosquitoes into the environment must be addressed and resolved satisfactorily before such a strategy is deployed. These include ensuring public understanding and acceptance of genetic modification. In some instances the understanding of the role of particular mosquito species in disease transmission and awareness of the possible unknown and potentially serious side effects of a release are needed. Furthermore, acceptance by one community, or even country, is likely to affect many of its neighbours, whether they agree with the decision to release or not. Accidental release may occur from an outdoor cage trial. Lessons can be learned from SIT programmes that encountered such ethical boundaries. In the Indian SIT trial the ambitious programme of releasing SIT mosquitoes sparked allegations of research supporting biological warfare by certain sections of the Indian media and politicians (Editorial, 1975; Sehgal, 1974). Accidental release of non-irradiated flies in SIT programmes is known to happen. An example was the accidental release of the New World Screwworm flies (*Cochlimyia hominivorax*) in SIT plant in Tuxtla Gutierrez, Chiapas, Mexico (del Valle, 2003). This potentially hazardous error could have been prevented if transgenic sexing strains were used. Concern about releasing transgenic insects is the potential of uncontrolled spread of transgenes since transgenic insects are not sterilised as in SIT which limits the spread of the released insects. Therefore, careful considerations and research into release sites are needed to ensure containment of transgenes. Like many genetically modified organisms, resistance against transgenes develop due to the natural selection process. If organisms carry deleterious genes, such genes

will eventually be eliminated in nature.

1.4 Benefits of transgenic sexing strains

Vector control via transgenic mosquitoes is a species-specific control method. Thus its environmental impact is more acceptable than the non-specific mosquito population controls. Transgenic sexing strains also provide a safeguard in case of accidental release, as strains cannot survive in the wild without additives such as tetracycline.

1.4.1 Cost effectiveness

The genetic sexing system is effective as only males are being released into the environment (Rendon et al., 2004). More males in reared populations means less killing and discarding of females can be difficult to separate by other means. This should improve production of transgenics and in the long term reduce expenditure on rearing of transgenics. However, a study into cost effectiveness of a medfly strain with a mutation called temperature sensitive lethal (*tsl*) compared to a bisex strain found that the cost of production per million male pupae was not different between bisex strain production and *tsl* strain production (Caceres et al., 2000).

Diet for mass-rearing insects for release is also a factor to consider. Larval conditions will directly impact the fitness of the adults they will develop into. It is often by trial and error that optimisation of diet to rear certain species is achieved (Benedict et al., 2009). The cost of high quality food for large-scale rearing could be high. A chemically defined diet was used to successfully rear several culicines (Dadd and Kleinjan, 1976) for SIT projects, but was

less successful in rearing more fastidious species such as *Anopheles freeborni* (Dadd et al., 1977). It was estimated that approximately 2 - 3 kg of larval food per day will be required for production of one million males per day in SIT programmes (Benedict et al., 2009). This problem is applicable to all mass-rearing facilities. Male biased transgenics may boost the output of 'useable' insects and so the overall cost of rearing on high quality diet, if need be, would be justified.

Production cost of insects reared with tetracycline is low (Marrelli et al., 2006), as only a small amount of tetracycline is needed for rearing. This may be advantageous compared to SIT where cost of running an irradiation facility is high. However, actual costs of using genetic sexing systems in comparison to SIT have not been compared. Mass rearing of *An. albimanus* in the early stages in SIT programmes in El Salvador was fraught with a high number of diseased larvae or "bad trays" (10%-16%) from which few or no pupae were harvested (Kaiser et al., 1979). Increased hygiene during rearing such as rinsing eggs with water prior to hatching, was able to reduce the infection in the facility. However, additional use of antibiotics such as tetracycline may have additional benefit to mass-rearing facilities, as it could keep bacterial infection at bay.

1.4.2 Fitness cost

Fitness can be defined as the relative success with which a genotype transmits its genes to the next generation (Marrelli et al., 2006). In many SIT programs, the reduction in fitness is a result of mass-rearing conditions, radiation-sterilisation and the rigours and inefficiencies of the release protocols. To compensate for fitness reduction, the number of transgenic insects released must be up to 100 times the number of wild insects in the target area (Marrelli

et al., 2006). For driving genes into the wild populations, fitness reductions must be compensated for by the development of a driver gene with sufficient force to overcome the fitness deficit (Gould and Schliekelman, 2004; Ribeiro and Kidwell, 1994).

Fitness reduction through insertional mutagenesis is not frequent (Marrelli et al., 2006). Studies on the effect of insertional mutagenesis of P elements in *Drosophila* has shown that the effects on viability are recessive (Lyman et al., 1996). In fact, in this study, the transgenic line generated by P element insertion marked with an eye colour gene, *rosy* (ry^+), appears to improve the viability - the heterozygous line were fitter than the wild type.

In mosquitoes, the impact of transgenes on fitness of mosquitoes is not always deleterious. Studies on the fitness of anopheline mosquitoes carrying two different transgenes that cause them to be unable to transmit *Plasmodium berghei* (Moreira et al., 2004) showed that one line (SM1) had no detectable fitness load compared to the non-transgenic controls (Ghosh et al., 2001; Ito et al., 2002). The second line, PLA2, however competed poorly with the non-transgenic strain in cage experiments, and the transgenic allele almost disappeared by fifth generation (Zieler et al., 2001; Moreira et al., 2002). These transgenic strains were kept heterozygous, therefore the direct relation of fitness cost and presence of transgene is unclear. The ability of transgenic *An. stephensi* to compete with wild-type mosquitoes of the same species in cage trials showed that transgenic allele frequency decreased sharply until extinction in the four homozygous lines examined (Catteruccia et al., 2003). Fitness of three homozygous transgenic lines of *Ae. aegypti* carrying GFP markers, showed that there were high fitness cost when compared to their non-transgenic counterpart (Irvin et al., 2004). Survivorship was significantly reduced in all life stages and mortality rate during transition from egg to

larva was higher. Fecundity and adult longevity was also decreased in two of the observed lines. Irvin et al (2004) found that one of the lines had an active *Hermes* transposase that catalysed somatic transposition of a *Hermes* element that could have caused the deleterious effects.

Hence, according to Marrelli et al. (2006), the decreased fitness in homozygous transgenic mosquitoes can be brought about by two mechanisms:

- Decreased fitness may be a consequence of either negative effects of the transgene product or 'insertional mutagenesis' during transgene transposition.
- Decreased fitness may be a consequence of the hitchhiking effect. Genetic hitchhiking is the process by which a neutral or deleterious allele or mutation may spread through the gene pool by virtue of being linked to a gene that is positively selected (Barton, 2000).

The LA513A strain from Oxitec ltd. was found to have little negative effects on fitness cost. In fact, recent findings using cage trials for mating competitiveness showed that when this strain is mass-reared to a certain size that is not too large, the fitness can be higher than that of the wild type strain Rockefeller (K. P. Hoang, personal communication).

1.5 Limitations of transgenic strains

In order for mosquito population control to be successful, the target population should ideally be isolated, so that the transgenic strains are able to mix into that population and induce a population crash. If the population has a high genetic flow by migrating mosquitoes, the effect of introducing transgenic strains will be lost. Moreover, the transgenes that were introduced could potentially spread into other populations in other areas. Targets for genetic control of mosquitoes are therefore most ideal on islands, or 'urban islands'. Examples of 'urban islands' include *An. arabiensis* populations in Southern Nigerian cities that are surrounded by rural *An. gambiae s.s.* (Kristan et al., 2003). Another example is *Ae. aegypti* populations in urban areas, with the distribution of *Ae. albopictus* in rural areas (Chan et al., 1971a; Chan et al., 1971b; Hawley, 1988; Tsuda et al., 2001; Braks et al., 2003).

Even so, releases may not be successful due to many unforeseen factors. For example, island trials previously carried out in Taiaro, a small atoll in French Polynesia, using *Ae. albopictus* to displace indigenous *Ae. polynesiensis* as a control measure for filariasis failed during the 1970s (Rosen et al., 1976). All four strains of *Ae. albopictus* released were eliminated on the island within four years. The exact cause of failure of the trial were unexplained, since cage trials and semi-field trials that were carried out prior to the field trials indicated that *Ae. albopictus* could displace the *Ae. polynesiensis*. Several factors were postulated: i) Unsuitable vertebrate host for blood meal ii) failure to adapt to new environmental conditions of the atoll that could be due to unfavourable strains used iii) The indiscriminate mating behaviour of male *Ae. albopictus*. Even though both male and female *Ae. albopictus* were used in these release programmes, the trial is still relevant to other programmes such as SIT and other transgenic releases. In short, survival, dispersal, fecundity, regeneration

of sperm, ratio of released mosquitoes to wild ones, and competitiveness of released males may all contribute to the efficacy of release operations (Lacroix et al., 2009).

Anopheles species are infamous for their association with malaria. Approximately 40% of the worlds population are at risk of malaria, with more than 500 million people affected by this disease annually. Most cases and deaths are in sub-Saharan Africa, although other tropical and subtropical regions are also affected (WHO, 2009b). The mosquito *An. gambiae* is the major vector of malaria in Africa. But the disease is also spread by other species of the same genus around the world (see figure 1.4), as well as sibling species complex. This causes complications for malaria control strategies when multiple vectors are present in the same area. This has been observed in Africa where up to 5 different anopheline species are found to be malarial vectors either simultaneously or seasonally (Fontenille and Simard, 2004). Using transgenic mosquito would not be effective, as all sibling species within the target area would need to be engineered and released simultaneously in a program. This method is obviously not feasible to be carried out.

Sibling species complexes also occur in culicines, where it has been known that *Culex* and *Aedes* have their own sibling species complexes. One example is the *Culex pipiens* complex, which are widely distributed around the world. These include *Culex pipiens*, *Culex molestus*, *Culex quinquefasciatus* and *Culex pallens* (Smith and Fonseca, 2004).

Reproductive isolation is also one factor to consider. For example, *Culex pipiens molestus* that was found to populate the London Underground railway system had two genetically distinct populations - the surface and subterranean populations (Bryne and Nichols, 1999). In the study, there was no evidence of gene flow between closely adjacent populations of the different forms (Bryne

and Nichols, 1999). There was little differentiation between the different populations of each form. Reproductive isolation such as this would be problematic given that it may mean that the release of transgenic strain may not mate with all of the wild population.

With all these potential limitations mentioned above, extensive surveys are needed to analyse all possible factors prior to release of transgenic insects.

The image originally presented here cannot be made freely available via ORA because of copyright. The image was sourced at Kiszewski et al., 2004 (weblink: www.ajtmh.org/content/70/5/485/F1.expansion.html

Figure 1.4: Global distribution of dominant or potentially important anopheline vectors of malaria (Kiszewksi et al., 2004)

1.5.1 Potential problems with using tetracycline

Tetracycline in ground and surface water due to agricultural usage may compromise the safety of using tetracycline in mosquito control programs using the tet system. This is because transgenic mosquitoes based on the tet system may be able to breed in the wild. However, wheter the concentration of the antibiotic in the environment is enough to cause an effect is unknown since no studies of this kind have been carried out.

A study reported that much as 0.2 μg per kg of tetracycline has been detected in soil due to manure from animal husbandry practices (Hamscher et al., 2002). The amount of tetracycline stated is well below the concentration needed for transgenic mosquitoes to survive. For example, *Ae. aegypti* LA513A strain survival threshold is 5 μg/l of tetracycline, but some could survive with a concentration of as little as 1 μg/l (K. P. Hoang, personal communication). The probability of tetracycline in outdoor water being a problem to tet system transgenic release is therefore considered unlikely. Nonetheless, there is still a lack of fundamental data on the occurrence, fate and effects of antimicrobials in the environment needed for proper risk assessment and risk management both for humans and the environment (Krümmerer, 2003). Tetracycline also affects mosquito growth since any natural gut flora would be eliminated by the antibiotic, which can become a fitness cost (Mourya and Soman, 1985).

1.6 Possible future developments in mosquito population control

1.6.1 Homing endonuclease genes

Homing endonuclease genes (HEGs) are elements that encode endonucleases which recognise a specific sequence flanking the HEG. If HEGs are introduced into a target sequence on one of the pairs of homologous chromosomes of a diploid organism, it will cleave the chromosome which does not contain the HEG sequence and then gets copied across to the broken chromosome as a byproduct of the repair process (Belfort and Roberts, 1997; Chevalier and Stoddard, 2001; Burt and Koufopanou, 2004) as repair mechanism of the cell would use the intact HEG containing chromosome as a template for repairs (see figure 1.5).

Because of these properties, HEGs are promising candidates to transfer genetic modifications from engineered laboratory mosquitoes to wild-type populations. It was shown that homing endonucleases I-SceI and I-PpoI could cleave their recognition sites with high efficiency in *An. gambiae* cells and embryos (Windbichler et al., 2007). In fact, it was found that I-PpoI cuts genomic rDNA located on the X chromosome in *An. gambiae*, though this proved to be lethal since all cells contain at least one copy of the X chromosome. HEG could be used in the future as a gene drive system to cleave novel DNA sequences, potentially sex-determining genes.

Figure 1.5: Diagram of how HEG works. A specific homing endonuclease gene (HEG) is typically found inserted between two specific sequences of DNA within the genome (light green). The HEG (dark green) codes for the production of an enzyme that recognises these two specific coding sequences when they are not interrupted by the presence of an HEG. In individuals that carry the HEG on only one of two homologous chromosomes, the enzyme catalyses a break within the DNA sequence of the chromosome that lacks the HEG (step 1), which is then naturally repaired using the HEG within the homolog as a template (step 2). Diagram from Sinkins and Gould (Sinkins and Gould, 2006).

1.6.2 Medea

One proposed strategy for controlling the transmission of insect-borne pathogens uses a drive mechanism to ensure the rapid spread of transgenes conferring disease refractoriness throughout wild populations (Chen et al., 2007a). A new promising system that is able to do so is Medea (Maternal effect dominant embryonic arrest). The design of this synthetic element is based on a naturally occurring selfish genetic element first discovered in the flour beetle *Tribolium castaneum*. Medea is composed of a toxin and an antidote: a mother with the gene will express the toxin in her germline and kills her progeny unless a copy

of the gene is inherited (see figure 1.6).

Medea's selfish behaviour gives it a selective advantage over other genes. If introduced into a population at sufficiently high levels, the Medea gene will spread rapidly, replacing entire populations of flies with flies carrying Medea (Wade and Beeman, 1994). A Medea-like effect was successfully produced in *D. melanogaster*: the toxin was a microRNA that blocked the expression of *myd88*; the antidote was an extra copy of *myd88* that was insensitive to the microRNA (Chen et al., 2007a). Maternal *myd88* is required for dorsal-ventral pattern formation in early embryo development.

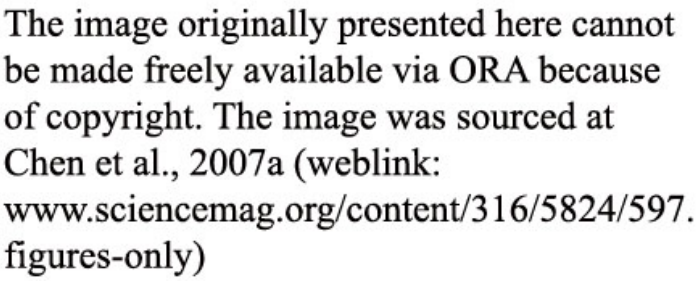

Figure 1.6: Diagram of how Medea works. It is postulated that females heterozygous for Medea (Medea/+) deposit a protoxin or toxin (red dots) into all oocytes. Embryos that do not inherit a Medea-bearing chromosome from either parent die because toxin activation or activity is unimpeded (bottom left square). Embryos that inherit Medea from the maternal genome (top left square), the paternal genome (bottom right square), or both (top right square) survive because zygotic expression of a Medea-associated antidote (green background) neutralises toxin activity. Adapted from Chen et. al, 2007a.

Medea is therefore a promising system that can be adapted to mosquitoes. By linking the Medea construct to a gene of interest such as disease refractoriness, Medea could be used to drive such genes into the population. For control programs, the release of male-only, non-biting mosquitoes would be important for the acceptability of this strategy. Medea can be used by releasing male-only transgenics to spread within the population. When wild type females inherit the Medea construct, the maternal toxin would activate to kill all progenies without the antidote. Therefore, the only embryos that inherit Medea from the maternal genome, paternal genome, or both can survive. Lab trials where 25% of the original members were homozygous for Medea, was able to spread the gene to the entire population within 10 to 12 generations. It is also resistant to recombination-mediated dissociation of drive and disease refractoriness functions (Chen et al., 2007a).

1.6.3 *Wolbachia*

Wolbachia are parasitic intracellular bacteria which are maternally inherited in arthropods, often resulting in manipulation of host reproduction to enhance transmission of infection. *Wolbachia* was first discovered in the 1924 by M. Hertig and S. B. Wolbach in the mosquito *Culex pipiens*, and thus named the bacterium *Wolbachia pipientis* (Hertig and Wolbach, 1924). *Wolbachia* could be used in alternative strategies for mosquito control by inducing Cytoplasmic Incompatibility (CI) in populations, and are able to use these patterns of sterility to spread themselves through populations (Sinkins, 2004). This is because *Wolbachia* induces embryonic lethality that results from crosses between infected males with uninfected females, to rapidly spread into insect populations (Hoffman and Turelli, 1997). Unidirectional CI occurs when the sperm from a *Wolbachia* infected male fertilises an uninfected egg which results in embryo mortality. However, the reciprocal cross (figure 1.7), between infected female and uninfected male is compatible. Unidirectional CI is important for its potential use in invasion of *Wolbachia*-uninfected mosquito populations. The first major success in mosquito control using unidirectional CI was against *Culex quinquefasciatus* in Myanmar (Laven, 1967), where *Cu. quinquefasciatus* carrying a different *Wolbachia* strain than the one carried by the isolated target population were introduced, causing CI and effectively eliminating the wild population. In a mixed population, females that contain *Wolbachia* will produce a greater number of offsprings. Therefore, carrying *Wolbachia* is advantageous for females, but is deleterious for males, as it reduces fertility in crosses with uninfected females (Turelli, 1994; Charlat et al., 2003).

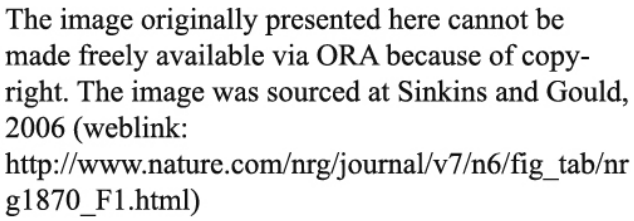

Figure 1.7: *Wolbachia* induced unidirectional CI in mosquitoes. Red denotes *Wolbachia* infected individuals. Non-infected individuals are in green. The crossing patterns result in non-viable eggs when a *Wolbachia* infected male mates with an uninfected female. *Wolbachia* infected females however produce infected progeny in all matings. Modified from Sinkins and Gould, 2006.

Wolbachia could be used to shorten lifespan of a host, as seen in a strain called *w*MelPop. It was found that *w*MelPop could shorten the lifespan of adult *D. melanogaster*, as well as causing CI (Min and Benzer, 1997). This strain was successfully transferred into *Ae. aegypti* (which naturally do not carry *Wolbachia*) and cause the same life-shortening effect; in fact, the lifespan was halved (McMeniman et al., 2009). Further study on *w*MelPop in *Ae. aegypti* found that the life-shortening effect could be due to the increase in up-regulation of innate immune response in the host caused by the bacteria itself,

and thus protecting the host from other invading pathogens and parasites such as filarial nematodes that cause lymphatic filariasis (Kambris et al., 2009), making *w*MelPop an alternative tool to eradicate lymphatic fliariasis, and possibly other mosquito-borne parasites and other diseases. Indeed, *Wolbachia* derived from *Ae. albopictus* called *w*AlbB that was transfected into *Ae. aegypti* were also able to inhibit dengue virus replication, dissemination and transmission (Bian et al., 2010). It has also been demonstrated that not only dengue, but Chikungunya, and *Plasmodium* parasite infection were also inhibited when *Ae. aegypti* were infected with *w*MelPop *Wolbachia* (Moreira et al., 2009).

Using CI caused by *w*MelPop infection that is able to drive itself through populations, as for other *Wolbachia* strains, makes this a very good candidate for an alternative disease control strategy. *Wolbachia* may also be more easily accepted by the public, since *Wolbachia* already exist in nature. For a release programme, due to the nature of *Wolbachia* being maternally inherited, a small percentage of the release will need to be female. Therefore, a male biased release programme would be ideal as the male infected with *Wolbachia* will still act as a control measure by causing sterility in wild type females, and the small percentage of infected females would serve to introduce the *Wolbachia* into the wild population. There is also a concern that *Wolbachia*, as well as *Medea* may be difficult or impossible to stop from spreading beyond the target population, perhaps even to all populations of the species worldwide (Alphey, 2009).

1.6.4 Conclusion on gene drive systems

The gene drive systems mentioned are not perfect and could not totally solve the problem of vector-borne diseases. Self-spreading systems will undoubtedly fail over time, due to mutation and pathogen evolution, and replacement versions will be required (Alphey, 2009). It has been argued that SIT based methods may be more effective especially if coupled with an integrated vector management programme, but it will require constant releases. All of these methods will have to be tested in order to determine where each is more or less valuable (Alphey, 2009). Nonetheless, these new technologies will provide chances to advance the field of vector control in an effort to eliminate major vector-borne diseases. A problem facing these release programmes is the sexing of the insects and discarding the unwanted females. Therefore, a reliable sexing system is needed in all release strategies since female release will most likely face controversy. The studies reported here provide a promising route to genetic sexing in *Ae. aegypti*.

1.7 Sex ratio distortion

Several of the strategies under development would benefit from an ability to genetically distort the ratio of the population towards males, since males do not bite people and thus can be released much more readily. Distortion of sex ratio could also be used to reduce female population size as a variant form of SIT. This could be one way forward to overcome some of the problems encountered in other types of genetic manipulation and will be emphasized in this study.

There are two possible approaches in theory:

- Use the natural male determining gene at the top of the male sex determination cascade, if it can be identified.
- Genetically inhibit female determining genes.

1.7.1 Sex determination in insects

The determination of sexual fate in an individual is a key event in the development of higher eukaryotes. It begins with a primary sex-determining signal, which is then relayed through a cascade of genes to a terminal regulator gene that activates or represses effector genes responsible for sexual phenotypes. In dipteran insects, it is found that they use the same basic strategy: a primary sex-determining signal, a key gene that responds to the primary signal, and a "double-switch" gene that selects between two alternative sexual programmes (Schütt and Nöthiger, 2000) and therefore dictates the sexual fates.

Primary signals are very diverse in nature, even within the taxonomic group of dipteran insects (Marin and Baker, 1998; Schütt and Nöthiger, 2000). The

most common and ancestral mechanism seems to be the use of a dominant male determining factor (Marin and Baker, 1998), located on any chromosomes. It can be found in many dipteran species, such as *Aedes, Anopheles, Calliphora, Ceratitis, Chironomus, Megaselia* and *Musca* (White, 1973). The chromosomal balance system of *Drosophila*, using a complex mechanism of X chromosome and autosome counting, seems to be a rather uncommon sex determining system (Cline and Meyer, 1996). Dominant female determiners are found in some populations of *Musca* (Dübendorfer et al., 1992) and in taxa such as the silkworm *Bombyx mori* (Hashimoto, 1933). There are also cases where sex of the offspring is determined by the genotype of the mother, such as is shown in the blowfly, *Chrysomya fufifaciens* (Ullerich, 1984), where some female individuals only produce sons (arrhenogenic females), and others only daughters (thelygenic females). Environment could also play a part in sex determination as found in the sub-artic mosquito *Aedes stimulans* (Anderson and Horsfall, 1961; Nöthiger and Steinmann-Zwicky, 1985), where environmental stress such as alterations in temperature changes the sex ratio to favour female development. In *Ae. aegypti*, it was thought that the presence of the *intersex (ix)* gene induces the possibility of sex inversion in the male (Craig, 1965). Efforts to understand the phenomenon of this female biased sex ratio that was also seen in lab-reared *Ochlerotatus triseriatus* were not successful (Graham et al., 2004).

1.7.2 Heteromorphic Y chromosomes

In dioecious animals the division between male and female development commonly depends on the segregation of X-Y chromosome, whether the zygote is carrying XX or XY chromosomes (Mather, 1953). Many animals adopt this system, where the Y chromosome is present only in males and is responsible

for male sex determination.

Though *Drosophila* possesses the Y chromosome, it does not rely on it to determine adult sexual phenotype. The *Drosophila* Y chromosome was the first entirely heterochromatic chromosome to be characterised (Carvalho, 2002). However, identification of the Y-linked genes has been slow. It was proven that the Y is essential for male fertility (Bridges, 1916). It was found that X0 males were morphologically normal but sterile, YY individuals however die. The Y chromosome appears to be a peculiarly specialised and gene-impoverished chromosome (Bridges, 1916). Several genes found in the Y chromosome were found to be important for spermatogenesis (Brosseau, 1960). An example is the first single copy gene named *kl-5* that was found to encode a dynein heavy chain (Gepner and Hays, 1993). Besides single copy genes linked to male fertility (*kl-1, kl-2, kl-3, kl-5*) the Y chromosome was found to contain repetitive genes such as rDNA (Gatti and Pimpinelli, 1992) and more recently, aldehyde reductase Y (*ARY*) that is a potential member of the aldo-keto reductase gene family, involved in conversion of glucose to fructose, and in the inactivation of cytotoxic metabolites; and WD40 Y (*WDY*), the function of which remains elusive (Vibranovski et al., 2008).

It is thought that in mosquitoes, the genetic difference is smaller, judging by the lack of cytological differentiation in mosquitoes (Gilchrist and Haldane, 1947). In anopheline mosquitoes, the Y chromosome can be differentiated from the X chromosome (figure 1.8). The Y chromosome is important in determining maleness in the mosquito. This was evident when XXY mutants that have been isolated in the mosquito *An. culicifacies* were males (Baker and Sakai, 1979). However, the mechanism of this is still not understood. The *An. gambiae* genome project has not been able to assemble the Y chromosome due to the presence of large amount of repetitive sequences present. Nevertheless,

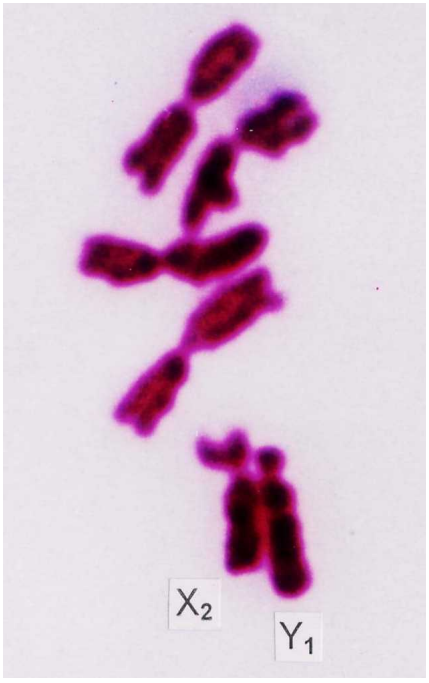

Figure 1.8: *Anopheles minimus* chromosomes at metaphase. An example of Giemsa stained heteromorphic Chromosome pairs. X_2 and Y_1 indicates the female and male forms of chromosome 1 respectively. Y_1 contains shorter p regions than X_2. Picture from Dr. Hoang Kim Phuc.

some Y-linked sequences were identified which could prove to be useful in assembling the chromosome (Krzywinski et al., 2004).

1.7.3 Homomorphic chromosomes

In culicine mosquitoes, including the *Aedes* and *Culex* genera, there are three pairs of homomorphic chromosomes in both sexes (Hickey and Craig, 1966), meaning that the sexual chromosomes seem to be indistinguishable. They are however distinguishable when Giemsa C-banding is used: when *Ae. aegypti* chromosomes are stained with Giemsa, there is a C- band in the centromere region located on one pair of the seemingly similar chromosome in females, but in males the C-band on one of the chromosomal arms is absent (Newton et al., 1974), as in figure 1.9. The un-banded chromosome arm is consistently found in males only, indicating that the sex locus occurs somewhere within a pericentric region, the minimum extent of which includes both the intercalary band and the centromere. In *Culex pipiens*, a recessive gene for white eye colour was approximately 6.3 crossover units from the sex locus. It was found that males were heterogametic (Gilchrist and Haldane, 1947). This gave rise to the theory that a single gene or a small chromosome segment determines sex. The sex locus was designated m, with females being mm and males Mm, and that M is the male determining factor. A similar mechanism in *Ae. aegypti* was found using a gene for red eye colour and is 6 to 7 units from the sex locus (McClelland, 1962). The precise position of the sex locus M/m in linkage group 1 is uncertain, but it is known to be situated towards the centre of the length accounted for (Bhalla and Craig, 1970). As one of the two sex chromosomes in males is consistently unbanded in the centromere region while both in females are exclusively banded, Newton et al. (1974) postulated that M and m are always linked with a particular type of centromere.

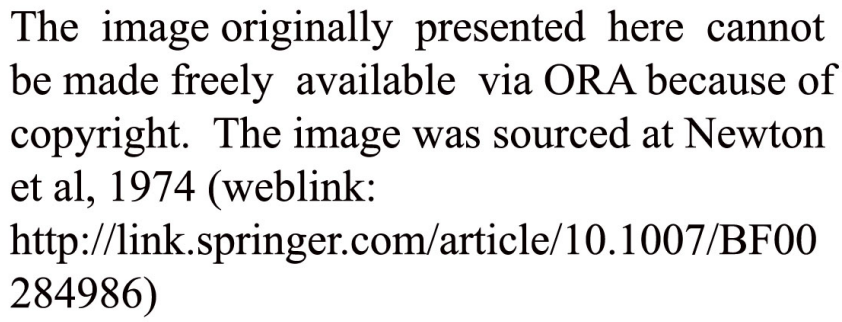

Figure 1.9: Giemsa stained mitotic metaphase chromosome showing distribution of centrometric and intercalary bands in *Aedes aegypti* (Newton et al., 1974). Numbers indicate chromosome type. It is clear here that the centrometric band on one of the chromosomal arms is absent in the male form of chromosome 1.

1.7.4 Sex determination in *Drosophila*

Drosophila is one of the most studied insects in the world. Its sex determination system has been widely used as a model and referenced for research in other insect species. Therefore, here the focus will be on the sex determination mechanism of *Drosophila* for understanding the mosquito sex determination. *Drosophila* does not utilise Y chromosomes directly for sexual differentiation, but the X chromosome number determines the sex of the insect. In *Drosophila*, there are four main genes in the sex determination cascade – *sex-lethal (Sxl)* being the top, *transformer (tra)* and *transformer-2 (tra-2)* in the middle, and *doublesex (dsx)* at the bottom (refer to figure 1.10).

There are three branches of the hierarchy - X chromosome dosage compensation, somatic sexual differentiation and male sexual behaviour. The primary signal in sex determination of *Drosophila* is the X chromosome to autosome ratio (X:A) ratio. Individuals with an X chromosome to autosome ratio of 1 usually develop into female, whereas an X chromosome to autosome ratio of 0.5 results in male development, thus generating equal levels of X-encoded gene

products (Baker et al., 1994). The decisive event occurs early in embryogenesis just prior to blastoderm stage and every somatic cell is cell-autonomously assessed on its X:A ratio (Schütt and Nöthiger, 2000). The basic helix-loop-helix product of the X-linked gene called *sisterless B (sisB or scute)* is a key indicator of the X dose and functions to activate the switch gene *sex-lethal (Sxl)* in females (Cline, 1978; Yang et al., 2001).

The image originally presented here cannot be made freely available via ORA because of copyright. The image was sourced at Gilbert, 2000 (weblink: http//www.ncbi.nlm.nih.gove/books/NBK10025/figure/A4141/?report=objectonly)

Figure 1.10: Proposed regulation cascade for *Drosophila* somatic sex determination. Arrows represent activation, while a block at the end of a line indicates suppression. The primary signal in sex determination is the X chromosome to autosome ratio (X:A) ratio. Individuals with an X chromosome to autosome ratio of 1 develops into female, whereas an X chromosome to autosome ratio of 0.5 results in male development, thus generating equal levels of X-encoded gene products such as the *sisterless (sis)* genes (Baker et al., 1994). The *msl* loci, under the control of the *Sxl* gene, regulate the dosage compensatory transcription of the male X chromosome (Gilbert, 2000). Diagram adapted from Gilbert, 2000.

1.7.5 *Daughterless (da)*

The *daughterless (da)* gene in *Drosophila* encodes a transcriptional regulator that functions to control sex determination and neurogenesis (Cummings and Cronmiller, 1994). The *Drosophila daugtherless (da)* gene only has two exons and encodes a protein of 710 amino acids long, which are not sex-specific. Mutation in *da* lead to an all male progeny (Bell, 1954). Characterisation of *da* mutants has lead to the hypothesis that *da* maternal function is needed in females for the the expression of the gene *Sex Lethal* (see below for more details of *Sex Lethal*) and thus sex determination (Cline, 1978). The *da* mRNA plays a major role in regulating follicle cell formation and maturation during oogenesis and that its expression is of a multiphasic pattern in both the soma and germline (Cummings and Cronmiller, 1994). Apart from being important in the maturation of follicle cells during egg chamber morphogenesis in females (Cummings and Cronmiller, 1994), it is also critical in both males and females for neural differentiation. It was found that *da* is a cofactor of *achaete-scute* complex genes (consists of *achaete, scute, lethal of scute, and asense*) that activates a variety of other neural genes such as *prospero, cyclin A* and *calmodulin* (Cline, 1978; Erickson and Cline, 1991; Vassin et al., 1994; Kovalick and Beckingham, 1992; Garcia-Bellido and de Celis, 2009). Daughterless protein (DA) also interacts with Dorsal to bring about the induction of *twist* and *snail* genes that are required for gastrulation. Maternal DA proteins form heterodimers with zygotically expressed sisterless-B (sisB) that binds E-box sites (a DNA sequence which usually lies upstream of a gene in a promoter region) and activates transcription of *Sxl-Pe* promoter that activates female-specific splicing of mRNAs expressed from the *Sxl* maintenance promoter, Sxl-Pm (Yang et al., 2001).

1.7.6 *Sex Lethal (Sxl)*

Sex Lethal (Sxl) encodes an RNA splicing enzyme and a translational regulator. It is activated in females where the X chromosome to autosome (X:A) ratio is 1. It remains "off" in males where the X:A ratio is 0.5. Once set, these two modes of *Sxl* expression are maintained through the rest of the life cycle (Cline, 1976; Cline, 1984).

Sex lethal has two transcripts, early and late, with different promoters and dramatically different splicing patterns. The *Sxl* early transcripts are activated transiently in early embryos by a female-specific promoter and have a unique 5' exon (E1) located between late exons 1 and 2. Exon E1 is spliced to exon 4, which is common to all SXL transcripts, skipping both exons 2 and 3 (Keyes et al., 1992). In contrast, the late SXL transcripts derive from an essentially constitutive promoter but are spliced sex specifically. The male-specific exon, exon 3, is included by default in all male transcripts and contains in-frame nonsense codons that block SXL protein production (Zhu et al., 1997). In the presence of SXL protein, the late transcripts skip exon 3 and splice in the female pattern. No embryo-specific splicing factors are needed for the early splice. Neither are sex-specific factors required. Instead, the early splicing pattern is dependent on whether the 5' splice site region originates from exon E1 or exon 2 (Zhu et al., 1997).

In females, the SXL proteins are produced and act upon their own pre-mRNA such that in exon 3 of the *Sxl*, pre-mRNA will be spliced out to remove the stop codons. Thus, *Sxl* is also self-regulating to ensure its own expression is maintained in females (Graham et al., 2002). In somatic sexual development (development in the soma, or outer body), *Sxl* promotes female differentiation by controlling the female specific splicing of the *transformer (tra)* gene (Boggs

et al., 1987; McKeown et al., 1987). Sex Lethal is also needed in oogenesis. Pole cell transplantation experiments have demonstrated that female germ cells require SXL for regulating mitosis. Germ cells that lack SXL develop as tumorous cysts of many small undifferentiated cells (Schüpbach, 1985), a phenotype that is shared by the female sterile alleles of *Sxl*. The germline targets of SXL are not known, but they are not located in the soma (Marsh and Wieschaus, 1978; Schüpbach, 1985).

The male mode of *Sxl* splicing is maintained by default. There are five male-specific lethal *(msl)* genes: *msl-1, msl-2, msl-3, mle* (Kuroda et al., 1991; Palmer et al., 1993; Gorman et al., 1995) and *mof* (Hilfiker et al., 1997) identified to be involved in dosage compensation and characterised at a molecular level. *Sex-lethal* directs somatic sexual development by controlling the dosage compensation system (Lucchesi, 1978) and the somatic sexual differentiation pathway (Cline, 1976; Cline, 1978). *Sex-lethal* turns off the dosage compensation system through both splicing regulation and translational repression (Bashaw and Baker, 1997). With the absence or presence of *Sxl*, a cascade of sex-specific genes is activated by a series of alternative splicing events. *Sex-lethal* acts directly upon the next gene down the cascade, *transformer (tra)* and so manufactures a specific *tra* product, which then regulates to produce gender-specific *dsx*.

1.7.7 *Transformer (tra)* and *Transformer- 2 (tra-2)*

The first major gene downstream regulated by *Sxl* is the *transformer* gene *(tra)*, which acts with another gene called *transformer-2 (tra-2)*. Together, they are necessary for sexual differentiation of the somatic cells in females (Sturtevant, 1945; Brown and King, 1961; Watanabe, 1975; Fujihara et al., 1978; Belote and Baker, 1982; Steinmann-Zwicky et al., 1990) by forming a splicing enhancer complex which consists of TRA, TRA-2 and serine/arginine-rich (SR) proteins (Boggs et al., 1987). This then controls the alternative splicing of *doublesex (dsx)* and *fruitless (fru)* transcripts that are lower down the sexual hierarchy. Sex-specific regulation of *tra* is mediated by the alternative splicing of its pre-mRNA (Boggs et al., 1987).

The SXL protein regulates the choice between two variants of *tra* pre-mRNA that differs in its 3' splice sites (Sosnowski et al., 1989). In females, SXL bind to *tra* pre-mRNA and therefore allows *tra* to be expressed. A functional TRA protein is made, and this variant of TRA is found to be sufficient in feminising male germ cells. TRA shuts off male-specific markers and activates female-specific markers, depending upon the expression of the *tra-2* gene (Waterbury et al., 2000), which encodes a second RNA-binding protein that alternatively splice the *doublesex* pre-mRNA in females (Amrein et al., 1988; Baker and Wolfner, 1988; Burtis and Baker, 1989; Goralski et al., 1989; Inoue et al., 1992). TRA-2 functions as a regulator of sexual differentiation in the female germline tissues as well as in the somatic tissue (Mattox and Baker, 1991). The TRA-2 protein central domain spans approximately 80 amino acids and shows similarities to a family of proteins that bind single-stranded nucleic acids. On either side of the RNA recognition motif are serine/arginine (SR) motifs similar to those found in SR proteins (Amrein et al., 1988; Goralski et al., 1989). *Tra-2* uses alternative promoters and splicing patterns, to generate 7

different mRNAs (types A to G), of which four (two in male - type C and E, and two in both sexes- type A and B, although type B is more predominant in female germ line) encode three putative RNA-binding polypeptides (Mattox and Baker, 1991). Only three putative RNA-binding polypeptides are encoded because mRNA type B and E encode the same polypeptide (figure 1.11). Type A and B variants appear to be functionally redundant in directing female differentiation and female specific *doublesex* splicing by regulating the selection of alternative 3' end exons in the *dsx* pre-mRNA (Mattox et al., 1996). Analysis of the effects of *tra-2* on the genital disc shows that a functional TRA-2 is necessary in females to prevent male sexual differentiation and to permit female differentiation (Belote and Baker, 1982).

In the male germline, *Sxl* is not expressed in male embryonic cells, and with the absence of SXL protein, *tra* pre-mRNA follows a default non-sex-specific splicing, leading to mRNAs encoding shortened and presumably non-functional TRA protein (Sosnowski et al., 1989; Inoue et al., 1992). *Transformer-2* however has an important role in spermatogenesis. The types of male specific *tra-2* mRNA (C and E) are determined by an intron known as M1 between exons 3 and 4 (Mattox et al., 1996), where type C has M1 and type E has not. It was found that type E performs all the necessary functions for spermatogenesis. This type also appears to regulate its own synthesis during spermatogenesis through a negative feedback mechanism involving retention of intron 3 (Mattox et al., 1996). An additional transcript (type D) is similar to type A, but codes for a C-terminally truncated protein and is expressed at very low levels in males (Mattox et al., 1990; Mattox and Baker, 1991). Type A and type D *tra-2* mRNAs are therefore thought to encode functionally similar proteins, but each may play separate roles in RNA metabolism (Mattox et al., 1990).

The other two *tra-2* transcripts - type F and G, are also present in both sexes

and their function is unknown (Amrein et al., 1988; Mattox et al., 1990).

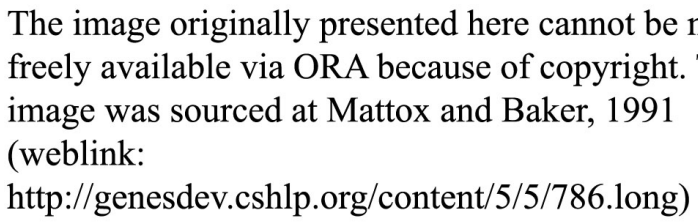

Figure 1.11: Alternative *tra-2* transcripts in *D. melanogaster*. The organisation of the *tra-2* gene is shown at the top. Exons are indicated by open boxes. The M1 intron is indicated by a hatched box. Other introns are indicated by blank regions between exons. The position of the male germ line-specific transcription start site is indicated by an arrow over exon 3. Shown below the gene are the differentially processed portions of each of the major *tra-2* mRNAs. Exons 5 - 7 are also present in each of these RNAs. Alternative splicing is predicted to result in translation of products initiating at three different AUG codons. The position of the predicted initiation codon in each mRNA is indicated along with the size of the expected polypeptide. Type B and E RNAs are expected to use the same initiation codon. This codon is formed in the splicing of exons 3 and 4. Diagram adapted from Mattox and Baker, 1991.

1.7.8 *Doublesex (dsx)*

Doublesex (dsx) is the master double-switch gene that determines the sex-specific expression of the genes down stream of *tra* and *tra-2*. *Doublesex* works in concert with other regulatory genes such as *engrailed, polycomb* and *extra sex combs* in *Drosophila* (Steinmann-Zwicky et al., 1990). As mentioned in the previous section, male or female specific *dsx* variants are produced as dictated by the TRA and TRA-2 proteins. The default *dsx* transcript produced in males encodes the protein that represses female development while the alternatively spliced *dsx* mRNA of females encodes a protein that represses male development (Walthour and Schaeffer, 1994). It has been shown that the male version of *dsx* is able to repress the genes that are involved in determining female fate such as *tra*, and vice versa (Waterbury et al., 2000). The female DSX protein (DSXF) has been ectopically expressed, and also plays an important role in sexual behaviour. Wild-type males ectopically expressing DSXF are actively courted by other males, thus feminine sex appeal acquired is likely due to the induction of female pheromones by female *dsx* (Waterbury et al., 1999). *Doublesex* is also involved in differential regulation of a gene in the fat body sex-specific transcription of the two yolk protein genes (An and Wesink, 1995).

1.7.9 *Fruitless (fru)*

The *fruitless* (*fru*) gene is one of the regulatory factors functioning downstream of the sex-determination hierarchy, and plays a prominent role in male sexual behaviour (Lee and Hall, 2000). Alternative splicing of *fru* transcripts produces sex-specific proteins belonging to the BTB-ZF (Bric- a-brac, Tramtrack, Broad-complex, and Zinc finger) family of transcriptional regulators. One *fru* cDNA codes for a protein of 855 amino acids, and another codes for a male-specific form of *fru* transcript class I that contains a long open reading frame that starts upstream of the male-specific 5 splice site and encodes a polypeptide of 776 amino acids (Ito et al., 1996). In females, a start codon downstream of the female-specific splice site is used and results in a polypeptide of 675 amino acids missing the first 101 amino acids (Ryner et al., 1996). The 101 amino acids specific to the male form contain a stretch of 12 histidines, alternating with neutral residues, followed by a proline rich stretch (Ryner et al., 1996).

Complex behaviour in flies such as singing, wing vibration, licking, and tapping have all been proven to be associated with this gene (Lee and Hall, 2000). It was shown that *fru* also dictates frequency of head-to-head interactions (a putative form of aggression) between males, which are to an extent, operationally separable from courtship behaviour. Certain loss-of-function alleles of the *fru* gene disrupt both male courtship behaviour and sexual orientation: courtship ritual performance is below par and it is directed indiscriminately at either sex (Anand et al., 2001; Ito et al., 1996; Lee and Hall, 2001; Ryner et al., 1996). Highly expressed *fru* alleles block courtship behaviour completely, but weaker *fru* alleles disrupt individual steps variously, with each step affected in some allelic combination (Anand et al., 2001; Lee and Hall, 2001). Therefore, *fru* is required for every step in the courtship ritual, and not for a single critical step (Demir and Dickson, 2005). The *fru*

gene also controls the neuronally determined feature of sex-specific anatomy, such as the muscle in the males abdomen in *Drosophila* called the Muscle of Lawrence (Taylor and Knittel, 1995). Muscle of Lawrence is a feature of male *Drosophila* development and is dictated by *fru*. Sex specific transcripts of *fru* are also found in groups of primary sensory interneurons in the antennal lobe involved in the processing of chemosensory information. These appear to be relay interneurons, known to project to higher brain centers such as the calyx of the mushroom body (Ryner et al., 1996). Males and females carry different versions of the FRU protein. In the absence of Transformer (TRA) proteins, *fru* mRNA is spliced and results in a male fate. Apart from FRU effects on the fly's abdomen, *fru* mutants show aberrant mating behaviour. Interestingly, when female *fru* is engineered to be expressed in males, the male mutants are sterile and are more interested in courting males instead of females (Demir and Dickson, 2005). Some mutant males show very little wing extension and generate no song pulse signals during wing displays. This defect is specific to courtship, as these mutants are normal in regarding flight and are able to flick their wings when rejecting advances made by another male. When *fru* mutant males are grouped together, they form male-male courtship chains in which each male is simultaneously both courting and being courted. All mutant combinations show some male-male chaining (Ryner et al., 1996; Ito et al., 1996). Male-specific splicing (*fruM*) induced in females produced the same male behaviour: courtship towards other females were observed in these mutants (Demir and Dickson, 2005).

1.8 Homologies and Evolution

Morphologically distinct sex chromosomes are believed to be derived from an initially identical chromosome pair, with morphological differentiation a by-product of the gradual loss of gene functions on the chromosome that is present only in the heterogametic sex (Charlesworth, 1996). However, once a chromosome becomes sex-limited, restriction of recombination with its homolog is likely to ensue. Restricted recombination in turn favours the accumulation of mutations on the sex-limited chromosome, although the precise evolutionary forces are still under discussion and may differ between species (Charlesworth and Charlesworth, 1997; Orr and Kim, 1998; Yi and Charlesworth, 2000). The result, after millions of years, is the loss of essentially all genes on the sex limited chromosome (Marin et al., 2000).

No sex-specific regulation of *Sxl* homologs has been observed in other dipterans (Saccone et al., 2002). A homolog of *Sxl* was found in *Chrysomya rufifacies* (Müller-Holtkamp, 1995), *Megaselia scalaris* (Sievert et al., 1997), and *Ceratitis capitata* (Saccone et al., 1998), and even in *An. gambiae* (Pannuti et al., 2000) but it did not qualify as a sex-determining gene in any of these Dipteran insects; SXL proteins of the same size were found in both sexes, but the expression pattern was inconsistent with a discriminatory role in sex determination. Despite a high degree of similarity, a transgene carrying a Musca-Sxl cDNA had no sex-transforming effect in *Drosophila* although the protein was properly expressed and localised to the nuclei (Meise et al., 1998). Therefore, although the SXL protein structure is well conserved (up to 85% of all amino acids), its function seems to have changed. TRA amino acid sequences were expected to be highly conserved amongst *Drosophila* species throughout evolution due to its importance in sex determination. However this was shown not to be the case. A study demonstrated that the TRA protein

has a high degree of divergence when the coding regions were compared among five *Drosophila* species (O'Neil and Belote, 1992). Therefore, the likelihood of *tra* being conserved across other Dipteran species is low. The high level of divergence in TRA protein sequences may be a result of neutral forces where amino acid substitutions accumulate by mutation and random genetic drift (Kimura, 1983). Another possibility is that the adaptive fixation of beneficial protein alleles over time may have contributed to the accumulation of amino acid differences among species (McDonald and Kreitman, 1991; Eanes et al., 1993). Comparing nucleotide sequence variation in eleven *tra* alleles within *D. melanogaster*, the latter view was supported as it was found that the lack of *tra* gene conservation is consistent with a recent selective sweep of a beneficial allele, that was alcohol dehydrogenase *(Adh)*, near the *tra* locus (Walthour and Schaeffer, 1994). The ortholog for *Adh* has been identified in *An. gambiae* (Bolshakov et al., 2002). However, as there are no reports of *tra* ortholog identified in mosquitoes, it is not possible to say whether a similar effect has occurred in mosquitoes.

Functional homologs of *dsx* are found amongst distantly related insects, which is highly conserved across species (Schütt and Nöthiger, 2000). Studies on the sex determination cascade showed that it has evolved from bottom to up (Wilkins, 1995), meaning that the genes at the bottom of the cascade are more ancient and highly conserved than those at the upper tiers. Recruitment of regulators at higher levels may evolve rapidly, but lower genes have pleiotropic effects, so it is difficult to replace them (Marin and Baker, 1998). *Doublesex* was found to be conserved in *An. gambiae* when compared with other insects such as *Drosophila melanogaster, Bombyx mori*, and *M. domestica* (Scali et al., 2005). It has also been shown that *dsx* homologs in *Megaselia scalaris* (Kuhn et al., 2000), *Ceratitis capitata* (Saccone et al., 1998), and

even *Caenorhabditis elegans* (Yi and Zarkower, 1999) are structurally and functionally conserved. Despite poor sequence conservation of *tra*, TRA in various species are functionally similar, as demonstrated by Pane et al. (2002). The *C. capitata tra* homolog was proven essential for female development (Pane et al., 2002). The regulatory genes at higher levels in the hierarchy can differ amongst various species (Graham et al., 2002). Although TRA has maintained its functional role, the gene upstream, *Sxl*, does not control *tra* in *C. capitata*. In fact, *tra* appears to have taken over the role of master switch gene that initiates the sex determination (Graham et al., 2002).

The *fruitless* gene was also thought to be highly conserved. The male-specific ortholog of *fru* in *An. gambiae* has been isolated and is shown to be conserved in its genomic sequence and function (Gailey et al., 2006). Moreover, when *fru* -mutant males develop, an abnormal extra fifth abdominal muscle segment that resembles the Muscle of Lawrence in *Drosophila* is apparent.

1.8.1 Sex determination pathway in *Aedes aegypti*

At the start of this study, almost nothing was known at a molecular level about sex determination in *Ae. aegypti*. During the study, the gene *dsx* had been identified (Mauro et al., 2006).

1.9 Meiotic Drive

Meiotic drive is a phenomenon that appears in nature where some alleles in the genome are over-represented in the gametes formed during meiosis, by disabling or destroying its alternate form, hence distorting the normal Mendelian segregation at the genetic locus (Lyttle, 1993; Braig and Yan, 2001). Examples of meiotic drive in male *Drosophila* can be categorized into two types: those caused by specific interactions between defined genetic elements, and those resulting from alterations in chromosome structure. The first class includes *Segregation Distorter (SD)/Responder (RSP)* interactions (Lyttle, 1993) and *Stellate/crystal* interactions (Palumbo et al., 1994). Males carrying SD and sensitive alleles of the Rsp locus are subject to high rates of dysfunction in their sperm. *Segregation Distorter* encodes a truncated form of RanGAP, which is a nuclear import regulator (Merrill et al., 1999), and *Rsp* corresponds to a middle-repetitive element (Wu et al., 1988; Pimpinelli and Dimitri, 1989), but the precise interaction between these two components and how this results in drive is not yet understood. Evidence suggests that there is a connection between XY pairing and sperm dysfunction in chromosomally-induced cases of meiotic drive. This link was first put forward by observations of meiotic drive in males bearing $In(1)sc^{4L}sc^{8R}$, an X chromosome that is deficient for the male meiotic pairing sites (Tomkiel, 2000). The frequency of XY nondisjunction in these males varies in different genetic backgrounds and at different temperatures, and the level of meiotic drive varies correspondingly (Peacock et al., 1975).

1.9.1 Meiotic drive in *Ae. aegypti*

Endogenous meiotic drive systems are described in *Ae. aegypti* (Hickey and Craig, 1966; Wood and Ouda, 1987) and *Cu. quinquefasciatus* (Sweeney and Barr, 1978). Both species have a sex-linked gene or genes linked to the male-determining allele at the sex determination locus on chromosome 1 that drives the normally equal sex ratio to be biased toward males. The driver gene, distorter (D) of *Ae. aegypti* is tightly linked to the male determining allele (M), and the responder locus is tightly linked to the female determining allele (m) on the homologous chromosome (Cha et al., 2006). Neither the driver gene linked to the female determining allele nor the responder linked to the male determining allele function properly to produce the drive phenotype (Newton et al., 1974). Males that are heterozygous for the male distorter driver (MD) gene cause fragmentation in gametes bearing the sensitive responder (ms) during spermatogenesis. Therefore, the observed phenotype of the meiotic drive system in *Ae. aegypti* is a highly male-biased sex ratio (Cha et al., 2006). A proportion of enlarged and misshapen MD spermatozoa was observed under light microscopy (Wood and Newton, 1977; Newton et al., 1978). Abnormally high levels of DNA were found in some MD spermatozoa (Newton et al., 1978), due to an inhibition of reduction division in spermatogenesis, leading to the formation of a restitution nucleus (Wood and Newton, 1991).

Endogenous meiotic drive is of interest because of its potential use as an alternative population control mechanism. This is because meiotic drive could cause a population crash when more males are produced than females. Moreover, it can be used in conjunction with a modified gene to rapidly spread them into natural populations, thus replacing it. However, past investigations of meiotic drivers to suppress *Ae. aegypti* populations had limited success. Hickey and Craig (1966) observed that the distorted sex ratio was quickly

reverted in cage populations where driving males were introduced to sensitive females, but driver alleles were still present and active in the population. This observation was later explained when the tolerance gene was identified near the red-eye (re) locus on chromosome 1 (Wood and Ouda, 1987). The tolerance gene locus is a suppressor that reduces the sensitivity of the responder, which accumulates in the population due to strong selective pressure caused by the driver. This gene has been designated as *t (tolerance of Distorter)* and results in reduction of sex ratio distortion. Investigations on the influence of an endogenous meiotic drive gene on *Ae. aegypti* population dynamics lead to the establishment of populations initiated with male mosquitoes carrying a strong specific meiotic driver (T37 strain) and drive-sensitive females (RED strain) that were highly male biased in early generations, with more than 60% male by the F15 (Cha et al., 2006). A genetic marker tightly linked with the meiotic driver on chromosome 1 showed strong selection for the T37 strain-specific allele.

A drawback of this mechanism is that not all strains of the species are sensitive to meiotic drive, making this control mechanism useless in some populations. Wood and Newton (Wood and Newton, 1991) speculated from their findings that meiotic driver Distorter (D):

- Is present in Australia, Sri Lanka, and some parts of Africa and America.
- Is absent in Asia and the Pacific Region.
- In some parts of Africa different "strengths" of distortion are evident.

Natural X chromosomes vary greatly in sensitivity to D, and most populations are polymorphic for this sensitivity (Suguna et al., 1977). Exceptions are western and central African strains, which are tested to be uniformly resistant

and individual strains from Surinam, Kenya, Taiwan, and Hawaii that are uniformly sensitive.

1.10 Aims

Sex determination in mosquitoes is not well understood. The identification of potential gene(s) involved in the sex determination cascade of the mosquito *Ae. aegypti* could help in understanding the mechanism of sexual fate in the mosquito. This could then be used potentially in vector population control by switching the sexual fate to enable an imbalance of gender ratio, and hence cause a population crash if linked to a drive mechanism, or enable male-only releases.

1.10.1 Identifying male specific regions and M

The mechanism which determines maleness of the mosquito is unknown. Therefore, the first major aim of this study is to develop novel approaches to identify the male determining gene that could be used in sex ratio distortion. Identifying male determining gene(s) will also contribute to the understanding of the mechanism of sex determination in mosquitoes.

1.10.2 Identification and characterisation of sex-determining genes in *Ae. aegypti*

Identification and characterisation of sex-determining genes are important to understand the workings of the mechanism. The second major aim is to identify and characterize the *transformer-2* homolog because of its potential use in sex ratio distortion by switching the female gender to male.

1.10.3 Inducing sex ratio distortion

The third major aim is to develop a reliable approach to induce sex ratio distortion by knocking down a female determining gene or a gene essential to female development. In this project, *tra-2* in *Ae. aegypti* is the main candidate for knockdown to produce distortion in the sex ratio. This in turn would also answer the question of whether the putative *tra-2* homolog of *Ae. aegypti* is functional or not in the sex determining cascade.

Chapter 2

General Materials and Methods

Some of the materials and methods common to many of the experiments performed are described below. All basic protocols were as follows unless stated otherwise.

2.1 Mosquito Rearing

2.1.1 Mosquito strains

Several strains of *Aedes aegypti* were used in these experiments. The colonies used are listed below:

Table 2.1: List of *Ae. aegypti* strains used in study with strain and colony origin.

Strain	Origin	Details
Rockefeller	USA	Laboratory strain originated from the Carribean in 1930s.
LA513A	United Kingdom	Strain is genetically modified from the Rockefeller strain, as developed by Oxford University (Phuc et al., 2007), also known as OX513A.

2.1.2 Mosquito rearing conditions

All strains of adult *Ae. aegypti* used in experiments were reared in an insectary with environmental temperature of 26-28°C, relative humidity of 60-75%, and a 12-hour day/ 10-hour dark cycle, with a one-hour crepuscular period at the beginning and end of each cycle. The adults were kept in a 30×30×30 cm plastic cage and were fed with water-soaked cotton wool and whole sugar cubes. All larvae were reared at low density (200-500 larvae/ 3 L) in water-filled trays. *Aedes aegypti* larvae were given a diet of pond fish food pellets (Omega Sea, Ltd., USA). *Aedes aegypti* strain LA513A larvae were reared in

the same condition but with the addition of tetracycline to the larvae water at a concentration of 0.03 g/l as required for their survival.

To obtain eggs, female adults were fed using artificial feeders (Hemotek Ltd., UK) with warmed defibrinated sheep's blood (TCS Biosciences, UK) and stretched Parafilm membranes for 2-4 hours during insectary day light to crepuscular cycle. An "oviposition-bowl" (a plastic bowl with wet cotton wool and lined with wet filter paper) was then left in the cage for the mosquito to lay their eggs. Eggs deposited on the wet filter paper were slowly air-dried for three days in the insectary environment. Eggs collected were either submerged again in water to hatch or kept in a sealed polythene bag at insectary conditions which can be hatched later up to 3 months (eggs that are kept longer than this period may result in lower hatch rate, and are therefore avoided if possible).

2.1.3 Problems with mosquito rearing

Although *Ae. aegypti* are relatively hardy and thus easily reared, several strains were particularly fastidious and significant problems arose. In addition, faults experienced with insectary humidifier several times throughout the year caused inconsistency in humidity and in some cases lead to colony loss. When colony experienced difficulty in blood feeding, purinergic phagostimulant ATP is added to the blood to the concentration of 5 mM to improve feeding (Hosoi, 1959; Galun et al., 1963).

2.2 Nucleic acid extraction

2.2.1 DNA extraction

DNA from single or multiple mosquito samples was extracted using a modified protocol of Livak buffer method (Collins et al., 1987) with sodium acetate-ethanol precipitation as described below.

- Mosquito samples were homogenised with a sterile pestle in 100 μl of Livak grind buffer pre-heated 65°C in 1.5 ml microcentrifuge tubes.
- Samples were incubated at 65°C for 30 minutes.
- 8 M potassium acetate was added to samples to obtain a 1.0 M solution. Samples were mixed gently by inverting the tubes,
- Samples were incubated on ice for 30 minutes, followed by centrifugation at 20,000 $\times$ g for 20 minutes at room temperature.
- The supernatant was collected and 200 μl of ice cold 100% ethanol was added to each sample, mixed and incubated at 20°C for 15 minutes.
- The samples were centrifuged at 4°C, at 20,000 $\times$ g for 25 minutes to pellet the DNA. The supernatant was removed after centrifugation.
- 250 μl of ice-cold 70% ethanol was added to the remaining DNA pellet and centrifuged at 20,000 $\times$ g for 10 min. to rinse the pellet, and then air-dried for 1 hour at room temperature or 15 min. at 37°C. The DNA pellet was resuspended in variable amounts of sterile water, depending on concentration needed for experiments.
- DNA samples were kept at -20°C or -80°C for long term storage.

2.2.2 RNA extraction

The protocol used for all RNA extractions was as described in Sigma's TRI reagent RNA extraction protocol, with some minor modifications (Paton et al., 2000).

- For mass extractions (i.e 5 or more mosquitoes at a time), 700 μl of Trizol (Invitrogen) was used to homogenise the mosquitoes in a sterile microcentrifuge tube with a sterile pestle.
- 350 μl of chloroform (Sigma) was added and mixed well by vortexing for about 20 seconds. The mixture was then centrifuged at 12000 $\times$ g, 4°C, for 20 minutes.
- The upper aqueous phase was then transferred to a new sterile microcentrifuge tube. 350 μl of cold isopropanol (Sigma) was added and mixed by vortex to precipitate the RNA. The tube was then centrifuged at 12000 $\times$ g at 4°C for 15 minutes.
- The supernatant was discarded and 250 μl of 75% ethanol was added to the pelleted RNA in the tube. The sample was then centrifuged at 8000 $\times$ g at 4°C for 5 minutes.
- The supernatant was removed and the pellet remaining was air-dried for about 10 minutes for the ethanol to evaporate.
- The RNA was resuspended in variable amounts of sterile Dimethyldicarbonate (DMDC) or Diethylpyrocarbonate (DEPC) treated water, depending on the need for experiments.
- Samples were stored at -80°C.

For smaller amounts of RNA extraction (1-4 mosquitoes), methods were the same except that the volume of reagents used were scaled down: 250 μl of Trizol, 50 μl of chloroform, and 125 μl of isopropanol.

2.2.3 RNA treatment to remove DNA

RNA samples were treated with DNase I (Invitrogen) to remove unwanted single- and double-stranded DNA that may remain after extraction:

- Every 80 μl of RNA was treated with 10 μl of 10× DNase buffer, 4 μl of DNase and 0.2 μl of RNase inhibitor.
- Samples were incubated for 15 to 20 minutes (depending on concentration) at 37°C.
- 4 μl of DNase stop solution was added to the mix and incubated at 70°C for 10 minutes to de-activate the DNase.

2.2.4 Phenol-chloroform extraction

Removal of protein and/or other impurities from nucleic acid samples (e.g. removal of digestion enzymes) was done by phenol-chloroform extraction method (Sambrook et al., 1989). The protocol for this method is briefly described as follows:

- The nucleic acid sample was transferred to a centrifuge tube and an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1), pH 7.8 - 8.0, was added.
- The contents of the tube was mixed until an emulsion forms.

- The mixture was centrifuged at 15000 $\times$ g for 1 to 2 minutes at room temperature.

- Using a pipette the top aqueous phase was transferred to a fresh tube. For small volumes (<200 μl), an automatic pipettor fitted with a disposable tip was used. The interface and lower organic phase was discarded.

- The phenol extraction was repeated until no protein was visible at the interface of the organic and aqueous phases.

- An equal volume of chloroform: isoamyl alcohol (24:1) was added and mixed.

- The mixture was centrifuged again for 1 to 2 minutes.

- The top phase was then transferred into a fresh tube.

- The nucleic acid was recovered by standard precipitation with 2.5 volumes of ethanol and 0.1 volume of sodium acetate (final concentration 0.3 M, pH 5.2).

- The solution was mixed well and incubated on ice for 15 to 30 minutes to allow the DNA to precipitate. When the size of the DNA was small (<100 nucleotides) or when it was present in small amounts (<0.1 μg/ml), the period of incubation was extended to at least 1 hour.

- The DNA was recovered by centrifugation at 4°C. For most purposes, centrifugation at maximum speed for 10 minutes in a microcentrifuge was sufficient. However, when low concentrations of DNA ($\leq$20 ng/ml) or very small fragments ($\leq$100 nucleotides) were being processed, more extensive centrifugation were required. Centrifugation at 100,000 $\times$ g

for 20 - 30 minutes allows the recovery of picogram quantities of nucleic acid in the absence of carrier.

- The supernatant was carefully removed with an automatic micropipettor taking care not to disturb the pellet of nucleic acid (which may be invisible).
- The supernatant was saved from valuable DNA samples until recovery of the precipitated DNA was verified.
- The tube was filled halfway with 70% ethanol and re-centrifuged at maximum speed for 2 minutes at 4°C in a microcentrifuge.
- The supernatant was removed without disturbing the pellet.
- The open tube was left on the bench at room temperature until the last traces of fluid had evaporated.
- The DNA pellet (which is often invisible) was then dissolved in the desired volume of buffer (usually TE [pH between 7.6 and 8.0]) or sterile water, rinsing the walls of the tube well with the buffer or water to ensure all DNA was dissolved.

2.2.5 Gel Extraction

Bands of interest in an agarose gel were excised with a clean scalpel and purified using QIAquick Gel Extraction Kit (Qiagen) according to the manufacturers protocol.

The sample of the extracted DNA was then tested for quality and quantity using a Nanodrop ND 100 Spectrophotometer.

2.2.6 Nucleic acid quality confirmation

DNA and RNA quality was assessed using the Nanodrop ND 100 spectrophotometer. The nucleic acid section of the software was selected. Depending on the samples to be tested, 1 μl of water or buffer that was used to make the DNA and RNA solution was loaded onto the cleaned column to set the 'blank'. The column was then cleaned and 1 μl of the sample was loaded for measurement. The sample type was specified in the sample type box according to the type of nucleic acid that was being measured. The nucleic acid was measured for the concentration and 260:280 ratio was used as a measure of quality of the nucleic acid. For DNA, samples showing less than 1.70 in 260/280 nm wavelength ratio were discarded to ensure the quality of the experiments. For RNA, samples below 1.80 in 260/280 nm wavelength ratio were discarded.

2.2.7 mRNA extraction

To extract mRNA from RNA, PolyATract mRNA Isolation System kit (Large-Scale mRNA Isolation: 1 to 5mg of total RNA) from Promega was used according to the manufacturers manual.

The mRNA was then tested on Nanodrop ND 100 Spectrophotometer to check the quality and quantity of the mRNA.

2.3 Polymerase Chain Reaction (PCR)

2.3.1 Standard PCR

Each DNA sample was amplified by PCR using Sigma 10× PCR Buffer, 2.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.4 mM forward and reverse primers, 0.75 units of Taq polymerase (Sigma-Aldrich, UK), and double distilled water to make up to 25 μl (DNA sample amount variable).

The basic amplification cycle condition used was as follows:

- Denature at 94°C for 5 minutes.
- 25 to 35 cycles: Denature at 94°C for 30 seconds, anneal at X°C for 30 seconds, and extension at 72°C for 45 seconds.
- Final extension at 72°C for 5 minutes.

Annealing temperature is denoted by X°C, as it is variable according to the primers used.

If amplification of DNA was problematic, the basic touch down cycle conditions was used as follows:

- Denature at 94°C for 5 minutes.
- 3 cycles: Denature at 94°C for 30 seconds, anneal at 60°C for 30 seconds, and extension 68°C for 45 seconds.
- 3 cycles: Denature at 94°C for 30 seconds, anneal at 57°C for 30 seconds, and extension at 68°C for 45 seconds.
- 35 cycles: Denature at 94°C for 30 seconds, anneal at 54°C for 30 seconds, and extension at 68°C for 45 minutes.

- Final extension at 68°C for 5 minutes.

2.3.2 Reverse transcriptase (RT) PCR

2.3.2.1 One step RT-PCR

One Step RT-PCR kit (Qiagen, UK) was used for analysing gene expression. This is a quick way to analyse gene expression using abundant RNA samples. All reactions were carried out according to manufacturers instructions. Cycle conditions below were used as recommended by the manufacturer:

- 50°C for 30 minutes
- 95°C for 15 minutes
- 94°C for 5 minutes
- 35 cycles: 94°C for 1 minute, X°C for 1 minute, and 72°C for 1 minute
- 72°C for 10 minutes

Annealing temperature is denoted by X°C, as it is variable according to the primers used. The annealing temperature was previously optimized for each primer pair on standard genomic DNA by standard PCR.

2.3.2.2 Two step RT-PCR

Two step RT-PCR requires generating first-strand cDNA separately. The cDNA would then be used to run a basic PCR amplification or a quantitative PCR (qPCR). This application is useful when multiple PCRs of the same limited RNA sample is carried out. SuperScript VILO cDNA Synthesis Kit

(Invitrogen) was used to make the cDNA synthesis. The protocol used were as follows:

For a single reaction, the following components were combined in a tube on ice. For multiple reactions, a master mix without RNA was made.

- 5X VILO reaction mix 4 μl
- 10X SuperScript enzyme mix 2 μl
- RNA (up to 2.5 μg) x μl
- DEPC-treated water to 20 μl

The tube contents were gently mixed and incubated in a thermal cycler with the following conditions:

- 25°C for 10 minutes.
- 42°C for 60 minutes.
- 85°C at 5 minutes to terminate the reaction.

The cDNA can then be used in a Quantitative PCR (qPCR) or regular PCR with the appropriate concentration, or stored at -20°C.

2.3.3 Gel electorphoresis of PCR products

Five to 25 μl of each PCR product and a DNA marker ladder (Promega/Bioline/Eurogentec) were analysed by gel electrophoresis using 1% to 2% w/v agarose gels in TAE buffer (see Appendix for recipe). Gels were stained with ethidium bromide and visualised under ultraviolet illumination. Images of gels were obtained using a Kodak DC40/DC120 camera and Kodak Image Analysis software.

2.3.4 Quantitative PCR (qPCR)

Quantitative PCR was carried out using SYBR Green qPCR Detection kit (Invitrogen, UK). SYBR Green is a light sensitive dye that binds to a single amplification product. The amount of target nucleic acid was determined by using external standards containing sequences that were identical to the target sequence with identical primer binding sites. Template cDNA was obtained from DNase treated RNA samples, using SuperScript VILO cDNA Synthesis Kit (Invitrogen) for transcription and then diluted accordingly for qPCR. The samples were amplified in a total volume of 10 μl per reaction, with the following set up:

- 5 μl SYBR Green mix
- 2 μl DEPC-treated water
- 0.5 μl Forward primer (10 μM)
- 0.5 μl Reverse primer (10 μM)
- 2 μl of cDNA

The reaction mix was loaded into qPCR tubes and incubated in an Opticon 2 continuous Fluorescence Detection System (GRI) with the following conditions:

- 95°C for 15 minutes.
- 45 cycles: 95°C for 10 sec., X°C* for 10 sec, 72°C for 20, and plate reading taken.
- melting curve from 50°C - 95°C, plate reading taken every 1°C with 1 second hold.

*X indicate the optimum annealing temperature for the primer sets used. Primers were tested using standard PCR to ascertain optimum temperature and amplified PCR product is specific.

2.4 Molecular cloning

2.4.1 Ligation reaction

DNA fragments of interest were inserted into pGEM-T Easy Vector (figure 2.1) plasmids (Promega) by first setting up a reaction for each sample:

- 5 μl 2X Rapid ligation buffer, T4 DNA ligase
- 1 μl (50 ng) pGEM-T easy
- X ng PCR product (volume depending on concentration)*
- 1 μl T4 ligase (3 Weiss unit/μl)
- Bring total volume of reaction to 10 μl with sterile water

* PCR product required for insertion into vector depending on size and optimum transformation ratio. A formula provided by Promega was used for calculating amount of PCR product:

$$\frac{\text{ng of vector} \times \text{ kb size of insert}}{\text{kb size of vector}} \times \text{insert:vector molar ratio} = \text{ng of insert}$$

The reaction was then incubated overnight at 4°C for obtaining maximum number of vectors inserted. The ligated vectors were then transformed into competent cells or stored in -20°C until needed.

2.4.2 Cell transformation

After ligation, DH10 competent cells (Invitrogen) were transformed using the following protocol:

- The ligation reactions were centrifuged to collect the contents at the bottom. 2 μl of each ligation was added into a sterile 1.5 ml microcentrifuge (17 $\times$ 100 mm) tube on ice.
- Competent cells were allowed to thaw in an ice bath. Cells were mixed by gently flicking the tube.
- 50 μl of cells were transferred into the 1.5 ml microcentrifuge tube containing the ligation reactions prepared earlier.
- The tubes were mixed by gentle flicking.
- The cells were heat-shocked for 30 seconds in a water bath at 42°C.
- The tubes were then immediately returned to ice for 2 minutes.
- 950 μl of room-temperature Super Optimal Broth with catabolite repression (SOC medium) were added to the tubes containing the transformed cells to encourage high tranformation efficiency of plasmids.
- The tubes were incubated at 37°C while shaking (about 150 rpm) for 1.5 hours.
- The cells were then spread onto LB ampicillin agar plates impregnated with X-gal and IPTG.
- The plates were incubated overnight at 37°C.

- White colonies grown on the plates were transferred into 200 μl LB media with 0.01% w/v ampicillin and cultured overnight at 37°C without shaking.

2.4.3 Screening of clones

Many of the clones produced were screened by direct PCR using primers designed to amplify the M13 sequence region of the plasmid vector (figure 2.1). 1 μl of each culture was used for standard PCR (with the annealing temperature of 50°C). M13 primers used were as follows:

- M13 Forward Primer: 5'-GTAAAACGACGGCCAGT-3'
- M13 Reverse Primer: 5'-CAGGAAACAGCTATGAC-3'

Cycle conditions were:

- 94°C for 4 minutes
- then repeat the following for 35 cycles: 94°C for 30 seconds, 50°C for 30 seconds, 72°C for 1 minute 30 seconds
- 72°C for 10 minutes

The PCR products were then visualised on an agarose gel and the right sized band was ascertained. Because of amplifying using M13 primers, the PCR products will be about 200 bp larger than the expected size from the insertional sequence.

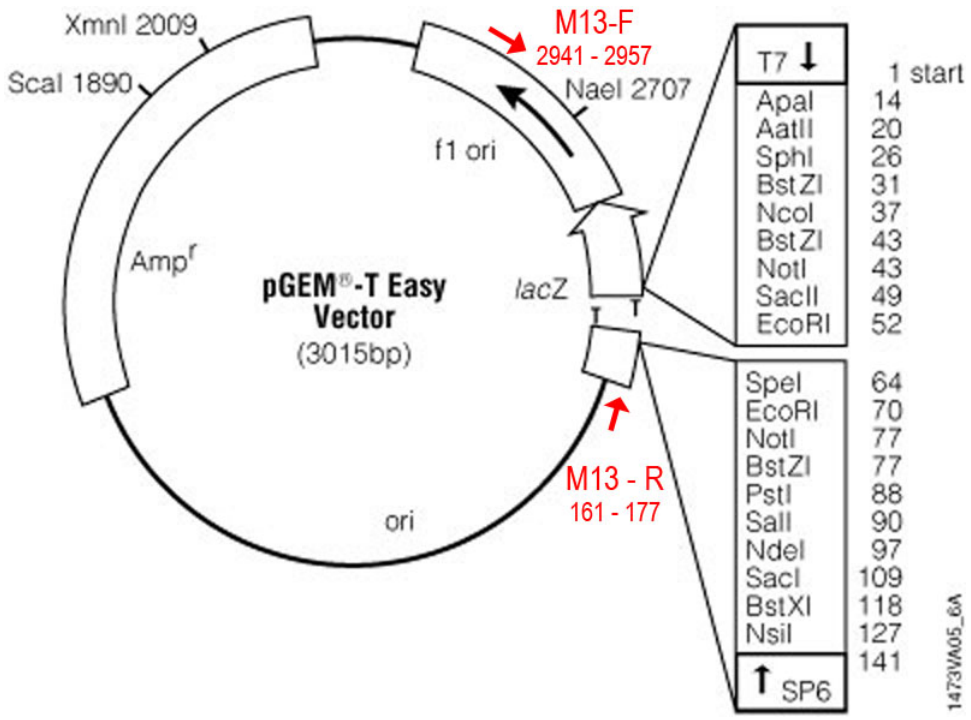

Figure 2.1: Map of pGEM-T Easy vector. Red arrows show M13 forward and reverse primer sequencing binding sites (M13-F and M13-R respectively). Red numbers indicate exact position of sequence (in bp) in the plasmid. DNA insert would be between the T overhangs at the lacZ region indicated. Diagram from Promega PGEM-T Easy Systems technical manual.

2.5 DNA sequencing

DNA to be sequenced was first amplified and purified to ensure good quality sequence. The DNA samples were then prepared for sequencing. The sample preparation was based on Perkin Elmer ABI Prism Big Dye Terminator Cycle Sequencing Ready Reaction Kit protocol:

- 1.0 μl BigDye Terminator reaction mix
- 2.0 μl $\times$5 Sequencing Buffer
- 1.0 μl Primer (3.2 ρmol)
- 100 ng PCR product to be sequenced

Bring up reaction volume to 10 μl with sterile water

Cycle conditions were:

- 96°C for 1 minute
- Then repeat the following for 24 cycles: 96°C for 10 seconds, 50°C for 5 seconds*, 60°C for 4 minutes

*Note that the annealing temperature used here is for M13 sequence. When other primers were used, the annealing temperatures were adjusted accordingly.

The samples were sent to Zoology's Sequencing Laboratory for sequencing.

Alternatively, purified DNAs were sent to GATC Biotech (www.gatcbiotech.com) for sequencing.

2.6 Sequence read and consensus alignment

Where possible, the DNAs for each sample were sequenced at least two times in both directions to resolve any ambiguity that might occur. The sequenced DNA trace data returned were aligned and analysed using Geneous software to achieve a consensus sequence (figure 2.2).

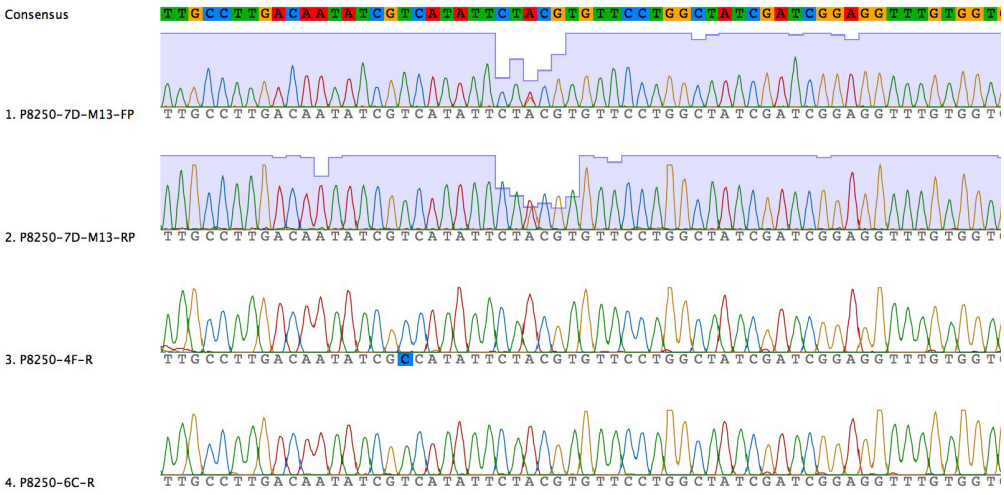

Figure 2.2: Achieving a consensus sequence using Geneious software. The four sequenced DNA trace data was aligned to achieve a consensus sequence. Nucleotides that were different from the majority of the sequence will be highlighted in colour. Here, there is a C base highlighted in blue that differs from the rest of the sequence trace, where it is a T base. Hence, the consensus sequence for this position is T.

Chapter 3

Towards the identification of male-specific chromosomal regions in *Aedes aegypti*

3.1 Introduction

A possible route to inducing male biased sex ratios in mosquitoes is to manipulate the primary male determining gene M. One route to identifying M in *Ae. aegypti* is to first identify male-specific regions of its genome. In this study, genomic subtractive hybridisation was attempted to isolate male-specific genes. Laser microdissection and microarrays were also used to detect male-specific sequences.

3.1.1 Genomic subtractive hybridisation

Subtraction, or subtractive hybridisation, is a technique that enables the comparison of two populations cDNA or genomic DNA in order to identify sequences that are present in one population but not in the other. Genomic DNA was used in this study to enable the identification of unexpressed regions in the male and female genome.

3.1.2 Laser microdissection and microarray analysis

Comparative Genomic Hybridization (CGH) microarrays provide a powerful tool for the genome-wide detection of copy number variations for sequenced genomes. This technique could be used to examine the differences between male and female DNA by competitive hybridization. Oligonucleotide CGH arrays for the Agilent platform were designed and used for this section of the study. These CGH arrays use a "two-colour" process to measure DNA copy number changes in an experimental sample relative to a reference sample.

In *Ae. aegypti* males the male-determining M factor is located in a region of chromosome 1 that lacks the constitutive heterochromatin (C-band) (Newton et al., 1974). At the time of writing the *Ae. aegypti* genome size was estimated to be 1.38 Gb from Vectorbase and the Broad Institute *Ae. aegypti* database (USA). The total unique contig length generated from the Broad Institute and Vectorbase stands at 1,310090344 bp, with 36206 contigs (4758 Supercontigs). However, the genome sequence database of *Ae. aegypti* is still incomplete as about 53% of the genome consists of repetitive sequences (Nene et al., 2007), making the process of assembling supercontigs into continuous sequence for each chromosome rather difficult. 15988 genes and 17402 peptides were predicted from the Vectorbase and TIGR collaboration. Only one-third (about 5000 genes) of the predicted gene set were annotated to a chromosomal location using genetic and physical mapping data (Nene et al., 2007). Thus, the majority have not been assigned to a region and/or chromosome. A procedure for micro-dissecting human chromosomes, and micro-cloning of its DNA sequences has been described, in which universal amplification of the dissected fragments by *Mbo I* linker adaptors and PCR was used to isolate sequences from chromosome 21 (Kao and Yu, 1991). It should be possible to amplify the mosquito chromosomes after micro-dissection using the same *Mbo*

I linker adaptor technique. Using CGH arrays, amplified chromosome from the micro-dissection data and total gDNA were used to identify supercontigs that are located on chromosome 1 and for which there is also evidence that they may be male specific.

3.1.3 Aims

The overall aim was to identify male-specific genes as candidates for the M-factor by genomic subtraction and library screening. The subtractive hybridisation technique was adapted and genomic DNA was used from the mosquito instead of cDNA as the aim was to identify male-specific region of the genome, not genes expressed only in males. By hybridising the male and female total gDNA, genes present in both sexes would be subtracted out and what is left would be enriched in sex-specific sequences, including, in theory, the M factor gene(s).

A second aim was to compare total gDNA from males and females to see if there is any male specific hybridisation that could lead to identification of male specific genes. An additional aim of this part of the study was to help determine which unassigned annotations in the genome are located on chromosome one by microdissection of individual chromosomes using a laser and then analysing the different sets of DNA by microarray hybridization. Arrays were designed with oligonucleotides based on the *Ae. aegypti* predicted gene set - it was not possible to design total gDNA probes for the microarray due to the genome size and large number of sequence repeats in the *Ae. aegypti* genome.

3.2 Methodology

3.2.1 Subtractive hybridisation of male to female DNA

Figure 3.1 gives a visual overview of the events that occur in a PCR-select cDNA subtractive hybridisation.

Six each of female and male mosquito pupae from the *Ae. aegypti* Rockefeller strain were separately mass DNA extracted, and then purified by phenol-chloroform extraction. Ten μg of DNA from the samples were digested with *Rsa I* restriction enzyme overnight at 37°C. The digestion was terminated by heating at 65°C for 10 minutes and the DNA was purified by a second phenol-chloroform extraction, and re-suspended in 10 μl of sterile water.

Samples were then prepared for subtraction using PCR-select cDNA Subtraction Kit (Clontech) protocol as briefly described:

Adaptor-ligation preparation

In a microcentrifuge tube, ligation master mix (per reaction) was prepared as follows:

- Sterile water - 3 μl
- 5X Ligation Buffer - 2 μl
- T4 DNA ligase (100 units/μl) - 1 μl

Ligation reactions were set up for each reaction as shown in table 3.1:

*Sequence for Adaptor 1 and 2R primers are provided by Clontech. See description of sequences in the Appendix.

The DNA samples that were to be the Tester (i.e. male DNA) were ligated with adaptors 1 and 2R (figure 3.2). A total of 3 reactions of the subtractive

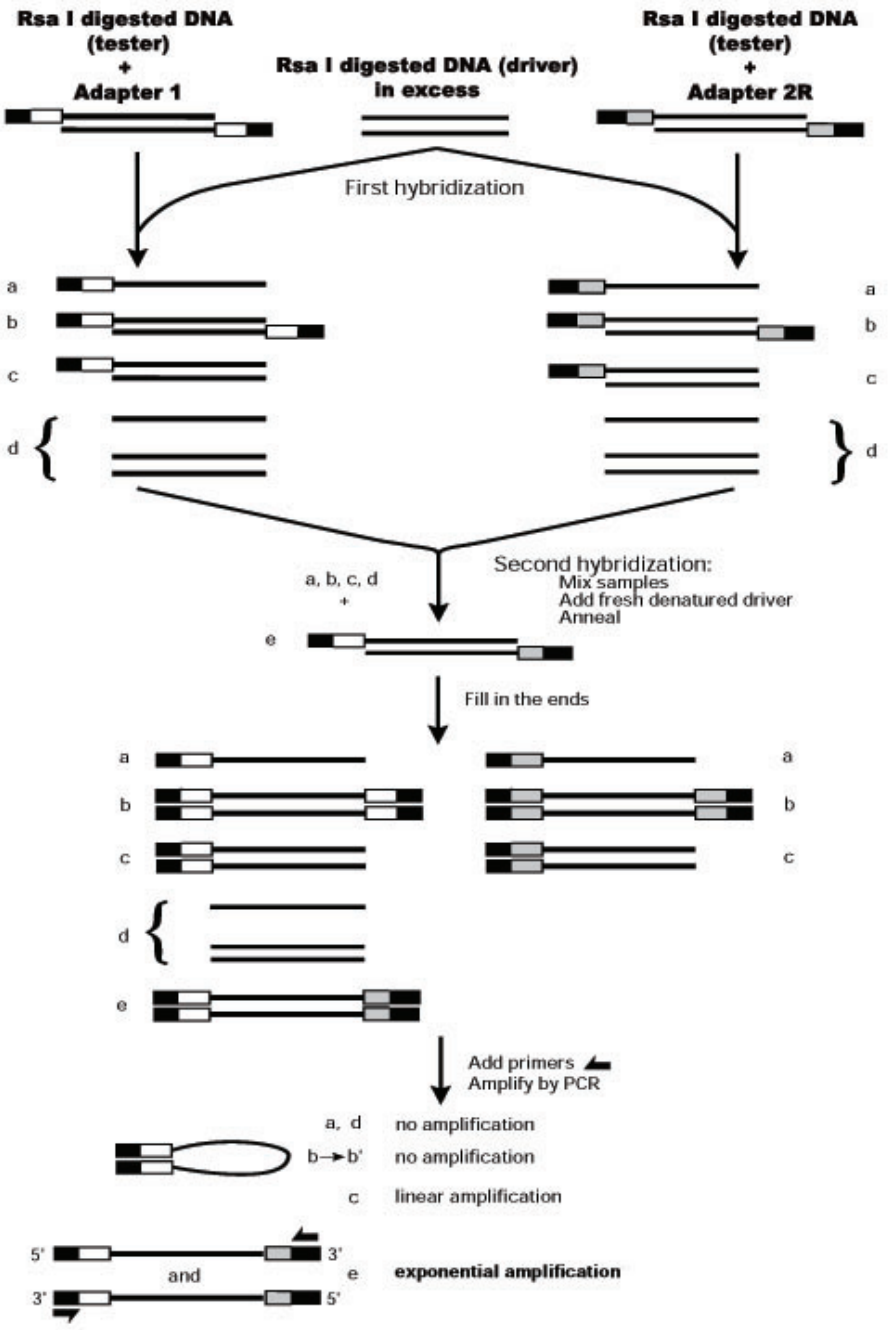

Figure 3.1: Overview of subtractive hybridisation using Clontechs PCR-Select procedure. Male total DNA is referred to as tester and female total DNA is referred to as driver. The first hybridisation generates the type a, b, c, and d molecules in each sample. Types a and b molecules are unique male sequences while type c molecules are not unique. Type d molecules are driver DNA in excess. After second hybridisation, type e molecules are formed by hybridising between the two tester DNA pools and are unique male sequences. After filling in ends by DNA polymerase, the type e molecules can be amplified exponentially by nested PCR using primers to amplify adaptors 1 and 2R. Refer to Appendix for detailed sequences of the primers and adaptors.

Table 3.1: Setting up the ligation reactions for subtraction.

Component	Tester 1	Tester 2
5X diluted tester DNA	2 μl	2 μl
Adaptor 1* (10 μM)	2 μl	-
Adaptor 2R* (10 μM)	-	2 μl
Master mix	6 μl	6 μl
Final volume	10 μl	10 μl

hybridisation were attempted with varying amounts of total *Rsa I* digested male DNA, which was then diluted 5X: a) 15.7 μg, b) 14.4 μg and, c) 5.0 μg. Amount in A and B are high concentration samples (B slightly lower due to lower amount of extracted DNA). Amount C is much less than amounts A and B to compare if carrying out subtraction at a lower concentration would give a more efficient subtraction. In a fresh microcentrifuge tube, 2 μl each of Tester 1 and Tester 2 were mixed, centrifuged briefly and incubated overnight at 16°C. This is the unsubtracted tester control. After incubation, 1 μl of EDTA/glycogen mix was added to stop the ligation reaction. The samples were then heated at 71°C for 5 min to inactivate the ligase and briefly spun in a centrifuge.

Analysis of ligation

Analysis of ligation was performed to verify that the DNAs have adaptors on both ends. This experiment was designed to amplify fragments that span the adaptor/DNA junctions of Tester 1 (figure 3.2). One μl of each ligated DNA was diluted into 200 μl of water and used to set up the PCR according to table 3.2:

*PCR Primer 1 amplifies the T7 promoter site within adaptors 1 and 2R. See description of sequence in the Appendix.

The master mix for the PCR were made as shown below and added to each of

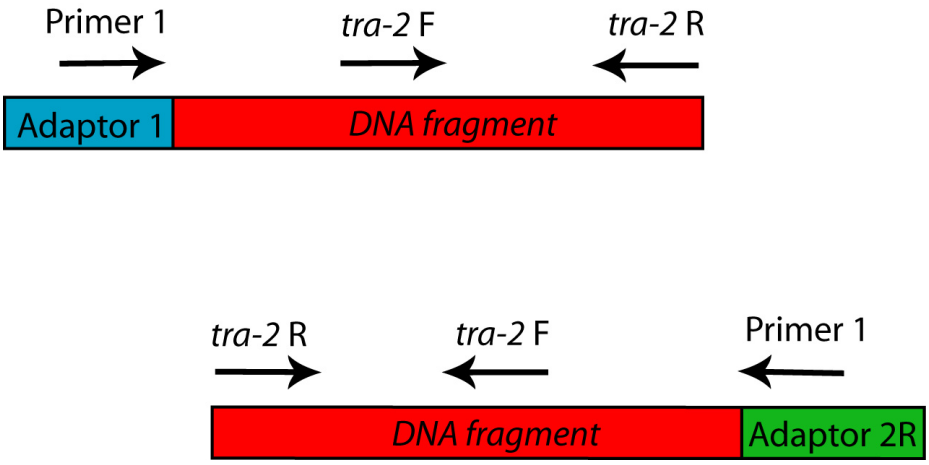

Figure 3.2: Diagram of how ligation analysis prior to the subtraction works. Primer 1 amplifies the adaptor/DNA junctions of Tester 1 (DNA fragment ligated to Adaptor 1) and Tester 2 (DNA fragment ligated to Adaptor 2R). *Tra-2* primers will amplify a fragment of 135 bp. When Primer 1 and *tra-2* R primer are used in a PCR, a larger than 135 bp tra-2 fragment will appear (size is depending on DNA fragment).

Table 3.2: Setting up the ligation analysis PCR

Component	tube 1	tube 2	tube 3	tube 4
Tester 1-1 (ligated to Adaptor 1)	1 μl	1 μl	-	-
Tester 1-2 (ligated to Adaptor 2R)	-	-	1 μl	1 μl
tra-2 forward primer (10 μM)	1 μl	1 μl	1 μl	1 μl
tra-2 reverse primer (10 μM)	-	1 μl	-	1 μl
PCR Primer 1* (10 μM)	1 μl	-	1 μl	-
Total volume	3	3	3	3

the ligation analysis tubes:

- Sterile water - 18.5 μl
- 10X PCR reaction buffer - 2.5 μl
- dNTP mix (10 mM) - 0.5 μl
- 50X Advantage cDNA Polymerase mix - 0.5 μl

The reaction mix were then incubated at 75° for 5 min in a thermal cycler

to extend the adaptors and the following cycling conditions were used for 20 cycles immediately:

- 94°C - 30 sec
- 65°C - 30 sec
- 68°C - 2.5 min

5 μl from each reaction were then analysed on a 2.0% agarose/EtBr gel to check the efficiency of the ligation.

First Hybridisation

If adaptor ligation was successful, first hybridisation was then performed.

For each subtraction, the reagents were combined as shown in table 3.3:

Table 3.3: Setting up the first hybridisation

Component	Tester 1-1	Tester 1-2
Rsa I-digested Driver (Female) DNA	1.5 μl	1.5 μl
Adaptor 2R-ligated Tester 1	1.5μl	-
Adaptor 2R-ligated Tester 2	-	1.5 μl
4X Hybridisation Buffer	1.0 μl	1.0 μl
Final Volume	4.0 μl	4.0 μl

A total 1.4 μg of driver DNA was added in each subtraction hybridisation. The samples were then overlaid with a drop of mineral oil.

Second Hybridisation

The following reagents were added to a sterile tube per reaction:

- Driver DNA - 1 μl
- 4X Hybridsation Buffer - 1 μl

- Sterile Water - 2 μl

The samples were then overlaid with a drop of mineral oil. The second hybridisation was then carried out by mixing the tester samples together quickly with 700 ng of freshly denatured driver (by incubating the driver at 98°C for 2 min) and left to incubate for 48 hours at 68°C, with 1 μl of new denatured driver (700 μg) added every 20 hours. The reaction was then diluted with 200 μl of dilution buffer (20 mM HEPES; pH 6.6, 20 mM NaCl, 0.2 mM EDTA; pH 8.0). and heated at 68°C for 7 min in a thermal cycler.

PCR Amplication of subtraction product

The following procedures to enrich the subtraction product were done as described in Clontech cDNA subtraction manual:

After the hybridisation, 1 μl of the subtraction product was aliquoted and diluted to 20μl. One μl of the diluted product was used for first round PCR amplification using PCR Primer 1. Second round of PCR amplification was done by diluting 3 μl of the first PCR product with 27 μl of sterile water and using Nested PCR Primers P1 and P2 (see Appendix for primer details). The primary PCR master mix were prepared as shown in table 3.4.

Table 3.4: Preparation of the primary PCR master mix

Reagent	Per Reaction
Sterile water	19.5 μl
10X PCR reaction buffer	2.5 μl
dNTP mix (10 mM)	0.5 μl
PCR Primer 1 (10 μM)	1.0 μl
50X Advantage cDNA Polymerase mix	0.5 μl
Total volume	24.0 μl

24 μl of the master mix was aliquoted into each of the reaction tubes and overlaid with 50 μl of mineral oil and incubated at 75°C for 5 minutes in a thermal cycler to extend the adaptors.

The following thermal cycling conditions were then commenced immediately after incubation for 27 cycles:

- 94°C - 30 sec
- 66°C - 30 sec
- 72°C - 1.5 min

3 μl of each primary PCR mixture were diluted in 27 μl of water. The secondary PCR master mix were then prepared as in table 3.5

Table 3.5: Preparation of the secondary PCR master mix

Component	Per Reaction
Sterile water	18.5 μl
10X PCR reaction buffer	2.5 μl
Nested PCR primer 1 (10 μM)	1.0 μl
Nested PCR primer 2R (10 μM)	1.0 μl
dNTP mix (10 μM)	0.5 μl
50X Advantage cDNA Polymerase mix	0.5 μl
Total Advantage cDNA Polymerase mix	24 μl

24 μl of master mix was then aliquoted into each reaction to 1 μl of each diluted primary PCR. One drop of mineral oil was overlaid onto each reaction, and put into a thermal cycler to commence the following conditions for 12 cycles:

- 94°C - 30 sec
- 68°C - 30 sec
- 72°C - 1.5 min

8 μl from each tube was then run on 2 % agarose/EtBr gel in 1X TAE buffer for analysis.

Subtraction efficiency test

PCR was used to estimate the efficiency of subtraction by comparing the amplification of one rare DNA segment of the genome. In this case, we used primers for Internal Transcribed Spacer 2 (ITS2) sequence, which amplifies ribosomal DNA, and *tra-2* for comparison. ITS2 is universal and abundant in the genome, whereas *tra-2* is rare, and probably a single-copy gene.

The following were combined in a microcentrifuge tube:

- Diluted subtracted DNA (2° PCR product) - 1.0 μl
- Forward primer (10 μM) - 1.2 μl
- Reverse primer (10 μM) - 1.2 μl
- Sterile water - 22.4 μl
- 10X PCR reaction buffer - 3.0 μl
- dNTP mix (10 mM) - 0.6 μl
- 50X Advantage cDNA Polymerase mix - 0.6 μl

The reaction was then overlaid with mineral oil and put in the thermal cycler with the following cycling conditions:

- 94°C - 30 sec
- 60°C - 30 sec
- 68°C - 2 min

18 cycles were first performed, and 5 μl was sampled from each reaction and placed in a clean tube. The rest of the reaction was put back into the thermal

cycler for 5 additional cycles for up to 38 cycles, 5 μl was sampled after each additional 5 cycles (i.e. 5 μl removed after cycles 18, 23, 28, 33, and 38. The aliquoted samples were then visualised on a 2.0% agarose/EtBr gel.

Dot blot

The PCR product was then cloned to make a DNA library by cloning the PCR products into pGEM-T easy vectors (Promega), as described in Chapter 2.

The DNA library was used in a dot blot to search, analyse and determine the difference between males and females. Library clones were screened in batches of 96 (96 well PCR plates) by PCR using M13 primers. Four μl of the PCR products were then diluted in 6 μl of denaturing buffer, covered with 20 μl of mineral oil and denatured at 100°C for 15 minutes. The mixtures were immediately chilled on ice. Two μl of each sample was dotted onto nitrocellulose membranes in duplicate. The dotted membrane was subjected to UV light for 1.5 minutes to cross-link the DNA to the membrane and allowed to dry overnight at 37°C. The membrane was hybridised with 10 μg of DIG-labelled *Rsa I* digested female DNA probes to detect DNA that were also present in females. Probe DNA was digested with *Rsa I* to maintain similarity as to the *Rsa I* digested DNA in the subtractive hybridization. Probes were made using DIG-Chem Link Label Set (Roche), according to manufacturer's protocol. Finally, the membrane was exposed to X-ray film for 3 - 6 hours at 37°C, and developed with Kodak film development reagents. Hybridisation results were reconfirmed by repeating the procedure using duplicate membranes.

The PCR products which showed no hybridisation to the probe, were sequenced and subjected to BLASTn search on NCBI and Broad Institute *Ae. aegypti* databases.

3.2.2 Laser microdissection and microarray

3.2.2.1 Laser microdissection of chromosome pairs

Chromosome preparation and laser microdissection

To obtain metaphase chromosomes, about 100 *Ae. aegypti* Rockefeller strain 4th instar larvae were treated with 0.1% colchicine solution for one hour. The larval neural ganglions were then dissected out and spread onto special polyester (POL)-membrane slides (Leica Microsystem Ltd.), pre-heated at 65°C, using methanol:glacial acetic acid (3:1) mix as a fixative. The slides were allowed to dry for 3 minutes at 65°C, then stained with 2% giemsa stain, pH 6.8 and using a laser microdissection machine (Leica Microsystem Ltd.), the chromosomes of interest were excised. A total of about 200 of each type of metaphase chromosome pairs were dissected for each type of chromosome from the POL-membrane slides.

Preparation of chromosomal DNA for microarray

Fifty to sixty pairs of each *Ae. aegypti* chromosomes were aliquoted and digested with *Mbo I* enzyme for 3 hours at 37°C, and then denatured at 75°C for 20 min. The chromosome samples were purified using the phenol-chlorform method mentioned in Chapter 2.

The subsequent steps were carried out primarily by Dr. K. P. Hoang, adapting the technique as described by Johnson (Johnson, 1990):

Mbo I linker adaptors were prepared by annealing two of the following sequences:

- 24-mer: 5' -GAT CTG TAC TGC ACC AGC AAA TCC- 3'
- 20-mer: 3' -... .AC ATG ACG TGG TCG TTT AGG- 5' with a dephosphorylated 5' end.

The overhanging *Mbo I* end of the adaptor will ligate to the *Mbo I* digested chromosomal DNA, while the dephosphorylated 5' termini at the blunt end of the adaptor will prevent them from self-ligation to form long repeating units that would interfere with amplification. The two oligonucleotides were annealed at 60°C for one hour at a concentration of 100 μg/ml. The phenol-extracted chromosomal DNA was ligated to the adaptor by adding 1/2 vol of ligation reaction mixture containing 0.25 Tris-HCl, (pH 7.6), 50 mM MgCl, 5 mM ATP, 5 mM dithiothreitol, 25% polyethylene glycol 8000, and an equal volume of T4 DNA ligase (8 units/μl) mixed with another 1/2 volume of adaptor (100 μg/ml). The ligation mixture was incubated at 12°C for 16 hours.

After ligation, the mixture was used for amplification with the 20-mer of the adaptor as primer. The PCR protocol is standard as in Chapter 2, except with

the following PCR cycle conditions:

- 1 cycle at 94°C for 5 min.
- 1 cycle at 45°C for 2 min.
- 1 cycle at 72°C for 3 min.
- 30 cycles at 94 at 94°C for 1.5 min., 45°C for 2 min., at 72°C for 3 min.
- extension at 72°C for 10 min.

The DNA was then purified by PCR clean-up kit (Qiagen) and tested to see if the amplification was successful by taking an aliquot of the sample and running on an agarose gel to visualise it.

Total DNA preparation for microarray

Total DNA from 4 adult females and males were also separately extracted by the Livak method, and purified by phenol-chloroform extraction, as described in Chapter 2. The sample was then suspended in TE buffer. DNA will be used in the 4^{th} microarray (Male total DNA versus female total DNA).

3.2.2.2 Microarray design and preparation

Sequences for the microarray design were gathered by downloading all the expressed sequence tags (ESTs with ends trimmed for polyA/T sequences) from VectorBase (http://aaegypti.vectorbase.org/GetData/Downloads/). The CGH array was designed by Dr. Peter Cook (postdoctoral member of the lab), using Agilent eArray online software to custom build a 4×44K microarray via the gene expression modules, but printed with CGH control features. A 60-mer probe was designed for each exon longer than 150 bp, which meant 26,645 exon sequences, representing 15,394 genes of 99.8% of the genome, were designed.

The chromosomal DNA and total gDNA samples were used for the CGH microarray hybridisation and analysed.

3.2.2.3 Universal Linkage System (ULS) labelling of CGH microarray

Samples for the array were labelled with individual reactions using genomic DNA ULS labelling kit (Agilent) according to the oligonucleotides array-based CGH for genomic DNA analysis protocol, with some minor changes. The protocol is briefly described below:

Step 1: Preparation of chromosomal and genomic DNA before labelling

500 ng of chromosomal DNA or total gDNA was required for each array. All samples were in 8 μl TE buffer, compatible with ULS labelling, with a minimum concentration of 62.5 ng/μl, in nuclease free PCR tubes.

Step 2: Fragmentation

As the *Mbo I* linker adaptor method had already fragmented the DNA, there was no use to heat fragment the DNA by heating at 95°C for 10 minutes as recommended. The samples were incubated at 95°C in the PCR machine with heated lid for 1 minute to denature the DNA. The sample tubes were then transferred to ice and incubated on ice for three minutes. The tubes were then spun in a microcentrifuge for 30 seconds at 6000 $\times$ g to drive the contents of the walls and lid. The tubes were stored on ice immediately ready for ULS labelling.

Step 3: ULS labelling

One Cy3 and one Cy5 labelling master mix were prepared by mixing the components shown in table 3.6.

Table 3.6: Preparation of labelling master mix for microarray

components	per reaction (μl)	per slide (μl)
nuclease free water	0.5	2.5
ULS-Cy3 or ULS-Cy5	0.5	2.5
10 × labelling solution	1	5
final volume of labelling master mix	2	10

Two μl of labelling master mix were added to each PCR tube containing the gDNA to make the total volume of 10 μl and mixed by pipetting up and down. The tubes were then transferred to a PCR machine with heated lid and incubated at 85°C for 30 minutes. The samples were transferred to ice and incubated for 3 minutes, and spun in a microcentrifuge for one minute at 6000 × g to drive the contents off the walls and lid. 10 μl of nuclease free water were added to each PCR tube to make the total volume of 20 μl.

Step 4: Removal of non-reacted Cy-ULS

KREApure columns (Agilent) were used to remove non-reacted ULS dye according to the manufacturer's protocol. A NanoDrop ND-1000 UV-VIS spectrophotometer (set to MicroArray Measurement menu with Sample Type DNA-50) was used to measure the gDNA quantitation, absorbance at A_{260} nm (DNA), A_{550} nm (Cy3), and A_{650} nm (Cy5) for calculating the degree of labelling according to the manufacturer's manual. After degree of labelling was calculated and deemed efficient (degree = 1.5% to 3.0%), the ULS-Cy5 labeled sample and ULS-CY3-labeled sample were combined as follows for a total volume of 37 μl in a 1.5 ml heat resistant microfuge tube:

- Array 1 - Chromosome 1 (Cy5) vs. Chromosome 2 (Cy3)
- Array 2 - Chromosome 1 (Cy5) vs. Chromosome 3 (Cy3)
- Array 3 - Chromosome 2 (Cy5) vs. Chromosome 3 (Cy3)

- Array 4 Adult male DNA (Cy5) vs. Adult female DNA (Cy3)

3.2.2.4 Microarray processing and feature extraction

Step 1: Preparation of labelled genomic DNA for hybridisation

The 100× blocking agent (supplied with Agilent oligo aCGH hybridisation kit) was prepared by adding 135 μl of nuclease free water. Water baths were equilibrated to 95°C and 37°C. Using a speed-vac, the combined Cy3 and Cy5-labelled gDNA mixture were concentrated to 22 μl. The hybridisation master mix was prepared by mixing the following components shown in table 3.7:

Table 3.7: Preparation of hybridisation master mix for microarray

components	volume per hybridisation (μl)	volume per slide (μl)
Salmon sperm DNA (1.0 mg/ml)	5	25
Agilent 100 X CGH blocking agent	1	5
Agilent 2X CGH hybridisation buffer	55	275
final volume of hybridisation master mix	61	305

Sixty one μl of the hybridisation master mix and was added to 1.5 mL microcentrifuge tube containing the labelled gDNA to make the total volume of 83 μl. The samples were mixed by pipetting up and down, and then spun quickly in a microcentrifuge to drive the contents of the walls and lid. The sample tubes were transferred to a circulating water bath at 95°C and incubated for three minutes. The samples were immediately transferred to a circulating water bath at 37°C for 30 minutes. The samples were quickly spun in a microcentrifuge to drive the content off the walls and lid. 27 μl of Agilent-CGH Block (supplied with the gDNA ULS labelling kit) was added to each sample to make the final volume of 110 μl hybridisation sample mixture.

Step 2: Hybridisation assembly

A clean gasket slide was loaded into the Agilent SureHyb chamber base with the gasket label facing up and aligned with the rectangular section of the chamber base. 100μl of each hybridisation sample mixture was slowly dispensed onto the gasket well in a "drag and dispense" manner. The microarray slide was then pressed down onto the gasket slide, forming a sandwich pair. The SureHyb chamber cover was then placed onto the sandwiched slides and clamped. The assembled slide chamber was placed in the rotator rack in an everyday station set to 65°C. The hybridisation rotator was set to rotate at 20 rpm. The microarray was left to hybridise for 24 hours.

Step 3: Wash preparation

Slide staining dish, slide rack and stir rod were thoroughly washed with Milli-Q water and then with acetonitrile to remove any remaining residue of Agilent stabilisation and drying solution. The oligo aCGH wash buffer 2 was pre-warmed overnight at 37°C. The Agilent stabilisation and drying solution was warmed slowly in a water bath set to a 37°C to 40°C in a close container with sufficient head space to allow for expansion.

Step 4: Microarray washing

Microarray wash procedure B in the Agilent protocol was used as described in table 3.8.

Table 3.8: List of wash procedure B with stabilisation in drying solution

-	dish	wash buffer	temperature	time
Disassembly	1	Oligo aCGH Wash Buffer 1	room temperature	-
First wash	2	Oligo aCGH Wash Buffer 1	room temperature	5 minutes
Second wash	3	Oligo aCGH Wash Buffer 2	37°C	1 minute
Acetonitrile wash	4	Acetonitrile	room temperature	1 minute
Third wash	5	stabilisation and drying solution	room temperature	30 seconds

Step 5: Microarray scanning using Agilent scanner

The microarray was sent to Oxford Gene Technology (OGT) Ltd. for scanning. The default scan settings were verified and set to scan with appropriate 4x44K microarray settings (Scan region = 61 × 21.6 mm, Scan resolution = 5 μm, Dye channel = Red & Green, Green PMT = 100%, Red PMT = 100%).

Step 6: Data extraction using Feature Extraction software

The Feature Extraction software v9.5 (Sanger Institute) was used to extract the data of the scanned CGH microarray. After the extraction was completed, the Quality Control (QC) report for each extraction was obtained. The results were analysed with the assistance of Ramona Schmid and Wolfgang Hubert from the European Bioinformatics Institute (EBI) in the Sanger Institute.

PCR of candidate genes

Candidate genes that were found after analysis were then tested by PCR to confirm male specificity, using 50 ng of total DNA for each reaction (see chapter 2 for PCR methodology).

3.3 Results

3.3.1 Subtractive hybridisation results

Adaptor ligation

Two of the three ligation attempts made were successful. When the ligation efficiency test was carried out, the first sample, labelled A did not show the expected bands, therefore this sample was not used for subtraction. Sample B, which is a repeat of Sample A, and C, a sample with low concentration, showed efficient ligation (figure 3.3 and figure 3.4). Samples B and C were therefore used for the subsequent steps of subtraction.

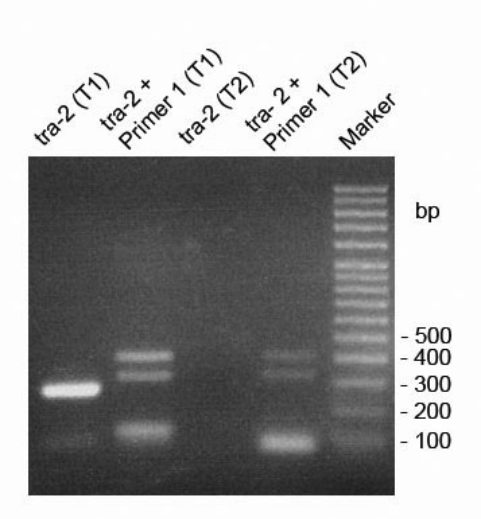

Figure 3.3: Adaptor ligation efficiency test of Sample A. *tra-2* sequence in tester 1 ligated to Adaptor 1 labelled *tra-2* (T1), was amplified strongly though of the wrong size; while no amplification was seen in the tester 1 ligated to Adaptor 2R, labelled *tra-2* (T2). The amplification between forward *tra-2* primer and PCR primer 1 in both testers showed double banding, indicating inadequate ligation.

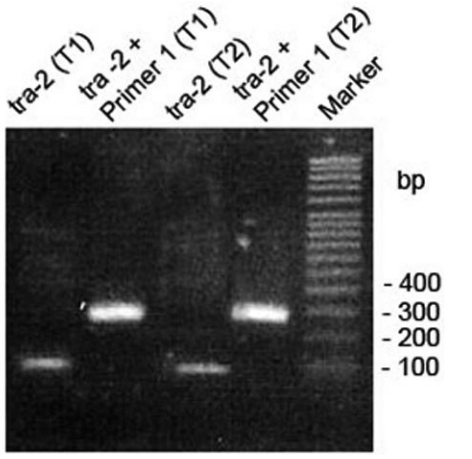

Figure 3.4: Adaptor ligation efficiency test of Sample B. *tra-2* sequence in Tester 1 ligated to Adaptor 1 labelled *tra-2* (T1), Tester 1 ligated to Adaptor 2R, labelled *tra-2* (T2) were amplified. The amplification between forward *tra-2* primer and PCR primer 1 in both testers showed single banding, indicating successful ligation.

Hybridisation

Subtractions from samples B and C were done and their efficiencies were compared. There was positive amplification of *tra-2* sequence in Subtraction B after 23 PCR cycles. ITS2 sequence did not amplify even after 38 cycles (figure 3.5). Subtraction C however showed positive amplification for *tra-2* after 28 cycles. Subtraction C ITS2 sequence was amplified after 33 cycles. Subtraction B was shown to be the most stringent of all with *tra-2* sequence amplified roughly 20 fold (5 PCR cycles) more than that in Subtraction C.

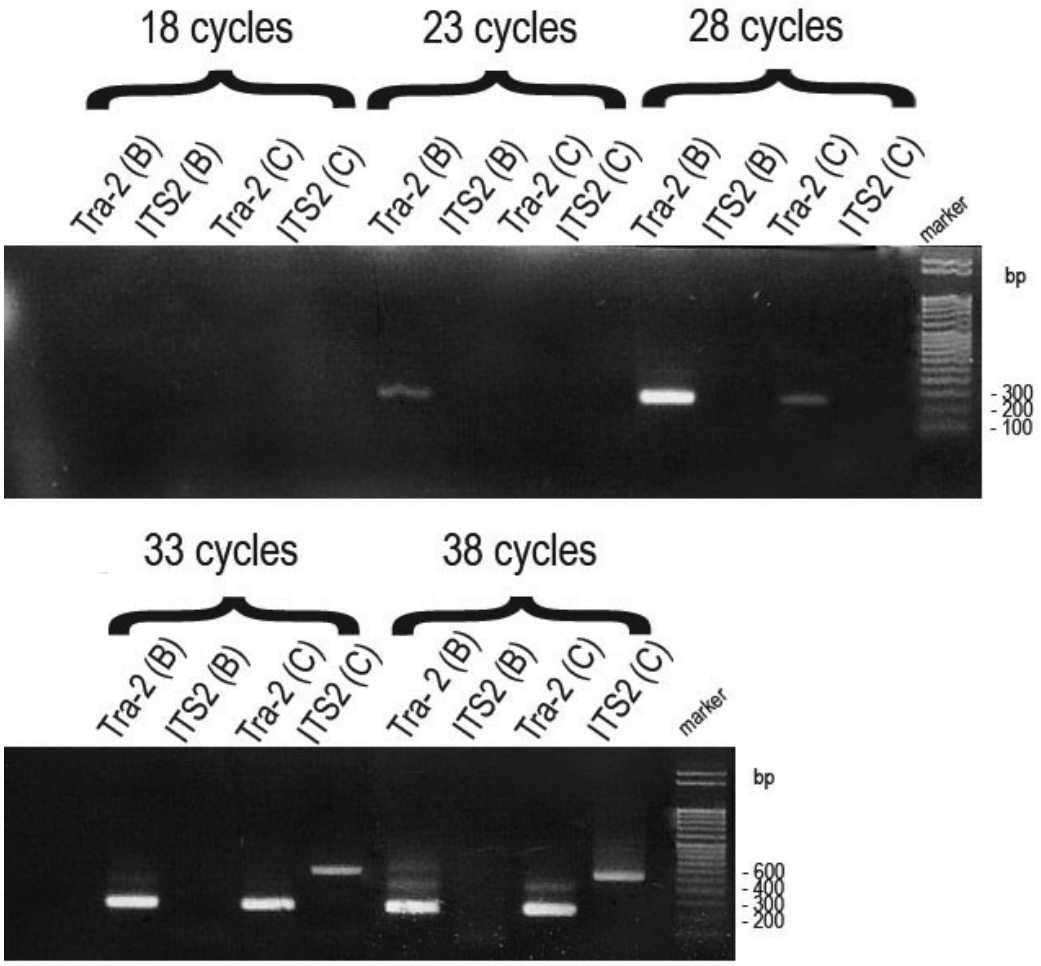

Figure 3.5: Subtraction efficiency test. 2 of 3 subtraction attempts were successful, which are labelled B and C in brackets. Subtraction B shows that *tra-2* sequences appear after 23 PCR cycles whereas Subtraction C *tra-2* appears after 28 cycles. No ITS2 sequences were amplified in Subtraction B whereas Subtraction C ITS2 was amplified after 33 cycles.

Dot blot

Both Subtraction B and C were cloned and screened by dot blot. A total of 6272 clones were picked to form a library in 96 well plates (64 plates in total). Twelve plates were randomly selected and screened using the dot blot method. Twenty eight candidate clones that did not hybridise strongly to the DIG-labelled female DNA probes (figure 3.6), were amplified using M13 primers and sequenced. Only 3 of the clones showed useful sequences, i.e. sequences that contained inserts and are matched in the *Ae. aegypti* genome when BLASTn searches were carried out (table 3.9).

Table 3.9: Candidate clones for male specificity

Clone ID	supercontig	position in contig
P02-D5	1.48	2290462 - 2290913
P03-B5	1.493	125551 - 125646
P01-A1	1.479	249065 - 249207

The sequenced clones that did not hybridise strongly to the probes were found to be present in both sexes when amplified by RT-PCR. RT-PCR was used instead of PCR to ensure that sequence was expressed. One sequence (clone ID: P02-D5) was found to be from an un-annotated region of supercontig 1.48 whereas Clone P03-B5 was an un-annotated region of supercontig 1.493. Clone P01-A1 was found best matched to an un-annotated region of supercontig 1.479, though this short sequence was also matched in many positions of the same supercontig, indicating that it is a non-specific, repetitive sequence (Refer to the Appendix for the full sequences of these clones).

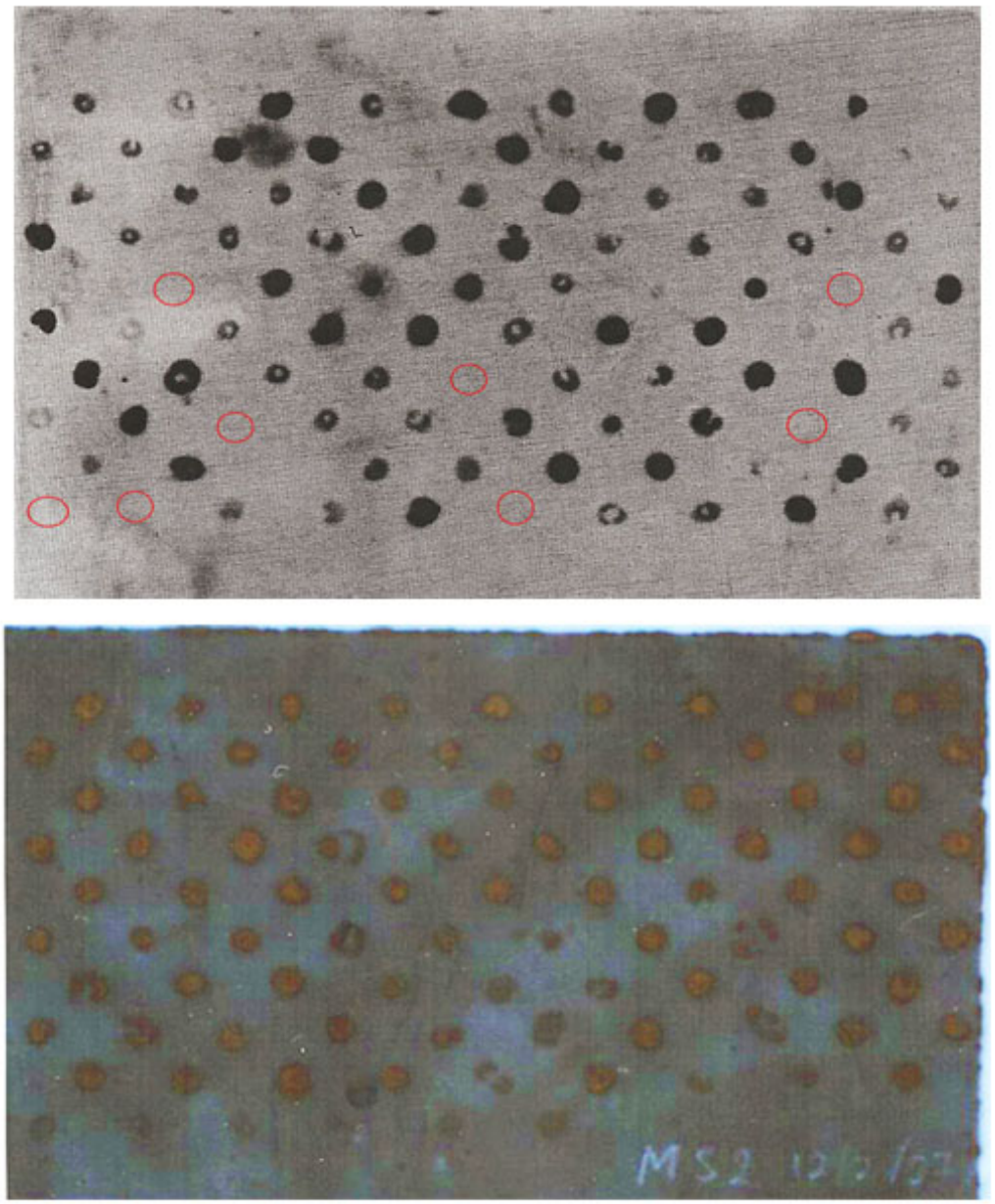

Figure 3.6: Dotblots after hybridisation of male gDNA to DIG-labelled *Rsa I* digested female gDNA probes and exposed to X-ray film to visualise the result. The top picture shows an example of a dotblot with several clones that were unhybridised to probes (circles red), while most of the dotblot clones were all strongly hybridised to the probes in the bottom picture.

3.3.2 Laser microdissection and microarray results

The chromosome pairs were successfully microdissected (see figures 3.7 and 3.8 for examples of microdissection) and pooled into chromosome 1, 2, and 3 to

achieve enough starting material (50-60 chromosome pairs in each pool).

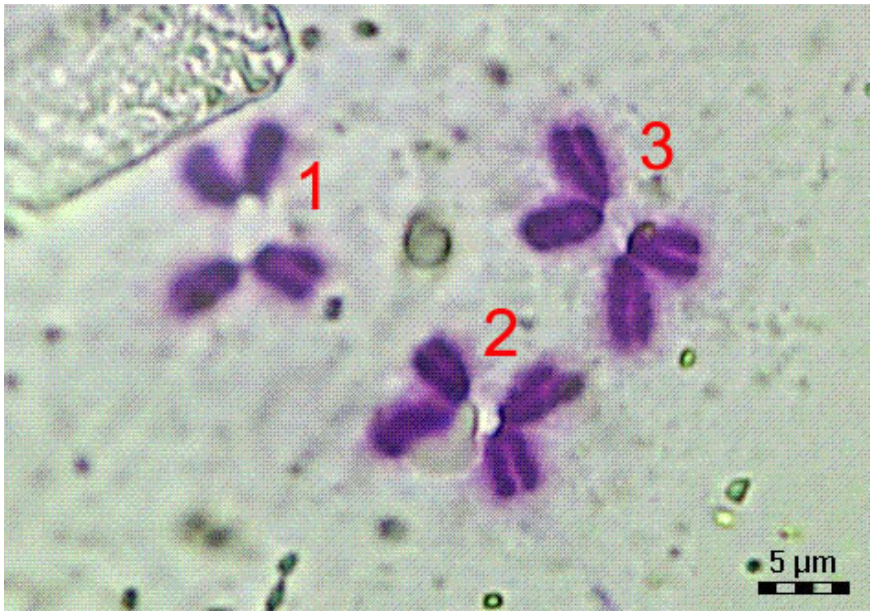

Figure 3.7: Giemsa stained *Aedes aegypti* chromosome at metaphase on POL-membrane slide ready for microdissection. Chromosome pairs are as indicated in red numerals. Note that chromosomes are not easily distinguished.

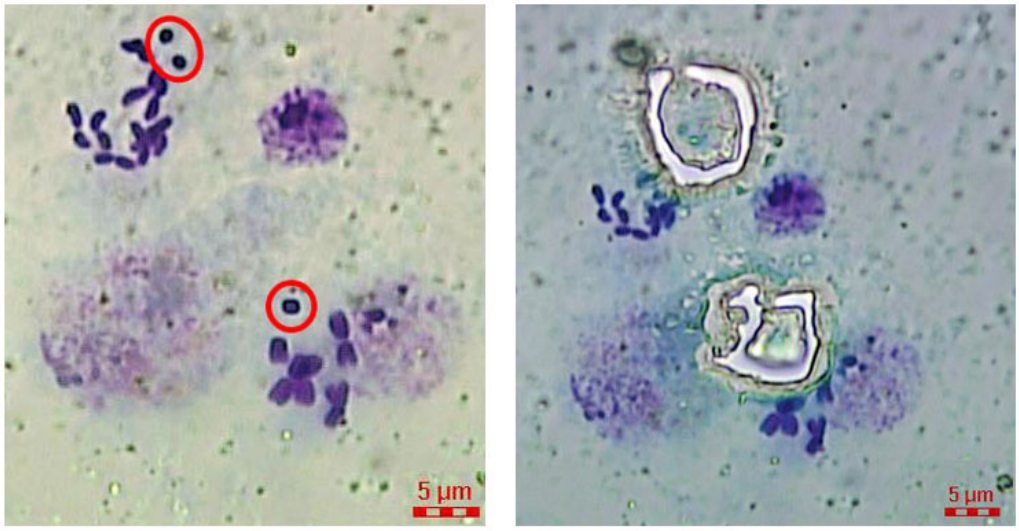

Figure 3.8: *Anopheles gambiae* Y chromosome microdissection (by Dr. Kim Phuc Hoang). Photo on the left shows metaphase chromosomes with the Y chromosome circled red. Y, or chromosome 1, is distinct in comparison to culicine mosquitoes (see figure 3.7. Photo on the right shows the same field while Y chromosomes are being microdissected with Leica microdissection laser at smallest laser diameter setting.

The DNA obtained after amplification of the dissected chromosome was loaded into an agarose gel (15 μl per sample) to visualise. Samples that gave a smear of fragmented DNA between 50 to 1100 base pairs were seen (range of *Mbo I* cleavage), indicating successful amplification (figure 3.9).

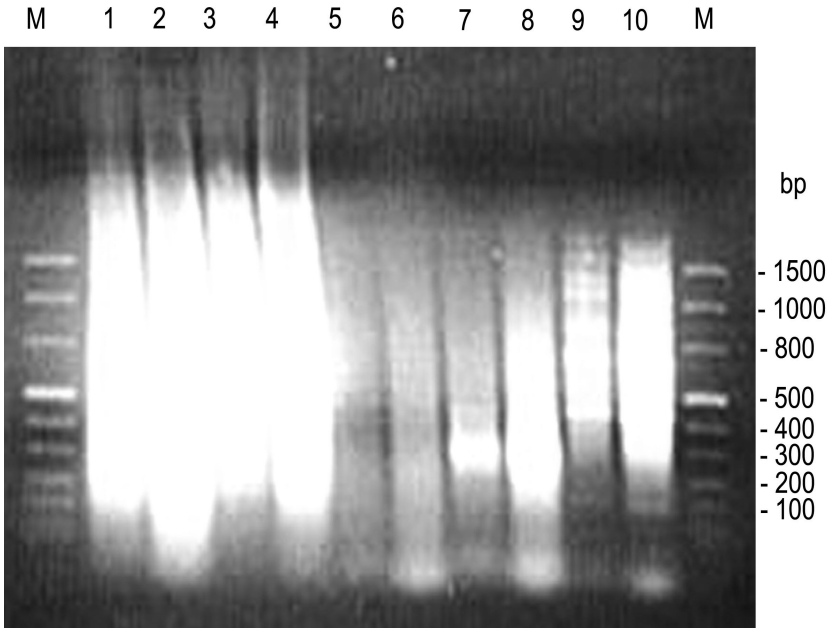

Figure 3.9: PCR of fragmented *Ae. aegypti* chromosomal DNA. Ten samples were selected and visualised on 2% agarose gel. Smears of 50 to 1100 bp indicates successful amplification of the chromosomal DNA. M = DNA marker ladder. Result by K.P. Hoang.

Both array 1 and array 2 hybridised strongly to chromosome 1 gDNA, displaying red (Cy5) signals rather than green (Cy3) or mixed signals (see figure 3.10). Array 3 did not show strong hybridisation to either chromosome 2 or 3 gDNA. Array 4 hybridised strongly to both female and male gDNA, as expected.

The array male DNA versus female DNA hybridisation suggests 220 possible candidates for male specific genes after analysis from Schmid and Hubert of EGI (see Appendix for further information). The number of candidates were further narrowed down by selecting the hybrised genes that showed high intensity signals to the male DNA array, and not surrounded by adjacent genes that are equally hybridised to female and male DNA on its corresponding supercontig by Schmid and Hubert. These selected candidates were then tested for male specificity by PCR. Genes tested are listed in table 3.10.

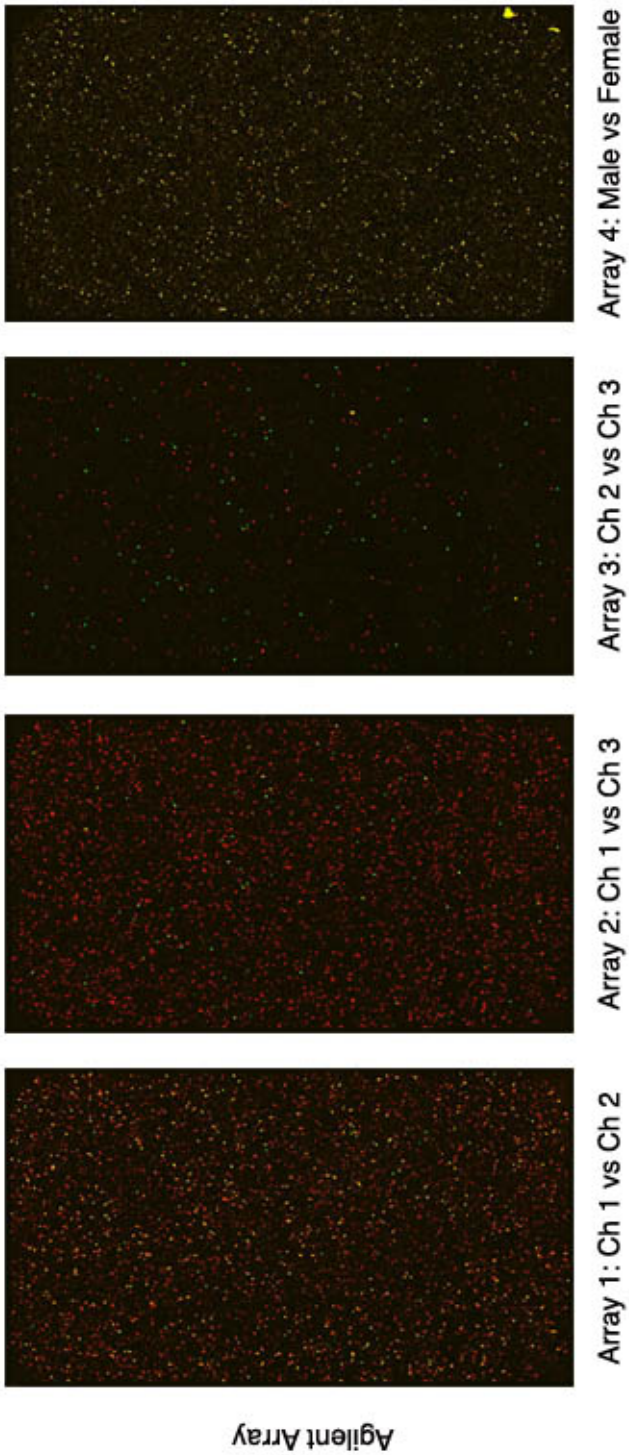

Figure 3.10: Arrays scanned after hybridisation. Array 1 and 2 show extensive red signals in comparison to array 3, indicating that array 1 and 2 hybridise strongly to chromosome 1 pooled DNA. Array 4 hybridises to both male and female gDNA resulting in mostly yellow signals.

3.3.2.1 PCR of candidate genes

Primers were designed by Dr. P. Cook to amplify the male candidate genes using PCR. Primers are as shown in table 3.10. All candidate gene primer sets were found to amplify both male and female gDNA (See figure 3.11).

Table 3.10: List of best candidates genes to be male specific, based on microarray 4 results. Listed are the primers used to amplify the gene candidates in samples. Primer design by Dr. P. Cook.

Gene	Supercontig	Forward primer	Reverse primer
AAEL015242	1.1647	AGGTCACGGACTGGCTCA	CTCCCAACCAACCACAGTTT
AAEL015048	1.1421	AAGTGCTACCGGACGATGAC	CAACGAATCTGCGTCGTCTA
AAEL015447	1.2305	ACATGGTCAACGTCCTCCTC	GCCCATGATGAACTTGTCCT
AAEL015308	1.1778	ACCGCTTGGGGTACGATT	CTACGGTCGGGTAATTTTGC
AAEL015365	1.1925	CGGTAGAAATGGCGAATGTT	CTACGAAAGCCTGACGGAAG
AAEL015394	1.2032	CATCTGTTACGGTGCAATCG	ACCTGGTCATCGGTGTTCTC
AAEL015492	1.2646	GCAAAGTGGAAGGAAGAACG	AGCGGGCATTTGATCTCTAC
AAEL015554	1.3221	GAAACATGCAGGAGGCATTT	CTGGCTCTGTTGGCAATGTA
AAEL015578	1.3489	GAGCGTTAGGCTTTTTGTGC	CAGTTTCGCTGTGGAGACAA
AAEL015597	1.3615	CTAGCACCGCACCTGAAAAT	TTTGCACACTGGCGAATATC
AAEL015631	1.3912	GTTGAGGGCAACTTGTGGAT	CAGAGGACTTTTTGCGGAAG

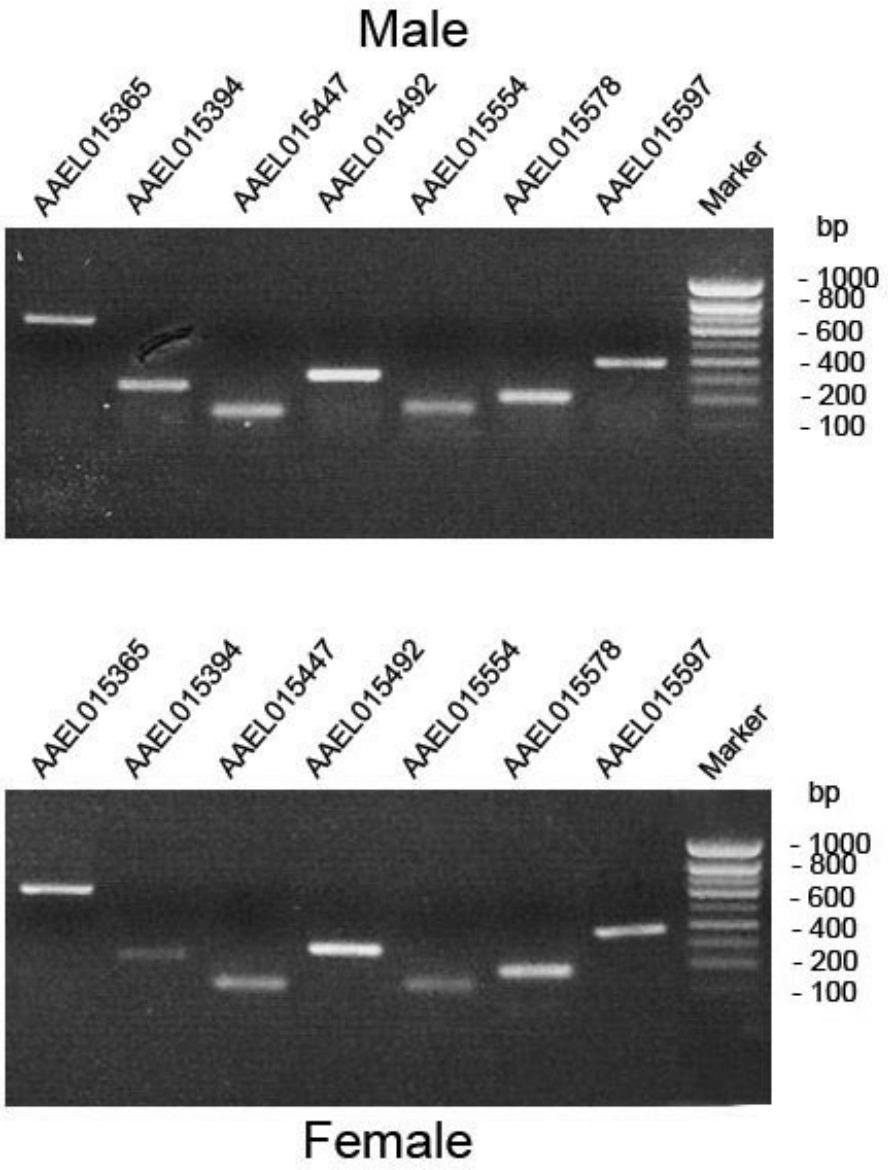

Figure 3.11: PCR of candidate sequences for male specificity. Genes tested were positive using male DNA (upper picture) and female DNA (lower picture) as template.

3.4 Discussion

3.4.1 Identifying male-specific genes

3.4.1.1 Subtractive hybridisation

From the results, it was concluded that this method was not stringent enough to isolate the small proportion of male specific sequences. It is possible that male and female versions of chromosome 1 in Culicine mosquitoes only differ slightly and not enough to allow detection by subtractive hybridisation. Another possibility is that the male-specific factor lies within a region of heterochromatin and hence was unable to be digested by the utilised *Mbo I* enzyme.

3.4.1.2 Laser microdissection and microarray

Chromosome dissection is a novel approach to isolate individual mosquito chromosomes. However contamination after microdissection could not be prevented due to the environment where the procedure was carried out. This is because the room where the laser microdissection machine was set up was a small room which meant the airflow was greatly affected everytime the door was opened and closed. Also, an air duct was right above the machine, which could introduce contaminants onto the samples, although efforts were made to minimize this by covering the duct with a board. As for isolating male chromosome 1, the fact that male and female chromosome 1 is difficult to distinguish from each other adds another hurdle to the identification. It was not possible to distinguish male and female versions of the culicine sex chromosomes (chromosome 1) without the C-banding procedure, as POL-membrane slides for the laser microdissection were too fragile to withstand the

treatment. Nevertheless, the attempt was deemed worth pursuing given the importance of the aims, since not much advancement had been undertaken in identifying male determining genes on chromosome 1. The size of chromosomes and numbers of chromosomes that were extracted were also small. Thus, it was necessary to perform many rounds of PCR amplification to obtain enough material for microarray hybridisation. This also meant amplifying any contaminating DNA which is likely have affected the results. In addition, though the size of the laser beam was small, it was still too thick for mosquito chromosome dissection (see figure 3.8 for example). There could be contaminants from neighbouring chromosomes that were caused by the laser beam. This would explain the strong hybridisation of pooled DNA from chromosome 1 to probes on array 1 and 2.

Chapter 4

Identification and characterisation of sex determining gene homologs in *Aedes aegypti*

4.1 Introduction

Studies of the mosquito sex determination pathway, based on comparisons with *D. melanogaster*, could provide targets for knockdown of genes that are essential for female development. The first aim of this section was to find potential sex determining gene targets that could be used as a switch to manipulate sex ratio. For this, it is necessary to characterise candidate genes, understand the intron-exon structure, and see if there are any expression differences between the sexes. Using *Drosophila* as a model for finding potential sex determining switches, several expression of genes in the sex determination pathway were explored: *tra-2, dsx, fruitless,* and *daughterless*. The gene *transformer-2 (tra-2)* was the main focus for characterising its putative homolog due to its potential for later utilisation as a sex-determining switch because of it being further up the sex determination cascade than the other genes mentioned and presumably more conserved than *tra*. Therefore, the major aim in this study was to characterise *tra-2* expression patterns, exon-intron structure and whether there are any differences between males and females.

Drosophila studies also suggested that *tra-2* can be used as a switch to distort sexual development to a female bias using RNAi (Fortier and Belote, 2000), as mentioned in Chapter 1. At the beginning of this study, *tra-2* was not annotated in any of the genome databases (i.e. TIGR, Ensembl, NCBI, Broad Institute). *Ae. aegypti tra-2* was partially identified using automated annotations, and it was reported to be 585 bp long with 2 exons (Nene et al., 2007). In VectorBase, AAEL004293 was only described as a gene that encodes a hypothetical conserved protein.

4.1.1 Aims

4.1.1.1 Sex determination pathway homologs

Using *Drosophila* as a model, several putative orthologs of genes in the sex determination pathway in *Ae. aegypti* were explored to find potential sex determining switches: *tra-2, dsx, fru*, and *da*.

4.1.1.2 Characterising *tra-2*

To characterise the full length of transcribed *tra-2*, a technique called Rapid Amplification of cDNA Ends (RACE) was used. RACE is a method to amplify cDNA copies of the RNA sequence of interest, produced through reverse transcription (see figure 4.1). The cDNAs were amplified by PCR using known sequences from the middle of the *tra-2* transcript obtained from the in silico search of *tra-2*. The cDNAs were then cloned and sequenced. If the sequence is long enough, it would be possible to map the full sequence of unique *tra-2* mRNA variants, if any. RACE can provide the sequence of an RNA transcript with a small known sequence within the transcript to the 5' end or 3' end of the RNA. In *D. melanogaster*, at least 7 *tra-2* mRNA transcripts were noted, and some were sex-specific (Mattox et al., 1990), as mentioned in Chapter 1. Hence it would be expected that *tra-2* splice variants would exist in *Ae. aegypti* and show some sex-specific variation, if this gene is indeed a functional ortholog of *tra-2* in *D. melanogaster*.

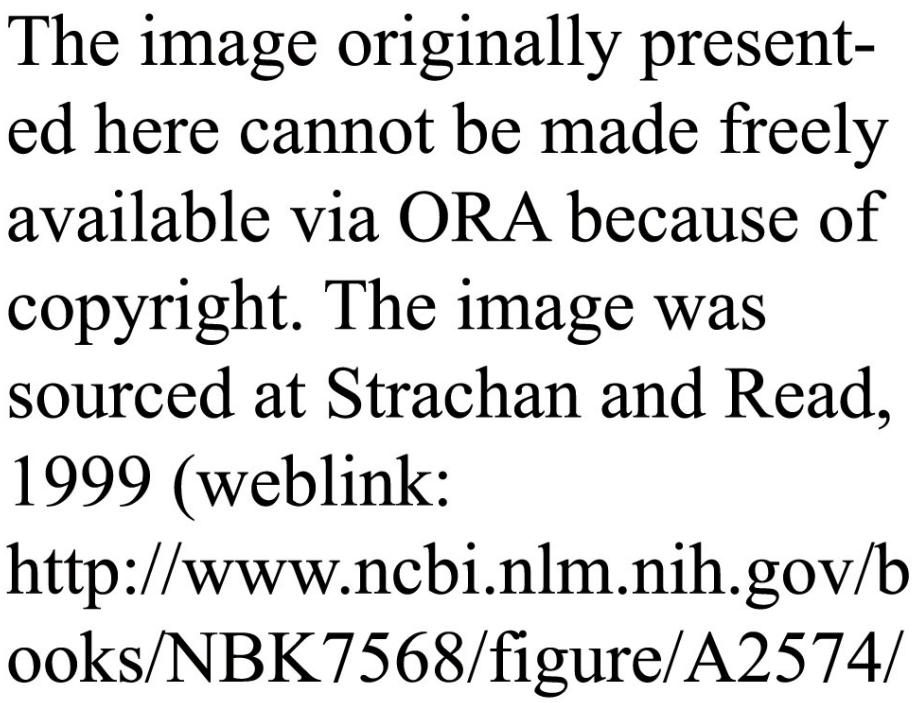

Figure 4.1: RACE-PCR facilitates the identification of 5' and 3' end sequences from cDNA. A preliminary step involves the introduction of a specific sequence at either the 3' or the 5' end by what is effectively a form of 5'-add-on mutagenesis. 3' RACE-PCR uses a starting antisense primer with a specific 5' extension sequence which becomes incorporated into the cDNA transcript at the reverse transcriptase step. An internal sense primer is then used to generate a short second strand ending in a sequence complementary to the original anchor sequence. Thereafter, PCR is initiated using the internal sense primer and an anchor sequence primer. 5' RACE-PCR. Here an internal antisense primer is used to prime synthesis from a mRNA template (blue) of a partial first cDNA strand (black). A poly(dA) is added to the 3' end of the cDNA using terminal transferase. Second strand synthesis is primed using a sense primer with a specific extension (anchor) sequence. This strand is used as a template for a further synthesis step using the internal primer in order to produce a complementary copy of the anchor sequence. PCR can then be accomplished using internal and anchor sequence primers. (Strachan and Read, 1999)

4.2 Methodology

4.2.1 Sex determining genes in *Ae. aegypti*

4.2.1.1 Database search for *dsx*

Putative *dsx* homologs were searched using tBLASTx (BLAST using a translated nucleotide query to search the translated nucleotide databases) against the *Anopheles gambiae dsx* nucleotide sequence: GenBank accession number DQ137802 (Scali et al., 2005). Nucleotide sequences that matched were then retrieved from the Broad Institute database. Sequences of *D. melanogaster*, *An. gambiae*, and the retrieved sequences of *Ae. aegypti* were then aligned with each other by using bioinformatics software Geneious.

4.2.1.2 RT-PCR of *dsx* regions

Two pairs of primers were designed for RT-PCR based on the retrieved sequence. The first pair of primers (A) were used to amplify the common region of both male and female domains by having the forward primer targeting the Oligomerisation Domain (OD), OD1; and the reverse, just after female specific part of OD2. The second pair of primers (B) however included the female specific OD2 (figure 4.2).

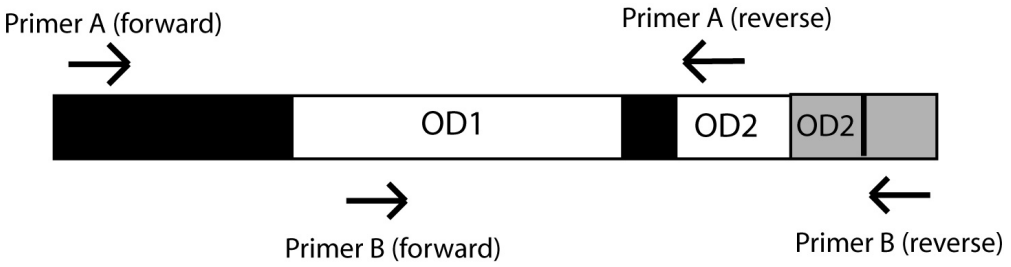

Figure 4.2: Diagram of *dsx* RT-PCR. Black box denotes exon region of *dsx* that does not encode any Oligomerisation Domain (OD) and are common in both sexes. Blank box denotes OD1 that is not sex-specific. Grey box denotes female specific sequence in *dsx*, and female specific part of OD2 indicated within the grey box. Primer pair A amplifies the common region of *dsx*, while primer pair B amplifies the regions OD1 and OD2, including the female specific region. Thus, primer pair B will only amplify in female samples.

Dsx primer pair A expected amplification size of 570 bp

- Forward: 5'- TGGTTTCGCAAGATCGCTGGATGGTAAA -3'
- Reverse: 5'- TCTCCGTCGGCGCCTTTCAGTATCACG -3'

Dsx primer pair B expected amplification size of 588 bp

- Forward: 5'- ACTGTGCCCGCTGCCGGAACCAC -3'
- Reverse: 5'- CGCGTCGTACTGCGCAACTCTACACC -3'

Approximately 100 ng of RNA was used in each RT-PCR (one step method) reaction. Each RT-PCR were repeated 4 times. The RT-PCR reactions were then loaded on a 1% agarose gel to visualise the products.

4.2.1.3 Database search for *da* and *fru* homologs

Putative homologs were searched by tBLASTn (BLAST using a protein query to search the translated nucleotide database) against the *D. melanogaster* sequences with GenBank accession number NM_001170177 for *da* and AE0142972

for *fru*. Nucleotide sequences matched were then retrieved from the Vectorbase. The sequences of the genes were then analysed using Geneious.

4.2.1.4 RT-PCR of *fru*

After identifying *da* and *fru* homologs in *Ae. aegypti*, primers were designed for amplifying *fru* in a RT-PCR.

Fruitless primers used - expected size of 101 bp (supercontig 1.199):

- Fru F (Forward): 5'- GGACCAGCAGTATTGCTTACG -3'
- Fru R (Reverse): 5'- CGAGCGTGACATCACACACAGT -3'

Cycling conditions were as mentioned in Chapter 2, with an annealing temperature of 56°C for One Step RT-PCR.

4.2.1.5 Database search for *tra-2*

The first method used was *in silico* searches using *D. melanogaster*, *An. gambiae*, and *Ae. aegypti* gene on databases to find the gene sequences. Protein sequences of *D. melanogaster tra-2* were first searched and retrieved from the *D. melanogaster* Ensembl database (website: http://www.ensembl.org/Drosophila_melanogaster/Info/Index/) by typing in 'transformer-2' in the gene search bar. Search using tBLASTn was performed to find similar sequences from the Broad Institute *Aedes aegypti* database (website: http://www.broadinstitute.org/annotation/genome/aedes_aegypti/Home.html), TIGR database (website: http://gsc.jcvi.org/projects/msc/aedes_aegypti/) and *Anopheles gambiae* Ensembl database (Now intergrated to VectorBase, website: http://agambiae.vectorbase.org/index.php).

4.2.1.6 *Tra-2* PCR and RT-PCR assays

Primers were designed based on the sequence result of the *in silico* search. PCR and RT-PCR were then performed on *Ae. aegypti* DNA and RNA respectively. Three pupae of each sex were used for DNA and RNA extraction and about 200 ng were used per amplification. Cycling conditions were as mentioned in Chapter 2, with an annealing temperature of 57°C for One Step RT-PCR.

The following primers were used to amplify *tra-2*. The expected amplification size was 135 bp:

- Tra-2F (Forward): 5'- ATGCCAAGACGAAGGTTTCCCG -3'
- Tra-2R (Reverse): 5'- CGGTCACCGAATAATCCACTCT -3'

4.2.2 *Tra-2* Northern Blot

Northern blot is a standard method for the detection and quantification of gene expression. Attempts were made to detect *tra-2* signals in males and females by Northern blot.

Northern Blot DIG probe synthesis

3 pupae were used to extract gDNA. About 100 ng of the extracted DNA was used to generate template *tra-2* gDNA for RNA probes by PCR with linker regions attached.

Primers for *tra-2* with linker attached were (underline indicates linker sequence):

- Forward: 5'- GGCCGCGGATGCCAAGACGAAGGTTTCCCG -3'
- Reverse: 5'- CCCGGGGCCGGTCACCGAATAATCCACTCT -3'

A second round of PCR was performed with the first round PCR as template but this time with T7 primers with linker regions attached. The product was then run on an electrophoresis gel to check the quality of the template. When satisfied with the quality, another 10 reactions of the same PCR were done to produce more templates. The products were then run on an electrophoresis gel, the band of interest excised, and purified using QIAQuick Gel Extraction Kit (Qiagen). The DNA template was then further purified by phenol-chloroform extraction and suspended in 13 μl of DMDC treated water. Labelled RNA probe was then made using Roches DIG Northern Starter Kit.

Primers for linker with T7 attached. Underlined indicates T7 sequence:

- Forward: 5'- GAGAATTCTAATACGACTCACTATAGGGGCCGCGG -3'
- Reverse: 5'- GAGAATTCTAATACGACTCACTATAGCCCGGGGC -3'

DIG-labelled RNA probe using actin 1 primers were also made to use as control.

Sample preparation

Total RNA from male and female pupae of *Ae. aegypti* were separately mass extracted and each aliquoted into 3 different concentrations: 20 ng/μl, 10 ng/μl and 1 ng/μl. Five μl of each RNA (1000 ng, 50 ng, and 5 ng total RNA) were then used for the Northern blot. The subsequent Northern blot technical procedures were carried out (Rueger et al., 1996), as described:

- 1.2% Formaldehyde agarose gel was prepared freshly.
- RNA target samples were denatured for 10 min. at 65°C in RNA loading buffer (10μl RNA loading buffer per 5μl of target RNA) and placed immediately on ice.

- The gel was placed in a gel tank that was sterilized with 10% hydrogen peroxide prior to use. The electrophoresis running buffer was added into the gel tank up to edges of the gel. The gel was pre-run for 5 min. at 5 V/cm. The dry gel was then loaded with the RNA samples and the RNA was allowed to move into the gel at high voltage briefly. Once samples were in the gel, more running buffer was added to submerge the gel.

- The gel was run at 25 V overnight under fume hood, adding 1mg/ml EtBr to the running buffer to ensure the gel remains stained uniformly with EtBr.

RNA transfer

- The gel was submerged in 20× saline-sodium citrate (SSC) at room temperature, shaking for 2 × 15 min. to remove formaldehyde.
- In a sterile transparent plastic box, 20× SSC was added to about 6 cm in depth.
- Two pieces of Whatman 1MM paper were placed on a smooth plate and wetted with 20× SSC. Any air bubbles were squeezed out by rolling a sterile pipette over the paper.
- The nylon membrane (positively charged) was placed on the gel surface, using 20× SSC as a transfer buffer.
- The surface of membrane with 20× SSC. Five sheets of Whatman 3MM paper were cut to same size as the gel and placed on top of the membrane.
- Clean paper towels were put on top of Whatman paper to about 6 cm, or to slightly above box rim. The box lid was closed carefully and weight was added to the box and left overnight.
- The flattened gel and membrane were recovered. The position of the wells on the membrane was marked to ensure correct orientation.
- The membrane was rinsed for 2 to 5 min. with 2× SSC, placed on Whatman 3MM paper (80°C) and baked dry for 30 min. in a drying oven.
- The RNA was fixed to the membrane by UV cross-linking for 1.3 min. (120 mJ).
- The membrane was washed for 30 min. in 1× SSC, 0.1% SDS at 65°C.

Hybridisation of blot

- In a sealed plastic bag, the blot was incubated in a 25 ml pre-hybridisation solution for 30 min. at 68°C.
- The probe was denatured for 5 min. at 100°C, then immediately chilled on ice.
- Three ml hybridisation buffer were pre-warmed to 68°C, then the probe was added to a concentration of 100 ng/ml.
- The probe mixture was added to the blot and was incubated overnight at 68°C in a water bath. The probe mixture was drained out of the blot and kept at -20°C for reuse.
- The blot was washed for 5 min. twice while shaking in low stringency buffer 1 at room temperature.
- The blot was washed for 15 min. twice while shaking in pre-warmed high stringency wash buffer 2 at 68°C. After this, the blot may be stored dry at 4°C or placed in wash buffer for immediate detection.

Chemiluminescent detection using Roche DIG Northern Starter kit

- The vial of antibody conjugate was centrifuged for 5 to 10 min. to pellet precipitates that form during storage. 0.5 μl from the surface of the antibody conjugate solution was pipetted and diluted in 2000 μl blocking buffer.
- The membrane was equilibrated in washing buffer (maleic acid, NaCl and Tween) for 5 min. in a transparent polythene bag.
- The membrane was washed for 30 min. in 1 $\times$ blocking buffer.

- The membrane was washed for 30 min. with antibody diluted 1:10000 in blocking buffer.
- The membrane was washed 2 × 15 min. in wash buffer.
- The membrane was then equilibrated for 2 min. in detection buffer.
- Disodium 3-(4-methoxyspiro {1,2-dioxetane-3,2-(5'-chloro)tricyclo [3.3.1.1^{3,7}]decan}-4-yl)phenyl phosphate (CSPD) substrate was diluted 1:100 in detection buffer.
- Approximately 500 μl of diluted chemiluminescent substrate were added to the membrane (per 100 cm^2 membrane).
- The membrane was incubated for 5 min. at room temperature.
- The excess liquid was then allowed to drip out of the polythene bag. The bag was then sealed without any air bubbles inside.
- The membrane was incubated for 10 min. at 37°C.
- The membrane was then placed in an X-ray cassette, and exposed to an X-ray film for 2 to 48 hours at 37°C.

4.2.3 Rapid Identification of cDNA Ends (RACE) PCR of *tra-2*

Using the previously identified region of *tra-2* (exon 1), RACE 3' and 5' were carried out to identify variants of *tra-2* in males and females (figure 4.3 shows a diagram of how RACE aims to identify the complete length of the transcript).

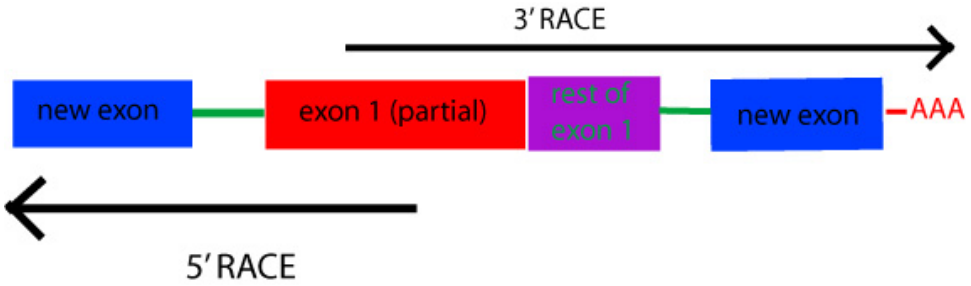

Figure 4.3: Overview of RACE to identify *tra-2* transcripts. Coloured boxes denotes exons. Green lines denotes introns. A red line followed by 'AAA' in red denotes the poly A tail. Note that 5' and 3' RACE overlaps in exon 1.

Clontech Marathon method

About 100 female and 100 male pupae were used for RNA extraction to obtain 0.8 to 0.9 mg of total RNA to extract about 4 μg of mRNA each. Three μg of each extracted mRNA were used for cDNA strand synthesis and ligated with Marathon Adapter ends from Marathon RACE-Ready cDNA kit (Clontech) according to manufacturers manual. The protocol is as briefly described:

First-strand cDNA synthesis

The following reagents were first combined in a 0.5 ml microcentrifuge tube:

- 3 μg (1-4 μl) RNA sample (poly A)
- 1 μl cDNA Synthesis Primer (10 μM)

Sterile water was added to a final volume of 5 μl, mixed and spun in a microcentrifuge briefly. The tube was then incubated at 70°C for 2 min.,

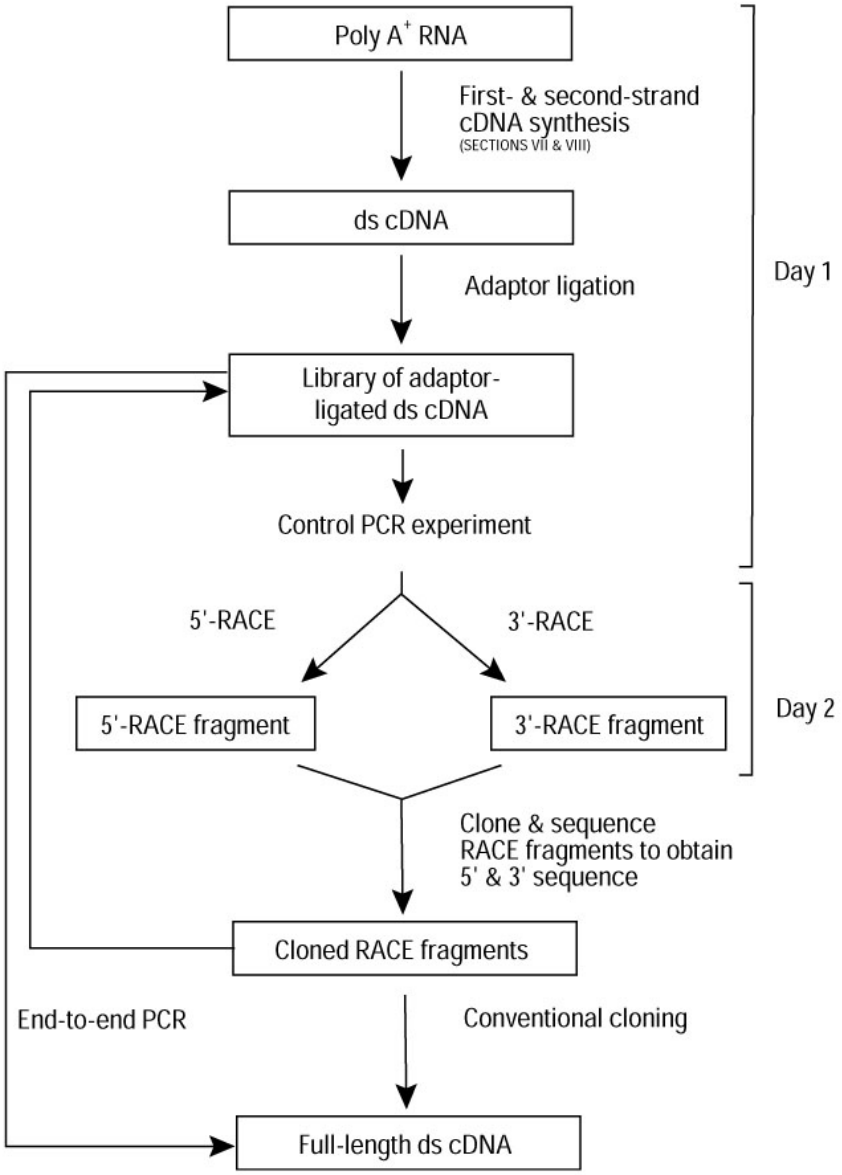

Figure 4.4: Overview of Marathon RACE procedure (adapted from Clontech Marathon RACE protocol). Note that conventional cloning was carried out instead of end-to-end PCR to obtain full-length ds cDNA sequence.

and cooled on ice for 2 min. The tube was spun briefly to collect the contents at the bottom. The following were added to each reaction tube:

- 2 μl 5$\times$ First-Strand Buffer
- 1 μl dNTP Mix (10 mM)

- 1 μl AMV reverse transcriptase (10 units/μl)

Sterile water was added to each reaction to make a final volume of 10 μl. The contents were mixed gently and spun briefly, and incubated at 42°C for 1 hour in an air incubator. The tube was placed on ice to terminate first-strand synthesis.

Second-strand cDNA synthesis

The following were combined in the reaction tube from the first-strand synthesis:

- 10 μl First-strand reaction
- 48.4 μl sterile water
- 1.6 μl dNTP Mix (10 mM)
- 4 μl 20× Second-strand enzyme cocktail

The contents were mixed gently and spun briefly to collect the contents at the bottom. The tube was incubated at 16°C for 1.5 hours. 2 μl (10 units) of T4 DNA polymerase were added and gently mixed. The tube was incubated for 45 min. 4 μl of EDTA Glycogen mix were added to terminate the second strand synthesis. The reaction was purified by phenol-chloroform extraction as described in chapter 2.

One half volume of 4 M Ammonium Acetate and 2.5 volumes of 95% ethanol were added to the purified reaction at room-temperature and mixed by vortexing. The tube was spun immediately at 14,000 rpm (about 10000 × g) at room temperature for 20 min. The supernatant was removed and the pellet was overlaid with 300 μl of 80% ethanol. The tube was spun again for 10 min. The supernatant was removed and the pellet was air-dried for 10 min.

approximately to evaporate residual ethanol. The precipitate was dissolved in 10 μl of sterile water.

Adaptor ligation

The following reagents were combined in a 0.5 ml microcentrifuge tube at room temperature and in the order shown:

- 5 μl ds cDNA (0.5 μg)
- 2 μl Marathon cDNA Adaptor (10 μM)
- 2 μl 5$\times$ DNA ligation buffer
- 1 μl T4 DNA ligase (400 units/μl)

The tube was then mixed by vortex and spun briefly in a microcentifuge and incubated at 16°C overnight. The tube was then heated at 70°C for 5 min. to inactivate the ligase. 1 μl of the adaptor-ligated ds cDNA was diluted with 250 μl of Tricine-EDTA buffer and heated at 94°C for 2 min. to denature the ds cDNA. The tube was cooled on ice for 2 min. and spun briefly to collect the content in the bottom of the tube.

RACE experiments

5' and 3' RACE PCR was performed on the synthesised cDNA according to the same manual but Advantage 1 Polymerase Mix was used instead of Advantage 2 Polymerase Mix (Clontech). Primers for the PCRs are as follows:

RACE 3' PCR primers:

- Forward (Tra-2F; GSP2): 5'- ATGCCAAGACGAAGGTTTCCCG -3'
- Reverse (AP1): 5'- CCATCCTAATACGACTCACTATAGGGC -3'

RACE 5' PCR primers:

- Forward (AP1): 5'- CCATCCTAATACGACTCACTATAGGGC -3'
- Reverse (TRA-2R; GSP1): 5'- CGGTCACCGAATAATCCACTCT -3'

The master mix for all PCR reactions was made as follows:

- 36 μl water
- 5 μl 10$\times$ cDNA PCR reaction buffer
- 1 μl dNTP mix (10 mM)
- 1 μl Advantage 2 polymerase mix (50X)

The tube was mixed by vortex (without introducing bubbles) and briefly spun in a microcentrifuge.

PCR reactions were prepared as shown in table 4.1 for 5' RACE PCR and 3' RACE PCR.

After the PCR, the products were run on a 1% agarose gel to visualise the result. Bands of interests were excised, purified and cloned into Promega

Table 4.1: Setting up 5' and 3' RACE PCR reactions

component	5' RACE sample	3' RACE sample
Diluted adaptor-ligated cDNA	5 μl	5 μl
AP1 Primer (10 μM)	1 μl	1 μl
GSP1 (antisense primer; 10 μM)	1 μl	-
GSP2 (sense primer; 10 μM)	-	1 μl
master mix	43 μl	43 μl
final volume	50 μl	50 μl

pGEM-T Easy vectors. The clones were then screened by PCR using *tra-2* primers as mentioned above, and candidate (*tra-2* positive) clones with inserts were then amplified using M13 primers, and sequenced.

Clontech SMART method

A more advanced method developed by Clontech was used to overcome the difficulties experienced in using Marathon RACE. About 50 pupae per sex, and 80 3^{rd} instar larvae, were used for RNA extraction and then used to extract its mRNA. SMART Race cDNA Amplification Kit was used in this RACE experiment. The protocol is briefly described as follows:

First-strand cDNA synthesis

The following were combined in separate microcentrifuge tubes: For 5' RACE:

- 3 μl RNA sample (1 μg)
- 1 μl 5'-CDS primer
- 1 μl SMART II A oligo

For 3' RACE:

- 4 μl RNA sample (1 μg)
- 1 μl 3'-CDS primer A

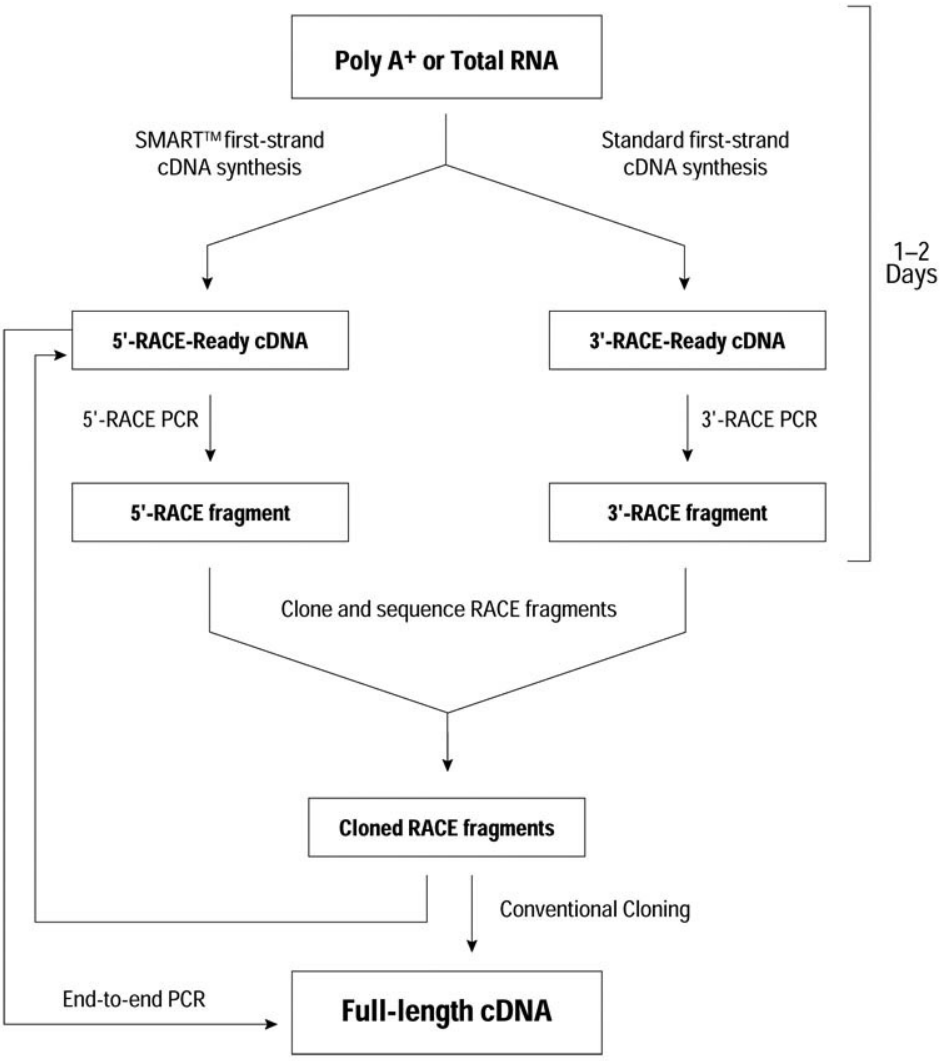

Figure 4.5: Overview of SMART RACE (adapted from Clontech SMART RACE protocol). Note that conventional cloning was carried out instead of end-to-end PCR to obtain full-length cDNA.

The contents were mixed and spun briefly in a microcentrifuge. The tubes were incubated at 70°C for 2 min. and left on ice for 2 min. The tubes were briefly spun to collect the contents at the bottom. The following were added to each reaction tube:

- 2 μl 5× first-strand buffer
- 1 μl DTT (20 mM)

- 1 μl dNTP Mix (10 mM)
- 1 μl MMLV Reverse transcriptase

The contents were mixed and spun briefly, then incubated at 42°C for 1.5 hours in a hot lid thermal cycler. The first-strand reaction product was diluted with 250 μl Tricine-EDTA buffer.

RACE experiment

PCR master mix was prepared according to the list below:

- 34.5 μl sterile water
- 5 μl 10× advantage 2 PCR buffer
- 1 μl dNTP mix (10 mM)
- 1 μl 50× advantage 2 polymerase mix

The PCR for RACE reactions were set up as shown in table 4.2

Table 4.2: Setting up of 5' and 3' RACE PCR reaction

component	5' RACE sample	3' RACE sample
5' RACE ready cDNA	2.5 μl	-
3' RACE ready cDNA	-	2.5 μl
UPM (10×)	5 μl	5 μl
GSP1 (E1R1; 10 μM)	1 μl	-
GSP2 (E1F1/E1F2;10 μM)	-	1 μl
Master mix	41.5 μl	41.5 μl

RACE 3' PCR primers were:

First primer pair (E1F1):

- Forward E1F1, 10 μM): 5'- GCCAAGACGAAGGTTTCCCGTGGGT TCG -3'

- 3' CDS primer A, 12 μM: 5'- AAGCAGTGGTATCAACGCAGAGTA $C(T)_30VN$ -3'

Second primer pair (E1F2):

- (Forward) E1F2, 10 μM): 5'- GGGTTCGGATTCGTGTACTTCCAG GAGC -3'
- 3' CDS primer A, 12 μM: 5'- AAGCAGTGGTATCAACGCAGAGTA $C(T)_30VN$ -3'

RACE 5' PCR primers:

- Smart II Oligo, 12 μM: 5'- AAGCAGTGGTATCAACGCAGAGTACG CGGG -3'
- (Reverse) E1R1, 10 μM: 5'- CCATGTAGACACCGGGCGTGGGCGT ATG -3'
- 5' RACE CDS Primer A, 12 μM: 5'- $(T)_25VN$ - 3'

Note: N = A, C, G, or T; V = A, G, or C.

The PCR was then run in a thermal cycler with the following conditions:

- 5 cycles:
  94°C for 30 sec.
  72°C for 3 min.

- 5 cycles:
  94°C 30 sec.

70°C 30 sec.

72°C 3 min.

- 35 cycles:

  94°C for 30 sec.

  68°C for 30 sec.

  72°C for 3 min.

After the PCR, the products were run on a 1% agarose gel to visualise the result. Bands of interests were excised, purified using QIAquick Gel Extraction Kit (Qiagen), and cloned into Promega pGEM-T Easy vectors (Promega) according to the procedure mentioned in Chapter 2. The clones were screened by PCR using the M13 primers, and candidate clones with the right size bands were then sequenced using the protocol mentioned in Chapter 2.

Confirmation of transcription from RACE using RT-PCR

The RACE sequences were confirmed by RT-PCR using pupae and adult RNA from both sexes. 200 ng of RNA were used per reaction and protocol of RT-PCR was as described in Chapter 2, with the annealing temperature at 58°C. The primers used were as follows:

- Forward (tra-2 Exon 1): 5'- CATTGACAAGGCGATGATTG -3'
- Reverse 1 (tra-2 variant 1): 5'- GACTGAATCAATTGCGATGC -3'
- Reverse 2 (tra-2 variant 2, from clone C2R3): 5'- TTAATGGTTGCC-CAAACGAC -3'

A standard PCR (described in Chapter 2) was then run to amplify the RT-PCR products by diluting the RT-PCR products by 400×, and 1μl of each

sample was used for each reaction. Primers used and annealing temperature were the same as in the RT-PCR above.

4.3 Results

4.3.1 Sex determining genes

4.3.1.1 *dsx* results

The protein sequence of *Ae. aegypti* gene *dsx* was found to be highly similar to *An. gambiae* DSX. The female variant of *dsx* (*dsxF*, 428 amino acids in length), including the common region shared by both sexes, was found in supercontig 1.370 from the Broad Institute database when compared to *An. gambiae* protein sequences. Nucleotide sequence from the proteins that matched was from supercontig 1.370 (positions 948570 948779 (84% similarity to *An. gambiae* DSX, e-value = 1.80799^{e-2}) and 543822 543929 (65% similarity to *D. melanogaster* DSX, e-value = 8.15342^{e-5}) on the minus strand. The nucleotide sequence matching on 2 regions of the supercontig indicated that there were at least 2 separate exons found in *Ae. aegypti dsx*.

4.3.2 RT-PCR of *dsx* regions

The amplified fragment sizes from the RT-PCR were larger than predicted by the BLAST results, suggesting that one section of the sequence of *dsx* was missing from the database sequence (figure 4.6). Explanation of why this region did not show up in the BLAST search result could be that it is poorly conserved, and therefore deviates from the protein sequence of *An. gambiae* used in the search. It is confirmed however that the *dsx* primers designed based on the female specific region that included the female specific region of OD2, only amplify female derived RNA and not that derived from males. Sequencing of the RT-PCR product were not successful. Later, sequences from the BLAST results were compared to submitted *Ae. aegypti* (Mauro et al.,

2006) and indeed it was found that there are regions in *Ae.aegypti dsx* not found in the *An. gambiae* protein sequence. When these regions were taken into account, the size adds up to the RT-PCR fragment size seen. Protein alignment using Geneious with the latest sequences of *D. melanogaster* DSX Isoform B (female-specific protein, NCBI reference sequence: NP_731198.1), *An. gambiae* and *Ae. aegypti* DSX, revealed that the product size obtained from the RT-PCR is correct (figure 4.7).

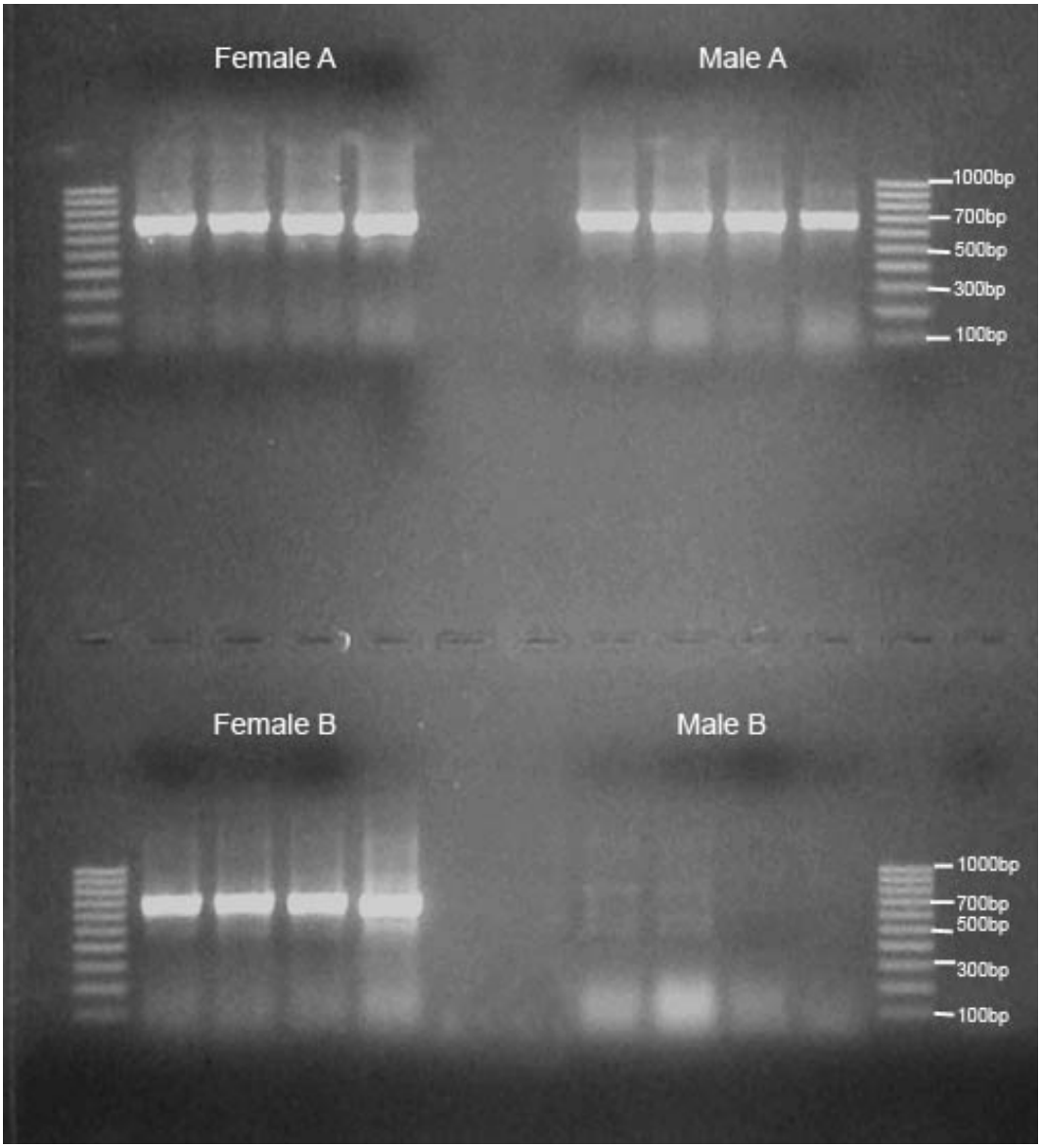

Figure 4.6: RT-PCR of *dsx* region. Upper row (Female and Male A) shows amplification products using primer pair A which amplifies the common region of Oligo Binding Domain for both sexes. Lower row shows amplification products using primer pair B, which amplifies not only the common region of *dsx* but also female specific part of OD2.

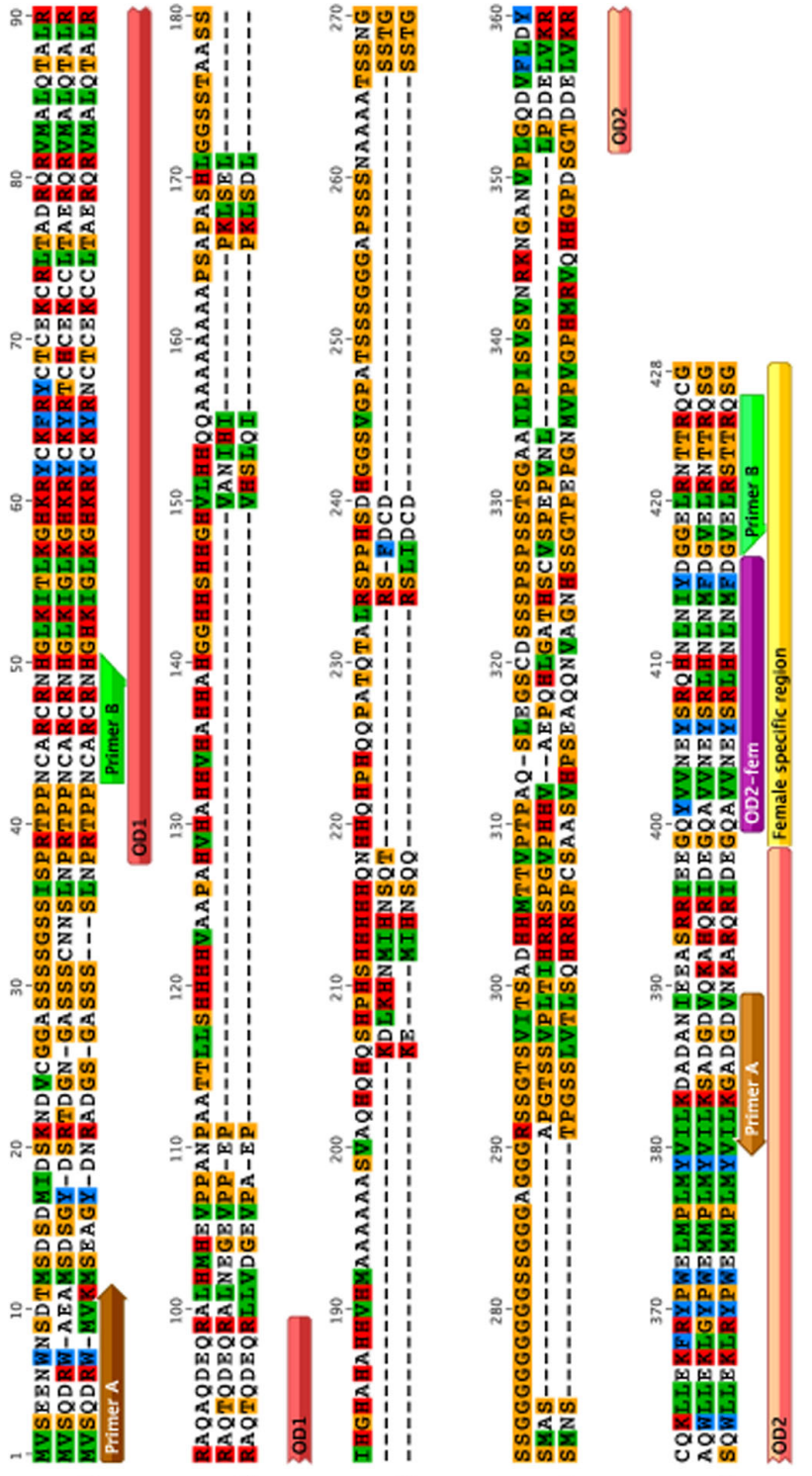

Figure 4.7: Multiple sequence alignment of DSX protein homologs from *D. melanogaster* (sequence 1), *An. gambiae* (sequence 2), and *Ae. aegypti* (sequence 3). The Oligo Binding Domain 1 (OD1) is indicated in dark red bar, common part of OD2 in light red bar, the female-specific region is indicated in yellow bar, and the female specific part of OD2 in purple bar. Brown and green arrows indicate primer pair A and B respectively.

4.3.3 *da* results

D. melanogaster da was annotated with only one transcript, CG5102-RA (flybase gene ID is FBgn0000413). BLASTp search using this transcript against the *Ae. aegypti* database at Vectorbase indicated that highest similarity to gene *D. melanogaster da* was AAEL010226 (genomic location is Supercontig 1.464, regions 381832 to 399092) with 52% similarity (e-value = $1e^{-59}$). BLASTp also showed that AAEL010226 was 73% similar to *An. gambiae da* (e-value= 0.0), AGAP008814. AAEL010226 has 10 exons, spanning a total of 2128 bp in length. Amino acid sequence alignment of these three species showed that the highest conservation of amino acid sequences lies towards the carboxy (C-) terminus (figure 4.8). Prosite search of the sequences indicated that the helix-loop-helix DNA binding motif is situated here which explains the conservation.

Since *da* transcription is not sex-specific in *D. melanogaster*, it was deemed likely that *da* for *Ae. aegypti* is not sex-specifically transcribed as well.

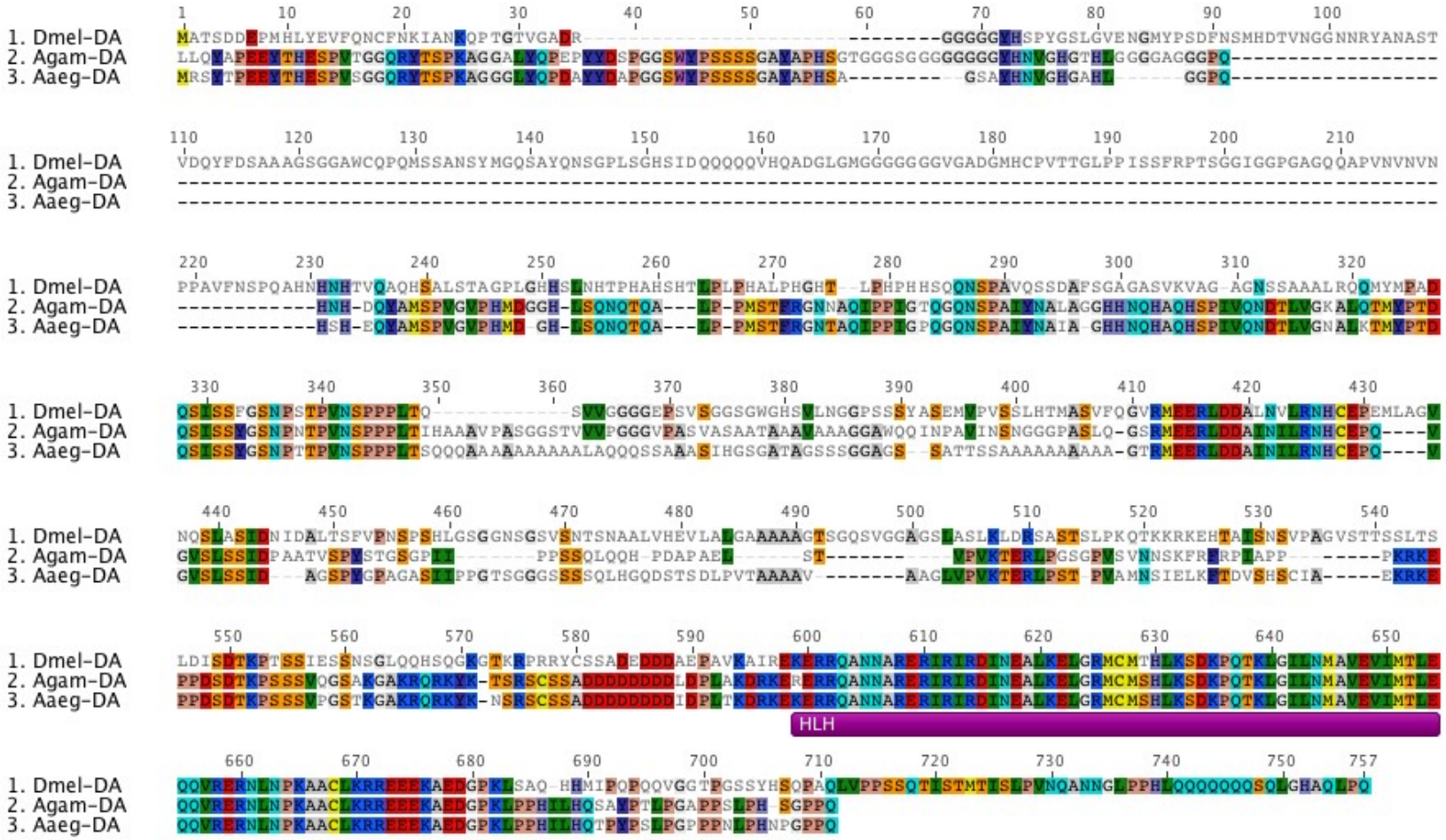

Figure 4.8: Daughterless protein alignment between three species. Species compared are *D. melanogaster* (Dmel), *An. gambiae* (Ang) and *Ae. aegypti* (Aeg). High conservation is seen in the C-termini of the sequences since it encodes the Helix-loop-helix (HLH) DNA binding motif, represented by the purple bar.

4.3.4 *fru* results

D. melanogaster fru (flybase gene ID: FBgn0004652) BLASTp search identified the highest match as AAEL006301 with a 22% similarity (e-value=2.63103^{e-39}) which is situated on Supercontig 1.199 at regions 701,824 to 808,545. BLASTp of AAEL006301 also showed 53% similarity (e-value= 3^{e-93}) to *An. gambiae fru*, AGAP000080. AAEL006301 has a total of 6 exons according to predictions in Vectorbase, spanning 1945 bp in length. Protein sequence alignment of FRU from *D. melanogaster*, *An. gambiae*, and *Ae. aegypti* shows highest similarities in the amine- (N-) and C- terminus of the sequence, due to the conservation of the BTB domain and the zinc finger (ZF) domain respectively (figure 4.9). Unfortunately the male specific region was not found in the *Ae. aegypti fru* homolog.

Using RT-PCR, the first 101 bp of the BTB region were amplified from both male and female cDNA samples at different concentrations (see figure 4.10).

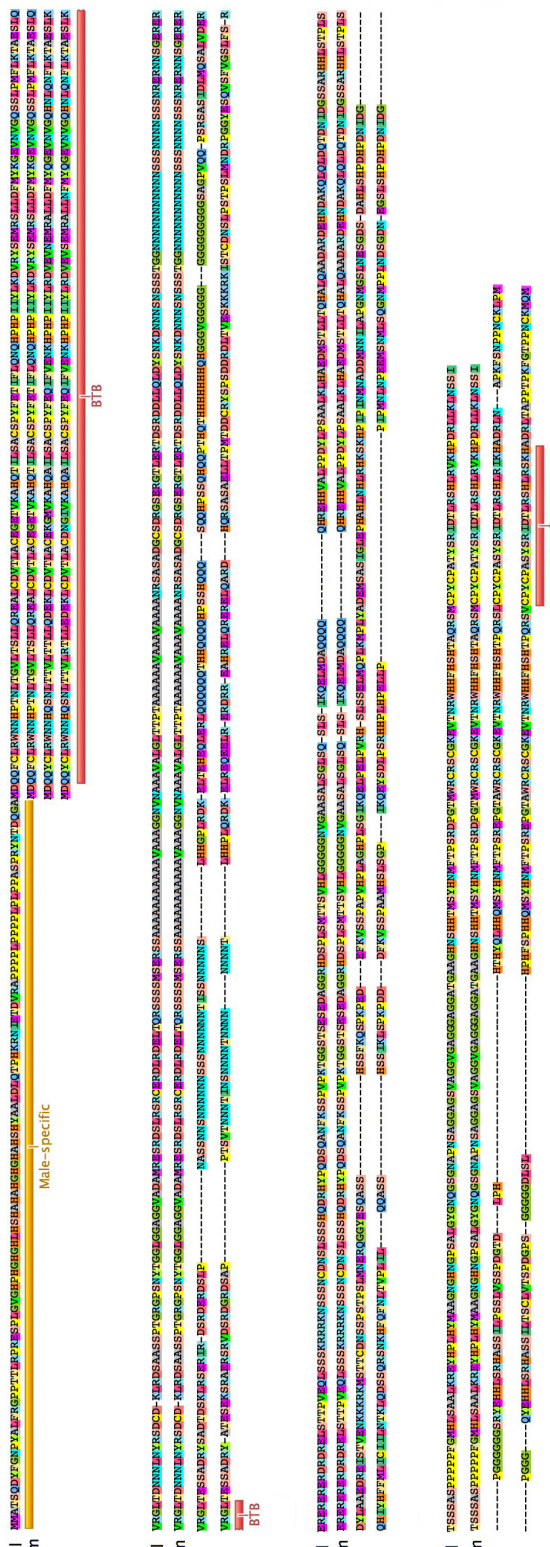

Figure 4.9: Fruitless protein alignment between three species. Species compared are *D. melanogaster* (Dmel) - male and female isoform (sequence 2), *An. gambiae* (sequence 3) and *Ae. aegypti* (sequence 4). High conservation is seen in the N- and C- termini of the sequences since they encode the BTB-ZnF domain, represented by the red bars. Note that difference in *D. melanogaster* FRU is in the N-termini (BTB) region, where the male isoform is longer than the female isoform (yellow bar).

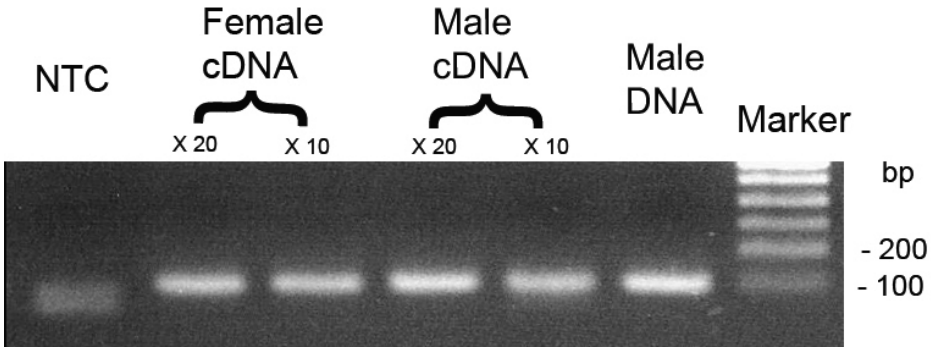

Figure 4.10: RT-PCR of *fruitless* first 101 bp of the BTB region. Female and male pupal cDNA were used to amplify the region at × 20 dilution (25 ng), and × 10 dilution (50 ng). Male DNA (20 ng) was also used as positive control. NTC = no template control.

4.3.5 *tra-2* results

4.3.5.1 *in silico* search for *tra-2*

When BLASTp was performed using *D. melanogaster* TRA-2 protein A (GenBank accession number AAF58232) against *An. gambiae* Ensembl database, the best match was a gene with accession number ENSANGP00000010299 on chromosome 3R (figure 4.11), with 41.4% identity (99 amino acids) and e value of 1.7337^{e-18}.

```
An. gambiae   90    ASRSRIRV----DSPEPSRCLGVFGLSVYTTEPYLNDIFCHFGTVEKSVVIYDAKTRLSR   145
                    +SR R R+    + P+ SRC+GVFGL+  T++  + ++F  +G +E+  ++ DA+T+ SR
D. melanogaster 79  SSRDRERMHKSREHPQASRCIGVFGLNTNTSQHKVRELFNKYGPIERIQMVIDAQTQRSR   138

An. gambiae   146   GFGFVYFKSQAEASIARANCNGLQIHGRRIRVDYSITDQ   184
                    GF F+YF+  ++A  A+ +C+G+++ GRRIRVD+SIT +
D. melanogaster 139 GFCFIYFEKLSDARAAKDSCSGIEVDGRRIRVDFSITQR   177
```

Figure 4.11: Alignment (BLASTp) of *D. melanogaster* TRA-2 protein A to *An. gambiae* protein ENSANGP00000010299. Only partial protein sequence is shown in alignment as there was low conservation between the homologs. Numbers indicate the position of amino acids aligned.

Using the TIGR *Ae. aegypti* EST database, BLASTp search with *D. melanogaster* TRA-2, retrieved a sequence (accession number TC189133)

which was 65% identical within a 40 amino acid stretch.

Since all mosquito databases used were integrated into Vectorbase over the course of the project, the sequences now matched to the gene with accession number AAEL004293 of the *Ae. aegypti* Vectorbase database. AAEL004293 is a hypothetical gene. ENSANGP00000010299 protein was the best match to AAEL004293 , with 101 out to the total predicted length of 194 amino acids being 60.40% similar (e value = 1^{e-35}) when BLASTp was performed (figure 4.12). Multiple sequence alignment of the three proteins mentioned (AAEL004293, ENSANGP00000010299 and AAF58232) showed that only small sections in the sequence shared some similarities in the three species compared (figure 4.13).

```
>AAEL004293-PA
           description|supercont1.113:1220838:1221484:-
           1|gene:AAEL004293
          Length = 194

 Score =  145 bits (365), Expect = 1e-35,   Method: Compositional matrix adjust.
 Identities = 61/101 (60%), Positives = 82/101 (81%), Gaps = 1/101 (0%)

Query: 97  VDSPEPSRCLGVFGLSVYTTEPYLNDIFCHFGTVEKSVVIYDAKTRLSRGFGFVYFKSQA 156
           VD P+ S+CLGVFGLS YT E  L D+F  +GT++K++++YDAKT++SRGFGFVYF+ Q+
Sbjct: 76  VDPPK-SKCLGVFGLSSYTNETSLMDVFAPYGTIDKAMIVYDAKTKVSRGFGFVYFQEQS 134

Query: 157 EASIARANCNGLQIHGRRIRVDYSITDQPHPPTPGVYMGRR 197
            A+ A+  CNG+ +H R IRVDYS+T++PH PTPGVYMG R
Sbjct: 135 AATEAKMQCNGMMLHERTIRVDYSVTERPHTPTPGVYMGAR 175
```

Figure 4.12: Protein alignment (BLASTp) result using Vectorbase BLAST tool. *An. gambiae* TRA-2 protein ENSANGP00000010299 showed highest match to protein AAEL004293 in the *Ae. aegypti* database. Only partial protein sequence is shown in alignment as there was low similarity between the protein sequences. Numbers indicate the position of amino acids aligned.

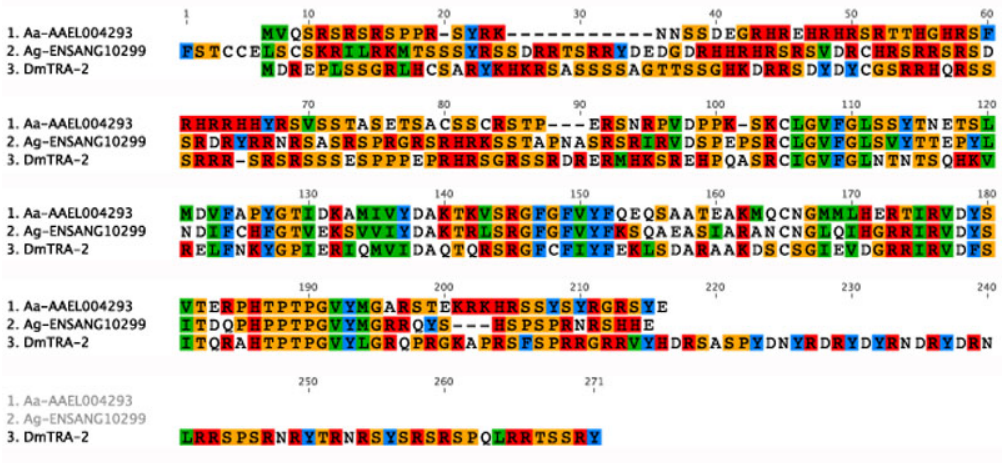

Figure 4.13: Protein alignment of *Ae. aegypti* AAEL004293 (sequence 1), *An. gambiae* ENSANG10299 (sequence 2) and *D. melanogaster* TRA-2 protein isoform A (sequence 3). Alignment shows that the three amino acid sequences share some similarities to each other in certain regions.

4.3.5.2 PCR and RT-PCR assays

PCR on DNA samples and RT-PCR on RNA samples showed a single band of expected size 135 bp in males and females, indicating that the region amplified is a common region for both sexes and in one exon (figure 4.14). This band was also amplified from cDNA of larvae and the pupal stage (figure 4.15). This band was termed exon 1 as no other exons were known at this point in the study.

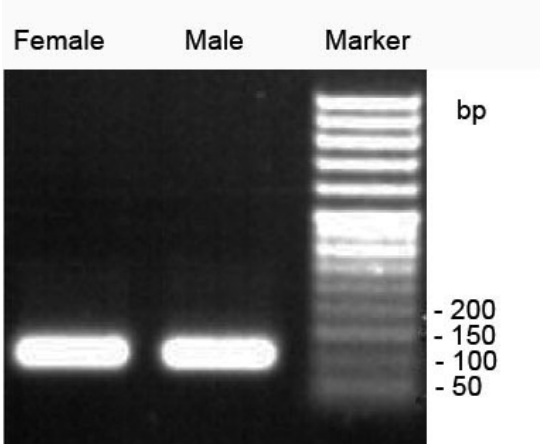

Figure 4.14: *tra-2* RT-PCR of male and female RNA from pupae (200 ng per reaction). Both sexes appear to have the same expected size band of 135 bp in the *tra-2* exon used (exon 1).

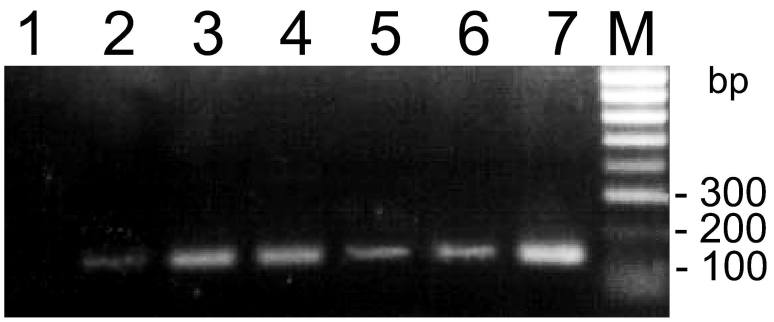

Figure 4.15: *tra-2* RT-PCR of single extracted RNA samples (100 ng) at different development stages. All samples showed amplification of the expected 135 bp sized DNA from exon 1. Lane: 1= negative control (water), 2= 1^{st} instar larva, 3= 2^{nd} instar larva, 4= 3^{rd} instar larva, 5= 4^{th} instars larva, 6= female pupa, 7= male pupa, M= size ladder marker.

4.3.6 Northern Blot

Northern Blot was unsuccessful in detecting *tra-2*, even when the x-ray film was exposed to the chemiluminescent membrane for 24 hours. Degradation of RNA was ruled out as control and RNA ladder was visible on the x-ray film. Unsuccessful labelling of the probe was also ruled out as a control RNA probe using Actin 1, gave a very strong positive signal.

4.3.7 5' RACE

Marathon 5' RACE gave a consistent fragment of about 200 bp in both male and female samples. The sequencing of this fragment after cloning produced sequences that were of low quality and those that were legible were not matched to the primers used. This meant that mispriming and thus non-specific amplification occurred in the Marathon 5' RACE. SMART 5' RACE however produced a fragment of about 650 bp in both male and female samples which was cloned and sequenced although "smearing" could be seen in the PCR products (figure 4.16). A total of 48 clones were sequenced (see Appendix for sequences of clones) and 44 were confirmed to be *tra-2*. Out of 44 sequences, 28 were from female samples and 16 were from male. Via BLASTn, the sequences from the RACE were matched to two regions (1220842 - 1221411 and 1221480 - 1221544) on the minus strand of 1.113 Supercontig.

Alignment of the 400 bp long sequence obtained from 5' RACE experiment with the relevant part of the genomic DNA revealed an intron of 62 bp in between this (exon 1) and a newly identified exon upstream (figure A-6 in Appendix). The new exon range from 45 to 70 bp long, while the majority was 50 bp in both males and females.

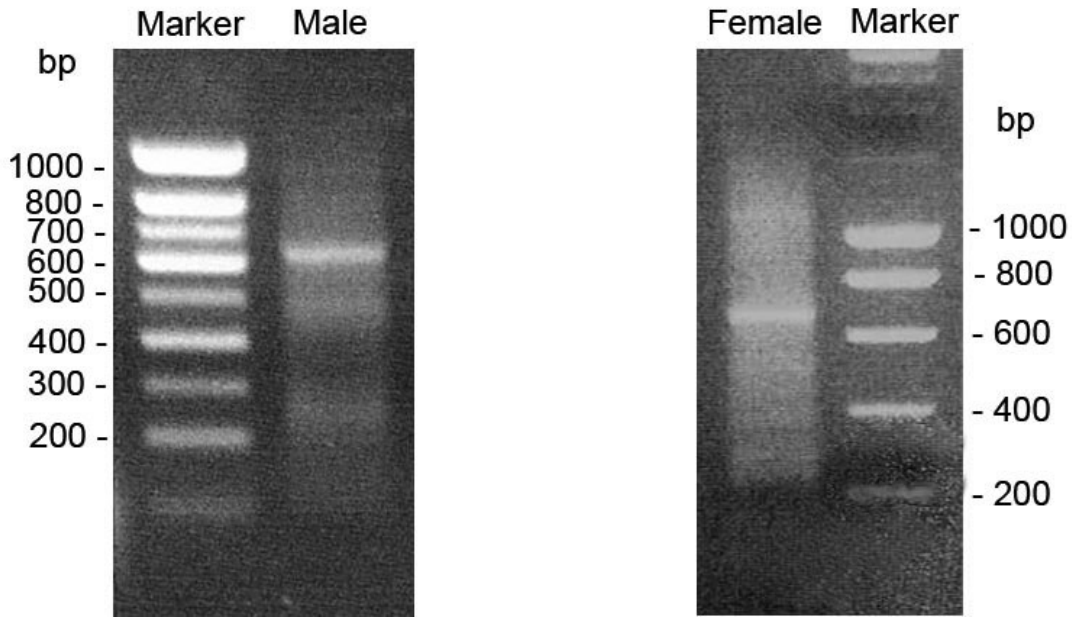

Figure 4.16: *tra-2* 5' RACE using Clontech's SMART cDNA Amplification on *Aedes aegypti* as visualised on a 1% agarose gel. Both female RACE (right) and male RACE (left) result in the amplification of about 650 bp cDNA fragment.

4.3.7.1 Analysis of 5' RACE

Only a partial sequence of exon 1 was known at the start of the study. The annotation of exons for AAEL004293 in Vectorbase was later published which consisted of 2 exons. The first exon in the annotation contained a start codon (ATG), and the second exon was 582 bp in length. The intron in between the two exons was 62 bp long. The partial sequence of exon 1 matched the second exon of the Vectorbase annotation of AAEL004293. The length of the intron from the 5' RACE was of the same site as the AAEL004293 intron. When the sequences 'Vectorbase exon 1' were translated in 3 frameshifts, they did not make a complete open reading frame. However, it was found that it contained a Kozak sequence (CAAAATGG), required for the initiation of translation, where ATG in the sequence was the start of the first exon of AAEL004293 by Vectorbase. The conclusion is that the fragment of sequence labelled exon 1

in the Vectorbase annotation is not an exon, but part of the 5' UTR where it contains a Kozak sequence pattern.

4.3.8 3' RACE

The sequence obtained via 3' Marathon RACE can only be read in one direction (5' to 3') due to the poly A tail at the end of the mRNA. To get the consensus sequence, sequencing in the same direction (5' to 3') was done several times. A total of 65 clones were sequenced, of which only 54 were of acceptable quality to be read. Out of the 54, 22 sequences were derived from the female mRNA, while 32 sequences were from male mRNA. Marathon RACE 3' produced different sized products, which could indicate the different size of *tra-2* transcripts in males and females (figure 4.17). Via BLASTn, the sequences from the RACE products were matched to the regions 1194263 - 1194397 and 1220842 - 1221076 on the minus strand of 1.113 supercontig (see table A-4 for full BLASTn result). This indicates that the transcribed RNA was spliced, with an intron of more than 26 kbp, suggesting the number of exons in this gene is at least two (not including the previously annotated first exon from Vectorbase). Sequences were aligned using Clustal W with the Geneious Software to regions of Supercontig 1.113, and this showed 2 different variants - 1 specific to males (variant G, 22 clones), and 1 common to both sexes (variant E - 22 clones in females, 10 clones in males). The splicing of exon 1 to exon 2 of the variants seem to be common. The two variants also seem to only differ in length, where the shorter sequences were simply truncated (variant G, figure A-8).The other sequence common in both sexes (variant E) were found to be about 200 bp longer than variant G.

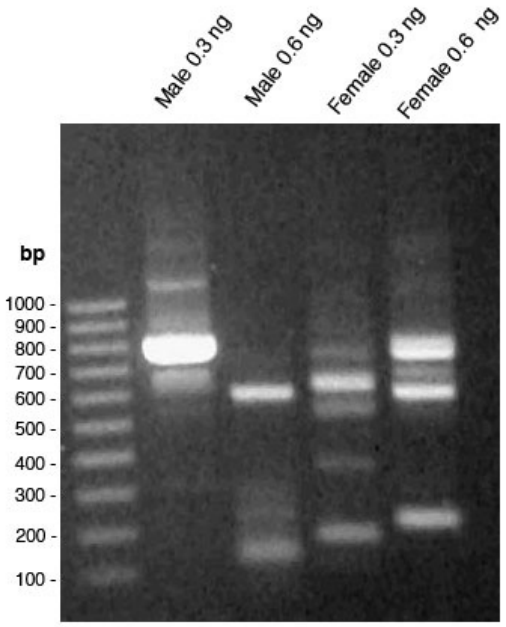

Figure 4.17: *Ae. aegypti tra-2* 3' RACE using Clontech's Marathon cDNA Amplification as visualised on a 1% agarose gel. The largest product (about 750 bp sized band in male 0.3 ng and female 0.6 ng samples) and 600 bp sized product in 0.6 ng male sample were variants of *tra-2*.

As in Marathon RACE, sequences obtained from 3' SMART RACE can only be read in one direction (5' to 3') due to the poly A tail at the end of the mRNA. Consensus sequence was achieved by sequencing a clone in the same direction several times. Via 3' SMART RACE, more variants were identified from both sexes than from Marathon RACE. SMART RACE produced different sized products in males and females (figure 4.18). SMART RACE using female sample of produce a product of about 650 bp, while RACE using male cDNA (right) gives a band of about 500 bp for E1F1 primer pair; and just over 500 bp for E1F2 primer pair.

Sixty candidate clones derived from SMART RACE of the female mRNA were sequenced. Multiple nucleotide sequence alignment showed that there were 5 variants (variant B, C, D, E and H). The span of the intron between exon 2 and 3, which was over 26 kbp long was the same as found in Marathon RACE (refer to figure A-9 in Appendix for the sequence alignment). However, sequence of exon 2 was unlike the one that was found via Marathon RACE. The exon 2 region 1220842-1221075 (233 bp long) in Supercontig 1.113 that was not found in the sequence derived from SMART RACE (figure A-10 in

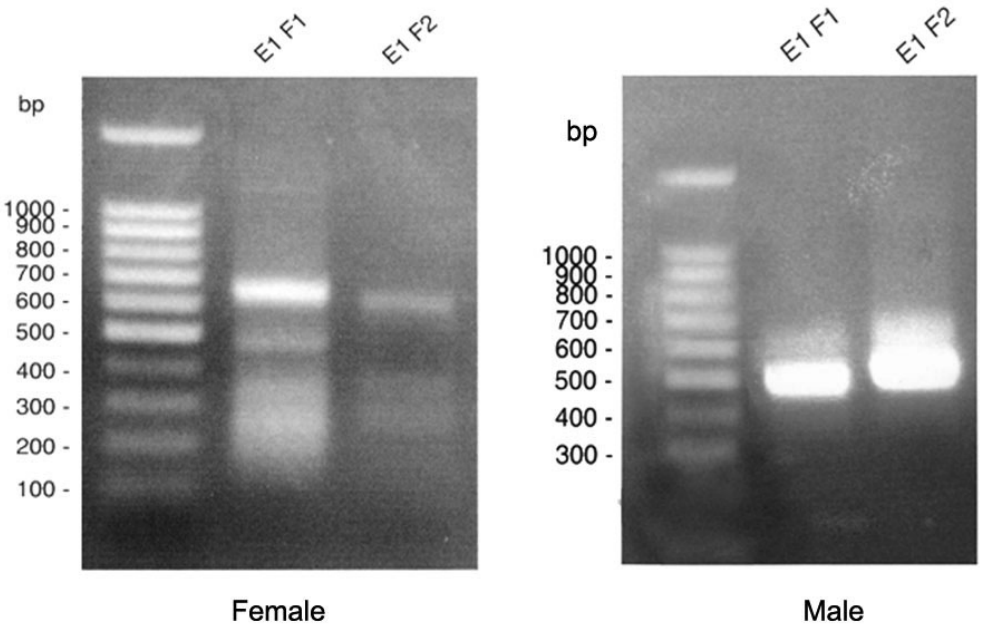

Figure 4.18: *Ae. aegypti tra-2* 3' RACE using Clontech's SMART cDNA Amplification as visualised on a 1% agarose gel. Two RACE experiment were run using two primer pairs - E1F1 and E1F2 (see methods for primer details). Female RACE (left) shows a band of about 650 bp in E1F1 and E1F2 samples. Male RACE (right) shows a band of about 500 bp for E1F1 primer pair sample; and a slightly larger band (over 500 bp) for E1F2 sample.

Appendix). Therefore, only sequences from Marathon RACE were able to detect this region. This indicates that the region is another splice region, raising the total number of exons to 4 (diagram 4.19). The sequences of 5 clones were different, as they have shorter exon 1 region than the rest of the clones and the exons following were spliced to exclude several regions in exon 2 and 3. This sequence hence forms variant H.

A total of 45 candidate clones derived from the 3' SMART RACE experiment using male mRNA were sequenced. From these clones, 6 variants were identified (variant A, D, E, F, and G) after aligning their sequences using the Geneious Clustal W software. The span of exon 1 to exon 2 was the same as in Marathon RACE where an intron of over 26 kbp separates the two exons (figure A-12). The exon 3 pattern was similarly found to have a section in the

middle that was not transcribed as seen in the female SMART RACE that may indicate an intron (figure A-13). However, there were variations in length in exons 2 and 3, where MalD5-1 is about 108 bp longer towards the end of exon 3, while MalC6-1R seemed to be the most truncated (ending before exon 4), similar to Clone B1M from Marathon RACE, but the poly A tail is short and therefore some ambiguity exits as to whether the sequence is truly a truncated sequence (figure A-13). For the sequences of clones, refer to the Appendix.

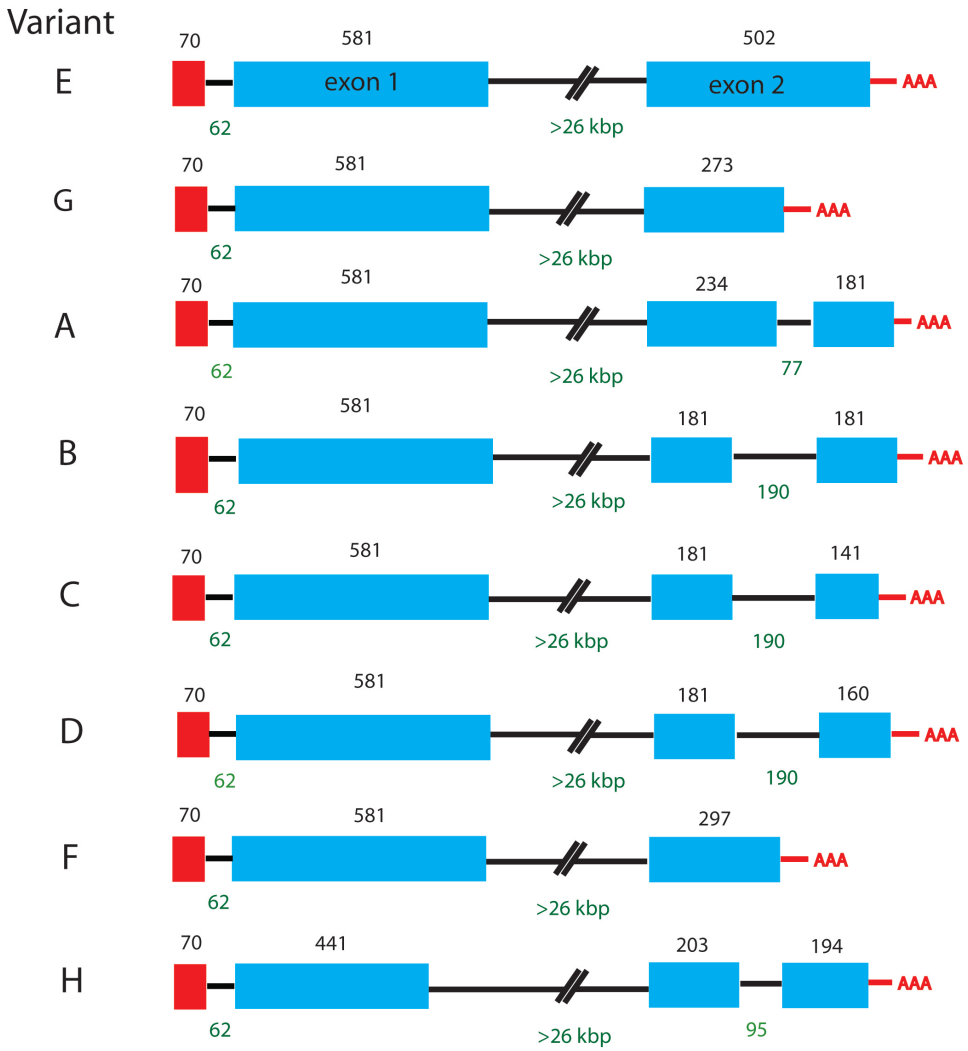

Figure 4.19: Schematic diagram of RACE result. Coloured boxes denotes exons. Red coloured boxes are sequences upstream of exon 1. Black bars denotes introns. Red bar with A letters denotes the poly A tail. Black numbers indicate exon length in bp. Green numbers indicate intron length in bp. From the Marathon RACE, only variants E and D were isolated, which differs in 3' end length. SMART RACE however showed that in addition to the differences in 3' length, sections in the exons could be spliced out to produce different variations of *tra-2* (variants A, B, C, D, E, F, G and H).

Certain variants were more abundant than others judging by the number of sequenced clones obtained (table 4.3). This may reflect the differences of transcription levels in each variants. Variant D which was found in both males and females seemed to be have the highest number of clones, followed by variant E (also found in both males and females). Variant G in males were

also abundant. Variant A that was found only in males seemed to be the least abundant, with only 2 clones sequenced.

Table 4.3: List of variants identified by 3' RACE sequences. Each variant was noted for type of RACE experiment, sequence length (minus exon upstream of exon 1) and gender. Numbers in brackets indicate number of clones detected.

Variant	Sequence length (bp)	Marathon RACE	Smart RACE
A	1059	-	Male (2)
B	943	-	Female (15)
C	903	-	Female (17)
D	922	-	Male (30), Female (21)
E	1083	Female (22), Male (10)	-
F	878	-	Male (7)
G	854	Male (22)	Male (6)
H	838	-	Female (7)

As for the 3^{rd} instar larvae 3' SMART RACE, 25 clones were sequenced. Sequence of *tra-2* variant D was detected in larvae samples. No variations were seen in the splice sites (see figure A-15) except for the ends of the mRNA, where only clone LG6-3LF was truncated at the 3' end (figure A-16 and A-5) by 22 bp.

4.3.8.1 Analysis of 3' RACE variants

When comparing the two 3' RACE experiments, it was found that several of the clones were similar and were thus regarded as the same variant. One particular difference was seen in the end of exon 4 region. Marathon RACE female clone C2R3 has several nucleotides that are different from the rest of the RACE samples. To investigate if this sequence was polymorphic and not an erroneous sequence, primers were designed for RT-PCR (see above for the methodology and below for results).

RT-PCR of RACE derived sequence to confirm region is transcribed

The first RT-PCR used primer pair "variant 1" amplifying *tra-2* that spans from mid exon 1 to exon 2. The results showed that products of different sizes and intensities were expressed in different sexes and the two final stages of development (figure 4.20). The sequences of 600, 550 and 450 bp PCR product fragments matched to Supercontig 1.113.

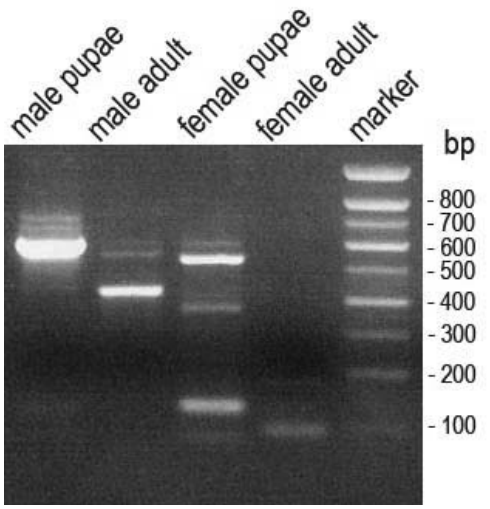

Figure 4.20: RT-PCR of tra-2 with first type of 3' variant containing exon E5. Female and male RNA (200 ng per reaction) in pupal and adult stages showed different sized transcription products. Product sized 600 bp and 450 bp in the male pupae and male adult samples matched to *tra-2* when sequenced. Product size 600 bp from the female pupae sample was also matched to *tra-2* when sequenced.

The RT-PCR using primer pair "variant 2" was performed to detect the nucleotide changes in the sequence of Variant A . The results showed both male pupal and adult RNA yielded products of about 550 bp, which matched with Supercontig 1.113 (figure 4.21). Again, this variant was not found to contain exon 3. This also suggests that the minor polymorphism seen in the sequence of clone C2R3 in Marathon RACE may have significance in sexual differentiation by different amount of expression. C2R3 sequence was thought to originate from a female pupal RNA sample. Yet, RT-PCR amplified this variant in males and not in females. Therefore, an error such as mis-labelling might have occurred when C2R3 was sequenced, and that the sequence was actually that of a male pupal RNA sample. It may also be possible that amount of this variant is so low in female samples that RT-PCR could not detect it.

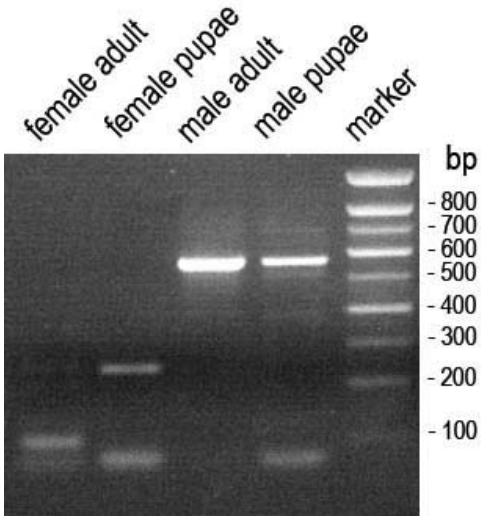

Figure 4.21: RT-PCR of tra-2 with the second type of 3' variation (variant A). Female and male RNA (200 ng per reaction) in pupal and adult stages showed different sized transcription products, but the male samples had higher intensity bands. The male RT-PCR products sequenced were matched to *tra-2*. Female products, when sequenced, were not matched to *tra-2*.

Alignment of the sequences (figure 4.22) showed that the first type of variant (variant B, which contains exon 5), was present in the pupal stage of both male and female samples as well as the adult male sample. Variant B contains all the exons except exon 3. However, another new variant was also found in the adult male sample, i.e. the 450 bp fragment size product. This variant (Variant I) does not contain exons 2 and 3.

The nucleotide sequence region at the end of exon 4 was translated using Geneious software. Comparison showed a change in two amino acids between variant A and other variants:

Sequence of other variants: TGCAACAATAAA

Amino acid translated: **G N N K**

Sequence of Variant A: GGCAACCATTAA

Amino acid translated: **G N H Stop**

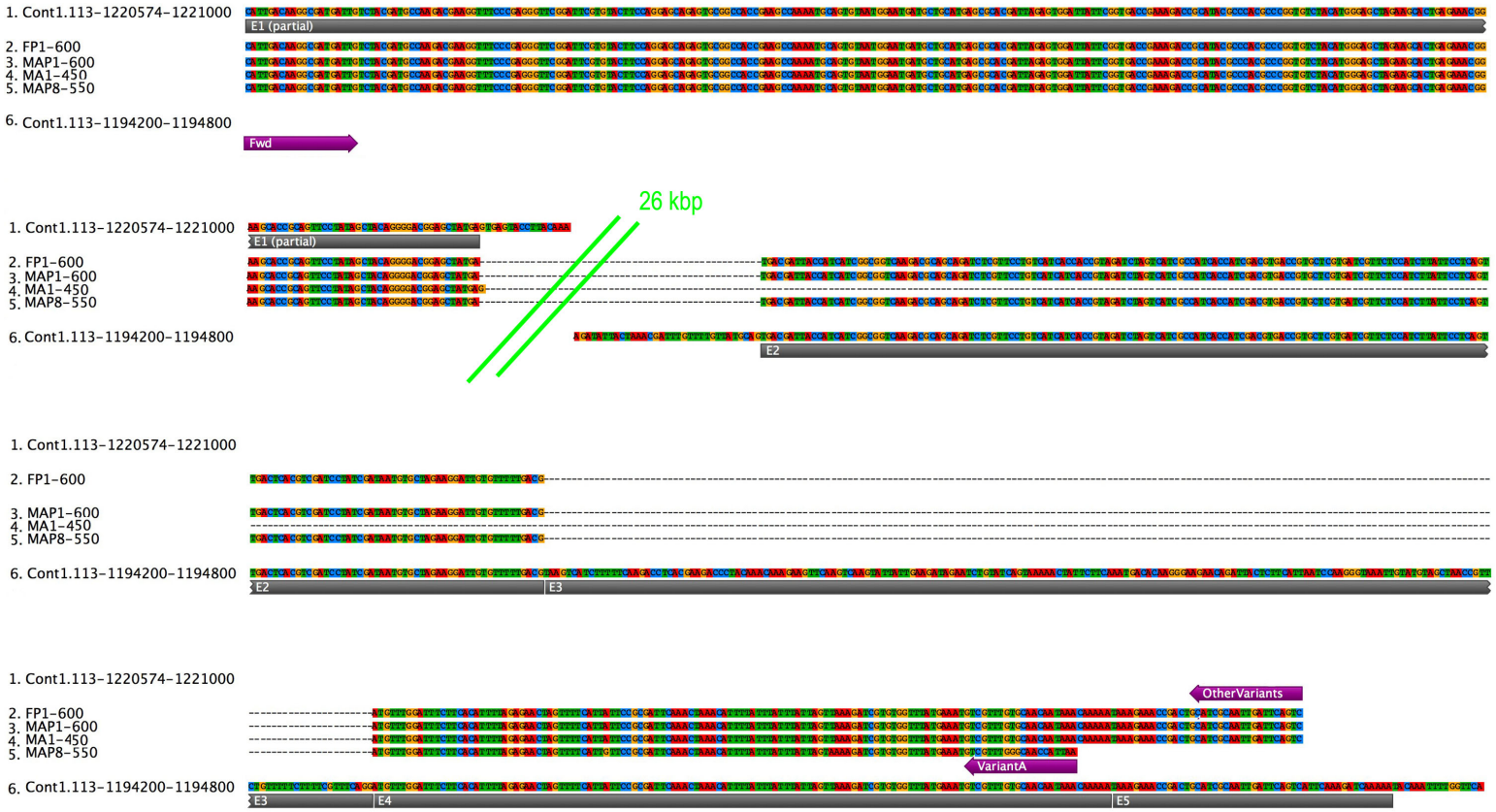

Figure 4.22: Multiple alignment sequence of OneStep RT-PCR of tra-2 products. Sequence 1 and 6 denotes the SuperContig 1.113 region. Sequence 2, 3, and 4 are samples using the first variant reverse primer for RT-PCR. Sequence 2 = the female pupal RNA sample (product size of 600 bp). Sequence 3 = representations male adult and pupal RNA samples (600 bp). Sequence 4 = male adult RNA sample (450 bp). Sequence 5 = representation of male RNA samples from the variant A primer RT-PCR. Arrow marked OtherVariants indicate the reverse primer sequence with the first type of variation used in the the RT-PCR. Arrow marked Variant A is the sequence in the reverse primer that was used to ascertain the nucleotide variation seen in C2R3 (Note the nucleotide differences). Grey bars indicate the exon regions. Double slanted green line indicates a gap in alignment of 26 kbp. Sequence 4 does not have regions exon 2 and 3 (marked E2 and E3).

4.3.9 Structure of *tra-2* and its variants

The size of exon 1 was estimated by using Clustal W in the Geneious Software to assemble to sequences obtained from 5' and 3' RACE. The total length of exon 1 is 581 bp. The longest length of 5' UTR was 70 bp. But as it is a 5' UTR, the length can be variable, as seen in some of the sequencing results. From the results, it was deduced that *tra-2* has a total of 4 exons: E1, E2, E3, and E4. The final transcribed region (E5) is probably the 3' UTR since ORF prediction does not extend into this region in any of its variants (see figure 4.23 for diagram of *tra-2* and its variants). From the RACE and RT-PCR of RACE-derived sequences, it was deduced that at least 9 variants of *tra-2* were transcribed. Variants A, F, G, and I were only detected in male samples. Variants B, C and H were only in female samples. Variants D and E were found in both sexes. The difference in nucleotide sequences in exon E4 seen in Variant A is not taken into account for the diagram of *tra-2* structure for simplicity.

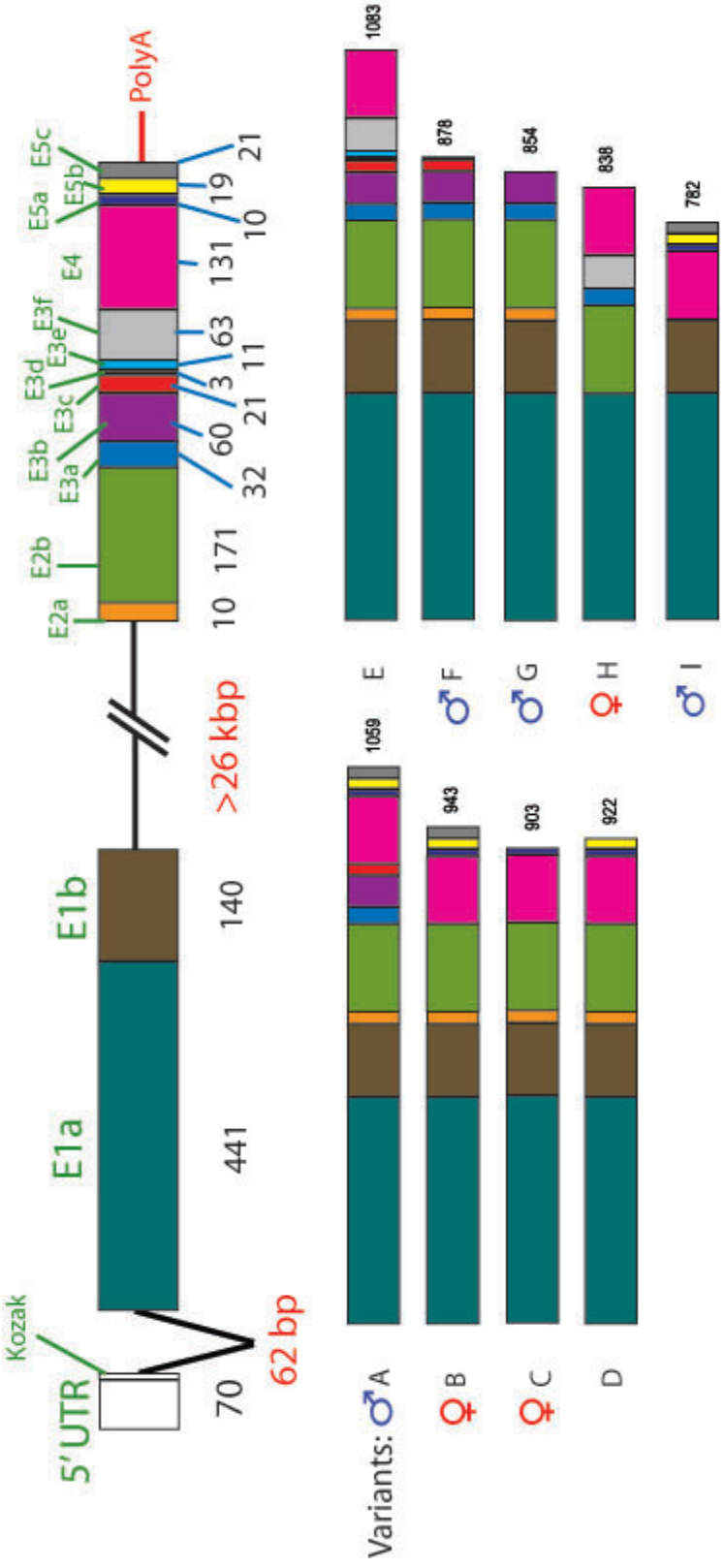

Figure 4.23: Structure of *tra-2* mRNA and its variants as deduced from comparison of genomic and RACE sequence data. Top diagram indicates the whole of *tra-2* structure. Boxes denote transcribed regions in Supercontig 1.113. Black lines denotes untranscribed regions (intron). Black numbers below the first diagram indicates the size of each region in bp. PolyA tail of the structure is indicated by the red line and the word 'PolyA'. Kozak sequence labelled is sequence CAAAATGG. Variant A, F, G and I are only seen in male samples. Variant B, C and H are only seen in female samples. Variant D and E are found in both sexes. Note that Variant I is deduced from RT-PCR of male adult sample. Each variant size is indicated in bp by the black numbers next to the variant diagram.

To understand its potential function, *tra-2* was translated to protein in all 3 possible frames. Using Prosite (http://www.expasy.ch/prosite/), the amino acid sequences were searched for any protein domains, families and functional sites. It was found that TRA-2 contains an RNA recognition motif (RRM) at 240 bp to 475 bp from the start of the exon 1. RRM is involved in most post-transcriptional gene expression processes. Since RRM region is across E1a and E1b, this suggests that the variant not containing the full exon 1 sequence (i.e. clone FemE7-BF, variant H) has an incomplete structure of RRM.

ORF prediction was also made using Geneious (figure 4.24). The variants with E1a, E1b, E2a and E2b have ORFs extending until the E2b region. Variant I which does not contain E2 and E3 has an ORF that extends until the E4 region. Variant H has a short ORF, terminating early in E1b. From the results, the differences in some of the variants were attributed to alternative splicing. The other reason for differences in mRNA length is due to alternative polyadenylation, where 3' UTR plays a part in determining mRNA variant lengths of *tra-2*. Several sites, AUUUA (ATTTA), which are thought to increase instability of mRNA (Malter, 1989; Stoecklin et al., 1994), are located in exon 4 (E4 in diagram 4.23). Signals containing sequence AAUAAA (AATAAA) between 10 and 35 nucleotides upstream of the actual cleavage and polyadenylation site is a consensus for polyadenylation (Lutz, 2008), and is also found in E4 and E5.

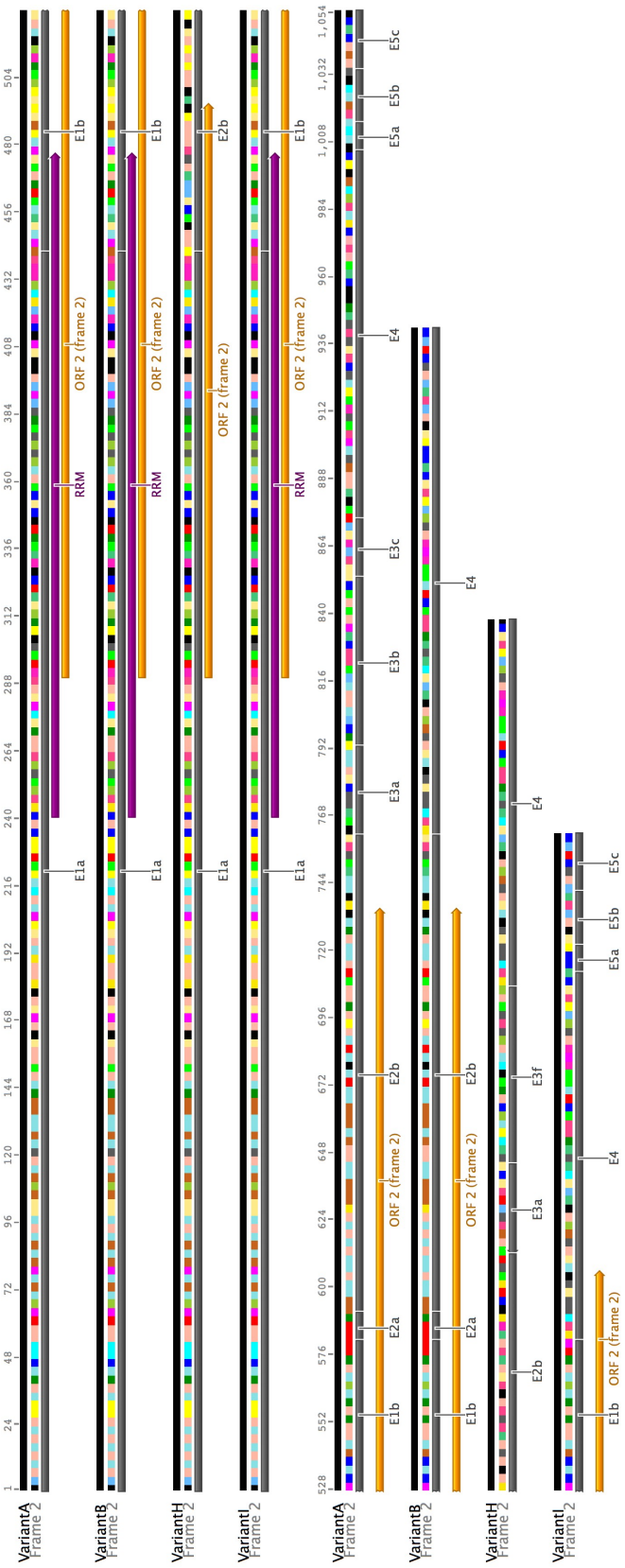

Figure 4.24: Predicted ORF and RRM position in *tra-2* variants. Grey bars indicate transcribed regions. Orange bars indicate ORFs. RRM is indicated by purple bars. Variant A and B which represents the variants with E1a, E1b, E2a and E2b, shows ORF extends to only E2b. Variant H shows a short ORF that only extends to E1b. Variant I has an ORF extending into E4 due to absence of E2 and E3 regions.

4.3.9.1 Potential transcription factor-binding motifs for corresponding elements upstream of *tra-2* exons

The genomic DNA sequence of *Ae. aegypti* spanning 10kb upstream of the *tra-2* exon 1 sequence was obtained from VectorBase, and analysed via ProScan promoter predictor (website: http://thr.cit.nih.gov/molbio/proscan) to search for putative promoter regions based on scoring homologies with eukaryotic Pol II (RNA polymerase II) promoter sequences. A number of potential promoters within the region was searched and a list was generated (Prestridge, 1995). From the list, only those on the minus strand were regarded as potential promoters since *tra-2* sequence used is on the minus strand. Not all promoters found were relevant to insects, and hence the Transcription Factor Database (TFD) outputs generated by ProScan were viewed to find insect-relevant promoters (http://www.ifti.org/ootfd/).

The predicted promoter region was 9141 to 8890 bp upstream of exon 1. It was found that a TATA-like element that is relevant to insects is located in between 8956 bp to 8949 bp (TTTATAA) upstream of exon 1. This element is known as transcription factor II (TFIID). The transcription start site is predicted at 8927 bp upstream of exon 1 (see Appendix for full sequence of region upstream of exon 1 to the predicted promoter region). Promoters predicted are listed in tables 4.4 and 4.5. In the TFD database, all promoters listed were from mammalian systems, except for TFIID (TFD No. S01540), which was described as a "eukaryotic promoter".

Another promoter region was also predicted 16862 to 16611 bp upstream of exon 2. The same TATA-like elements as the ones upstream of exon 1 were also found at 16657 bp to 16652 bp (TATAA) upstream of exon 3, with the estimated transcription start site at 16686 bp upstream of the exon. Promoters

predicted are listed in table 4.5. All promoters were again from mammalian systems, except for TFIID promoter with TFD No. S01540, which was described as an "eukaryotic promoter" and not solely a mammalian promoter. These promoters described above are very likely to be the promoters for *tra-2*. Vectorbase gene map shows that the position of the promoters, as well as the new exons found in 3' RACE experiment does not coincide with any other gene (figure 4.25).

Table 4.4: Predicted promoter sites of *tra-2* upstream of exon 1 using Proscan (Prestridge, 1995). "TFD No." denotes the transcription factor database number. "Strand" indicates the direction of the DNA strand, and "Weight" indicates the proscan relative weighting based upon that particular signal's ability to discriminate promoter from non-promoter sequences, and is based upon the relative frequency with which that signal is found in promoter versus non-promoter sequences (e.g. if SP1 promoter has a weight of 6, it means that SP1 is found about 6 times more frequently in promoter than in non-promoter sequences.)

Promoter	TFD No.	Strand	Weight
CREB	S00144	minus	1.912000
HSV IE repeat	S01565	plus	1.363000
EivF/CREB	S00104	plus	1.564000
JCV repeated sequence	S01193	minus	1.427000
GArC3	S02066	minus	7.376000
NRE Box2 CS	S01555	plus	1.229000
GArC1	S02065	plus	7.376000
CTF	S00780	minus	1.704000
CP1	S00098	plus	1.388000
TFIID	S00615	minus	2.92000
TFIID	S01540	minus	1.971000
TFIID	S00087	minus	2.618000

Table 4.5: Predicted promoter sites of *tra-2* upstream of exon 3 using Proscan (Prestridge, 1995). The "TFD No." denotes the transcription factor database number. "Strand" indicates the direction of the DNA strand, and "Weight" indicates the proscan relative weighting based upon that particular signal's ability to discriminate promoter from non-promoter sequences, and is based upon the relative frequency with which that signal is found in promoter versus non-promoter sequences (e.g. if SP1 promoter has a weight of 6, it means that SP1 is found about 6 times more frequently in promoter than in non-promoter sequences.)

Promoter	TFD No.	Strand	Weight
CTF	S00780	plus	1.448000
element II rs-3	S01506	plus	1.087000
TFIID	S00087	minus	2.618000
TFIID	S01540	minus	1.971000
HSV IE repeat	S01565	plus	1.363000
GCF	S01964	minus	2.361000
UCE.2	S00437	plus	1.216000
AP-2	S00180	minus	1.863000

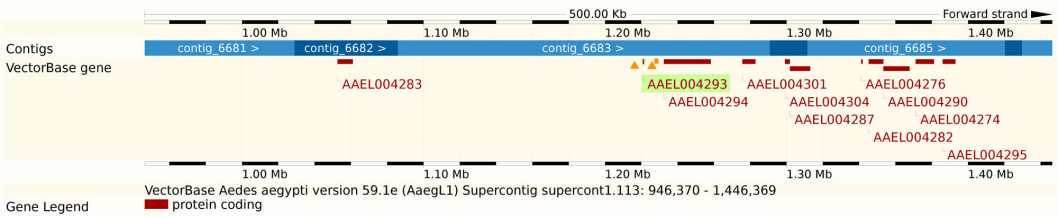

Figure 4.25: *tra-2* (AAEL004293) position in a Vectorbase-generated gene map. Orange bar denotes the rest of *tra-2* as gathered by 3' RACE experiment. Orange triangles mark out the positions of where the possible promoters lie. No other genes were seen overlapping the *tra-2* gene position.

4.4 Discussion

4.4.1 Genes *dsx, da* and *fru*

The gene doublesex is a highly conserved gene across species as expected and has sex-specific regions. The high conservation of protein sequences in OD1 and OD2 strongly suggests that the gene found in *Ae. aegypti* is orthologous to *An. gambiae* and *D. melanogaster dsx*. *Doublesex* could be used as a target for sex ratio manipulation. Silencing of sex-specific *dsx* in young adults may induce sterility in individuals such as reported in *D. melanogaster* due to disruption of the transcription of yolk protein genes (Chen et al., 2007b).

The genes *daughterless* and *fruitless* are highly conserved across the species examined. There is a male-specific sequence in *D. melanogaster fru*, but this was not found in *Ae. aegypti*. This is probably because the male-specific sequence is not conserved and that the region is absent in *Ae. aegypti*, or the *Ae. aegypti* database is not assembled completely. More primers would need to be designed in order to ensure that sex-specific regions would not be missed, but based on the result of the protein alignment, the probability of finding sex-specific sequences is low. A method to find this region is to design degenerate primers to amplify the putative male-specific region in *Ae. aegypti fru*. Since no information is currently published about the genes *fru* and *da* in *Ae. aegypti*, conservation of sex determination function of these two is not yet proven.

4.4.2 *tra-2* homolog search

Searches *in silico* provided a candidate *tra-2* ortholog in *Ae. aegypti* though the conservation of the gene was relatively low. This may indicate that sex

determination genes that are higher in the hierarchy than *tra-2* would be difficult to identify as conservation is likely to be even less, if the theory that the sex determination cascade at the upper tiers are less conserved than the lower tiers (Wilkins, 1995). *Drosophila melanogaster* genome databases proved to be valuable in finding sex determining genes due to the very extensive research done in the fly, making it easier to locate the homologs in *Ae. aegypti* by *in silico* methods. *Anopheles gambiae* is also more closely related to *Ae. aegypti* than *D. melanogaster* and may give better results in searching for the homologs wanted. Since *tra-2* is higher up in the sex determining cascade, it is a good candidate for sex ratio manipulation.

4.4.2.1 Northern Blot

The failure to detect *tra-2* could be explained by the possibility that the RNA probe was degraded prematurely. Therefore, it is possible that this experiment might work better if a more stable DNA probe was used instead of RNA. Another possibility is that the amount of total RNA per lane was relatively low for the detection of rare transcript. Failure of detecting *tra-2* could be due to the very low levels of *tra-2* transcription. Indeed, this was the case when transcription level of *tra-2* was compared to others via a microarray by Dr. Z. Kambris (see figure 4.26), that incorporates two distinct 60-mer probes for each annotated *Ae. aegypti* gene and were used to investigate transcript abundance in pools of adult female mosquito (Klasson et al., 2009). Due to time restrictions the northern blot experiments were not repeated.

4.4.2.2 5' and 3' RACE experiments

5' RACE results suggest that upstream of the exon 1 site is the 5' UTR which includes the Kozak sequence. The Kozak sequence is a sequence which

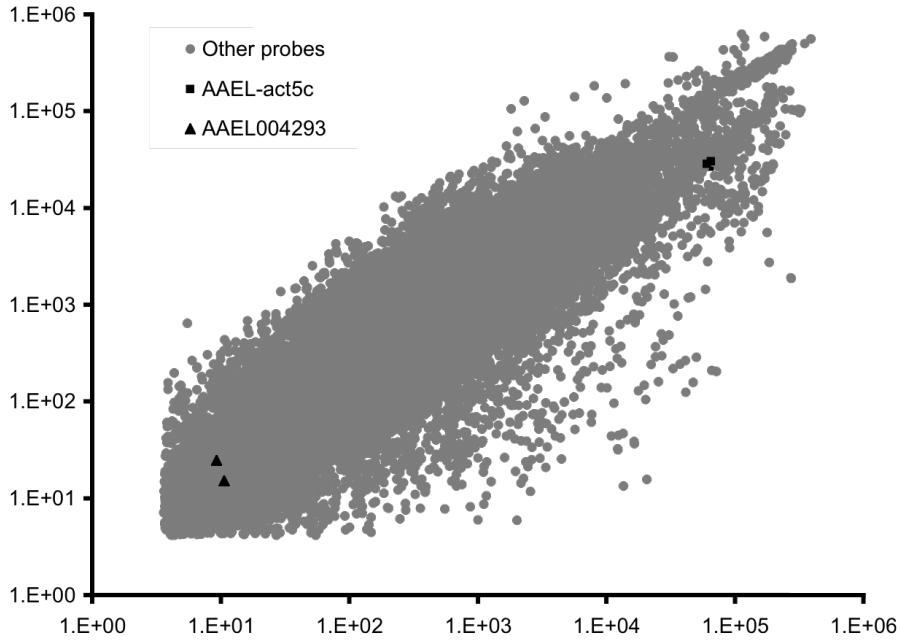

Figure 4.26: Transcript levels of *Ae. aegypti* gene AAEL004293 (*tra-2*) relative to whole genome transcription. Microarray data (Klasson et al., 2009) adapted to show transcription of AAEL004293, in relative to act5c gene (a gene that is transcribed in abundance). Two points of each gene shows the two array samples hybridised: array 1 was hybrised with cRNA from young mosquitoes (0-2 days post-eclosion) and array 2 with cRNA from older mosquitoes (14-16 days post-eclosion). x and y axis denotes log cyanine-3 signal intensity.

occurs on eukaryotic mRNA and has the consensus (gcc)gccRccAUGG, where R is a purine (adenine or guanine) three bases upstream of the start codon (AUG), which is followed by another 'G' (Kozak, 1987). The Kozak consensus sequence plays a major role in the initiation of the translation process. There are variations to the Kozak sequence, and *Drosophila* species have the Kozak sequence of CAAAATG. This fits in with the sequence of *Ae. aegypti tra-2* at the 5' end of the first exon. Variable 5' ends may play a role in translation efficiency, though it is usually UTR sequence less than 20 nucleotides that significantly affects the efficiency (Kozak, 1987). Therefore, the variable length that was obtained in the 5' RACE should not affect the translation of the mRNA variant, as they were at least 20 nucleotides upstream of the Kozak

sequence.

3' RACE proved to be useful in discovering unknown sequences towards the 3' end of *tra-2*. It was interesting to find that an intron that was over 26 kbp long separates exon 1 and 2. This may be why the predicted sequence of *tra-2* in Vectorbase only ends at exon 1. Large introns such as found here however are not uncommon in eukaryotes. There are at least two types of transcripts for male *tra-2* from Marathon 3' RACE (variant F and G), but it cannot be confirmed to be male-specific, as the sole female transcript successfully isolated is the same as the longer male transcript of variant F. 3' SMART RACE produced different variants and revealed that the second exon was actually further spliced. RT-PCR of *tra-2* suggested that pupal and adult expression levels of *tra-2* are different. It is likely that more *tra-2* variants exist and are expressed at different levels during different developmental stages. *In silico* translation of the so called exon 5 showed many stop codons in all possible frames, and that ORF predictions do not go beyond exon 4 in variant I, indicating that this is in actual fact the 3' UTR (figure 4.24).

The difference in sequence at the end of exon E4 in Variant A that causes two amino acids to change may be important, but since the ORF prediction does not extend past exon E2b, this mutation could be redundant in terms of function. The E1b region of Variant H is spliced out, which suggested that this variant is non-functional, since RRM is situated across E1a and E1b. The protein of Variant H may be unable to interact with other RNA or protein molecules. Variant I was only detected when male adult sample was used in RT-PCR, which could explain its absence in the RACE experiments since only pupal samples were used in the experiments. This may indicate that variant I is needed in male adults, possibly playing a role in spermatogenesis, like in *D. melanogaster tra-2* variants C and E, which are expressed in the

germline. In short, alternative splicing, as well as alternative polyadenylation were responsible in producing different *tra-2* variants.

In summary, there are sex specific variants of the *tra-2* mRNA, as well as variants that are common to both males and females of *Ae. aegypti*. There may be more variants of *tra-2* expressed in different levels. Further RACE experiment may be able to reveal more sequence variants if a bigger library can be generated from the RACE transcripts. An alternative is to generate a cDNA library of *Ae. aegypti* and screen for *tra-2* sequences. This work is extremely time consuming, but will have a higher probability in obtaining variants that are transcribed at lower levels. This could also confirm whether certain variants were truly sex-specific. For the purpose of this project however, the number of variants found was deemed sufficient to get a general picture of the workings of *tra-2*.

Quantitative PCR (qPCR) may be used to analyse expressions levels in different tissues and at different developmental stages. This may also give an insight to the workings of *tra-2* variants. Detecting variants of a lowly expressed gene is very difficult and thus quantification of this kind may not be successful. Several attempts were made to achieve this with specific primers but were not successful (data not shown). However, qPCR of all *tra-2* variants by targeting the most common transcribed site (i.e. exon 1) is much more straightforward (see Chapter 5).

Chapter 5

Sex Ratio Distortion in *Aedes aegypti* by RNA Interference

5.1 Introduction

5.1.1 RNA interference

RNA interference (RNAi) is a mechanism that is now frequently used for assessing gene function by knocking down its expression, and could be used to understand the role of candidate genes in sex determination. The introduction of a homologous double stranded RNA (dsRNA) into an organism causes a down-regulation of the expression of a gene that is complementary to the dsRNA (Fire et al., 1998). The dsRNA is cleaved into short interfering RNAs (siRNA) of about 23 base pairs in size by a ribonuclease enzyme that is called a Dicer (Bernstein et al., 2001; Hamilton and Baulcombe, 1999). They then form an siRNA-Dicer complex that recruits additional components to form RNA-induced silencing complexes (RISCs). The siRNA unwinds with the complementary mRNA, guiding the RISCs to the complementary RNA molecules and thus destroying the mRNA by cleavage (Hammond et al., 2000). See figure 5.1 for a diagram RNAi mechanism (Karpala et al., 2005).

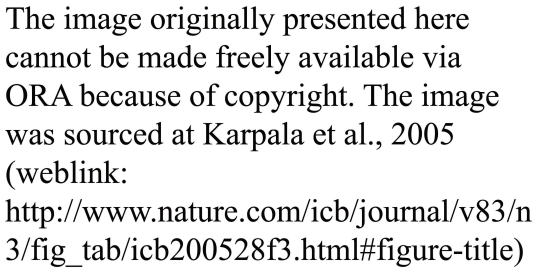

Figure 5.1: Mechanism of RNAi. dsRNA initiates RNAi by activating protein Dicer, which binds and cleaves dsRNAs to produce double-stranded fragments of 21 - 25 nts long. Modified from Karpala et. al, 2005.

5.1.2 Sex ratio distortion by RNAi

The use of heritable RNAi, mediated by transgenes exhibiting dyad symmetry was capable of directing suppression of gene expression in *Anopheles* mosquitoes (Brown et al., 2003). The suppression was acheived using DNA constructs designed to transcribe dsRNA as extended hairpin-loop RNAs (Brown et al., 2003). This indicates that dsRNA-mediated silencing of a target gene via plasmid DNA can be achieved at high levels in mosquito cell lines and larvae. RNAi was used successfully to knockdown pre-membrane protein expression in dengue virus by introducing dsRNA in *Ae. aegypti*, thus making the mosquito resistant to dengue virus infection (Franz et al., 2006). Studies in *D. melanogaster* using RNAi to knock down *tra-2* have also been reported. It has been shown that an expressed inverted repeat of a portion of *tra-2*, driven by a GAL4-dependent promoter, could genetically repress the endogenous wild-type *D. melanogaster tra-2* function, producing a dominant loss of function mutant sexual phenotype on homozygous *tra-2* chromosomal females (XX) (Fortier and Belote, 2000). GAL4-promoter is temperature sensitive, therefore the transformant is also temperature-sensitive, where the phenotypic consequences could only be observed at 29°C (which is the optimum temperature for GAL4-promoter), but not at 22°C. They also discovered that different expression levels of the transgene or the endogenous tra2+ loci can vary the effects of mutant phenotypes, ranging from male-like intersexes to the usual male phenotype.

It is also possible to induce RNAi without introducing a DNA construct, by direct injection of dsRNA. This was successfully done in the silencing the *Toll* pathway components in *D. melanogaster* (Goto et al., 2003). This transient RNAi method was also successfully carried out in mosquito models where gene silencing of two c-type lectin genes and one leucine rich-repeat immune

gene was achieved, which affects *Plasmodium* development within the mosquito (Osta et al., 2004). Silencing of the *TEP1* gene (encoding thioester-containing protein) in adult *An. gambiae* was also achieved by direct injection of dsRNA (Blandin et al., 2004). The silencing of this gene, essential for parasite killing, led to an increase of *Plasmodium* parasite numbers in the mosquito. In regards to sex specific silencing, direct injections of female-specific *dsx* dsRNA has also been performed in adult oriental fruit flies, *Bactrocera dorsalis* which resulted in deformed ovipositors in the female progeny (Chen et al., 2007b). Transient RNAi targeting *tra* in *Ceratitis capitata* embryos resulted in normal males and XX females that transformed into fertile phenotypic males (Pane et al., 2002). No transient knockdown of the gene *tra-2* has been reported in *Drosophila* or mosquitoes so far.

5.1.3 The tet system

Tetracycline-Controlled Transcriptional Activation is a method where transcription is reversibly turned on or off in the presence of the antibiotic tetracycline (figure 5.2). This system was identified in *E. coli* Tn10 bacteria, which makes use of the tetracycline transactivator (tTA) protein created by fusing TetR (tetracycline repressor) protein with another protein called HSV VP16, a powerful transcriptional transactivator, produced by the Herpes Simplex Virus (Gossen and Bujard, 1992). The tTA protein binds to a tetO operator. Once bound the tetO operator will activate a promoter coupled to the tetO operator, activating the transcription of a gene of interest constructed near the promoter. Transcription can be inactivated by presence of tetracycline as TetR will preferentially bind to tetracycline instead of tetO (Orth et al., 2000).

The tet system is now frequently used for conditional expression in insect

vectors, and has also been found to provide a conditional lethal system by over-expression of tTA when no tetracycline is present. Accumulation of tTA within the host cell is thought to cause 'transcriptional squelching', hence interferes with the functioning of the cell (Strathdee et al., 1999). Constructs from this method were over 99% lethal in the Medfly, with just one copy of a positive feedback construct. The tTA produced acts as a driver to increase tTA production, in the absence of tetracycline during developmental stages (Gong et al., 2005).

Figure 5.2: Tetracycline repressible system(a,b) Two-component system as previously published (Thomas et al., 2000; Horn and Wimmer, 2003; Heinrich and Scott, 2000). tTA (Gossen and Bujard, 1992) is placed under the control of a suitable promoter, for example, constitutive, female-specific, embryo-specific. In the absence of tetracycline (Tc) tTA binds tetO, driving expression of an effector molecule and leading, in the case of a lethal effector, to death (a). In the presence of tetracycline, tTA binds tetracycline; the tetracycline-bound form does not bind DNA, therefore does not activate expression of the effector, and the system is inactivated (b). (c,d) A simplified one-component system. In the absence of tetracycline, basal expression of tTA leads to the synthesis of more tTA, which accumulates to high level (c). This level can be regulated by modifying the stability and translational efficiency of the tTA mRNA. At the highest levels, expression is lethal, so tTA is both the driver and the effector. In the presence of tetracycline, tTA is inactivated by tetracycline and is therefore expressed only at basal levels (d). Tc, tetracycline. Figure taken from Gong et al., 2005.

In *Ae. aegypti* strain LA513A, the condition was lethal when two copies of a similar positive feedback construct were inherited as indicated in figure 5.3

(Phuc et al., 2007).

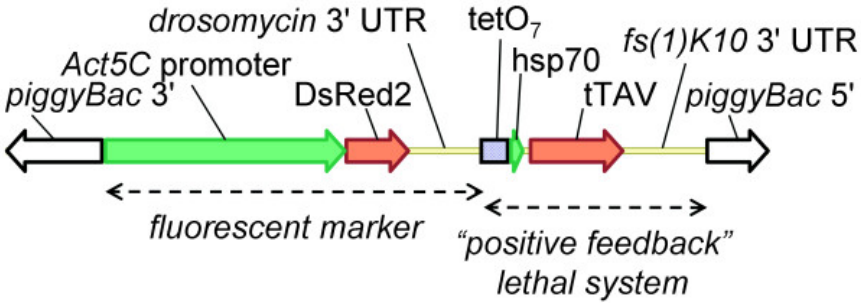

Figure 5.3: Structure and function of transposon LA513 to create LA513A *Ae. aegypti* strain. LA513 is a non-automous *piggyBac*-based transposon. tTAV is a tetracycline-repressible transcriptional activator with a independent binding site, tetO, a minimal promoter from *Drosophila* hsp70, and a 3' UTR sequence from *Drosophila* fs(1)K10. DsRed2 reporter gene contains its own promoter of Act5C, so knockdown of the fluorescent marker will not affect the tetracycline expression. Diagram taken from Phuc et al., 2007.

5.1.4 Aims

The aim of this study was to investigate if *tra-2* silencing could distort the development of chromosomally female individuals to a male phenotype. The first approach was to investigate if direct injection of dsRNA *tra-2* was enough to produce that distortion. Evidence of sexual differentiation is present by the 4^{th} instar larvae stage (Timmermann and Briegel, 1999). Therefore, silencing of *tra-2* up to the 4^{th} instar stage may suffice to distort sexual development. The second stage was to determine if germline transformation with a *tra-2* dsRNA construct could be established. The germline transformants would then be studied to determine if long term RNAi expression could distort the sex ratio.

5.2 Methodology

5.2.1 Transient knockdown: Direct injection of dsRNA

5.2.1.1 Microinjection of *tra-2* dsRNA into *Ae. aegypti* Rockefeller strain embryos

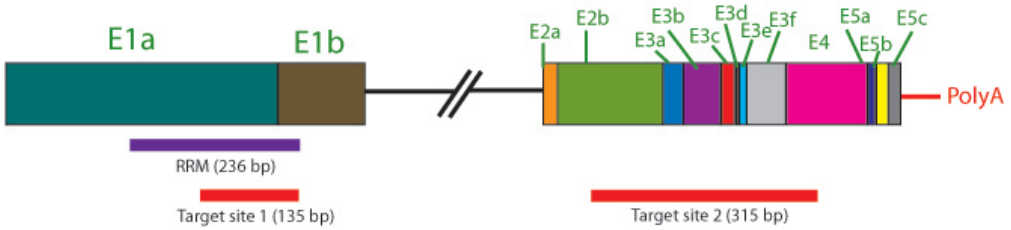

Figure 5.4: RNAi target sites for *tra-2*. The thin red bars underneath the diagram of *tra-2* structure indicate the position of the RNAi target site for which dsRNA was produced. The thin purple bar underneath the *tra-2* diagram indicates where the RNA Recognition Motif (RRM) is situated. Note that RNAi target site 1 is within the RRM site.

dsRNA synthesis of *tra-2* target sites

Two RNAi target sites of *tra-2* were selected based on the results of Chapter 4 and primers were designed. The first target site was based on exon 1. Exon 1 sequence and T7 sequence attachment method was done as in the Northern Blot experiment described in Chapter 4 (see Appendix for the sequence of exon 1 target site). The second RNAi target site was based on the exons 2 to 4 as found by 3' RACE PCR (see Chapter 4). As the sequence of exon 2 - 4 was longer than exon 1 alone, it was thought that the former was more optimal for designing primer pairs than that from the first exon. Therefore, this would give a higher chance of successful interference (see figure 5.4). Sequence of *tra-2* exon 2 - 4 region (position 1194346-1194646 in supercontig 1.113, see Appendix for the sequence) was used for primer design. Primers for *tra-2* exons 2 to 4 with T7 for dsRNA synthesis were as follows:

- Forward with T7: 5'- TAATACGACTCACTATAGGGTCCTCAGTT GACTCACGTCG -3'
- Reverse with T7: 5'- TAATACGACTCACTATAGGGTGTTTAGTT TGAATCGCGGA -3'

For dsRNA synthesis, pupal DNA from the *Ae. aegypti* Rockefeller strain was extracted as described in Chapter 2 and used to synthesise *tra-2* target site templates with T7 sequences attached by standard PCR (refer to Chapter 2 for standard PCR method) to create a total amount of 1 - 2 μg product. The PCR product was then run on 1.0% agarose gel. The band of interest was cut and purified using Qiagen Gel Extraction Kit (see Chapter 2 for the standard protocol).

Using MEGAScript RNAi kit (Ambion), dsRNA was synthesised from 8 μl (200 ng) of PCR template DNA to create as much as 100 μg of dsRNA according to the manufacturer's manual.

Nanodrop was used to check the concentration and purity of dsRNA. $1/400^{th}$ of the product was run on a 1% TAE agarose gel to examine the integrity and efficiency of duplex formation. The dsRNA was stored at -20°C until further use.

dsRNA preparation for microinjection

The dsRNA was diluted in 1× injection buffer (5 mM KCl; 0.1 mM $NaPO_4$, pH 7.2) to the appropriate concentration needed.

Microinjection of *tra-2* dsRNA

Microinjection procedure was adapted from Jasinskiene (Jasinskiene et al., 1998) and Lobo (Lobo et al., 2006) with the following changes:

- *Aedes aegypti* embryos collected at age 20 to 45 minutes and were lined

up in batches of 30-50 on wet filter paper (Whatman) with the narrow posterior ends facing the same direction and the dorsal surface facing up.

- The embryos were allowed to desiccate slightly at room temperature by drying the wet filter paper for 40 seconds to 1 minute.
- Embryos were transferred to double-sided tape (3M) on a plastic microscope cover slip (TAAM) by gently pressing the coverslip against the dorsal surface of the embryos.
- The embryos were then quickly covered in water-saturated hydrocarbon oil mix, which is 2 parts halocarbon 700 oil to 1 part halocarbon oil 27 (Sigma-Aldrich) to prevent further desiccation.
- The coverslip was attached to a glass slide using a small drop of halocarbon oil mix, and then mounted onto the microinjection stage.
- dsRNA in injection buffer was thawed and spun briefly in a micro-centrifuge. Using a microloader pipette tip (Eppendorf), about 4 μl of dsRNA solution was loaded into a femtotip 2 microinjection needle (Eppendorf).
- At 1-2 hours after laying, embryos were injected through the posterior pole at an angle of about 15 degrees using Femtojet Microinjector (Eppendorf, UK). Roughly 0.5 to 1 nl dsRNA (with concentration of 1, 3, or 5 μg) was injected into each embryo (see figure 5.5).
- The embryos were then left on the cover slip in coverslip rack inside a 100% humidity chamber (water-tight container lined with damp cotton and filter paper) at insectary conditions for 3-5 days before allowing them to hatch.

- Injected eggs were hatched by submerging in deoxygenated water with several pellets of fish food (Omega) for two weeks. The number of hatched eggs on the cover slip was counted.

- The larvae were allowed to grow until pupal stage before males and females were counted. To ensure that any sex ratio distortion observed is statistically significant, Chi square tests were performed.

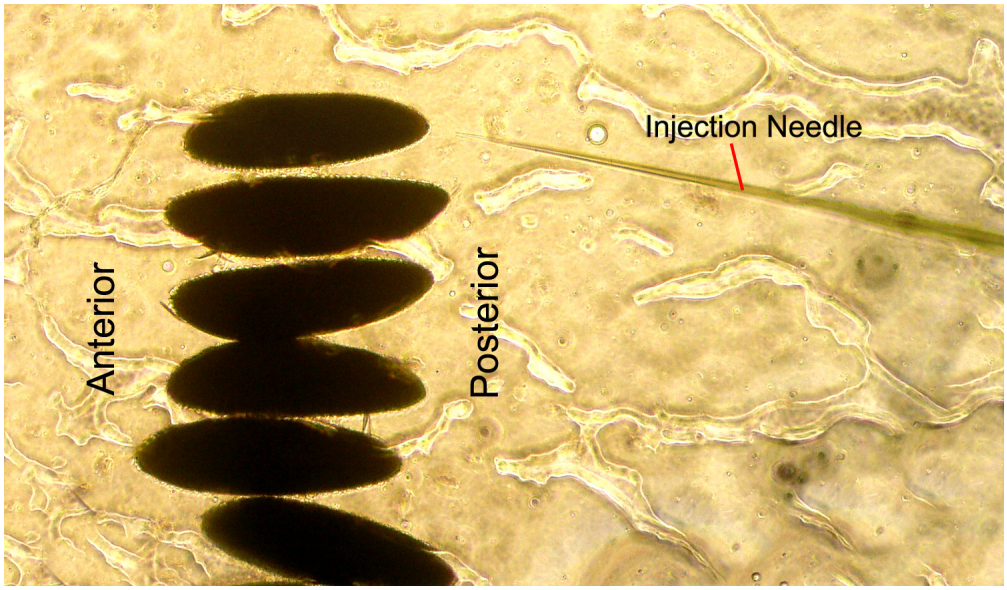

Figure 5.5: Microinjection of *Ae. aegypti* embryos. Embryos were lined up in an anterior to posterior direction and pressed onto double-sided tape on a coverslip and covered in halocarbon oil. Injection takes place in the posterior end of the embryos at an angle of 15 degrees.

5.2.1.2 RT-PCR to confirm *tra-2* knockdown

Individual RNA extractions were carried out on injected survivors at different stages. RT-PCR was carried out using 100 ng of each RNA sample to confirm if RNAi knockdown of *tra-2* was successful. The primers used to amplify a 120 bp fragment that is 55 bp upstream of the *tra-2* RNAi target site in exon 1 were:

- Forward: 5'- GAAACATCGGCCTGTTCATC -3'
- Reverse: 5'- GCTGGTTTCGTTGGTGTAGC -3'

For primers in exon 2 to 5:

- Forward (tra-2 exon 2): 5'- AGATCTCGTTCCTGTCATCATC -3'
- Reverse (tra-2 variant 1, exon 5): 5'- GACTGAATCAATTGCGATGC -3'

5.2.2 LA513A strain rearing

Transgenic *Ae. aegypti* strain LA513A containing an insert for the red fluorescent marker DsRed2 (a modification of the original marker DsRed), was used as a control to to ensure knockdown was successful and to estimate the time it lasted. Rearing conditions for this colony were the same as the Rockefeller strain, except that 0.03 g/l of tetracycline hydrochloride (Sigma-Aldrich, USA) was dissolved into the water prior to egg hatching. The tetracycline is essential for larvae development as it suppresses the lethal construct of LA513A strain.

5.2.3 Microinjection of strain LA513A

Injection procedures were the same as for the Rockefeller strain. To observe if RNAi knockdown of the DsRed2 gene was successful, larvae of different stages from the injected eggs were observed under a Olympus SZX-12 microscope equipped with red fluorescence filter: 510-550 excitation, 590LP emission.

5.2.3.1 Microinjection of dsRNA DsRed2 into *Ae. aegypti* LA513 strain embryos

Microinjection procedure of dsRNA DsRed2 was as described above for microinjection of *tra-2* dsRNA.

dsRNA synthesis of DsRed2 target sites

Primers for DsRed2 dsRNA with T7 for dsRNA synthesis

- Forward: 5'- TAATACGACTCACTATAGGGTGGTGTAGTCCTCG TTGTGG -3'

- Reverse: 5'- TAATACGACTCACTATAGGGAAGGTGTACGTGAA GCACCC -3'

dsRNA synthesis were made using the MEGAScript RNAi kit (Ambion). Procedure was as previously described in dsRNA synthesis of *tra-2* target sites.

5.2.4 Germline transformation

Germline transformation may provide a more reliable/longlasting way of inducing RNAi. Germline transformation will provide a consistent dsRNA transcription which means RNAi levels will be maintained throughout the lifespan of the transformed individual.

5.2.4.1 Preparation of Vector Plasmid Construct

Plasmid construct Zoo-2 contained the genetic marker DsRed2 and 135 bp of *tra-2* exon 1 sequence of target site 1 (see figure 5.4) that contained an inverted repeat, joined by fragment of *tra-2* intron-rabbit beta-globulin 3', in order to form a duplex when expressed (figure 5.6). Rabbit beta-globulin 3' was chosen as an insulator sequence to aid control and optimizes the transcription of the transgene (K. P. Hoang, personal communication). The expression is regulated under CMV-tetO7. CMV is a human cytomegalovirus promoter and is ligated to tetO7 (tetracycline operator). Therefore, when tTA is present, the CMV promoter would activate the transcription of the transgene i.e. *tra-2* dsRNA.

Plasmids were introduced into DH10 competent cells. A *piggyBac* 'helper plasmid' phsp-pBac in competent cells were also provided. These cells were cultured overnight in LB with ampicillin and harvested. Plasmids were

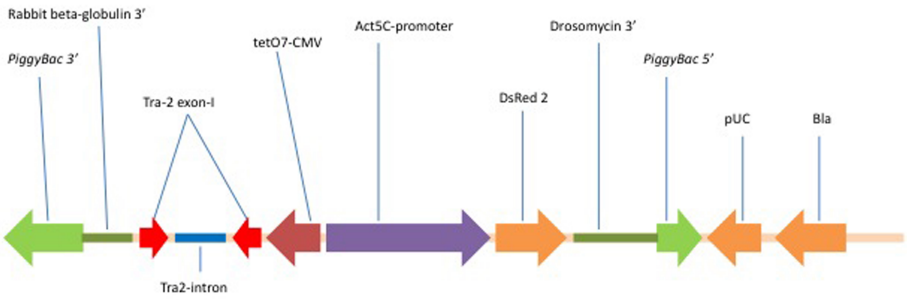

Figure 5.6: Diagrammatic representation of plasmid Zoo-2, linearized at the 3 end of the *piggyBac* transposon. There are two functional segments within the ends of the transposon : the marker (Act5C promoter-DsRed2-Drosomycin 3') to allow detection of transgenic individuals by fluorescence and a *tra-2* repression RNAi cassette (tetO7-CMV minimal promoter two *tra-2* Exon 1 inverted repeats, joined by a fragment of *tra-2* intron-rabbit beta-globulin 3). The *tra-2* RNAi cassette acts by a tTA protein binding into the tetO7 to free CMV minimal promoter to drive. The tTA protein will be expressed after crossing line bearing Zoo-2 construct with a line bearing a construct which produces tTA protein. Construct design and assembly primarily by Dr. K. P. Hoang.

purified using Endofree Plasmid Purification Maxi Kit (Qiagen) according to manufacturer's protocol.

To determine the yield, the DNA concentration was determined by both UV spectrophotometry Nanodrop at 260 nm and quantitative analysis on an agarose gel. For reliable spectrophotometric DNA quantification, A260 readings should lie between 0.1 and 1.0.

5.2.4.2 Microinjection of embryos

Germline microinjection was undertaken with Dr. Kim Phuc Hoang. The protocol was as described previously in the microinjection of dsRNA. Vector plasmid pLA513 concentration of 600 ng/μl was co-injected with a *piggyBac* 'helper plasmid' phsp-pBac at 300 ng/μl as a source of *piggyBac* transposase (Handler et al., 1998). After injection, embryos were stored for 3 days in 100% humidity chamber in insectary conditions before submerging in deoxygenated

water for hatching. The survivors were reared to adult stages, and these (G0) survivors were back-crossed to the wild-type colony: individual G0 males were crossed to 10 wild-type females and G0 females were allowed to cross with 1 wild-type male. G1 eggs were then collected and hatched following the same procedure as for the LA513 strain described earlier. The pupae were then screened for fluorescence using Olympus SZX-12 microscope with red fluorescence filter: excitation 510-550, emission 590LP.

5.2.5 Molecular analysis of the germline transformant

5.2.5.1 Inverse PCR

Inverse PCR was carried out to identify genomic sequence adjacent to the insertion site of the *tra-2* construct.

Five transformed *Ae. aegypti* that showed DsRed were used to extract DNA using the Livak method described in Chapter 2. The DNA was then digested with restriction enzyme as described:

- 10 μl (2 μg in total) of DNA each were placed in 1.5 ml microfuge tubes separately. 15.8 μl of sterile water was added to each tube.
- 3 μl of buffer 3 (NEB) were added to one of the tubes together with 0.2 μl of BSA and 1 μl of *Taq*$^{\alpha}$ *I* restriction enzyme to make up a total volume of 30 μl. The tube was then overlaid with mineral oil.
- 3 μl of buffer 2 (NEB) were added to the second tube together with 0.2 μl of BSA and 1 μl of *Msp I* restriction enzyme to make up a total volume of 30 μl.
- The reaction was mixed gently by flicking the tubes.
- The tubes were then spun briefly by micro-centrifuge to collect the reaction at the bottom of the tubes.
- Tube containing *Taq*$^{\alpha}$ *I* restriction enzyme was incubated at 65°C for 16 hours. Tube containing *Msp I* restriction enzyme was incubated at 37°C for 16 hours.
- After incubation, tubes were heated at 80°C for 15 minutes to stop the reaction.

- The digested DNA samples were purified by phenol-chloroform extraction as described in Chapter 2.
- 10 μl of each DNA sample were run in 1% agarose gel to check if DNA were digested completely.

Each remaining digested DNA sample was then ligated to form circular DNAs using T4 as follows:

- 20 μl digested DNA
- 249 μl sterile water
- 30 μl $\times$10 T4 ligase buffer (NEB)
- 1 μl of T4 ligase (NEB)

The samples were mixed gently by flicking the tubes and spun briefly in a microcentrifuge. The tubes were incubated at 4°C for 24 hours. The reactions were then terminated by heating at 70°C for 10 minutes.

Two μl of each ligated DNA were then used for inverse PCR. The primers used were as follows:

For *Taq*$^{\alpha}$ *I* digested DNA:

- *Taq I* Forward: GATGAGGTACATGAAGTGCAGCCA
- *Taq I* Reverse: CATGCGTCATTTTGACTCACGC

For *Msp I* digested DNA:

- *Msp I* Forward: GCCTGCCGTGAATACCGTATA
- *Msp I* Reverse: CTGGCTCTTCAGTACTGTCATC

The sequence of the flanking region were compared to those found in the genome databases to see where the insert lies and if the position of the insert could be disrupting another gene. The flanking sequences were used as a diagnostic tool to identify specifically if transformed individuals are truly homozygous by simple PCR amplification method.

5.2.6 Producing homozygous mosquitoes

5.2.6.1 Crosses

Crosses to produce and test the effectiveness of the *tra-2* construct were as follows. Figure 5.7 is a diagram of the crosses.

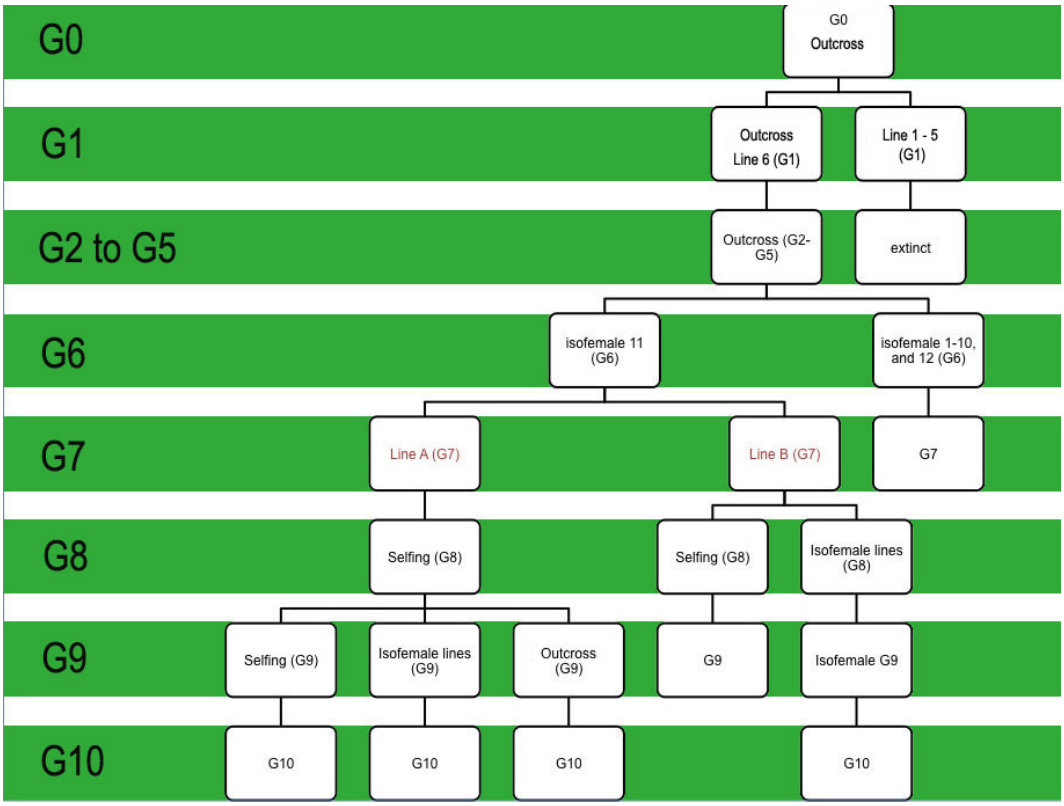

Figure 5.7: Diagram showing transgenic line No. 6 strain crosses in each generation (G) and other test crosses. At each generation, DsRed+ individuals were selected for the next cross. Selfing refers to mating of individuals from the same generation. Outcross refers to DsRed+ males crossed with wild-type females. Isofemale lines refer to single DsRed+ male that was allowed to mate with several DsRed+ females. Line A refers to the DsRed+ population with higher fluorescent intensity. Line B population has a lower fluorescent intensity than Line A.

Initial cross to maintain transgenic survivors

The surviving embryos of the injections (G0) that were DsRed2 positive were crossed to wild-type Rockefeller strain. Non-fluorescent individuals were discarded. The progeny of G0 (G1) would consist of mostly heterozygous

individuals. As numbers of injection survivors were low, each G1 male from each strain were outcrossed with wild-type strain Rockefeller females to produce a more robust line. The lines were then screened for DsRed2 and outcrossed again for 3 more generations. At G5, the strains were crossed, as described below for each generation, to achieve a homozygous line:

G5 crosses

Eggs from the progeny of G4 crosses (G5) were hatched and selected for DsRed2 fluorescence at the pupal stage. The non- fluorescent individuals were discarded. Twelve crosses were then set up using the DsRed2 positive individuals, mating every male with 6 to 12 virgin females to achieve 12 lines. The females were then isolated singly in vials containing damp cotton lined with small filter paper disc two days after blood meal and allowed to lay eggs. Each iso-female line eggs were separately hatched in 250 ml of deoxygenated water.

G6 crosses

G6 progenies showing strong fluorescence at pupal stage were selected. The crosses were then set up as in the G5 crosses. Each batch of eggs from individual females were hatched separately and screened for fluorescence. Sex ratio of the offspring was determined. Hatch and pupation rate were also noted.

G7 crosses

G6 progenies with 100% fluorescence were used for G7 crosses. Individuals with strong DsRed2 fluorescence distributed around the body of the pupae from iso-female lines were selected to examine if sex ratios show any correlation with the observed location of fluorescence (see figure 5.11):

- Cross A displays strong fluorescence distributed around the thorax and small areas of the head and abdominal area.
- Cross B shows fluorescence mainly in the thorax region and with a weaker intensity than individuals in Cross A.

The sex ratio of progeny resulting from these crosses (G6) were determined.

G8 crosses

Progenies of Cross A and B were allowed to mate with each other. Progenies were then screened for DsRed2 and sex ratio was determined. Hatch and pupation rates were also noted.

Viability of homozygous line

DsRed2 fluorescence was checked in each generation to evaluate whether the homozygous line is viable, or if germline transformation could interfere with fitness. At each generation the population was screened for DsRed2 and the numbers of DsRed2+ were recorded.

Male phenotype observation

Dissections of male adult mosquitoes were performed on Rockefeller strain and Line 6 strain at G9 to compare the development of the sexual organs.

Crossing Line 6 with Rockefeller

Single pair matings of Line 6 male to Rockefeller virgin females were carried out in 25 individual containers. Also 2 mass crosses were carried out where 30 Line 6 males were allowed to mate with 30 virgin Rockefeller females. The progenies of these crosses were then screened for DsRed2 marker, hatch rate , pupation rate and sex ratio.

Crossing Line 6 with LA513A Strain

To determine if sex ratios observed in the homozygous Line 6 population were influenced by the presence of tTA, Line 6 was crossed with the LA513A strain which carries tTA construct. Two crosses were made with 5 male LA513A adults to 5 Line 6 virgin females. The eggs produced from the crosses were hatched in deoxygenated water, with or without 0.10 μg/l of tetracycline.

5.2.7 Quantitative PCR

Quantitative RT-PCR was carried out to analyse the level of expression of *tra-2* following knockdown experiments. Total RNA from Rockefeller (wild-type) and the transgenic Line 6 strain A were isolated in pools of 5 individuals for male pupae in 3 replicates. One μg of each sample were then used to synthesise cDNA using VILO cDNA Synthesis kit, as described in Chapter 2. Quantitative RT-PCR was performed using 1 to 20 dilutions of the cDNAs using SYBR Green I as described in Chapter 2, with the annealing temperature of 54°C for *tra-2* primers, and 59°C for the housekeeping gene Actin5C.

The primers used for Actin5C, which amplify a 125 bp fragment, were:

- Forward: 5'- ATCGTACGAACTTCCCGATG -3'
- Reverse: 5'- ACAGATCCTTTCGGATGTCG -3'

Primers used for detecting *tra-2*, which amplify a 120 bp fragment that is 55 bp upstream of the *tra-2* RNAi target site in exon 1, were:

- Forward: 5'- GAAACATCGGCCTGTTCATC -3'
- Reverse: 5'- GCTGGTTTCGTTGGTGTAGC -3'

Two more assay sets were carried out to ensure the results were consistent and an average of the expression levels were obtained. Expression levels were

calculated by the relative standard curve method, as described in Technical Bulletin #2 of the ABI Prism 7700 Manual (Applied Biosystems), using (i) Act5C as an endogenous reference and (ii) an exogenous calibrator appropriate for the experiment (in this case, the wild-type Rockefeller strain).

5.3 Results

5.3.1 Transient RNAi knockdown results

5.3.1.1 Microinjection of *tra-2*

dsRNA *tra-2* exon 1

A total of 2239 embryos were injected with dsRNA targeting *tra-2* exon 1 at a concentration of 3 μg/μl. A total of 282 embryos hatched (11%). The total number that survived to pupal stage was 234 individuals, giving the survival rate to pupal stage of 83%, where 55% of the pupae were males (table 5.1). A Chi square test showed that the percentage skew was not significant (χ^2= 0.526, p= 0.4683).

Table 5.1: Result of *tra-2* dsRNA injection using Exon 1 region. Concentration of dsRNA was 3 μg/μl.

Batch no.	Total embryos	Percentage Hatched (larvae/embryos)	Percentage Pupation (pupae/larvae)	Percentage Female (female/pupae)	Percentage Male (male/pupae)
1	438	9% (39/438)	69% (27/39)	30% (8/27)	70% (19/27)
2	514	7% (38/514)	86% (33/38)	30% (10/33)	70% (23/33)
3	558	50% (115/228)	86% (99/115)	53% (52/99)	47% (47/99)
4	217	14% (31/217)	97% (30/31)	50% (15/30)	50% (15/30)
5	368	8% (30/368)	83% (25/30)	44% (11/25)	56% (14/25)
6	224	7% (16/224)	88% (14/16)	43% (6/14)	57% (8/14)
7	320	4% (13/320)	62% (8/13)	37% (3/8)	63% (5/8)
Total	2639	11% (282/2639)	83% (236/282)	45% (105/236)	55% (129/236)

To analyse the expression of the target gene, RT-PCR was performed on total RNA extracted from different life stages of the injected strain (5 individuals per stage). Approximately 100 ng of total RNA was used for each sample. It was found that the injected strains still show positive expression of *tra-2* target gene (figure 5.8).

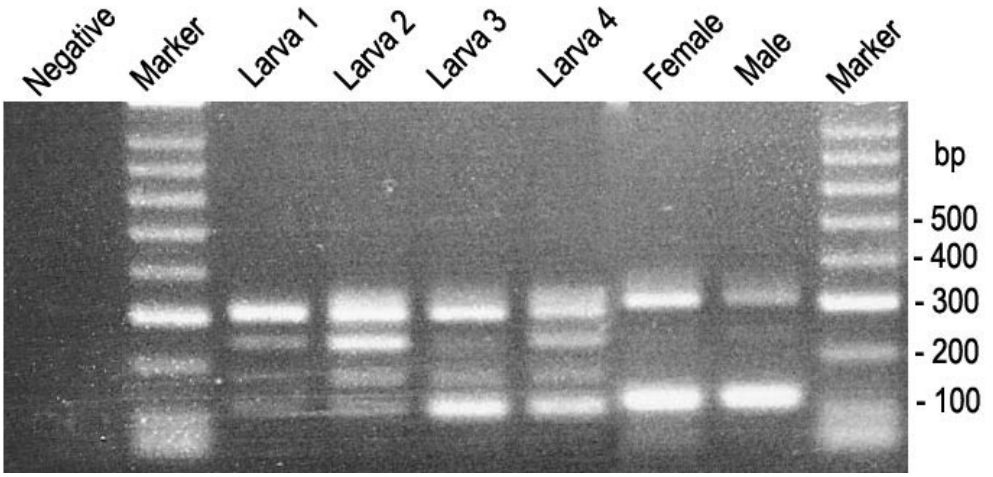

Figure 5.8: *Tra-2* expression at different developmental stages of mosquitoes. Embryos were injected with *tra-2* dsRNA targeting exon 1. Samples were analysed via RT-PCR, using *tra-2* primers designed from exon 2 to 4 to avoid overlap with the dsRNA target site. Negative control (Negative) = no template, Larva 1= 1^{st} instars larva, Larva 2 = 2^{nd} instars larva, Larva 3 = 3^{rd} instars larva, Larva 4 = 4^{th} instars larva, Female = female pupae, Male = male pupae, Marker = 100 bp size marker ladder. Note that RT-PCR yields more than one PCR product due to splicing of the variants of *tra-2* in exons amplified. 100 ng of total RNA that was extracted from 5 individuals were used for each sample.

tra-2 exons 2 to 4 dsRNA

A total of 551 *Ae. aegypti* embryos (Rockefeller strain) were injected with *tra-2* exons 2 to 4 dsRNA at a concentration of 1 μg/μl. Three hundred and forty seven embryos were injected at a concentration of 3 μg/μl and 72 eggs at a concentration of 5 μg/μl. Out of the 551 injected Rockefeller embryos at 1 μg/μl concentration, 226 hatched successfully with 41% hatch rate. A total of 191 pupae reached adulthood, giving a survival rate of 85%. Out of the 191 pupae, 88 (46%) were females (table 5.2). A Chi square test showed that the sex ratio is not significantly different from the 1:1 ratio (χ^2= 0.308, p = 0.5789).

Table 5.2: Result of *tra-2* dsRNA injection using exon 2 to 4 region (target site 2). Concentration of dsRNA was 1 μg/μl.

Batch no.	Total embryos	Percentage Hatched (larvae/embryos)	Percentage Pupation (pupae/larvae)	Percentage Female (female/pupae)	Percentage Male (female/pupae)
1	125	41% (51/125)	100% (51/51)	49% (25/51)	51% (26/51)
2	32	16% (5/32)	100%(3/3)	100% (3/3)	0% (0/3)
3	82	10% (8/82)	100% (8/8)	25% (2/8)	75% (6/8)
4	22	5% (1/22)	100% (1/1)	0% (0/1)	100% (1/1)
5	82	38% (31/82)	97% (30/31)	13% (4/30)	87% (26/30)
6	208	63% (131/208)	75% (98/131)	55% (54/98)	45% (44/98)
total	551	42% (229/551)	85% (191/226)	46% (88/191)	54% (103/191)

For the embryos that were injected at 3 μg/μl, 47 of 347 hatched (14%). The number that survived to pupal stage was 35 (74%), of which 49% were male (table 5.3). A Chi square test indicated that the sex ratio was not significantly different from the expected value of 50% (χ^2= 0.041, p= 0.8395).

Table 5.3: Result of *tra-2* dsRNA injection using exons 2 to 4 region. Concentration of dsRNA was 3 μg/μl.

Batch no.	Total embryos	Percentage Hatched (larvae/embryos)	Percentage Pupation (pupae/larvae)	Percentage Female (female/pupae)	Percentage Male (female/pupae)
1	41	10% (4/41)	50% (2/4)	0% (0/2)	100% (2/2)
2	101	16%(16/101)	75% (12/16)	67% (8/12)	33% (4/12)
3	140	10% (14/140)	71% (10/14)	60% (6/10)	40% (4/10)
4	65	2% (13/65)	85% (11/13)	36% (4/11)	64% (7/11)
total	347	14% (47/347)	74% (35/47)	51% (18/35)	49% (17/35)

Of the 72 eggs injected at 5 μg/μl, only 16 hatched, giving a hatch rate of 22%. Out of those, 10 pupae (63%) were male (table 5.4). The low number of injected embryos at 5 μg/μl was mainly due to the technical difficulty of injecting at a high concentration as the fluid is viscous and thus has more tendency to block the needle. The survival of these mosquito embryos is expected to be low since the injection material is more concentrated which will cause mortality and also more pressure was needed to inject. Due to the low number of injected individuals, there was not enough evidence to conclude that the percentage

Table 5.4: Result of *tra-2* dsRNA injection using exon 2 region. Concentration of dsRNA was 5 μg/μl.

Total embryos	Percentage Hatched (larvae/embryos)	Percentage Pupation (pupae/larvae)	Percentage Female (female/pupae)	Percentage Male (female/pupae)
72	22% (16/72)	100% (16/16)	38% (6/16)	63% (10/16)

of females observed is significantly reduced. Indeed, a Chi square test on the result showed that χ^2= 3.125, (p = 0.0771).

RT-PCR was performed on total RNA extracted from different life stages of the injected strain (5 individuals per stage). Approximately 100 ng of total RNA was used for each sample. It was found that the injected strains still show positive expression of *tra-2* target gene (figure 5.9).

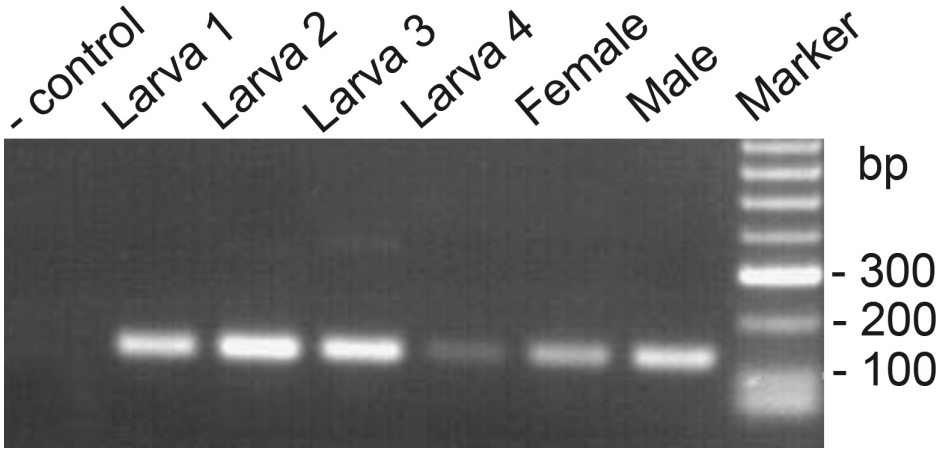

Figure 5.9: *Tra-2* expression in samples injected with *tra-2* dsRNA targeting exon 2 analysed at different developmental stages via RT-PCR. The *tra-2* primers used were from the first exon to avoid overlap with the target site (*tra-2* exon 1 primers as in chapter 3). Negative control (- control) = no template, Larva 1= 1^{st} instars larva, Larva 2 = 2^{nd} instars larva, Larva 3 = 3^{rd} instars larva, Larva 4 = 4^{th} instars larva, Female = female pupae, Male = male pupae, Marker = 100 bp size marker ladder. 100 ng of total RNA that was extracted from 5 individuals per sample was used.

5.3.1.2 DsRed2 Knockdown

A total of 707 of LA513A embryos were injected with DsRed2 dsRNA at a concentration of 3 μg/μl and 62 embryos (9%) hatched. 77% of those larvae survived to pupal stage, of which 60% (table 5.5) were female (χ^2= 2.167, p= 0.141 indicating that it is not significantly different from the 1:1 ratio. When knockdown individuals were viewed under DsRed2 filter, it was found that no red fluorescence was visible, suggesting complete knockdown had occurred, in 2^{nd} and 3^{rd} instar larvae observed (figure 5.10). At 4^{th} instar larval stage (Day 7-8), some red fluorescence or full red fluorescence was observed, indicating that transient knockdown effect of the RNAi had faded by this stage.

Table 5.5: Result of DsRed2 dsRNA injection. Concentration of dsRNA was 3 μg/μl.

Batch no.	Total embryos	Percentage Hatched (larvae/embryos)	Percentage Pupation (pupae/larvae)	Percentage Female (female/pupae)	Percentage Male (male/pupae)
1	284	14% (38/284)	87% (33/38)	64% (21/33)	30% (10/33)
2	90	8% (7/90)	86% (6/7)	67% (4/6)	33% (2/6)
3	91	3% (3/91)	100% (3/3)	33% (1/3)	67% (2/3)
4	242	6% (14/242)	43% (6/14)	50% (3/6)	50% (3/6)
total	707	9% (62/707)	77% (48/62)	60% (29/48)	40% (19/48)

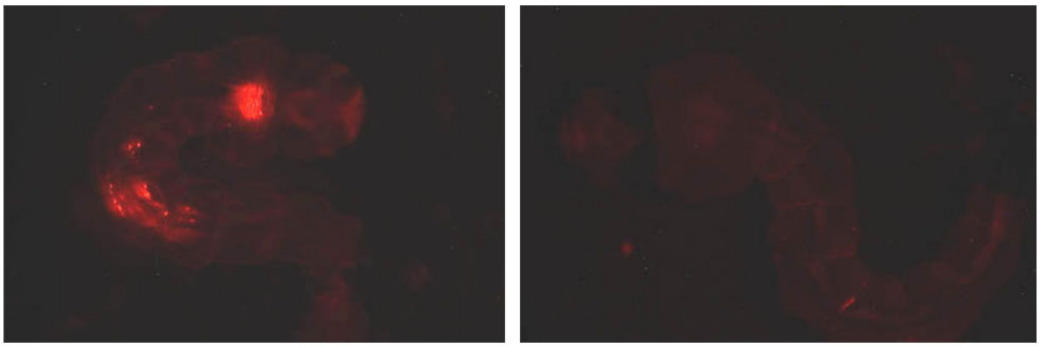

Figure 5.10: DsRed2 transient knockdown in LA513A. LA513A 3^{rd} instar larvae showing DsRed2 in non-injected line (left box). While DsRed2 is completely knocked down even at 3^{rd} instar stage by injection of DsRed2 dsRNA in LA513A embryos (right box).

5.3.2 Germline transformation

Six lines were generated with Dr. K. P. Hoang from the injection survivors. Only Lines 5 and 6 were found to be DsRed2 positive. However, only Line 6 survived, while Line 5 was gradually lost due to a fungal contamination.

5.3.3 Line 6 crosses

G7 population selection for homozygous lines

Sibling isofemale lines (DsRed2 positive Line 6) showed that not all individuals from the G7 population are homozygous since some crosses did not show 100% DsRed2 positive progeny (table 5.6). Only cross sets 1 and 11 had 100% DsRed2 positive progeny. Of these two sets, set 11 showed higher male bias (average of 74% male) than set 1 (average of 54% male), therefore it was chosen to be used for the next generation of crossing.

G8 population selection for homozygous lines

Progeny of Line 6 G7 set 11 population from the mass cross with its siblings that were DsRed2 positive was hatched in two sets. In batch 1 and batch 2, 73% and 32% of the offsprings were DsRed2 positive respectively (table 5.7). This suggests that although the male parents of G7 were homozygous, the female parents were heterozygous. Thus, the G7 population is a mix of homozygous and heterozygous individuals, producing some G8 individuals that carry no construct. Both batches of G8 showed an average of 62% male bias. This was not significantly different from the 1:1 ratio (χ^2= 6.131, p= 0.0133).

Table 5.6: Number of transformed individuals and sex ratio of G7. G7 individuals were progeny of G6 showing DsRed2 marker and were mass-crossed together. Cross ID denotes the individual male crossed with females which then laid individually. (i.e. 1- 1 = Cross number one with male 1, egg batch from single mother no. 1.

Cross	Total embryos	Percentage Hatch (larvae/embryos)	Percentage Pupation (pupae/larvae)	Percentage DsRed2+ pupae (DsRed2+ pupae/ total pupae)	Percentage male DsRed2+ (DsRed2+ male/DsRed2+ pupae)
1-1	30	73% (22/30)	86% (19/22)	100% (19/19)	42% (8/19)
1-3	42	90% (38/42)	95% (36/38)	100% (36/36)	61% (22/36)
1-4	57	54% (31/57)	87% (27/31)	100% (27/27)	59% (16/27)
5-1	55	96% (53/55)	100% (53/53)	60% (32/53)	41% (13/32)
5-3	90	94% (85/90)	96% (82/85)	55% (45/82)	49% (22/45)
5-5	61	97% (59/61)	98%(58/59)	53% (31/58)	55%(17/31)
6-1	62	89% (55/62)	96% (53/55)	51% (27/53)	48% (13/27)
6-2	72	94% (68/72)	96% (65/68)	46% (30/65)	53% (16/30)
6-3	57	88% (50/57)	94% (47/50)	55% (26/50)	69% (18/26)
7-1	53	92% (49/53)	88% (43/49)	47% (20/43)	65% (13/20)
7-4	83	96% (80/83)	100% (80/80)	43% (34/80)	56% (19/34)
7-5	95	93% (88/95)	89% (78/88)	69% (54/78)	67% (36/54)
8-1	58	97% (56/58)	98% (55/58)	71% (39/55)	41% (16/39)
8-2	62	97% (60/62)	95% (57/60)	68% (39/60)	69% (27/39)
8-3	69	94% (65/69)	100% (65/65)	65% (42/65)	60% (25/42)
8-4	78	96% (75/78)	97% (73/75)	64% (47/73)	67% (36/54)
11-1	63	92% (58/63)	100% (58/58)	100% (58/58)	76% (44/58)
11-2	17	88% (15/17)	100% (15/15)	100% (15/15)	60% (9/15)
11-3	9	56% (5/9)	100% (5/5)	100% (5/5)	100% (5/5)
11-4	57	96% (55/57)	91% (50/55)	100% (50/50)	60% (30/50)

Table 5.7: Number of transformed individuals and sex ratio of G8. Progeny were from G7 sibling cross number 11.

Cross	Total embryos	Percentage Hatch (larvae/embryos)	Percentage Pupation (pupae/larvae)	Percentage DsRed2+ pupae (DsRed2+ pupae/total pupae)	Percentage Male DsRed2+ (DsRed2+ male/ DsRed2+ pupae)
1	73	100% (73/73)	100% (73/73)	73% (53/73)	62% (33/53)
2	81	99% (80/81)	94% (75/80)	32% (24/75)	63% (15/24)

G9 population

Progeny of G8 that were sibling mass crossed in two sets, A and B (A had stronger DsRed2 fluorescence than B, figure 5.11) and they were hatched in two batches (table 5.8). Both sets of the cross showed a high hatch rate (Set A = 87%, Set B = 87%) and pupation rate (Set A = 96%, Set B= 98%). 94% of set A were DsRed2 positive, while Set B on average contained 22% fewer DsRed2 positive individuals (average of 72%). 94% of the population in set A on average were males. This is significantly skewed from the normal ratio (χ^2= 60.197, p< 0.0001). Set B had a lower male percentage compared to set A, averaging 70%, but this skew was still significant from the expected 1:1 ratio (χ^2= 23.121, p< 0.0001).

About 40 adult male mosquitoes from the Set A population were dissected and the sexual organs were compared to Rockefeller strain adult males. No visible differences were seen in the male organs of the two strains.

Table 5.8: Number of transformed individuals and sex ratio of G9. Progeny were from G8 sibling cross A (strong DsRed2 marker) and B (weak DsRed2 marker).

Cross	Total embryos	Percentage Hatch (larvae/embryos)	Percentage Pupation (pupae/larvae)	Percentage DsRed2+ pupae (DsRed2+ pupae/total pupae)	Percentage Male DsRed2 (DsRed2+ male/ DsRed2+ pupae)
A -1	95	88% (84/95)	95% (80/84)	93% (74/80))	84 (62/74)
A -2	238	87% (206/238)	97% (200/206)	96% (192/200)	93% (179/192)
Mean	-	87%	96%	94%	94%
B -1	95	83% (79/95)	99% (78/79)	59% (46/78)	63% (29/46)
B -2	238	90% (215/238)	98% (211/215)	86% (181/211)	77% (140/181)
Mean	-	87%	98%	72%	70%

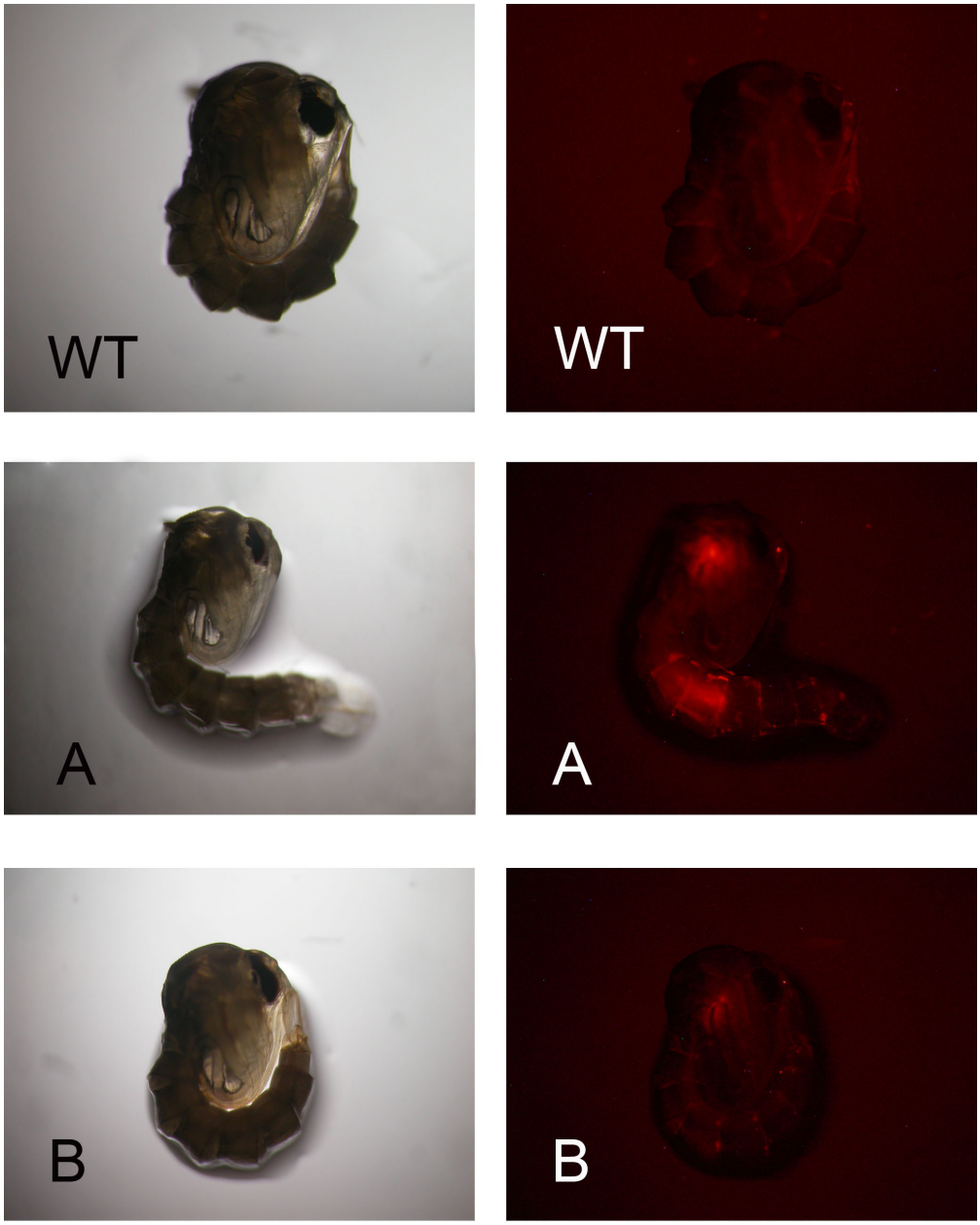

Figure 5.11: Wild type *Ae. aegypti* pupa in comparison to DsRed2 positive individuals from Line 6 at G8. Pupa under normal light (left boxes) and DsRed filter (right boxes). "WT" indicates wild-type (Rockefeller) strain, "A" indicates Line 6 strain A and "B" indicates Line 6 strain B. Photos were captured with Nikon Coolpix 4500 on Nikon SMZ1500.

G10 Isofemale lines

Single crosses of individuals of G9 population Set A isofemale lines gave 100% DsRed2 positive progenies (see table 5.9). Crosses A5, A7, A10 and A11 have less than 10 embryos each in total and are therefore regarded as insignificant due to the small sample size. In cross A3, the DsRed2+ pupae that survived were 100% male. Crosses A2, A4, A6 and A8 had over 90% male DsRed2+ pupae. Cross A12 showed a smaller skew in sex ratio, with 67% of the survived DsRed2 pupae male. However, cross A9 had 50% of DsRed2 males. When the total embryos (subtracting non-DsRed2 pupae from the total) were taken into account, the male ratios in crosses A6 and A9 were near equal. Only 8% male was observed in cross A12, and the hatch rate of this cross was only 12%. Crosses A2, A3, and A8 still showed a male biased sex ratio. Crosses A6 and A9 however showed lower male percentage, with 41% of the population being male.

Set B showed that not all pupae observed were DsRed2 positive. A low a hatch rate was observed in crosses B1, B2, B3, B4, B5, B6, and B12, which was less than 40% (table 5.9). Only B3 and B6 had 100% male DsRed2 pupae, but the surviving pupae were only 2 and 3 individuals respectively so the results were discounted. There was a high percentage of males in the surviving DsRed2 pupae in crosses B7 and B10 (97% and 92%, respectively). Crosses B1, B2, B5, B8, B9 and B12 have significant sex ratio skew although not as high as B7 and B10, with percentage of males ranging from 62 to 77%. Cross B11 showed equal sex ratio, while cross B4 showed female bias (29% male). When total embryos were taken into account, crosses B6, B7, B8, B9, and B10 had equal sex ratios. Crosses B1, B2, B4, B5, B8, B11, and B12 resulted in less than 50% males.

Table 5.9: Number of transformed individuals and sex ratios of isofemale lines (G10). Progenies were from G9 Cross A (crosses labelled A2 to A12) and G9 Cross B (cross labelled B1 to B12). *crosses with less than 10 embryos. **Normal sex ratio in total DsRed individuals.

Cross	Total embryos	Percentage Hatch (larvae/embryos)	Percentage Pupation (pupae/larvae)	Percentage DsRed2+ (DsRed2+ pupae/DsRed2+ pupae)	Percentage DsRed2+ Male (DsRed2+ male/DsRed2+ pupae)	Percentage Total DsRed2+ Male (DsRed2+ male/ embryos)	χ^2 value for percentage total DsRed2+ male
A2	69	83% (57/69)	93% (53/57)	100% (53/53)	94% (50/53)	72% (50/69)	9.69 (p = 0.0019)
A3	49	100% (49/49)	100% (49/49)	100% (49/49)	100% (49/49)	100% (49/49)	50 (p= 1.0)
A4	73	95% (69/73)	95% (68/69)	100% (68/68)	95% (64/68)	88% (64/73)	28.88 (p< 0.0001)
A5*	9	56% (5/9)	100% (5/5)	100% (5/5)	100% (5/5)	100% (5/5)	0.00 (p=1.0)
A6**	58	48% (28/58)	93% (26/28)	100% (26/26)	92% (24/26)	41% (24/58)	1.62 (p = 0.2031)
A7*	6	83% (5/6)	100% (5/5)	100% (5/5)	80% (4/5)	67% (4/6)	5.78 (p = 0.0162)
A8	48	100% (48/48)	100%(48/48)	100% (48/48)	98% (47/48)	98% (47/48)	46.08 (p<0.0001)
A9**	39	82%(32/39)	100% (32/32)	100% (32/32)	50% (16/32)	41% (16/39)	1.62 (p= 0.2031)
A10*	3	67% (2/3)	100% (2/2)	100% (2/2)	100% (2/2)	67% (2/3)	5.78 (p= 0.0162)
A11*	3	100% (2/2)	100% (2/2)	100% (2/2)	100% (2/2)	67% (2/3)	5.78 (p=0.0162)
A12*	26	12% (3/26)	100% (3/3)	100% (3/3)	67% (2/3)	8% (2/26)	35.28 (p< 0.0001)
B1	44	55% (24/44)	96% (23/24)	74% (17/23)	76% (13/17)	34% (13/38)	5.12 (p= 0.0237)
B2	62	37% (23/62)	96% (22/23)	82% (18/22)	72% (13/18)	22% (13/58)	15.68 (p<0.0001)
B3	37	5% (2/37)	100% (2/2)	100% (2/2)	100% (2/2)	100% (2/2)	0.00 (p= 1)
B4	62	18% (11/62)	100% (11/11)	64% (7/11)	29% (7/11)	12% (7/58)	28.88 (p< 0.0001)
B5	110	21% (23/110)	96% (22/23)	81% (18/22)	72% (13/18)	12% (13/106)	28.88 (p< 0.0001)
B6	46	7% (3/46)	100% (3/3)	100% (3/3)	100% (3/3)	7% (3/46)	36.98 (p< 0.0001)
B7**	50	62% (31/50)	97% (30/31)	100% (30/30)	97% (29/30)	58% (29/50)	1.28 (p= 0.2579)
B8**	45	69% (31/45)	97%(30/31)	100% (30/30)	67% (20/30)	44% (20/45)	0.72 (p= 0.3961)
B9**	33	73% (24/33)	97% (22/24)	100% (22/22)	77% (17/22)	52% (17/33)	0.08 (p= 0.7773)
B10**	37	68% (25/37)	96% (24/25)	100% (24/24)	92% (22/24)	59% (22/37)	1.62 (p= 0.2031)
B11	40	75% (30/40)	100% (30/30)	100% (30/30)	50% (15/30)	38% (15/40)	2.88 (p= 0.0897)
B12	33	39% (13/33)	100% (13/13)	100% (13/13)	62% (8/13)	24% (8/33)	13.52 (p= 0.0002)

Line 6 A outcross to Rockefeller strain

An outcross to the wild-type strain was undertaken to reduce the deleterious effects of bottlenecking. The progeny of the outcross were then observed for sex ratio distortion. Of the 25 crosses, only 14 crosses produced eggs (table 5.10). All progeny of the crosses were DsRed2 positive. As for sex ratio, each cross resulted in a different ratio. Crosses 1, 4, 5, 6, 7, 8 and 13 produced a 1:1 ratio. However, cross numbers 7, 9 and 14 produced too few eggs or larvae to be considered significant. Cross numbers 2, 3, 5, 11 and 12 produced significantly male biased ratios (88%, 71%, 100% male, respectively). From some of the results, it might seem that when hatch rate is low (below 70%), the male population is significantly above the expected 50% (i.e. over 60%). However, cross number 5, 10 , 11 and 12 suggested otherwise. The first mass cross of Line 6 to the Rockefeller strain also gave a low hatch rate (32%), but the population was not highly male biased, that is only 60% of the population were male (χ^2= 2.19, p= 0.1389). In the second mass cross however, which also had a low hatch rate of 24%, the percentages of males in the survivors were 90% (χ^2= 32.77, p <0.0001, n= 21).

Table 5.10: Number of transformed individuals and sex ratio of Line no 6 to Rockefeller. Progeny were from G9 Cross A. All pupae observed were DsRed2 positive. Cross numbers 1 - 14 were single pair crosses.

Cross	Total embryos	Percentage Hatch (larvae/embryos)	Percentage Pupation (pupae/larvae)	Percentage Male DsRed2 (DsRed2+ male/pupae)	χ^2 value
1	39	87% (34/39)	100% (34/34)	50% (17/34)	0.000 (p=1.0000)
2	49	69% (34/49)	94% (32/34)	13% (4/32)	28.125 (p<0.0001)
3	70	36% (25/70)	84% (21/25)	71% (15/21)	9.184 (p=0.0024)
4	28	93% (26/28)	96% (25/26)	56% (14/25)	0.720 (p=0.3961)
5	35	94% (33/35)	100% (33/33)	67% (22/33)	5.557 (p=0.0184)
6	44	89% (39/44)	100% (39/39)	56% (22/39)	0.821 (p=0.3649)
7	2	100% (2/2)	100% (2/2)	100% (2/2)	-
8	35	97% (34/35)	100% (34/34)	53% (18/34)	0.172 (p=0.6783)
9	74	4% (3/74)	100% (3/3)	100% (3/3)	-
10	56	96% (54/56)	96% (52/54)	62% (32/52)	2.663 (p=0.1027)
11	25	92% (23/25)	96% (22/23)	68% (15/22)	6.610 (p=0.0101)
12	27	89% (24/27)	96% (23/24)	100% (23/23)	50.00 (p<0.0001)
13	76	86% (65/76)	100% (65/65)	58% (38/65)	1.431 (p=0.2316)
14	61	10% (6/61)	100% (6/6)	33% (2/6)	-
Mass Cross 1	140	32% (45/140)	96% (43/45)	60% (26/43)	2 (p=0.1573)
Mass Cross 2	90	24% (22/90)	95%(21/22)	90% (19/21)	32 (p<0.0001)

Line 6 cross with LA513A strain

An outcross to the LA513A strain was undertaken to examine whether LA513A could act as a driver (source of tTAV), to control expression of the inserted construct via the tet system. At total of 246 eggs were hatched without tetracycline (Tet Off), and 163 eggs with tetracycline (Tet On). Both hatch rates of both crosses were high (97% in Tet On and 96% in Tet Off), suggesting no adverse effects of the cross on these two strains (table 5.11). Crosses of Line 6 to LA513A did not produce a gender ratio that was significantly biased. When tetracycline was not present, and the presence of tTA should be allowing the expression of the construct RNAi, the population was 70% male. Chi square test showed that the population skew was significantly different from the expected 50% (At Tet Off, χ^2= 7.904, p= 0.0049). When tetracycline was to suppress the expression of RNAi, the population is 52% male. A Chi square test showed that at Tet On, the population was not significantly different from the expected value of 50% (χ^2= 0.072, p= 0.951). However, accumulation of tTA in the mosquito can cause lethality at late 4^{th} instars and pupal stages. The pupation rates in the two crosses were different, the Tet Off cross pupation rate was significantly lower (53%) then the Tet On cross (99%).

Table 5.11: Number of transformed individuals and sex ratio of Line 6 to LA513A strain. Progenies were from G9 Cross A.

Cross	Total embryos	Percentage Hatch (larvae/embryos)	Percentage Pupation (pupae/larvae)	Percentage male (male/pupae)	χ^2 value
Tet On	246	97% (238/246)	99% (235/238)	52% (122/235)	0.072 (p=0.951)
Tet Off	163	96% (156/163)	53% (83/156)	70% (58/83)	7.904 (p=0.0049)

5.3.4 Inverse PCR of germline transformants

The aim of this experiment was to identify the site of construct integration in the genome. Inverse PCR of *Taq*$^{\alpha}$ *I* did not yield any specific product, but inverse PCR of *Msp I* resulted in a product of about 600 bp. The PCR product was purified and sequenced using the same *Msp I* primers twice in both the 5' and 3' directions. A BLASTn search was carried out using the consensus sequence, which was 613 bp long, against the Broad Institute database. It was found that the sequence had a 98% similiarity to Supercontig 1.6 at position 775690 - 776277 (588 bp) on the minus strand. BLASTn of the same sequence was repeated against the Vectorbase *Ae. aegypti* database. The highest match was to gene ID DV300635.1, with 97.6% identity along 482 bp of the sequence. This DV300635.1 is a gene with unknown functions. FISH-mapping showed that this sequence on Supercontig 1.6 (BAC clone K6.2D1) is mapped to position 23.6 on chromosome 3 (D. Severson, personal communication.) Since the flanking regions expected to be amplified were not found in the above sequence obtained, it was concluded that the amplified product was due to non-specific priming. Technical difficulties therefore prevented the identification of the insertion site of the construct.

5.3.5 Quantitative RT-PCR

The aim of this experiment was to confirm that the expression of *tra-2* is reduced in the transgenic versus the wild-type lines. Normalised delta Ct value in each experimental set was averaged and converted to percent expression. The average *tra-2* expression in Line 6 was reduced by 48.7% (sd= 4.98), relative to wild-type (figure 5.12).

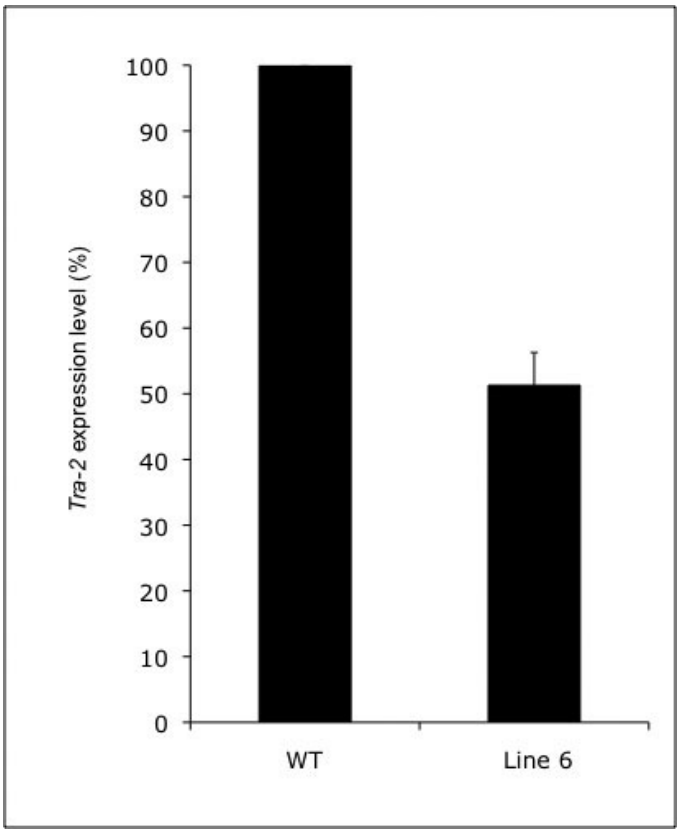

Figure 5.12: Relative expression of *tra-2* in wild-type strain Rockefeller (WT) and transgenic strain Line 6. Error bar shows standard deviation value between 3 biological repeats. Expression in the Line 6 strain was 51.29% of the value in the Rockefeller strain.

5.4 Discussion

5.4.1 Sex ratio distortion by RNAi

5.4.1.1 Transient RNAi

A 14.2% hatch rate for the transformation of *Ae. aegypti* using the *Mariner* transposon was reported (Coates et al., 1998), whereas transformation in the same species using *Hermes* transposon had a 12.97% hatch rate for (Jasinskiene et al., 1998). Therefore, hatching rate of 9% to 42% in the injected embryos for transient RNAi experiments were considered to be fairly good when compared to other mosquito injection experiments.

For DsRed2 transient RNAi experiment, silencing of DsRed2 could be observed phenotypivally, suggesting that RNAi was successful. The RNAi for *tra-2* target site 1 and 2 did not significantly knock down *tra-2* or cause sex ratio distortion. One reason that the knockdown was not successful in *tra-2* target site 2 could be that only certain variant(s) were knocked down. DsRNA against target site 1 however is expected to reduce expression of all variants of the gene. Another possibility is that the amount of dsRNA injected was not enough to knock down the gene to distort the sex ratio. It could also be that knockdown did not last long enough to influence sex determination, given that the RT-PCR carried out still detected *tra-2* transcript. The variation observed in sex ratio amongst different injection batches in the transient RNAi experiments using target 1 site indicates that the effects of this dsRNA on the expression of the target gene were inconsistent. Chi-square tests indicated that the results were not statistically significant. More injections are required for statistically significant skews in the sex ratios at the same concentration and RT-PCR should be done on early instars of these injected individuals that showed no

DsRed2 signal. Quantitative RT-PCR might be useful to assess if there were any difference in levels of expression in the target genes in the injected strains at different developmental stages, but as discussed in Chapter 4, this proved to be difficult since *tra-2* expression is very low.

Direct dsRNA injection is a fairly quick way of determining whether RNAi would work, as no lengthy construction of a plasmid, transformation and selection to deliver the RNAi construct is needed. The problem with direct dsRNA injection however is that the interference effect is not distributed always evenly throughout the organism and dsRNA degrades after some time. In DsRed2 injections, it was concluded that the knockdown effect only lasts up to 3^{rd} instars (roughly 6 days). This is a disadvantage of direct dsRNA injection in comparison to dsRNA-mediated silencing from intergrated DNA, where transcribe dsRNA is transcribed constantly. It is not known for how long *tra-2* expression would be needed to determine sex, but evidence of sexual differentiation is present by the 4^{th} instar (Timmermann and Briegel, 1999). Another disadvantage of direct injection is that the exact amount of dsRNA is not equal in all individuals injected and therefore individuals with partial knockdown are likely to be produced.

5.4.1.2 Germline transformation

Germline transformation was successful in producing a distortion of sex ratio, but the degree of distortion varied, and it is not known why the penetrance of the phenotype varied. By looking at the number of DsRed2 positive individuals, a homozygous line could be selected, although the number of generations that was needed to achieve this was higher than expected, suggesting that there may be a fitness cost. As the line was selected for DsRed2, the sex ratio was more distorted, with some cross progeny have over

90% males. Single pair mating (G10 isofemale) of Line 6 showed that even 100% male could be observed in a population. It was also observed that sex ratios in the crosses were variable. The fact that Line 6 sex ratio was already distorted even without tTAV presence suggests that the construct of Line 6 was leaky. This might be a consequence of the position of insertion in the genome. Therefore, this construct only serves as a test if RNAi of *tra-2* could distort sex, but not an application tool for use in population control as yet.

Could the male biased populations in these crosses be explained by increased female mortality alone? Some of the G10 isofemale line crosses show fairly high mortality between the embryo and adult stages, such that if the ratio of males to the total number of embryos is counted, there is no longer a significant skew in sex ratio toward males; therefore the hypothesis that the male bias could be produced by increased female mortality alone cannot be rejected for these crosses. However, in other crosses the hatch rate / rate of development to the pupal stage was very high, and in these cases female mortality alone could not account for the sex ratio distortion observed. Therefore, the hypothesis that female mortality alone accounts for the male bias can be rejected. The most convincing explanation for the male biased sex ratios observed might be that some females are developing as phenotypic males, as was observed in the *tra-2* heritable RNAi experiments of Fortier and Belote (2000) in *D. melanogaster*.

It was thought that male sexual organs in genotypic females with male phenotype would be sterile and thus may have incomplete or odd development of the sexual organs. Dissections of the male adult mosquitoes indicated however that the sexual organs developed normally. Genotypic females that are phenotypically males could be detected by the absence of male DNA markers. Unfortunately no markers for the *Ae. aegypti* Y / male version of chromosome 1 are available to test this possibility.

None of the crosses produced individual egg batches that had no hatch, which suggests that the putative genetic females that developed as phenotypic males were fertile. In the *D. melanogaster tra-2* heritable knockdown experiments of Fortier and Belote (2000), only male-like phenotypes that retained female genitalia plates were produced, which strongly suggests that the knockdown (which was temperature sensitive) was incomplete. Therefore the question of whether chromosomally female *tra-2* knockdown males would be fertile remained unaddressed; the data reported here suggest that these 'trans-sex' males can indeed be fertile. It may be that RNAi knockdown in the transformed individuals was adequate, given that the CMV promoter was used, which promotes high expression, and not the GAL4 promoter used by Fortier and Belote (2000).

Knockdown of *M. domestica tra-2* gave rise to fertile XX male phenotype (Dübendorfer et al., 2002). Medfly females with *tra* gene knockdown developed as fertile males (Pane et al., 2002), thus suggesting *M. domestica tra-2* also took the over the role of *tra*.

The cross between Line 6 and Rockefeller was to reduce the deleterious effects of bottlenecking and to see if Line 6 crossed to a wild-type strain could still distort the sex ratio. The crosses gave varied results. More crosses need to be carried out at a larger scale to get a deeper and more thorough insight.

The cross between Line 6 and the LA513A line (the latter provides a source of tTAV), produced progeny that did show a difference in sex ratio in the absence of tetracycline. However, when both tTAV and tetracycline were present, the sex ratio was not significantly distorted. It appears, interestingly, tetracycline and tTAV were able to suppress the tetO7 promoter and thus *tra-2* dsRNA expression in the Line 6 construct. But since only one copy was inherited, the sex ratio was only distorted to a small degree. The fact that *tra-2* is a lowly

expressed gene means that the leakiness of the construct was enough to distort the sex ratio even without a source of tTAV, but the data from these crosses suggest that tTAV can still affect the ability of tetO7 to drive gene expression.

The site of insertion of the transgene can affect the levels of protein expression and the phenotype observed. Therefore, it is essential to identify the position of the inserted transgene in the *Ae. aegypti* genome. However, the site of construct integration proved to be difficult to locate, as inverse PCRs failed to specifically amplify the flanking sequences.

5.4.2 Knockdown of *tra-2* expression

Quantitative RT-PCR showed a consistent trend of reduced *tra-2* expression in Line 6 strain A, indicating that the *tra-2* transcription was indeed knocked down by almost 50% in the male pupae in comparison to the wild-type strain. This observed knockdown effect suggests that the male biased ratios observed in Line 6 strain A, are due to *tra-2* RNAi in *Ae. aegypti* and thus that wild-type levels of transcription of this gene are needed to determine a female fate. Moreover, the transcriptional knockdown has a larger knockdown effect at the protein level. The transient knockdown of DsRed2 showed that the transcript levels at 3^{rd} instars were reduced by 50% in a qPCR assay, but the DsRed2 fluorescence in those individuals were undetectable under fluorescent light (data not shown).

5.4.3 Future work

The RNAi by germline transformation was deemed successful, although the exact location of the construct insertion is unknown. Molecular analysis of progeny from the germline transformation is therefore needed. Inverse

PCR could be attempted again by designing more specific primers to the *piggyBac* sequence. Alternatively, specific adaptors could be built to internal *piggyBac* sequences at the 5' and 3' ends to create a highly specific sequence for amplification. Further transformation experiments with the same construct could be undertaken to determine if the insertion site was responsible for the expression leakiness observed, possibly through proximity to other regulatory sequences. Alternatively a new construct could be designed to attempt to eliminate the leaky expression of the dsRNA in the absence of tTAV.

Chapter 6

Summary and General Conclusions

6.1 Identification of male determining genes

Male-specific markers are difficult to identify, because the non-recombining Y chromosome consists mostly of repetitive elements that share high similarity with sequences on other chromosomes. However, Y chromosome DNA sequences have been identified in *An. gambiae* (Krzywinski et al., 2004). Only a few genes vital for male fertility remain functional on the heterochromatic Y in putative 'older' sex chromosome systems such as *Drosophila* (Carvalho et al., 2000; Carvalho et al., 2001). *Aedes aegypti* may have homomorphic chromosomes, but undoubtedly there are differences between the male and female version of chromosome 1 in the M region. Microdissection and microarray techniques carried out in the study were novel, but still were not sufficient to isolate the desired gene(s). The M region is difficult to identify due to the heterochromatin structure in both arms of the female chromosome and one arm of the male chromosome. It may be that the M region prevents chromosomal crossover, keeping the male genes on the male chromosome. With advances in technology such as ever-increasing genome sequencing and assembly capacities, these difficulties may be overcome in the near future. The M gene remains a high priority and its identification would be of great basic and applied interest.

6.2 Sex determining genes in mosquitoes

The identification of mosquito orthologs of *Drosophila* sex determining genes as reported here provided further evidence that the lower level genes in the sex determination cascade were found to be highly conserved, when *An. gambiae, Ae. aegypti* and *D. melanogaster* were compared, further strengthening the theory that the cascade is conserved from the bottom up (Wilkins, 1995). It is hoped that sex determination studies undertaken here in the mosquito *Ae. aegypti* will contribute to the understanding of its sex determination mechanisms.

6.3 Characterisation of *tra-2*

Understanding of the transcript splice pattern exhibited in the variants will contribute to the genome database of *Ae. aegypti* and may be useful to other research in the future. The most important use of this information is utilising sex ratio distortion in the effort to control mosquito populations that cause diseases in humans and animals, which is the main aim of this research. In addition, sex determination in *Ae. aegypti* contributes to understanding the evolution of developmental genetic pathways dictated by sex determination that control key aspects of animal morphology and behaviour. In this project, *tra-2* was found to be a low transcript copy number gene, creating problems in identifying the full transcript sequences of the transcript. It also meant that quantification of this gene was particularly problematic. *D. melanogaster tra-2* was also shown to produce a variety of mRNAs, but only a few have been shown to be functional (Amrein et al., 1988; Mattox et al., 1990). It is therefore probable that *Ae. aegypti tra-2* is as highly complex as *D. melanogaster tra-2*. More studies are needed to understand the functions of the variants identified in this project.

6.4 The potential role of sex ratio distortion in vector control

6.4.1 RNAi in suppressing *tra-2*

Sex ratio distortion could prove to be a way of improving the current strategies for vector control. It provides an alternative to dominant lethal genes for the release of transgenic males that are sterile, as a variant of the Sterile Insect Technique. Suppressing the female phenotype via knocking down the *tra-2* gene is ideal because it does not cause female embryo lethality, as a dominant lethal gene would do in the RIDL strategy; by changing the gender it would boost the number of males that can be artificially reared for release. It also has the advantage of producing all-male broods from crosses with wild females rather than simply sterilizing them, thus producing a continued effect in the next generation. More ambitious strategies might ultimately be envisaged to create an artificial version of the MD phenotype, which would have the potential to drive a population extinct for lack of females from seeding rather than inundative releases. One possibility would be to link the RNAi construct to a meiotic driver, and another would be to design homing endonuclease genes (HEGs) to recognize and disrupt this sequence. However both these strategies are more speculative because they would depend on the identification or creation of a suitable meiotic drive system or protein re-engineering (Burt and Koufopanou, 2004).

6.4.2 Future work utilising the knockdown of *tra-2*

RNAi constructs utilising the *tra-2* gene in this study were not as optimal as expected. Future improvements could be made to the construct. More research

is needed into the workings of *tra-2* in order to see if other parts of the gene could be used as target sequences for RNAi. For example, interaction of TRA-2 with other proteins within the sex determination cascade could be studied. The RNAi construct could also be modified to include the tet system such as done in the LA513A strain. This would then remove the need of crossing the RNAi construct carrying individual to another that carries the tet system, as it was done in this study by crossing strain Line 6 to strain LA513A. The new construct would then ideally be homozygous and RNAi expression would be repressible by using tetracycline.

Adaptation of *tra-2* knockdown to other mosquito species that are medically important such as *An. gambiae*, *Cu. quincifaciatus* and *Ae. albopictus* could also be carried out. *Ae. albopictus* is most likely to succeed using the information from *Ae. aegypti* since it is evolutionarily closer to *Ae. aegypti*. However, unlike the other two mosquito species mentioned, *Ae. albopictus* genome database is very limited. This may prove to be a hurdle when trying to search for the *tra-2* homolog, although degenerate primers may be successful at this distance of species relationship. *Aedes albopictus* is becoming an important mosquito to control due to its wide range of spread. This includes Asia, North and Central America, and many European countries. In 2006, *Ae. albopictus* has been recorded in Gabon (Krueger and Hagen, 2007). *Ae. albopictus* can displace its cousin, *Ae. aegypti*, as has been documented in the USA (Juliano et al., 2004). *Ae. albopictus* is a vector for many diseases, including dengue, filariasis, Chikungunya, several encephalitis viruses and is also feared to be able to carry West Nile Virus. Thus, *Ae. albopictus* is receiving rapid international attention concerning its role as a vector for disease and hence may pose a serious threat to public health in many countries.

6.5 Final conclusion

The well characterized *D. melanogaster* sex determination system has allowed the identification of sex determination genes in the mosquito. This will not only enable us to better understand the long unsolved puzzle of mosquito sex determination and its evolution, but also to utilise it for vector control. This will help us develop long term viable solutions to the global problem of controlling vector borne diseases.

Bibliography

Alphey, L. (2009). Natural and engineered mosquito immunity. *Journal of Biology*, 8(4):40.

Alves, W. and Blair, D. M. (1955). Malaria control in Southern Rhodesia. *Journal of Tropical Medicine and Hygiene*, 58:273 – 280.

Amrein, H., Gorman, M., and Nöthiger, R. (1988). The sex-determining gene *tra-2* of *Drosophila* encodes a putative RNA binding protein. *Cell*, 55:1025 – 1035.

An, W. and Wesink, P. C. (1995). Three protein binding sites form an enhancer that regulates sex and fat body-specific transcription of *Drosophila melanogaster* yolk protein genes. *EMBO Journal*, 14:1221 – 1230.

Anand, A., Villella, A., Ryner, L. C., Carlo, T., Goodwin, S. F., Song, H., Gailey, D. A., Morales, A., Hall, A., Baker, J. C., and Taylor, B. J. (2001). Molecular genetic dissection of the sex-specific and vital functions of the /textitDrosophila melanogaster sex determination gene *fruitless*. *Genetics*, 158:1569 – 1595.

Anderson, J. F. and Horsfall, W. R. (1961). Thermal stress and anomalous development of mosquitoes. I. Effect of constant temperature on dimorphism of adults of *Aedes stimulans*. *Journal of Experimental Zoology*, 154:67 – 107.

Atkinson, M. P., Su, Z., Alphey, N., Alphey, L. S., Coleman, P. G., and Wein, L. M. (2007). Analyzing the control of mosquito-borne diseases by a dominant lethal genetic system. *Proceedings of the National Academy of Sciences*, 104(22):9540–9545.

Bai, A. R. K. and Reddy, C. R. (1977). Laboratory toxicity of some insecticides to honeybee *Apis cerana indica* F. *Journal of Apicultural Research*, 16:161 – 162.

Baker, B. S., Gorman, M., and Martin, I. (1994). Dosage compensation in *Drosophila*. *Annual Review of Genetics*, 28:491 – 521.

Baker, B. S. and Wolfner, M. F. (1988). A molecular analysis of *doublesex*, a bifunctional gene that controls both male and female sexual differentiation in *Drosophila melanogaster*. *Genes and Development*, 2:477 – 489.

Baker, R. H. and Sakai, R. K. (1979). Triploids and male determination in the mosquito, *Anopheles culicifacies*. *Journal of Heredity*, 70:345 – 346.

Barry, J. D., Vargas, R. I., Miller, N. W., and Morse, J. G. (2003). Feeding and foraging of wild and sterile mediterranean fruit flies (diptera: Tephritidae) in the presence of spinosad bait. *Journal of Economic Entomology*, 96:1405 – 1411.

Barton, N. H. (2000). Genetic hitchhiking. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 355(1403):1553 – 1562.

Baselt, R. (2008). *Disposition of toxic drugs and chemicals in man*. Biomedical Publications, Foster City, CA, 8th edition.

Bashaw, G. J. and Baker, B. S. (1997). The regulation of the *Drosophila msl-2* gene reveals a function for sex-lethal in translational control. *Cell*, 89(5):789 – 798.

Belfort, M. and Roberts, R. J. (1997). Homing endonucleases: keeping the house in order. *Nucleic Acids Research*, 25:4626 – 4638.

Bell, A. E. (1954). A gene in *Drosophila melanogaster* that produces all male progeny. *Genetics*, 39:958 – 959.

Belote, J. M. and Baker, B. S. (1982). Sex determination in *Drosophila melanogaster*: Analysis of *transformer-2*, a sex-transforming locus. *Genetic Proceeding of National Academy of Science U.S.A.*, 79:1568 – 1572.

Beltrame, A., Angheban, A., Bisoffi, Z., Monteiro, G., Marocco, S., Calleri, G., Lipani, F., Gobbi, F., Canta, F., Castelli, F., Gulletta, M., Bigoni, S., Punta, V. D., Iacovazzi, T., Romi, R., Nicoletti, L., Clufolini, M. G., Rorato, G., Negri, C., and Viale, P. (2007). Imported chikungunya infection, Italy [letter]. *Emerging Infectious Diseases (http://www.cdc.gov/EID/content/13/8/1264.htm)*.

Bender, M. E. (1969). The toxicity of the hydrolysis and breakdown products of malathion to the fathead minnow (*Pimephales Promelas*, Rafinesque). *Water Research*, 3(8):571 – 582.

Benedict, M. Q., Knols, B. G. J., Bossin, H. C., Howell, P. I., Mialhe, E., Caceres, C., and Robinson, A. S. (2009). Colonisation and mass rearing: learning from others. *Malaria Journal*, 8(Suppl 2):S4.

Benedict, M. Q. and Robinson, A. S. (2003). The first releases of transgenic mosquitoes: an argument for the sterile insect technique. *Trends in Parasitology*, 19(8):349 – 355.

Benjamin, S., Rath, A., Chiang, Y. F., and Lim, H. L. (2005). Efficacy of a *Bacillus thitingiensis israelensis* tablet formulation, vectobac dt for control of dengue mosquito vectors in potable water containers. *Southeast Asian Journal of Tropical Medicine and Public Health*, 36(4):879 – 892.

Bernstein, E., Caudy, A. A., Hammond, S. M., and Hannon, G. J. (2001). Role for a bidentate ribonuclease in the initiation step of rna interference. *Nature*, 409:363 – 366.

Besansky, N. J. and Fahey, G. T. (1997). Utillity of the white gene in estimating phylogenetic relationships among mosquitoes (diptera: Culicidae). *Molecular Biology and Evolution*, 14:442 – 454.

Bhalla, S. C. and Craig, G. B. (1970). Linkage analysis of chromosome of *Aedes aegypti*. *Canadian Journal of Genetics and Cytology*, 12:425 – 435.

Bian, G., Xu, Y., Lu, P., Xie, Y., and Xi, Z. (2010). The endosymbiotic bacterium *Wolbachia* induces resistance to dengue virus in *Aedes aegypti*. *PLoS Pathogens*, 6(4):e1000833.

Binka, F. N., Kubaje, A., Adjuik, M., Williams, L. A., Lengeler, C., Maude, G. H., Armah, G. E., Kajihara, B., Adiamah, J. H., and Smith, P. G. (1996). Impact of permethrin impregnated bednets on child mortality in Kassena-Nankana district, Ghana: a randomized controlled trial. *Tropical Medicine & International Health*, 1(2):137 – 138.

Blandin, S., Shiao, S., Moita, L. F., Janse, C. J., Waters, A. P., Kafatos, F. C., and Levashina, E. A. (2004). Complement-like protein tep1 is a determinant of vectorial capacity in the malaria vector *Anopheles gambiae*. *Cell*, 116:661 – 670.

Bloomquist, J. R., Barlow, R. L., Gilette, J. S., Li, W., and Kirby, M. L. (2002). Selective effects of insecticides on nigrostriatal dopominergic nerve pathways. *Neurotoxicology*, 23(4-5):537 – 544.

Boggs, R. T., Gregor, P., Idriss, S., Belote, J. M., and McKeown, M. (1987). Regulation of sexual differentiation in *D. melanogaster* via alternative splicing of RNA from the *transformer* gene. *Cell*, 50:739 – 747.

Bolshakov, V. N., Topalis, P., Blass, C., Kokoza, E., della Torre, A., Kafatos, F. C., and Louis, C. (2002). A comparative genomic analysis of two distant Diptera, the fruit fly, *Drosophila melanogaster*, and the malaria mosquito, *Anopheles gambiae*. *Genome*, 12:57 – 66.

Braig, H. R. and Yan, G. (2001). The spread of genetic constructs in natural insect populations. In Letourneau, D. and Burrows, B. E., editors, *Genetically engineered organims assessing environmental and human health effects*, page 438. CRC Press, Boca Raton, FL.

Braks, M. A. H., Honorio, N. A., Lourenco-De-Oliveira, R., Juliano, S. A., and Lounibos, L. P. (2003). Convergent habitat segregation of *Aedes aegypti* and *Aedes albopictus* (Diptera: Culicidae) in southeastern Brazil and Florida. *Journal of Medical Entomology*, 40:785 – 794.

Bridges, C. B. (1916). Non-disjunction as proof of the chromosome theory of heredity. *Genetics*, 1:1 – 52.

Brink, C. J. H. (1958). Malaria control in the northern Transvaal. *South African Medical Journal*, 32:800 – 808.

Broderick, N. A., Raffa, K. F., and Handelsman, J. (2006). Midgut bacteria required for *Bacillus thuringiensis* insecticidal activity. *Proceedings of the National Academy of Science USA*, 103(41):15196 – 15199.

Brogdon, W. G. and McAllister, J. C. (1998). Insecticide resistance and vector control. *Emerging Infectious Diseases*, 4(4):605 –613.

Brosseau, G. E. (1960). Genetic analysis of the male fertility factors on the Y chromosome of *Drosophila melanogaster*. *Genetics*, 45:257 – 274.

Brown, A. E., Crisanti, A., and Catteruccia, F. (2003). Comparative analysis of DNA vectors at mediating RNAi in *Anopheles* mosquito cells and larvae. *The Journal of Experimental Biology*, 206:1817 – 1823.

Brown, E. H. and King, R. C. (1961). Studies on the expression of the *transformer* gene in *Drosophila melanogaster*. *Genetics*, 46:143 – 156.

Bryne, K. and Nichols, R. A. (1999). *Culex pipiens* in London Underground tunnels: differentiation between surface and subterranean populations. *Heredity*, 82:7 – 15.

Burdick, G. E. (1967). Use of bioassays in determining levels of toxic wastes harmful to aquatic organisms. *American Fishery Society Special Publication*, 4:3 – 12.

Burt, A. and Koufopanou, V. (2004). Homing endonuclease genes: the rise and fall and rise again of a selfish elements. *Current Opinion Genetical Development*, 14(6):609 – 615.

Burtis, K. C. and Baker, B. S. (1989). *Drosphila doublesex* gene controls somatic sexual differentiation by producing alternatively spliced mRNAs encoding related sex-specific polypeptides. *Cell*, 56:997 – 1010.

Caceres, C., Fisher, K., and Rendon, P. (2000). *Mass rearing of the medfly temperature sensitive lethal genetic sexing strain in Guatemala.*, chapter Area-wide management of fruit flies and other major insect pests., pages 551 – 558. Universiti Sains Malaysia Press.

Carvalho, A. B. (2002). Origin and evolution of the *Drosophila* Y chromosome. *Genetics & Development*, 12:664 – 668.

Carvalho, A. B., Dobo, B. A., Vibranovski, M. D., and Clark, A. G. (2001). Identification of five new genes on the y chromosome of *Drosophila melanogaster*. *Proceedings of the National Academy of Science USA*, 98:13225 – 13230.

Carvalho, A. B., Lazzaro, B. P., and Clark, A. G. (2000). Y chromosomal fertility factors *kl-2* and *kl-3* of *Drosophila melanogaster* encode dynein heavy chain polypeptides. *Proceedings of the National Academy of Science USA*, 97:13239 – 13244.

Catteruccia, F., Benton, J. P., and Crisanti, A. (2005). An *Anopheles* transgenic sexing strain for vector control. *Nature Biotechnology*, 23:1414 – 1417.

Catteruccia, F., Godfray, C. J., and Crisanti, A. (2003). Impact of genetic manipulation on the fitnesss of *Anopheles stephensi* mosquitoes. *Science*, 299(1610):1225 – 1227.

Catteruccia, F., Nolan, T., Loukeris, T. G., Blass, C., Savakis, C., Kafatos, F. C., and Crisanti, A. (2000). Stable germline transformation of the malaria mosquito *Anopheles stephensi*. *Nature*, 405:959 –962.

Centre for Disease Control and Prevention (2006a). Chikungunya fever in India. travelers' health outbreak notice. *http://wwwnc.cdc.gov/travel/content/outbreak-notice/chikungunya-fever.aspx*.

Centre for Disease Control and Prevention (2006b). History of dengue. *http://www.cdc.gov/ncidod/dvbid/dengue/index.htm#history*.

Cha, S., Mori, A., Chadee, D. D., and Severson, D. W. (2006). Cage trails using an endogenous meiotic drive gene in the mosquito *Aedes aegypti* to promote population replacement. *American Journal of tropical medicine and hygiene*, 74(1):62 – 68.

Chan, K. L., Ho, B. C., and Chan, Y. C. (1971a). *Aedes aegypti* (L.) and *Aedes albopictus* Skuse in Singapore City. 2. larval habitats. *Bulletin World Health Organisation*, 44:629 – 633.

Chan, Y. C., Chan, K. L., and Ho, B. C. (1971b). *Aedes aegypti* (L.) and *Aedes albopictus* Skuse in Singapore City. 1. distribution and density. *Bulletin World Health Organisation*, 44:617 – 627.

Charlat, S., Hurst, G. D. D., and Mercot, H. (2003). Evolutionary consequences of *Wolbachia* infections. *Trends in Genetics*, 19:217 – 223.

Charlesworth, B. (1996). The evolution of chromosomal sex determinationa and dosage compensation. *Current Biology*, 6:149 – 162.

Charlesworth, B. and Charlesworth, D. (1997). Rapid fixation of deleterious alleles can be caused by Muller's ratchet. *Genetic Research*, 70:63 – 73.

Chayajabera, S., Sobti, S. K., Payne, D., and Braga, F. (1975). Malaria situation in Botswana: report on visit (December 1973 - October 1974). Technical report, WHO, Geneva.

Chen, C., Huang, H., Ward, C. M., Su, J. T., Schaeffer, L. V., Guo, M., and Hay, B. A. (2007a). A synthetic maternal-effect selfish genetic element drives population replacement in drosophila. *Science*, 316:597 – 600.

Chen, C. D., Nazni, W. A., Lee, H. L., and Sofian-Azirun, M. (2005). Susceptibility of *Aedes aegypti* and *Aedes albopictus* to temephos in four study sites in Kuala Lumpur City Center and Selangor State, Malaysia. *Tropical Biomedicine*, 22(2):207 – 216.

Chen, S., Dai, S., Lu, K., and Chang, C. (2007b). Female-specific *doublesex* dsRNA interrupts yolk protein gene expression and reproductive ability in oriental fruit fly, *Bacrocera dorsalis* (Hendel). *Insect Biochemistry and Molecular Biology*, 38:155 – 165.

Chevalier, B. S. and Stoddard, B. L. (2001). Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. *Nucleic Acids Research*, 29:3757 – 3774.

Chou, W. H., Huber, A., Schulz, S., Bentrop, J., Paulsen, R., and Britt, S. G. (1998). A binary switch signal from the R7 cell that control opsin expression in the R8 cell. *Annual Drosophila Research Conference*, 39:120.

Chow, V. T. K., Chan, Y. C., Yong, R., Lee, K. M., Lim, L. K., Chung, Y. K., Lam-Phua, S. G., and Tan, B. T. (1998). Monitoring of dengue viruses in field-caught *Aedes aegypti* and *Aedes albopictus* mosquitoes by a type-specific polymerase chan reaction and cycle sequencing. *American Journal of Tropical Medicine and Hygieneournal of tropical medicine and hygiene*, 58(5):578 – 586.

Chua, K. B., Chua, I. L., Chua, I. E., and Chua, K. H. (2005). Effect of chemical fogging on immature *Aedes* mosquitoes in natural field conditions. *Singapore Medical Journal*, 46(11):639 – 644.

Chuah, M. L. K. and Yap, H. H. (1984). Studies on biological control potentials of *Toxorhynchites splendens* Diptera: Culicidae. *Tropical Biomedicine*, 1:145 – 150.

Cline, T. W. (1976). A sex-specific, temperature-sensitive maternal effect of the *daugtherless* mutation of *Drosophila melanogaster*. *Genetics*, 84:723 – 742.

Cline, T. W. (1978). Two closely linked mutations in *Drosophila melanogaster* that are lethal to opposite sexes and interact with *daughterless*. *Genetics*, 1990:683 – 698.

Cline, T. W. (1984). Autoregulatory functioning of a *Drosophila* gene product that establishes and maintains the sexually determined state. *Genetics*, 107:231 – 277.

Cline, T. W. and Meyer, B. J. (1996). Vive la difference: males vs females in flies vs worms. *Annual Review of Genetics*, 30:637 – 702.

Coates, C. J., Jasinskiene, N., Miyashiro, L., and James, A. A. (1998). Mariner transposition and transformation of the yellow fever mosquito, *Aedes aegypti*. *Proceedings of the National Academy of Science USA*, 95:3748 – 3751.

Collins, F. H., Mendez, M. A., Rasmussen, M. O., Mehaffey, P. C., Besansky, N. J., and Finnerty, V. (1987). A ribosomal RNA gene probe differentiates member's species of *Anopheles gambiae* complex. *American Journal of tropical medicine and hygiene*, 37:37 – 41.

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

Coto, M. M., Lazcano, J. A., de Fernandez, D. M., and Soca, A. (2000). Malathion resistance in *Aedes aegypti* and Culex quinquefasciatus after its use in *Aedes aegypti* control programs. *Journal of America Mosquito Control Association*, 16:324 – 330.

Craig, G. B. (1965). Genetic control of thermally-induced sex reversal in *Aedes aegypti*. *Proceedings of International Congress of Entomology*, 12th:263.

Crickmore, N. (2010). *Bacillus thuringiensis* toxin nomenclature. *Website: http://www.lifesci.sussex.ac.uk/home/Neil_Crickmore/Bt/index.html. Retrieved: 2010-9-30.*

Cui, F., Raymond, M., and Qiao, C. L. (2006). Insecticide resistance in vector mosquitoes in China. *Pest Management Science*, pages 1013 – 1022.

Cummings, C. A. and Cronmiller, C. (1994). The *daughterless* gene functions together with Notch and Delta in the control of ovarian follicle development in *Drosophila*. *Development*, 120(2):381 – 394.

Curtis, C. F. (2002). Should the use of DDT be revived for malaria vector control? *Biomedica*, 22(4):455 – 461.

Cyranoski, D. (2008). Sterile mosquitoes near take-off. malaysia prepares to fight dengue with GM insects. *Nature News*, 453(435):http://www.nature.com/news/2008/080521/full/453435a.html.

da Cunha, M. P., Lima, J. B., Brogdon, W. G., Moya, G. E., and Valle, D. (2005). Monitoring of resistance to the pyrethroid cypermethrin in Brazilian *Aedes aegypti* (Diptera: Culicidae) populations collected between 2001 and 2003. *Memorias do Instituto Oswaldo Cruz*, 100(4):441 – 444.

Dadd, R. H. and Kleinjan, J. E. (1976). Chemically defined dietary media for larvae of the mosquito *Culex pipiens* (Diptera: Culicidae): effects of colloid texturizers. *Journal of Medical Entomology*, 13:285 – 291.

Dadd, R. H., Kleinjan, J. E., and Sneller, V. P. (1977). Development of several species of mosquito larvae in fully defined dietary media: preliminary evaluation. *Mosquito News*, 37:699 – 703.

D'Alessandro, U., Olaleye, B., McGuire, W., Langerock, P., Bennett, S., Aikins, M. K., Thomson, M. C., Cham, M. K., Cham, B. K., and Greenwood, B. M. (1995). Mortality and morbidity from malaria in Gambian children after introduction of an impregnated bednet programme. *Lancet*, 345(8948):479 – 483.

del Valle, J. (2003). New word screwworm *(Cochliomyia hominivorax)* in Mexico. *Disease information*, 16:49 – 50.

Demir, E. and Dickson, B. J. (2005). *fruitless* splicing specifies male courtship behaviour in *Drosophila*. *Cell*, 121(5):785 – 94.

Dübendorfer, A., Hediger, M., Burghardt, G., and Bopp, D. (2002). *Musca domestica*, a window on the evolution of sex-determining mechanisms in insects. *International Journal of Developmental Biology*, 46:75 – 79.

Dübendorfer, A., Hilfiker-Kleiner, D., and Nöthiger, R. (1992). Sex determination mechanisms in dipteran insects: the case of *Musca domestica*. *Seminar Developmental Biology*, 3:349 – 356.

Eanes, W. F., Kirchner, M., and Yoon, J. (1993). Evidence for adaptive evolution of the G6pd gene in the *Drosophila melanogaster* and *Drosophila simulans* lineages. *Proceedings of the National Academy of Science USA*, 90:7475 – 7479.

Editorial (1975). Oh, New Delhi; oh Geneva. *Nature*, 256:355 – 357.

Edwards, D. (2006). Reregistration eligibility decision for malathion. *US Environmental Protection Agency - Prevention, Pesticides and Toxic Substances EPA 738-R-06-030 journal*, 9.

Edwards, F. W. (1932). Genera Insectorium. Diptera, Fam. Culicidae. Fascicle 194.

Erickson, J. W. and Cline, T. W. (1991). Molecular nature of the *Drosophila* sex determination signal and its link to neurogenesis. *Science*, 251:1071 – 1074.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998). Potent and specific genetic interference by double stranded RNA in *Caenorhabditis elegans*. *Nature*, 391(6669):806 – 811.

Focks, D. A., Brenner, R. J., Hayes, J., and Daniels, E. (2000). Tansmission thresholds of dengue in terms of *Aedes aegypti* pupae per person with discussion of their utility in source reduction efforts. *American Journal of Tropical Medicine and Hygiene*, 62:11 – 18.

Fontenille, D. and Simard, F. (2004). Unraveling complexities in human malaria transmission dynamics in Africa through a comprehensive knowledge of vector populations. *Comparative Immunology, Microbiology & infectious Diseases*, 27:257 – 375.

Fortier, M. and Belote, J. M. (2000). Temperature-dependent gene silencing by an expressed inverted repeat in *Drosophila*. *Genesis*, 26:240 – 244.

Franz, A. W. E., Sanchez-Vargas, I., Adelman, Z. N., Blair, C. D., Beaty, B. J., James, A. A., and Olson, K. E. (2006). Engineering RNA interference-based resistance to dengue virus type 2 in genetically modified *Aedes aegypti*. *Proceedings of the National Academy of Science USA*, 103(11):4198 – 4203.

Fu, G., Lees, R. S., Nimmo, D., Aw, D., Jin, L., Gray, P., Berendonk, T. U., White-Cooper, H., Scaife, S., Phuc, H. K., Marionotti, O., Jasinskiene, N., James, A. A., and Alphey, L. (2010). Female-specific flightless phenotype for mosquito control. *Proceedings of the National Academy of Sciences*, 107(10):4550 – 4554.

Fujihara, T., Kawabe, M., and Oishi, K. (1978). A sex transformation gene in *Drosophila melanogaster*. *Journal of Heredity*, 69:229 – 236.

Gailey, D. A., Billeter, J. C., Liu, J. H., Bauzon, F., Allendorfer, J. B., and Goodwin, S. F. (2006). Functional conservation of the *fruitless* male sex-determination gene across 250 myr of insect evolution. *Molecular Biology and Evolution*, 23(3):633 – 643.

Galun, R., Avi-Dor, Y., and Bar-Zeev, M. (1963). Feeding response in *Aedes aegypti*: stimulation by adenosine triphosphate. *Science*, 142:1674 – 1675.

Garcia-Bellido, A. and de Celis, J. F. (2009). The complex tale of the *achaete-scute* complex: A paradigmatic case in the analysis of gene organization and function during development. *Genetics*, 182(3):631 – 639.

Gatti, M. and Pimpinelli, S. (1992). Functional elements in *Drosophila melanogaster* heterochromatin. *Annual Review of Genetics*, 26:239 – 275.

Gepner, J. and Hays, T. S. (1993). A fertility region on the Y chromosome of *Drosophila melanogster* encodes a dynein microtublue motor. *Proceedings of the National Academy of Science USA*, 90:11132 – 11136.

Ghosh, A. K., Ribolla, P. E. M., and Jacobs-Lorena, M. (2001). Targeting plasmodium ligands on mosquito salivary glands and midgut with a phage display peptide library. *Proceedings of the National Academy of Science USA*, 98(23):13278 – 13281.

Gilbert, S. F. (2000). *Developmental Biology.*, chapter 17: Sex determination., page 547. Sinauer Associates, sixth edition edition.

Gilchrist, B. M. and Haldane, J. B. S. (1947). Sex lineage and sex determination in a mosquito, *Culex molestus*. *Hereditas (Lund)*, 33:175 – 190.

Goeldi, E. A. (1905). Os mosquitos do pará.reunião de quatro trabalhos sobre os mosquitos indígenas, principalmente as espécies que molestam o homem. *Memorias do Museu Goeldi*, 4:1 – 152.

Goh, K. T. (1998). Dengue - a remerging infectious disease in Singapore. In Goh, K. T., editor, *Dengue in Singapore*, page 3349. Singapore: Institute of environmental epidemiology, Ministry of the environment.

Gong, P., Epton, M., Fu, G., Scaife, S., Hiscox, A., Condon, K. C., Condon, G. C., Morrison, N. I., Kelly, D. W., Dafa'alla, T., Coleman, P. G., and Alphey, L. (2005). A dominant lethal genetic system for autocidal control of the mediterranean fruitfly. *Nature Biotechnology*, 23:453 – 456.

Goralski, T. J., Edström, J., and Bruce, S. B. (1989). The sex determination locus *transformer-2* of Drosophila encodes a polyopeptide with similarity to RNA binding proteins. *Cell*, 56:1011 – 1019.

Gorman, M., Franke, A., and Baker, B. S. (1995). Molecular characterization of the *male-specific lethal-3* gene and investigations of the regulation of dosage compensation in Drosophila. *Development*, 121:463 – 475.

Gossen, H. and Bujard, H. (1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proceedings of the National Academy of Sciences*, 89:5547 – 5551.

Goto, A., Blandin, S., Royet, J., Reinchhart, J., and Levashina, E. A. (2003). Silencing of Toll pathway components by direct injection of double-stranded RNA into Drosophila adult flies. *Nucleic Acids Research*, 31(22):6619 – 6623.

Gould, F. and Schliekelman, P. (2004). Population genetics of autocidal control and strain replacement. *Annual Review of Entomology*, 49:193 – 217.

Graham, D. H., Holmes, J. L., and IV, W. C. B. (2004). Identification of quantitative trait loci affecting sex determination in the eastern treehole mosquito (*Ochlerotatus triseriatus*). *Journal of Heredity*, 95(1):35 – 45.

Graham, P., Penn, J. K., and Schedl, P. (2002). Masters change, slaves remain. *Bioessays*, 25:1 – 4.

Gubler, D. J. (1981). Transmission of Ross River virus by *Aedes polynesiensis* and *Aedes aegypti*. *American Journal of Tropical Medicine and Hygiene*, 30(6):1303 – 1306.

Habluetzel, A., Diallo, D. A., Esposito, F., Lamizana, L., Pagnoni, F., Lengeler, C., Traorè, C., and Cousens, S. N. (1997). Do insecticide-impregnated curtains reduce all- cause child mortality in Burkina Faso? *Tropical Medicine & International Health*, 2(9):855 – 862.

Hamilton, A. J. and Baulcombe, D. C. (1999). A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science*, 286:950 – 952.

Hammond, S. M., Bernstein, E., Beach, D., and Hannon, G. J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature*, 202:293 – 296.

Hamscher, G., Sczesny, S., Hoper, H., and Nau, H. (2002). Determination of persistent tetracycline residues in soil fertilized with liquid manure by high-performance liquid chromatography with electrospray ionization tandem mass spectrometry. *Analytical Chemistry*, 74(7):1509 – 1518.

Handler, A. M., McComb, S. D., Fraser, M. J., and Saul, S. H. (1998). The lepidopteran transposon vector, piggyBac, mediates germ-line transformation in the Mediterrranean fruit fly. *Proceedings of the National Academy of Science USA*, 95(13):7520 – 7525.

Harbach, R. E. (2007). *The Culicidae (Diptera): a review of taxnomy, classification and phylogeny*. Magnolia Press.

Harbach, R. E. and Kitching, I. J. (1998). Phylogeny and classification of the Culicidae (Diptera). *Systematic Entomology*, 23:327 – 370.

Hashimoto, H. (1933). The role of the W-chromosome in the sex determination of *Bombyx mori* comment: (in japanese). *Japanese Journal of Genetics*, 8:245 – 247.

Hawley, W. A. (1988). The biology of *Aedes albopictus*. *Journal of America Mosquito Control Association*, 4 (Suppl 1):1 – 40.

Heinrich, J. and Scott, M. A. (2000). A repressible female-specific lethal genetic system for making transgenic insect strains suitable for a sterile-release program. *Proceedings of the National Academy of Science USA*, 97:8229 – 8232.

Helinski, M., Parker, A., and Knols, B. (2006). Radiation-induced sterility for pupal and adult stages of the malaria mosquito *Anopheles arabiensis*. *Malaria Journal*, 5(1):41.

Hemingway, J., Hawkes, N. J., McCarroll, L., and Ranson, H. (2004). The molecular basis of insecticide resistance in mosquitoes. *Insect Biochemistry and Molecular Biology*, 34:653 – 665.

Hendrichs, J., Franz, G., and Rendon, P. (1995). Increased effectiveness and applicability of the sterile insect technique through male-only releases for control of Mediterranean fruit-flies during fruiting seasons. *Journal of Applied Entomology*, 119:371–377.

Hendrix, C. M., Brunner, C. J., and Bellamy, L. K. (1986). Natural transmission of *Dirofilaria immitis* of *Aedes aegypti*. *Journal of America Mosquito Control Association*, 2(1):48 – 51.

Hertig, M. and Wolbach, S. B. (1924). Studies on Rickettsia-like microorganisms in insects. *Journal of Medical Research*, 44:329 – 374.

Hickey, W. A. and Craig, G. B. (1966). Genetic distortion of sex ratio in a mosquito *Aedes aegypti*. *Genetics*, 52:1177 – 1196.

Hilfiker, A., Hilfiker-Kleiner, D., Pannuti, A., and Lucchesi, J. C. (1997). *mof*, a putative acetyl transferase gene related to the Tip60 and MOZ human genes and to the SAS genes of yeast, is required for dosage compensation in *Drosophila*. *EMBO Journal*, 16:2054 – 2060.

Hoch, A. L., Gargan, T. P., and Bailey, C. L. (1985). Mechanical transmission of Rift Valley fever virus by hematophagous diptera. *American Journal of Tropical Medicine and Hygiene*, 34:188 – 193.

Hoffman, A. A. and Turelli, M. (1997). *Influential Passengers: Inherited Microorganisms and Arthropod Reproduction*, pages 42 – 80. Oxford University Press.

Hofte, H., de Greve, H., Seurinck, J., Jansens, S., Mahillon, J., Ampe, C., Vandekerckhove, J., Vanderbruggen, H., and van Montagu, M. (1986). Structural and functional analysis of a cloned delta endotoxin of *Bacillus thuringiensis berliner* 1715. *European Journal of Biochemistry*, 161(2):273 – 280.

Horn, C. and Wimmer, E. A. (2003). A transgene-base, embryo-specific lethality system for insect pest management. *Nature Biotechnology*, 21:64 – 70.

Hosoi, T. (1959). Indentification of blood components which induce gorging of the mosquito. *Journal of Insect Physiology*, 3:191 – 218.

Inoue, K., Hoshijima, K., Higushi, I., Sakamoto, H., and Shimura, Y. (1992). Binding of the *Drosophila* transformer and transformer-2 proteins to the regulatory elements of *doublesex* primary transcript for sex-specific RNA processing. *Proceedings of the National Academy of Science USA*, 89:8092 – 9692.

Irvin, N., Hoddle, M. S., O'Brochta, D. A., Carey, B., and Atkinson, P. W. (2004). Assessing fitness costs for transgenic *Aedes aegypti* expressing the GFP marker and transposase genes. *Proceedings of the National Academy of Science USA*, 101(3):891 – 896.

Ito, H., Fujitani, K., Usui, K., Shimizu-Nishikawa, K., Tanaka, S., and Yamamoto, D. (1996). Sexual orientation in *Drosophila* is altered by the satori mutation in the sex-determination gene *fruitless* that encodes a zinc finger protein with a BTB domain. *Proceedings of the National Academy of Science USA*, 93:9687 – 9692.

Ito, J., Ghosh, A. K., Moreira, L. A., Wimmer, E. A., and Jacobs-Lorena, M. (2002). Transgenic anopheline mosquitoes impaired in transmission of a malaria parasite. *Nature*, 417:452 – 455.

Jasinskiene, N., Coates, C. J., Benedict, M. Q., Cornel, A. J., Rafferty, C., Salazar-Rafferty, C., James, A. A., and Collins, F. H. (1998). Stable, transposon mediated transformation of the yellow fever mosquito, *Aedes aegypti*, using the Hermes element from the housefly. *Proceedings of the National Academy of Science USA*, 95(5):486 – 498.

Jirakanjanakit, N., Rongnoparut, P., Saengtharatip, S., Chareonviriyaphap, T., Duchon, S., Bellec, C., and Yoksan, S. (2007). Insecticide susceptible/resistance status in *Aedes* (Stegomyia) *aegypti* and Aedes (Stegomyia) *albopictus* (Diptera: Culicidae) in Thailand during 2003 - 2005. *Journal of Economic Entomology*, 100:545 – 550.

Johnson, D. H. (1990). Molecular cloning of DNA from specific chromosomal regions by microdissection and sequence-independent amplification of DNA. *Genomics*, 6(2):243 – 251.

Juliano, S. A., Lounibos, L. P., and O'Meara, G. F. (2004). A field test for competitive effects of *Aedes albopictus* on *A. aegypti* in South Florida: differences between sites of coexistence and exclusion? *Oecologia*, 139(4):583 – 593.

Kaiser, P. E., Bailey, D. L., Lowe, R. E., Seawright, J. A., and Dame, D. A. (1979). Mating competitiveness of chemosterilized males of a genetic sexing strain of *Anopheles albimanus* in laboratory and field tests. *Mosquito News*, 39:768 – 775.

Kambris, Z., Cook, P. E., Phuc, H. K., and Sinkins, S. P. (2009). Immune activation by life-shortening *Wolbachia* and reduced filarial competence in mosquitoes. *Science*, 326(5949):134 – 136.

Kao, F. and Yu, J. (1991). Chromosome microdissection and cloning in human genome and genetic disease analysis. *Proceedings of the National Academy of Science USA*, 88:1844 – 1848.

Karpala, A. J., Doran, T. J., and Bean, A. G. D. (2005). Immune responses to dsRNA: Implications for gene silencing technologies. *Immunology and Cell Biology*, 83:211 – 216.

Kay, B. H. (1980). Transovarial transmission of Murray Valley encephalitis virus by *Aedes aegypti* (L.). *Australian Journal of Experimental Biology and Medical Sciences*, 58(5):1 – 4.

Keyes, L. N., Cline, T. W., and Schedl, P. (1992). The primary sex detemination signal of *Drosophila* acts at the level of transcription. *Cell*, 68:933 – 943.

Kimura, M. (1983). *The neutral theory of molecular evolution*. Cambridge University Press, New York.

Kiszewksi, A., Mellinger, A., Spielman, A., Malaney, P., Sachs, S. E., and Sachs, J. (2004). A global index representing the stability of malaria transmission. *American Journal of tropical medicine and hygiene*, 70(5):486 – 498.

Klasson, L., Kambris, Z., Cook, P. E., Walker, T., and Sinkins, S. P. (2009). Horizontal gene transfer between *Wolbachia* and the mosquito *Aedes aegypti*. *BMC Genomics*, 10(33):1471 – 2164.

Knipling, E. F. (1955). Possiblities of insect population control through the use of sexually sterile males. *Journal of Economic Entomology*, 48:459 – 462.

Koella, J. C., Lorenz, L., and Bargielowski, I. (2009). Chapter 12: Microsporidians as evolution - proof agents of malaria control? In Webster, J. P., editor, *Natural History of Host-Parasite Interactions*, volume 68 of *Advances in Parasitology*, pages 315 – 327. Academic Press.

Kovalick, G. E. and Beckingham, K. (1992). Calmodulin transcription is limited to the nervous system during *Drosophila* embryogenesis. *Development Biology*, 150(1):33 – 46.

Kozak, M. (1987). An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Research*, 15(20):8125 – 8148.

Kraaijeveld, K. and Chapman, T. (2004). Effects of male sterility on female remating in the Mediterranean fruitfly, *Ceratitis capitata*. *Proceedings of Royal Society London. Biology. (Supplementary)*, 71:S209 – S211.

Kristan, M., Fleischmann, H., della Torre, A., Stich, A., and Curtis, C. F. (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*, 17:326 – 332.

Kroeger, A., Lenhart, A., Ochoa, M., Villegas, E., Levy, M., Alexander, N., and McCall, P. J. (2006). Effective control of dengue vectors with curtains and water container covers treated with insecticide in Mexico and Venezuela: cluster randomised trials. *British Medical Journal*, 332:1247 – 1252.

Krueger, A. and Hagen, R. M. (2007). Short communication: First record of *Aedes albopictus* in Gabon, Central Africa. *Tropical Medicine & International Health*, 12(9):1105 – 1107.

Krümmerer, K. (2003). Significance of antibiotics in the environment. *The British Society for Antimicrobial Chemotherapy*, 52:5 – 7.

Krzywinski, J., Nusskern, D. R., Kern, M. K., and Besansky, N. J. (2004). Isolation and characterisation of Y chromosome sequences from african malaria mosquito *Anopheles gambiae*. *Genetics*, 166:1291 – 1302.

Kuhn, S., Sievert, V., and Traut, W. (2000). The sex-determining gene *doublesex* in the fly *Megaselia scalris*:conserved structure and sex-specific splicing. *Genome*, 43:1011 – 1020.

Kuroda, M. I., Kernan, M. J., Kreber, R., Ganetzky, B., and Baker, B. S. (1991). The *maleless* protein associates with the X chromosome to regulate dosage compensation in *Drosophila*. *Cell*, 66:936 – 947.

Lacroix, R., Delatte, H., Hue, T., and Reiter, P. (2009). Dispersal and survival of male and female *Aedes albopictus* (diptera: Culicidae) on Reunion Island. *Journal of Medical Entomology*, 46(5):1117 – 1124.

LaMer, V. K., Hochberg, S., Hodges, K., Wilson, I., Fales, J. A., and Latta, R. (1947). The influence of the particle size of homogeneous insecticidal aerosols on the mortality of mosquitoes in confined atmospheres. *Journal of Colloidal Science*, 2:539 – 549.

Lance, D. R., McInnis, D. O., Rendon, P., and Jackson, C. G. (2000). Courtship among sterile and wild *Ceratitis capitata* (diptera: Tephritidae) in field cages in Hawaii and Guatemala. *Annual Entomological Society of America*, 93(5):1179 – 1185.

Laven, H. (1967). Eradication of *Culex pipiens fatigans* through cytoplasmic incompatibility. *Nature*, 216:383 – 384.

Lee, G. and Hall, J. C. (2000). A newly uncovered phenotype associated with the *fruitless* gene of *Drosophila melanogaster*: Aggression-like head interactions between mutant males. *Behavioural Genetics*, 30:263 – 275.

Lee, G. and Hall, J. C. (2001). Abnormalities of male-specific FRU protein and serotonin expression in the CNS of fruitless mutants in *Drosophila*. *Journal of Neuroscience*, 21:513 – 526.

Lemaux, P. (2008). Genetically engineered plants and foods: A scientist's analysis of the issues (part i). *Annual Review of Plant Biology*, 59:771 – 812.

Lengeler, C. (2004). Insecticide-treated bed nets and curtains for preventing malaria. *The Cochrane Database of Systematic Reviews*, Issue 2.

Lima, J. B., Da-Cunha, M. P., Silva, R. C. D., Galardo, A. K., Sda, S. S., Ramos, R. P., and Valle, D. (2003). Resistance of *Aedes aegypti* to organophosphates in several municipalities in the State of Rio de Janeiro and Espírito Santo, Brazil. *American Journal of Tropical Medicine and Hygiene*, 68(3):329 – 333.

Lindsay, S. W., Snow, R. W., Broomfield, G. I., Janneh, M. S., Wirtz, R. A., and Greenwood, B. M. (1989). Impact of permethrin-treated bednets on malaria transmission by the *Anopheles gambiae* complex in The Gambia. *Medical and Veterinary Entomology*, 3(4):263 – 271.

Lobo, N. F., Clayton, J. R., Fraser, J. F., Kafatos, F. C., and Collins, F. H. (2006). High efficiency germ-line transformation of mosquitoes. *Nature Protocols*, 1:1312 – 1317.

Lofgren, C. S. (1970). Ultralow volume applications of concentrated insecticides in medical and veterinary entomology. *Annual Review of Entomology*, 15:321 – 342.

Lucchesi, J. C. (1978). Gene dosage compensation and the evolution of sex chromosomes. *Science*, 202:711 – 716.

Lutz, C. S. (2008). Alternative polyadenylation: A twist on mRNA 3' end formation. *ACS Chemical Biology*, 3(10):609 – 617.

Lyman, R. F., Lawrence, F., Nuzhdin, S. V., and Mackay, T. F. C. (1996). Effects if single P-element insertions on bristle number and viability in *Drosophila melanogaster*. *Genetics*, 143(1):277 – 292.

Lyttle, T. W. (1993). Cheaters sometimes prosper distortion of mendelian segregation by meiotic drive. *Trends in Genetics*, 9:205 – 210.

Malter, J. S. (1989). Identification of an AUUUA-specific messenger RNA binding protein. *Science*, 246(4930):664 – 666.

Marin, I. and Baker, B. S. (1998). The evolutionary dynamics of sex determination. *Science*, 281:1990 1994.

Marin, I., Siegal, M. L., and Baker, B. S. (2000). The evolution of dosage-compensation mechanisms. *Bioessays*, 22:1106–1114.

Marrelli, M. T., Moreira, C. K., Kelly, D., Alphey, L., and Jacobs-Lorena, M. (2006). Mosquito transgenesis: what is the fitness cost? *Trends in Parasitology*, 22(5):197 – 202.

Marsh, J. L. and Wieschaus, E. (1978). Is the sex determination in germ-line and soma controlled by separate genetic mechanisms? *Nature*, 272:249 – 251.

Mastbaum, O. (1954). Observations of two epidemic malaria seasons (1946 and 1953) before and after malaria control in Swaziland. *Transactions in the Royal Society of Tropical Medicine and Hygiene*, 48:325 – 331.

Mather, K. (1953). Genetical control of stability in development. *Heredity*, 7:297 – 336.

Mattox, W. and Baker, B. S. (1991). Autoregulation of the splicing of transcripts from the *transformer-2* gene of *Drosophila*. *Genes and Development*, 5:786 – 796.

Mattox, W., Griffin, E. M., and Baker, B. S. (1996). A negative feedback mechanism revealed by funtional analysis of the alternative isoforms of the *Drosophila* splicing regulator *transformer-2*. *Genetics*, 143:303 – 314.

Mattox, W., Palmer, M. J., and Baker, B. S. (1990). Alternative splicing of the sex determination gene *transformer-2* is sex-specific in the germ line but not in the soma. *Genes and Development*, 4:789 – 805.

Mauro, U., Salvemini, M., Arca, B., Lonbardo, F., Polito, C. L., and Saccone, G. (2006). The *doublesex* gene of *Aedes aegypti*. *Direct NCBI submission. Retrieved: 2006-05-16*.

Mbogo, C., Baya, N. M., Ofulla, A. V., Githure, J. I., and Snow, R. W. (1996). The impact of permethrin-impregnated bednets on malaria vectors of the Kenya coast. *Medical and Veterinary Entomology*, 10(3):251 –259.

McClelland, G. A. H. (1962). Sex-linkage in *Aedes aegypti*. *Transactions in the Royal Society of Tropical Medicine and Hygiene*, 56:4.

McDonald, J. H. and Kreitman, M. (1991). Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature*, 351:652 – 654.

McKeown, M., Belote, J. M., and Baker, B. S. (1987). A molecular analysis of transformer, a gene in *Drosophila melanogaster* that controls female sexual differentiation. *Cell*, 48:489 – 499.

McMeniman, C. J., Lane, R. V., Cass, B. N., Fong, A. W. C., Sidhu, M., Wang, Y., and O'Neill, S. L. (2009). Stable introduction of a life-shortening *Wolbachia* infection into the mosquito *Aedes aegypti*. *Science*, 323(5910):141 – 144.

Meisch, M. V. (1985). *Gambusia affinis affinis*. In Chapman, H. C., editor, *Biological Control of mosquitoes*, pages 3 – 17. American Mosquito Control Association.

Meise, M., Hilfiker-Kleiner, D., Dübendorfer, A., and Brunner, C. (1998). *Sex-lethal*, the master sex-determining gene in *Drosophila*, is not sex-specifically regulated in *Musca domestica*. *Development*, 125:1487 – 1494.

Merrill, C., Bayraktaroglu, L., Kusano, A., and Ganetzky, B. (1999). Truncated RanGAP encoded by the *Segregation Distorter* locus of *Drosophila*. *Science*, 283:1742 – 1745.

Metcalf, R. L. (1973). A century of *DDT*. *Journal of Agriculture and Food Chemistry*, 21:511 – 519.

Min, K. T. and Benzer, S. (1997). *Wolbachia*, normally a symbiont of *Drosophila* can be virulent, causing degeneration and early death. *Proceedings of the National Academy of Science USA*, 94:10792 – 10796.

Moreira, L. A., Ito, J., Ghosh, A. K., Davenport, M., Zieler, H., Abraham, E. A., Crisanti, A., Nolan, T., Catteruccia, F., and Jacobs-Lorena, M. (2002). Bee venom phospholipase inhibits malaria parasite development in transgenic mosquitoes. *The Journal of Biological Chemistry*, 277:40839 – 40843.

Moreira, L. A., Iturbe-Ormaetxe, I., Jeffery, J. A., Lu, G., Pyke, A. T., Hedges, L. M., Rocha, B. C., Hall-Mendelin, S., Day, A., Riegler, M., Hugo, L. E., Johnson, K. N., Kay, B. H., McGraw, E. A., van den Hurk, A. F., Ryan, P. A., and O'Neill, S. L. (2009). A *Wolbachia* symbiont in *Aedes aegypti* limits infection with dengue, Chikungunya, and *Plasmodium*. *Cell*, 139:1268 – 1278.

Moreira, L. A., Wang, J., Collins, F. H., and Jacobs-Lorena, M. (2004). Fitness of anopheline mosquitoes expressing transgenes that inhibit plasmodium development. *Genetics*, 166:1337 – 1341.

Mourya, D. T. and Soman, R. S. (1985). Effect of gregarine parasite, *Ascogregarina culicis* and tetracycline on the susceptibility of *Culex bitaeniorhynchus* to JE virus. *Indian Journal of Medical Research*, 81:247 – 250.

Moyou-Somo, R., Lehman, L. G., Awahmukalah, S., and Ayuk-Enyong, P. (1995). Deltamethrin impregnated bednets for the control of urban malaria in Kumba Town, South-West Province of Cameroon. *Journal of Tropical Medicine and Hygiene*, 98(5):319 – 324.

Msangi, A., Saleh, K. M., Kiwia, N., Malele, I. I., Mussa, W. A., Mramba, F., Juma, K. G., Dyck, V. A., Vreysen, M. J. B., Parker, A. G., Feldmann, U., Zhu, Z. R., and Pan, H. (2000). *Success in Zanzibar: eradication of tsetse*. Penerbit Universiti Sains Malaysia, Penang, Malaysia.

Müller-Holtkamp, F. (1995). Ther sex-lethal gene in *Chrysomya rufifacies* is highly conserved in sequence and exon-intron organization. *Jounal of Molecular Evolution*, 42:467 – 477.

Nam, V. S., Marchand, R., Tien, T. V., and Binh, N. V. (1997). Dengue vector control in Vietnam using Mesocyclops through community participationesocyclops through community participation. *Dengue Bulletin*, 21:96 – 104.

Nene, V., Wortman, J. R., Lawson, D., Haas, B., Kodira, C., Tu, Z., Loftus, B., Xi, Z., Megy, K., Grabherr, M., Ren, Q., Zdobnov, E. M., Lobo, N. F., Campbell, K. S., Brown, S. E., Bonaldo, M. F., Zhu, J., Sinkins, S. P., Hogenkamp, D. G., Amedeo, P., Arensburger, P., Atkinson, P. W., Bidwell, S., Biedler, J., Bierney, E., Bruggner, R. V., Costas, J., Coy, M. R., Crabtree, J., Crawford, M., deBruyn, B., DeCaprio, D., Eiglmeier, K., Eisenstadt, E., El-Dorry, H., Gelbart, W. M., Gomes, S. L., Hammond, M., Hannick, L. I., Hogan, J. R., Holmes, M. H., Jaffe, D., Johnston, J. S., Kennedy, R. C., Koo, H., Kravitz, S., Kriventseva, E. V., Kulp, D., LaButti, K., Lee, E., Li, S., Lovin, D. D., Mao, C., Mauceli, E., Menck, C. F. M., Miller, J. R., Montgomery, P., Mori, A., Nascimento, A. L., Naveira, H. F., Nusbaum, C., O'Leary, S., Orvis, J., Pertea, M., H.Quesneville, Reidenbach, K. R., Rogers, Y., Roth, C. W., Schneider, J. R., Schatz, M., Shumway, M., Stanke, M., Stinson, E. O., Tubio, J. M. C., VanZee, J. P., Verjovski-Almeida, S., Werner, D., White, O., Wyder, S., Zheng, Q., Zhao, Q., Hill, Y., Hill, C. A., Raikhel, A. S., Soares, M. B., Knudson, D. L., Lee, N. H., Galagan, J., Salzberg, S. L., Paulsen, I. T., Dimopoulos, G., Collins, F. H., Birren, B., Fraser-Liggett, C. M., and Severson, D. W. (2007). Genome sequence of *Aedes aegypti*, a major arbovirus vector. *Science*, 316(5832):1718 – 1723.

Newton, M. E., Southern, D. I., and Wood, R. J. (1974). X and Y chromosomes of *Aedes aegypti* (L) distinquished by Giemsa C-banding. *Chromosoma*, 49:41 – 49.

Newton, M. E., Wood, R. J., and Southern, D. I. (1978). Relative DNA content of normal and sex-ratio distorting spermatozoa of the mosquito, *Aedes aegypti* (L.). *Chromosoma*, 67:253 – 261.

Nöthiger, R. and Steinmann-Zwicky, M. (1985). A single principle of sex determination in insects. *Cold Spring Harbour. Symposium of Quantitative Biology*, 50:615 – 621.

Oladunmade, M. A., Feldmann, U., and Takken, W. (1990). Eradication of *Glossia palpalis palpalis* (Robineau-Desvoidy) (Diptera: Glossinidae) from agropastoral land in central Nigeria by means of the sterile insect technique. In *In Eighteenth meeting of the international scientific council for trypanosomiasis research and control (ISCTRC) (Harare, Zimbabwe)*, pages 351 – 363, OAU/BAR, Nairobi.

O'Neil, M. T. and Belote, J. M. (1992). Interspecific comparison of the transformer gene of *Drosophila* reveals an unusually high degree of evolutionary divergence. *Genetics*, 131:113 – 128.

Ooi, E. E., Hart, T. J., Tan, H. C., and Chan, S. H. (2001). Dengue sero-epidemiology in Singapore. *Lancet*, 357:685 – 686.

Orr, H. A. and Kim, Y. (1998). An adaptive hypothesis for the evolution of the Y chromosome. *Genetics*, 150:1693 –1698.

Orth, P., Schnappinger, D., Hillen, W., Saeger, W., and Hinrichs, W. (2000). Structural and basis of gene regulation by the tetracycline inducible tet repressor-operator system. *Nature Structural Biology*, 7:215 – 219.

Osta, M. A., Christophides, G. K., and Kafatos, F. C. (2004). Effects of mosquito genes on plasmodium development. *Science*, 303:2 – 30.

Paeporn, P., Ya-umphan, P., Supaphathom, K., Savanpanyalert, P., Wattanachai, P., and Patimaprakorn, R. (2004). Insecticide susceptibility and selection for resistance in a population of *Aedes aegypti* from Ratchaburi Province, Thailand. *Tropical Biomedicine*, Suplement:1 – 6.

PAHO (1982). *Emergency Vector Control after Natural Disaster - A Study Guide*. Pan American Health Organization.

Pal, R. (1974). WHO/ICMR programme of genetic control of mosquitoes in India. *The Use of Genetics in Insect Control*, pages 75 –95.

Palmer, M. J., Mergener, V. A., Richman, R., Manning, J. E., Kuroda, M. I., and Lucchesi, J. C. (1993). The *male-specific lethal-one (msl-1)* gene of *Drosophila melanogaster* encodes a novel protein that associates with the X chromosome in males. *Genetics*, 134:545 – 557.

Palumbo, G., Bonnacorsi, S., Robbins, L., and Pimpinelli, S. (1994). Genetic analysis of stellate elements of *Drosophila melanogaster*. *Genetics*, 138:1181 – 1197.

Pane, A., Salvemini, M., Bovi, P. D., Polito, C., and Saccone, G. (2002). The *transformer* gene in *Ceratitis capitata* provides a genetic basis for selecting and remembering the sexual fate. *Development*, 129:3715 – 3725.

Pannuti, A., Kocacitak, T., and Lucchesi, J. C. (2000). *Drosophila* as a model for the study of sex determination in Anopheline and Aedine mosquitoes. In Tan, K. H., editor, *Area-wide control of fruit flies and other insect pests*, pages 263 – 269. Penerbit Universiti Sains Malaysia, Penang.

Paton, M. G., Karunaratne, S. H. P., Giakoumaki, E., Roberts, N., and Hemingway, J. (2000). Quantitative analysis of gene amplification in insecticide-resistant *Culex* mosquitoes. *Journal of Biochemistry*, 346:17 – 24.

Patterson, R. S., Weidhaas, D. E., Ford, H. R., and Lofgren, C. S. (1970). Suppression and elimination of an island population of *Culex pipiens quinquefasciatus* with sterile males. *Science*, 168:1368 – 1370.

Payne, D., Grab, B., and Fontaine, R. E. (1976). Impact of control measures on malaria transmission and general mortality. *Bulletin of the World Health Organisation*, 54:369 – 377.

Peacock, W. J., Miklos, G. L. G., and Goodchild, D. J. (1975). Sex chromosome meiotic drive systems in *Drosophila melanogaster*: I. abnormal spermatid development in males with a heterochromatin deficient X chromosome (sc4sc8). *Genetics*, 79:623 – 634.

Phuc, H. K., Andreasen, M. H., Burton, R. S., Vass, C., Epton, M. J., Pape, G., Fu, G., Condon, K. C., Scaife, S., Donnelly, C. A., Coleman, P. G., White-Cooper, H., and Alphey, L. (2007). Late-acting dominant lethal genetic systems and mosquito control. *BioMed Central Biology*, 5(1):11.

Pimpinelli, S. and Dimitri, P. (1989). Cytogenetic analysis of segregation distortion in *Drosophila melanogaster*: the cytological organization of the *Responder (Rsp)* locus. *Genetics*, 121:765 – 772.

Politzar, H. and Cuisance, D. (1984). An integrated campaign against riverine tsetse, *Glossina palpalis gambiensis* and *Glossina tachinoides*, by trapping and the release of sterile males. *Insect Science Application*, 5:439 – 442.

Prestridge, D. (1995). Predicting Pol II promoter sequences using transcription factor binding sites. *Jounal of Molecular Biology*, 249:923 – 932.

Rao, T. R. (1984). *The Anophelines of India*. Delhi Press, India.

Rawlins, S. C. (1998). Spatial distribution of insecticide resistance in Caribbean populations of *Aedes aegypti* and its significance. *Rev Panam Salud Publica/Pan American Journal of Public Health*, 4(4):243 – 251.

Reisin, W. K. (1981). *Anopheles culicifacies* Giles: mating behaviour and competitiveness in nature of chemosterilised males carrying a genetic sexing system. *Annals of the Entomological Society of America*, 74:395 – 401.

Rendon, P., McInnis, D., Lance, D., and Stewart, J. (2004). Medfly (diptera: Tephritidae) genetic sexing: Large-scale field comparison of males-only and bisexual sterile fly releases in Guatemala. *Journal of Economic Entomology*, 97:1547 – 1553.

Ribeiro, J. M. and Kidwell, M. G. (1994). Transposable elements as population drive mechanisms: specification of critical parameter values. *Journal of Medical Entomology*, 31:10 – 16.

Rodriguez, M. H., Gonzalez-Ceron, L., Hernandez, J. E., Nettel, J. A., Villarreal, C., Kain, K. C., and Wirtz, R. A. (2000). Different prevalences of *Plasmodium vivax* phenotypes VK210 and VK247 associated with the distribution of *Anopheles albimanus* and *Anopheles pseudopunctipennis* in Mexico. *American Journal of Tropical Medicine and Hygiene*, 62(1):122 – 127.

Rodriguez, M. M., Bisset, J. A., Armas, Y. D., and Ramos, F. (2005). Pyrethroid insecticide-resistant strain of *Aedes aegypti* from Cuba induced by deltamethrin selection. *Journal of America Mosquito Control Association*, 21(4):437 – 445.

Rosen, L., Rozeboom, L. E., Reeves, W. C., Saugrain, J., and Gubler, D. J. (1976). A field trial of competitive displacement of *Aedes polynesiensis* by *Aedes albopictus* on a pacific atoll. *American Journal of Tropical Medicine and Hygiene*, 25(6):906 – 913.

Rubin, G. M. and Spradling, A. C. (1982). Genetic transformation of *Drosophila* with transposable element vectors. *Science*, 218:348 – 353.

Rueger, B., Thalhammer, J., Obermaier, I., and Gruenewald-Janho, S. (1996). Experimental procedure for the detection of a rare human mRNA with the DIG sytem. *Biochemica*, 3.

Ryner, L. C., Goodwin, S. F., Castrillon, D. H., Anand, A., Villella, A., Baker, B. S., Hall, J. C., Taylor, B. J., and Wasserman, S. A. (1996). Control of male sexual behaviour and sexual orientation in *Drosophila* by the *fruitless* gene. *Cell*, 87(6):1079 – 1089.

Saccone, G., Pane, A., and Polito, L. C. (2002). Sex determination in flies, fruitflies, and butterflies. *Genetica*, 116(1):15 – 23.

Saccone, G., Peluso, I., Artiaco, D., Giordano, E., Bopp, D., and Polito, L. C. (1998). The *Ceratitis capitata* homologue of the *Drosophila* sex-determinting gene *Sex-lethal* is structurally conserved, but not sex-specifically regulated. *Development*, 125:1495 – 1500.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular cloning: A laboratory manual*. Cold Spring Harbour Laboratory Press, Cold Spring Harbour, New York, second edition.

Scali, C., Catteruccia, F., Li, Q., and Crisanti, A. (2005). Identification of sex-specific transcripts of the *Anopheles gambiae doublesex* gene. *Journal of Experimental Biology*, 208:3701 – 3709.

Schüpbach, T. (1985). Normal female germ cell differentiation requires the female X-chromosome to autosome ratio and expression of sex-lethal in *Drosophila melanogaster*. *Genetics*, 109:529 – 548.

Schütt, C. and Nöthiger, R. (2000). Structure, function and evolution of sex-determining systems in dipteran insects. *Development*, 127:667 – 677.

Scott, T. W., Chow, E., Strickman, D., Kittayapong, P., Wirtz, R. A., Lorenx, L. H., and Edman, J. D. (1993). Blood– feeding patterns of *Aedes aegypti* (diptera: Culicidae) collected in a rural Thai village. *Journal of Medical Entomology*, 30:922 – 927.

Seawright, J. A., Kaiser, P. E., Dame, D. A., and Lofgren, C. S. (1978). Genetic method for the prefential elimination of female *Anopheles albimanus*. *Science*, 200:1303 – 1304.

Sehgal, N. K. (1974). Doubts over US in India. *Nature*, 251:177 – 178.

Sexton, J. D., Ruebush, T. K., Brandling-Bennett, A. D., Bremen, J. G., Roberts, J. M., Odera, J. S., and Were, J. B. (1990). Permethrin-impregnated curtains and bed-nets prevent malaria in western Kenya. *American Journal of Tropical Medicine and Hygiene*, 43(1):11 – 18.

Sharma, D. and Abrol, D. P. (2005). Contact toxicity of some insecticides to honeybee *Apis mellifera* (L.) and *Apis cerana* (F.). *Journal of Asia-Pacific Entomology*, 8(1):113 – 115.

Sharp, B., van Wyk, P., Sikasote, J. B., Banda, P., and Kleinschmidt, I. (2002). Malaria control by residual insecticide spraying in Chingola and Chililabombwe, Copperbelt Province, Zambia. *Tropical Medicine & International Health*, 7(9):732 – 736.

Shelly, T. E., Whittier, T. S., and Kaneshiro, K. Y. (1994). Sterile insect release and the natural mating system of the mediterranean fruit fly *Ceratitis capitata* (diptera: Tephritidae). *Annual Entomological Society of America*, 87:470 – 481.

Shepard, J. J., Andreadis, T. G., and Vossbrinck, C. R. (2006). Molecular phylogeny and evolutionary relationships among mosquitoes (Diptera: Culicidae) from the northeastern United States based on small subunit ribosomal DNA (18S rDNA) sequences. *Journal of Medical Entomology*, 43:443 – 454.

Sievert, V., Kuhn, S., and Traut, W. (1997). Expression of the sex determining cascade genes sex-lethal and doublesex in the phorid fly *Megaselia scalaris*. *Genome*, 40:211 – 214.

Sinkins, S. P. (2004). *Wolbachia* and cytoplasmic incompatibility in mosquitoes. *Insect Biochemistry and Molecular Biology*, 34(7):723 – 729.

Sinkins, S. P. and Gould, F. (2006). Gene drive systems for insect disease vectors. *Nature Review Genetics*, 7:427 – 435.

Smith, J. L. and Fonseca, D. M. (2004). Rapid assays for identification of members of the *Culex pipiens* complex, their hybrids, and other sibling species (diptera: Culicidae). *American Journal of Tropical Medicine and Hygiene*, 70:339 – 345.

Snow, R. W., Rowan, K. M., Lindsay, S. W., and Greenwood, B. M. (1988). A trial of bed nets (mosquito nets) as a malaria control strategy in a rural area of The Gambia, West Africa. *Transactions in the Royal Society of Tropical Medicine and Hygiene*, 82(2):212 – 215.

Sosnowski, B. A., Belote, J. M., and McKeown, M. (1989). Sex-specific alternative splicing of RNA from the *transformer* gene results from sequence-dependent splice site blockage. *Cell*, 58:449 – 459.

Stahly, D. P. (1984). Biochemical genetics of the bacterial insect-control agent *bacillus thuringiensis*: basic principles and prospects engineering. *Biotechnology & Genetic Engineering Review*, 84:341 – 363.

Steinmann-Zwicky, M., Amrein, H., and Nöthiger, R. (1990). Genetic control of sex determination in *Drosophila*. *Advance Genetics*, 27:189 – 237.

Stoecklin, G., Hahn, S., and Moroni, C. (1994). Functional hierarchy of AUUUA motifs in mediating rapid interleukin-3 mRNA decay. *Journal of Biological Chemistry*, 269:28591 – 28597.

Strachan, T. and Read, A. P. (1999). *Human Molecular Genetics 2*. Garland Science.

Strathdee, C. A., McLeod, M. R., and Hall, J. R. (1999). Efficient control of tetracycline-responsive gene expression from an autoregulated bi-directional expression vector. *Gene*, 229:21 – 29.

Sturtevant, A. H. (1945). A gene in *Drosophila melanogaster* that tranforms females into males. *Genetics*, 30(297 - 299).

Su, T. and Mulla, M. S. (2004). Documentation of high-level *Bacillus sphaericus* 2362 resistance in field populations of *Culex quinquefasciatus* breeding in polluted water in Thailand. *Journal of America Mosquito Control Association*, 20(4):405 – 411.

Suguna, S. G., Wood, R. J., Curtis, C. F., Whitelaw, A., and Kazmi, S. J. (1977). Resistance to meiotic drive at the MD locus in an Indian wild population of *Aedes aegypti*. *Genetic Reseach (Cambridge)*, 29:123 – 132.

Sweeney, T. L. and Barr, A. R. (1978). Sex-ratio distortion caused by meiotic drive in a mosquito, *Culex pipiens* L.. *Genetics*, 88:427 – 446.

Taylor, B. J. and Knittel, L. M. (1995). Sex-specific differentiation of a male-specific abdominal muscle, the Muscle of Lawrence, is abnormal in hydroxyurea-treated and in *fruitless* male flies. *Development*, 121:3079 – 3088.

Thomas, D., Donnelly, C., Wood, R., and Alphey, L. (2000). Insect population control using a dominant, repressible, lethal genetic system. *Science*, 287:2474 – 2476.

Thomson, M. C., Adiamah, J. H., Connor, S. J., Jawara, M., Bennett, S., D'Alessandro, U., Quinones, M., Langerock, P., and Greenwood, B. M. (1995). Entomological evaluation of the Gambia's National Impregnated Bednet Programme. *Annals of Tropical Medicine and Parasitology*, 89(3):229 – 241.

Timmermann, S. E. and Briegel, H. (1999). Larval growth and biosynthesis of reserves in mosquitoes. *Journal of Insect Physiology*, 5:461 – 470.

Tomkiel, J. E. (2000). Chromosomally-induced meiotic drive in *Drosophila* males: checkpoint or fallout? *Genetica*, 109:95 – 103.

Tsuda, Y., Kobayashi, J., Nambanya, S., Miyagi, I., Toma, T., Phompida, S., and Manivang, K. (2001). An ecological survey of dengue vector mosquitoes in central Lao PDR. *Southeast Asian Journal of Tropical Medicine and Public Health*, 33:63 – 67.

Turelli, M. (1994). Evolution of incompatibility-inducing microbes and their hosts. *Evolution*, 48:1500 – 1513.

Ullerich, F. H. (1984). Analysis of sex determination in the monogenic blowfly *Chrysoma rufifacies* by pole cell transplantation. *Molecular and General Genetics*, 193:479 – 487.

Vassin, H., Brand, M., Jan, L. Y., and Jan, Y. N. (1994). *Daughterless* is essential for neuronal precursor differentiation but not for initiation of neuronal precursor formation in *Drosophila* embryo. *Development*, 120(4):935 – 945.

Vibranovski, M. D., Koerich, L. B., and Carvalho, A. B. (2008). Two new Y-linked genes in *Drosophila melanogaster*. *Genetics*, 179:2325 – 2327.

Wade, M. J. and Beeman, R. W. (1994). The population dynamics of maternal-effect selfish genes. *Genetics*, 138:1309 – 1314.

Walthour, C. S. and Schaeffer, S. W. (1994). Molecular population genetics of sex determination genes: The *transformer* gene of *Drosophila melanogaster*. *Genetics*, 136:1367 – 1372.

Wang, N. C. (1994). Control of dengue vectors in Singapore. *Gaoxiong Yi Xue Ke Xue Za Zhi*, 10:33 – 38.

Watanabe, T. K. (1975). A new sex-transforming gene on the second chromosome of *Drosophila melanogaster*. *Japanese Journal of Genetics*, 50:269 – 271.

Waterbury, J. A., Horabi, J. I., Bopp, D., and Schedl, P. (2000). Sex determination in the *Drosophila* germline is dictated by the sexual identity of the surrounding soma. *Genetics*, 155(4):1741 – 1756.

Waterbury, J. A., Jackson, L. L., and Schedl, P. (1999). Analysis of the *doublesex* female protein in *Drosophila melanogaster*: role on sexual differentiation and behavior and dependence on intersex. *Genetics*, 152(4):1653 – 1667.

Weidhass, D. E., Breeland, S. G., Lofgren, C. S., Dame, D. A., and Kaiser, R. (1974). Release of chemosterilized males for the control of *Anopheles albimanus* in El Salvador. IV. Dynamics of the test population. *American Journal of Tropical Medicine and Hygiene*, 23:298 – 308.

White, M. J. D. (1973). *Animal cytology and evolution*. Cambridge University Press, Cambridge, 3 edition.

WHO (Retrieved: April 2009b). Fact sheet no. 117: Dengue and dengue hemorrhagic fever. *http://www.who.int/mediacentre/factsheets/fs117/en/*.

WHO (Retrieved: July 2009a). Fact sheet no. 100: Yellow fever. *http://www.who.int/mediacentre/factsheets/fs100/en/*.

WHO (Retrieved: October 2008a). Chikungunya fever. Chikungunya in South-East Asia. *http://www.searo.who.int/en/Section10/Section2246_13975.htm*.

WHO (Retrieved: October 2008b). Fact sheet no. 327: Chikungunya. *http://www.who.int/mediacentre/factsheets/fs327/en/index.html*.

Wilkins, A. S. (1995). Moving up the hierarchy: A hypothesis on the evolution of a genetic sex determination pathway. *BioEssays*, 17:71 – 77.

Windbichler, N., Papathanos, P. A., Catteruccia, F., Ranson, H., Burt, A., and Crisanti, A. (2007). Homing endonuclease mediated gene targeting in *Anopheles gambiae* cells and embryos. *Nucleic Acids Research*, 35(17):5922 – 5933.

Wood, R. J. and Newton, M. E. (1977). Meiotic drive and sex ratio distortion in the mosquito *Ae. aegypti*. In *Proceedings of the 15th International Congress of Entomology*, pages 97 – 105, Washington, DC.

Wood, R. J. and Newton, M. E. (1991). Sex ratio distortion caused by meiotic drive in mosquitoes. *American Nature*, 137:379 – 191.

Wood, R. J. and Ouda, N. A. (1987). The genetic basis of resistance and sensitivity to the meiotic drive gene D in the mosquito *Aedes aegypti* L. *Genetica*, 72:69 – 79.

Wu, C. I., Lyttle, T. W., Wu, M. L., and Lin, G. F. (1988). Association between a satellite DNA sequence and the Responder of Segregation Distorter in *D. melanogaster*. *Cell*, 54:179 – 189.

Yang, D., Lu, H., Hong, Y., Jinks, T. M., Estes, P. A., and Erickson, J. W. (2001). Interpretation of X chromosome dose at *sex-lethal* requires Non-E-Box sites for the basic helix-loop-helix proteins SISB and *daughterless*. *Molecular and Cellular Biology*, 21(5):1581 – 1592.

Yi, S. and Charlesworth, B. (2000). Contrasting patterns of molecular evolution of the genes on the new and old sex chromosomes of *Drosophila Miranda*. *Molecular Biology and Evolution*, 17(5):703 –717.

Yi, W. and Zarkower, D. (1999). Similarity of DNA binding and transcriptional regulation by *C. elegans* MAB-3 and *Drosophila melanogaster* DSX suggests conservation of sex determining mechanisms. *Development*, 126:873 – 881.

Zhu, C., Urano, J., and Bell, L. R. (1997). The *Sex-lethal* early splicing pattern uses a default mechanism dependent on the alternative 5' splice sites. *Molecular and Cellular Biology*, 17(3):1674 – 1681.

Zieler, H., Keister, D. B., Dvorak, J. A., and Ribeiro, J. M. C. (2001). A snake venom phospholipase A2 blocks malaria parasite development in the mosquito midgut by inhibiting ookinete association with the midgut surface. *The Journal of Experimental Biology*, 204:4157 – 4167.

Appendix

A-1 List of buffers

A-1.1 Livak buffer

- 80 mM NaCl
- 0.16 M sucrose
- 130 mM Tris Base
- 50.8 mM EDTA
- 5mM SDS

A-1.2 TAE buffer

- 40 mM Tris-Acetate
- 1 mM EDTA

A-1.3 RNA loading buffer

- 250 μl formamide, deionised
- 83 μl formaldehyde, 37% (w/v)
- 50 μl 10X MOPS buffer
- 0.01% (w/v) bromophenol blue

A-1.4 10X MOPS buffer

- 200 mM morpholinopropan-sulfonic acid (41.1 g to make 1l)
- 50 mM sodium acetate (4.12 g to make 1l)
- 10 mM EDTA (3.72 g for 1l) pH 7.0
- 1l DEPC/DMDC treated water

A-1.5 20X SSC

- 3M NaCl (175.2 g for 1l)
- 300 mM sodium citrate (88.2 g for 1l) pH 7.0

A-1.6 Running buffer For RNA gel (1l)

- 100 ml 10X MOPS
- 52. 6 ml of formaldehyde (0.2M final concentration)
- 847. 4 ml DEPC/DMDC treated water

A-1.7 Low stringency wash buffer 1

- 2X SSC (10X dilution) 20 ml for 200 ml or 50 ml 20X SSC for 500 ml
- 0.1 % sodium dodecyl sulfate (SDS) 0.5 g for 500 ml solution, or 2 ml 10% SDS for 200ml solution
- 178 ml DEPC water for 200 ml solution

A-1.8 High stringency wash buffer 2

- 0.1X SSC (200X dilution) 1ml for 200 ml or 2.5 ml 20X SSC for 500 ml
- 0.1% SDS (0.5 g for 500 ml solution) 2 ml 10% SDS for 200 ml solution
- 193 ml DEPC water for 200 ml solution

A-1.9 0.37 M Formaldehyde agarose gel (1.2% agarose)

- 1.08 g agarose
- 78.3 ml DEPC/DMDC water
- 9 ml 10X MOPS
- 2.7 ml 37% formaldehyde

A-2 Genetic cross for meiotic drive sensitivity

A preliminary study was conducted into meiotic drive sensitivity in *Ae. aegypti*, to examine the effectiveness of using meiotic drive for distorting sex ratio. Several strains of *Ae. aegypti* were tested for drive susceptibility with a strain carrying driver gene M^D and insensitive responder (m^i) strain named T37 (selected to express high rates of sex ratio distortion in F1 males by Mori et al., 2004). The Rockefeller strain was chosen to investigate whether it was drive sensitive as it was important in many field trials of transgenic *Ae. aegypti* lines. The Hanoi strain (collected in Hanoi, Vietnam in 2006 by the Vietnam National University, Hanoi) and Bora-Bora strain (Collected in 1990 from Bora-bora, French Polynesia and maintained in the Medical Entomology Laboratory, Institut Louis Malardè, Tahiti) were used to investigate whether drive sensitivity was present in the population targeted for vector control. *Ae. aegypti* RED strain, a Red-eyed strain that is highly sensitive to meiotic driver bred by Wood and Ouda 1987, was used as a positive meiotic drive control.

All cage populations were maintained as described in Chapter 2. Fifteen adults (3-5 days old) of each T37 males and tester females were crossed. The females were then blood-fed for 4 hours per day for two consecutive days. Two days after blood-feeding, females were allowed to oviposit in an egg bowl. The eggs collected were dried for three days in insectary conditions to mature them. One thousand to one thousand five hundred eggs were then hatched in deoxygenated water. The hatch rate (number of larvae/eggs) and sex ratio of F1 pupae developed were counted. The experiment was repeated on the second generation (F2), which was progeny from mass mating of F1. Chi-square tests were performed to determine if any deviation from the 1:1 ratio was statistically significant. It is known that even in normal populations, sex ratio tend to be slightly biased towards males (Hickey and Craig, 1966). Therefore the accepted threshold for significance is only accepted when the p-value of the Chi-square test is less than 0.05.

In the cross between T37 and test strains, all F1 progeny showed a relatively equal sex ratio as expected (table A-1). Chi square test of each cross was also calculated and it was determined that variations in ratio were not significant and therefore F1 crosses have a normal sex ratio. Hatch rate of each cross was over 90%, except for the T37-Red cross strain, which had 43% to 67% hatch rate. Mass cross of T37 to various populations when it was observed at F2 (see table A-2) showed that Rockefeller and Hanoi strains are not significantly sensitive to drive, with an average of 58% and 50% males in the population respectively. However the hatch rate of T37 to Rockefeller crosses showed disparity between the two replicates, where the first hatch rate was 98% and the second replicate was 25% lower. The percentage of males in each cross yet showed a 3% difference, which was within the normal sex ratio. The Bora-Bora strain had an average of 64% male, which indicates a skew in the sex ratio. The hatch rate of the strain were very high (96% on average). Red strain

showed a lower average hatch rate of 74%. Red strain had the highest amount of skew, with an average of 84% of the population being male.

When Chi square tests were performed on the crosses, only the result from RED and Bora-Bora strain crosses were statistically significant, with p values of less than 0.005 (table A-2). This means that the null hypothesis (where null hypothesis is that percentage male is 50%) for the RED and Bora-Bora strain was rejected. It was concluded that the Bora-Bora and Hanoi strains were not sufficiently insensitive to MD meiotic drive to represent suitable targets for a population control strategy using MD. The meiotic drive component of the project was therefore discontinued, in order to concentrate on sex ratio distortion by manipulation of sex determination.

Table A-1: Result of F1 from mass cross tester strains with strain T37. Each cross was made in duplicate. Numbers in brackets denote actual numbers of individuals. Hatch rate was calculated by number of 1^{st} instar larvae out of the number of embryos hatched. χ^2 value was calculated for the null hypothesis, where the null hypothesis is that percentage male is 50%.

Tester Strain	Total embryos	Hatch rate (larvae/embryos)	Pupation rate (pupae/larvae)	Percentage Male (male/pupae)	χ^2 value (p-value)
Rockefeller -1	1300	95% (1229/1300)	98% (1202/1229)	52% (622/1202)	-
Rockefeller -2	1100	94% (1035/1100)	99% (1028/1035)	60% (616/1028)	-
Mean	1200	94%	99%	56%	2.503 (p= 0.1136)
Bora-Bora - 1	1002	97% (973/1002)	91% (883/973)	53% (471/883)	-
Bora-Bora-2	1213	97% (1175/1213)	86% (1013/1175)	53% (538/1013)	-
Mean	1108	97%	88%	53%	0.416 (p=0.5189)
Hanoi - 1	1119	91% (1013/1119)	99% (1001/1013)	52% (518/1001)	-
Hanoi - 2	1357	98% (1328/1357)	99% (1316/1328)	54% (712/1316)	-
Mean	1238	94%	99%	53%	0.400 (p=0.5271)
RED-1	1067	67% (717/1067)	96% (685/717)	59% (404/685)	-
RED-2	1139	43% (495/1139)	85% (419/1139)	53%(221/419)	-
Mean	1103	55%	90%	56%	1.763 (p=0.1843)

Table A-2: Result of F2 from mass cross of tester strains with strain T37. Numbers in brackets denote actual numbers of individuals. Hatch rate was calculated by number of 1^{st} instar larvae out of the number of embryos hatched. χ^2 value was calculated for the null hypothesis, where the null hypothesis is that percentage male is 50%.

Tester Strain	Total embryos	Hatch rate (larvae/embryos)	Pupation rate (pupae/larvae)	Percentage Male (male/pupae)	χ^2 value (p-value)
Rockefeller - 1	1350	98% (1319/1350)	98% (1306/1319)	60% (782/1306)	-
Rockefeller - 2	1382	73% (1015/1382)	98% (995/1015)	57% (563/995)	-
Mean	1366	86%	98%	58%	2.81 (p=0.0937)
Bora-Bora - 1	2589	96% (2473/2589)	98% (2428/2473)	65% (1574/2428)	-
Bora-Bora -2	1070	96% (1031/1070)	99% (101/1031)	62% (636/1019)	-
Mean	1985	96%	99%	64%	7.48 (p=0.0062)
Hanoi - 1	1104	99% (1093/1104)	97% (1061/1093)	52% (556/1061)	-
Hanoi- 2	1036	99% (1028/1036)	99% (1022/1028)	48% (490/1022)	-
Mean	1120	99%	98%	50%	0.199 (p=0.6555)
RED- 1	1059	71% (749/1059)	99% (738/749))	82% (604/738)	-
RED-2	1136	78% (882/1136)	93% (816/882)	86% (694/816)	-
Mean	1098	74%	96%	84%	46.123 (p=<0.0001)

A-3 Microarray analysis by Schmid and Hubert (for Chapter 3)

Quality Control reports generated by the Agilent Feature Extraction software confirms the observation when the spatial distribution of all outliers on the array. The green channel of arrays 2, 3, and 4 showed no clear signal. Arrays 1 and 2 were displaying a high proportion of red dye (chromosome 1) (see figure A-1). Moreover, the local background score of the red channel was over 50, indicating poor performance as it is too high. In array 3, there was some background clustering in the upper right corner of the red channel (chromosome 2), which may be a sign of contamination. Local background score was less than 50, indicating good hybridisation. Spatial distribution of outliers were also quite evenly distributed on the array. Array 4 also had a local background score less then 50. Spatial distribution of outliers were even, except for the right lower corner of the array, which had two small clusters of both red and green dye.

The scatterplots of red and green background corrected signals generated from the QC report showed distribution was spread out in arrays 1, 2 and 3 (see figure A-2). This suggests that there were little significant signals in the array. It is therefore concluded that no significant difference was observed between the hybridisation to these chromosomes. The array 4 scatterplot however showed a linear distribution as expected.

Scatterplots of log intensities of the red and green channel expression values were generated by Schmid and Hubert (EBI), and it indicated that arrays were odd with respect to the expected intensities (figure A-3). Thresholds were calculated using the upperBoundNull() function supplied by Ringo. This calculates a cutoff than can be used for estimating probed intensities that are more indicative for a binding of fragments of one of the chromosomes. The intensities are assumed to arise from a mixture of two distributions, a symmetric null distribution and an an alternative distribution that is stochastically larger than the null. upperBoundNull() tries to pinpoint the minimum of data values that are more likely to arise from the alternative distribution, for the values following the null distribution. Based on the density distribution of the intensities the mode of the null distribution were calculated for each channel on each array and the thresholds were calculated and are shown in figure A-3. Based on these thresholds the number of probes which can be assigned to bind to one of the chromosomes were calculated. The results are shown in the table A-3.

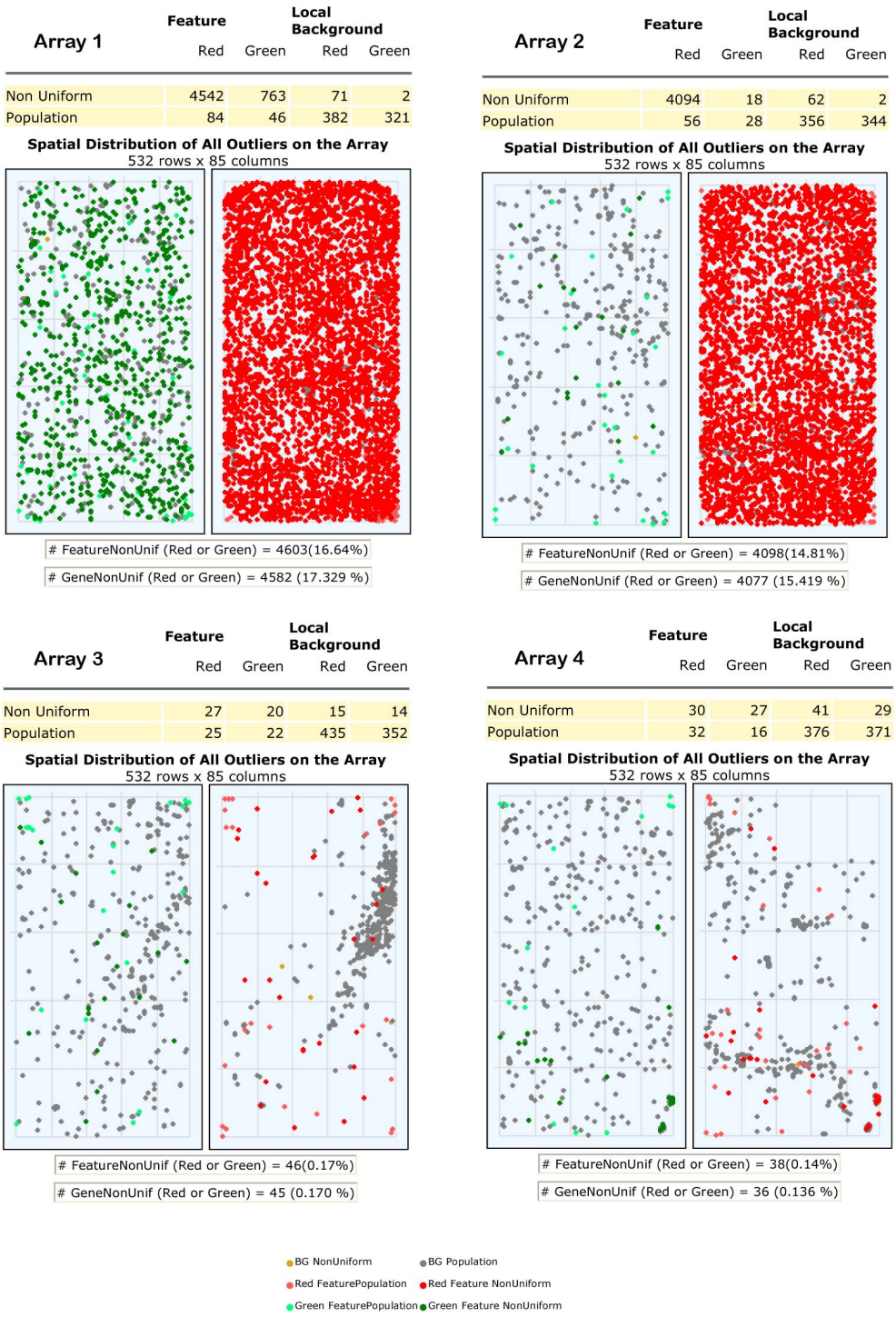

Figure A-1: Spatial Distribution of outliers in the *Aedes aegypti* CGH microarray CGH Quality Control (QC) report generated from Feature Extraction software. Local Background non uniform score of red dye in Arrays 1 and 2 (Chromosome 1 vs 2 and 3 respectively) were more than 50, indicating high background noise.

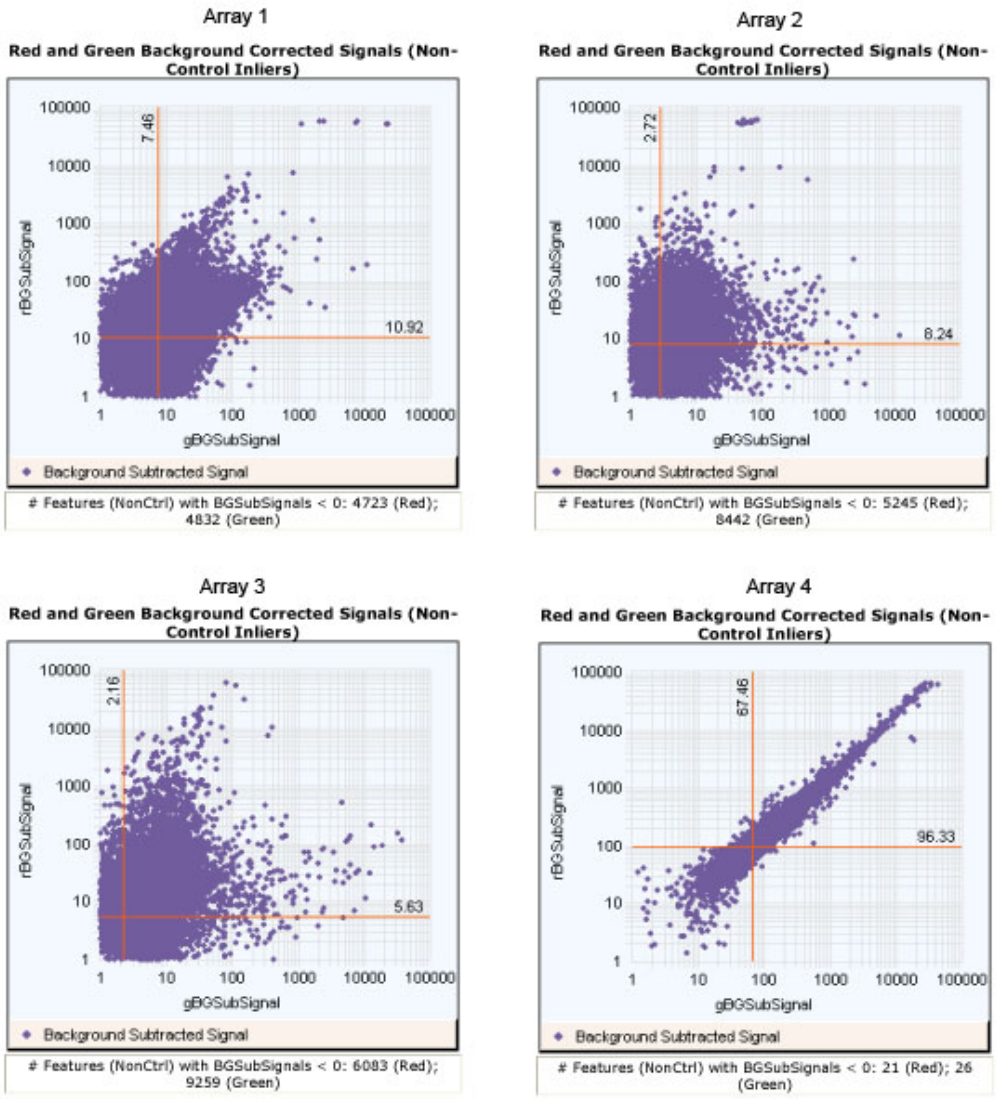

Figure A-2: Scatterplot of red and green background corrected signals (non-control inliers) in the *Aedes aegypti* CGH microarray CGH Quality Control (QC) report generated from Feature Extraction software. Array 4 scatterplot showed a linear distribution, while the rest of the arrays showed a spread-out distribution.

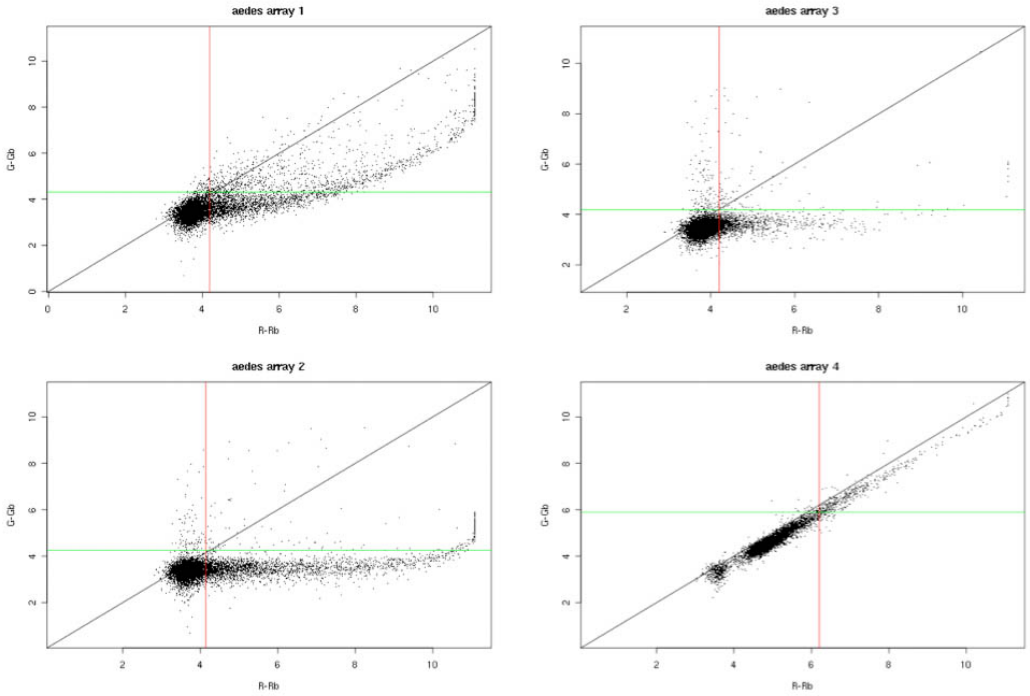

Figure A-3: Analysis of microarray by Ramona Schmid and Wolfgang Hubert from the European Bioinformatics Institute in Sanger Institute. The axes show the red channel, with red intensity-red background intensity, the y-axes the respective results of the green channel. The red line indicates the threshold calculated for the red channel intensities and the green line the threshold for the green channel intensities, respectively. Spots with intensities higher than the calculated threshold are assumed to be significant, that is, they would be mapped to the respective chromosome. Using threshold for the red and green intensities, the plots can be divided into four areas, *upper left* (probes assigned to chromosome labelled green), *lower left* (probes assigned to no chromosome), *lower right* (probes assigned to chromosome labelled red), *upper right* (unsure probes which cannot be clearly assigned to one of the chromosomes).

Table A-3: Number of spots lying in the four quadrants based on the calculated threshold, by Schmid and Hubert (EBI). For explanation of the different quadrants, see caption of figure A-3.

	Array 1: Chr 1 vs. Chr 2	Array 2: Chr 1 vs. Chr 3	Array 3: Chr 2 vs. Chr 3	Array 4: Male vs. Female
significant green (upper left)	1004	549	754	342
significant red (lower right)	5440	8539	3332	220
not significant (lower left	15318	17195	2995	24835
unsure (upper right)	5896	1385	587	2271

A-4 Sequences for Chapter 4

A-4.1 *An. gambiae* protein ENSANGP00000010299

ENSANGP00000010299 protein sequence matching to *D. melanogaster* TRA-2 are as follows (sequence in bold denotes the area matching with tBLASTn search to *Ae. aegypti*):

FSTCCELSCSKRILRKMTSSSYRSSDRRTSRRYDEDGDRHHRHRSRSVD
RCHRSRRSRSDSRDRYRRNRSASRSPRGRSRHRKSSTAPNAS**RSRIRVD-
SPEPSRCLGVFGLSVYTTEPYLNDIFCHFGTVEKSVVIYDAKTRL-
SRGFG**

FVYFKSQAEASIARANCNGLQIHGRRIRVDYSITDQPHPPTPGV
YMGRRQYSHSPSPRNRSHHE

A-4.2 tBLASTn result from *tra-2 D. melanogaster*

Ae. aegypti superconting 1.113 of *Ae aegypti* [DNA] 1220850-1221212

CGTTCCAACCGGCCGGTGGATCCCCCGAAGAGTAAGTGCCTCGGT
GTGTTCGGCCTAAGCAGCTACACCAACGAAACCAGCCTGATGGAC
GTTTTCGCACCGTACGGAACCATTGACAAGGCGATGATTGTCTAC
GATGCCAAGACGAAGGTTTCCCGAGGGTTCGGATTCGTGTACTTC
CAGGAGCAGAGTGCGGCCACCGAAGCCAAAATGCAGTGTAATGGA
ATGATGCTGCATGAGCGCACGATTAGAGTGGATTATTCGGTGACC
GAAAGACCGCATACGCCCACGCCCGGTGTCTACATGGGAGCTAGA
AGCACTGAGAAACGGAAGCACCGCAGTTCCTATAGCTACAGGGGA
CGG

A-4.3 BLAST result from sequences obtained from RACE 3'

The following are the sequence of the longer product matching with two regions of supercontig 1.113 (using nucleotide BLAST, BLASTn):

Position: 1194296 to 1194767 in *Ae. aegypti* supercontig 1.113 (minus strand)

CATAAACCACACGATCTTTAACTAATAAATAAATAAAATGTTTAG
TTTGAATCGCGGAATAATGAAAAAAATGTGAAGAAATCCAAACAT
CCTGAAACGAAAAGAAAAACAGAACGGTTAGCTCATACAATTTAC
CCTTGGATTAATGAAGAGTAATCTGTTCTTCCCTTGTGTCATTTG
AAGAATAGTTTTTACTGATACAGATTCTATCTTCAATAATACTTG
ACTTGAACTTCTTTGTTTGTAGGGTCTTCGTGAGGTCTTGAAAAA
GATGACTTACGTCAAAAACACAATCCTTCTAGCACATTATCGATA
GGATCGACGTGAGTCAACTGAGGAATAAGATGGAGAACGATCACG

AGCACGGTCACGTCGATGGTGATGGCGATGACTAGATCTACGGTG
ATGATGACAGGAACGAGATCTGCTGCGTCTTGACCGCCGATGATG
GTAATCGTCA

Position: 1220842 to 1221076 in *Ae. aegypti* supercontig 1.113 (minus strand)

TCATAGCTCCGTCCCCTGTAGCTATAGGAACTGCGGTGCTTCCGT
TTCTCAGTGCTTCTAGCTCCCATGTAGACACCGGGCGTGGGCGTA
TGCGGTCTTTCGGTCACCGAATAATCCACTCTAATCGTGCGCTCA
TGCAGCATCATTCCATTACACTGCATTTTGGCTTCGGTGGCCGCA
CTCTGCTCCTGGAAGTACACGAATCCGAACCCTCGGGAAACCTTC
GTCTTGGCAT

A-4.4 Subtractive hybridisation primers

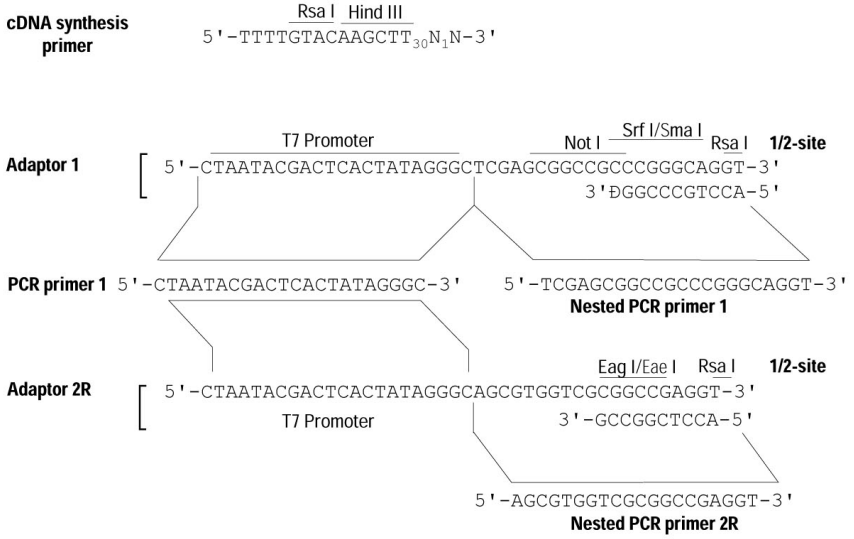

Figure A-4: Sequences of the PCR-Select cDNA synthesis primer, adaptors, and PCR primers. When Adaptors 1 and 2R are ligated to *Rsa I*-digested cDNA/DNA, the *Rsa I* site is restored. Adaptor 1 possess restriction sites *Not I, Srf I/Sma I and Rsa I.* Adaptor 2R possess restriction sites *Eag I/ Eae I.*

A-5 RACE alignments for *tra-2*

The multiple nucleotide sequence alignment of the RACE experiments were aligned using Clustal W in the Geneious Software.

Figure A-5: Multiple nucleotide sequence alignment from 3' RACE result. Grey boxes denotes exon positions. Sequence 1 and 22 are sequence regions from contig 1.113. Sequence 2 and 3 are sequence from larvae samples. Sequence 4 to 12 are sequences from male pupae samples. Sequence 13 to 21 are sequences from female pupae samples. Alignment shows that the RACE experiments detects many variants due to alternative splicing.

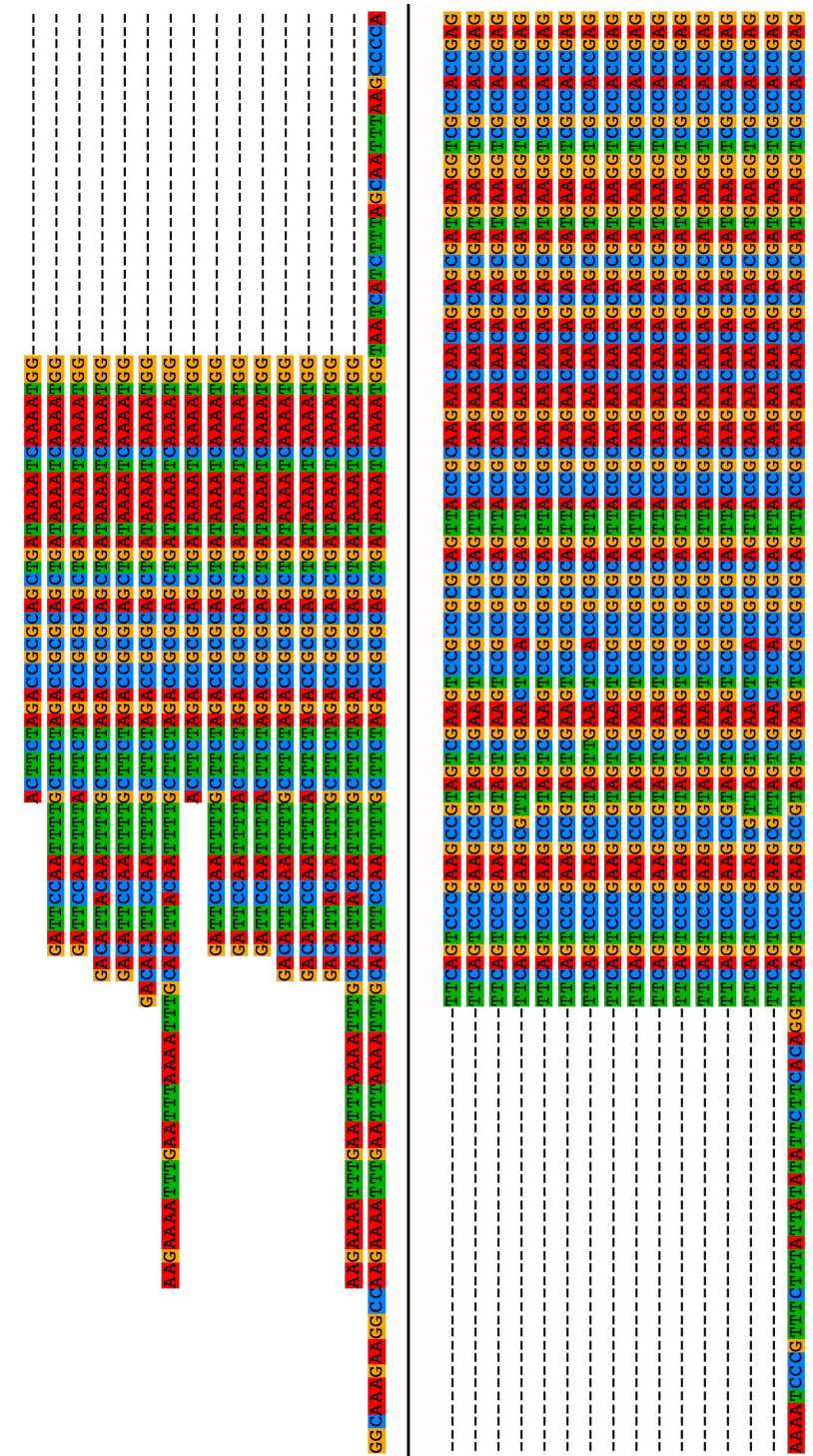

Figure A-6: Partial multiple nucleotide sequence alignment from female (Fem) and male (Mal) SMART 5' RACE. 5' end sequence shown at the start of alignment. Sequence number 16 denotes part region of Supercontig 1.113. Sequences 1 to 15 aligns with 2 areas of Supercontig 1.113, hence a gap in the alignment is seen, indicating an intron (62 nucleotides long). Purple bar denotes gap in alignment.

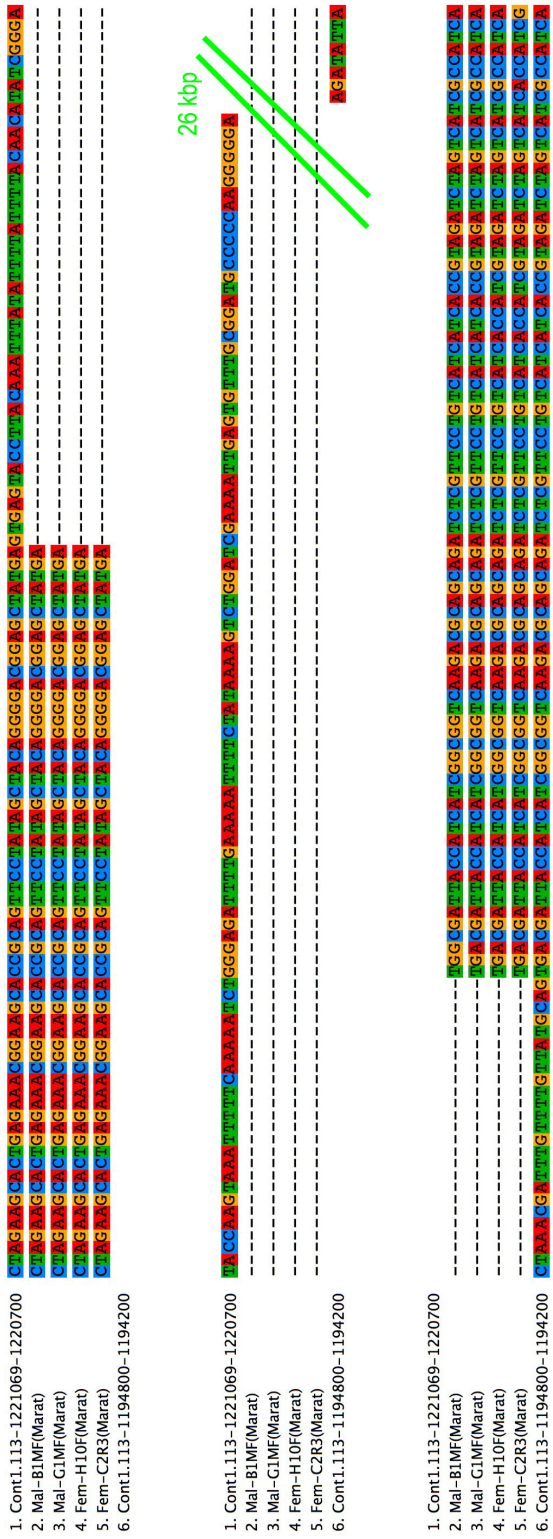

Figure A-7: Partial multiple alignment sequence from Marathon 3' RACE showing a gap of more than 26 kbp (denoted by the double-slanted green bars) in comparison to the genome database. The sequence spans from part of exon 1 to exon 2. The sequences 2 and 3 (Mal-B1MF and Mal-G1F) were obtained from male mRNA and sequences 4 and 5 (Fem-C2R3 and Fem-H10F) were obtained from female mRNA. Supercontig 1.113 regions (sequence numbers 1 and 6) were used to compare the result of the RACE sequences. The current annotation of this gene does not include the second region of this gene possibly due to the large intron in between the two exons.

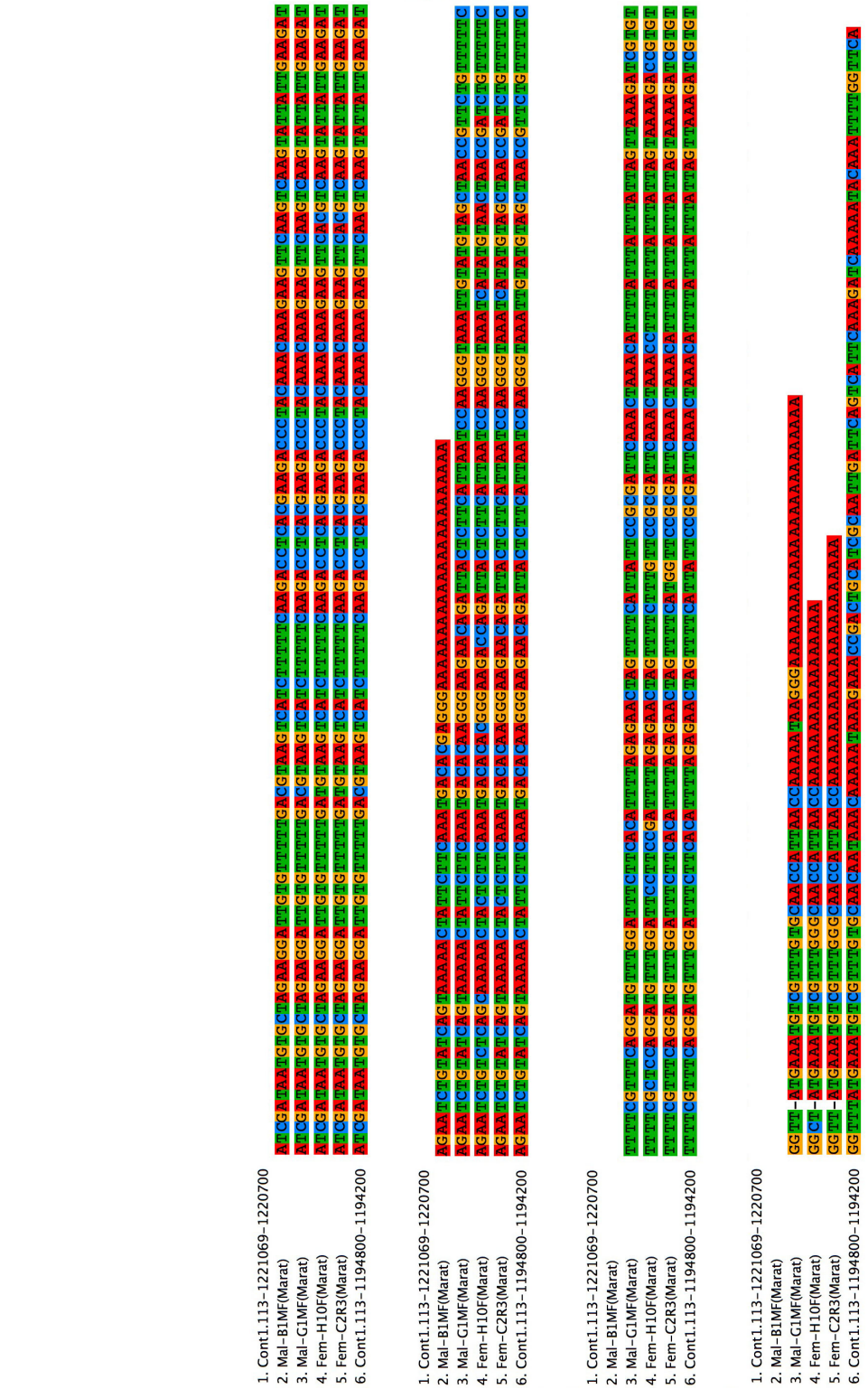

Figure A-8: Partial multiple alignment sequence from Marathon RACE showing 3' end variations in comparison to the genome database. The sequences 2 and 3 (Mal-B1MF and Mal-G1F) were obtained from male mRNA and the bottom sequences 4 and 5 (Fem-C2R3 and Fem-H10F) were obtained from female mRNA. Supercontig 1.113 regions (sequence numbers 1 and 6) were used to compare the result of the RACE sequences. Purple bar denotes gap in alignment.

Figure A-9: Partial multiple alignment sequence of *tra-2* sequences from female SMART 3' RACE to Supercontig 1.113 positions in the *Ae. aegypti* genome database. The sequence spans from part of exon 1 to exon 2. The top and bottom sequences denotes the Supercontig 1.113 positions. Note the gap in sequence denoted by the double-slanted green bars (26 kbp) is not shown in the alignment.

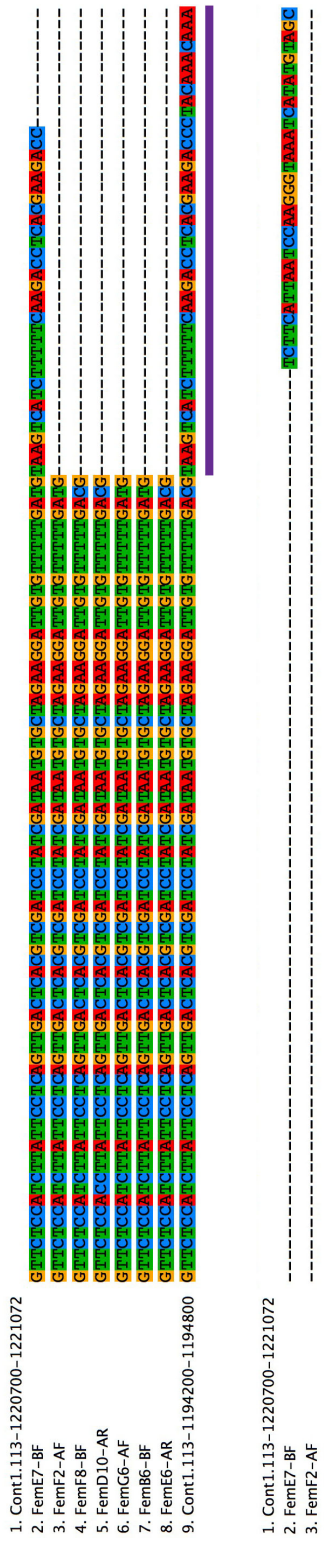

Figure A-10: Partial multiple alignment sequence of female SMART 3' RACE *tra-2* ends to Supercontig 1.113 positions in the *Ae. aegypti* genome database (i.e. second splice site). The sequences spans from part of exon 2 to end of the mRNA. The bottom sequence denotes the Supercontig 1.113 positions. Purple bar denotes gap in alignment. Note that sequence FemE7-BF has a smaller gap in sequence than the rest of the sequences.

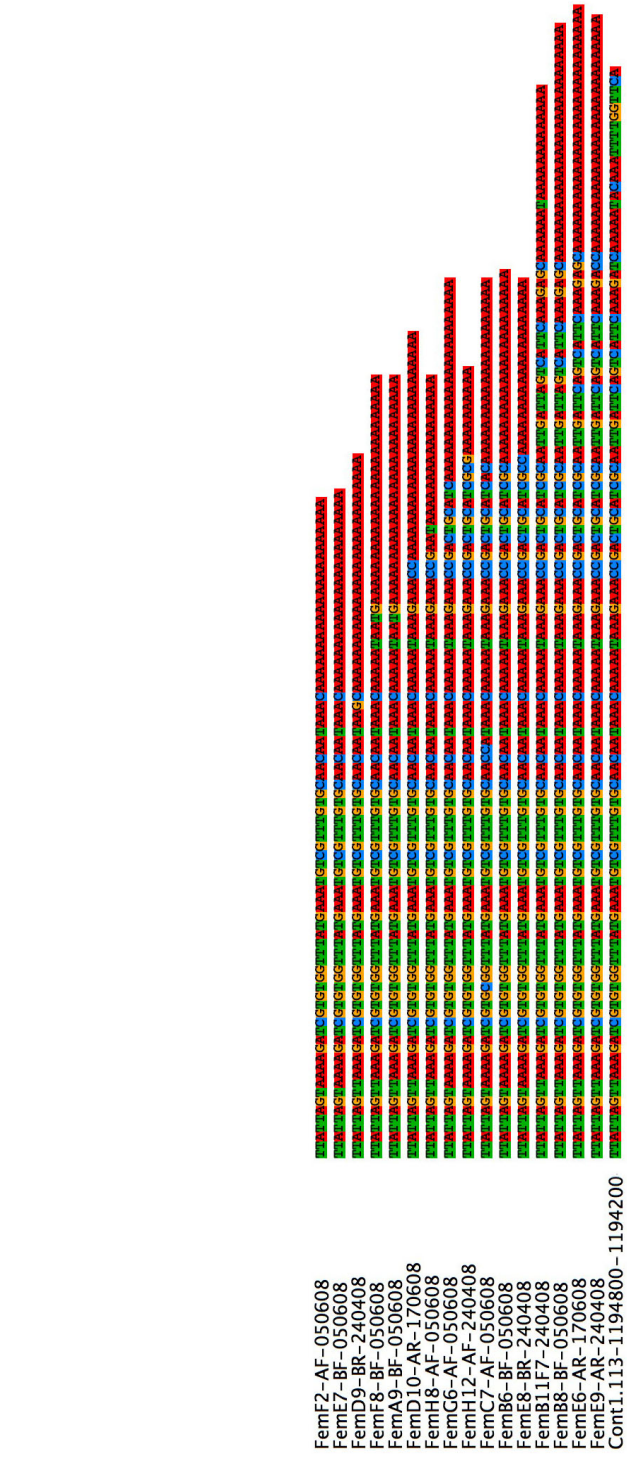

Figure A-11: Partial multiple alignment sequence of female SMART 3' RACE *tra-2* 3' end sequences to the Supercontig 1.113 positions in the *Ae. aegypti* genome database. The bottom sequence denotes the Supercontig 1.113 positions. Note the gap in sequences end differ in length.

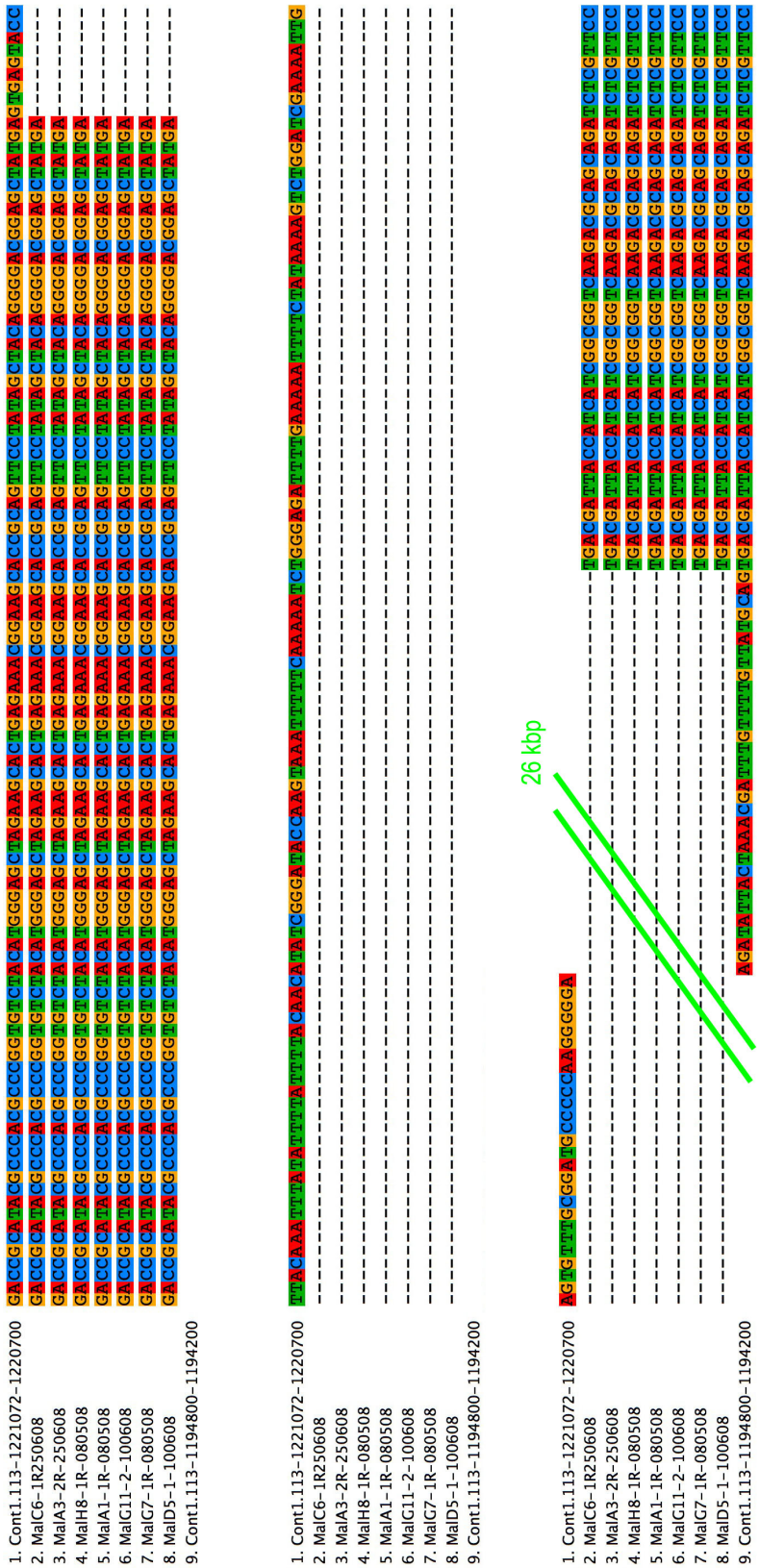

Figure A-12: Partial multiple alignment sequence of male SMART 3' RACE *tra-2* sequences to the Supercontig 1.113 positions in the *Ae. aegypti* genome database. The sequence spans from part of exon 1 to exon 2. The bottom sequence denotes the Supercontig 1.113 positions.

Figure A-13: Partial multiple alignment sequence of male SMART 3' RACE *tra-2* 3' end sequences to the Supercontig 1.113 positions in the *Ae. aegypti* genome database. The sequence spans from part of exon 2 to the end of the mRNA. The bottom sequence denotes the Supercontig 1.113 positions. The sequence ends differ in length. MalC6-1R seemed to be the most truncated, but the poly A tail is short and therefore there is some ambiguity whether the sequence is truly truncated. Purple bar denotes gap in alignment.

Figure A-14: Partial multiple alignment sequence of male and female SMART 3' RACE *tra-2* 3' sequences to the Supercontig 1.113 positions in the *Ae. aegypti* genome database to compare the differences in the two sexes at the second splice site. The bottom sequence denotes the Supercontig 1.113 region. Some male and female sequences have some similar exon-intron structure, except for FemE7-BF which has a smaller gap than the rest of the clones by having 33 bp more at the end of the first exon and 54 bp more a the beginning of the second exon shown here. Purple bar denotes gap in alignment.

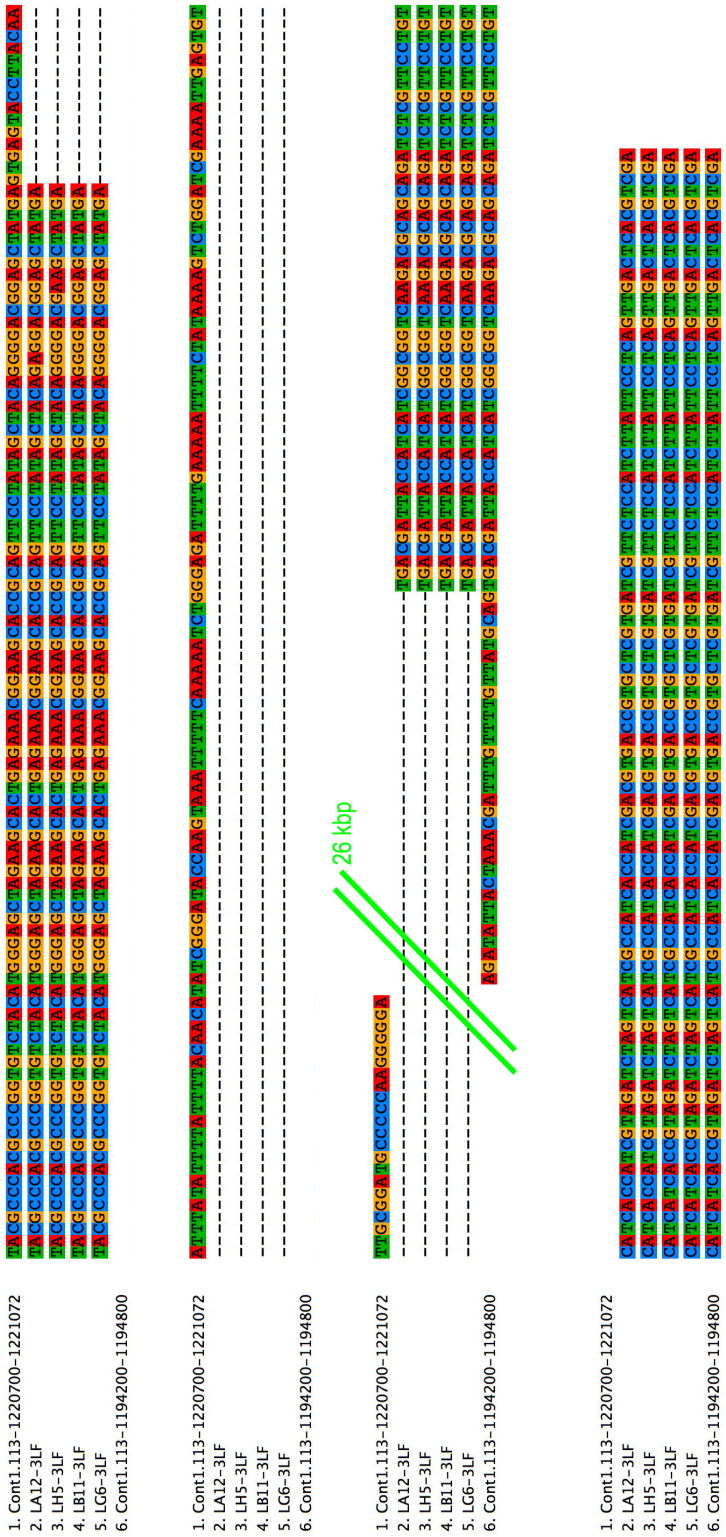

Figure A-15: Partial multiple alignment sequence of 3rd instar larvae SMART 3' RACE *tra-2* 3' sequences to Supercontig 1.113 positions in the *Ae. aegypti* genome database to compare the differences in the two sexes at the first splice site. The sequence spans from part of exon 1 to exon 2. The bottom sequence denotes the Supercontig 1.113 region. Double-slanted green bars denotes position of 26 kbp gap in the sequence.

Figure A-16: Partial multiple alignment sequence of 3^{rd} instar larvae SMART 3' RACE *tra-2* 3' sequences to Supercontig 1.113 positions in the *Ae. aegypti* genome database. Sequence alignment spans from partial exon 2 to the 3' end. The bottom sequence denotes the Supercontig 1.113 region. Purple bar denotes gap in alignment.

Table A-4: Clone position on Supercontig 1.113. BLASTn result of candidate clones from RACE experiment corresponding to the positions of Supercontig 1.113

Clone	Positions on Supercontig 1.113
Mal-B1MF	1220842-1221072, 1194473-1194764
Mal-G1MF	1220842-1221072, 1194278-1194767
Fem-H10F	1220842-1220985, 1194278-1194767
Fem-C2R3	1220842-1221075, 1194278-1194767
LA12-3LF	1220842-1221006, 1194589 - 1194767, 1194221-1194397
LH5-3LF	1220842-1221006, 1194221-1194397, 1194587-119467
LB11-3LF	1220842-1221006, 1194587-1194767, 1194223-1194397
LG6-3LF	1220842-1221006, 1194587-1194767, 1194244-1194397
MalC6-1R	1220842-1221006,1194481-1194767
MalA3-2R	1220842-1221006, 1194589-1194767, 1194281-1194397
MalH8-1R	1220842-1221006, 1194586-1194767, 1194271-1194395
MalG11-2	1220842-1221006, 1194589-1194767, 1194255-1194395
MalG7-1R	1220842-1221006, 1194587-1194767, 1194223-1194397
MalD5-1	1220842-1221006, 1194474-1194767, 1194226-1194399
FemE7-BF	1194555-1194757, 1194271-1194460
FemF2-AF	1220842-1221006, 1194589-1194767, 1194271-1194395
FemF8-BF	1220842-1221006, 1194586-1194767, 1194263-1194395
FemD10-AR	1220842-1221006, 1194586-1194767, 1194256-1194395
FemG6-AF	1220842-1221006, 1194589-1194767, 1194247-1194397
FemB6-BF	1220842-1221006, 1194589-1194767, 1194245-1194397
FemE6-AR	1220842-1221006,1194586-1194767, 1194223-1194395

A-5.1 Sequenced clones from RACE

A-5.1.1 5' RACE clones

Mal-35-DF:

ACTTCTAGACCGCGCAGCTGATAAAATCAAAATGGTTCAGTCCCG
AAGCCGTAGTCGAAGTCCGCCGCGCAGTTACCGCAAGAACAACAG
CAGCGATGAAGGTCGCCACCGAGAGCATCGCCATCGCAGCCGTAC
GACGCATGGTCACCGTTCGTTCCGGCACCGTCGTCATCACTACCG
TAGCGTTAGCAGCACTGCCAGCGAAACATCGGCCTGTTCATCCTG
CCGGTCCACTCCGGAACGTTCCAACCGGCCGGTGGATCCCCCGAA
GAGTAAGTGCCTCGGTGTGTTCGGCCTAAGCAGCTACACCAACGA
AACCAGCCTGATGGACGTTTTCGCACCGTACGGAACCATTGACAA
GGCGATGATTGTCTACGATGCCAAGACGAAGGTTTCCCGAGGGTT
CGGATTCGTGTACTTCCAGGAGCAGAGTGCGGCCACCGAAGCCAA
AATGCAGTGTAATGGAATGATGCTGCATGAGCGCACGATTAGAGT
GGATTATTCGGTGACCGAAAGACCGCATACGCCCACGCCCGGTGT
CTATCATGG

Mal-35-FF:

GATTCCAATTTTGCTTCTAGACCGCGCAGCTGATAAAATCAAAAT
GGTTCAGTCCCGAAGCCGTAGTCGAAGTCCGCCGCGCAGTTACCG
CAAGAACAACAGCAGCGATGAAGGTCGCCACCGAGAGCATCGCCA
TCGCAGCCGTACGACGCATGGTCACCGTTCGTTCCGGCACCGTCG
TCATCACTACCGTAGCGTTAGCAGCACTGCCAGCGAAACATCGGC
CTGTTCATCCTGCCGGTCCACTCCGGAACGTTCCAACCGGCCGGT
GGATCCCCCGAAGAGTAAGTGCCTCGGTGTGTTCGGCCTAAGCAG
CTACACCAACGAAACCAGCCTGATGGACGTTTTCGCACCGTACGG
AACCATTGACAAGGCGATGATTGTCTACGATGCCAAGACGAAGGT
TTCCCGAGGGTTCGGATTCGTGTACTTCCAGGAGCAGAGTGCGGC
CACCGAAGCCAAAATGCAGTGTAATGGAATGATGCTGCATGAGCG
CACGATTAGAGTGGATTATTCGGTGACCGAAAGACCGCATACGCC
CACGCCCGGTGTCTACATGG

Mal-35-GR:

GATTCCAATTTTGCTTCTAGACCGCGCAGCTGATAAAATCAAAAT
GGTTCAGTCCCGAAGCCGTAGTCGAAGTCCGCCGCGCAGTTACCG
CAAGAACAACAGCAGCGATGAAGGTCGCCACCGAGAGCATCGCCA
TCGCAGCCGTACGACGCATGGTCACCGTTCGTTCCGGCACCGTCG
TCATCACTACCGTAGCGTTAGCAGCACTGCCAGCGAAACATCGGC
CTGTTCATCCTGCCGGTCCACTCCGGAACGTTCCAACCGGCCGGT
GGATCCCCCGAAGAGTAAGTGCCTCGGTGTGTTCGGCCTAAGCAG
CTACACCAACGAAACCAGCCTGATGGACGTTTTCGCACCGTACGG
AACCATTGACAAGGCGATGATTGTCTACGATGCCAAGACGAAGGT
TTCCCGAGGGTTCGGATTCGTGTACTTCCAGGAGCAGAGTACGGC

CACCGAAGCCAAAATGCAGTGTAATGGAATGATGCTGCATGAGCG
CACGATTAGAGTGGATTATTCGGTGACCGAAAGACCGCATACGCC
CACGCCCGGTGTCTACATGG

Mal-35-GF:

GATTCCAATTTTACTTCTAGACCGCGCAGCTGATAAAATCAAAAT
GGTTCAGTCCCGAAGCCGTAGTCGAAGTCCGCCGCGCAGTTACCG
CAAGAACAACAGCAGCGATGAAGGTCGCCACCGAGAGCATCGCCA
TCGCAGCCGTACGACGCATGGTCACCGTTCGTTCCGGCACCGTCG
TCATCACTACCGTAGCGTTAGCAGCACTGCCAGCGAAACATCGGC
CTGTTCATCCTGCCGGTCCACTCCGGAACGTTCCAACCGGCCGGT
GGATCCCCCGAAGAGTAAGTGCCTCGGTGTGTTCGGCCTAAGCAG
CTACACCAACGAAACCAGCCTGATGGACGTTTTCGCACCGTACGG
AACCATTGACAAGGCGATGATTGTCTACGATGCCAAGACGAAGGT
TTCCCGTGGGTTCGGATTCGTGTACTTCCAGGAGCAGAGTGCGGC
CACCGAAGCCAAAATGCAGTGCAATGGGATGATGCTGCATGAGCG
CACGATTAGAGTGGATTATTCGGTGACCGAAAGACCGCATACGCC
CACGCCCGGTGTCTACATAAATG

Mal-55-CF:

GACATTACAATTTTGCTTCTAGACCGCGCAGCTGATAAAATCAAA
ATGGTTCAGTCCCGAAGCGTTAGTCGAACTCCACCGCGCAGTTAC
CGCAAGAACAACAGCAGCGATGAAGGTCGCCACCGAGAGCATCGC
CATCGCAGCCGTACGACGCATGGTCACCGTTCGTTCCGGCACCGT
CGTCATCACTACCGTAGCGTTAGCAGCAGTGCCAGCGAAACATCG
GCCTGTTCATCCTGCCGGTCCACTCCGGAACGTTCCAACCGGCCG
GTGGATCCCCCGAAGAGTAAGTGCCTCGGTGTGTTCGGCCTAAGC
AGCTACACCAACGAAACCAGCCTGATGGACGTTTTCGCACCGTAC
GGAACCATTGACAAGGCGATGATTGTCTACGATGCCAAGACGAAG
GTTTCCCGTGGGTTCGGATTCGTGTACTTCCAGGAGCAGAGTGCG
GCCACCGAAGCCAAAATGCAGTGCAATGGGATGATGCTGCATGAG
CGCACGATTAGAGTGGATTATTCGGTGACCGAAAGACCGCATACG
CCCACGCCCGGTGTCTACATAGAAT

Mal-35-BR:

GACACATTCCAATTTTGCTTCTAGACCGCGCAGCTGATAAAATCA
AAATGGTTCAGTCCCGAAGCCGTAGTCGAAGTCCGCCGCGCAGTT
ACCGCAAGAACAACAGCAGCGATGAAGGTCGCCACCGAGAGCATC
GCCATCGCAGCCGTACGACGCATGGTCACCGTTCGTTCCGGCACC
GTCGTCATCACTACCGTAGCGTTAGCAGCACTGCCAGCGAAACAT
CGGCCTGTTCATCCTGCCGGTCCACTCCGGAACGTTCCAACCGGC
CGGTGGATCCCCCGAAGAGTAAGTGCCTCGGTGTGTTCGGCCTAA
GCAGCTACACCAACGAAACCAGCCTGATGGACGTTTTCGCACCGT
ACGGAACCATTGACAAGGCGATGATTGTCTACGATGCCAAGACGA
AGGTTTCCCGAGGGTTCGGATTCGTGTACTTCCAGGAGCAGAGTG
CGGCCACCGAAGCCAAAATGCAGTGTAATGGAATGATGCTGCATG

AGCGCACGATTAGAGTGGATTATTCGGTGACCGAAAGACCGCATA
CGCCCACGCCCGGTGTCTACATGG

Mal-55-BF:

GACATTCCAATTTTGCTTCTAGACCGCGCAGCTGATAAAATCAAA
ATGGTTCAGTCCCGAAGCCGTAGTCGAAGTCCGCCGCGCAGTTAC
CGCAAGAACAACAGCAGCGATGAAGGTCGCCACCGAGAGCATCGC
CATCGCAGCCGTACGACGCATGGTCACCGTTCGTTCCGGCACCGT
CGTCATCACTACCGTAGCGTTAGCAGCACTGCCAGCGAAACATCG
GCCTGTTCATCCTGCCGGTCCACTCCGGAACGTTCCAACCGGCCG
GTGGATCCCCCGAAGAGTAAGTGCCTCGGTGTGTTCGGCCTAAGC
AGCTACACCAACGAAACCAGCCTGATGGACGTTTTCGCACCGTAC
GGAACCATTGACAAGGCGATGATTGTCTACGATGCCAAGACGAAG
GTTTCCCGAGGGTTCGGATTCGTGTACTTCCAGGAGCAGAGTGCG
GCCACCGAAGCCAAAATGCAGTGTAATGGAATGATGCTGCATGAG
CGCACGATTAGAGTGGATTATTCGGTGACCGAAAGACCGCATACG
CCCACGCCCGGTGTCTACATGGAATG

Mal-55-DR:

AAGAAAATTTGAATTTAAAATTTGCACATTACAATTTTGCTTCTA
GACCGCGCAGCTGATAAAATCAAAATGGTTCAGTCCCGAAGCCGT
AGTTGAACTCCACCGCGCAGTTACCGCAAGAACAACAGCAGCGAT
GAAGGTCGCCACCGAGAGCATCGCCATCGCAGCCGTACGACGCAT
GGTCACCGTTCGTTCCGGCACCGTCGTCATCACTACCGTAGCGTT
AGCAGCAGTGCCAGCGAAACATCGGCCTGTTCATCCTGCCGGTCC
ACTCCGGAACGTTCCAACCGGCCGGTGGATCCCCCGAAGAGTAAG
TGCCTCGGTGTGTTCGGCCTAAGCAGCTACACCAACGAAACCAGC
CTGATGGACGTTTTCGCACCGTACGGAACCATTGACAAGGCGATG
ATTGTCTACGATGCCAAGACGAAGGTTTCCCGTGGGTTCGGATTC
GTGTACTTCCAGGAGCAGAGTGCGGCCACCGAAGCCAAAATGCAG
TGCAATGGGATGATNCTGCATGAGCGCACGATTAGAGTGGATTAT
TCGGTGACCGAAAGACCGCATACGCCCACGCCCGGTGTCTATCAT
GG

Fem-55-AF:

GGTACGCGGGGAGAGCATCGCCATCGCAGCCGTACGACGCATGGT
CACCGTTCGTTCCGGCACCGTCGTCATCACTACCGTAGCGTTAGC
AGCACTGCCAGCGAAACATCGGCCTGTTCATCCTGCCGGTCCACT
CCGGAACGTTCCAACCGGCCGGTGGATCCCCCGAAGAGTAAGTGC
CTCGGTGTGTTCGGCCTAAGCAGCTACACCAACGAAACCAGCCTG
ATGGACGTTTTCGCACCGTACGGAACCATTGACAAGGCGATGATT
GTCTACGATGCCAAGACGAAGGTTTCCCGAGGGTTCGGATTCGTG
TACTTCCAGGAGCAGAGTGCGGCCACCGAAGCCAAAATGCAGTGT
AATGGAATGATGCTGCATGAGCGCACGATTAGAGTGGATTATTCG
GTGACCGAAAGACCGCATACGCCCACGCCCGGTGTCTACATGGAA
TGC

Fem-55-AR:

GTACGCGGGGAGAGCATCGCCATCGCAGCCGTACGACGCATGGTC
ACCGTTCGTTCCGGCACCGTCGTCATCACTACCGTAGCGTTAGCA
GCACTGCCAGCGAAACATCGGCCTGTTCATCCTGCCGGTCCACTC
CGGAACGTTCCAACCGGCCGGTGGATCCCCCGAAGAGTAAGTGCC
TCGGTGTGTTCGGCCTAAGCAGCTACACCAACGAAACCAGCCTGA
TGGACGTTTTCGCACCGTACGGAACCATTGACAAGGCGATGATTG
TCTACGATGCCAAGACGAAGGTTTCCCGAGGGTTCGGATTCGTGT
ACTTCCAGGAGCAGAGTGCGGCCACCGAAGCCAAAATGCAGTGTA
ATGGAATGATGCTGCATGAGCGCACGATTAGAGTGGATTATTCGG
TGACCGAAAGACCGCATACGCCCACGCCCGGTGTCTACATGG

Fem-35-DF:

ACTTCTAGACCGCGCAGCTGATAAAATCAAAATGGTTCAGTCCCG
AAGCCGTAGTCGAAGTCCGCCGCGCAGTTACCGCAAGAACAACAG
CAGCGATGAAGGTCGCCACCGAGAGCATCGCCATCGCAGCCGTAC
GACGCATGGTCACCGTTCGTTCCGGCACCGTCGTCATCACTACCG
TAGCGTTAGCAGCACTGCCAGCGAAACATCGGCCTGTTCATCCTG
CCGGTCCACTCCGGAACGTTCCAACCGGCCGGTGGATCCCCCGAA
GAGTAAGTGCCTCGGTGTGTTCGGCCTAAGCAGCTACACCAACGA
AACCAGCCTGATGGACGTTTTCGCACCGTACGGAACCATTGACAA
GGCGATGATTGTCTACGATGCCAAGACGAAGGTTTCCCGAGGGTT
CGGATTCGTGTACTTCCAGGAGCAGAGTGCGGCCACCGAAGCCAA
AATGCAGTGTAATGGAATGATGCTGCATGAGCGCACGATTAGAGT
GGATTATTCGGTGACCGAAAGACCGCATACGCCCACGCCCGGTGT
CTACATGG

Fem-35-GR:

GATTCCAATTTTGCTTCTAGACCGCGCAGCTGATAAAATCAAAAT
GGTTCAGTCCCGAAGCCGTAGTCGAAGTCCGCCGCGCAGTTACCG
CAAGAACAACAGCAGCGATGAAGGTCGCCACCGAGAGCATCGCCA
TCGCAGCCGTACGACGCATGGTCACCGTTCGTTCCGGCACCGTCG
TCATCACTACCGTAGCGTTAGCAGCACTGCCAGCGAAACATCGGC
CTGTTCATCCTGCCGGTCCACTCCGGAACGTTCCAACCGGCCGGT
GGATCCCCCGAAGAGTAAGTGCCTCGGTGTGTTCGGCCTAAGCAG
CTACACCAACGAAACCAGCCTGATGGACGTTTTCGCACCGTACGG
AACCATTGACAAGGCGATGATTGTCTACGATGCCAAGACGAAGGT
TTCCCGAGGGTTCGGATTCGTGTACTTCCAGGAGCAGAGTACGGC
CACCGAAGCCAAAATGCAGTGTAATGGAATGATGCTGCATGAGCG
CACGATTAGAGTGGATTATTCGGTGACCGAAAGACCGCATACGCC
CACGCCCGGTGTCTACATGG

Fem-35-FF:

GATTCCAATTTTGCTTCTAGACCGCGCAGCTGATAAAATCAAAAT
GGTTCAGTCCCGAAGCCGTAGTCGAAGTCCGCCGCGCAGTTACCG

CAAGAACAACAGCAGCGATGAAGGTCGCCACCGAGAGCATCGCCA
TCGCAGCCGTACGACGCATGGTCACCGTTCGTTCCGGCACCGTCG
TCATCACTACCGTAGCGTTAGCAGCACTGCCAGCGAAACATCGGC
CTGTTCATCCTGCCGGTCCACTCCGGAACGTTCCAACCGGCCGGT
GGATCCCCCGAAGAGTAAGTGCCTCGGTGTGTTCGGCCTAAGCAG
CTACACCAACGAAACCAGCCTGATGGACGTTTTCGCACCGTACGG
AACCATTGACAAGGCGATGATTGTCTACGATGCCAAGACGAAGGT
TTCCCGAGGGTTCGGATTCGTGTACTTCCAGGAGCAGAGTGCGGC
CACCGAAGCCAAAATGCAGTGTAATGGAATGATGCTGCATGAGCG
CACGATTAGAGTGGATTATTCGGTGACCGAAAGACCGCATACGCC
CACGCCCGGTGTCTACATGG

Fem-35-GF:

GATTCCAATTTTACTTCTAGACCGCGCAGCTGATAAAATCAAAAT
GGTTCAGTCCCGAAGCCGTAGTCGAAGTCCGCCGCGCAGTTACCG
CAAGAACAACAGCAGCGATGAAGGTCGCCACCGAGAGCATCGCCA
TCGCAGCCGTACGACGCATGGTCACCGTTCGTTCCGGCACCGTCG
TCATCACTACCGTAGCGTTAGCAGCACTGCCAGCGAAACATCGGC
CTGTTCATCCTGCCGGTCCACTCCGGAACGTTCCAACCGGCCGGT
GGATCCCCCGAAGAGTAAGTGCCTCGGTGTGTTCGGCCTAAGCAG
CTACACCAACGAAACCAGCCTGATGGACGTTTTCGCACCGTACGG
AACCATTGACAAGGCGATGATTGTCTACGATGCCAAGACGAAGGT
TTCCCGTGGGTTCGGATTCGTGTACTTCCAGGAGCAGAGTGCGGC
CACCGAAGCCAAAATGCAGTGCAATGGGATGATGCTGCATGAGCG
CACGATTAGAGTGGATTATTCGGTGACCGAAAGACCGCATACGCC
CACGCCCGGTGTCTACATGGGTG

Fem-35-EF:

GATTCCAATTTTACTTCTAGACCGCGCAGCTGATAAAATCAAAAT
GGTTCAGTCCCGAAGCCGTAGTCGAAGTCCGCCGCGCAGTTACCG
CAAGAACAACAGCAGCGATGAAGGTCGCCACCGAGAGCATCGCCA
TCGCAGCCGTACGACGCATGGTCACCGTTCGTTCCGGCACCGTCG
TCATCACTACCGTAGCGTTAGCAGCACTGCCAGCGAAACATCGGC
CTGTTCATCCTGCCGGTCCACTCCGGAACGTTCCAACCGGCCGGT
GGATCCCCCGAAGAGTAAGTGCCTCGGTGTGTTCGGCCTAAGCAG
CTACACCAACGAAACCAGCCTGATGGACGTTTTCGCACCGTACGG
AACCATTGACAAGGCGATGATTGTCTACGATGCCAAGACGAAGGT
TTCCCGTGGGTTCGGATTCGTGTACTTCCAGGAGCAGAGTGCGGC
CACCGAAGCCAAAATGCAGTGCAATGGGATGATGCTGCATGAGCG
CACGATTAGAGTGGATTATTCGGTGACCGAAAGACCGCATACGCC
CACGCCCGGTGTCTACATGG

Fem-55-FF:

GATTCCAATTTTGCTTCTAGACCGCGCAGCTGATAAAATCAAAAT
GGTTCAGTCCCGAAGCCGTAGTCGAAGTCCGCCGCGCAGTTACCG
CAAGAACAACAGCAGCGATGAAGGTCGCCACCGAGAGCATCGCCA

TCGCAGCCGTACGACGCATGGTCACCGTTCGTTCCGGCACCGTCG
TCATCACTACCGTAGCGTTAGCAGCACTGCCAGCGAAACATCGGC
CTGTTCATCCTGCCGGTCCACTCCGGAACGCTCCAACCGGCCGGT
GGATCCCCCGAAGAGTAAGTGCCTCGGTGTGTTCGGCCTAAGCAG
CTACACCGACGAAACCAGCCTGATGGACGTTTTCGCACCGTACGG
AACCATTGACAAGGCGATGATTGTCTACGATGCCAAGACGAAGGT
TTCCCGTGGGTTCGGATTCGTGTACTTCCAGGAGCAGAGTGCGGC
CACCGAAGCCAAAATGCAGTGCAATGGGATGATGCTGCATGAGCG
CACGATTAGAGTGGATTATTCGGTGACCGAAAGACCGCATACGCC
CAGGCCCGGTGTCTACATGGG

Fem-55-CF:

GACATTACAATTTTGCTTCTAGACCGCGCAGCTGATAAAATCAAA
ATGGTTCAGTCCCGAAGCGTTAGTCGAACTCCACCGCGCAGTTAC
CGCAAGAACAACAGCAGCGATGAAGGTCGCCACCGAGAGCATCGC
CATCGCAGCCGTACGACGCATGGTCACCGTTCGTTCCGGCACCGT
CGTCATCACTACCGTAGCGTTAGCAGCAGTGCCAGCGAAACATCG
GCCTGTTCATCCTGCCGGTCCACTCCGGAACGTTCCAACCGGCCG
GTGGATCCCCCGAAGAGTAAGTGCCTCGGTGTGTTCGGCCTAAGC
AGCTACACCAACGAAACCAGCCTGATGGACGTTTTCGCACCGTAC
GGAACCATTGACAAGGCGATGATTGTCTACGATGCCAAGACGAAG
GTTTCCCGTGGGTTCGGATTCGTGTACTTCCAGGAGCAGAGTGCG
GCCACCGAAGCCAAAATGCAGTGCAATGGGATGATGCTGCATGAG
CGCACGATTAGAGTGGATTATTCGGTGACCGAAAGACCGCATACG
CCCACGCCCGGTGTCTACATGGAAT

Fem-55-BF:

GACATTCCAATTTTGCTTCTAGACCGCGCAGCTGATAAAATCAAA
ATGGTTCAGTCCCGAAGCCGTAGTCGAAGTCCGCCGCGCAGTTAC
CGCAAGAACAACAGCAGCGATGAAGGTCGCCACCGAGAGCATCGC
CATCGCAGCCGTACGACGCATGGTCACCGTTCGTTCCGGCACCGT
CGTCATCACTACCGTAGCGTTAGCAGCACTGCCAGCGAAACATCG
GCCTGTTCATCCTGCCGGTCCACTCCGGAACGTTCCAACCGGCCG
GTGGATCCCCCGAAGAGTAAGTGCCTCGGTGTGTTCGGCCTAAGC
AGCTACACCAACGAAACCAGCCTGATGGACGTTTTCGCACCGTAC
GGAACCATTGACAAGGCGATGATTGTCTACGATGCCAAGACGAAG
GTTTCCCGAGGGTTCGGATTCGTGTACTTCCAGGAGCAGAGTGCG
GCCACCGAAGCCAAAATGCAGTGTAATGGAATGATGCTGCATGAG
CGCACGATTAGAGTGGATTATTCGGTGACCGAAAGACCGCATACG
CCCACGCCCGGTGTCTACATGGAATG

Fem-35-BFR:

GACACATTCCAATTTTGCTTCTAGACCGCGCAGCTGATAAAATCA
AAATGGTTCAGTCCCGAAGCCGTAGTCGAAGTCCGCCGCGCAGTT
ACCGCAAGAACAACAGCAGCGATGAAGGTCGCCACCGAGAGCATC
GCCATCGCAGCCGTACGACGCATGGTCACCGTTCGTTCCGGCACC

GTCGTCATCACTACCGTAGCGTTAGCAGCACTGCCAGCGAAACAT
CGGCCTGTTCATCCTGCCGGTCCACTCCGGAACGTTCCAACCGGC
CGGTGGATCCCCCGAAGAGTAAGTGCCTCGGTGTGTTCGGCCTAA
GCAGCTACACCAACGAAACCAGCCTGATGGACGTTTTCGCACCGT
ACGGAACCATTGACAAGGCGATGATTGTCTACGATGCCAAGACGA
AGGTTTCCCGAGGGTTCGGATTCGTGTACTTCCAGGAGCAGAGTG
CGGCCACCGAAGCCAAAATGCAGTGTAATGGAATGATGCTGCATG
AGCGCACGATTAGAGTGGATTATTCGGTGACCGAAAGACCGCATA
CGCCCACGCCCGGTGTCTACATGG

Fem-35-JF:

TCAACGCATAGTACGCGGGGACATTCCAATTTTACTTCTAGACCG
CGCAGCTGATAAAATCAAAATGGTTCAGTCCCGAAGCCGTAGTCG
AAGTCCGCCGCGCAGTTACCGCAAGAACAACAGCAGCGATGAAGG
TCGCCACCGAGAGCATCGCCATCGCAGCCGTACGACGCATGGTCA
CCGTTCGTTCCGGCACCGTCGTCATCACTACCGTAGCGTTAGCAG
CACTGCCAGCGAAACATCGGCCTGTTCATCCTGCCGGTCCACTCC
GGAACGTTCCAACCGGCCGGTGGATCCCCCGAAGAGTAAGTGCCT
CGGTGTGTTCGGCCTAAGCAGCTACACCAACGAAACCAGCCTGAT
GGACGTTTTCGTACCGTACGGAACCATTGACAAGGCGATGATTGT
CTACGATGCCAAGACGAAGGTTTCCCGAGGGTTCGGATTCGTGTA
CTTCCAGGAGCAGAGTGCGGCCACCGAAGCCAAAATGCAGTGTAA
TGGAATGATGCTGCATGAGCGCACGATTAGAGTGGATTATTCGGT
GACCGAAAGACCGCATACGCCCACCCCCGGTGTCGACATGG

Fem-55-DR:

ACGTGGGGAAGAAAATTTGAATTTAAAATTTGCACATTACAATTT
TGCTTCTAGACCGCGCAGCTGATAAAATCAAAATGGTTCAGTCCC
GAAGCCGTAGTTGAACTCCACCGCGCAGTTACCGCAAGAACAACA
GCAGCGATGAAGGTCGCCACCGAGAGCATCGCCATCGCAGCCGTA
CGACGCATGGTCACCGTTCGTTCCGGCACCGTCGTCATCACTACC
GTAGCGTTAGCAGCAGTGCCAGCGAAACATCGGCCTGTTCATCCT
GCCGGTCCACTCCGGAACGTTCCAACCGGCCGGTGGATCCCCCGA
AGAGTAAGTGCCTCGGTGTGTTCGGCCTAAGCAGCTACACCAACG
AAACCAGCCTGATGGACGTTTTCGCACCGTACGGAACCATTGACA
AGGCGATGATTGTCTACGATGCCAAGACGAAGGTTTCCCGTGGGT
TCGGATTCGTGTACTTCCAGGAGCAGAGTGCGGCCACCGAAGCCA
AAATGCAGTGCAATGGGATGATGCTGCATGAGCGCACGATTAGAG
TGGATTATTCGGTGACCGAAAGACCGCATACGCCCACGCCCGGTG
TCTACATGG

Fem-55-DF:

CTAGTACGTGGGGAAGAAAATTTGAATTTAAAAATTTGCACATTA
CAATTTTGCTTCTAGACCGCGCAGCTGATAAAATCAAAATGGTTC
AGTCCCGAAGCGTTAGTCGAACTCCACCGCGCAGTTACCGCAAGA
ACAACAGCAGCGATGAAGGTCGCCACCGAGAGCATCGCCATCGCA

GCCGTACGACGCATGGTCACCGTTCGTTCCGGCACCGTCGTCATC
ACTACCGTAGCGTTAGCAGCAGTGCCAGCGAAACATCGGCCTGTT
CATCCTGCCGGTCCACTCCGGAACGTTCCAACCGGCCGGTGGATC
CCCCGAAGAGTAAGTGCCTCGGTGTGTTCGGCCTAAGCAGCTACA
CCAACGAAACCAGCCTGATGGACGTTTTCGCACCGTACGGAACCA
TTGACAAGGCGATGATTGTCTACGATGCCAAGACGAAGGTTTCCC
GTGGGTTCGGATTCGTGTACTTCCAGGAGCAGAGTGCGGACACCG
AAGCCAAAATGCAGTGCAATGGGATGATGCTGCATGAGCGCACGA
TTAGAGTGGATTATTCGGTGACCGAAAGACCGCATACGCCCACGC
CCGGTGTCTACATGG

A-5.1.2 3' RACE Clones

Mal-B1MF:

CAAGACGAAGGTTTCCCGAGGGTTCGGATTCGTGTACTTCCAGGA
GCAGAGTGCGGCCACCGAAGCCAAAATGCAGTGTAATGGAATGAT
GCTGCATGAGCGCACGATTAGAGTGGATTATTCGGTGACCGAAAG
ACCGCATACGCCCACGCCCGGTGTCTACATGGGAGCTAGAAGCAC
TGAGAAACGGAAGCACCGCAGTTCCTATAGCTACAGGGGACGGAG
CTATGATGGCGATTACCATCATCGGCGGTCAAGACGCAGCAGATC
TCGTTCCTGTCATCATCACCGTAGATCTAGTCATCGCCATCACCAT
CGACGTGACCGTGCTCGTGATCGTTCTCCATCTTATTCCTCAGTT
GACTCACGTCGATCCTATCGATAATGTGCTAGAAGGATTGTGTTT
TTGACGTAAGTCATCTTTTTCAAGACCTCACGAAGACCCTACAAA
CAAAGAAGTTCAAGTCAAGTATTATTGAAGATAGAATCTGTATCA
GTAAAAACTATTCTTCAAATGACACGAGGGAAAAAAAAAAAAAAA
AAAAAAAA

Mal-G1MF:

CAAGACGAAGGTTTCCCGAGGGTTCGGATTCGTGTACTTCCAGGA
GCAGAGTGCGACCACCGAAGCCAAAATGCAGTGTAATGGAATGAT
GCTGCATGAGCGCACGATTAGAGTGGATTATTCGGTGACCGAAAG
ACCGCATACGCCCACGCCCGGTGTCTACATGGGAGCTAGAAGCAC
TGAGAAACGGAAGCACCGCAGTTCCTATAGCTACAGGGGACGGAG
CTATGATGACGATTACCATCATCGGCGGTCAAGACGCAGCAGATC
TCGTTCCTGTCATCATCACCGTAGATCTAGTCATCGCCATCACCAT
CGACGTGACCGTGCTCGTGATCGTTCTCCATCTTATTCCTCAGTT
GACTCACGTCGATCCTATCGATAATGTGCTAGAAGGATTGTGTTT
TTGACGTAAGTCATCTTTTTCAAGACCTCACGAAGACCCTACAAA
CAAAGAAGTTCAAGTCAAGTATTATTGAAGATAGAATCTGTATCA
GTAAAAACTATTCTTCAAATGACACAAGGGAAGAACAGATTACTC
TTCATTAATCCAAGGGTAAATTGTATGTAGCTAACCGTTCTGTTT
TTCTTTTCGTTTCAGGATGTTTGGATTTCTTCACATTTTAGAGAA
CTAGTTTTCATTATTCCGCGATTCAAACTAAACATTTTATTTATTT

ATTAGTTAAAGATCGTGTGGTTATGAAATGTCGTTTGTGCAACCA
TTAACCAAAAATAAGGGAAAAAAAAAAAAAAAAAAAAAAA

Fem-H10F:

GATGCTGCATGAGCGCACGATTAGAGTGGATTATTCGGTGACCGA
AAGACCGCATACGCCCACGCCCGGTGTCTACATGGGAGCTAGAAG
CACTGAGAAACGGAAGCACCGCAGTTCCTATAGCTACAGGGGACG
GAGCTATGATGACGATTACCATCATCGGCGGTCAAGACGCAGCAG
ATCTCGTTCCTGTCATCACCATCGTAGATCTAGTCATCGCCATCAC
CATCGACGTGACCGTGCTCGTGGTCGTTCTCCATCTTATTCCTCA
GTTGACTCACGTCGATCCTATCGATAATGTGCTAGAAGGATTGTG
TTTTTGATGTAAGTCATCTTTTTCAAGACCTCACGAAGACCCTAC
AAACAAAGAAGTTCACGTCAAGTATTATTGAAGATAGAATCTGTC
TCAGCAAAAACTACTCTTCAAATGACACACGGGAAGACCAGATTA
CTCTTCATTAATCCAAGGGTAAATCATATGTAACTAACCGATCTG
TTTTTCTTTTCGCTCCAGGATGTTTGGATTCCTTCCGATTTTAGA
GAACTAGTTTTCTTTGTTCCGCGATTCAAACTAAACCTTTTATTTA
TTTATTAGTAAAAGACCGTGTGGCTATGAAATGTCGTTTGGGCAA
CCATTAACCAAAAAAAAAAAAAAAA

Fem-C2R3:

TGCCAAGACGAAGGTTTCCCGTGGGTTCGGATTCGTGTACTTCCA
GGAGCAGAGTGCGGCCACCGAAGCCAAAATGCAGTGCAATGGGAT
GATGCTGCATGAGCGCACGATTAGAGTGGATTATTCGGTGACCGA
AAGACCGCATACGCCCACGCCCGGTGTCTACATGGGAGCTAGAAG
CACTGAGAAACGGAAGCACCGCAGTTCCTATAGCTACAGGGGACG
GAGCTATGATGACGATTACCATCATCGGCGGTCAAGACGCAGCAG
ATCTCGTTCCTGTCATCACCATCGTAGATCTAGTCATCACCATCGC
CATCGACGTGACCGTGCTCGTGATCGTTCTCCATCTTATTCCTCA
GTTGACTCACGTCGATCCTATCGATAATGTGCTAGAAGGATTGTG
TTTTTGATGTAAGTCATCTTTTTCAAGACCTCACGAAGACCCTAC
AAACAAAGAAGTTCACGTCAAGTATTATTGAAGATAGAATCTGTA
TCAGTAAAAACTACTCTTCAAATGACACAAGGGAAGAACAGATTA
CTCTTCATTAATCCAAGGGTAAATCATATGTAGCTAACCGATCTG
TTTTTCTTTTCGTTTCAGGATGTTTGGATTTCTTCACATTTTAGA
GAACTAGTTTTCATGGTTCCGCGATTCAAACTAAACATTTTATTT
ATTTATTAGTAAAAGATCGTGTGGTTATGAAATGTCGTTTGGGCA
ACCATTAACCAAAAAAAAAAAAAAAAAAAAA

LA12-3LF:

CAAAATGCAGTGCAATGGGATGATGCTGCATGAGCGCACGATTAG
AGTGGATTATTCGGTGACCGAAAGACCGCATACGCCCACGCCCGG
TGTCTACATGGGAGCTAGAAGCACTGAGAAACGGAAGCACCGCAG
TTCCTATAGCTACAGAGGACGGAGCTATGATGACGATTACCATCA
TCGGCGGTCAAGACGCAGCAGATCTCGTTCCTGTCATCACCATCG
TAGATCTAGTCATCGCCATCACCATCGACGTGACCGTGCTCGTGA

TCGTTCTCCATCTTATTCCTCAGTTGACTCACGTCGATCCTATCGA
TAATGTGCTAGAAGGATTGTGTTTTTGATGATGTTTGAATTTCTT
CACATTTTAGAGAACTAGTTTTCATTGTTCCGCGATTCAAACTAA
ACATTTTATTTATTTATTAGTTAAAGATCGTGTGGTTTATGAAAT
GTCGTTTGTGCAACAATAAACAAAAATAAAGAAACCGACTGCATC
GCAATTGATTCAGTCATTCAAAGATCAAAAAAAAAAAAAAAAAAA
AAAAAAAAA

LH5-3LF:

CAAAATGCAGTGTAATGGAATGATGCTGCATGAGCGCACGATTAG
AGTGGATTATTCGGTGACCGAAAGACCGCATACGCCCACGCCCGG
TGTCTACATGGGAGCTAGAAGCACTGAGAAACGGAAGCACCGCAG
TTCCTATAGCTACAGGGGACGAAGCTATGATGACGATTACCATCA
TCGGCGGTCAAGACGCAGCAGATCTCGTTCCTGTCATCACCATCG
TAGATCTAGTCATCGCCATCACCATCGACGTGACCGTGCTCGTGA
TCGTTCTCCATCTTATTCCTCAGTTGACTCACGTCGATCCTATCGA
TAATGTGCTAGAAGGATTGTGTTTTTGACGATGTTTGGATTTCTT
CACATTTTAGAGAACTAGTTTTCATTATTCCGCGATTCAAACTAA
ACATTTTATTTATTTATTAGTTAAAGATCGTGTGGTTTATGAAAT
GTCGTTTGTGCAACAATAAACAAAAATAAAGAAACCGACTGCATC
GCAATTGATTCAGTCATTCAAAGATCAAAAAAAAAAAAAAAAAAA
AAAAAAAA

LB11-3LF:

CAAAATGCAGTGTAATGGAATGATGCTGCATGAGCGCACGATTAG
AGTGGATTATTCGGTGACCGAAAGACCGCATACGCCCACGCCCGG
TGTCTACATGGGAGCTAGAAGCACTGAGAAACGGAAGCACCGCAG
TTCCTATAGCTACAGGGGACGGAGCTATGATGACGATTACCATCA
TCGGCGGTCAAGACGCAGCAGATCTCGTTCCTGTCATCATCACCG
TAGATCTAGTCATCGCCATCACCATCGACGTGACCGTGCTCGTGA
TCGTTCTCCATCTTATTCCTCAGTTGACTCACGTCGATCCTATCGA
TAATGTGCTAGAAGGATTGTGTTTTTGACGATGTTTGGATTTCTT
CACATTTTAGAGAACTAGTTTTCATTATTCCGCGATTCAAACTAA
ACATTTTATTTATTTATTAGTTAAAGATCGTGTGGTTTATGAAAT
GTCGTTTGTGCAACAATAAACAAAAATAAAGAAACCGACTGCATC
GCAATTGATTCAGTCATTCAAAGAGCAAAAAAAAAAAAAAAAAAA
AA

LG6-3LF:

CAAAATGCAGTGTAATGGAATGATGCTGCATGAGCGCACGATTAG
AGTGGATTATTCGGTGACCGAAAGACCGCATACGCCCACGCCCGG
TGTCTACATGGGAGCTAGAAGCACTGAGAAACGGAAGCACCGCAG
TTCCTATAGCTACAGGGGACGGAGCTATGATGACGATTACCATCA
TCGGCGGTCAAGACGCAGCAGATCTCGTTCCTGTCATCATCACCG
TAGATCTAGTCATCGCCATCACCATCGACGTGACCGTGCTCGTGA
TCGTTCTCCATCTTATTCCTCAGTTGACTCACGTCGATCCTATCGA

TAATGTGCTAGAAGGATTGTGTTTTTGACGATGTTTGGATTTCTT
CACATTTTAGAGAACTAGTTTTCATTATTCCGCGATTCAAACTAA
ACATTTTATTTATTTATTAGTTAAAGATCGTGTGGTTTATGAAAT
GTCGTTTGTGCAACAATAAACAAAAATAAAGAAACCGACTGCATC
GCAGAAAAAA

MalC6-1R:

CAAAATGCAGTGTAATGGAATGATGCTGCATGAGCGCACGATTAG
AGTGGATTATTCGGTGACCGAAAGACCGCATACGCCCACGCCCGG
TGTCTACATGGGAGCTAGAAGCACTGAGAAACGGAAGCACCGCAG
TTCCTATAGCTACAGGGGACGGAGCTATGATGACGATTACCATCA
TCGGCGGTCAAGACGCAGCAGATCTCGTTCCTGTCATCATCACCG
TAGATCTAGTCATCGCCATCACCATCGACGTGACCGTGCTCGTGA
TCGTTCTCCATCTTATTCCTCAGTTGACTCACGTCGATCCTATCGA
TAATGTGCTAGAAGGATTGTGTTTTTGACGTAAGTCATCTTTTTC
AAGACCTCACGAAGACCCTACAAACAAAGAAGTTCAAGTCAAGTA
TTATTGAAGATAGAATCTGTATCAGTAAAAACTATTCTTCAAATG
AAAAA

MalA3-2R:

CAAAATGCAGTGTAATGGAATGATGCTGCATGAGCGCACGATTAG
AGTGGATTATTCGGTGACCGAAAGACCGCATACGCCCACGCCCGG
TGTCTACATGGGAGCTAGAAGCACTGAGAAACGGAAGCACCGCAG
TTCCTATAGCTACAGGGGACGGAGCTATGATGACGATTACCATCA
TCGGCGGTCAAGACGCAGCAGATCTCGTTCCTGTCATCACCATCG
TAGATCTAGTCATCGCCATCACCATCGACGTGACCGTGCTCGTGA
TCGTTCTCCATCTTATTCCTCAGTTGACTCACGTCGATCCTATCGA
TAATGTGCTAGAAGGATTGTGTTTTTGATGATGTTTGAATTTCTT
CACATTTTAGAGAACTAGTTTTCATTGTTCCGCGATTCAAACTAA
ACATTTTATTTATTTATTAGTAAAAGATCGTGTGGTTTATGAAAT
GTCGTTTGTGCAAAAAAAAAAAAAAAAAAAAA

MalH8-1R:

CAAAATGCAGTGTAATGGAATGATGCTGCATGAGCGCACGATTAG
AGTGGATTATTCGGTGACCGAAAGACCGCATACGCCCACGCCCGG
TGTCTACATGGGAGCTAGAAGCACTGAGAAACGGAAGCACCGCAG
TTCCTATAGCTACAGGGGACGGAGCTATGATGACGATTACCATCA
TCGGCGGTCAAGACGCAGCAGATCTCGTTCCTGTCATCATCACCG
TAGATCTAGTCATCGCCATCACCATCGACGTGACCGTGCTCGTGA
TCGTTCTCCATCTTATTCCTCAGTTGACTCACGTCGATCCTATCGA
TAATGTGCTAGAAGGATTGTGTTTTTGACGTGTTTGGATTTCTTC
ACATTTTAGAGAACTAGTTTTCATTATTCCGCGATTCAAACTAAA
CATTTTATTTATTTATTAGTTAAAGATCGTGTGGTTTATGAAATG
TCGTTTGTGCAACAATAAACAAAAAAAAAAAAAAAAAAAAAAAA
A

MalA1-1R:

CAAAATGCAGTGTAATGGAATGATGCTGCATGAGCGCACGATTAG
AGTGGATTATTCGGTGACCGAAAGACCGCATACGCCCACGCCCGG
TGTCTACATGGGAGCTAGAAGCACTGAGAAACGGAAGCACCGCAG
TTCCTATAGCTACAGGGGACGGAGCTATGATGACGATTACCATCA
TCGGCGGTCAAGACGCAGCAGATCTCGTTCCTGTCATCACCATCG
TAGATCTAGTCATCGCCATCACCATCGACGTGACCGTGCTCGTGA
TCGTTCTCCATCTTATTCCTCAGTTGACTCACGTCGATCCTATCGA
TAATGTGCTAGAAGGATTGTGTTTTTGATGTGTTTGAATTTCTTC
ACATTTTAGAGAACTAGTTTTCATTGTTCCGCGATTCAAACTAAA
CATTTTATTTGTTTATTAGTAAAAGATCGTGTGGTTTATGAAATG
TCGTTTGTGCAACAATAAACAAAAATAAAGAAACCGAAAAAAAAA
AAAAAAAAAAAAAAAAAA

MalG11-2:

CAAAATGCAGTGTAATGGAATGATGCTGCATGAGCGCACGATTAG
AGTGGATTATTCGGTGACCGAAAGACCGCATACGCCCACGCCCGG
TGTCTACATGGGAGCTAGAAGCACTGAGAAACGGAAGCACCGCAG
TTCCTATAGCTACAGGGGACGGAGCTATGATGACGATTACCATCA
TCGGCGGTCAAGACGCAGCAGATCTCGTTCCTGTCATCATCACCG
TAGATCTAGTCATCGCCATCACCATCGACGTGACCGTGCTCGTGA
TCGTTCTCCATCTTATTCCTCAGTTGACTCACGTCGATCCTATCGA
TAATGTGCTAGAAGGATTGTGTTTTTGACGTGTTTGGATTTCTTC
ACATTTTAGAGAACTAGTTTTCATTATTCCGCGATTCAAACTAAA
CATTTTATTTATTTATTAGTTAAAGATCGTGTGGTTTATGAAATG
TCGTTTGTGCAACAATAAACAAAAATAAAGAAACCGACTGCATCG
CAAAAAAAAAAAAAAAAAAAAAAAAA

MalG7-1R:

CAAAATGCAGTGTAATGGAATGATGCTGCATGAGCGCACGATTAG
AGTGGATTATTCGGTGACCGAAAGACCGCATACGCCCACGCCCGG
TGTCTACATGGGAGCTAGAAGCACTGAGAAACGGAAGCACCGCAG
TTCCTATAGCTACAGGGGACGGAGCTATGATGACGATTACCATCA
TCGGCGGTCAAGACGCAGCAGATCTCGTTCCTGTCATCATCACCG
TAGATCTAGTCATCGCCATCACCATCGACGTGACCGTGCTCGTGA
TCGTTCTCCATCTTATTCCTCAGTTGACTCACGTCGATCCTATCGA
TAATGTGCTAGAAGGATTGTGTTTTTGACGATGTTTGGATTTCTT
CACATTTTAGAGAACTAGTTTTCATTATTCCGCGATTCAAACTAA
ACATTTTATTTATTTATTAGTTAAAGATCGTGTGGTTTATGAAAT
GTCGTTTGTGCAACAATAAACAAAAATAAAGAAACCGACTGCATC
GCAATTGATTCAGTCATTCAAAGACCAAAAAAAAAAAAAAAAAA
AAAAA

MalD5-1:

CAAAATGCAGTGTAATGGAATGATGCTGCATGAGCGCACGATTAG
AGTGGATTATTCGGTGACCGAAAGACCGCATACGCCCACGCCCGG
TGTCTACATGGGAGCTAGAAGCACTGAGAAACGGAAGCACCGCAG
TTCCTATAGCTACAGGGGACGGAGCTATGATGACGATTACCATCA
TCGGCGGTCAAGACGCAGCAGATCTCGTTCCTGTCATCATCACCG
TAGATCTAGTCATCGCCATCACCATCGACGTGACCGTGCTCGTGA
TCGTTCTCCATCTTATTCCTCAGTTGACTCACGTCGATCCTATCGA
TAATGTGCTAGAAGGATTGTGTTTTTGACGTAAGTCATCTTTTTC
AAGACCTCACGAAGACCCCACAAACAAAGAAGTTCAAGTCAAGTA
TTATTGAAGATAGAATCTGTATCAGTAAAAACTATTCTTCAAATG
ACACAAGGATGTTTGGATTTCTTCACATTTTAGAGAACTAGTTTT
CATTATTCCGCGATTCAAACTAAACATTTTATTTATTTATTAGTTA
AAGATCGTGTGGTTTATGAAATGTCGTTTGTGCAACAATAAACAA
AATAAAGAAACCGCCTGCATCGCATGATTCAGTCATTCAAGACCA
AAAAAAAAAAAAAAAAAAAAA

FemE7-BF:

CAAAATGCAGTGCAATGGGATGATGCATCATCGGCGGTCAAGACG
CAGCAGATCTCGTTCCTGTCATCACCATCGAAGATCTAGTCATCG
CCATCACCATCGACGTGACCGTGCTCGTGATCGTTCTCCATCTTA
TTCCTCAGTTGACTCACGTCGATCCTATCGATAATGTGCTAGAAG
GATTGTGTTTTTGATGTAAGTCATCTTTTTCAAGACCTCACGAAG
ACCTCTTCATTAATCCAAGGGTAAATCATATGTAGCTAACCGATC
TGTTTTTCTTTTCGTTTCAGGATGTTTGAAGGTCTTCACATTTTA
GAGAACTAGTTTTCATTGTTCCGCGATTCAAACTAAACATTTTAT
TTATTTATTAGTAAAAGATCGTGTGGTTTATGAAATGTCGTTTGT
GCAACAATAAACAAAAAAAAAAAAAAAAAAAAA

FemF2-AF:

CAAAATGCAGTGCAATGGGATGATGCTGCATGAGCGCACGATTAG
AGTGGATTATTCGGTGACCGAAAGACCGCATACGCCCACGCCCGG
TGTCTACATGGGAGCTAGAAGCACTGAGAAACGGAAGCACCGCAG
TTCCTATAGCTACAGGGGACGGAGCTATGATGACGATTACCATCA
TCGGCGGTCAAGACGCAGCAGATCTCGTTCCTGTCATCACCATCG
TAGATCTAGTCATCGCCATCACCATCGACGTGACCGTGCTCGTGA
TCGTTCTCCATCTTATTCCTCAGTTGACTCACGTCGATCCTATCGA
TAATGTGCTAGAAGGATTGTGTTTTTGATGTGTTTGAATTTCTTC
ACATTTTAGAGAACTAGTTTTCATTGTTCCGCGATTCAAACTAAA
CATTTTATTTATTTATTAGTAAAAGATCGTGTGGTTTATGAAATG
TCGTTTGTGCAACAATAAACAAAAAAAAAAAAAAAAAAAAA

FemF8-BF:

CAAAATGCAGTGTAATGGAATGATGCTGCATGAGCGCACGATTAG
AGTGGATTATTCGGTGACCGAAAGACCGCATACGCCCACGCCCGG

TGTCTACATGGGAGCTAGAAGCACTGAGAAACGGAAGCACCGCAG
TTCCTATAGCTACAGGGGACGGAGCTATGATGACGATTACCATCA
TCGGCGGTCAAGACGCAGCAGATCTCGTTCCTGTCATCATCACCG
TAGATCTAGTCATCGCCATCACCATCGACGTGACCGTGCTCGTGA
TCGTTCTCCATCTTATTCCTCAGTTGACTCACGTCGATCCTATCGA
TAATGTGCTAGAAGGATTGTGTTTTTGACGTGTTTGGATTTCTTC
ACATTTTAGAGAACTAGTTTTCATTATTCCGCGATTCAAACTAAA
CATTTTATTTATTTATTAGTTAAAGATCGTGTGGTTTATGAAATG
TCGTTTGTGCAACAATAAACAAAAATAATGAAAAAAAAAAAAAAA
AAAAAAAAAAA

FemD10-AR:

CAAAATGCAGTGTAATGGAATGATGCTGCATGAGCGCACGATTAG
AGTGGATTATTCGGTGGCCGAAAGACCGCATACGCCCACGCCCGG
TGTCTACATGGGAGCTAGAAGCACTGAGAAACGGAAGCACCGCAG
TTCCTATAGCTACAGGGGACGGAGCTATGATGACGATTACCATCA
TCGGCGGTCAAGACGCAGCAGATCTCGTTCCTGTCATCACCATCG
TAGATCTAGTCATCGCCATCACCATCGACGTGACCGTGCTCGTGA
TCGTTCTCCACCTTATTCCTCAGTTGACTCACGTCGATCCTATCGA
TAATGTGCTAGAAGGATTGTGTTTTTGACGTGTTTGGATTTCTTC
ACATTTTAGAGAACTAGTTTTCATTATTCCGCGATTCAAACTAAA
CATTTTATTTATTTATTAGTTAAAGATCGTGTGGTTTATGAAATG
TCGTTTGTGCAACAATAAACAAAAATAAAGAAACCAAAAAAAAAA
AAAAAAAAAAAAAAA

FemG6-AF:

CAAAATGCAGTGCAATGGGATGATGCTGCATGAGCGCACGATTAG
AGTGGATTATTCGGTGACCGAAAGACCGCATACGCCCACGCCCGG
TGTCTACATGGGAGCTAGAAGCACTGAGAAACGGAAGCACCGCAG
TTCCTATAGCTACAGGGGACGGAGCTATGATGACGATTACCATCA
TCGGCGGTCAAGACGCAGCAGATCTCGTTCCTGTCATCACCATCG
TAGATCTAGTCATCGCCATCACCATCGACGTGACCGTGCTCGTGA
TCGTTCTCCATCTTATTCCTCAGTTGACTCACGTCGATCCTATCGA
TAATGTGCTAGAAGGATTGTGTTTTTGATGATGTTTGAATTTCTT
CACATTTTAGAGAACTAGTTTTCATTGTTCCGCGATTCAAACTAA
ACATTTTATTTATTTATTAGTAAAAGATCGTGTGGTTTATGAAAT
GTCGTTTGTGCAACAATAAACAAAAATAAAGAAACCGACTGCATC
AAAAAAAAAAAAAAAAAAAAAA

FemB6-BF:

CAAAATGCAGTGCAATGGGATGATGCTGCATGAGCGCACGATTAG
AGTGGATTATTCGGTGACCGAAAGACCGCATACGCCCACGCCCGG
TGTCTACATGGGAGCTAGAAGCACTGAGAAACGGAAGCACCGCAG
TTCCTATAGCTACAGAGGACGGAGCTATGATGACGATTACCATCA
TCGGCGGTCAAGACGCAGCAGATCTCGTTCCTGTCATCACCATCG
TAGATCTAGTCATCGCCATCACCATCGACGTGACCGTGCTCGTGA

TCGTTCTCCATCTTATTCCTCAGTTGACTCACGTCGATCCTATCGA
TAATGTGCTAGAAGGATTGTGTTTTTGATGATGTTTGAATTTCTT
CACATTTTAGAGAACTAGTTTTCATTGTTCCGCGATTCAAACTAA
ACATTTTATTTATTTATTAGTAAAAGATCGTGTGGTTTATGAAAT
GTCGTTTGTGCAACAATAAACAAAAATAAAGAAACCGACTGCATC
GCAAAAAAAAAAAAAAAAAAAAA

FemE6-AR:

CAAAATGCAGTGTAATGGAATGATGCTGCATGAGCGCACGATTAG
AGTGGATTATTCGGTGACCGAAAGACCGCATACGCCCACGCCCGG
TGTCTACATGGGAGCTAGAAGCACTGAGAAACGGAAGCACCGCAG
TTCCTATAGCTACAGGGGACGGAGCTATGATGACGATTACCATCA
TCGGCGGTCAAGACGCAGCAGATCTCGTTCCTGTCATCATCACCG
TAGATCTAGTCATCGCCATCACCATCGACGTGACCGTGCTCGTGA
TCGTTCTCCATCTTATTCCTCAGTTGACTCACGTCGATCCTATCGA
TAATGTGCTAGAAGGATTGTGTTTTTGACGTGTTTGGATTTCTTC
ACATTTTAGAGAACTAGTTTTCATTATTCCGCGATTCAAACTAAA
CATTTTATTTATTTATTAGTTAAAGATCGTGTGGTTTATGAAATG
TCGTTTGTGCAACAATAAACAAAAATAAAGAAACCGACTGCATCG
CAATTGATTCAGTCATTCAAAGAGCAAAAAAAAAAAAAAAAAAA
AAAAAAAA

A-5.2 Sequence upstream of *tra-2* exon 1 to the predicted promoter site.

Promoter region indicated in underline, TFIID in italics, and the start of exon 1 indicated in bold. *Ae. aegypti* supercontig 1.113 of Aedes aegypti [DNA] 1221483-1230441

TCGTCAGTTTTCTTCCGCCTGCCGCGTTTTGTTGCTACAAGGTTTT
TTAAAAGTGTCTCATAGTATGGCTTGTTCGTAACGTAGGGGAGGGT
TGATCGGGCTCAAATGTGGAAGTGGAGCTTCTAAAAGGGAGTCTTG
CATGCGAAGTGGAATTGATGATAGTTTTCATCTGCGGGATTGG *TTT*
*ATAA*AAAACCGTGCCCAAAGCACTCGGCCGATACTCGAAAACCAT
AATTTCCTTCAGTTTAAGAACCTCTCCAACACTATTTGTCTTTGGA
CGTCCATATCTGTTTTCCAGCATCTTGAGGTTAACGAAATCGTTC
AAACTCCTCTCGATCACTTTGAACGGCACACGTCTTCGAATTTGG
GAGATAAACACAGCCCAATCACGTGGGGTTTCAATATCGATATGC
GTATTTTTTTTTGATCTTTCAATGGCAGCATGGACCGAATCAACT
TCCATATGGGAGTGTCCTGATACCATAAGATTATGTGTTATGACT
ACCTGGTAGTTTTTTGCCTTTAACCATTCTCGCAGGTAAATAAAG
GTCATAACGATGACATAATTTAGGTTTTGTCCGCTGCATCTATCT
GAATACAGGTTCAATTGTTGGATAGGAACTCCGTCTGAATGCGCC
TTTTCAACGAGGTTTTGCACATAGTGCAATACGCATGAGCCAATT
TCTTGTGCTCCTCTACGAGCCTTTGATTCATCCCACAGATAACAAT

GGGCTTCGTTATGTCCTCTGTGGGTTTGAAAAATTGTGAAGTTCA
ACGTATAGAGCTGCCTACTGTAGAATGCCTGACCGCATGTTAAAT
GCGGTGTGGGAAGGCACTTTTGCAAGTCAAATGAAACGGTGCACA
CACCTTCATTTCTCTTCGCCGACACGACATCCCTTTTCTTACATTA
ATACACCTTTTCTGCCCTCGCTTGGTGTGCTTCTTGCGCTTTTTTT
AAAATTTCCAGCTCTGCATTGTCTTTCGCCTGTTTTACTTTTACAT
TAAGCTTGTCACACTCCCCACAAGTATCCGCTTTAGGTTTGCGGA
ATCCAATGTTGTACGATCGAAATATATTCCGATATGTATTATAAT
GTACCGGCGAGAAGAGCTTGTCTTTGCATTGAATGCAGTATAAAT
CATACATTTTTGAAACATTCAAATCAGCGCTTAAATACATTTTCG
ACGATTTTTTACGGGCGTAATGAGATGTATACGCTGGGAATGATT
CAATATGACGTTCTATAAACTTTTTTTCCTCAGCGCTTATCGTCTG
CCTTCTAGTGTTGTCTTTTCGAGCGTCTTCGGACACGATTCCAAG
TCCCTTGATCTTCTTGCGCACTAAGCAACGTGCCTTTTTCTCCTTG
ATGTCAAACGTGGACATGAACATTGTTTTGCACACCCGCACCATT
CCACCATTACCAGGTAGGTAATACTTATAGGCGTAAACTCGTTTC
GAATTGACTACCTAAAACAATGAATATGAAGTTGTTATAGAAAAC
AGTAAAACATATATATTGTTTAGTCTTTACCCTGCTTGATGCCGT
GTATGAAATAGTGATATGATTTGCCAAAAATTGGTTTTGAGCCGA
TGTTGAGAGTAGCGAAAAATTCCTCAACATGGTCTTTCGCATCTC
TATGTTGTACTTAGTATCACATTTCAACTTGACACAATAGCAAGG
ATCCCGACAGCTTCTTGGCGGTGCTTTAGTTCCATTAGTACGCGT
ATACGGCATTCCTCTTTGTTTATAAAATTTGTTTAGGAGCTTTGG
GTCTGCATTTTTCAAAACCTTATTGTTCTGTAGCGCAACATCATC
ATAGTCGTCATCTGAAATAGATATTATTATTGAAAGTCAGCTAGA
ATCTGTGTTGTAACATCTTATCAGAGACAAAAAGGCGTAACACTA
AGAAAAGCAGCGAACAAATTGGCGGACGGTGGGGTGTCGATGAC
TATTTCAATATTTTTGGAATCGCTATTGCTAAATAACTGGTAATT
CGTGTCAAAAACCACTAATGATAAGAATTCTCTTTGCCGATGACA
TGTGTCTCAAGCACGTGTCCCAAAGGTTCTGGTAATGTTTGGCGC
TAACAAGCACACAAAAAGTTTTTTGATTGATTTGGATATTTGATA
TTAGGTGGGTAAAAAATATGGTTCATATTACTCTCAGCGTACTCC
ATATAACTCATGACTAATTTTCGAATCACTCAACTACCGCTGCAC
AGACTTTTGACTAATTCTTGACCAATCATCAACTCTCTGTAAAGT
TTCGACTAACTCTTTATATAACATAATGGAGAGTTAATCGAAAAA
TAAAGATTTTTCTTGGCGTTACATTAGATTTGCTCTCAATTACAT
CATGACCTATTATTGATTTATTTTTGATTAACTCTCTATTATTCTT
TTTTTTTCTAGAATTTCTGTCCTTTAGACACATTTTTAAATCACTC
TTCATTTTCTTTAACTGAATCTCGGTTCTCTTTACTTACACTTTGG
TCTCAGATGTAATTAGGCTTTCACTAATGCAATTCTCAAGTAATC
CTTATTTAATTTCGATGAAAATTTTGCTCTCGAAAAATTAAACTC
ACTTTCGGTACAATATATACGTATATTATTCTTTAATTTCTATTAA
TTTACTTTAGAATTATTTTCGACTTACTGCATTTTTCCATTTTAAC
CAAAGACGAATTAAAGACTTCCATGTTCCATGTAATTGCCCAACA
TTTTATGGCACATGAGGTTATAGAGCAAAATCTAGAATACCTTGA

AAAGTATACTGCGATCTGTAAAACTTGGGTACCTTTTGCAGTACC
AAATGCAGTACTAGAAGTGGTACCCAATTTGAACCCGAGTGGGAC
GGACCTGTTTCAACTCTAGTTAAATTTTGCGCACTATGCAATATT
TGATTTATATTTACTCCGTTTAGCTTTTTCAAACGATATGCGAAA
TGTTGGCAAAGCTTTTAGATTAAAACATTAAATATATTTATAAAA
ATAACAATAAATTTTTATTTTTTTTATATAATAAAATCATAATTTA
TTTTCTATTTACTCTCGAGTACACATTCAAGTTCAGTAAAAATTA
ATGGGTATTTGGACTACCGAAAAAGTCAGTGAACTAAAAAAATGT
AATAAATCGAAAAAACAGATTTTTCTAAGAAATTCCTATCAGAAT
TATTTTCTGACTATTCACCCAATATGTGAGACGAGTTGTACAAAA
TTTCAGCTTCATATAAGCATTCTTGAATTCTCAGTAATTTATTGA
AATTTAGGGGCCTTCTCATACTGACGAGAAATTTACATTTGGTAT
TCCATTTACTCAGTGCCTACTTTTGTCGAATGTTACAAATATGCA
GTTATGGGCAGTTTAGTAATCTAAGTTATTTCTCAAAACATTTTT
TTTCGCTCTGTTTAATTTTTGTTTTAGTGGTGAACAACTCCTTACC
ATAATTTGAAATTTCTAATTTTTACACTTTCTTTGGAAAGTCACAT
TTCCAAAATGGAGATAATTGTTTCAAAATCTGTTTGCTGTTTTTT
TAATTTGATTTTTTTAATTTAATTTGATTTAATTTGATTGGGGGA
GAGGGGAGTAGGCTCAAGCGTTACGGCTCATACAAAAAAATGAAA
TTTTTCATACAAAAAGCGTTACGGAGGGGGTCAAAAATTTCCAAT
TTTAGCGTTACGTAATAAATGGACGCTGTCTTAATGATTATCCAA
AGAATGATCGCGCAGAAGACAACGATTTGAAGTAACTTCGACAAT
CGACGATGATCCTTACTGCATTTTAAGATCGACTGCGTGAGATTT
GTTAGCGCCTCACATAATATTTGAAATAATCGTTAAGACAGTGGC
CAAAAACGAGAGTGTGATGCATGTTTGCTTTGAGTATAAAACATA
CCAGAGGAAGAAATATCCATGTCATCACTAGCTTCATTTGATTTT
CGCTTTCGATTTTCAGGCAATCCTTCCGATAATGCACCTAAAATG
AAAAAAATGGAAATGCATGACTCAAATATTCCTTTTTTAATTAAT
GTGTACCGTAAAGTGCCCTAATTCAGCGCATTGTCTAAATTCGCG
TAGTTGGTACACGAATTTTATGAAATTTTCTACTATTTGTCCATA
TTGTTATCTGTATATAATGCTAATACATCACTATTGCACAATTGA
TTATTGTTATCAATCATTTACAATTTCTTAGAAATTTTCAAAAATT
ACAGAACTTTACGGTAGCCGTATATGCGTATATAATTTGCCCACT
TCCATTTGATTTTTCTGTAAATGGTCAAATTATGTACGCGTTATG
GTCTGTCCAGAACACTTTGATGACCATAGTATTCAAATGCTTCTT
ATTGCAAAGATTTTAAAAAATGGTCTAAACCACCGTGCCAAAGTG
AATCATGGAAGTGTCCAAGAACAGTCTGGTCTCCTAAAAGACGCT
GCCACCCACTTTCCTCCCACCGCAAAATCCCAGCTGTTTTTAATCC
AGTGAAGCTGATTTTTTGTTTGTGACTGAGGCATTGTATAAGCTA
ATTACCTAGCAAATTACACTAGTATTAATTGAATGACAGCTAAGA
AATCGCGGTGGACTTAAAGTGGGTGGCAATGTCTAATAAGAGACC
TGACTGTAGTTTTTTTTACCCTACTTACCTTGAGAATCCAAATTTT
CTTTATTTGAGTCCACATCGCAGCAACCAACGACCGAGGCATTCA
AGTAGGGCTCAAAATATTCGCGGATGGAAAAATTATTCAACAATG
TCACGTTTCCAATCTCAATGTCCTCGTCATCATCGGAATCATCATC

ACAATCTTCTGTATGTTCCTGTGAACTTGAGGGACCAACAAAAGG
CTGCGATGTGGAGGGTCCAGCAAACGGGCACGAGGTTGAAGGTCC
AGCGCACGCTTGAAAATTTGAGAGACCAGCAATGGGCTGCAAAGT
TGAGGGTCCTGGCTGGCTGTGAACGGTTGTCATTTGAGAGGGATC
CGTACAGCCGTCATCTTCCTCTGAAGGATTATATGTCGGACAAGC
GAGTTGACTTCTCATAGGCGATTTTTTTATTTCAAATTGATGCAC
AAGCTGACTTTCTTCGACAAGGATTTCACTAGTAGTTAAATATAT
GCTTTGATCGTTGGGTGTCGTCAGTAGTCGACTGACCCTGGCTGC
AATCTCGGTGTGCTCACCCTTCATCCCTTCCATTTTGCTAGGAGC
TAGCACTTGGAAGTTTTGCTCGGTACGTAAAGCAGCAATGGTCGC
CGAAGTTGACAAAAACTCGGGTTTGTTGTTTCCACATTCGGCTTC
CGGAATCGGAACATTACCATCATCTGCTTCATCTGAATCCTTTTT
TAAATTTAACGGACAGCAACTGTATAAAGAAGCCGCAGCAACGGA
TGGAGGCGCTTTTGCCAATGCCAAGTTGAGAGGTCGATATTCATC
CTCCAAATCGAAATCTCTGATCGATAAATTAATTGGCGAGTCTTC
CTGCATGCTCCTTGATAGATCCATGGGAATTGAACAACTTGGATC
CTCTACGGCCTTCATCGGAGTAGATGCTCTAACAGCCGAGAATGA
GAGCTCCAAACTATGTGATGGCATCGTACTGCTATGGGCGGCCAC
CTCTTCAAATCTTGCCACCGACACATTGCGACCATCGCTTTCGTCC
TGGTCAACCGATTCAGCAGCAGCTGAAATTTAAGAATAATGAATT
GTATTTTGTATTGAATAGGTTCTATATGACGGAGGACTTCTTTCG
CTTACCTAACAGCATTTCAATTCTGCTATTGAAATCCATTCGAAAT
AGCAGTTCTCAGATTTTTCTGATCCGTCGATAAATTTCACTGCTT
TACAGGAACCTGCTCTTATCTTCGACGAGTTATCAGCTGGTGGAG
ATTTTATTTTTTGGCACAAACAACGAGAGCAAAATAATACTTCTT
TACATTGTGTTGATCTTGAAATGTTTGTTTTGAGAACAATTTTCT
GGAAGGCCTCAACACCGAAATAAAAAATTTTGACATGAGGCTTTT
TCAATCAAAGGCCTCATCACAAACAGAAAAAAATGCTGTTGAGGC
CTTTTAAGAAATTGATGTTTATTAGTATAAATGCGATGTTTTCCA
TAAATTTTCAACTCAGACAACTAAAACTAATAGCAATCTATTGAA
CTATAGAAAAAAAATATTTGTTTACATTTTGGACTTAGGGCCTTT
TCAATTGTTGGTCTTGAAATATTTCATTTGAATTGACGCCAATAA
ATTTCTTCTCATCGTTCATTATTTTTATTTTTATTTTTAACAGAAC
GTAATGAATTAATCAAAATATTCGGATATTGATTTTTATTTATTG
ATCACAGGGTGGGTGTACGGCTGTCAACTACAAACAAAGCTCGCC
TAACAATCATCTCTCGGGCGGGCCGTCCCAAACAGCATCATCTCT
CACACGGGGCGACCCAAACCAGGGGGGGTATACGCAACGAGATGC
GCCTACCGTTGATGTTTTGTGGGTGTATTCGTTTGGCTCACAACG
CTGCTAGGCAGGGTGGGTGTACGGCTGTCAACTACAAACAAAGCT
CGCCTAACAATCATCTCTCGGGCGGGCCGTCCCAAACAGCATCAT
CTCTCACACGGGGCGACCCAAACCAGGGGGGGTATACGCAACGAG
ATGCGCCTACCGTTGATGTTTTGTGGGTGTATTCGTTTGGCTCAC
AACGCTGCTAGGCAACCCGCTGTTTGAGTGTTGAAATGCATCAGC
ATTTGCTGCCTGGCATGGCCCGCGTTCGACCTGTGCTGTTTGGGG
TCGGCCCGCGTGAGAGATGATGCTGTTTAAGACGGCCCGCCCGAG

AGATGATTGTTAGGCGAGCTTTGTTTGTAGTTGACAGCCGTACAC
CCACCCTGACCCAAACAGCACAGGTCGAACGCGGGCCATGCCAGA
CAGCAAATGCTGATGCATTTCAACACTCAAACAGCGGGATGCCTA
GCAGCGTTGTGAGCCAAACGAATACACCCACAAAACATGTACGGT
AGGCAAACCTGGTTGCGTATACCCCCCCTGGCGCCGAGTTGTGCG
CCAGCCAAAAAGTAAATAAAGAAATGTCAAAATATCTCGACGAGG
AAGTGAATCACTTTTAGGGCTCGAGTCCTAAACTGGAAAGCGATC
ATATTGAGCTTCTCGAGAATTTTTAGCATTTCGATCACTTCCTCTT
TACTTTGAATGTTACGCGCCATGGTCTGGTACTAGTTGGGCATTT
TGCGAATCTTTTCTCGAAACCGCAGGCCACCAAATAAGTTGCGCG
CCGATCGCCGAGTTGTGGATCGGTCAAAACGATTTTAATCGCTTG
CAACTTGAACACAAGAAGGTACGAACTAATTTATTCCATATGTTT
TGTATTGATAAAACAATTATGATCGCTTTCCAGTTTAGGACTCGA
GCCCTAAAAGTGATTCACTTCCTAGGCGAGATATTTTGACATTTC
TTTATTTACTTTTTGGCTGGCGCACAACTCGGCGCATGGGCCACC
GGCGCGCAACTTGTTGGACAGTCTGCGGATTTAAATAAAGATTCG
CAATAATTGTTTTATCAATACAAAACATATGGAATAAATTAGTTC
GTACTAGAAATCCTTTATAAGAGAGATTTGCGGAAGCTGACATTT
CTGTCAAAGTGTAAGCCAATCGAGCAGCGAGAGCTGTCAAAGTGT
GAGCCAAACGAACATGCGTTATGTTTGTTTTGAACTTTTCTTGAG
GAGTAATGTTATCCACTTGAAATTAATTCGGTAAAAGTAATTTTG
CATGAAGCAAAATAAATGAAATATTTTTCTTTTTTTAATTTCTTTC
AATGCGATTTTTAGTTGTTGATTTTTTGGGCTGTTTATTAGTTTG
AATGTGTAGCTCAAAGATCTATAACGAAACAAAATAATATTTTTA
GCGATGATGAAGTCATGTCCGTGTTACAATATTGTTCTCGACGCC
GAGCAAAATCAGCACTGCTGGTCTGTTGCCAAATATTTCCATTTT
ACTTCCGCGTATCTCTCTATATAGGATCACTAGTTCGTACCTTCTT
GTGTTCAAGTTGCAAGCGATTGAAATCGTTTTGACCGATCCACAA
CTCGGCGATCGGCGCGCAACTTATTTGGTGGCCTGCGGTTTCGAG
AAAAGATTCGCAAAATGCCCAACTAGTACCAGACCATGGCGCGTA
ACATTCAAAGTAAAGAGGAAGTGATCGAAATGCTAAAAATTCTCG
AGAAGCTCAATAAATGGTTGGACAAGTCTATCTTAAATTTTATGA
CTCAAAGAAATGCTTGTAAGCTATACAAGTTAGGGTGCCGGTGCC
ATTAGTGGACTACCTAAGCTATAAAAAATCATAACAAAATAACAA
AAAGCCTTAGCAGAGATCTTTTGGCGTCAACAGAAAGATTTCAAC
CTCTAATATATGGGAAAAATATAAAAAGAGTGATAAAACTAATGT
TTAACATAAAATCGTTTGAGCCACTATTGGTACACGTGTTCCAGT
AGTTGCGCTAGTGCTCCAGTAGTAGACGTTCTGGAATTATACAAG
GAATTTTAAACAAAATGGCAAATTTTTACAAAGGTTTAACATTTT
TCCCTCGGAAAAGCAGCTAATATCTTTCGAATGAAGCATAAAGAT
CATCTCTAGGAATTTTCTCCATTTTATACCTCAATTTTAAAAACTG
ATTTCATCCATTGATCCACTACTGGAACACGACGCCAACTATTGG
TGCAAGGGAGCAAATTATTTGCGTAAACTTAATTATTTATATGAC
TTTTTGATGAAATAAAAAGCTGAAATTTGGCAAATCGACTAAACT
AGAGGCGATCTATCCATCAAAAATAGCACATGTCGTTTTTATCGA

AAAATTTACGTTTTGTTCCATAGTCCAATCTACTGGTACATTACA
ACCATTGGCACCAGCACCCTATGTAATGACCACGGATGTATCATC
TGCGCTGCTACGTCGTGTAGTAAATACCGTAAATACAAAGTTGTG
TCCAGAGTTTGAATCTAAACGGTACCACCCTAGCACGAAATATTT
CATCTCATCCGGCACGGCAAGCGCAAGGCAAAGAAGGCCAAGAAA
ATTTGAATTTAAAAATTTGCACATTCCAATTTTGCTTCTAGACCG
CGCAGCTGATAAAATCAAA**ATG**

A-6 Sequences for Chapter 5

A-6.1 Sequence of *tra-2* exon 1 used in RNAi

Partial sequence of *tra-2* exon 1 amplified region (position 1220942 - 1221076 in supercontig 1.113) used in RNAi experiments.

ATGCCAAGACGAAGGTTTCCCGAGGGTTCGGATTCGTGTACTTCC
AGGAGCAGAGTGCGGCCACCGAAGCCAAAATGCAGTGTAATGGAA
TGATGCTGCATGAGCGCACGATTAGAGTGGATTATTCGGTGACCG

A-6.2 Sequence of *tra-2* exon 2 used in RNAi

Partial sequence of *tra-2* exon 2 amplified region (position 1194346-1194646 in supercontig 1.113) used in RNAi experiments.

TCCTCAGTTGACTCACGTCGATCCTATCGATAATGTGCTAGAAGG
ATTGTGTTTTTGACGTAAGTCATCTTTTTCAAGACCTCACGAAGA
CCCTACAAACAAAGAAGTTCAAGTCAAGTATTATTGAAGATAGAA
TCTGTATCAGTAAAAACTATTCTTCAAATGACACAAGGGAAGAAC
AGATTACTCTTCATTAATCCAAGGGTAAATTGTATGTAGCTAACC
GTTCTGTTTTTCTTTTCGTTTCAGGATGTTTGGATTTCTTCACATT
TTAGAGAACTAGTTTTCATTATTCCGCGATTCAAACTAAACA

A-6.3 Sequence of DSRed2 marker used in RNAi

CTTGGCCATGTAGATGGACTTGAACTCCACCAGGTAGTGGCCGCC
GTCCTTCAGCTTCAGGGCCTTGTGGGTCTCGCCCTTCAGCACGCC
GTCGCGGGGGTACAGGCGCTCGGTGGAGGCCTCCCAGCCCATGGT
CTTCTTCTGCATCACGGGGCCGTCGGAGGGGAAGTTCACGCCGAT
GAACTTCACCTTGTAGATGAAGCAGCC