IDENTIFICATION AND CHARACTERIZATION OF GERMLINE-SPECIFIC PROMOTERS FOR REMOBILIZATION OF TRANSGENES IN THE MOSQUITOES, *AEDES AEGYPTI* AND *ANOPHELES GAMBIAE*

A Dissertation

by

DARREN ERICH HAGEN

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2007

Major Subject: Genetics

IDENTIFICATION AND CHARACTERIZATION OF GERMLINE-SPECIFIC PROMOTERS FOR REMOBILIZATION OF TRANSGENES IN THE MOSQUITOES, *AEDES AEGYPTI* AND *ANOPHELES GAMBIAE*

A Dissertation

by

DARREN ERICH HAGEN

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Craig J. Coates
Committee Members,	Christine Elsik
	Thomas Ficht
	Clare Gill
Chair of Genetics Faculty,	James Wild

August 2007

Major Subject: Genetics

ABSTRACT

Identification and Characterization of Germline-specific Promoters for Remobilization of Transgenes in the Mosquitoes, *Aedes aegypti* and *Anopheles gambiae*. (August 2007)

Darren Erich Hagen, B.S., Angelo State University

Chair of Advisory Committee: Dr. Craig J. Coates

The development of genetic transformation systems in insects has revolutionized the field of entomology. Transgenic insects provide tools to identify, isolate and analyze insect genes and to genetically modify insects for the purposes of insect control or disease vector modification. When transformation frequencies are high, multiple transgenic lines can be generated with relative ease. However, in most mosquito species, the results of transformation experiments have been suboptimal. Increased mosquito transformation efficiency is a research priority. Additionally, incorporation of refractory transgenes will not be sufficient to modify natural populations. A gene drive system will be required to allow transgenes to proliferate throughout populations and potentially reach fixation. This study proposes the use of germline-specific regulatory elements to promote confined, regulated transposase expression within the germ tissue. Creation of helper constructs utilizing endogenous promoters will potentially increase genetic transformation frequencies. The generation of lab strains of mosquitoes expressing an endogenous source of transposase within the germline will also serve as a powerful research tool. Endogenous sources of transposase will allow for comparative analysis of integration rates using different donor plasmids. Finally, the generation of

autonomous transposable elements will provide a gene drive mechanism to move a tightly-linked refractory gene into a population.

Four genes have been identified, cloned, and characterized, revealing expression patterns expected of germline-specific genes. Transcription profiles and *in situ* hybridization data support these conclusions. Putative *cis*-acting regulatory elements have been cloned and incorporated into DNA plasmid constructs. These elements are cloned in a manner such that they will regulate fluorescent gene expression. Additionally, similar elements have been cloned upstream of the *Mos1* open reading frame, within the inverted terminal repeats of the *mariner* transposable element, thus creating autonomous elements and a potential gene drive mechanism.

DEDICATION

Opie and Omie

I dedicate this publication of my doctoral studies to you. I watched for many years while you worked so hard to provide a better life for your children and grandchildren. You provided us with the work ethic, motivation, and stubbornness required to achieve all we desire. While I will never be able to appropriately thank you, all that I achieve is your achievement too.

Mom and Dad

Everything I am and the entirety of my accomplishments are direct results of all that you have given to me; your time, patience, values, and love. Your unwillingness to allow me to accept less, much less quit, has sent me on a path in life that has allowed me to achieve all my dreams. You pushed me when I was unwilling to push myself. Although you allowed me to believe I could and would do it on my own, I have always known that you were behind me, silently helping me forge ahead. While no words will ever do it justice, please know that I love you and am eternally grateful to you.

ACKNOWLEDGEMENTS

Many people have made vital contributions to my life and education during my time at Texas A&M. I am forever indebted to them for all they have provided me.

Thank you to Dr. Craig J. Coates, my committee chair. You took a huge risk on me and simply provided me with the chance when others might not have. I will never forget the opportunity you afforded me. I will carry your scientific ideology with me throughout my career. The "hands off" approach proved invaluable for my education, teaching self reliance and independent work.

I must express my appreciation to Dr. Clare Gill. You were always very supportive of me and my ideas. I thank you for the discussions we had of both science and life. Your technical expertise and willingness to share it with me gave me the boost I often needed to move ahead with my project.

Thank you to Dr. Christine Elsik. Without the guidance you provided me, my project would have never taken the turn it has. The computational training you have provided me will benefit me throughout my future endeavors.

I must mention Dr. Thomas Ficht. You gave numerous hours providing guidance and assistance with my project as well as my basic understanding of molecular biology. I thank you for all that you gave so that I could achieve the goals I had set.

Aside from my committee, there were numerous faculty and staff members who gave of themselves to get me to this endpoint. I owe a debt of gratitude to Dr. James Olson for spending countless hours teaching me about mosquitoes from an organismal

and field perspective. While never truly succeeding in his quest to turn me into an entomologist, I thank Dr. Olson for his persistence and fresh views. I thank Dr. Mulenga and Dr. Pietrantonio for the fruitful discussions and the creation of a molecular entomology based journal group. Not to be forgotten, the Entomology Dept. support staff supported my work with plenty of coffee and candy, but primarily for a place to escape to when I needed a breather.

I absolutely must thank the past and present members of the Coates Lab. Without your support and friendships, I would have never survived the rigors of graduate school. I arrived in the lab without any experience or much to offer, but you aided me every step of the way. I am most appreciative of Dr. Christine Gray. I am not sure I could ever fully express my appreciation. From the scientific discussions to the reagents and assistance, you were always willing to mentor me. Undoubtedly, I will forever be grateful for the life discussions we had. Your willingness to listen to me with an open mind and heart and offer advice, both pragmatic and encouraging, helped me grow personally to a level I never imagined. Ironically, perhaps my greatest professional growth came as a result of your leaving the lab. The mentor I had come to rely on was no longer over my shoulder. Lara Butler has become one of the closest friends I have ever been blessed with. The time you spent providing support, both mental and emotional, provided me with the motivation to persist in my career. It was always a joy to come to work and interact with you. Paul Barron's in depth theoretical discussions (and sense of humor) allowed me to think outside the box I often found myself stuck in. Ahmed Mohammed, thank you for your friendship. I am grateful for having met you

during a time when I was ignorant about other cultures. Haisheng Tian and KJM Vally gave me insight into protocols and theories that opened my mind to new possibilities as well as introductions to new cultures.

My friends have been consistently supportive throughout my time in graduate school. I am particularly grateful to Ryan Huebinger. A friend in the truest sense, Ryan was there whenever I needed help whether it was in the lab, around the house, or simply to escape from the rigors of life. GW Yandle and Charles Jones have constantly provided support over the past eight years of my education career. I must also acknowledge those who stood by me through some of the tougher times in my life. I will never be able to fully thank Raina Tucker for the emotional support she provided me. The "vet school" crowd was always there to provide me relief from myself as well. Kristina Anthony gave me a breath of fresh air when I needed it most.

Finally, I thank my family. My parents watched me struggle and fret, knowing all along I would pull through. They have proven to be right more times than not and many times have more faith in me than I have in myself. I must thank my siblings for their poking and prodding. The years of "torment" they endowed me with by calling me Dr. Darren served as motivation to achieve. My brother Garrett is one of my closest confidants. Life works in amazing ways. I never would have imagined how much I would appreciate our lives in Aggieland overlapping. Kirk and Stacy have always taken the opportunity to keep me grounded. I could not imagine a better family than the one given to me.

TABLE OF CONTENTS

Page

Page

Page

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

1. INTRODUCTION

1.1. Defining the Problem

1.1.1. Mosquito-borne pathogens

Arthropod-borne pathogens sicken and kill millions annually around the world. Among arthropods, mosquitoes are the most responsible for the spread of these diseases. Greater than forty percent of the world's population lives in regions where malaria is endemic. *Anopheles gambiae*, the primary human malaria vector, is responsible for 100-500 million cases of malaria each year. It is estimated that between 1 and 2.7 million of these cases result in death, with most of these deaths being children (Breman *et al.*, 2001). *Aedes aegypti* is the primary vector for the pathogens that cause Yellow Fever and Dengue Fever. Ten to fifty million people contract Dengue Fever each year, leading to thousands of deaths. An additional 30,000 people die annually as a result of Yellow Fever (Centers for Disease Control and Prevention website, http://www.cdc.gov). Added to the costs of human lives, mosquito-borne pathogens have a large economic impact in countries as a result of lost work days due to illness. The cost of providing healthcare, vector control, or simply infrastructure to aid in protection and prevention is often more than many governments can afford.

1.1.2. *Aedes aegypti*, the Yellow Fever mosquito

The mosquito, *Ae. aegypti*, traditionally inhabits the tropical and subtropical regions of the world. Domestic, day-biting females preferentially feed on humans (Centers for Disease Control, 2005). Breeding sites include the use of artificial

---

This dissertation follows the style and format of Gene.

containers, making the species highly cosmopolitan (Vezzani, 2007). Increased population growth and urbanization of tropical regions in conjunction with expansion of favorable climates resulting from global weather changes, has further increased the *Ae. aegpyti* range (Gubler, 2002). Female aedine mosquitoes require vertebrate bloodmeals for ovarian development and maturation prior to mating and oviposition (Eldridge, 2005). Consumption of blood from a viremic host results in the infection of the female mosquito. For transmission from vertebrate to vertebrate to succeed, the viral particles must infect the cells lining the midgut. The viral particles amplify, move from the midgut into the salivary glands, and one eventually passed to a new host during bloodfeeding. Once infected with a virus, a female mosquito will be retain a level of virermia for the duration of her life (Blair *et al.*, 2000).

Aedes aegypti is the principle vector for two medically important arboviruses, the causative agents of Dengue Fever (DF) and Yellow Fever (YF). DF is a far-reaching vector-borne flavivirus with more than 2.5 billion people living in dengue infection risk areas (Mackenzie *et al.*, 2004). With 10-50 million individual cases of DF, 500,000 cases of the more severe Dengue Haemorrhagic Fever (DHF), and casualties in the tens of thousands, dengue is perhaps the most important arboviral disease of humans (World Health Organization, 2003). Dengue virus consists of four closely related serotypes. There is no crossprotective immunity in humans for these viral serotypes. An individual living in an endemic area can have as many as four infections either concurrently or in series (Mackenzie *et al.*, 2004). No effective vaccine is available to date. For a vaccine to be effective, it must protect against all four serotypes. Several candidates are now in various stages of development (Halstead, 2002). Rather than mitigate the effects,

infection by other serotypes often exacerbates the symptoms due to antibody-mediated enhancement (Stephenson, 2005).

Over 600 million people reside in regions of risk to YF infection. Approximately 6,000 cases of YF are reported annually. The number of cases is estimated to be much higher, the discreprency a result of under-reporting. The WHO believes 30,000 deaths occur annually as a result of YF infection. Urban reinfestation by *Ae. aegypti* has resulted in an increasing number of YF outbreaks (Barrett and Higgs, 2007). Unlike dengue, an effective and affordable YF vaccine exists, however, the expense to those in developing countries often makes the vaccine unattainable to those who will benefit.

1.1.3. *Anopheles gambiae*, the malaria mosquito

Malaria is transmitted by members of the Anopheline family. Grouped into 41 genera and approximately 430 species, only 30-40 *Anopheles* species are capable of malaria transmission in nature. Transmission via Anopheline mosquitoes poses a risk to over forty percent of the global population (Centers for Disease Control, 2005). Of the 300-500 million cases of malaria occurring annually across the globe, greater than one million result in patient death. Many of these deaths are young children in sub-Saharan Africa (World Health Organization, 2003). Malaria is caused by members of the protozoan genus, *Plasmodium*. With greater than one hundred members of the genus, only four species are capable of infecting humans; *Plasmodium falciparum*, *P. malariae*, *P. ovale*, and *P. vivax*. *Plasmodium falciparum* leads to the most severe disease manifestation and is the most prevalent species in Africa (Centers for Disease Control, 2005). Most Anopheline mosquitoes do not have host preferences. However, members of the *An. gambiae* species complex are unique in their preference for human blood.

Anthropophily by *An. gambiae* makes this mosquito one of the most efficient malaria vectors worldwide. Furthermore, females previously feeding on infected blood will take more blood meals during her life than non-infected individuals (Koella *et al.*, 1998). *Plasmodium* exhibit a complex lifecycle which requires both the vertebrate host and mosquito to complete maturation and reproduction. Sporozoites are released from the salivary glands of the mosquito into the vertebrate host. Upon invasion of the liver, growth and division lead to the release of merozoites into the blood stream where they will eventually form gametocytes. Mature gametes will form in the bloodmeal taken by a female mosquito (Phillips, 2001).

1.1.4. Current control strategies

Mosquito control currently includes the use of chemical pesticides, biological control, and environmental management. These methods of control are costly and require public education and community involvement. Public education and community participation have presented real stumbling blocks to those desiring a reduction in disease incidence. Many individuals living in endemic areas are impoverished and uneducated. The lack of formal education leads to skepticism of "western knowledge" and priority to traditional beliefs. Even with the knowledge of disease transmission via arthropods, individuals and communities often feel as if there is little they can do personally without outside involvement (Opiyo *et al.*, 2007).

The cheapest and often most effective way to combat mosquito-borne disease transmission is through the use of insecticides. Because of economic factors, most insecticides are developed for use on arthropods of agricultural importance. These agrochemicals are eventually formulated for use in public health (Hemingway and

Ranson, 2005). The reliance on insecticides used also in agriculture effectively means that many of these insects may have previously been exposed to these chemicals while breeding or resting in areas of agricultural use. Reoccurring exposure of the vector insect to these insecticides eventually leads to the selection of alleles that confer tolerance and/or resistance to the chemical insecticide. Many examples of this phenomenon are described (Hemingway and Ranson, 2005, Chandre *et al.*, 1999, Ranson *et al.*, 2000). The cheapest and most effective chemical, DDT, is no longer a desirable pesticide. Due to the stability of the compound, DDT persists in the environment, accumulating in the food chain. Although not generally toxic to humans, there is some evidence that using the chemical at application rates high enough for malaria control might harm the very young (Rogan and Chen, 2005). The mosquito eradication programs using primarily DDT to fight against mosquito borne illnesses were eventually discontinued in the early 1970s. This lapse in mosquito eradication had important consequences for Dengue, Yellow Fever, and malaria. The combination of the re-infestation of the areas by mosquito vectors combined with the unplanned urbanization and increased travel and commerce played a critical role in the increase of disease incidence. Unfortunately, alternative insecticides are more costly and often unrealistic for poor nations.

Biological control involves the reduction of a target pest population by a predator, pathogen, parasite, competitor, or microbial toxin (Hemingway, 2005). Biological control focuses on host-specific agents as opposed to broad-spectrum insecticides, resulting in little or no disruption to nontarget organisms or the environment. For mosquitoes, the introduction of the mosquito fish, *Gambusia affinis*, in the early 1900's was the first successful attempt at biological control. *Gambusia* was introduced in many

countries as a means of larval control. Their use was greatly diminished in the 1940's with the advent of prolific insecticide use, particularly DDT (Hemingway, 2005). Interest in biological control of mosquitoes reemerged in the 1960's after many concerns arose about the environmental impact of chemical pesticides. The entomopathogenic nematode, *Romanomermis culicivorax*, and the protozoan, *Nosema algerae*, were studied extensively because they are capable of infecting and killing mosquito larvae (Kaya and Gaugler, 1993, Chapman, 1974, Legner, 1995). Large-scale trials with these agents had varied results due to the adverse effects of environmental factors. High production costs, problems with production and storage, and the advent of additional microbes ultimately made these organisms unacceptable for use in mosquito control (Hemingway, 2005). The most effective biological control agent to date remains the bacterium, *Bacillus thuringiensis israelensis*. This species produces a toxin highly specific for mosquito larvae. The ability to incorporate the bacterium into briquettes, granules, and other dry products has increased the persistence of the organism in the environment. These advantages make this product economically feasible and practical enough for widespread use (Becker and Ascher, 1998). The release of sterile males as competitors of wild-type males has previously been successful in the eradication of the screw worm fly from all of North America and much of Central America (Wyss, 2000). Attempted sterile insect releases of mosquitoes have been marginally successful on smaller scales. The lack of success is due to the high density and massive immigration/emigration of the target vector insects. The inability to produce enough sterile males on a daily basis to achieve sufficient ratios of sterile males to wild-type males, coupled with inefficient sexing techniques, reduces the efficiency of the sterile insect technique in mosquito vectors

(Wood, 2005). Recent advances in sexing technology by fluorescent protein expression in the gonads of developing males mosquitoes may prove vital for any future SIT success (Catteruccia *et al.*, 2005). Sexing technologies such as sex-lethal systems provide powerful tools that eliminate the need to screen individuals to determine sex (Phuc *et al.*, 2007). Biological control methods hold great potential, however current techniques and technologies are not sufficient to reduce the incidence of mosquito-borne disease.

Environmental manipulation to alter the habitat of insect vectors is another means to control the incidence of vector-borne diseases. As far back as the 1st century, the Romans loosely associated mosquitoes with disease and advocated the draining of marshes in the area as a way to reduce the risk of disease (Small, 2005). It wasn't until Ronald Ross made the connection of anopheline mosquitoes with malaria in the 1890's and Finley and Reed connected *Ae. aegypti* with Yellow Fever in the early 1900's that concrete evidence existed. Armed with the knowledge of the environmental requirements for these vectors, environmental control programs were put in place prior to and during World War II (Small, 2005). With the advent of chemical pesticides, environmental management programs have been largely neglected. All mosquitoes require water at some point in development. However, different species of mosquitoes each require unique conditions for survival and development. Each require water with specific organic content, oxygen content, salinity, temperature, turbidity, and shade density (Small, 2005). Alteration of aqueous environments might include draining or filling of ponds and marshes or impounding streams and rivers into canals. In areas of high shade, vegetation can be cut back or cleared completely to alter the light characteristics. Prolific urbanization has introduced new breeding sites, bringing the mosquito populations closer

to the homes of those living in endemic areas. Artificial breeding sites include rubbish, water cisterns, animal and human footprints, and tires (Small, 2005). Sufficient breeding sites have been found in cemeteries where loved one's place water and flower filled vases as memorials to the deceased. These new breeding areas then subject future mourners to risks, bringing the mosquito and people in close proximity (Vezzani, 2007). Perhaps the greatest environmental change will occur for populations living in endemic areas. Education of indigenous people will assist in controlling the spread of disease. Often feeling like they cannot personally do anything to aid the fight against disease, communities are armed with little or no scientific knowledge (Opiyo *et al.*, 2007). The lack of awareness and education play a vital role in the disease cycle.

1.2. Proposed Solution

1.2.1. Creation of refractory strains of mosquitoes

With the failure of eradication and control programs, the focus of vector control has shifted from population elimination and reduction to population modification and replacement. The replacement of wild-type mosquitoes with genetically engineered stains that are incapable of disease transmission would reduce the incidence of disease (Collins and Besansky, 1994, Curtis, 1994, James, 2000, Olson *et al.*, 2002). This idea was demonstrated in a proof-of-principle experiment where the Sindbis virus system was used to transiently express antisense RNA to part of the Dengue viral genome, creating the first refractory mosquito strain (Olson *et al.*, 1996). While functional, the Sindbis system is not hereditary and its effectiveness is lost in one generation. The desire to stably integrate exogenous genes into the genomes of *Aedes aegypti* and the Anophelines was realized with the creation of stable germline transformants (Coates *et al.*, 1998,

Jasinskiene *et al.*, 1998, Catteruccia *et al.*, 2000, Grossman *et al.*, 2001). A strain of refractory mosquitoes with stable and heritable characteristics was produced using a transposable element vectored synthetic gene, reducing the transmission of *Plamodium* parasites (Ito *et al.*, 2002). Harnessing the power of RNA interference (RNAi), a strain of *Ae. aegypti* was recently generated that targets transcripts specific to DEN-2 virus replication. The genetically manipulated *Ae. aegypti* mosquitoes expressed inverted-repeat sequences derived from DENV-2 genomic RNA under the control of the *CarboxypeptidaseA* promoter. Expression of viral-specific dsRNA was localized to the midgut of female mosquitoes. When challenged with the DEN-2 serotype, the transgenic mosquitoes poorly supported viral replication (Franz *et al.*, 2006). Similarly, heritable gene silencing via RNAi was demonstrated in *An. stephensi* (Brown *et al.*, 2003). Experiments have shown the plausibility of competence reduction as a result of RNAi techniques(Osta *et al.*, 2004, Arrighi *et al.*, 2005, Michel *et al.*, 2005, Franz *et al.*, 2006). Stable, heritable integration of transgenes harnessing the power of RNAi has been successfully revealed. This technology exhibits promise in the creation of refractory mosquitoes with heritable traits reducing but not completely eliminating pathogen transmission. It is difficult to determine the efficacy of RNAi techniques under real world conditions. In typical lab experiments, viral titers found in bloodmeal sources are generally much higher than viral titers found in natural hosts. In the absence of complete pathogen elimination the door remains open for selective pressure to restore competency. A multi-faceted approach, combining multiple gene targets of both the vector and pathogen directed at different stages of pathogen infection and mosquito response may be the best option for eliminating pathogen transmission (Nirmala and James, 2003).

Beyond the vector-pathogen interaction, host-seeking behaviors could also be used as potential targets (Zwiebel and Takken, 2004).

1.2.2. Proliferation of refractory genes into natural mosquito populations

Upon identification and incorporation of genes or alleles that confer a refractory phenotype, these alleles must then be introduced in an attempt to replace the natural population. Standard Mendelian inheritance of antipathogen genes is expected to take too long to lead to allele fixation (Collins and James, 1996, O'Brochta and Atkinson, 1998). There would be continual breeding of heterozygous individuals that would produce true-breeding offspring capable of unimpeded pathogen transmission. This would prevent complete interruption of the pathogen transmission cycle. Driving genes to fixation in natural populations by linking antipathogen genes to meiotic drive loci, employing the use of transposable elements (TEs), or using symbionts or other infectious agents may hold the greatest potential (James, 2000, Sinkins and Gould, 2006).

Transduction uses viral vectors to package and deliver genes of interest through their natural route of infection. Densonucleosis viruses have the potential to serve as transducing agents, introducing and spreading genes through a population. Densoviruses primarily infect insects and certain viruses exhibit a narrow host range (Carlson *et al.*, 2000). A narrow host range makes the viruses less likely to infect non-target species. Stability in the environment means that viral particles can persist in larval habitats for extended periods of time, allowing viral accumulation. In addition, vertical transmission via transovarial transmission makes the use of densoviruses attractive due to the fact that a single female could potentially infect multiple sites during a gonotrophic cycle. This could lead to viral introduction in areas that are difficult to find and treat (Carlson *et al.*,

2000). Infectious clones of the *Ae. aegypti* densovirus (*Ae*DNV) have been engineered to carry and express genes of interest in the mosquito. Furthermore, vertical transmission resulted in the infection of developing embryos and larvae (Olson *et al.*, 2005, Carlson *et al.*, 2000).

Paratransgenesis using an endosymbiont is another means of gene introduction into a population. *Wolbachia*, an obligately intracellular bacterium, is found in many arthropods. It rapidly spreads itself through populations by means of cytoplasmic incompatibility (CI). CI is an induced crossing sterility within or between populations. Unidirectional CI occurs in crosses between *Wolbachia*-infected and uninfected individuals. *Wolbachia*-infected males release modified sperm that are unable to complete fertilization in eggs produced by wild-type females. Infected females are fully able to successfully mate with both infected and uninfected males (Sinkins and O'Neill, 2000). CI thus provides a reproductive advantage to *Wolbachia*-infected females, resulting in an increased percentage of infected individuals in succeeding generations (Sinkins and Gould, 2006). Cage studies with the mosquito, *Culex quiquefasciatus*, the urban filariasis vector, established the proof of principle that cytoplasmic replacement is possible (Curtis, 1976). Population invasion has been observed in *Drosophila simulans* in nature. Infection frequencies increased from nearly 30% to greater than 80% of the fly population in approximately 15 generations (Turelli and Hoffmann, 1991). *Wolbachia* infections in a diverse number of tissues have been observed in numerous *Aedes* and *Culex* species. Recent studies demonstrate that *Wolbachia* is capable of infecting immunocompetent *An. gambiae* cells *in vitro* (Rasgon *et al.*, 2006). This data suggests that there is no intrinsic genetic block to *Wolbachia* infection in Anophelines,

showing that whole tissue infection in *An. gambiae* is likely possible. Many *Wolbachia* infections result in a wide tissue distribution within the insect host (Dobson *et al.*, 2002). These mechanisms could be used to express antipathogen molecules in mosquitoes (Sinkins and O'Neill, 2000). Superinfected insects with two or more strains of *Wolbachia* commonly occur as well, resulting in bidirectional CI. A male with two strains will not produce offspring when mated with a female infected with only one of the strains, however, the reciprocal cross will produce offspring (Sinkins, 2004). Superinfections could be utilized to drive multiple transgenes into a population. The use of *Wolbachia* as an effective genetic drive mechanism is attractive. The introduction and proliferation of transformed *Wolbachia* can alter a population by resulting in the unanimous infection of mosquitoes. Further study of *Wolbachia* would likely result in the isolation of the genes responsible for inducing CI in insects. These genes, linked with an antipathogenic gene, could be stably integrated into mosquitoes and driven to fixation in a population (Sinkins, 2004).

New alleles generally spread through a population because their presence confers an advantage. Underdominance, a form of selection that reduces the fitness of heterozygote genotypes relative to homozygous genotypes, has been proposed as a means of driving transgenes into a population (Curtis, 1968). Originally based on the fact that meiosis events in which one of the chromosomes suffered a translocation usually produce semi-sterile individuals, the focus has shifted to genetically engineering an underdominance system (Davis *et al.*, 2001, Magori and Gould, 2006). Engineered underdominance is based upon the integration of two constructs, each with four components. Each construct would have a *trans*-acting suppressor, *cis*-acting promoter, a

toxin-coding gene and an anti-pathogen gene with its own promoter. When an insect carries at least one copy of each construct it will survive. If only one construct is present in a genome, the trans-acting suppressor will not be present to inactivate transcription of the toxin gene (Sinkins and Gould, 2006). Engineering a system capable of underdominance that does not have major fitness costs is not a trivial task and some leakage in suppression of the lethal gene seems likely. The number of individuals released that carry the construct will have to exceed the frequencies previously discussed (Magori and Gould, 2006). The use of an underdominant system for large-scale replacement of disease vectors are not feasible because the levels of release needed would not be economically viable (Sinkins and Gould, 2006).

Meiotic drive occurs when a particular heterozygous locus segregates at a frequency greater than the expected 0.5 Mendelian ratio, as a result of the destruction or disabling of the homologous chromosome. A driver locus targets a responder locus on the homologous chromosome but is itself closely linked to an insensitive allele of the responder so that it is not suicidal. The resulting meiosis products form a non-functional gamete for the homologous chromosome with the sensitive responder locus (Lyttle, 1991). The best characterized example of meiotic drive is *Segregation Distorter* (SD) in *D. melanogaster*, where males heterozygous for the SD chromosome and the wild-type SD+ chromosome typically produce 95-99% SD-carrying sperm (Hartl *et al.*, 1967). The components of the complex include three highly conserved, closely-linked loci on chromosome 2; SD, E(SD) (*Enhancer of Segregation Distortion*) and the highly insensitive Rsp (*Responder*) (Kusano *et al.*, 2003). A meiotic drive mechanism has been previously described in *Ae. aegypti* (Hickey and Craig, 1966). A sex-linked gene coupled

with the male-determining allele at the sex determination locus drives the normally equal sex ratio in favor of males. Sex is determined by a single autosomal gene on chromosome 1, with the dominant allele producing males (Gilchrist and Haldane, 1947). A meiotic drive locus (M^D) product acts in *trans* with a locus believed to be tightly linked to the *m* allele at the sex locus, causing breakage of the chromosome carrying the m^s (*m* sensitive) allele (Newton *et al.*, 1976). Unfortunately, strains carrying the endogenous meiotic driver have not been maintained. Recently a new strain (T37) was established from field collections in Trinidad that carries a strong meiotic driver locus, most likely the M^D gene (Mori *et al.*, 2004). Meiotic drive presents a potentially strong mechanism for driving transgenes into native mosquito populations. Furthermore, much study is still required, however, to better understand the interactions between all loci, both known and unknown.

Homing endonuclease genes (HEG) encode endonucleases that recognize a specific sequence that flanks the HEG only when the sequence is not already interrupted by the HEG itself. In heterozygous individuals the HEG- chromosome is cleaved, inducing a chromosomal break. The organism reverts to the recombinatorial repair system, using the HEG^+ chromosome as a template converting HEG- chromosomes into HEG^+ chromosomes in the process (Burt and Koufopanou, 2004). Invading HEG- chromosomes increases the frequency of HEGs in successive generations above the expected Mendelian frequency of 50%. Primarily found in bacteria and phages, HEGs have also been found in eukaryotic organisms such as fungi, plants, and recently the cnidarians (Jacquier and Dujon, 1985, Wessler, 2005, Goddard *et al.*, 2006). HEGs have not been described in insects. Current research investigating whether HEGs function in

D. melanogaster is ongoing. Artificially created HEGs designed to affect specific target sequences may also be an option to drive transgenes into native populations (Burt, 2003, Bibikova *et al.*, 2001). If targeting a highly conserved area of the mosquito genome, HEGs designed to carry an anti-pathogen gene should be able to invade a population when introduced at low frequencies (Sinkins and Gould, 2006).

Introducing a transgenic strain of mosquitoes with an inherent fitness advantage is another means of driving a transgene throughout a population. The tight linkage of a transgene with a male-specific insecticide resistance gene would provide a fitness advantage to male mosquitoes significant enough to offset any genetic load incurred by female carriers (Sinkins and Hastings, 2004). The locus would be propagated in both sexes, but transcribed using a male-specific promoter. This strategy will require the release of only a small number of transgenic individuals. The capability for male only releases is desired as males mosquitoes are non-biting, unlikely to produce any effect on overall population size, and is more acceptable to members of the communities where releases occur (Sinkins and Hastings, 2004). Selection pressure in the presence of insecticide enables the control of spreading and use of a second class of insecticides will facilitate the loss of the transgenic strain if the need arises.

Finally, perhaps the best choice of mechanisms to drive transgenes into populations is the use of transposable elements (TEs). First described in the 1950s in maize(McClintock, 1950), TEs are mobile elements capable of rapidly spreading through entire populations. The relative speed at which TEs move through populations was witnessed when the *P*-element spread throughout the world's natural populations of *D. melanogaster* in approximately 30 years (Kidwell *et al.*, 1983, Anxolabehere *et al.*,

1988). Wild populations caught prior to 1950 and maintained as laboratory stocks contained no *P*-elements, while all populations sampled since 1978 carry multiple copies of *P*-elements arising as a result of a horizontal transfer from *D. willistoni* (Anxolabehere *et al.*, 1988, Ribeiro and Kidwell, 1993). Hybrid dysgenesis, a reduction in reproductive fitness upon initial *P*-element introduction in strains with an M cytotype, is due to an increase in genomic instability in the germline (Engels, 1989, Rio, 1990). In P cytotype strains, however, *P*-elements have replicated to a much higher number of copies. Class II transposable elements, like *P*-elements, invade non-homologous chromosomes by targeting short, frequently occurring sites. While there is not an exact target site, those with close matches to the octamer GGCCAGAC are most likely to receive *P*-element insertions (O'Hare *et al.*, 1992, O'Hare and Rubin, 1983). Copy numbers ultimately stabilize as a balance must be reached between selection for higher copy number and selection against constant genome instability caused by TE replication (Kidwell and Lisch, 2000). TEs provide a very powerful tool as a gene drive mechanism, however continued study must lead to greater understanding of TE regulation to prevent deleterious events occurring in the host genome.

1.3. The Current State of Mosquito Transgensis

1.3.1. The *P*-element paradigm

Germline transformation of *Drosophila melanogaster* in the early 1980s (Rubin and Spradling, 1982) has facilitated the discovery and characterization of many genes, the study of regulatory mechanisms in the genome, and interactions at the DNA, RNA, and protein level (Coates, 2005). The *P*-element transposon has been used to successfully integrate transgenes into the *Drosophila* genome and has been modified for transposon

tagging, enhancer trapping, and targeted transposition (Lukacsovich and Yamamoto, 2001). While a highly efficient means of incorporating transgenes into the *D. melanogaster* genome, the *P*-element seemingly lacks activity in non-drosophilid species (Handler *et al.*, 1993). Host-specific factors that both mediate and regulate *P*-element activity are primarily confined to *D. melanogaster* and closely-related sibling species. Additionally, the structure of the cleaved termini generated during the transposition process are atypical of other class II TEs (Atkinson and James, 2002). While *P*-elements are lacking appropriate activity for transgenesis in non-drosophilid species, the basic binary system of transformation utilized in *P*-element transformation remains the paradigm for subsequently identified TEs.

The lack of *P*-element activity in other species drove the search for other class II TEs that could be harnessed for transgene integration in insects of medical and agricultural importance. Class II TEs transpose by a "cut and paste" mechanism utilizing a DNA intermediate (Finnegan, 1985). Generating transgenic insects involves microinjecting a mixture of two plasmids, the donor and helper, into the posterior end of preblastoderm embryos where the germline cells are initially formed. The donor plasmid consists of the inverted terminal repeats (ITRs) of the transposon and a marker gene (for identifying transgenics), its promoter, and any nucleotide sequences of interest. The helper plasmid carries the transposase open reading frame (ORF) driven by either an inducible or constitutive promoter, allowing for transposase expression in the cells of the embryo. Translated transposase protein acts *in trans* to cut and move the ITRs and internal sequences from the donor plasmid and integrate them into target sites within the genome. Assuming integration into the presumptive germ cells, the subsequent progeny

will inherit the transgene in a disproportionate manner, as only a subset of the germ cells are transformed. Integration events that occur in somatic cells will only affect the current generation. The helper plasmid does not contain ITRs and therefore cannot integrate into the genome. The effect of transposase is lost after degradation of the helper plasmid or dilution as a result of cell division. This will lead to stable transgene integration into the genome of the host organism, assuming no endogenous transposase is capable of remobilization.

1.3.2. Mosquito transgenesis

As a result of numerous years spent pursuing *P*-element transposition in non-drosophilids, research tools were developed to quickly and efficiently assess transposon function in other hosts (Atkinson and O'Brochta D, 2000). Excision assays measure the ability of the transposable element to excise from a donor plasmid molecule. First used to study excision in bacteria (Berg *et al.*, 1983), their use was extended into *D. melanogaster* (Rio *et al.*, 1986) and non-drosophilid insects (O'Brochta D and Handler, 1988) to assess the excision activity of the *P*-element. Interplasmid transposition assays measure the ability of the transposase to mediate the movement of the transposable element located on a donor plasmid to a target plasmid (Sarker *et al.*, 1997, Atkinson and O'Brochta D, 2000). Testing the mobility of TEs in the target species *in vivo* reduces the initial investment and effort required to generate transgenic lines by eliminating TEs that will not transpose within the host.

Identification of new class II TEs brought hope of cross-species functionality. TEs discovered in a range of insect species; *Hermes* element from *Musca domestica* (Atkinson *et al.*, 1993, Warren *et al.*, 1994), *mariner* from *Drosophila mauritiana*

(Medhora *et al.*, 1991), *piggybac* from *Trichoplosia ni* (Fraser *et al.*, 1995), and *Minos* from *Drosophila hydei* (Franz and Savakis, 1991), have led to successful transformation of non-drosophilid insects. Each of these class II TEs have successfully been used to generate stable germline transformants in a number of medically important mosquito species (Table 1). The TE *Herves* was recently identified in the malaria vector, *An. gambiae* and is capable of genetically transforming *D. melanogaster*. *Herves* transposition within the mosquito is expected (Arensburger *et al.*, 2005). Recent population studies suggest that *Herves* is indeed currently active, or at least very recently active in populations of *An. gambiae, An. arabiensis* and *An. merus* (O'Brochta et al., 2006). Whole genome sequencing projects have also revealed a diverse range of other TEs in mosquitoes.

While genetic transformation in mosquitoes has become commonplace, the process is still inefficient and laborious when compared to transformation of *Drosophila*. The space and time requirements for mosquito colony maintenance are generally much more significant than those for *Drosophila*. Mosquitoes need specific environmental conditions for each life stage, requiring the isolation of life stages. Maintenance of vertebrates to serve as hosts for the bloodmeals necessary for egg production is also required. For most mosquito species, colonies must be continually maintained, rendering the perpetuation of independent transgenic lines for extended periods of time impractical. Transformation efficiencies in the mosquito are typically very low. Survival of injected embryos (G_0) is generally very low, only 10-20%. Additionally, the transformation frequency, the number of G_1 transgenic individuals/number of G_0 adults, is generally low as well, making this work costly and time consuming in the mosquitoes.

Table 1. Successfully transformed mosquitoes. Class II TEs have successfully been used to generate stable germline transformants in a number of medically important mosquito species.

Species	TE (host species)	Transformation Efficiency[a]	Reference[b]
Aedes aegypti	*Hermes* (*Musca domestica*)	~8%	(Jasinskiene *et al.*, 1998)
Aedes aegypti	*MosI* (*Drosophila mauritiana*)	~4%	(Coates *et al.*, 1998)
Aedes aegypti	*piggyBac* (*Trichoplusia ni*)	~5-10%	(Kokoza *et al.*, 2001)
Aedes aegypti	*Tn5* (*Escherichia coli*)	0.22%	(Rowan *et al.*, 2004)
Anopheles albimanus	*piggyBac*	20-43%	(Perera *et al.*, 2002)
Anopheles gambiae	*piggyBac*	~0.5%	(Grossman *et al.*, 2001)
Anopheles stephensi	*Minos* (*Drosophila hydei*)	~7%	(Catteruccia *et al.*, 2000)
Anopheles stephensi	*piggyBac*	~4%	(Nolan *et al.*, 2002)
Culex quinquefasciatus	*Hermes*	~12%	(Allen *et al.*, 2001)

TEs do not always behave as expected when integrating into the genomes of non-host species. Integration events have been observed where the entire donor cassette along with flanking plasmid sequence has been inserted into the germline of *Ae. aegypti* by the TEs *Mos1*, *piggyBac*, and *Hermes* (Tu and Coates, 2004) and in *Culex quinquefasciatus*

by *Hermes* (Allen *et al.*, 2001). Unstable tandem arrays of *piggyBac* are commonly observed in *Ae. aegypti* transgenic lines (Adelman *et al.*, 2004). Of the elements transformed into *Aedes*, somatic remobilization occurs infrequently and none have shown remobilization within the germline of the mosquito (O'Brochta *et al.*, 2003). Deficient transposition within the germline is possibly due to a requirement of specific spatiotemporal transposase expression. A second concern is the transposase copy number within the cells of the germline. In traditional transposition experiments, millions of helper plasmids may be co-injected with the donor plasmid. Upon stable integration of an endogenous transposase gene only one copy per cell is expected within the germline. It is plausible that this level of transposase is insufficient for proper integration or remobilization.

1.3.3. Endogenous promoters for transgene remobilization in the germline of the mosquitoes

Typically, a well-studied *Drosophila* promoter is used to drive transposase expression in non-drosophilids (Atkinson and Michel, 2002). These promoters are usually heat-inducible promoters. However, it has been shown that some *Drosophila* promoters do not display the same expression characteristics when transformed into the mosquito (Morris *et al.*, 1991). It has also been noted that inducible promoters, such as heat-shock promoters initially experience high levels of expression. This higher level of transposase from heat-induced promoters may actually reduce the rate of excision of a target element. It has been shown that high levels of transposase protein production reduces the overall effectiveness of the protein by overproduction inhibition in the *mariner* family of TEs (Lohe and Hartl, 1996). Similar phenomena have not been revealed in studies of the *piggyBac* TE (Wilson *et al.*, 2007). By replacing the

Drosophila promoter with an endogenous promoter and regulatory sequences, it is hypothesized that there will be higher, more consistent levels of transposase expressed (Adelman *et al.*, 2002). By selecting promoters that are responsible for the transcription of tissue and stage-specific expression, it is argued that one will reduce the potential negative fitness load that the transposable element's cut-and-paste method levies on the genome (James, 2002).

Previous studies in both *Aedes* and *Anopheles* mosquitoes have identified tissue-specific promoters (Coates *et al.*, 1999, Arca *et al.*, 1999, Kokoza *et al.*, 2000, Moreira *et al.*, 2000). Continued research to identify and characterize endogenous tissue-specific promoters led to the utilization of the *cis*-acting region upstream from an *An. gambiae* salivary gland gene to drive *LacZ* expression in the salivary glands of *An. stephensi* (Lombardo *et al.*, 2005). Cho and colleagues utilized the promoter region from a vitellogenin receptor gene to drive DsRed expression in the ovaries of female *Ae. aegypti* (Cho *et al.*, 2006). In these studies, the 5' *cis*-acting regulatory region from known tissue-specific genes was fused to a cDNA reporter gene. This new synthetic construct was then placed into a transposable element and then injected into embryos. The reporter gene allowed for not only the confirmation of transposition, but also for the characterization of the promoter and regulatory region.

Remobilization within the germline is integral to the use of transposable elements for future studies involving mosquitoes, and as a genetic drive mechanism for the creation of refractory strains capable of spreading genes through wild populations. Transposon mutagenesis and enhancer trap studies have aided the discovery and characterization of many genes within *Drosophila*. Similar studies in mosquitoes would

be of great importance to gene discovery and genome exploration. Such studies would require remobilization of the transposon for efficient genome coverage. Remobilization is also a requirement necessary for the fulfillment of the final goal: the creation of a refractory strain of mosquitoes by the introduction of exogenous DNA. A genetic drive mechanism is required to spread the introduced DNA into wild mosquito populations. Transposon replication relies heavily upon the remobilization of the element. Replication is thought to occur through template-directed gap repair (Figure 1). After remobilization of the integrated element, the homologous DNA serves as a template for gap repair. Replication will also occur if an element found behind the replication fork of S-phase nuclei transposes ahead of the fork, replicating within the chromosome once more (Wilson *et al.*, 2003).

By using an endogenous, germline-specific promoter to express transposase an increase in transposition frequency is expected due to several factors. It is believed that by confining the transposase to the germline, the mobility of the transposable element will also be strictly confined. This will increase the number of germ cells that incorporate the exogenous DNA, resulting in a greater number of transgenic offspring. Secondly, for an autonomous element (Figure 2A) with germline-specific transposase expression, remobilization will result in an increase in copy number and subsequently an

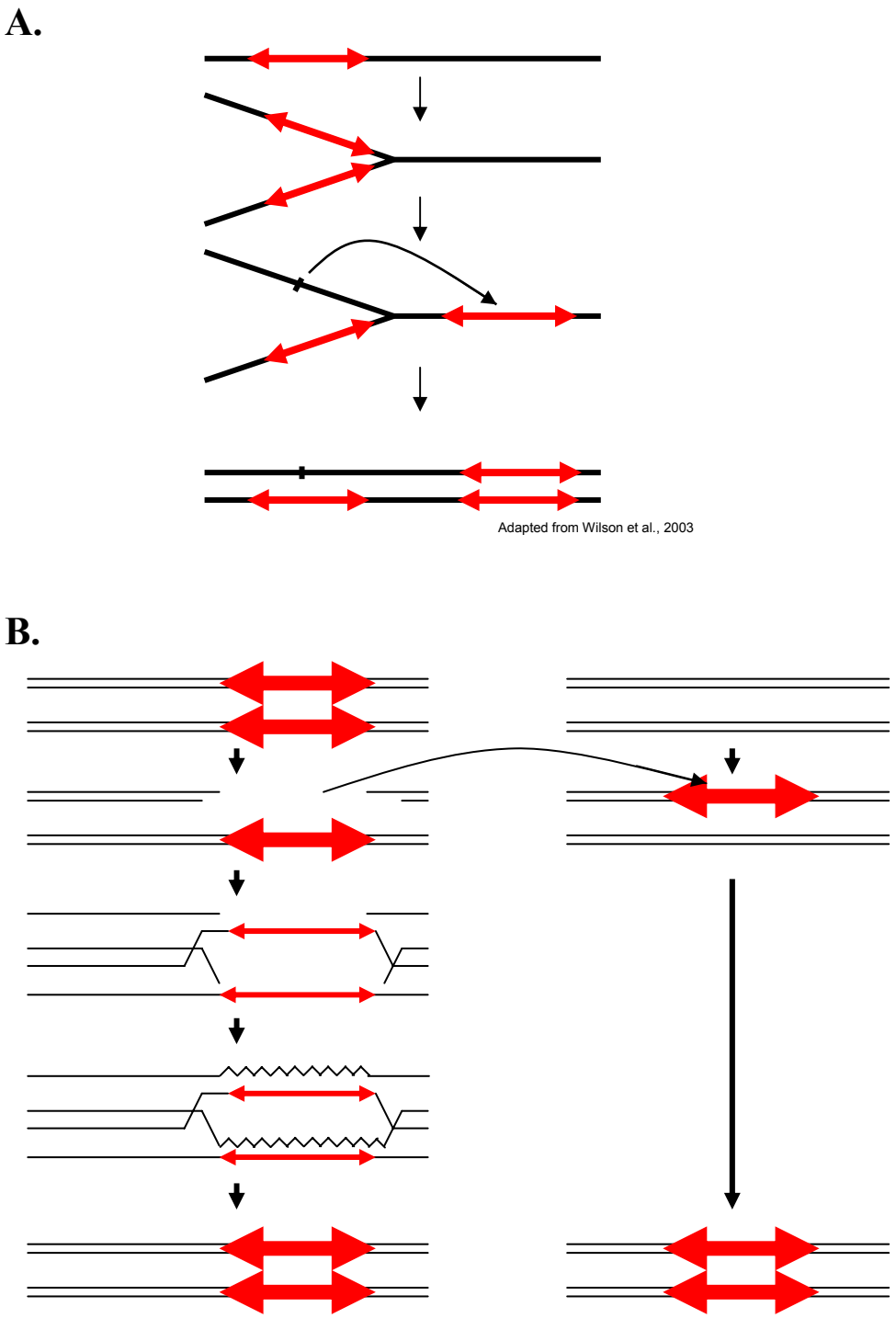


Figure 1. Illustration of replicative transposition. A) Replicative transposition from behind the replication fork to an unreplicated region of the chromosome. As replication is completed, one homologous chromosome maintains a single copy while the other now has a second TE. B) Replicative transposition followed by template-directed gap repair. The TE excises from a region of the genome and is inserted into a separate region. The resulting gap is then repaired via template-directed gap repair mechanism with a Holiday junction intermediate. The gap is resolved, resulting in an increased TE copy number.

increase in transposase concentration within germline cells. Additionally, by restricting the production of transposase, a reduction in lethal phenotypes caused by double strand breaks is expected. Finally, higher levels of transposase expression are expected using an endogenous promoter. It is this restricted but increased transposase activity that will lead to higher transposition frequency.

The creation of a line of transgenic mosquitoes that will be capable of producing germline-restricted transposase will be of great importance to the disease vector research community (Atkinson and James, 2002). A helper line developed with stable germline-specific transposase expression would require the injection of only a donor plasmid. By creating a helper strain of mosquitoes with germline-specific transposase expression, the nuclei of highly replicative germ tissue is targeted. Integration of the donor DNA would occur within the germline, resulting in a higher transposition frequency. This method would also allow for study of exogenous DNA integration by reducing the number of variables associated with current techniques of insect transformation; timing of transposase expression and transposase concentration (Adelman *et al.*, 2002).

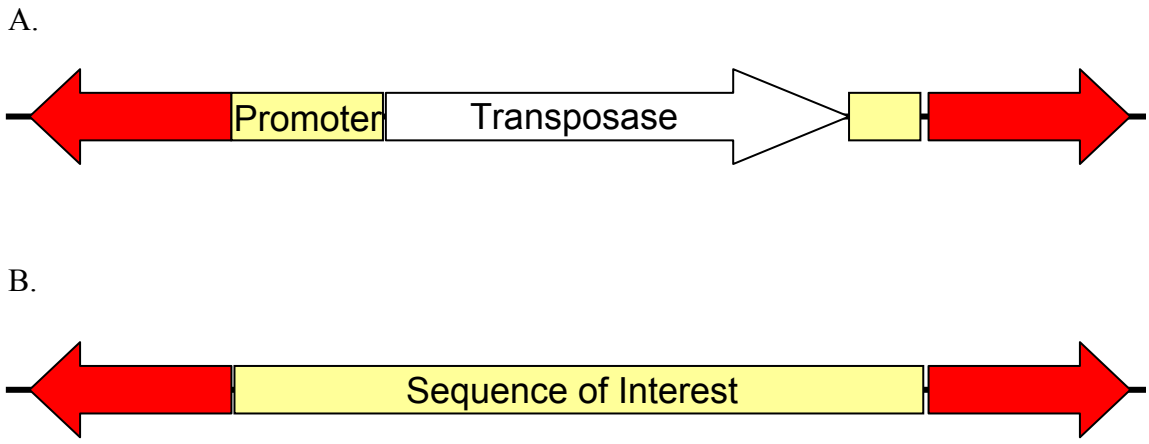


Figure 2. Transposable element composition referred to in the text. A) Autonomous TE contains an endogenous source of transposase that enables the TE to excise itself from the genome and re-insert elsewhere. B) A stably integrated TE. A sequence of interest is located within the inverted terminal repeats (red). Without an external supply of transposase, the TE cannot remobilize.

2. IDENTIFICATION OF A *VASA* HOMOLOG IN *Aedes aegypti* AND *Anopheles gambiae*

2.1. Introduction

Primordial germ cell (PGC) formation depends on the presence of pole plasm. Assembled at the posterior of the egg during oogenesis, pole plasm consists of multiple components (Saffman and Lasko, 1999). One such component, *vasa* (*vas*), is an ancient gene found in a number of widely divergent species. Originally described in *Drosophila melanogaster*, homologs have subsequently been found in planarians, arthropods, teleosts and mammals (Hay B., 1988a, Lasko and Ashburner, 1988, Shibata *et al.*, 1999, Nakao, 1999, Dearden *et al.*, 2003, Chang *et al.*, 2002, Olsen *et al.*, 1997, Yoon *et al.*, 1997, Fujiwara *et al.*, 1994, Komiya and Tanigawa, 1995). In *Drosophila*, *vas* is essential for proper PGC development (Mahowald, 2001).

Studies of embryonic development in *Aedes aegypti* and *Anopheles albitarsis* provide some evidence of potential pole plasm and future PGC development (Raminani LN, 1975, Raminani LN, 1978, Monnerat *et al.*, 2002). It will be of great interest to utilize a molecular marker capable of delineating the developing germ line from somatic tissues. Sequence and expression analysis data are consistent with the conclusion that we have cloned *vasa*-like genes from *Ae. aegypti* (*AaeVLG*) and *An. gambiae* (*AgVLG*). It is proposed that these genes can be used as markers to better understand germ line formation in developing mosquito embryos.

Germ line transformation efficiency in *Ae. aegypti* and *An. gambiae* is currently 5-10 fold lower than that of *D. melanogaster* (Handler, 2000). It is hypothesized that

replacing the exogenous constitutive promoters currently used for transposase expression with an endogenous germline-specific promoter and regulatory sequences will increase genetic transformation efficiency. Such a system would potentially result in localized expression of transposase at high levels in the appropriate cells (Adelman *et al.*, 2002). Identification of an endogenous germline-specific gene and subsequently regulatory regions would be of great importance to this field. A 40bp regulatory region responsible for localized *vas* expression within the germline has been identified in *D. melanogaster* (Sano *et al.*, 2002). With the identification of *AaeVLG* and *AgVLG*, it is now possible to clone the corresponding regulatory regions from the mosquito genes. These regulatory regions can then be used to result in localized transposase expression within the ovaries and developing embryos, potentially improving genetic transformation efficiency within mosquitoes.

2.2. Materials and Methods

2.2.1. Mosquito rearing

Aedes aegypti (Liverpool strain) were raised at 25^{o}C and 80% (±5%) relative humidity. Larvae were maintained in plastic containers and fed a steady diet of fish food. Pupae were transferred to paper containers and adults fed a 10% sucrose diet. Adult females were allowed to bloodfeed on mice 72 hours prior to ovary dissection.

Anopheles gambiae (pink-eye strain) were hatched and reared to the fourth instar larvae stage in plastic containers. *Anopheles gambiae* mosquitoes (G3 strain) were reared using standard protocols (Benedict, 1997) by Jennifer Juhn at University of California, Irvine.

2.2.2. Preparation of *Ae. aegypti* and *An. gambiae* genomic DNA

Pigment found in the eyes of *Ae. aegypti* and *An. gambiae* mosquitoes inhibits *Taq* polymerase activity and the heads must therefore be removed prior to homogenization. Clean razor blades were used to remove the heads of 14 male *Ae. aegypti* mosquitoes (Liverpool strain). A mixture of *An. gambiae* mosquito larvae (pink-eye strain) were used for the isolation of genomic DNA. The mosquitoes were homogenized in 80 µL of Bender Buffer (0.1 M NaCl; 0.2 M sucrose; 0.1 M Tris-HCl, pH 9.1; 0.05 M EDTA; 0.5% SDS) in a 1.7 mL sterile microfuge tube with a Kontes pestle and Pellet Pestle Motor (Kontes). The pestle was then washed with an additional 120 µL of Bender Buffer prior to adding 20 µL of 20 mg/mL Proteinase K (Promega). The solution was incubated overnight at 50^{o}C. The sample was gently extracted twice with Phenol/Chloroform/Isoamyl alcohol (25:24:1) and once with Chloroform/Isoamyl alcohol (24:1). The DNA was precipitated with the addition of 4 µL of 3 M NaOAc (pH 4.8) and 200 µL of isopropanol, mixing for 5 minutes and centrifugation for 10 minutes at 14,000 x g. The DNA pellet was washed with 70% ethanol, centrifuged for 5 minutes and air dried. The pellet was resuspended in 100 µL of nuclease-free water and stored at -20 ^{o}C.

2.2.3. *AaeVLG* and *AgVLG* PCR amplification and sequence isolation from genomic DNA

Degenerate primers were designed to the highly conserved regions, CAQTGSGT (CAQTG primer) and YVRIGH (YVRIGH primer) (Invitrogen, Carlsbad, CA). All primers used in this section can be found in Table 2. Genomic DNA was amplified using

Table 2. Primers found in Section 2. Names and corresponding nucleotide sequences of oligonucleotide primers used to amplify *AaeVLG* and *AgVLG* in Section 2.

Primer Name	Sequence (5'-3')
CAQTG	CARACIGGNWSNGGNAARAC
YVRIGH	CCDATNCKRTGNACRTANTC
DEADRM-F	TNGAYGARGCNGAYMGNATG
DEADRM-R	ARCATNCKRTCNGCYTCRTC
Ae-vasa1F	TGGAACCGTTCTCAAGGTTTGCG
5RACE-REV1	TGCTCCACATCCGTGCTGGCACCTCCGA
Ae-vasa1R	CGCGAATTTTACCCGTTCAAACG
Ae-vasa 1R2	GGGTGCCACATTCTTGTCGCCACGCC
cDNAanchor dT	GCTAATACGTACGATCGGTCGACAAGT18V
cDNAprimer2	GCTAATACGTACGATCGGTCGAC
An-gambiae 1R	TTGGGATTGCCGCCGTTCAGGATC
An-gambiae 5RN	CTTCAGCTTGGTGCCGTGCGCG
An-gambiae 1F	CGCCTTCATGCTGCCGATGATCCA
An-gambiae 3RN	CGGCGGAGCGTGTGCCGATGTG

these primers and nested in both directions using sense and antisense primers complementary to the VLDEADRM region (DEADRM-F and DEADRM-R).

Known amino acid sequences from *D. melanogaster*, *B. mori*, and *Ae. aegypti* were queried (TBLASTN) against the *An. gambiae* whole genome sequence (ENSEMBL). The match with the highest score was then used as a query (BLASTX) against the NCBI non-redundant protein database. Expecting high conservation, the BLOSUM80 scoring matrix was used. Specific primers (An-gambiae 1F and An-gambiae 1R) were used to amplify a fragment of the predicted *AgVLG* using genomic DNA from the pink-eye strain.

PCR reactions were performed under the following reaction conditions; 2 mM MgCl2, 0.2 μM each primer, 10 mM dNTPs, 1.0 μL genomic DNA, and 5.0 units of *Taq* polymerase, Promega, Madison, WI). The reaction parameters were as follows: 96 ^{o}C for 2'; 10 cycles of 96 ^{o}C for 30", 45 ^{o}C for 30", and 72 ^{o}C for 2' and 30 cycles of 96 ^{o}C for 30", 55 ^{o}C for 30", and 72 ^{o}C for 2', followed by a final extension at 72 ^{o}C for 10'. The amplified product was visualized on a 1% agarose gel to verify amplification and product size. Nested PCR reactions were performed under the conditions: 96 ^{o}C for 1'; 30 cycles of 96 ^{o}C for 30", 58 ^{o}C for 30", and 72 ^{o}C for 2'. The amplified product was once again visualized on an agarose gel to confirm amplification and the amplified product was then cloned into the TOPO 2.1 TA cloning vector (Invitrogen). Single pass sequence reads were generated from each end using the universal M13 primer set and the sequence determined using an ABI 3100 capillary sequencer.

2.2.4. Isolation of RNA and preparation of cDNA by reverse-transcriptase PCR

Total RNA was isolated from approximately 30 mg each of *Ae. aegypti* and *An. gambiae* larvae using the RNeasy Mini Kit (Qiagen, Valencia, CA), treated with DNase I with DNA-free (Ambion, Austin, TX) and subsequently used to synthesize first strand cDNA using SuperScript II reverse transcriptase (Invitrogen, Carlsbad, CA) per manufacturer's instructions. Briefly, 0.2 µM cDNAanchor dT-primer (Table 2), 1.0 µg total RNA, 1 µl of 10 mM dNTPs were brought to a total volume of 12 µl with ddH_2O, heated to 65 ^{o}C for 5 minutes and quick chilled on ice. Additionally, 5X 1st strand buffer, 10 mM DTT and 40 units of RNase OUT Recombinant Ribonuclease Inhibitor (Invitrogen) were added and the mixture incubated at 42 ^{o}C for 2 minutes. To increase the efficiency of the reverse-transcriptase reaction, 150 ng/µl of T4 Gene 32 protein (Roche Diagnostics, Indianapolis, IN) (Villalva *et al.*, 2001) was added to the mixture. After the addition of 200 units of SuperScript II (Invitrogen) the reaction was incubated for 1 hour at 42 ^{o}C, followed by heat inactivation at 70 ^{o}C for 15 minutes. The resulting product was digested with RNase H (Promega, Madison, WI) to produce single-stranded cDNA and stored at -20 ^{o}C.

2.2.5. Rapid amplification of cDNA ends (RACE) of *AaeVLG* and *AgVLG*

Aedes aegypti 4th instar larvae and *An. gambiae* 3rd-4th instar larvae were collected, rinsed in PBS, and snap-frozen. Total RNA was extracted using the RNeasy Mini Kit (Qiagen) and treated with DNase I using DNA-free (Ambion) according to the manufacturers' instructions. First-strand cDNA was synthesized using the SMART cDNA RACE Kit (Clontech, Palo Alto, CA) and 5' and 3' RACE reactions were performed according to manufacturer's instructions. Gene specific primers used for 5'

RACE were; Ae-vasa 1R and nested with Ae-vasa 1R2 for *AaeVLG*, and An-gambiae 1R and nested with An-gambiae 5RN for *AgVLG*. 3' RACE reactions were performed using the gene specific primers Ae-vasa1F and nested with 5RACE-REV1 for *AaeVLG* and the primers An-gambiae 1F and An-gambiae 3RN for *AgVLG*. RACE reaction conditions were outlined by the manufacturer and are as follows: 94 ^{o}C for 5'; 5 cycles of 94 ^{o}C for 10", 72 ^{o}C for 3'; 5 cycles of 94 ^{o}C for 10", 70 ^{o}C for 10", 72 ^{o}C for 3'; 25 cycles of 94 ^{o}C for 10", 68 ^{o}C for 10", 72 ^{o}C for 3'; and a final extension time of 10' at 72 ^{o}C. The resulting products were visualized on a 1% agarose gel, gel purified, cloned into the TOPO 2.1 vector (Invitrogen), and the DNA sequence determined using an ABI 3100 sequencer. For sequencing reactions, universal M13 primers were used followed by species specific sequencing primers, when required, to completely sequence the inserts. Sequences were edited to remove vector ends and collapsed into corresponding contigs using the Vector NTI Suite 8 software package (InforMax, Inc., 1999) to assemble full-length cDNAs for each candidate gene. Candidates were translated in silico using the Vector NTI programs to identify start and stop codons.

2.2.6. *Aedes* RNA isolation and Northern blot analysis

Total RNA was extracted from whole organisms or tissues using the Trizol method (Invitrogen, Carlsbad, CA). Embryos were aged either 1 hr or 24hrs at 25^{o}C and approximately 80% relative humidity. Embryos were collected 1hr after oviposition and at 24hrs after oviposition. Embryos were aged 24hrs at 25^{o}C and approximately 80% relative humidity. Fourth instar larvae (mixed sex) and pupae (sexed) were collected for RNA isolation one to two days after emergence. Adult male and female mosquitoes were collected. Sugarfed (SF) ovaries were dissected from sugarfed adult females and

bloodfed (BF) ovaries from bloodfed adult females 72 hours post bloodfeeding. Female carcasses were the result of SF ovary dissection. All dissections were performed in PBS and immediately frozen in liquid nitrogen. 10μg of total RNA was electrophoresed, transferred and fixed to a charged nylon NytranN membrane (Schleicher & Schuell BioScience, Keene, NH) and hybridized at 42^{o}C to a *vasa* probe corresponding to the 3' RACE amplification product.

2.2.7. Reverse-transcriptase polymerase chain reaction (RT-PCR) developmental profile in *Ae. aegypti*

Total RNA was isolated from approximately 30 mg each of *Ae. aegypti* and *An. gambiae* larvae using the RNeasy Mini Kit (Qiagen), treated with DNase I with DNA-free (Ambion). The concentration and 260/280 nm ratio of each RNA sample was determined using a spectrophotometer. Total RNA was subsequently used to synthesize first strand cDNA using SuperScript II reverse transcriptase (Invitrogen, Carlsbad, CA) per manufacturer's instructions. Briefly, 0.2 uM c-anchor dT-primer (Table 2), 1.0 μg total RNA, 1 μl of 10 mM dNTPs were brought to a total volume of 12 μl with ddH_2O, heated to 65 ^{o}C for 5 minutes and quick chilled on ice. Additionally, 5X 1st strand buffer, 10 mM DTT and 40 units of RNase OUT Recombinant Ribonuclease Inhibitor (Invitrogen) were added and the mixture incubated at 42 ^{o}C for 2 minutes. After the addition of 200 units of SuperScript II (Invitrogen) the reaction was incubated for 1 hour at 42 ^{o}C, followed by heat inactivation at 70 ^{o}C for 15 minutes. The resulting product was digested with RNase H (Promega, Madison, WI) to produce single-stranded cDNA and stored at -20 ^{o}C. Great care was taken to ensure that input RNA mass was identical

for each developmental stage. Upon final treatment of the single-strand cDNA with RNase H, DNA concentration was determined.

Polymerase chain reactions were assembled using precisely 100 ng 1st strand cDNA for each developmental stage. Assembled reactions also contained 10X *Taq* buffer, 1.5 µl 10 mM dNTPs, 0.2 µM each primer (Table 2) and 1 µl Advantage2 Taq Polymerase (BD Biosciences Clontech, Palo Alto, CA) and ddH_2O was added to a final volume of 50 µl. Reaction conditions were as follows: 95 °C for 5', 20-35 cycles of 95°C for 15", 60 °C for 15", 72 °C for 1'; final extension time was 5' at 72 °C. Resulting products were visualized on a 1% agarose gel.

2.2.8. *Anopheles* RNA isolation and Northern blot analysis

Total RNA was extracted using the Trizol (Life Technologies, Rockville, MD) method from whole organisms and tissues in a manner similar to that described for *Ae. aegypti*. However, for *An. gambiae*, the ovaries were extracted 48hrs after a bloodmeal. 5ug of Total RNA was electrophoresed in a 1.2% agarose-1X MOPS/2% formaldehyde gel and blotted by capillary action to a positively charged nylon Zeta-Probe GT membrane (Bio-Rad, Hercules, CA). Hybridization was performed at 65°C in Church's buffer (0.5M sodium phosphate, pH7.2, 7% SDS) with a *vasa* probe corresponding to a 5' RACE gene amplification product.

2.2.9. Whole mount *in situ* hybridizations

Whole-mount *in situ* hybridizations were performed using digoxigenin-labelled probes essentially as described in previous literature (Juhn and James, 2006). Sense or anti-sense dig-labeled probes were transcribed with SP6 or T7 RNA polymerase. Runoff transcription reactions were performed using 1µg of linearized plasmid DNA in 20 µl

reactions, supplemented 2 µl 10x digoxigenin RNA labeling mix (Roche Diagnostics, Indianapolis, IN) for unlabeled nucleotides, using MAXIscript *In vitro* Transcription Kit (Ambion, Austin, TX). Reactions were incubated for 2 hours at 37^{o}C, transcription verified on an agarose gel and precipitated with ethanol in the presence of 0.8M ammonium acetate. Embryos were collected at varying time points ranging from 1 hr. after oviposition to 20 hrs. post oviposition. *Aedes aegypti* embryos were washed in 1.3125% sodium hypochlorite for 30-35 seconds while *An. gambiae* embryos were washed in the same solution for 1.25 minutes. Rinsing with ddH_20 commenced immediately for 2 minutes and the embryos were then placed in fresh ddH_20 before being transferred to heptane (Goltsev *et al.*, 2004).

Embryos were fixed in 9.25% formaldehyde (FA) solution (10mL 37% formaldehyde, 30mL ddH_20, and 10µl 10M NaOH) for 1 hour while gently mixing on a nutator. FA solution was removed and embryos rinsed twice in ddH_20, the second rinse lasting 30 minutes. Water was removed and embryos were submerged in boiling water for 30 seconds before the hot water was then removed. Prechilled ddH_20 was added to the embryos and incubated on ice for 10 minutes. Water and remaining heptane were removed and fresh heptane was added. To crack the exochorion and vitelline membrane, methanol was added and the embryos allowed to settle for 15-20 minutes. After washing several times with fresh methanol, embryos were transferred to a Petri dish with double-sided stick tape. Once the embryos were allowed to settle and stick to the tape, ethanol and water were added before the embryos were coaxed out of the vitelline membrane with the aide of a fine-tipped paintbrush or 25 gauge needle. Freed embryos were rinsed

2-3 times in ethanol and carried directly to the hybridization steps or placed at -20 °C for storage.

Prehybridization, hybridization, and staining steps were carried out as described by Juhn and James (Juhn and James, 2006). Embryos were washed in 90% P-xylene/10% ethanol and rinsed in ethanol, methanol, and then rehydrated. Embryos were postfixed in 4% FA solution in PBT. After treatment with proteinase K, embryos were washed in PBT and postfixed in 4% FA solution before being transferred step-wise into hybridization solution (50% formamide, 5x SSC, 100 µg/mL salmon sperm DNA, 50 µg heparin and 1% Tween-20). After incubation of at least 1 hour, dig-labeled probe was added to a final concentration of 100ng/mL and incubated overnight at 55 °C. Embryos were subjected to six 10 minute PBT washes, treated for 30 minutes in 20 µg/mL RNase A at 37 °C, washed again in PBT, and then incubated for 2 hours in blocking solution (5% goat serum, 1% Roche Western Blocking reagent (Roche Diagnostics, Indianapolis, IN), and PBT). Embryos were incubated at 4 °C overnight with anti-digoxigenin-AP Fab fragments (Roche Diagnostics, Indianapolis, IN). Embryos were washed in PBT following antibody incubation then rinsed in alkaline phosphate buffer (50mM $MgCl_2$, 100mM NaCl, 100mM Tris pH 9.5, 0.01% Tween20). NBT/BCIP color substrates were used for probe detection. Embryos were then rinsed in PBT before being mounted in 70% glycerol.

2.2.10. Amino acid alignment

AaeVLG and *AgVLG* amino acid predictions were aligned with amino acid sequences from *B. mori*, *D. melanogaster*, *Schistocerca gregaria*, and *Tribolium castaneum* (Nakao, 1999, Hay B., 1988a, Lasko and Ashburner, 1988, Chang *et al.*,

2002), AAW78361) using CLUSTALW v. 1.83 (Thompson *et al.*, 1994). Alignments were visualized using BioEdit (Hall, 2001).

2.2.11. Phylogenetic analysis

Amino acid predictions for *AaeVLG* and *AgVLG* were aligned with protein sequences from *B. mori*, *D. melanogaster*, *S. gregaria*, *T. castaneum*, *Danio rerio*, *Gallus gallus*, *Oncorhynchus mykiss*, and *Rattus* sp. (Nakao, 1999, Hay B., 1988a, Lasko and Ashburner, 1988, Chang *et al.*, 2002, Yoon *et al.*, 1997, Yoshizaki *et al.*, 2000, Komiya and Tanigawa, 1995, Tsunekawa *et al.*, 2000) using ClustalW v 1.83 (Thompson *et al.*, 1994). Rascal was used to review the alignment and Gblocks removed the highly divergent regions of the alignment (Thompson JD, 2003, Castresana, 2000). Maximum likelihood analysis was performed using Phylip v 3.63 and the dendrogram was created by Drawgram (Felsenstein, 1989).

2.3. Results

2.3.1. Isolation and expression of *AaeVLG*

Degenerate oligonucleotide primers were designed to amplify genomic DNA between the known conserved Vas domains CAQTGSGT, VLDEADRM, and YVHRIGHR (Linder *et al.*, 1989). A 363bp fragment was isolated spanning the region from CAQTGSGT to VLDEADRM, enabling the design of gene-specific primers for use in 3' Rapid Amplification of cDNA Ends (RACE) and 5' RACE. This provided a full length *vas* sequence approximately 2.3Kb in length with a 191bp 5' UTR and 215bp 3' UTR (Accession # AY957502). A potential upstream ORF is found in the 5' UTR with an AUG initiation codon at nucleotides 9-11 and a termination codon at nucleotides 174-176. The occurrence of AUG triplets in 5' UTRs (uAUG) is not uncommon. *In silico*

studies determined the frequency of uAUG in mRNA sequences to be approximately 40% in vertebrates and more than 50% in *Drosophila sp.* (Peri and Pandey, 2001, Rogozin *et al.*, 2001). Experimental evidence has shown that uAUG preceded by short leader sequences are translated inefficiently, resulting in initiation of translation from a downstream AUG via leaky scanning (Kozak, 1991). An upstream sequence with a minimum of 20 nucleotides is required for efficient translation to be initiated at the uAUG.

Sequence analysis revealed *AaeVLG* has a predicted open reading frame of 1914 nucleotides, encoding a protein of 638 amino acids in length. The predicted amino acid sequence shares approximately 52% identity with the *Drosophila* Vas sequence. AaeVLG contains the eight expected conserved motifs unique to members of the DEAD-box family of proteins (Linder *et al.*, 1989). Additionally, a N-terminal arginine-glycine-glycine repeat region was also present, as found in *D. melanogaster* Vas (Lasko and Ashburner, 1988, Hay B., 1988b). An ARKF motif unique to *Vasa*, AN3, and PL10 proteins is found at residues 308 to 311. Motifs common to insect *Vasa* proteins are also found (GIVGGA, residues 428 to 433, and a EXEEXW motif at the C-terminal end) (Chang *et al.*, 2002). Figure 3 shows an alignment of the predicted *AaeVLG* protein with other insect Vas or Vasa-like proteins (Nakao, 1999, Chang *et al.*, 2002, Hay B., 1988a, Lasko and Ashburner, 1988), AAW78361).

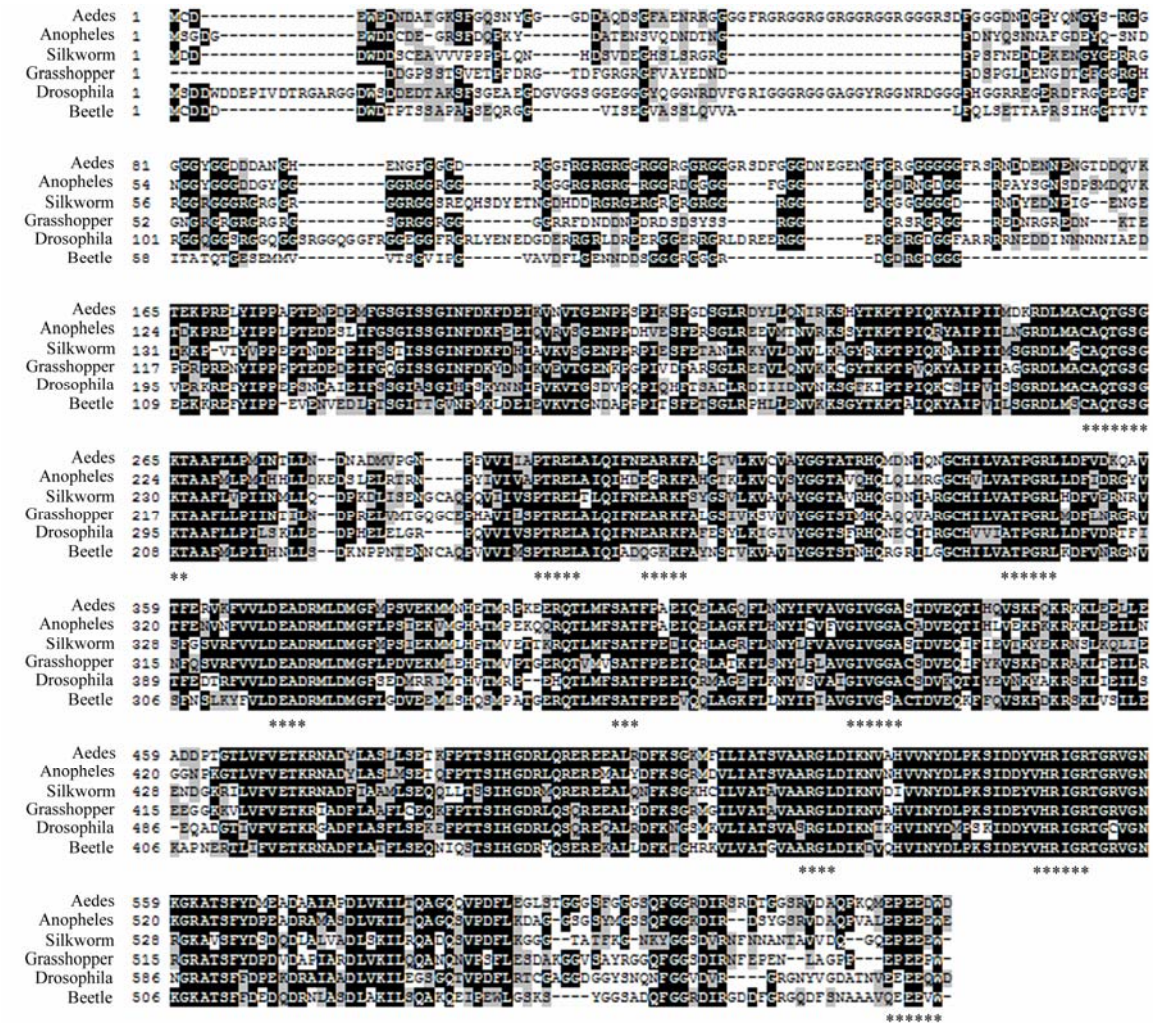
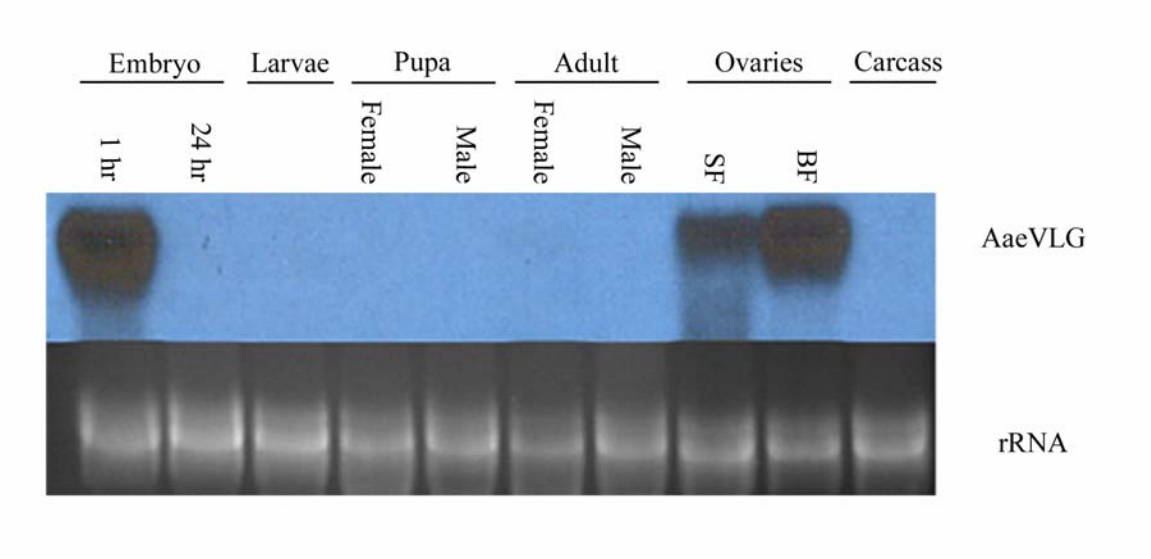


Figure 3. Multiple sequence alignment of *AaeVLG* and *AgVLG* predicted amino acid sequences and homologous insect Vasa proteins. Asterisks mark positions of conserved amino acids in the DEAD-box family and of insect Vasa proteins.

Northern Blot analysis showed that *AaeVLG* is highly expressed in early embryos and adult ovaries (Figure 4a). Ovaries undergoing vitellogenesis also showed additional upregulation of *vas*. Northern Blot analysis using longer exposure times also showed low levels of expression in 24 hour old embryos, larvae, and adult females (Figure 4b). This is consistent with the *Drosophila* data which shows near complete maternal transcript degradation essentially coinciding with blastoderm formation (Hay B., 1988b). In *Ae. aegypti*, blastoderm formation occurs 5-10 hrs after egg deposition (Raminani LN, 1975). The low levels of transcript in 24hr embryos suggest that maternally derived transcript is systematically degraded prior to this time point. More sensitive expression analysis (RT-PCR) revealed low levels of *vas* RNA in pupae and male adults (Figure 5). This is similar to the results found in *Drosophila* where *in situ* hybridization data show transcript in developing testes as well as ovaries (Lasko and Ashburner, 1988, Hay B., 1988b).

A.



B.

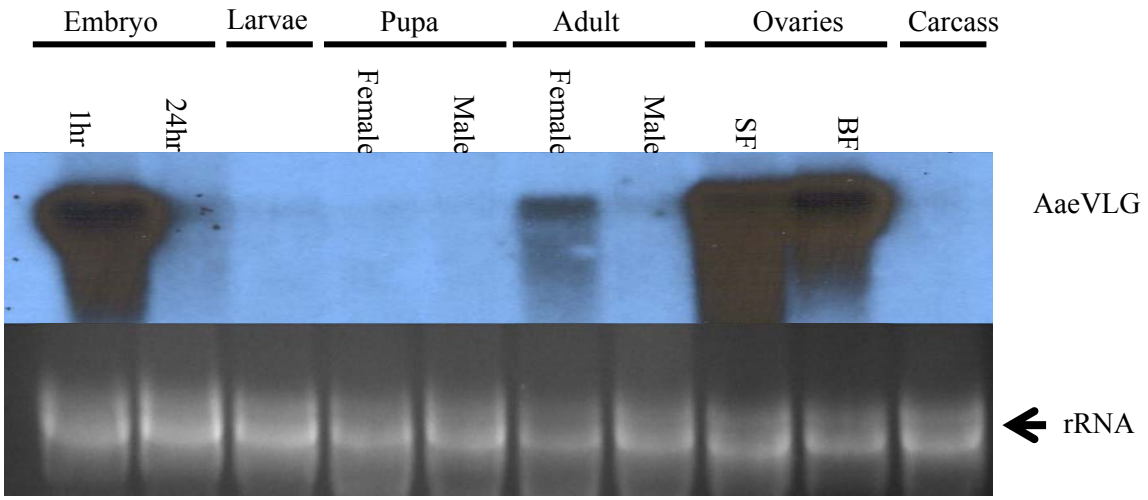


Figure 4. Northern blot analysis of *AaeVLG* expression. Developmental expression profile is from 1 hour embryo to adult. Ovaries were dissected from either sugar-fed (SF) adult females or blood-fed (BF) adult females dissected 72 hours after bloodmeal. Carcasses are of adult females with ovaries dissected out. 2A) 4 hour exposure. 2B) 18 hour exposure. Ribosomal RNA loading controls are shown in the lower panel.

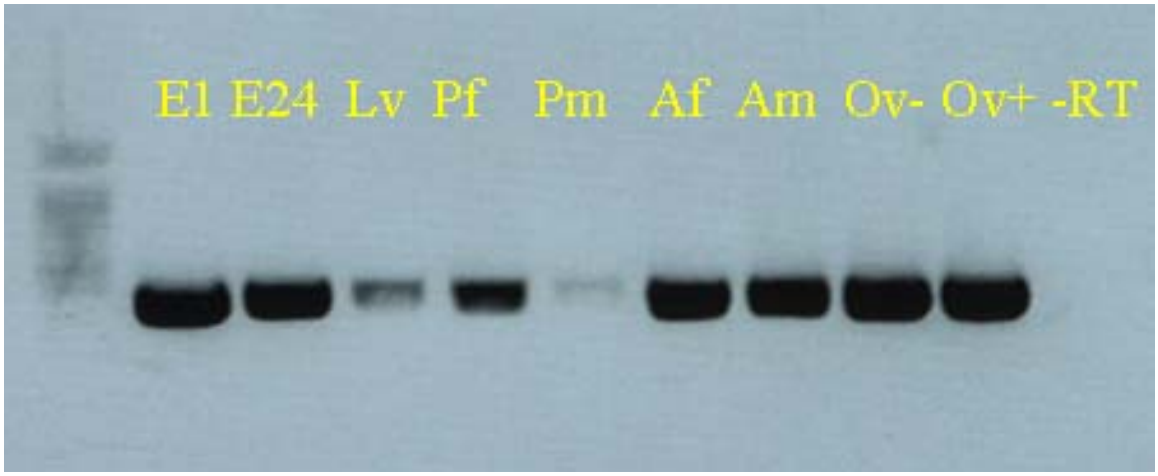


Figure 5. Inverse image of RT-PCR expression profile of *AaeVLG*. Developmental stages include 1 hr. embryos (E1), 24 hr. embryos (E24), larvae (Lv), female and male pupae (Pf/Pm), female and male adults (Af/Am), dissected ovaries from sugar-fed females (Ov-) and blood-fed females (Ov+), and negative control (-RT).

2.3.2. Isolation and expression of *AgVLG*

Amino acid sequences from the *vasa* genes of three insect species, *D. melanogaster*, *Bombyx mori* and *Ae. aegypti* were used to query the *Anopheles* whole genome sequence annotated by ENSEMBL (Holt *et al.*, 2002). TBLASTN database searches for each query sequence resulted in a single sequence with significant similarity. Queries matched a region of the whole genome sequence on chromosome 3R, approximately from nucleotides 13601000 to 13605000. Gene-specific primers were used to amplify an internal region from *An. gambiae* larval cDNA to provide evidence of transcription. Subsequently, overlapping primers amplified a full length cDNA by means of 3' RACE and 5' RACE. The 5' terminus of *AgVLG* was extended by 9 nucleotides beyond that of the EST (ENSANGESTT00000358823) provided by ENSEMBL. All

amplification and sequencing efforts resulted in a single sequence, suggesting that it is likely that we have identified the primary transcript. The *AgVLG* cDNA is approximately 2.3Kb in length with a 230bp 5' UTR and 266bp 3' UTR and a 1788bp open reading frame encoding a predicted protein of 596 amino acids. Each of the eight conserved motifs common to DEAD-box proteins are found in *AgVLG*. The arginine-glycine-glycine repeat region and the GIVGGA and EXEEXW insect motifs also are present. Unlike most other *vasa* sequences, *AgVLG* does not completely maintain the expected ARKF domain. However, the alanine to glycine substitution within this motif in *An. gambiae* is also seen in the flour beetle, *Tribolium castaneum* (AAW78361).

Northern blot hybridization shows high levels of *AgVLG* expression in early embryos and ovaries (Figure 6). Unlike *Ae. aegypti*, there is no evidence of a dramatic upregulation of mRNA expression from unfed ovaries to bloodfed ovaries. Embryos aged 24hrs show low levels of transcript, suggesting that maternal transcript is degraded by this time. An embryological study of *An. albitarsis* revealed that blastoderm formation occurs 4-6hrs after egg deposition (Monnerat *et al.*, 2002). However, evidence suggests that maternal *AgVLG* degrades rapidly approximately 24hrs after deposition (data not shown). This rapid degradation of *vas* transcript is consistent with the *Aedes* and *Drosophila* data (Lasko and Ashburner, 1988).

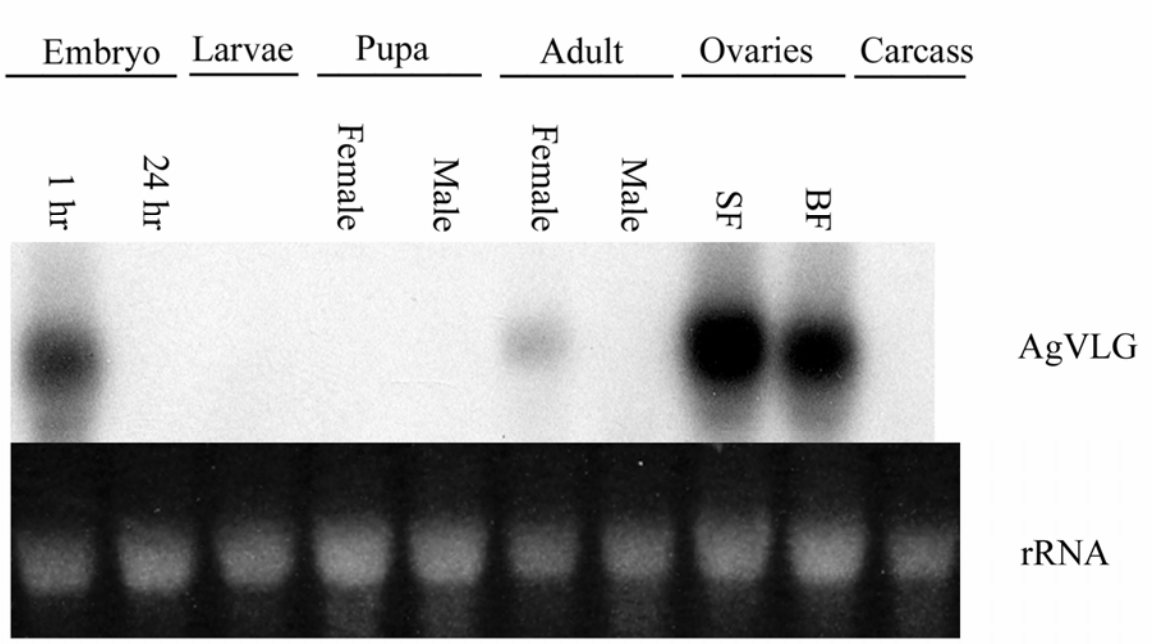


Figure 6. Northern blot analysis of *AgVLG* expression. Developmental expression profile is from 1 hour embryo to adult. Ovaries were dissected from either sugar-fed (SF) adult females or blood-fed (BF) adult females dissected 48 hours after bloodmeal. Carcasses are of adult females with ovaries dissected out. Ribosomal RNA loading controls are shown in the lower panel.

2.3.3 Whole-mount embryo *in situ* hybridizations

In situ hybridization of digoxigenin-labelled anti-sense RNA to whole-mount embryos showed *AaeVLG* mRNA localized to the posterior end of presyncytial blastoderm-stage embryos (mitotic divisions 4-7) as well as the posterior end and pole buds of the syncytial blastoderm (mitotic divisions 11-12) (Figure 7). These results are consistent with *vasa* RNA accumulation in the posterior pole of developing *Drosophila* embryos as well as non-drosophilid embryos (Schroder, 2006, Nakao, 1999). *AaeVLG* localized to the presumptive germ tissue of gastrula stage embryos (mitotic division 14-16), mimicking results from studies of the germline-specific gene, *oskar* (J. Juhn, personal conversation).

Figure 7. *In situ* hybridization of *Aedes aegypti* embryos using digoxigenin-labelled anti-sense RNA. Posterior poles are oriented to the left side of the image. A) Early stage syncytial blastoderm embryo corresponding to mitotic division 11-12. *AaeVLG* transcript localization is found in the budding pole cells in the posterior pole. B) Gastrulation stage embryo (approximately mitotic division 14-16) shows *AaeVLG* localized to the presumptive germ tissue. Arrows show *AaeVLG* transcript localization

2.3.4. Phylogenetic analysis of *AaeVLG* and *AgVLG*

Global alignments using ClustalW v.1.83 were performed using insect and vertebrate Vasa protein homologs. GBlocks removed all but the most highly conserved regions, reducing the sequences to 440 amino acids (Castresana, 2000). Removing the repetitive or divergent regions such as the arginine-glycine-glycine repeat regions allowed for more sensitive phylogenetic analysis. Maximum likelihood analysis of this

alignment using Phylip v3.63 shows that both *AaeVLG* and *AgVLG* cluster with other insect *vasa* homologs and not the closely related *D. melanogaster* DEAD-box protein, Belle (Felsenstein, 1989, Johnstone *et al.*, 2005)(Figure 8). This analysis suggests that we have cloned the *vasa* homolog in *Ae. aegypti* and *An. gambiae*.

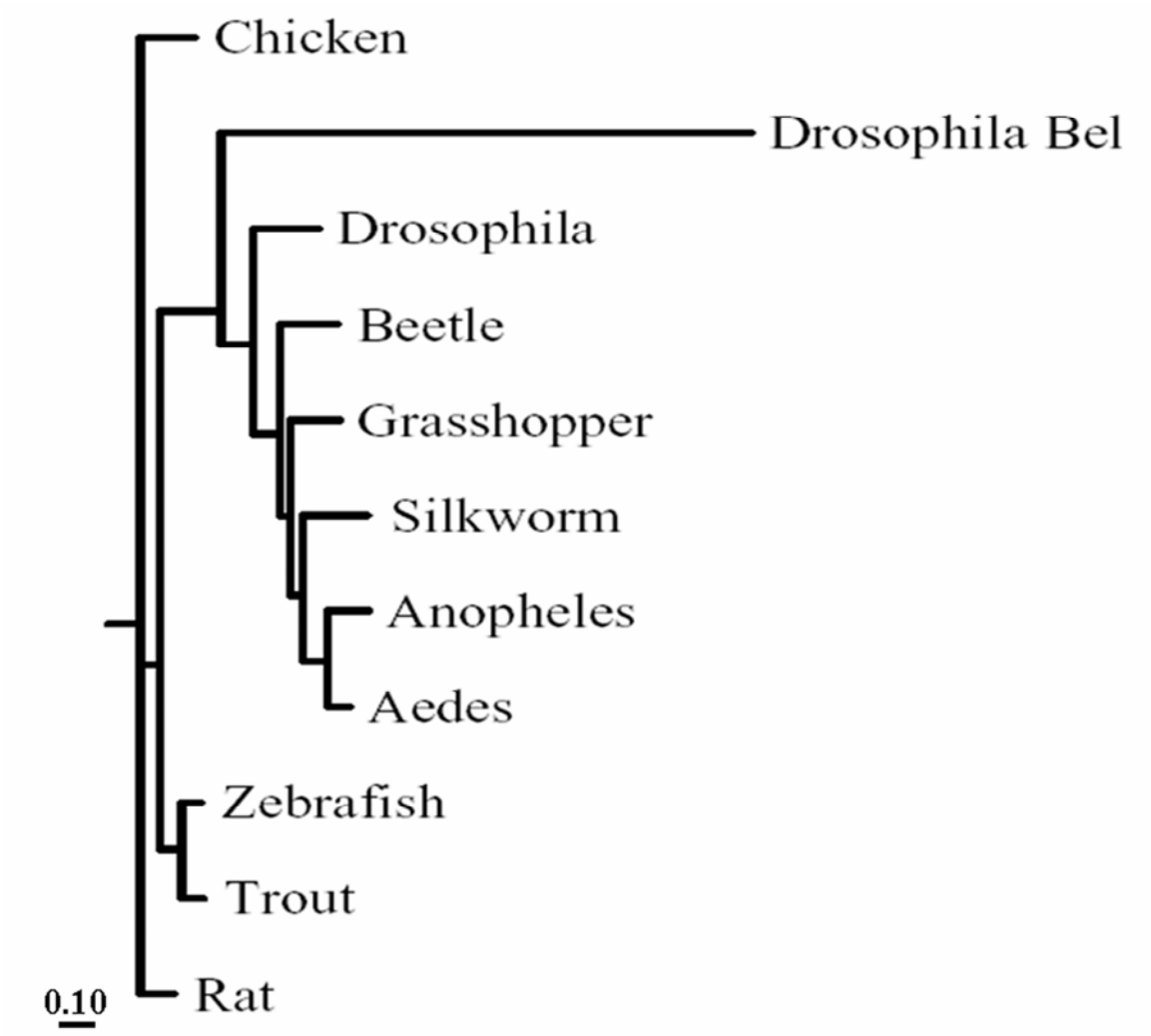


Figure 8. Dendrogram of maximum likelihood analysis of multiple sequence alignment of *AaeVLG*. Multiple sequence alignment of *AaeVLG*, *AgVLG*, vertebrate and invertebrate Vasa homologs, and the closely related Drosophila DEAD-box protein, Belle. *AaeVLG* and *AgVLG* cluster with other invertebrate Vasa-like proteins and not with Belle.

2.4. Discussion

2.4.1. *AaeVLG* and *AgVLG* are *vasa* orthologs

Significant evidence including sequence similarity, gene structure, spatio-temporal expression and mRNA accumulation patterns of *AaeVLG* and *AgVLG* lead to the conclusion that the mosquito orthologs of *vasa* have successfully been identified and cloned.

2.4.2 Future of *AaeVLG* and *AgVLG*

With the identification of *AaeVLG* and *AgVLG*, it is now possible to use the mosquito *vas* transcripts as a means of establishing molecular time points, following the germline as it develops. In addition to *in situ* hybridizations the focus is now on cloning the regulatory regions that confer germline-specific expression. The identification of regulatory sequences will allow the expression of transgenic fluorescent reporter genes, identifying the cells of the germline as it develops and proving correct regulatory region is cloned and complete.

The identification of regulatory regions responsible for conferring localized expression will also be of great importance in the creation of a more efficient mosquito transformation system. Transposase expression driven by a well-defined endogenous germline-specific promoter will result in a localized, highly expressed protein. This will assist in the optimization of transpositional activity in mosquitoes (Adelman *et al.*, 2002). Low levels of transposition within the germline are possibly due to a requirement for specific spatiotemporal transposase expression. By using an endogenous, germline-specific promoter an increase in transposition frequency is expected due to several factors. It is believed that by confining the transposase to the germline, the mobility of

the transposable element will also be strictly confined. For primary transformation experiments, this will increase the number of germ cells that incorporate the exogenous DNA vectors, resulting in a greater number of transgenic offspring.

It has also been proposed that autonomous transposons can be harnessed for spreading refractory transgenes throughout populations (Adelman *et al.*, 2002, Robinson *et al.*, 2004). For the transposable elements introduced into the *Aedes* genome thus far, somatic remobilization of previously integrated transposons occurs infrequently and at very low frequencies within the germline (O'Brochta *et al.*, 2003). Remobilization of these transposons using germline specific regulatory sequences will result in an increase in copy number and subsequently an increase in transposase concentration within germline cells. Using a germline-specific promoter will also decrease the likelihood of any fitness costs as a result of somatic remobilization.

3. IDENTIFICATION OF GERMLINE-SPECIFIC GENES IN THE DENGUE MOSQUITO, *Aedes aegypti*

3.1. Introduction

Model organisms in the post-genomics era are more important than ever. While an ever increasing number of completed genomes are being sequenced for a variety of organisms, it is impractical to expect the same level of experimental coverage as provided to model organisms. It is therefore of great interest to utilize expression and functional data from related organisms to identify homologous, or orthologous, genes within the species of interest. With the limited amount of expression, genomic, and particularly functional data available for non-drosophilid insects, *Drosophila* data can provide a wealth of information. Identification of genes exhibiting germline specificity within *D. melanogaster* embryos has previously led to the identification of putative orthologs in several mosquito species (Juhn and James, 2006).

The Berkeley Drosophila Genome Project database (http://www.fruitfly.org/cgi-bin/ex-insitu.pl) (Tomancak *et al.*, 2002) was queried to identify genes unique to the germline and pole plasm. Additionally, genes described in the literature to be germline-specific were included in the study. Literature reviews were performed in an attempt to validate expression profiles as well as germ cell and posterior pole cell function. Of the genes identified; *maelstrom*, *tudor*, *piwi*, and *aubergine* were well described in the current literature. CG10125, henceforth known as *zpg*, was originally identified in a *D. melanogaster P*-element mutagenesis study. *zpg* has since been well studied and found to localize to the posterior pole in developing embryos as well as differentiated germ cells and is required early in both oogenesis and spermatogenesis (Tazuke *et al.*, 2002, Gilboa

et al., 2003). The gene, CG5292 (*shu*), was originally identified in a screen identifying mutants affecting oogenesis that were located on the second chromosome (Schupbach and Wieschaus, 1991). *shu* RNA also accumulates in the germplasm of blastoderm-stage embryos in pole cells and the embryonic gonads. s*hu* was successfully cloned and the expression pattern at the RNA and protein level is consistent with a function in the germline stem cells (Munn and Steward, 2000). The gene, CG10050, has an unknown molecular function and the biological processes in which it is involved are also not known. Whole-mount embryo *in situ* hybridization data shows that the CG10050 mRNA localizes to the pole plasm of the *D. melanogaster* embryo, suggesting a role in the development of germ cells.

Identification of endogenous germline-specific genes and the regulatory regions that drive expression will allow for the creation of autonomous TEs to potentially serve as a gene drive mechanism (Atkinson and James, 2002). The use of endogenous regulatory sequences that are species-specific may also mitigate the potential of TEs to invade non-target species via horizontal transfer. Three *Ae. aegypti* cDNAs encoding polypeptides with significant sequence similarity to *D. melanogaster* protein sequences believed to have germline-specific expression have been cloned. RT-PCR and mRNA localization reveal that *Aezpg*, *Aeshu*, and *AeEwald* are expressed in presumptive germ tissue.

3.2. Materials and Methods

3.2.1. Screening of the Berkeley Drosophila Genome Project (BDGP) Patterns of Gene Expression Database

The Berkeley Drosophila Genome Project database (http://www.fruitfly.org/cgi-bin/ex-insitu.pl) (Tomancak *et al.*, 2002) was queried to identify genes unique to the germline and pole plasm. The BDGP database consists of a series of mRNA *in situ* hybridization images documenting the expression patterns of over 3,500 genes. Candidate gene selection was performed using key word queries consisting of the terms; germ cell, gonad, pole cell, and pole plasm. Additional searches used the terms, subset and localized, to identify non-specified mRNA localization. All searches were performed using the 'Basic Query' function. Candidate genes with defined somatic expression were removed from the results using the 'Advanced Query' function and the 'NOT' operator. The hybridization images for the resulting candidate genes were each visually inspected to confirm unique germline expression.

3.2.2. Identification of *Ae. aegypti* germline-specific candidate genes

Drosophila candidate gene sequences identified from the BDGP database were queried using TIGR Gene Indicies BLAST (Quackenbush *et al.*, 2001) and the National Center for Biotechnology Information (NCBI) website (http://www.ncbi.nlm.nih.gov) against *Ae. aegypti* EST and genomic sequence databases utilizing the tBLASTx function. Putative *Ae. aegypti* sequences were reciprocally queried against the Drosophila EST database.

3.2.3. Isolation of RNA and preparation of cDNA by reverse-transcriptase PCR

Total RNA was isolated from approximately 30 mg of *Ae. aegypti* using the RNeasy Mini Kit (Qiagen, Valencia, CA), treated with DNase I with DNA-free (Ambion, Austin, TX) and subsequently used to synthesize first strand cDNA using SuperScript II reverse transcriptase (Invitrogen, Carlsbad, CA) per manufacturer's instructions. Briefly, 0.2 μM c-anchor dT-primer (Table 3), 1.0 μg total RNA, 1 μl of 10 mM dNTPs were brought to a total volume of 12 μl with ddH_2O, heated to 65 ^oC for 5 minutes and quick chilled on ice. Additionally, 5X 1st strand buffer, 10 mM DTT and 40 units of RNase OUT Recombinant Ribonuclease Inhibitor (Invitrogen) were added and the mixture incubated at 42 ^oC for 2 minutes. To increase the efficiency of the reverse-transcriptase reaction, 150 ng/μl of T4 Gene 32 protein (Roche Diagnostics, Indianapolis, IN) (Villalva *et al.*, 2001) was added to the mixture. After the addition of 200 units of SuperScript II (Invitrogen) the reaction was incubated for 1 hour at 42 ^oC, followed by heat inactivation at 70 ^oC for 15 minutes. The resulting product was digested with RNase H (Promega, Madison, WI) to produce single-stranded cDNA and stored at -20 ^oC.

3.2.4. Amplification of *Ae. aegypti* germline-specific gene candidates: Confirmation of expression

Gene specific primers were designed to amplify a region of approximately 1 Kb for each of the candidate genes to confirm expression (Table 3). Internally nested primers were also designed. PCR amplifications were performed using standard gene amplification protocols. Reactions were performed with 2 mM $MgCl_2$, 0.2 μM of each primer, 2.5 mM dNTPs, 100 ng 1st strand cDNA product and 2.0 units of *Taq* polymerase (Continental Lab Products, San Diego, CA). PCR conditions used to amplify gene

fragments were as follows: 95 °C for 2 minutes; 30 cycles of 95 °C for 30 seconds, 62 °C for 30 seconds, and an extension phase of 72 °C for 1 minute; followed by a final extension phase of 72 °C for 5 minutes. Amplified products were visualized on a 1% agarose gel. Gene fragments successfully amplified from 1st strand cDNAs were subsequently cloned into the pGEM-T vector (Promega Corp., Madison, WI) and sequenced using standard M13 primers and analyzed using an ABI 3100 capillary sequencer.

3.2.5. Rapid amplification of cDNA ends (RACE) of gene candidates

Aedes aegypti 4th instar larvae were collected, rinsed in PBS, and snap-frozen. Total RNA was extracted using the RNeasy Mini Kit (Qiagen) and treated with DNase I using DNA-free (Ambion) according to the manufacturers' instructions. First-strand cDNA was synthesized using the SMART cDNA RACE Kit (Clontech, Palo Alto, CA) and 5' and 3' RACE reactions were performed according to manufacturer's instructions. Primers used for RACE reactions were the same as those used to confirm expression of each candidate gene (Table 3) with the primers noted "F" used for 3' RACE and those noted "R" used for 5' RACE. RACE reaction conditions were outlined by the manufacturer and are as follows: 94 °C for 5'; 5 cycles of 94 °C for 10", 72 °C for 3'; 5 cycles of 94 °C for 10", 70 °C for 10", 72 °C for 3'; 25 cycles of 94 °C for 10", 68 °C for 10", 72 °C for 3'; and a final extension time of 10' at 72 °C. The resulting products were visualized on a 1% agarose gel, gel purified, cloned into the TOPO 2.1 vector (Invitrogen), and the DNA sequence determined using an ABI 3100 sequencer. For sequencing reactions, universal M13 primers were used followed by species specific

sequencing primers, when required, to completely sequence the inserts. Sequences were edited to remove vector ends and collapsed into corresponding contigs using the Vector

Table 3. Primers used in Section 3.

Primer Name	Sequence (5'-3')
cDNAanchor dT	GCTAATACGTACGATCGGTCGACAAGT18V
cDNAprimer2	GCTAATACGTACGATCGGTCGAC
AeagMAEL-F	GGAAGCTTTCGTACCGGCGGAG
AeagMAEL-R	CGCATTCGGCGGTGTCGGAAGTTCG
AeagTUD-F	TCCGTCGGAGCAACGCTGCTTCACC
AeagTUD-R	GCCGCTAGTCCGGGGTTGTGCAGCT
AedesCG10050F	CCATGTTATCGGTCGGATTGGC
AedesCG10050R	CCACGTCCCGTCGATGAGAATG
AedesCG4735F	GAATACGAACGGGCCAAAACAAGC
AedesCG4735R	CCAGGTTCTTCTTCATCGTGTTGG
AedesCG10125F	AGGAACGCACATGGGTCACATTCCC
AedesCG10125R	TTACGCTTCCGAGTCCACGCACCA
AedesCG5292F	TATTCATACGGCCCAATAACTAGG
AedesCG5292R	CATAGAGCCATTTTACGAGACGT
CG10125-1	ACTCTACCAGAAGTACTACCAGTGG
CG10125-2	GTGGTACCTTCTCTCCTTGGGCCAC
CG4735-1	CTGAAAGAAGCGCTCAACATCAGCG
CG4735-2	CCTCTTGGCGATGTTTTTCCCACCG
CG10050-1	GTGAAAATGACGGATTTGTTCGCCG
CG10050-2	AACTGACAGAAACACACACCAACCC

NTI Suite 8 software package (InforMax, Inc., 1999) to assemble full-length cDNAs for each candidate gene. Candidates were translated in silico using the Vector NTI programs to identify start and stop codons.

3.2.6. Multiple sequence alignment and analysis of innexins

Amino acid sequences of seven members of the *D. melanogaster* innexin gene family were aligned with seven putative innexin genes from *An. gambiae* and six putative *Ae. aegypti* innexin genes, including the *zpg*-like candidate gene using the ClustalW algorithm in the Vector NTI Suite 8 (InforMax, Inc., 1999) software package. A phylogenetic tree was generated by the Vector NTI program using the nearest-neighbor joining method.

3.2.7. Reverse-transcriptase polymerase chain reaction (RT-PCR) developmental profile in *Ae. aegypti*

Total RNA was isolated from approximately 30 mg of *Ae. aegypti* larvae using the RNeasy Mini Kit (Qiagen), treated with DNase I with DNA-free (Ambion). The concentration and 260/280 nm absorption ratio of each RNA sample was determined using a spectrophotometer. RT-PCR reactions were performed as described in 3.2.3. Great care was taken to ensure that input RNA mass was identical for each developmental stage. Upon final treatment of the single-strand cDNA with RNase H, the DNA concentration was determined.

Polymerase chain reactions were assembled using precisely 100 ng 1st strand cDNA for each developmental stage. Assembled reactions also contained 10X *Taq* buffer, 1.5 µl 10 mM dNTPs, 0.2 µM each primer (Table 3) and 1 µl Advantage2 Taq Polymerase (BD Biosciences Clontech, Palo Alto, CA) and ddH_2O was added to a final

volume of 50 µl. Reaction conditions were as follows: 95 ^{o}C for 5', 20-35 cycles of 95^{o}C for 15", 60 ^{o}C for 15", 72 ^{o}C for 1'; final extension time was 5' at 72 ^{o}C. Resulting products were visualized on a 1% agarose gel. The constitutively expressed *Ae. aegypti S17* gene was used as a control.

3.2.8. Whole mount *in situ* hybridizations

Whole-mount *in situ* hybridizations were performed using digoxigenin-labelled probes essentially as described in previous literature (Juhn and James, 2006). Sense or anti-sense dig-labeled probes were transcribed with SP6 or T7 RNA polymerase. Runoff transcription reaction were performed using 1µg of linearized plasmid DNA in 20 µl reactions, supplementing 2 µl 10x digoxigenin RNA labeling mix (Roche Diagnostics, Indianapolis, IN) for unlabeled nucleotides, using the MAXIscript *In vitro* Transcription Kit (Ambion, Austin, TX). The plasmid DNA templates were from the gene fragments amplified to confirm expression in section 3.2.4. Reactions were incubated for 2 hours at 37^{o}C, verified on an agarose gel and precipitated with ethanol in the presence of 0.8M ammonium acetate. Embryos were collected at varying time points ranging from 1 hr. after oviposition to 20 hrs. post oviposition. *Aedes aegypti* embryos were washed in 1.3125% sodium hypochlorite for 30-35 seconds. Rinsing with ddH_20 commenced immediately for 2 minutes and the embryos were then placed in fresh ddH_20 before being transferred to heptane (Goltsev *et al.*, 2004).

Embryos were fixed in 9.25% formaldehyde (FA) solution (10mL 37% formaldehyde, 30mL ddH_20, and 10µl 10M NaOH) for 1 hour while gently mixing on a nutator. FA solution was removed and embryos rinsed twice in ddH_20, the second rinse lasting 30 minutes. Water was removed and embryos were submerged in boiling water

for 30 seconds before the hot water was then removed. Prechilled ddH_20 was added to the embryos and incubated on ice for 10 minutes. Water and remaining heptane were removed and fresh heptane was added. To crack the exochorion and vitelline membrane, methanol was added and the embryos allowed to settle for 15-20 minutes. After washing several times with fresh methanol, embryos were transferred to a Petri dish with double-sided stick tape. Once the embryos were allowed to settle and stick to the tape, ethanol and water were added before the embryos were coaxed out of the vitelline membrane with the aide of a fine-tipped paintbrush or 25 gauge needle. Freed embryos were rinsed 2-3 times in ethanol and carried directly to the hybridization steps or placed at -20 ^{o}C for storage.

Prehybridization, hybridization, and staining steps were carried out as described by Juhn and James (Juhn and James, 2006). Embryos were washed in 90% P-xylene/10% ethanol and rinsed in ethanol, methanol, and then rehydrated. Embryos were postfixed in 4% FA solution in PBT. After treatment with proteinase K, embryos were washed in PBT and postfixed in 4% FA solution before being transferred step-wise into hybridization solution (50% formamide, 5x SSC, 100 µg/mL salmon sperm DNA, 50 µg heparin and 1% Tween-20). After incubation of at least 1 hour, dig-labeled probe was added to a final concentration of 100ng/mL and incubated overnight at 55 ^{o}C. Embryos were subjected to a series of PBT washes, treated for 30 minutes in 20 µg/mL RNase A at 37 ^{o}C, washed again in PBT, and then incubated for 2 hours in blocking solution (5% goat serum, 1% Roche Western Blocking reagent (Roche Diagnostics, Indianapolis, IN), and PBT). Embryos were incubated at 4 ^{o}C overnight with anti-digoxigenin-AP Fab fragments (Roche Diagnostics, Indianapolis, IN). Embryos were washed in PBT

following antibody incubation then rinsed in alkaline phosphate buffer (50mM $MgCl_2$, 100mM NaCl, 100mM Tris pH 9.5, 0.01% Tween20). NBT/BCIP color substrates were used for probe detection. Embryos were then rinsed in PBT before being mounted in 70% glycerol.

3.3. Results

3.3.1. *Drosophila* candidate gene identification

A keyword search of the Berkeley Drosophila Genome Project database (http://www.fruitfly.org/cgi-bin/ex-insitu.pl) (Tomancak *et al.*, 2002) using the terms germ cell, gonad, pole cell, and pole plasm returned 197 candidate cDNA clones. A second series of searches using the terms, subset or localized, returned 214 candidates (Table 4). Additional analysis confirmed a subset of the results were returned by both searches. A second search removing all candidates with defined somatic expression was performed and the remaining 97 *D. melanogaster* candidates were subjected to visual screening of the digital *in situ* hybridization images. Eighteen genes (Table 4) showed mRNA expression specific to the developing germline. Additional genes were added due to previous knowledge of expression patterns described in literature.

3.3.2. Putative gemline-specific gene identification in *Ae. aegypti*

Genomic DNA database searches were performed to identify orthologous gene candidates in *Ae. aegypti*. Of the twenty-two probable germline-specific genes identified in *D. melanogaster*, fourteen were found to have significantly similar sequences in *Ae. aegypti* databases (Table 4). *Vasa* was not pursued further due to previous work (Section 2). Table 5 lists the thirteen remaining *Drosophila* genes and their respective sequence matches in the mosquito. Literature reviews reveal a number of the genes to have

biological functions pertaining to germ cell differentiation or pole determination. For example, the gene, CG10520 (tube), is known to be involved in zygotic dorsal/ventral axis determination (Galindo *et al.*, 1995). CR32885 (*polar granule component* (*pgc*)) is a non-coding RNA responsible for the regulation of transcriptional repression in *Drosophila* germ cells. Germ cells lacking *pgc* express genes responsible for differentiation of nearby somatic cells (Martinho *et al.*, 2004).

Table 4. Process used to identify germline-specific gene candidates. Putative germline-specific genes with unique germ tissue expression patterns were identified by searching the Berkeley Drosophila Genome Project (BDGP) Patterns of Gene Expression Database. After initial identification of embryos with germline staining, those showing expression in somatic tissues were removed. Remaining *Drosophila* genes were queried against *Aedes aegypti* EST and genome sequence databases. *In vivo* expression was confirmed using RT-PCR techniques.

Identification of *Drosophila melanogaster* germline-specific candidate genes and putative mosquito orthologs		
Screening Criteria	Genes Meeting Criteria	Candidate Genes
BDGP expression database: germ cell, gonad, pole cell, and pole plasm/BDGP search terms subset or localized	197/214	Not listed
Manual removal of all candidates with somatic expression	97	Not listed
Visual screening of *in situ* images	18	CG10520, CG9925, CG10125, CG17658, CG6072, CG1034, CG5292, CG9699, CG15737, CR32885, CG4735, CG6122(*piwi*), CG10050, CG3273, CG12218, CG4755, CG7097, CG8418
Addition of genes from literature	—	*vasa*, *maelstrom*, *tudor*, *aubergine*, *piwi*
Sequence similarity search for mosquito ortholog candidates	14	CG10520, CG10125, CG17658, CG6072, CG5292, CG15737, CR32885, CG4735, CG6122(*piwi*), CG10050, *vasa*, *mael*, *tud*, *aub*
Expression confirmed in *Ae. aegypti*	10	CG10520, CG10125, CG5292, CG4735, CG6122(*piwi*), CG10050, *vasa*, *mael*, *tud*, *aub*
RT-PCR/Northern blot expression analysis	4	CG10125, CG4735, CG10050, *vasa*
		Aubergine and piwi were eliminated because high sequence similarity between these genes made cloning and sequencing difficult

piwi and *aubergine*, both germline-specific gene and members of the Argonaute family, may have roles in RNA silencing mechanisms in the germ cells (Saito *et al.*, 2006, Vagin *et al.*, 2006).

Table 5. *Drosophila melanogaster* germline-specific candidate genes. Candidate genes were queried against the *Aedes aegypti* genomic DNA database (NCBI).

Drosophila melanogaster* germline-specific gene candidates and similar sequences found in the genome of *Ae. aegypti			
***D. melanogaster* gene candidate**	***Ae. aegypti* genomic DNA search result***	**e-value/score****	**D. melanogaster citations*****
CG10520(*tub*)	AAGE02012852	1e-19/80.3	Galindo et al., 1995
CG17658	AAGE02018042	1e-3/40.9	NA
CG10125(*zpg*)	AAGE02030575	4e-52/149	Tazuke et al., 2002
CG6072(*sra*)	AAGE02031625	5e-83/309	Ejima et al., 2004
CG5292	AAGE02025184	2e-46/186	NA
CG15737	AAGE02010183	1e-57/228	NA
CR32885(*pgc*)	AAGE02005726	2e-43/179	Martinho et al., 2004
CG4735(*shu*)	AAGE02021134	1e-70/269	Munn & Steward, 2000
CG6122(*piwi*)	AAGE02010239	0.0/353	Cox et al., 1998
CG10050	AAGE02004274	2e-53/210	NA
mael	AAGE02001407	6e-24/114	Clegg et al., 1997
tud	AAGE02013357	2e-68/264	Bardsley et al., 1993
aub	AAGE02010239 AAGE02025028	0.0/362 0.0/355	Harris & McDonald, 2001

* Accession number of *Ae. aegypti* genomic DNA contig best hit match.
** E-value and bit score from tBLASTn search of *Ae. aegypti* whole genome sequence database
*** *Drosophila melanogaster* candidate gene references provided for further detail

3.3.3. Confirmation of gene expression in *Ae. aegypti* as a result of EST database search

As a result of the identification of putative gene orthologs using the *Ae. aegypti* genomic database, a search was conducting to determine whether the candidate gene is known to be expressed, by using the *Ae. aegypti* Expressed Sequence Tag (EST) database housed at TIGR (http://www.tigr.org). Ten of the candidate *Aedes* genes matched sequences from the EST database. Again, *vasa* was removed from further study (Section 2). Table 6 lists the nine candidate genes with confirmed expression in *Ae. aegypti*. Although the data was originally collected using the TIGR Gene Indices database, it is presented here using the NCBI EST database accession numbers. TIGR has released the data to NCBI where the information can be readily obtained. Visual assessment of sequence similarities and alignments for CG10520 led to the conclusion that pursuit of this gene would not be beneficial. High sequence similarity between the proteins *tub* and *pelle* made it difficult to ascertain which gene was a best match to the sequence isolated from genomic and EST databases. Similarly, *piwi* and *aub* are highly related genes with a high degree of sequence similarity. As illustrated in Table 5 and Table 6, sequence similarity searches using *piwi* and *aub* amino acid sequences as a query ultimately presented us with the same *Ae. aegypti* sequence results as the most significant match (genomic DNA- AAGE02010239, EST DNA- DV251570). At the time of initial analysis, incomplete genomic DNA sequences and partial EST fragments yielded ambiguous results. As a result, *tub*, *aub*, and *piwi* have been omitted from subsequent studies.

3.3.4. Using RT-PCR as confirmation of germline-specific candidate gene expression in *Ae. aegypti*

PCR amplification of candidate genes using single strand cDNA as a template revealed that six genes identified *in silico* were present in the transcriptome of 4th instar larvae.

Table 6. Identification of putative endogenous germline-specific genes in *Ae. aegypti*. EST sequence databases assembled by TIGR were queried using hypothetical translations of D. melanogaster candidate genes and best hit matches from the *Aedes* genomic sequence database search. Nine putative genes had significant sequence similarity to known EST sequencing, confirming expression.

Expression confirmation using EST database sequence similarity searches			
***D. melanogaster* gene candidate**	***Ae. aegypti* genomic DNA search result***	***Ae. aegypti* EST search result****	**e-value/score*****
CG10520(*tub*)	AAGE02012852	DV325890	1e-58/207
CG10125(*zpg*)	AAGE02030575	DV390927	5e-44/158
CG5292	AAGE02025184	DV290675	2e-48/189
CG4735(*shu*)	AAGE02021134	DW209593	5e-49/175
CG6122(*piwi*)	AAGE02010239	DV251570	2e-83/290
CG10050	AAGE02004274	DV293342	7e-83/305
mael	AAGE02001407	DV352557	5e-20/97.4
tud	AAGE02013357	DV373352	8e-38/159
aub	AAGE02010239 AAGE02025028	DV251570 DW201793	2e-87/322 2e-85/315

* Accession number of *Ae. aegypti* genomic DNA contig best hit match
** Accession number of *Ae. aegypti* EST sequences found in NCBI EST database, current as of May 2007. TIGR is shifting their future focus and has reported EST sequences to NCBI where they will be maintained.
*** E-value and bit score from tBLASTn search of *Ae. aegypti* expressed sequence tag database

DNA sequencing of the amplification products confirmed that the correct gene had been amplified. Larval RNA expression was confirmed for the *Aedes* putative orthologs; *Aezpg*, *Aeshu*, *AeEwald*, *Aemael*, and *Aetud* (Figure 9). We were not able to confirm gene expression for CG5292.

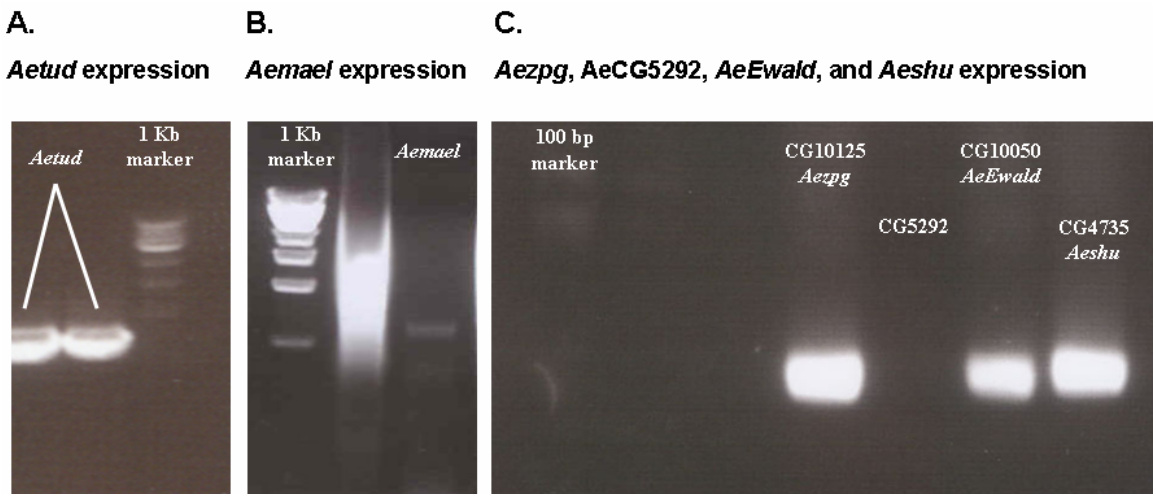


Figure 9. Amplification of putative germline-specific gene expression from *Ae. aegypti* 4th instar larvae. RT-PCR yielded a 1.8 Kb amplified gene fragment of *Aetud* (A), a 1.2 Kb product of *Aemael* (B), and an approximate 200 bp gene fragment from the genes; *Aezpg*, *AeEwald*, and *Aeshu* (C). Failure to amplify AeCG5292 eliminated the gene from further study.

3.3.5. Cloning of full-length cDNA from the *Ae. aegypti maelstrom* gene

RT-PCR amplification of larval cDNA using gene specific primers amplified a gene fragment of 1.1 Kb in length. Successful amplification confirmed *Aemael* expression in the mosquito. Subsequent 5' and 3' RACE (rapid amplification of cDNA ends) using *Ae. aegypti* larval cDNA as template yielded a probable full-length cDNA of 1919 bp in length with a 1560 bp coding region. This is translated into a protein of 520 amino acids. Figure 10 shows the putative protein has 27% sequence identity and 49% similarity when aligned with the *D. melanogaster maelstrom* amino acid sequence using the ClustalW algorithm (Vector NTI Suite 8, Informax, Inc., 1999). *Aemael* has a 76 bp 5' untranslated region (UTR) and a 264 3' UTR. *in silico* analysis suggests that *Aemael* is a single copy gene within the *Ae. aegypti* genome. Genomic DNA data shows *Aemael* to have three introns and four exons spanning roughly 6.4 Kb.

Additional *in silico* analysis identified a putative ortholog in the mosquito *An. gamibae* that is 33.9% and 48.5% identical and similar in sequence, respectively.

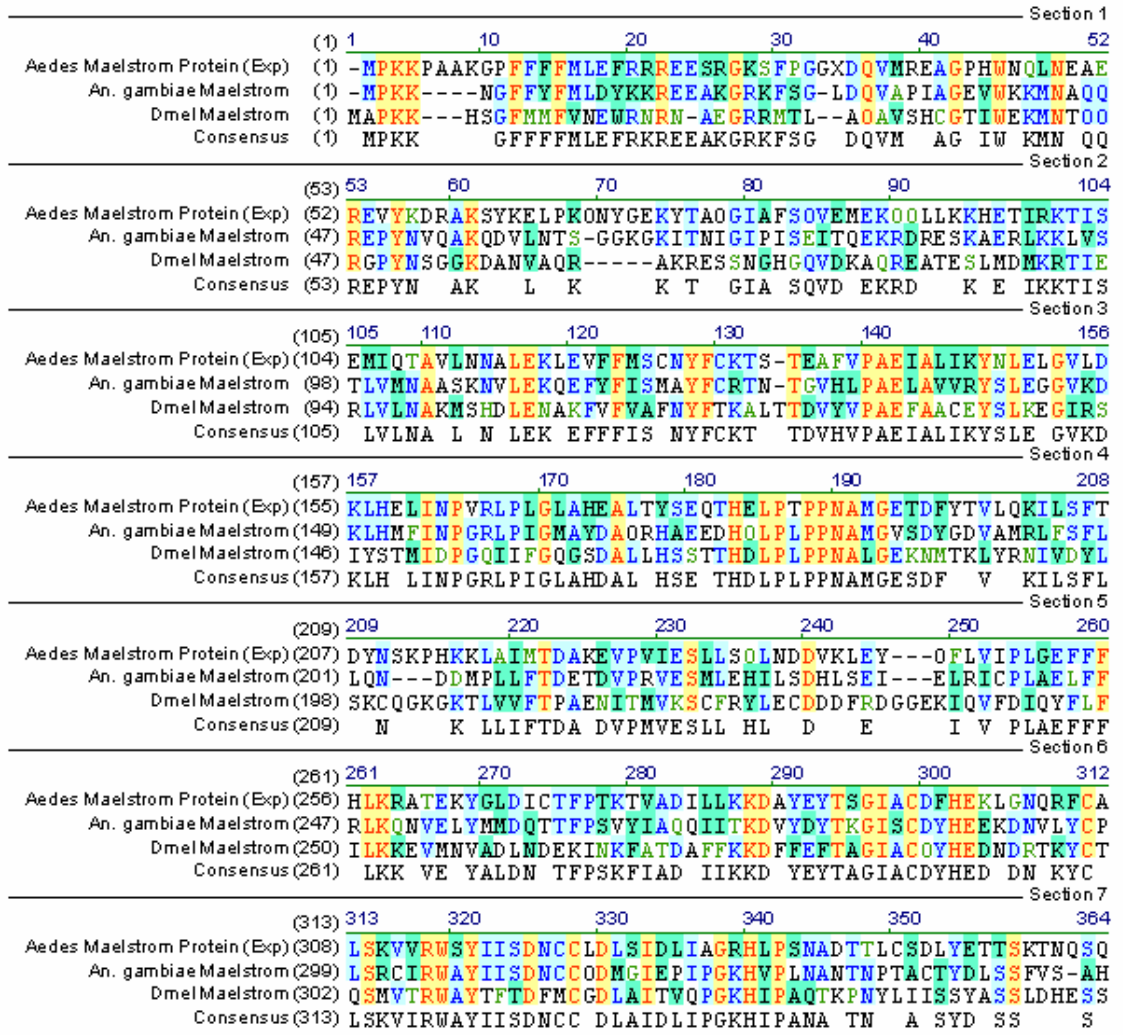


Figure 10. ClustalW alignment of full-length, translated *maelstrom* sequence for *Ae. aegpyti*, *D. melanogaster*, and *An. gambiae*. The A*n. gambiae* protein is a conceptual translation of gene, ENSANGP00000011239.

3.3.6. *Aemael* is expressed in all mosquito lifestages

Reverse-transcriptase PCR (RT-PCR) indicates *Aemael* is expressed in all mosquito life-stages (Figure 11). While not a quantitative analysis, there appears to be noticeably higher levels of mRNA present in early stage embryos and dissected ovaries from bloodfed adult females. Expression levels in female carcasses may be a result of incomplete ovarian dissection or may indicate some level of somatic expression.

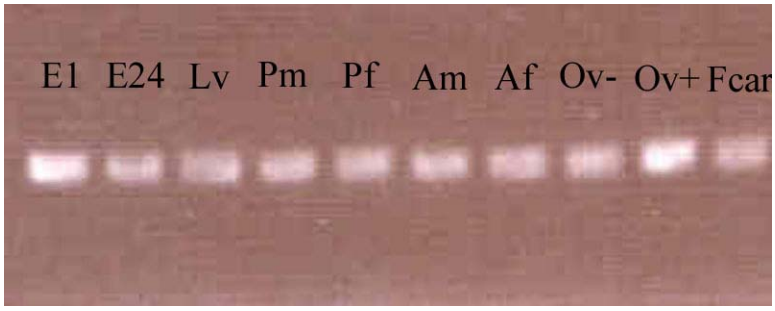


Figure 11. RT-PCR expression profile of *Aemael* reveals mRNA expression in all life-stages and both sexes. Transcript is present in ovarian tissues as well as potentially in somatic tissues as well. Aemael is upregulated in embryos 1 hour post oviposition (E1) and ovaries dissected from bloodfed females (Ov+).

3.3.7. Cloning of *Ae. aegypti shutdown* full-length cDNA

Using primers based on genomic DNA and partial cDNA sequences, a fragment of *Aeshu* was amplified that was approximately 200 bp in length. Standard 5' and 3' RACE resulted in a putative full-length cDNA sequence of 1561 bp in length. *Aeshu* is composed of three exons spread out over 1.6 Kb in the genome. The structure of *Aeshu* is similar to the *D. melanogaster* shutdown gene which also has three exons. A coding sequence of 1.326 Kb is translated into a 442 amino acid protein with 36.1% identity and 47.4% sequence similarity to the shutdown protein in *D. melanogaster* when aligned (Figure 12). The translated polypeptide retains a peptidyl-prolyl cis-trans isomerase

(PPIASE) domain and a tetraticopeptide repeat (TPR) domain previously described for *shu* in *Drosophila* (Munn and Steward, 2000). PPIASE has been shown to catalyze the cis-trans isomerization of the peptide bond or praline residues, altering protein folding (Fischer and Schmid, 1990). The TPR domain is believed to mediate protein-protein interactions (Lamb *et al.*, 1995).

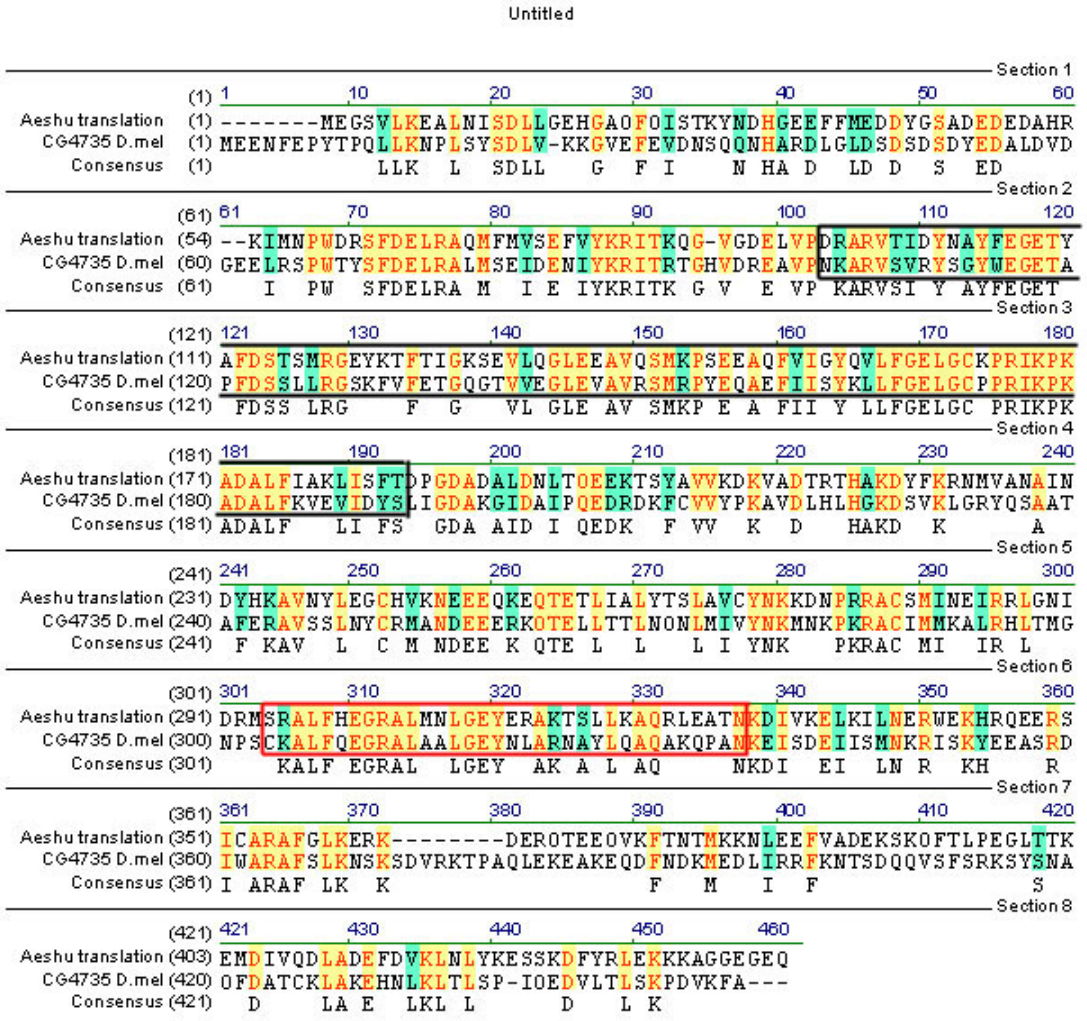


Figure 12. ClustalW alignment of full-length, translated *shutdown* sequence for *Ae. aegpyti* and *D. melanogaster*. The conserved PPIASE domain is identified by the black box and the TPR domain is identified by the red box.

3.3.8. *Aeshu* is expressed in all life-stages and is upregulated in early embryos and the ovaries of blood-fed females

Analysis of RT-PCR gene expression over multiple mosquito life stages reveals that *Aeshu* is highly upregulated in ovaries dissected from bloodfed females and 1 hour old embryos with a reduction in transcript level in embryos 24 hours post deposition (Figure 13). Low expression levels are observed in female pupae, female adults, and sugar-fed dissected ovaries, a likely result of basal expression levels. Additionally, low level expression of *Aeshu* appears in female carcasses where the ovaries were previously dissected. This is likely a result of incomplete ovarian dissection and the high sensitivity of RT-PCR.

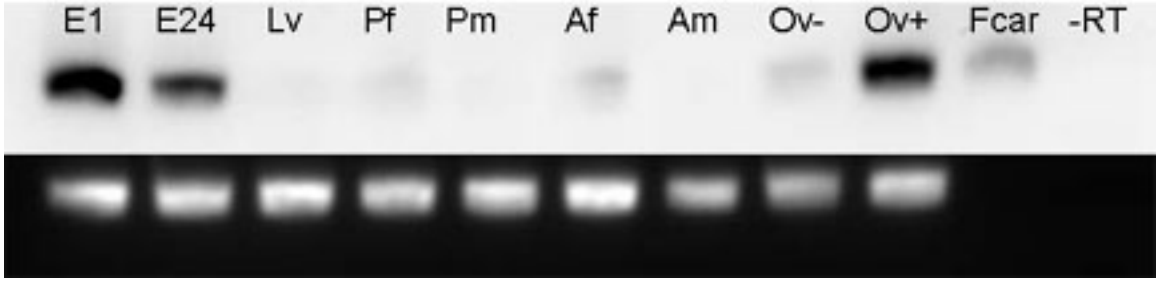


Figure 13. Inverted image of RT-PCR expression profile of *Ae. aegypti* life-stages. *Aeshu* is highly expressed in 1 hour old embryos (E1) and bloodfed dissected ovaries (Ov^{+}). Lower levels of expression are visible in female pupae and adults as well as dissected ovaries from sugar females. Expression in female carcasses (sugar fed females with ovaries dissected out) is likely a result of incomplete dissection and experimental sensitivity. The constitutively expressed S17 gene serves as a control as shown in the lower panel.

3.3.9. *Aeshu* mRNA localizes to the posterior end of early stage embryos

In situ hybridization of digoxigenin-labeled antisense RNA to whole mount embryos reveals that *Aeshu* mRNA localizes to the posterior pole of early stage *Ae. aegypti* embryos (mitotic divisions 5-8) (Figure 14). A small amount of *Aeshu* mRNA appears to also localize at the anterior pole of early embryos. Similar unexpected anterior localization has been reported previously for the *An. gambiae* germline-specific *nos* gene (Goltsev *et al.*, 2004).

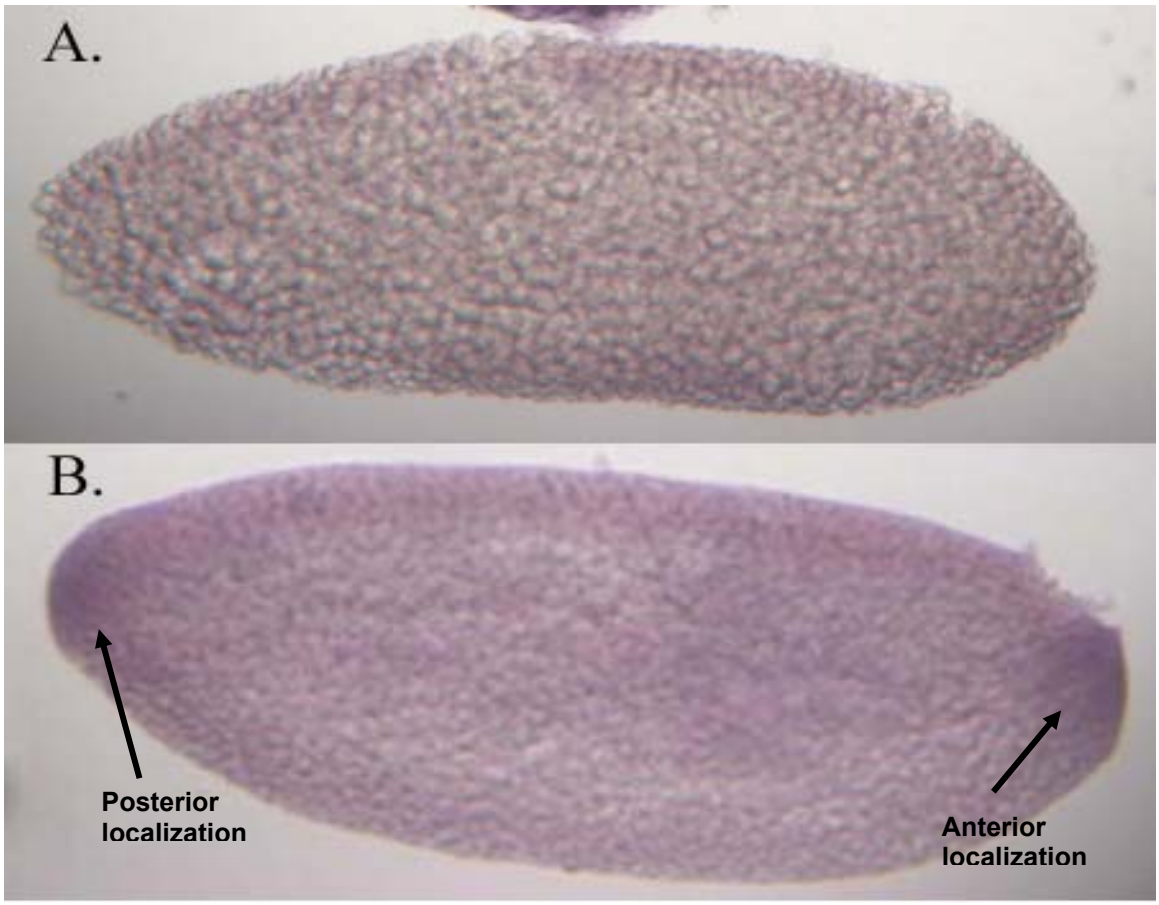


Figure 14. *In situ* hybridization of *Aedes aegypti* embryos using digoxigenin-labelled sense and anti-sense RNA. Posterior poles are oriented to the left side of the image. A-B) Early stage syncytial blastoderm embryos. A) Negative control of syncitial blastoderm embryo labeled with sense-strand RNA. B) *Aeshu* transcript localization is found in the posterior pole plasm region as well and the anterior pole.

3.3.10. Cloning of *Ae. aegypti zero population growth* (*zpg*) full-length cDNA

Primers based on partial cDNA sequences were used to amplify a fragment of *Aezpg* that was approximately 200 bp in length. Standard 5' and 3' RACE resulted in a full-length cDNA clone with a sequence of 1542 bp in length. *Aezpg* is composed of three exons encompassing 2.95 Kb of the *Ae. aegypti* genome. The structure of *Aezpg* is similar to the *D. melanogaster* zero population growth gene which also has three exons. A coding sequence of 1.167 Kb is translated into a 389 amino acid protein with 32% identity and 46.8% sequence similarity to the *zpg* protein in *D. melanogaster* when aligned (Figure 15). However, when the translated protein product from *Aezpg* is used as a query to search known *D. melanogaster* proteins, the resulting best match is the *Drosophila innexin2* gene (CG4735). Drosophila *zpg* and *innexin2* are highly related members of the innexin family of proteins. Global alignment using ClustalW and phylogenetic analysis using the nearest neighbor joining method of six *in silico* predicted innexin family members from *Ae. aegypti*, seven predicted members from *An. gambiae*, and seven known innexin family members from *D. melanogaster* reveals that *Aezpg* is more closely related to *D. melanogaster zpg* than to *D. melanogaster* innexin2 (CG4735). This is illustrated in figure 16.

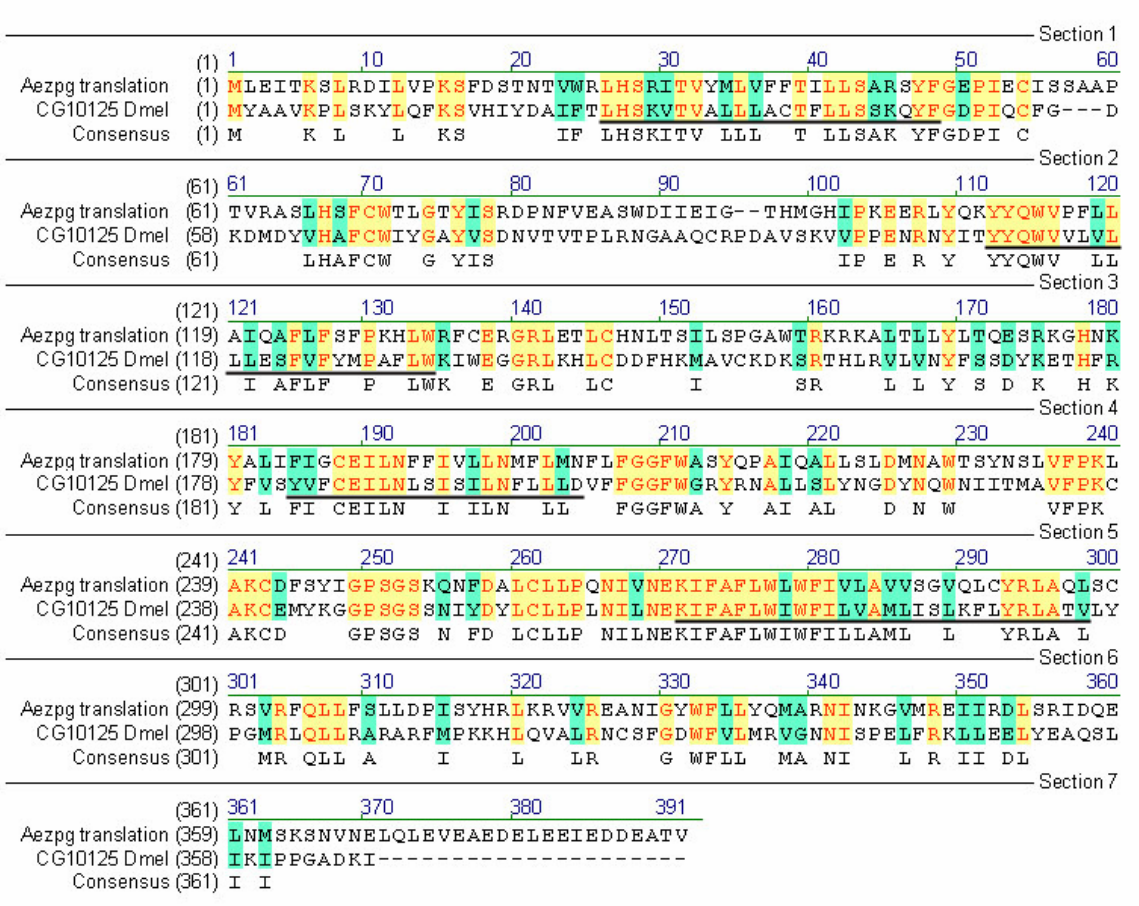


Figure 15. ClustalW alignment of full-length, translated zero population growth sequence for *Ae. aegpyti* and *D. melanogaster*. The four transmembrane regions predicted in *zpg* (underlined regions) show significant conservation.

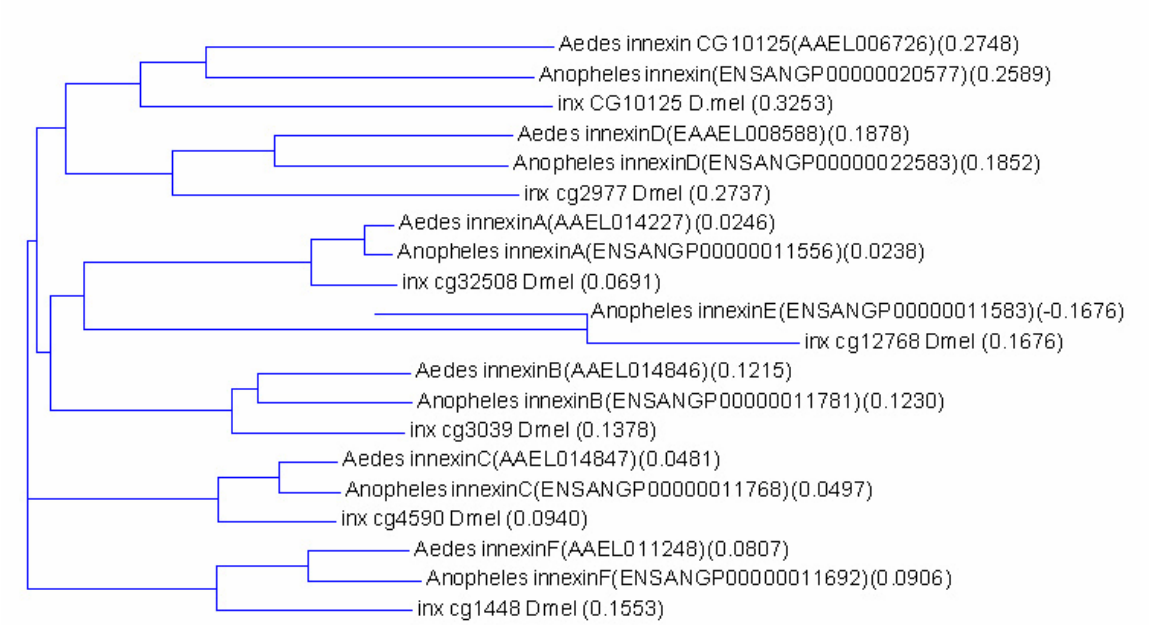


Figure 16. Phylogenetic tree shows that *Aedes* innexin CG10125 (*Aezpg*) is most closely related to *Drosophila* innexin CG10125 (*zpg*). Nearest neighbor joining method was employed to determine relatedness of predicted *Ae. aegypti* innexin gene to innexin family members known in *D. melanogaster* and predicted in *An. gambiae*. *An. gambiae* gene name follows innexin notion in parenthesis.

3.3.11. *Aezpg* is expressed in all life-stages and is upregulated in early embryos and the ovaries of blood-fed females

RT-PCR gene expression analysis reveals that *Aezpg* is likely expressed in all mosquito life stages and is highly upregulated in ovaries dissected from bloodfed females and 1 hour old embryos, with a drastic reduction in the level of transcript in embryos 24 hours post egg deposition (Figure 17). Basal levels of expression are visible in female adults and sugar-fed dissected ovaries. This expression profile is consist with other putative germline-specific genes. Additionally, low expression of *Aezpg* appears in female carcasses where the ovaries were previously extracted. This is likely a result of incomplete ovarian dissection and the high sensitivity of RT-PCR.

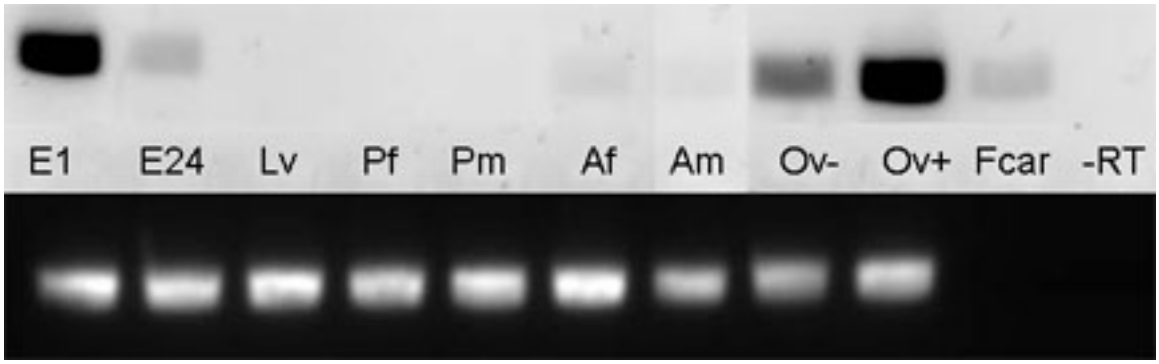


Figure 17. Inverted image of RT-PCR expression profile of *Aezpg* in *Ae. aegypti* life-stages. *Aezpg* is highly expressed in 1 hour old embryos (E1) and bloodfed dissected ovaries (Ov^{+}). The transcript appears to be rapidly degraded as it is less prominent in embryos 24 hours post oviposition. Lower levels of expression are visible in female adults as well as dissected ovaries from sugar females. Expression in female carcasses (sugar fed females with ovaries dissected out) is likely a result of incomplete dissection and experimental sensitivity. The constitutively expressed S17 gene serves as a control as shown in the lower panel.

3.3.12. *Aezpg* mRNA localizes to the posterior end of early stage embryos and presumptive germ tissue of gastrula stage embryos

In situ hybridization of digoxigenin-labeled antisense RNA to whole mount embryos reveals *Aezpg* mRNA localized to the posterior pole of early stage *Ae. aegypti* embryos (mitotic divisions 5-8) and presumptive germ tissue of gastrula stage embryos (mitotic divisions 13-15) (Figure 18). This expression pattern is consistent with *D. melanogaster zpg* expression. The expression pattern revealed by *in situ* hybridization is unique from the expected pattern of *D. melanogaster* innexin2 gene (CG4735) providing further support of our conclusion, that *Aezpg* is the *Ae. aegypti* ortholog of the *D. melanogaster zero population growth* gene.

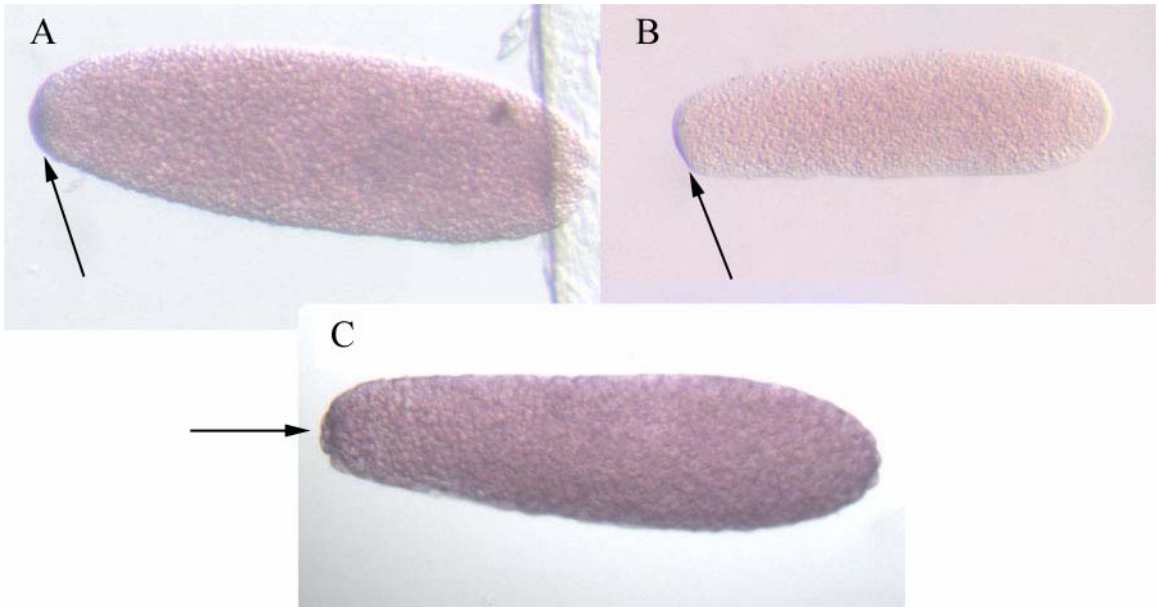


Figure 18. *In situ* hybridization of *Aedes aegypti* embryos using digoxigenin-labelled anti-sense RNA. Posterior poles are oriented to the left side of the image. A-B) Presynctial blastoderm shows transcript localized to the pole plasm. C) Early stage syncytial blastoderm embryo corresponding to mitotic division 11-12. *Aezpg* transcript localization is found in the budding pole cells in the posterior pole. The budding pole cells are shown at the tip of the arrow.

3.3.13. Cloning of *Ae. aegypti* CG10050 (*AeEwald*) full-length cDNA

Using oligonucleotide primers based on *Ae. aegypti* cDNA sequences from EST projects we amplified gene fragments with sequence similarity to CG10050, an uncharacterized gene in the *D. melanogaster* genome. Utilizing standard 5' and 3' RACE techniques we successfully cloned and sequenced the full-length cDNA for the *Ae. aegypti* ortholog (*AeEwald*) of *D. melanogaster* CG10050. The full-length cDNA is 1.153 Kb in length with a 170 bp 5' UTR, a 786 bp CDS, and a 2.5 Kb 3' UTR. The putative protein is a 262 amino acid polypeptide with 55.7% sequence identity and 67.4% sequence similarity to D. melanogaster CG10050 (Figure 19). This is a high degree of conservation despite 250 million years of divergence (Gaunt and Miles, 2002) which suggests the protein is involved in a critical role during development.

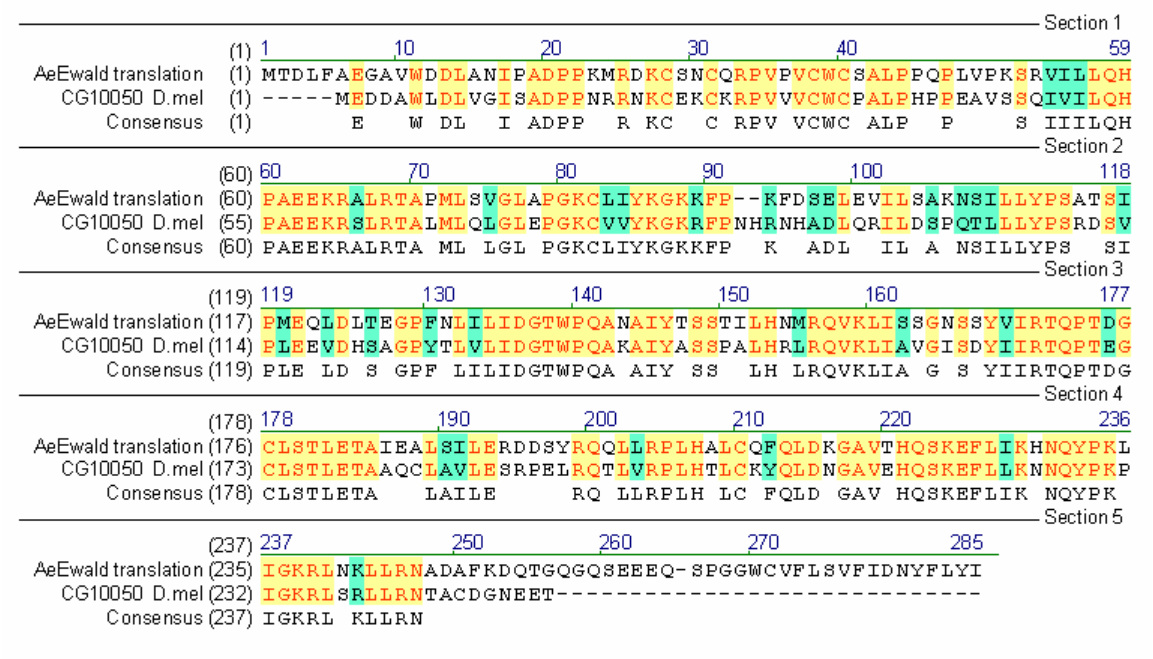


Figure 19. ClustalW alignment of the full-length *AeEwald* and *D. melanogaster* CG10050. The putative polypeptides share 55.7% sequence identity.

3.3.14. *AeEwald* is a germline-specific gene and is upregulated in early embryos and the ovaries of blood-fed females

RT-PCR gene expression reveals *AeEwald* is expressed in the presumptive developing germ tissue and ovaries of female mosquitoes. *AeEwald* is upregulated in ovaries dissected from bloodfed females and 1 hour old embryos. Embryos aged 24 hours after oviposition show little or no mRNA transcript (Figure 20). Low levels of expression are visible in female adults and sugar-fed dissected ovaries. The expression profile is consistent with other putative germline-specific genes. *AeEwald* is a tightly regulated germline-specific gene with minimal basal expression. The transcriptional

profile data suggests that *AeEwald* may be a maternally expressed gene that is transported into the egg or early embryo.

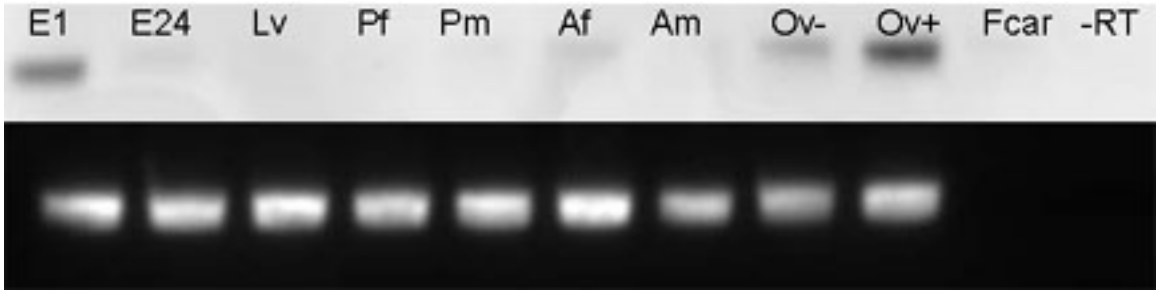


Figure 20. Inverted image of RT-PCR expression profile of *Ae. aegypti* life stages. *AeEwald* is a germline-specific and shows higher expression in 1 hour old embryos (E1) and bloodfed dissected ovaries (Ov^{+}). Lower levels of basal expression are visible in female adults as well as dissected ovaries from sugar females. The constitutively expressed S17 gene serves as a control as shown in the lower panel.

3.3.15. *AeEwald* mRNA localizes to the posterior end of early stage embryos

In situ hybridization of digoxigenin-labeled antisense RNA to whole mount embryos reveals that *AeEwald* mRNA localizes to the posterior pole of early stage *Ae. aegypti* embryos (mitotic divisions 3-4) (Figure 21). mRNA expression or localization has not been identified in more developed embryos, consistent with findings in *D. melanogaster* of CG10050 showing expression in early stage embryos (Tomancak *et al.*, 2002). It should also be noted that in situ hybridization data analyzing gene expression in Drosophila ovaries does not exist.

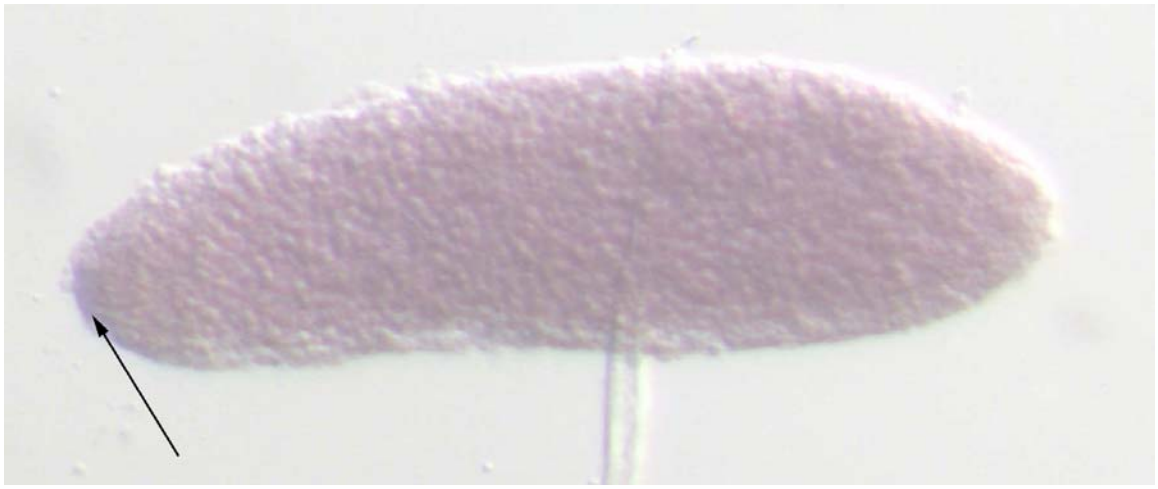


Figure 21. *In situ* hybridization of *Aedes aegypti* embryos using digoxigenin-labelled anti-sense RNA. Posterior poles are oriented to the left side of the image. Presynctial blastoderm shows transcript localizes to the pole plasm.

3.4. Discussion

3.4.1. *Aemael* is the mosquito ortholog of *D. melanogaster maelstrom*

Maelstrom is known to localize to the nuage of developing embryos in association with the Vasa protein (Findley *et al.*, 2003). Maelstrom protein is also shown to localize to the nuage in mouse, thereby implying a conserved role in the germ line throughout evolution (Costa *et al.*, 2006). More recently it has been suggested that *maelstrom* may be involved in RNAi, or micro-RNA, suppression mechanisms. The local alignment algorithm, BLAST, and the global alignment algorithm ClustalW reveal that *Aemael* is the mosquito ortholog of the *D. melanogaster maelstrom* gene. Using a reciprocal-best-BLAST-hits (RBH) approach (Altschul *et al.*, 1990), with *Aemael* as a query sequence against the *D. melanogaster* protein database, *mael* is returned as the most significant sequence match. Polypeptide sequence alignments show 27% identity and 49% sequence similarity, significant conservation of amino acid sequences after 250 million years of divergence. Initial analysis of *Drosophila in situ* hybridization images in the BDGP suggested that the mRNA is localized exclusively to the pole plasm and developing

germline. Subsequent evaluation of additional images, however, revealed somatic expression. As a result, we have chosen not to pursue *Aemael* as a germline-specific gene candidate and will not pursue the isolation of cis-acting DNA.

Further *Aemael* gene characterization may prove beneficial for future germ cell development research. In *Drosophila*, *mael* is known to localize with *vasa* in the nuage of developing embryos and *mael* mutants fail to develop properly. Indications of potential involvement in gene regulation via RNAi or micro-RNA pathways make *Aemael* a critical player in the developing germline.

3.4.2. *Aeshu* is a germline-specific gene and an ortholog of *D. melanogaster shutdown*

Shutdown, originally identified in *D. melanogaster*, is essential for the normal function of the germline stem cells. Loss-of-function alleles revealed that *shutdown* is also required at later stages of oogenesis. In *Drosophila*, *shutdown* functions autonomously in the germline. The RNA and protein are strongly expressed in the germline stem cells and the RNA is also present in the germ cells throughout embryogenesis (Munn and Steward, 2000). Identification of a novel full-length cDNA in *Ae. aegypti* using RT-PCR revealed a putative translated polypeptide with 36% sequence identity and conserved PPIASE and TPR amino domains. *Shu* and *Aeshu* proteins show significant homology to immunophilins, an evolutionarily conserved class of proteins containing PPIASE and TPR domains. While their actual function is unclear, related proteins have been proposed to regulate translation by modulating the binding of translation factors to the 5' end of the mRNA. Interestingly, *AaeVLG* encodes a germline-specific homolog of the translation initiation factor eIF4A. It is plausible that *Aeshu* and *AaeVLG* interact to regulate gene expression within the developing germline.

The expression pattern for *Aeshu* is consistent for genes that function in germline stem cells. Localization of mRNA by *in situ* hybridization reveals transcript at the posterior pole and developing germline. This is consistent with *shu* expression in developing *Drosophila* embryos.

The similarity of putative polypeptide sequences suggest that *Aeshu*, a homolog of the *D. melanogaster* gene, *shutdown* has been successfully cloned. While *shu* has not been well characterized, two conserved domains have previously been identified. *Aeshu* has similar domains, indicating conserved function. The mRNA expression pattern of *Aeshu* is similar to that of *shu*, showing localization to the posterior pole of early stage embryos. It is in the posterior pole that the developing germ cells originate. The collective data suggests that *Aeshu* is indeed the ortholog *shutdown*.

3.4.3. *Aezpg* is a member of the innexin family of genes and is an ortholog of *zero population growth*, a germline-specific *Drosophila melanogaster* gene

The gap junction protein *zero population growth* is required for an early step of both oogenesis and spermatogenesis. Typical of gap junction proteins, *zpg* has four transmembrane proteins. Sequence similarity searches utilizing the BLAST function revealed *Aezpg* is similar to *Drosophila innexin2*. However, Global alignment of full-length polypepetide sequences and phylogenetic analysis indicates that *Aezpg* is more similar to *zpg* than it is to innexin2. This suggests that *Aezpg* is an ortholog of the gap junction protein *zpg*. *In situ* hybridization analysis also supports the sequence alignment data. According to the BDGP the *Drosophila innexin2* gene is expressed ubiquitously early in development, an expression pattern unique from that of *zpg* which is localized to

the presumptive germ tissue. *Aezpg* was found to have an expression pattern similar to that of *zpg*.

3.4.4. The uncharacterized *D. melanogaster* gene, CG10050, is similar to *AeEwald*

CG10050 is an uncharacterized gene found in the germline of *D. melanogaster* embryos. It appears to be expressed very early in development and the transcript quickly degrades. This gene may be crucial to proper development and differentiation in ovarian or perhaps oocyte development. The gene identified as *AeEwald* shows remarkable sequence conservation. Given the high level of sequence conservation after 250 million years of divergence, CG10050 and *AeEwald* likely have critical roles in oogenesis and mutation is heavily selected against.

The *AeEwald* promoter is an ideal candidate for germline-specific transposase expression. *AeEwald* appears to have low levels of transcript that are isolated at the pole plasm of developing embryos. Very precise spatio-temporal isolation of transposase expression would be beneficial to driving transgene spread without a significant reduction in fitness. *AeEwald* appears to be expressed in bloodfed ovaries and early stage embryos with much lower levels of basal expression in female mosquitoes. It is plausible that *AeEwald* is a maternal gene expressed in ovaries and developing ovarioles. While mRNA is present in the early embryos, it may have originated in the nurse cells and not as a result of embryonic expression.

3.5. Conclusions

It is not likely that mosquitoes will receive the biological research attention afforded to *D. melanogaster*. Therefore, approaches to basic research must be devised to utilize existing data from model organisms in order to infer gene presence and/or function

in the mosquitoes. A novel approach to identify putative gene orthologs across distantly related species proved useful. Utilizing genomic and expression data from a well studied model organism, homologous genes can be found in other species. The method employed here will continue to be of use as an increasing number of non-model organisms, specifically insects, are investigated. Gene discovery and characterization of function within a model insect will at the very least provide preliminary insight into the role and function of an orthologous candidate in other insects.

Worth noting in this experiment is the fact that only maternal and very early embryonic gene expression was assessed. The BDGP database does not include expression data for larvae through adult life stages. It is unknown if expression patterns found in early embryos will resemble the expression patterns in maturing individuals.

Using existing data from the dipteran, *D. melanogaster*, three putative germline-specific genes in the mosquito, *Ae. aegypti* were cloned. Computational and *in situ* hybridization data presented here suggests that each of the genes are orthologous to germline-specific genes in *D. melanogaster*. *Aeshu*, *Aezpg*, and *AeEwald* are all upregulated and highly expressed in early embryos as well as ovaries dissected from females post bloodfeeding. While the roles of *zpg* and *shu* in the establishment and development of germ cells have been described in some detail, the function of the highly conserved gene, CG10050, is unknown. The function of *Aezpg* and *Aeshu* in mosquitoes can be inferred based upon *D. melanogaster* data. The role of *AeEwald* in the establishment of germ cells will require functional studies employing both forward and reverse genetics.

4. UTILIZING ENDOGENOUS *CIS*-ACTING DNA TO DRIVE TRANSGENE EXPRESSION IN THE MOSQUITO GERMLINE

4.1. Introduction

The current genetic transformation efficiency in *Ae. aegypti* is 5-10 fold lower than that of *D. melanogaster* transformation (Handler, 2000). The binary transformation system uses a helper plasmid to provide transposase. These protocols typically utilize a transposase gene cloned under the control of a *Drosophila* heat shock promoter or a constitutive promoter. Fly promoters do not always display the expression characteristics of a heat inducible promoter in mosquitoes and can exhibit high basal activity (Morris *et al.*, 1991). Constitutive promoters many not contain regulatory elements that are well enough conserved to function in mosquitoes.

Previous studies in both *Aedes* and *Anopheles* mosquitoes have proven the utility and functionality of using tissue-specific promoters to drive transgenes within the mosquito (Arca *et al.*, 1999, Coates *et al.*, 1999, Kokoza *et al.*, 2000, Moreira *et al.*, 2000). In brief, the 5' *cis* acting regulatory region from known tissue-specific genes was fused to a cDNA reporter gene. This new synthetic construct was then placed into a transposable element and then injected into embryos, transgenics lines obtained, and promoter function was analyzed via reporter gene expression.

Replacing an exogenous promoter on the helper plasmid with an endogenous mosquito promoter may provide consistent, high level transposase expression within the mosquito. Additionally, the successful stable integration of endogenous promoters driving transposase expression may lead to the generation of a helper strain of mosquitoes requiring only the injection of donor plasmid, as opposed to the current

method of dual plasmid injection (Atkinson and James, 2002). The tight confinement of transposase expression in a spatio-temporal manner will allow the study of transgene integration and remobilization while substantially reducing the number of variables involved (Adelman *et al.*, 2002).

Moving transgenes through a naïve population will necessitate the use of a gene drive system. As stated previously (Section 1) TEs may provide a mechanism for gene drive. However, the minimization of risks associated with the potential horizontal transfer of transgenes will be an important consideration. Using species-specific endogenous regulatory DNA may mitigate the risks associated with horizontal transfer.

The use of transposons as gene drive mechanisms requires their adaptation to regulate the timing and level of transposition such that transposons are spread rapidly within the target species while minimizing detrimental fitness costs. Transposon mobilization through replicative transposition would lead to an increase in transposon number. If transposon mobilization occurs via replicative transposition in the germline, the frequency of progeny inheriting the transgene will increase in each subsequent generation (Atkinson and James, 2002). This increases our interest in genes whose control sequences will permit the expression of an exogenous gene in the developing germline. The identification of these genes and their subsequent control sequences will permit the use of this *cis*-acting DNA to drive transposase expression in the developing germarium of *Ae. aegypti* mosquitoes.

Genes believed to be expressed exclusively in the developing germ tissue of the early embryo have previously been identified. The upstream and downstream *cis*-acting regulatory DNA elements have been cloned in attempts to drive expression of reporter

genes and transposase. Novel plasmids have been constructed with dual reporter genes to follow protein expression and localization as a result of germline-specific promoter function. The plasmids utilize either the *Mos1 mariner* transposable element or the φC31 integrase system (Nimmo *et al.*, 2006). Additionally, novel plamsids containing autonomous *Mos1* transposable elements have been generated.

4.2. Materials and Methods

4.2.1. Preparation of genomic DNA

Pigment found in the eyes of *Ae. aegypti* mosquitoes inhibits *Taq* polymerase activity and the heads must therefore be removed prior to homogenization. Clean razor blades were used to remove the heads of 14 male mosquitoes (Liverpool strain). The mosquitoes were homogenized in 80 µL of Bender Buffer (0.1 M NaCl; 0.2 M Sucrose; 0.1 M Tris-HCl, pH 9.1; 0.05 M EDTA; 0.5% SDS) in a 1.7 mL sterile microfuge tube with a Kontes pestle and Pellet Pestle Motor (Kontes). The pestle was then washed with an additional 120 µL of Bender Buffer prior to adding 20 µL of 20 mg/mL Proteinase K (Promega). The solution was incubated overnight at 50^{o}C. The sample was gently extracted twice with Phenol/Chloroform/Isoamyl alcohol (25:24:1) and once with Chloroform/Isoamyl alcohol (24:1). The DNA was precipitated with the addition of 4 µL of 3 M NaOAc (pH 4.8) and 200 µL of isopropanol, mixing for 5 minutes and centrifugation for 10 minutes at 14,000 x g. The DNA pellet was washed with 70% ethanol, centrifuged for 5 minutes and air dried. The pellet was resuspended in 100 µL of nuclease-free water and stored at -20 ^{o}C.

4.2.2. 1st strand cDNA synthesis by RT-PCR

Total RNA was isolated from approximately 30 mg of *Ae. aegypti* larvae using the RNeasy Mini Kit (Qiagen, Valencia, CA), treated with DNase I with DNA-free (Ambion, Austin, TX) and subsequently used to synthesize first strand cDNA using SuperScript II reverse transcriptase (Invitrogen, Carlsbad, CA) per manufacturer's instructions. Briefly, 0.2 µM cDNA anchor primer (Table 7), 1.0 µg total RNA, 1 µl of 10 mM dNTPs were brought to a total volume of 12 µl with ddH_2O, heated to 65 ^oC for 5 minutes and quick chilled on ice. Additionally, 5X 1st strand buffer, 10 mM DTT and 40 units of RNase OUT Recombinant Ribonuclease Inhibitor (Invitrogen) were added and the mixture incubated at 42 ^oC for 2 minutes. After the addition of 200 units of SuperScript II (Invitrogen) the reaction was incubated for 1 hour at 42 ^oC, followed by heat inactivation at 70 ^oC for 15 minutes. The resulting product was digested with RNase H (Promega, Madison, WI) to produce single-stranded cDNA and stored at -20 ^oC.

4.2.3. *In silico* putative regulatory region identification

Drosophilid genomic sequence datasets were queried to identify putative *vasa* orthologs. We extracted 5.0 Kb of upstream sequence and saved the sequence data as a text file. The *Ae. aegypti* and *An. gambiae* genomic DNA databases were queried using the 5' end of the cDNA and transcription start site of *AaeVLG* and *AgVLG,* respectively. Using the existing database information, we extracted 5.0 Kb of upstream sequence and saved the sequence in a text file. The text files containing 5' upstream putative regulatory regions were subjected to the VISTA comparative analysis program (Frazer *et al.*, 2004) to identify conserved sequences.

Full length cDNA sequences were used to query existing *Ae. aegypti* genomic DNA sequence databases, both from Ensembl and NCBI. Sequence information beginning approximately 5.0 Kb upstream of the transcription start site and ending about 1.0 Kb downstream of the poly-adenylation site was copied into the Vector NTI Suite 8 software package (InforMax, Inc., 1999). To assist in determining the length of upstream DNA to clone, the 5.0 Kb upstream region identified *in silico* was submitted to the Transfac 7.0 Public database (BioBase Biological Databases) for putative transcription factor binding site identification.

4.2.4. Amplification of PAaeVLG

Efforts to amplify the cis-acting regulatory DNA responsible for *AaeVLG* mRNA expression and localization were undertaken prior to the development of the *Ae. aegypti* genomic sequence database. As a result, adaptor-ligated PCR (Siebert *et al.*, 1995) was utilized to amplify unknown upstream genomic DNA. The oligonucleotide adaptor was generated by annealing two oligonucleotides, ALPCRAdaptor1 (5'-CTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAGGT -3') and ALPCRAdaptor2 (5'- ACCTGCCCGG -3'). An amine group is added to the 3' end of ALPCRAdaptor2 to confer direction specific PCR amplification. Oligonucleotide annealing conditions were as follows; 200pmol/µL each adaptor and 10X annealing buffer (100 mM Tris-HCL, pH 7.5; 1 M NaCl; 10 mM EDTA), heated at 65 ^{o}C for 10 minutes and then cooled at room temperature for 2 hours. Genomic DNA was digested with either *Eco*RV, *Sca*I, *Dra*I, or *Pvu*II and annealed adaptor ligated to the ends using a 5:1 ratio of adaptor to genomic DNA using standard ligation techniques. Regulatory DNA amplification was performed using 1 µL adaptor-ligated DNA, 0.2 µM ALPCR-

AP1 primer, 0.2 µM gene specific primer, 10 mM dNTPs, 5.0 µL 10X Advantage2 Polymerase, and 2.5 units of Advatage2 Taq DNA polymerase. Nested PCR reactions were performed using essentially the same techniques, however ALPCR-AP2 and a nested gene specific primer were used. The resulting amplified products were cloned into pGEM-T vector (Promega) and sequenced. The sequence data was transferred to the Vector NTI Suite 8 software program and collapsed into a contig. Gene specific primers were designed to amplify an approximate 2.5 Kb DNA fragment.

4.2.5. PCR amplification of putative 5' *cis*-acting regulatory DNA

Utilizing known DNA sequence information, specific oligonucleotide primers were designed to amplify a region of genomic DNA approximately 2.6 Kb in length under the following reaction conditions; 2 mM MgCl2, 0.2 µM each primer, 10 mM dNTPs, 1.0 µL genomic DNA, and 2.5 units of Advantage2 *Taq* polymerase BD Biosciences). The reaction parameters were as follows: 96 ^{o}C for 2'; 35 cycles of 96 ^{o}C for 30", 65 ^{o}C for 30", and 72 ^{o}C for 4', followed by a final extension at 72 ^{o}C for 10'. The primers and their corresponding sequences can be found in Table 7. The amplified product was visualized on a 1% agarose gel to verify amplification and product size. Nested PCR reactions were performed under the conditions described above. Primers homologous to the 5' end of the amplified DNA were synthesized so as to incorporate an *Eco*RI restriction site while those homologous to the 3' end included a unique *Kpn*I restriction site. The amplified product was once again visualized on an agarose gel to confirm amplification and the amplified product was then cloned into the TOPO 2.1 TA cloning vector (Invitrogen). Single pass sequence reads were generated from each end

using the universal M13 primer set and the sequence determined using an ABI 3100 capillary sequencer.

Table 7. Primers used in Section 4.

Primer Name	Sequence (5'-3')
AaeVLG3UTR	GCGGCCGCGATTAAGGTTAGGGCCATTTAATGC
CG10125-3utrNOT1	GCAACGCGGCCGCTGATCGATAAAAGTATCGTCC
CG4735-3utrNOT1	GAGAGTGCGGCCGCTAGAGAAGTGGACGAACTATG
CG10050-3utrNOT1	GCAATGCGGCCGCTGATCTCCGGGGGGTTGGTGTG
AaeVLG5up1	CCCACCATCAAGCAGCACCCACAAGCCGCC
AaeVLG5up1N	GAATTCTAGCACATTTTACGATGTTGTTCTG
AaeVLG5upMUT	GGATCCCGCACTTGCAGTTTCTTTCAGAGGAGAATA ACGAATGAGAAAACACGC
AaeCG4735-5upF	TCTGATAAGAAATCTATACTGGTCGTCGCC
AaeCG4735-5upF2	GAATTCGCTTGAGTTCTGATTTTTCGGCAACTG
AaeCG4735-5upR	ACCTGATGTTGAGCGCTTCTTTCAGCACCG
AaeCG4735-5upR2	GGTACCTGTTTGAAGATTTAAGTTTTTGGAATG
AaeCG10125-5upF	GGCGCCACTATATATGGGATTTAACTTGTG
AaeCG10125-5upF2	GAATTCCATTGCTGTTCTGTCAAACACACAAGG
AaeCG10125-5upR	CAAACCGTGTTGGTCGAGTCGAATGATTTC
AaeCG10125-5upR2	GGTACCGATGATTTAGGGGTTTGAACGTTTC
AaeCG10050-5upF	GATGAGTTTATTAGTTCAATAACCATGCGG
AaeCG10050-5upF2	GAATTCCCGTGGAAATTCTAGGAGGAACTTCTG
AaeCG10050-5upR	GTCCCAAACGGCTCCTTCGGCGAACAAATC
AaeCG10050-5upR2	GGTACCTTTCACTATTCAAAAGGACTTTCTTTC

4.2.6. Modification of existing DNA plasmids for generation of dual-reporter constructs

An oligonucleotide fragment (AscFse) was designed to include a unique *Asc*I and *Fse*I restriction sites for easy integration of DNA from plasmids using the pSLfa1180fa shuttle vector. The oligonucleotides AscFse1 (5'- GGCCGCGCTGCGCAAGG CCGGCCTCTCTTCGCTATTACGAGGCGCGCCGACGTTGTAA -3') and AscFse2 (5'- CTAGTTACAACGTCGGCGCGCCTCGTAATAGCGAAGAGAGGCCGGCCTT GCGCAGCGC -3') were mixed in equal molar ratios of 0.1 M oligonucleotide. 12.5 µL of AscFse1, 12.5 µL AscFse2, 5.0 µL 10X annealing buffer (100 mM Tris-HCl, pH 7.5; 1 M NaCl; and 10 mM EDTA), and 20.0 µL double distilled water were heated at 65 ^{o}C for 10 minutes and then cooled at room temperature for 2 hours. The resulting double-strand DNA fragment was treated with T4 Polynucleotide Kinase (Promega) per manufacturer's instructions and extracted twice with phenol, precipitated with ethanol, and resuspended in 50 µL of water. The plasmid DNA construct, pBattB[3xP3-DsRed2nls]lox66 (Nimmo *et al.*, 2006) was digested with *Spe*I and *Not*I. The oligonucleotide fragment, AscFse, was cloned into the *Spe*I/*Not*I site to generate the plasmid, pBattBfa[3xP3-DsRed2nls]. Sequencing reactions were performed using the pBattB-F and pBattB-R primers to verify oligonucleotide insertion. The *Not*I site was restored while the *Spe*I restriction site was not. pBattBfa[3xP3-DsRed2nls] was digested with *Age*I and *Hpa*I to remove the DsRed2nls fragment and the DsRed2 fragment from pDsRed2.1 was cloned in, eliminating the nuclear localization sequence.

The *Mos1* transposon donor plasmid, pMos[3xP3-EGFPaf] (Horn and Wimmer, 2000) was digested with *Age*I and *Not*I to remove the EGFP fluorescent reporter gene

which was replaced with the DsRed2 fragment removed from pDsRed2.1 (Clontech) using *Age*I/*Not*I to create the plasmid, pMos[3xP3-DsRed2af].

4.2.7. PCR amplification and cloning of endogenous 3' untranslated regions (3' UTR) into the shuttle vector, pSLfa1180fa

Gene specific primers were designed to amplify the 3' UTR for each gene of interest. Used in conjunction with the primer, cDNA primer2, the 3' UTR was amplified to include the polyA$^+$ tail. The gene specific primers were designed with a unique *Not*I site on the 5' end. PCR reactions were carried out using standard gene amplification protocols. Reactions were performed with 2 mM MgCl2, 0.2 μM of each primer, 2.5 mM dNTPs, 100 ng 1st strand cDNA product and 2.0 units of *Taq* polymerase (Continental Lab Products). PCR conditions used to amplify gene fragments were as follows: 95 ^oC for 2 minutes; 30 cycles of 95 ^oC for 30 seconds, 62 ^oC for 30 seconds, and an extension phase of 72 ^oC for 1 minute; followed by a final extension phase of 72 ^oC for 5 minutes. Amplified products were visualized on a 1% agarose gel. Gene fragments successfully amplified from 1st strand cDNAs were subsequently cloned into the pGEM-T vector (Promega) and screened for appropriate fragment insertion as well as directionality using the gene specific primer and the universal M13 reverse primer. Plasmid DNA with the correct insert size and orientation was sequenced using an ABI 3100 capillary sequencer. The 3' UTR fragments were digested with the restriction enzymes *Not*I and *Spe*I and cloned into the *Not*I/*Spe*I sites of the shuttle vector, pSLfa1180fa (Horn and Wimmer, 2000) to create the plasmids; AaeVLG3'utr:pSLfa1180fa (*AaeVLG*), AeCG10125 3'utr:pSLfa1180 (*Aezpg*), AeCG4735 3'utr:pSLfa1180fa (*Aeshu*), and AeCG10050 3'utr:pSLfa1180fa (*AeEwald*).

4.2.8. Construction of DNA plasmids utilizing endogenous 5' *cis*-acting regulatory regions

The 5' *cis*-acting DNA for the *Ae*. *aegypti* genes, *AaeVLG*, *Aezpg*, *Aeshu*, and *AeEwald* were excised from PAaeVLG:pGEM-T, PCG10125:pGEM-T, PCG4735:pGEM-T, and PCG10050:pGEM-T, respectively. Each was digested with the restriction enzymes *Eco*RI and *Kpn*I. PCG10050:pGEM-T, however was digested with *Pst*I and *Kpn*I due to internal *Eco*RI restriction sites. The fragments; PAeVLG, PCG10125, and PCG4735, approximately 2.5 Kb each in length, were cloned into the *Eco*RI/*Kpn*I sites of the plasmid pEGFP.1. pEGFP.1 was generated by replacing the DsRed2 gene of pDsRed2.1 (Clontech) with EGFP from pEGFP-1 (Clontech) using the *Bam*HI and *Not*I sites. This generated the plasmids PAaeVLG:pEGFP.1, PCG10125:pEGFP.1, and PCG4735:pEGFP.1. The fragment, PCG10050, was excised from PCG10050 and cloned into pEGFP.1 using the unique *Pst*I/*Kpn*I restriction sites to generate PCG10050:pEGFP.1. PAaeVLG:pEGFP.1, PCG10125:pEGFP.1, and PCG4735:pEGFP.1 were each cut with the restriction enzymes *Eco*RI and *Not*I and the subsequent promoter fragments cloned into the plasmids AaeVLG3'utr:pSLfa1180fa, AeCG10125 3'utr:pSLfa1180, AeCG4735 3'utr:pSLfa1180fa to generate the novel DNA plasmid constructs; Pvas:EGFP:pSLfa1180fa, Pzpg:EGFP:pSLfa1180fa, and Pshu:EGFP:pSLfa1180fa, respectively. PCG10050:pEGFP.1 was digested with *Pst*I and *Not*I and the 3.4 Kb fragment containing the putative promoter inserted into the *Pst*I/*Not*I sites of PCG10050 3'utr:pSLfa1180fa to create PCG10050:EGFP:pSLfa1180fa.

To generate plasmid DNA capable of transposase expression the *Mos1 mariner* transposase was cloned into each existing construct currently using regulatory DNA.

Mos1 from the plasmid pIE1-3marORF (Mohammed, 2003) was digested with *Bam*HI and site filled with T4 DNA Polymerase to generate a blunt end before subsequent *Sac*II digestion. The germline-specific DNA plamsids; Pzpg:EGFP:pSLfa1180fa, Pshu:EGFP:pSLfa1180fa, and PCG10050:EGFP:pSLfa1180fa were digested with *Not*I, the ends blunted with T4 DNA Polymerase, and then digested with *Sac*II. The *Mos1* open reading frame was directionally cloned into the resulting vectors to generate the novel plasmids; Pzpg:Mos1:pSLfa1180fa, Pshu:Mos1:pSLfa1180fa, and PCG10050:Mos1:pSLfa1180fa. Pvas:EGFP:pSLfa1180fa was digested with *Bam*HI and *Not*I and the resulting backbone ends made blunt using T4 DNA Polymerase. The *Mos1* open reading frame was extracted from pIE1-3marORF (Mohammed, 2003) using *Sac*II and *Bam*HI and the ends site filled with T4 DNA Polymerase prior to the fragment being directionally cloned into the Pvas:pSLfa1180fa backbone to generate Pvas:Mos1:pSLfa1180fa.

All of the germline-specific expression plasmids were digested with the restriction enzyme *Asc*I to extract the synthetic constructs from the pSLfa1180fa vector. pBattBfa[3xP3-DsRed2nls] and pMos[3xP3-DsRed2af] were linearized with *Asc*I and the germline-specific constructs inserted into the open *Asc*I site. The synthetic *Mos1* insertions into pMos[3xP3-DsRed2af] created an autonomous element potentially capable of self-mobilization.

4.2.9. φC31 integrase

C31 integrase mRNA was transcribed from pET11phiC31poly(A) (Groth *et al.*, 2004) using MessageMAX™T7 Capped Message Transcription Kit (Epicentre Biotechnologies, Madison, WI). Prior to transcription, pET11phiC31poly(A) was

linearized with *Bam*HI. To extend the length of the polyA+ tail, the product resulting from the transcription reaction was subjected to the A-Plus™ Poly(A) Polymerase Tailing Kit (Epicentre Biotechnologies, Madison, WI). Approximately 32 µg of capped, A-tailed mRNA was produced. The resulting mRNA was treated to remove buffers, salts, and unincorporated nucleotides using the RNeasy Mini Kit (Qiagen, Valencia, CA). The mRNA was quantified using a spectrophotometer and precipitated using $1/10^{th}$ volume of 4M LiCl and 100% EtOH. The precipitated pellet was washed with 70% EtOH and resuspended in DEPC treated water to a final concentration of 1.5 µg/µL.

The DNA plasmid pIE1-3:Int was generated by digesting the pIE1-3 plasmid (Novagen, San Diego, CA) with *Bam*HI. The integrase open reading frame was digested out of pTA-Int (Groth *et al.*, 2000) using *Bam*HI and ligated into the linearized pIE1-3 vector DNA.

4.2.10. *Aedes aegypti* embryo microinjections

Microinjections were performed essentially as described (Morris, 1997). To increase throughput the coverslips were supported vertically in water to cover the eggs and allow the halocarbon oil to run off. The embryos were then carefully removed with a fine brush and placed on damp filter paper in a humid container.

For site-specific transformation, the #2 targeting strain (provided by Paul Eggleston, Keele University) used carries the attP site and expresses ECFP in the eyes (Nimmo *et al.*, 2006). Embryos were co-injected with mRNA (750 ng/µL) and either pBattBfa[3xP3-DsRed2nls]:Pvas:EGFP or pBattBfa[3xP3-DsRed2]: Pvas:EGFP at a final concentration of 500 ng/µL. Additional injection experiments utilized the pBattBfa[3xP3-DsRed2]: PAaeVLG:EGFP construct along with either pIE1-3:Int or

pHSC70/IntN-NLS (Nakayama *et al.*, 2006) at final concentrations of 500 ng/µL. G_1 larvae were screened using the Rhodamine filter on a UV stereo-microscope.

4.3. Results

4.3.1. Identification of putative *cis*-acting regulatory regions

It has been demonstrated in *D. melanogaster* that a 40 bp region of the DNA upstream of the transcription start site directs germline-specific *vasa* expression (Sano *et al.*, 2002). To identify putative regulatory regions in related Drosophilids, the *Drosophila* genomic sequence databases were queried for *vasa* orthologs and extracted public genomic sequence information. When compared to the 40 bp region of DNA, it was revealed that closely related species such as *D. simulans* and *D. sechellia* showed remarkable sequence conservation. When the 40 bp regulatory region sequence was used to query the upstream regions of other Drosophilids and the mosquitoes, sequences similar to the 40 bp region in *D. melanogaster* were present. Similar 40 bp regions from seven Drosophilids, the honeybee, *An. gambiae*, *Ae. aegypti* were isolated and aligned. The alignment was entered into the WebLogo program (Crooks *et al.*, 2004) to produce an image representing nucleotide frequencies in specific positions (Figure 22). An interesting pattern emerges where there is a seemingly consistent consensus sequence. While the significance of this sequence in the mosquito is uncertain, it is of interest to note. Oligonucleotide primers were designed to amplify a genomic region that would in fact include the putative regulatory region.

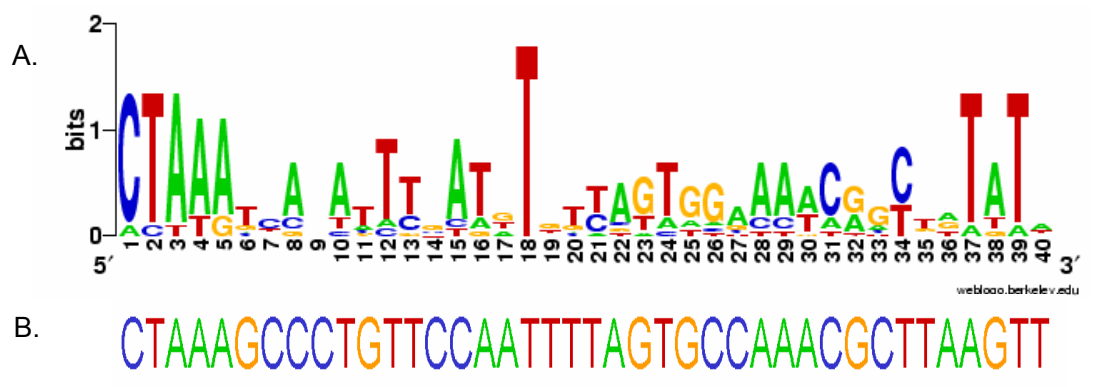


Figure 22. Putative *vasa* regulatory regulatory region. A) WebLogo representation of sequences similar to the 40 bp regulatory region found upstream of the *Drosophila melanogaster vasa* gene. Seven species of Drosophilids and *Ae. aegypti* genomic DNA sequences were mined for sequences similar to the one defined for *D. melanogaster*. A consensus sequence appears to be present in the 5' end. Sequences aligned included *D. simulans*, *D. sechellia, D. melanogaster, D. yakuba, D. ananassae, D. erectus,* and *D. psuedoobscura*. B) Aedes aegypti putative regulatory region for comparison against the consensus sequence found in A.

Utilizing the Transfac public database, we subjected each upstream DNA sequence to a search identifying transcription factor binding sites. The *Aezpg* upstream analysis reveal a transcriptional activator/repressor binding site that is putatively bound by Hunchback and Bicoid. The *Aeshu* upstream genomic analysis identified putative binding sites for the transcription factors, BR-C, which is involved in pupal development, and Elf, a regulator of developmental control genes such as tailless and decapentaplegic. These regions are of particular interest to this project as they may contain specific regulatory elements conferring germline-specific expression. Sequence specific oligonucleotide primers were designed such that the resulting amplified product would include these regions with putative transcription factor binding sites.

4.3.2. Mosquito microinjection

Several attempts were made to transform mosquitoes using the φC31 integrase system. Due to a lack of success co-injecting plasmid DNA with integrase mRNA,

plasmid DNA constructs were generated to express integrase. pIE1-3:Int and pHSC70/IntN-NLS were co-injected with a plasmid containing the attB recombination site. Table 8 provides an illustration of the results from microinjection experiments. 7,917 embryos were injected and 890 successfully hatched, resulting in a hatch rate of 11.2%. However, no successful integrations were observed.

Table 8. Summary of transformation experiments. Donor plasmid is as described with PAaeVLG:EGFP cloned into *AscI* site.

Donor	**Helper**	**# of injected embryos**	**# hatched**	**# transformed**
pBattBfa[3xP3-DsRed2nls]	φC31 integrase mRNA	2931	273 (9.31%)	0
pBattBfa[3xP3-DsRed2]	φC31 integrase mRNA	1789	186 (10.4%)	0
pBattBfa[3xP3-DsRed2]	pIE1:int	987	101 (10.2%)	0
pBattBfa[3xP3-DsRed2]	pHSC70/IntN-NLS	2210	330 (14.93%)	0

4.4. Discussion

The advent of fluorescent reporter protein technologies has greatly enhanced the ability to identify transformed individuals as well as assess promoter function and protein localization. By manipulating existing plasmid DNA to generate novel constructs the power of fluorescent reporter genes to identify successful transformation events,

promoter function, or both has been harnessed. The promoter regions for four putative germline-specific genes have been cloned so that they will drive expression of either a fluorescent reporter gene or *Mos1* transposase. The successful, stable integration of a transposase gene whose expression is restricted to the germline will potentially result in an increased frequency of transformation events and/or to enhance the spread of transgenes through naïve populations. There is likely to be an optimal transposase concentration that will yield ideal activity of transposable element. Further study of promoter function both spatio-temporally and in terms of promoter strength will be important in understanding transposase activity within the developing embryos. The successful isolation of four putative germline-specific promoters will allow for testing of their efficacy *in vivo*.

While embryonic injections are on going, no successful transgene integrations have been recorded to date. Using the φC31 integrase system (Nimmo *et al.*, 2006), an increased frequency of transformed individuals is expected. Using traditional techniques of transformation with TEs and an estimated transformation frequency of 0.05%, there would be an expected four individual trangenics lines produced.

It has recently been suggested that mRNA kit choice and DNA/RNA fidelity and purity is of utmost importance. The kits used to generate and purify mRNA are being reassessed. Additionally, current efforts are underway to generate a plasmid DNA construct using endogenous mosquito promoters to drive integrase expression. Work will continue to generate transgenic lines expressing EGFP reporter protein under the control of *cis*-acting regulatory elements.

Transformation by TEs is non-specific and results in random integration events. This makes promoter characterization difficult to assess due to the possibility of position effect variegation. The φC31 integrase system is extremely powerful as it results in site-specific integration events. The site-specificity eliminates the randomness of the integration and reduces the variability between generated transformed mosquito lines. This reduction in variation will allow for the direct assessment of gene expression and importantly, promoter function, between lines generated with different promoter elements.

5. GENERAL CONCLUSIONS AND FUTURE DIRECTIONS

5.1. General Conclusions

The development of efficient germ-line transformation systems in non-drosophilid insects has been a goal in the field of medical entomology for many years. The high levels of human morbidity and mortality caused by mosquitoes in particular makes the incorporation of refractory transgenes an attractive proposal. The diversity of the mosquitoes is great; in habitat, in pathogens transmitted, in their human impact, and in genome organization. The complexity of the target genomes coupled with the precise interaction of the pathogen, the mosquito, and human host makes the tasks at hand more daunting. While the generation of transgenic mosquitoes is now routine it is by no means a trivial task. To generate stably transformed insects incapable of disease transmission, we must increase the frequency with which we have success at the initial stages. Enhanced transformation will not only aid in the integration or proliferation of transgenes but will also play a vital role in the future of gene and genome characterization in the mosquitoes.

It is proposed that utilizing endogenous promoters to drive transposase expression will greatly increase the frequency of transgene integration. If the power of germline-specific promoters is appropriately harnessed, it will allow the creation of helper lines of mosquitoes to further enhance the integration and remobilization of transgenes. This study has identified four germline-specific gene candidates with evidence of mRNA localization in the presumptive germ tissue. Work presented here also supports the idea that large scale datasets from well described model organisms can shed light on similar processes occurring in non-model systems. The evidence provided here should

encourage those who perform research in insects with challenges such as large and complex genome size and architecture, or complex life cycles.

The germline-specific *vasa* gene was successfully cloned in the mosquitoes, *Ae. aegypti* and *An. gambiae*. *Vasa* was shown to be highly expressed in early embryos and bloodfed ovaries in both species. Transient expression was shown in later, gastrula stage embryos while reduced levels of expression were shown in ovarian tissue dissected from sugar-fed females. *In situ* hybridization and putative amino acid alignment confirm the mosquito *vasa* orthologs have been cloned. A region of DNA upstream of the *Ae. aegypti vasa* transcription start site as well as the 3' UTR has been cloned to drive EGFP or *Mos1* expression and localization. A potential autonomous element has been constructed as a means of integrating transgenes into the mosquito genome as well as to provide a gene drive mechanism. A series of sequences similar to the 40 bp regulatory region known in *D. melanogaster* to cause germ cell localization has been identified in the mosquitoes. Further study may conclude that this short sequence is a conserved regulatory region.

This study has identified germline-specific gene orthologs in *Ae. aegypti* using existing database information from the model organism, *D. melanogaster*. *zpg* and *shu* have been characterized to some extent in *D. melanogaster* while CG10050 has an unknown function. The respective orthologs, *Aezpg*, *Aeshu*, and *AeEwald* have been cloned from the yellow fever mosquito. *Aezpg* and *Aeshu* are expressed primarily at high levels in early embryos and the ovaries of bloodfed females with lesser levels of transcript found in all female lifestages. *AeEwald* appears to be unique to the ovaries of females during oogenesis and early embryo pole plasm. *In situ* hybridization data

suggests mRNA localization to the presumptive germ tissue, similar to the pattern found in *D. melanogaster*.

Novel plasmids have been generated that will allow us to test promoter function and protein localization. Furthermore, autonomous elements were created to assess transgene remobilization within the germline.

5.2. Future Consideration

Continued use of exogenous promoter sequences has proven beneficial to the mosquito community. However, exogenous promoter activity is unreliable within the species of interest and must therefore be independently tested within each host. The identification of endogenous promoters with characterized spatio-temporal expression patterns will provide the opportunity to more precisely regulate transgene expression. The restricted regulation of transposase might serve to increase transformation frequencies as well as create a gene drive system when using an autonomous element. The identification of endogenous genes and their regulatory region has become significantly more efficient with the release of the *An. gambiae* and *Ae. aegypti* genome sequences in 2002 and 2007, respectively.

While a variety of genes have been identified in *D. melanogaster* with germline-specific expression, few have been identified in mosquitoes. It is reasonable to assume that orthologous genes will exist within mosquito genomes due to the similarity of embryonic development amongst dipterans. The identification of orthologous genes and *in situ* hybridization evidence confirms the similarity of early development in the mosquitoes and fruit fly. While this study originated prior to the release of genomic datasets, the subsequent releases of draft and final genome sequences facilitated the

discovery of additional putative germline-specific gene orthologs and the cloning of regulatory regions.

To properly assess promoter/enhancer activity, the production of plasmid constructs and their incorporation into the genome is vital. Current technologies produce transgenic lines a small fraction of the time, typically with 1-10% efficiency. Consensus theory suggests one reason for reduced transformation frequency is the great cost incurred by injected embryos. Injected embryos hatch at a much lower rate than those not subjected to injection, typically 20-30% as opposed to roughly 90%. The need to slightly desiccate embryos prior to injection is a cause of the reduced hatch rate. The vast size of the *Ae. aegypti* genome (approximately five fold larger than *An. gambiae*) may be another reason for reduced transformation frequency. Nearly 50% of the genome is composed of TEs. The *Ae. aegypti* genome has a similar number of genes as compared to other dipterans. The large size of the genome compared to the number of genes found within suggests that a large portion of the genome is not transcribed. It is likely that much of these areas are composed of heterochromatin. With large fields of heterochromatin within the genome, there are fewer areas with acceptable TE integration sites that will allow for transgene expression. The size and repetitiveness of the *Ae. aegypti* genome may result in the increased likelihood a TE is integrated into an area unable to support efficient gene transcription.

The generation of stable helper lines of mosquitoes may serve to improve initial transformation frequencies. By providing transposase expression through an endogenous source, only the injection of donor plasmids will be required, potentially reducing the

volume injected into the early embryos. The reduced volume injected will reduce the level of desiccation required prior to injection.

In order for a transgene to spread through a population it must be able to remobilize. Previous studies have revealed that the TEs commonly used in mosquito transformation rarely, if ever, remobilize in the germ tissue. The creation of an autonomous element capable of excising and remobilizing itself may alleviate this issue. Restricting the transposition activity to the germline may also reduce the fitness costs incurred by transgenic individuals while stimulating remobilization of the TE within the germline. However, high levels of transposition activity may not be beneficial due to the possibility of overproduction inhibition and perhaps an increased likelihood that the host genomes will negatively regulate highly active TEs.

This study was originally limited by the technology available at the time. The lack of sequence and functional data for endogenous genes was a driving force to use data from related species to identify orthologs *in silico*. Subsequent releases of both EST and genome sequence databases revolutionized the approach previously taken. While the databases are incomplete, specifically for *Ae. aegypti*, they still provide a wealth of information useful in basic science.

The publication of the whole mount embryo *in situ* hybridization protocol by Jennifer Juhn at UC-Irvine was a first for *Ae. aegypti*. While the techniques works well for the *oskar* and *nanos* genes, similar success was not obtained with the genes studied here. This may be a result of higher expression levels of *osk* and *nos* mRNA relative to *AaeVLG*, *Aeshu*, *Aezpg*, and *AeEwald*. However, the lack of rich transcript staining found here might also be the result of the lack of optimization of this protocol (although

significant *osk* staining was observed with this protocol). A number of probes were generated for each gene to potentially resolve the issues and reduce background. Future interest will involve the optimization of this protocol. The creation of plasmid DNA constructs utilizing endogenous promoters might alleviate these concerns. As the focus of this study is the localization of transposase protein, the localization of fluorescent reporter proteins to specific tissues will be a significant measure of success and also a means to observe the developing germline at multiple developmental timepoints.

The difficulties of *in situ* hybridization and transformation presented by the mosquito will continue to create challenges. However, the identification of endogenous germline-specific promoters for improved transformation will enhance the current tools used to study and manipulate mosquito species. Knowledge gained from promoter/enhancer trap studies as well as forward and reverse genetic studies will go a long way in shaping the field of vector biology and mosquito-borne diseases. Improved transformation will play a vital role in the creation and success of new control strategies.

REFERENCES

Adelman, Z. N., Jasinskiene, N. and James, A. A., 2002 Development and applications of transgenesis in the yellow fever mosquito, *Aedes aegypti*. Mol Biochem Parasitol. 121, 1-10.

Adelman, Z. N., Jasinskiene, N., Vally, K. J., Peek, C., Travanty, E. A., Olson, K. E., Brown, S. E., Stephens, J. L., Knudson, D. L., Coates, C. J. and James, A. A., 2004 Formation and loss of large, unstable tandem arrays of the *piggyBac* transposable element in the yellow fever mosquito, *Aedes aegypti*. Transgenic Res. 13, 411-25.

Allen, M. L., O'Brochta, D. A., Atkinson, P. W. and Levesque, C. S., 2001 Stable, germ-line transformation of *Culex quinquefasciatus* (Diptera: Culicidae). J Med Entomol. 38, 701-10.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J., 1990 Basic local alignment search tool. J Mol Biol. 215, 403-10.

Anxolabehere, D., Kidwell, M. G. and Periquet, G., 1988 Molecular characteristics of diverse populations are consistent with the hypothesis of a recent invasion of *Drosophila melanogaster* by mobile *P* elements. Mol Biol Evol. 5, 252-69.

Arca, B., Lombardo, F., Capurro, M., della Torre, A., Spanos, L., Dimopoulos, G., Louis, C., James, A. A. and Coluzzi, M., 1999 Salivary gland-specific gene expression in the malaria vector *Anopheles gambiae*. Parassitologia. 41, 483-7.

Arensburger, P., Kim, Y. J., Orsetti, J., Aluvihare, C., O'Brochta, D. A. and Atkinson, P. W., 2005 An active transposable element, *Herves*, from the African malaria mosquito *Anopheles gambiae*. Genetics. 169, 697-708.

Arrighi, R. B., Lycett, G., Mahairaki, V., Siden-Kiamos, I. and Louis, C., 2005 Laminin and the malaria parasite's journey through the mosquito midgut. J Exp Biol. 208, 2497-502.

Atkinson, P. W. and James, A. A., 2002 Germline transformants spreading out to many insect species. Adv Genet. 47, 49-86.

Atkinson, P. W. and Michel, K., 2002 What's buzzing? Mosquito genomics and transgenic mosquitoes. Genesis. 32, 42-8.

Atkinson, P. W. and O'Brochta D, A., 2000. *Hermes* and other hAT elements in gene vectors in insects In: Handler A. M., James A. A. (Eds.), Insect Transgenesis, CRC Press, Boca Raton, FL, pp. 219-235.

Atkinson, P. W., Warren, W. D. and O'Brochta, D. A., 1993 The *hobo* transposable element of *Drosophila* can be cross-mobilized in houseflies and excises like the Ac element of maize. Proc Natl Acad Sci U S A. 90, 9693-7.

Barrett, A. D. and Higgs, S., 2007 Yellow fever: a disease that has yet to be conquered. Annu Rev Entomol. 52, 209-29.

Becker, N. and Ascher, K. R. S., 1998 The use of *Bacillus theringiensis* subsp. *israelensis* (Bti) against mosquitoes, with special emphasis on the ecological impact. Isr. J. Entomol. 32, 63-69.

Benedict, M. (1997) Care and Maintenance of Anopheline Mosquito Colonies, Chapman and Hall, London.

Berg, D. E., M.A., S. and Lowe, J. B., 1983 Specificty of transposon Tn5 insertion. Genetics. 105, 813-828.

Bibikova, M., Carroll, D., Segal, D. J., Trautman, J. K., Smith, J., Kim, Y. G. and Chandrasegaran, S., 2001 Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. Mol Cell Biol. 21, 289-97.

Blair, C. D., Adelman, Z. N. and Olson, K. E., 2000 Molecular strategies for interrupting arthropod-borne virus transmission by mosquitoes. Clin Microbiol Rev. 13, 651-61.

Breman, J. G., Egan, A. and Keusch, G. T., 2001 The intolerable burden of malaria: a new look at the numbers. Am J Trop Med Hyg. 64, iv-vii.

Brown, A. E., Bugeon, L., Crisanti, A. and Catteruccia, F., 2003 Stable and heritable gene silencing in the malaria vector *Anopheles stephensi*. Nucleic Acids Res. 31, e85.

Burt, A., 2003 Site-specific selfish genes as tools for the control and genetic engineering of natural populations. Proc Biol Sci. 270, 921-8.

Burt, A. and Koufopanou, V., 2004 Homing endonuclease genes: the rise and fall and rise again of a selfish element. Curr Opin Genet Dev. 14, 609-15.

Carlson, J. O., Afanasiev, B. and Suchman, E., 2000. Densonucleosis viruses as transducing vectors for insects In: Handler A. M., James A. A. (Eds.), Insect Transgenesis, CRC Press, Boca Raton, FL, pp. 139-159.

Castresana, J., 2000 Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol. 17, 540-52.

Catteruccia, F., Benton, J. P. and Crisanti, A., 2005 An *Anopheles* transgenic sexing strain for vector control. Nat Biotechnol. 23, 1414-7.

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, 959-62.

Centers for Disease Control, 2005. Dengue Fever. http://www.cdc.gov/ncidod/dvbid/dengue/.

Chandre, F., Darrier, F., Manga, L., Akogbeto, M., Faye, O., Mouchet, J. and Guillet, P., 1999 Status of pyrethroid resistance in *Anopheles gambiae* sensu lato. Bull World Health Organ. 77, 230-4.

Chang, C. C., Dearden, P. and Akam, M., 2002 Germ line development in the grasshopper *Schistocerca gregaria*: *vasa* as a marker. Dev Biol. 252, 100-18.

Chapman, H. C., 1974 Biological control of mosquito larvae. Annu Rev Entomol. 19, 33-59.

Cho, K. H., Cheon, H. M., Kokoza, V. and Raikhel, A. S., 2006 Regulatory region of the vitellogenin receptor gene sufficient for high-level, germ line cell-specific ovarian expression in transgenic *Aedes aegypti* mosquitoes. Insect Biochem Mol Biol. 36, 273-81.

Coates, C. J., 2005. Stable transformation of vector species In: Marquardt, W. C., Black, I., William C., Freier, J. E., Hagedorn, H. H., Hemingway, J., Higgs, S., James, A. A., Kondratieff, B., Moore, C. G. (Eds.), Biology of Disease Vectors, Elsevier Academic Press, San Diego, CA, pp. 613-624.

Coates, C. J., Jasinskiene, N., Miyashiro, L. and James, A. A., 1998 *Mariner* transposition and transformation of the yellow fever mosquito, *Aedes aegypti*. Proc Natl Acad Sci U S A. 95, 3748-51.

Coates, C. J., Jasinskiene, N., Pott, G. B. and James, A. A., 1999 Promoter-directed expression of recombinant fire-fly luciferase in the salivary glands of *Hermes*-transformed *Aedes aegypti*. Gene. 226, 317-25.

Collins, F. H. and Besansky, N. J., 1994 Vector biology and the control of malaria in Africa. Science. 264, 1874-5.

Collins, F. H. and James, A. A., 1996 Genetic modification of mosquitoes. Science Med. 3, 52-61.

Costa, Y., Speed, R. M., Gautier, P., Semple, C. A., Maratou, K., Turner, J. M. and Cooke, H. J., 2006 Mouse MAELSTROM: the link between meiotic silencing of unsynapsed chromatin and microRNA pathway? Hum Mol Genet. 15, 2324-34.

Crooks, G. E., Hon, G., Chandonia, J. M. and Brenner, S. E., 2004 WebLogo: a sequence logo generator. Genome Res. 14, 1188-90.

Curtis, C. F., 1968 Possible use of translocations to fix desirable genes in insect pest populations. Nature. 218, 368-9.

Curtis, C. F., 1994 Approaches to vector control: new and trusted. 4. Appropriate technology for vector control: impregnated bed nets, polystyrene beads and fly traps. Trans R Soc Trop Med Hyg. 88, 144-6.

Davis, S., Bax, N. and Grewe, P., 2001 Engineered underdominance allows efficient and economical introgression of traits into pest populations. J Theor Biol. 212, 83-98.

Dearden, P., Grbic, M. and Donly, C., 2003 *Vasa* expression and germ-cell specification in the spider mite *Tetranychus urticae*. Dev Genes Evol. 212, 599-603.

Dobson, S. L., Marsland, E. J., Veneti, Z., Bourtzis, K. and O'Neill, S. L., 2002 Characterization of *Wolbachia* host cell range via the *in vitro* establishment of infections. Appl Environ Microbiol. 68, 656-60.

Eldridge, B. F., 2005. Mosquitoes, the Culicidae In: Marquardt, W. C., Black, I., William C., Freier, J. E., Hagedorn, H. H., Hemingway, J., Higgs, S., James, A. A., Kondratieff, B., Moore, C. G. (Eds.), Biology of Disease Vectors, Elsevier Academic Press, San Diego, CA, pp. 95-111.

Engels, W. R., 1989. *P*-elements in *Drosophila melanogaster* In: Berg, D. E., Howe, M. M. (Eds.), Mobile DNA, American Society for Microbiology, Washington DC, pp. 437-489.

Felsenstein, J., 1989 PHYLIP - Phylogeny Inference Package (Version 3.2). Cladistics. 5, 164-166.

Findley, S. D., Tamanaha, M., Clegg, N. J. and Ruohola-Baker, H., 2003 *Maelstrom*, a *Drosophila* spindle-class gene, encodes a protein that colocalizes with *Vasa* and RDE1/AGO1 homolog, *Aubergine*, in nuage. Development. 130, 859-71.

Finnegan, D. J., 1985 Transposable elements in eukaryotes. Int Rev Cytol. 93, 281-326.

Fischer, G. and Schmid, F. X., 1990 The mechanism of protein folding. Implications of *in vitro* refolding models for *de novo* protein folding and translocation in the cell. Biochemistry. 29, 2205-12.

Franz, A. W., 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*. Proc Natl Acad Sci U S A. 103, 4198-203.

Franz, G. and Savakis, C., 1991 Minos, a new transposable element from *Drosophila hydei*, is a member of the Tc1-like family of transposons. Nucleic Acids Res. 19, 6646.

Fraser, M. J., Cary, L., Boonvisudhi, K. and Wang, H. G., 1995 Assay for movement of Lepidopteran transposon IFP2 in insect cells using a baculovirus genome as a target DNA. Virology. 211, 397-407.

Frazer, K. A., Pachter, L., Poliakov, A., Rubin, E. M. and Dubchak, I., 2004 VISTA: computational tools for comparative genomics. Nucleic Acids Res. 32, W273-9.

Fujiwara, Y., Komiya, T., Kawabata, H., Sato, M., Fujimoto, H., Furusawa, M. and Noce, T., 1994 Isolation of a DEAD-family protein gene that encodes a murine homolog of *Drosophila vasa* and its specific expression in germ cell lineage. Proc Natl Acad Sci U S A. 91, 12258-62.

Galindo, R. L., Edwards, D. N., Gillespie, S. K. and Wasserman, S. A., 1995 Interaction of the *pelle* kinase with the membrane-associated protein *tube* is required for transduction of the dorsoventral signal in *Drosophila* embryos. Development. 121, 2209-18.

Gaunt, M. W. and Miles, M. A., 2002 An insect molecular clock dates the origin of the insects and accords with palaeontological and biogeographic landmarks. Mol Biol Evol. 19, 748-61.

Gilboa, L., Forbes, A., Tazuke, S. I., Fuller, M. T. and Lehmann, R., 2003 Germ line stem cell differentiation in *Drosophila* requires gap junctions and proceeds via an intermediate state. Development. 130, 6625-34.

Gilchrist, B. M. and Haldane, J. B. S., 1947 Sex-linkage and sex-determination in a mosquito, *Culex molestus*. Heredity. 33, 175-190.

Goddard, M. R., Leigh, J., Roger, A. J. and Pemberton, A. J., 2006 Invasion and persistence of a selfish gene in the cnidaria. PLoS ONE. 1, e3.

Goltsev, Y., Hsiong, W., Lanzaro, G. and Levine, M., 2004 Different combinations of gap repressors for common stripes in *Anopheles* and *Drosophila* embryos. Dev Biol. 275, 435-46.

Grossman, G. L., Rafferty, C. S., Clayton, J. R., Stevens, T. K., Mukabayire, O. and Benedict, M. Q., 2001 Germline transformation of the malaria vector, *Anopheles gambiae*, with the *piggyBac* transposable element. Insect Mol Biol. 10, 597-604.

Groth, A. C., Fish, M., Nusse, R. and Calos, M. P., 2004 Construction of transgenic *Drosophila* by using the site-specific integrase from phage phiC31. Genetics. 166, 1775-82.

Groth, A. C., Olivares, E. C., Thyagarajan, B. and Calos, M. P., 2000 A phage integrase directs efficient site-specific integration in human cells. Proc Natl Acad Sci U S A. 97, 5995-6000.

Gubler, D. J., 2002 Epidemic dengue/dengue hemorrhagic fever as a public health, social and economic problem in the 21st century. Trends Microbiol. 10, 100-3.

Hall, T. (2001) BioEdit, version 5.0.9. North Carolina State University Raleigh, NC.

Halstead, S. B., 2002 Vaccines aplenty. Curr Opin Infect Dis. 15, 461-3.

Handler, A., 2000. An introduction to the history of insect gene transfer In: Handler A. M., James A. A. (Eds.), Insect Transgenesis: Methods and Applications, CRC Press, Boca Raton, pp. 3-28.

Handler, A. M., Gomez, S. P. and O'Brochta, D. A., 1993 A functional analysis of the *P*-element gene-transfer vector in insects. Arch Insect Biochem Physiol. 22, 373-84.

Hartl, D. L., Hiraizumi, Y. and Crow, J. F., 1967 Evidence for sperm dysfunction as the mechanism of segregation distortion in *Drosophila melanogaster*. Proc Natl Acad Sci U S A. 58, 2240-5.

Hay B., Ackerman, L., Barbel S., Jan L. Y., Jan Y. N., 1988a Identification of a component of *Drosophila* polar granules. Development. 103, 625-640.

Hay B., Jan L. Y., Jan Y. N., 1988b A Protein Component of *Drosophila* Polar Granules Is Encoded by *vasa* and Has Extensive Sequence Similarity to ATP-Dependent Helicases. Cell. 55, 577-87.

Hemingway, J., 2005. Biological control of mosquitoes In: Marquardt, W. C., Black, I., William C., Freier, J. E., Hagedorn, H. H., Hemingway, J., Higgs, S., James, A.A., Kondratieff, B., Moore, C. G. (Eds.), Biology of Disease Vectors, Elsevier Academic Press, San Diego, CA, pp. 649-660.

Hemingway, J. and Ranson, H., 2005. Chemical control of vectors and mechanisms of resistance In: Marquardt, W. C., Black, I., William C., Freier, J. E., Hagedorn, H. H., Hemingway, J., Higgs, S., James, A. A., Kondratieff, B., Moore, C. G. (Eds.), Biology of Disease Vectors, Elsevier Academic Press, San Diego, CA, pp. 627-637.

Hickey, W. A. and Craig, G. B., 1966 Genetic distortion of sex ratio in a mosquito, *Aedes aegypti*. Genetics. 53, 1177–1196

Holt, R. A., Subramanian, G. M., Halpern, A., Sutton, G. G., Charlab, R., Nusskern, D. R., Wincker, P., Clark, A. G., Ribeiro, J. M., Wides, R., Salzberg, S. L., Loftus, B., Yandell, M., Majoros, W. H., Rusch, D. B., Lai, Z., Kraft, C. L., Abril, J. F., Anthouard, V., Arensburger, P., Atkinson, P. W., Baden, H., de Berardinis, V., Baldwin, D., Benes, V., Biedler, J., Blass, C., Bolanos, R., Boscus, D., Barnstead, M., Cai, S., Center, A., Chaturverdi, K., Christophides, G. K., Chrystal, M. A., Clamp, M., Cravchik, A., Curwen, V., Dana, A., Delcher, A., Dew, I., Evans, C. A., Flanigan, M., Grundschober-Freimoser, A., Friedli, L., Gu, Z., Guan, P., Guigo, R., Hillenmeyer, M. E., Hladun, S. L., Hogan, J. R., Hong, Y. S., Hoover, J., Jaillon, O., Ke, Z., Kodira, C., Kokoza, E., Koutsos, A., Letunic, I., Levitsky, A., Liang, Y., Lin, J. J., Lobo, N. F., Lopez, J. R., Malek, J. A., McIntosh, T. C., Meister, S., Miller, J., Mobarry, C., Mongin, E., Murphy, S. D., O'Brochta, D. A., Pfannkoch, C., Qi, R., Regier, M. A., Remington, K., Shao, H., Sharakhova, M. V., Sitter, C. D., Shetty, J., Smith, T. J., Strong, R., Sun, J., Thomasova, D., Ton, L. Q., Topalis, P., Tu, Z., Unger, M. F., Walenz, B., Wang, A., Wang, J., Wang, M., Wang, X., Woodford, K. J., Wortman, J. R., Wu, M., Yao, A., Zdobnov, E. M., Zhang, H., Zhao, Q., et al., 2002 The genome sequence of the malaria mosquito *Anopheles gambiae*. Science. 298, 129-49.

Horn, C. and Wimmer, E. A., 2000 A versatile vector set for animal transgenesis. Dev Genes Evol. 210, 630-7.

Ito, J., Ghosh, A., Moreira, L. A., Wimmer, E. A. and Jacobs-Lorena, M., 2002 Transgenic anopheline mosquitoes impaired in transmission of a malaria parasite. Nature. 417, 452-5.

Jacquier, A. and Dujon, B., 1985 An intron-encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. Cell. 41, 383-94.

James, A. A., 2000. Control of disease trasnmission through genetic modification of mosquitoes In: Handler A. M., James A. A. (Eds.), Insect Transgenesis, CRC Press, Boca Raton, FL, pp. 319-333.

James, A. A., 2002 Engineering mosquito resistance to malaria parasites: the avian malaria model. Insect Biochem Mol Biol. 32, 1317-23.

Jasinskiene, N., Coates, C. J., Benedict, M. Q., Cornel, A. J., Rafferty, C. S., James, A. A. and Collins, F. H., 1998 Stable transformation of the yellow fever mosquito, *Aedes aegypti*, with the *Hermes* element from the housefly. Proc Natl Acad Sci U S A. 95, 3743-7.

Johnstone, O., Deuring, R., Bock, R., Linder, P., Fuller, M. T. and Lasko, P., 2005 *Belle* is a *Drosophila* DEAD-box protein required for viability and in the germ line. Dev Biol. 277, 92-101.

Juhn, J. and James, A. A., 2006 *oskar* gene expression in the vector mosquitoes, *Anopheles gambiae* and *Aedes aegypti*. Insect Mol Biol. 15, 363-72.

Kaya, H. K. and Gaugler, R., 1993 Entomopathogenic nematodes. Annu Rev Entomol. 38, 181-206.

Kidwell, M. G., Frydryk, T. and Novy, J. B., 1983 The hybrid dysgenesis potential of *Drosophila melanogaster* strains of diverse temporal and geographic origin. Drosophila Inf. Serv. 59, 63-69.

Kidwell, M. G. and Lisch, D. R., 2000 Transposable elements and host genome evolution. Trends in Ecology and Evolution. 15, 95-99.

Koella, J. C., Sorensen, F. L. and Anderson, R. A., 1998 The malaria parasite, *Plasmodium falciparum*, increases the frequency of multiple feeding of its mosquito vector, *Anopheles gambiae*. Proc Biol Sci. 265, 763-8.

Kokoza, V., Ahmed, A., Cho, W. L., Jasinskiene, N., James, A. A. and Raikhel, A., 2000 Engineering blood meal-activated systemic immunity in the yellow fever mosquito, *Aedes aegypti*. Proc Natl Acad Sci U S A. 97, 9144-9.

Kokoza, V., Ahmed, A., Wimmer, E. A. and Raikhel, A. S., 2001 Efficient transformation of the yellow fever mosquito *Aedes aegypti* using the *piggyBac* transposable element vector pBac[3xP3-EGFP afm]. Insect Biochem Mol Biol. 31, 1137-43.

Komiya, T. and Tanigawa, Y., 1995 Cloning of a gene of the DEAD box protein family which is specifically expressed in germ cells in rats. Biochem Biophys Res Commun. 207, 405-10.

Kozak, M., 1991 A short leader sequence impairs the fidelity of initiation by eukaryotic ribosomes. Gene Expr. 1, 111-5.

Kusano, A., Staber, C., Chan, H. Y. and Ganetzky, B., 2003 Closing the (Ran)GAP on segregation distortion in *Drosophila*. Bioessays. 25, 108-15.

Lamb, J. R., Tugendreich, S. and Hieter, P., 1995 Tetratrico peptide repeat interactions: to TPR or not to TPR? Trends Biochem Sci. 20, 257-9.

Lasko, P. F. and Ashburner, M., 1988 The product of the *Drosophila* gene *vasa* is very similar to eukaryotic initiation factor-4A. Nature. 335, 611-7.

Legner, E. F., 1995 Biological control of Diptera of medical and veterinary importance. J. Vector Ecology. 20, 59-120.

Linder, P., Lasko, P. F., Ashburner, M., Leroy, P., Nielsen, P. J., Nishi, K., Schnier, J. and Slonimski, P. P., 1989 Birth of the D-E-A-D box. Nature. 337, 121-2.

Lohe, A. R. and Hartl, D. L., 1996 Autoregulation of *mariner* transposase activity by overproduction and dominant-negative complementation. Mol Biol Evol. 13, 549-55.

Lombardo, F., Nolan, T., Lycett, G., Lanfrancotti, A., Stich, N., Catteruccia, F., Louis, C., Coluzzi, M. and Arca, B., 2005 An *Anopheles gambiae* salivary gland promoter analysis in *Drosophila melanogaster* and *Anopheles stephensi*. Insect Mol Biol. 14, 207-16.

Lukacsovich, T. and Yamamoto, D., 2001 Trap a gene and find out its function: toward functional genomics in *Drosophila*. J Neurogenet. 15, 147-68.

Lyttle, T. W., 1991 Segregation distorters. Annu Rev Genet. 25, 511-57.

Mackenzie, J. S., Gubler, D. J. and Petersen, L. R., 2004 Emerging flaviviruses: the spread and resurgence of Japanese encephalitis, West Nile and dengue viruses. Nat Med. 10, S98-109.

Magori, K. and Gould, F., 2006 Genetically engineered underdominance for manipulation of pest populations: a deterministic model. Genetics. 172, 2613-20.

Mahowald, A. P., 2001 Assembly of the *Drosophila* germ plasm. Int Rev Cytol. 203, 187-213.

Martinho, R. G., Kunwar, P. S., Casanova, J. and Lehmann, R., 2004 A noncoding RNA is required for the repression of RNApolII-dependent transcription in primordial germ cells. Curr Biol. 14, 159-65.

McClintock, B., 1950 The origin and behavior of mutable loci in maize. Proc Natl Acad Sci U S A. 36, 344-55.

Medhora, M., Maruyama, K. and Hartl, D. L., 1991 Molecular and functional analysis of the *mariner* mutator element *Mos1* in *Drosophila*. Genetics. 128, 311-8.

Michel, K., Budd, A., Pinto, S., Gibson, T. J. and Kafatos, F. C., 2005 *Anopheles gambiae* SRPN2 facilitates midgut invasion by the malaria parasite *Plasmodium berghei*. EMBO Rep. 6, 891-7.

Mohammed, A. 2003 In *Germline Transformation and Isolation of Midgut Related Genes from the Potato Tuber Moth, Phthoramiaea operculella, (Lepidoptera: Gelechiidae)*. Dissertation, Texas A&M University, College Station, TX.

Monnerat, A. T., Machado, M. P., Vale, B. S., Soares, M. J., Lima, J. B., Lenzi, H. L. and Valle, D., 2002 *Anopheles albitarsis* embryogenesis: morphological identification of major events. Mem Inst Oswaldo Cruz. 97, 589-96.

Moreira, L. A., Edwards, M. J., Adhami, F., Jasinskiene, N., James, A. A. and Jacobs-Lorena, M., 2000 Robust gut-specific gene expression in transgenic *Aedes aegypti* mosquitoes. Proc Natl Acad Sci U S A. 97, 10895-8.

Mori, A., Chadee, D. D., Graham, D. H. and Severson, D. W., 2004 Reinvestigation of an endogenous meiotic drive system in the mosquito, *Aedes aegypti* (Diptera: Culicidae). J Med Entomol. 41, 1027-33.

Morris, A. C., 1997. Microinjection of mosquito embryos In: Crampton, J. M., Beard, C. B. and Louis, C. (Eds.), Molecular Biology of Insect Disease Vectors, Chapman and Hall, London, pp. 423-429.

Morris, A. C., Schaub, T. L. and James, A. A., 1991 FLP-mediated recombination in the vector mosquito, *Aedes aegypti*. Nucleic Acids Res. 19, 5895-900.

Munn, K. and Steward, R., 2000 The *shut-down* gene of *Drosophila melanogaster* encodes a novel FK506-binding protein essential for the formation of germline cysts during oogenesis. Genetics. 156, 245-56.

Nakao, H., 1999 Isolation and characterization of a *Bombyx vasa*-like gene. Dev Genes Evol. 209, 312-6.

Nakayama, G., Kawaguchi, Y., Koga, K. and Kusakabe, T., 2006 Site-specific gene integration in cultured silkworm cells mediated by phiC31 integrase. Mol Genet Genomics. 275, 1-8.

Newton, M. E., Wood, R. J. and Southern, D. I., 1976 A cytogenetic analysis of meiotic drive in the mosquito *Aedes aegypti* (L.). Genetica. 48, 137-143.

Nimmo, D. D., Alphey, L., Meredith, J. M. and Eggleston, P., 2006 High efficiency site-specific genetic engineering of the mosquito genome. Insect Mol Biol. 15, 129-36.

Nirmala, X. and James, A. A., 2003 Engineering Plasmodium-refractory phenotypes in mosquitoes. Trends Parasitol. 19, 384-7.

Nolan, T., Bower, T. M., Brown, A. E., Crisanti, A. and Catteruccia, F., 2002 *piggyBac*-mediated germline transformation of the malaria mosquito *Anopheles stephensi* using the red fluorescent protein dsRED as a selectable marker. J Biol Chem. 277, 8759-62.

O'Brochta, D. A. and Atkinson, P. W., 1998 Building the better bug. Sci Am. 279, 90-5.

O'Brochta D, A. and Handler, A. M., 1988 Mobility of *P* elements in drosophilids and nondrosophilids. Proc Natl Acad Sci U S A. 85, 6052-6056.

O'Brochta, D. A., Sethuraman, N., Wilson, R., Hice, R. H., Pinkerton, A. C., Levesque, C. S., Bideshi, D. K., Jasinskiene, N., Coates, C. J., James, A. A., Lehane, M. J. and Atkinson, P. W., 2003 Gene vector and transposable element behavior in mosquitoes. J Exp Biol. 206, 3823-34.

O'Brochta, D. A., Subramanian, R. A., Orsetti, J., Peckham, E., Nolan, N., Arensburger, P., Atkinson, P. W. and Charlwood, D. J., 2006 *hAT* element population genetics in Anopheles gambiae s.l. in Mozambique. Genetica. 127, 185-98.

O'Hare, K., Driver, A., McGrath, S. and Johnson-Schiltz, D. M., 1992 Distribution and structure of cloned *P* elements from the *Drosophila melanogaster* P strain pi 2. Genet Res. 60, 33-41.

O'Hare, K. and Rubin, G. M., 1983 Structures of *P* transposable elements and their sites of insertion and excision in the *Drosophila melanogaster* genome. Cell. 34, 25-35.

Olsen, L. C., Aasland, R. and Fjose, A., 1997 A *vasa*-like gene in zebrafish identifies putative primordial germ cells. Mech Dev. 66, 95-105.

Olson, K. E., Adelman, Z. N., Travanty, E. A., Sanchez-Vargas, I., Beaty, B. J. and Blair, C. D., 2002 Developing arbovirus resistance in mosquitoes. Insect Biochem Mol Biol. 32, 1333-43.

Olson, K. E., Carlson, J. O. and Blair, C. D., 2005. Virus-induced, transient expression of genes in mosquitoes In: Marquardt, W. C., Black, I., William C., Freier, J. E., Hagedorn, H. H., Hemingway, J., Higgs, S., James, A. A., Kondratieff, B., Moore, C. G. (Eds.), Biology of Disease Vectors, Elsevier Academic Press, San Diego, CA, pp. 601-611.

Olson, K. E., Higgs, S., Gaines, P. J., Powers, A. M., Davis, B. S., Kamrud, K. I., Carlson, J. O., Blair, C. D. and Beaty, B. J., 1996 Genetically engineered resistance to dengue-2 virus transmission in mosquitoes. Science. 272, 884-6.

Opiyo, P., Mukabana, W. R., Kiche, I., Mathenge, E., Killeen, G. F. and Fillinger, U., 2007 An exploratory study of community factors relevant for participatory malaria control on Rusinga Island, western Kenya. Malar J. 6, 48.

Osta, M. A., Christophides, G. K. and Kafatos, F. C., 2004 Effects of mosquito genes on *Plasmodium* development. Science. 303, 2030-2.

Perera, O. P., Harrell, I. R. and Handler, A. M., 2002 Germ-line transformation of the South American malaria vector, *Anopheles albimanus*, with a *piggyBac*/EGFP transposon vector is routine and highly efficient. Insect Mol Biol. 11, 291-7.

Peri, S. and Pandey, A., 2001 A reassessment of the translation initiation codon in vertebrates. Trends Genet. 17, 685-7.

Phillips, R. S., 2001 Current status of malaria and potential for control. Clin Microbiol Rev. 14, 208-26.

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. BMC Biol. 5, 11.

Quackenbush, J., Cho, J., Lee, D., Liang, F., Holt, I., Karamycheva, S., Parvizi, B., Pertea, G., Sultana, R. and White, J., 2001 The TIGR Gene Indices: analysis of gene transcript sequences in highly sampled eukaryotic species. Nucleic Acids Res. 29, 159-64.

Raminani LN, C. E., 1975 Early Embryology of *Aedes aegypti* (L.) (Diptera: Culicidae). Int. J. Insect Morphol. & Embryol. 4, 517-528.

Raminani LN, C. E., 1978 Embryology of *Aedes aegypti* (L.) (Diptera: Culicidae): Organogenesis. Int. J. Insect Morphol. & Embryol. 7, 273-296.

Ranson, H., Jensen, B., Wang, X., Prapanthadara, L., Hemingway, J. and Collins, F. H., 2000 Genetic mapping of two loci affecting DDT resistance in the malaria vector *Anopheles gambiae*. Insect Mol Biol. 9, 499-507.

Rasgon, J. L., Ren, X. and Petridis, M., 2006 Can *Anopheles gambiae* be infected with *Wolbachia pipientis*? Insights from an *in vitro* system. Appl Environ Microbiol. 72, 7718-22.

Ribeiro, J. M. and Kidwell, M. G., 1993 Transposable elements as population drive mechanisms: specification of critical parameter values. J. Med. Entomol. 31, 10-16.

Rio, D. C., 1990 Molecular mechanisms regulating *Drosophila P* element transposition. Annu Rev Genet. 24, 543-78.

Rio, D. C., Laski, F. A. and Rubin, G. M., 1986 Identification and immunochemical analysis of biologically active *Drosophila P*-element transposase. Cell. 44, 21-32.

Robinson, A. S., Franz, G. and Atkinson, P. W., 2004 Insect transgenesis and its potential role in agriculture and human health. Insect Biochem Mol Biol. 34, 113-20.

Rogan, W. J. and Chen, A., 2005 Health risks and benefits of bis(4-chlorophenyl)-1,1,1-trichloroethane (DDT). Lancet. 366, 763-73.

Rogozin, I. B., Kochetov, A. V., Kondrashov, F. A., Koonin, E. V. and Milanesi, L., 2001 Presence of ATG triplets in 5' untranslated regions of eukaryotic cDNAs correlates with a 'weak' context of the start codon. Bioinformatics. 17, 890-900.

Rowan, K. H., Orsetti, J., Atkinson, P. W. and O'Brochta, D. A., 2004 Tn5 as an insect gene vector. Insect Biochem Mol Biol. 34, 695-705.

Rubin, G. M. and Spradling, A. C., 1982 Genetic transformation of *Drosophila* with transposable element vectors. Science. 218, 348-53.

Saffman, E. E. and Lasko, P., 1999 Germline development in vertebrates and invertebrates. Cell Mol Life Sci. 55, 1141-63.

Saito, K., Nishida, K. M., Mori, T., Kawamura, Y., Miyoshi, K., Nagami, T., Siomi, H. and Siomi, M. C., 2006 Specific association of *Piwi* with rasiRNAs derived from retrotransposon and heterochromatic regions in the *Drosophila* genome. Genes Dev. 20, 2214-22.

Sano, H., Nakamura, A. and Kobayashi, S., 2002 Identification of a transcriptional regulatory region for germline-specific expression of vasa gene in *Drosophila melanogaster*. Mech Dev. 112, 129-39.

Sarker, A., Coates, C. J., Whyard, S., Willhoeft, U. and Atkinson, P. W., et al., 1997 The *Hermes* element from *Musca domestica* can trasnpose in for families of cyclorrhaphan flies. Genetica. 99, 15-29.

Schroder, R., 2006 *vasa* mRNA accumulates at the posterior pole during blastoderm formation in the flour beetle *Tribolium castaneum*. Dev Genes Evol. 216, 277-83.

Schupbach, T. and Wieschaus, E., 1991 Female sterile mutations on the second chromosome of *Drosophila melanogaster*. II. Mutations blocking oogenesis or altering egg morphology. Genetics. 129, 1119-36.

Shibata, N., Umesono, Y., Orii, H., Sakurai, T., Watanabe, K. and Agata, K., 1999 Expression of *vasa*(*vas*)-related genes in germline cells and totipotent somatic stem cells of planarians. Dev Biol. 206, 73-87.

Siebert, P. D., Chenchik, A., Kellogg, D. E., Lukyanov, K. A. and Lukyanov, S. A., 1995 An improved PCR method for walking in uncloned genomic DNA. Nucleic Acids Res. 23, 1087-8.

Sinkins, S. P., 2004 *Wolbachia* and cytoplasmic incompatibility in mosquitoes. Insect Biochem Mol Biol. 34, 723-9.

Sinkins, S. P. and Gould, F., 2006 Gene drive systems for insect disease vectors. Nat Rev Genet. 7, 427-35.

Sinkins, S. P. and Hastings, I. M., 2004 Male-specific insecticide resistance and mosquito transgene dispersal. Trends Parasitol. 20, 413-6.

Sinkins, S. P. and O'Neill, S. L., 2000. *Woolbachia* as a vehicle to modify insect populations In: Handler A. M., James A. A. (Eds.), Insect Transgenesis, CRC Press, Boca Raton, FL, pp. 271-287.

Small, G. J., 2005. Environmental management for vector control In: Marquardt, W. C., Black, I., William C., Freier, J. E., Hagedorn, H. H., Hemingway, J., Higgs, S., James, A. A., Kondratieff, B., Moore, C. G. (Eds.), Biology of Disease Vectors, Elsevier Academic Press, San Diego, CA, pp. 639-647.

Stephenson, J. R., 2005 Understanding dengue pathogenesis: implications for vaccine design. Bull World Health Organ. 83, 308-14.

Tazuke, S. I., Schulz, C., Gilboa, L., Fogarty, M., Mahowald, A. P., Guichet, A., Ephrussi, A., Wood, C. G., Lehmann, R. and Fuller, M. T., 2002 A germline-specific gap junction protein required for survival of differentiating early germ cells. Development. 129, 2529-39.

Thompson, J. D., Higgins, D. G. and Gibson, T. J., 1994 CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673-80.

Thompson JD, T. J., Poch O, 2003 RASCAL: rapid scanning and correction of multiple sequence alignments. Bioinformatics. 19, 1155-1161.

Tomancak, P., Beaton, A., Weiszmann, R., Kwan, E., Shu, S., Lewis, S. E., Richards, S., Ashburner, M., Hartenstein, V., Celniker, S. E. and Rubin, G. M., 2002 Systematic determination of patterns of gene expression during *Drosophila* embryogenesis. Genome Biol. 3, RESEARCH0088.

Tsunekawa, N., Naito, M., Sakai, Y., Nishida, T. and Noce, T., 2000 Isolation of chicken *vasa* homolog gene and tracing the origin of primordial germ cells. Development. 127, 2741-50.

Tu, Z. and Coates, C., 2004 Mosquito transposable elements. Insect Biochem Mol Biol. 34, 631-44.

Turelli, M. and Hoffmann, A. A., 1991 Rapid spread of an inherited incompatibility factor in California *Drosophila*. Nature. 353, 440-2.

Vagin, V. V., Sigova, A., Li, C., Seitz, H., Gvozdev, V. and Zamore, P. D., 2006 A distinct small RNA pathway silences selfish genetic elements in the germline. Science. 313, 320-4.

Vezzani, D., 2007 Review: artificial container-breeding mosquitoes and cemeteries: a perfect match. Trop Med Int Health. 12, 299-313.

Villalva, C., Touriol, C., Seurat, P., Trempat, P., Delsol, G. and Brousset, P., 2001 Increased yield of PCR products by addition of T4 gene 32 protein to the SMART PCR cDNA synthesis system. Biotechniques. 31, 81-3, 86.

Warren, W. D., Atkinson, P. W. and O'Brochta, D. A., 1994 The *Hermes* transposable element from the house fly, *Musca domestica*, is a short inverted repeat-type element of the hobo, Ac, and Tam3 (hAT) element family. Genet Res. 64, 87-97.

Wessler, S. R., 2005 Homing into the origin of the AP2 DNA binding domain. Trends Plant Sci. 10, 54-6.

Wilson, M. H., Coates, C. J. and George, A. L., Jr., 2007 *PiggyBac* Transposon-mediated Gene Transfer in Human Cells. Mol Ther. 15, 139-45.

Wilson, R., Orsetti, J., Klocko, A. D., Aluvihare, C., Peckham, E., Atkinson, P. W., Lehane, M. J. and O'Brochta, D. A., 2003 Post-integration behavior of a *Mos1 mariner* gene vector in *Aedes aegypti*. Insect Biochem Mol Biol. 33, 853-63.

Wood, R. J., 2005. Genetic control of vectors In: Marquardt, W. C., Black, I., William C., Freier, J. E., Hagedorn, H. H., Hemingway, J., Higgs, S., James, A. A., Kondratieff, B., Moore, C. G. (Eds.), Biology of Disease Vectors, Elsevier Academic Press, San Diego, CA, pp. 661-669.

World Health Organization, 2003. Malaria in Africa Infosheet. http://www.rbm.who.int/cmc_upload/0/000/015/370/RBMInfosheet_3.htm.

Wyss, J. H., 2000 Screwworm eradication in the Americas. Ann N Y Acad Sci. 916, 186-93.

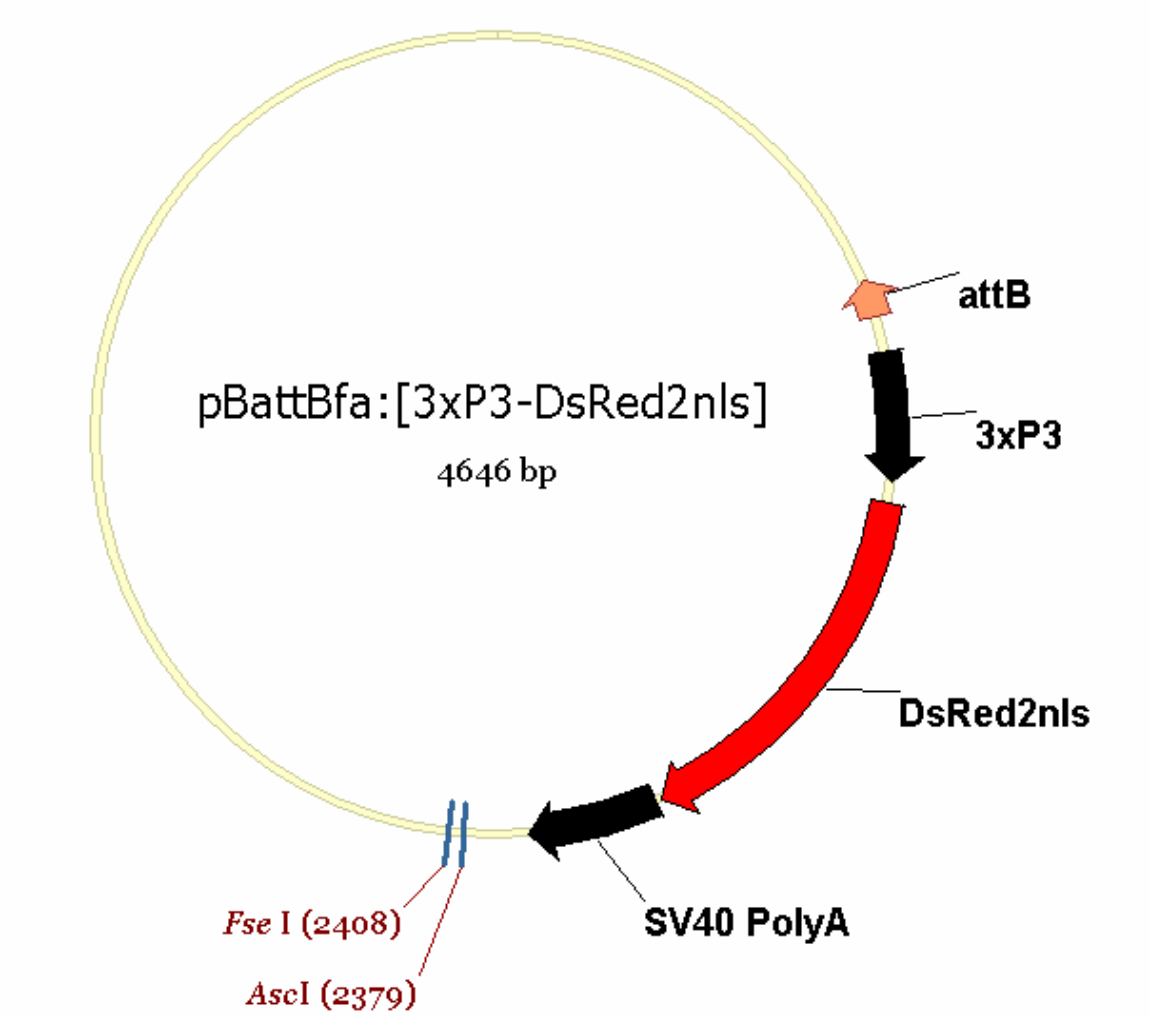
Yoon, C., Kawakami, K. and Hopkins, N., 1997 Zebrafish *vasa* homologue RNA is localized to the cleavage planes of 2- and 4-cell-stage embryos and is expressed in the primordial germ cells. Development. 124, 3157-65.

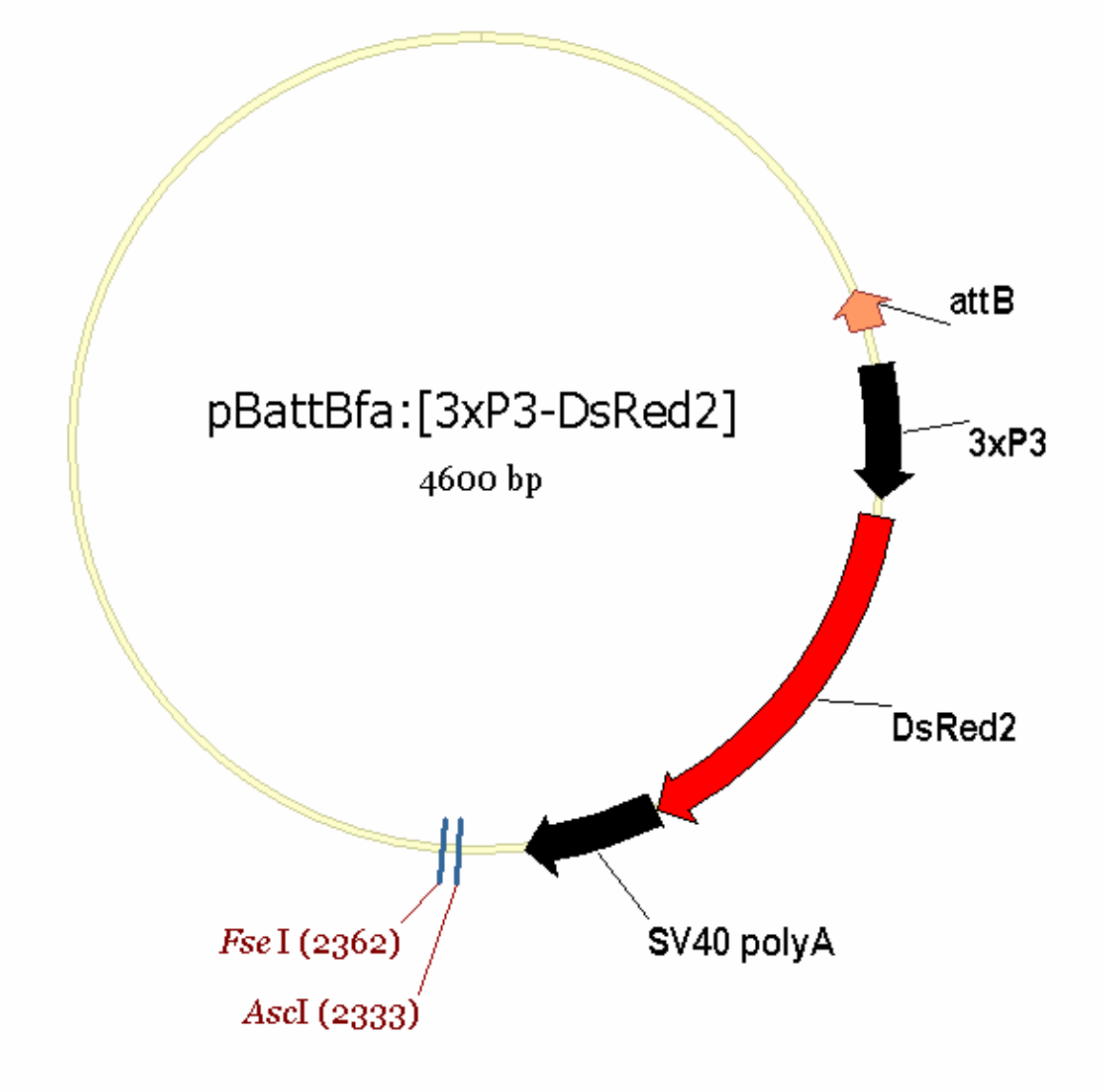
Yoshizaki, G., Sakatani, S., Tominaga, H. and Takeuchi, T., 2000 Cloning and characterization of a *vasa*-like gene in rainbow trout and its expression in the germ cell lineage. Mol Reprod Dev. 55, 364-71.

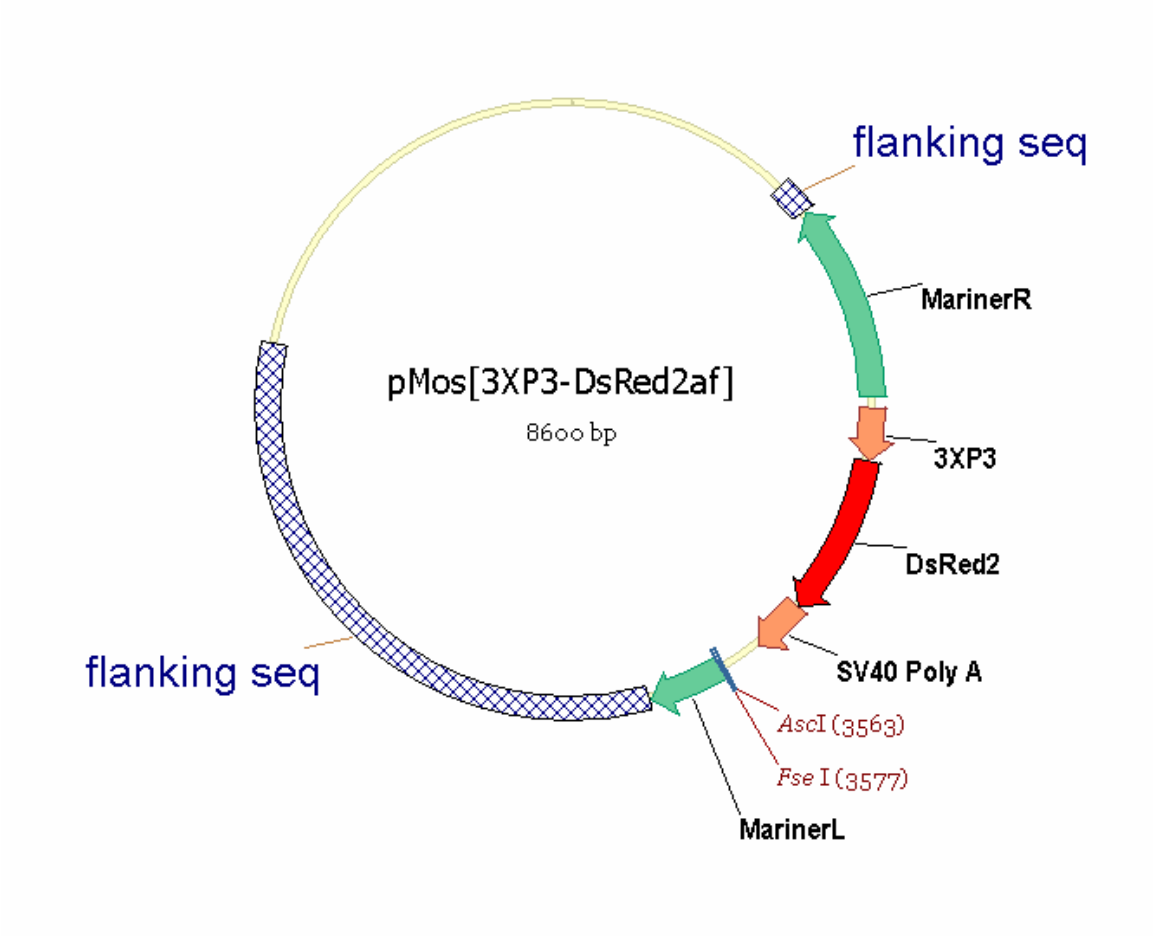
Zwiebel, L. J. and Takken, W., 2004 Olfactory regulation of mosquito-host interactions. Insect Biochem Mol Biol. 34, 645-52.

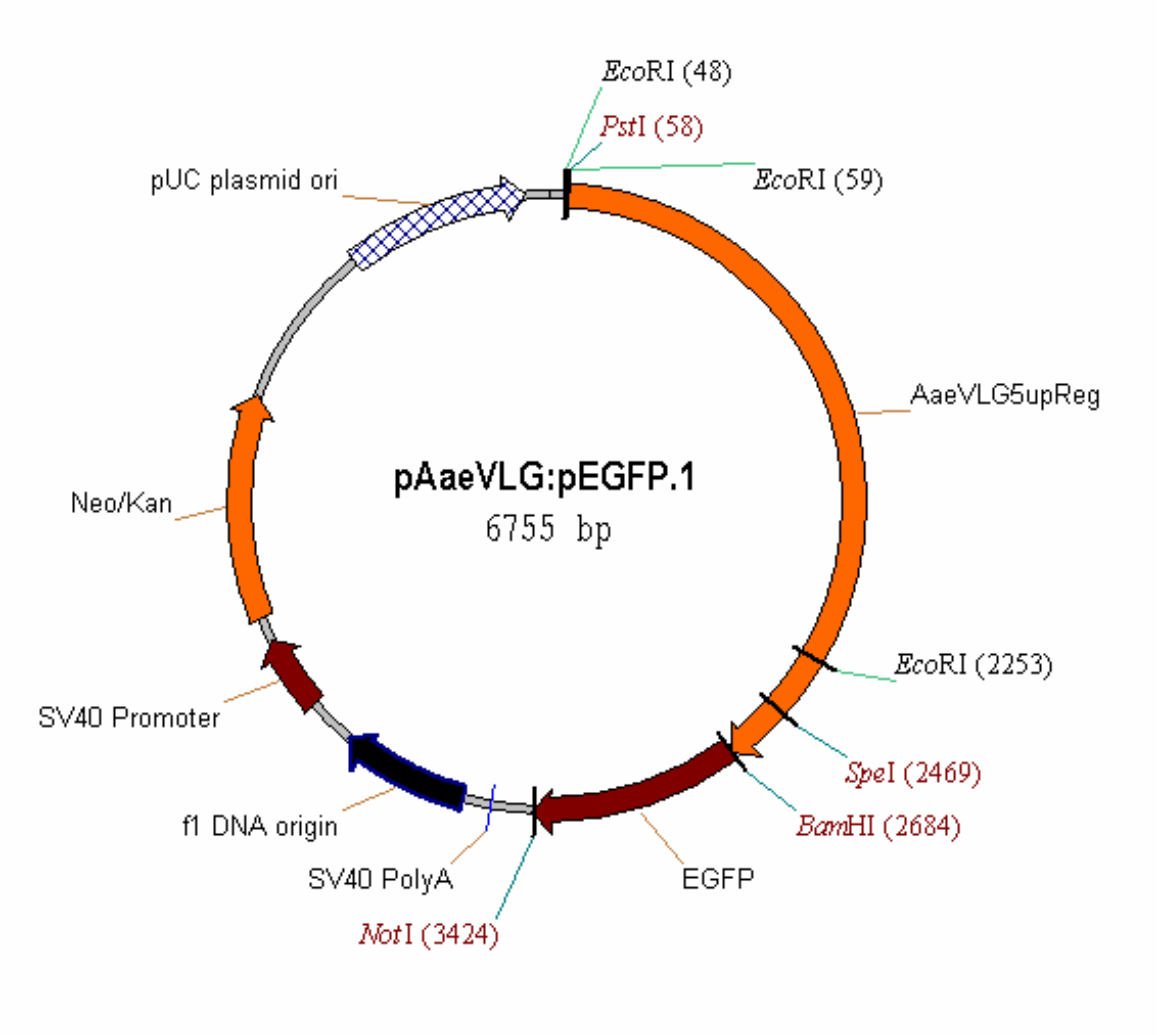
APPENDIX

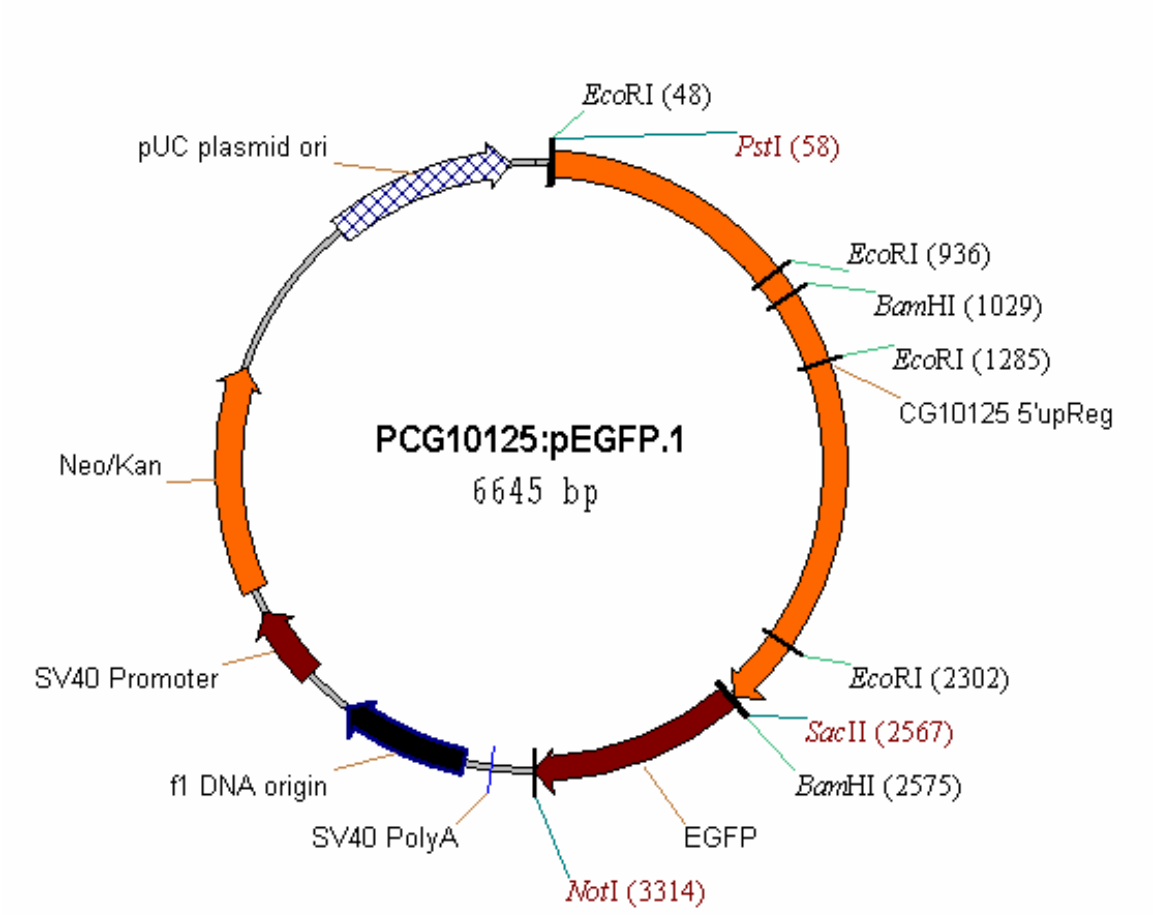
Plasmid Maps

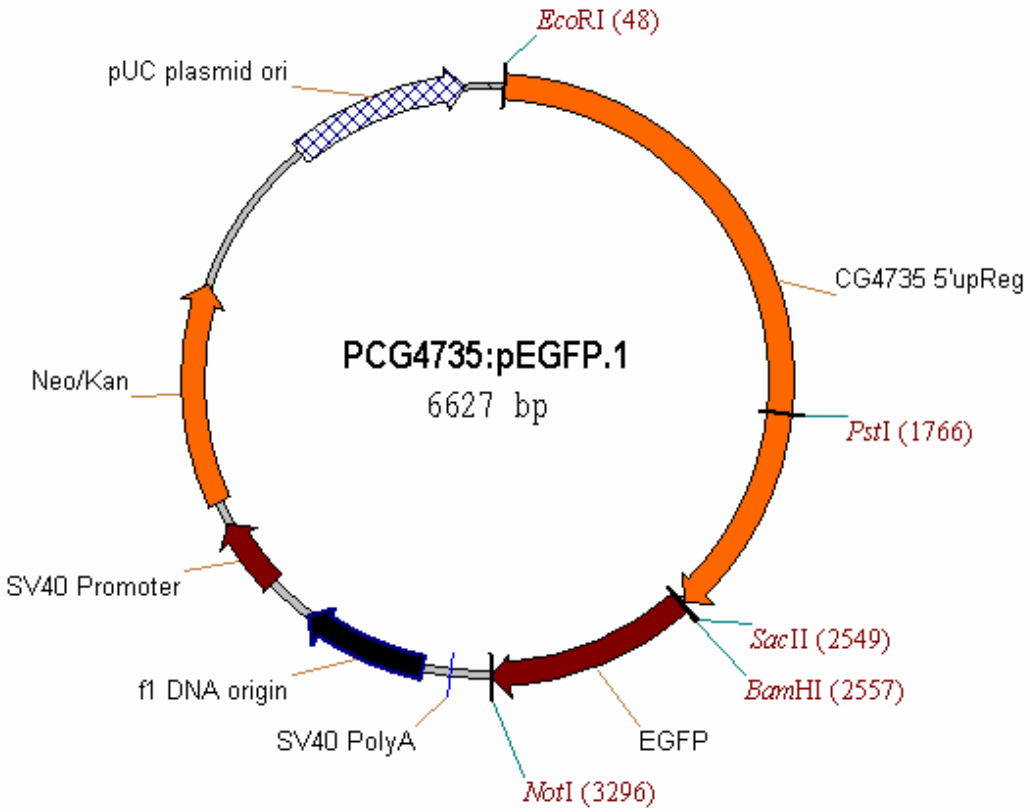

pBattBfa:[3xP3-DsRed2nls]
4646 bp
attB
3xP3
DsRed2nls
SV40 PolyA
Fse I (2408)
AscI (2379)

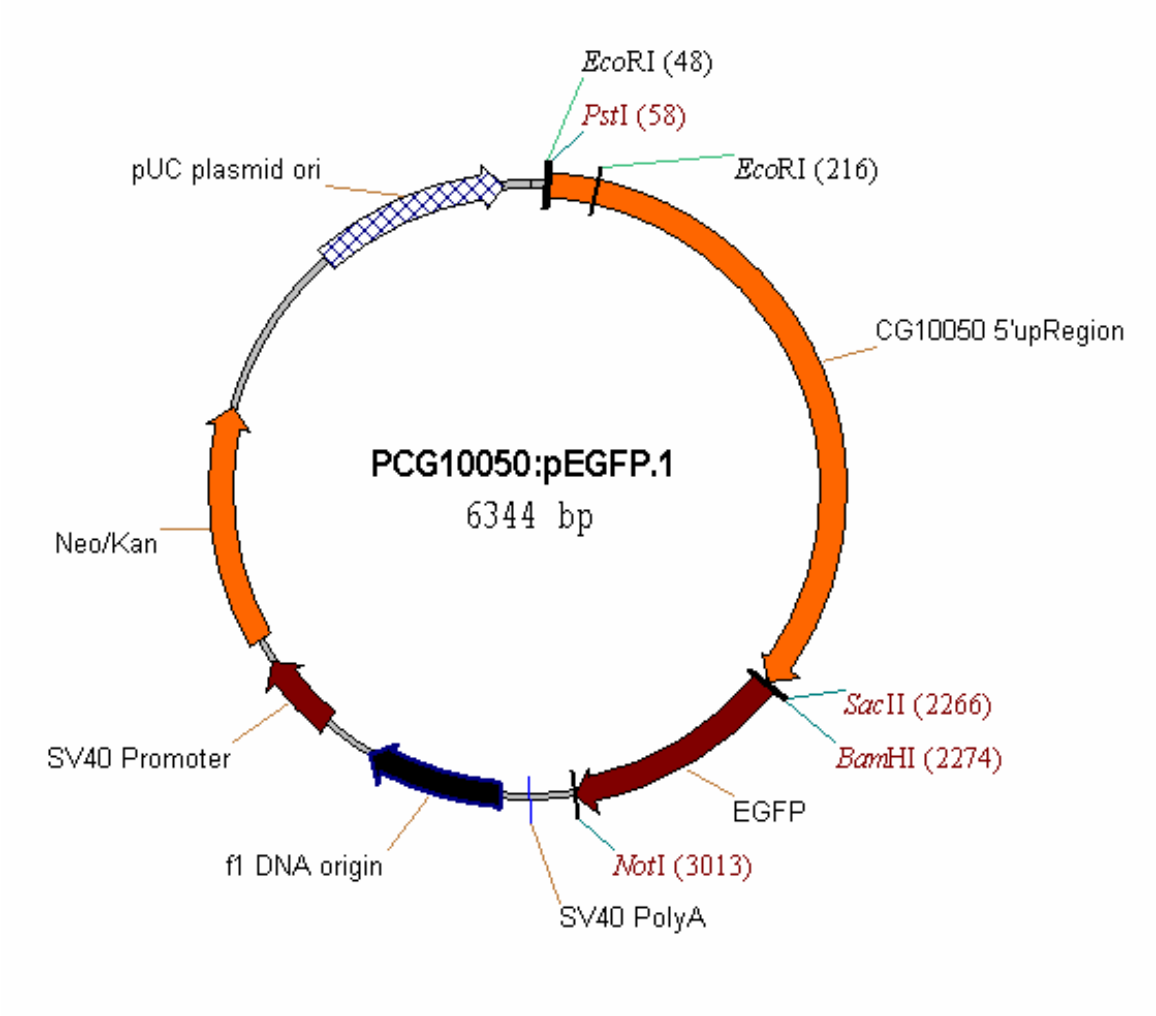

pBattBfa:[3xP3-DsRed2]
4600 bp
attB
3xP3
DsRed2
SV40 polyA
FseI (2362)
AscI (2333)

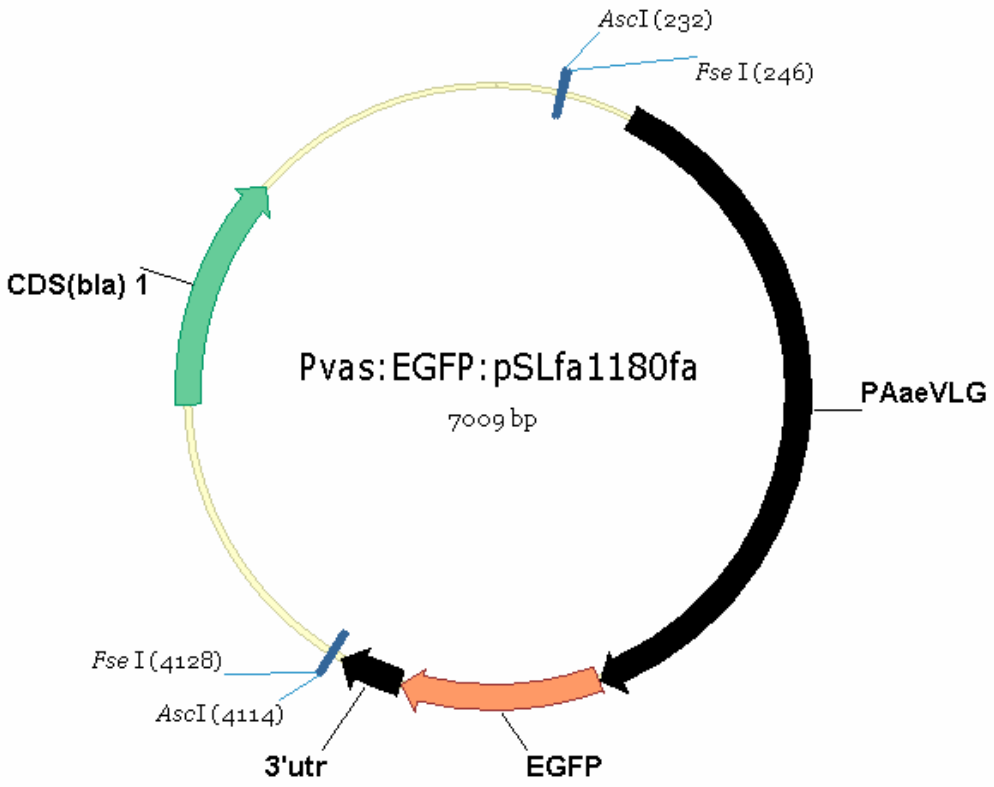

flanking seq
MarinerR
pMos[3XP3-DsRed2af]
8600 bp
3XP3
DsRed2
SV40 Poly A
flanking seq
AscI (3563)
Fse I (3577)
MarinerL

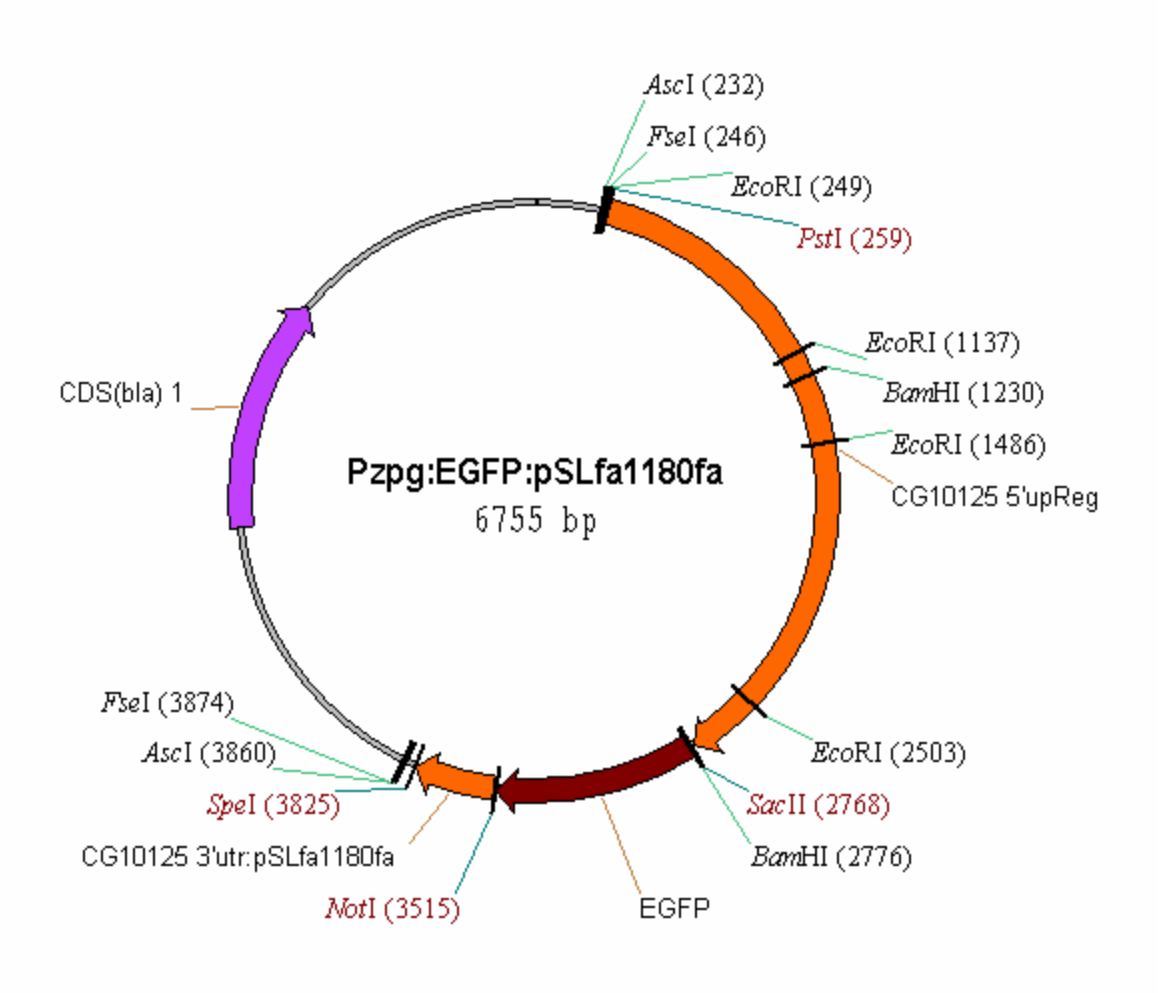

EcoRI (48)
PstI (58)
EcoRI (59)
pUC plasmid ori
AaeVLG5upReg
pAaeVLG:pEGFP.1
6755 bp
Neo/Kan
EcoRI (2253)
SpeI (2469)
BamHI (2684)
SV40 Promoter
f1 DNA origin
SV40 PolyA
EGFP
NotI (3424)

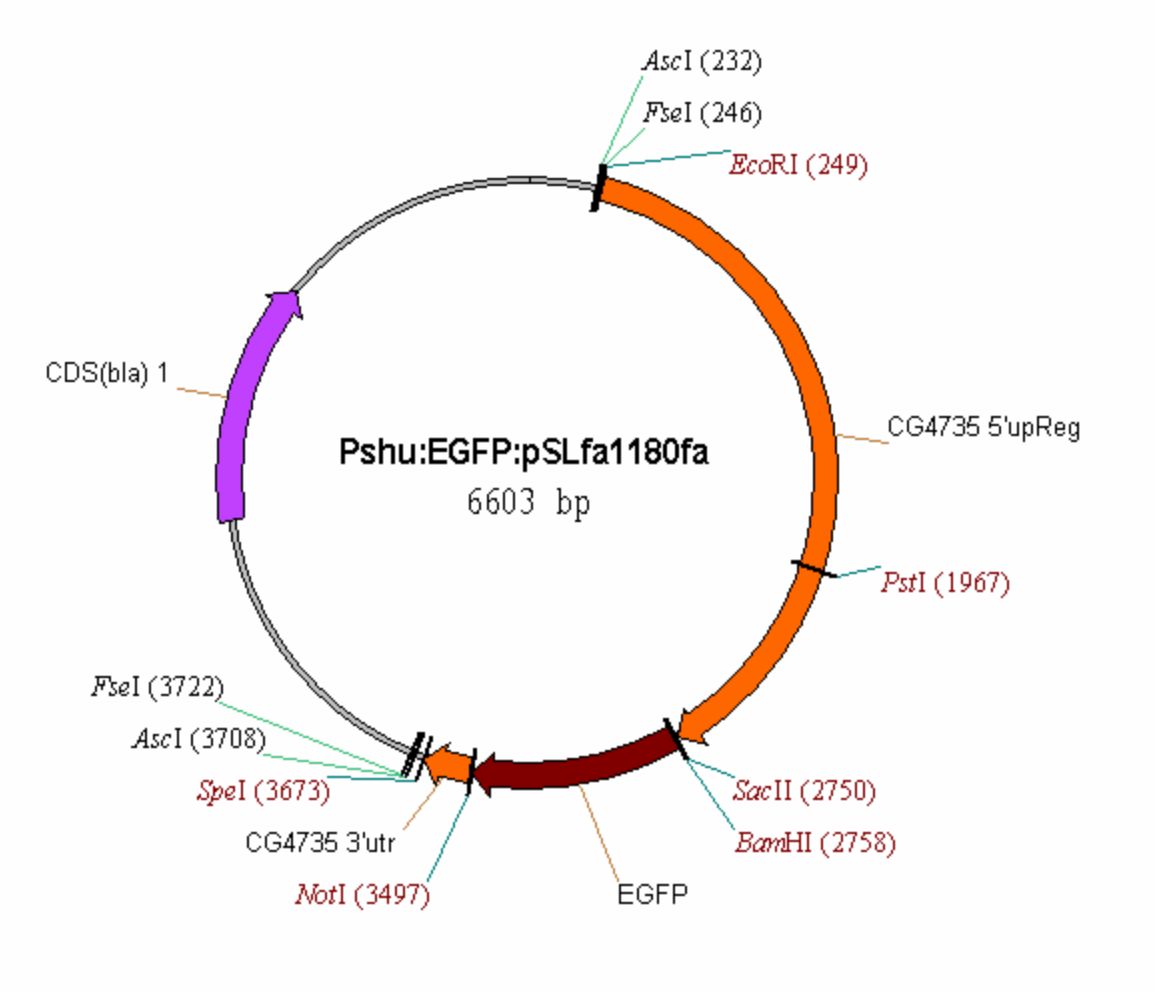

EcoRI (48)
PstI (58)
EcoRI (936)
BamHI (1029)
EcoRI (1285)
CG10125 5'upReg
EcoRI (2302)
SacII (2567)
BamHI (2575)
EGFP
NotI (3314)
SV40 PolyA
f1 DNA origin
SV40 Promoter
Neo/Kan
pUC plasmid ori
PCG10125:pEGFP.1
6645 bp

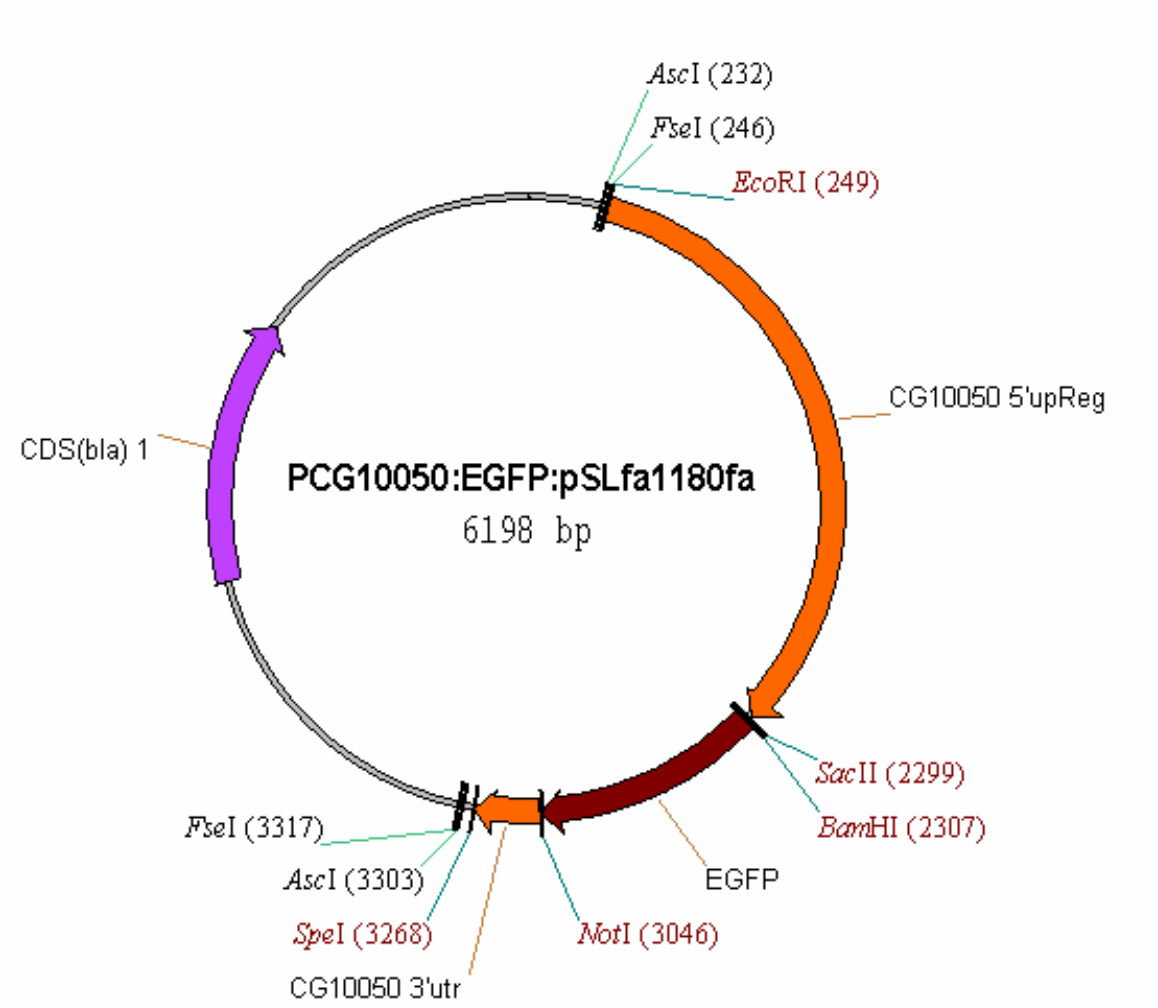

EcoRI (48)
pUC plasmid ori
CG4735 5'upReg
PCG4735:pEGFP.1
6627 bp
Neo/Kan
PstI (1766)
SV40 Promoter
SacII (2549)
BamHI (2557)
f1 DNA origin
SV40 PolyA
EGFP
NotI (3296)

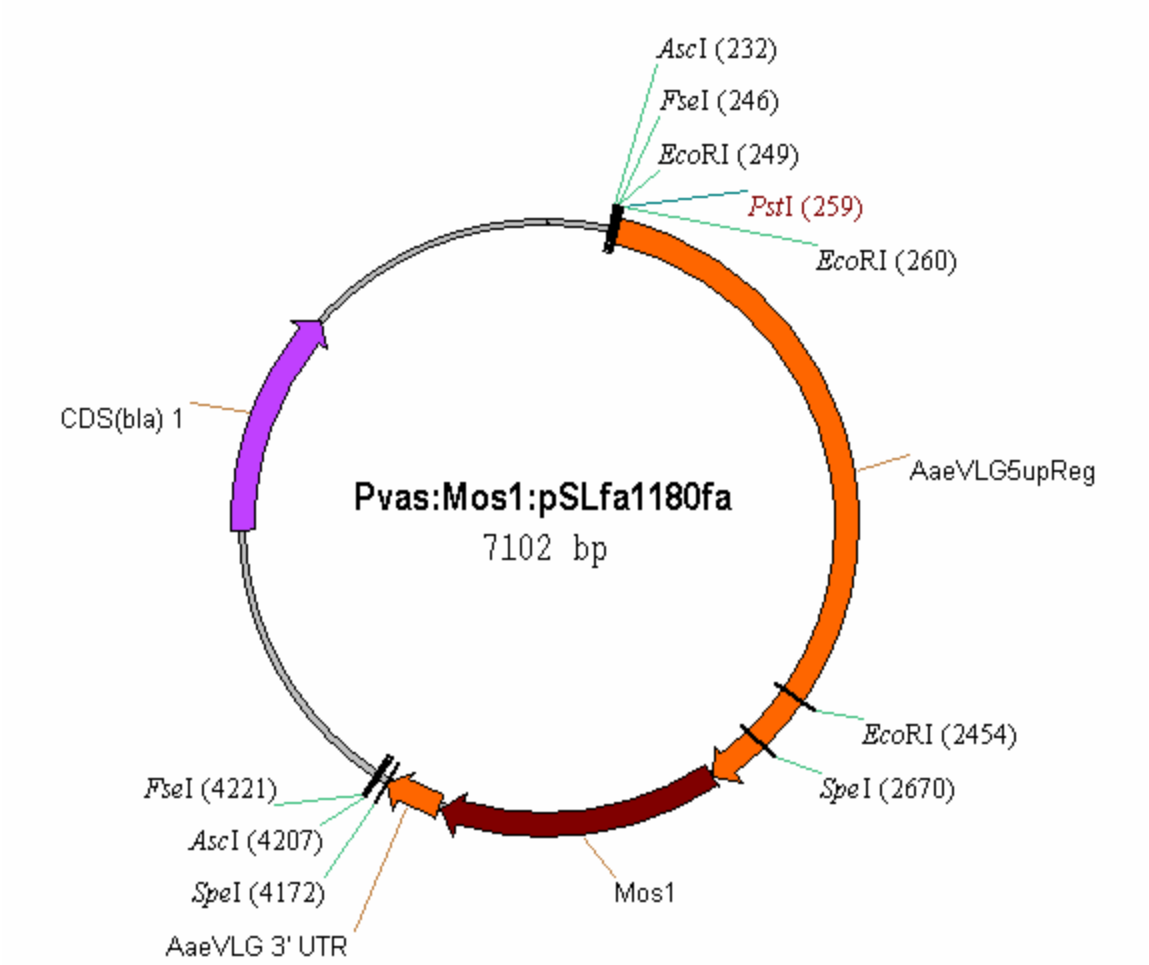

EcoRI (48)
PstI (58)
EcoRI (216)
pUC plasmid ori
CG10050 5'upRegion
PCG10050:pEGFP.1
6344 bp
Neo/Kan
SacII (2266)
BamHI (2274)
SV40 Promoter
EGFP
f1 DNA origin
NotI (3013)
SV40 PolyA

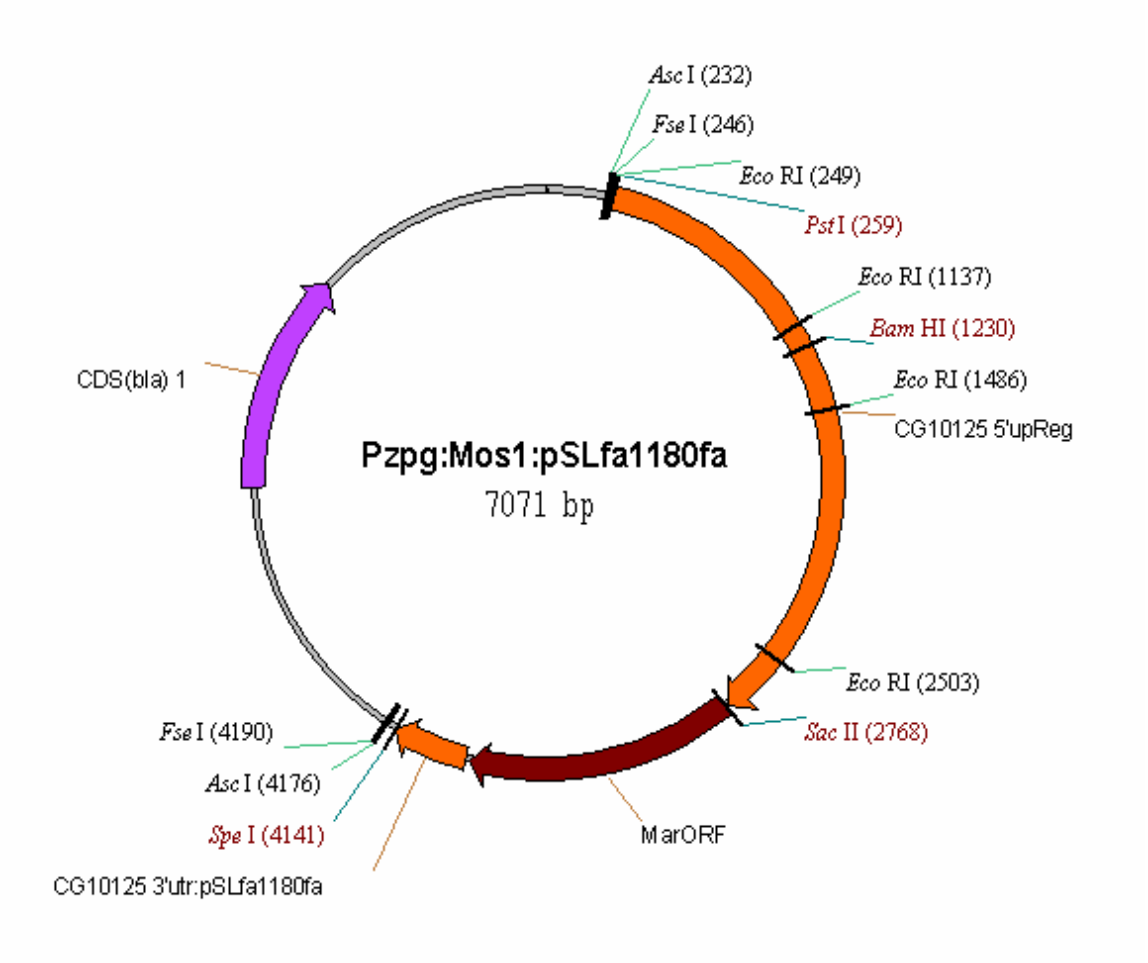

AscI (232)
Fse I (246)
CDS(bla) 1
Pvas:EGFP:pSLfa1180fa
7009 bp
PAaeVLG
Fse I (4128)
AscI (4114)
3'utr
EGFP

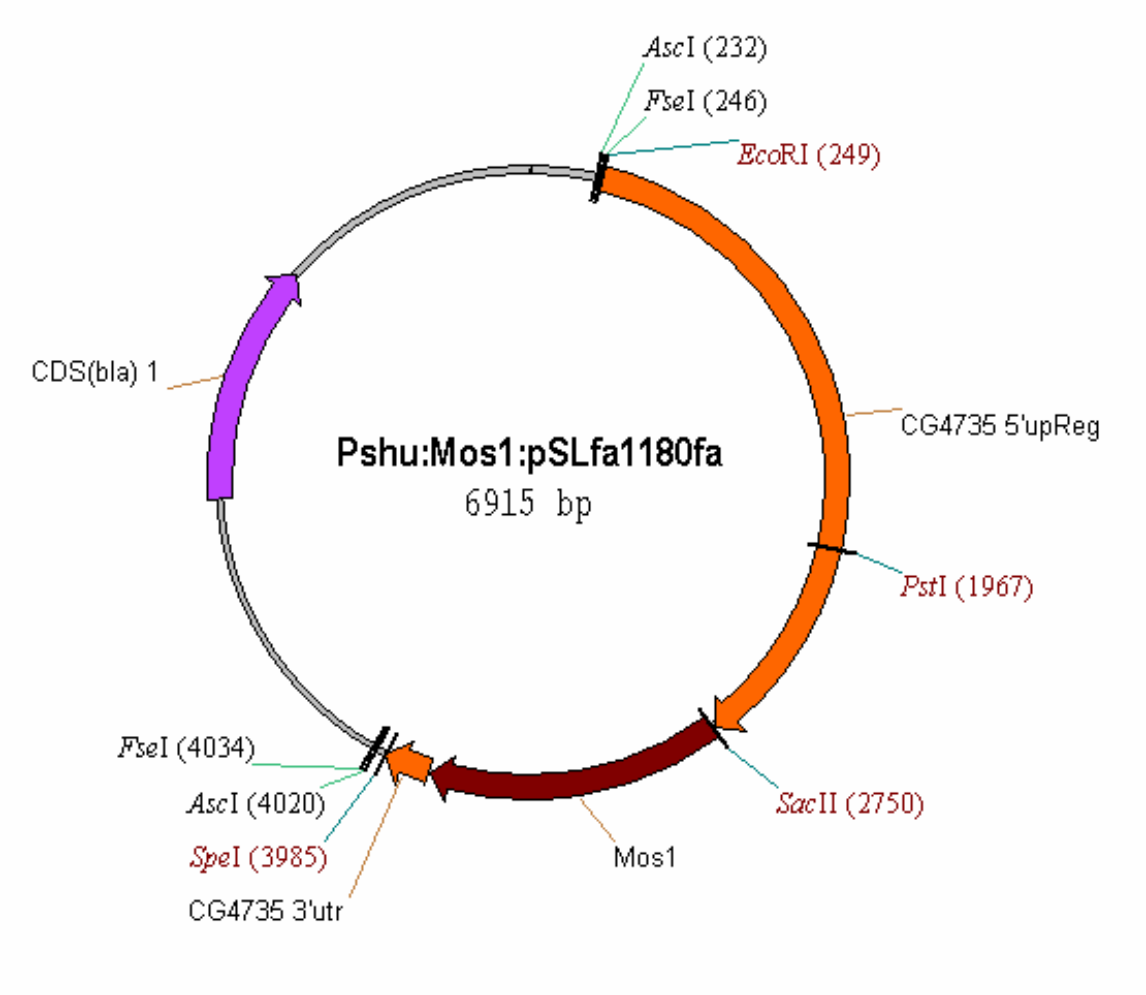

AscI (232)
FseI (246)
EcoRI (249)
PstI (259)
EcoRI (1137)
BamHI (1230)
EcoRI (1486)
CG10125 5'upReg
EcoRI (2503)
SacII (2768)
BamHI (2776)
EGFP
NotI (3515)
CG10125 3'utr:pSLfa1180fa
SpeI (3825)
AscI (3860)
FseI (3874)
CDS(bla) 1
Pzpg:EGFP:pSLfa1180fa
6755 bp

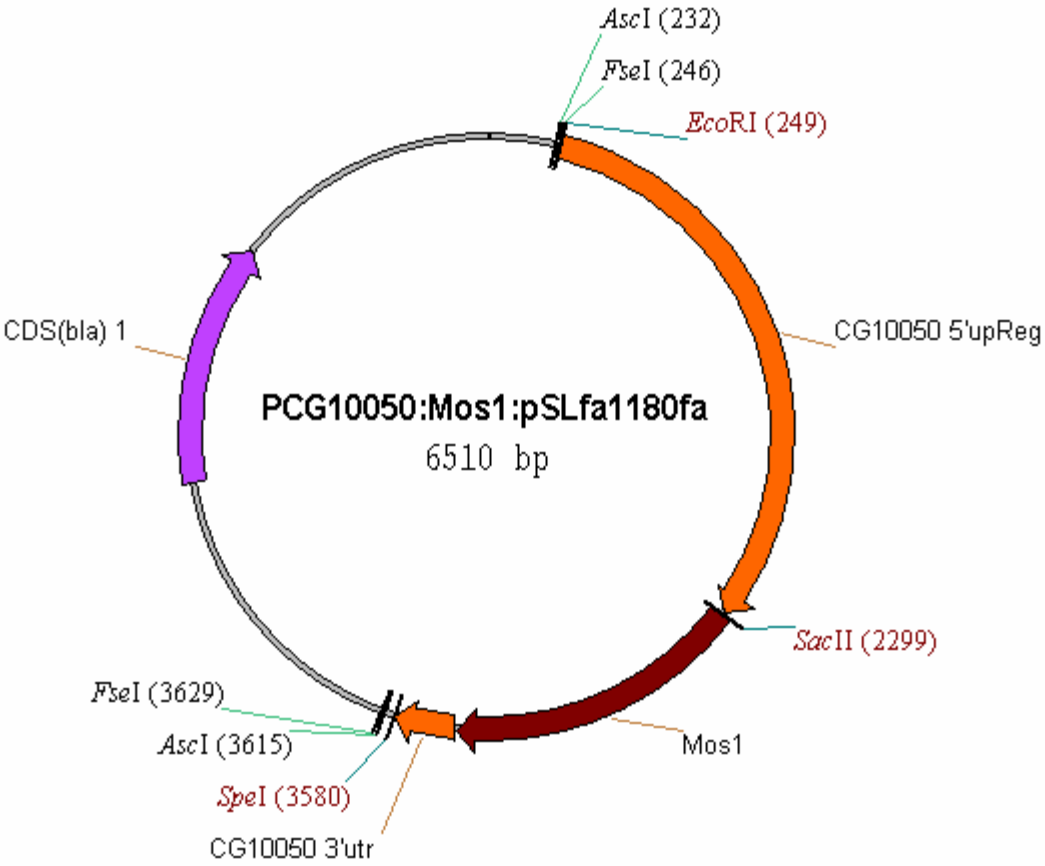

AscI (232)
FseI (246)
EcoRI (249)
CDS(bla) 1
CG4735 5'upReg
Pshu:EGFP:pSLfa1180fa
6603 bp
PstI (1967)
FseI (3722)
AscI (3708)
SpeI (3673)
SacII (2750)
CG4735 3'utr
BamHI (2758)
NotI (3497)
EGFP

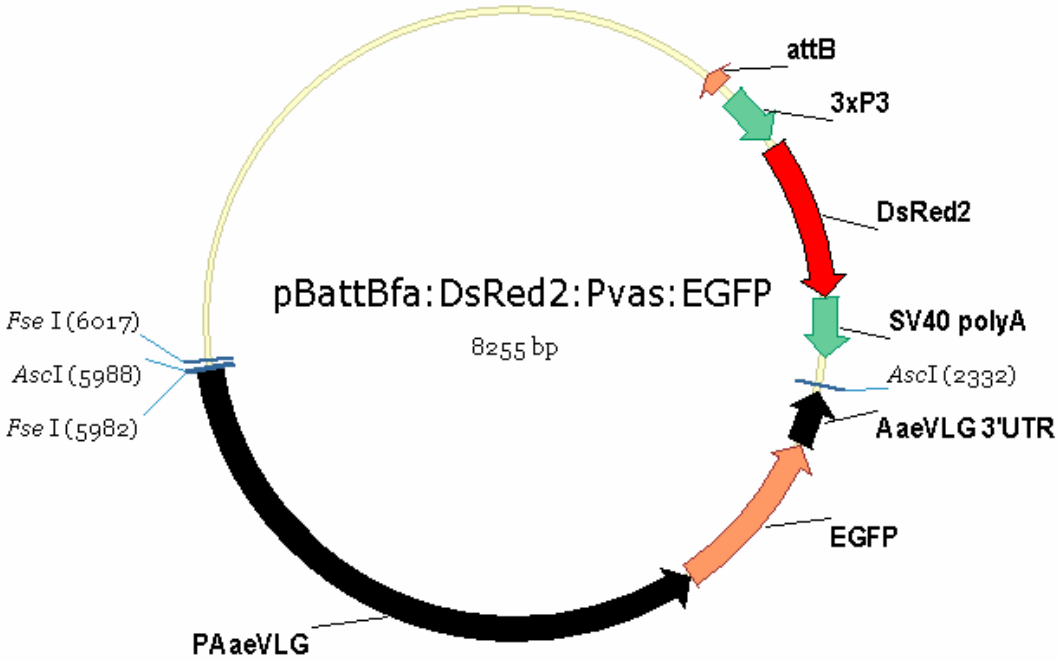

AscI (232)
FseI (246)
EcoRI (249)
CG10050 5'upReg
CDS(bla) 1
PCG10050:EGFP:pSLfa1180fa
6198 bp
SacII (2299)
BamHI (2307)
FseI (3317)
AscI (3303)
SpeI (3268)
CG10050 3'utr
NotI (3046)
EGFP

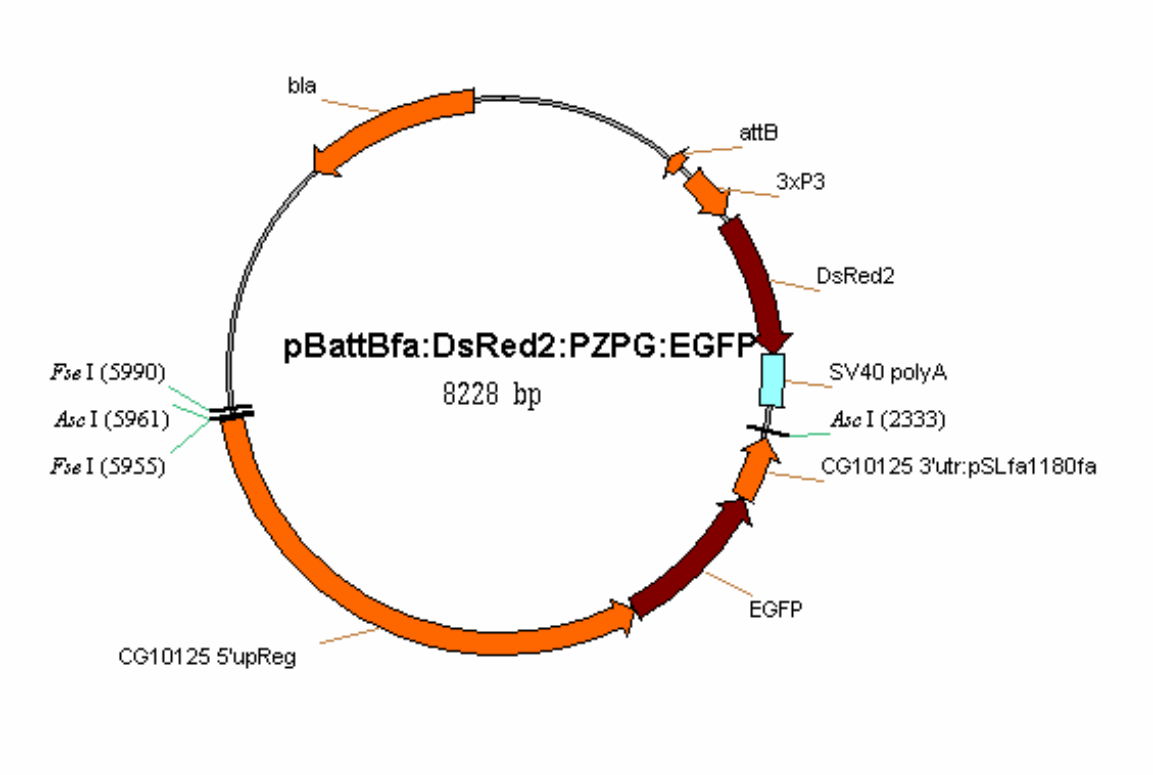

AscI (232)
FseI (246)
EcoRI (249)
PstI (259)
EcoRI (260)
AaeVLG5upReg
EcoRI (2454)
SpeI (2670)
Mos1
AaeVLG 3' UTR
SpeI (4172)
AscI (4207)
FseI (4221)
CDS(bla) 1
Pvas:Mos1:pSLfa1180fa
7102 bp

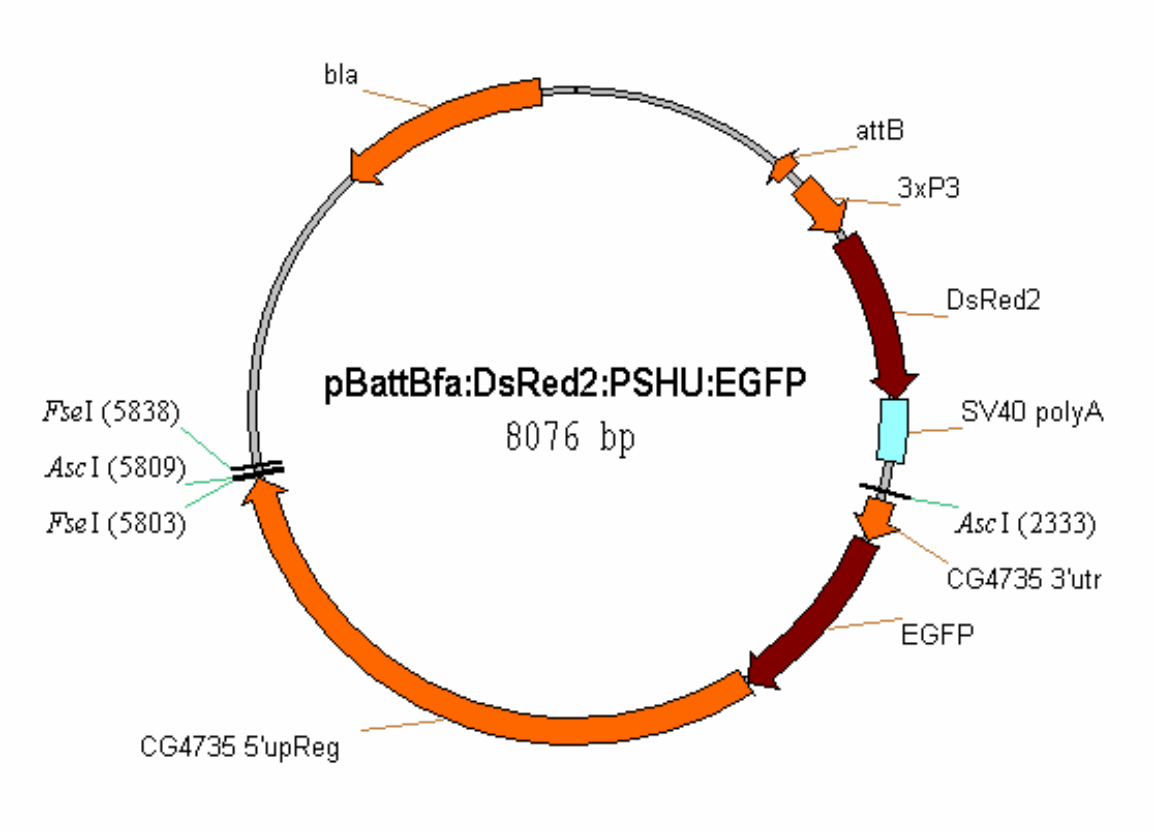

Asc I (232)
Fse I (246)
Eco RI (249)
Pst I (259)
Eco RI (1137)
Bam HI (1230)
Eco RI (1486)
CG10125 5'upReg
Eco RI (2503)
Sac II (2768)
MarORF
Fse I (4190)
Asc I (4176)
Spe I (4141)
CG10125 3'utr:pSLfa1180fa
CDS(bla) 1
Pzpg:Mos1:pSLfa1180fa
7071 bp

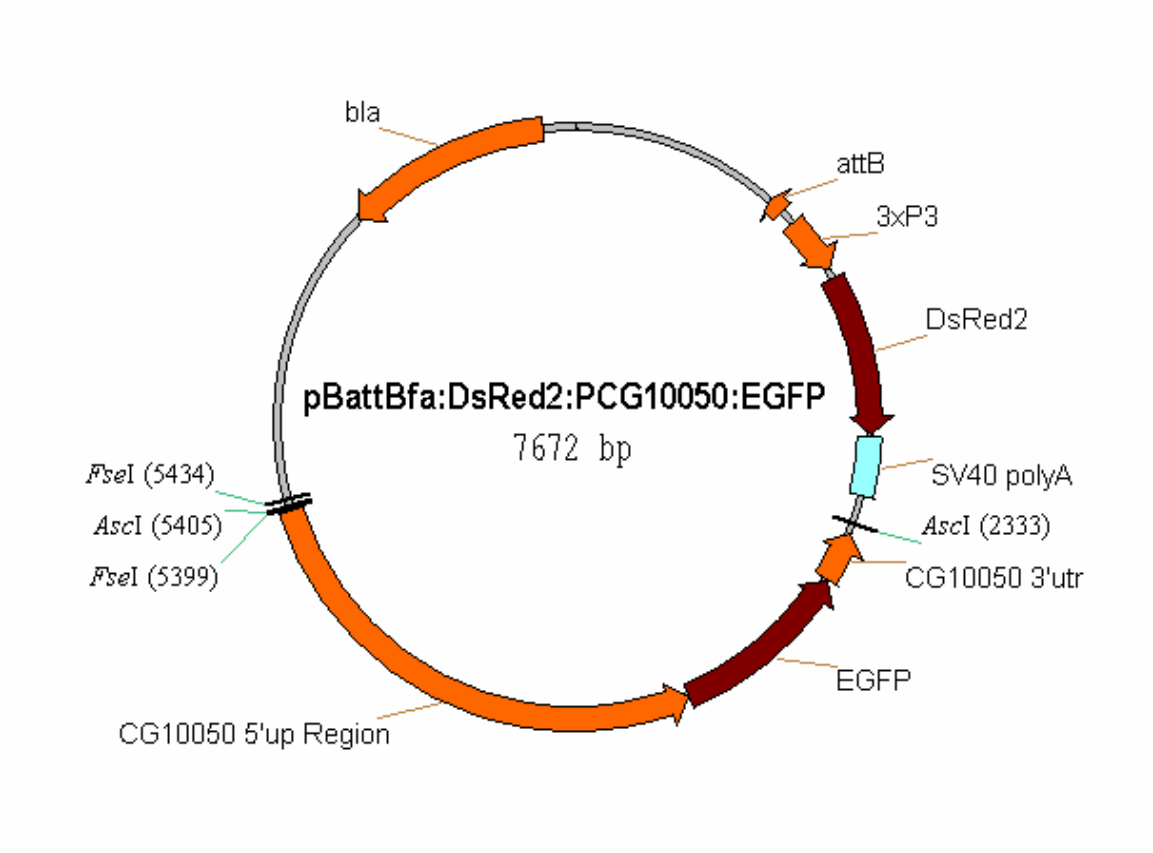

AscI (232)
FseI (246)
EcoRI (249)
CG4735 5'upReg
PstI (1967)
SacII (2750)
Mos1
CG4735 3'utr
SpeI (3985)
AscI (4020)
FseI (4034)
CDS(bla) 1
Pshu:Mos1:pSLfa1180fa
6915 bp

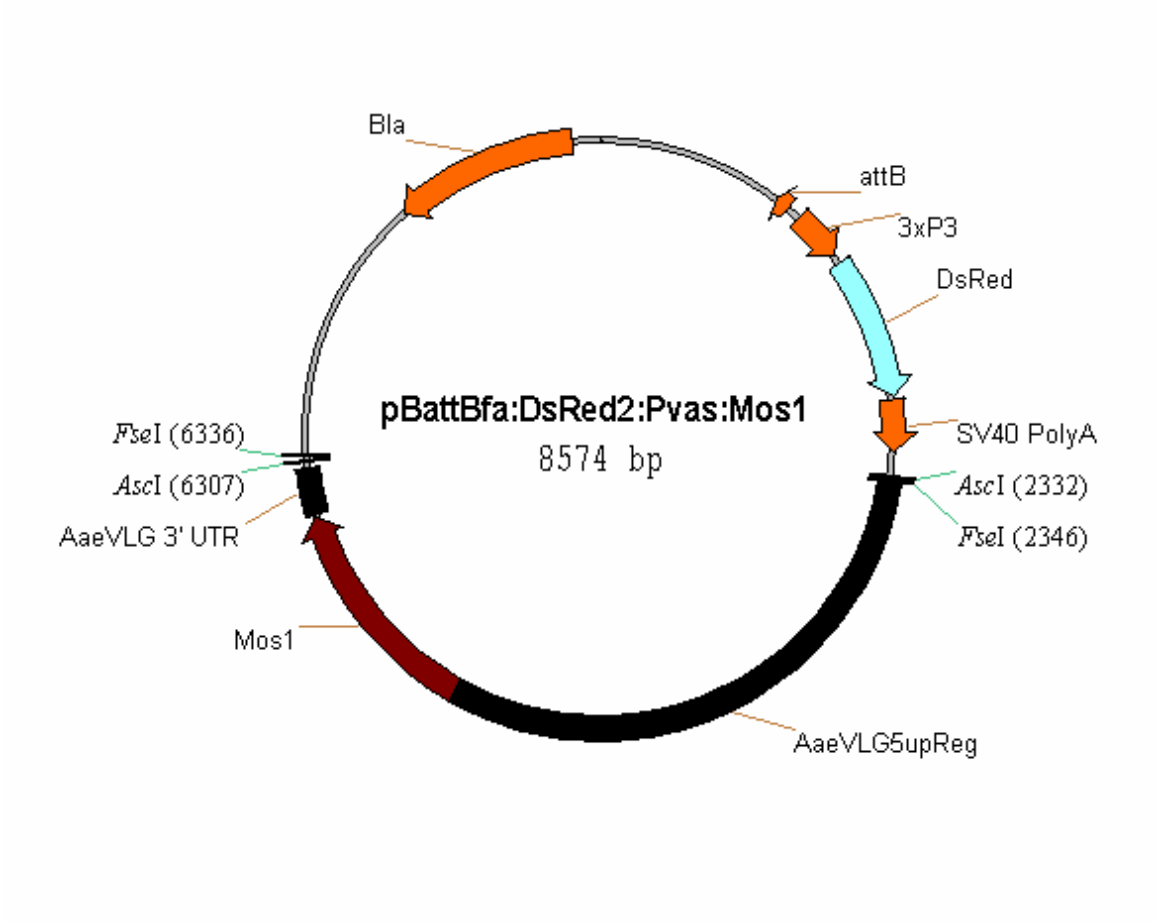

AscI (232)
FseI (246)
EcoRI (249)
CG10050 5'upReg
SacII (2299)
Mos1
SpeI (3580)
CG10050 3'utr
AscI (3615)
FseI (3629)
CDS(bla) 1
PCG10050:Mos1:pSLfa1180fa
6510 bp


attB
3xP3
DsRed2
SV40 polyA
AscI (2332)
AaeVLG 3'UTR
EGFP
PAaeVLG
Fse I (6017)
AscI (5988)
Fse I (5982)
pBattBfa:DsRed2:Pvas:EGFP
8255 bp


bla
attB
3xP3
DsRed2
pBattBfa:DsRed2:PZPG:EGFP
8228 bp
SV40 polyA
Fse I (5990)
Asc I (5961)
Fse I (5955)
Asc I (2333)
CG10125 3'utr:pSLfa1180fa
EGFP
CG10125 5'upReg


bla
attB
3xP3
DsRed2
SV40 polyA
AscI (2333)
CG4735 3'utr
EGFP
CG4735 5'upReg
FseI (5838)
AscI (5809)
FseI (5803)
pBattBfa:DsRed2:PSHU:EGFP
8076 bp


bla
attB
3xP3
DsRed2
pBattBfa:DsRed2:PCG10050:EGFP
7672 bp
SV40 polyA
AscI (2333)
CG10050 3'utr
EGFP
CG10050 5'up Region
FseI (5434)
AscI (5405)
FseI (5399)


Bla
attB
3xP3
DsRed
SV40 PolyA
AscI (2332)
FseI (2346)
AaeVLG5upReg
Mos1
AaeVLG 3' UTR
AscI (6307)
FseI (6336)
pBattBfa:DsRed2:Pvas:Mos1
8574 bp

VITA

Name:	Darren Erich Hagen
Address:	Department of Entomology, TAMU 2475, Texas A&M University, College Station, TX 77843-2475
E-mail Address:	dehagen@tamu.edu
Education:	B.S., Biology, Angelo State University, May 2001 Ph.D., Genetics, Texas A&M University, August 2007