DISSERTATION

AEDES AEGYPTI AND DENGUE VIRUS INVESTIGATION OF ANATOMIC, GENOMIC, AND MOLECULAR DETERMINANTS OF VECTOR COMPETENCE

Submitted by

Scott Arthur Bernhardt

Department of Microbiology, Immunology and Pathology

In partial fulfillment of the requirements

for the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, CO

Spring 2009

COLORADO STATE UNIVERSITY

April 6, 2009

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER SUPERVISION BY SCOTT ARTHUR BERNHARDT ENTITLED "*AEDES AEGYPTI* AND DENGUE VIRUS INVESTIGATION OF ANATOMIC, GENOMIC, AND MOLECULAR DETERMINANTS OF VECTOR COMPETENCE" BE ACCEPTED AS FULFILLING IN PART THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Ph.D. Graduate Committee

Carol D Blair
(Co-Advisor)

William C. Black IV
(Co-Advisor)

Barry Beaty

Edward A. Hoover
(Department Head)

ABSTRACT OF DISSERTATION

AEDES AEGYPTI AND DENGUE VIRUS INVESTIGATION OF ANATOMIC, GENOMIC, AND MOLECULAR DETERMINANTS OF VECTOR COMPETENCE

Dengue (DENV) causes one of the most rapidly expanding diseases in the tropics. Vector competence (VC) in *Aedes aegypti* for DENV-2 is a quantitative trait and has been shown to be highly variable. Questions remain as to whether variation in VC continues to exist after the primary field observation. What genetic factors contribute to VC and do these factors evolve from arbovirus exposure remain unclear.

To determine whether variation in VC phenotype is temporally stable which had previously been characterized, I collected *Ae. aegypti* eggs in 2005 from 19 sites in Mexico where VC had been studied previously and determined VC for DENV-2. VC in these collections ranged from 38-83%, the midgut infection barrier rates ranged from 9-30% and the midgut escape barrier (MEB) rates ranged from 7-36%. Spatial variations in VC continued to exist between collection sites; however, similar relative VC rates were found in specific sites when analyzed over time. During this study I also detected the presence of nuclear copies of mitochondrial DNA in *Ae. aegypti*.

I hypothesized that the RNAi and apoptosis pathways significantly contribute to DENV-2 VC in *Aedes aegypti*. We constructed an F_1 intercross family using a P_1 female from *Aedes aegypti aegypti* (Aaa) and a P_1 male *Aedes aegypti formosus* (Aaf). Recombination rates in the F_2 offspring were significantly reduced and genotype ratios suggested a deleterious recessive allele on chromosome 3. The F_2 linkage map was

incongruent with the established map for Aaa. Recombination rates and gene order were consistent with the presence of multiple chromosomal inversions in Aaf on all three chromosomes. Genotype-phenotype analyses demonstrated that the three RNAi pathway genes we examined significantly contribute to the MEB for DENV-2.

The final research aim was to determine the rate at which RNAi genes are evolving and to determine whether arbovirus infections contribute to RNAi evolution in *Aedes aegypti*. Sequence data were obtained from three RNAi genes and their three microRNA paralogs from 94 mosquitoes from five geographically widespread collection sites. Rates of non-synonymous to synonymous amino acid substitutions and phylogenetic analysis showed that the RNAi genes are evolving faster than their microRNA paralogs. A significant correlation was also identified between *Aedes aegypti dcr2* average number of nucleotide differences and DENV-2 MEB.

Scott Arthur Bernhardt
Department of Microbiology,
Immunology and Pathology
Colorado State University
Fort Collins, CO 80523
Spring 2009

ACKNOWLEDGMENTS

First and foremost, I would like to acknowledge and thank my wife, Korryn. She is my best friend, an example, one who I can confide in and one I have turned to for strength. I thank her for the guidance, unconditional love, and support as countless hours were spent in the laboratory and behind a computer. I must also acknowledge and thank my children Brigham Scott, Spencer Reid and Leigha Elizabeth. Many nights and weekends were spent working on this research and away from home. They were understanding and continue to demonstrate unconditional love.

I would like to thank my advisors, Drs. Carol D. Blair and William C. Black for their guidance and support. Countless hours were spent mentoring, developing databases, conducting and discussing analyses, and reading and revising drafts. Their dedication and direction has helped me to become a better scientist. I would also like to thank Dr. Barry J. Beaty. Without his guidance, mentoring, and recruitment I would have never considered attending Colorado State University. I would also like to acknowledge Drs. Kenneth E. Olson and Michael Antolin for their time, commitment and insightfulness while serving on my graduate committee.

Finally, I would like to thank the many research associates, post-docs, graduate and undergraduate students in the Arthropod and Infectious Disease Laboratory and the Black Laboratory. They provided numerous hours of assistance, support and moments of laughter to make this research happen and memorable. Thank you!

TABLE OF CONTENTS

Chapter 1

Literature Review

Introduction

Arthropod-borne diseases significantly impacted human civilization long before arthropods were known to be efficient vectors of disease. Malaria has been recorded in human populations in China as far back as 2700 BC (Cox, 2002). Plague, most notably the Black Death, occurred as multiple pandemics beginning in the mid-fourteenth century (Pollitzer, 1960; Gage and Kosoy, 2005). Dengue virus-like symptoms were recorded during the Chin Dynasty (265 to 420 A.D.) as a water poison that was thought to be connected with flying insects (Nobuchi, 1979). Yellow fever, filariasis, louse-borne typhus, trypanosomiasis, and leishmaniasis are just a few vector-borne diseases that have been responsible for significant human disease and death (Gubler, 1991). Even with sizeable medical and technological advances, emerging and re-emerging arthropod-borne diseases continue to cause mortality and significantly impact human populations.

Mosquito-borne parasites have had a considerable resurgence since the 1970s. Malaria is the most important vector-borne disease due to its global distribution and the number of people affected worldwide. Malaria causes illness in at least 300 million people each year, with an annual mortality of at least 1 million. Ninety percent of all deaths due to malaria occur in Africa (WHO 2009, http://www.rbm.who.int).

Mosquito-borne viruses as a group, however, are as important, if not more so than malaria in terms of numbers of infections, varieties of vector species and global distribution. Arboviruses, particularly dengue virus, yellow fever virus, chikungunya, and West Nile virus, have experienced a global resurgence in the last three decades. Dengue is the second most important tropical disease with approximately 50 to 100 million cases of dengue fever and 500,000 cases of dengue hemorrhagic fever (DHF)

each year (Gubler, 1998). Cases have been recorded throughout the tropics and impact more than 2 billion people worldwide. Mosquito management and public health education are the only methods available for dengue control. Prior to the first yellow fever vaccine becoming available in the 1930s, thousands of deaths were recorded worldwide (Barrett and Higgs, 2007). But even with a highly effective and safe vaccine available for yellow fever, present day epidemics still occur in parts of Africa and South America incurring significant mortality (Gubler, 2004; Roberts, 2007; Monath, 2006). In the last decade, West Nile virus has made a resurged in Europe and emerged in North America (Hubalek and Halouzka, 1999; Lanciotti *et al.*, 1999; Nash *et al.*, 2001). West Nile virus has traditionally only caused severe disease in the elderly, the young, and the immunocompromised. This recent re-emergence is unique in that some previously healthy individuals affected by the virus experienced severe neurological complications and, in some cases, acute flaccid paralysis (Sejvar *et al.*, 2006; Saad *et al.*, 2005).

Numerous factors have contributed to the re-emergence of vector-borne diseases. Increases in insecticide and drug resistance, virus evolution, and alteration in global climate have all directly impacted transmission of vector-borne diseases (Vora *et al.*, 2008; Harrus *et al.*, 2005; Gubler, 2004; Guzman and Kouri, 2003). Additional problems are lack of a pro-active, prevention-based public health policy with substantial infrastructure, as well as the decline of effective mosquito control programs in the past 40 years (Gubler, 2004, 1997; Gubler and Trent, 1994). Global demographic and societal changes have also significantly affected emergence and resurgence of vector-borne diseases. Unprecedented global population growth, along with unplanned and uncontrolled urbanization in tropical developing countries, has created ideal conditions

for increased vector densities and vector-borne disease transmission potential. Relative ease of global travel has also contributed to the transmission potential of vectors and disease pathogens.

Pro-active mosquito management programs are proven to effectively reduce arbovirus transmission. Insecticides are often the first approach to mosquito management. Insecticide resistance, however, must always be considered in any mosquito control program (Saavedra-Rodriguez *et al.*, 2008; Rodriguez *et al.*, 2007). Effective control programs require unique and innovative approaches such as changes to urban landscape and pro-active management of mosquito larval populations. For this reason, it is important to continue arbovirus vector research. Research leading to an understanding of the genetics of the vector and arbovirus with respect to infection, dissemination, and transmission is particularly important as it may provide avenues to new methods of disease reduction.

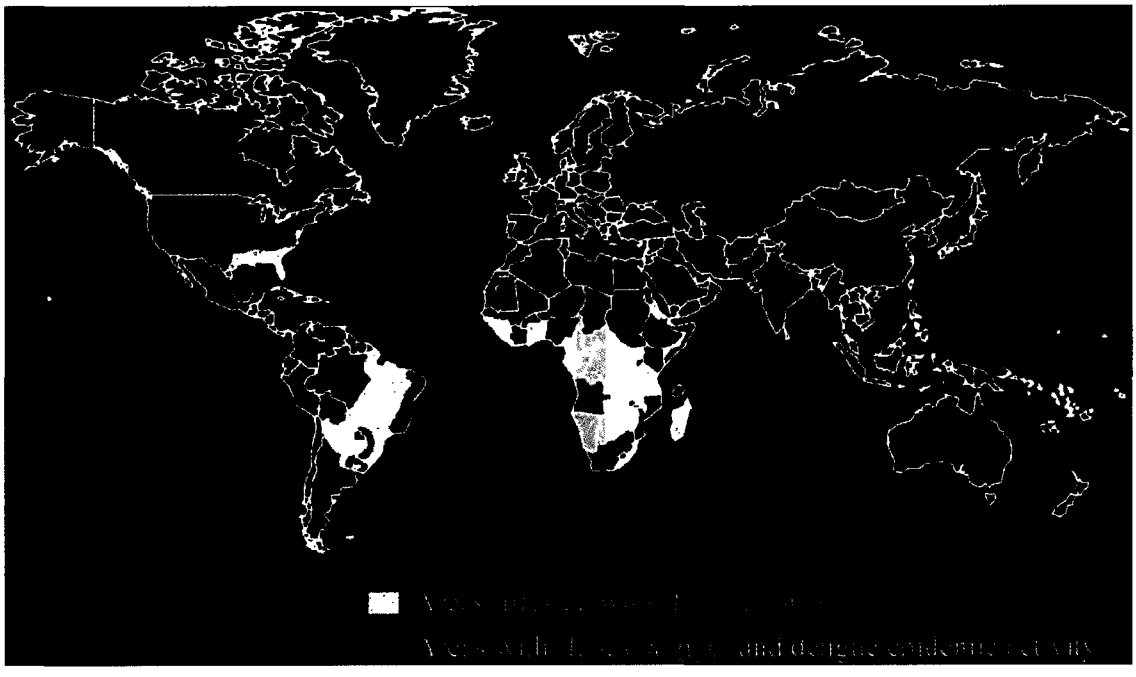
Recent advances in molecular genetic techniques have given researchers new ways for studying arbovirus transmission potential in mosquitoes (Carlson *et al.*, 1995; Beaty, 2000; Black *et al.*, 2001; Saavedra-Rodriguez *et al.*, 2008; Obbard *et al.*, 2005; Urdaneta-Marquez *et al.*, 2008; Mori *et al.*, 2008). Understanding the mechanisms of mosquito vector competence for specific viruses can lead to the development of potential targets for use in vector control. For example, researchers have determined that mosquitoes use RNA silencing as an innate immune response (Sanchez-Vargas *et al.*, 2009). From these findings, genetically modified mosquitoes have been developed that are resistant to DENV-2 (Franz *et al.*, 2006).

Numerous questions remain concerning genetically modified organism strategies, including the ethical issue of releasing transgenic mosquitoes and effectiveness of this approach to lowering vector competence in an actual field population. Limited understanding is also available as to why vector competence is variable among mosquito populations. Investigators would like to know which genetic loci are involved in the quantitative traits of vector infection and dissemination and what drives this variation. These questions, along with many others, demonstrate the importance of conducting research on vectors at the molecular level and within natural field populations. An understanding of how mosquito mechanisms and pathways work within and among populations is necessary to develop new and effective mosquito and disease control tools.

Dengue epidemiology

Dengue is a re-emerging disease and is currently the most important arthropod-borne viral disease (Figure 1.1) (Gubler, 2004; Guzman and Kouri; 2003). Dengue viruses and the vector, *Aedes aegypti*, have a worldwide distribution in the tropics and have the potential to significantly impact more than 2.5 billion humans (Gubler, 2004; Gubler and Clark, 1995; Halstead, 1980). Numerous epidemics of dengue-like illness have been recorded for the French West Indies (1635), Panama (1699), Indonesia, and Egypt (1779) in the 17^{th} and 18^{th} centuries (Howe, 1977; McSherry, 1982). The first recognized and well-recorded dengue epidemics occurred nearly simultaneously in Asia, Africa, and North America in the 1780s, shortly after the identification and naming of the disease in 1779 (Carey, 1971). Benjamin Rush, of Philadelphia, definitively described dengue in a case report as "break-bone fever" (Rush, 1809). Dengue fever has traditionally been considered a tropical disease that seriously

Figure 1.1. World map demonstrating global geographic distribution of dengue/dengue hemorrhagic fever and *Ae. aegypti* (Gubler, 2004).



impacts developing countries. Recently, however, changes in urbanization, travel, and climate have contributed to dengue epidemics in countries such as Australia, Easter Island, Chile, Taiwan, Pakistan and India (Caceres *et al.*, 2008; Hanna *et al.*, 1998; Paul *et al.*, 1998; Chen *et al.*, 2008; Malla *et al.*, 2008).

Dengue virus infections in humans can cause illnesses ranging from mild or asymptomatic infection to severe and fatal hemorrhagic disease (Guzman and Kouri, 2002; Gubler, 1998). Dengue virus includes four genetically related, but antigenically distinct serotypes (DENV-1, DENV-2, DENV-3, DENV-4). Infection with any of these four serotypes can cause a similar clinical presentation that may vary in severity. The incubation period ranges from three to 14 days and is often clinically difficult to diagnose in dengue endemic regions (Sabin, 1952; Siler *et al.*, 1926; Gubler, 2004). Dengue fever is the classic and mild form of the disease. It is primarily a disease of older children and adults. Dengue fever is characterized by a sudden onset of fever and variety of other symptoms ranging from frontal headache, retro-orbital pain, body aches, nausea, vomiting, joint pain, weakness, and rash (Hayes and Gubler, 1992; Waterman and Gubler, 1989). The fever temperature often rises to 102° to 105°F and may last for two to seven days. During this time, the fever may drop after a few days, only to rebound 12 to 24 hours later.

Dengue hemorrhagic fever and dengue shock syndrome (DHF/DSS) are the more severe, potentially fatal forms of the disease. Their onset is often indistinguishable to the mild form, dengue fever, beginning with sudden onset of fever that may last for two to seven days with a variety of additional symptoms as described with dengue fever. No identifying clinical characteristics of DHF/DSS usually occur during the acute stage of

disease. As the fever begins to drop, characteristic manifestations of circulatory failure and plasma leakage appear (Gubler, 1998). Bleeding occurs in the skin (petechiae) and occasionally in the gums and nose. Dengue shock syndrome (DSS) is caused by plasma leakage and loss of blood volume into the extravascular spaces. Plasma leakage, hemorrhages and complement activation are all distinguishing features of DHF/DSS (Gubler and Clark, 1995; Halstead, 1983; Rothman and Ennis, 1999). The World Health Organization has classified DHF into four grades of severity, where grades III and IV are considered to be DSS (WHO, 1997). The presence of thrombocytopenia with concurrent hemo-concentration differentiates grades I and II DHF from dengue fever.

Grade I: Fever accompanied by non-specific constitutional symptoms; the only hemorrhagic manifestation is a positive tourniquet test and/or easy bruising.
Grade II: Spontaneous bleeding in addition to the manifestations of Grade I patients, usually in the forms of skin or other hemorrhages.
Grade III: Circulatory failure manifested by a rapid, weak pulse and narrowing of pulse pressure or hypotension, with the presence of cold, clammy skin and restlessness.
Grade IV: Profound shock with undetectable blood pressure or pulse.

Limited understanding exists as to why some individuals who become infected progress to DHF/DSS, while others do not. Currently two bodies of literature describe dengue pathogenesis, and these are not mutually exclusive. Halstead (1970) originally described a model, known as antibody-dependent enhancement (ADE), where patients are at significantly higher risk for developing DHF/DSS if they acquire a secondary dengue infection with a heterologous dengue virus serotype (Halstead, 1970, 1980, 1988). During the second dengue infection, pre-existing heterologous dengue antibodies recognize the new infecting virus and form antigen-antibody complexes. The hypothesis is that the first dengue infection, acts to enhance (ADE) the secondary infection and

replication of dengue in cells of the mononuclear cell lineage (Brandt *et al.*, 1982; Halstead, 1988; Halstead *et al.*, 1984). The virus is not neutralized and is free to replicate inside the macrophages because original recognizing antibody is heterologous. These cells produce and secrete vascular mediators in response to infection, which ultimately increase vascular permeability and potential for shock (Kurane, 2007; Kurane and Takasaki, 2001).

The second model describing increased pathogenesis is specific to virus virulence and evolution. Dengue viruses vary and change genetically, due to selection pressures, as they replicate in humans and mosquitoes. Through virus evolution and virulence studies, Rico-Hesse (2003) has identified specific dengue virus genotypes more commonly associated with DHF/DSS (Rico-Hesse, 2003; Rico-Hesse *et al.*, 1997). Genetic changes can amount to phenotype changes such as increased virus replication, virulence, and epidemic potential. Many researchers believe that DHF/DSS is a combination of both theories. The process is initiated by a second infection with a more virulent dengue virus, in the presence of antibodies that enhances the dengue virus infection. This second dengue infection triggers an intense immune response due to ADE from antibodies resulting from the primary infection, as well as being infected with a more virulent dengue genotype virus.

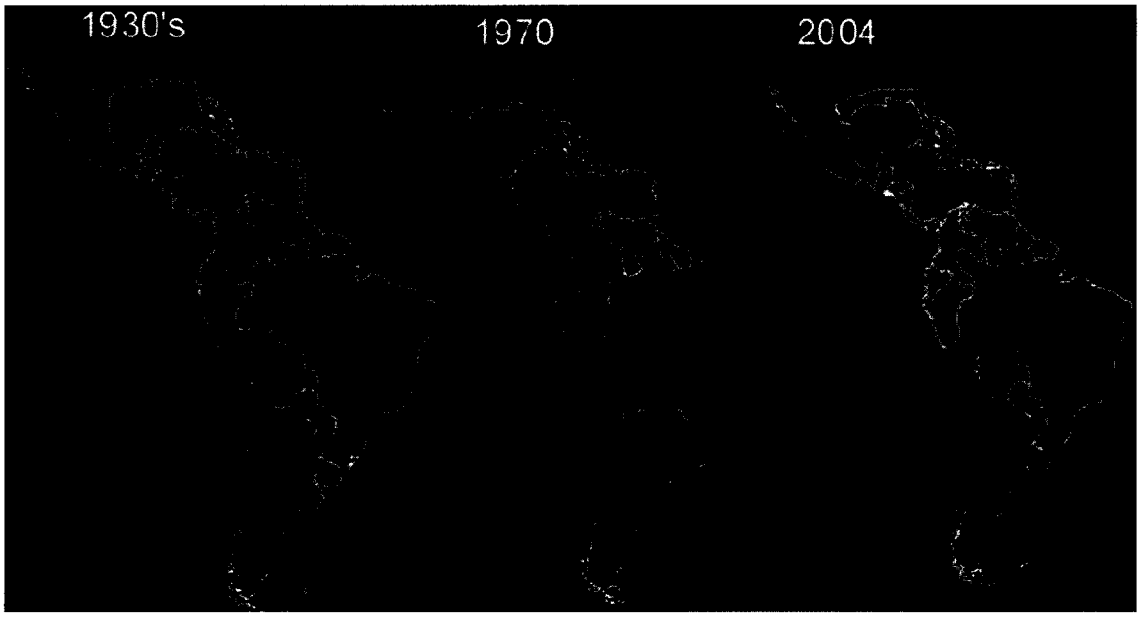
There are currently two recognized *Ae. aegypti* dengue transmission cycles. Sylvatic dengue viruses are transmitted in an enzootic cycle between nonhuman primates (*Macaca*, *Presbytis*, *Erythrocebus patas* monkeys) and arboreal *Aedes* mosquitoes in Southeast Asia and West Africa (Vasilakis *et al.*, 2007a and 2007b). Even though little evidence exists that these viruses move out of the forest to urban areas, the potential

remains as a source for dengue re-emergence (Rico-Hesse, 1990; Vasilakis and Weaver, 2008). The second transmission cycle is the urban human *Ae. aegypti* cycle. This is the more important public health transmission cycle because the virus is maintained in urban settings and has been linked to endemic and epidemic dengue disease.

Humans are infected with dengue virus by the bite of an infected *Ae. aegypti* or *Ae. albopictus* mosquito. *Ae. aegypti*, the principal vector in the domestic urban transmission cycle, is a highly domesticated tropical mosquito that prefers to lay its eggs in artificial containers found in and around homes. The female mosquito preferentially feeds on human blood to promote vitellogenesis (Harrington *et al.*, 2001; Scott *et al.*, 1997). *Ae. aegypti* are nervous feeders and often feed on several persons in a short time during a single blood meal (Platt *et al.*, 1997; Putnam and Scott, 1995). This feeding behavior makes *Ae. aegypti* a very efficient vector during epidemics. The mosquito has the opportunity to acquire the virus from an infected human during the acute febrile period, at which time the virus is circulating in the peripheral blood (Gubler *et al.*, 1981). The virus then undergoes an extrinsic incubation period of 10 to 14 days in the mosquito (Sanchez-Vargas *et al.*, 2009). If the mosquito survives beyond this extrinsic incubation period, the mosquito has the opportunity to release the virus in the saliva when taking a new blood meal or probing a naïve human.

Dengue virus infections had virtually vanished in the Americas from 1946 to the late 1970s as a result of successful *Ae. aegypti* eradication campaigns conducted by the Pan American Health Organization (PAHO) (Figure 1.2). These programs focused on mosquito larval control using source reduction and insecticide spraying, primarily dichlorodiphenyltrichloroethane (DDT). Changes in economic, environmental, and

Figure 1.2. Map of the Americas demonstrating geographic distribution of *Ae. aegypti* in 1930, 1970 and 2004 (Gubler, 2004).



public health policies in the 1970s resulted in discontinuation of the mosquito eradication programs. *Ae. aegypti* has since been able to quickly re-infest and expand from its pre-1940s original habitat (Gubler, 2004).

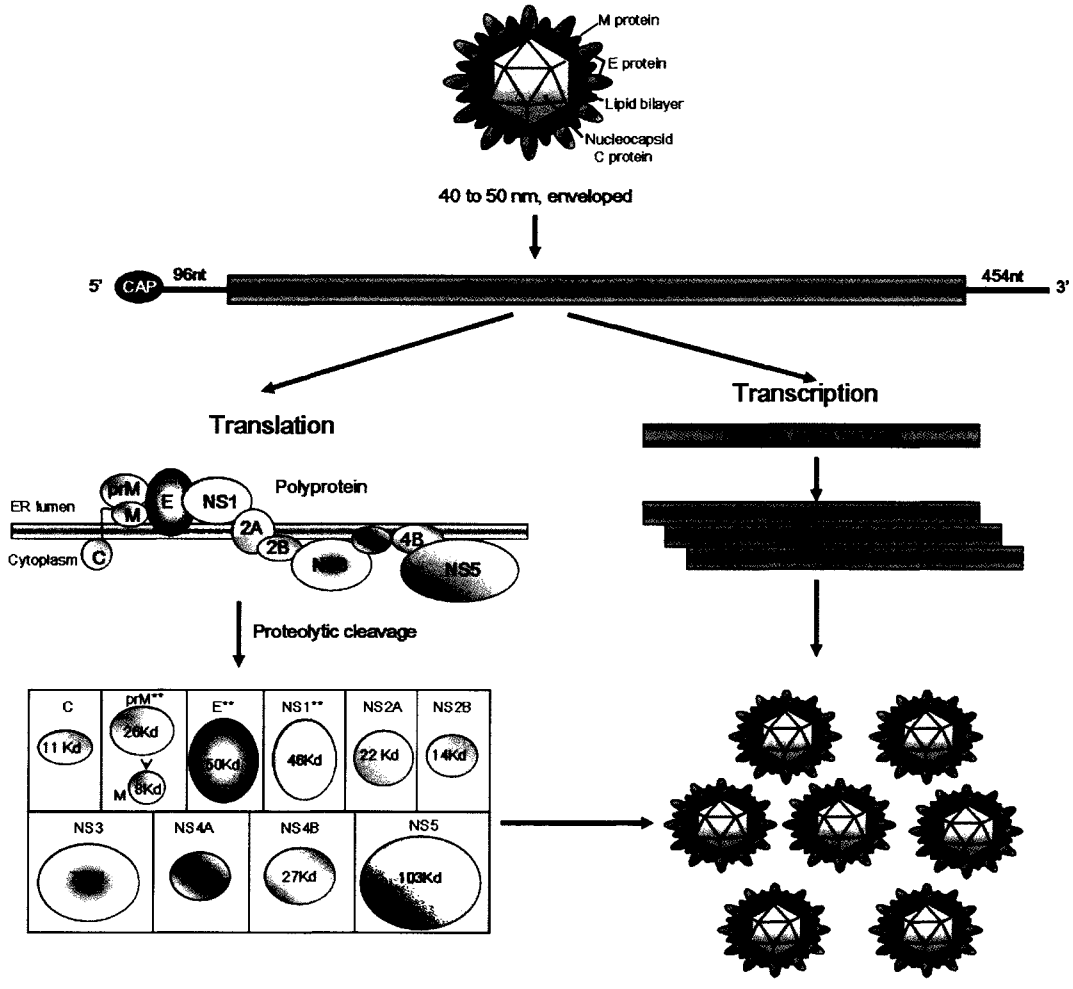
This re-infestation in turn led to renewed dengue and yellow fever outbreaks throughout central and South America (Gubler, 1989). No effective vaccines or therapies are currently available for dengue infections. Effective mosquito control is the only method to reduce disease incidence. History has demonstrated that mosquito control can be an effective disease reduction strategy. New mosquito reduction techniques, such as controlled indoor application methods, have suggested being effective and efficient at reducing mosquito populations and the spread of dengue virus (Garcia-Rejon *et al.*, 2008).

Dengue virus

Dengue viruses belong to the genus Flavivirus in the family Flaviviridae. The Flaviviridae family contains three genera, *Flavivirus*, *Pestivirus* and *Hepacivirus*. The genus Flavivirus includes important human pathogens such as dengue, yellow fever, West Nile and Japanese encephalitis viruses. Dengue, as with other arboviruses, must infect two phylogenetically distinct organisms -- vertebrates and arthropods -- to maintain its transmission cycle.

The dengue genome is approximately 11,000 bases in length and is a single strand positive sense RNA (Monath and Heinz, 1996, Burke and Monath, 2001). The RNA genome is associated with a capsid (C) protein in an icosahedral nucleocapsid. A lipid bilayer membrane containing two virus-encoded proteins, M and E, surrounds the nucleocapsid. The mature virion obtains a lipid envelope from a host cell membrane but

Figure 1.3. General diagram of the dengue virus replication cycle (Salazar Sanchez, 2005).



adds its own virus encoded membrane protein and envelope glycoprotein (Henschal and Putnak, 1990; Monath and Heinz, 1996).

The RNA genome contains a single open reading frame and is capped at the 5' end and lacks a 3' poly-A tail (Figure 1.3). The genome codes for three structural and seven nonstructural proteins. The seven non-structural proteins are coded at the 3' end of the genome and consist of NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5. The three structural proteins are coded at the 5' end of the genome and consist of capsid protein (C), the premembrane/membrane protein (prM/M) and the envelope protein (E).

A host signal peptidase is responsible for the cleavage between C-prM, prM-E, and E-NS1. The virus-encoded NS2B-NS3 serine protease is responsible for the cleavages between NS2A/NS2B, NS2B/NS3, NS3/NS4A, NS4A/NS4B, NS4B/NS5, and also an internal cleavage in C protein. The enzyme responsible for NS1-NS2A cleavage remains unknown (Lindenbauch and Rice, 2003). The entire genome is translated as a single polyprotein that is cleaved co-translationally and post-translationally by host and viral proteases into functional viral proteins. Formation of intracellular prM:E heterodimers occurs rapidly after translation and is important for the assembly and secretion of virus particles (Pryor *et al.*, 2004). Virus particles assemble by budding into the rough endoplasmic reticulum (Duebel *et al.*, 1981; McBride *et al.*, 2000, Parquet *et al.*, 2001). The assembled virus particles pass through the host ER-Golgi pathway where glycosylation and maturation take place, and the vesicles fuse to the plasma membrane for release of virions from the host cell.

Dengue virus, genetic diversity

As previously described, dengue virus exists as four antigenically distinct serotypes (DENV 1-4) that have all been associated with dengue fever and DHF/DSS (Monath and Heinz, 1996; Burke and Monath, 2001). Each of these serotypes contains considerable genetic variation, resulting in phylogenetically defined genotypes (Holmes and Burch, 2004). The Rico-Hesse (1990) study was one of the first to define intra-serotype variation in dengue viruses. A 240 nt fragment of the E/NS1 genes was used to measure genetic diversity and define different genotypes within serotypes as having no more than six percent sequence divergence. With advances in technology, additional genotype studies have been conducted, mostly relying on complete E gene sequences (~1485 bp) as the phylogenetic marker (Diaz *et al.*, 2006; Araujo *et al.*, 2008; Twiddy *et al.*, 2002; Holmes and Twiddy, 2003; Rico-Hesse *et al*, 1997, 1998, 2007).

Numerous genetic diversity studies have been completed for DENV-2 and DENV-3. DENV-2 is made up of at least six genotypes (Holmes and Twiddy, 2003). One genotype contains the sylvatic strains, two genotypes are found in Asian populations (Asian 1 and Asian 2), and one genotype, known as the "Cosmopolitan" genotype, is found in much of the tropical world (Twiddy *et al.*, 2002). The remaining two recognized genotypes in DENV-2 are the American/Asian genotype, linked to significant disease and epidemics, and the less pathogenic American genotype. DENV-3 is made up of at least three distinct genotypes (I through III). Genotype I consists of viruses from Indonesia that branched out to other regions in Southeast Asia, including East Timor, Malaysia, and the Philippines. Genotype II consists of viruses from Thailand that branched out to Bangladesh, Myanmar, Singapore, and Vietnam. Genotype III has a

much broader distribution that includes isolations from Asia, East Africa, Sri Lanka, and the Americas. It is unknown whether the genotype III lineage came from East Africa or Asia (Araujo *et al.*, 2008).

Genetic diversity data available for DENV-1 and DENV-4 are limited. DENV-1 contains three distinct clusters of sequences, which are regarded as genotypes even though they have not been formally defined. One cluster contains the sylvatic DENV-1 strains. DENV-4 contains three genetically diverse genotypes. One distinct cluster of sequences contains the sylvatic strains of DENV-4. Genotype I includes viruses from the Philippines, Thailand, and Sri Lanka. Genotype II consists of viruses from Indonesia, Tahiti, the Caribbean islands, Central America, and South America (Lanciotti *et al.*, 1997).

Earlier studies suggested that certain dengue viral genetic variants significantly contribute to severe pathogenesis (Rosen, 1977; Gubler *et al.*, 1978). The recent combinations of viral evolution and epidemiologic analyses have made it easier to determine whether certain genotypes within a serotype are associated with greater disease severity (Rico-Hesse, 1990 and 1997; Messer *et al.*, 2003). For example, DENV-2 and DENV-3 genotypes originating from Southeast Asia and the Indian subcontinent have caused more outbreaks with DHF/DSS than the American genotype (Watts et al., 1999).

This abundance of genetic variants likely occurs because dengue viruses have RNA genomes that are synthesized by a RNA-dependent RNA polymerase. This mode of replication lacks proofreading enzymes that DNA-based systems use to maintain high fidelity of genome replication (Drake, 1993). Dengue viruses also have a rapid rate of replication and immense population size (Drake and Holland, 1999). Dengue virus

evolution research has demonstrated a mean nucleotide substitution rate (for all genome sites) in the range of 4.55×10^{-4} (DENV-1) to 9.01×10^{-4} substitutions per site, per year (DENV-3) (Twiddy *et al.*, 2003). These rates are similar to those reported in other vector-borne RNA viruses, and are slightly lower than those in RNA viruses transmitted by other mechanisms. The lower evolution rates may be a result of the constraints of being required to replicate in a two-host (vertebrates and invertebrates) system (Jenkins *et al.*, 2002).

Viral recombination may also contribute to viral diversity. Research has demonstrated that dengue genotypes are not fixed entities and the opportunity for recombination events exist among them (Holmes *et al.*, 1999; Worobey *et a*l., 1999; Holmes and Twiddy, 2003). This recombination likely occurs through a copy-choice mechanism in which the polymerase switches between parental viral molecules during replication. It is unclear how widespread this mechanism is in nature. The necessary components are available since the *Ae. aegypti* vector is known to take multiple blood meals from different hosts, providing the potential of mixed infections (Lorono-Pino *et al.*, 1999). Persistent infections in mosquitoes have also been documented as a potential source for mixed infections (Burivong *et al.*, 2004). Data suggest though that recombination between genotypes are probably an infrequent event. If these mechanisms were to occur frequently in nature, one would expect to find virus sequences to be much more homogeneous and for recombinants to be difficult to identify (Holmes and Burch, 2000).

Error prone replication, recombination events, migration, and lack of gene flow are all contributing factors in viral genetic diversity and the potential for increase of

pathogenesis. The ease of human and mosquito travel over long distances in a relatively short period of time allows for the introduction of different DENV serotypes and genotypes into new regions. Virus trafficking can potentially contribute to ADE by introducing a different serotype within a DENV endemic population. Virus trafficking can also introduce different genotypes of a similar DENV serotype, allowing for a potential recombination event to form a more efficient and/or virulent DENV.

Aedes aegypti

The *Ae. aegypti* mosquito is one of the most important arbovirus vectors worldwide as the principal vector of yellow fever and dengue fever (Gubler *et al.*, 2004). In 1897, Ronald Ross demonstrated that malaria parasites could be transmitted from infected patients to mosquitoes (Ross, 1923). Major Walter Reed and the US Army Yellow Fever Commission demonstrated shortly thereafter that the yellow fever virus was transmitted by mosquitoes. Later, *Ae. aegypti* was identified by James Carroll and Carlos Finlay as the mosquito vector for yellow fever (de Kruif, 1926). Within this same timeframe, dengue virus was also determined to be transmitted by *Ae. aegypti* (Graham, 1903). No licensed vaccine is available for dengue and mosquito management is the only way to prevent transmission. Radical sanitation methods implemented by William Gorgas in 1901 effectively eliminated yellow fever in Cuba. Aggressive management of *Ae. aegypti* may also be necessary to control dengue virus transmission in the future.

Ae. aegypti includes two forms or subspecies that are distinguished by morphologic, ecologic and behavioral differences (Diallo *et al.*, 2005). *Ae. aegypti formosus* (Aaf) is the African sylvatic subspecies. It is believed that this subspecies is the ancestor to *Ae. aegypti aegypti* (Aaa). Aaf in West Africa preferentially oviposit in tree

holes, away from urban settings (Powell *et al.*, 1980; Tabachnick and Powell, 1979). Aaf is less susceptible to DENV and yellow fever infection than Aaa (Bosio *et al.*, 1998; Vazeille-Falcoz *et al.*, 1999). Aaa is a synanthropic, container-breeding mosquito with a global tropical and subtropical distribution (Gubler, 1979). Aaa is well adapted to cohabitation with humans in an urban environment and is the vector most associated with dengue transmission in epidemics. Increased interaction with humans has resulted in females feeding almost exclusively on human blood (Tabachnick, 1991; Scott *et al.*, 1993 and 2000a). This close association with feeding on human blood may benefit *Ae. aegypti aegypti* by providing a selective fitness advantage by providing increased energy reserves and fecundity (Harrington *et al.*, 2001a and 2001b; Costero *et al.*, 1998a, 1998b and 1999; Scott *et al.*, 1997). Evidence also suggests that this close proximity to humans facilitated global dispersion of *Ae. aegypti* from its ancestral roots of Central Africa (Christophers, 1960).

Substantial research has been conducted to further understand *Ae. aegypti* genomics. The availability of a draft sequence of the *Anopheles gambiae* genome has accelerated research to develop new mosquito and malaria control strategies (Holt *et al.*, 2002; Coluzzi *et al.*, 2002). In 2007, Nene *et al.* presented a draft sequence of the genome of *Ae. aegypti*. The genome is approximately 1376 million base pairs and is about five times the size of the *Anopheles gambiae* genome. It has been difficult to assemble an *Ae. aegypti* physical map from the draft sequence due to size, content of transposable elements (i.e. ~47% of the genome) and other "non-coding" sequences. Prior to the availability of a draft sequence, *Ae. aegypti* genomics relied heavily on comparisons of genome sequences from *Anopheles* and *Drosophila*. Even with these

assembly difficulties, current genomics studies have facilitated further understanding of dengue virus interaction with midgut receptors, differences in vector competence, and the mechanisms of immune-related genes and pathways (Waterhouse *et al.*, 2007; Mercado-Curiel *et al.*, 2006; Bosio *et al.*, 2000).

Population genetics of *Aedes aegypti*

The *Ae. aegypti* mosquito has a significant distribution worldwide between the 40°N and 40°S latitudes, is phenotypically polymorphic, varies in gene frequencies as detected by biochemical and molecular genetic markers, and exhibits variation in vector competence for arboviruses (Black *et al.*, 2002). As described in the previous sections *Ae. aegypti* includes two forms or subspecies that are distinguished by morphologic, ecologic, and behavioral differences (Diallo *et al.*, 2005). *Ae. aegypti formosus* is the African black sylvatic subspecies and primarily oviposits in tree holes. A light colored subspecies, *Ae. aegypti aegypti*, primarily oviposits in artificial containers (e.g. tires and discarded cups and containers) and can be found in tropical and subtropical regions throughout the world. Using allozymes as markers, population genetic studies identified eight distinct *Ae. aegypti* genetic groups worldwide (Tabachnick, 1991; Tabachnick and Powell, 1979). These groups are divided into two sylvan *formosus* subspecies in East and West Africa, one group in southeast United States, one group in southwest United States and Mexico, one in Central and South America, one in the Caribbean and another in Southeast Asia-Pacific. A recent study found that the West African, as well as East African, sylvan *formosus* subspecies continue to demonstrate reduced gene flow and retain genetic differentiation between sylvan and domestic *Ae. aegypti* populations (Paupy *et al.*, 2008).

Ae. aegypti genetic variation among these worldwide populations is relatively small, on the basis of sequence analyses, compared to that observed in other insect populations (Tabachnick and Powell, 1979; Mousson *et al.*, 2005). These data suggest that *Ae. aegypti* originated from Africa, and that the division into sylvatic and domestic subspecies, as well as its development as a vector for human disease transmission, evolved recently. The low level of genetic diversity within Aaa suggests the spread and establishment of Aaa throughout the world is a recent event and potentially was a result of human commerce (Black *et al.*, 2002).

Unlike earlier *Ae. aegypti* population studies that focused on worldwide analyses, most recent population genetic studies of *Ae. aegypti* focus on identifying localized or regional differentiation. The new focus allows researchers to better understand the spread of arthropod-borne pathogens and allows them to connect disease severity to vector genetics (Apostol *et al.*, 1996; Gorrochotegui-Escalante *et al.*, 2002; Urdaneta-Marquez *et al.*, 2008; Wong *et al.*, 2008; da Costa-Ribeiro *et al.*, 2007). Advances in molecular technology have allowed researchers to detect single nucleotide variations within individual mosquitoes and among mosquito collections.

Recent *Ae. aegypti* population studies in Mexico demonstrate that gene flow among regions varies a great deal (Gorrochotegui-Escalante *et al.*, 2000 and 2002). Availability of breeding sites and commerce, as well as barriers to natural migration through flight, all have the potential to affect gene flow. These studies demonstrated three major *Ae. aegypti* groups in Mexico corresponding to geographic regions in Northeastern Mexico, the Pacific coast, and the Yucatan peninsula. Differing rates of gene flow, genetic diversity, and genetic isolation existed among each of these groups.

Lozano-Fuentes *et al.* (2009) recently discovered an abrupt barrier to gene flow in *Ae. aegypti* associated with the Neovolcanic axis (NVA) along the Gulf Coast of Mexico. The mountains associated with the NVA were shown to limit distribution of *Ae. aegypti* north and south in a small geographic section between the mountains and the Atlantic Ocean. Lozano-Fuentes *et al.* noted that barriers to gene flow apparently affected the frequency of mosquitoes with a midgut infection barrier to DENV-2 on either side of the NVA. These unique findings demonstrate an example of how barriers to gene flow can also influence the frequencies of alleles and thereby contribute to differences in vector competence.

Besides understanding variations in vector competence, mosquito population genetic studies provide vital information that can be used when developing disease transmission studies or mosquito control programs that may require the introduction of genetically modified organisms. For example, if researchers understand the level of gene flow and the genetic diversity of a specific vector when developing disease transmission studies and mosquito control programs, energy can be appropriately focused in resource-poor environments while also using appropriate and effective pesticides (Garcia-Rejon *et al.*, 2008). Population genetic studies also provide much needed data as to the potential and real limitations that can be avoided and/or overcome with specific mosquito control programs.

Physiological genetics, quantitative genetics and environmental factors affecting vector competence (MIB, MEB) in *Ae. aegypti*

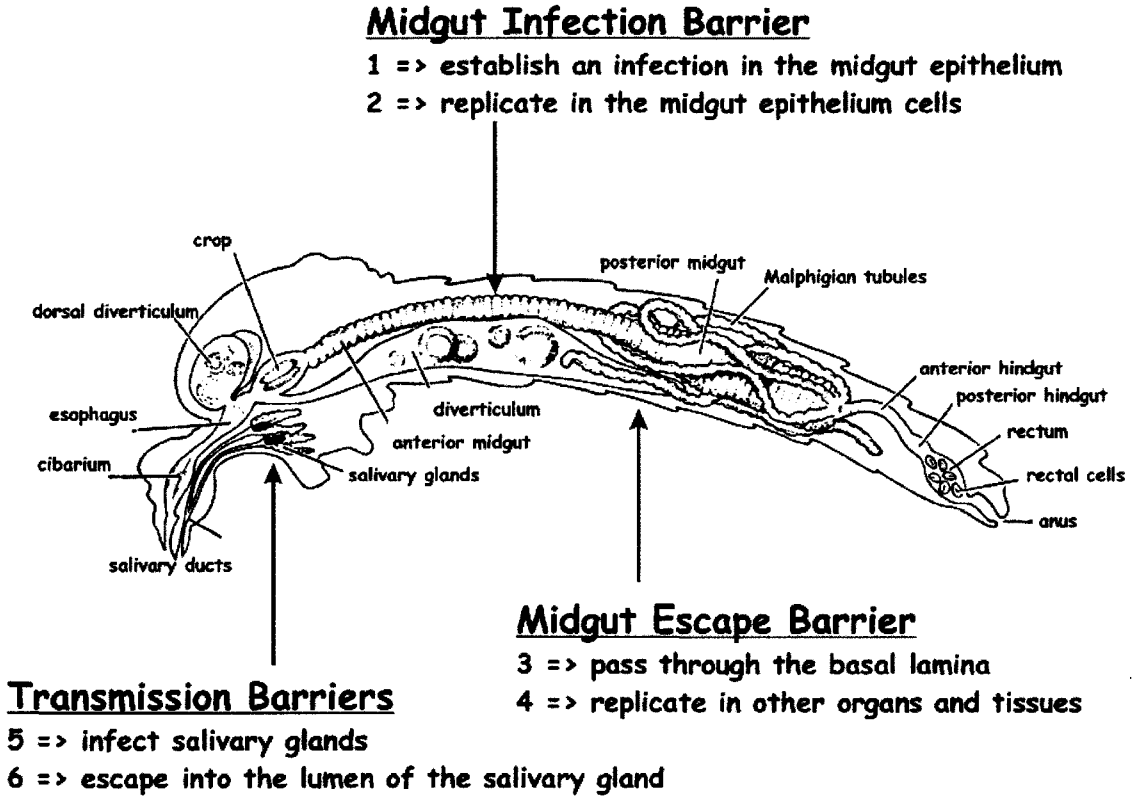
Vector competence is the intrinsic permissiveness of an arthropod vector to infection, replication, and transmission of a pathogen (Hardy, 1988; Woodring *et al.*, 1996). Studies have demonstrated that *Ae. aegypti* exhibits significant variation in vector

competence for flaviviruses (Bennett *et al.*, 2002; Diallo *et al.*, 2008; Knox *et al.*, 2003). For example, the *Ae. aegypti* subspecies *formosus* from sub-Saharan Africa has demonstrated a very low vector competence for dengue virus due to a midgut infection barrier. *Ae. aegypti aegypti* collected from specific sites in the Yucatan peninsula of Mexico (e.g. Chetumal, Cancun), however, are highly susceptible to infection and dissemination of dengue virus.

Once the mosquito host ingests the arbovirus during a blood meal, the virus must overcome several potential barriers if it is to be transmitted to a subsequent vertebrate host. First, the virus must be able to successfully establish an infection in the mosquito midgut by overcoming the midgut infection barrier (MIB) (Figure 1.4). A number of flavivirus vector competence studies in *Ae. aegypti* have identified that MIB is a major determinant of flavivirus transmission (Schoepp *et al.*, 1990; Gubler *et al*, 1979; Tabachnick *et al.*, 1985; Black *et al.*, 2002). Research has not yet identified the receptors in the midgut epithelia that is responsible for dengue virus attachment and entry. Presence of a 67 kDa protein (R67/R64), shown to be a potential receptor, has been found to be significantly associated with vector competence (Mercado-Curiel *et al.*, 2008). Limited information is available concerning the identity of this possible receptor. Mapping studies have identified quantitative trait loci (QTL) for MIB on each of the three *Ae. aegypti* chromosomes (Bosio *et al.*, 2000; Gomez-Machorro *et al.*, 2004).

Following replication in the midgut epithelium, the virus must overcome the midgut escape barrier (MEB) and replicate in other tissues, such as the hemocoele and salivary glands (Figure 1.4). Bennett *et al.* (2005) identified three distinct QTLs on each of the three chromosomes in *Ae. aegypti* specific to MEB for DENV-2. Virus

Figure 1.4. Diagram demonstrating six potential barriers to transmission that an arbovirus faces in being transmitted by a mosquito (Black *et al.*, 2002).



disseminates through the mosquito body via hemolymph. The virus must then ultimately be able to infect and replicate in the salivary gland tissue and be shed into the lumen and saliva to be transmitted to a vertebrate host. Salivary gland infection or escape barriers have also been identified and can prevent infection (Whitfield *et al.*, 1971 and 1973; Takahashi and Suzuki, 1979; Sriurairatna and Bhamarapravati, 1977).

A number of quantitative genetic studies have been conducted in *Ae. aegypti* to further understand susceptibility and resistance factors (Gubler *et al*, 1979; Miller and Mitchell, 1991; Wallis *et al.*, 1983, 1984, 1985). Quantitative genetics estimates the genetic and environmental effects associated with variation as quantitative rather than discrete traits (Falconer and MacKay, 1996). Examples of quantitative traits include size, weight, fecundity, and vector competence. The assumption that vector competence is a quantitative trait suggests that it is under the control of more than one gene and subject to environmental effects. Bosio *et al.* (1998) conducted one of the more detailed studies in *Ae. aegypti* and the results suggested that sets of genes control MIB and MEB.

The primary goal of quantitative genetics is to estimate the extent to which variation in a phenotypic trait (V_P) is conditioned by the environment (environmental variance, V_E) and genetics (V_G). V_G can be further divided into additive (V_A), dominance (V_D), and epistatic (V_I) variance. Specific genotypes may perform differently in different environments and this must be taken into consideration (V_{GxE}). Statistical formulas have been developed to analyze these factors in the context of a specific system (e.g. *Ae. aegypti*) (Becker 1992, Falconer and Mackay, 1996; Mather and Jinks, 1982). The general formula for genetic variance is:

$$V_G = V_A + V_D + V_I$$

The formula for phenotypic variance is:

$$V_P = V_A + V_D + V_I + V_E + V_{GxE}$$

Narrow sense heritability is also an important measure that is generated when analyzing a quantitative trait such as vector competence (Severson *et al.*, 1994, 1995; Bosio *et al.*, 1998). Narrow sense heritability, also termed additive genetic heritability, is the proportion of variation in a trait attributed to additive genetic effects. It is calculated as:

$$h^2 = V_A/V_P$$

This variable is important because it predicts the response to selection on an identified trait (Falconer and Mackay, 1996). Dominance heritability can also be calculated, but it is not a good predictor of response to selection. This variable takes into account one or more alleles when determining the expression of a trait and can mask the expression of other alleles.

Genes that control the ability of DENV-2 to propagate in *Ae. aegypti* tissues act additively and account for 39-41% of the total phenotypic variation (Bosio *et al.*, 1998). This finding also demonstrates that vector competence can be substantially influenced by outside environmental factors, or environmental variation (V_E). For example, MIBs can be overwhelmed by high virus titers (Woodring *et al.*, 1996). Laboratory experiments have shown that when *Ae. aegypti* mosquitoes ingest a blood meal containing a very low DENV-2 titer, dissemination rates tend to also be lower (Bernhardt, 2009 unpublished). Low doses of virus may not be able to overcome the midgut infection threshold or may increase the time it takes for a productive infection to be established and to disseminate.

Viremia titers naturally vary among infected humans, which consequently could influence virus dose and overcoming the MIB.

Environmental factors can also significantly impact extrinsic incubation period (EIP) in the mosquito, which ultimately can alter vector competence. The EIP is the time interval between the ingestion of an infectious blood meal by a mosquito and the oral transmission of virus to a vertebrate host. Variations in environmental temperature have been shown to affect EIP. For example, variations in temperature have been shown to alter virus dissemination rates of West Nile virus in *Culex pipiens* (Richards *et. al*, 2007; Dohm *et. al*., 2002). Richards *et al*. (2007) fed *Culex pipiens quinquefasciatus* an infectious blood meal containing West Nile virus and held the mosquitoes for 13 days at 25° C, 28° C, or 30° C. Each of these three groups exhibited significantly different rates of virus dissemination to head tissues (33%, 22% and 81%). Another example, with *Ae. aegypti* and dengue, demonstrated that the EIP was decreased from 12 days at 30° C to 7 days at 32-35° C (Watts *et al*., 1987). Numerous other studies have demonstrated that the EIP is generally inversely proportional to the temperature the vector is exposed to following ingestion of an infectious blood meal (Bates and Roca-Garcia, 1946; Chamberlain and Sudia, 1955; Kramer, 1983). These EIP variations with respect to environmental temperature can vary among vectors and viruses (Hardy, 1988). Incubation below or above an ideal EIP temperature can also either hinder or significantly limit virus titer within a mosquito host (Hurlbut, 1973; Kramer *et al*., 1983; Hardy *et al*., 1988).

Larval nutrition has also been shown to impact vector competence. For example, Alto *et al*. (2008a, 2008b) varied *Ae. aegypti* and *Ae. albopictus* larval diet to produce

adults of differing sizes. Small females from nutritionally deprived larvae were significantly more likely to become infected and to disseminate virus to the head than larger individuals. Similar findings have been demonstrated in *Ae. triseriatus* with LaCrosse virus and in *Culex tritaeniorhynchus* with Japanese encephalitis virus and West Nile virus (Grimstad and Haramis, 1984; Grimstad and Walker, 1991; Takahashi, 1976; Baqar *et al.*, 1988). The hypothesis is that smaller adults consume larger blood meals, with respect to their size, resulting in infection of proportionally more cells. Understanding how larval nutrition affects virus infection and replication is complicated by contradictory findings. For example, researchers have demonstrated that body size among *Ae. aegypti* does not significantly affect DENV-2 infection (Schneider *et al.*, 2007; Sumanochitrapon *et al.*, 1998). Additional environmental factors can affect vector competence and arbovirus transmission. As indicated in the formula for vectorial capacity, aspects such as daily survivorship, mosquito density and number of times a vector feeds on a host in one day can all impact arbovirus transmission.

Physical mapping, linkage mapping and QTL mapping

Physical mapping determines the actual physical location of a gene on a chromosome. Physical maps are important in studies of genome organization and evolution, for localization and isolation of genes of interest, and when assembling sequence data obtained in systematic genome-sequencing (Black and Severson, 2005). Physical maps are conceptually more useful than genetic linkage maps in determining the actual size of chromosomes and the nucleotide distances between markers.

Physical maps have been generated for a number of organisms to clarify gene position and sequence order. Work on *Anopheles gambiae* is a good example of physical

mapping in vectors. The genome of *Anopheles gambiae* was sequenced by a whole-genome shotgun approach (Holt *et al.*, 2002). Physical mapping of the genome was conducted using in situ hybridization of cloned sequences to the polytene chromosome. This technique included cloned RAPD fragments, microsatellites, cDNAs, cosmids, and BAC clones (della Torre *et a*l., 1996; Zheng *et al.*, 1996). This provided some of the first physical maps related to genetic maps of *Anopheles gambiae* (Dimopoulos *et al*, 1996). Since the introduction of these preliminary physical maps, additional techniques have been used to improve resolution of gene order (Sharakhova *et al.*, 2007).

Physical mapping in culicine mosquitoes has been very difficult to accomplish due to the poor resolution of their polytene chromosomes. In 2007, Nene *et al.* presented a draft nucleotide sequence of the genome of *Ae. aegypti*. Due to the immense genome size, and the significant proportion of transposable elements in the genome (~47%), it has been difficult to assemble an *Ae. aegypti* physical map from the draft sequence. A preliminary physical map has been generated for *Ae. aegypti* using *fluorescent in situ hybridization* (*FISH*) combined with digital-imaging microscopy to position random cosmid clones on metaphase chromosomes (Brown and Knudson, 1997). This technique provided preliminary insight to the physical map. Physical mapping has been limited due to complexities of assembling the *Ae. aegypti* sequence scaffold. For this reason, researchers have continued to rely heavily on linkage mapping in *Ae. aegypti*. The ability to generate probes based on single nucleotide polymorphisms (SNPs), cDNA markers, restriction fragment length polymorphism and random amplified polymorphic DNA markers have increased the resolution of linkage map.

Linkage mapping determines the frequency with which markers on a chromosome segregated by recombination during meiosis, providing a locus for each marker (Black and Severson, 2005). A linkage map is often developed within a species based on determining how a specific phenotype (e.g. vector competence) is under genetic control and how alleles are inherited and contribute to a specific quantitative trait. Fortunately, advances in molecular technology and bioinformatics have greatly reduced the intense labor and computation requirements that are necessary when developing these linkage maps.

Linkage maps are usually generated using an F_1 intercross design (a recombinant inbred line) or recurrent backcrosses. An F_1 intercross is constructed by having F_1 siblings interbreed from parents of known alleles. Whereas, a backcross to a recurrent parent is crossing of a F_1 offspring with one of its parents, in order to achieve offspring with a genetic identity which is closer to that of the parent. Recombination frequencies between alleles at a pair of loci within a single round of meiosis are determined and mapped using centimorgan (cM) units. Hundreds of markers can be mapped simultaneously in a single cross. Parents within the F_1 intercross are chosen based on the phenotype trait of interest (e.g. insecticide resistance). Each parent is usually homozygous at a specific locus and represents each of the extreme phenotypes. This allows tracing the origin of a given allele back to the parent and identifying points of recombination.

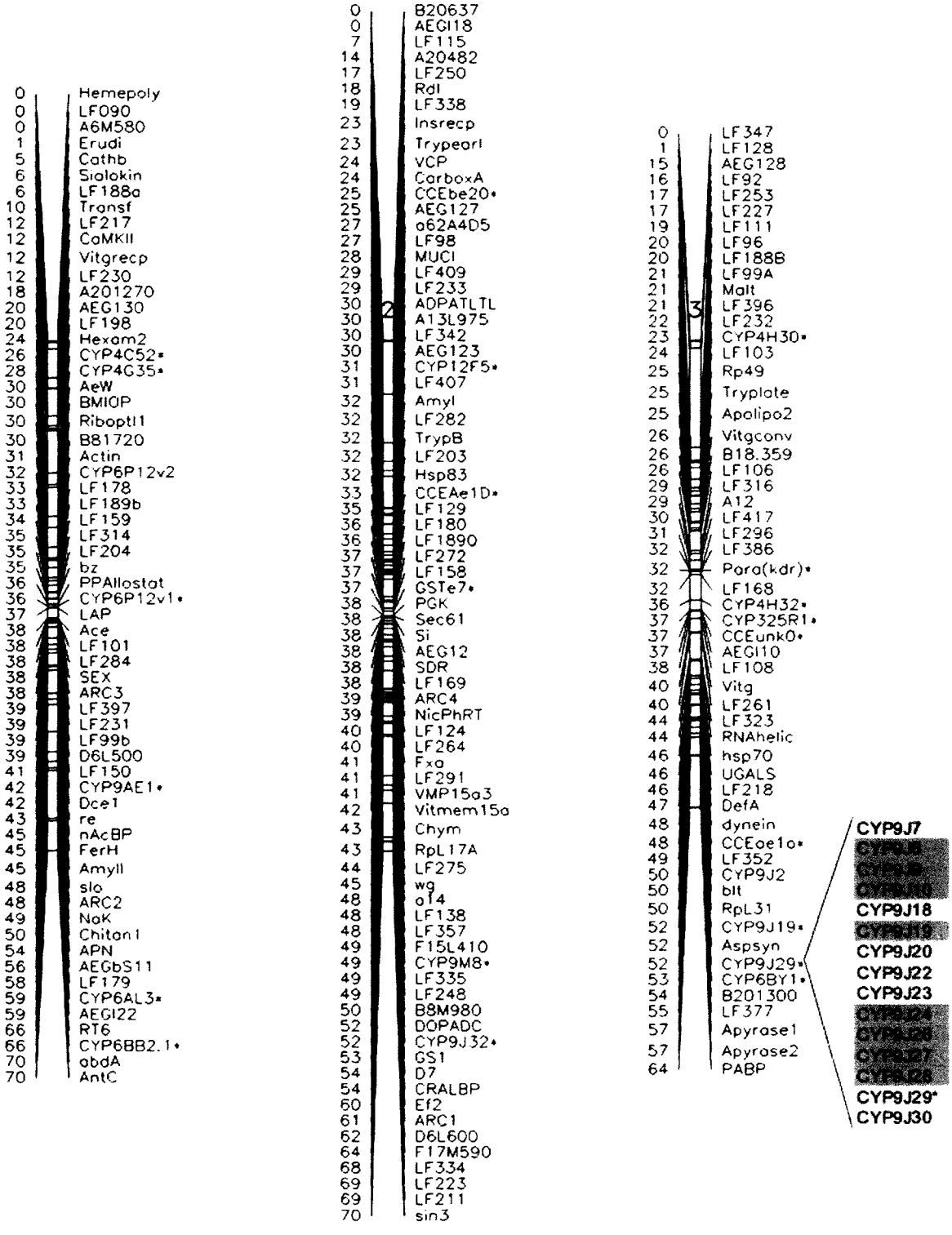
Linkage mapping is often performed using the F_2 generation. QTL mapping is often performed in an advanced intercross line to allow for additional recombination. SNPs are highly abundant and have been used as effective genetic markers (Landegren *et*

al., 1998; Brookes 1999). Techniques such as single strand conformation polymorphism analysis and SNP melting curve analysis take advantage of detecting these SNPs to generate recombination data (Black and DeTeau, 1997; Urdaneta-Marquez *et al.*, 2008). Statistical computer programs have been developed to analyze the recombination frequencies among all the loci within the mapping family. JoinMap statistical software has been developed to determine, for each pair of loci, which genotypes are informative, to estimate the pairwise distances from informative genotypes, and to estimate the logs odds ratio scores for each of the estimates (Stam and van Ooijen, 1995). Linkage maps developed in specific studies are often less dense, use fewer markers, compared to the example of the current linkage map of *Ae. aegypti* that is provided in Figure 1.5.

Quantitative trait loci (QTL) mapping can be performed once a linkage map is generated and a specific phenotypic characteristic is determined to be a quantitative trait or found to be under control of more than one gene or set of genes. The phenotypic trait is determined for each of the offspring in the advanced intercross line and "attached" to the genotype data generated to construct the linkage map. Contingency χ^2 analyses are often first performed to determine whether a phenotype (e.g. MIB, MEB) is equal within a specific genotype class. This test allows one to determine whether a specific phenotype is linked to a certain locus. A problem with this analysis is that loci do not segregate independently and this analysis treats each locus as an independent factor. This analysis provides an idea as to which loci contribute to a specific phenotype, but one must use a more precise method to identify the presence of a QTL.

QTL Cartographer (http://statgen.ncsu.edu/qtlcart/WQTLCart.htm) software has been developed to conduct interval and composite interval mapping (Lander and

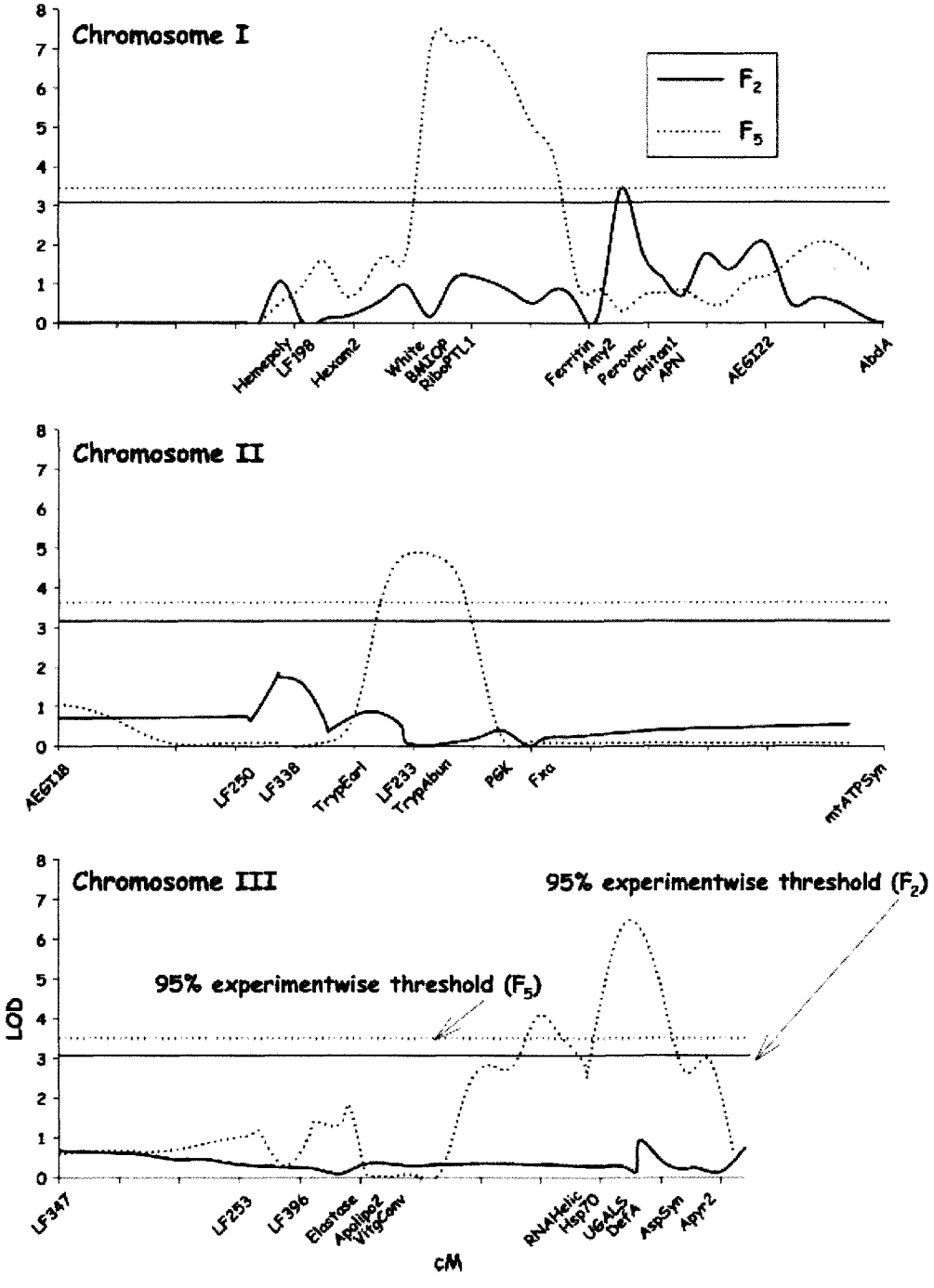
Figure 1.5. Current linkage map of *Ae. aegypti*. 62, 75 and 71 markers exist, respectively, on chromosomes 1, 2, and 3. Markers were mapped using SNP variation, STAR loci and RFLP variation. Data used to construct this map came from numerous independent studies (Saavedra-Rodriquez et al., 2008; Black and Severson, 2005).



Botstein, 1989). Interval mapping is a method that assesses pairs of adjacent markers to test for the statistical association of the interval with the phenotype. Composite interval mapping takes into account certain biases within interval mapping by adjusting for the effects of intervals in proximity to the interval under analysis. Comparisonwise 95% thresholds are often applied to interval mapping and confidence interval mapping to identify LOD (logs odds distance) scores that are significant with respect to a QTL. An example of a QTL map for DENV-2 MEB in *Ae. aegypti* from Bennett *et al.* (2005) is shown in Figure 1.6.

Numerous linkage mapping and QTL mapping studies have been performed in various vectors (Zheng *et al.*, 1997; Mori *et al.*, 2008; Bosio *et al.,* 2000; Bennett *et al.*, 2005; Beernsten *et al.*, 1995; Severson *et al.*, 1995; Saavedra-Rodriquez *et al.*, 2008; Wondji *et al.*, 2007; Gomez-Machorro *et al.*, 2004). Most studies in mosquitoes focus on aspects specific to infectious disease transmission or vector control methods (e.g. dengue vector competence). A number of QTL mapping studies have been conducted for *Ae. aegypti* to understand midgut infection and dissemination barriers to DENV-2. Susceptibility to dengue viruses has been shown to be under control of more than one genetic locus, as well as various environmental factors. Gomez-Machorro *et al.* (2004) identified two distinct QTLs in *Ae. aegypti* for midgut susceptibility (MIB) to DENV-2. They found that the alleles at these QTLs contributed additively in determining susceptibility and accounted for 24% of the phenotypic variance. Bennett *et al.* (2005) identified three distinct QTLs in *Ae. aegypti* for MEB for DENV-2. They also found that alleles at these QTLs were additive or dominant in determining rates of DENV-2 dissemination and accounted for 45% of the phenotypic variance.

Figured 1.6. Plot of logs odds distance (LOD) values associated with MEB for DENV-2 along all three chromosomes in *Ae. aegypti* (Bennett *et al.*, 2005). Names of markers are listed to orient QTL positions relative to the linkage map. LOD estimated in the F_2 by composite-interval mapping appears as a solid line and the 95% experimentwise threshold is represented by a straight solid line. LOD and the 95% experimentwise thresholds estimated in the F_5 appear as dotted lines.



Mosquito fitness and arbovirus impact

The impact of arbovirus infections on mosquito vector fitness is an ongoing discussion. Few cytopathic effects in the vector upon arbovirus infection and transmission were observed in early studies (Lam and Marshall, 1968). Researchers believed that a co-evolutionary state existed where the host (vector) would become increasingly resistant to parasite (arbovirus) infections. Through resistance and attenuation, a benign, or commensal relationship would result (Burnet and White, 1972). More recently, however, researchers have demonstrated that arboviruses can cause cytopathic effects in cultured mosquito cells (Wang *et al.*, 2008; Burivong *et al.*, 2004). A number of studies in vectors (e.g. light microscopy and transmission electron microscopy, feeding behavior, survivorship and reproduction studies of arbovirus-infected vectors) in the last 30 years have also challenged the notion that there are no vector fitness effects of arbovirus infections (Weaver *et al.*, 1988, 1992; Bowers *et al.*, 2003; Platt *et al.*, 1997; Grimstad *et al.* 1980; Faran *et al.*, 1987; Mahmood *et al.*, 2004). Researchers recently described virus-induced cell death in the salivary glands and midgut epithelia of *Culex pipiens* associated with West Nile virus infections (Vaidyanathan and Scott, 2006; Girard *et al.*, 2005 and 2007). Also survivorship and fecundity studies have shown reductions in vector fitness when certain mosquito species (i.e. *Culex pipiens*, *Culex tarsalis*, *Culiseta melanura*, *Ae. triseriatus*) were infected with arboviruses such as LaCrosse, West Nile, western equine encephalomyelitis, eastern equine encephalomyelitis, and Rift Valley fever virus (McGaw *et al*, 1998; Vaidyanathan *et al.*, 2008; Lee *et al.*, 2000; Scott and Orenz, 1998; Faran *et al.*, 1987).

Apoptosis in mosquito vectors

Apoptosis is a genetically regulated mechanism of programmed cell death that is critical during normal development, tissue homeostasis, and disease processes in metazoans (Evan and Littlewood, 1997; Saikumar *et al.*, 1999). Several stereotypic morphological markers characterize it. In the nucleus, chromatin condensation occurs and the DNA undergoes breakdown into nucleosomal fragments. Cell shrinkage and membrane blebbing are observed and apoptotic bodies are quickly engulfed by phagocytes. The caspases, a family of cysteine proteases, have been shown to cause the majority of biochemical and morphological changes associated with apoptosis (Thornberry and Lazebnik, 1998). Some parasites, viruses, and autoimmune disorders induce or inhibit host cell death via this genetically controlled process (Thompson, 1995; Peter *et al.*, 1997).

Morphological features indicative of apoptosis in mosquito midgut lumens have been described for invading unicellular eukaryotes such as *Leishmania* spp., *Trypanosoma* spp. and *Plasmodium* spp. (Arnoult *et al.*, 2002; Hurd and Carter, 2004; Barillas-Mury and Kumar, 2005). Zieler *et al.* (1998 and 1999) developed an *in vitro* invasion assay with which to study the interaction between *Plasmodium gallinaceum* and *Ae. aegypti* midgut epithelium and reported an intracellular route of invasion. Upon invasion by the *Plasmodium* species, the *Ae. aegypti* midgut epithelium underwent cell death. Loss of plasma membrane integrity was demonstrated as well as nuclear swelling, movement to an apical position, and cell surface blebbing.

Limited information is available on apoptotic regulation in mosquitoes, despite the role that these vectors play in disease transmission. Preliminary work by Blitvich *et.*

al. (2002) identified and described an inhibitor of apoptosis protein 1 homologue from *Ochlerotatus triseriatus* mosquitoes. They detected this protein in various mosquito life stages and tissues, such as the ovaries and salivary glands. Additional data have demonstrated that the apoptosis pathway and proteins exist in additional mosquito species and at different life stages (Bryant *et al.*, 2008; Pridgeon *et al.*, 2008; Beck *et al.*, 2007).

Questions exist as to whether arbovirus infections in vectors trigger the apoptosis pathway. Arbovirus infections were generally thought to cause persistent infections, with only moderate cytopathic, apoptotic effects (Karpf and Brown, 1998). Recently, researchers have described virus-induced cell death in the salivary glands and midgut epithelia of *Culex pipiens* infected with West Nile virus (Vaidyanathan and Scott, 2006; Girard *et al.*, 2005 and 2007). Wang *et al.* (2008) developed a recombinant Sindbis virus that expressed either pro-apoptotic or anti-apoptotic genes that they used to infect cultured mosquito cells. The goal was to test the effects of inducing or inhibiting apoptosis on Sindbis virus replication in mosquito cells. Even though the recombinant Sindbis virus caused extensive apoptosis in the cell monolayer, high levels of virus replication were observed initially. Virus replication precipitously dropped after the initial phase, suggesting that apoptosis had some effects on viral replication.

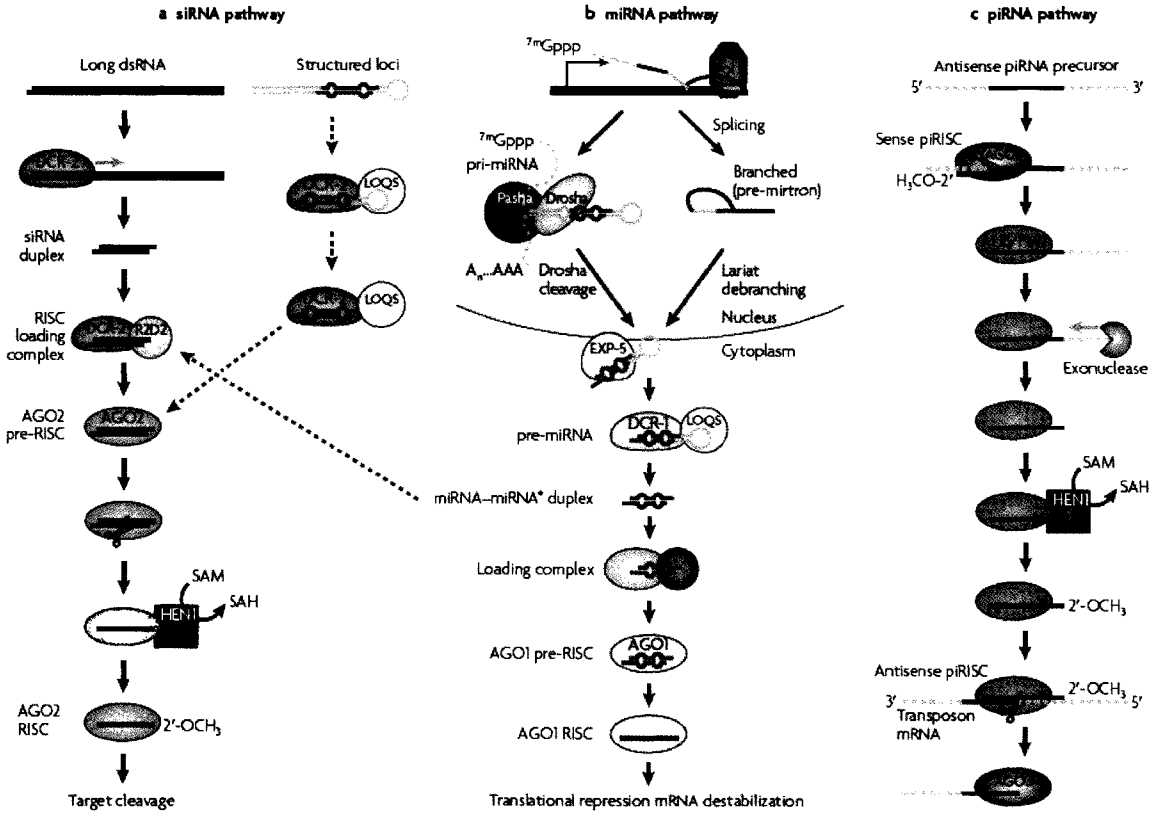
RNA interference in mosquito vectors

RNA interference (RNAi), an RNA silencing mechanism, is an important antiviral innate immune response in invertebrates. *Drosophila melanogaster* has consistently been used as a RNAi model (Zambon *et al.*, 2006; Wang *et al.*, 2006; Obbard *et al.*, 2009). The post-transcriptional gene silencing (PTGS)/RNAi system was initially described in plants by van der Krol *et al.* (1990a and b). Shortly thereafter researchers discovered that

PTGS was able to induce an immune response against plant viruses (Lindbo *et al.*, 1993). The initiating factor in this silencing mechanism was determined, using *Caenorhabditis elegans*, to be double-stranded RNA (dsRNA) and the process was renamed RNAi (Fire *et al.*, 1999). Three RNA silencing pathways in *Drosophila* (Figure 1.7) have been described in the literature and depend on sequence specific binding between small (~20-30 nucleotides) RNAs and their target sequence (Baulcombe, 2004; Ding and Voinnet, 2007; Waterhouse *et al.*, 2001a and 2001b). The antiviral small interfering RNA (siRNA) pathway (Figure 1.7a) and the microRNA pathway (Figure 1.7b) will be described in this review.

The RNAi pathway is initiated through recognition of exogenous and/or endogenous dsRNA structures formed by some RNA viruses in host cells (Zambon *et al.*, 2006; Wang *et al.*, 2006; Hannon, 2002; Saleh *et al.*, 2009; Deddouche *et al.*, 2008). Exogenous dsRNA are recognized and processed by Dicer-2 into siRNAs of about 21-23 bp (Liu *et al.*, 2006; Liu *et al.*, 2003; Galiana-Arnoux *et al.*, 2006). The Dicer proteins are members of the RNAse III gene family. The dimerized RNAse III protein domains ultimately determine the cleaving lengths (21-23 bps) of the dsRNA (Blaszczyk *et al.*, 2001). The dsRNA, with the aid of Dicer-2 and R2D2, is loaded into the Argonaute-2 (Ago2)-containing effector complex, RNA-induced silencing complex (RISC) (Schwarz *et al.*, 2002; Liu *et al.*, 2003). The RISC loaded siRNA strand acts as the targeting mechanism to guide the RISC to recognize complementary mRNAs. Ago2, within the RISC, is involved with RNA binding and cleaving the recognized complimentary mRNA strand (Matranga *et al.*, 2005; Hammond *et al.*, 2001; Miyoshi *et al.*, 2005). The *Ae.*

Figure 1.7. RNA silencing pathways in *Drosophila* (Ghildiyal and Zamore, 2009). (a) The antiviral siRNA pathway. (b) The miRNA pathway. (c) The piwi interacting or piRNA pathway for transposable element silencing in the *Drosophila* germ line.



aegypti genome has orthologs to all of these components (Franz *et al.*, 2006; Sanchez-Vargas *et al.*, 2004, 2009; Nene *et al.*, 2007).

The endogenous microRNA pathway is a fundamental component of post-transcriptional gene regulation in plants and animals (Carthew 2006; Bushati and Cohen, 2007; Chapman and Carrington, 2007). microRNAs are encoded by the genome and are transcribed by RNA polymerase II. The primary microRNAs are processed in the nucleus by Drosha family members to form short stem loops (Du and Zamore, 2005; Chapman and Carrington; 2007). The stem-loop microRNAs are transported to the cytoplasm and undergo further processing as described for siRNA. Endogenous dsRNA stem-loop microRNAs are recognized and processed by Dicer-1 into microRNAs of about 21-23 bp. The miRNA, with the aid of Dicer-1 and R3D1, and loaded into the Argonaute 1 (Ago1)-containing effector complex (RISC). The microRNA RISC interacts with target mRNAs. This interaction can result in cleavage and degradation of the target transcript or translation inhibition (Pillai *et al.*, 2007). The microRNA pathway was once thought only to regulate host gene expression. Recent findings, however, suggest that this pathway potentially can play a role in defense against viruses by expressing microRNAs that target viral RNA (Lecellier *et al.*, 2005; Pedersen *et al.*, 2007).

RNA silencing is an antiviral defense mechanism and has been shown to be an innate immune pathway to protect adult flies from infection and unregulated viral replication. For example, expression of long DENV dsRNA genome-derived in mosquito cells or adult females resulted in inhibition of later dengue virus infection and replication (Olson *et al.*, 1996; Adelman *et al.*, 2001, 2002). Keene *et al.* (2004) demonstrated that RNAi acts as a natural antiviral defense to alphavirus infection of *Anopheles gambiae*.

They silenced the Ago2 gene in *Anopheles* and demonstrated increased viral replication. Another recent example showed that the RNAi mechanism is active during alphavirus infection and that a transient increase in alphavirus replication occurred by silencing RNAi pathway components in *Ae. aegypti* (Campbell *et al.*, 2008).

Additional findings also demonstrate that an exogenous siRNA pathway is a vital innate immune response necessary for mosquito survival. Cirimotich *et al.*, (2009) and Myles *et al.* (2008) both demonstrated that the RNAi pathway in *Ae. aegypti* is active in suppressing unrestricted alphavirus replication. They demonstrated that if virus-derived siRNA production was suppressed by expression of the FHV B2 protein, the result was an increase of viral replication and mosquito mortality. Sanchez-Vargas *et al.* (2009) showed that persistent DENV-2 infections of *Ae. aegypti* cells and mosquitoes generate long dsRNA as well as small DENV-2 specific RNA consistent in size and sequence with siRNAs. When they impaired the RNAi pathway in mosquitoes infected with DENV-2, titers of infectious virus also increased while the extrinsic incubation period decreased.

Recent research also has demonstrated that the RNAi pathway in insects and plants may impose selective pressures on the virus to maintain efficient replication. Suppressors of RNA silencing have been identified in many insect and plant viruses. For example, Flock House virus and Nodamura virus were found to encode for an RNAi suppressor proteins (B2 for FHV) (Johnson *et al.*, 2004; Li *et al.*, 2002). The B2 protein, one of the best characterized suppressors of RNAi, protects the virus from host RNAi responses and promotes replication as well as pathogenesis. B2 acts by preventing RNAi pathway components access to dsRNA associated with Flock House virus replication. No examples have been found to date of RNAi suppressors associated with or encoded by

arboviruses. Questions still remain with respect to the role of the RNAi pathway in arbovirus infections of mosquitoes and viral mechanism of RNAi evasions.

Pathogenic mosquito viruses

Most viral pathogens that cause disease in mosquitoes are divided into four major groups. In two groups the virion is embedded in a protein matrix, or occluded viruses (Becnel and White, 2007). The most common of the occluded viruses are the baculoviruses (Baculoviridae: Deltabaculovirus), cytoplasmic polyhedrosis viruses (CPVs) (Reoviridae: Cypovirus) and entomopoxviruses (Jehle *et al.*, 2006; Federici, 1977, 1980, 1995; Federici and Lowe, 1972). The main two non-occluded virus groups are densoviruses (Parvoviridae: Brevidensovirus) and the iridoviruses (Iridoviridae: Chloriridovirus). Baculoviruses, densoviruses and iridoviruses are DNA viruses. Cypoviruses are double-stranded RNA viruses.

The *Nucleopolyhedrovirus* group (NPV) are the only baculoviruses that are known to infect mosquitoes (Federici, 1995; Becnel *et al.*, 2001; Andreadis *et al.*, 2003, Shapiro *et al.*, 2004). More than 20 mosquitoes, including *Aedes*, *Anopheles*, *Culex*, *Ochlerotatus*, *Uranotaenia* and *Wyeomiyia* have been shown to be susceptible to NPV infections (Murphy *et al.*, 1995). NPV infections mostly occur in larval midgut tissues, but have been also found in adult midgut epithelium. Larvae are typically infected by ingestion of viral occlusion bodies. The occlusion bodies attach to the microvilli membrane and enter the cytoplasm of midgut epithelial cells. The nucleocapsid moves to the nuclear envelope where the double-stranded circular viral DNA is released into the nucleoplasm through the nuclear pores. Viral DNA replication occurs in the nucleus. Infections are characterized by hypertrophic nuclei of the midgut cells. Death of the

larvae usually occurs within 72 - 96 hours post infection (Becnel *et al.*, 2001; Moser *et al.*, 2001; Alfonso *et al.*, 2001).

Ae. aegypti densovirus (AaeDNV) and *Ae. albopictus* densovirus (AalDNV) are the most common mosquito densoviruses (MDVs) of the genus *Brevidensovirus*, family Parvoviridae (Buchen-Osmond, 2003). *Culex pipiens* densovirus (CpDNV) is categorized in the genus *Densovirus* (Jousset *et al.*, 2000). MDVs have been shown to infect a number of species within *Aedes*, *Anopheles*, *Toxorhynchites*, *Haemagogus*, *Ochlerotatus*, *Psorophora*, *Culex*, and *Culiseta* genera (Rwegoshora *et al.*, 2000; Lebedinets and Zelenko, 1975; Buchatsky, 1989; Barreau *et al.*, 1996). Horizontal transmission occurs during the larval stage and virus enters via the anal papillae and bristle cells (Ward *et al.*, 2001). Horizontal transmission is dose and age dependent. Higher doses result in higher larval mortality (up to 90%) (Barreau *et al.*, 1996). Large numbers of larvae survive at lower doses and allow adults to transmit the virus vertically (Ledermann *et al.*, 2004; Rwegoshora and Kittayapong, 2004). The genome is composed of linear, single-stranded DNA and viral replication occurs in the nuclei. Depending on the virus and host, virions are assembled either in the nuclei (AalDNV) or within inclusion bodies that are formed in the cytoplasm (AaeDNV) without destruction of the nuclear envelope (Buchatsky and Raikova, 1979; Barreau *et al.*, 1996). Almost all tissues, except the midgut epithelium, are susceptible to infection (Ward *et al.*, 2001; Jousset *et al.*, 1993). Infected larvae become lethargic, exhibit body contortions, develop malformed parts as pupae and have inclusion bodies within their tissues (Clark and Chapman, 1969; Lebedeva *et al.*, 1975). Large scale production of AaeDNV has been

attempted with varying success, as a mosquito control method (Buchatsky *et al.*, 1987; Suchman and Carlson, 2004).

Mosquito iridescent viruses (MIVs) and invertebrate iridescent virus 6 (ILV-6), from the family Iridoviridae, have been shown to infect mosquitoes. MIVs are very host specific and have been reported to infect *Aedes*, *Ochlerotatus*, *Psorophora*, *Culex*, and *Culiseta* species (Becnel and White, 2007; Woodard and Chapman, 1968, Becnel and Fukuda, 1989). Infections by this virus have been proposed to occur either by horizontal transmission in the larval stage or vertical transovarial transmission (Linley and Nielson, 1968). The genome is composed of a single, linear double stranded DNA. The virus has been shown to replicate and assemble in the cytoplasm of fat body cells of the host. Only 20% of larvae exposed horizontally to MIV become infected, whereas very high levels of vertical transmission occur (Woodard and Chapman, 1968, Undeen and Fukuda, 1998; Fukuda and Clark, 1975). Infected larvae usually die due to destruction of the fat body. *Ae. aegypti* have been shown to be susceptible to infection by ILV-6, which was originally isolated from a Lepidopteran species. Infections are sub-lethal and result in a loss of fitness that leads to a shortened lifespan and reduced fecundity (Marina *et al.*, 1999, 2003).

Mosquito cypoviruses (MCPVs) are one of many insect cytoplasmic polyhedrosis viruses in the genus *Cypovirus*, family Reoviridae (Belloncik and Mori, 1998, Shapiro *et al.*, 2005). Besides MCPVs that are pathogenic to the mosquito vector, other notable arboviruses like Colorado tick fever virus and bluetongue virus are members of the family Reoviridae. MCPVs have the ability to infect a wide range of hosts that include, but are not limited to the genera *Culex*, *Aedes* (*aegypti*), *Ochlerotatus*, *Uranotaenia*, and

Anopheles. MCPVs have a double-stranded, ten-segment RNA genome and assemble in an inclusion body (Shapiro *et al.*, 2005). Replication and assembly take place in the cytoplasm of cells of the gastric ceca and posterior midgut (Andreadis, 1981, 1986; Shapiro *et al.*, 2005). Upon entry into the cytoplasm, the reovirus occlusion body is relatively stable and contains virus-encoded transcriptase (Banerjee *et al.*, 1970; Drayna *et al.*, 1982; Levin *et al.*, 1970). The products of transcription are single-stranded mRNAs that undergo capping and methylation at their 5' ends prior to being released into the cytoplasm for translation (Furuichi *et al.*, 1977; Yamakawa *et al.*, 1982). Replication and packaging of the reoviral genome RNAs are not fully understood, but each packaged plus strand RNA serves as a template for synthesis of one minus strand copy and thereby regenerates genomic double-stranded RNA (Acs *et al.*, 1971; Sakuma *et al.*, 1971, 1972). This mode of replication allows the viral genome to be protected from mosquito host recognition and triggering of the RNAi innate immune response. The mechanism for spread of MCPVs within the insect is unknown. Shapiro *et al.* (2005) identified non-occluded virions that accumulated in the microvilli in both adults and larvae that might be involved in lateral transmission within the midgut (Andreadis, 1981). Horizontal transmission from larva to larva has been demonstrated through direct exposure of larvae to viral inclusion bodies (Andreadis, 1981; Federici, 1973; Shapiro *et al.*, 2005). Very little virus-induced mortality in infected larvae has been demonstrated. Virus infection is confined to the cytoplasm of epithelial cells of the gastric ceca and posterior midgut and does not noticeably affect development, feeding, or behavior.

The newly characterized *Aedes pseudoscutellaris* reovirus has a 9 segmented, non-occluded double-stranded RNA genome (Attoui *et al.*, 2005). This virus chronically

infects mosquitoes, suggesting the possibility of co-evolution between the virus and the innate immune response of the host.

The Nodamura virus has the ability to lethally infect insects and mammals (Venter and Schneemann, 2008). It is the type species of the genus *Alphanodavirus* of the Nodaviridae family and has a bipartite, positive-strand RNA genome. This virus was first isolated in 1956 from a pool of *Culex tritaeniorhynchus* mosquitoes collected in Japan (Scherer and Hurlbut, 1967). It has been shown to produce flaccid ascending paralysis and death in mammals (Scherer *et al.*, 1968). The virus has also been shown to infect and cause varying levels of pathogenesis in *Aedes*, *Toxorhynchites*, and *Culex* mosquito species (Tesh, 1980). Since it has the ability to infect and cause pathogenesis in both vertebrates and invertebrates, questions remain as to whether Nodamura virus should be considered an arbovirus or pathogenic mosquito virus. RNA viral replication complexes of the Nodamura virus are associated with modified intracellular membranes (Salonen *et al.*, 2005). Spherules are formed within the outer membrane of mitochondria of infected cells to provide a protective environment for viral RNA replication (Miller *et al.*, 2002). Positive-strand viral mRNAs are translated and amplified by virus-encoded RNA-dependent RNA polymerase through negative-strand intermediates. These RNA replication intermediates have the potential to trigger the insect's RNA interference innate immune response. In addition to sequestering their replication complexes in spherules during replication, these viruses encode the B2 protein, which is a viral suppressor of RNAi and protects the virus from host RNAi responses and promote replication and pathogenesis (Johnson *et al.*, 2004; Aliyari *et al.*, 2008).

Objectives and goals

Dengue virus is a serious global health concern (Malavige *et al.*, 2004). The *Ae. aegypti* mosquito is the principal vector of yellow fever and dengue virus (Gubler et al., 2004; Monath 1994, 1991). Mosquito management programs and public health education are the only effective control methods available to reduce dengue transmission and disease burden. A better understanding of virus-vector interactions is necessary to develop more effective mosquito control methods. Limited information exists as to what genetic components contribute to variations in vector competence. The purpose of this dissertation was to investigate *Aedes aegypti* and dengue virus anatomic, genomic, and molecular determinants of vector competence.

Chapter 2 describes "*Aedes aegypti* vector competence for DENV-2 in the Veracruz state and Yucatan peninsula of Mexico". Three previous studies documented variations in vector competence. Changes in human travel, urbanization, and weather events all have the potential for altering mosquito population sizes, breeding structure, and vector competence. The null hypothesis of this chapter is that minimal variation in vector competence, midgut infection barrier, and midgut escape barrier for DENV-2 exist in populations of *Ae. aegypti* in the states of Veracruz, Campeche, Yucatan, and Quintana Roo, Mexico. Rejection of the null hypothesis indicates that population genetic components (e.g. gene flow, isolation by distance) and/or molecular mechanisms contribute to extensive vector competence variation within specific populations.

Chapter 3 describes "Abundant nuclear copies of mitochondrial origin (NUMTs) in the *Aedes aegypti* genome". Upon completion of the vector competence study in Chapter 2, we attempted to complete a population genetics gene flow study with the 19

Ae. aegypti collections. As we amplified a portion of the *Aedes aegypti* mitochondrial NADH dehydrogenase subunit 4 gene (ND4), we identified that the amplified fragments were similar to ND4 but contained heterozygous sites. We suspected that nuclear copies of mitochondrial origin (NUMTs) exist in the *Ae. aegypti* genome. Further blast searches and analyses were completed to fully describe the existence of NUMTs *Ae. aegypti.*

Chapter 4 describes "Evidence of multiple chromosomal inversions in *Ae. aegypti formosus* from Senegal". The original outline of this study was to identify, at a molecular level, the potential contribution of the RNAi, microRNA, and apoptosis pathway with the DENV-2 virus dissemination phenotype in *Ae. aegypti.* An intercross mosquito family was generated to develop a linkage map with these genes of interest. As the linkage map data was collected and analyzed, it became apparent that multiple inversions existed with one of the mosquito subspecies used for generating the intercross family. These inversions are described in this chapter.

Chapter 5 describes "Linkage mapping of RNAi, microRNA, and apoptosis genes to identify associations with midgut infection (MIB) and midgut escape barriers (MEB) to infection with DENV-2 in *Aedes aegypti*". The null hypothesis was that DENV-2 infections are equal in each genotype class or that RNAi and/or apoptosis markers will randomly segregate. Rejection of the null hypothesis would indicate possible linkage between the marker locus and a significant association between the susceptibility for DENV-2 phenotype and a specific allele inheritance. When a significant χ^2 statistic was detected, we examined the inheritance of the alleles at that locus. Our *a priori* hypothesis is that an excess of F_2 and F_5 individuals with an allele inherited from the susceptible

D2S3 P_1 parent will become infected while an excess of F_2 individuals with an allele inherited from the PK10 P_1 parent will not become infected.

Chapter 6 describes "Analysis of evolutionary rates of RNAi and microRNA genes in *Aedes aegypti*". The RNAi pathway was identified in Chapter 5 to significantly contribute to DENV-2 dissemination phenotype in *Ae. aegypti*. Previous research has also identified that the genes in this pathway are some of the fastest evolving genes in *Drosophila*. Our final research question was to determine the rate at which RNAi genes are evolving and to identify whether arboviruses contribute to RNAi evolution in *Ae. aegypti*. The null hypothesis is that rates of non-synonymous versus synonymous changes between RNAi and miRNA genes are the same within *Ae. aegypti* collections from different world geographical regions. Rejection of the null hypothesis indicates significant differences in the rate of non-synonymous versus synonymous changes between RNAi and miRNA within *Ae. aegypti* collections from different world geographical collections and a possible association with various selective pressures.

Chapter 2

Aedes aegypti vector competence for DENV-2 in the Veracruz state and the Yucatan peninsula of Mexico

Abstract

Southern Mexico has historically recorded many dengue cases due to transmission by *Aedes aegypti* mosquitoes. Vector competence (VC) is defined as the intrinsic permissiveness of an arthropod vector to infection, replication, and transmission of a virus. Mosquito populations collected in Vera Cruz, Mexico in 2003 showed significant variance in VC for DENV-2 north and south of the Neovolcanic axis, a geological barrier to gene flow. To determine whether variation in VC phenotype is temporally stable which had previously been characterized, I collected *Ae. aegypti* eggs in 2005 from 19 sites in the Vera Cruz state and Yucatan peninsula of Mexico and determined VC for DENV-2. Immunofluorescence assays (IFA) were performed on the heads and midguts to determine the presence of DENV-2 antigen. VC was defined as the proportion of total IFA positive head tissues in each collection. VC for DENV-2 (JAM1409) in these collections ranged from 38% to 83%, the midgut infection barrier rates ranged from 9-30% and the midgut escape barrier (MEB) rates ranged from 7-36%. Spatial variations in VC continued to exist between collection sites; however, similar relative VC rates were found in specific sites when analyzed over time.

Introduction

Dengue has become a major international public health concern with two fifths of the world population at risk of infection. It is estimated that 50-100 million cases of classical dengue fever and 500,000 cases of dengue hemorrhagic fever (DHF) occur annually (Monath, 1994; Gubler, 1998).

Aedes aegypti is the most common vector of yellow fever and dengue (DENV) flaviviruses (Gubler, 2002). *Ae. aegypti* is a synanthropic, container-breeding mosquito

with a global tropical and subtropical distribution (Gubler, 1979). Vector competence (VC) refers to the intrinsic permissiveness of an arthropod vector to infection, replication and transmission of a virus (Hardy, 1988; Woodring *et al.*, 1996). Environmental conditions, viral and vector genetics are important factors that contribute to VC (Myles *et al.*, 2004). Because of its public health importance with regards to flavivirus transmission, *Ae. aegypti* has been the subject of many VC and population genetic studies (Aitken *et al.*, 1977,; Tabachnick *et al.*, 1985; Rosen *et al.*, 1985; Gubler *et al.*, 1979; Tabachnick and Powell, 1979; Tardieux *et al.*, 1990; Miller and Mitchell, 1991; Apostol *et al.*, 1996; Bosio *et al.*, 1998 and 2000; Vazeille-Falcoz *et al.*, 1999; Gorrochotegui-Escalante *et al.*, 2000 and 2002).

Vector competence is associated with a number of anatomic or physiological barriers to productive vector infection. These include a midgut infection barrier (MIB), a midgut escape barrier (MEB) and a salivary gland barrier (SGB) (Woodring, 1996; Bosio *et al.*, 1998; Black *et al.*, 2002). In vectors that exhibit a MIB, virus is unable to infect and/or replicate in the mosquito midgut cells. A vector that has a MEB is permissive to viral infection and replication within the midgut, but does not allow the virus to establish a disseminated infection. Barriers to infection and dissemination can significantly vary in frequency and magnitude among *Ae. aegypti* populations.

Lozano-Fuentes *et al.* (2009) previously discovered that the intersection of the Neovolcanic axis (NVA) with the Atlantic coast in the state of Veracruz, Mexico acts as a discrete barrier to gene flow among *Ae. aegypti* populations north and south of the NVA. The mosquito populations north and south of the NVA also differed in their VC for DENV-2. The average VC rate for *Ae. aegypti* mosquitoes from populations north of the

NVA was 55%; in contrast the average VC rate for mosquitoes from populations south of the NVA was 20%. Most of this variation was attributable to a MIB; 21.5% of *Ae. aegypti* north of the NVA and 45.2% mosquitoes south of the NVA failed to develop midgut infections. However, Lozano-Fuentes *et al.* (2009) only examined VC in a single year (2004).

The purpose of this study was to re-evaluate the DENV-2 VC patterns in *Ae. aegypti* in the Veracruz state and the Yucatan peninsula and to determine their stability two years later. I hypothesize that the NVA will continue to act as a geological barrier to gene flow, confirmed by finding that northern and southern *Ae. aegypti* populations continue to exhibit differences in DENV-2 VC. I hypothesize that *Ae. aegypti* populations south of the NVA will continue to exhibit low VC while mosquito collections from the Yucatan peninsula will exhibit high DENV-2 VC as a result of limited gene flow patterns previously observed (Gorrochotegui-Escalante *et al.,* 2002; Lozano-Fuentes *et al.*, 2009). I also hypothesize that comparing analyses over time will show significant variation in VC attributable to differences in collection sites as well as year to year variation.

To test these hypotheses, *Ae. aegypti* were collected from 19 sites in the states of Veracruz, Campeche, Yucatan and Quintana Roo, Mexico, and DENV-2 JAM1409 VC of each was determined by analysis of MIB and MEB.

Material and methods

Mosquito collections

Ae. aegypti eggs were collected from the sites in Mexico shown in Figure 2.3. F_1 eggs were shipped to Colorado State University's Arthropod-borne and Infectious

Diseases Laboratory. Table 2.1 describes the state and city where the eggs were collected, the longitude and latitude of each collection site, month of collection and the number of mosquitoes used from each site to acquire VC data. The susceptible *D2S3* laboratory strain was also used as a positive control in VC experiments (Bennett *et al.*, 2005).

Mosquito rearing

Ae. aegypti field collections were raised at a constant temperature of 27°C and 80% relative humidity in an insectary with a 14 hour photoperiod. Eggs were hatched on day 0 by immersion in tap water. On day 1, approximately 1000 newly emerged larvae were removed from the hatch pan and placed into a large plastic pan containing ~10 liters of room temperature tap water and ground fish/rabbit food. On day 6, pupae were removed from the larvae pan and placed into small plastic cups. The pupal cups were placed in cages for emergence to occur. Adult females were offered a non-infectious defibrinated sheep blood meal every 3 days to maximize egg production, starting on day 12-13 post hatch. After a sufficient number of eggs were acquired from the F_1 adults, all viable mosquitoes were collected and individually separated into 0.5 ml microcentrifuge tubes and stored at -70°C for DNA extraction. If the F_2 females were to be used for VC studies, the female pupae were separated and transferred into 1 quart cardboard cartons in preparation for artificial infectious blood meal.

Virus propagation

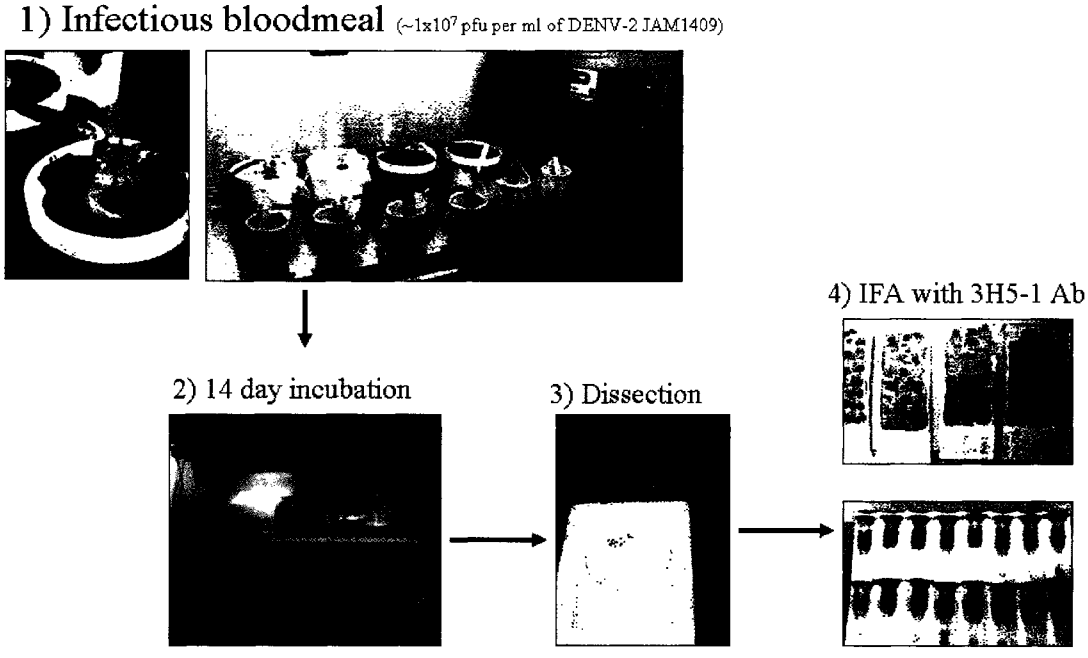
DENV-2 JAM1409 was first isolated in 1983, as described in Deubel *et al.* (1986). It was subsequently been passaged multiple times in C6/36 mosquito cells. Stock virus was maintained at -70°C. C6/36 cells are used to amplify virus for VC

studies. Cells were maintained as a monolayer at 28°C in Leibowitz L-15 medium supplemented with 10% fetal bovine serum (FBS), 2 mM L-glutamine (Gibco), 1X nonessential amino acids solution (Cellgro), 100 U/ml of penicillin, and 100 µg/ml of streptomycin. C6/36 cells in 25 cm^2 flasks were infected at a multiplicity of infection (MOI) of 0.01 in 1 ml of L-15 (2% FBS) medium at room temperature on a slow rocker for 45 minutes. Following incubation, 6 ml of L-15 (2% FBS) medium were added to each flask. Infected flasks were incubated for 7 days at 28°C, when medium was replaced with fresh medium. At 9-12 days post infection when the C6/36 cell monolayer demonstrated significant syncytia, virus was harvested by scraping the monolayer into suspension and transferring the medium, cells and virus as one suspension.

***Per os* infection with DENV-2 JAM1409**

Approximately 200 female mosquitoes, 12 days post hatch, were aspirated into 1 quart cardboard cartons with double mesh organdy covering the top and sealed shut with masking tape. The sugar source was removed ~24 hours and the water source was removed ~6 hours prior to blood feed. Infected cells and medium were mixed 1:1 with defribrinated sheep blood and placed in membrane feeders covered with hog gut. The blood meal was maintained at 37°C via circulating water in the jacket of the membrane feeders, which were placed on top of the cartons for one hour, allowing the mosquitoes to take a blood meal (Figure 2.1). The one hour feeding schedule was enforced to maintain consistency and to prevent any significant viral titer decrease during the feed. Engorged mosquitoes were immobilized by lowering the temperature to 4° C, sorted, counted, placed into cartons, and returned to insectary for 14 days. Infected cell suspension and the 1:1 cell suspension and sheep blood mixture were stored at -70°C until virus plaque

Figure 2.1. Mosquito head/midgut infection assay. Mosquito artificial infection and immunofluorescence assay for DENV-2 in *Ae. aegypti*.



titrations could be performed.

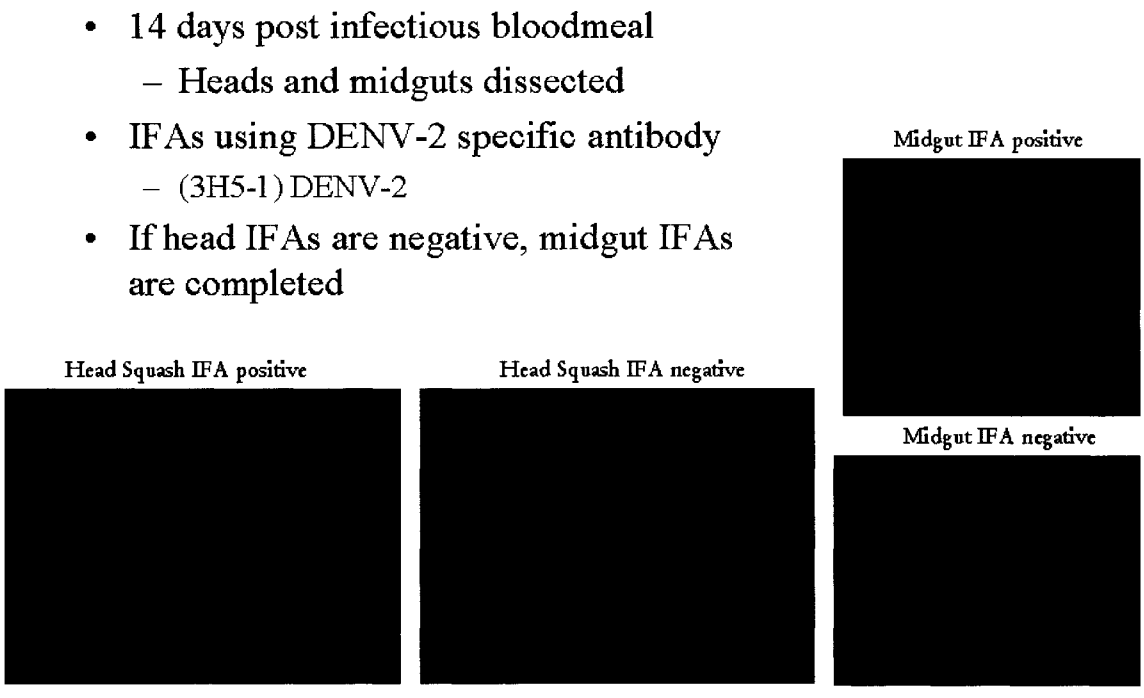
Plaque titrations

Infected cell suspension and 1:1 blood/cell suspension were titrated using protocols developed by the Centers for Disease Control and Prevention in the *Manual on Dengue Diagnostic Laboratory Procedures for the Americas* (1981). Ten-fold serial dilutions of infected cell suspension and blood/cell suspensions were overlaid on LLC-MK_2 cell monolayers in 6-well plates containing agar and cell medium. Plaques were counted and scored approximately 12 days following the infection.

Vector competence

Mosquitoes were phenotyped for VC 14 days post oral infection using indirect immunofluorescence assay (IFA). Mosquitoes were dissected to sever the heads and remove the midguts. The heads were squashed onto a VECTABOND (Vector Laboratories, Burlingame, CA) treated slide and fixed in acetone. Midguts were placed individually in microtubes containing 4% paraformaldehyde in PBS. IFA for heads and midguts followed the protocol described by Bennett *et al.* (2002). Slides containing squashed mosquito heads were incubated for 60 minutes at 37°C with a 1:200 dilution of DENV-2 serotype specific mouse derived monoclonal antibody 3H5-1 (ATCC number HB-46) in PBS. After primary incubation, slides were rinsed 3 times in PBS. Slides were incubated for 45 minutes at 37°C with a 1:200 dilution of biotinylated sheep anti mouse secondary antibody (Amersham/GE), , and 1% Evans blue counterstain. Slides were rinsed 3 times, then incubated for 30 minutes at 37°C with 1:200 streptavidin fluorescein (Amersham/GE) in PBS. Slides were then rinsed two times in PBS and once in water. VectaShield (Vector Laboratories, Burlingame, CA) was applied to the head

Figure 2.2. DENV-2 immunofluorescence assay. Immunofluorescence detection assay for DENV-2 in the heads and midguts of *Ae. aegypti*.



squashes and coverslips were overlaid. Heads were examined at 200x using an Olympus BH2-RFCA fluorescence microscope (Figure 2.2). Mosquito midguts underwent IFA if virus antigen was not detected in their respective heads. The protocol for IFA of mosquito midguts is the same as for mosquito head squashes with the following modifications: Midguts were incubated and washed individually in microtubes and the rinses for midguts contained 0.1% PBT solution (0.2% Triton X-100, 1% BSA in 1X PBS-Ashburners). This reduced midgut loss from slides while improving IFA results by enhancing antibody-antigen interaction.

IFA data were used to determine midgut infection rate (MIR), midgut infection barrier (MIB), midgut escape barrier (MEB), and vector competence (VC). This was calculated by determining virus infections in the midgut with respect to virus dissemination to the mosquito head. The number of positive midguts was divided by number of mosquitoes exposed to infectious blood meal to determine the MIR. DIR was calculated as the number of mosquitoes with virus antigen in the head divided by the number of mosquitoes with virus antigen in the midgut. MIB was determined by calculating 1-MIR. MEB was determined by calculating MIR – DIR. VC was calculated as the number of mosquitoes with virus antigen in the head divided by the number of mosquitoes exposed to infectious blood meal. VC was assumed to be the same as the rate of virus dissemination to the head tissues. To maintain consistency with previous studies, mosquito salivary glands were not dissected or assayed and actual virus transmission studies were not performed.

Analysis of variance

Analysis of variance was performed to study variation in VC, MIB and MEB with respect to the geological barrier of the NVA, collection site, and year of collection. Analysis of variance computations were performed using SAS software (SAS Institute).

Results

Variation in vector competence and anatomic barriers to transmission

All VC results are based on sample sizes greater than 111, with a mean of 152 mosquitoes. A minimum of 2 oral infection repetitions, with at least 30 mosquitoes per replicate were conducted for each collection site (Table 2.2). Approximately 40 *Ae. aegypti D2S3* were also included in every oral challenge as a positive control. This laboratory colony has consistently demonstrated a VC >80%. VC data were collected for every oral challenge but were only used in the analysis when the DENV-2 viral titer in the blood meal was greater than $1x10^6$ plaque forming units/ml and the *D2S3* positive control group had a dissemination rate above 80%.

The VC of mosquitoes collected from the 19 geographic sites is shown in Figures 2.4, 2.5 and Table 2.2. Significant variation in VC was determined by χ^2 analysis ($P < 0.05$). The VC ranged from 38.0% in the Lerdo de Tejada, Veracruz collection to 83.0% in the Chetumal, Quintana Roo collection. The mosquito collections from Yucatan and Quintana Roo states exhibited a greater mean VC (75.4%) than collections from Veracruz state (54.6%) ($P < 0.05$). The VC was also more variable and had a larger range between sites within Veracruz state (38.0%-70.3%) than between sites within Yucatan and Quintana Roo states (71.7%-81.3%).

Anatomic basis of transmission barriers

To determine anatomic barriers that condition transmission, mosquitoes from the 19 *Ae. aegypti* collections and the *D2S3* colony were phenotyped for MIB and MEB (Table 2.2, Figures 2.4, 2.5). The collections exhibited a wide range of expression in MIB and MEB for DENV-2. MIB rates ranged from 9.3% in Chetumal, Quintana Roo to 30.1% in Paso del Cedro, Veracruz. MEB rates ranged from 7.6% in Chetumal, Quintana Roo to 36.4% in Lerdo de Tejada, Veracruz (Figure 2.4).

Correlation between MIB and MEB

Figure 2.9 is a plot of the percentage of mosquitoes exhibiting a MIB versus those exhibiting a MEB to determine whether MIB and MEB are independent of one another or whether a correlation exists between them. Additional data from previous VC studies at the same sites, conducted in 1998-1999, 2001, and 2003 are included in this analysis (Bennett *et al.*, 2002; Lozano-Fuentes *et al.*, 2005 and 2009). Regression analysis was conducted to determine correlation for MIB and MEB (i.e. correlation reported as an r value). A correlation ($r = 0.63$) existed between MIB and MEB among the 19 collections of 2005. The correlation between MIB and MEB was substantially higher (r ranging from 0.82 - 0.90) within the 6 sites also in the 1998-1999 and 2001 VC studies (Figure 2.9). No correlation ($r = 0.03$) was found between MIB and MEB from the 7 sites included in the 2003 VC study.

Analysis of variance with respect to the Neovolcanic axis

Analysis of variance was performed using all four collection years (1998-1999, 2001, 2003, 2005) for VC, MIB and MEB with respect to the NVA (Table 2.3). Most of the variation respectively in VC and MIB (49.79% and 49.74%) are attributed to year of

collections. Smaller variations in VC and MIB (18.98% and 4.96%) were attributed to whether collections were made north or south of the NVA. Whereas, most variation (34.68%) in MEB was due to whether collections were made north or south of the NVA, compared with smaller variation (4.22%) attributed to year of collections. Finally, interaction between the two terms, NVA and collection year, is a substantial contributor to variations in VC, MIB and MEB (ranging from 31.23% - 61.11%).

Analysis of variance with respect to collection site

Analyses of variance in VC, MIB and MEB with respect to collection site and year of collection are shown in Table 2.3, Figures 2.6, 2.7, and 2.8. Most of the variation in VC and MEB (44.35% and 33.8%) are attributed to differences between collection sites, compared to variation (29.4% and -4.45%) attributed to year of collections (Figures 2.6, 2.8). Most variation (45.39%) in MIB is attributed to year of collection compared to smaller variation (21.46%) attributed to collection sites (Figure 2.7). Finally, interaction between NVA and collection year is a substantial contributor to variation in VC, MIB and MEB (ranging from 26.25% - 66.20%).

Discussion

During the Spring/Summer of 2005, *Ae. aegypti* eggs were collected at 19 sites in the states of Veracruz, Campeche, Yucatan and Quintana Roo, Mexico to analyze VC, MIB and MEB for DENV-2. This was the fourth study over the course of eight years to evaluate VC in *Ae. aegypti* populations within these areas where DENV-2 transmission occurs. This study tested for stability of VC within and between populations of mosquitoes in 2005, and compared findings to similar studies in 1998-1999, 2001, and 2003.

Ae. aegypti collected in 2005 exhibited considerable variation in VC, which ranged from 38.0% to 83.0%. Collections from the Yucatan peninsula demonstrated higher VC (Figures 2.4, 2.5). Collections from Chetumal continued to consistently have a very high VC due to low MIB and MEB for DENV-2. In 2003, Lozano-Fuentes *et al.* encountered a region around the city of Veracruz (e.g. Alvarado) where *Ae. aegypti* were highly refractory (11%) to DENV-2 dissemination. Mosquitoes collected two years later (2005) from Lerdo de Tejada and Alvarado continued to exhibit low rates of DENV-2 VC (38% and 45%). Even though these findings in these two collection sites were not as divergent as those observed in 2003, this study demonstrated that VC remained variable and that a DENV-2 refractory mosquito population continued to exist around Veracruz, Mexico.

Ae. aegypti collections in this study differed substantially in their MIB and MEB for DENV-2. Mosquitoes exhibited MIB that ranged from 9.3% in the Chetumal, Quintana Roo collection to 30.1% in the Paso del Cedro, Veracruz collection. MEB rates ranged from 7.6% in Chetumal, Quintana Roo, to 37.0% in the Lerdo de Tejada, Veracruz, collection. In almost every situation, MIB and MEB followed an inverse trend with VC, suggesting that VC, MIB and MEB have genetic determinants.

Lozano-Fuentes et al. (2009) previously discovered the NVA to be a discrete barrier to gene flow, potentially contributing to differences in DENV-2 VC among *Ae. aegypti* populations. Most of the variation they observed was attributed to MIB. Poza Rica was the only site north of the NVA where mosquito eggs were collected in 2005 and they had a VC of 60%, MIB of 21% and MEB of 19%. Mosquitoes collected in 2005 south of the NVA continued to exhibit lower VC (38%-45%) and higher MIB (25%) and

MEB (30%-37%). An analysis of variance was performed for VC, MIB and MEB with respect to the NVA and the year of collection. Most of the variation (49.7%) in VC and MIB with respect to the NVA was attributed to differences in the collection year (i.e. 2003 and 2005). Most of the variation (34.68%) in MEB was attributed to whether the site of collection was north or south of the NVA. MEB variance analysis confirms the previous findings that the NVA acts as a discrete barrier to gene flow and has been maintained over numerous years. It is difficult though to interpret the observed variation specific to the NVA when taking MIB and VC variance analyses also into consideration and their direct relationship to MEB. One analysis (MEB) suggests the existence of a discrete barrier, while the other two analyses (VC, MIB) suggest substantial north and south gene flow or experimental variation.

VC, MIB and MEB of mosquitoes from the 1998-1999, 2001 and 2005 collections maintained similar variability within and between sites over eight years (Table 2.2). Data varied between years but no collection site population followed a directional trend and data were not erratic; however, between 2005 and 2003 there were substantial variations in VC and MIB. VC was higher in populations from 6 of 7 collection sites in 2005 compared to 2003. MIB was substantially higher in populations from 5 of 7 collection sites in 2003 compared to 2005.

An analysis of variance in VC, MIB and MEB with respect to collection sites and year of collections was performed (Table 2.3, Figures 2.6, 2.7, 2.8). For this analysis, VC, MIB and MEB were averaged by year and separately by collection site. Most variations in VC and MEB (44.35% and 33.8%) were due to differences in collection sites (Table 2.2, Figures 2.6, 2.8). Differences between collection years also contributed

substantially to VC variation (29.4%), whereas there was little MEB variation (-4.45%) observed between year of collection. Most variation (45.39%) in MIB was due to year of collection (Figure 2.7). The major variation in MIB was due to the 2003 collection. The collection from 2005, 2001 and 1998-1999 were relatively uniform with respect to MIB (16.2% to 25.8%), whereas MIB for 2003 was substantially elevated (41.6%). Differences in populations between collection sites also contributed substantially to MIB variation (21.46%). Substantial MIB variation was found between collections within the refractory region around Veracruz, whereas collections on the periphery, north of or near the NVA, and on the eastern side of the Yucatan peninsula had MIB rates that were relatively uniform between sites.

Analysis of variance by year identified substantial differences in VC of populations from the same collection sites between 2003 and the other three years of study. For example, the Alvarado collection had the lowest VC in 2003 at 11%, whereas in 2005 it had a VC of 45%. VC in study sites, such as Chetumal, demonstrated relative consistency between three years of study with the exception of 2003 (e.g. 1998-1999--81%; 2001--83%; 2005--83%). Regression analysis of MIB and MEB also identified substantial correlative differences in the 2003 study compared to the other three studies (Figure 2.9). For example, 2005, 2001 and 1998-1999 data exhibited a significant degree of correlation ($r = 0.69$ to 0.90) between MIB and MEB, whereas no correlation ($r = 0.03$) was demonstrated with the 2003 study data.

One potential reason for the variation in VC in 2003 as compared to the other three collection years may be the study design. The 2005 study design followed that of Bennett *et al.* (2002) for the 2001 study, and used large sample sizes and multiple

repetitions to generate VC data. Sample sizes ranged from 111 to 199 with repetitions ranging from 2 to 4. The 2003 study had sample sizes that ranged from 47 to 77 mosquitoes with an unstated number of repetitions. Another key difference in the 2005 study was that a positive control *Ae. aegypti* strain (*D2S3*) with high VC was included. In repetitions in which the *D2S3* sample did not have a dissemination rate greater than 80%, the experiment data from that entire blood feed were eliminated from the analysis. This ensured consistency across numerous blood feeds. In a region where *Ae. aegypti* DENV-2 susceptibility has been identified as being low, it is important to have large sample sizes, numerous repetitions and consistent controls to reduce errors from experimental variation.

Another factor in variation between years might involve gene flow in the *Ae. aegypti* collections. The 2005 and 2003 collection sites in Veracruz state were very close to one another. For example, two collection sites (Minatitlan and Cosoleacaque) were within two kilometers of each other. Gorrochotegui-Escalante *et al.* (2002) demonstrated that substantial gene flow existed among *Ae. aegypti* populations in Mexico and significantly varied based on geographic region. Lozano-Fuentes (2009) also studied gene flow among populations and found correlation between phylogenies and infection rates suggesting the potential importance of geographical barriers.

Haplotype data were collected for individual mosquitoes in the 2005 study, using DNA sequences from the mitochondrial ND4 gene, to conduct gene flow analysis. Data from preliminary analyses suggested that the ND4 gene, as well as other mitochondrial DNA, has potentially been integrated into the nuclear genome (Nuclear copies of mitochondrial sequences - NUMTs). NUMTs occur in other insect species such as

Drosophila and honeybees but have not been observed in *Ae. aegypti*; however, numerous ND4 gene BLASTn searches suggested that certain mitochondrial DNA sequences have integrated into the nuclear genome (Pamilo *et al.,* 2007; Black WC and Bernhardt SA, Submitted 02/2009). This integration may cause difficulties in interpreting gene flow based on the mitochondrial gene due to the potential presence of multiple variable sequence copies. A different *Ae. aegypti* marker might be needed to conduct gene flow analysis.

A number of quantitative trait loci (QTL) studies have been conducted to further understand the genetic determinants that condition VC, MIB and MEB (Bosio *et al.,* 2000; Gomez-Machorro *et al.,* 2004; Bennett *et al.,* 2005). The correlation of MIB and MEB to VC, as well as the variation in VC observed in these studies, demonstrate that VC is a quantitative trait in natural *Ae. aegypti* populations. RNA interference (RNAi) has been shown to act as an innate immune response to infection with arboviruses in mosquitoes (Keene *et al.*, 2004; Campbell *et al.*, 2008; Sanchez-Vargas *et al.*, 2009). Data suggest that the RNAi pathway genes may contribute to MEB to arboviruses. Questions remain as to whether variations in VC in natural *Ae. aegypti* populations are due to genetic variations in the RNAi pathway genes. The natural variation in VC and the correlation to MIB and MEB supports linkage mapping determination of association of RNAi pathway genes with viral transmission.

Figure 2.3. Map of 19 *Ae. aegypti* collection sites in Mexico.

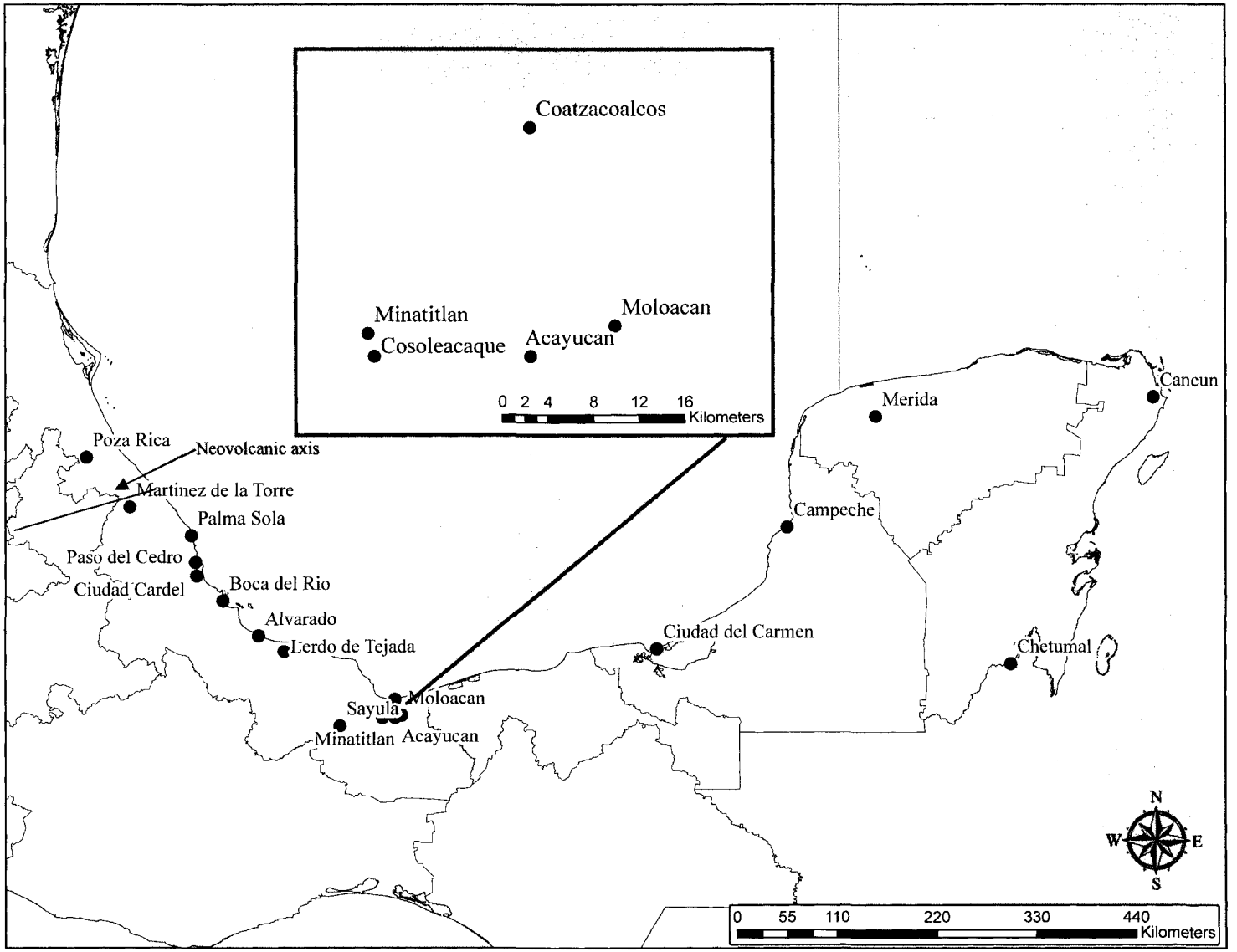


Table 2.1. Locations, dates of collections, coordinates, and sample sizes of *Ae. aegypti* collections in Mexico.

State	City	Date	Latitude (°N)	Longitude (°W)	Number of individuals analyzed for vector competence in the F_2 generation
Vera Cruz	Acayucan	06/2005	17.96196	-94.41255	199
Vera Cruz	Alvarado	06/2005	18.77422	-95.76356	112
Vera Cruz	Boca del Rio	06/2005	19.12331	-96.11798	148
Vera Cruz	Ciudad Cardel	06/2005	19.37113	-96.37649	120
Vera Cruz	Coatzacoalcos	06/2005	18.14081	-94.41310	111
Vera Cruz	Cosoleacaque	06/2005	17.96196	-94.53605	120
Vera Cruz	Lerdo de Tejada	06/2005	18.61932	-95.51033	142
Vera Cruz	Martinez de la Torre	06/2005	20.04999	-97.03883	134
Vera Cruz	Minatitlan	06/2005	17.97972	-94.54083	157
Vera Cruz	Moloacan	06/2005	17.98583	-94.34611	124
Vera Cruz	Palma Sola	06/2005	19.76936	-96.43028	199
Vera Cruz	Paso del Cedro	06/2005	19.50444	-96.3875	135
Vera Cruz	Poza Rica	06/2005	20.54366	-97.47079	181
Vera Cruz	Saylula	06/2005	17.88150	-94.95941	162
Campeche	Campeche	05/2005	19.90000	-90.60000	130
Campeche	Ciudad del Carmen	05/2005	18.59998	-91.80000	167
Yucatan	Merida	05/2005	20.95000	-89.64000	240
Quintana Roo	Cancun	05/2005	21.14000	-86.88000	130
Quintana Roo	Chetumal	05/2005	18.50000	-88.30000	171

Table 2.2. Comparison of VC, MIB and MEB in this study to previous studies.

State	Site	Total No. of mosquitoes fed with complete data	No. of repetitions	Total pos. heads	Vector competence or DIR (Pos heads/total mosquitoes	Previous VC	MIB (1-MIR)	Previous MIB	MEB (MIR-DIR)	Previous MEB
Vera Cruz	Poza Rica	181	4	109	0.60	0.74	0.21	0.13	0.19	0.02
	Martinez de la Torre	134	3	80	0.60	0.51	0.14	0.16	0.26	0.24
	Palma Sola	199	4	109	0.55		0.21		0.24	
	Paso del Cedro	135	2	71	0.53		0.30		0.17	
	Ciudad Cardel	120	2	57	0.48		0.28		0.25	
	Boca del Rio	148	4	70	0.47		0.25		0.27	
	Alvarado	112	2	50	0.45	0.11	0.25	0.64	0.30	0.06
	Lerdo de Tejada	142	3	54	0.38		0.25		0.37	
	Saylula	162	3	64	0.40		0.28		0.33	
	Minatitlan	157	3	96	0.61	0.2	0.20	0.52	0.19	0.07
	Cosoleacaque	120	4	77	0.64	0.13	0.16	0.48	0.20	0.28
	Acayucan	199	4	116	0.58	0.11	0.16	0.52	0.25	0.25
	Moloacan	124	4	86	0.69	.73, .57	0.12	0.16	0.18	0.11
	Coatzacoalcos	111	2	78	0.70	0.31	0.12	0.35	0.18	0.17
Campeche	Ciudad del Carmen	167	3	110	0.66	0.43	0.16	0.33	0.18	0.24
	Campeche	130	2	57	0.44	0.43	0.28	0.30	0.28	0.27
Yucatan	Merida	240	4	172	0.72	.90, .68	0.10	0.20	0.19	0.08
Quintana Roo	Cancun	130	2	94	0.72	.72, .75	0.12	0.25	0.16	0.02
	Chetumal	171	4	142	0.83	.81, .83	0.09	0.19	0.08	0.02
Lab Colony	D2S3	599	15	508	0.85	.86-100	0.07	0.06	0.08	0.09

Previous data cited from: Gorrochotegui-Escalante *et al.*, 2002; Lozano-Fuentes *et al.*, 2009; Bennett *et al.*, 2002 and 2005

Figure 2.4. 2005 vector competence, MIB and MEB of all 19 *Ae. aegypti* mosquito collections made in 2005 from Mexico.

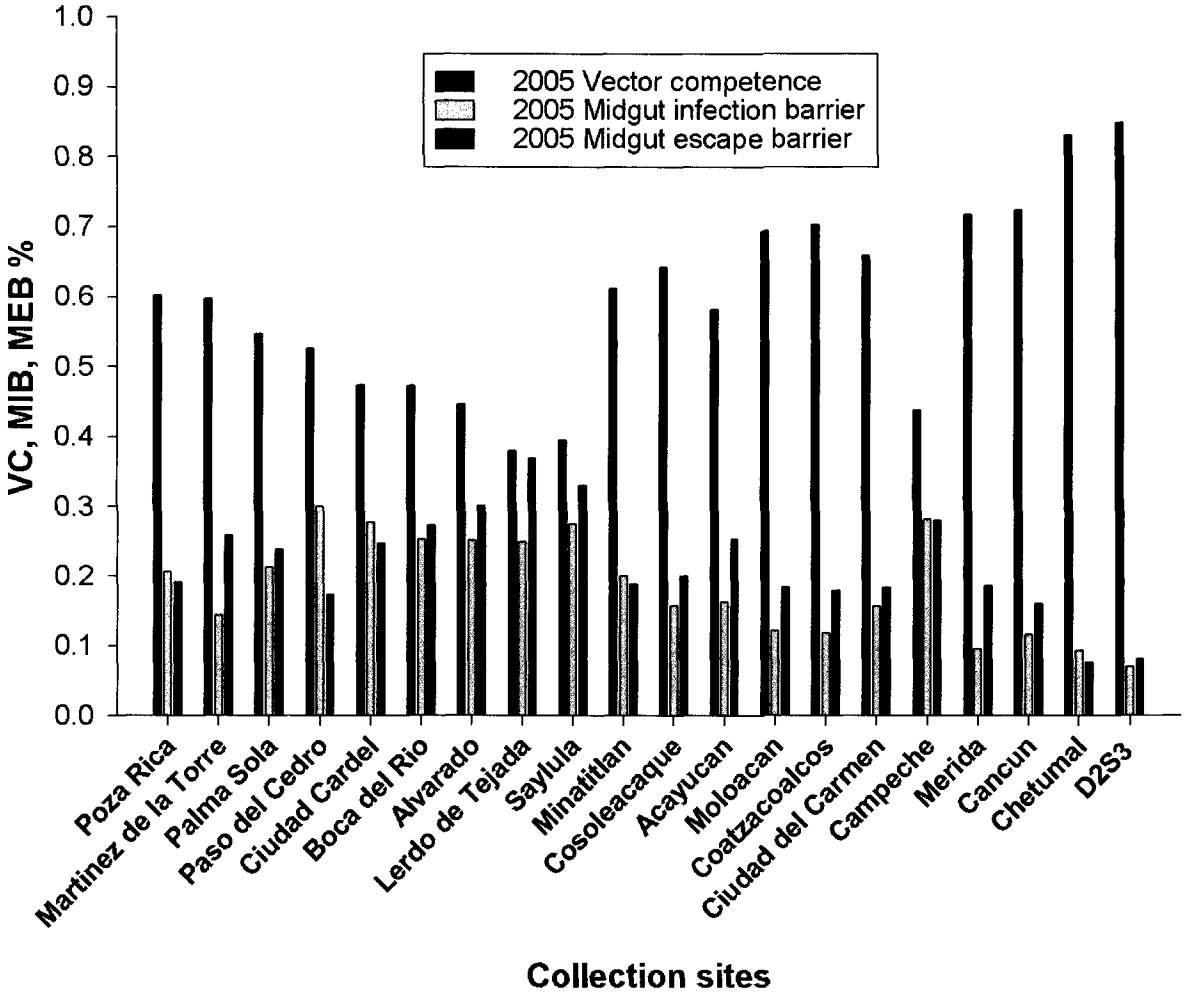


Figure 2.5. Geospatial analysis of vector competence, MIB and MEB of all 19 *Ae. aegypti* collections made in 2005 from Mexico.

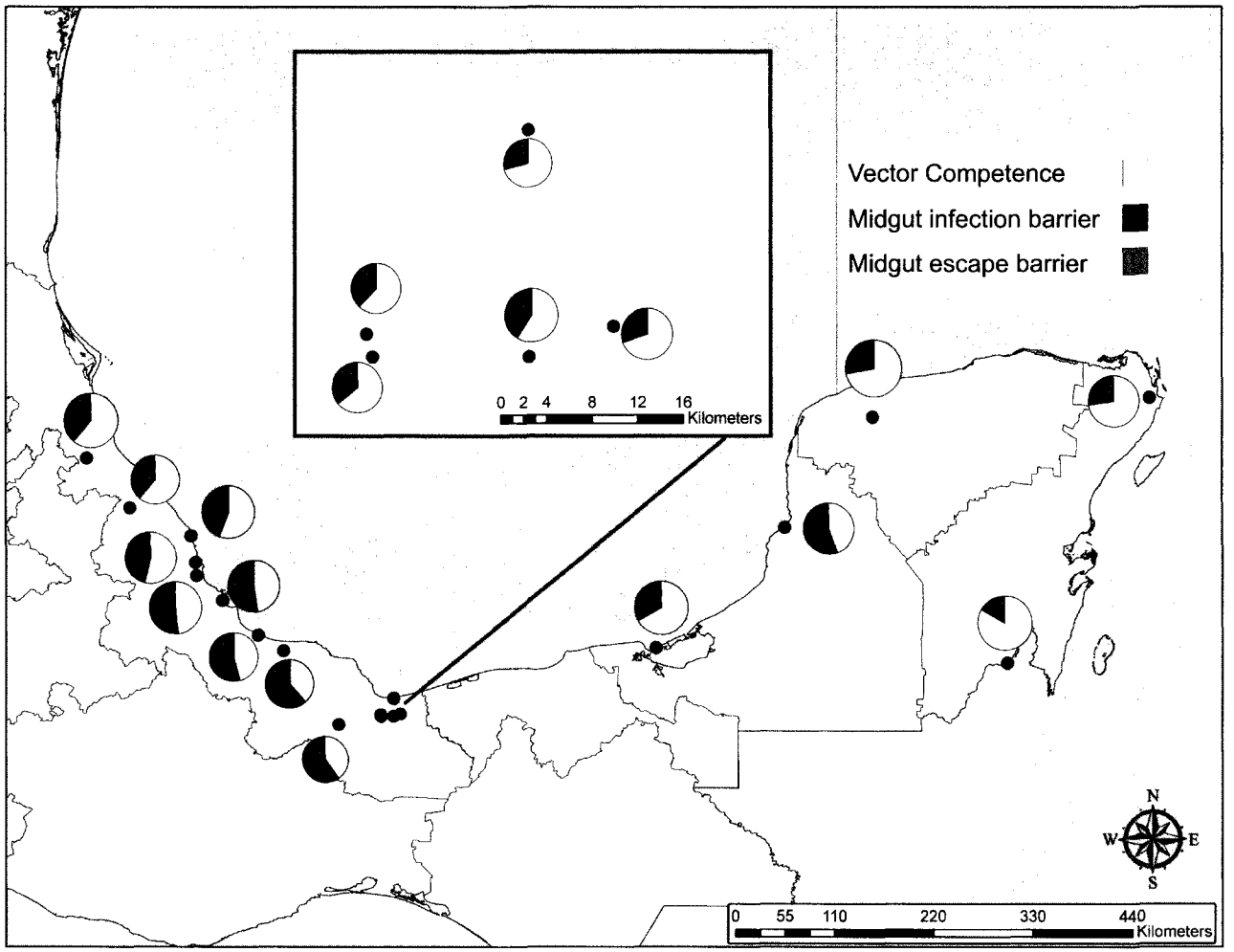


Table 2.3. Analysis of variance among *Ae. aegypti* collections in Mexico for VC, MIB, and MEB with regard to the Neovolcanic axis, collection sites and years.

Source of variation	Degree of Freedom	Sum of squares	Variance components	Variation (%)
Neovolcanic axis (NVA) by collection year				
Vector competence by NVA	1	0.0758	0.0144	18.98
Vector competence by Year	3	0.1521	0.0378	49.79
Vector competence Error (Interaction)	12	0.0237	0.0237	31.23
Midgut infection barrier by NVA	1	0.0251	0.0019	4.96
Midgut infection barrier by Year	3	0.2504	0.0194	49.74
Midgut infection barrier Error (Interaction)	12	0.2115	0.0176	45.3
Midgut escape barrier by NVA	1	0.016	0.003	34.68
Midgut escape barrier by Year	3	0.0198	0.0004	4.22
Midgut escape barrier Error (Interaction)	12	0.064	0.0053	61.11
Collection site by collection year				
Vector competence by Collection Site	12	0.8746	0.0223	44.35
Vector competence by Year	3	0.321	0.0148	29.4
Vector competence Error (Interaction)	16	0.2115	0.0132	26.25
Midgut infection barrier by Collection Site	12	0.3002	0.0053	21.46
Midgut infection barrier by Year	3	0.2391	0.0113	45.39
Midgut infection barrier Error (Interaction)	16	0.1318	0.0082	33.15
Midgut escape barrier by Collection Site	12	0.2042	0.0039	33.8
Midgut escape barrier by Year	3	0.0132	-0.0005	-4.45
Midgut escape barrier Error (Interaction)	16	0.1221	0.0076	66.2

Figure 2.6. Analysis of variance in vector competence with respect to collection site and collection year for *Aedes aegypti* collections in Mexico.

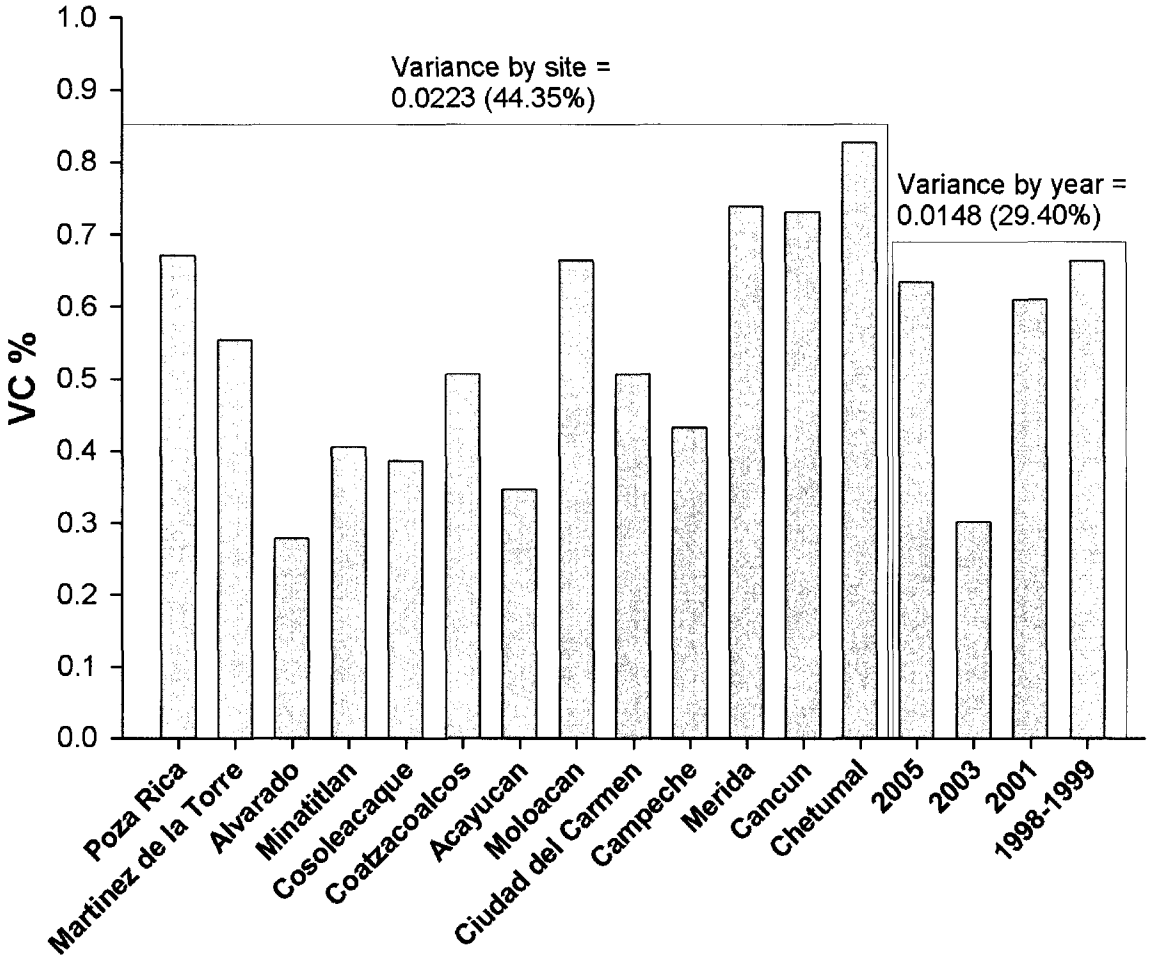


Figure 2.7. Analysis of variance in midgut infection barrier with respect to collection site and collection year for *Aedes aegypti* collections in Mexico.

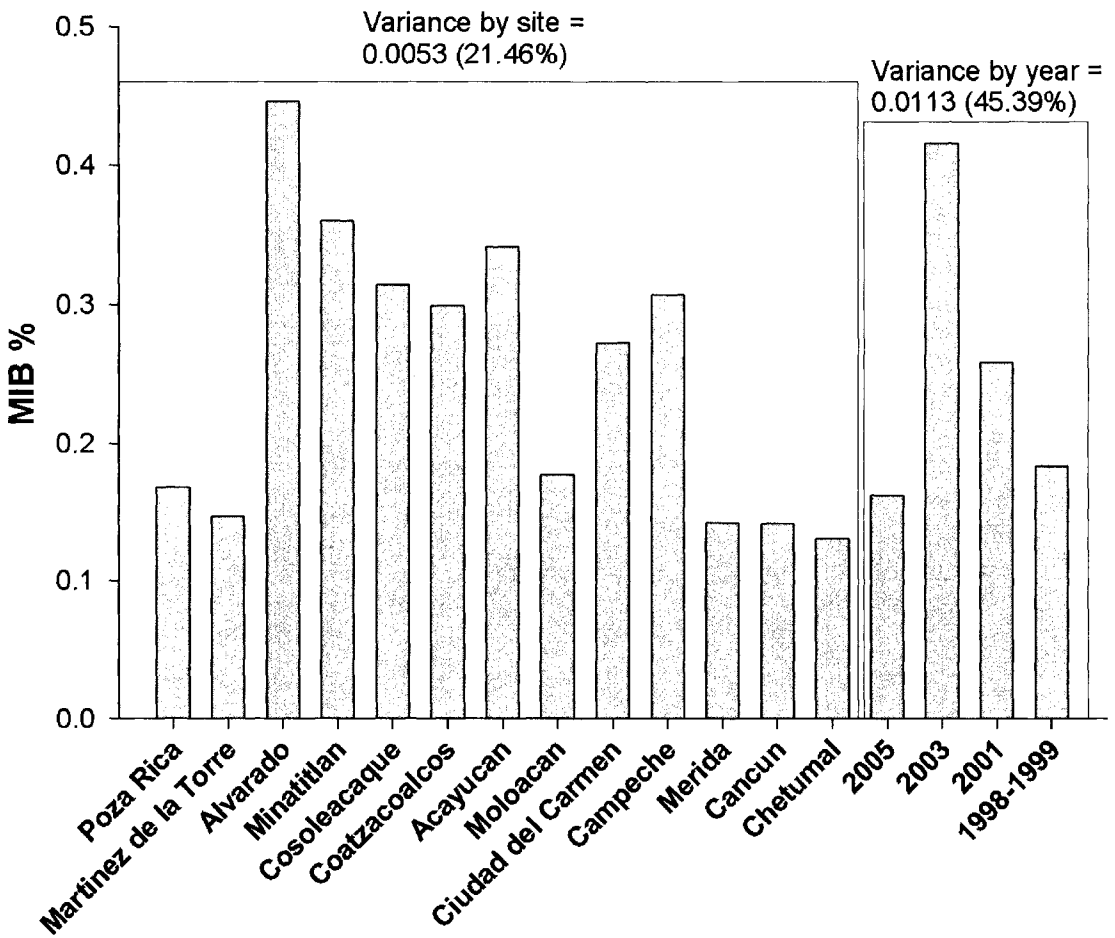


Figure 2.8. Analysis of variance in midgut escape barrier with respect to collection site and collection year for *Aedes aegypti* collections in Mexico.

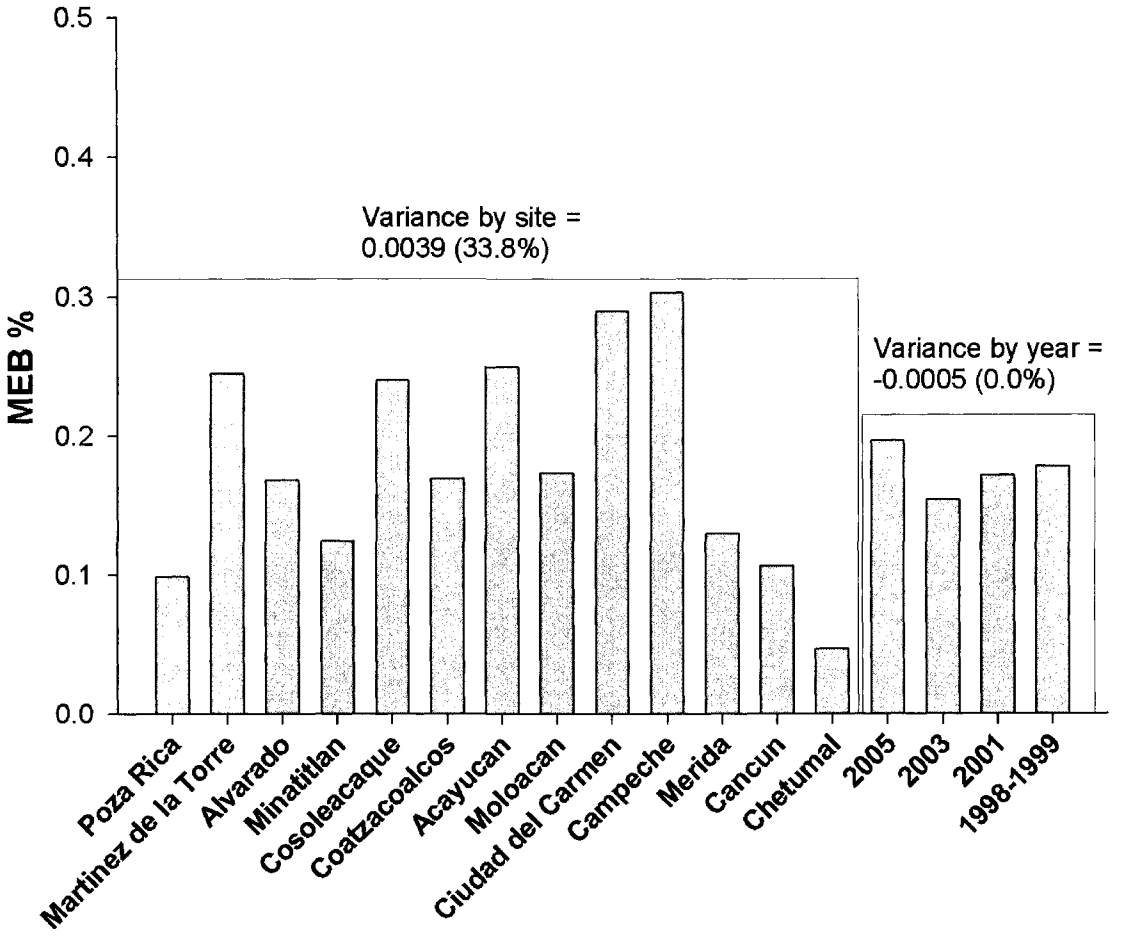
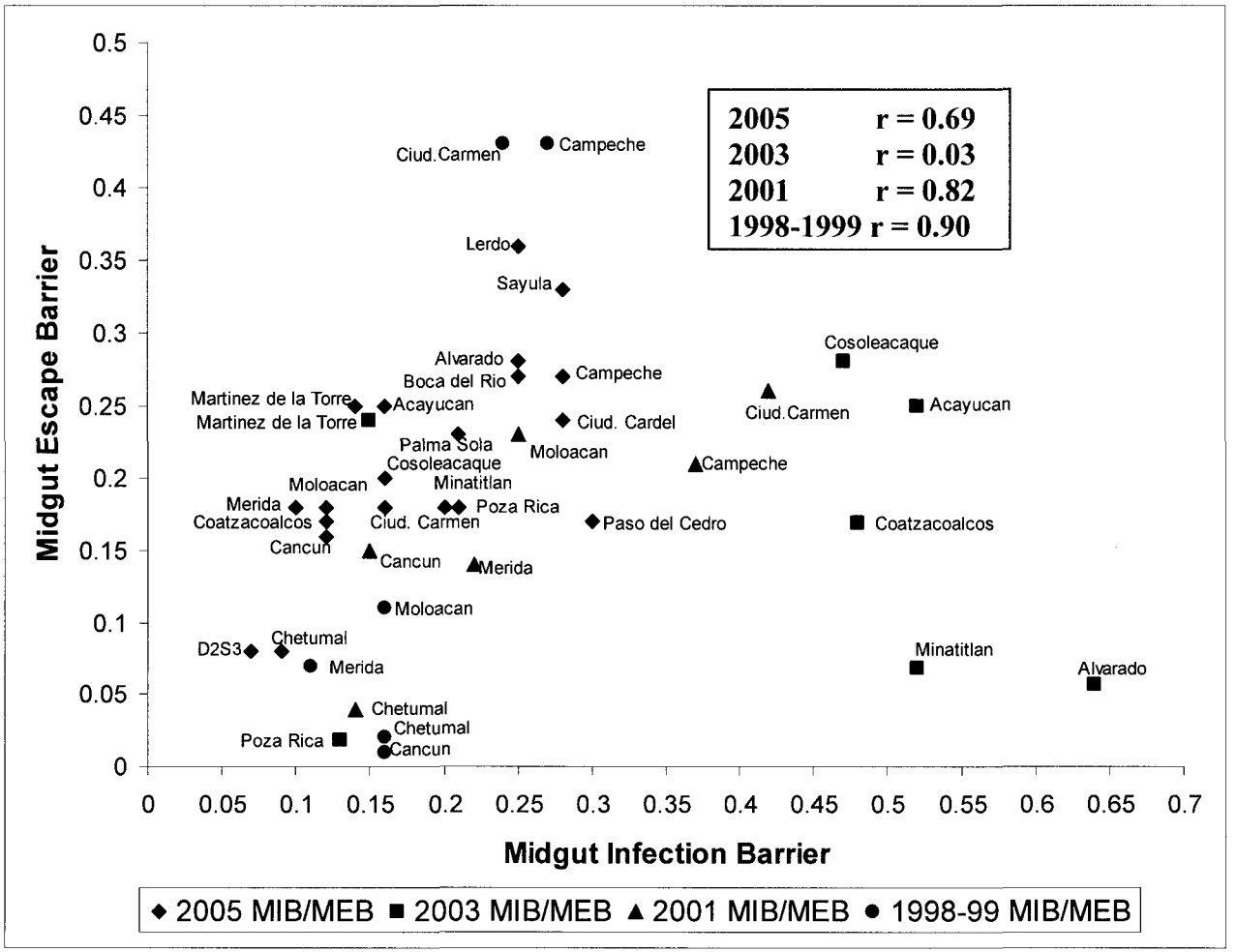


Figure 2.9. Regression analysis of MIB vs MEB. MIB versus MEB is plotted for populations from each site. Regression analysis was performed for all sites within the 2005 collection. Data are included for past collections (1998-2003) at sites that corresponded to the 2005 geographic collection sites.



Chapter 3

Abundant nuclear copies of mitochondrial origin (NUMTs) in the *Aedes aegypti* genome

This chapter has been accepted for publication in Insect Molecular Biology

William C. Black IV and Scott A. Bernhardt (2009) Abundant nuclear copies of mitochondrial origin (NUMTs) in the *Aedes aegypti* Genome. Insect Mol Biol. **In Press**.

Abstract

A portion of the *Aedes aegypti* mitochondrial NADH dehydrogenase subunit 4 gene (ND4) was amplified using the Polymerase Chain Reaction at a 42°C annealing temperature. Amplified fragments were similar to ND4 but contained heterozygous sites. We suspected that nuclear copies of mitochondrial origin (NUMTs) exist in the *Ae. aegypti* genome. A BlastN search in VectorBase with the entire *Ae. aegypti* mitochondrial genome identified 233 NUMTs comprising 110,178 bp in 145 supercontigs. At a density of 0.080 bp/kb, this represents the second highest density of NUMTs in an insect genome and the highest in Diptera. Analyses of flanking sequences suggested that *Ae. aegypti* NUMTs arise through mtDNA breakage and nonhomologous recombination rather than through duplicative processes such as transposition or molecular drive.

Introduction

The presence of mitochondrial DNA sequences (mtDNA) in the nuclear genome has been documented in more than 100 eukaryotic species (Bensasson *et al.,* 2001; Lopez *et al.,* 1994; Ricchetti *et al.,* 1999; Richly and Leister, 2004; Woischnik and Moraes, 2002; Zhang and Hewitt, 1996). These sequences probably originated from the invasion of nuclear DNA by mtDNA and were given the acronym "NUMTs" (nuclear mtDNA) (Lopez *et al.,* 1994). Several reviews of NUMTs have been written and all agree that NUMTs probably arose from nonhomologous recombination of nuclear DNA with mtDNA fragments that leaked from damaged mitochondria (Henze and Martin, 2001; Mourier *et al.,* 2001; Woischnik and Moraes, 2002)

NUMTs vary in their similarity to the mtDNA genome of a species. Highly similar sequences probably represent recent introgressions while degenerate NUMTs reflect ancient introgression, possibly older than the species (e.g. Hazkani-Covo and Graur, 2007; Kolokotronis *et al.,* 2007). NUMTs differ widely in size and often become rearranged, fragmented and tandemly duplicated, but otherwise tend to be uniformly distributed throughout a nuclear genome. Phylogenetic analyses of NUMTs suggest that mtDNA fragments tend to introgress continuously and are either gradually lost or accumulated in a nuclear genome (Bensasson *et al.,* 2003; Chung and Steiper, 2008; Lascaro *et al.,* 2008; Lough *et al.,* 2008; Mourier *et al.*, 2001; Qu *et al.,* 2008; Triant and DeWoody, 2008; Woischnik and Moraes, 2002).

NUMTs have been well documented in insects including aphids (Sunnucks and Hales, 1996), grasshoppers (Bensasson *et al.,* 2000) and ants (Martins *et al.,* 2007). Most recently, insect genome sequencing projects have enabled extensive searches for NUMTs. In general these studies have found very few or no copies of mitochondrial genes (Pamilo *et al.,* 2007; Richly and Leister, 2004). The *Anopheles gambiae* genome apparently has no NUMTs (Richly and Leister, 2004). The *Drosophila melanogaster* genome has only eight NUMTs at 5 nuclear loci and the total amount of NUMTs is 777bp (Bensasson *et al.,* 2001). The density of NUMTs in *D. melanogaster* is therefore 0.005 bp/kb of nuclear DNA. The flour beetle (*Tribolium castaneum*) has 91 NUMTs in 57 nuclear loci, a total NUMT content of 8,821 bp and a NUMT density of 0.056 bp/kb (Pamilo *et al.,* 2007). A fascinating exception is the genome of *Apis mellifera* (Pamilo *et al.,* 2007; Pereira and Baker, 2004; Richly and Leister, 2004). Although statistics vary among these three studies, they all show that the honey bee genome has >2,000 NUMTs

at > 800 loci with as much as 275,000 bp of NUMTs and a density of >1.0 bp NUMTs/kb. This is the highest density in any animal studied, almost ten times higher than in humans and comparable to the densities in some plant genomes. Pamilo *et al.* (2007) suggested that one possible explanation for this high density is that two-thirds of *A. mellifera* NUMTs are flanked by "mariner" type transposons that may have caused secondary transpositions.

Herein we report on the serendipitous discovery of NUMTs in the *Ae. aegypti* genome. During routine amplification of the mitochondrial NADH dehydrogenase subunit 4 gene (ND4) using the Polymerase Chain Reaction (PCR), we inadvertently set the thermal cycler to use an annealing temperature of 42°C. Products were of the predicted size but subsequent single strand conformation polymorphism (SSCP) analyses suggested that every mosquito had a unique SSCP pattern. Subsequent sequencing revealed an abundance of heterozygous sites in single mosquitoes. We suspected the existence of NUMTs because this amount of diversity was inconsistent with heteroplasmy (Hale and Singh, 1986; Paduan and Ribolla, 2008; Townsend and Rand, 2004) and far greater than the amount of diversity observed in six prior studies in our laboratory that cumulatively examined the ND4 in 5,840 mosquitoes worldwide.

The publication and release of the *Ae. aegypti* genome (Nene *et al.*, 2007) allowed us to confirm this suspicion. Herein, we report the results of a BlastN search of the 4,758 supercontigs in the assembled *Ae. aegypti* genome with the entire 16,655 bp of the *Ae. aegypti* mtDNA. We demonstrate that *Ae. aegypti* has the second highest density of NUMTs found in an insect genome and the highest in Diptera and that, unlike *A. mellifera*, most of these NUMTs appear to have arisen independently.

Material and methods

DNA isolation and PCR amplification

We isolated DNA from mosquitoes using a salt extraction method (Black and DuTeau, 1997). The final DNA pellet was resuspended in 200 µl sterile TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA). A 50-µl aliquot of DNA was stored at 4°C for daily use in PCR. The remainder was stored in plastic screw-top vials at -80°C. The primers used in Gorrochetegui-Escalante *et al.* (2000) were degenerate ND4+ (GTDYAT TTATGATTRCCTAA) and ND4- (CTTCGDCTTCCWADWCGTTC) because at the time we only had mitochondrial sequences from *Anopheles gambiae* and *An. albimanus*. Underlined nucleotides in the two primers were subsequently changed to produce the more precise primers listed below. The thermal cycler heated reaction tubes to 95°C for 5 min and then cooled them to 80°C prior to the addition of 1 unit of Taq DNA polymerase. The subsequent program consisted of 10 cycles of 1 min at 92°C (denaturation 1), 1 min at 48°C (annealing 1), and 2 min at 72°C (extension 1). This was followed by 32 cycles of 1 min at 92°C (denaturation 2), 35 sec at 52°C (annealing 2), and 2 min at 72°C (extension 2). A final extension reaction was carried out for 7 min at 72°C and the samples were then cooled to 4°C.

Interestingly the primers used to amplify the sequences in Figure 1b were designed from the *Ae. aegypti* mitochondrial genome (EU352212) and were therefore more precise than those used by Gorrochetegui-Escalante et al. (2000). They were ND4sb+ (TTATGATTGCCAAAGGCTCAT), and ND4sb- (CTTCGTCTTCCTA TTCGTTC). Underlined nucleotides are those that were changed from Gorrochetegui-Escalante et al. (2000). The thermal cycling conditions used to obtain the sequences in

Figure 1b were the same except that denaturation 2 was 42°C rather than 52°C. The ND4 PCR products were purified using a QIAquick spin column (QIAGEN, Chatsworth, CA) and sequenced along both strands on the ABI sequencer at Macromolecular Resources at Colorado State University Proteomics and Metabolomics Facility.

NUMT discovery

The *Ae. aegypti* mitochondrial genome (EU352212) was downloaded and subjected to a BlastN search in VectorBase (http://aaegypti.vectorbase.org/index.php) against the 4,758 supercontigs in the *Ae. aegypti* whole genome assembly. The cutoff was set to an Expected value (E-value) $< 10^{-4}$ and word size was varied among 11, 15, 30 and 60.

The E-value is a parameter that describes the number of alignments expected by chance when searching a database of a particular size. An E-value of 1 calculated for an alignment suggests that when searching a database of a given size, one match with a similar score is expected by chance. The lower the E-value the more unlikely that the alignment arose at random. The E-value is frequently treated as a significance threshold for reporting results. Following Pamilo *et al.* (2007) and Richly and Leister (2004) we adopted a significance threshold of $E < 10^{-4}$ when searching for NUMTs.

Word size is the size of fragments from the query sequence that BlastN generates to test for alignments to database sequences. The penalty for using large word sizes is that BlastN only finds large uninterrupted alignment blocks between the query and subject sequences. Sensitivity for NUMTs bearing substitutions, insertions or deletions is lowered. The penalty for using smaller word sizes is that BlastN fragments an alignment of a single NUMT into many smaller alignments when individual words have $E > 10^{-4}$

due to substitutions or small insertions/deletions. This was corrected using bl2seq (http://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch and PROG_DEF=BlastNandBLAST_PROG_DEF=megaBlastandSHOW_DEFAULTS=onandBLAST_SPEC=blast2seqandLINK_LOC=align2seq) to align EU352212 with the sequence of the supercontig using a word size of 32.

The beginning and ending nucleotide positions of the mtDNA nucleotide match was recorded for all BlastN alignments with an expected value of $E < 10^{-4}$ as were the E-values. The beginning and ending of the corresponding match in each supercontig were also recorded. The gene identity of the beginning and end of a NUMT was determined using a lookup table created in EXCEL that contained the precise 5' and 3' nucleotide positions for each mtDNA gene. Each bl2seq alignment was also inspected for insertions, deletions and duplications and recorded, and a summary of the NUMT was recorded. The resulting database is included as Supplement 2.

Analysis of flanking genomic sequences

A FORTRAN program FLANKS (written by WCB4) identified 1 kb of DNA 5' and 3' to the insertion site in the supercontig sequences and wrote these to separate files. These 1-kb flanking sequences were then used in a BlastN search against the nr NCBI database to determine the identity and number of alignments with $E < 10^{-4}$. The accession number and gene name of the best alignment for each flanking sequence were recorded. The nucleotide content of the flanking sequences was also recorded. The gene names for the best alignment 5' and 3' of the NUMT were then concatenated and these names were sorted to identify NUMTs with the same or similar flanking sequences. The flanking sequence database is included as Supplement 3.

Results

Discovery of NUMTs

The primers and PCR conditions used to amplify the ND4 gene are published (Gorrochotegui-Escalante *et al.,* 2000). Figure 3.1a shows the variable sites among the 40 unique haplotypes in the 387 bp amplicon found in six separate gene flow studies of *Ae. aegypti*. The first study took place in northeast Mexico and found seven haplotypes in 574 mosquitoes (Gorrochotegui-Escalante *et al.,* 2000). Gorrochotegui-Escalante *et al.* (2002) examined gene flow throughout the remainder of Mexico and found 18 additional haplotypes in 1,977 mosquitoes. Bosio *et al.* (2005) performed a gene flow study in Thailand and found seven additional haplotypes in 1,346 mosquitoes. A study in Venezuela found three additional haplotypes in 619 mosquitoes (Urdaneta *et al.,* 2008), and in an unpublished study in Senegal we detected five additional haplotypes in 250 mosquitoes (Sylla et al., unpublished). Lozano-Fuentes *et al.* (submitted) analyzed the ND4 in 1,074 mosquitoes from the state of Veracruz, Mexico in 2003-2004, and no new haplotypes were found. In total we detected the 40 novel haplotypes in Figure 1a in 5,840 mosquitoes, a rate of one novel haplotype every 146 mosquitoes.

In a subsequent study, we used less degenerate primers and inadvertently changed the annealing temperature to 42°C. In doing this we detected 42 new and unique ND4 haplotypes in 986 mosquitoes in Veracruz, Mexico (Fig. 1b) finding a novel haplotype in every 23 mosquitoes. Of these 42 haplotypes, 32 contained one or more heterozygous sites. There were a total of 103 heterozygous sites consisting of 67 Y (C/T), 29 (A/G) and seven (A/T) genotypes. Furthermore, the two heterozygous sites at position 17 (Fig. 1a) in the first 40 haplotypes were not found in the genotypes amplified at 42°C.

BlastN search of the *Ae. aegypti* genome for NUMTs

The complete *Ae. aegypti* mitochondrial genome (EU352212 = 16,655 bp) was downloaded and then subject to a BlastN search in VectorBase against the 4,758 supercontigs in the *Ae. aegypti* whole genome assembly. The significance threshold was set to $E < 10^{-4}$ and word size was varied among 11, 15, 30 and 60. At a word size of 60, 30, 15 and 11, BlastN found, respectively, 143 alignments on 30 supercontigs, 286 alignments on 87 supercontigs, 409 alignments on 148 supercontigs and 417 alignments on 151 supercontigs.

We chose to work with a word size of 15 to increase the likelihood of finding most of the supercontigs containing NUMTs (results in Supplement 1). The penalty for using smaller word sizes was that BlastN fragmented an alignment of a single NUMT into several smaller alignments because individual words had an $E > 10^{-4}$ due to substitutions or small insertions/deletions. We used the program bl2seq at the National Center for Bioinformatics NCBI website (http://www.ncbi.nlm.nih.gov/) to align EU352212 with the sequence of the supercontig using a word size of 32. In every case this reduced the number of NUMTs. For example the number of NUMTs in supercontig 1.600 decreased from 20 to one and 36 NUMTs on supercontig 1.363 decreased to two. Once this was completed there were 233 NUMTs comprising 110,178 bp on 145 supercontigs (Supplement 2).

Table 3.1 characterizes the NUMTs in these 145 supercontigs. Most (125) supercontigs had a single NUMT, 19 had two to three NUMTs and supercontig 1.98 had six identical tandem NUMT duplications. Most (117) of the NUMTs were uninterrupted by deletions or insertions. Table 2 indicates the size distribution of NUMTs in the *Ae.*

aegypti genome. This is a strongly leptokurtotic distribution with 76% of NUMTs <300 bp in size, 63% of NUMTs <200 bp and 46% of NUMTs < 100 bp. The distribution of $-\log_{10}(E)$ as a function of $\log_{10}$(NUMT size (bp)) is displayed in Fig. 3.2. All zero E-values were converted to 200. There was a significant ($P < 0.0001$) and expected positive correlation between $-\log_{10}(E)$ and $\log_{10}$(NUMT size) ($R^2 = 0.77$) because longer alignments will have lower values of E.

The mtDNA nucleotide positions were recorded at the beginning and end of each NUMT. The frequency distribution of these 466 NUMT ends was then plotted as a histogram with 100 bp intervals (Fig. 3.3a) alongside a linear map of the mtDNA along the abscissa. Assuming a homogeneous frequency of NUMT ends along the mtDNA genome, the expected frequency of ends per 100 bp interval was 466 NUMT ends/168 100-bp interval = 2.77 (dotted line in Fig. 3a). A heterogeneity χ^2 test was performed to test the hypothesis that NUMT ends were homogenously distributed along the mtDNA molecule. This hypothesis was rejected ($P = 5.06 \times 10^{-26}$).

The distribution of ends was independently tested by treating each mtDNA gene as a target in which the likelihood of a NUMT end in a gene was assumed to be proportional to the relative size of that gene. A goodness-of-fit χ^2 test was performed to test the hypothesis that NUMT ends were distributed according to gene size (Fig. 3.3b). This hypothesis was also rejected ($P = 1.34 \times 10^{-13}$). Observed and expected numbers of breakpoints are listed in parentheses for those genes with the largest deviations from expected. More NUMT ends occurred in the COX1, Cytochrome B (CytB) and 16S genes than expected based upon the sizes of these genes. Fewer than expected NUMT

ends occurred in the mitochondrial NADH subunit 5 dehydrogenase (ND5) and ND4 genes, and in the D-loop region.

BlastN analysis of sequences flanking NUMTs

For each NUMT we extracted and analyzed 1kb of sequence 5' and 3' to the NUMT to test for patterns in the genomic sites in which NUMT insertions occurred. Specifically we wished to determine if NUMTs were flanked by identical sequences suggesting replication through transposition or tandem duplications. For each of the 466 (2 x 233 Blast hits), BLASTCl3 was used to search the non-redundant (nr) NCBI database to determine the identity and number of alignments with $E < 10^{-4}$.

To avoid detecting the same NUMT more than once, the top alignments from the 5' and 3' flanking sequences of all 233 NUMTs were compared. There were 15 cases in which identical flanking sequences were found. However in each case, with the exception of supercontig 1.98 (six tandem duplications), the NUMT sequences did not match. Thus, the same NUMT was never detected more than once.

A flanking sequence could have an alignment with an E-value greater than the 10^{-4} threshold for two reasons. Either a NUMT might occur at the 5' or 3' end of the supercontig or the flank might fail to align with any of the *Ae. aegypti* genome entries in the nr NCBI database even though the NUMT was detected in VectorBase. The first situation occurred with six alignments and the second occurred in 69 of the 466 alignments. In six supercontigs (1.357, 1.556, 1.440, 1.314, 1.359 and 1.96) both of the 5' or 3' flank sequences failed to align with anything in the nr NCBI database. Thirty-nine sequences aligned with all or part of the original *Ae. aegypti* mtDNA (EU352212),

and these were also excluded to avoid analyzing the same flanking sequence in a supercontig more than once.

With the remaining 352 (=466-69-6-39) flanking sequences, if the NUMT inserted into a unique region of the genome then it aligned once in the BLASTCl3 search, but if the NUMT inserted into a repetitive genome region then it aligned more than once. The 352 flanking sequences aligned 7,477 times to *Ae. aegypti* sequences in the nr NCBI database at a cutoff of $E < 10^{-4}$. The number of times the same sequences occurred in these 7,477 alignments was determined, and the frequency distribution is shown in Figure 3.4. There were 32 flanking sequences that aligned only once, 18 that aligned twice, 16 sequences that occurred thrice, etc.

This distribution was then compared to a Poisson distribution to determine if the number of alignments/flank was a randomly distributed variable. The average alignment density was 21.24 (7,477/352) alignments/flanking sequence, and the predicted Poisson distribution is shown as a dotted line in Figure 3.4. The numbers of alignments/flank are not randomly distributed (χ^2 goodness of fit : $P = 1.74 \times 10^{-24}$). An excess of flanking sequences that only aligned with one or a few *Ae. aegypti* nr sequences were noted. In addition, a large number of flanking sequences aligned with *Ae. aegypti* repetitive elements. Forty-six flanking sequences occurred more than 50 times in the nr NCBI database.

Discussion

At a density of 0.080 bp/kb (110,178 bp NUMTs/1,376,000 kb haploid genome) *Ae. aegypti* has the second highest density of NUMTs found to date in any insect genome and the highest in Diptera. Assuming that NUMTs arise by rare nonhomologous

recombination of nuclear DNA with mtDNA fragments that have leaked from damaged mitochondria (Henze and Martin, 2001; Mourier *et al.,* 2001; Woischnik and Moraes, 2002), we initially hypothesized that this large number of NUMTS arose through duplicative processes in the nuclear genome such as transposition or tandem duplication through unequal crossing over or gene conversion. Instead, with the exception of supercontig 1.98 which had six identical tandem NUMT duplications, the remaining 232 NUMTs were unique and never surrounded by the same elements. This suggests that transposition or tandem duplication are not common mechanisms for spreading NUMTs in the *Ae. aegypti* genome. Apparently mtDNA fragments frequently undergo independent nonhomologous recombination with the nuclear *Ae. aegypti* genome.

Analysis of the 5' and 3' ends of NUMTs demonstrated a biased distribution in the composition of NUMT ends. More than the expected number of NUMT ends occurred in the COX1, Cytochrome B (CytB) and 16S genes and fewer ends than expected occurred in the ND5 gene, in the ND4 gene and in the D-loop region. This can be interpreted in two ways. The CO1 and 16S regions may have been more prone to breakage and the ND5, ND4 and D-loop less subject to fracture. Alternatively, it is possible that nonhomologous recombination occurred most successfully when mtDNA fragments ended in the CO1 or 16S genes or was less likely when mtDNA fragments ended in ND5, ND4 and D-loop sequences.

We considered the possibility that insertion sites were more AT-rich and, therefore, more similar to the high AT-content of the *Ae. aegypti* mtDNA (79% AT) than the remainder of the genome (61.8% AT). However, flanking sequences had an AT content of 63.1%, which was significantly, albeit slightly, greater than the overall genome

rate. No obvious relationship was seen when examining flanking regions more proximate (10-100 bp) to the insert (analyses not shown). Conversely, we tested for a relationship between AT content and the frequency of NUMT ends using the 100 bp intervals in Figure 3.3a. The regression model (NUMT ends = -7.01(AT content) + 8.3, $R^2 = 0.056$, $P = 0.0021$) was significant, suggesting that NUMT ends tended to have a lower AT content. Note, however, that this model explains very little of the overall variance.

Our primary motivation for identifying and characterizing the NUMTs in *Ae. aegypti* was concern that our four previous published studies (Gorrochetegui-Escalante et al., 2000; 2002; Bosio *et al.,* 2004; Urdaneta *et al.,* 2007) and two unpublished studies (Sylla *et al.,* unpublished; Lozano-Fuentes *et al.,* submitted) might have erroneously amplified and identified variation at NUMTs rather than detecting variation in a valid mtDNA marker. The many NUMTs that we have described here are paralogous and would have yielded uninterpretable data or severely misleading results. Furthermore, our population genetic analyses would have been invalid because they were based upon an assumption of haploidy.

Heteroplasmy is the presence of a mixture of more than one type of an mtDNA genome within a cell or individual. The rate of heteroplasmy in *Drosophila melanogaster* is between 0.015 (3 flies/194 sampled) (Townsend and Rand, 2004) and 0.17 (24/144) (Hale and Singh, 1986). In a study of heteroplasmy in *Ae. aegypti*, Paduan and Ribolla (2008) detected a rate of 0.024 (3/125 mosquitoes). Out of the 5,840 mosquitoes that we have characterized, only two haplotypes (MexHap13 and, -20, Fig. 1a) contained a heterozygous site and these occurred at a frequency of 3.42×10^{-4} and 1.71×10^{-4}, respectively. Because of their low frequencies, our interpretation of the

heterozygous sites in haplotypes MexHap13 and -20 is that these represent valid heteroplasmy. In contrast, the 0.76 (32/42) rate of heterozygotes in PCR products amplified at a 42°C annealing temperature (Table 1b) far exceeds any of these previous results and is extremely unlikely to represent heteroplasmy. Furthermore, in a current study involving extensive analyses of the mtDNA ND4 (amplified at 52°C annealing temperature) in Mexico, we have failed to re-encounter any of the sequences in Table 1b (Cassidy, unpublished). We are therefore confident that our prior studies did not involve NUMTs.

Table 3.1. Numbers and characteristics of NUMTs in the 145 supercontigs.

Supercontigs that had 1 NUMT	NUMT characteristics	Number
	TOTAL	125
	w/o insertion/deletions	104
	Insertions	10
	Deletions	7
	Duplications	2
	insertion and deletion	1
	deletion and duplication	1
Supercontigs that had >1 NUMT	TOTAL	19
	2 NUMTs	14
	3 NUMTs	5
	w/o insertion/deletions	13
	Deletions	3
	insertion and deletion	3
Supercontigs that had 6 NUMT duplications		1
		145

Table 3.2. Size distribution of the 233 NUMTs in the *Aedes aegypti* genome.

Size range (bp)	Frequency	Average size(bp)
37 - 100	108	72
101 - 200	40	145
201 - 300	28	245
301 - 400	18	339
401 - 500	8	454
501 - 600	3	558
601 - 700	4	620
701 - 800	6	764
801 - 900	1	826
901 – 1,000	2	909
1,001 – 1,100	1	1,076
1,101 – 1,200	1	1,187
1,201 – 1,300	1	1,216
1,501 – 1,600	1	1,600
1,601 – 1,700	1	1,629
1,801 – 1,900	1	1,895
2,401 – 2,500	1	2,404
2,501 – 2,600	2	2,576
3,901 – 4,000	1	3,976
4,901 – 5,000	1	4,917
7,701 – 7,800	1	7,772
8,901 – 9,000	1	8,909
12,801 – 12,900	1	12,878
13,501 – 13,600	1	13,589
Mean size (bp)		475
Modal size(bp)		91
Median size(bp)		119

Figure 3.1. a) The variable sites among the 40 unique haplotypes in the 387 bp ND4 amplicon found in six separate gene flow studies of *Ae. aegypti* and b) the 42 new and unique ND4 haplotypes detected in 986 mosquitoes in Veracruz, Mexico.

```
          Position in the 387 bp amplicon
                        1111111222222222333333333333
          1123334455889912256781344589990023334557888
          7846892824340371932190706581273980353165O2
a)
MexHap1   CTTCCTCGAAGATCGGTTCGACGTTCGATTGTTATCTTCAGA
MexHap2   ..................T.......................
MexHap3   ...T..T......T...CTA.TA..TAT..A...C.......
MexHap4   T..T..T......T...CTA.TA..TAT..A...C.......
MexHap5   ...T..T......T...CTA.T...TAT..A...C.......
MexHap6   ...................A.........CA...........
MexHap7   ...T..T......T...CTA.TA..TA...A...C.......
MexHap8   T..T..T......T...CTA.TA..TAT.AA...C.......
MexHap9   ...T..T...A..T...CTA.TA..TAG..A...C.......
MexHap10  T..T..T......T...CTA.TA..TAT..A...C..A....
MexHap11  T..T..T......T...CTA.TA..TAT..A...C.G.....
MexHap12  ...T..T...A..T...CTA.TA..TAT..A...C.......
MexHap13  Y..T.AT......T...CTA.T...TA...A...C.......
MexHap14  ...T..T...A..T...C.A.T...TAG.CA...C.......
MexHap15  ....A.T......T...CTA.TA..TAT..A...C.......
MexHap16  T.....................................
MexHap17  T.........A...........................
MexHap18  T..................T..................
MexHap19  ...................T........G...........
MexHap20  Y.....................................C.......
MexHap21  ......T.........................C...A.A...
MexHap22  ......A.........................C...A...
MexHap23  ......T.....................................
MexHap24  ..........G.......A.........CA...........
MexHap25  T..................A.......GACA...........
VenHap1   ...T..............A.....T....AC...........
VenHap2   ..............................A...........
VenHap3   ......T...........A.........CA...........
ThaiHap1  ...T..T......T...CTA.T...TAT..A...C.......
ThaiHap2  ...T.........................A...C.......
ThaiHap3  ...T..T......T...CTA.T..CTAT..A...C.......
ThaiHap4  ...T...A.......AC..A.....T....A...........
ThaiHap5  ....................G.........A...........
ThaiHap6  ...T...........AC..A.....T....A...........
ThaiHap7  ...T..T......T...CTA.T.G.TAT..A...C.......
SenHap1   .A.T..T......T...CTA.T...TA...A...C.......
SenHap2   .AAT..T......T...CTA.T...TA...A...C.......
SenHap3   .A.T..T......T...CTA.T...TA...A...........
SenHap4   .A.T..........A....A.....T....AC...........
SenHap5   .A.T..T.....CT...CTA.T...TA...A..G........
```

Figure 3.1. ***Continued.***

b)

```
MexNUMT01 ..................T.....................A.
MexNUMT02 ...T..T......Y...YYR.Y...YR...A...Y.......
MexNUMT03 ........WW........T......Y.........Y......
MexNUMT04 ...Y..TR.....Y...YTR.YR..YRW..A...C.......
MexNUMT05 ......T...........T.......................
MexNUMT06 ...T..T......T...YTA.T...TA...A...Y.......
MexNUMT07 ...T.....................Y....R...........
MexNUMT08 ......T...........Y..Y...Y....R...........
MexNUMT09 ......T......T...YTA.T...TR...A...........
MexNUMT10 ..................T.......A...A...C.......
MexNUMT11 ...T..T......T...CTA.T...TR...A...C.......
MexNUMT12 ......T....R.Y....YA.Y...Y...CA...........
MexNUMT13 ......Y...........T......Y................
MexNUMT14 ......Y...........T......Y....R...........
MexNUMT15 ......Y...........T.......................
MexNUMT16 ..................T.......................C
MexNUMT17 ...T..T......T...CTA.TA..TAT..A...C....T..
MexNUMT18 ...T..T......T...CTA.TR..TA...A...C.......
MexNUMT19 ......T......Y....YA.Y...Y...YA.........TA.
MexNUMT20 ...T..T......T...CTA.T...TAT..A...C....WA.
MexNUMT21 ...T..T...........TA.....T....A.........TA.
MexNUMT22 ......Y...........A.........CA........W..
MexNUMT23 ...T..T......Y....YR.Y...YR...R...........
MexNUMT24 ...Y..T......Y...YYA.Y...YR..CA...Y.......
MexNUMT25 ......T......Y...Y.......Y....A...........
MexNUMT26 ...T..T......T...CTA.T...TAW..A...C.......
MexNUMT27 ......T...........R.........R...........
MexNUMT28 ...T..T......T...CYA.T...TA...A........T..
MexNUMT29 ......T.......................R...........
MexNUMT30 ...T..T......T...CTA.T...TA...A...C.......
MexNUMT31 ......T....R.Y...YYA.........CA...........
MexNUMT32 ......Y...........R.........R...........
MexNUMT33 ..................A.........CA........TA.
MexNUMT34 ...Y..Y......Y....YA.Y...Y...CA.........A.
MexNUMT35 ..........................................A.
MexNUMT36 ...T..T......Y....Y..Y...Y....R.........A.
MexNUMT37 ......T...........TR.....Y....R........T..
MexNUMT38 ..................T.....................T..
MexNUMT39 ...T..T......T...CTA.TA..TAW..A...C.......
MexNUMT40 ......T....R.......A.........CA........T..
MexNUMT41 ..................T..............C.......
MexNUMT42 ......Y...........R.........R...........
```

Figure 3.2. The distribution of -$\log_{10}$(E) as a function of $\log_{10}$(NUMT size (bp)).

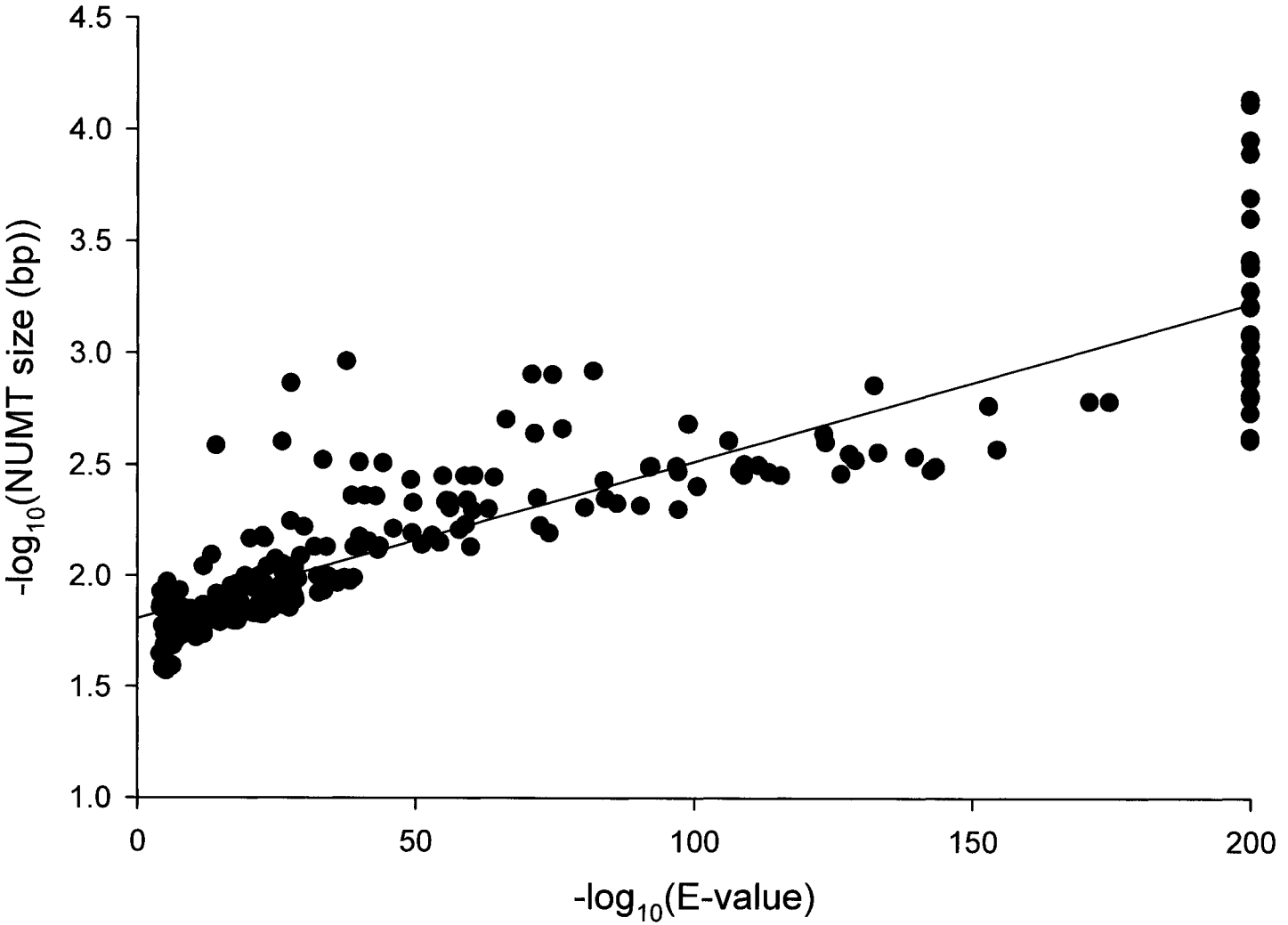
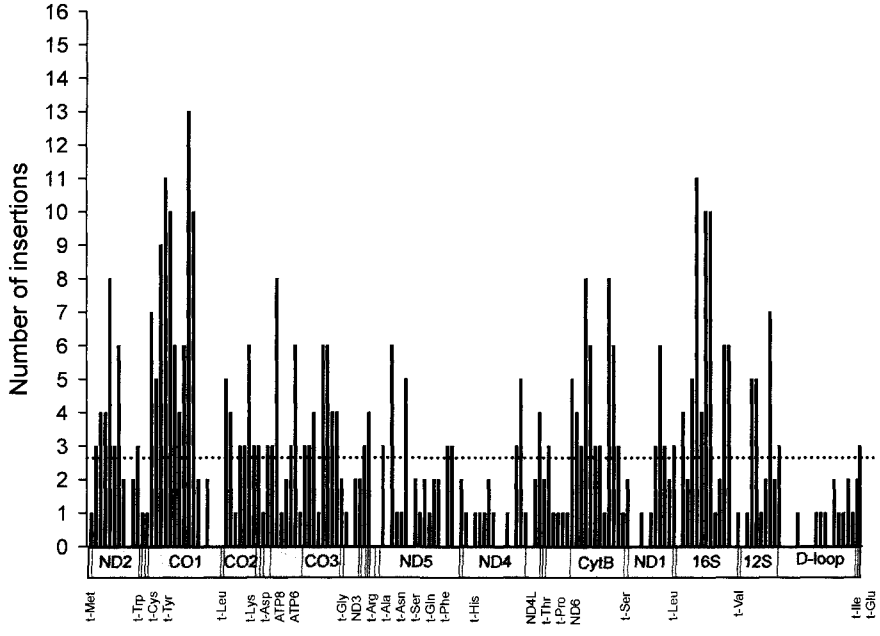


Figure 3.3. a) The observed frequency of breakpoints per 100 bp interval. The expected number was 2.77 (466 ends/168 100-bp interval) and appears as a dotted line. B) each mtDNA gene was treated as a target in which the likelihood of a NUMT end in a gene was assumed to be proportional to the relative size of the gene. Observed and expected numbers of breakpoints are listed in parentheses for those genes with the largest deviations from expected. For example, there were 81 NUMTs that began or ended in COX1 but 41 were expected based upon the size of COX1 relative to other genes in the mtDNA.

a)



b)

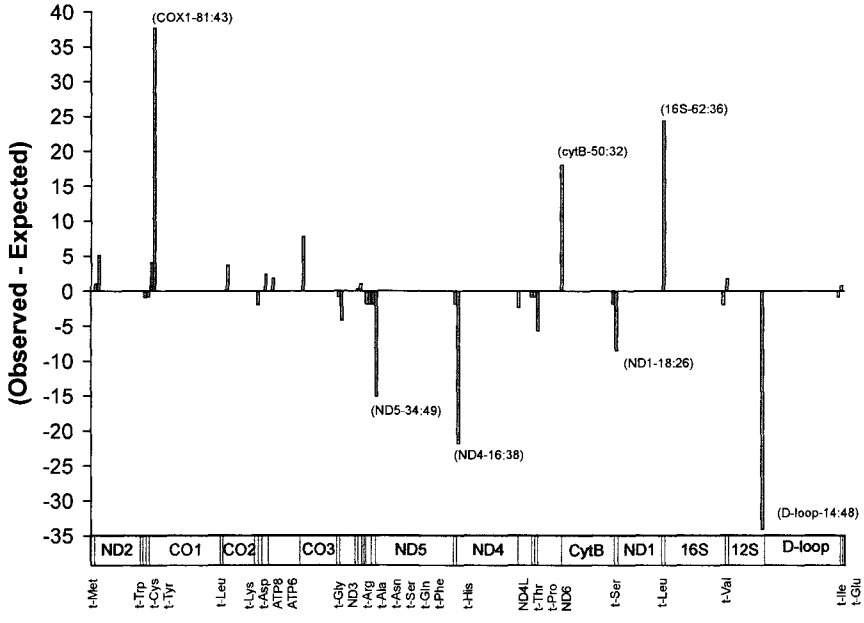
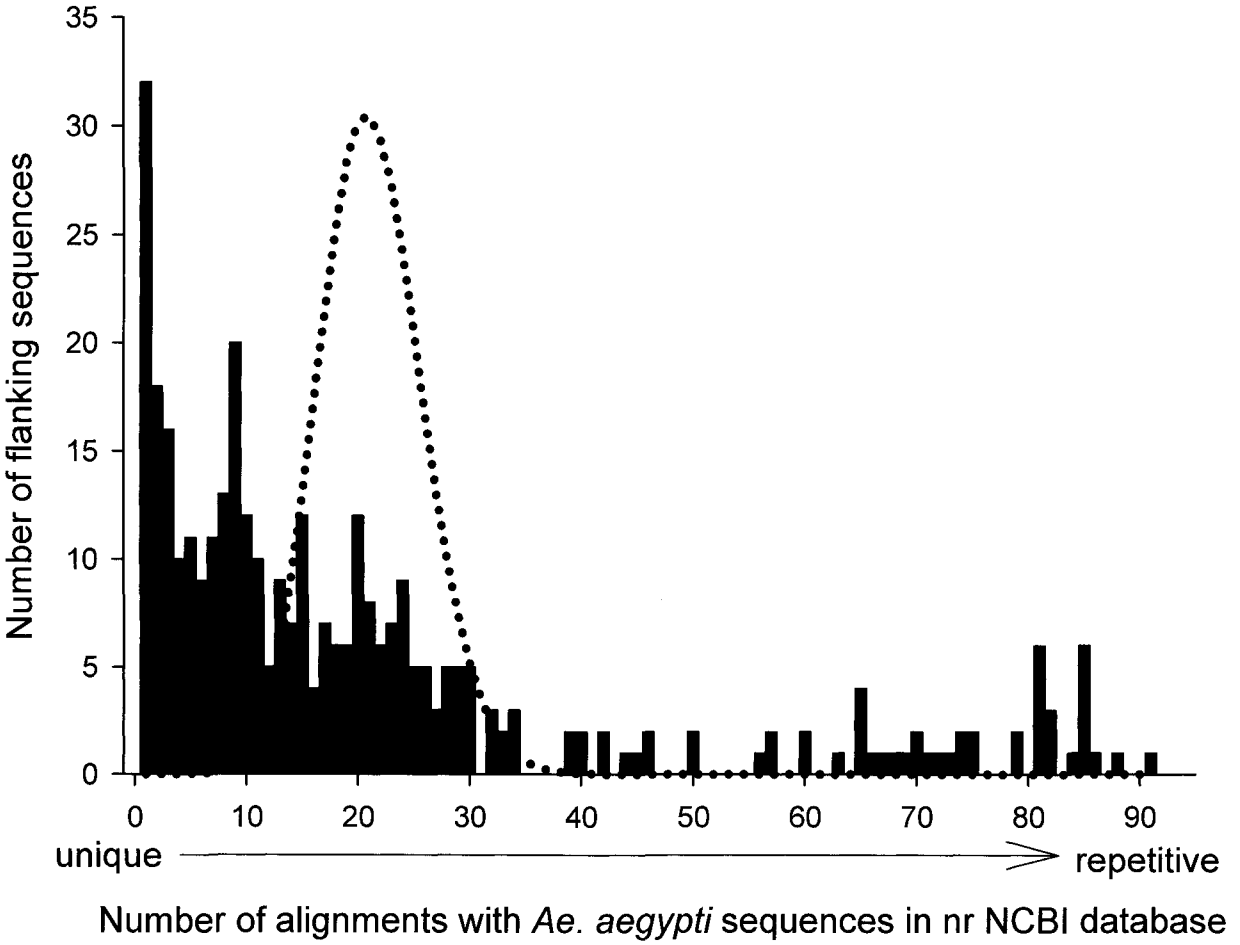


Figure 3.4. The number of times that a NUMT flanking sequence aligned with an *Ae. aegypti* sequence in the NCBI nr database. The predicted Poisson distribution (λ=19.15 alignments/flanking sequence) is plotted as a dotted line to determine whether the number of alignments/flank was a randomly distributed variable.



Chapter 4

Evidence of multiple chromosomal inversions in *Aedes aegypti formosus* from Senegal

This chapter has been accepted for publication in Insect Molecular Biology

Bernhardt SA, Blair CD, Sylla M, Bosio B, Black WC IV (2009) Evidence of multiple chromosomal inversions in *Aedes aegypti formosus* from Senegal. Insect Mol Biol. **In Press**.

Abstract

Chromosomal inversions are prevalent in mosquito species but polytene chromosomes are difficult to prepare and visualize in species in the tribe Aedinii and thus there exists only indirect evidence of inversions. We constructed an F_1 intercross family using a P_1 female from a laboratory strain of *Aedes aegypti aegypti (Aaa)* and a P_1 male *Aedes aegypti formosus (Aaf)* from a strain collected from southeastern Senegal. Recombination rates in the F_2 offspring were severely reduced and genotype ratios suggested a deleterious recessive allele on chromosome 3. The F_2 linkage map was incongruent in most respects with the established map for *Aaa*. Furthermore, no increased recombination was detected in F_5 offspring. Recombination rates and gene order were consistent with the presence in *Aaf* of at least four large inversions on chromosome 1, a single small inversion on chromosome 2 and three inversions on chromosome 3.

Introduction

Many mosquito species exhibit extensive chromosomal polymorphisms (Coluzzi and Kitzmiller, 1975; Kitzmiller, 1976). Paracentric inversions are the most common type of polymorphisms and are usually detected through examination of polytene chromosomes from ovarian nurse cells, larval salivary glands, or Malphigian tubules in heterokaryotypic mosquitoes. Paracentric inversions have been most extensively documented in *Anopheles* species because of the ease and consistency with which their polytene chromosome can be resolved (Atrie *et al.,* 1999; Coluzzi *et al.,* 2002; Cornel and Collins, 2000; Dideco *et al.,* 1978; Green *et al.,* 1992; Michel *et al.,* 2005; Pili and Marchi, 1999; Poopittayasataporn and Baimai, 1995; Ramirez and Dessen, 2000a;b;

Subbarao *et al.,* 1994; 2000; Suguna *et al.,* 1981;1994). Prior to the 1980's, it was even speculated that the chromosomes of anopheline mosquitoes were simply more permissive or tolerant of inversions than the chromosomes of culicines (Rai and Hartberg, 1975). However, by the late 1970's, cytogenetic techniques had improved sufficiently for researchers to document abundant inversions in the Culicine genera *Culex* (Tewfik and Barr, 1974), *Wyeomyia* (Moeur and Istock, 1982), *Orthopodomyia* (Munstermann *et al.,* 1985) and *Sabethes* (Munstermann and Marchi, 1986). Techniques and knowledge have continued to improve such that excellent and consistent polytene preparations of *Culex pipiens* sensu lato are now routinely obtained (Achary, 1994; Campos *et al.,* 2003a; McAbee *et al.,* 2007; Zambetaki *et al.,* 1998).

Paracentric inversions are also probably prevalent in the tribe Aedinii. Polytene chromosomes have been visualized in *Aedes aegypti* (Campos *et al.,* 2003b; Sharma *et al.,* 1978) but band resolution was poor and inconsistent and chromosomes did not cleanly separate away from one another ("spread"). Campos *et al.* (2003b) reported that polytene chromosomes could only be resolved in 4% (58/1,383) of preparations. Thus evidence for inversions in the tribe Aedinii has instead been indirect. Macdonald and Sheppard (1965) examined recombination rates among the *Brugia malayi* susceptibility (f^m), SEX (*M*), and red eye (*re*) loci on chromosome I (see Figure 4.1). They estimated recombination rates in crosses among the Liverpool filaria refractory, Liverpool filaria susceptible, red eye refractory and red eye susceptible strains. They found significantly different recombination rates in male mosquitoes between *re* and *M* (2.1 vs. 5.7 cM) and in females between *re* and f^m (33.7 vs. 15.9 cM). They postulated at least three arrangements, two on the X-chromosome, and one on the Y-chromosome (the

chromosome carrying the dominant M "male" allele). Kuzoe *et al.* (1966) made squash preparations of the testes from *Ae. aegypti* pupae and reported chromosome fragments in 5 of 278 preparations. They interpreted these fragments as the acentric products of crossing over within inversions. Severson and Kassner (1995) used a graphical genotyping technique to show recombination suppression suggestive of inversions on chromosomes 1 and 3.

We generated an F_1 intercross family between the *D2S3* laboratory strain of *Aedes aegypti aegypti (Aaa)* that had been selected for susceptibility to Dengue Serotype 2 Virus (DENV-2) (Bennett *et al.* 2005a) and an *Aedes aegypti formosus* (*Aaf*) strain called PK10 that was collected from the Forest Gallery region of southeastern Senegal. All PK10 mosquitoes exhibited solid black scaling on the first abdominal tergite, the character considered definitive for subspecies *formosus* (McClelland, 1974). Recombination rates in the F_2 offspring were severely reduced and the linkage map was incongruent in most respects to the established *Aaa* linkage map. Furthermore, no increase in recombination was detected in F_5 offspring. We report here evidence for multiple inversions on all three chromosomes of *Aaf* from southeastern Senegal.

Material and methods

F1 intercross family

Twenty P_1 crosses were made with a virgin *Aaa* female from the *D2S3* strain (Bennett *et al.*, 2005a) and a virgin male from the *Aaf* PK10 strain (Sylla *et al.*, 2008) in individual 1-pint ice cream cartons covered and sealed with fine organdy mesh. Twenty reciprocal crosses were also made. Adults were fed raisins and tap water. An adult mouse was offered to the mosquitoes twice a week as a source of blood for egg

development. F_1 eggs from the largest of these 40 crosses were hatched and larvae were reared in plastic shoeboxes and fed a standardized liver powder suspension. Mosquitoes were maintained in an insectary at 26°C and 80% relative humidity. The light cycle was 14 hr light : 10 hr dark.

Upon emergence, twenty pairs of virgin F_1 brothers and sisters were aspirated into individual 1-pint ice cream cartons, fed raisins and tap water and bloodfed on mice twice a week. F_2 eggs from the largest F_1 family were hatched and reared to adults. These adults were allowed to mate at random for 4 days and then 120 males were collected and frozen at -80°C. Females were allowed to oviposit for three days and then 107 females were collected and frozen at -80°C.

Genomic DNA extraction

We isolated DNA from the two P_1 parents, the two F_1 parents and the 227 F_2 offspring using a salt extraction method (Black and DuTeau, 1997). The final DNA pellet was resuspended in 200 µl sterile TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA). A 50 µl aliquot of DNA was stored at 4°C for daily use in PCR. The remainder was stored in plastic screw-top vials at -80°C.

PCR amplification

We amplified 30 loci of known linkage location in the *Ae. aegypti* genome in the P_1 and F_1 parents. We used 20 µL reactions using 1 µL of template DNA; all runs include a negative control (all reagents except template). Routine amplification programs consisted of: 2 min at 95°C, 40 cycles of 30 sec at 95°C, 30 sec at T_a, and 40 sec at 72°C. This was followed by an extension of 5 min at 72°C. The optimal primer annealing

temperatures (T_a) had previously been determined on a gradient thermal cycler. PCR products were checked by electrophoresis in 1.5% agarose gels.

Single Strand Conformation Polymorphism (SSCP) analysis

PCR products from the P_1 and F_1 parents were subjected to SSCP analysis to identify those loci whose genotype inheritance patterns were potentially informative. Polyacrylamide gels 38 x 50 cm in size and 0.4 mm in thickness were used as described in Black and DuTeau (1997).

SNP discovery and development

The 20 loci shown in Table 4.1 appeared to have informative genotypes in the P_1 and F_1 parents. PCR products from the P_1 parents were sequenced in these 20 loci to identify informative SNP alleles segregating in the F_1 family. A pair of allele-specific primers and an opposite primer that amplified both alleles were designed from the P_1 sequences. Primers were selected that: 1) did not individually contain hairpins, 2) did not anneal to one another, and 3) generated a PCR product $\leq$ 70 bp. Each of the two allele-specific primers were also modified to contain a 3' nucleotide corresponding to one of the two alleles and an intentional mismatch introduced three bases in from the 3' end that differed from the wild type nucleotide by a transversion (Table 4.2). At this third position the alternate allele-specific primers differed by a transition. In addition the allele-specific primers were manufactured with 5' tails designed to enable discrimination between SNP alleles on the basis of melting temperature. These were (5'-GCGGGCAGGGCGGCGGGGGCGG GGCC [primer 1]-3' or 5'-GCGGGC [primer 2]-3'. These 20 SNP loci were then amplified and genotyped in the 227 F_2 offspring using

the allele-specific primers in Table 4.2 in melting curve PCR as described by Urdaneta-Marquez *et al.* (2008).

Intensive linkage mapping

Linkage maps were constructed using the computer program JoinMap (Stam and Van Ooijen, 1995). A JoinMap locus x offspring matrix was first constructed in the "cross pollinator" format. JMSLA (a program in JoinMap) checked genotype ratios for conformity to expected Mendelian ratios. JMGRP calculated pairwise recombination distances and associated LOD values (equations in Table 4.3 of Maliepaard *et al.,* 1997). LOD scores varied among locus pairs depending upon recombination distances and the numbers of informative genotypes. We ran JMGRP from an LOD = 0.0 - 8.0 at intervals of 0.1. We then spliced out the portion of the output file that had 3 linkage groups to create a *.grp file. JMSPL physically split the original locus x offspring matrix into three separate linkage group files according to the *.grp file. JMREC estimated pairwise recombination distances and the associated LOD scores among all locus pairs in a linkage group. JMMAP estimated the maximum likelihood solution for the linear order of loci as well as their recombination distances using the approach of Lander and Green (1987).

Genome Rearrangements In Man and Mouse (GRIMM) analysis of rearrangements

GRIMM is a computational tool for analyzing rearrangements of gene orders in pairs of unichromosomal and multichromosomal genomes. GRIMM estimates the minimum possible number of rearrangement steps and then determines a possible sequence of rearrangements that use this number of steps. The order of genes in *Aaa* (Figure 4.1) were entered as the *source genome* (1 - 9 $, 10 - 16 $; 17 - 22 $) while

signed gene orders of the same genes in *Aaf* (Figure 4.1) were entered as the *destination genome* (-4 -2 -1 -6 -3 -5 7 8 9 $, 10 -13 -12 -11 14 15 16 $. -18 -19 -21 -17 -20 22 $).

Results

F_2 recombination rates

We bred an F_1 intercross family with the intention of mapping Quantitative Trait Loci (QTL) that condition the large differences in midgut escape barriers between *Aaa* and *Aaf*. We crossed an *Aaa* female from the *D2S3* (DENV-2 Susceptible on 3 chromosomes) strain (Bennett *et al.* 2005a) with an *Aaf* male from the PK10 strain (Sylla *et al.*, 2008). We isolated DNA from the two P_1 parents, the two F_1 parents and the 227 F_2 offspring. We then amplified 30 loci of known linkage location in the *Ae. aegypti* genome in the P_1 and F_1 parents and subjected these to Single Stranded Conformation Polymorphism (SSCP) analysis to identify the loci with genotype inheritance patterns that were potentially informative. The 20 loci shown in Table 4.1 appeared to have informative genotypes in the P_1 and F_1 parents, and PCR products from these were sequenced to identify informative Single Nucleotide Polymorphisms (SNP) alleles that were segregating in the family. These 20 SNP loci were then amplified and genotyped in the 227 F_2 offspring with allele-specific primers (Table 4.2) in melting curve PCR.

Genotypes were then entered in the "cross pollinator" format for analysis by the JOINMAP package (Stam and Van Ooijen, 1995). The program JMSLA in JOINMAP performed χ^2 goodness-of-fit tests on observed versus expected Mendelian genotype numbers (Table 4.3). Mendelian ratios were observed in all markers on chromosomes 1 and 2. However, all loci on chromosome 3 failed to fit Mendelian expectations. In every case, fewer than expected homozygotes for alleles from the *Aaf* father were observed.

The program JMREC in JOINMAP estimated recombination distances between all pairs of markers and the associated LOD values. For pairs of loci in which LOD $\geq$ 3.0, Table 4.4 lists the recombination distance estimates in the F_2 offspring and the associated LOD values. Expected distances in Table 4.4 were obtained by subtracting the published cM locations (Table 4.1) between each pairwise combination of the 20 loci. Minimum and maximum expected cM distances were obtained by calculating the absolute value of the differences among all four cM locations for a pair of loci in Table 4.1 (i.e. |minimum cM locus 1 - minimum cM locus 2|, |minimum cM locus 1 - maximum cM locus 2|, |maximum locus 1 - minimum locus 2|, |maximum locus 1 - maximum locus 2|). The smallest and largest values of these four values represented the minimum and maximum published recombination rates for a pair of loci.

The observed numbers of recombinant and parental dilocus genotypes were compared to the expected numbers using a Fisher's Exact Test (FET). The "LOD (difference)" values in Table 4.4 are –log10 transformed probabilities from the FET. On chromosome 1, fifteen of a potential 36 dilocus comparisons had an LOD $\geq$3.0. In fourteen of these fifteen comparisons, the observed recombination rate was less than the expected rate and this difference was significant in thirteen of the fourteen.

On chromosome 2, three of ten potential dilocus comparisons had an LOD $\geq$3.0. In two of the three comparisons, the observed recombination rate was significantly less than expected. Three FETs were performed to test the three observed cM distances against the minimum expected distances and in two of these the observed distance was significantly less than expected.

On chromosome 3, nine of a potential fifteen dilocus comparisons had an LOD $\geq$ 3.0. In six of the nine comparisons, the observed recombination rate was significantly less than expected. Twelve FETs were performed to compare the six observed numbers with the minimum and maximum expected values. Observed recombination rates were not significantly less than the minimum expected distances but were all significantly less than the maximum expected values.

We considered the possibility that the presumptive deleterious allele(s) on chromosome 3 inherited from the *Aaf* father might reduce the estimates of cM distances and therefore be incorrectly interpreted as evidence of an inversion. Equation 7 in Table 3 of Maliepaard *et al.* (1997) for the '*cxc*' phase was used to perform iterative estimation of *r* (cM distance) under conditions when the fitness of the for dilocus gametes (a-a, a-b, b-a, b-b) were all equal (Excel spreadsheet available from wcb4).

'a-a' and 'b-b' indicate that both alleles at locus 1 and 2 were inherited from the father or mother (i.e. parentals), respectively; while 'a-b' and 'b-a' gametes are recombinant. The equation returned the correct value of r when values of r = 0.01, 0.1, 0.25, 0.45 or 0.5 were used. To model the fitnesses among genotypes on chromosome 3, we divided their observed frequencies by the expected frequencies. For example for the VC locus, the fitnesses of 'a-a', 'a-b', 'b-a', 'b-b' gametes were respectively 1.39 (79/57), 1.30, 1.30, and 0.02. This analysis was done for all locus pairs and in every case the cM distance was substantially increased rather than decreased by the presence of linked deleterious recessive alleles. For example, LF253 and UGALS are 29 cM in the traditional *Aaa* map but when linked to deleterious recessive alleles on chromosome 3, the predicted distance was 37 cM rather than the 1 cM that was observed.

F_5 recombination rates

Reduced recombination was further tested for by estimating recombination rates in the F_5 generation of an advanced intercross line (AIL). The eggs of the F_2 parents, were raised to F_3 adults, and allowed to randomly intermate to produce F_4 eggs. The same process was followed with the F_4 eggs to produce F_5 eggs. A total of 426 F_5 females were then frozen. Males were not examined because we had only anticipated using the F_5 offspring for QTL mapping. DNA was isolated from these females and used to genotype the loci in Table 4.1.

Assuming 1) random mating, 2) no deleterious or lethal recessives, and 3) no regions of cross-over enhancement or suppression then three additional rounds of meiosis in the F_2, F_3, and F_4 parents should have increased the recombination distances among markers by an amount predicted by the recurrence equations:

$$P_{t+1} = (1-r)\, P_t \qquad [1]$$

$$R_{t+1} = R_t + r\, P_t \qquad [2]$$

$$cM = R / (R + P) \qquad [3]$$

where P_t and R_t are the number of parental and recombinant gametes in generation t, respectively, and r is the recombination rate between a pair of loci measured in the F_2. Equation 1 predicts that the number of parental gametes at time t+1 will decline at a rate (1-r) while equation 2 predicts that the number of recombinant gametes at time t+1 will increase by $r\, P_t$. Equation 3 predicts the recombination rate (cM) at time t. For example, loci that are estimated to be 1, 5, or 10 cM apart in the F_2 are predicted respectively to be 4, 19 and 34 cM apart by the F_5.

Table 4.5 lists the cM distances between markers in F_2 (as in Table 4.4) and F_5 offspring along with the LOD values from JOINMAP. Locus pairs shaded in gray indicate where recombination rates remained the same or actually decreased between the F_2 and F_5 generations. On every chromosome F_2 and F_5 recombination rates were not significantly different. Table 4.5 also lists the predicted F_5 cM distances obtained by equations [1-3]. Some cells are empty because when $cM_{F2} = 0$ then $cM_{F5} = 0$ which downwardly biases estimates of recombination. On chromosome 1, seven of eleven comparisons were significant and all indicated that observed F_5 recombination rates were less than the expected rates. On chromosome 2, all four comparisons were significant and three indicated that the observed F_5 recombination rates were less than the expected rates. On chromosome 3, six of nine comparisons were significant and four indicated that the observed F_5 recombination rates were less than the expected rates. Recombination between linked loci did not increase in the F_5 generation.

Estimating inversion breakpoint positions

The precise locations of inversion breakpoints cannot be ascertained from the current dataset. That requires intensive linkage mapping between two homokaryotypic *Aaf* strains. However, to provide a rough approximation of breakpoint locations, a linkage map was constructed from the F_2 recombination rates listed in Table 4.1. JMMAP in JoinMap was used to estimate the maximum likelihood solution for the linear order of loci as well as their recombination distances in *Aaa/Aaf* hybrids. For each chromosome, Figure 4.1 shows on the left the current linkage map of *Aaa* as presented in Severson *et al.* (2002; 2004). On the right side is the maximum likelihood linkage map in *Aaa/Aaf* F_2 hybrids. To facilitate comparison of the two maps, we aligned each pair of

maps on a common locus. The uppermost marker on the hybrid map always had a map position of 0.0 cM in the JMMAP output. On chromosome 1, we therefore added 38 cM to all locus positions to align both maps on the SEX locus. On chromosome 2, we added 7 cM to align on the LF115 locus. On chromosome 3, we added 21 cM to align on the *Malt* locus.

GRIMM (Tesler, 2002) is a software package that estimates the minimum number of steps needed to explain gene arrangements between a pair of genomes and then determines a possible sequence of these rearrangements. The beginnings and ends of the curved arrows on the *Aaa* map in Figures 4.1 and 4.2 represent GRIMM's estimates of the most parsimonious sequence of inversions and their respective breakpoints to explain the gene order and locations in *Aaa/Aaf* hybrids. Loci in dashed boxes were not mapped in this study but are included for later discussion. GRIMM proposed four sequential inversions on chromosome 1. A single inversion on chromosome 2 containing TrypEarl and PGK would have inverted the positions of these two loci and suppressed recombination from 15 cM to 1.1 cM. GRIMM's estimated three sequential inversions on chromosome 3.

Discussion

The original designation of *Ae. aegypti s.l.* subspecies arose from observations made in East Africa in the late 1950's that the frequency of pale "forms" of *Ae. aegypti* was higher in populations in and around human dwellings than in those of the nearby bush. The implied correlation between color and behavior prompted Mattingly (1957) to revisit the biology and taxonomy of *Ae. aegypti s.l.*. He described *formosus* (Walker) as a subspecies of *Ae. aegypti* that was restricted to sub-Saharan Africa and in West Africa

"is the only form known to occur except in coastal districts and in one or two areas of limited island penetration." He also suggested that it most frequently breeds in natural containers such as tree holes and feeds on wild animals. Mattingly also stated that in addition to the dark-scaled parts of the body being generally blacker, "*ssp. formosus* never has any scales on the first abdominal tergite." The type form of *Aaa* was alternatively defined as "either distinctly paler and browner (at least in the female) than *ssp. formosus* or with pale scaling on the first abdominal tergite or both." He also suggested that *Aaa* breeds in artificial containers provided by humans, will breed indoors, and has a preference for feeding on human blood (Mattingly, 1967). McClelland (1974) made a comprehensive study of differences in scale patterns in the abdominal dorsum in 74 *Ae. aegypti* collected from 69 different worldwide locations. He concluded that many of Mattingly's distinctions between the subspecies were not always distinct in Africa, the only region in the world where both forms are found. In East Africa, pure *Aaa* or *Aaf* collections, as defined by both color and behavior, were found but there were also collections where the subspecies were mixed. In areas of sympatry, McClelland (1974) found intermediate forms with peridomestic habits and a wide range of pale scaling. Populations widely overlapped in the extent of pale scaling. He concluded that, with a large enough sample size, populations could be distinguished on the basis of body color, although peridomestic populations overlapped with the distributions of both *Aaa* and *Aaf* populations. But, body color alone was unreliable as a means to assign individuals to a particular subspecies and, instead, he recommended use of a scoring system.

Later, mark-release-recapture studies in Kenya (Trpis and Hausermann, 1975) demonstrated that immature mosquitoes collected from sylvan, peridomestic, or domestic

breeding containers showed an overwhelming preference for their respective habitat as adults. In contrast, in West Africa, mosquitoes morphologically consistent with *Aaf* were found breeding domestically indoors in Nigeria (Nasidi *et al.,* 1989) and Gabon (Vazeille-Falcoz *et al.,* 1999). Therefore, the classic behavioral/habitat descriptions given by Mattingly (1967) for these two subspecies were not valid throughout Africa. In eastern Kenya, genetic crosses between *Aaf* and *Aaa* showed that preferences for endophily had a strong genetic component (Trpis and Hausermann, 1978). Laboratory experiments crossing *Aaa* and *Aaf* from Kenya showed no evidence of assortative mating (Moore, 1979). Furthermore, there was no decrease in fecundity in hybrids, nor any morphological defects.

Subsequently, scaling on the first abdominal tergite was found to be under control of a single genetic locus on chromosome 2 (see Bt in Figure 4.1) (Munstermann and Craig, 1979). Endophily also appeared to be under genetic control. Trpis and Hausermann (1978) collected *Ae. aegypti* either inside houses (D), near a village (P), or from tree holes in a forest (F). These were interbred in an insectary to generate six hybrid strains (DP, PD, DF, FD, PF, and FP). The original and hybrid strains were then labeled with fluorescent dusts and released near houses. Mosquitoes were then recaptured in houses and most (45%) were of the original D type, 14% were DP and PD, 10% were P, and 6% were DF and FD. Very few PF or FP (2%) were collected in houses. Opposite trends were seen in mosquitoes recaptured outdoors in the village. Tabachnick *et al.* (1985) found a strong association between the frequency of specific allozymes (Idh-2 and Pgm-1; Figure 4.1) and susceptibility to Yellow Fever Virus (YFV). In artificial selection experiments, Lorenz *et al.* (1984) found a strong association between the

frequency of Mdh-2 allozymes (Figure 4.1) and YFV susceptibility. If the genes controlling these traits are located in inversions then the traits that they condition would be maintained as correlated characters. Inversions in *Aaf* would explain why, despite continued gene flow between the two forms in East and West Africa, there exists a correlation between lower vector competence, solid black scaling on the first abdominal tergite, isozymes, and exophilic behavior in East Africa.

Several findings in the current study warrant additional comment. Table 4.1 lists 14 published linkage mapping studies in *Aaa*. Many additional QTL mapping studies have also been published (e.g. Beerntsen *et al.,* 1995; Severson *et al.,* 1995). Table 4.1 demonstrates that it is very common in these studies to find variation in the estimated rate of recombination between pairs of markers. This is attributed to localized, strain-specific recombination enhancers (Ercan *et al.,* 2005; Houston *et al.,* 2004; Wei *et al.,* 2002; 2005) and suppressors (Behura *et al.,* 2004; Downs *et al.,* 2003; Labonne *et al.,* 2007; Ma *et al.,* 2004; Yang *et al.,* 2007). Deleterious and lethal recessive genes can also cause inaccuracies in recombination rate estimates by changing the numbers of specific parental or recombinant dilocus genotypes. Our analysis of the presumptive deleterious recessive allele on chromosome 3, showed that cM distances are overestimated by deleterious alleles. Gomez-Machorro *et al.* (2004) found a large excess of heterozygotes in F_2 progeny of an F_1 intercross family, which suggested that balanced lethal and/or deleterious alleles were segregating in the family. Matthews and Craig (1987; 1989) analyzed the progeny of two inbred strains of the treehole mosquito, *Ochlerotatus triseriatus*, at several isozyme loci to determine which factors maintained heterozygosity at these loci after twelve generations of full-sib mating. They showed that heterozygosity

was enforced by lethal alleles at loci to which all the polymorphic isozyme loci were linked and that the lethal alleles formed a balanced lethal system. Similarly, Munstermann (1994) showed that while loss in genetic variability accompanies the colonization process, this loss can be partially masked by the development of balanced polymorphisms. He also concluded with the possibility that balanced lethals maintain heterozygosity in inbred lines.

We don't believe that balanced lethals explain the extensive changes in gene order seen in the present study. Major discrepancies in gene order are relatively rare and are usually due to inaccurate estimates of recombination rates upon which intensive multipoint linkage mapping algorithms are dependent. The discrepancies in recombination rates and gene order that we have detected among the F_2 progeny of this *Aaa* x *Aaf* cross are not minor, and in most cases extended to the F_5. Furthermore, the suppression of recombination did not involve one or a few marker loci; rather on chromosomes 1 and 3 suppression appears to involve many loci simultaneously. GRIMM analysis suggests that, minimally, gene order in *Aaf* compared with *Aaa* are consistent with eight large paracentric inversions; four in chromosome 1, a single small inversion on chromosome 2 and three inversions on chromosome 3.

In examining the range of variation of locus positions in Table 4.1, especially in chromosome 3, we also propose that prior mapping studies may have also unknowingly involved these inversions. Bosio *et al.* (2000) and Fulton *et al.,* (2001) (references 5 and 7 in Table 4.1) used an *Aaf* strain from Nigeria. Table 4.4 shows that the average observed recombination frequency (12.3 cM) was <u>not</u> significantly different than the minimum expected distances (11.0 cM). Note also the large discrepancy in the position

of LF253, originally mapped in the Liverpool strain mosquitoes (Severson *et al.,* 1993). The strain of *Ae. aegypti* used in the original selection experiments was the Liverpool strain from West Africa and has been maintained in the Liverpool School of Tropical Medicine since 1936 (Macdonald, 1962a; b).

In no cases did we find evidence for complete suppression of crossing over. Similar results were reported by (Stump *et al.,* 2007) who estimated recombination rates on chromosome arm 2L of *Anopheles gambiae* in the backcross progeny of 2La/+(a) heterokaryotypes and as a control, from 2L+(a) homokaryotypes. In homokaryotypes, the recombination rate was a uniform 2 cM/megabase (Mb). In heterokaryotypes, recombination within the rearranged region was reduced, but not eliminated, to < 0.5 cM/Mb, with slightly higher but nevertheless reduced levels (< 1.0 cM/Mb) flanking the rearrangement breakpoints. Similar results have been reported in *Drosophila* species (Navarro *et al.,* 1997; Schaeffer and Anderson, 2005). Unless our markers were very close to the inversion breakpoints, we would not expect complete suppression of crossing over. The proximity of the Bt locus and loci associated with flavivirus vector competence or endophily await final assembly of the *Ae. aegypti* genome.

Table 4.1. Map positions (in cM) of markers used in the present study. Maximum and minimum positions are listed when discrepancies existed in the literature. Loci whose position differs by more than 10 cM in the literature are highlighted in gray. References indicate the studies in which cM was estimated (first number) and subsequent studies that assumed the same position.

Chromosome Locus	Minimum position (cM)	References*	Maximum position (cM)	References*
Chromosome 1				
CathB	5.0	(10,14)	7.0	(7)
LF198	18.0	(2,3,6,8,13)	20.0	(7,8,12)
BMIOP	30.0	(12)	32.0	(7)
SEX	**34.0**	**(1)**	**49.0**	**(7)**
AmylI	44.0	(4,8,12)	52.0	(7)
FerH	45.0	(8,10,11,12)		
CHT2	50.0	(8,12)	53.0	(7)
APN	54.0	(8,10,12,14)		
ANTC	69.0	(8,10,12)	79.0	(7)
Chromosome 2				
LF115	0.0	(6)	7.0	(2,8,10,11)
TrypEarl	19.0	(7)	23.0	(12)
PGK	38.0	(8,10,11,12)		
Ef2	60.0	(8,10,11)		
Sin3b	**56.0**	**(7)**	**70.0**	**(10,11,12,14)**
Chromosome 3				
UGALS	**7.0**	**(7)**	**46.0**	**(14)**
LF253	**17.0**	**(2,8,9,10,12)**	**40.0**	**(3)**
apoLP-II	**17.0**	**(7)**	**53.0**	**(5)**
Malt	**21.0**	**(8,11,12)**	**65.0**	**(5)**
VC	**26.0**	**(8,10,11,12)**	**38.0**	**(14)**
PABP	64.0	(8,11)		

*Reference numbers: 1. Munstermann and Craig, 1979; 2. Severson *et al.*, 1993; 3. Severson and Zhang, 1996; 4. Grossman *et al.*, 1997; 5. Bosio *et al.*, 2000; 6. Brown *et al.*, 2001; 7. Fulton *et al.*, 2001; 8. Severson *et al.*, 2002; 9. Chambers *et al.*, 2003; 10. Gomez-Machorro *et al.*, 2004; 11. Severson *et al.*, 2004; 12. Bennett *et al.*, 2005b; 13. Chambers *et al.*, 2007; 14. Saavedra-Rodriguez *et al.*, 2008.

Table 4.2. A) PCR primer sequences, annealing temperature (T_a), and anticipated PCR product size for Sin3b and Malt loci. Genotypes at these loci were determined by SSCP. B) Oligonucleotides used for allele-specific PCR. The sequence of the long tail is 5'-GCGGGCAGGGCGGCGGGGGCGGGGCC-3' while the sequence of the short tail is 5'-GCGGGC-3'. The 3' allele specific nucleotide is bold and the transition mismatches at the third nucleotide from the 3' end appear in lower case.

A)

Locus name	Forward Primer Reverse Primer	PCR product size	T_a
Sin3b	5'-TTCGGCATCCACGCCTACATT-3'	497	55.8
	5'-CGTAGTTGCTGTAGTTTTGG-3'		
Malt	5'-CACATCGTTATTCATTTGCT-3'	488	50.4
	5'-CACTTATCGGACAACCGCTGGAA-3'		

B)

Locus name	Oligonucleotide	PCR product size	SNP position
Chromosome 1			
CathB	5'-ATCCGCAGCATCCTTGA-3'		402
	5'-[long tail]CATTCGACGCTCGTGAAcA**G**-3'	98	
	5'-[short tail]CATTCGACGCTCGTGAAtA**A**-3'	78	
LF198	5'-TCTGTTACCATGAGCGATT-3'		306
	5'-[long tail]CACCAAACACATACGAgA**G**-3'	79	
	5'-[short tail]CACCAAACACATACGAaA**A**-3'	59	
BMIOP	5'-TGTCCCATTGGATCAGGT-3'		116,963
	5'-[long tail]GAGAGAAGCAGGAAAGcG**G**-3'	105	
	5'-[short tail]GAGAGAAGCAGGAAAGtG**A**-3'	85	
Amy2	5'-CGTGACCTCGTTGGTTATCG-3'		888
	5'-[long tail]TCCTTCCGTCGCACTTGcC**G**-3'	86	
	5'-[short tail]TCCTTCCGTCGCACTTGtC**A**-3'	66	
FerH	5'-CGAGTACGCTCTTATGC-3'		8,030
	5'-[long tail]GTCTCWTCACCGATGTcA**G**-3'	111	
	5'-[short tail]GTCTCWTCACCGATGTtA**C**-3'	91	
CHT2	5'-AACTCTGGGTCAACTACGG-3'		8,893
	5'-[long tail]GTACTGGCACTCAAAGTGcA**C**-3'	114	
	5'-[short tail]GTACTGGCACTCAAAGTGtA**T**-3'	94	
APN	5'-GAATGAGTTTTAGAGTGATGTCGT-3'		2,062
	5'-[long tail]GCTGATTGATGACTCGaT**G**-3'	97	
	5'-[short tailGCTGATTGATGACTCGtT**C**-3'	77	
ANTC	5'-ACRAATCAGACGATCATTCAAAC-3'		251
	5'-[long tail]AAGCTAGAAAATTATTTgT**G**-3'	104	
	5'-[short tail]AAGCTAGAAAATTATTTaT**T**-3'	84	
Chromosome 2			
LF115	5'-GCAGAACCAAYCAGCCACCA-3'		285
	5'-[long tail]TCCGACAACGACGGCAcTA-3'	121	
	5'-[short tail]TCCGACAACGACGGCAtTG-3'	101	

Table 4.2. *Continued.*

Trypearl	5'-ACARCGTGTTMATCTCAAAA-3'		499
	5'-[long tail]AAACCGTGGCAGATGGAcC**G**-3'	104	
	5'-[short tail]AAACCGTGGCAGATGGAtC**T**-3'	84	
PGK	5'-CTTCAACCTCAGCACC-3'		356
	5'-[long tail]GTCGTGATGTGACCTTCaT**G**-3'	75	
	5'-[short tail]GTCGTGATGTGACCTTCgT**T**-3'	55	
Ef2	5'-AACAGGTTCTTGCCGTCC-3'		5,781
	5'-[long tail]AGAAGATTAAGGTCACCaT**G**-3'	84	
	5'-[short tail]AGAAGATTAAGGTCACCgT**T**-3'	64	
Chromosome 3			
UGALs	5'-GAGCGGTCATGGTCTTGGA-3'		433
	5'-[long tail]TGGATGCCGAACTACcA**G**-3'	81	
	5'-short tail]TGGATGCCGAACTACtA**A**-3'	61	
LF253	5'-CGGCATCGGTGATCGACAGGG-3'		611
	5'-[long tail]CGGAGTCCACCACTACcA**C**-3'	118	
	5'-[short tail]CGGAGTCCACCACTACtA**T**-3'	98	
apoLP-II	5'-AATCTGGAAACAATCAAGC-3'		1769
	5'-[long tail]CAGTTCGGTGAAGTGGgT**T**-3'	122	
	5'-[short tail]CAGTTCGGTGAAGTGGaT**C**-3'	102	
VC	5'-CATAACATTCTCYCGCTGAG-3'		5,408
	5'-[long tail]GTGGCATAGACAAAACGtT**A**-3'	97	
	5'-[short tail]GTGGCATAGACAAAACGcT**T**-3'	77	
PABP	5'-CAAGAACCACTCACGGC-3'		5,018
	5'-[long tail]ACATCTGCTTTTGTTCAcC**G**-3'	87	
	5'-[short tail]ACATCTGCTTTTGTTCAtC**A**-3'	67	

Table 4.3. Analysis of genotype frequencies among F_2 offspring. Expected numbers of genotypes are listed in parentheses for those loci whose genotypes fail to fit Mendelian expectations.

Locus	aa*	ab*	bb*	a-*	b-*	Missing	χ^2	d.f.	LOD
Chromosome 1									
CathB	64	120	43	0	0	0	4.63	2	1.01
LF198	124	102	0	0	0	1	2.14	1	0.84
BMIOP	57	0	0	0	169	1	0.01	1	0.04
SEX	107	120	0	0	0	0	0.74	1	0.41
AmyII	56	124	47	0	0	0	2.66	1	0.99
FerH	63	111	53	0	0	0	0.99	1	0.50
CHT2	58	121	48	0	0	0	1.87	2	0.41
APN	0	0	49	177	0	1	1.33	1	0.60
ANTC	53	0	0	0	174	0	0.33	1	0.25
Chromosome 2									
LF115	105	122	0	0	0	0	1.27	1	0.59
TrypEarl	113	114	0	0	0	0	0	1	0.00
PGK	118	109	0	0	0	0	0.36	1	0.26
Ef2	113	114	0	0	0	0	0	1	0.00
Sin3b	56	116	54	0	0	1	0.19	2	0.04
Chromosome 3									
UGALS	0	0	72(57)	155(170)	0	0	5.46	1	1.71
LF253	77(57)	0	0	0	150(170)	0	9.63	1	2.72
apoLP-II	87(113)	139(114)	0	0	0	1	11.96	1	3.26
Malt	77(57)	146(113)	4(57)	0	0	0	65.56	2	14.24
VC	79(57)	147(113)	1(57)	0	0	0	73.38	2	15.93
PABP	76(57)	119(113)	32(57)	0	0	0	17.59	2	3.82

*aa = both F_2 alleles inherited from P_1 mother; ab = one F_2 allele inherited from P_1 mother and one from P_1 father; bb = both F_2 alleles inherited from P_1 father; a- = P_1 father shared an allele with P_1 mother, inheritance can only be assessed in bb F_2; b- = P_1 mother shared an allele with P_1 father, inheritance can only be assessed in aa F_2.

† LOD (difference) is $-\log_{10}$(probability) where the probability is calculated from the χ^2 distribution with the indicated χ^2 test value and the degrees of freedom.

Table 4.4. Observed and expected recombination distances (cM) between pairs of marker loci in F_2 offspring of the Aaa x Aaf F_1 intercross. Rows highlighted in gray indicate when the observed recombination distance was less than the minimum or maximum expected distances. Rows in white indicate when the observed recombination distance was ≥ published values.

Locus 1	Locus 2	Observed cM (F_2)[a]	LOD(F_2)[a]	Minimum expected cM[b]	Maximum expected cM[b]	LOD
Chromosome 1						
FerH	CHT2	0	55.4	5.3	8.4	3.69
CathB	LF198	6.5	20.0	11.2	15.0	NS
FerH	APN	0	7.3	9.0		6.23
BMIOP	ANTC	22.9	9.3	37.2	49.0	3.00
CathB	ANTC	31.6	4.2	62.2	74.0	10.21
BMIOP	FerH	0.9	50.6	12.6	14.6	5.93
BMIOP	AmyII	0.9	50.3	12.0	22.0	5.64
CathB	FerH	18.6	15.0	37.6	39.6	4.96
CathB	AmyII	18.6	23.5	37.0	47.0	4.59
APN	ANTC	0	6.5	15.6	25.4	11.19
BMIOP	CHT2	0.9	50.1	17.9	23.0	9.54
CathB	CHT2	19.4	22.6	42.9	48.0	6.93
BMIOP	APN	0	6.5	21.6	23.6	15.66
CathB	APN	18	14.5	46.6	48.6	10.09
BMIOP	SEX	30.5	3	2.0	19.0	15.66
Chromosome 2						
LF115	Ef2	33	6.5	52.5	59.8	4.55
TrypEarl	PGK	1.1	50.3	14.7	18.7	7.40
LF115	TrypEarl	30.4	7.9	11.7	23.0	5.95
Chromosome 3						
apoLP-II	UGALS	4.2	15.1	7.0	46.0	15.66
VC	UGALS	3.8	44.2	8.0	31.0	14.57
apoLP-II	VC	3.8	15.5	8.3	27.3	11.71
LF253	UGALS	3.6	46.8	6.5	32.5	15.57

[a] Observed cM distances and LOD (F_2) are from the JMREC output. LOD is $-\log_{10}$(probability) that the cM estimate arose by random chance.

[b] Expected values were calculated by subtracting the cM distances in Table 4.1 and then using the minimum and maximum differences.

[c] LOD (difference) is $-\log_{10}$(probability) from a Fisher's Exact Test comparing observed vs. expected rates among 227 F_2 offspring. The two LOD values correspond to the results of comparing the minimum and maximum expected values.

Table 4.5. Observed recombination distances (cM) between pairs of marker loci in F_2 and F_5 offspring of the Aaa x Aaf F_1 intercross. Rows highlighted in gray indicate when the recombination distance increased in the F_5.

Locus 1	Locus 2	Observed cM (F_2)[a]	Observed cM (F_5)[a]	LOD (F_5)[a]	Expected cM (F_5)[b]	LOD (Difference)[a]
Chromosome 1						
FerH	CHT2	0.0	0.3	142.2	-	-
CathB	**LF198**	**6.5**	**5.2**	**60.9**	**24.0**	**14.68**
FerH	APN	0.0	8.2	122.2	-	-
BMIOP	**ANTC**	**22.9**	**6.3**	**16.4**	**50.0**	**15.66**
CathB	**ANTC**	**31.6**	**12.1**	**2.2**	**50.0**	**15.66**
BMIOP	FerH	0.9	1.8	89.1	4.0	0.99
BMIOP	AmyII	0.9	3.6	46.0	4.0	0.07
CathB	**FerH**	**18.6**	**12.4**	**1.0**	**50.0**	**15.66**
CathB	**AmyII**	**18.6**	**12**	**1.8**	**50.0**	**15.66**
APN	ANTC	0.0	12.3	30.5	-	-
BMIOP	CHT2	0.9	1.9	86.0	4.0	0.99
CathB	**CHT2**	**19.4**	**12.8**	**0.9**	**50.0**	**15.66**
BMIOP	APN	0.0	7.5	89.5	-	-
CathB	**APN**	**18.0**	**16.4**	**0.3**	**50.0**	**15.66**
BMIOP	SEX	30.5	50	0.0	50.0	0.00
Chromosome 2						
LF115	**Ef2**	**33**	**14.6**	**0**	**50.0**	**15.66**
TrypEarl	PGK	1.1	19.2	8.4	4.0	11.87
TrypEarl	**Sin3b**	**27.3**	**14.2**	**0.2**	**50.0**	**15.66**
LF115	**TrypEarl**	**30.4**	**12.7**	**0**	**50.0**	**15.66**
Chromosome 3						
VC	PABP	22.6	50.0	0.0	50.0	0.00
Malt	**UGALS**	**9.8**	**8.3**	**8.8**	**34.0**	**15.66**
Malt	PABP	24.8	50.0	0.0	50.0	0.00
apoLP-II	UGALS	4.2	16.4	19.1	16.0	0.03
VC	UGALS	3.8	50.0	0.0	14.0	15.66
LF253	**PABP**	**25.3**	**13.6**	**0.1**	**50.0**	**15.66**
apoLP-II	VC	3.8	50.0	0.0	14.0	15.66
Malt	**apoLP-II**	**12.5**	**7.6**	**32.3**	**41.0**	**15.66**
LF253	**UGALS**	**3.6**	**2.3**	**75**	**14.0**	**9.78**

[a] Observed cM distances and LOD (F_5) are from the JMREC output. LOD is $-\log_{10}$ (probability) that the cM estimate arose by random chance.
[b] From equations [1-3]

Figure 4.1. Linkage maps of Aaa as presented in Severson et al. (2002; 2004) (left) and Aaa/ Aaf F_2 hybrids (right). Loci in dashed boxes were not mapped in this study but are included for discussion.

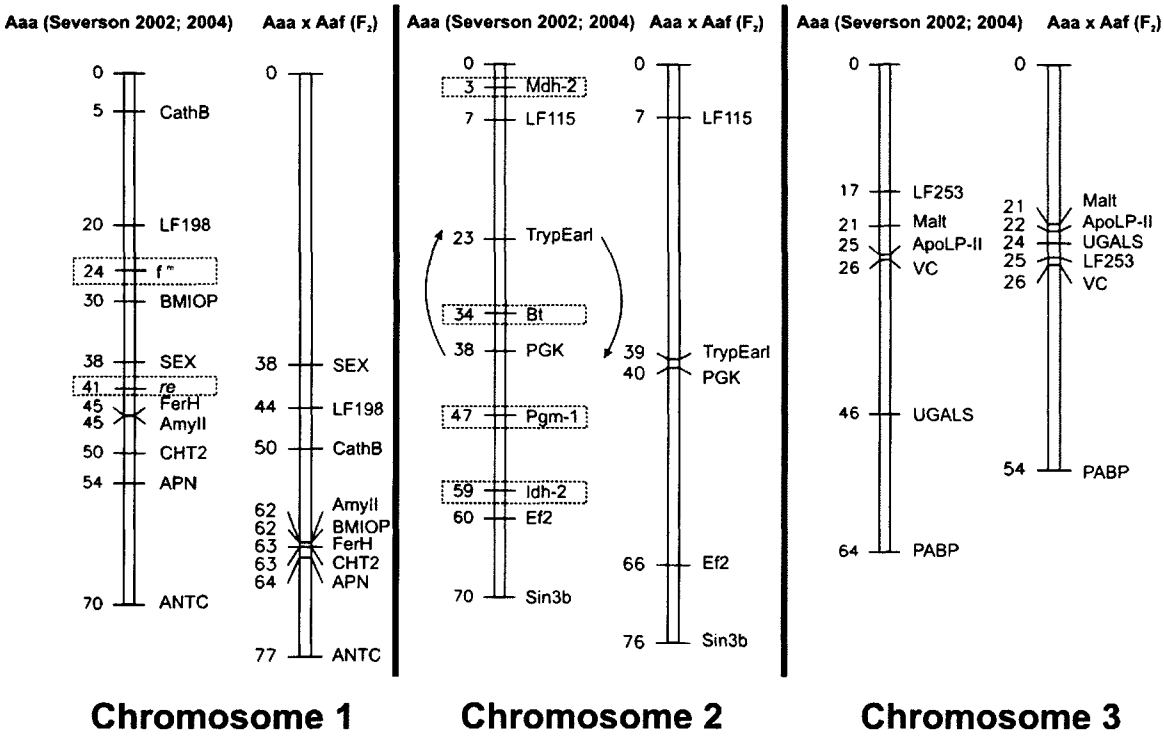
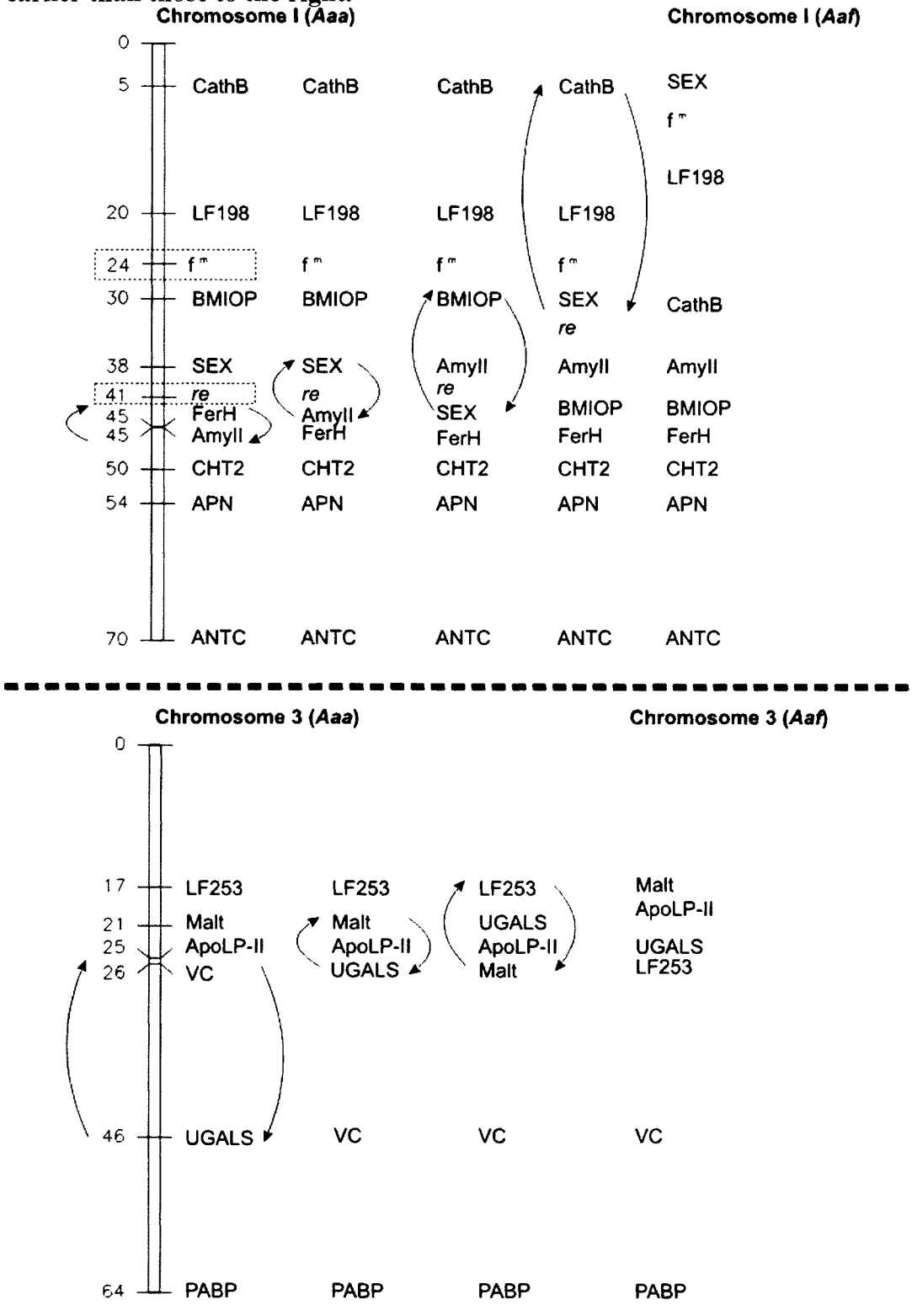


Figure 4.2. GRIMM analysis of gene rearrangements in chromosomes 1 and 3. The beginnings and ends of the curved arrows on the Aaa map represent GRIMM's estimate of the most parsimonious locations of inversion breakpoints to explain the gene order and locations in Aaa/Aaf hybrids. Rearrangements to the left occurred earlier than those to the right.



Chapter 5

Linkage mapping of RNAi, microRNA, and apoptosis genes to identify associations with midgut infection (MIB) and midgut escape barriers (MEB) to infection with DENV-2 in *Aedes aegypti*

Abstract

Aedes aegypti is the most important vector of yellow fever and dengue fever flaviviruses. Vector competence (VC) is defined as the intrinsic permissiveness of an arthropod vector to infection, replication, and transmission of a virus. Researchers have demonstrated global variation in VC for flaviviruses. The RNAi and apoptosis machinery have both been suggested to play a significant role in the mosquito innate immune response and I hypothesized that the RNAi and apoptosis pathways significantly contribute to DENV-2 VC in *Aedes aegypti*. I constructed an F_1 intercross family using a P_1 female from *Aedes aegypti aegypti* (Aaa) and a P_1 male *Aedes aegypti formosus* (Aaf). Recombination rates and gene order were consistent with the presence of multiple inversions in Aaf on all three chromosomes. Even though I was unable to construct an accurate linkage and QTL map in the F_1 intercross *Ae. aegypti* family because of these chromosomal inversions, I conducted genotype-phenotype analysis, using χ^2–goodness-of-fit test, on each of the loci that were used to generate the linkage map (Chapter 5). I determined that *dcr2*, *r2d2*, *r3d1*, *ago2*, and IAP2 all significantly contribute to the midgut escape barrier (MEB) for DENV-2. *dcr2*, *r2d2*, and *ago2* have been identified as major components of the RNAi pathway and my study showed that they are significant contributors to MEB for DENV-2.

Introduction

Vector competence (VC) is the intrinsic permissiveness of a vector to infection, replication and transmission of a virus and is associated with a number of anatomic barriers to productive vector infection (Woodring *et al*, 1996; Hardy, 1988). *Aedes aegypti* populations exhibit extensive genetic variability in VC for flaviviruses,

suggesting association with barriers to infection (Gubler *et al.*, 1979; Rosen *et al.*, 1985). Analyses in Chapter 2 demonstrated that VC in *Ae. aegypti* to dengue virus type 2 (DENV-2) is variable within and between geographic sites, as well as between collection years. It was also observed that certain geographical populations maintain either very high or low VC for DENV-2. These findings suggest the need for detailed genetic analysis of potential mosquito determinants of VC.

Emerging evidence indicates that RNA interference (RNAi) and related pathways function in many fundamental processes, including antiviral defense, development, and maintenance of genomic stability (Tomari and Zamore, 2005; Hannon *et al.*, 2002). RNAi silences gene expression through small interfering RNAs (siRNAs), providing an antiviral mechanism in the innate immune pathway to protect adult flies from infection and unregulated viral dissemination (Waterhouse *et al.*, 2007; Campbell *et al.*, 2008; Keene *et al.*, 2004; Franz *et al.*, 2006; Sanchez-Vargas *et al.*, 2009). Data suggest that the RNAi pathway may contribute to the regulation of virus midgut infection and dissemination. The natural variation of DENV-2 VC in *Ae. aegypti* and the identified antiviral role of RNAi emphasize the need to determine the potential contribution of innate immune pathways to quantitative traits such as VC.

The entire genome sequence of *Ae. aegypti* has been published (Nene *et. al.*, 2007; Lawson, 2009; Sharakhova *et al.*, 2007; Coluzzi *et al.*, 2002). The *Ae. aegypti* genome is roughly 1376 million bp and about 5 times the size of the genome of *Anopheles gambiae* (Nene *et al.*, 2007) Due to this difference in size, as well as complex features such as content of transposable elements (~47%), the *Ae. aegypti* genome

physical map has not been assembled. For this reason it is important to construct genetic linkage maps and identify genes associated with quantitative trait loci (QTL).

Phenotypic genetic markers, such as eye color, scale patterns and insecticide resistance were first identified in the early 1960s (Craig *et al.,* 1961; Craig and Hickey, 1967). Allozyme markers, RFLP, and single strand conformation polymorphism (SSCP) analysis of randomly amplified polymorphic DNA (RAPD) markers have been successfully used to construct *Ae. aegypti* linkage maps (Munstermann and Craig, 1979; Severson *et al.*, 1993, 1994, 995; Antolin *et al.,* 1996; Black and DuTeau, 1997; Bosio *et al.*, 2000; Zhong *et al.*, 2006). RAPD analysis data unfortunately are inconsistent among intercross families.

Fulton *et al.* (2001) explored single nucleotide polymorphisms (SNPs) in PCR products with the use of SSCP analysis. They developed an intensive linkage map by using a large diversity of *Ae. aegypti* cDNA sequences that were available in GenBank. Gomez-Machorro *et al.* (2004) and Bennett *et al.* (2005) used this technique to determine additional loci for the *Ae. aegypti* linkage map and to identify MEB and MIB QTL specific to DENV-2. Saavedra-Rodriguez *et al.* (2008) further refined SNP analysis using allele specific real time melting curve PCR to construct a linkage map and identify QTLs specific to permethrin resistance. This particular methodology allows for linkage mapping of specific genes.

We generated an F_1 intercross family between the laboratory strains of *Ae. aegypti aegypti* (*Aaa*) that had been selected for susceptibility to DENV-2 (*D2S3*) and an *Ae. aegypti formosus* (*Aaf*) strain (PK10) collected from the Forest Gallery region of southeastern Senegal (Bennett *et al.*, 2005; McClelland, 1974). The original goal of

creating this intercross family was to add to the current linkage map of QTL for midgut infection barrier (MIB) and midgut escape barrier (MEB) in *Ae. aegypti* specific to DENV-2. RNAi is an antiviral mechanism to protect mosquitoes from infection and unregulated viral dissemination (Campbell *et al.*, 2008; Keene *et al.*, 2004; Sanchez-Vargas *et al.*, 2009). Limited data also suggest a potential role for apoptosis in the limitation of viral replication in mosquitoes (Wang H. *et al.,* 2008). Genes specific to the RNAi, microRNA and apoptosis induction and implementation pathways were chosen, in addition to previously mapped housekeeping genes, for determination of their map locations and potential contribution to the DENV-2 MIB and MEB quantitative traits. Phenotype-genotype analyses were performed to identify loci that were significantly associated with DENV-2 MIB and MEB. We were unable to complete QTL mapping due to the incongruence in the F_2 linkage map with previously developed linkage maps as a result of the genomic inversions found in Aaf (Bernhardt *et al.*, 2009).

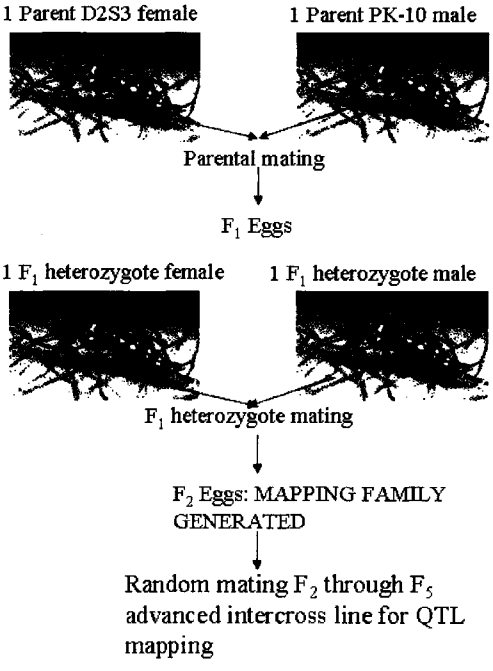
Materials and methods

Mosquito strains

Two *Ae. aegypti* strains were used to generate the F_1 intercross family: DENV-2 susceptible on chromosome 3 (*D2S3*) mosquitoes, of which 80-90% have a disseminated infection with DENV-2 after ingesting an infectious artificial blood meal, and *Ae. aegypti formosus* (PK10) mosquitoes, which originated from eggs collected during the summer of 2006 from the Forest Gallery region in Southeastern Senegal, Africa (Bennett *et al.*, 2005; Sylla *et al.*, 2009). Of the PK10 mosquitoes from this field collection, 92% developed a DENV-2 midgut infection, and 16% of these developed a disseminated

Figure 5.1. F_1 intercross design in mosquitoes. F_1 intercross using *Ae. aegypti* with high (*D2S3*) and low (PK10) vector competence for DENV-2.

- ***D2S3* colony – *Aedes aegypti aegypti* dengue 2 susceptible on chromosome 3**
  - Lab selected colony (Bennett *et al.*, 2005)
  - Vector Competence = **~80-90%**
  - MEB = ~8-10%
- **PK10 – *Aedes aegypti formosus***
  - Field collection obtained from Senegal, Africa
  - Vector competence = **~16%**
  - MEB = ~73%
- F_1 intercross design
  - One female/one male in P_1
  - F_1 intermate to generate F_2 family
  - Random intermating continued >F_2



infection, thus demonstrating a natural MEB (Sylla *et al.*, 2009).

Breeding design

A *D2S3* and PK10 *Ae. aegypti* F_1 intercross family was developed by mating *D2S3*(♀) x PK10(♂) described in Chapter 4 and in Figure 5.1.

***Per os* infection**

The F_2 and F_5 generations of the *D2S3*(♀)xPK10(♂) advanced intercross line were orally challenged with DENV-2 JAM1409 to obtain VC phenotype data. *Per os* infections followed the same protocol as described in Chapter 2.

Vector competence, midgut infection barrier and midgut escape barrier

Mosquitoes were phenotyped for VC 14 days post oral challenge by indirect immunofluorescence assay (IFA) for DENV-2 E antigen in heads and midguts, following the same protocol as described in Chapter 2. IFA data were obtained for 45 F_2 and 426 F_5 generation mosquitoes.

Genomic DNA extraction

DNA was extracted from single mosquitoes from the *D2S3*(♀)xPK10(♂) family. This family was renamed D2PK at the time of DNA extraction to simplify labeling and future analysis. P_1 male and female, F_1 male and female, 227 F_2 males and females, and 426 F_5 females underwent DNA extraction using the Qiagen DNeasy Blood and Tissue Kit (Valencia, CA).

Single mosquitoes were placed in 1.7 ml tubes with 180 µl of PBS (8 g NaCl, 0.2 g KCl, 1.15 g sodium phosphate dibasic, 0.2 g potassium phosphate monobasic, 1 L ddH_2O, pH 7.2). Mosquitoes were manually triturated using a plastic pestle until no distinguishable body parts were visible, then 180 µl of ATL buffer (Qiagen) and 20 µl of

proteinase K (Qiagen) were added and the tube was thoroughly vortexed. The homogenate was then incubated at 56°C for 30 minutes.

The homogenate was quickly vortexed prior to adding 200 µl of Buffer AL (Qiagen) and 200 µl of ethanol (100%), then the homogenate was vortexed again and pipetted into the DNeasy Mini spin column provided by Qiagen. The column was then centrifuged at 8,000 rpm for 1 minute and flow-through was discarded. 500 µl of Buffer AW1 (Qiagen) was added to the spin column, centrifuged at 8,000 rpm for 1 minute, and flow-through was discarded. 500 µl of Buffer AW2 (Qiagen) was added to the spin column and centrifuged for 3 minutes at 14,000 rpm. The flow-through was discarded and the spin column was transferred to a clean 1.7 ml tube. 200 µl of Buffer AE (Qiagen) were pipetted directly onto the DNeasy membrane inside the spin column and allowed to incubate for 1 minute at room temperature. The spin column inside the 1.7 ml tube was centrifuged for 1 minute at 8,000 rpm. The eluate contained the concentrated DNA from a single mosquito. 100 µl of the eluate was removed, placed in a separate tube and stored long-term at -70°C, while the remaining 100 µl of eluate was stored at -20°C for PCR amplification.

PCR amplification

PCR was performed with primers to amplify specific genetic loci for SSCP analysis. The list of loci and their primer sets are in Tables 4.2, 5.1 and 5.2. DNA extracted from the P_1 grandparents, F_1 parents, all F_2 offspring and F_5 females from the D2PK family were used as templates for PCR.

PCR was performed using the Promega GoTaq Green Master Mix 2x protocol (Madison, WI). A single batch of PCR reaction mixture sufficient for 100 single 50 µl

reactions was prepared. The reaction mixture consisted of 2500 µl of GoTaq Green Master Mix 2x, 1900 µl of nuclease-free water (provided by Promega), 200 µl of 10 mM forward primer and 200 µl of 10 mM reverse primer. The reaction mixture was distributed into a 96-well polycarbonate plate (Eppendorf, Westbury, NY) at 48 µl per well. 2 µl of template DNA from individual mosquitoes were then added to each well. Each plate underwent PCR amplification on an Eppendorf thermocycler. Thermal cycling conditions were: (1) 95°C for 2 min (first denature); (2) 95°C for 30 s (denature in cycle); (3) 50-58°C (depending on annealing temperature of primers) for 1 min (anneal); (4) 72°C for 30 s (extension); (5) cycle to step (2) 34 times; (6) 72°C for 5 min (final extension); and (7) 4°C final cooling for storage.

Single strand conformation polymorphism analysis

Single strand conformational polymorphism (SSCP) analysis was performed on all PCR products as described by Black and DuTeau (1997) and Bosio *et al.* (2000). SSCP gels were scored based on the banding patterns of DNA from the P_1 and F_1 males and females. Homozygous alleles inherited from the P_1 female would have a specific SSCP banding pattern and were coded as a "1". Homozygous alleles inherited from the P_1 male would have a specific SSCP banding pattern and were coded as a "2". F_1 parents that had distinct banding pattern different from the P_1 male or female were considered heterozygotes and coded as a "3". Locus specific PCR and SSCP analyses were performed to determine rates of allele specific inheritance in each mosquito in the F_2 and F_5 D2PK family.

Allele-specific PCR

SNPs were also detected using allele-specific real time PCR following the protocol described in Urdaneta-Marquez *et al.* (2008). Optimal PCR primers were designed using Primer Premier software (Premier Biosoft International, Palo Alto, CA) and are listed in Tables 4.2 Set B and 5.2. The 3' ends of the initial 25-nucleotide (nt) primers contained one of the alternate nucleotides at the SNP locus in the antisense strand and the primer was then examined for stable ($\Delta G < -6.0$ kcal/mol) self-annealing or secondary structures. If either structure appeared on the 5' end, we trimmed the sequence to ≥ 18 nt in an attempt to remove the problem nucleotides. Otherwise, we repeated the primer design process but with the SNP at the 3' end of the 25 nt primer located on the sense strand. Once optimal primers were confirmed, the third nucleotide from the 3' end was replaced with a transversion mismatch. This step was originally recommended by Okimoto and Dodgson (1996) and greatly improved the specificity of the assay. In addition, we used different mismatched purines or pyrimidines in the two allele-specific primers (see Table 5.2). Next, we used Primer Premier software to design a reverse primer to the allele-specific primers. Criteria were set to identify reverse primers that would amplify a product < 100 bp as recommended by Wang *et. al.* (2005). The reverse primer was also examined for stable self-annealing or secondary structures both with itself and with its two allele-specific primer pairs.

Genotypes were determined in a single-tube PCR using the two different "allele-specific" primers and the reverse primer. Allele-specific primers were manufactured (Operon Inc., Huntsville, AL, USA) with the 5' tails originally described by Germer and Higuchi (1999) and Wang *et al.* (2005) to discriminate amongst the amplification

products either based on size or melting temperature. These were (5'-GCGGGCAGGGCGGCGGGGGCGGGGCC [allele-specific primer 1]-3' or 5'-GCGGGC [allele-specific primer 2]-3').

PCR was performed in a 25 µl volume in 96-well Hard Shell plates with white wells (Bio-Rad Laboratories, Hercules, CA, USA). Each reaction contained 12.5 µl of 2x IQ SYBR Green Supermix (Bio-Rad Laboratories) (final concentration = 50 mM KCl, 20 mM Tris-HCl, pH 8.4, 0.2 mM of each dNTP, 0.625 units iTaq DNA polymerase, 3 mM $MgCl_2$, 1 x SYBR Green I, 10 nM fluorescein), 25 pm of each primer, ~100 ng of template DNA, and sterile filtered ddH_2O added to make a final 25 µl volume. PCR wells were covered with Flat Cap Strips (Bio-Rad Laboratories) and placed into the BioRad iCycler 5 (Bio-Rad Laboratories, Hercules, CA, USA). Thermal cycling conditions were: (1) 95°C for 12 min (first denature); (2) 95°C for 20 s (denature in cycle); (3) 60°C for 1 min (anneal); (4) 72°C for 20 s (extension); (5) cycle to step (2) 39 times; (6) 72°C for 5 min (final extension); and (7) ramp from 65°C to 95°C at a rate of 0.2 °C/s (melting curve).

Intensive linkage mapping

Linkage mapping was performed following the protocol described in Bosio *et al.* (2000). Genotypes at each locus were scored and entered into a JoinMap data file with "cross pollination" format (Stam and Van Ooigen, 1995). These were tested for conformity to Mendelian genotype ratios with a χ^2 goodness of fit analysis using the JMSLA procedure in JoinMap. Loci at which Mendelian genotype ratios were observed were separated into individual linkage groups using the JMGRP and JMSPL procedures with a starting LOD threshold of 0.0 increasing to 8.0 in increments of 0.1. Pairwise

distances were estimated among loci in each of the three linkage groups using the JMREC procedure and the maximum likelihood map was estimated using the JMMAP procedure. The linkage map was drawn with DrawMap 1.1 (Van Ooigen, 1994).

Genotype-phenotype analysis

Associations between genotypes at each marker locus and DENV-2 infection phenotypes were initially assessed by a goodness-of-fit χ^2 analysis. The null hypothesis was that the number of mosquitoes with midgut and disseminated infections were equal in each genotype class. When a significant χ^2 was detected, we examined the inheritance of the alleles at that locus. Our *a priori* hypothesis was that an excess of F_5 individuals with an allele inherited from the *D2S3* P_1 parent would have disseminated infections, while an excess of F_5 individuals with an allele inherited from the PK10 P_1 parent would not have disseminated infections.

Additive and dominance effects were calculated relative to alleles at each marker locus that was significant for either MIB and MEB quantitative trait based on methods described by Mori *et al.* (2008) and Babu *et al.* (2006). Mode of gene action was determined as the absolute value of dominance/additive where additive = 0 - 0.20, partial dominance = 0.21 - 0.80, dominance = 0.81 - 1.20 and overdominance > 1.20.

Results

Mapping family

An F_1 intercross family was bred with the intention of mapping QTLs that condition the large differences in midgut escape barriers between *Aaa* and *Aaf*. We crossed an *Aaa* female from the *D2S3* strain with an *Aaf* male from the PK10 strain and isolated DNA from the two P_1 parents, two F_1 parents, 227 F_2 offspring and 426 F_5

offspring (Bennett *et al.*, 2005; Sylla *et al.*, 2009). Thirty-five loci were selected for mapping and underwent PCR amplification. Twenty of these are fully described in Chapter 4 and 15 loci were added specifically for this study. Three loci were selected because of their involvement in the RNAi pathway (*r2d2*, *dcr2*, *ago2*), three for their involvement in the microRNA pathway (*r3d1*, *dcr1*, *ago1*), and four were selected for their involvement in the apoptotic pathway (IAP1, IAP2, Casp8, VIAP) (Tables 5.1, 5.2). These markers were subjected to SSCP analysis to determine informative genotype inheritance patterns. Loci with informative genotypes in the P_1 and F_1 parents underwent PCR product sequencing to identify SNP alleles that were segregating in the family. Genotype data for 4 loci were collected using SSCP analysis in the 227 F_2 and 426 F_5 offspring. Thirty-one SNP loci were amplified and genotyped with allele-specific primers (Tables 4.2 Set B, 5.2) and real time melting curve PCR.

Preliminary linkage map

A D2PK F_2 linkage map was constructed as described in Bernhardt *et al.* (2009) (Figure 5.3). The linkage map encompassed 13 loci on chromosome 1 covering 97 cM, 9 loci on chromosome 2 covering 81 cM, and 13 loci on chromosome 3 covering 54 cM. Twenty-three of the 35 loci were mapped in previous studies and are identified with red arrows in the most current intense linkage map (Figure 5.2).

Vector competence, MIB and MEB

The *D2S3* line had a VC of 85%, MIB of 6% and a MEB of 9% (Chapter 2). The PK10 line had a VC of 16%, MIB of 11% and a MEB of 73% (Bernhardt SA, unpublished). The 45 female F_2 offspring from the D2PK F_1 intercross had a VC of 26.67%, MIB of 26.67% and a MEB of 46.67%. These numbers were low due to low

titer of DENV-2 JAM1409 ($1x10^4$ plaque forming units/ml) that was used in the artificial infectious blood meal. Positive control *D2S3* mosquitoes were fed at the same time as the F_2 family and they also had a low VC of 63%. The 426 female F_5 offspring from the D2PK F_1 intercross had a VC of 45.31%, MIB of 13.15% and a MEB of 41.55%. The DENV-2 virus suspension used to create the infectious blood meal had a titer of $1x10^7$ plaque forming units/ml.

Genotype-phenotype associations

χ^2–goodness-of-fit results indicate whether each locus is distributed independently of a particular phenotype (e.g. MIB or MEB; Figure 5.3). Significant P-values from analysis of the F_5 offspring are labeled to the right of the chromosomes. Loci that were statistically associated with MIB and/or MEB are labeled with a P-value from the χ^2–goodness-of-fit test. Loci that were statistically associated with MEB were *r2d2*, Amy2, *r3d1*, UGALs, LF253, IAP2, *ago2* and *dcr2*. Loci that were statistically associated with MIB are Coe1 and Sin3.

The number of alleles inherited from each of P_1 parents was plotted against MIB to test the *a priori* assumption that F_5 alleles associated with a higher rate of midgut infections are inherited from the *D2S3* parent (Figure 5.4). Figures 5.5 and 5.6 illustrate two loci that were significantly associated with MIB. These figures include the plots for both MIB and MEB, and show the number of mosquitoes with each allele pair (labeled as n) and the Fisher exact test results for both phenotypes. Mosquitoes in the F_5 generation were less likely to become infected with DENV-2 when they were homozygous or heterozygous for alleles from the PK10 parent at the Coe1 and/or Sin3 loci.

The number of alleles inherited from each of the P_1 parents was plotted against MEB to test the *a priori* assumption that F_5 alleles associated with a higher rate of disseminated infections were inherited from the *D2S3* parent (Figure 5.7). Loci that were significantly associated with MEB, by χ^2–goodness-of-fit test, were individually graphed and can be found in Figures 5.8-5.15. Figures 5.8-5.15 include the plots for both MIB and MEB for each locus, and show the number of mosquitoes with each allele pair (labeled as n) and the Fisher exact test results for both phenotypes.

UGALs, LF253, IAP2, *dcr2*, Amy2 and *ago2* were statistically significantly associated with MEB, by Fisher exact test $P<0.05$, and graphs demonstrate that a mosquito is less likely to allow DENV-2 dissemination when it inherited one or two alleles from the PK10 parent in the F_1 intercross family. *r2d2* and *r3d1* also are significantly associated with MEB, by Fisher exact test ($P<0.05$), but do not follow the predicted trend. *r2d2* is found on chromosome 1 and is closely linked with the SEX locus, and this particular study used only females in the F_5 analysis. Analysis of *r2d2* demonstrated that 98.3% of mosquitoes in the F_5 generation inherited two alleles from the *D2S3* female grandparent. *r3d1* is found on chromosome 2 and analysis demonstrated that 99.4% of mosquitoes in the F_5 generation inherited at least one allele from the *D2S3* female parent.

Additive and dominance effects were calculated for each of the loci that was statistically significantly associated with MIB or MEB by Fisher exact test (Table 5.3). Values highlighted in bold represent the phenotype for which association of the gene was determined to be statistically significant. Two loci had insufficient data available (*r2d2*, Coe1) and effects were not calculated. Sin3 was determined to be statistically significant

for partial dominance in conditioning midgut infections (MIB). Amy2, *r3d1*, UGALs, IAP2, *ago2* and *dcr2* were statistically significant for conditioning dominance in MEB. Amy2 and *r3d1* were overdominant, IAP2 was dominant and UGALs, *ago2* and *dcr2* were partially dominant in conditioning disseminated infections.

Discussion

Specific loci significantly contribute to the MIB and MEB phenotypes for DENV-2 in *Ae. aegypti* and thus to DENV-2 infection and dissemination. The MIB rate increased as a mosquito inherited one and two alleles of the Sin3 (unknown function) and Coe1 (Carboxylesterase) loci from the PK10 refractory parent (Figure 5.4). The Sin3 allele demonstrated partial dominance in conditioning midgut infections in the F_5. Small sample size prevented calculation of additive or dominance effects specific to Coe1. Coe1 is found on chromosome 1 and is linked to the SEX locus. Since no males were included in the F_5 analysis, no PK10 Coe1 alleles were inherited.

The MEB rate increased as mosquitoes inherited one or two alleles at loci UGALs (UGALs vitellogenin), IAP2, *dcr2*, Amy2 (alpha-Amylase 2) and *ago2* from the PK10 parent (Figure 5.7). Amy2 alleles demonstrated overdominance effects, IAP2 alleles demonstrated dominance effects and UGALs, *ago2* and *dcr2* alleles demonstrated partial dominance effects in conditioning disseminated DENV-2 infections in the F_5 generation. *r2d2* alleles significantly contributed to MEB when even one allele was inherited from the PK10 parent, even though the MEB rate did not increase in the allele specific analysis. *r2d2* is on chromosome 1 and is also linked to the SEX locus. No males were analyzed from the F_5 generation, which in turn resulted in very few mosquitoes with PK10 alleles specific to *r2d2* being identified. It is likely that *r3d1* is within the

inversion in chromosome 2 due to the low number mosquitoes containing two PK10 alleles, which helps explain the misrepresentation of the number of PK10 alleles specific to this locus.

The RNAi pathway genes (*ago2*, *r2d2*, *dcr2*) significantly contributed to DENV-2 dissemination in *Ae. aegypti* mosquitoes. The microRNA pathway genes that are paralogs to *ago2* and *dcr2* (*ago1* and *dcr1*), did not significantly contribute to DENV-2 dissemination. We examined these genes based on their potential antiviral role (Waterhouse *et al.*, 2007; Campbell *et al.*, 2008; Keene *et al.*, 2004; Franz *et al.*, 2006; Sanchez-Vargas *et al.*, 2009).

r3d1 (loquacious) alleles were found to significantly contribute to DENV-2 dissemination. Both *r2d2* and *r3d1* belong to a family of double stranded RNA binding proteins that function in tandem with Dicer enzymes in the siRNA and/or microRNA pathways (Paroo *et al.* 2007). A substantial amount of research has found that Dicer-2 partners with R2D2 for loading siRNA into siRISC (Liu *et al.* 2003, 2006; Tomari *et al*, 2004). R3D1 has been found to play an important role in the microRNA pathway, partnering with Dicer-1 for assembly of the microRISC (Förstemann *et al.*, 2005; Jiang *et al.*, 2005; Saito *et al.*, 2005). These data suggest that R3D1 may also play a role in RNAi. Recent findings with *Drosophila* have identified a new small RNA class, endogenous small interfering RNAs (Czech *et al.*, 2008; Chung *et al.*, 2008). This pathway is similar to the antiviral pathway that requires Ago2 and Dcr2, but differs in that loquacious (R3D1), not R2D2, plays the RISC-loading role for some siRNA (Obbard *et al.*, 2009). These observations in *Drosophila*, and potentially in *Ae. aegypti*, suggest redundancy of functions between several pathways.

IAP1, IAP2, Casp8 (caspase 8) and VIAP (viral inhibitor of apoptosis) loci were selected to determine possible association of apoptosis with VC. Apoptosis is a genetically regulated mechanism of programmed cell death that is critical during normal development, tissue homeostasis and disease processes and innate immunity in metazoans (Evan *et al.* 1998; Saikumar *et al.*, 1999). Some parasites, viruses and autoimmune disorders induce or inhibit host cell apoptotic death via genetically controlled mechanisms (Thompson *et al.*, 1995; Peter *et al.* 1997). Apoptosis upon infection has been thoroughly described in vertebrates. Morphological features indicative of apoptosis in mosquito midgut lumen cells have been described after invasion by unicellular eukaryotes such as *Leishmania* spp., *Trypanosoma* spp. and *Plasmodium* spp. (Arnoult *et al.*, 2002; Hurd *et al.*, 2004).

Limited descriptions of apoptosis exist in vectors as a result of arbovirus infection. One example describes potential signs of apoptosis in the midguts of *Culex pipiens* as a result of West Nile virus infection (Viadyanathan *et al.*, 2006). Wang *et al.* (2008) infected cultured *Ae. albopictus* cells (C6/36) with an alphavirus expressing *Ae. aegypti* pro-apoptotic genes. The cells demonstrated apoptosis and the infection changed from persistent to lytic. Research is just beginning to uncover potential immune related genes specific to apoptosis in the *Anopheles gambiae* and *Ae. aegypti* genomes (Christophides *et al.*, 2002; Bartholomay *et al.*, 2004; Zhou *et al.*, 2005; Waterhouse *et al.*, 2007).

IAP2 was the only apoptotsis-associated locus that was found to be significantly associated with DENV-2 dissemination in *Ae. aegypti* mosquitoes. The inhibitor of apoptosis protein (IAP) family were shown to be cell death inhibitors (Hay *et al.*, 1995;

Wang *et al.*, 1999; Hawkins *et al.*, 1998; Meier *et al.*, 2000). Recent evidence suggests that IAPs are multifunctional signal transducers that influence diverse biological processes (Bryant *et al.*, 2008). For example, Leulier *et al.* (2006) demonstrated that IAP2 in *Drosophila* functions in the host immune response to gram-negative bacteria. It is difficult to hypothesize the potential role of IAP2 from this study since no other apoptosis loci (IAP1, VIAP, Casp8) were found to be associated with DENV-2 dissemination. Ultimately, it is important to conduct QTL analyses beyond the χ^2–goodness-of-fit test, since the locus and phenotype of interest are not usually independent of other nearby loci.

The primary purposes of Chapters 4 and 5 were to produce linkage and QTL maps for RNAi, microRNA and apoptosis genes to determine whether alleles at these loci contribute to DENV-2 infection and transmission in *Ae. aegypti*. Although chromosomal inversions precluded mapping, my findings contribute to knowledge of the importance of RNAi and/or apoptosis as potential contributors to VC, but also demonstrates the need to develop linkage and QTL maps, using an *Aedes aegypti aegypti* cross, to show possible associations of RNAi, microRNA and apoptosis genes with arbovirus infection and dissemination.

Table 5.1. Primer sequences, annealing temperature (T_a), and anticipated PCR product size for each of the loci used to construct the linkage map in *Ae. aegypti*.

Locus name	Gene name or description	Forward Primer	Reverse Primer	Start location	PCR product size	Ta
Dicer-1	microRNA pathway	5'-ATCTGCTGCGATTGAGGAA-3'	5'-TCGGAGTCACCGTCACCACT-3'	45321	675	58.5
Dicer-2	RNAi pathway	5'-ATGCTAAGACTGGCGTTGAGG-3'	5'-TGTTTGCGGATGGTCATACTTGA-3'	16959	595	55.8

Table 5.2. Sequences of oligonucleotides used in this study for allele-specific PCR. The sequence of the long tail is 5'-GCGGGCAGGGCGGCGGGGGCGGGGCC-3' while the sequence of the short tail is 5'-GCGGGC-3'. The 3' allele-specific nucleotide is in bold and the mismatch and the third nucleotide from the 3' end is underlined.

Locus name	Gene name or description	Oligonucleotide	PCR product size	SNP position (according to gene sequence)
CYP6P12V1	Cytochrome P450s	5'-GTTCATCTTTGCGTCGTTG-3'		762
		5'-[long tail]AGCATCCGTAATCTTAACCCCCCAC-3'	90	
		5'-[short tail]AGCATCCGTAATCTTAACCCCCTAT-3'	70	
R2D2	RNAi pathway	5'-TCGTCGGGCGGCGGAAA-3'		12996
		5'-[long tail]CATCAAGAAAAARCTGGGCGTACAA-3'	114	
		5'-[short tail]CATCAAGAAAAARCTGGGCGTATAG-3'	94	
Coe1	Carboxylesterase	5'-TACCAGGTGTTCAGACGGG 3'		1851
		5'-[long tail]CCAAAATCAAAACATGGCTTTCGTC-3'	113	
		5'[short tail]CCAAAATCAAAACATGGCTTTCATT-3'	93	
IAP1	Inhibitor of apoptosis	5'-TGGATTAGCGTTAGCGTAG-3'		201
		5'-[long tail]TTCAAATAAGTAATTGCGTCTAGCG -3'	101	
		5'-[short tail]TTCAAATAAGTAATTGCGTCTAACA-3'	81	
Mucin	Mucin	5'-CACCACAACGGCACCAGC-3'		621
		5'-[long tail]GTTGAAGAGGGAGGAGCAATA-3'	74	
		5'-[short tail]GTTGAAGAGGGAGGAGCAGTG-3'	54	

Table 5.2. *Continued.*

Locus name	Gene name or description	Oligonucleotide	PCR product size	SNP position (according to gene sequence)
R3D1	microRNA pathway	5'-CGATTTGCCCGTACCCAT-3'		12653
		5'-[long tail]ATGTTTAAGGTAATGAAC**T**G-3'	78	
		5'-[short tail]ATGTTTAAGGTAATGAA**T**T**A**-3'	58	
PGM	Phosphoglucomutase	5'-CTGTTGGCATACTTCTGGC-3'		1093
		5'-[long tail]GTCATTGCTCACTACG**T**C-3'	105	
		5'-[short tail]GTCATTGCTCACTACG**T**A-3'	85	
VIAP	Viral inhibitor of apoptosis	5'-AACAAGGTTGATATCGTTCTT-3'		91
		5'-[long tail]TTCGGTGTCYTCATTGGGGTCC**A**A**G**-3'	76	
		5'-[short tail]TTCGGTGTCYTCATTGGGGTCC**G**A**A**-3'	56	
IAP2	Inhibitor of apoptosis	5'-CACATTTGACCTGATCCGCATTG-3'		60
		5'-[long tail]ACTGGCCCATTTCGCAC**T**C-3'	110	
		5'-[short tail]ACTGGCCCATTTCGCAC**T**T**T**-3'	90	
Ago2	RNAi pathway	5'-CATTCTCTCGTACTCCTTG-3'		44221
		5'-[long tail]TTTACCAACCAACATCT**T**G-3'	103	
		5'-[short tail]TTTACCAACCAACATCT**T**T**G**-3'	83	
CYP4H32	Cytochrome P450s	5'-CTATCCAGATCCAGAACG-3'		1299
		5'-[long tail]GCTGAACGGAATGTAATCGTAY**C**G**G**-3'	114	
		5'-[short tail]GCTGAACGGAATGTAATCGTAY**T**G**A**-3'	94	
Ago1	microRNA pathway	5'-GCGTCATCACACGAGACTATTC-3'		14655
		5'-[long tail]TTGCCCGATTTGCCGCTC**G**G**C**-3'	86	
		5'-[short tail]TTGCCCGATTGCCGCTC**A**G**T**-3'	66	
Casp8	Caspase 8 - apoptosis	5'-TCACCAACCTCGGATACG-3'		982
		5'-[long tail]CCACCTGGTTCATTATC**A**C**T**-3'	95	
		5'-[short tail]CCACCTGGTTCATTATC**G**C**C**-3'	75	

Figure 5.2. Current linkage map of *Ae. aegypti*. Loci indicated with red arrows represent the previously mapped markers that were used to generate the linkage map from the D2PK F_1 intercross (Black WC and Severson D, 2005).

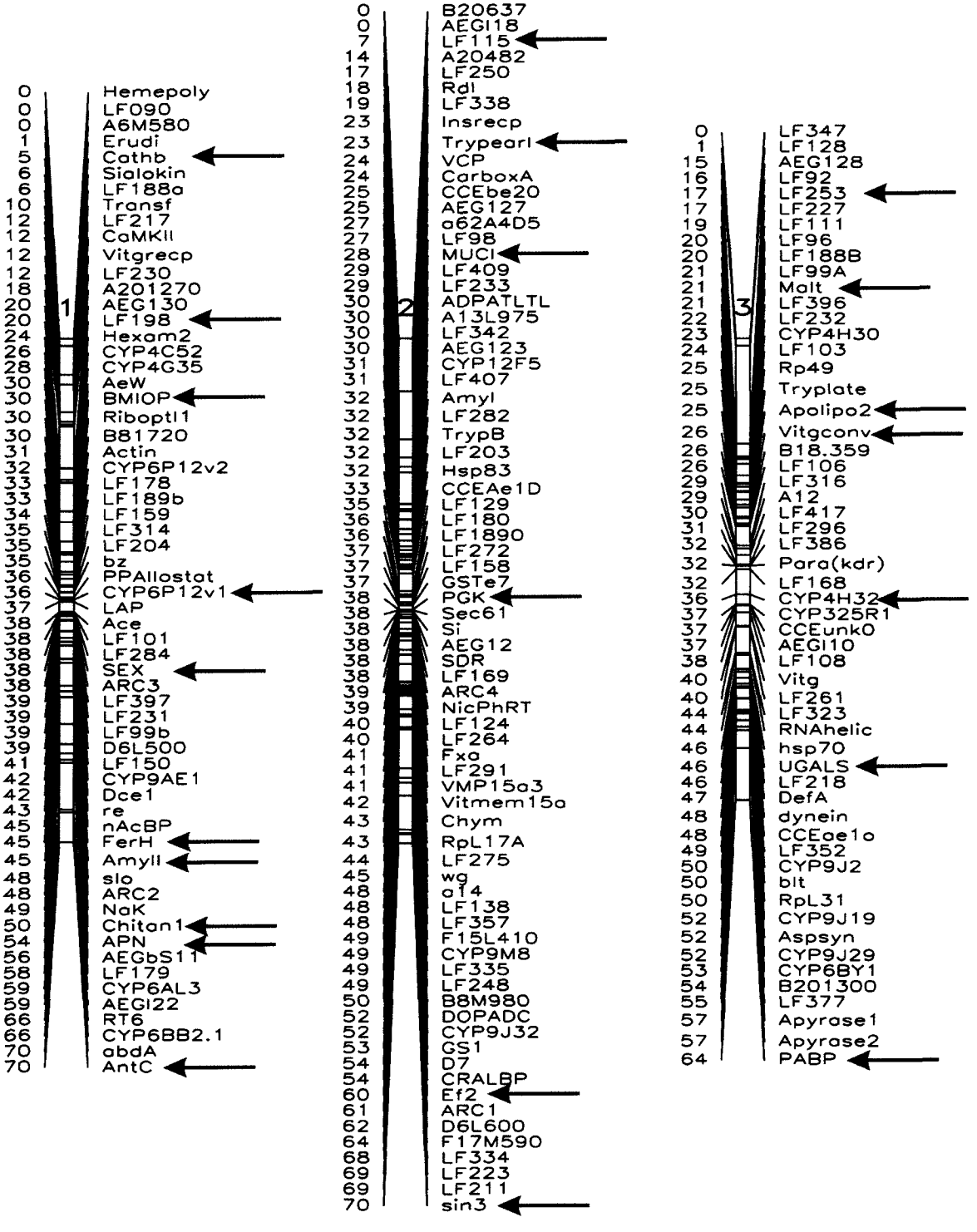


Figure 5.3. D2PK F_1 intercross *Ae. aegypti* linkage map (Aaa x Aaf (F_2)). The significant χ^2–goodness-of-fit test P-values for MIB and MEB are listed to the right of specific loci.

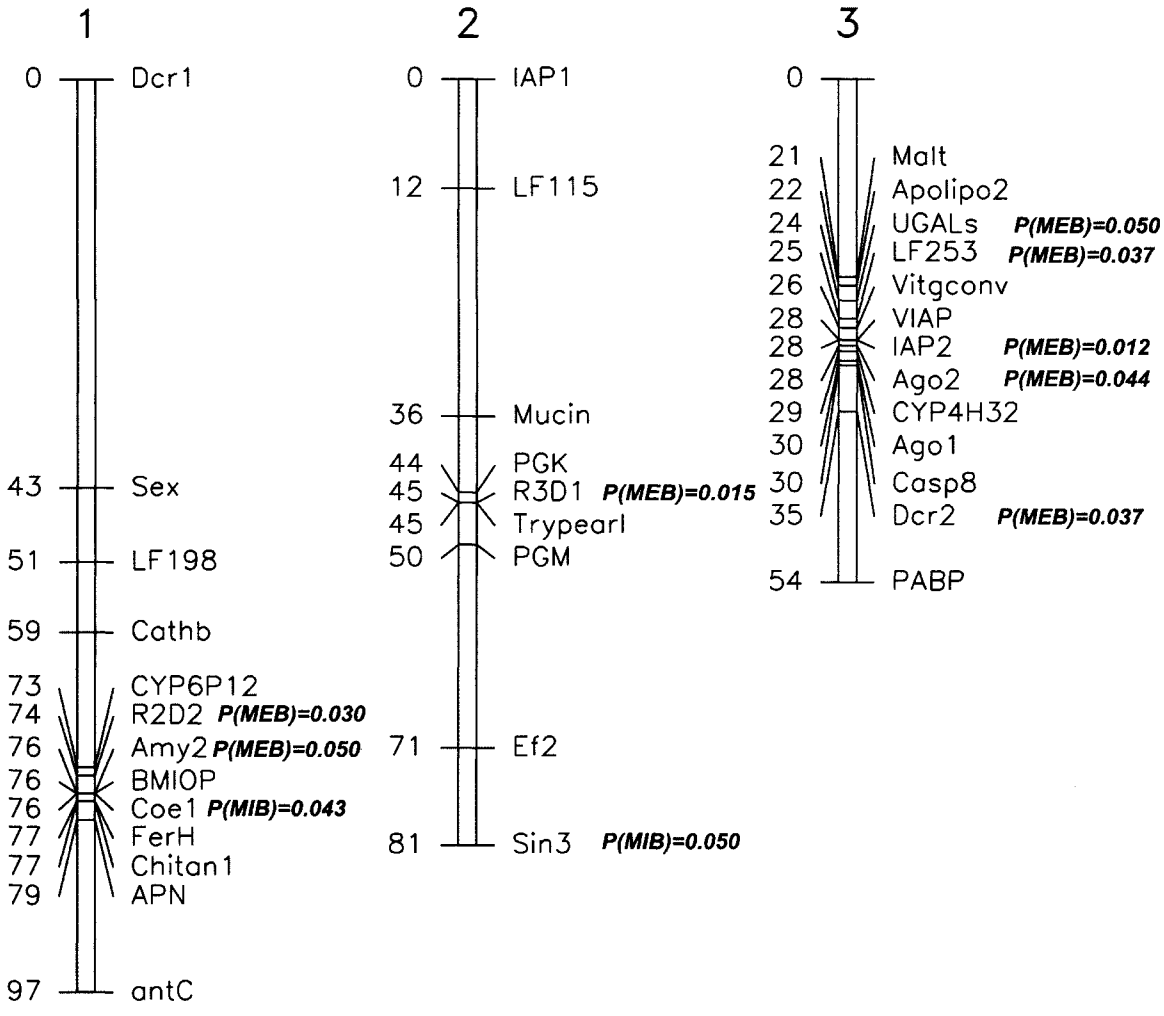


Figure 5.4. Plot of MIB as a function of the number of alleles inherited from either the *D2S3* or PK10 P_1 parent in the F_5 D2PK family. Only significant loci, by χ^2–goodness-of-fit, specific to DENV-2 MIB were plotted.

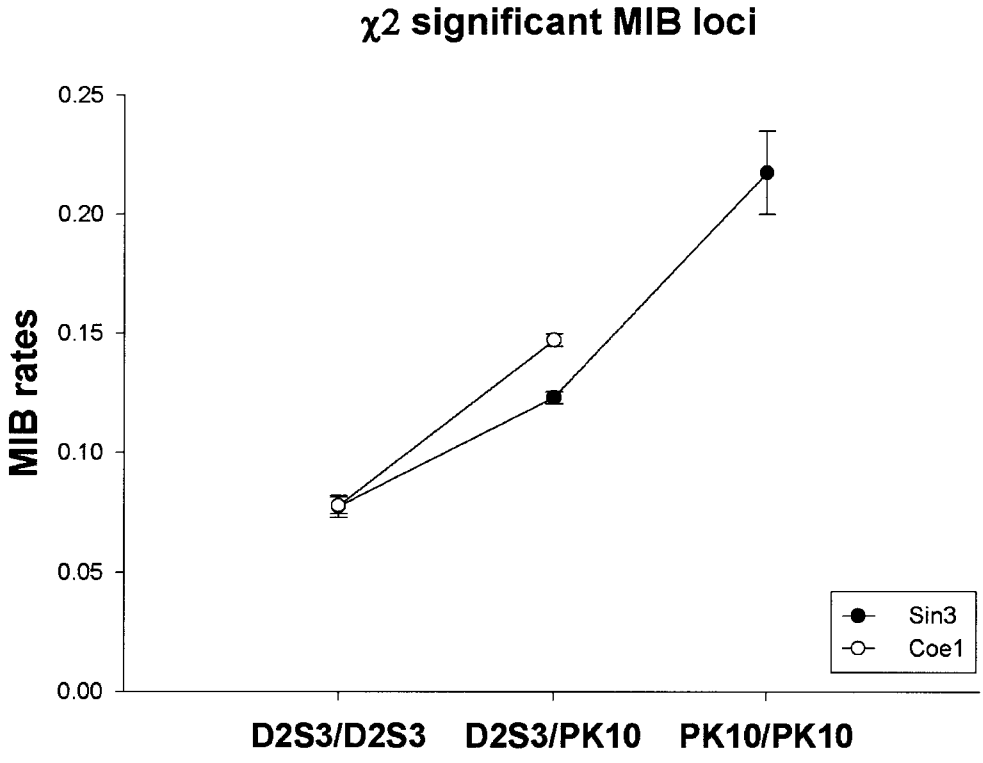


Figure 5.5. Plot of MIB and MEB rates as a function of the Coe1 locus alleles inherited from the *D2S3* and PK10 P_1 parents. Plot shows number (n) of mosquitoes with each set of alleles and the P-value for the Fisher exact test (FET) for each allele pair.

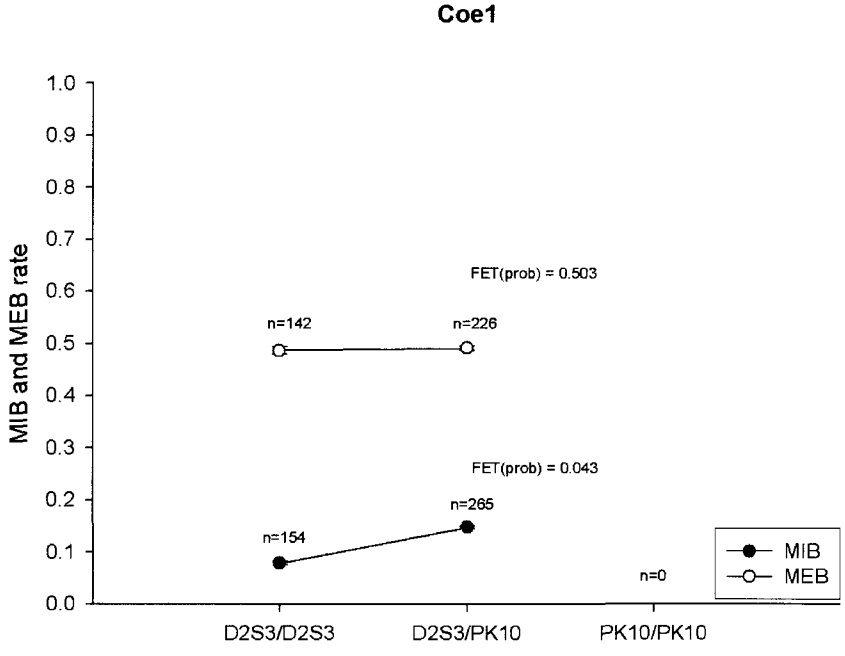


Figure 5.6. Plot of MIB and MEB rates as a function of the Sin3 locus alleles inherited from the *D2S3* and PK10 P_1 parents. Plot shows number (n) of mosquitoes with each set of alleles and the P-value for the Fisher exact test (FET) for each allele pair.

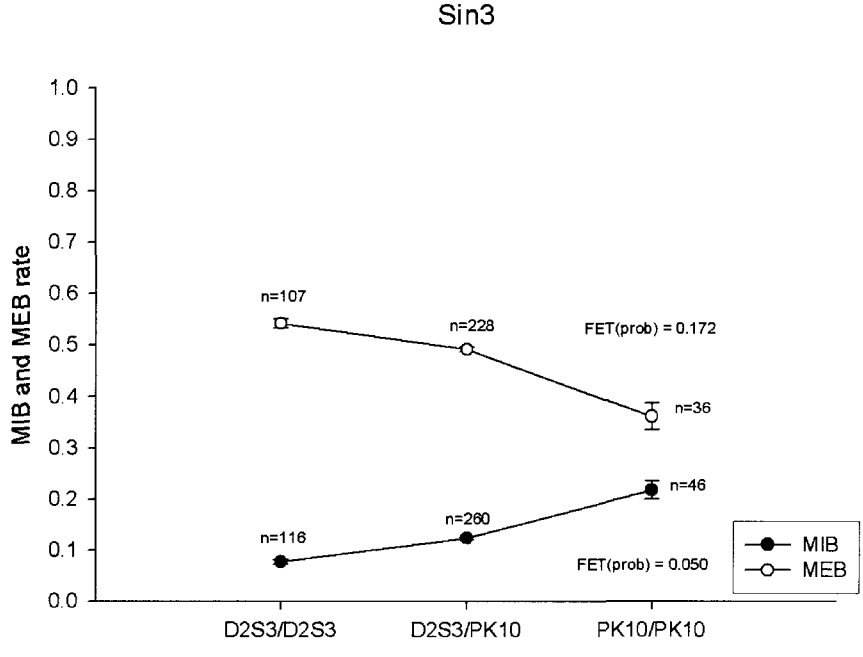


Figure 5.7. Plot of MEB as a function of the number of alleles inherited from either the *D2S3* or PK10 P_1 parent in the F_5 D2PK family. Only significant loci, by χ^2–goodness-of-fit, specific to DENV-2 MEB were plotted.

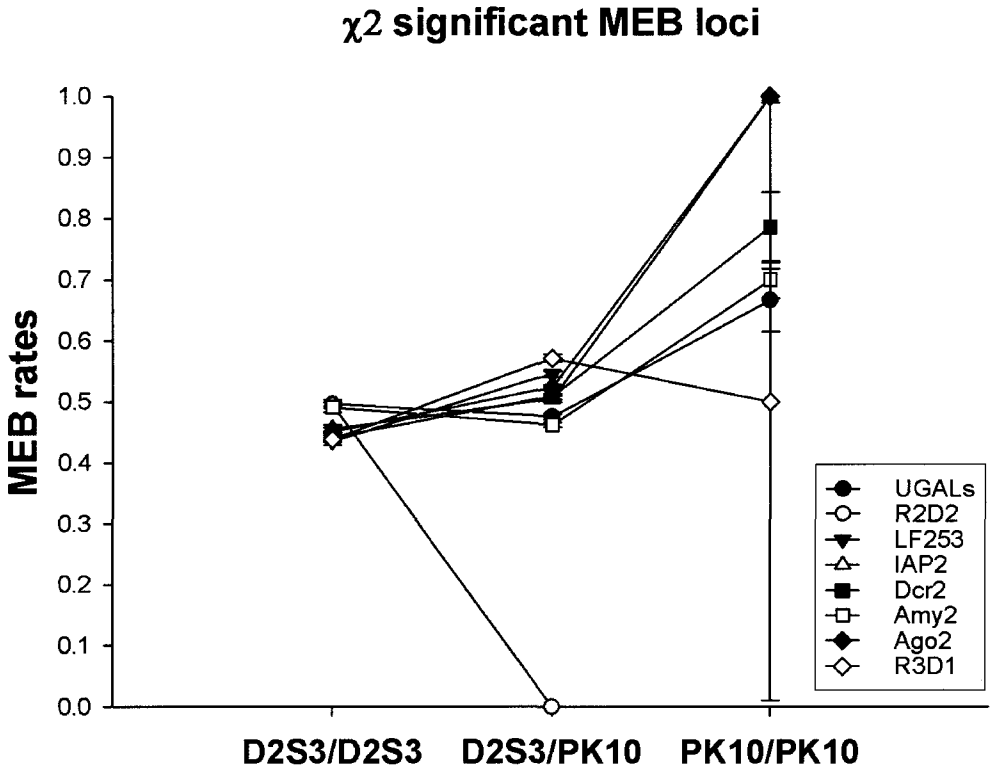


Figure 5.8. Plot of MIB and MEB rates as a function of the UGALs locus alleles inherited from the *D2S3* and PK10 P_1 parents. Plot shows number (n) of mosquitoes with each set of alleles and the P-value for the Fisher exact test (FET) for each allele pair.

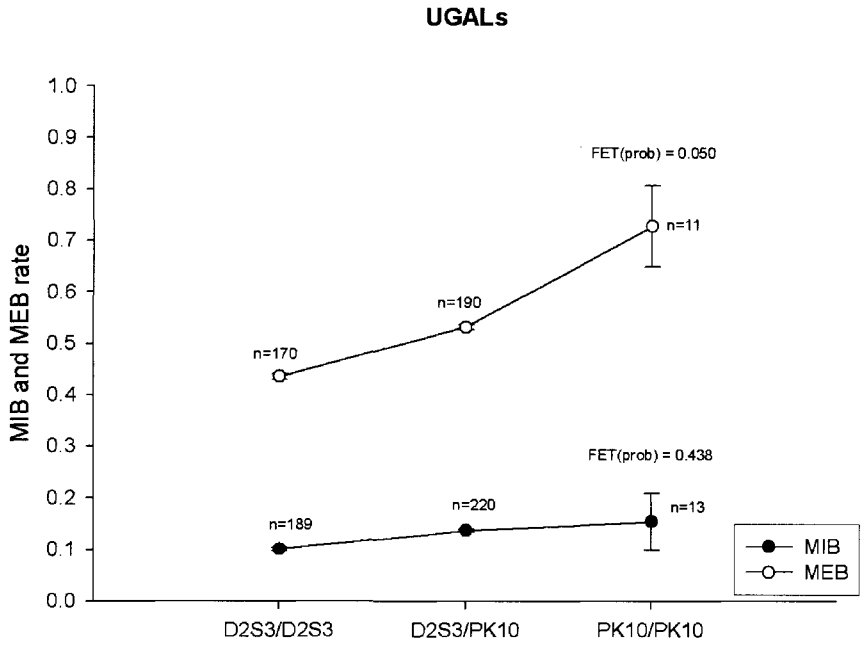


Figure 5.9. Plot of MIB and MEB rates as a function of the *r2d2* locus alleles inherited from the *D2S3* and PK10 P_1 parents. Plot shows number (n) of mosquitoes with each set of alleles and the P-value for the Fisher exact test (FET) for each allele pair.

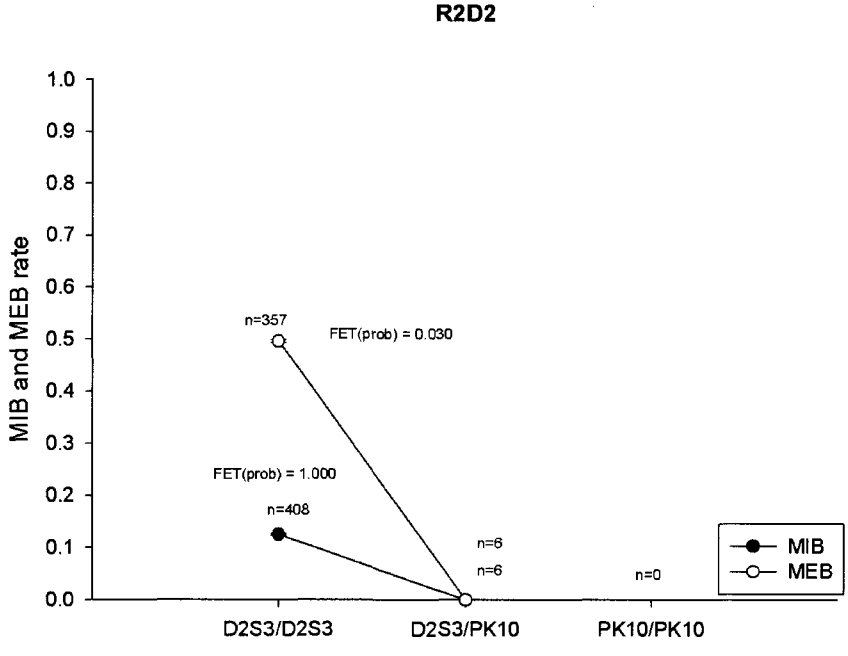


Figure 5.10. Plot of MIB and MEB rates as a function of the LF253 locus alleles inherited from the *D2S3* and PK10 P_1 parents. Plot shows number (n) of mosquitoes with each set of alleles and the P-value for the Fisher exact test (FET) for each allele pair.

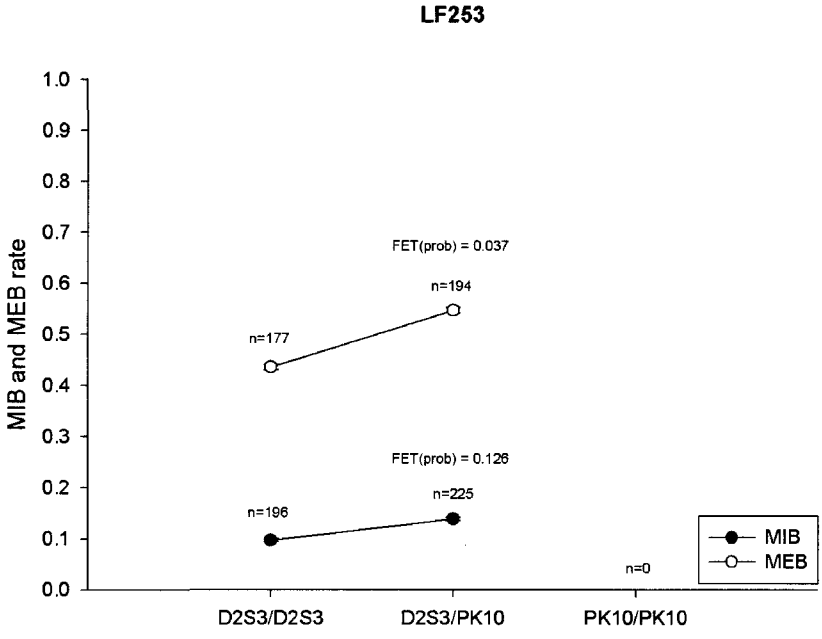


Figure 5.11. Plot of MIB and MEB rates as a function of the IAP2 locus alleles inherited from the *D2S3* and PK10 P_1 parents. Plot shows number (n) of mosquitoes with each set of alleles and the P-value for the Fisher exact test (FET) for each allele pair.

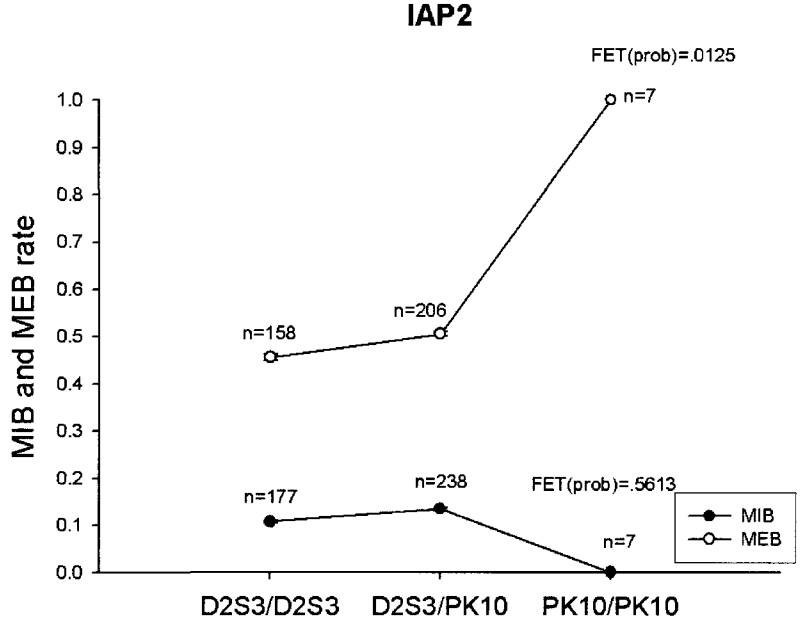


Figure 5.12. Plot of MIB and MEB rates as a function of the *dcr2* locus alleles inherited from the *D2S3* and PK10 P_1 parents. Plot shows number (n) of mosquitoes with each set of alleles and the P-value for the Fisher exact test (FET) for each allele pair.

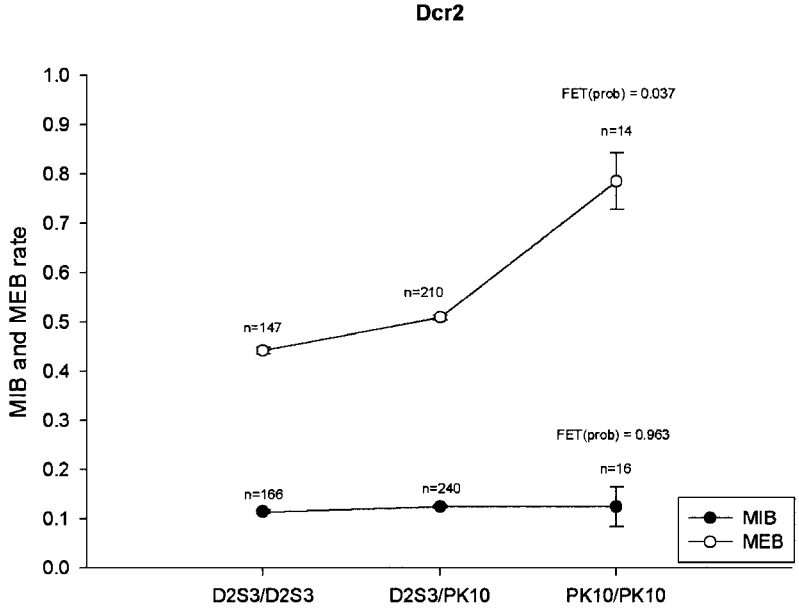


Figure 5.13. Plot of MIB and MEB rates as a function of the Amy2 locus alleles inherited from the *D2S3* and PK10 P_1 parents. Plot shows number (n) of mosquitoes with each set of alleles and the P-value for the Fisher exact test (FET) for each allele pair.

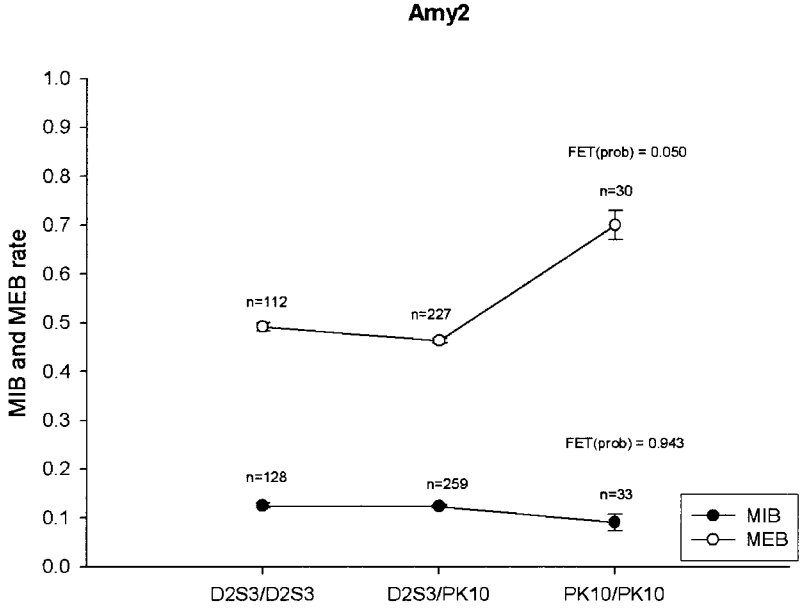


Figure 5.14. Plot of MIB and MEB rates as a function of the *ago2* locus alleles inherited from the *D2S3* and PK10 P_1 parents. Plot shows number (n) of mosquitoes with each set of alleles and the P-value for the Fisher exact test (FET) for each allele pair.

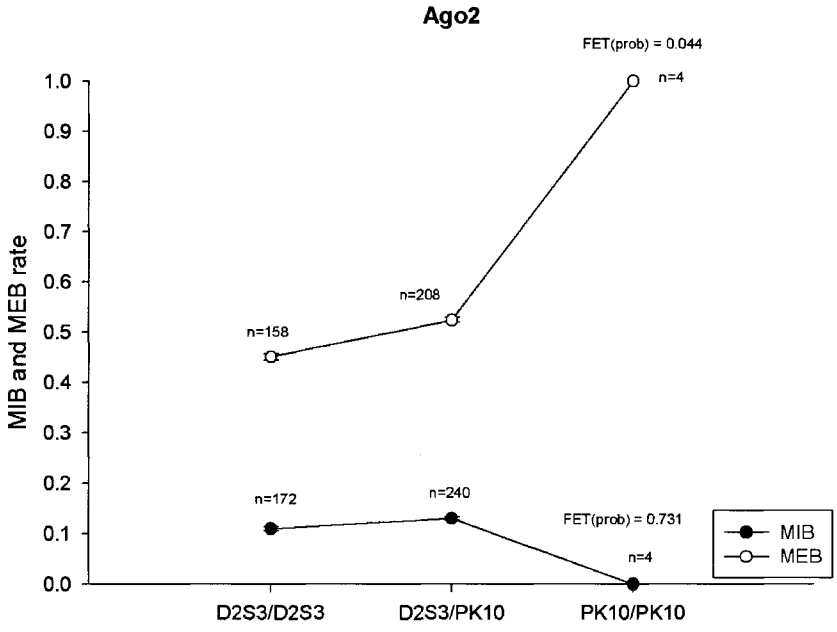


Figure 5.15. Plot of MIB and MEB rates as a function of the *r3d1* locus alleles inherited from the *D2S3* and PK10 P_1 parents. Plot shows number (n) of mosquitoes with each set of alleles and the P-value for the Fisher exact test (FET) for each allele pair.

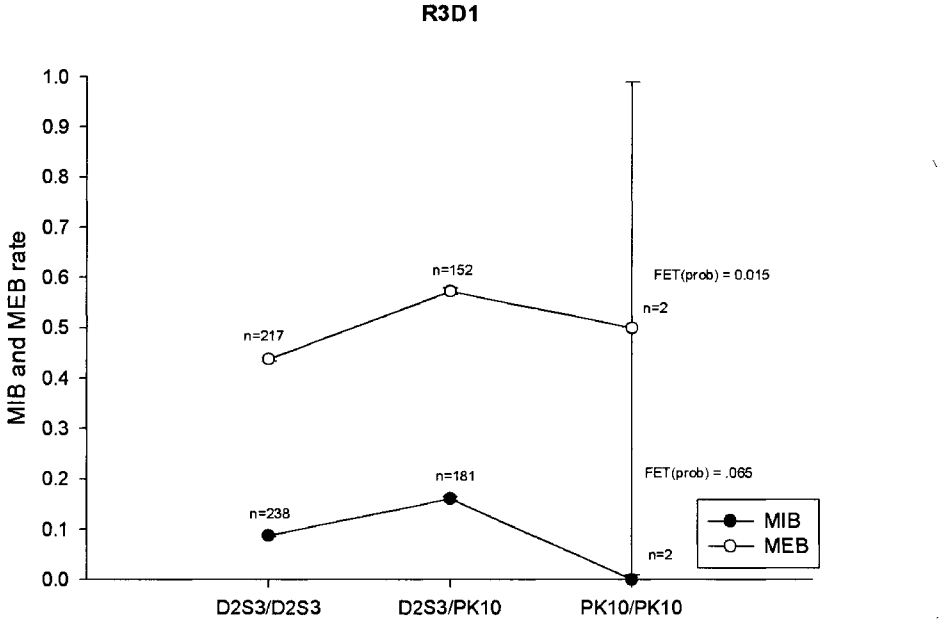


Table 5.3. Additive and dominance effects of loci with significant Fisher exact test value specific to phenotype (MIB, MEB).

Gene	Chromosome	Linkage map position[a] (cM)	MIB gene effect		MIB degree of dominance[b]	MIB gene action[c]	MEB gene effect		MEB degree of dominance[b]	MEB gene action[c]
			Additive	Dominant			Additive	Dominant		
r2d2	1	74	NA	NA	NA	NA	**NA**	**NA**	**NA**	**NA**
Amy2	1	76	-0.0171	0.1406	8.2484	OD	**0.1045**	**-0.1329**	**1.2729**	**OD**
Coe1	1	76	**NA**	**NA**	**NA**	**NA**	NA	NA	NA	NA
r3d1	2	45	-0.0441	0.2043	4.6317	OD	**0.0311**	**0.1035**	**3.3265**	**OD**
Sin3	2	81	**0.0699**	**0.0532**	**0.7607**	**PD**	-0.0905	0.0396	0.4382	PD
UGALs	3	24	0.0267	0.1097	4.1152	OD	**0.1459**	**-0.0497**	**0.3405**	**PD**
LF253	3	25	NA	NA	NA	NA	**NA**	**NA**	**NA**	**NA**
IAP2	3	28	-0.0537	0.1881	3.5051	OD	**0.2721**	**-0.2229**	**0.8194**	**D**
ago2	3	28	-0.0552	0.1886	3.414	OD	**0.2745**	**-0.2014**	**0.7339**	**PD**
dcr2	3	35	0.0053	0.1197	22.7143	OD	**0.1718**	**-0.1044**	**0.6079**	**PD**

Abbreviations: a, additive; d, dominant; od, overdominance; pd, partial dominance; NA, not applicable

Values in bold represent significant Fisher exact test statistic specific to MIB and/or MEB

[a]Chromosome position in cM

[b]Absolute value of dominance divided by additive gene effect (d/a)

[c]Mode of gene action determined as the absolute value of d/a where additive = 0.0-0.20, partial dominance = 0.21-0.80, dominance = 0.81-1.20 and overdominance => 1.20 (after Stuber *et al.*, 1987; Babu *et al.* 2006).

Chapter 6

Analysis of evolutionary rates of RNAi and microRNA genes in *Aedes aegypti*

Abstract

RNA silencing has been shown to act as an effective antiviral mechanism and as an innate immune response to protect adult mosquitoes from alphavirus and flavivirus infection and unregulated replication. Questions remain as to the rate at which RNAi pathway genes are evolving and whether arbovirus infections contribute to RNAi evolution in *Aedes aegypti*. Sequence data were obtained from three RNAi pathway genes and their three microRNA paralogs from 94 mosquitoes from five geographically widespread collection sites. Ratios of non-synonymous to synonymous amino acid substitutions and phylogenetic analysis showed that the RNAi genes are evolving faster than their microRNA paralogs. I also hypothesized that we would see different rates of RNAi evolution in *Ae. aegypti* populations that demonstrate different VC rates for DENV-2. Through a phenotype correlation analysis, I found the average number of nucleotide differences in *dcr2* to be significantly correlated to the DENV-2 MEB in *Ae. aegypti*. A strong inverse correlation was also shown for nucleotide diversity (π) in *ago2* and DENV-2 MEB in *Ae. aegypti*. Even though I found the existence of significant correlations between DENV-2 MEB and RNAi pathway gene evolution, my findings are not conclusive regarding relationship of susceptibility to dengue viruses and rates of RNAi pathway gene evolution in the mosquito. Questions remain as to what drives this significant rate of evolution, the overall significance of this evolution and what infection with arboviruses may contribute to evolution.

Introduction

It is generally believed that a co-evolutionary arms race exists between viruses and their arthropod hosts (Galiana-Arnoux *et al.*, 2006; Aliyari *et al.*, 2008; Reese *et al.*,

2008; Dawkins and Krebs, 1979). Understanding how arboviruses such as dengue and yellow fever viruses evade the mosquito vector innate immune system and cause persistent infections is an important goal. Researchers are beginning to better understand this balance between the ability of the virus to maintain its natural transmission cycle while the host limits unrestricted replication.

The availability of a draft sequence of the *Anopheles gambiae* genome has allowed innovative research approaches and has accelerated the development of new mosquito and malaria control techniques (Holt *et al.*, 2002; Coluzzi *et al.* 2002). An unassembled draft genome sequence of *Aedes aegypti* has been completed (Nene *et al.*, 2007). The genome is approximately 1376 million base pairs and is about 5 times the size of the genome of *An. gambiae*. Due to this immense genome size and content of transposable elements (i.e. ~47% of the genome) and other "non-coding" sequences, it has been difficult to assemble the *Ae. aegypti* scaffolds and to develop physical maps of gene loci (Nene *et al.* 2007). Even with this inaccessible nature, the availability of significant genome sequence data has allowed researchers to understand quantitative traits such as dengue vector competence (VC) (Mercado-Curiel *et al.*, 2006; Bosio *et al.*, 2000, Bennett *et al.*, 2005; Gomez-Machorro *et al.*, 2004). Researchers are able to acquire sequence data to identify potential candidate genes that may be studied and manipulated to determine mechanisms of protein function (Saavedra-Rodriguez *et al.*, 2008; Gorrochotegui-Escalante *et al.*, 2005; Brackney *et al.*, 2008; Mercado-Curiel *et al.*, 2008).

RNA silencing has been shown to act as an effective antiviral mechanism and as an innate immune response to protect adult mosquitoes from alphavirus and flavivirus

infection and unregulated replication (Sanchez-Vargas *et al.*, 2009; Campbell *et al.*, 2008; Keene *et al.*, 2004). Cirimotich *et al.* (2009) and Myles *et al.* (2008) demonstrated that the RNA interference (RNAi) pathway in *Ae. aegypti* is active at suppressing unrestricted alphavirus replication. They also demonstrated that viral replication and mosquito mortality increased if production of exogenous virus-derived small interfering RNAs (siRNAs) were suppressed. Sanchez-Vargas *et al.* (2009) showed that persistent dengue 2 virus (DENV-2) infections of *Ae. aegypti* cells and mosquitoes generate long dsRNA and DENV-2 specific small RNA that are consistent in size and sequence with siRNAs. When they impaired the RNAi pathway in mosquitoes infected with DENV-2, titers of infectious virus also increased while the extrinsic incubation period decreased.

In Chapters 4 and 5, we described the construction of an F_1 intercross family of *Ae. aegypti* to develop a linkage map containing genes from the RNAi and microRNA pathways. The goal was to determine whether these genes mapped to QTLs specific to the DENV-2 midgut infection barriers (MIB) and/or midgut escape barriers (MEB). We found through genotype-phenotype analysis that the loci that contained the RNAi pathway genes (*ago2*, *dcr2*, and *r2d2*) and one microRNA pathway gene (*r3d1* or loquacious) correlated with DENV-2 dissemination within the *Ae. aegypti* mosquito.

Determining the relationship between functional gene properties and evolutionary patterns is one goal of evolutionary genetics. Obbard *et al.* (2006, 2008) determined that the rate of evolution of amino acid sequences is substantially elevated in genes related to antiviral RNAi function in *Drosophila* compared to their paralog genes in the microRNA pathway and to the average rate in the entire genome. They showed that the RNAi genes were among the fastest evolving 3% of all *Drosophila* genes. They hypothesized that the

interactions between the RNAi mechanism in *Drosophila* and viral genes that encode RNAi suppressors could be co-evolutionary and adaptive (Schelenke *et al.*, 2003).

The question is whether these observations in *Drosophila* can also be seen in other dipteran arbovirus-vectors. Advances in the knowledge of genetic and biochemical immune pathways have allowed us to use the *Drosophila* genome *in silico* to identify highly conserved gene orthologs in *Ae. aegypti* (Schlenke *et al.* 2003). The question addressed in this chapter is whether a correlation exists between RNAi pathway gene evolution in *Ae. aegypti* mosquito and VC for DENV-2 (e.g. MIB and MEB). A high rate of evolution in the RNAi pathway does not automatically mean arbovirus infections in mosquitoes are driving the evolutions of the innate immune response.

A number of factors need to be considered when analyzing arbovirus influence on the evolution of the mosquito innate immune response. First and foremost, can arboviruses cause pathogenesis within the mosquito vector and affect the fitness? Early belief was that there was a state where the host (vector) would become increasingly resistant to parasite (arbovirus) infections. Through resistance and attenuation, a benign or commensal relationship would result (Burnet and White, 1972). Recent research has begun to challenge this notion and has found evidence of arbovirus pathogenesis and altered vector fitness such as changes in feeding behavior, survivorship and fecundity due to arbovirus infections (Weaver *et al.*, 1988, 1992; Bowers *et al.*, 2003; Platt *et al.*, 1997; Grimstad *et al.* 1980; Faran *et al.*, 1987; Mahmood *et al.*, 2004). Myles *et al.* (2008) found virus derived small RNAs in mosquitoes that potentially have the ability to modulate mosquito survival and maintenance of the virus in nature. These findings, along with the *Drosophila* data, suggest that the necessary components are in place in the

mosquito to limit virus replication and to maintain mosquito fitness, while continuing a co-adaptation process of arbovirus infection, dissemination and transmission.

We generated sequence data for almost the entire exon regions of three RNAi genes (*ago2*, *dcr2*, *r2d2*) and of three microRNA genes (*ago1*, *dcr1*, *r3d1*) from the *Ae. aegypti* genome from 94 mosquitoes. Three mosquito collections were acquired from Mexico, two collections were made in Senegal, Africa, and one collection was from Thailand. Collection sites were chosen based on their diverse geographic distribution, VC for DENV-2 and subspecies (*Ae. aegypti aegypti* or *Ae. aegypti formosus*). Our goal was to compare the rate of evolution between RNAi and microRNA genes and to determine whether the VC for DENV-2 correlates with rates of evolution within and between mosquito populations within the same species.

We hypothesized that the evolutionary rates in the RNAi pathway genes are significantly higher compared the microRNA pathway genes, regardless of the mosquito population. The logic is that RNA viruses are quickly evolving (e.g. ranging from 4.55 x 10^{-4} (DENV-1) to 9.01 x 10^{-4} (DENV-3) substitutions per site, per year (Twiddy *et al.*, 2003)) and thus maintaining the ability to infect and replicate within the mosquito. The fitness of the mosquito therefore could be compromised if virus replication is not limited. We understand from previous research that the RNAi machinery acts as part of the innate immune pathway to protect adult mosquitoes from alphavirus and flavivirus infection and unregulated replication. It would be necessary for the mosquito to maintain high rates of evolution in specific RNAi pathway genes to effectively recognize viruses and maintain fitness. In this study we found that the rate of amino acid evolution is substantially elevated in genes related to RNAi function compared to their paralogs in the microRNA

pathway. These findings are similar to those recorded in an inter-species study with *Drosophila* (Obbard *et al.* 2006).

We also hypothesize that the rates of RNAi pathway gene evolution are significantly correlated to DENV-2 VC in *Ae. aegypti*. VC is a quantitative trait and is highly variable in different populations (Chapter 2). We found, in Chapter 5, that the RNAi pathway genes significantly contribute to DENV-2 VC in *Ae. aegypti*. We therefore hypothesize that populations that have low VC will have high rates of RNAi gene evolution. Confirmation of this hypothesis would suggest that high rates of RNAi evolution will facilitate virus recognition and significantly contribute to MEB. Through a phenotype correlation analysis with RNAi gene evolution variability in the different collection sites, we found that the average number of nucleotide differences in Dicer 2 is significantly correlated to the DENV-2 MEB in *Ae. aegypti*. We also found a strong inverse correlation with nucleotide diversity in *ago2* and DENV-2 MEB in *Ae. aegypti*. Interestingly, we found no significant correlations between RNAi gene evolution variability and DENV-2 MIB in *Ae. aegypti*.

Materials and methods

Mosquito strains

Six geographically distinct *Ae. aegypti* populations were used in this study and locations where they were collected are illustrated in Figure 6.1. Three mosquito collections were acquired from Mexico as described in Chapter 2. These three sites were selected based on assayed DENV-2 VC phenotypes and geographical distances. F_2 generation adults were used for DNA isolation. Mosquito eggs from Mindin and PK10 populations in Senegal were collected and transported by Dr. Christopher Bosio

(Colorado State University, Fort Collins) in the summer of 2006. F_2 generation adults were used for DNA isolation. *Ae. aegypti* collected in Pai Lom and Lao Bao, Thailand, were used to generate a laboratory colony by Dr. L.C. Harrington at Cornell University, Ithaca, New York. Eggs were shipped to Colorado State University's Arthropod-borne and Infectious Disease Laboratory for an unrelated study. The Thailand laboratory colony, at unknown generation, was used for DNA isolation.

Mosquito rearing

Mosquito rearing followed the same protocol as described in Chapter 2.

DNA extraction

DNA extraction followed the same protocol as described in Chapter 5, using the Qiagen DNeasy Blood and Tissue Kit (Valencia, CA). DNA was extracted from 104 mosquitoes: 18 from Poza Rica, 18 from Lerdo de Tejada, and 18 from Chetumal, Mexico; 20 from PK10 and 10 from Mindin, Senegal; and 20 from Pai Lom and Lao Bao, Thailand. These numbers were sufficient to provide statistical power while maintaining efficiency.

PCR amplification

F_2 mosquito DNA from Poza Rica, Lerdo de Tejada and Chetumal, Mexico, F_4 mosquito DNA from PK10, Senegal, F_2 mosquito DNA from Mindin, Senegal, and Thailand laboratory colony DNA were amplified by following the PCR protocol described in Chapter 2, to provide samples for SSCP analysis. Primers were designed to amplify entire exon regions of the 6 genes in the study, *argonaute* 1 (*ago1*) and 2 (*ago2*), *dicer* 1 (*dcr1*) and 2 (*dcr2*), *r2d2* and *r3d1*. The genes amplified and primer sets used are

in Table 6.1. A total of 56 amplicons (i.e. *ago1*: 7, *ago2*: 9, *dcr1*: 20, *dcr2*: 14, *r2d2*: 2, *r3d1*: 4) were produced for each of the 104 mosquito DNAs.

Single strand conformational polymorphism gel electrophoresis and scoring

Single strand conformational polymorphism (SSCP) analysis was performed on all PCR products as described in Chapter 2. SSCP gels were scored and sequences were obtained from all PCR products as described in Chapter 2. Gels were scored based on the banding patterns observed for the 56 amplicons from each mosquito in the study. Typically, 47 individual mosquito PCR products were run on each SSCP gel. Unique haplotypes were identified and recorded from each SSCP gel. The unique haplotypes for each of the 56 amplicons were sequenced at least once on both strands on the ABI synthesizer at the Proteomics and Metabolomics facility at Colorado State University.

Haplotype identification, sequence alignment and consensus sequence construction

Individual haplotype sequence data for each of the six genes were analyzed to remove any duplicates. Sequences were aligned using Clustal W2 (http://www.ebi.ac.uk/Tools/clustalw2/index.html). Forward and reverse strand sequences for each unique haplotype were analyzed with SeqMan 2 (DNAStar, Madison, WI) to determine a consensus sequence.

Molecular evolution analysis

Molecular evolution analyses were completed for each of the 6 genes (*ago1*, *ago2*, *dcr1*, *dcr2*, *r2d2*, *r3d1*) in each of the 94 individual mosquitoes. Entire exon sequences for each gene from each mosquito were built from the shorter amplicon sequence data previously obtained. The program DNAsp 4.50.3 (http://www.ub.es/dnasp, Department of Genetics, University of Barcelona, Spain) was

used to determine polymorphic sites (S), nucleotide diversity(π), synonymous (π(s)) or non-synonymous (π(a)) nucleotide sequence diversity, synonymous (*Ks*) versus non-synonymous (*Ka*) amino acid substitutions, ratio of non-synonymous to synonymous amino acid differences (*Ka/Ks*), average number of nucleotide differences (k), Fu and Li's F test statistic, linkage disequilibrium, gene flow and genetic differentiation.

Analysis of polymorphic sites (S) identifies variable nucleotides compared to the rest of the nucleotides from each of the sequences. These polymorphic sites are used to determine nucleotide diversity (π). Nucleotide diversity is average number of nucleotide differences per site between two presumably identical sequences (Nei, 1987). Analysis of synonymous (*Ks*) (i.e. silent) and non-synonymous (*Ka)* (i.e. replacement) mutations are conducted to recognize changes in nucleotides that may directly and indirectly affect amino acid sequence. This is calculated by dividing the number of (non-) synonymous substitutions by (non-) synonymous sites. Synonymous (π(s)) and non-synonymous (π(a)) diversity analysis identifies nucleotide diversity specific to synonymous and/or non-synonymous amino acid substitutions. The ratio of non-synonymous to synonymous amino acid differences (*Ka/Ks*) is commonly used as a measure for sequence evolution. The average number of nucleotide differences (k) is a measure that provides a reasonable estimate of the coefficient of variation (Tajima, 1983). Output was also obtained for the Fu and Li's F* test statistic (Simonsen *et al.*, 1993). This statistic tests the hypothesis that all mutations are selectively neutral and bases the statistic on the number of single nucleotides appearing only once among sequences by (k), the average number of nucleotide differences between pairs of sequences.

Linkage disequilibrium analysis measures the degree of nonrandom associations between nucleotide variants at different polymorphic sites and has the power to detect valid genotype-phenotype associations (Black *et al.*, 2008). Linkage disequilibrium is detected using a Hill and Robertson (1966) correlation coefficient $R_{ijk} = D_{ij}/((p_i(1-p_i))(p_j(1-p_j))(p_k(1-p_k))^{1/2}$. The squared correlation coefficient $(R_{ijk})^2$ was used as a metric of disequilibrium because it ranges from zero (linkage equilibrium) to one (linkage disequilibrium). The degree of linkage disequilibrium was calculated and an estimate of relationship with physical distance was determined by regression analysis. Gene flow and genetic differentiation analysis measure DNA divergence and gene flow among and between populations with respect to specific gene or sequence of interest (e.g. RNAi, microRNA pathway genes). The analysis output includes the average number of synonymous nucleotide differences for each gene between collection sites (Ks), coefficient of gene differentiation (G_{st}) or gene differentiation relative to the total population, and the average number of nucleotide substitutions per site, or nucleotide diversity, between populations (Dxy).

Phenotype correlation analysis

Pearson correlation coefficient and Fisher exact tests were calculated using SAS software (SAS Institute), to perform phenotype correlation analysis. Phenotype data, collected for analyses conducted in Chapters 2 and 4, consisted of DENV-2 MIB and MEB rates from Poza Rica (MIB: 21%, MEB: 18%), Lerdo de Tejada (MIB: 25%, MEB: 36%), and Chetumal, Mexico (MIB: 9%, MEB: 8%), and PK10, Senegal (MIB: 8%, MEB: 76%) collection sites. Phenotype data were not collected for the Thailand colony and were not included in the correlation analysis. Evolutionary analysis components for

each RNAi and microRNA gene that were used in the correlation calculation were average number of nucleotide differences (k), number of synonymous substitutions by synonymous sites (*Ks*), evolutionary rate (*Ka/Ks*) and nucleotide diversity (π). Pearson correlation coefficients and Fisher exact tests were performed for each variable of the six genes to each phenotype (i.e. MIB, MEB).

Functional domain analysis

Amino acid sequences for each of the six genes were acquired from Vectorbase (http://www.vectorbase.org) and individually entered into Pfam web software (http://pfam.sanger.ac.uk/) to recognize functional regions, or domains. Pfam web software is a database that contains a large collection of protein families. Amino acid sequences can be uploaded to recognize functional protein domains. Once functional domains were determined, the sequence data for each of the 94 mosquitoes were reviewed to ensure completeness. Molecular evolution analyses were then completed for each individual domain using DNAsp 4.50.3. This program allows analysis of specific regions instead of the entire sequence. Polymorphic sites, haplotype and nucleotide diversity, synonymous (*Ks*) versus non-synonymous (*Ka)* substitutions, and a Fisher exact test (i.e. to recognize significant differences in the number of variable polymorphic sites within the domain) were performed on each of the domains within each gene.

Phylogenetic analysis

Equally weighted parsimony tree searches were conducted for each gene data matrix using 2,000 random addition tree-bisection-reconnection (TBR) searches in PAUP* version 4.0b10 (Swofford *et al.*, 2001) with a maximum of ten trees held per replicate. Parsimony jackknife analyses were conducted using PAUP* with the removal

probability set to approximately e^{-1} (36.7879%), and "jac" resampling emulated. Two-thousand jackknife replicates were performed with 100 random addition TBR searches (each with a maximum of ten trees) per replicate.

jModeltest version 0.1.1 (Posada *et al.*, 2008) was used to select the best-fit likelihood model for each data matrix using the Akaike Information Criterion (AIC; Akaike *et al.*, 1974). Maximum likelihood (Felsenstein *et al.*, 1973) analyses of nucleotide characters from each of the molecular data matrices were performed as (not infallible; Gaut and Lewis *et al.*, 1995; Sanderson and Kim, 2000) tests for long-branch attraction (Felsenstein *et al.*, 1978). Likelihood analyses were conducted using RAxML version 7.03 (Stamatakis *et al.*, 2006). Given that RAxML only implements GTR Q-matrices for nucleotide characters, more restrictive variants of the GTR matrix were not used when selected by the AIC. Optimal likelihood trees were searched for using 1,000 independent searches starting from randomized parsimony trees with the GTRGAMMA model and four discrete rate categories. Likelihood bootstrap (BS) analyses (Felsenstein *et al.*, 1985) were conducted with 2,000 replicates with ten searches per replicate using the "–f i" option, which "refine[s] the final BS tree under GAMMA and a more exhaustive algorithm" (Stamatakis *et al.*, 2008 and 2009).

Results

Gene sequence determination

We determined almost complete sequence data from 94 *Ae. aegypti* for the exon regions of *ago2*, *dcr2*, *r2d2*, *ago1*, *dcr1*, *r3d1*. Analyses were conducted and compared within each geographic population collection site, as well as an analysis of populations from all collection sites combined.

Evolutionary rates of RNAi and microRNA genes in *Ae. aegypti* mosquitoes

The evolutionary rates of RNAi and microRNA genes were first analyzed (Figure 6.2, Tables 6.2, 6.3, 6.6, 6.7, 6.10, 6.11) by examining the ratio of *Ka*, non-synonymous mutations to *Ks*, synonymous mutations. Fisher exact tests were performed for each gene paralog pair to test for differences ($P<0.05$) in rates of evolution. *ago1* had a *Ka/Ks* ratio of 0.031 while *ago2* had a ratio of 0.412. *dcr1* had a *Ka/Ks* ratio of 0.183 and *dcr2* had a ratio of 0.286. Substantially more non-synonymous nucleotide substitutions were found in *ago2* (58.1%) and *dcr2* (40.5%) than in their paralogs, *ago1* (8.9%) and *dcr1* (37.8%). The *Ka/Ks* ratio for *r3d1* and *r2d2* followed a similar trend (Figure 6.2). *r3d1* had a *Ka/Ks* ratio of 0.117 while *r2d2* had a ratio of 0.147; however, these ratios are not significantly different as determined by Fisher exact test.

Evolutionary rates of RNAi and microRNA genes between *Ae. aegypti* populations

The six RNAi and microRNA genes were also analyzed between populations from each collection site to identify variations. Tables 6.2 and 6.3 focus on collection site data for *ago1* and *ago2*, Tables 6.6 and 6.7 focus on collection site data for *dcr1* and *dcr2* and Tables 6.10 and 6.11 focus on collection site data for *r3d1* and *r2d2*, showing data used for evolutionary analysis.

The evolutionary rate of *ago1* and *ago2*, *Ka/Ks* ratio, for each collection site population is diagrammed in Figure 6.3. The *Ka/Ks* ratios for each of the 5 collection sites followed similar trends as for the populations as a whole and were significantly higher for *ago2* (0.282 - 0.454) compared to *ago1* (0.00 - 0.036). The evolutionary rates of *dcr1* and *dcr2*, *Ka/Ks* ratio, for populations from each collection site are compared in Figure 6.12. The *Ka/Ks* ratios for *dcr2* ranged from 0.158 to 0.294 compared to 0.107 to

0.158 for *dcr1*; however, the *dcr1* and *dcr2 Ka/Ks* ratios in Thailand were almost identical (0.158 - 0.159). The evolutionary rates of *r3d1* and *r2d2*, *Ka/Ks* ratio, for each collection site population are compared in Figure 6.21. Of the 3 gene paralog pairs analyzed, *r3d1* and *r2d2* demonstrated the lowest evolutionary differences for the collections as a whole; however, *r2d2* and *r3d1* demonstrated the most variability of the 3 gene paralog pairs within each site. The *Ka/Ks* ratios for *r2d2* ranged from 0.042 to 0.196 compared to a range of 0.00 to 0.088 for *r3d1*. The *r3d1 Ka/Ks* ratio for Chetumal was actually slightly higher (0.062) than its *r2d2* paralog (0.042). The *r2d2 Ka/Ks* ratios were substantially higher in Poza Rica (0.196) and Thailand (0.098) than the *r3d1* ratios in Poza Rica (0.044) and Thailand (0.00).

DNA polymorphism analysis

DNA polymorphism analysis was conducted for each of the 6 genes in the 94 mosquitoes from all 5 collection sites (Figures 6.4, 6.8, 6.13, 6.17, 6.22, 6.26). These figures demonstrate the significant differences in sizes of the genes analyzed.. *r2d2* and *r3d1* are 900 bp and 978 bp in length respectively, compared to *dcr2* with a size of 4746 bp and *dcr1* with 6183 bp. Even though gene size is not always associated with nucleotide diversity, *dcr1* and *dcr2*, as well as *ago1* (2376 bp) and *ago2* (2901 bp), have substantial amounts of nucleotide diversity across the genes. *r2d2* and *r3d1* were much more conserved than the other 4 genes. These figures also demonstrate that in certain regions of the genome substantially more DNA polymorphisms exist and in other sections there is almost complete sequence conservation.

Gene flow and genetic differentiation analysis

Gene flow and genetic differentiation were performed for each of the six gene paralog pairs for the five geographical collections. Tables 6.4, 6.5, 6.8, 6.9, 6.12 and 6.13 show the average number of synonymous amino acid differences (Ks) for each gene (*ago1*, *ago2*, *dcr1*, *dcr2*, *r3d1* and *r2d2*) between collection sites, coefficient of gene differentiation (G_{st}) or gene differentiation relative to the total population and the average number of nucleotide substitutions per genome site, or nucleotide diversity (Dxy), between populations. Relatively little nucleotide diversity (Dxy) was observed among collection sites in each of the six genes analyzed. Nucleotide diversity for all genes and comparison between collection sites ranged from 0.002 to 0.009. Gene differentiation (G_{st}) varied within each of the six genes and between each of the five collection sites, ranging from 0.00 to 0.267. G_{st} values were never large for *ago2*, *dcr1* and *dcr2*, suggesting extensive gene flow between mosquito populations. *dcr1* contained the least gene differentiation (0.00 - 0.002). G_{st} values varied among *ago1*, *r3d1* and with *r2d2*, containing the greatest gene differentiation between each of the five population collections (0.002 - 0.267).

Linkage disequilibrium analysis

Linkage disequilibrium analysis was performed for each of the six genes of all 94 mosquitoes from all five collection sites. Figures 6.5, 6.9, 6.14, 6.18, 6.23, 6.27 illustrate the pairwise comparisons and the formula for the regression analysis to discover significant linkages for each of the 6 analyzed genes. Overall, each of the 6 genes demonstrated linkage disequilibrium, or nonrandom association, between nucleotide variants at different polymorphic sites with respect to their physical distance. Every

gene, except *r3d1*, contained alleles that were in complete linkage disequilibrium (R^2 = 1.0).

Phenotype correlation analysis

Phenotype correlation analysis was performed to recognize potential relationships of DENV-2 MIB and MEB in *Ae. aegypti* with evolutionary rate variations in the RNAi and microRNA genes. Pearson correlation coefficient and Fisher exact tests were used to identify significant correlations. No significant DENV-2 MIB correlations were found with the evolutionary analysis rate variables for either RNAi or microRNA genes. A significant correlation was found between DENV-2 MEB and (k), the average number of nucleotide differences in *dcr2* (Figure 6.30). A significant inverse correlation was identified in *ago2* between DENV-2 MEB and π, nucleotide diversity (per site) (Figure 6.31).

Functional domain analysis

Amino acid sequences for each of the six genes were entered into Pfam web software to discover functional regions, or domains. Unique domains were overlaid over the total DNA polymorphism analysis diagram for each of the 6 genes (Figures 6.6, 6.10, 6.15, 6.19, 6.24, 6.28). The *ago1* and *ago2* domains identified were DUF1785 (domain of unknown function), PIWI and PAZ (Figures 6.6, 6.10). The *dcr1* domains identified were helicase C, dsRNA binding domain, PAZ, ribonuclease III #1 and #2 and dsrm (double stranded RNA binding motif) (Figure 6.15). The *dcr2* domains identified were DEAD, helicase C, dsRNA binding motif, PAZ, ribonuclease III #1 and #2 (Figure 6.19). The *r3d1* domains identified were dsrm #1, #2 and #3 (Figure 6.24). The *r2d2* domains identified were dsrm #1 and #2 (Figure 6.28). Some domains, such as those in *r2d2*, are

highly conserved. Most domains that were analyzed had substantial variability in both microRNA and RNAi pathway genes, as seen in *dcr1* and *dcr2*.

The rate of amino acid sequence evolution (*Ka/Ks*) ratio was calculated for each of the domains identified for each of the six genes (Figures 6.7, 6.11, 6.16, 6.20, 6.25, 6.29). Fisher exact tests were also applied to determine significant differences in the number of polymorphic sites within each domain with respect to the rest of the gene. A number of domains, as seen in *ago1* and *r2d2*, were found in which the *Ka/Ks* ratio was zero due to the existence of no non-synonymous (*Ka*) changes.

The PIWI domain in *ago2*, the two ribonuclease III domains in *dcr1* and one ribonuclease III domain in *dcr2* were all significantly different from the rest of the gene in number of polymorphic sites by Fisher exact test. The *Ka/Ks* ratio for the PIWI domain in *ago2* was 0.389 compared to 0.419 for the rest of the gene. The *Ka/Ks* ratio for the ribonuclease III #1 and #2 domains in *dcr1* were 0.231 and 0.122 compared to 0.183 for the rest of the gene. The *Ka/Ks* ratio for ribonuclease III #1 domain in *dcr2* was 0.152 compared to 0.286 for the rest of the gene.

Phylogenetic analysis

Maximum parsimony phylogenetic trees (Figures 6.32, 6.34, 6.36, 6.38, 6.40, 6.42) and maximum likelihood phylogenetic trees (Figures 6.33, 6.35, 6.37, 6.39, 6.41, 6.43) were constructed for each of the six genes using consensus sequences from each of the 94 mosquitoes. Preliminary analyses contained data from only five geographical collections (Poza Rica, Lerdo de Tejada, Chetumal, PK10 and Thailand). Phylogenetic analyses, by both maximum parsimony and maximum likelihood, demonstrate that substantial variation in both RNAi and microRNA pathway genes was observed between

populations. Each mosquito population trended to occupy a separate clade. Numerous exceptions were seen with respect to mosquitoes from Poza Rica, Chetumal, Lerdo de Tejada and Thailand. PK10, Senegal, was the one major exception to these general findings. Maximum likelihood analyses with all 6 genes demonstrated that only PK10, Senegal, regularly segregated to form a distinct clade. It is unclear whether this segregation is a result of geographical diversity or because it is a different *Ae. aegypti* subspecies (*formosus*) compared to the rest of the geographic populations (*aegypti*). For this reason that an additional site, Mindin, Senegal, was added to the study. The Mindin, Senegal, collection contains both *Ae. aegypti* subspecies. We hypothesize that the Mindin, Senegal collection will either separately group based on subspecies (i.e. some with PK10 *formosus,* some with other *aegypti* collections) or they will group based on geographic collection site.

Discussion

Understanding the relationship between functional properties of genes and patterns of evolution can provide insight into potential evolutionary forces (Schlenke and Begun, 2003). Evolutionary forces may be manifested in DNA sequence variation within gene paralogs, between species, and among geographical collections (Tanaka and Nei, 1989; Hughes *et al.*, 1990; Bishop *et al.*, 2000; Stahl and Bishop, 2000). Population genetic data for genes encoding proteins with known functions can provide information on evolutionary forces.

Emerging evidence indicates that RNA silencing functions in several fundamental processes, including development, antiviral defense and maintenance of genomic stability (Tomari and Zamore, 2005; Hannon *et al.*, 2002). RNAi silencing of gene expression can

act through small interfering RNAs (siRNAs) and microRNAs (Wang *et al.*, 2006). RNAi provides an antiviral mechanism and has been recognized as part of the innate immune pathway to protect adult flies from infection and unregulated viral replication (Waterhouse *et al.*, 2007; Campbell *et al.*, 2008; Keene *et al.*, 2004; Franz *et al.*, 2006; Sanchez-Vargas *et al.*, 2009).

Obbard *et al.* (2006, 2009) recently demonstrated that the antiviral RNAi genes were among the fastest evolving of all *Drosophila* genes. Their study design was based on comparing RNAi and microRNA pathway genes with the remainder of the genome among different *Drosophila* species. These findings raised the question whether the RNAi genes are also evolving significantly faster than microRNA pathway genes in *Ae. aegypti* that are vectors for RNA arboviruses. This led us to address whether variations in the RNAi gene evolution rates exist between *Ae. aegypti* populations that varied in DENV-2 VC.

We generated almost complete sequence data for the exons of three RNAi pathway genes (*ago2*, *dcr2*, *r2d2*) and three microRNA pathway paralog genes (*ago1*, *dcr1*, *r3d1*) from 94 *Ae. aegypti* mosquitoes. We showed that the *ago2*, *dcr2*, and *r2d2* genes are evolving faster than their microRNA paralogs (*ago1*, *dcr1*, *r3d1*) (Figure 6.2). Fisher exact tests indicated that the rates of evolution (*Ka*/*Ks*) within *ago2* and *dcr2* are significantly higher than their *ago1* and *dcr1* paralogs. Data from the *r2d2* and *r3d1* paralogs demonstrated a trend toward more rapid evolution of *r2d2* than its paralog *r3d1*, although the differences were not statistically significant.

The purpose of using the intra-species study design with *Ae. aegypti* was to determine whether differences exist in the rate of RNAi pathway evolution between

populations. If evolutionary differences between populations were apparent, we could begin to make inferences about the antiviral response. Our hypothesis was that certain *Ae. aegypti* populations that are more refractory for DENV-2 disseminated infection would also have higher rates of RNAi gene evolution compared to DENV-2 susceptible populations. RNA viruses are known to be rapidly and thus maintaining the ability to infect and replicate within the mosquito (Twiddy et al., 2003). The RNAi machinery acts as part of an innate immune response towards virus infections and has been shown to significantly contribute to DENV-2 VC in *Ae. aegypti*. Genes in this pathway might be evolving at a high rate to effectively recognize the intruding pathogens and maintain mosquito fitness. The relative rates of evolution in the RNAi pathway genes would reflect the variable DENV-2 VC in mosquito populations.

We consistently saw higher *dcr2* and *r2d2* evolution rates in the DENV-2 refractory populations, Lerdo de Tejada and PK10, compared to lower evolution rates in DENV-2 susceptible Chetumal mosquitoes (Figures 6.12 and 6.21). Collection site specific evolution rates for *dcr2* and *r2D2* suggest an inverse trend between geographic populations and the VC phenotype. Through a phenotype correlation analysis with RNAi evolution variables in the different mosquito populations, we found that the average number of nucleotide differences in *dcr2* was significantly correlated to the DENV-2 MEB (Figure 6.30). *dcr2* encodes the RNAi sensor protein that recognizes and cleaves long dsRNA (viral), producing 21-25 bp short interfering RNAs (siRNAs) (Bernstein *et al.*, 2001; Lee *et al.*, 2004). The dsRNA sensor function of Dcr2 needs to recognize highly evolving (viral) dsRNA secondary structures, which potentially explains why higher substitution rates exist in *dcr2* from refractory populations (Figures 6.12).

Figure 6.3 describes the evolutionary rates of *ago1* and *ago2* of each of the five *Ae. aegypti* collections. Consistently high evolution rates in *ago2* were found in each of the five geographic populations. When comparing *ago2* variation between highly refractory and susceptible populations for DENV-2, we saw relatively minor variation in evolution rates. We did find a significant inverse correlation with nucleotide diversity (π) in *ago2* and MEB to DENV-2 (Figure 6.31). Ago2 has been shown to be responsible for target RNA cleavage (Slicer) activity in the RNAi mechanism (Miyoshi et al., 2005). The inverse correlation suggests that it is important for Ago2 to remain evolutionarily constant to effectively contribute to a high DENV-2 MEB in *Ae. aegypti*. Ultimately, it is difficult to interpret *ago2* evolutionary rates since these findings suggest that high nucleotide diversity in *Ae. aegypti ago2* is highly correlated with low MEB (high VC) while high *ago2* evolution rates (*Ka/Ks*) exist with little variation in *Ae. aegypti* populations with differing DENV-2 VC.

Even though the collection sites were selected based on population differences in DENV-2 VC, our findings are not conclusive regarding susceptibility to DENVs and correlations to RNAi pathway gene evolution in the mosquito. It is unknown whether *Ae. aegypti* frequency of exposure to arboviruses is sufficient to influence RNAi evolution. A number of field studies have demonstrated that very few mosquitoes (1.5%-6.9%) within a population in areas endemic for certain arboviruses (e.g. dengue, EEE virus) are actually infected at any given time (Urdaneta *et al.*, 2004; Chow *et al.*, 1998; Chung *et al.*, 2002; Scott *et al.*, 1998). If arbovirus infections were selectively driving RNAi pathway evolution, significantly higher rates of evolution would be expected exist in RNAi pathway genes in collection sites where DENV-2 VC is extremely low (e.g. PK10,

Lerdo de Tejada). Ultimately, it is difficult to understand the impact, or the lack thereof, that other arboviruses (i.e. yellow fever, chikungunya) or mosquito pathogenic viruses contribute to RNAi pathway gene evolution.

Functional protein domains were identified using Pfam for each of the six genes. Nucleotide variability and evolution analyses were conducted (Figures 6.7, 6.11, 6.16, 6.20, 6.25, 6.29) to determine whether specific functional domains undergo different rates of evolution compared to the entire gene. Substitution rates in the PIWI domain (carries out cleavage of single stranded RNA) in *ago2*, the two ribonuclease III domains (i.e. double stranded RNA-specific cleavage) in *dcr1* and one ribonuclease III domain in *dcr2* were found to be significantly different using Fisher exact test. Even though these domains have significantly higher numbers of polymorphic sites, they do not necessarily have different evolutionary rates (*Ka/Ks*). Many of these polymorphisms may be in synonymous sites rather than non-synonymous sites, which would not influence the amino acid sequence and evolution rates. The *Ka/Ks* ratio for the highly polymorphic *ago2* PIWI domain (0.379) was similar to the *Ka/Ks* ratio for the rest of the gene (0.419). This domain is an example of a region with a statistically significant number of polymorphisms that has a similar evolutionary rate (*Ka/Ks*) to the rest of the gene. The *Ka/Ks* ratio for the *dcr2* ribonuclease III domain (0.152) was substantially lower compared to the *Ka/Ks* ratio for the rest of the gene (0.286). This is an example of a domain with statistically significant polymorphisms that has substantially lower evolutionary rate (*Ka/Ks*) to the rest of the gene.

Phylogenetic analyses (maximum parsimony, maximum likelihood) were performed to further understand the evolutionary relationships of RNAi and microRNA

genes from populations at the six collection sites. PK10, Senegal, mosquitoes formed distinct clades from the other collections. The Thailand and Mexico collections formed multiple blended clades (Figures 6.33, 6.35, 6.37, 6.39, 6.41, 6.43). PK10, Senegal, *Ae. aegypti formosus* was included in this study based on its very low VC for DENV-2 (16%). As phylogenetic data were obtained it became evident that this subspecies population was an outlier. Therefore, the Mindin, Senegal, population was added to the study because it contains both subspecies and thus would allow testing for subspecies associations. Integrating this population into the analysis will assist in determining whether the unique phylogenetic PK10, Senegal, clade was the result of geographic isolation or a different subspecies.

Limited research has been conducted to determine what other mechanisms exist (e.g. environmental pressures) that could selectively drive RNAi gene evolution rates and/or impact variation in VC. One approach would be to examine evolutionary pressures from viruses known to be pathogenic to mosquitoes. Information concerning pathogenic mosquito viruses is limited. *Ae. aegypti* has been shown to be susceptible to cypovirus (Reoviridae) infections as larvae and maintain persistent infections through adulthood. Nodamura virus (Nodaviridae) also has the ability to infect and produce pathogenesis in *Ae. aegypti* (Tesh, 1980; Venter and Schneemann, 2008). It has a bipartite, positive-strand RNA genome that develops dsRNA intermediates during the replication cycle. Unique features of this virus are that the replication complex is sequestered in spherules during replication and it encodes the B2 protein, a viral suppressor of RNAi. This protein has been demonstrated to protect the virus from the host RNAi response and enhance replication and pathogenesis (Johnson *et al.*, 2004;

Aliyari *et al.*, 2008; Li *et al.*, 2002). To date, no known RNAi suppressors have been found to be encoded by arboviruses. It is unclear what is the level of exposure of *Ae. aegypti* to these pathogenic viruses and whether this exposure is sufficient to drive RNAi gene evolution. Questions also remain as to how many undiscovered pathogenic viruses infect mosquito populations. Little research is being conducted in this area and if one is not looking for unidentified causes of pathogenesis, it is difficult to evaluate potential evolutionary impact. Much has yet to be discovered and understood.

In summary, ratios of non-synonymous to synonymous amino acid substitutions and phylogenetic analysis showed that the RNAi genes are evolving faster than their microRNA paralogs. I also demonstrated, through phenotype correlation analysis, the average number of nucleotide differences in *dcr2* to be significantly correlated to the DENV-2 MEB in *Ae. aegypti*. A strong inverse correlation was also shown for nucleotide diversity (π) in *ago2* and DENV-2 MEB in *Ae. aegypti*. Even though I found the existence of significant correlations between DENV-2 MEB and RNAi pathway gene evolution, my findings are not conclusive regarding relationship of susceptibility to dengue viruses and rates of RNAi pathway gene evolution in the mosquito. Questions remain as to what drives this significant rate of evolution, the overall significance of this evolution and what infection with arboviruses may contribute to evolution.

Figure 6.1. Collection sites of *Ae. aegypti* used in molecular evolution study. DNA was extracted from mosquitoes collected from Poza Rica MX, Lerdo de Tejada MX, Chetumal MX, PK10 Senegal, Mindin Senegal and Thailand.

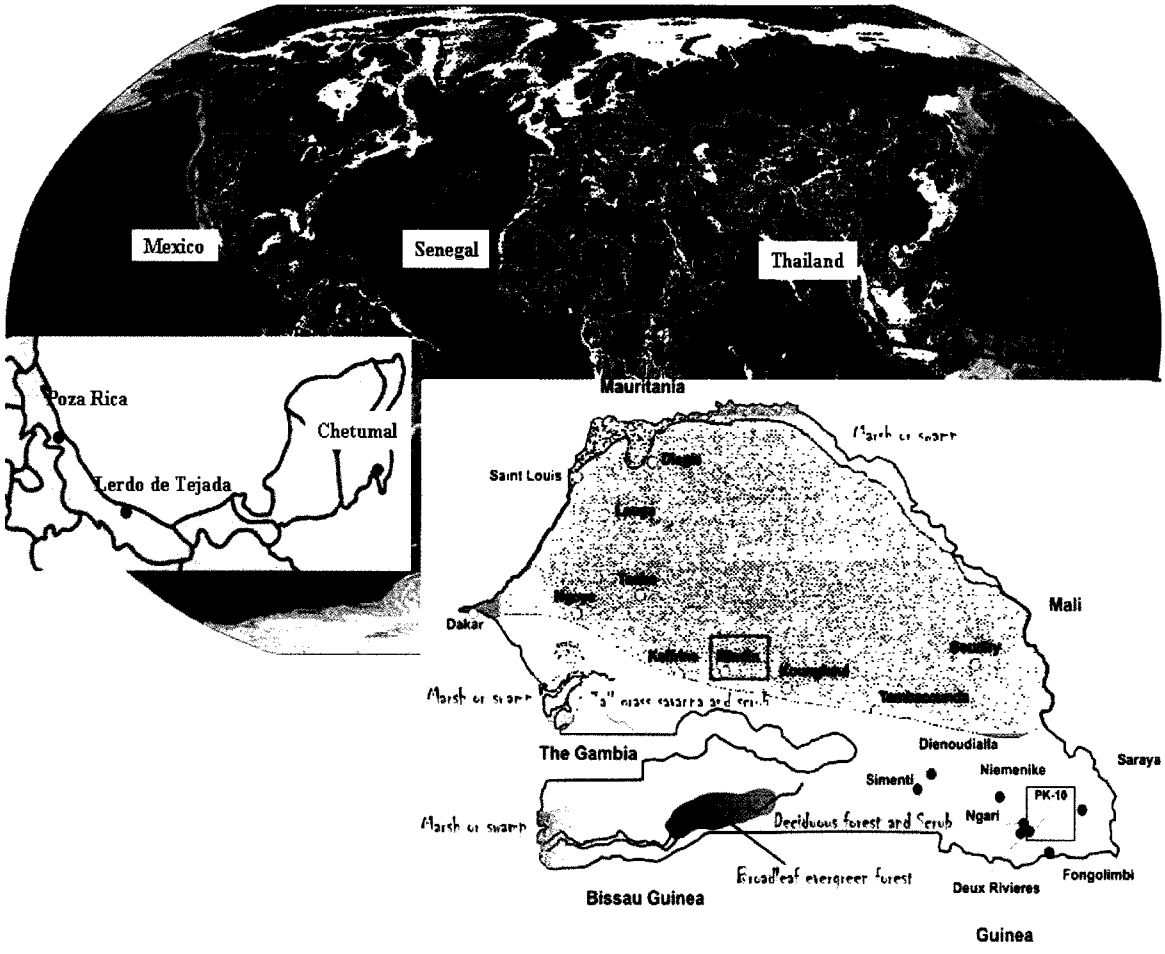


Table 6.1. SSCP primer sequences, annealing temperature (Ta), and anticipated PCR product size for each amplicon specific to the gene in *Ae. Aegypti.*

Gene	Amplicon ID	Forward Primer	Reverse Primer	Start location	PCR product size	T_a
Ago1	Ago1#1	5'-TCGTGCTGCGTGCCAATCACTTCCA-3'	5'-AATRACTTAYCCATCGGGCGAACTG-3'	-32	455	55.8
	Ago1#2	5'-CCTGGGCGGAGGTCGTG-3'	5'-AGGAAATAGAATCAAGAAGGGGAGT-3'	486	440	55.8
	Ago1#3	5'-GCTTCCCTCTGCAGCTAGAAA-3'	5'-GCGTCCTCCGTACTGTAACTTC-3'	927	455	53.0
	Ago1#4	5'-GTGGCACGTTAAAATCAGC-3'	5'-TACCAGGAACACCCCAAT-3'	13,584	402	53.0
	Ago1#5	5'-TTCAACAGCGGAAGTCAA-3'	5'-GCAACTCCCGAACCATAC-3'	14,006	428	53.0
	Ago1#6	5'-GCAGCAGCACCGCCAGG-3'	5'-CTTTCCCCGCATCAAACTCA-3'	14,378	403	55.8
	Ago1#7	5'-GGTTGTAACATCATTTGC-3'	5'-CTTTAAGCGAAGTACA-3'	21,712	405	48.2
Ago2	Ago2#1	5'-ACTGTATGATACTAAACGC-3'	5'-CGACGGTGACGACGAAT-3'	-5	428	50.4
	Ago2#2	5'-GCATTCGTCGTCACCGTCGCA-3'	5'-AGCCGCATAGGCATTTTT-3'	404	337	54.4
	Ago2#3	5'- TAGCGAGTTCACCAAGC-3'	5'-AACGACTAAGGTTATCAT-3'	688	345	48.2
	Ago2#4	5'-TTCATATTTCTCTCTATTGC-3'	5'-AGGATACCGAAGTTGTTTGT-3'	24,092	427	48.2
	Ago2#5	5'-CAACTTCGGTATCCTTCT-3'	5'-TTCCCGTCTTGTAATCTCC-3'	24,405	399	50.4
	Ago2#6	5'-GTACGGAGATTACAAGACGGGAAT-3'	5'-CCAACAAAACTTACCTTGAGCCA-3'	24,781	412	58.5
	Ago2#7	5'-GCATTTTACAGATCAACGCCA-3'	5'-ACGAAGTTCTATGGTCAGTA-3'	42,910	368	48.2
	Ago2#8	5'-CTGACCATAGAACTTCGTGC-3'	5'-GCTACTCACTCTTTGATGTAGACGC-3'	43,260	428	53.0
	Ago2#9	5'-CGTCCTCTGAACATGAACAACCT-3'	5'-AGTATCCTAGAGCCAACAA-3'	43,755	389	48.2
Dcr1	Dcr1#1	5'-GTCCTACGACGAGAATGGCTTAC-3'	5'-ATCACCAGCAGATTTACTT-3'	-14	430	53.0
	Dcr1#2	5'-TTGTGGCTATTTGGATCT-3'	5'-AACCTGTTAGTGGACCTTA-3'	374	292	48.2
	Dcr1#3	5'-ATTTGCTACCAGCCCTAA-3'	5'-CATTGGAATACTGTTGGAAA-3'	16,011	444	48.2
	Dcr1#4	5'-TTTTCCAACAGTATTCCAAT-3'	5'-CATTCTGGTTGTAATAGTTTG-3'	16,434	453	48.2
	Dcr1#5	5'-CAACCAGAATGATCCAGATGCTTTA-3'	5'-CATGCGTTCGCTCAGATCCTCAC-3'	16,876	448	53.0
	Dcr1#6	5'-CTTCCACCGCATATCATATTCTC-3'	5'-TCTCCATATAGATAGC-3'	17,235	315	48.0

Table 6.1. *Continued.*

Gene	Amplicon ID	Forward Primer	Reverse Primer	Start location	PCR product size	T_a
	Dcr1#7	5'-TCGTTAGCGTAATTTGATAACAC-3'	5'-TCTTACTTACCACAATGTCTTCCT-3'	17,602	347	55.8
	Dcr1#8	5'-GGCTTGCCCATGCCTACGGAGAT-3'	5'-GATGTTACAAGAATAGGTACT-3'	28,916	264	48.2
	Dcr1#9	5'-AAACTCATTTCTTTCCCTCAGATC-3'	5'-TTCCAATCAACGGTAACAACACT-3'	44,076	442	55.8
	Dcr1#10	5'-ACGAGAAAGTGTTGTTACCG-3'	5'-CTTGTAGGAAGAGCC-3'	44,488	448	48.2
	Dcr1#11	5'-GTGAATCGGAAAGGGGTGGCTCT-3'	5'-ATCGCAGCAGATTTGTCA-3'	44,905	428	55.8
	Dcr1#12	5'-ATCTGCTGCGATTGAGGAA-3'	5'-TAGTGTAAACTTGTCATGTATAAAT-3'	45,321	259	48.2
	Dcr1#13	5'-GCCAGTGTTTCTCAACCTGTATT-3'	5'-GTCCCACAACTCCTAATGCGTT-3'	45,813	450	55.8
	Dcr1#14	5'-AACGCATTAGGAGTTGTGGGACG-3'	5'-GGTTGTCGGTCAGATTGTTGAGA-3'	46,241	450	53.0
	Dcr1#15	5'-TCTCAACAATCTGACCGACAACC-3'	5'-CGGAGTTCACTTACCTT-3'	46,668	410	50.4
	Dcr1#16	5'-GTATGCTTGTAGTTTGCTAGTCC-3'	5'-TTCACTTTTAACCATGTAGA-3'	47,079	199	50.4
	Dcr1#17	5'-GACAATGAAGAGGGCGAAAC-3'	5'-AAGGCACTGTAACCATCCAAGA-3'	47,814	425	55.8
	Dcr1#18	5'-TCTTGGATGGTTACAGTGCCTTC-3'	5'-GGTGCCTGAAATACTTGTGG-3'	48,217	297	53.0
	Dcr1#19	5'-GATTTCCACAAGTATTTCAG-3'	5'-CTTACCTTCCACACAGCGTCCA-3'	48,489	306	53.0
	Dcr1#20	5'-ACCGAAGTATGATGGGACC-3'	5'-GTGATATTTCAACGACGTTTGTT-3'	61,861	348	50.4
Dcr2	Dcr2#1	5'-CCGCTTGACAAATTTTTCAG-3'	5'-GCCCATACTCACTTATCCAG-3'	-53	312	48.2
	Dcr2#2	5'-CCATTAACTGAAGGTG-3'	5'-ATAGAGTAACATTCCAGAAAGACCG-3'	16,100	441	48.2
	Dcr2#3	5'-GTGTAATCGGTCTTTCTG-3'	5'-ACGCCAGTCTTAGCATTG-3'	16,509	394	53.0
	Dcr2#4	5'-GCAGTTTGCGAAAGCCTG-3'	5'-AATGAAAGACATCACCCTTCTATCC-3'	17,046	342	48.2
	Dcr2#5	5'-TAAAAAGAATGAAACAAACG-3'	5'-CCTTGGCGGTGAAAAACGGCG-3'	17,450	405	48.2
	Dcr2#6	5'-ATTCCGCCGTTTTTCA-3'	5'-AAATCCTTCCAATGACG-3'	17,830	365	50.4
	Dcr2#7	5'-TCAGACCTCGGCAAACTA-3'	5'-AGAAATTCCTTCCACAGT-3'	26,750	398	50.4
	Dcr2#8	5'-CCGAAATAGAACTTGCTC-3'	5'-CGTAACATAACTTACCCGTAC-3'	27,069	373	48.2
	Dcr2#9	5'-CATTTGCTTTTCTTTTCAGTTCCTA-3'	5'-TGCTTTCCTTTCCGCTCCTTG-3'	38,271	394	53.0

Table 6.1. *Continued.*

Gene	Amplicon ID	Forward Primer	Reverse Primer	Start location	PCR product size	T_a
	Dcr2#10	5'-GAGCGGAAAGGAAAGCAG-3'	5'-CTACGGGTACATCCAAGA-3'	38,648	394	50.4
	Dcr2#11	5'-CAATCTTGGATGTACCCGTAG-3'	5'-GAGGTGGTTGCCAGTCGT-3'	39,021	357	55.8
	Dcr2#12	5'-CAACCACCTCTAGCAACG-3'	5'-ATCTACTTCACGAATATCAA-3'	39,368	411	48.2
	Dcr2#13	5'-CGCACAATGTCCTGAAGC-3'	5'-GATCGGTAATTTCGTGTT-3'	39,699	456	53.0
	Dcr2#14	5'-TTACCGATCAGGTGAAC-3'	5'-AAACGAAACATTACTTAGCAC-3'	38,446	435	50.4
R3D1	R3D1-#1	5'-AAAATCATTTCAGATGAGTA-3'	5'-GACTTACCATCCTTGTCCTG-3'	-14	343	50.4
	R3D1-#2	5'-GAAATATGTCATGTTTAAGG-3'	5'-CGCAGATGGAATAACTTACT-3'	12,624	182	48.2
	R3D1-#3	5'-GTTCAATATGCTCACTATCA-3'	5'-TAAATCCCTACCACC-3'	12,814	357	48.2
	R3D1-#4	5'-AGCGGGAAAACTATTACCAA-3'	5'-CCAGACAAGTACATAGACAT-3'	16,388	360	50.4
R2D2	R2D2-5'	5'-CTTGTGGTGTAGAATAATGG-3'	5'-TAATCATCTCGTTGC-3'	-16	603	48.2
	R2D2-3'	5'-ATACTTTAGATACCTCCCATTGACA-3'	5'-AGTGTAGTAAGCCGAATA-3'	12,697	659	50.4

Figure 6.2. The evolutionary rates of RNAi and microRNA pathway genes from 5 *Ae. aegypti* mosquito populations from Mexico, Senegal and Thailand. Rate of evolution was calculated as the ratio of amino acid sequence evolution (Ka, non-synonymous sites) relative to neutral mutations (*Ks*, synonymous sites). Fisher exact tests were performed for each gene paralog to identify significant differences (P<0.05) in rates of evolution. The number of mosquitoes (n) used in each analysis is shown above each gene bar.

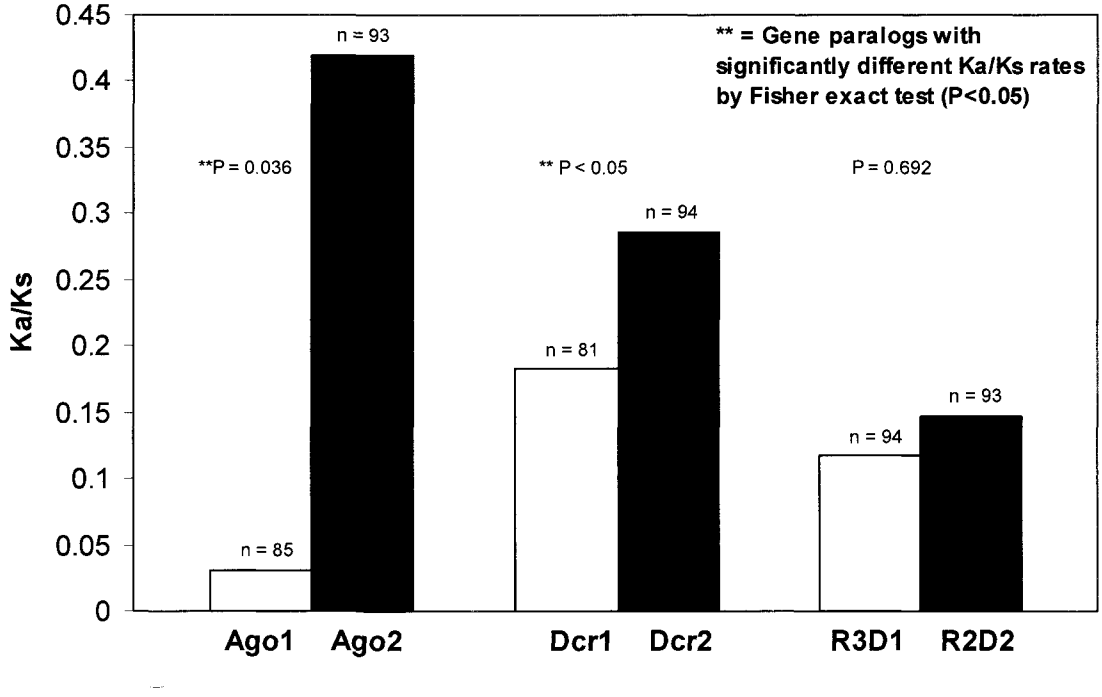


Table 6.2. DNAsp 4.50.3 output for *ago1* containing analysis for all collection sites combined, as well analyzing for each specific collection site. This table shows number of variable sites (S), whether these resulted in synonymous or non-synonymous substitutions, number of haplotypes identified (h), haplotype diversity (Hd), nucleotide diversity (P_i), synonymous ($P_i(s)$) or non-synonymous (P_i (a)) diversity, average number of nucleotide differences (k), output for Fu and Li's F* test statistic, ratio of non-synonymous to synonymous nucleotide differences (*Ka/Ks*) and pairwise comparisons to determine a linkage disequilibrium model.

Site	# of variable sites, S=	Synonymous substitutions	Non-synonymous substitutions	# of haplotypes, h	Haplotype (gene) diversity, Hd	Nucleotide diversity (per site), Pi	Pi(s)	Pi(a)	Pi(a)/Pi(s)	Average # of nucleotide differences, k
All Sites	77	71	7	97	0.978	0.007	0.030	0.000	0.008	17.150
Poza Rica, MX	37	37	1	27	0.984	0.006	0.023	0.000	0.003	13.169
Chetumal, MX	26	26	0	8	0.631	0.002	0.007	0.000	0.000	3.890
Lerdo de Tejada, MX	29	29	0	19	0.932	0.004	0.017	0.000	0.000	9.660
Thailand	42	40	2	21	0.953	0.006	0.026	0.000	0.009	14.682
PK10, Senegal	49	44	5	28	1.000	0.008	0.030	0.001	0.023	17.915

Site	Fu and Li's F* test statistic, F*	Ks	Ka	Ka/Ks	# of pairwise comparisons	# of significant pairwise comparisons by Fisher's exact test	Linkage disequilibrium R^2 model
All Sites	1.368	0.125	0.004	0.031	2850	1194	y = 0.0914 - 0.0273x
Poza Rica, MX	1.572	0.065	0.001	0.008	630	209	y = 0.2375 - 0.1035x
Chetumal, MX	0.210	0.046	0.000	0.000	325	166	y = 0.4110 + 0.0264x
Lerdo de Tejada, MX	**P<05, 1.690**	0.051	0.000	0.000	406	117	y = 0.2313 - 0.0891x
Thailand	**P<.02, 1.873**	0.070	0.001	0.016	861	351	y = 0.2255 - 0.0802x
PK10, Senegal	1.606	0.077	0.003	0.036	1176	157	y = 0.1194 - 0.0383x

Table 6.3. DNAsp 4.50.3 output for *ago2* containing analysis for all collection sites combined, as well analyzing for each specific collection site. This table shows number of variable sites (S), whether these resulted in synonymous or non-synonymous substitutions, number of haplotypes identified (h), haplotype diversity (Hd), nucleotide diversity (P_i), synonymous ($P_i(s)$) or non-synonymous (P_i (a)) diversity, average number of nucleotide differences (k), output for Fu and Li's F* test statistic, ratio of non-synonymous to synonymous nucleotide differences (*Ka/Ks*) and pairwise comparisons to determine a linkage disequilibrium model.

Site	# of variable sites, S=	Synonymous substitutions	Non-synonymous substitutions	# of haplotypes, h	Haplotype (gene) diversity, Hd	Nucleotide diversity (per site), Pi	Pi(s)	Pi(a)	Pi(a)/Pi(s)	Average # of nucleotide differences, k
All Sites	62	26	36	121	0.988	0.005	0.012	0.003	0.222	12.349
Poza Rica, MX	29	15	14	29	0.983	0.003	0.008	0.002	0.221	7.856
Chetumal, MX	31	13	18	30	0.984	0.003	0.007	0.002	0.243	7.371
Lerdo de Tejada, MX	29	12	18	27	0.984	0.003	0.008	0.002	0.264	8.484
Thailand	28	11	17	15	0.863	0.002	0.004	0.001	0.302	5.537
PK10, Senegal	16	7	9	24	0.963	0.002	0.005	0.001	0.259	5.050

Site	Fu and Li's F* test statistic, F*	Ks	Ka	Ka/Ks	# of pairwise comparisons	# of significant pairwise comparisons by Fisher's exact test	Linkage disequilibrium R^2 model
All Sites	1.254	0.045	0.019	0.419	1830	527	y = 0.0829 - 0.0113x
Poza Rica, MX	1.361	0.026	0.007	0.282	406	52	y = 0.1406 - 0.0597x
Chetumal, MX	0.468	0.022	0.009	0.419	465	68	y = 0.1479 - 0.0911x
Lerdo de Tejada, MX	1.069	0.021	0.009	0.454	378	59	y = 0.1285 - 0.0581x
Thailand	0.475	0.017	0.008	0.448	378	52	y = 0.2595 - 0.1471x
PK10, Senegal	**P<.05, 1.664**	0.012	0.005	0.389	120	31	y = 0.2169 - 0.1017x

Table 6.4. *ago1* gene flow and genetic differentiation analysis. Comparisons of the average number of synonymous amino acid differences for this gene (Ks) (Hudson *et al.*, 1992), coefficient of gene differentiation (G_{st}) or gene differentiation relative to the total population (Nei, 1973 Equation 9) and the average number of nucleotide substitutions per site, or nucleotide diversity (Dxy), between pairs of collection populations (Nei, 1987 Equation 10.20).

Population 1	Population 2	Ks	Gst	Dxy
Poza Rica	**Chetumal**	8.531	0.101	0.006
Poza Rica	**Lerdo de Tejada**	11.414	0.021	0.007
Poza Rica	**Thailand**	13.987	0.015	0.006
Poza Rica	**PK10**	15.313	0.004	0.011
Chetumal	**Lerdo de Tejada**	6.776	0.121	0.004
Chetumal	**Thailand**	9.725	0.114	0.006
Chetumal	**PK10**	10.226	0.103	0.010
Lerdo de Tejada	**Thailand**	12.374	0.025	0.007
Lerdo de Tejada	**PK10**	13.388	0.017	0.011
Thailand	**PK10**	16.013	0.012	0.011

Table 6.5 *ago2* gene flow and genetic differentiation analysis. Comparisons of the average number of synonymous amino acid differences for this gene (Ks) (Hudson *et al.*, 1992), coefficient of gene differentiation (G_{st}) or gene differentiation relative to the total population (Nei, 1973 Equation 9) and the average number of nucleotide substitutions per site, or nucleotide diversity (Dxy), between pairs of collection populations (Nei, 1987 Equation 10.20).

Population 1	Population 2	Ks	Gst	Dxy
Poza Rica	**Chetumal**	7.613	0.008	0.003
Poza Rica	**Lerdo de Tejada**	8.170	0.008	0.004
Poza Rica	**Thailand**	6.299	0.040	0.004
Poza Rica	**PK10**	6.418	0.014	0.009
Chetumal	**Lerdo de Tejada**	7.928	0.002	0.003
Chetumal	**Thailand**	6.069	0.040	0.003
Chetumal	**PK10**	6.182	0.013	0.009
Lerdo de Tejada	**Thailand**	6.596	0.031	0.003
Lerdo de Tejada	**PK10**	6.723	0.013	0.008
Thailand	**PK10**	4.974	0.046	0.008

Figure 6.3. The *ago1* and *ago2* evolutionary rates from 5 *Ae. aegypti* mosquito populations from Mexico, Senegal and Thailand. Rate of evolution was calculated as the ratio of amino acid sequence evolution (*Ka*, non-synonymous sites) relative to neutral mutations (*Ks*, synonymous sites). The number of mosquitoes (n) used in each analysis is listed above each gene bar.

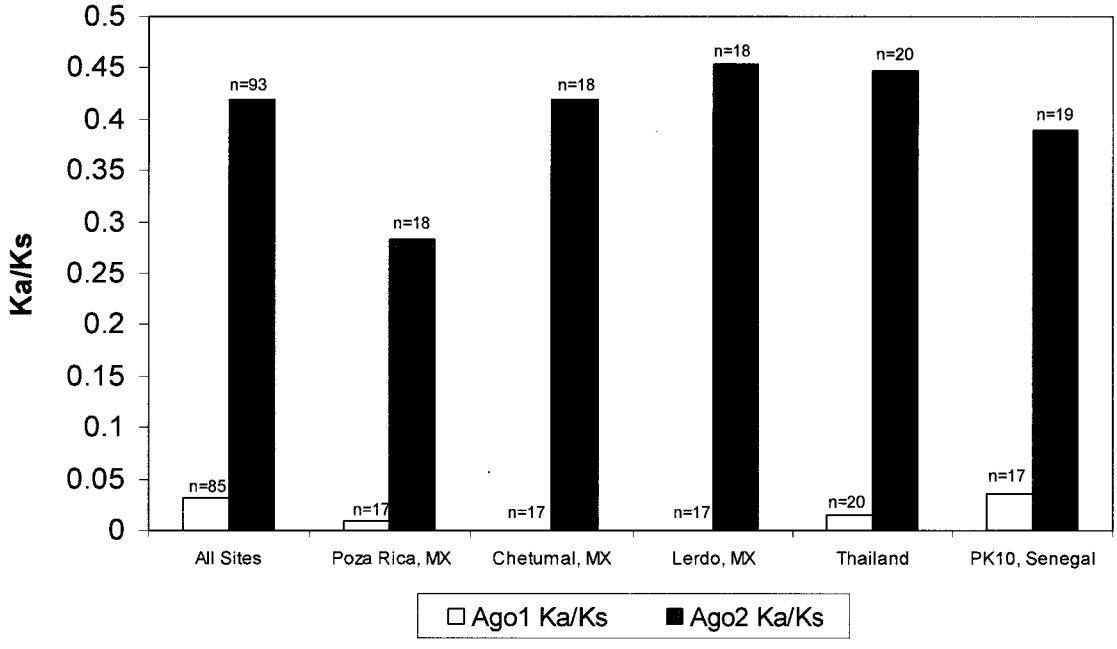


Figure 6.4. *ago1* DNA polymorphism analysis for all 94 mosquitoes from all 5 collection sites. π is the average number of nucleotide differences per site between sequences, or nucleotide diversity (Nei, 1987, equation 10.5).

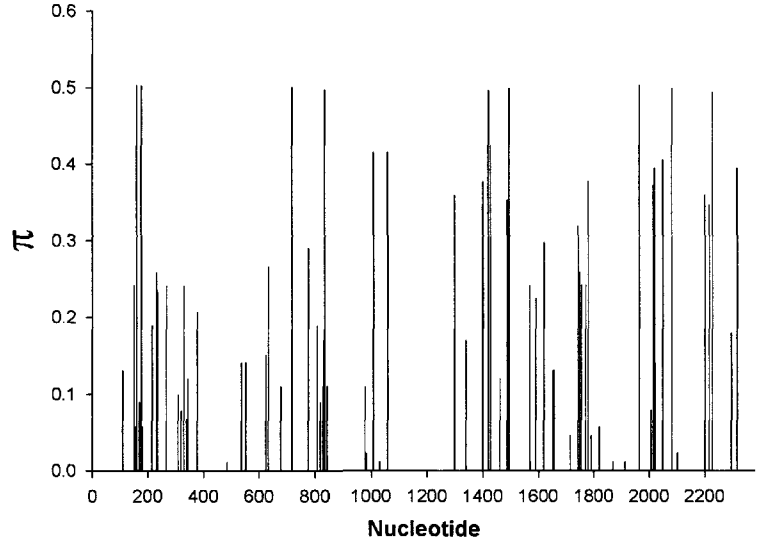


Figure 6.5. *ago1* linkage disequilibrium analysis for all 94 mosquitoes from all 5 collection sites. The degree of linkage disequilibrium, or nonrandom association between nucleotide variants at different polymorphic sites is calculated and an estimate of relationship with physical distance is determined by regression analysis (R^2) (Hill and Robertson, 1966). The linear regression equation is at the top right of the graph and is demonstrated as a red line.

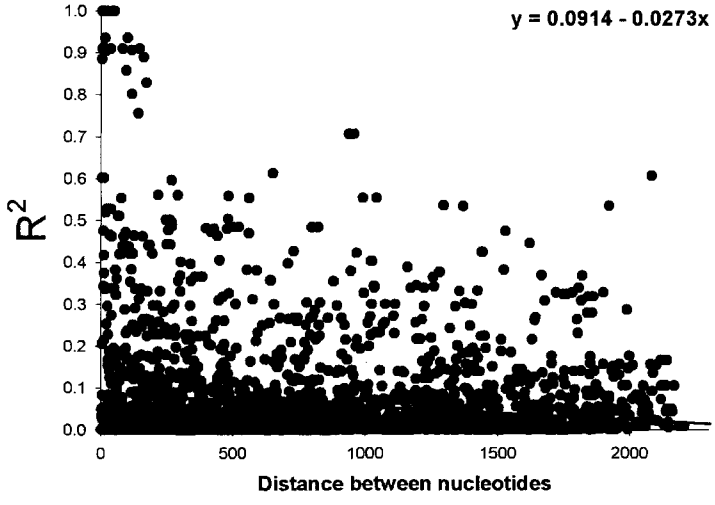


Figure 6.6. *ago1* functional domains identified with Pfam overlaid on the DNA polymorphism analysis for all 94 mosquitoes from all 5 collection sites. π is the average number of nucleotide differences per site between sequences, or nucleotide diversity.

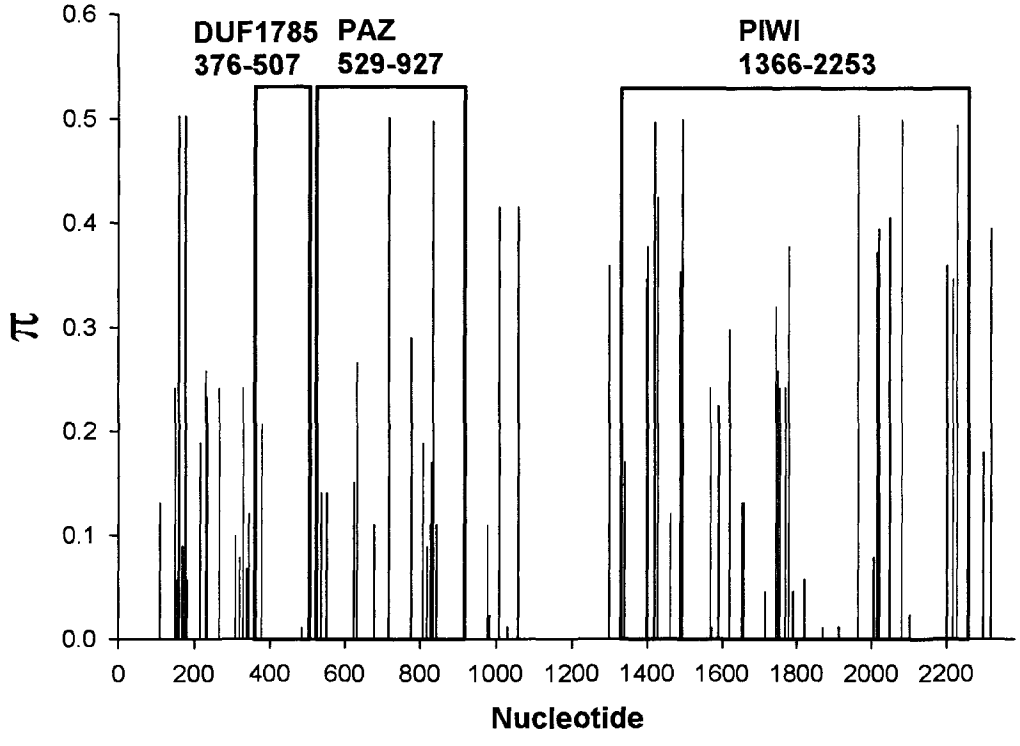


Figure 6.7. *ago1* functional domain evolutionary rate analysis. All 5 *Ae. aegypti* mosquito collections from Mexico, Senegal and Thailand were used in this analysis. Rate of evolution was calculated as the ratio of amino acid sequence evolution (*Ka*, non-synonymous sites) relative to neutral mutations (*Ks*, synonymous sites). Fisher exact tests were performed for each domain with respect to the rest of the gene to identify significant differences ($P<0.05$) in the number of variable polymorphic sites. Domains that were found to be significantly different by Fisher exact test are highlighted by **. No significant differences within the domains were identified.

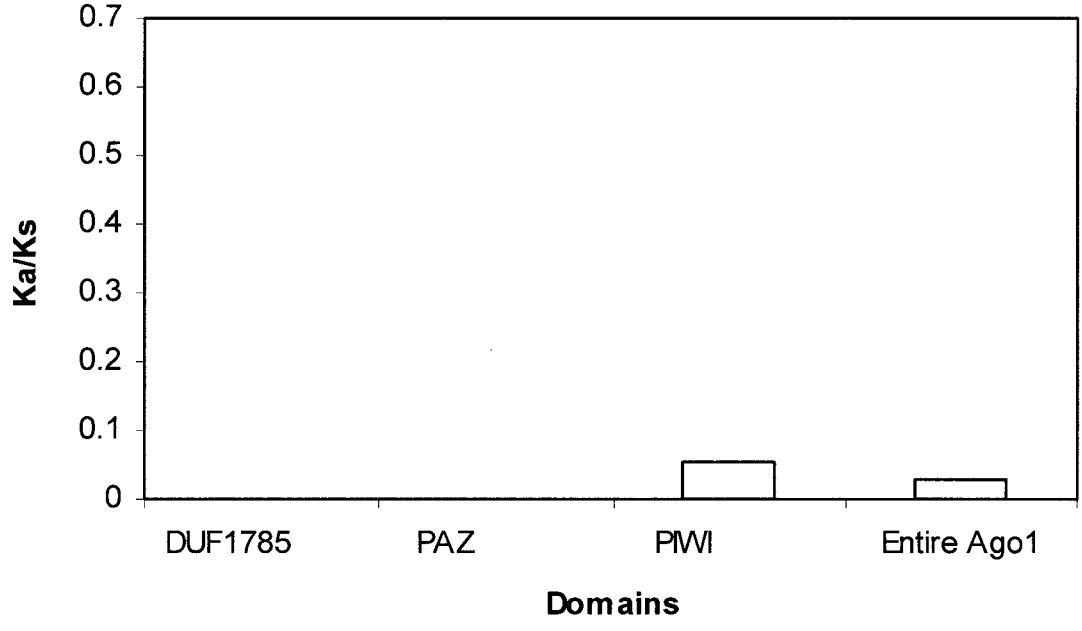


Figure 6.8. *ago2* DNA polymorphism analysis for all 94 mosquitoes from all 5 collection sites. π is the average number of nucleotide differences per site between sequences, or nucleotide diversity (Nei, 1987, equation 10.5).

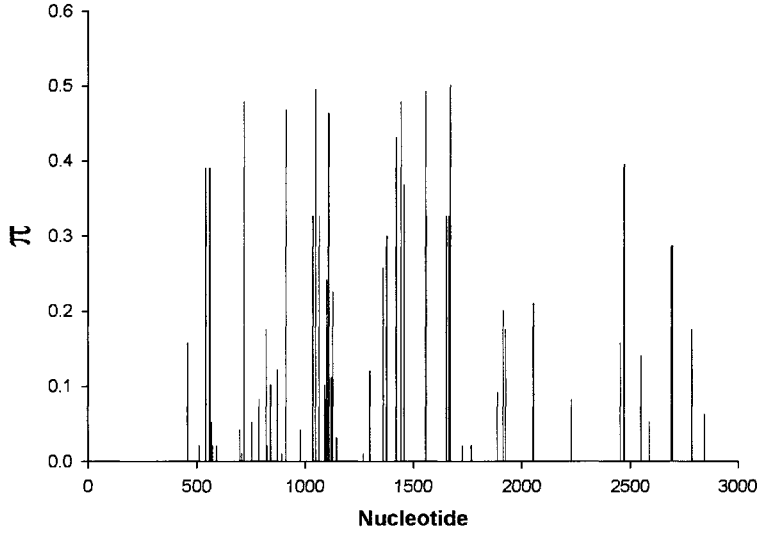


Figure 6.9. *ago2* linkage disequilibrium analysis for all 94 mosquitoes from all 5 collection sites. The degree of linkage disequilibrium, or nonrandom association between nucleotide variants at different polymorphic sites is calculated and an estimate of relationship with physical distance is determined by regression analysis (R^2) (Hill and Robertson, 1966). The linear regression equation is at the top right of the graph and is demonstrated as a red line..

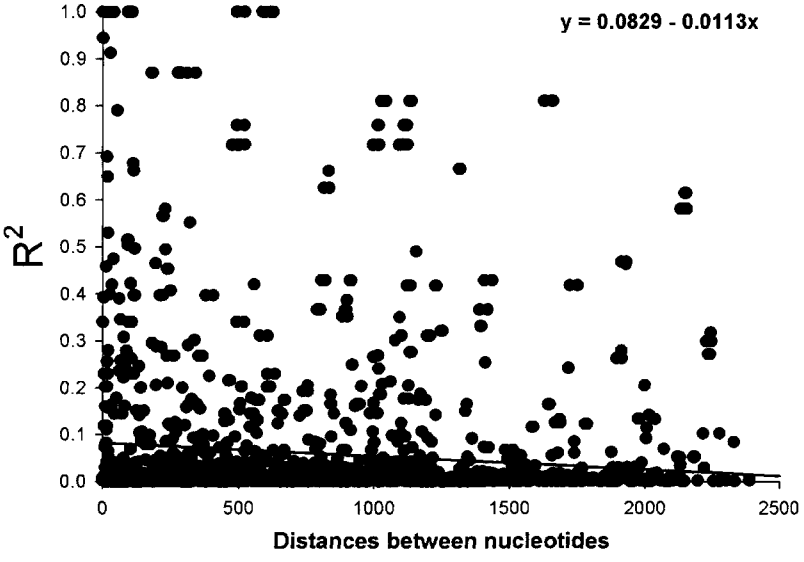


Figure 6.10. *ago2* functional domains identified from web software Pfam overlaid on the DNA polymorphism analysis for all 94 mosquitoes from all 5 collection sites. π is the average number of nucleotide differences per site between sequences, or nucleotide diversity.

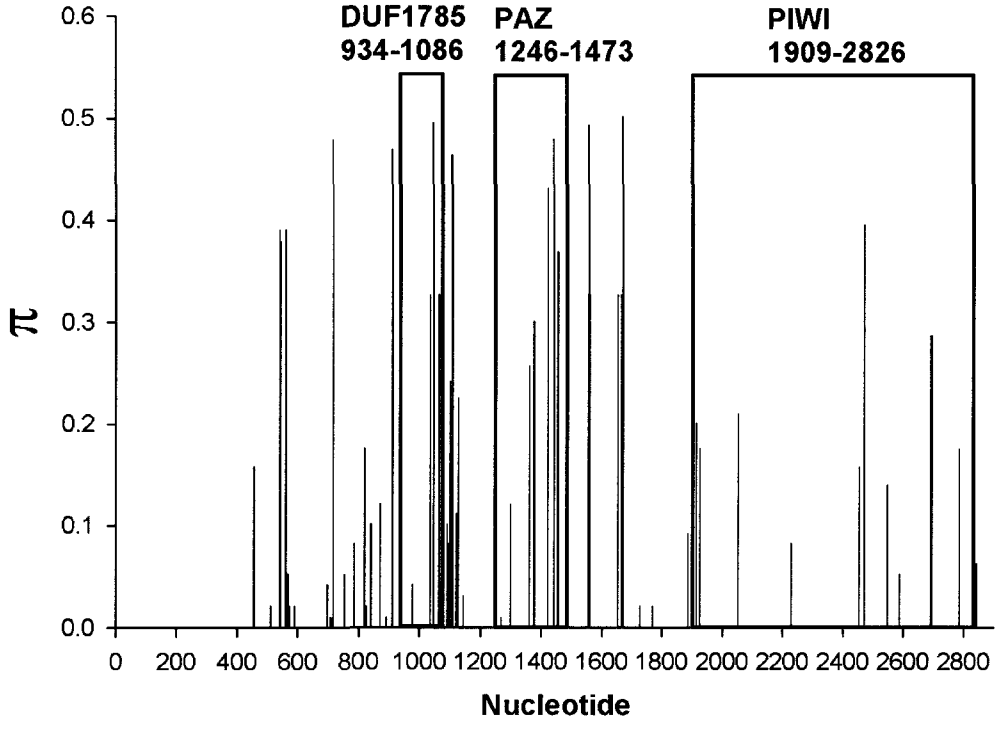


Figure 6.11. *ago2* functional domain evolutionary rate analysis. All 5 *Ae. aegypti* mosquito collections from Mexico, Senegal and Thailand were used in this analysis. Rate of evolution was calculated as the ratio of amino acid sequence evolution (*Ka*, non-synonymous sites) relative to neutral mutations (*Ks*, synonymous sites). Fisher exact tests were performed for each domain with respect to the rest of the gene to identify significant differences ($P<0.05$) in the number of variable polymorphic sites. Domains that were found to be significantly different by Fisher exact test are highlighted by **.

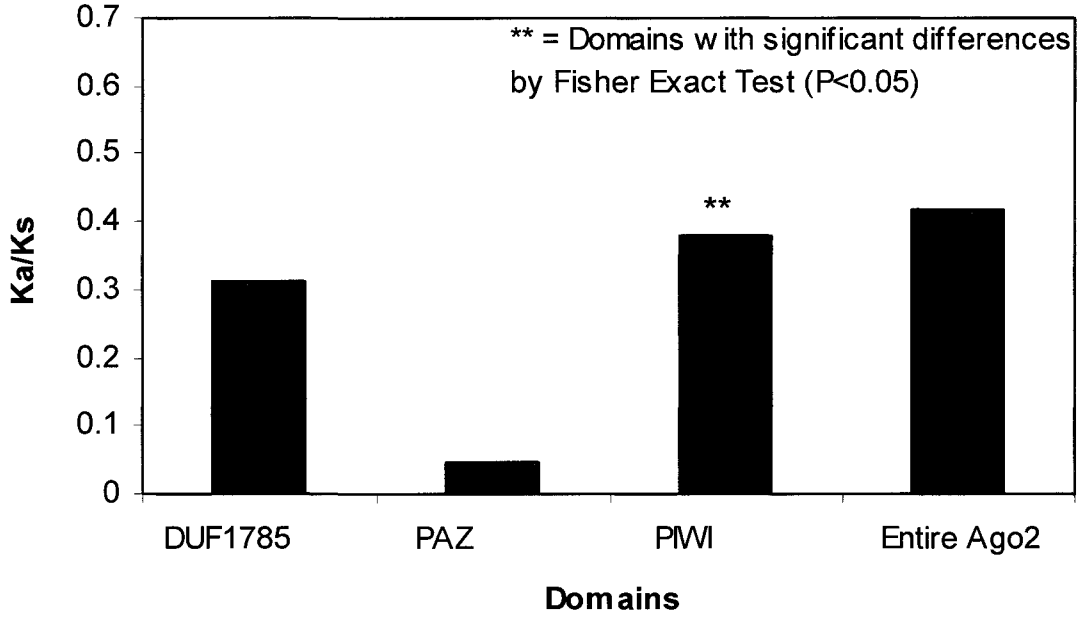


Table 6.6. DNAsp 4.50.3 output for *dcr1* containing analysis for all collection sites combined, as well analyzing for each specific collection site. This table shows number of variable sites (S), whether these resulted in synonymous or non-synonymous substitutions, number of haplotypes identified (h), haplotype diversity (Hd), nucleotide diversity (P_i), synonymous ($P_i(s)$) or non-synonymous (P_i (a)) diversity, average number of nucleotide differences (k), output for Fu and Li's F* test statistic, ratio of non-synonymous to synonymous nucleotide differences (*Ka/Ks*) and pairwise comparisons to determine a linkage disequilibrium model.

Site	# of variable sites, S=	Synonymous substitutions	Non-synonymous substitutions	# of haplotypes, h	Haplotype (gene) diversity, Hd	Nucleotide diversity (per site), Pi	Pi(s)	Pi(a)	Pi(a)/Pi(s)	Average # of nucleotide differences, k
All Sites	196	122	74	159	1	0.006	0.019	0.002	0.113	35.861
Poza Rica, MX	121	79	42	30	1	0.006	0.019	0.002	0.113	35.860
Chetumal, MX	113	82	31	35	0.998	0.005	0.015	0.001	0.091	28.024
Lerdo de Tejada, MX	82	57	25	36	1	0.004	0.013	0.001	0.105	25.249
Thailand	83	54	29	36	1	0.004	0.010	0.001	0.142	21.625
PK10, Senegal	79	58	21	22	0.993	0.004	0.014	0.002	0.112	26.924

Site	Fu and Li's F* test statistic, F*	Ks	Ka	Ka/Ks	# of pairwise comparisons	# of significant pairwise comparisons by Fisher's exact test	Linkage disequilibrium R^2 model
All Sites	0.466	0.085	0.016	0.183	18721	3470	y = 0.0446 - 0.0060x
Poza Rica, MX	1.110	0.056	0.009	0.158	7260	858	y = 0.1055 - 0.0111x
Chetumal, MX	0.816	0.058	0.007	0.112	6328	1067	y = 0.1435 - 0.0153x
Lerdo de Tejada, MX	1.134	0.040	0.005	0.130	3321	483	y = 0.1111 - 0.0138x
Thailand	0.602	0.038	0.006	0.159	3321	491	y = 0.1352 - 0.0181x
PK10, Senegal	1.263	0.041	0.004	0.107	3081	676	y = 0.1825 - 0.0037x

Table 6.7. DNAsp 4.50.3 output for *dcr2* containing analysis for all collection sites combined, as well analyzing for each specific collection site. This table shows number of variable sites (S), whether these resulted in synonymous or non-synonymous substitutions, number of haplotypes identified (h), haplotype diversity (Hd), nucleotide diversity (P_i), synonymous ($P_i(s)$) or non-synonymous (P_i (a)) diversity, average number of nucleotide differences (k), output for Fu and Li's F* test statistic, ratio of non-synonymous to synonymous nucleotide differences (*Ka/Ks*) and pairwise comparisons to determine a linkage disequilibrium model.

Site	# of variable sites, S=	Synonymous substitutions	Non-synonymous substitutions	# of haplotypes, h	Haplotype (gene) diversity, Hd	Nucleotide diversity (per site), Pi	Pi(s)	Pi(a)	Pi(a)/Pi(s)	Average # of nucleotide differences, k
All Sites	138	70	68	157	0.998	0.005	0.013	0.002	0.163	21.978
Poza Rica, MX	56	32	24	25	0.976	0.002	0.007	0.001	0.164	10.921
Chetumal, MX	34	22	12	31	0.992	0.002	0.006	0.001	0.137	10.033
Lerdo de Tejada, MX	60	30	30	30	0.995	0.004	0.009	0.002	0.249	16.984
Thailand	40	26	14	35	0.994	0.003	0.009	0.001	0.111	13.405
PK10, Senegal	72	38	34	34	0.992	0.005	0.014	0.002	0.138	22.187

Site	Fu and Li's F* test statistic, F*	Ks	Ka	Ka/Ks	# of pairwise comparisons	# of significant pairwise comparisons by Fisher's exact test	Linkage disequilibrium R^2 model
All Sites	0.814	0.065	0.019	0.286	9180	2035	y = 0.0530 - 0.0100x
Poza Rica, MX	0.131	0.030	0.007	0.221	1540	282	y = 0.2142 - 0.0511x
Chetumal, MX	1.224	0.020	0.003	0.160	561	117	y = 0.1634 - 0.0342x
Lerdo de Tejada, MX	0.822	0.028	0.008	0.294	1653	230	y = 0.1360 - 0.0316x
Thailand	**P<.02, 2.060**	0.024	0.004	0.158	780	173	y = 0.1566 - 0.0346x
PK10, Senegal	1.450	0.035	0.009	0.263	2485	458	y = 0.1065 - 0.0113x

Table 6.8. *dcr1* gene flow and genetic differentiation analysis. Comparisons of the average number of synonymous amino acid differences for this gene (Ks) (Hudson *et al.*, 1992), coefficient of gene differentiation (G_{st}) or gene differentiation relative to the total population (Nei, 1973 Equation 9) and the average number of nucleotide substitutions per site, or nucleotide diversity (Dxy), between pairs of collection populations (Nei, 1987 Equation 10.20).

Population 1	Population 2	Ks	Gst	Dxy
Poza Rica	**Chetumal**	31.586	0.001	0.006
Poza Rica	**Lerdo de Tejada**	30.186	0	0.006
Poza Rica	**Thailand**	28.096	0	0.006
Poza Rica	**PK10**	31.888	0.002	0.007
Chetumal	**Lerdo de Tejada**	26.741	0	0.005
Chetumal	**Thailand**	24.825	0	0.007
Chetumal	**PK10**	27.584	0.003	0.007
Lerdo de Tejada	**Thailand**	23.542	0	0.005
Lerdo de Tejada	**PK10**	26.045	0.002	0.007
Thailand	**PK10**	23.745	0.002	0.007

Table 6.9. *dcr2* gene flow and genetic differentiation analysis. Comparisons of the average number of synonymous amino acid differences for this gene (Ks) (Hudson *et al.*, 1992), coefficient of gene differentiation (G_{st}) or gene differentiation relative to the total population (Nei, 1973 Equation 9) and the average number of nucleotide substitutions per site, or nucleotide diversity (Dxy), between pairs of collection populations (Nei, 1987 Equation 10.20).

Population 1	Population 2	Ks	Gst	Dxy
Poza Rica	**Chetumal**	10.477	0.007	0.002
Poza Rica	**Lerdo de Tejada**	13.952	0.007	0.004
Poza Rica	**Thailand**	12.228	0.008	0.004
Poza Rica	**PK10**	16.85	0.008	0.007
Chetumal	**Lerdo de Tejada**	13.509	0.003	0.004
Chetumal	**Thailand**	11.808	0.004	0.004
Chetumal	**PK10**	16.43	0.004	0.007
Lerdo de Tejada	**Thailand**	15.1	0.003	0.004
Lerdo de Tejada	**PK10**	19.723	0.003	0.007
Thailand	**PK10**	17.796	0.004	0.007

Figure 6.12. The *dcr1* and *dcr2* evolutionary rates from 5 *Ae. aegypti* mosquito collections from Mexico, Senegal and Thailand. Rate of evolution was calculated as the ratio of amino acid sequence evolution (*Ka*, non-synonymous sites) relative to neutral mutations (*Ks*, synonymous sites). The number of mosquitoes (n) used in each analysis is listed above each gene bar.

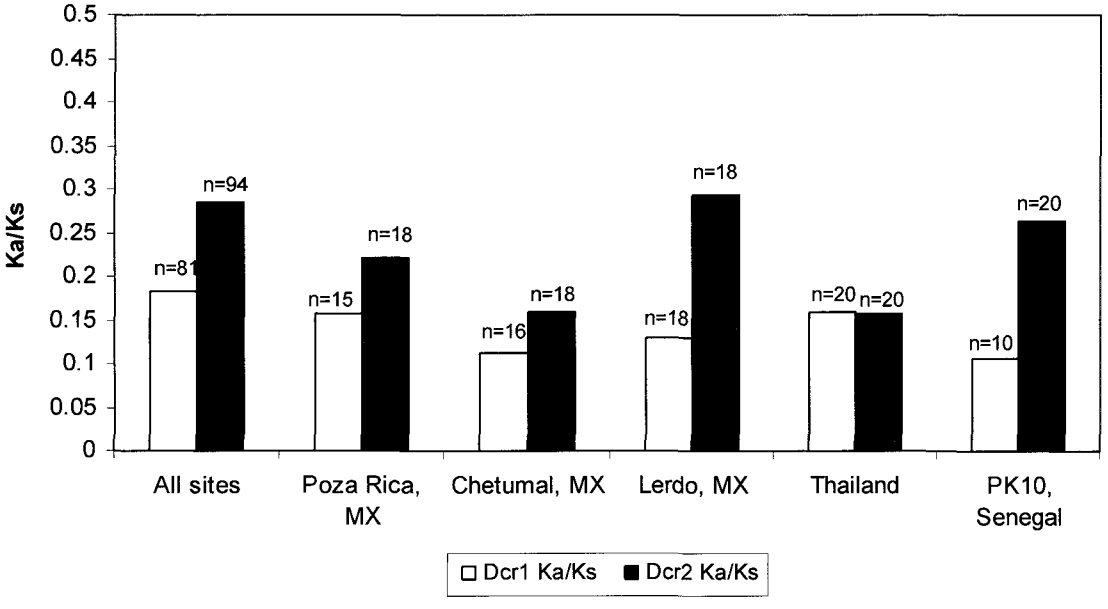


Figure 6.13. *dcr1* DNA polymorphism analysis for all 94 mosquitoes from all 5 collection sites. π is the average number of nucleotide differences per site between sequences, or nucleotide diversity (Nei, 1987, equation 10.5).

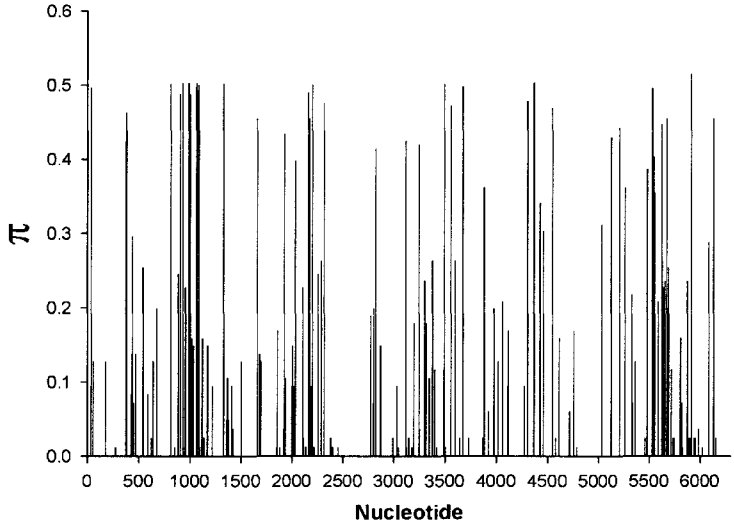


Figure 6.14. *dcr1* linkage disequilibrium analysis for all 94 mosquitoes from all 5 collection sites. The degree of linkage disequilibrium, or nonrandom association between nucleotide variants at different polymorphic sites is calculated and an estimate of relationship with physical distance is determined by regression analysis (R^2) (Hill and Robertson, 1966). The linear regression equation is at the top right of the graph and is demonstrated as a red line.

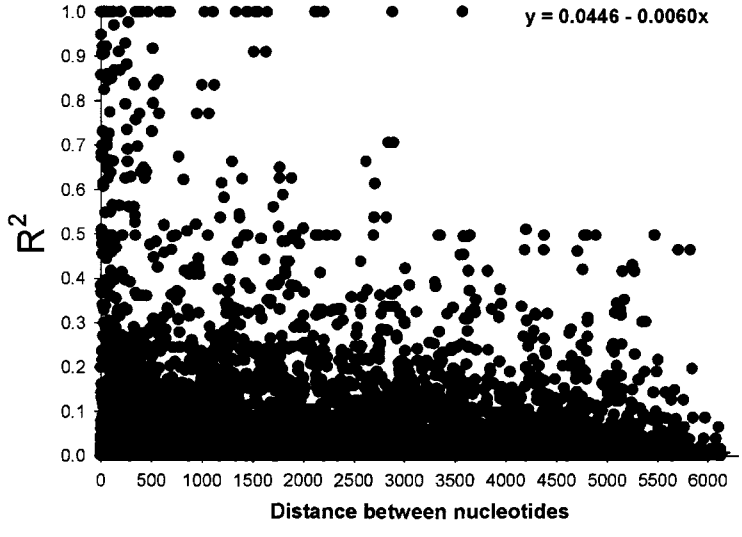


Figure 6.15. *dcr1* functional domains identified with Pfam overlaid on the DNA polymorphism analysis for all 94 mosquitoes from all 5 collection sites. π is the average number of nucleotide differences per site between sequences, or nucleotide diversity.

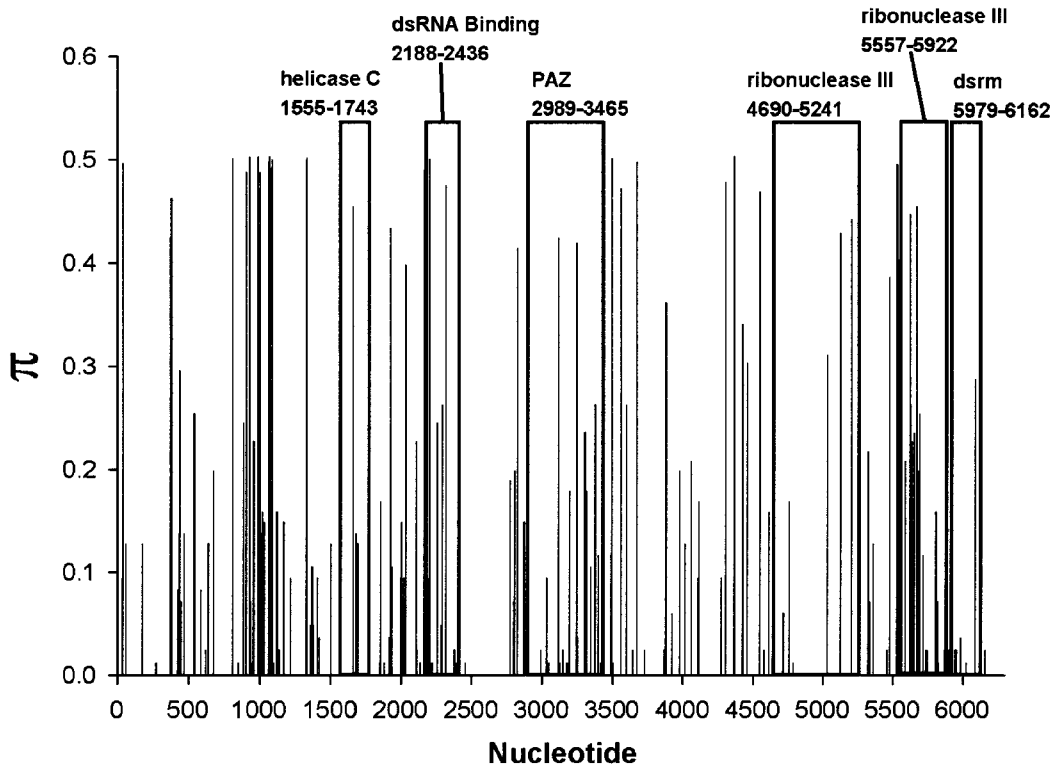


Figure 6.16. *dcr1* functional domain evolutionary rate analysis. All 5 *Ae. aegypti* mosquito collections from Mexico, Senegal and Thailand were used in this analysis. Rate of evolution was calculated as the ratio of amino acid sequence evolution (*Ka*, non-synonymous sites) relative to neutral mutations (*Ks*, synonymous sites). Fisher exact tests were performed for each domain with respect to the rest of the gene to identify significant differences ($P<0.05$) in the number of variable polymorphic sites. Domains that were found to be significantly different by Fisher exact test are highlighted by **.

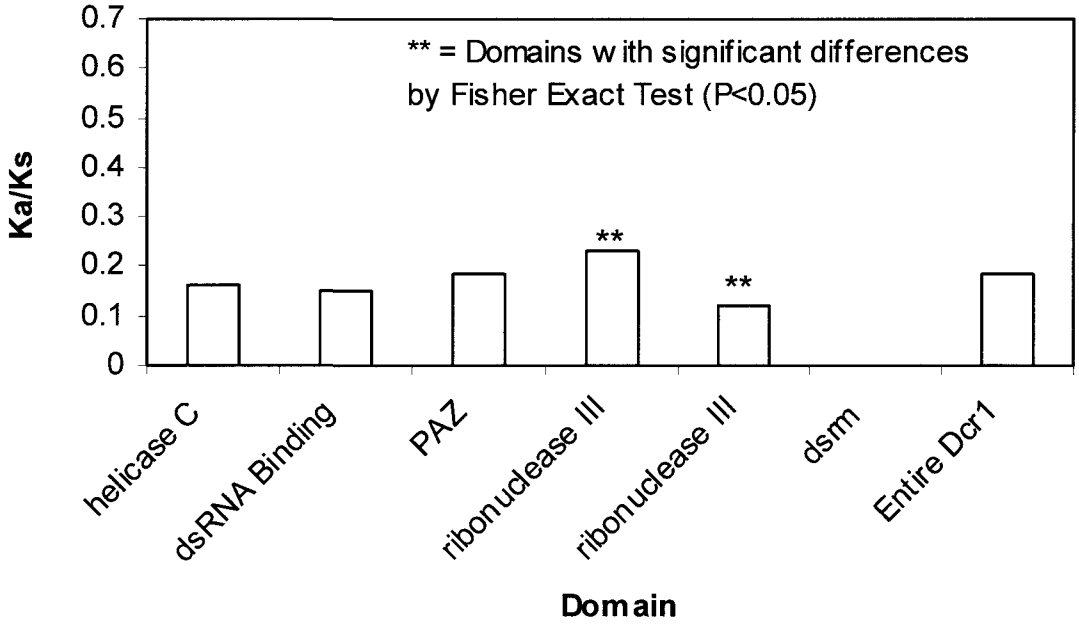


Figure 6.17. *dcr2* DNA polymorphism analysis for all 94 mosquitoes from all 5 collection sites. π is the average number of nucleotide differences per site between sequences, or nucleotide diversity (Nei, 1987, equation 10.5).

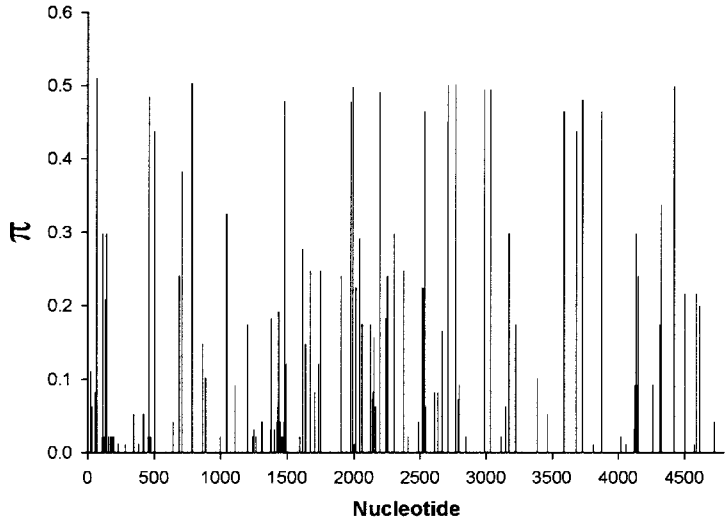


Figure 6.18. *dcr2* linkage disequilibrium analysis for all 94 mosquitoes from all 5 collection sites. The degree of linkage disequilibrium, or nonrandom association between nucleotide variants at different polymorphic sites is calculated and an estimate of relationship with physical distance is determined by regression analysis (R^2) (Hill and Robertson, 1966). The linear regression equation is at the top right of the graph and is demonstrated as a red line.

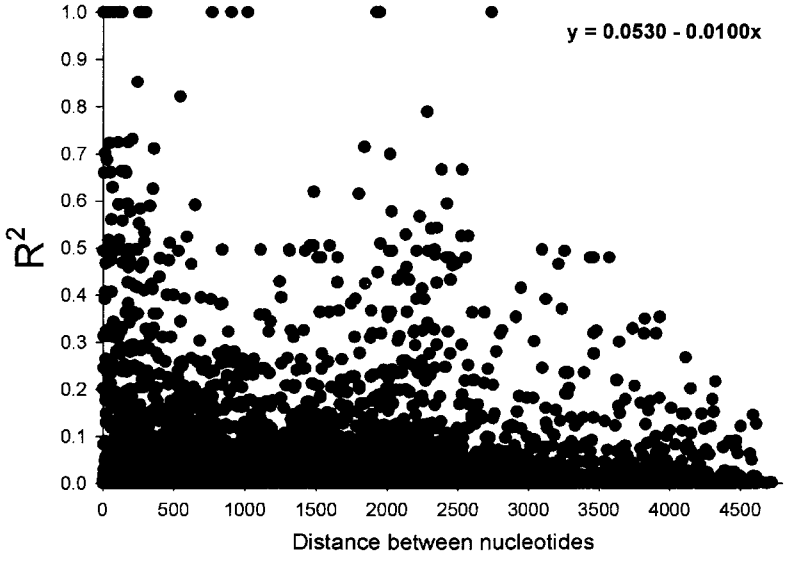


Figure 6.19. *dcr2* functional domains identified with Pfam overlaid on the DNA polymorphism analysis for all 94 mosquitoes from all 5 collection sites. π is the average number of nucleotide differences per site between sequences, or nucleotide diversity.

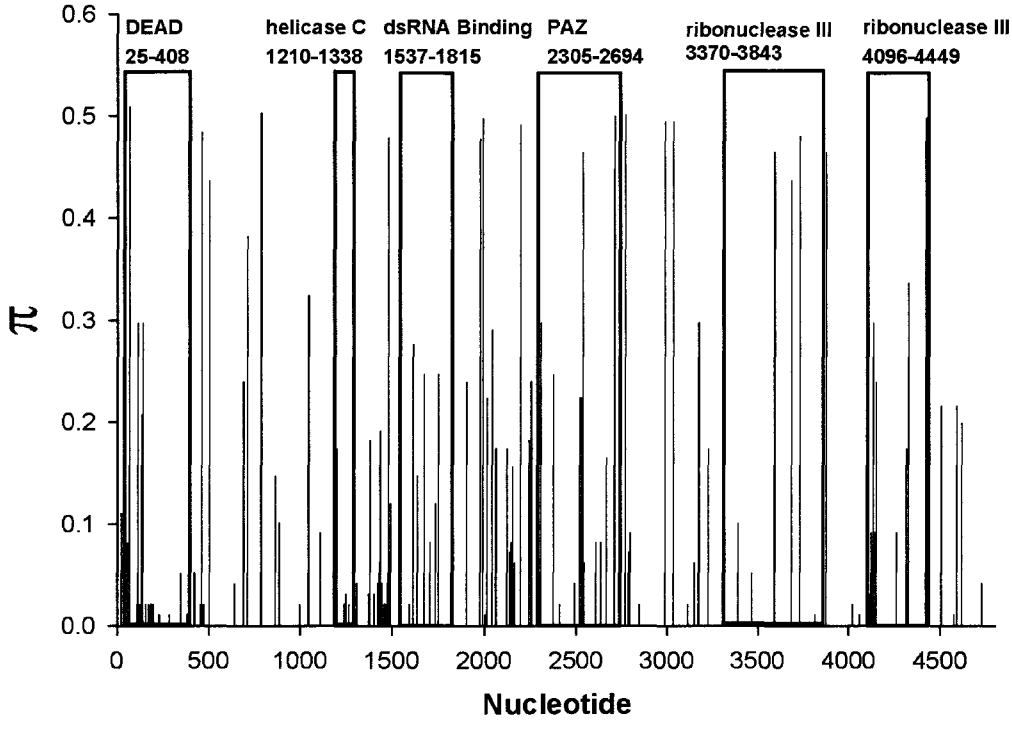


Figure 6.20. *dcr2* functional domain evolutionary rate analysis. All 5 *Ae. aegypti* mosquito collections from Mexico, Senegal and Thailand were used in this analysis. Rate of evolution was calculated as the ratio of amino acid sequence evolution (*Ka*, non-synonymous sites) relative to neutral mutations (*Ks*, synonymous sites). Fisher exact tests were performed for each domain with respect to the rest of the gene to identify significant differences ($P<0.05$) in the number of variable polymorphic sites. Domains that were found to be significantly different by Fisher exact test are highlighted by **.

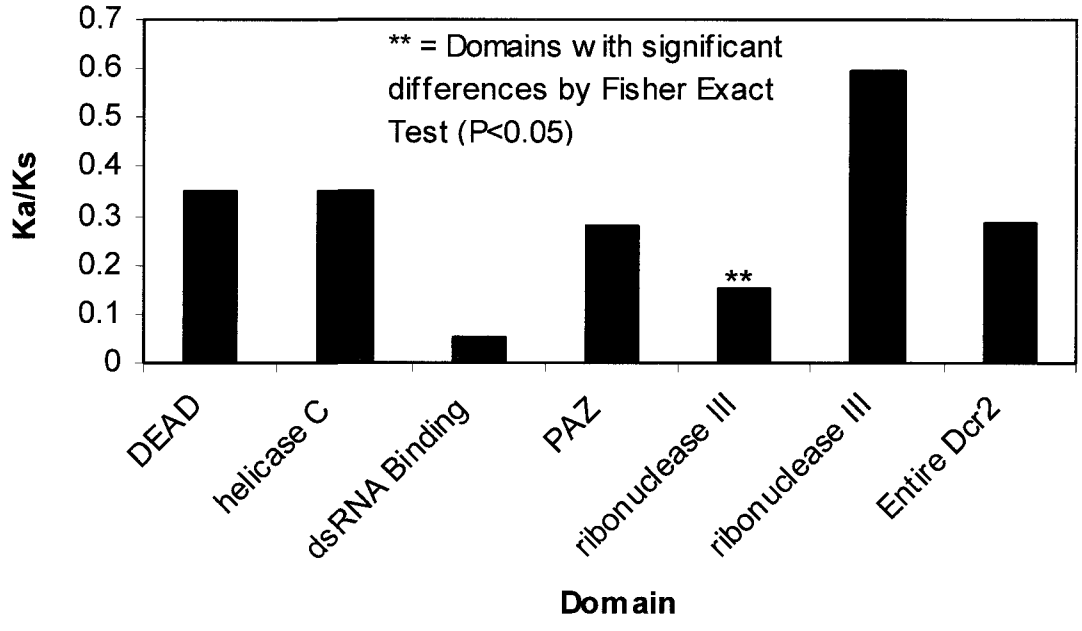


Table 6.10. DNAsp 4.50.3 output for *r3d1* containing analysis for all collection sites combined, as well analyzing for each specific collection site. This table shows number of variable sites (S), whether these resulted in synonymous or non-synonymous substitutions, number of haplotypes identified (h), haplotype diversity (Hd), nucleotide diversity (P_i), synonymous ($P_i(s)$) or non-synonymous (P_i (a)) diversity, average number of nucleotide differences (k), output for Fu and Li's F* test statistic, ratio of non-synonymous to synonymous nucleotide differences (*Ka/Ks*) and pairwise comparisons to determine a linkage disequilibrium model.

Site	# of variable sites, S=	Synonymous substitutions	Non-synonymous substitutions	# of haplotypes, h	Haplotype (gene) diversity, Hd	Nucleotide diversity (per site), Pi	Pi(s)	Pi(a)	Pi(a)/Pi(s)	Average # of nucleotide differences, k
All Sites	29	21	8	55	0.826	0.004	0.016	0.001	0.060	4.371
Poza Rica, MX	16	14	2	15	0.838	0.004	0.013	0.001	0.091	3.783
Chetumal, MX	11	10	2	8	0.614	0.002	0.006	0.001	0.158	2.021
Lerdo de Tejada, MX	17	15	2	15	0.833	0.004	0.077	0.000	0.006	4.367
Thailand	4	4	0	4	0.514	0.001	0.005	0.000	0.000	1.142
PK10, Senegal	18	14	4	20	0.954	0.005	0.018	0.001	0.070	4.953

Site	Fu and Li's F* test statistic, F*	Ks	Ka	Ka/Ks	# of pairwise comparisons	# of significant pairwise comparisons by Fisher's exact test	Linkage disequilibrium R^2 model
All Sites	0.652	0.091	0.011	0.117	378	70	y = 0.0691 - 0.0662x
Poza Rica, MX	1.171	0.061	0.003	0.044	15	22	y = 0.2009 - 0.2500x
Chetumal, MX	0.880	0.043	0.003	0.062	11	16	y = 0.5651 - 1.0900x
Lerdo de Tejada, MX	0.750	0.065	0.003	0.041	17	28	y = 0.2859 - 0.3481x
Thailand	1.016	0.017	0.000	0.000	6	3	N/A
PK10, Senegal	1.501	0.061	0.005	0.088	18	33	y = 0.2274 - 0.3043x

Table 6.11. DNAsp 4.50.3 output for *r2d2* containing analysis for all collection sites combined, as well analyzing for each specific collection site. This table shows number of variable sites (S), whether these resulted in synonymous or non-synonymous substitutions, number of haplotypes identified (h), haplotype diversity (Hd), nucleotide diversity (P_i), synonymous ($P_i(s)$) or non-synonymous (P_i (a)) diversity, average number of nucleotide differences (k), output for Fu and Li's F* test statistic, ratio of non-synonymous to synonymous nucleotide differences (*Ka/Ks*) and pairwise comparisons to determine a linkage disequilibrium model.

Site	# of variable sites, S=	Synonymous substitutions	Non-synonymous substitutions	# of haplotypes, h	Haplotype (gene) diversity, Hd	Nucleotide diversity (per site), Pi	Pi(s)	Pi(a)	Pi(a)/Pi(s)	Average # of nucleotide differences, k
All Sites	15	10	5	32	0.884	0.004	0.015	0.001	0.089	3.997
Poza Rica, MX	5	3	2	5	0.392	0.001	0.004	0.000	0.081	1.040
Chetumal, MX	8	7	1	13	0.827	0.003	0.009	0.001	0.066	2.317
Lerdo de Tejada, MX	5	4	1	10	0.813	0.002	0.008	0.001	0.079	1.998
Thailand	8	6	2	9	0.756	0.002	0.006	0.001	0.143	1.733
PK10, Senegal	7	5	2	5	0.774	0.002	0.004	0.001	0.187	1.377
Mindin, Senegal										

Site	Fu and Li's F* test statistic, F*	Ks	Ka	Ka/Ks	# of pairwise comparisons	# of significant pairwise comparisons by Fisher's exact test	Linkage disequilibrium R^2 model
All Sites	0.907	0.049	0.007	0.147	105	52	y = 0.1586 - 0.1289x
Poza Rica, MX	0.791	0.015	0.003	0.196	10	5	y = 0.5188 - 0.6578x
Chetumal, MX	0.205	0.034	0.001	0.042	28	5	y = 0.1218 - 0.0746x
Lerdo de Tejada, MX	1.520	0.020	0.001	0.074	10	3	y = 0.1582 - 0.1159x
Thailand	0.971	0.029	0.003	0.098	28	10	y = 0.4133 - 0.6970x
PK10, Senegal	1.218	0.024	0.003	0.118	6	3	y = 0.4688 - 0.7894x

Table 6.12. *r3d1* gene flow and genetic differentiation analysis. Comparisons of the average number of synonymous amino acid differences for this gene (Ks) (Hudson *et al.*, 1992), coefficient of gene differentiation (G_{st}) or gene differentiation relative to the total population (Nei, 1973 Equation 9) and the average number of nucleotide substitutions per site, or nucleotide diversity (Dxy), between pairs of collection populations (Nei, 1987 Equation 10.20).

Population 1	Population 2	Ks	Gst	Dxy
Poza Rica	**Chetumal**	2.902	0.016	0.003
Poza Rica	**Lerdo de Tejada**	4.075	0.005	0.004
Poza Rica	**Thailand**	2.393	0.044	0.003
Poza Rica	**PK10**	4.398	0.055	0.007
Chetumal	**Lerdo de Tejada**	3.194	0.025	0.004
Chetumal	**Thailand**	1.558	0.02	0.002
Chetumal	**PK10**	3.564	0.12	0.007
Lerdo de Tejada	**Thailand**	2.67	0.046	0.003
Lerdo de Tejada	**PK10**	4.675	0.056	0.008
Thailand	**PK10**	3.047	0.153	0.006

Table 6.13. *r2d2* gene flow and genetic differentiation analysis. Comparisons of the average number of synonymous amino acid differences for this gene (Ks) (Hudson *et al.*, 1992), coefficient of gene differentiation (G_{st}) or gene differentiation relative to the total population (Nei, 1973 Equation 9) and the average number of nucleotide substitutions per site, or nucleotide diversity (Dxy), between pairs of collection populations (Nei, 1987 Equation 10.20).

Population 1	Population 2	Ks	Gst	Dxy
Poza Rica	**Chetumal**	1.679	0.059	0.002
Poza Rica	**Lerdo de Tejada**	1.519	0.073	0.002
Poza Rica	**Thailand**	1.405	0.267	0.007
Poza Rica	**PK10**	1.213	0.262	0.008
Chetumal	**Lerdo de Tejada**	2.158	0.002	0.002
Chetumal	**Thailand**	2.01	0.116	0.006
Chetumal	**PK10**	1.835	0.111	0.007
Lerdo de Tejada	**Thailand**	1.859	0.121	0.006
Lerdo de Tejada	**PK10**	1.679	0.115	0.007
Thailand	**PK10**	1.56	0.133	0.004

Figure 6.21. The *r3d1* and *r2d2* evolutionary rates from 5 *Ae. aegypti* mosquito collections from Mexico, Senegal and Thailand. Rate of evolution was calculated as the ratio of amino acid sequence evolution (*Ka*, non-synonymous sites) relative to neutral mutations (*Ks*, synonymous sites). The number of mosquitoes (n) used in each analysis is listed above each gene bar.

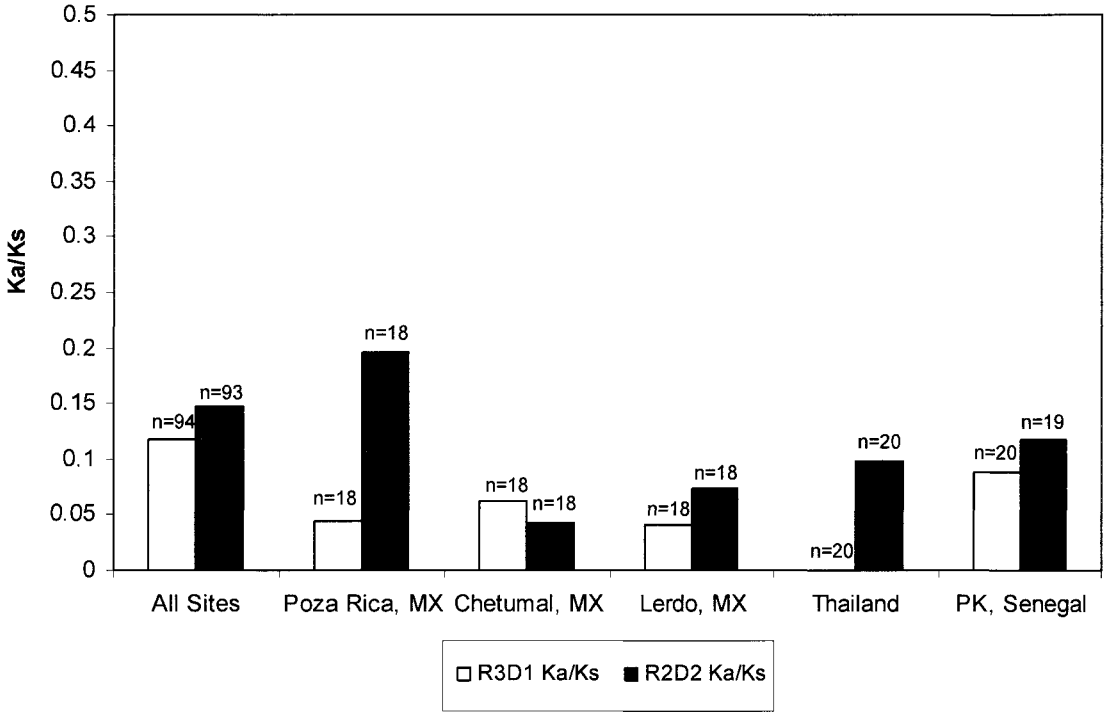


Figure 6.22. *r3d1* DNA polymorphism analysis for all 94 mosquitoes from all 5 collection sites. π is the average number of nucleotide differences per site between sequences, or nucleotide diversity (Nei, 1987, equation 10.5).

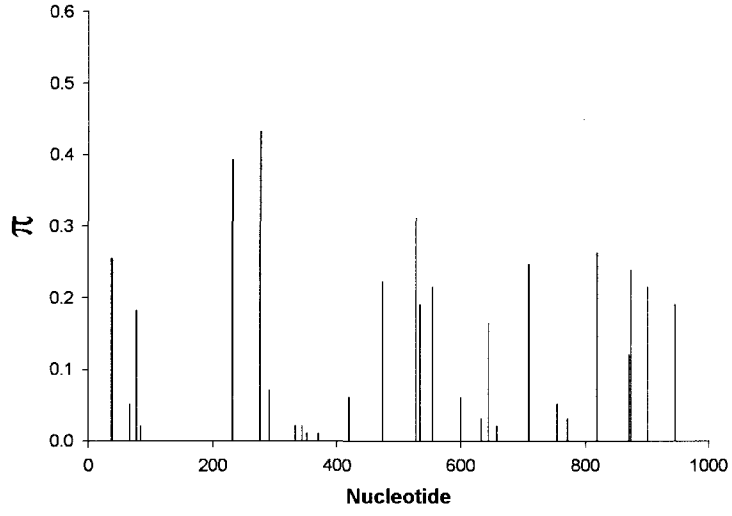


Figure 6.23. *r3d1* linkage disequilibrium analysis for all 94 mosquitoes from all 5 collection sites. The degree of linkage disequilibrium, or nonrandom association between nucleotide variants at different polymorphic sites is calculated and an estimate of relationship with physical distance is determined by regression analysis (R^2) (Hill and Robertson, 1966). The linear regression equation is at the top right of the graph and is demonstrated as a red line.

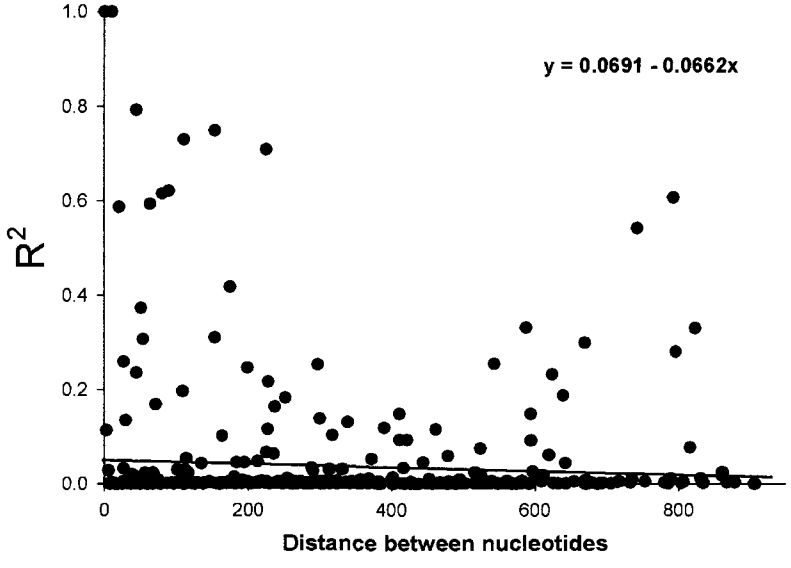


Figure 6.24. *r3d1* functional domains identified with Pfam overlaid on the DNA polymorphism analysis for all 94 mosquitoes from all 5 collection sites. π is the average number of nucleotide differences per site between sequences, or nucleotide diversity.

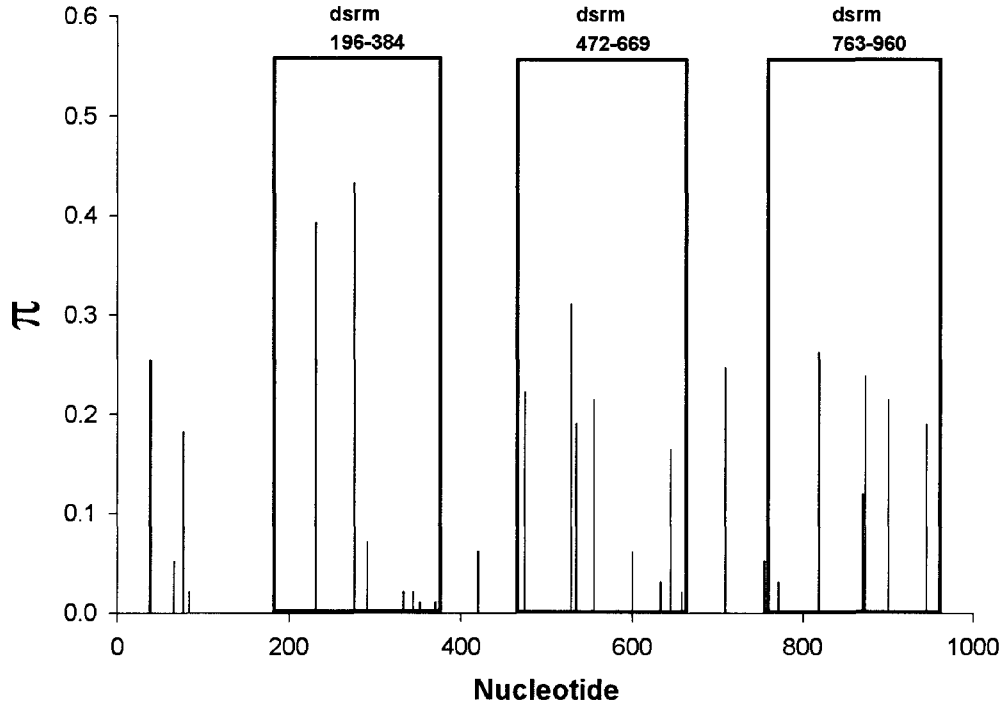


Figure 6.25. *r3d1* functional domain evolutionary rate analysis. All 5 *Ae. aegypti* mosquito collections from Mexico, Senegal and Thailand were used in this analysis. Rate of evolution was calculated as the ratio of amino acid sequence evolution (*Ka*, non-synonymous sites) relative to neutral mutations (*Ks*, synonymous sites). Fisher exact tests were performed for each domain with respect to the rest of the gene to identify significant differences ($P<0.05$) in the number of variable polymorphic sites. Domains that were found to be significantly different by Fisher exact test are highlighted by **. No significant differences within the domains were identified.

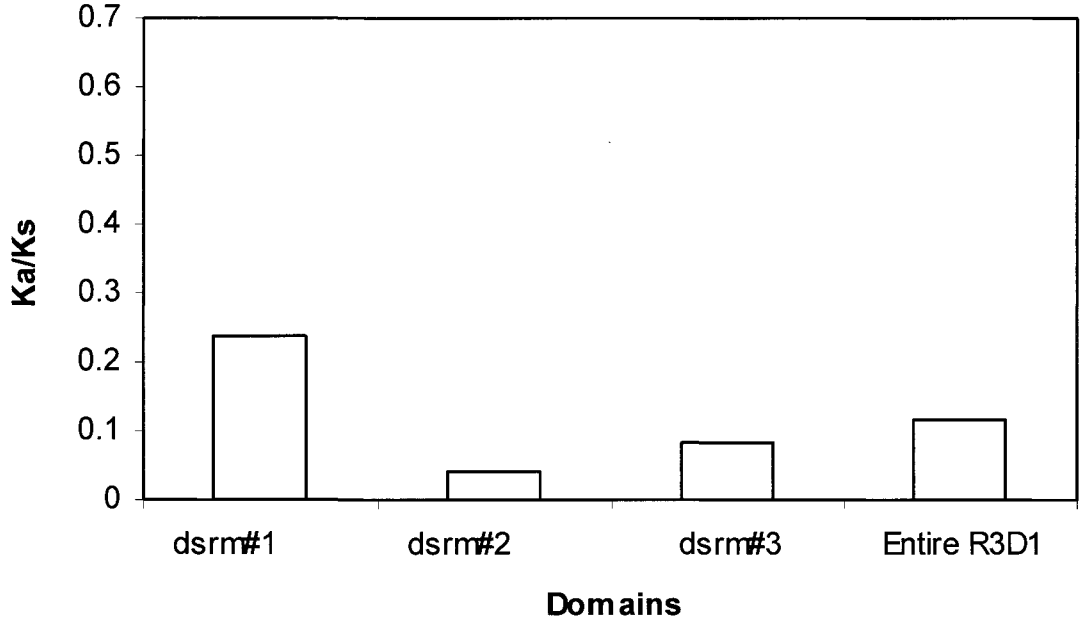


Figure 6.26. *r2d2* DNA polymorphism analysis for all 94 mosquitoes from all 5 collection sites. π is the average number of nucleotide differences per site between sequences, or nucleotide diversity (Nei, 1987, equation 10.5).

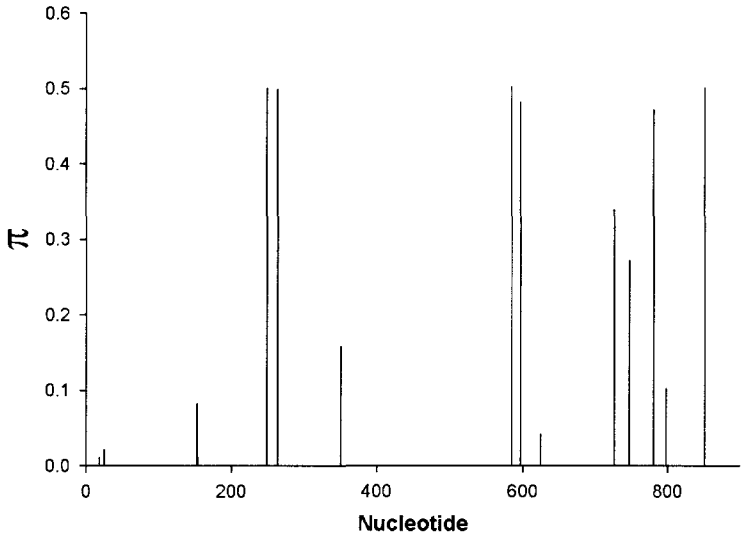


Figure 6.27. *r2d2* linkage disequilibrium analysis for all 94 mosquitoes from all 5 collection sites. The degree of linkage disequilibrium, or nonrandom association between nucleotide variants at different polymorphic sites is calculated and an estimate of relationship with physical distance is determined by regression analysis (R^2) (Hill and Robertson, 1966). The linear regression equation is at the top right of the graph and is demonstrated as a red line.

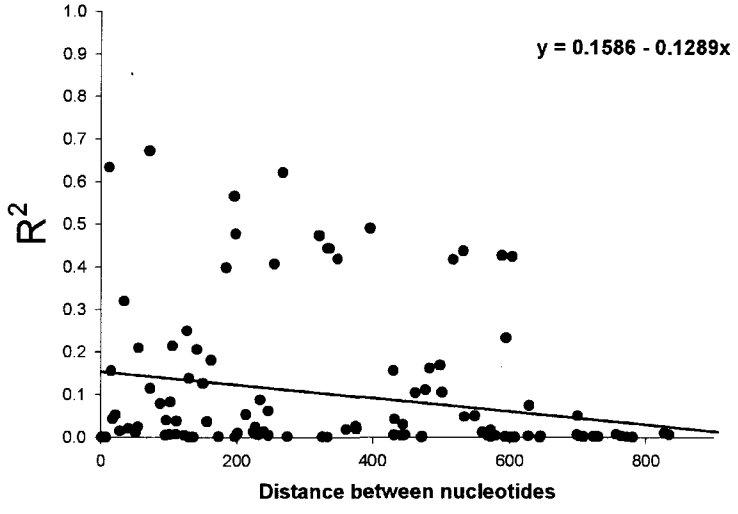


Figure 6.28. *r2d2* functional domains identified with Pfam overlaid on the DNA polymorphism analysis for all 94 mosquitoes from all 5 collection sites. π is the average number of nucleotide differences per site between sequences, or nucleotide diversity.

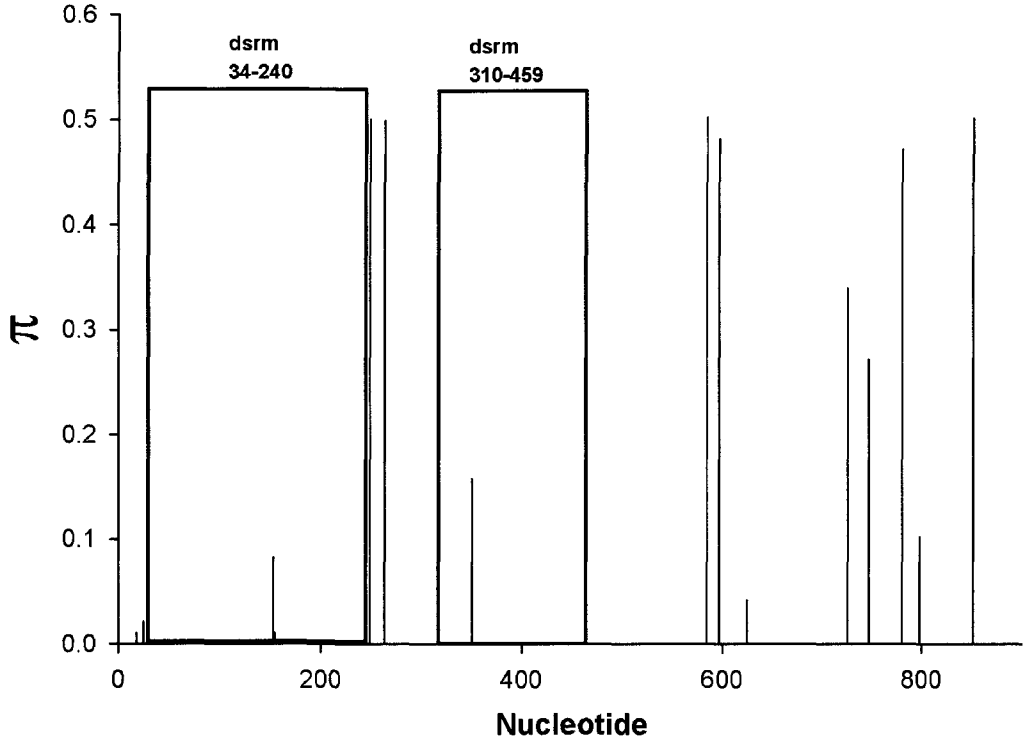


Figure 6.29. *r2d2* functional domain evolutionary rate analysis. All 5 *Ae. aegypti* mosquito collections from Mexico, Senegal and Thailand were used in this analysis. Rate of evolution was calculated as the ratio of amino acid sequence evolution (*Ka*, non-synonymous sites) relative to neutral mutations (*Ks*, synonymous sites). Fisher exact tests were performed for each domain with respect to the rest of the gene to identify significant differences ($P<0.05$) in the number of variable polymorphic sites. Domains that were found to be significantly different by Fisher exact test are highlighted by **. No significant differences within the domains were identified.

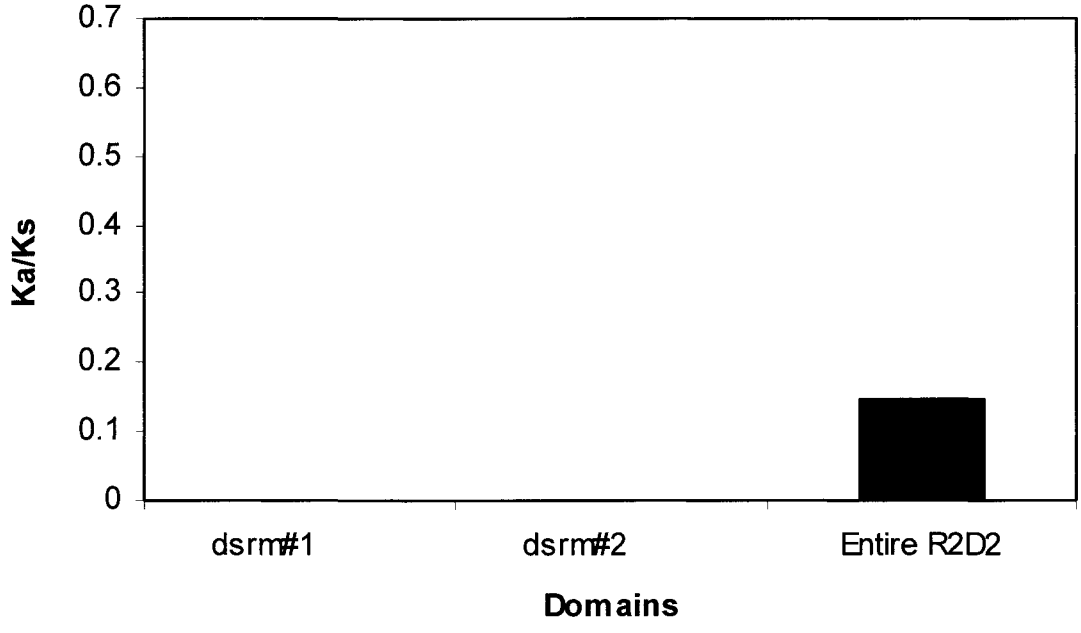


Figure 6.30. Dicer 2 MEB correlation analysis. Significant correlation was found between MEB rate and average number of nucleotide differences in *dcr2*. FET(prob) is Fisher exact test with P-value.

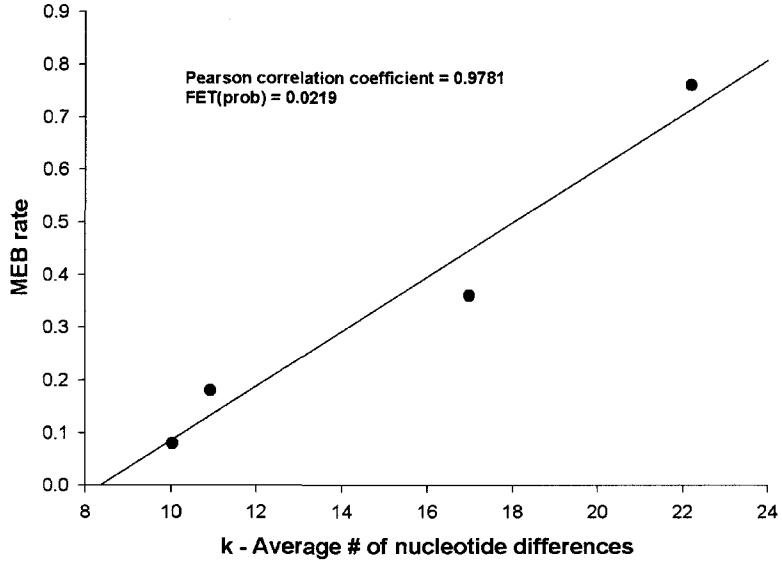


Figure 6.31. Argonaute 2 MEB correlation analysis. Significant inverse correlation was found between MEB rate and π (nucleotide diversity per site) in *ago2*. FET(prob) is Fisher exact test with P-value.

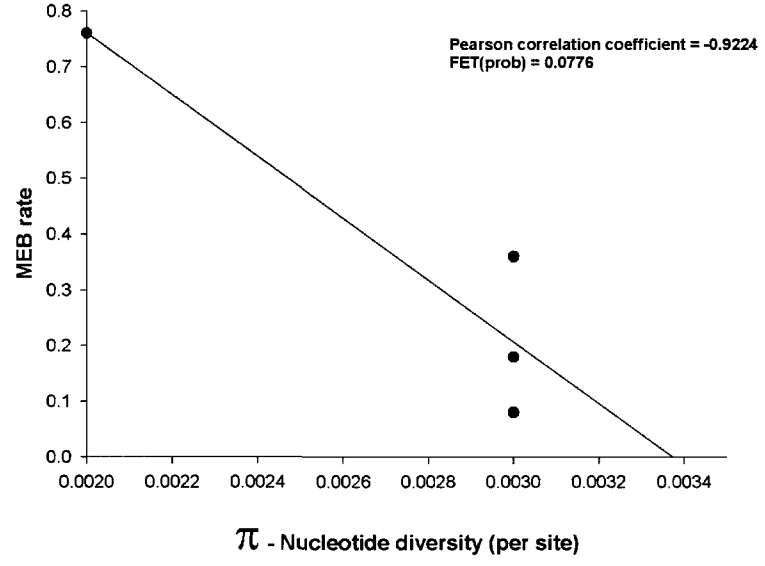


Figure 6.32. *ago1* maximum parsimony phylogenetic tree. Equally weighted parsimony tree searches were conducted using 2,000 random addition tree-bisection-reconnection (TBR) searches in PAUP* version 4.0b10 (Swofford *et al.*, 2001).

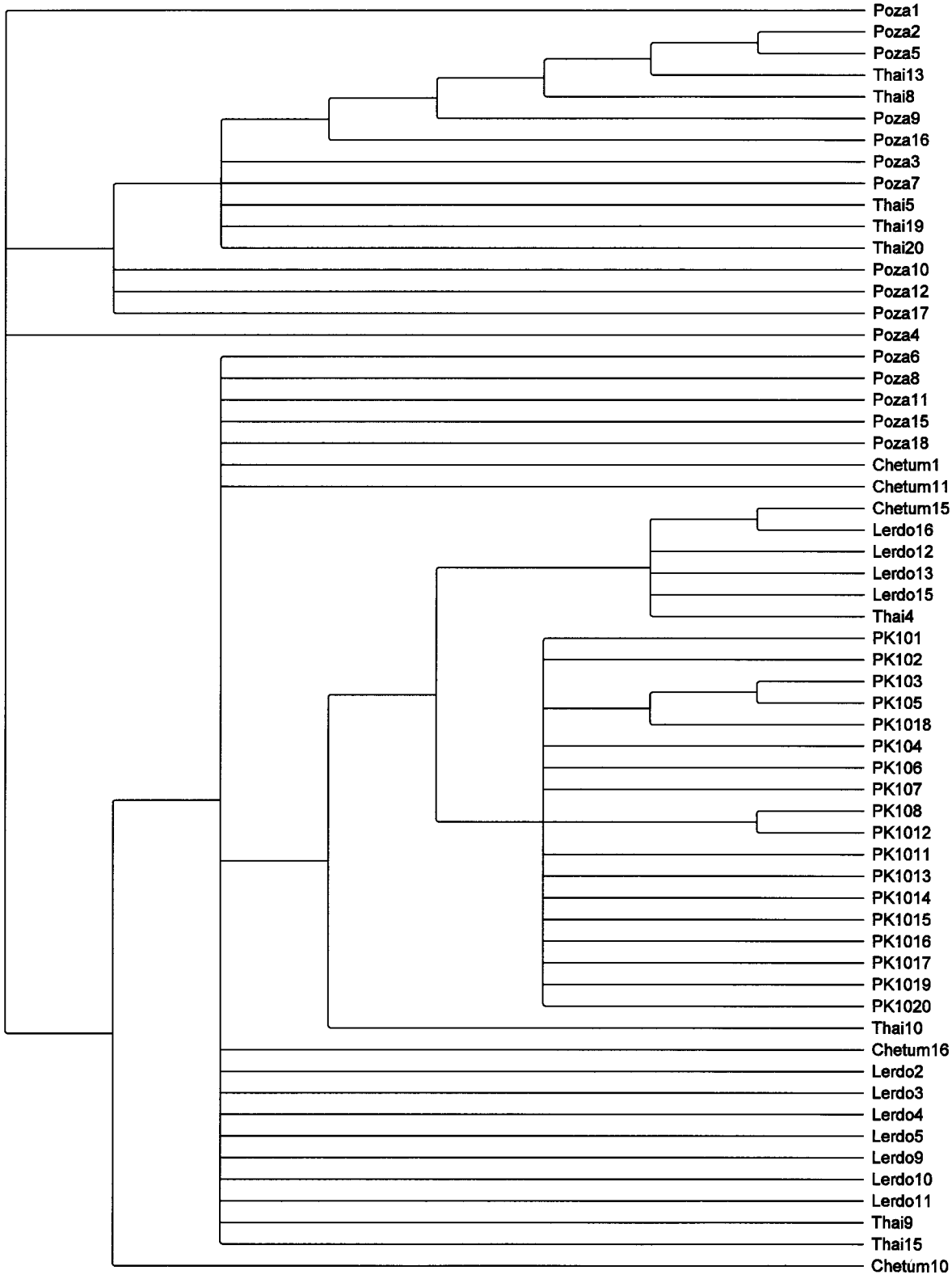


Figure 6.33. *ago1* maximum likelihood phylogenetic tree. Likelihood analyses were conducted using RAxML version 7.03 (Stamatakis *et al.*, 2006). Optimal likelihood trees were searched for using 1,000 independent searches starting from randomized parsimony trees with the GTRGAMMA model and four discrete rate categories. The numbers on each of the branches represent bootstrap values.

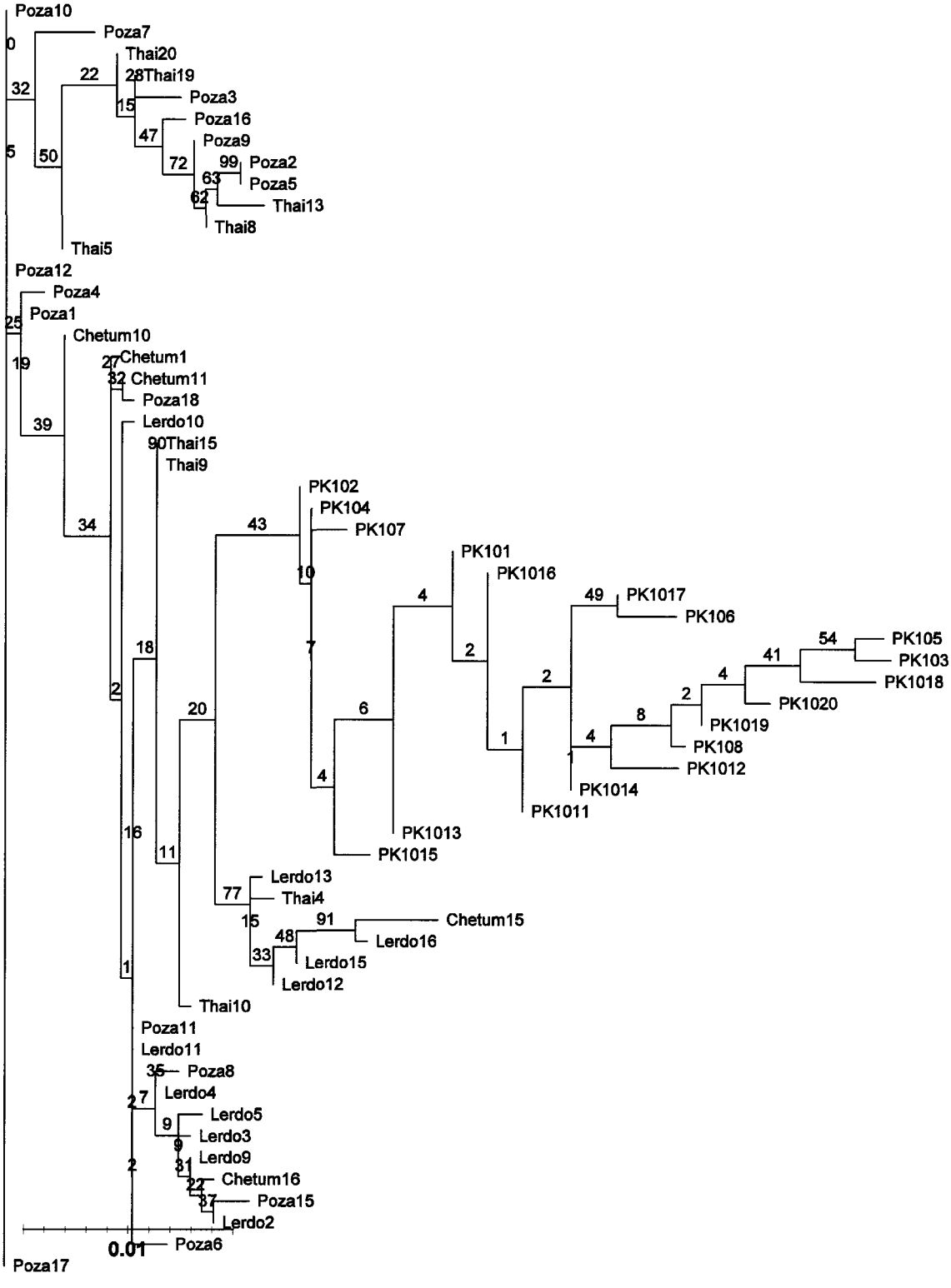


Figure 6.34. *ago2* maximum parsimony phylogenetic tree. Equally weighted parsimony tree searches were conducted using 2,000 random addition tree-bisection-reconnection (TBR) searches in PAUP* version 4.0b10 (Swofford *et al.*, 2001).

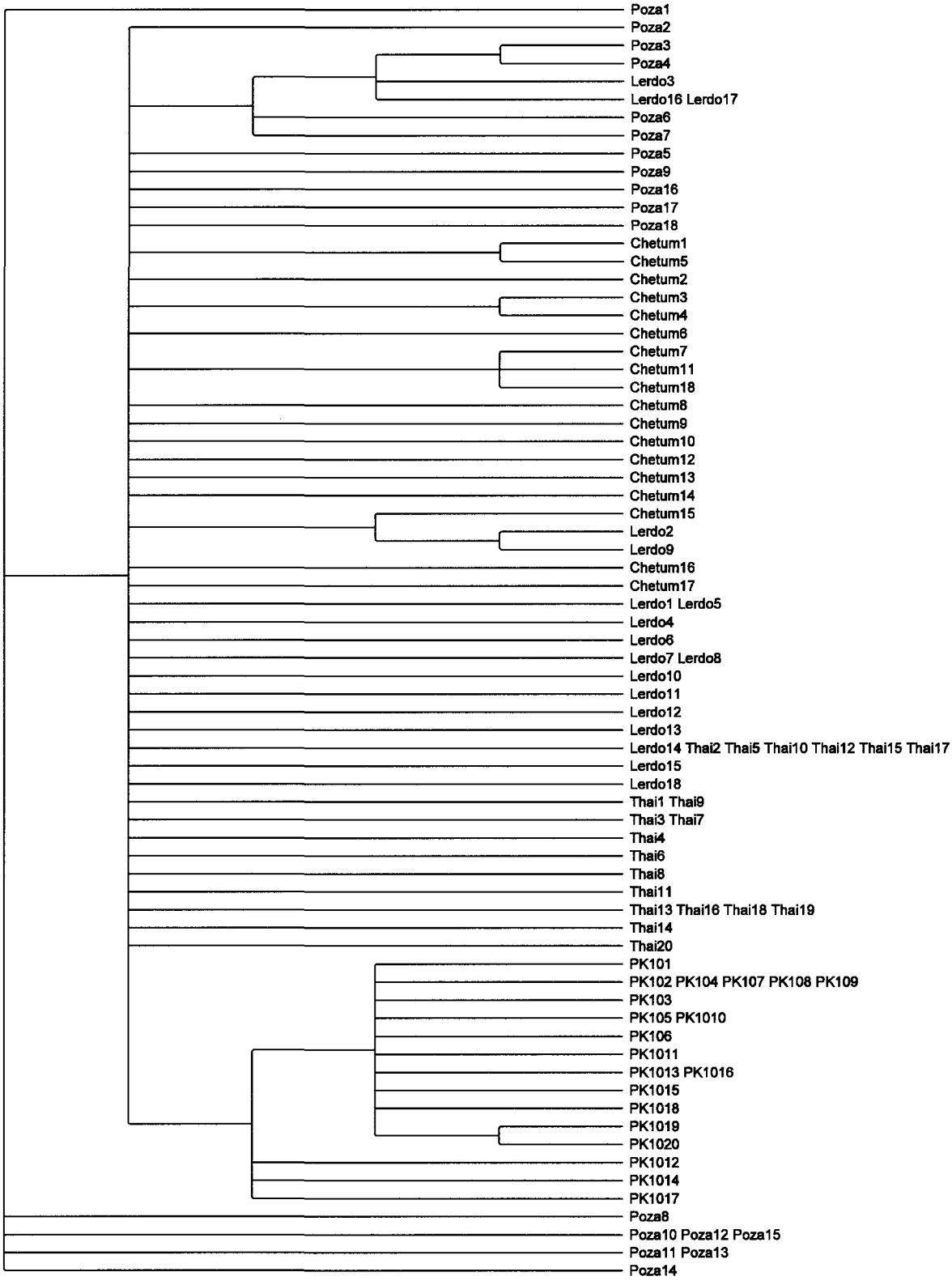


Figure 6.35. *ago2* maximum likelihood phylogenetic tree. Likelihood analyses were conducted using RAxML version 7.03 (Stamatakis *et al.*, 2006). Optimal likelihood trees were searched for using 1,000 independent searches starting from randomized parsimony trees with the GTRGAMMA model and four discrete rate categories. The numbers on each of the branches represent bootstrap values.

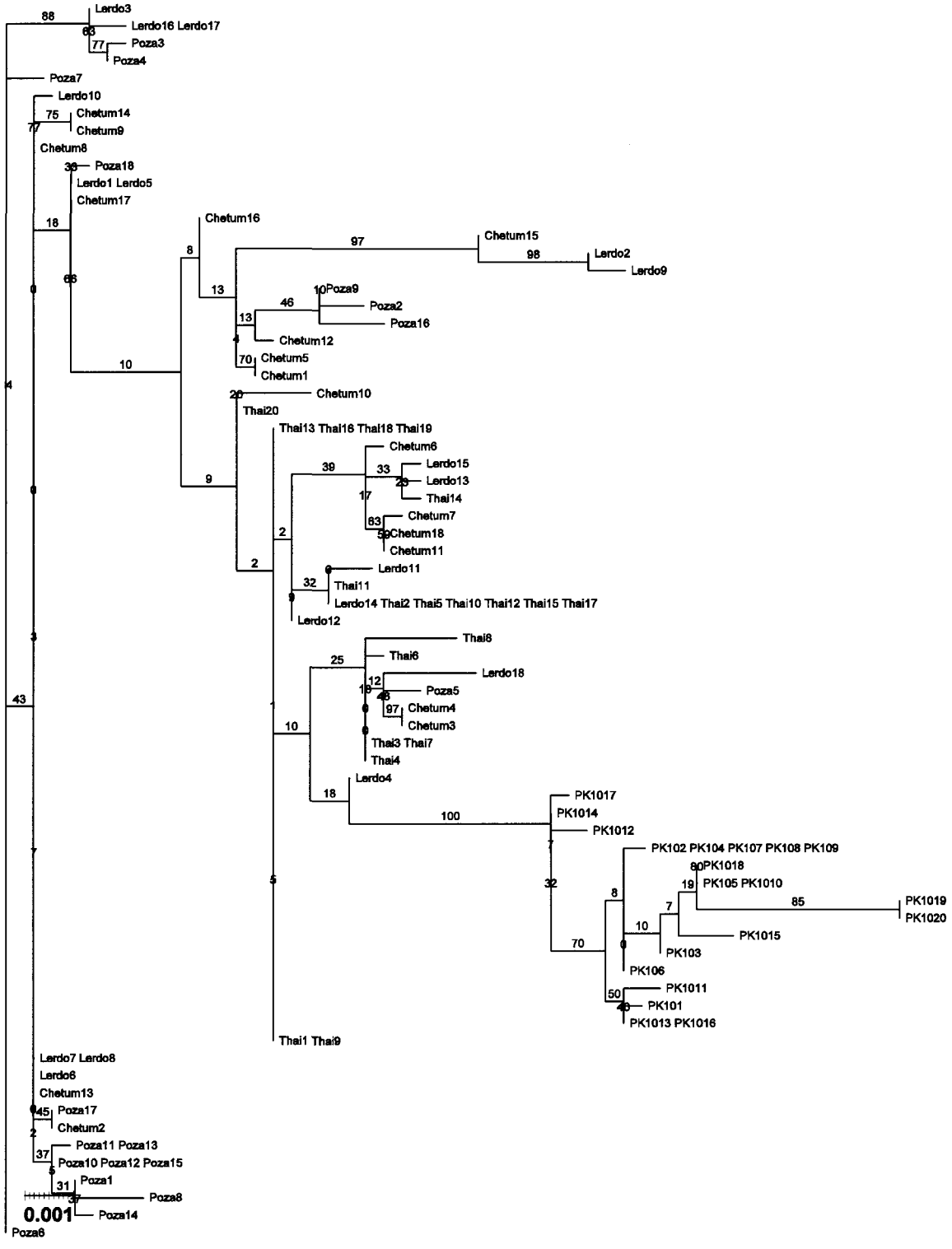


Figure 6.36. *dcr1* maximum parsimony phylogenetic tree. Equally weighted parsimony tree searches were conducted using 2,000 random addition tree-bisection-reconnection (TBR) searches in PAUP* version 4.0b10 (Swofford *et al.*, 2001).

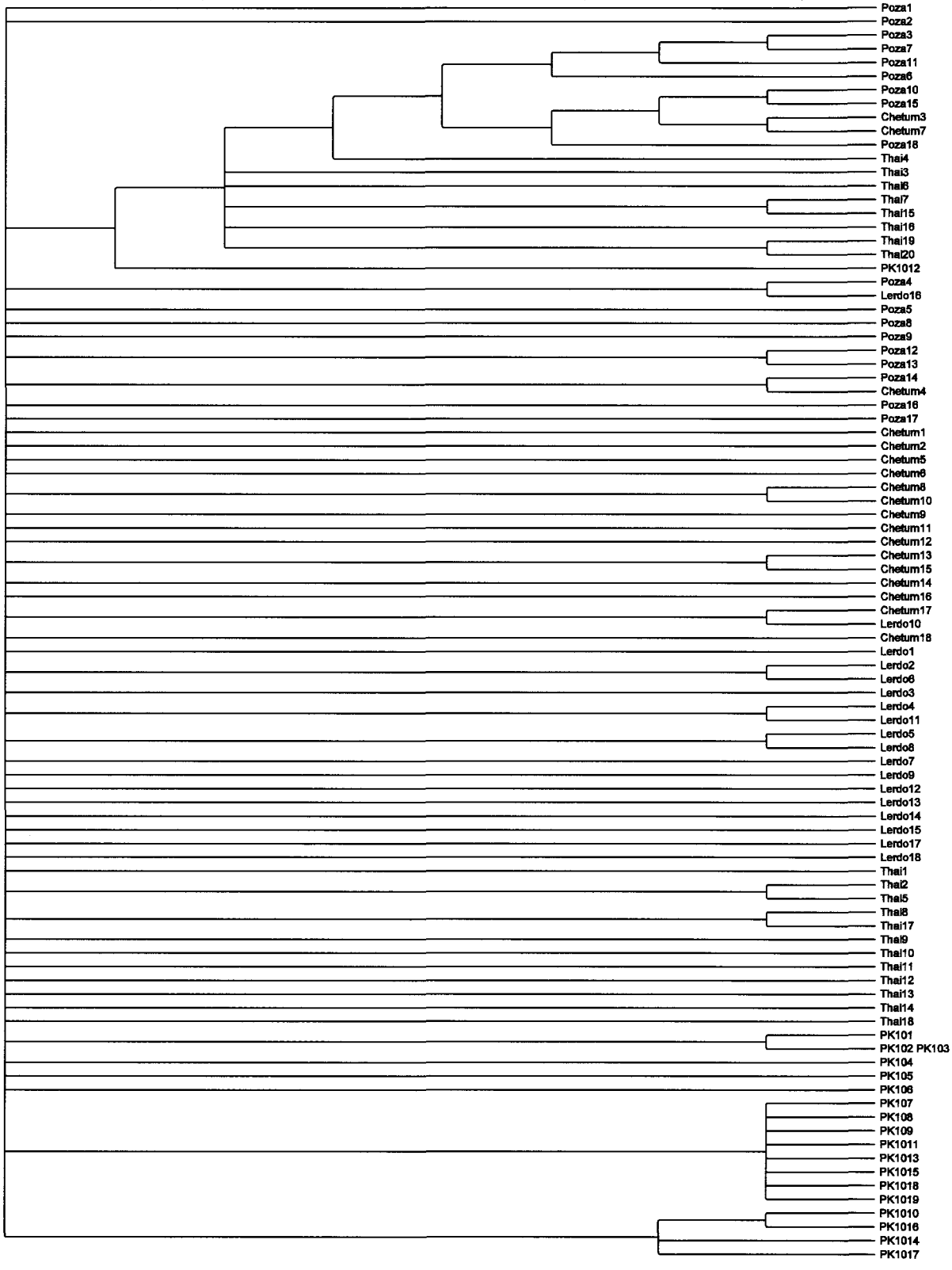


Figure 6.37. *dcr1* maximum likelihood phylogenetic tree. Likelihood analyses were conducted using RAxML version 7.03 (Stamatakis *et al.*, 2006). Optimal likelihood trees were searched for using 1,000 independent searches starting from randomized parsimony trees with the GTRGAMMA model and four discrete rate categories. The numbers on each of the branches represent bootstrap values.

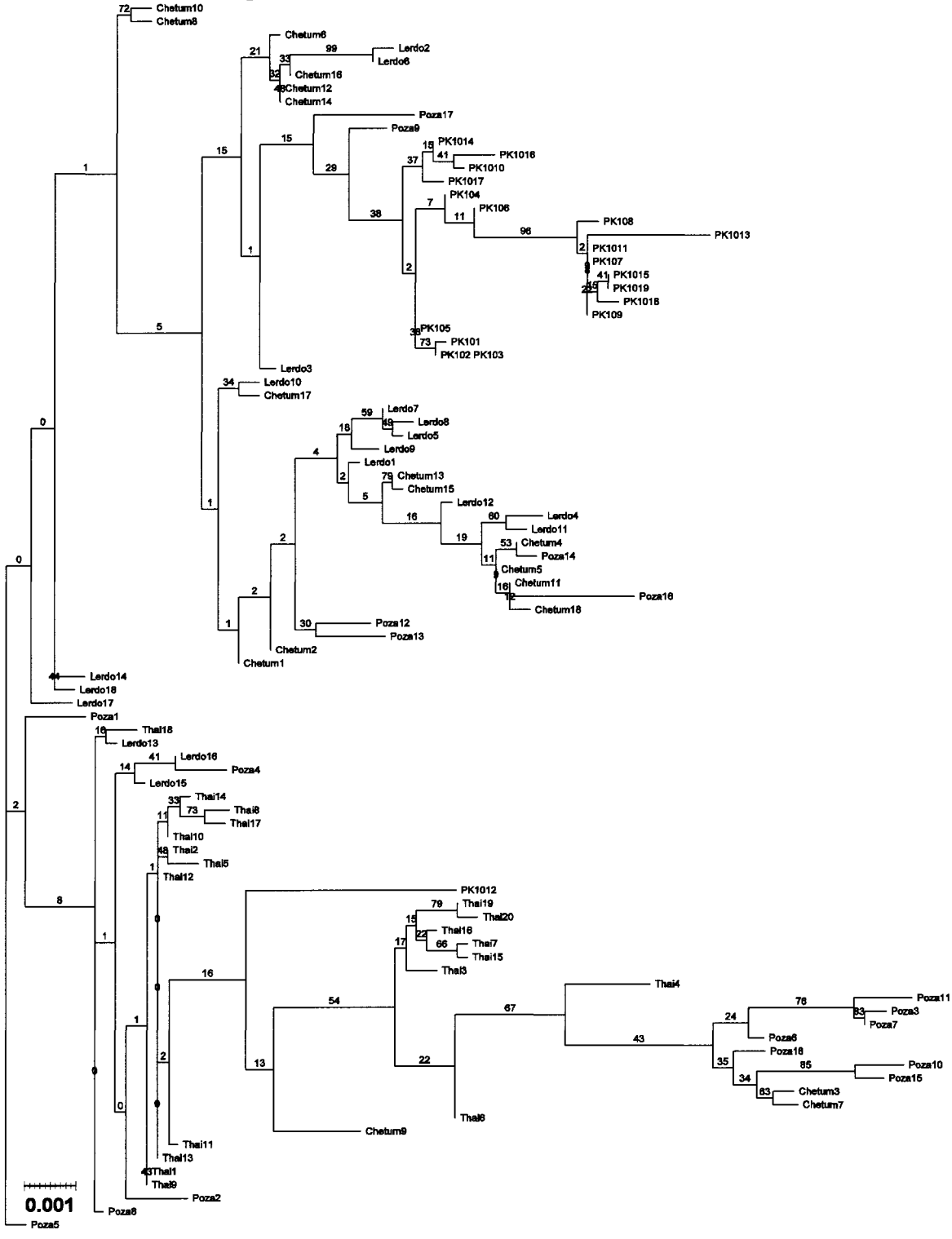


Figure 6.38. *dcr2* maximum parsimony phylogenetic tree. Equally weighted parsimony tree searches were conducted using 2,000 random addition tree-bisection-reconnection (TBR) searches in PAUP* version 4.0b10 (Swofford *et al.*, 2001).

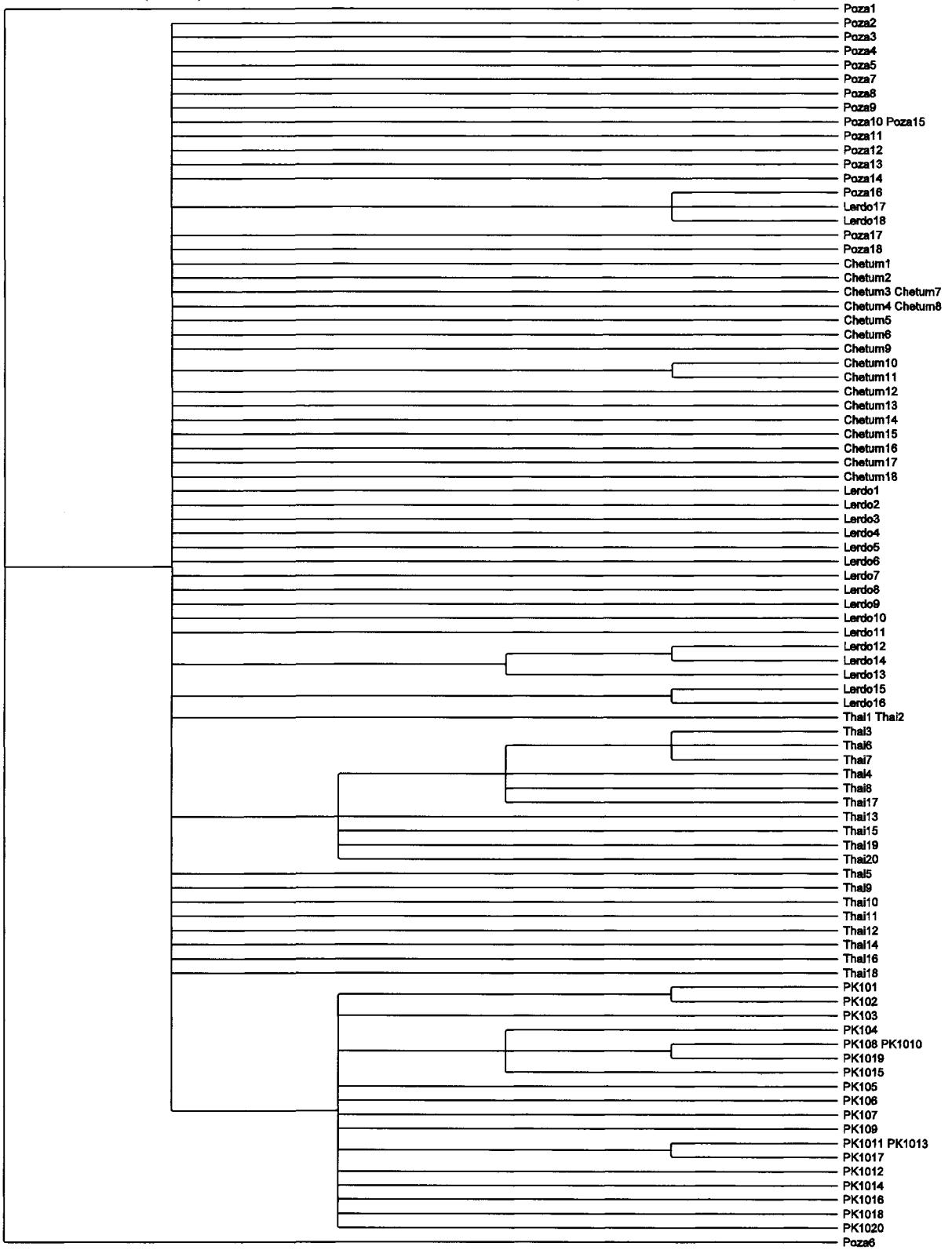


Figure 6.39. *dcr2* maximum likelihood phylogenetic tree. Likelihood analyses were conducted using RAxML version 7.03 (Stamatakis *et al.*, 2006). Optimal likelihood trees were searched for using 1,000 independent searches starting from randomized parsimony trees with the GTRGAMMA model and four discrete rate categories. The numbers on each of the branches represent bootstrap values.

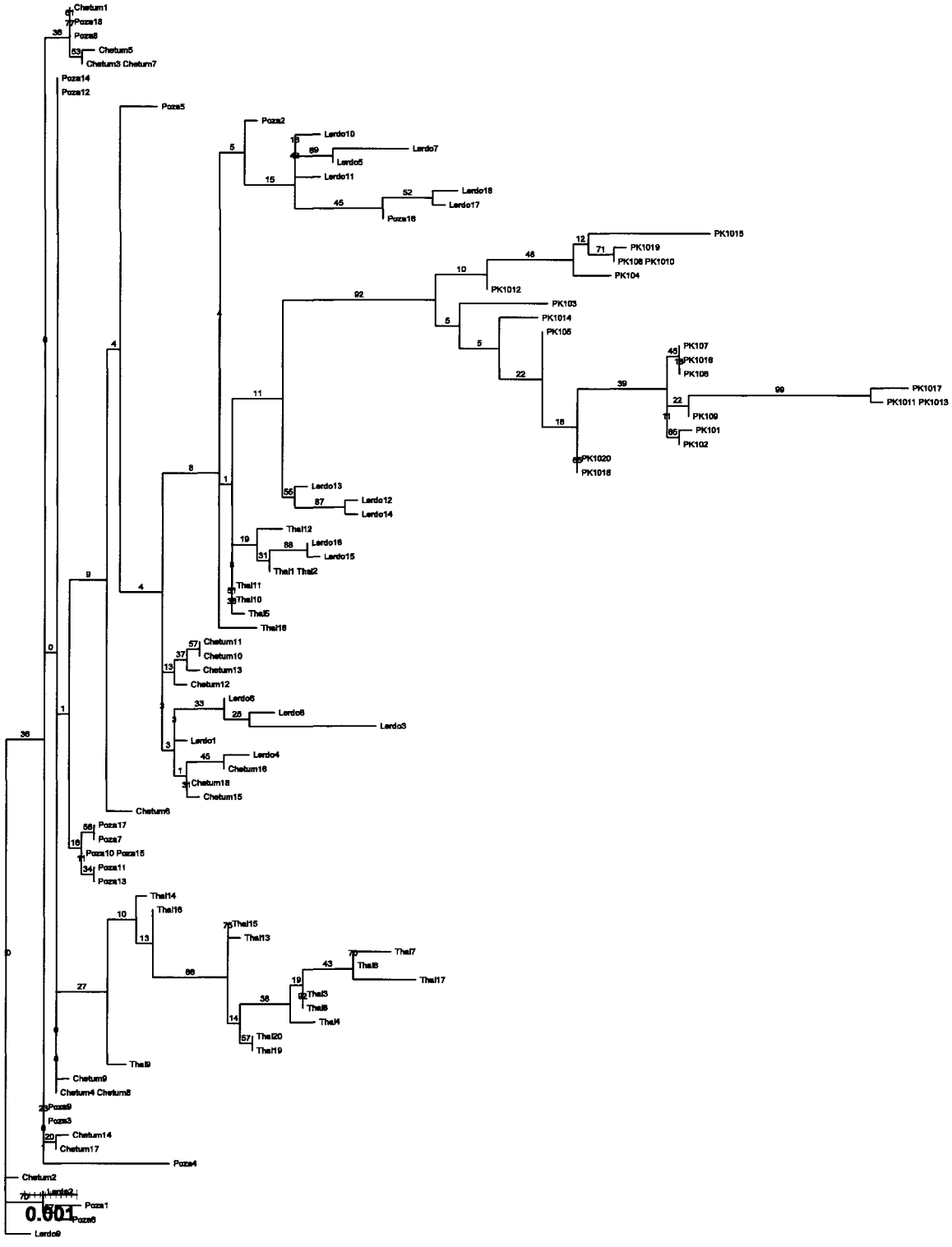


Figure 6.40. *r3d1* maximum parsimony phylogenetic tree. Equally weighted parsimony tree searches were conducted using 2,000 random addition tree-bisection-reconnection (TBR) searches in PAUP* version 4.0b10 (Swofford *et al.*, 2001).

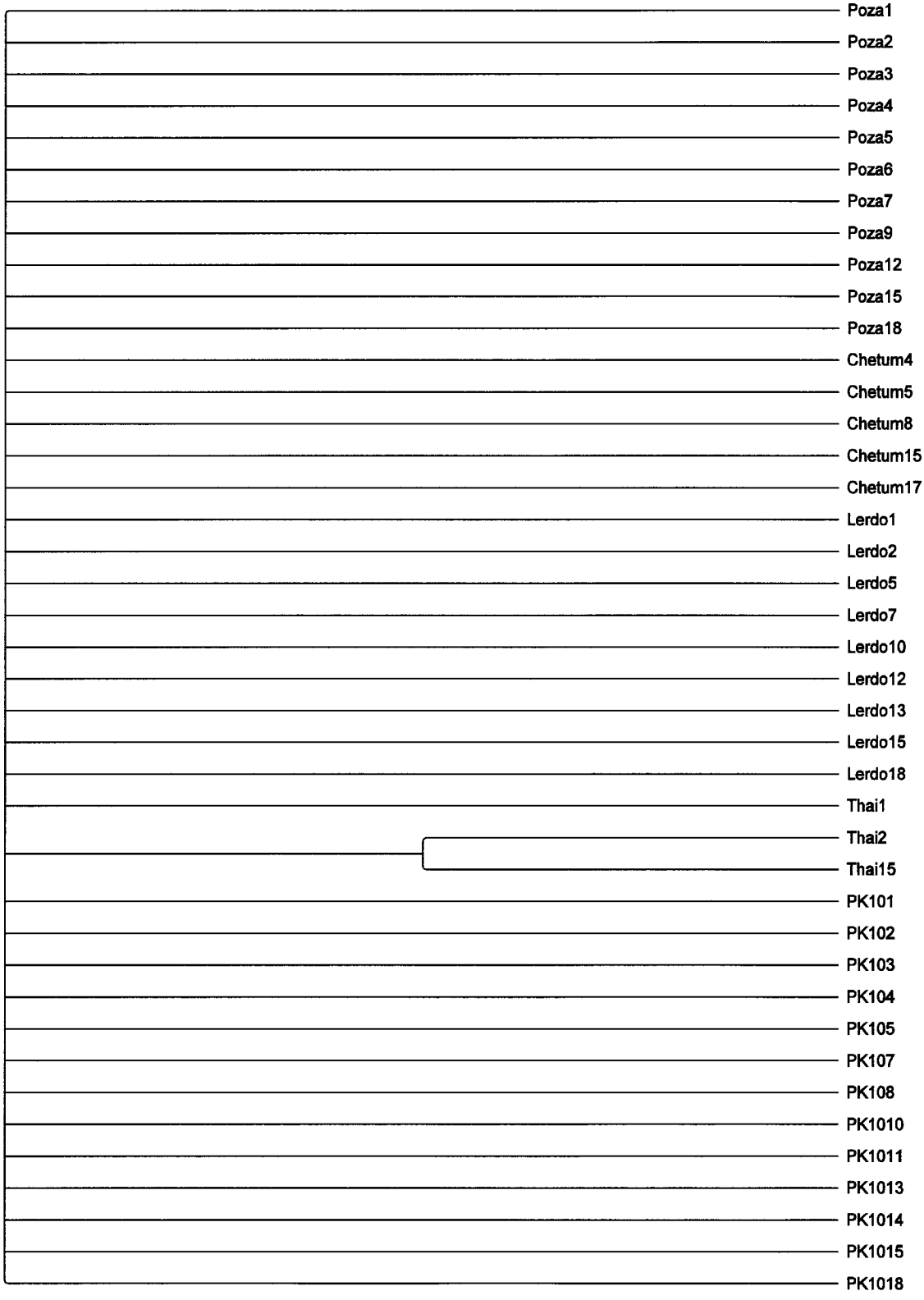


Figure 6.41. *r3d1* maximum likelihood phylogenetic tree. Likelihood analyses were conducted using RAxML version 7.03 (Stamatakis *et al.*, 2006). Optimal likelihood trees were searched for using 1,000 independent searches starting from randomized parsimony trees with the GTRGAMMA model and four discrete rate categories. The numbers on each of the branches represent bootstrap values.

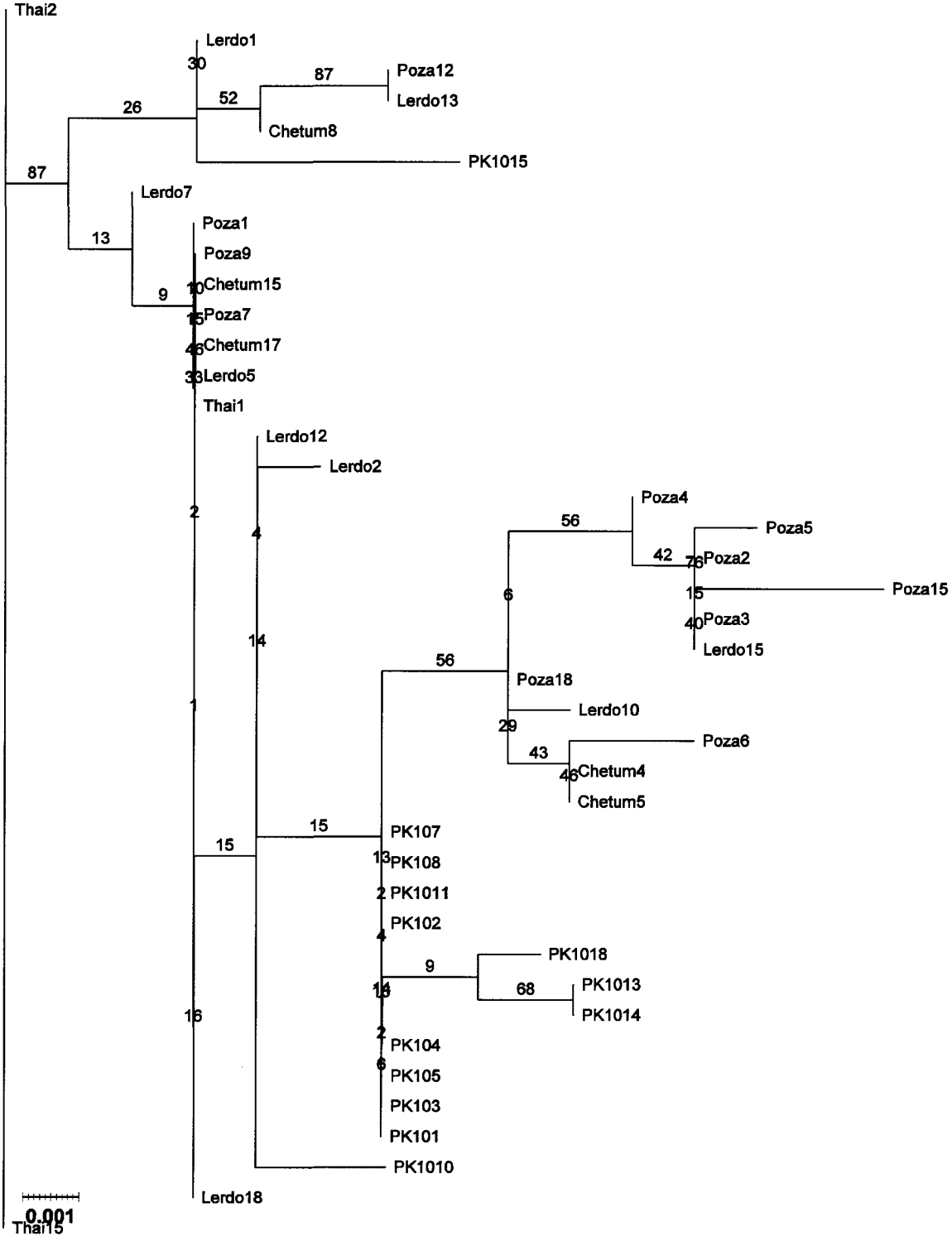


Figure 6.42. *r2d2* maximum parsimony phylogenetic tree. Equally weighted parsimony tree searches were conducted using 2,000 random addition tree-bisection-reconnection (TBR) searches in PAUP* version 4.0b10 (Swofford *et al.*, 2001).

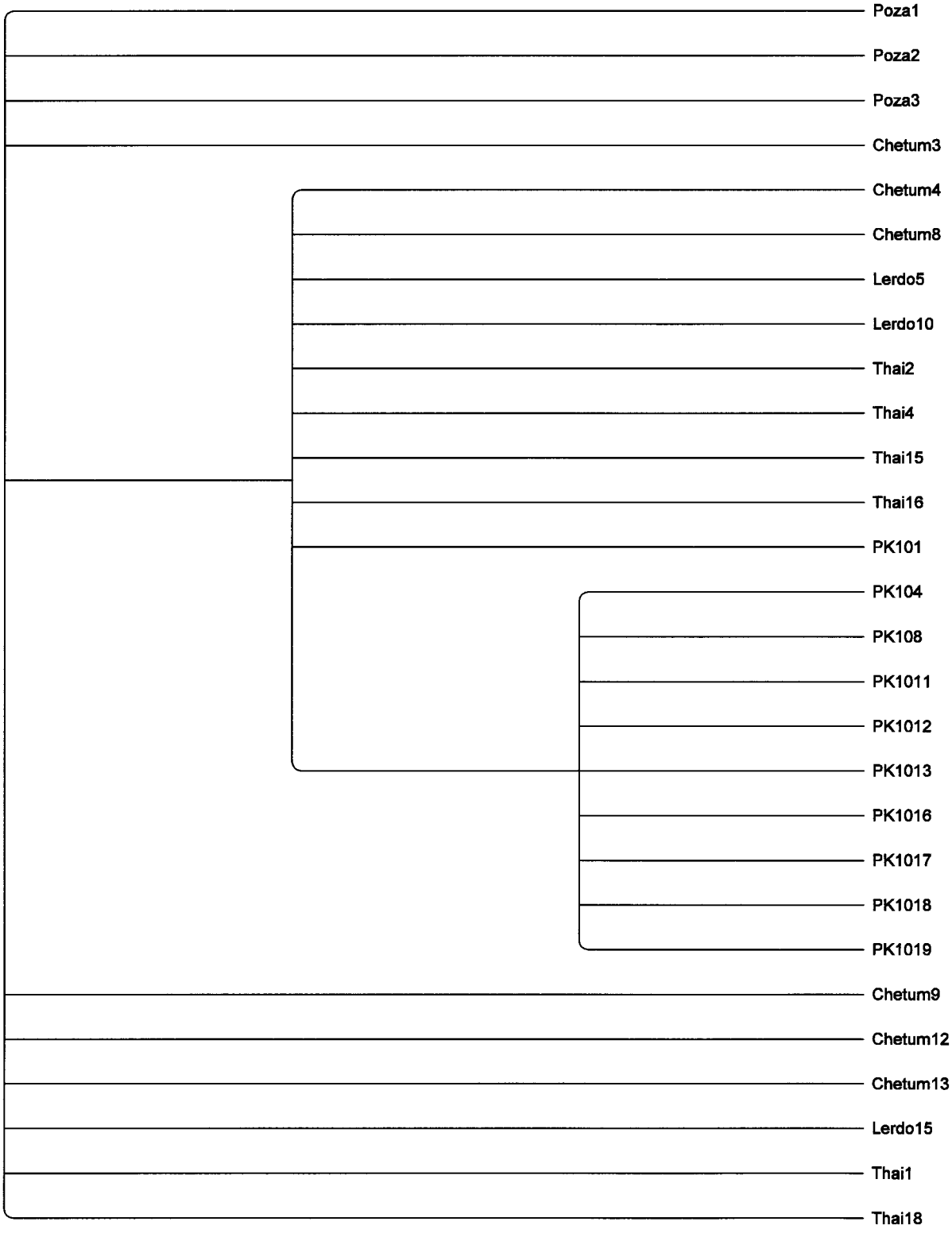
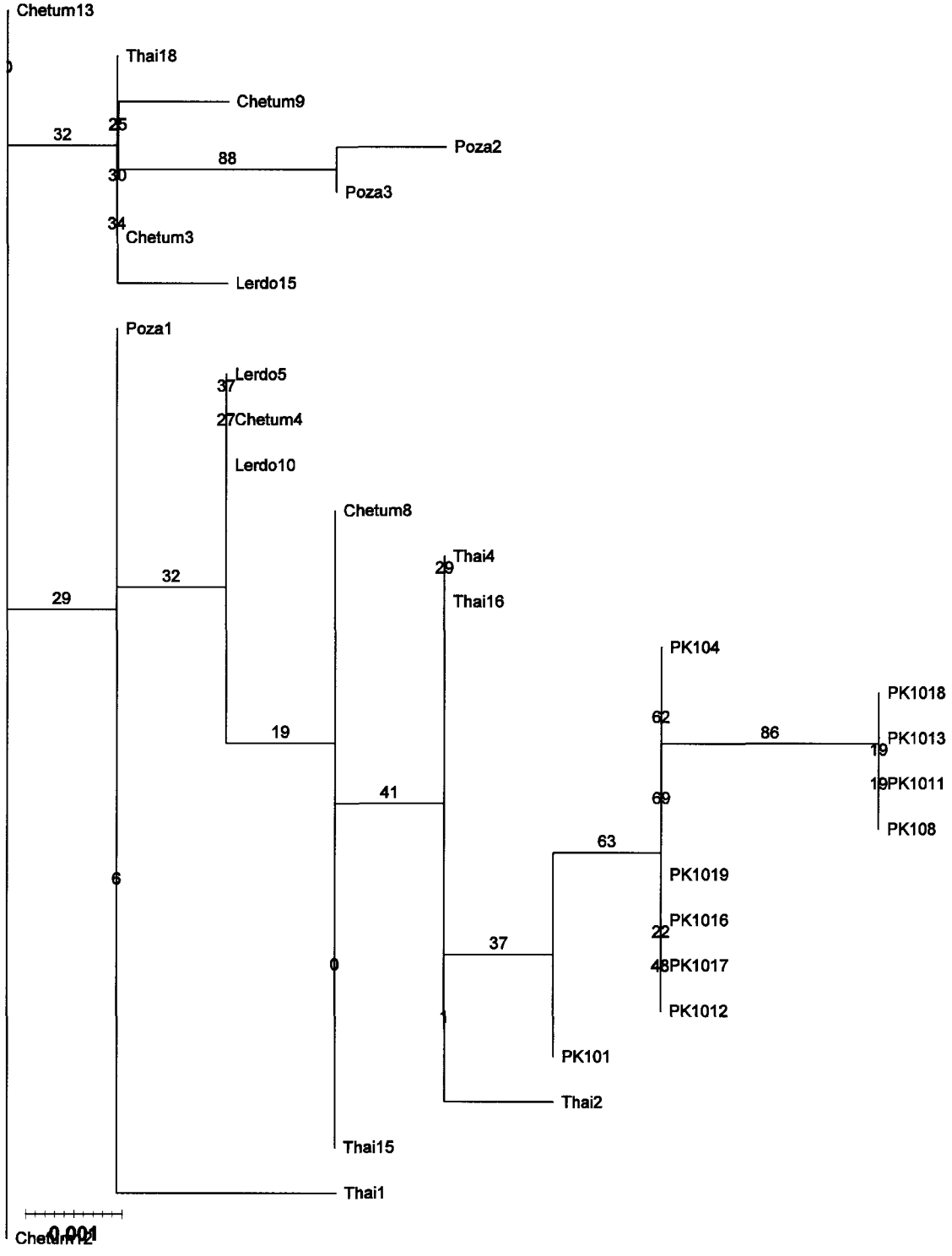


Figure 6.43. *r2d2* maximum likelihood phylogenetic tree. Likelihood analyses were conducted using RAxML version 7.03 (Stamatakis *et al.*, 2006). Optimal likelihood trees were searched for using 1,000 independent searches starting from randomized parsimony trees with the GTRGAMMA model and four discrete rate categories. The numbers on each of the branches represent bootstrap values.



Chapter 7

Conclusions

Summary and future directions

Dengue viruses continue to be among the most rapidly expanding agents of infectious disease in the tropics and have been estimated to cause 50 to 100 million cases of dengue fever and 500,000 cases of DHF each year (Gubler, 2004; Guzman and Kouri; 2003). It is essential to understand the interactions between DENVs and their arthropod vector, *Aedes aegypti* to develop new strategies to reduce transmission and disease burden. Understanding reasons for variation in *Ae. aegypti* with respect to vector competence (VC) for DENV-2 can provide important information and can guide effective vector control implementation in an economically limited environment. Molecular components of VC variation can provide researchers the necessary tools and approaches to develop new vector control methods, such as genetically modified resistant organisms. Understanding how the vector innate immune system affects VC may provide insight as to how effective or ineffective specific transgenic technologies may be within a natural population. In this dissertation, I investigated VC variability within natural populations over time, the genetic factors that contribute to VC, and the rate at which these factors are evolving in different populations.

My goal in Chapter 2 was to determine *Ae. aegypti* VC variation in populations from specific collection sites in Mexico where analyses were previously conducted. We confirmed that VC in *Ae. aegypti* for DENV-2 is highly variable between natural field collections in Mexico. VC in these collections ranged from 38-83%, the MIB rates ranged from 9-30%, and the MEB rates ranged from 7-36%. We also confirmed that mosquitoes collected on the east side of the Yucatan peninsula (e.g. Chetumal and Cancun) are highly susceptible to DENV-2. Lozano-Fuentes *et al.* (2009) identified an

area in Veracruz (Alvarado) where mosquitoes collected in 2003 exhibited refractoriness to DENV-2 infection. I confirmed that mosquitoes from this area (e.g. Alvarado, Lerdo de Tejada) continue to exhibit this phenotype, even though it is not as distinct from other populations as previously reported. These results reinforce the findings from previous studies that the VC phenotype is a quantitative trait that is probably highly polymorphic and demonstrates a wide range of variability.

Analysis of variance was for VC, MIB and MEB with respect to the Neovolcanic axis (NVA). Most of the variation in VC and MIB were attributed to year of collection, rather than position north or south of the NVA. Most variation in MEB was due to whether the collections were made north or south of the NVA, rather than year of collection.

Analysis of variance was also conducted for the three previously published VC studies (e.g. 1998-1999, 2001, 2003) to compare findings in this study for VC, MIB and MEB with respect to collection sites and years of collection. Most of the variations in VC and MEB were attributed to differences in collection sites rather than year of collection. Most variation in MIB was attributed to year of collection rather than differences in collection sites. A number of variables, ranging from genetics to environment, have been proposed as possible contributors to variation in VC.

Future directions should include a follow-up population genetics gene flow study of *Ae. aegypti* from the same collection sites to extend comparisons over time. We began to do gene flow analysis with the data in Chapter 2 but encountered an interesting finding regarding the mitochondrial gene marker, nicotinamide adenine dinucleotide dehydrogenase subunit 4 (ND4), that has traditionally been used in mosquito population

genetic studies (Black and Duteau; 1997; Bosio *et. al.*; 2000, Bennett *et. al.*, 2002; Gorrochotegui-Escalante *et. al.*, 2002). In Chapter 3 I describe how sequences from this gene, as well as numerous other mitochondrial DNA (mtDNA) sequences, were found to be integrated into the nuclear genome. Nuclear copies of mitochondrial genome sequences (NUMTs) have been found in other insect species such as *Drosophila* and honeybees (Pamilo *et al.*, 2007). They have never been observed in *Ae. aegypti* before. Numerous BLAST searches and analyses confirmed that certain mitochondrial DNA sequences have integrated into the *Ae. aegypti* nuclear genome, representing the second highest density of NUMTs in an insect genome and highest in Diptera reported to date. Analyses of flanking sequences suggest that *Ae. aegypti* NUMTs arose through mtDNA breakage and nonhomologous recombination, rather than through duplicative processes such as transposition or molecular drive.

My goal in Chapter 4 was to identify specific genes or mechanisms that contribute to MIB and MEB, quantitative traits of DENV-2 susceptibility in *Ae. aegypti*. We hypothesized that RNAi, microRNA and/or apoptosis pathway genes significantly impact DENV-2 infection and transmission in *Ae. aegypti*. We constructed an F_1 intercross family using a P_1 female from a highly susceptible to DENV-2 *Ae. aegypti aegypti* (Aaa) laboratory strain (*D2S3*) and a P_1 male from a highly refractory to DENV-2 *Ae. aegypti formosus* (Aaf) field collection (PK10). We observed that recombination rates in the F_2 offspring were severely reduced and the genotype ratios suggested a deleterious recessive allele on chromosome 3. The F_2 linkage map was incongruent with the established map for Aaa. Furthermore, no increased recombination was detected in F_5 offspring. Recombination rates and gene order were consistent with the presence of at least three

large inversions on chromosome 1, a single small inversion on chromosome 2, and a large inversion on chromosome 3 in Aaf.

Even though I was unable to construct an accurate linkage and QTL map of RNAi pathway genes in the generated F_1 intercross *Ae. aegypti* family, I conducted genotype-phenotype analysis, using χ^2–goodness-of-fit test, on each of the loci that were used to generate the linkage map (Chapter 5). I determined that *dcr2*, *r2d2*, *r3d1*, *ago2*, and IAP2 all significantly contribute to the MEB for DENV-2. *dcr2*, *r2d2*, and *ago2* have been identified as important components of the RNAi pathway and my study showed that they are significant contributors to MEB for DENV-2. A number of future directions can be drawn from the findings of Chapter 5. First, it is important to construct an accurate linkage map for the RNAi genes and use it to conduct QTL mapping. The χ^2–goodness-of-fit test is a good preliminary analysis that is specific to the locus and phenotype of interest and is independent of other nearby loci, but to understand the role of RNAi in VC, it will be necessary to construct a QTL map with these genes of interest.

Another future direction suggested by these findings is to quantitate expression levels of the RNAi genes. Previous findings demonstrated that manipulation of the RNAi pathway alters DENV-2 infection and dissemination in *Ae. aegypti*. Our findings confirm that RNAi genes contribute to the MEB phenotype. Examining gene expression levels will add to understanding why DENV-2 VC varies in different mosquito populations. Another future study will be to determine whether the inversions on the *Ae.aegypti formosus* chromosomes significantly contribute to MIB and/or MEB. Data suggest that the inversion break points on all three chromosomes are in close proximity to some of the RNAi pathway genes.

My final research goal (Chapter 6) was to determine the rate at which RNAi genes are evolving relative to microRNA pathway genes and to determine whether differences in VC are related to variation in rates of RNAi evolution in *Ae. aegypti*. We determined through non-synonymous to synonymous amino acid substitution rates and phylogenetic analyses that the RNAi genes are evolving faster than their microRNA paralogs. These results confirmed previous findings in various *Drosophila* species. I also hypothesized that we would see different rates of RNAi pathway gene evolution in *Ae. aegypti* populations that demonstrate different VC rates for DENV-2. Through a phenotype correlation analysis, I found the average number of nucleotide differences in *dcr2* to be significantly correlated to the DENV-2 MEB in *Ae. aegypti*. A strong inverse correlation was shown between nucleotide diversity (π) in *ago2* and DENV-2 MEB in *Ae. aegypti*. Even though I found the existence of significant correlations between DENV-2 MEB and RNAi pathway gene evolution, my findings are not conclusive regarding relationship of susceptibility to dengue viruses and rates of RNAi pathway gene evolution in the mosquito. Questions remain as to what drives this high rate of evolution (i.e. internal mechanism or external exposure), the overall significance of this evolutionary rate and what infection with arboviruses may contribute to evolution.

Literature cited

Achary, P. M. R. (1994) A simple technique for the preparation of polytene chromosomes from *Culex quinquefasciatus*. *J Am Mosq Cont Association* 10: 112-114.

Acs G, Klett H, Schonberg M, Christman J, Levin DH, Silverstein SC (1971) Mechanism of reovirus double-stranded ribonucleic acid synthesis in vivo and in vitro. *J Virol.* Nov;8(5):684-9.

Adelman ZN, Sanchez-Vargas I, Travanty EA, Carlson JO, Beaty BJ, Blair CD, Olson KE (2002) RNA silencing of dengue virus type 2 replication in transformed C6/36 mosquito cells transcribing an inverted-repeat RNA derived from the virus genome. *J. Virol.* Dec;76(24):12925-33.

Adelman ZN, Blair CD, Carlson JO, Beaty BJ, Olson KE (2001) Sindbis virus-induced silencing of dengue viruses in mosquitoes. *Insect Mol Biol.* Jun;10(3):265-73.

Afonso CL, Tulman ER, Lu Z, Balinsky CA, Moser BA, Becnel JJ, Rock DL, Kutish GF (2001) Genome sequence of a baculovirus pathogenic for *Culex nigripalpus*. *J. Virol.* 75:11157-11165.

Aitken THG, Downs WG, Shope RE (1977) *Aedes aegypti* strain fitness for yellow fever virus transmission. *Am J Trop Med Hyg.* 26(5):985-990.

Akaike H (1973) Maximum likelihood identification of Gaussian autoregressive moving average models. *Biometrika.* 60(2):255-265.

Aliyari R, Wu Q, Li HW, Wang XH, Li F, Green LD, Han CS, Li WX, Ding SW (2008) Mechanism of induction and suppression of antiviral immunity directed by virus-derived small RNAs in *Drosophila*. *Cell Host Microbe.* Oct 16;4(4):387-97.

Alto BW, Reiskind MH, Lounibos LP. (2008) Size alters susceptibility of vectors to dengue virus infection and dissemination. *Am J Trop Med Hyg.* Nov;79(5):688-95.

Alto BW, Lounibos LP, Mores CN, Reiskind MH (2008) Larval competition alters susceptibility of adult *Aedes* mosquitoes to dengue infection. *Proc Biol Sci.* Feb 22;275(1633):463-71.

Andreadis TG (1981) A new cytoplasmic polyhedrosis virus from the salt-marsh mosquito, *Aedes cantator* (Diptera: Culicidae). *J. Invertebr. Pathol.* 37:160-167.

Andreadis TG (1986) Characterization of a cytoplasmic polyhedrosis virus affecting the mosquito *Culex restuans*. *J. Invertebr. Pathol.* 47:194-202.

Andreadis TG, Becnel JJ, White SE (2003) Infectivity and pathogenicity of a novel baculovirus, CuniNPV from *Culex nigripalpus* (Diptera: Culicidae) for thirteen species and four genera of mosquitoes. *J Med Entomol.* Jul;40(4):512-7.

Antolin MF, Bosio CF, Cotton J, Sweeney W, Strand MR, Black WC 4th (1996) Intensive linkage mapping in a wasp (*Bracon hebetor*) and a mosquito (*Aedes aegypti*) with single-strand conformation polymorphism analysis of random amplified polymorphic DNA markers. *Genetics.* Aug;143(4):1727-38.

Apostol BL, Black WC 4th, Reiter P, Miller BR (1996) Population genetics with RAPD-PCR markers: the breeding structure of *Aedes aegypti* in Puerto Rico. *Heredity.* Apr;76 (Pt 4):325-34.

Araújo JM, Nogueira RM, Schatzmayr HG, Zanotto PM, Bello G (2008) Phylogeography and evolutionary history of dengue virus type 3. *Infect Genet Evol.* Oct 25. [Epub ahead of print].

Arnoult D, Akarid K, Grodet A, Petit PX, Estaquier J, Ameisen JC (2002) On the evolution of programmed cell death: apoptosis of the unicellular eukaryote *Leishmania major* involves cysteine proteinase activation and mitochondrion permeabilization. *Cell Death Differ.* 9: 65-81.

Atrie, B., Subbarao, S. K., Pillai, M. K. K., Rao, S. R. V. and Sharma, V. P. (1999) Population cytogenetic evidence for sibling species in *Anopheles annularis* (Diptera : Culicidae). *Ann Entomol Soc Am* 92: 243-24

Attoui H, Mohd Jaafar F, Belhouchet M, Biagini P, Cantaloube JF, de Micco P, de Lamballerie X (2005) Expansion of family Reoviridae to include nine-segmented dsRNA viruses: isolation and characterization of a new virus designated *Aedes pseudoscutellaris* reovirus assigned to a proposed genus (Dinovernavirus). *Virology.* Dec 20;343(2):212-23.

Babu R, Nair SK, Kumar A, Rao HS, Verma P, Gahalain A, Singh IS, Gupta HS (2006) Mapping QTLs for popping ability in a popcorn x flint corn cross. *Theor Appl Genet.* May;112(7):1392-9.

Banerjee AK, Shatkin AJ (1970) Transcription in vitro by reovirus-associated ribonucleic acid-dependent polymerase. *J Virol.* Jul;6(1):1-11.

Baqar S, Hayes CG, Ahmed T (1980) The effect of larval rearing conditions and adult age on the susceptibility of *Culex tritaeniorhynchus* to infection with West Nile virus. *Mosquito News.* 40:165.

Barreau C, Jousset FX, Bergoin M (1996) Pathogenicity of the *Aedes albopictus* parvovirus (AaPV), a denso-like virus for *Aedes aegypti* mosquitoes. *J. Invertebr. Pathol.* 68:299-309.

Barrett AD, Higgs S (2007) Yellow fever: a disease that has yet to be conquered. *Annu Rev Entomol.* 52:209-29.

Barillas-Mury C, Kumar S (2005) Plasmodium-mosquito interactions: a tale of dangerous liaisons. *Cell Microbiol.* Nov;7(11):1539-45.

Bartholomay LC, Cho WL, Rocheleau TA, Boyle JP, Beck ET, Fuchs JF, Liss P, Rusch M, Butler KM, Wu RC, Lin SP, Kuo HY, Tsao IY, Huang CY, Liu TT, Hsiao KJ, Tsai SF, Yang UC, Nappi AJ, Perna NT, Chen CC, Christensen BM (2004) Description of the transcriptomes of immune response-activated hemocytes from the mosquito vectors *Aedes aegypti* and *Armigeres subalbatus*. *Infect Immun.* Jul;72(7):4114-26.

Bates M., Roca-Garcia M (1946) The development of the virus of yellow fever in *Haemagogus* mosquitoes. *Am. J. Trop. Med. Hyg.* 26:585.

Baulcomb D (2004) RNA silencing in plants. *Nature* 431:356-363.

Beaty BJ (2000) Genetic manipulation of vectors: A potential novel approach for control of vector-borne diseases. *Proc Natl Acad Sci* Sep 12;97(19):10295-7.

Beck ET, Blair CD, Black WC 4th, Beaty BJ, Blitvich BJ (2007) Alternative splicing generates multiple transcripts of the inhibitor of apoptosis protein 1 in *Aedes* and *Culex* spp. mosquitoes. *Insect Biochem Mol Biol.* Nov;37(11):1222-33.

Becker W.A. (1992) *Manual of Quantitative Genetics*, 5th ed. Academic Enterprises, Pullman, WA.

Becnel JJ, Fukuda T (1989) Natural occurrence of a mosquito iridescent virus in *Aedes sollicitans*. *J. Am. Mosq. Control Assoc.* 5:610-611.

Becnel J, White S, Moser B, Fukuda T, Rotstein M, Undeen A, Cockburn A (2001) Epizootiology and transmission of a newly discovered baculovirus from the mosquitoes *Culex nigripalpus* and *C. quinquefasciatus*. *J Gen Virol.* Feb;82(Pt 2):275-82.

Becnel JJ, White SE (2007) Mosquito pathogenic viruses--the last 20 years. *J Am Mosq Control Assoc.* 23(2 Suppl):36-49.

Beerntsen, B. T., Severson, D. W., Klinkhammer, J. A., Kassner, V. A. and Christensen, B. M. (1995) Aedes aegypti: A quantitative trait locus (QTL) influencing filarial worm intensity is linked to QTL for susceptibility to other mosquito-borne pathogens. *Exper Parasit* 81: 355-362.

Behura, S. K., Valicente, F. H., Rider, S. D., Shun-Chen, M., Jackson, S. and Stuart, J. J. (2004) A physically anchored genetic map and linkage to avirulence reveals recombination suppression over the proximal region of hessian fly chromosome A2. *Genetics* 167: 343-355.

Belloncik S, Mori H (1998) Cypoviruses. In: Miller LK, Ball LA, eds. *The Insect Viruses*. New York: The Insect Viruses. pp. 337-370.

Bennett KE, Olson KE, Muñoz Mde L, Fernandez-Salas I, Farfan-Ale JA, Higgs S, Black WC 4th, Beaty BJ (2002) Variation in vector competence for dengue 2 virus among 24 collections of *Aedes aegypti* from Mexico and the United States. *Am J Trop Med Hyg.* Jul;67(1):85-92.

Bennett, K. E., Beaty, B. J. and Black, W. C. (2005a) Selection of D2S3, an *Aedes aegypti* (Diptera : Culicidae) strain with high oral susceptibility to Dengue 2 virus and D2MEB, a strain with a midgut barrier to Dengue 2 escape. *J Med Entomol* 42: 110-119.

Bennett, K. E., Flick, D., Fleming, K. H., Jochim, R., Beaty, B. J. and Black, W. C. (2005b) Quantitative trait loci that control dengue-2 virus dissemination in the mosquito *Aedes aegypti*. *Genetics* 170: 185-194.

Bensasson, D., Feldman, M. W. and Petrov, D. A. (2003) Rates of DNA duplication and mitochondrial DNA insertion in the human genome. *J. Mol. Evol.* 57: 343-354.

Bensasson, D., Zhang, D. X., Hartl, D. L. and Hewitt, G. M. (2001) Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends Ecol. and Evol.* 16: 314-321.

Bensasson, D., Zhang, D. X. and Hewitt, G. M. (2000) Frequent assimilation of mitochondrial DNA by grasshopper nuclear genomes. *Mol. Biol. and Evol.* 17: 406-415.

Bernstein E, Denli AM, Hannon GJ (2001) The rest is silence. *RNA.* Nov;7(11):1509-21.

Bishop JG, Dean AM, Mitchell-Olds T (2000) Rapid evolution in plant chitinases: molecular targets of selection in plant-pathogen coevolution. *Proc Natl Acad Sci* U S A. May 9;97(10):5322-7.

Black, W. C. and DuTeau, N. M. (1997) RAPD-PCR and SSCP Analysis for insect population genetic studies. The Molecular Biology of Insect Disease Vectors: A Methods Manual 361-373.

Black WC 4th, Gorrochetegui-Escalante N, Randle NP, DonnellyMJ (2008) The Yin and Yange of linkage disequilibrium: mapping of genes and nucleotides conferring insecticide resistance in insect disease vectors. *Adv Exp Med Biol.* 627:71-83.

Black WC 4th, Baer CF, Antolin MF, DuTeau NM (2001) Population genomics: genome-wide sampling of insect populations. *Annu Rev Entomol.* 46:441-69.

Black WC 4th, Bennett KE, Gorrochótegui-Escalante N, Barillas-Mury CV, Fernández-Salas I, de Lourdes Muñoz M, Farfán-Alé JA, Olson KE, Beaty BJ (2002) Flavivirus susceptibility in *Aedes aegypti*. *Arch Med Res.* Jul-Aug;33(4):379-88.

Black W.C. and Severson D (2005) Genetics of Vector Competence. In *The Biology of Disease Vectors*. University Press of Colorado., pp. 389-464.

Blaszczyk J, Tropea JE, Bubunenko M, Routzahn KM, Waugh DS, Court DL, Ji X (2001) Crystallographic and modeling studies of RNase III suggest a mechanism for double-stranded RNA cleavage. *Structure*. Dec;9(12):1225-36.

Blitvich BJ, Blair CD, Kempf BJ, Hughes MT, Black WC, Mackie RS, Meredith CT, Beaty BJ and Rayms-Keller A (2002) Developmental- and tissue-specific expression of an inhibitor of apoptosis protein 1 homologue from *Aedes triseriatus* mosquitoes. *Insect Molecular Biology* 11(5): 431-442.

Bosio CF, Beaty BJ, Black WC 4th (1998) Quantitative genetics of vector competence for dengue-2 virus in *Aedes aegypti*. *Am J Trop Med Hyg.* Dec;59(6):965-70.

Bosio, C. F., Harrington, L. C., Jones, J. W., Sithiprasasna, R., Norris, D. E. and Scott, T. W. (2005) Genetic structure of *Aedes aegypti* populations in Thailand using mitochondrial DNA. *Am. J. Trop. Med. and Hyg.* 72: 434-442.

Bosio, C. F., Fulton, R. E., Salasek, M. L., Beaty, B. J. and Black, W. C. (2000) Quantitative trait loci that control vector competence for dengue-2 virus in the mosquito *Aedes aegypti*. *Genetics* 156: 687-698.

Bowers DF, Coleman CG, Brown DT. (2003) Sindbis virus-associated pathology in *Aedes albopictus* (Diptera: Culicidae). *J Med Entomol.* Sep;40(5):698-705.

Brackney DE, Foy BD, Olson KE (2008) The effects of midgut serine proteases on dengue virus type 2 infectivity of *Aedes aegypti*. *Am J Trop Med Hyg.* Aug;79(2):267-74.

Brandt WE, McCown JM, Gentry MK, Russell PK (1982) Infection enhancement of dengue-2 virus in the U937 human monocyte cell line by antibodies to flavivirus cross-reactive determinant. *Infect. Immun.* 36:1036-1041.

Brookes AJ (1999) The essence of SNPs. *Genetics* Jul 8;234(2):177-86.

Brown SE, Knudson DL (1997) FISH landmarks for *Aedes aegypti* chromosomes. *Insect Mol. Biol.* 6:197-202.

Brown SE, Severson DW, Smith LA, Knudson DL (2001) Integration of the *Aedes aegypti* mosquito genetic linkage and physical maps. *Genetics* 157: 1299-1305.

Bryant B, Blair CD, Olson KE, Clem RJ (2008) Annotation and expression profiling of apoptosis-related genes in the yellow fever mosquito, *Aedes aegypti*. *Insect Biochem Mol Biol.* Mar;38(3):331-45.

Bushati N, Cohen SM (2007) MicroRNA functions. *Annu. Rev. Cell Dev. Biol.* 23:175-205.

Buchatsky LP, Kuznetsova MA, Lebedinets NN, Konoko AG (1987) Development and basic properties of the viral preparation viroden. *Voprosy Virusologii* 32:729-799.

Buchatsky LP (1989) Densonucleosis of bloodsucking mosquitoes. *Dis. Aquat. Org.* 6:145-150.

Buchatsky LP, Raikova AP (1979) Electron microscope study of mosquito densonucleosis virus maturation. *Acta. Virol.* 23:170-172.

Buchen-Osmond C (2003) The universal virus database ICTVdB. *Computing Sci. Eng.* 5:16-25.

Burivong P, Pattanakitsakul SN, Thongrungkiat S, Malasit P, Flegel TW (2004) Markedly reduced severity of Dengue virus infection in mosquito cell cultures persistently infected with *Aedes albopictus* densovirus (AalDNV). *Virology.* 24;329(2):261-9.

Burke D.S., Monath T.P. (2001) Flaviviruses, p. 1043-1126, In D.M. Knipe, and P.M. Howely (ed), *Fields Virology 4th* ed. Lippincott Williams and Wilkins, Philadelphia, PA.

Cáceres C, Yung V, Araya P, Tognarelli J, Villagra E, Vera L, Fernández J (2008) Complete nucleotide sequence analysis of a Dengue-1 virus isolated on Easter Island, Chile. *Arch Virol.* 153(10):1967-70.

Campbell CL, Keene KM, Brackney DE, Olson KE, Blair CD, Wilusz J, Foy BD (2008) *Aedes aegypti* uses RNA interference in defense against Sindbis virus infection. *BMC Microbiol.* Mar 17;8:47.

Campos, J., Andrade, C. F. S. and Recco-Pimentel, S. M. (2003a) Malpighian tubule polytene chromosomes of *Culex quinquefasciatus* (Diptera, Culicinae). *Memorias Do Instituto Oswaldo Cruz* 98: 383-386.

Campos, J., Andrade, C. F. S. and Recco-Pimentel, S. M. (2003b) A technique for preparing polytene chromosomes from *Aedes aegypti* (Diptera, Culicinae). *Memorias Do Instituto Oswaldo Cruz* 98: 387-390

Carey D.E. (1971) Chikungunya and dengue: a case of mistaken identity? *J. Hist. Med. Allied Sci.* 26:243-262.

Carlson J, Olson K, Higgs S, Beaty BJ (1995) Molecular genetic manipulation of mosquito vectors. *Annu Rev Entomol.* 40:359-388.

Carthew RW (2006) Gene regulation by microRNAs. Curr. *Opin. Genet. Dev.* 16:203-208.

Centers for Disease Control and Prevention (1981) *Manual on dengue diagnostic laboratory procedures for the Americas.*

Chamberlain RW, Sudia WD (1955) The effects of temperature upon the extrinsic incubation of eastern equine encephalitis in mosquitoes. *Am. J. Hyg.* 62:295.

Chambers EW, Lovin DD, Severson DW (2003) Utility of comparative anchor-tagged sequences as physical anchors for comparative genome analysis among the Culicidae. *Am J Trop Med Hyg* 69: 98-104.

Chambers EW, Meece JK., McGowan JA., Lovin DD., Hemme RR., Chadee DD, McAbee K, Brown SE, Knudson DL, Severson DW (2007) Microsatellite isolation and linkage group identification in the yellow fever mosquito *Aedes aegypti*. *J Hered* 98: 202-210.

Chapman EJ, Carrington JC (2007) Specialization and evolution of endogenous small RNA pathways. *Nat. Rev. Genet.* 8:884-896.

Chen HL, Lin SR, Liu HF, King CC, Hsieh SC, Wang WK (2008) Evolution of dengue virus type 2 during two consecutive outbreaks with an increase in severity in southern Taiwan in 2001-2002. *Am J Trop Med Hyg.* Oct;79(4):495-505.

Christophers SR (1960) *Aedes aegypti (L.)* The yellow fever mosquito. Cambridge University Press, London.

Christophides GK, Zdobnov E, Barillas-Mury C, Birney E, Blandin S, Blass C, Brey PT, Collins FH, Danielli A, Dimopoulos G, Hetru C, Hoa NT, Hoffmann JA, Kanzok SM, Letunic I, Levashina EA, Loukeris TG, Lycett G, Meister S, Michel K, Moita LF, Müller HM, Osta MA, Paskewitz SM, Reichhart JM, Rzhetsky A, Troxler L, Vernick KD, Vlachou D, Volz J, von Mering C, Xu J, Zheng L, Bork P, Kafatos FC (2002) Immunity-related genes and gene families in *Anopheles gambiae. Science.* 2002 Oct 4;298(5591):159-65.

Chow VT, Chan YC, Yong R, Lee KM, Lim LK, Chung YK, Lam-Phua SG, Tan BT (1998) Monitoring of dengue viruses in field-caught *Aedes aegypti* and *Aedes albopictus* mosquitoes by a type-specific polymerase chain reaction and cycle sequencing. Am J Trop Med Hyg. May;58(5):578-86.

Chung W-J, Okamura K, Martin R, Lai EC (2008) Endogenous RNA interference provides a somatic defense against Drosophila transposons. *Curr. Biol.* 18:795-802.

Chung, W. K. and Steiper, M. E. (2008) Mitochondrial COII Introgression into the Nuclear Genome of *Gorilla gorilla. Internat. J. Primat.* 29: 1341-1353.

Chung YK, Pang FY (2002) Dengue virus infection rate in field populations of female Aedes aegypti and Aedes albopictus in Singapore. *Trop Med Int Health.* Apr;7(4):322-30.

Cirimotich C, Scott JC, Phillips AT, Geiss BJ, Olson KE (2009) Suppression of RNA interference increases alphavirus replication and virus-associated mortality in *Aedes aegypti* mosquitoes. *BMC Micro.* 5:9(1)49 Epub ahead of print.

Clark TB, Chapman HC (1969) A polyhedrosis in *Culex salinarius* of Louisiana. *J. Invertebr. Pathol.* 13:312.

Coluzzi M, Kitzmiller JB (1975) Anopheline mosquitoes. *Handbook of genetics.* Vol. 3. Invertebrates of genetic interest. pp. 285-309.

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

Cornel AJ, Collins FH (2000) Maintenance of chromosome arm integrity between two Anopheles mosquito subgenera. *J Hered* 91: 364-370.

da Costa-Ribeiro MC, Lourenço-de-Oliveira R, Failloux AB (2007) Low gene flow of *Aedes aegypti* between dengue-endemic and dengue-free areas in southeastern and southern Brazil. *Am J Trop Med Hyg.* Aug;77(2):303-9.

Costero A, Attardo GM, Scott TW, Edman JD (1998) An experimental study on the detection of fructose in *Aedes aegypti. J Am Mosq Control Assoc.* Sep;14(3):234-42.

Costero A, Edman JD, Clark GG, Scott TW (1998) Life table study of *Aedes aegypti* (Diptera: Culicidae) in Puerto Rico fed only human blood versus blood plus sugar. *J Med Entomol.* Sep;35(5):809-13.

Cox FE (2002) History of human parasitology. *Clin Microbiol Rev.* Oct;15(4):595-612.

Craig GB Jr., Vandehey RC, Hickey WA (1961) Genetic variability in populations of *Aedes aegypti*. *Bull World Health Organ.* 24:527-39.

Craig GB Jr, Hickey WA (1967) Current status of the formal genetics of *Aedes aegypti*. *Bull World Health Organ.* 36(4):559-62.

Czech B, Malone CD, Zhou R, Stark A, Schlingeheyde C, Dus M, Perrimon N, Kellis M, Wohlschlegel JA, Sachidanandam R, Hannon GJ, Brennecke J (2008) An endogenous small interfering RNA pathway in *Drosophila*. *Nature*. 453:729-731.

Darvasi A, Soller M (1995) Advanced intercross lines, an experimental population for fine genetic mapping. *Genetics*. Nov;141(3):1199-207.

Dawkins R, Krebs JR (1979) Arms races between and within species. *Proc R Soc Lond B Biol Sci*. Sep 21;205(1161):489-511.

Deddouche S, Matt N, Budd A, Mueller S, Kemp C, Galiana-Arnoux D, Dostert C, Antoniewski C, Hoffman JA, Imler JL. (2008) The DExD/H-box helicase Dicer-2 mediates the induction of antiviral activitiy in *drosophila*. *Nature Immunology* 9:12;1425-1432.

De Kruif P (1926) *Microbe Hunters*. Harcourt, Brace, NY.

Deubel V, Kinney RM, Trent DW (1986) Nucleotide sequence and deduced amino acid sequence of the structural proteins of dengue type 2 virus, Jamaica genotype. *Virology*. Dec;155(2):365-77.

Deubel V, Digoutte JP, Mattei X, Pandare D (1981) Morphogenesis of yellow fever virus in *Aedes aegypti* cultured cells. II. An ultrastructural study. *Am J Trop Med Hyg.* 30 (5): 1071-7.

della Torre A, Favia G, Mariotti G, Coluzzi M, Mathiopoulos KD (1996) Physical map of the malaria vector *Anopheles gambiae*. *Genetics*. Jul;143(3):1307-11.

Diallo M, Ba Y, Faye O, Soumare ML, Dia I, Sall AA (2008) Vector competence of Aedes aegypti populations from Senegal for sylvatic and epidemic dengue 2 virus isolated in West Africa. *Trans R Soc Trop Med Hyg*. May;102(5):493-8.

Diallo M, Sall AA, Moncayo AC, Ba Y, Fernandez Z, Ortiz D, Coffey LL, Mathiot C, Tesh RB, Weaver SC (2005) Potential role of sylvatic and domestic African mosquito species in dengue emergence. *Am J Trop Med Hyg*. Aug;73(2):445-9.

Díaz FJ, Black WC 4th, Farfán-Ale JA, Loroño-Pino MA, Olson KE, Beaty BJ (2006) Dengue virus circulation and evolution in Mexico: a phylogenetic perspective. *Arch Med Res*. Aug;37(6):760-73.

Dideco M, Cancrini G, Coluzzi M, Bullini APB, Cianchi R, Bullini L (1978) Linkage studies between chromosome inversions and enzyme loci in mosquito *Anopheles stephensi*. *Heredity* 40: 457-458.

Dimopoulos G, Zheng L, Kumar V, della Torre A, Kafatos FC, Louis C (1996) Integrated genetic map of *Anopheles gambiae*: use of RAPD polymorphisms for genetic, cytogenetic and STS landmarks. *Genetics*. 1996 Jun;143(2):953-60.

Ding S-W, Voinnet O (2007) Antiviral immunity directed by small RNAs. *Cell* 130:413.

Dohm DJ, Sardelis MR, Turell MJ (2002) Experimental vertical transmission of West Nile virus by *Culex pipiens* (Diptera: Culicidae). *J Med Entomol*. Jul;39(4):640-4.

Downs JA., Kosmidou E, Morgan A, Jackson SP (2003) Suppression of homologous recombination by the *Saccharomyces cerevisiae* linker histone. *Molecular Cell* 11: 1685-1692.

Drake JW (1993) Rates of spontaneous mutation among RNA viruses. *Proc Natl Acad Sci.* May 1;90(9):4171-5.

Drake JW, Holland JJ (1999) Mutation rates among RNA viruses. *Proc Natl Acad Sci U S A.* Nov 23;96(24):13910-3.

Drayna D, Fields BN (1982) Activation and characterization of the reovirus transcriptase: genetic analysis. *J Virol.* Jan;41(1):110-8.

Du TT, Zamore PD (2005) microPrimer: the biogenesis and function of microRNA. *Development* 132:4645-4652.

Ercan S., Reese JC, WorkmanJL, Simpson RT (2005) Yeast recombination enhancer is stimulated by transcription activation. *Mol Cell Biol* 25: 7976-7987.

Erlanger TE, Keiser J, Utzinger J (2008) Effect of dengue vector control interventions on entomological parameters in developing countries: a systematic review and meta-analysis. *Med Vet Entomol.* Sep;22(3):203-21.

Evan G, Littlewood T (1998) A matter of life and cell death. *Science* 281: 1317-1322.

Falconer DS, Mackay TFC (1996) *Introduction to Quantitative Genetics*, 4th ed. Longman Group Ltd., Essex U.K.

Faran ME, Turell MJ, Romoser WS, Routier RG, Gibbs PH, Cannon TL, Bailey CL (1987) Reduced survival of adult *Culex pipiens* infected with Rift Valley fever virus. *Am J Trop Med Hyg.* Sep;37(2):403-9.

Federici BA (1995) The future of microbial insecticides as vector control agents. *J Am Mosq Control Assoc.* Jun;11(2 Pt 2):260-8.

Federici BA (1980) Mosquito baculovirus: sequence of morphogenesis and ultrastructure of the virion. *Virology.* Jan 15;100(1):1-9.

Federici BA (1977) Virus pathogens of Culicidae (mosquitos). *Bull World Health Organ.* 55 Suppl 1:25-36.

Federici BA, Lowe RE (1972) Studies on the pathology of a baculovirus in *Aedes triseriatus. J Invertebr Pathol.* Jul;20(1):14-21.

Federici BA (1973) Preliminary studies of cytoplasmic polyhedrosis virus of *Aedes taeniorhynchus.* 5th Int. Colleq. Insect Pathol. Microbiol. Contr. 1:34 (Abstract)

Felsenstein J (1973) Maximum-likelihood estimation of evolutionary trees from continuous characters. *Am J Hum Genet.* Sep;25(5):471-92.

Felsenstein J (1978) Cases in which parsimony or compatibility methods will be positively misleading. *Syst. Zool.* 27:401-410.

Felsenstein J (1985) Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39:783-791.

Fisher, R. A. (1935) *The Design of Experiments.* Edinburgh: Oliver and Boyd.

Fire A (1999) RNA-triggered gene silencing. *Trends Genet.* Sep;15(9):358-63.

Förstemann K, Tomari Y, Du T, Vagin VV, Denli AM, Bratu DP, Klattenhoff C, Theurkauf WE, Zamore PD (2005) Normal microRNA maturation and germ-line stem cell maintenance requires Loquacious, a double-stranded RNA-binding domain protein. PLoS Biol. 2005 Jul;3(7):e236.

Franz AW, Sanchez-Vargas I, Adelman ZN, Blair CD, Beaty BJ, James AA, Olson KE (2006) Engineering RNA interference-based resistance to dengue virus type 2 in genetically modified *Aedes aegypti*. *Proc Natl Acad Sci U S A*. Mar 14;103(11):4198-203.

Fukuda T, Clark TB (1975) Transmission of the mosquito iridescent virus (RMIV) by adult mosquitoes of *Aedes taeniorhynchus* to their progeny. *J. Invertebr. Pathol.* 25:275-276.

Fulton, R. E., Salasek, M. L., DuTeau, N. M. and Black, W. C. (2001) SSCP analysis of cDNA markers provides a dense linkage map of the *Aedes aegypti* genome. *Genetics* 158: 715-726.

Furuichi Y, LaFiandra A, Shatkin AJ (1977) 5'-Terminal structure and mRNA stability. *Nature*. Mar 17;266(5599):235-9.

Gage KL, Kosoy MY (2005) Natural history of plague: perspectives from more than a century of research. *Annu Rev Entomol*. 50:505-528.

Galiana-Arnoux D, Dostert C, Schneemann A, Hoffmann JA, Imler JL (2006) Essential function in vivo for Dicer-2 in host defense against RNA viruses in *Drosophila*. *Nat Immunol*. Jun;7(6):590-7.

Garcia-Rejon J, Loroño-Pino MA, Farfan-Ale JA, Flores-Flores L, Del Pilar Rosado-Paredes E, Rivero-Cardenas N, Najera-Vazquez R, Gomez-Carro S, Lira-Zumbardo V, Gonzalez-Martinez P, Lozano-Fuentes S, Elizondo-Quiroga D, Beaty BJ, Eisen L (2008) Dengue virus-infected *Aedes aegypti* in the home environment. *Am J Trop Med Hyg*. Dec;79(6):940-950.

Garrett-Jones C. (1964) Prognosis for the interruption of malaria vectors in relation to epidemiological assessment. *Bull. WHO* 30:241-261.

Gaut BS, Lewis PO (1995) Success of maximum likelihood phylogeny inference in the four-taxon case. *Mol Biol Evol.* 1995 Jan;12(1):152-62.

Germer S, Higuchi R (1999) Single-tube genotyping without oligonucleotide probes. *Genome Res*. Jan;9(1):72-8.

Ghildiyal M, Zamore PD (2009) Small silencing RNAs: an expanding universe. *Nat Rev Genet.* Feb;10(2):94-108.

Girard YA, Popov V, Wen J, Han V, Higgs S (2005) Ultrastructural study of West Nile virus pathogenesis in *Culex pipiens quinquefasciatus* (Diptera: Culicidae). *J Med Entomol.* May;42(3):429-44.

Girard YA, Schneider BS, McGee CE, Wen J, Han VC, Popov V, Mason PW, Higgs S (2007) Salivary gland morphology and virus transmission during long-term cytopathologic West Nile virus infection in *Culex* mosquitoes. *Am J Trop Med Hyg*. Jan;76(1):118-28.

Gomez-Machorro C, Bennett KE, Munoz MD, Black WC (2004) Quantitative trait loci affecting dengue midgut infection barriers in an advanced intercross line of Aedes aegypti. *Insect Mol Biol* 13: 637-648.

Gorrochotegui-Escalante N, Munoz ML, Fernandez-Salas I, Beaty BJ, Black WC 4th (2000) Genetic isolation by distance among Aedes aegypti populations along the northeastern coast of Mexico. *Am J Trop Med Hyg.* Feb;62(2):200-9.

Gorrochotegui-Escalante N, Gomez-Machorro C, Lozano-Fuentes S, Fernandez-Salas L, De Lourdes Munoz M, Farfan-Ale JA, Garcia-Rejon J, Beaty BJ, Black WC 4th

(2002) Breeding structure of *Aedes aegypti* populations in Mexico varies by region. *Am J Trop Med Hyg.* Feb;66(2):213-22.
Graham H (1903) The Dengue. A study of its pathology and mode of propagation. *J. Trop. Med. Hyg.* Londan 6:209-214.
Green CA, Rattanarithikul R, Charoensub A (1992) Population genetic confirmation of species status of the malaria vectors *Anopheles willmori* and *An. pseudowillmori* in Thailand and chromosome phylogeny of the Maculatus Group of Mosquitos. *Med Vet Entomol* 6: 335-341.
Grimstad PR, Ross QE, Craig GB Jr (1980) *Aedes triseriatus* (Diptera: Culicidae) and La Crosse virus. II. Modification of mosquito feeding behavior by virus infection. *J Med Entomol.* Jan 31;17(1):1-7.
Grimstad PR, Haramis LD (1984) *Aedes triseriatus* (Diptera: Culicidae) and La Crosse virus. III. Enhanced oral transmission by nutrition-deprived mosquitoes. *J. Med. Entomol.* 21:249-256.
Grimstad PR, Walker ED (1991) *Aedes triseriatus* (Diptera: Culicidae) and La Crosse virus IV. Nutritional deprivation of larvae effects adult barriers to infection and transmission. *J. Med. Entomol.* 28:378-386.
Grossman GL, Campos Y, Severson DW, James AA (1997) Evidence for two distinct members of the amylase gene family in the yellow fever mosquito, *Aedes aegypti. Insect Biochem Mol Biol* 27: 769-781.
Gubler DJ (1987) Dengue and dengue hemorrhagic fever in the Americas. *P R Health Sci J.* Aug;6(2):107-11.
Gubler DJ (1989) *Aedes aegypti* and *Aedes aegypti*-borne disease control in the 1990s: top down or bottom up. *Am. J. Trop. Med. Hyg.* 40:571-578
Gubler DJ (1991) Insects in Disease Transmission. In: Strickland GT, editor. *Hunter tropical medicine*, 7th edition. Philadelphia (PA): W. B. Saunders; 1991. p. 981-1000.
Gubler DJ (1997) Dengue and dengue hemorrhagic fever: its history and resurgence as a global health problem, p. 1-22 *In* D.J. Gubler and G. Kuno (ed.), *Dengue and dengue hemorrhagic fever.* CAB International, Londan, United Kingdom.
Gubler DJ (1998) Dengue and Dengue Hemorrhagic Fever. *Clinical Microbiology* Jul:11(3): 480-496.
Gubler DJ (2002) The global emergence/resurgence of arboviral diseases as public health problems. *Arch Med Res.* Jul-Aug;33(4):330-42.
Gubler DJ (2004) The changing epidemiology of yellow fever and dengue, 1900 to 2003: full circle? *Comp Immunol Microbiol Infect Dis.* Sep;27(5):319-330.
Gubler DJ, Clark GG (1994) Community-based integrated control of *Aedes aegypti*: a brief overview of current programs. *Am J Trop Med Hyg.* 50(6 Suppl):50-60.
Gubler DJ, Clark GG (1995) Dengue/dengue hemorrhagic fever: the emergence of a global health problem. *Emerg Infect Dis.* Apr-Jun:1(2):55-7.
Gubler DJ, Nalim S, Tan R, Saipan H, Sulianti Saroso J (1979) Variation in susceptibility to oral infection with dengue viruses among geographic strains of *Aedes aegypti. Am J Trop Med Hyg.* Nov;28(6):1045-52.
Gubler DJ, Reed D, Rosen L, Hitchcock JR Jr. (1978) Epidemiologic, clinical, and virologic observations on dengue in the Kingdom of Tonga. *Am J Trop Med Hyg.* May;27(3):581-9.

Gubler D.J., Suharyono W., Sumarmo H., Wulur H., Jahja E., Sulianti Soroso J. (1981) Viremia in patients with naturally acquired dengue infection. *Bull. W.H.O.* 59:623-630.

Gubler D.J., Trent D.W. (1994) Emergence of epidemic dengue/dengue hemorrhagic fever as a public health problem in the Americas. *Infect. Agents Dis.* 2:383-393.

Guzmán MG, Kourí G (2002) Dengue: an update. *Lancet Infect Dis.* Jan;2(1):33-42.

Guzman MG, Kouri G (2003) Dengue and dengue hemorrhagic fever in the Americas: lessons and challenges. *Journal of Clinical Virology* 27:1-13.

Hale, L. R. and Singh, R. S. (1986) Extensive variation and heteroplasmy in size of mitochondrial-DNA among geographic populations of *Drosophila melanogaster*. Proc Natl Acad Sci USA 83: 8813-8817.

Halstead SB (1970) Observations related to pathogenesis of dengue hemorrhagic fever. IV. Hypotheses and discussion. *Yale J. Biol. Med.* 42:350-362

Halstead SB (1983) Dengue haemorrhagic fever--a public health problem and a field for research. *Bull World Health Organ.* 58(1):1-21.

Halstead SB (1983) Dengue haemorrhagic fever. *Trans. R. Soc. Trop. Med. Hyg.* 77(5):739-740.

Halstead SB (1984) Selective primary health care: strategies for control of diseases in the developing countries. *Rev. Inf. Dis.* 6:257-264

Halstead SB (1988) Pathogenesis of dengue: challenges to molecular biology. *Science* 239:476-481.

Hammond SM, Boettcher S, Caudy AA, Kobayashi R, Hannon GJ (2001) Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science.* Aug 10;293(5532):1146-50.

Hanna JN, Ritchie SA, Merritt AD, van den Hurk AF, Phillips DA, Serafin IL, Norton RE, McBride WJ, Gleeson FV, Poidinger M (1998) Two contiguous outbreaks of dengue type 2 in north Queensland. *Med J Aust.* Mar 2;168(5):221-5.

Hannon GJ (2002) RNA interference. *Nature.* Jul 11;418(6894):244-51.

Hardy JL (1988) Susceptibility and resistance of vector mosquitoes. Monath T.P., ed. *Arboviruses. Epidemiology and Ecology*, Volume I. Boca Raton, FL: CRC press., pp 87-126.

Harrington LC, Edman JD, Scott TW (2001) Why do female *Aedes aegypti* (Diptera: Culicidae) feed preferentially and frequently on human blood? *J Med Entomol.* May;38(3):411-22.

Harrington LC, Buonaccorsi JP, Edman JD, Costero A, Kittayapong P, Clark GG, Scott TW (2001) Analysis of survival of young and old *Aedes aegypti* (Diptera: Culicidae) from Puerto Rico and Thailand. *J Med Entomol.* Jul;38(4):537-47.

Harrus S, Baneth G (2005) Drivers for the emergence and re-emergence of vector-borne protozoal and bacterial diseases. *Int J Parasitol.* Oct;35(11-12):1309-18.

Hawkins CJ, Ekert PG, Uren AG, Holmgreen SP, Vaux DL (1998) Anti-apoptotic potential of insect cellular and viral IAPs in mammalian cells. *Cell Death Differ.* Jul;5(7):569-76.

Hayes EB, Gubler DJ (1992) Dengue and dengue hemorrhagic fevers of children in the Philippines and Thailand. *Trans. Assoc. Am. Physicians* 73:140-155.

Hay BA, Wassarman DA, Rubin GM (1995) Drosophila homologs of baculovirus inhibitor of apoptosis proteins function to block cell death. *Cell.* Dec 29;83(7):1253-62.

Hazkani-Covo, E. and Graur, D. (2007) A comparative analysis of numt evolution in human and chimpanzee. *Mol. Biol. and Evol.* 24: 13-18.

Henschal EA, Putnak JR (1990) The dengue viruses. *Clin Microbiol. Rev.* 3:376-396.

Henze, K. and Martin, W. (2001) How do mitochondrial genes get into the nucleus? *Trends in Genetics* 17: 383-387.

Hill WG, Robertson A (1966) The effect of linkage on limits to artificial selection. *Genet Res.* Dec;8(3):269-94.

Holmes EC, Worobey M, Rambaut A (1999) Phylogenetic evidence for recombination in dengue virus. *Mol Biol Evol.* Mar;16(3):405-9.

Holmes EC, Burch SS (2000) The causes and consequences of genetic variation in dengue virus. *Trends in Microbiology.* 8(2):74-77.

Holmes EC, Twiddy SS (2003) The origin, emergence and evolutionary genetics of dengue virus. *Infect Genet Evol.* May;3(1):19-28.

Holt RA, Subramanian GM, Halpern A, Sutton GG, Charlab R, Nusskern DR, Wincker P, Clark AG, Ribeiro JM, Wides R, Salzberg SL, Loftus B, Yandell M, Majoros WH, Rusch DB, Lai Z, Kraft CL, Abril JF, Anthouard V, Arensburger P, Atkinson PW, 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 GK, Chrystal MA, Clamp M, Cravchik A, Curwen V, Dana A, Delcher A, Dew I, Evans CA, Flanigan M, Grundschober-Freimoser A, Friedli L, Gu Z, Guan P, Guigo R, Hillenmeyer ME, Hladun SL, Hogan JR, Hong YS, Hoover J, Jaillon O, Ke Z, Kodira C, Kokoza E, Koutsos A, Letunic I, Levitsky A, Liang Y, Lin JJ, Lobo NF, Lopez JR, Malek JA, McIntosh TC, Meister S, Miller J, Mobarry C, Mongin E, Murphy SD, O'Brochta DA, Pfannkoch C, Qi R, Regier MA, Remington K, Shao H, Sharakhova MV, Sitter CD, Shetty J, Smith TJ, Strong R, Sun J, Thomasova D, Ton LQ, Topalis P, Tu Z, Unger MF, Walenz B, Wang A, Wang J, Wang M, Wang X, Woodford KJ, Wortman JR, Wu M, Yao A, Zdobnov EM, Zhang H, Zhao Q, Zhao S, Zhu SC, Zhimulev I, Coluzzi M, della Torre A, Roth CW, Louis C, Kalush F, Mural RJ, Myers EW, Adams MD, Smith HO, Broder S, Gardner MJ, Fraser CM, Birney E, Bork P, Brey PT, Venter JC, Weissenbach J, Kafatos FC, Collins FH, Hoffman SL (2002) The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science*. Oct 4;298(5591):129-49.

Houston P, Simon PJ, Broach JR (2004) The *Saccharomyces cerevisiae* recombination enhancer biases recombination during interchromosomal mating-type switching but not in interchromosomal homologous recombination. *Genetics* 166: 1187-1197.

Howe G.M. (1977) *A world geography of human diseases.* Academic Press, Inc., New York, N.Y.

Hubálek Z, Halouzka J (1999) West Nile fever--a reemerging mosquito-borne viral disease in Europe. *Emerg Infect Dis.* Sep-Oct;5(5):643-50.

Huber K, Loan LL, Hoang TH, Tien TK, Rodhain F, Failloux AB (2002) Temporal genetic variation in Aedes aegypti populations in Ho Chi Minh City (Vietnam). *Heredity*. Jul;89(1):7-14.

Huber K, Mousson L, Rodhain F, Failloux AB (1999) Short report: microsatellite sequences as markers for population genetic studies of the mosquito *Aedes aegypti*, the vector of dengue viruses. *Am J Trop Med Hyg*. Dec;61(6):1001-3.

Hughes AL, Nei M (1990) Evolutionary relationships of class II major-histocompatibility-complex genes in mammals. *Mol Biol Evol*. Nov;7(6):491-514.

Hurd H, Carter V (2004) The role of programmed cell death in *Plasmodium*-mosquito interactions. *International Journ for Parasitology* 34: 1459-1472.

Hurlbut HS (1973) The effect of environmental temperature upon the transmission of St. Louis encephalitis virus by *Culex pipiens quinquefasciatus*. *J. Med. Entomol.* 10:1-12.

Jehle JA, Blissard GW, Bonning BC, Cory JS, Herniou EA, Rohrmann GF, Theilmann DA, Thiem SM, Vlak JM (2006) On the classification and nomenclature of baculoviruses: a proposal for revision. *Arch Virol*. Jul;151(7):1257-66.

Jenkins GM, Rambaut A, Pybus OG, Holmes EC (2002) Rates of molecular evolution in RNA viruses: a quantitative phylogenetic analysis. *J Mol Evol*. Feb;54(2):156-65.

Jiang F, Ye X, Liu X, Fincher L, McKearin D, Liu Q (2005) Dicer-1 and R3D1-L catalyze microRNA maturation in *Drosophila*. *Genes Dev*. Jul 15;19(14):1674-9.

Johnson KL, Price BD, Eckerle LD, Ball LA (2004) Nodamura virus nonstructural protein B2 can enhance viral RNA accumulation in both mammalian and insect cells. *J Virol*. Jun;78(12):6698-704.

Joussett FX, Barreau C, Boublik Y, Cornet M (1993) A parvo-like virus persistently infecting a C6/36 clone of *Aedes albopictus* mosquito cell line and pathogenic for Aedes aegypti larvae. *Virus. Res*. 29:99-114.

Jousset FX, Baquerizo E, Bergoin M (2000) A new densovirus isolated from the mosquito *Culex pipeans* (Diptera: Culicidae). *Virus Res*. 67:11-16.

Karpf AR, Brown DT (1998) Comparison of Sindbis virus-induced pathology in mosquito and vertebrate cell cultures. Virology. Jan 20;240(2):193-201.

Keene KM, Foy BD, Sanchez-Vargas I, Beaty BJ, Blair CD, Olson KE (2004) RNA interference acts as a natural antiviral response to O'nyong-nyong virus (Alphavirus; Togaviridae) infection of *Anopheles gambiae*. *Proc Natl Acad Sci U S A*. Dec 7;101(49):17240-5.

Kitzmiller JB (1976) Genetics, cytogenetics, and evolution of mosquitoes. *Advances in Genetics* 315-433.

Knox TB, Kay BH, Hall RA, Ryan PA (2003) Enhanced vector competence of *Aedes aegypti* (Diptera: Culicidae) from the Torres Strait compared with mainland Australia for dengue 2 and 4 viruses. J *Med Entomol*. Nov;40(6):950-6.

Kolokotronis SO., MacPhee, R. D. E. and Greenwood, A. D. (2007) Detection of mitochondrial insertions in the nucleus (NuMts) of Pleistocene and modern muskoxen. *Bmc Evolutionary Biology* 7:-.

Kramer LD, Hardy JL, Presser SB (1983) Effect of temperature of extrinsic incubation on the vector competence of *Culex tarsalis* for western equine encephalomyelitis virus. *Am. J. Trop. Med. Hyg.* 32:1130-1139.

Kurane I (2007) Dengue hemorrhagic fever with special emphasis on immunopathogenesis. *Comp Immunol Microbiol Infect Dis.* Sep;30(5-6):329-40.

Kurane I, Takasaki T (2001) Dengue fever and dengue haemorrhagic fever: challenges of controlling an enemy still at large. *Rev Med Virol.* Sep-Oct;11(5):301-11.

Kuzoe FAS, Gill JJB, Macdonald WW (1966) Evidence of inversions in chromosomes in testes of *Aedes aegypti* and *Culex pipiens fatigans. Trans Royal Soc Trop Med Hyg* 60: 21.

Labonne, JDJ, Hilliker AJ, Shore JS. (2007) Meiotic recombination in Turnera (Turneraceae): extreme sexual difference in rates, but no evidence for recombination suppression associated with the distyly (S) locus. *Heredity* 98: 411-418.

Lam KS, Marshall ID (1968)Dual infections of *Aedes aegypti* with arboviruses. I. Arboviruses that have no apparent cytopathic effect in the mosquito. *Am J Trop Med Hyg.* Jul;17(4):625-36.\

Lanciotti RS, Gubler DJ, Trent DW (1997) Molecular evolution and phylogeny of dengue-4 viruses. *J Gen Virol.* Sep;78 (Pt 9):2279-84.

Lanciotti RS, Roehrig JT, Deubel V, Smith J, Parker M, Steele K, Crise B, Volpe KE, Crabtree MB, Scherret JH, Hall RA, MacKenzie JS, Cropp CB, Panigrahy B, Ostlund E, Schmitt B, Malkinson M, Banet C, Weissman J, Komar N, Savage HM, Stone W, McNamara T, Gubler DJ (1999) Origin of the West Nile virus responsible for an outbreak of encephalitis in the northeastern United States. *Science* Dec 17;286(5448):2333-7.

Landegren U, Nilsson M, Kwok PY (1998) Reading bits of genetic information: methods for single-nucleotide polymorphism analysis. *Genome Res.* Aug;8(8):769-76.

Lander ES, Green P (1987) Construction of Multilocus Genetic-Linkage Maps in Humans. *Proc Natl Acad Sci* (USA) 84: 2363-2367.

Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121:185-199.

Lawson D, Arensburger P, Atkinson P, Besansky NJ, Bruggner RV, Butler R, Campbell KS, Christophides GK, Christley S, Dialynas E, Hammond M, Hill CA, Konopinski N, Lobo NF, MacCallum RM, Madey G, Megy K, Meyer J, Redmond S, Severson DW, Stinson EO, Topalis P, Birney E, Gelbart WM, Kafatos FC, Louis C, Collins FH (2009) VectorBase: a data resource for invertebrate vector genomics. *Nucleic Acids Res.* Jan;37(Database issue):D583-7.

Lascaro, D., Castellana, S., Gasparre, G., Romeo, G., Saccone, C. and Attimonelli, M. (2008) The RHNumtS compilation: Features and bioinformatics approaches to locate and quantify Human NumtS. *Bmc Genomics 9:* -.

Lebedeva OP, Topchii MK, Buchatskii LP, Lebedinets NN, Gonchar NW (1975) A rapid method for the diagnosis of densonucleosis in *Aedes aegypti* larvae and its use in search for insects susceptible to this virus. *Med. Parazitol* (Mosk) 44:612-615.

Lebedinets NN, Zelenko AP (1975) Infectivity of *Aedes aegypti* L. mosquito densonucleosis virus for larvae of other species of blood-sucking mosquitoes. *Vopr. Virusol.* 2:192-196.

Lecellier CH, Dunoyer P, Arar K, Lehmann-Che J, Eyquem S, Himber C, Saib A, Voinnet O (2005) A cellular MicroRNA mediates antiviral defense in human cells. *Science* 308:557-560.

Ledermann JP, Suchman EL, Black WC, Carlson JO (2004) Infection and pathogenicity of the mosquito densoviruses AeDNV, HeDNV, and APeDNV in *Aedes aegypti* mosquitoes (Diptera: Culicidae). *J. Econ. Entomol.* 97:1828-1836.Lee YS, Nakahara K, Pham JW, Kim K, He Z, Sontheimer EJ, Carthew RW (2004) Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. *Cell.* Apr 2;117(1):69-81.

Lee JH, Rowley WA, Platt KB (2000) Longevity and spontaneous flight activity of *Culex tarsalis* (Diptera: Culicidae) infected with western equine encephalomyelitis virus. *J Med Entomol.* Jan;37(1):187-93.

Leulier F, Lhocine N, Lemaitre B, Meier P The *Drosophila* inhibitor of apoptosis protein DIAP2 functions in innate immunity and is essential to resist gram-negative bacterial infection. *Mol Cell Biol.* Nov;26(21):7821-31.

Levin DH, Acs G, Silverstein SC (1970) Chain initiation by reovirus RNA transcriptase in vitro. *Nature.* Aug 8;227(5258):603-4.

Li H, Li WX, Ding SW (2002) Induction and suppression of RNA silencing by an animal virus. *Science* May 17;296(5571):1319-21.

Lindenbach BD, Rice CM (2001) Flaviviridae: the viruses and their replication. *Fields Fundamental Virology.* 4th. Lippincot Williams and Wilkins. Philadelphia, PA.,USA.

Lindbo JA, Silva-Rosales L, Proebsting WM, Dougherty WG (1993) Induction of a Highly Specific Antiviral State in Transgenic Plants: Implications for Regulation of Gene Expression and Virus Resistance. *Plant Cell.* Dec;5(12):1749-1759.

Linley JR, Nielsen HT (1968) Transmission of a mosquito iridescent virus in *Aedes taeniorhynchus I.* Laboratory experiments. *J. Inverteb. Pathol.* 12:7-16.

Liu Q, Rand TA, Kalidas S, Du F, Kim HE, Smith DP, Wang X (2003) R2D2, a bridge between the initiation and effector steps of the *Drosophila* RNAi pathway. Science. Sep 26;301(5641):1921-5.

Liu X, Jiang F, Kalidas S, Smith D, Liu Q (2006) Dicer-2 and R2D2 coordinately bind siRNA to promote assembly of the siRISC complexes. *RNA.* Aug;12(8):1514-20.

Lopez, J. V., Yuhki, N., Masuda, R., Modi, W. and Obrien, S. J. (1994) Numt, a recent transfer and tandem amplification of mitochondrial-DNA to the nuclear genome of the domestic cat. *J. Mol. Evol.* 39: 174-190.

Lorenz L, Beaty BJ, Aitken THG, Wallis GP, Tabachnick WJ (1984) The effect of colonization upon *Aedes aegypti* - susceptibility to oral infection with yellow-fever virus. *Am J Trop Med Hyg* 33: 690-694.

Loroño-Pino MA, Cropp CB, Farfán JA, Vorndam AV, Rodríguez-Angulo EM, Rosado-Paredes EP, Flores-Flores LF, Beaty BJ, Gubler DJ (1999) Common occurrence

of concurrent infections by multiple dengue virus serotypes. *Am J Trop Med Hyg.* Nov;61(5):725-30.

Lough, A. N., Roark, L. M., Kato, A., Ream, T. S., Lamb, J. C., Birchler, J. A. and Newton, K. J. (2008) Mitochondrial DNA transfer to the nucleus generates extensive insertion site variation in maize. *Genetics* 178: 47-55.

Lozano-Fuentes S (2005) *Aedes aegypti* vector competence and gene flow in Mexico. Association mapping software for testing candidate genes associated with a phenotype. Ph.D Dissertation. Colorado State University, Fort Collins, CO.

Lozano-Fuentes S, Fernandez-Salas I, Munoz ML, Farfan-Ale JA, Olson KE, Beaty BJ, William WC 4th (2009) The neovolcanic axis is a barrier to gene flow among *Aedes aegypti* populations in Mexico that differ in vector competence for dengue 2 virus. *Public Library o Science, Neglected Tropical Diseases* (Submitted 01/2009).

Ludwig GV, Christensen BM, Yuill TM, Schultz KT (1989) Enzyme processing of LaCrosse viruse glycoprotein G1: a bunyavirus-vector infection model. *Virology.* 171:108.

Ludwig GV, Israel BA, Christensen BM, Yuill TM, Schultz KT (1991) Role of LaCrosse virus glycoproteins in attachment of virus to host cells. *Virology.* 181:564.

Ma H, Moore PH, Liu ZY, Kim MS, Yu QY, Fitch MMM, Sekioka T, Paterson AH, Ming R (2004) High-density linkage mapping revealed suppression of recombination at the sex determination locus in papaya. *Genetics* 166: 419-436.

Macdonald WW, Sheppard PM (1965) Cross-over values in sex chromosomes of mosquito *Aedes aegypti* and evidence of presence of inversions. *Ann Trop Med Parasit* 59: 74-87.

Macdonald WW (1962a) Selection of a strain of *Aedes aegypti* susceptible to infection with semi-periodic *Brugia malayi*. *Ann Trop Med Parasit* 56: 368-372

Macdonald WW (1962b) Genetic Basis of Susceptibility to Infection with Semi-Periodic Brugia malayi in Aedes aegypti. *Ann Trop Med Parasit* 56: 373-382

Macdonald G (1957) *The Epidemiology and Control of Malaria.* Oxford University Press, London.

Mackenzie JS, Gubler DJ, Petersen LR (2004) Emerging flaviviruses: the spread and resurgence of Japanese encephalitis, West Nile and dengue viruses. *Nat Med.* Dec;10(12 Suppl):S98-109.

Mahmood F, Reisen WK, Chiles RE, Fang Y (2004) Western equine encephalomyelitis virus infection affects the life table characteristics of *Culex tarsalis* (Diptera: Culicidae) *J. Med. Entomol.* Sep;41(5):982-986.

Malavige GN, Fernando S, Fernando DJ, Seneviratne SL (2004) Dengue viral infections. *Postgrad Med J.* Oct;80(948):588-601.

Maliepaard C, Jansen J, Van Ooijen JW (1997) Linkage analysis in a full-sib family of an outbreeding plant species: overview and consequences for applications. *Genet. Res* 70: 237-250.

Malla S, Thakur GD, Shrestha SK, Banjeree MK, Thapa LB, Gongal G, Ghimire P, Upadhyay BP, Gautam P, Khanal S, Nisaluk A, Jarman RG, Gibbons RV (2008)

Identification of all dengue serotypes in Nepal. *Emerg Infect Dis.* Oct;14(10):1669-70.
Marina CF, Arredondo-Jimenz JI, Castillo A, Williams T (1999) Sublethal effects of iridovirus disease in a mosquito. *Oecologia.* 119:383-388.
Marina CF, Ibarra JE, Arredondo-Jimenez JI, Fernandez-Salas I, Valle J, Williams T (2003) Sublethal iridovirus disease of the mosquito *Aedes aegytpi* is due to viral replication not cytotoxicity. Med. *Vet. Entomol.* 17:187-194.
Marquardt W, Black WC, Freier JE, Hagedorn H, Hemingway J, Higgs S, James AA, Kondratieff B, Moore CG (2005) *The Biology of Disease Vectors* 2nd ed, Elsevier Academic Press, London.
Martins, J., Solomon, S. E., Mikheyev, A. S., Mueller, U. G., Ortiz, A. and Bacci, M. (2007) Nuclear mitochondrial-like sequences in ants: evidence from *Atta cephalotes* (Formicidae : Attini). *Ins. Mol. Biol.*16: 777-784.
Mather K., Jinks J.L. (1982) *Biomettrical Genetics*, 3rd ed. Chapman and Hall, London.
Matranga C, Tomari Y, Shin C, Bartel DP, Zamore PD (2005) Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. *Cell.* Nov 18;123(4):607-20.
Matthews TC, Craig GB (1987) Heterozygosity in inbred strains of the tree-hole mosquito *Aedes triseriatus. Biochem Genet* 25:647-655.
Matthews, TC, Craig GB (1989) Isozyme polymorphisms maintained by lethal loci in inbred strains of *Aedes triseriatus. J Heredity* 80:53-57.
Mattingly PF (1957) Genetical aspects of the *Aedes aegypti* problem, I. Taxonomy and bionomics. Ann *Trop Med Parasit* 51:392-408.
Mattingly PF (1967) Taxonomy of *Aedes aegypti* and related species. *Bulletin of the World Health Organization* 36:552.
McAbee RD, Christiansen JA, Cornel AJ (2007) A detailed larval salivary gland polytene chromosome photomap for *Culex quinquefasciatus* (Diptera : Culicidae) from Johannesburg, South Africa. *J Med Entomol* 44:229-237.
McBride WJ, Bielefeldt-Ohmann H (2000) Dengue viral infections; pathogenesis and epidemiology. *Microbes Infect.* 2 (9):1041-50.
McClelland GAH. (1974) Worldwide survey of variation in scale pattern of abdominal tergum of *Aedes aegypti* (L) (Diptera-Culicidae). *Trans Royal Entomol Soc* London 126:239-259.
McClelland GAH (1960) A prelimnary study of the genetics of abdominal color variations in *Aedes aegypti* (L.) (Diptera: Culicidae). Ann *Trop Med Parasit* 54:305-320.
McGaw MM, Chandler LJ, Wasieloski LP, Blair CD, Beaty BJ (1998) Effect of La Crosse virus infection on overwintering of *Aedes triseriatus. Am J Trop Med Hyg.* Feb;58(2):168-75.
McSherry JA (1982) Some medical aspects of the Darien schema: was it dengue? *Scot. Medi. J.* 27:183-184.
Meier P, Finch A, Evan G. (2000) Apoptosis in development. *Nature.* Oct 12;407(6805):796-801.
Mercado-Curiel RF, Esquinca-Avilés HA, Tovar R, Díaz-Badillo A, Camacho-Nuez M, Muñoz Mde L (2006) The four serotypes of dengue recognize the same putative

receptors in *Aedes aegypti* midgut and Ae. albopictus cells. *BMC Microbiol.* Oct 2;6:85.

Messer WB, Gubler DJ, Harris E, Sivananthan K, de Silva AM (2003) Emergence and global spread of a dengue serotype 3, subtype III virus. *Emerg Infect Dis.* Jul;9(7):800-9.

Michel AP, Guelbeogo WM, Grushko O, Schemerhorn BJ, Kern M, Willard MB, Sagnon N, Costantini C, Besansky NJ (2005) Molecular differentiation between chromosomally defined incipient species of *Anopheles funestus*. *Ins Mol Biol* 14: 375-387.

Mielke PW, Berry KJ (2001) Permutation Methods: *A Distance Function Approach.* Springer Series in Statistics. Springer, New York.

Miller BR, Mitchell CJ (1991) Genetic selection of a flavivirus-refractory strain of yellow fever mosquito *Aedes aegypti*. *Am. J. Trop. Med. Hyg.* 45:399-407.

Miller DJ, Ahlquist P (2002) Flock house virus RNA polymerase is a transmembrane protein with amino-terminal sequences sufficient for mitochondrial localization and membrane insertion. *J Virol.* Oct;76(19):9856-67.

Mitchell CJ, Miller BR, Gubler DJ (1987) Vector competence of *Aedes albopictus* from Houston, Texas, for dengue serotypes 1 to 4, yellow fever and Ross River viruses. *J Am Mosq Control Assoc.* Sep;3(3):460-5.

Miyoshi K, Tsukumo H, Nagami T, Siomi H, Siomi MC (2005) Slicer function of *Drosophila* Argonautes and its involvement in RISC formation. *Genes Dev.* Dec 1;19(23):2837-48.

Moeur JE, Istock CA (1982) Chromosomal Polymorphisms in the Pitcher-Plant Mosquito, *Wyeomyia smithii*. *Chromosoma* 84:623-651.

Monath TP (1991) Yellow fever: Victor, Victoria? Conqueror, conquest? Epidemics and research in the last forty years and prospects for the future. *Am J Trop Med Hyg.* Jul;45(1):1-43.

Monath TP (1994) Dengue: the risk to developed and developing countries. *Proc. National Acad. Sci.* 91:2395-2400.

Monath TP (2006) Yellow fever as an endemic/epidemic disease and priorities for vaccination. *Bull Soc Pathol Exot.* Dec;99(5):341-7.

Monath TP, Heinz FX (1996) *Flaviviruses. Fields Virology* 3rd ed. Philadelphia: Lippincott-Raven Publishers., 961-1034.

Moore, D. F. (1979) Hybridization and mating-behavior in *Aedes aegypti* (Diptera, Culicidae). *J Med Entomol* 16: 223-226.

Mori A, Romero-Severson J, Black WC 4th, Severson DW (2008) Quantitative trait loci determining autogeny and body size in the Asian tiger mosquito (*Aedes albopictus*). *Heredity.* Jul;101(1):75-82.

Morrison AC, Costero A, Edman JD, Clark GG, Scott TW (1999) Increased fecundity of *Aedes aegypti* fed human blood before release in a mark-recapture study in Puerto Rico. *J Am Mosq Control Assoc.* Jun;15(2):98-104.

Moser BA, Becnel JJ, White SE, Afonso C, Kutish G, Shanker S, Almira E (2001) Morphological and molecular evidence that *Culex nigripalpus* baculovirus is an unusual member of the family Baculoviridae. *J. Gen. Virol.* 82:283-297.

Mourier, T., Hansen, A. J., Willerslev, E. and Arctander, P. (2001) The human genome project reveals a continuous, transfer of large mitochondrial fragments to the nucleus. *Mol. Biol. and Evol.* 18: 1833-1837.

Mousson L, Dauga C, Garrigues T, Schaffner F, Vazeille M, Failloux AB (2005) Phylogeography of *Aedes* (*Stegomyia*) *aegypti* (L.) and *Aedes* (*Stegomyia*) *albopictus* (Skuse) (Diptera: Culicidae) based on mitochondrial DNA variations. *Genet Res.* Aug;86(1):1-11.

Munstermann LE (1994) Unexpected genetic consequences of colonization and inbreeding - allozyme tracking in Culicidae (Diptera). *Ann Entomol Soc Am* 87: 157-164.

Munstermann LE and Craig GB (1979) Genetics of *Aedes aegypti* - Updating the Linkage Map. *J Heredity* 70: 291-296.

Munstermann LE, Marchi A (1986) Cytogenetic and isozyme profile of *Sabethes cyaneus* - a mosquito of the neotropical canopy. *J Heredity* 77: 241-248.

Munstermann LE, Marchi A, Sabatini A, Coluzzi M (1985) Polytene chromosomes of *Orthopodomyia pulcripalpis* (Diptera, Culicidae). *Parassitologia* 27: 267-277.

Murphy FA, Fauguet CM, Bishop DHL, Ghabrial SA, Jarvis AW, Martelli GP, Mayo MA, Summers MD (1995) *Virus Taxonomy Sixth Report of the International Committee on Taxonomy of Viruses*. NewYork: Springler-Verlag.

Myles KM, Pierro DJ, Olson KE. (2004) Comparison of the transmission potential of two genetically distinct Sindbis viruses after oral infection of *Aedes aegypti* (Diptera: Culicidae). *J Med Entomol.* Jan;41(1):95-106.

Myles KM, Wiley MR, Morazzani EM, Adelman ZN (2008) Alphavirus-derived small RNAs modulate pathogenesis in disease vector mosquitoes. *Proc Natl Acad Sci* U S A. Dec 16;105(50):19938-43.

Nash D, Mostashari F, Fine A, Miller J, O'Leary D, Murray K, Huang A, Rosenberg A, Greenberg A, Sherman M, Wong S, Layton M; 1999 West Nile Outbreak Response Working Group (2001) The outbreak of West Nile virus infection in the New York City area in 1999. *N Engl J Med.* Jun 14;344(24):1807-14.

Nasidi A, Monath TP, Decock K, Tomori O, Cordellier R, Olaleye OD, Harry TO, Adeniyi JA, Sorungbe AO, Ajosecoker AO, Vanderlaan G, Oyediran ABO (1989) Urban yellow-fever epidemic in Western Nigeria, 1987. *Trans Royal SocTrop Med Hyg* 83:401-406.

Navarro A, Betran E, Barbadilla A, Ruiz A (1997) Recombination and gene flux caused by gene conversion and crossing over in inversion heterokaryotypes. *Genetics* 146:695-709.

Nei, M (1987) *Molecular Evolutionary Genetics*. Columbia Unv. Press, New York.

Nene V, Wortman JR, Lawson D, Haas B, Kodira C, Tu ZJ, Loftus B, Xi Z, Megy K, Grabherr M, Ren Q, Zdobnov EM, Lobo NF, Campbell KS, Brown SE, Bonaldo MF, Zhu J, Sinkins SP, Hogenkamp DG, Amedeo P, Arensburger P, Atkinson PW, Bidwell S, Biedler J, Birney E, Bruggner RV, Costas J, Coy MR, Crabtree J, Crawford M, Debruyn B, Decaprio D, Eiglmeier K, Eisenstadt E, El-Dorry H, Gelbart WM, Gomes SL, Hammond M, Hannick LI, Hogan JR, Holmes MH, Jaffe D, Johnston JS, Kennedy RC, Koo H, Kravitz S, Kriventseva EV, Kulp D, Labutti K, Lee E, Li S, Lovin DD, Mao C, Mauceli E, Menck CF, Miller JR, Montgomery P, Mori A, Nascimento AL, Naveira HF, Nusbaum C, O'leary S,

Orvis J, Pertea M, Quesneville H, Reidenbach KR, Rogers YH, Roth CW, Schneider JR, Schatz M, Shumway M, Stanke M, Stinson EO, Tubio JM, Vanzee JP, Verjovski-Almeida S, Werner D, White O, Wyder S, Zeng Q, Zhao Q, Zhao Y, Hill CA, Raikhel AS, Soares MB, Knudson DL, Lee NH, Galagan J, Salzberg SL, Paulsen IT, Dimopoulos G, Collins FH, Birren B, Fraser-Liggett CM, Severson DW (2007) Genome sequence of *Aedes aegypti*, a major arbovirus vector. *Science.* Jun 22;316(5832):1718-23.

Nobuchi, H (1979) The symptoms of a dengue-like illness recorded in a Chinese medical encyclopedia. *Kampo Rinsho* 26:422-425.

Obbard DJ, Gordon KH, Buck AH, Jiggins FM (2009) The evolution of RNAi as a defence against viruses and transposable elements. *Philos Trans R Soc Lond B Biol Sci.* Jan 12;364(1513):99-115.

Obbard DJ, Jiggins FM, Halligan DL, Little TJ (2006) Natural selection drives extremely rapid evolution in antiviral RNAi genes. *Curr Biol.* Mar 21;16(6):580-5.

Okimoto R, Dodgson JB (1996) Improved PCR amplification of multiple specific alleles (PAMSA) using internally mismatched primers. *Biotechniques.* Jul;21(1):20-2, 24, 26.

Olson KE, Higgs S, Gaines PJ, Powers AM, Davis BS, Kamrud KI, Carlson JO, Blair CD, Beaty BJ (1996) Genetically engineered resistance to dengue-2 virus transmission in mosquitoes. *Science*. May 10;272(5263):884-6.

Paduan, K. D. and Ribolla, P. E. M. (2008) Mitochondrial DNA polymorphism and heteroplasmy in populations of *Aedes aegypti* in Brazil. *J. Med. Entomol.* 45: 59-67.

Pamilo P, Viljakainen L, Vihavainen A (2007) Exceptionally high density of NUMTs in the honeybee genome. *Mol Biol Evol.* Jun;24(6):1340-6.

Paroo Z, Liu Q, Wang X (2007) Biochemical mechanisms of the RNA-induced silencing complex. *Cell Res.* 2007 Mar;17(3):187-94.

Parquet MC, Kumatori A, Hasebe F, Morita K and Igarashi A (2001). West Nile virus-induced bax-dependent apoptosis. *FEBS Lett.* 500 (1-2): 17-24.

Paul RE, Patel AY, Mirza S, Fisher-Hoch SP, Luby SP (1998) Expansion of epidemic dengue viral infections to Pakistan. *Int J Infect Dis.* Apr-Jun;2(4):197-201.

Paupy C, Brengues C, Kamgang B, Hervé JP, Fontenille D, Simard F (2008) Gene flow between domestic and sylvan populations of *Aedes aegypti* (Diptera: Culicidae) in North Cameroon. *J Med Entomol.* 2008 May;45(3):391-400.

Pedersen IM, Cheng G, Wieland S, Volinia S, Croce CM, Chisari FV, David M (2007) Interferon modulation of cellular microRNAs as an antiviral mechanism. *Nature* 449:919-U913.

Pereira, S. L. and Baker, A. J. (2004) Low number of mitochondrial pseudogenes in the chicken (*Gallus gallus*) nuclear genome: implications for molecular inference of population history and phylogenetics. *Bmc Evolutionary Biology 4:* -.

Peter ME, Heufelder AE and Hengartner MO (1997). Advances in apoptosis research. *Proc Natl Acad Sci* USA 4:12736-12737.

Pili E, Marchi A (1999) Chromosome evolution in treehole breeding *Anopheles* (Diptera, Culicidae). *Italian J Zool* 66: 33-37.

Pillai RS, Bhattacharyya SN, Filipowicz W (2007) Repression of protein synthesis by miRNAs: how many mechanisms? *Trends Cell Biol.* 17:118-126.

Pitman EJG (1938) Significance tests which can be applied to samples from any populations, III. The analysis of variance test. *Biometricka* 29: 322-335.
Platt KB, Linthicum KJ, Myint KS, Innis BL, Lerdthusnee K, Vaughn DW (1997) Impact of dengue virus infection on feeding behavior of *Aedes aegypti*. *Am J Trop Med Hyg.* Aug;57(2):119-25.
Pollitzer R (1960) A review of recent literature on plague. *Bull World Health Organ.* 23:313-400.
Poopittayasataporn A, Baimai V (1995) Polytene chromosome relationships of 5 species of the *Anopheles dirus* complex in Thailand. *Genome*. 38: 426-434.
Posada D (2008) jModelTest: phylogenetic model averaging. *Mol Biol Evol.* Jul;25(7):1253-6.
Powell JR, Tabachnick WJ, Arnold J (1980) Genetics and the origin of a vector population: *Aedes aegypti*, a case study. *Science.* Jun 20;208(4450):1385-7.
Pridgeon JW, Zhao L, Becnel JJ, Clark GG, Linthicum KJ (2008) Developmental and environmental regulation of AaeIAP1 transcript in *Aedes aegypti*. *Med Entomol.* Nov;45(6):1071-9.
Pryor MJ, Azzola L, Wright PJ and Davidson AD (2004). Histidine 39 in the dengue virus type 2 M protein has an important role in virus assembly. *J Gen Virol.* 85 (Pt 12): 3627-36.
Putnam JL, Scott TW (1995) The effect of multiple host contacts on the infectivity of dengue-2 virus-infected *Aedes aegypti*. *J Parasitol.* Apr;81(2):170-4.
Qu, H. Y., Ma, F. and Li, Q. W. (2008) Comparative analysis of mitochondrial fragments transferred to the nucleus in vertebrate. *J. Genetic and Genomics* 35: 485-490.
Rai KS, Hartberg WK (1975) *Aedes*. Handbook of genetics. Vol. 3. *Invertebrates of genetic interest* 311-345.
Rajatileka S, Black WC 4th, Saavedra-Rodriguez K, Trongtokit Y, Apiwathnasorn C, McCall PJ, Ranson H (2008) Development and application of a simple colorimetric assay reveals widespread distribution of sodium channel mutations in Thai populations of *Aedes aegypti*. *Acta Trop.* Oct;108(1):54-7.
Ramirez CCL, Dessen EMB (2000a) Chromosomal evidence for sibling species of the malaria vector *Anopheles cruzii*. *Genome* 43:143-151.
Ramirez CCL, Dessen EMB. (2000b) Chromosome differentiated populations of Anopheles cruzii: evidence for a third sibling species. *Genetica* 108:73-80.
Reese SM, Blitvich BJ, Blair CD, Geske D, Beaty BJ, Black WC 4th (2008) Potential for La Crosse virus segment reassortment in nature. Virol J. Dec 30;5(1):164.
Ricchetti, M., Fairhead, C. and Dujon, B. (1999) Mitochondrial DNA repairs double-strand breaks in yeast chromosomes. *Nature* 402: 96-100.
Richards SL, Mores CN, Lord CC, Tabachnick WJ (2007) Impact of extrinsic incubation temperature and virus exposure on vector competence of *Culex pipiens quinquefasciatus* Say (Diptera: Culicidae) for West Nile virus. *Vector Borne Zoonotic Dis.* Winter;7(4):629-36.
Richly, E. and Leister, D. (2004) NUMTs in sequenced eukaryotic genomes. *Mol. Biol. and Evol.* 21: 1081-1084.
Rico-Hesse R (2007) Dengue virus evolution and virulence models. *Clin Infect Dis.* Jun 1;44(11):1462-6.

Rico-Hesse R (2003) Microevolution and virulence of dengue viruses. *Adv Virus Res.* 59:315-41.

Rico-Hesse R. (1990) Molecular evolution and distribution of dengue viruses type 1 and 2 in nature. *Virology.* Feb;174(2):479-93.

Rico-Hesse R, Harrison LM, Nisalak A, Vaughn DW, Kalayanarooj S, Green S, Rothman AL, Ennis FA (1998) Molecular evolution of dengue type 2 virus in Thailand. *Am J Trop Med Hyg.* Jan;58(1):96-101.

Rico-Hesse R, Harrison LM, Salas RA, Tovar D, Nisalak A, Ramos C, Boshell J, de Mesa MT, Nogueira RM, da Rosa AT (1997) Origins of dengue type 2 viruses associated with increased pathogenicity in the Americas. *Virology.* Apr 14;230(2):244-51.

Roberts L (2007) Infectious disease. Resurgence of yellow fever in Africa prompts a counterattack. *Science.* May 25;316(5828):1109.

Rodríguez MM, Bisset JA, Fernández D (2007) Levels of insecticide resistance and resistance mechanisms in *Aedes aegypti* from some Latin American countries. *J Am Mosq Control Assoc.* Dec;23(4):420-9.

Rothman AL, Ennis FA (1999) Immunopathogenesis of dengue hemorragic fever. *Virology.* 257:1-6.

Rosen L (1977) The Emperor's New Clothes revisited, or reflections on the pathogenesis of dengue hemorrhagic fever. *Am J Trop Med Hyg.* May;26(3):337-43.

Rosen L, Roseboom LE, Gubler DJ, Lein JC, Chaniotis BN (1985) Comparative susceptibility of mosquito species and strains to oral and parenteral infection with dengue and Japanese encephalitis viruses. *Am J Trop Med Hyg* 34(3):603-615.

Ross Ronald (1923) Memoirs. John Murray, London, 1923.

Roukens AH, Visser LG (2008) Yellow fever vaccine: past, present and future. Expert *Opin Biol Ther.* Nov;8(11):1787-95.

Rush B. (1809) *Medical Inquiries and Observations.* Published by Johnson and Warner Vol II. (Third ed.) 383-431.

Rwegoshora RT, Baisley KJ, Kittayapong P (2000) Seasonal and spatial variation in natural densovirus infection in *Anopheles minimus* S.L. in Thailand. Southeast Asian J. Trop. *Med. Publ. Health* 31:3-9.

Rwegoshora RT, Kittayapong P (2004) Pathogenicity and infectivity of the Thai-strain densovirus (AThDNV) in *Anopheles minimus* S.L. *Southeast Asian J. Trop. Med. Publ. Health* 35:630-634.

Saad M, Youssef S, Kirschke D, Shubair M, Haddadin D, Myers J, Moorman J (2005) Acute flaccid paralysis: the spectrum of a newly recognized complication of West Nile virus infection. *J Infect.* Aug;51(2):120-7.

Saavedra-Rodriguez K, Strode C, Suarez AF, Fernandez-Salas I, Ranson H, Hemingway J and Black WC (2008) QTL mapping of genome regions controlling permethrin resistance in the mosquito *Aedes aegypti*. *Genetics* 180:1137–1152.

Sabin AB (1952) Research on Dengue During World War II. *Am. J. Trop. Med. Hyg.* 1:30-50.

Saikumar P, Dong Z, Mikhailov V, Denton M, Weinberg JM and Venkatachalam MA (1999) Apoptosis: definition, mechanisms, and relevance to disease. *Am J Med.* 107: 489-506.

Saito K, Ishizuka A, Siomi H, Siomi MC (2005) Processing of pre-microRNAs by the Dicer-1-Loquacious complex in *Drosophila* cells. *PLos Biol.* Jul;3(7):e235.

Sakuma S, Watanabe Y (1971) Unilateral synthesis of reovirus double-stranded ribonucleic acid by a cell-free replicase system. *J Virol.* Aug;8(2):190-6.

Sakuma S, Watanabe Y (1972) Reovirus replicase-directed synthesis of double-stranded ribonucleic acid. *J Virol.* Oct;10(4):628-38.

Salazar Sanchez MI (2005) Determinants of dengue type 2 virus infection in the mosquito *Aedes aegypti*. Ph.D. Dissertation Colorado State University.

Saleh MC, Tassetto M, van Rij RP, Goic B, Gausson V, Berry B, Jacquier C, Antoniewski C, Andino R. (2009) Antiviral immunity in *Drosophila* requires systemic RNA interference spread. *Nature*. Feb 8; EPub Ahead of Print.

Salonen A, Ahola T, Kääriäinen L (2005) Viral RNA replication in association with cellular membranes. *Curr Top Microbiol Immunol.* 285:139-73.

Sanchez-Vargas I, Travanty EA, Keene KM, Franz AW, Beaty BJ, Blair CD, Olson KE (2004) RNA interference, arthropod-borne viruses, and mosquitoes. *Virus Res.* Jun 1;102(1):65-74.

Sanchez-Vargas I., Scott J.C., Poole-Smith B.K., Franz A.W.E., Barbosa-Solomieu V., Wilusz J., Olson K.E., Blair C.D. (2009) Dengue virus 2 infections of *Aedes aegypti* are modulated by the mosquito's RNA interference pathway. *PLoS Pathogen* (In Press).

Sanderson MJ, Kim J (2000) Parametric phylogenetics? *Syst Biol.* Dec;49(4):817-29.

Schaeffer SW, Anderson WW (2005) Mechanisms of genetic exchange within the chromosomal inversions of *Drosophila pseudoobscura. Genetics* 171:1729-1739.

Scherer WF, Hurlbut HS (1967) Nodamura virus from Japan: a new and unusual arbovirus resistant to diethyl ether and chloroform. *Am J Epidemiol.* Sep;86(2):271-85.

Scherer WF, Verna JE, Richter W (1968) Nodamura virus, an ether- and chloroform-resistant arbovirus from Japan: physical and biological properties, with ecologic observations. *Am J Trop Med Hyg.* Jan;17(1):120-8.

Schlenke TA, Begun DJ (2003) Natural selection drives *Drosophila* immune system evolution. *Genetics.* Aug;164(4):1471-80.

Schneider JR, Mori A, Romero-Severson J, Chadee DD, Severson DW Investigations of dengue-2 susceptibility and body size among *Aedes aegypti* populations. *Med Vet Entomol.* Dec;21(4):370-6.

Schoepp R.J., Beaty B.J., Eckels K.H. (1990) Dengue 3 virus infection of *Aedes albopictus* and *Aedes aegypti*: comparison of parent and progeny candidate vaccine viruses. *Am. J. Trop. Med. Hyg.* 42(1):89-96.

Schwarz DS, Hutvágner G, Haley B, Zamore PD Evidence that siRNAs function as guides, not primers, in the *Drosophila* and human RNAi pathways. *Mol Cell.* Sep;10(3):537-48.

Scott TW, Amerasinghe PH, Morrison AC, Lorenz LH, Clark GG, Strickman D, Kittayapong P, Edman JD (2000) Longitudinal studies of *Aedes aegypti* (Diptera: Culicidae) in Thailand and Puerto Rico: blood feeding frequency. *J Med Entomol.* Jan;37(1):89-101.

Scott TW, Chow E, Strickman D, Kittayapong P, Wirtz RA, Lorenz LH, Edman JD (1993) Blood-feeding patterns of *Aedes aegypti* (Diptera: Culicidae) collected in a rural Thai village. *J Med Entomol.* Sep;30(5):922-7.

Scott TW, Lorenz LH (1998) Reduction of *Culiseta melanura* fitness by Eastern equine encephalomyelitis virus. *Am J Trop Med Hyg.* Aug;59(2):341-6.

Scott TW, Naksathit A, Day JF, Kittayapong P, Edman JD (1997) A fitness advantage for *Aedes aegypti* and the viruses it transmits when females feed only on human blood. *Am J Trop Med Hyg.* Aug;57(2):235-9.

Sejvar JJ, Bode AV, Marfin AA, Campbell GL, Pape J, Biggerstaff BJ, Petersen LR. (2006) West Nile Virus-associated flaccid paralysis outcome. *Emerg Infect Dis.* Mar;12(3):514-6

Severson DW, deBruyn B, Lovin DD, Brown SE, Knudson DL, Morlais I (2004) Comparative genome analysis of the yellow fever mosquito *Aedes aegypti* with *Drosophila melanogaster* and the malaria vector mosquito *Anopheles gambiae*. J *Heredity* 95:103-113.

Severson, DW, Kassner VA (1995) Analysis of mosquito genome structure using graphical genotyping. *Insect Mol Biol* 4:279-286.

Severson, DW, Meece JK, Lovin DD, Saha G, Morlais I (2002) Linkage map organization of expressed sequence tags and sequence tagged sites in the mosquito, *Aedes aegypti*. *Insect Mol Biol* 11:371-378.

Severson, DW, Mori A, Zhang Y, Christensen BM (1993) Linkage map for *Aedes aegypti* using restriction-fragment-length-polymorphisms. *J Heredity* 84:241-247.

Severson DW, Mori A, Zhang Y, Christensen BM (1994) Chromosomal mapping of two loci affecting filarial worm susceptibility in *Aedes aegypti*. *Insect Mol. Biol.* 3:67.

Severson, DW, Thathy V, Mori A, Zhang Y, Christensen BM (1995) Restriction-fragment-length-polymorphism mapping of quantitative trait loci for malaria parasite susceptibility in the mosquito *Aedes aegypti*. *Genetics* 139:1711-1717.

Severson DW, Zhang Y (1996) Generation of expressed sequence tags and sequence-tagged sites as physical landmarks in the mosquito, *Aedes aegypti* genome. *Genome* 39: 224-229.

Shapiro AM, Becnel JJ, White SE (2004) A nucleopolyhedrovirus from *Uranotaenia sapphirina* (Diptera: Culicidae). *J Invertebr Pathol.* Jul;86(3):96-103.

Shapiro A, Green T, Rao S, White S, Carner G, Mertens PC, Becnel J (2005) Morphological and molecular characterization of a cypovirus (Reoviridae) from the mosquito *Uranotaenia sapphirina* (Diptera: Culicidae). *J. Virology* 79:9430-9438.

Sharakhova MV, Hammond MP, Lobo NF, Krzywinski J, Unger MF, Hillenmeyer ME, Bruggner RV, Birney E, Collins FH (2007) Update of the *Anopheles gambiae* PEST genome assembly. *Genome Biol.* 8(1):R5.

Sharma GP, Mittal OP, Chaudhry S, Pal V (1978) Preliminary map of the salivary-gland chromosomes of *Aedes* (*Stegomyia*) *aegypti* (Culicidae, Diptera). *Cytobios* 22: 169-178.

Siler JF, Hall MW, Hitchens A (1926) Dengue, its History, Epidemiology, Mechanism of Transmission, Etiology, Clinical Manifestation, Immunity and Prevention. *Philipp J. Sci.* 29:1-304.

Simonsen KL, Churchill GA, Aquadro CF (1995) Properties of statistical tests of neutrality for DNA polymorphism data. *Genetics*. Sep;141(1):413-29.

Snowden FM (2008) Emerging and reemerging disease: a historical perspective. *Immunol Rev.* Oct;225:9-26.

Sriurairatna S, Bhamarapravati N (1977) Replication of dengue-2 virus in *Aedes albopictus* mosquitoes. An electron microscopic study. *Am. J. Trop. Med. Hyg.* 26:119.

Stahl EA, Bishop JG (2000) Plant-pathogen arms races at the molecular level. *Curr Opin Plant Biol.* Aug;3(4):299-304.

Stam P, Van Ooijen JW (1995) JoinMap (tm) Version 2.0: *Software for the Calculation of Genetic Linkage Maps.*

Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML Web servers. *Syst Biol.* Oct;57(5):758-71.

Stamatakis A, Ott M (2009) Efficient computation of the phylogenetic likelihood function on multi-gene alignments and multi-core architectures. *Philos Trans R Soc Lond B Biol Sci.* Oct 7, Ahead of print.

Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*. Nov 1;22(21):2688-90.

Stump AD, Pombi M, Goeddel L, Ribeiro JM C, Wilder JA, Torre AD, Besansky NJ (2007) Genetic exchange in 2La inversion heterokaryotypes of *Anopheles gambiae*. *Insect Mol Biol* 16:703-709.

Subbarao SK, Kumar KV, Nanda N, Nagpal BN, Dev V, Sharma VP (2000) Cytotaxonomic evidence for the presence of *Anopheles nivipes* in India. *J Am Mosq Cont Assoc* 16:71-74.

Subbarao SK, Nanda N, Vasantha K, Dua VK, Malhotra MS, Yadav RS, Sharma VP (1994) Cytogenetic evidence for 3 sibling species in *Anopheles fluviatilis* (Diptera, Culicidae). *Ann Entomol Soc Am* 87:116-121.

Suchman E, Carlson J (2004) Production of mosquito densonucleosis viruses by *Aedes albopictus* C6/36 cells adapted to suspension culture in serum-free protein-free media. *In Vitro Cell Dev. Biol – Animal* 40:74-75.

Suguna SG, Rathinam KG, Rajavel AR, Dhanda V (1994) Morphological and Chromosomal Descriptions of New Species in the Anopheles subpictus Complex. *Med Vet Entomol* 8:88-94.

Suguna, SG, Seawright JA, Joslyn DJ, Rabbani MG (1981) Homozygous pericentric inversions and other studies of inversions on chromosome-3 in the mosquito *Anopheles albimanus*. *Canadian J Genet Cytol* 23:57-64.

Sumanochitrapon W, Strickman D, Sithiprasasna R, Kittayapong P, Innis BL (1998) Effect of size and geographic origin of *Aedes aegypti* on oral infection with dengue-2 virus. *Am J Trop Med Hyg.* Mar;58(3):283-6.

Sunnucks, P. and Hales, D. F. (1996) Numerous transposed sequences of mitochondrial cytochrome oxidase I-II in aphids of the genus *Sitobion* (Hemiptera: Aphididae). *Mol. Biol. and Evol.*13: 510-524.

Swofford DL, Waddell PJ, Huelsenbeck JP, Foster PG, Lewis PO, Rogers JS (2001) Bias in phylogenetic estimation and its relevance to the choice between parsimony and likelihood methods. *Syst Biol.* Aug;50(4):525-39.

Sylla M, Bosio C, Urdaneta-Marquez L, Black WC (2009) Gene flow, subspecies composition, and dengue virus-2 susceptibility among *Aedes aegypti* collections in Senegal. *PLOS Neglected Tropical Diseases* (IN PRESS).

Tabachnick WJ (1991) The evolutionary relationships among arboviruses and the evolutionary relationships of their vectors provides a method for understanding vector-host interactions. *J Med Entomol.* May;28(3):297-8.

Tabachnick WJ, Powell JR (1979) A world-wide survey of genetic variation in the yellow fever mosquito, Aedes aegypti. *Genet Res.* Dec;34(3):215-29.

Tabachnick WJ, Wallis GP, Aitken THG, Miller BR, Amato GD, Lorenz L, Powell JR, Beaty BJ (1985) Oral Infection of *Aedes aegypti* with Yellow-Fever Virus - Geographic-Variation and Genetic Considerations. *Am J Trop Med Hyg* 34:1219-1224.

Tajima F (1983) Evolutionary relationshyip of DNA sequences in finite populations. *Genetics.* 105:437-460.

Takahashi M (1976) The effects of environmental and physiological conditions of *Culex tritaeniorhynchus* on the pattern of transmission of Japanese encephalitis virus. *J. Med. Entomol.* 13:275-284.

Takahashi M, Suzuki K (1979) Japanese encephalitis virus in mosquito salivary glands. *Am J. Trop. Med. Hyg.* Jan:28(1):122-35.

Tanaka T, Nei M (1989) Positive Darwinian selection observed at the variable region genes of immunoglobulins. *Mol Biol Evol.* Sep;6(5):447-59.

Tardieux I, Poupel O, Lapchin L, Rodhain F (1990) Variation among strains of *Aedes aegypti* in susceptibility to oral infection with dengue virus type 2. *Am J Trop Med Hyg.* Sep;43(3):308-13.

Tardieux I, Poupel O, Lapchin L, Rodhain R (1991) Analysis of inheritance of oral susceptibility of *Aedes aegytpi* (Diptera: Culicidae) to dengue-2 virus using isofemale lines. *J Med Entomol.* 28:518-521.

Tesh RB (1980) Infectivity and pathogenicity of Nodamura virus for mosquitoes. *J. Gen. Virol.* 48:177-182.

Tewfik HR, Barr AR (1976) Paracentric inversion in *Culex pipiens*. *J Med Entomol* 13:147-150.

Thompson CB (1995) Apoptosis in the pathogenesis and treatment of disease. *Science* 267:1456-1462.

Thornberry NA, Lazebnik Y (1998) Capases: enemies within. *Science* 281: 1312-1316.

Tomari Y, Zamore PD (2005) Perspective: machines for RNAi. *Gene Dev.* Mar 1;19(5):517-29.

Tomari Y, Matranga C, Haley B, Martinez N, Zamore PD (2004) A protein sensor for siRNA asymmetry. *Science*. Nov 19;306(5700):1377-80.

Tomori O (2004) Yellow fever: the recurring plague. *Crit Rev Clin Lab Sci*. 41(4):391-427.

Townsend, J. P. and Rand, D. M. (2004) Mitochondrial genome size variation in New World and Old World populations of *Drosophila melanogaster*. *Heredity* 93: 98-103.

Triant, D. A. and DeWoody, J. A. (2008) Molecular analyses of mitochondrial pseudogenes within the nuclear genome of arvicoline rodents. Genetica 132: 21-33

Trpis M, Hausermann W (1975) Demonstration of differential domesticity of *Aedes aegypti* (L) (Diptera, Culicidae) in Africa by mark release recapture. *Bull Entomol Res* 65:199-208.

Trpis M, Hausermann W (1978) Genetics of House-Entering Behavior in East-African Populations of Aedes aegypti (L) (Diptera-Culicidae) and Its Relevance to Speciation. *Bull Entomol Res* 68:521-532.

Twiddy SS, Farrar JJ, Vinh Chau N, Wills B, Gould EA, Gritsun T, Lloyd G, Holmes EC (2002) Phylogenetic relationships and differential selection pressures among genotypes of dengue-2 virus. *Virology*. Jun 20;298(1):63-72.

Twiddy SS, Holmes EC, Rambaut A (2003) Inferring the rate and time-scale of dengue virus evolution. *Mol Biol Evol*. Jan;20(1):122-9.

Undeen AH, Fukuda T (1994) Effects of host resistance and injury on the susceptibility of *Aedes taeniorhynchus* to mosquito iridescent virus. *J. Am. Mosq. Control Assoc*. 10:64-66.

Urdaneta-Marquez L, Bosio C, Herrera F, Rubio-Palis Y, Salasek M, Black WC (2008) Genetic relationships among *Aedes aegypti* collections in Venezuela as determined by mitochondrial DNA variation and nuclear single nucleotide polymorphisms. *Am J Trop Med Hyg* 78:479-491.

Urdaneta L, Herrera F, Pernalete M, Zoghbi N, Rubio-Palis Y, Barrios R, Rivero J, Comach G, Jiménez M, Salcedo M (2005) Detection of dengue viruses in field-caught *Aedes aegypti* (Diptera: Culicidae) in Maracay, Aragua state, Venezuela by type-specific polymerase chain reaction. *Infect Genet Evol*. Mar;5(2):177-84.

Vaidyanathan R, Scott TW (2006)Apoptosis in mosquito midgut epithelia associated with West Nile virus infection. *Apoptosis*. Sep;11(9):1643-51.

Vaidyanathan R, Fleisher AE, Minnick SL, Simmons KA, Scott TW (2008) Nutritional stress affects mosquito survival and vector competence for West Nile virus. *Vector Borne Zoonotic Dis*. Dec;8(6):727-32.

van der Krol AR, Mur LA, Beld M, Mol JN, Stuitje AR (1990) Flavonoid genes in petunia: addition of a limited number of gene copies may lead to a suppression of gene expression. *Plant Cell*. Apr;2(4):291-9.

van der Krol AR, Mur LA, de Lange P, Mol JN, Stuitje AR (1990) Inhibition of flower pigmentation by antisense CHS genes: promoter and minimal sequence requirements for the antisense effect. *Plant Mol Biol*. Apr;14(4):457-66.

Van ooijen JW (1994) DrawMap: a computer program for drawing genetic linkage maps. *J. Hered*. 85:66.

Vasilakis N, Holmes EC, Fokam EB, Faye O, Diallo M, Sall AA, Weaver SC (2007) Evolutionary processes among sylvatic dengue type 2 viruses. *J Virol.* Sep;81(17):9591-5.

Vasilakis N, Shell EJ, Fokam EB, Mason PW, Hanley KA, Estes DM, Weaver SC (2007) Potential of ancestral sylvatic dengue-2 viruses to re-emerge. *Virology.* Feb 20;358(2):402-12.

Vasilakis N, Weaver SC (2008) The history and evolution of human dengue emergence. *Adv Virus Res.* 72:1-76.

Vazeille-Falcoz M, Failloux AB, Mousson L, Elissa N, Rodhain F (1999) Oral receptivity of *Aedes aegypti formosus* from Franceville (Gabon, Central Africa) for dengue type 2 virus. *Bulletin de la Societe de Pathologie Exotique* 92:341-342.

Venter PA, Schneemann A (2008) Recent insights into the biology and biomedical applications of Flock House virus. *Cell Mol Life Sci.* Sep;65(17):2675-87.

Vora N (2008) Impact of anthropogenic environmental alterations on vector-borne diseases. *Medscape J Med.* 10(10):238.

Wallis GP, Aitken TH, Beaty BJ, Lorenz L, Amato GD, Tabachnick WJ (1985) Selection for susceptibility and refractoriness of *Aedes aegypti* to oral infection with yellow fever virus. *Am J Trop Med Hyg.* Nov;34(6):1225-31.

Wallis GP, Tabachnick WJ, Powell JR (1983) Macrogeographic genetic variation in a human commensal: *Aedes aegypti*, the yellow fever mosquito. *Genet. Res.* 41:241-258.

Wallis GP, Tabachnick WJ, Powell JR (1984) Genetic heterogeneity among Caribbean populations of *Aede aegypti. Am. J. Trop. Med. Hyg.* 33:492-798.

Wang H, Blair CD, Olson KE, Clem RJ (2008) Effects of inducing or inhibiting apoptosis on Sindbis virus replication in mosquito cells. *J Gen Virol.* Nov;89(Pt 11):2651-61.

Wang J, Chuang K, Ahluwalia M, Patel S, Umblas N, Mirel D, Higuchi R, Germer S (2005) High-throughput SNP genotyping by single-tube PCR with Tm-shift primers. *Biotechniques.* Dec;39(6):885-93.

Wang SL, Hawkins CJ, Yoo SJ, Müller HA, Hay BA (1999) The *Drosophila* caspase inhibitor DIAP1 is essential for cell survival and is negatively regulated by HID. *Cell.* Aug 20;98(4):453-63.

Wang XH, Aliyari R, Li WX, Li HW, Kim K, Carthew R, Atkinson P, Ding SW (2006) RNA interference directs innate immunity against viruses in adult *Drosophila. Science* 312:452-454.

Ward TW, Kimmick MW, Afanasiev BN, Carlson JO (2001) Characterization of the structural gene promoter of *Aedes aegypti* densovirus. *J. Virol.* 75:1325-1331.

Waterman SH, Gubler DJ (1989) Dengue fever. *Clin. Dermatol.* 7:117-122.

Waterhouse RM, Kriventseva EV, Meister S, Xi Z, Alvarez KS, Bartholomay LC, Barillas-Mury C, Bian G, Blandin S, Christensen BM, Dong Y, Jiang H, Kanost MR, Koutsos AC, Levashina EA, Li J, Ligoxygakis P, Maccallum RM, Mayhew GF, Mendes A, Michel K, Osta MA, Paskewitz S, Shin SW, Vlachou D, Wang L, Wei W, Zheng L, Zou Z, Severson DW, Raikhel AS, Kafatos FC, Dimopoulos G, Zdobnov EM, Christophides GK. (2007) Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes. *Science.* Jun 22;316(5832):1738-43.

Waterhouse PM, Wang MB, Finnegan EJ (2001) Role of short RNAs in gene silencing. *Trends Plant Sci.* Jul;6(7):297-301.

Waterhouse PM, Wang MB, Lough T (2001) Gene silencing as an adaptive defence against viruses. *Nature.* Jun 14;411(6839):834-42.

Watts DM, Burke DS, Harrison BA, Whitmire RE, Nisalak A (1987) Effect of temperature on the vector efficiency of *Aedes aegypti* for dengue 2 virus. *Am J Trop Med Hyg.* Jan;36(1):143-52.

Watts DM, Porter KR, Putvatana P, Vasquez B, Calampa C, Hayes CG, Halstead SB (1999) Failure of secondary infection with American genotype dengue 2 to cause dengue haemorrhagic fever. *Lancet.* Oct 23;354(9188):1431-4.

Weaver SC, Lorenz LH, Scott TW (1992) Pathologic changes in the midgut of *Culex tarsalis* following infection with Western equine encephalomyelitis virus. *Am J Trop Med Hyg.* Nov;47(5):691-701.

Weaver SC, Scott TW, Lorenz LH, Lerdthusnee K, Romoser WS (1988) Togavirus-associated pathologic changes in the midgut of a natural mosquito vector. *J Virol.* Jun;62(6):2083-90.

Wei, XC, Dohkan J, Kishi H, Wu CX, Kondo S, Muraguchi A (2005) Characterization of the proximal enhancer element and transcriptional regulatory factors for murine recombination activating gene-2. *European J Immuno* 35:612-621.

Wei, XC, Kishi H, Jin ZX, Zhao WP, Kondo S, Matsuda T, Saito S, Muraguchi A (2002) Characterization of chromatin structure and enhancer elements for murine recombination activating gene-2. *J Immuno* 169:873-881.

Whitfield SG, Murphy FA, Sudia WD (1973) St. Louis encephalitis virus: an ultrastructural study of infection in a mosquito vector. *Virology.* 56:70.

Whitfield SG., Murphy FA, Sudia WD (1971) Easter equine encephalomyelitis virus: an electron microscopic study of *Aedes triseriatus* (Say) salivary gland infection. *Virology.* 43:110.

Woischnik, M. and Moraes, C. T. (2002) Pattern of organization of human mitochondrial pseudogenes in the nuclear genome. *Genome Res.* 12: 885-893.

Wondji CS, Morgan J, Coetzee M, Hunt RH, Steen K, Black WC 4th, Hemingway J, Ranson H (2007) Mapping a quantitative trait locus (QTL) conferring pyrethroid resistance in the African malaria vector *Anopheles funestus. BMC Genomics.* 2007 Jan 29;8:34.

Wong J, Tripet F, Rasgon JL, Lanzaro GC, Scott TW (2008) SSCP analysis of scnDNA for genetic profiling of *Aedes aegypti. Am J Trop Med Hyg.* Oct;79(4):511-7.

Woodard DB, Chapman HC (1968) Laboratory studies with the mosquito iridescent virus (MIV). *J. Invertebr. Pathol.* 11:296-301.

Woodring JL, Higgs S, Beaty BJ (1996) Natural cycles of vector-borne pathogens. *The Biology of Disease Vectors.* University Press of Colorado., pp. 51-72.

World Health Organization (2009) WHO, Dengue and Dengue Haemorrhagic Fever, 2009. http://www.rbm.who.int/cmc_upload/0/000/015/372/RBMInfosheet_1.htm

Yamakawa M, Furuichi Y, Shatkin AJ (1982) Priming of reovirus transcription by GppppG and formation of CpG(5')GpC. *Proc Natl Acad Sci* U S A. Oct;79(20):6142-6.

Yang QY, Zhang DF, Li Q, Cheng ZK, Xue YB (2007) Heterochromatic and genetic features are consistent with recombination suppression of the self-incompatibility locus in Antirrhinum. *Plant* J 51:140-151.

Zambetaki A, Pasteur N, Mavragani-Tsipidou P (1998) Cytogenetic analysis of Malpighian tubule polytene chromosomes of *Culex pipiens* (Diptera : Culicidae). *Genome* 41:751-755.

Zambon RA, Vakharia VN, Wu LP (2006) RNAi is an antiviral immune response against a dsRNA virus in *Drosophila melanogaster*. *Cell Microbiol.* May;8(5):880-9.

Zhang, D. X. and Hewitt, G. M. (1996) Highly conserved nuclear copies of the mitochondrial control region in the desert locust *Schistocerca gregaria:* Some implications for population studies. *Mol. Ecol.* 5: 295-300.

Zheng L, Benedict MQ, Cornel AJ, Collins FH, Kafatos FC (1996) An integrated genetic map of the African human malaria vector mosquito, *Anopheles gambiae*. *Genetics*. Jun;143(2):941-52.

Zhong D, Menge DM, Temu EA, Chen H, Yan G (2006) Amplified fragment length polymorphism mapping of quantitative trait loci for malaria parasite susceptibility in the yellow fever mosquito *Aedes aegypti*. *Genetics*. Jul;173(3):1337-45.

Zhou L, Jiang G, Chan G, Santos CP, Severson DW, Xiao L (2005) Michelob_x is the missing inhibitor of apoptosis protein antagonist in mosquito genomes. *EMBO Rep*. Aug;6(8):769-74.

Zieler H, Garon CF, Fischer ER, Shahabuddin M (1998) Adhesion of *Plasmodium gallinaceum* ookinetes to the *Aedes aegypti* midgut: sites of parasite attachment and morphological changes in the ookinete. *J. Eukaryot. Microbiol.* 45:512-520.

Zieler H, Nawrocki JP, Shahabuddin M (1999) *Plasmodium gallinaeceum* ookinetes adhere specifically to the midgut epithelium of *Aedes aegypti* by interaction with a carbohydrate ligand. *J. Exp. Biol.* 202:485-495.