DISSERTATION

EVIDENCE FOR RNA INTERFERENCE AS AN ANTIVIRAL IMMUNE RESPONSE TO ALPHAVIRUS INFECTION IN THE MOSQUITOES *ANOPHELES GAMBIAE* AND *AEDES AEGYPTI*

Submitted by

Kimberly Marie Keene

Department of Microbiology, Immunology, and Pathology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2006

COLORADO STATE UNIVERSITY

April 3, 2006

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY KIMBERLY MARIE KEENE ENTITLED "EVIDENCE FOR RNA INTERFERENCE AS AN ANTIVIRAL IMMUNE RESPONSE TO ALPHAVIRUS INFECTION IN THE MOSQUITOES A*NOPHELES GAMBIAE* AND *AEDES AEGYPTI"* BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work

Co-advisor

Advisor

Department Head

Abstract of Dissertation

"RNA interference as an antiviral response to alphavirus infection in the mosquitoes *Anopheles gambiae* and *Aedes aegypti*"

Mosquito-borne viruses remain a serious threat to public heath worldwide. A lack of effective vaccines and vector control programs have led to the geographical expansion of the vectors and their associated agents, thereby increasing the percent of the global population that are affected. The studies described here are designed to determine whether RNA interference occurs in the mosquito and functions to naturally target replicating viruses in the mosquito. The results will provide information that will lead to the development of new strategies to control arbovirus transmission.

First, potential genes in the mosquito RNAi pathway were identified in the *Anopheles gambiae* genome. Genes in the *dicer* and *argonaute* gene families were tested for their involvement in the RNAi pathway in *An. gambiae* cells using a dual-luciferase assay and dsRNA knockdown techniques. These experiments identified the *Dcr2*, *Ago2*, and *Ago3* genes from *An. gambiae* as genes involved in the RNAi pathway.

To test the hypothesis that the RNAi pathway is a part of the mosquitoe's innate defenses against virus infection, genes in the pathway were targeted using RNAi techniques to reduce gene expression. Simultaneous injection of dsRNA with sequences homologous to *dcr* and *ago* genes and infection with o'nyong-nyong virus (ONNV) (*Alphavirus*; *Togaviridae*) in *An. gambiae* resulted in a significant increase in viral titers and increased dissemination of the virus in those mosquitoes. Additionally, injection of

the virus along with dsRNA targeting a viral gene resulted in decreased viral replication in the mosquito. These studies confirmed that ONNV replication in *An. gambiae* is naturally targeted by the RNAi pathway.

To test the hypothesis that the RNAi response would interfere with Sindbis virus replication in *Aedes aegypti,* dsRNA with sequence homologous to a gene from the TR339 strain was injected into mosquitoes three days prior to the infectious bloodfeed. This demonstrated that RNA interference could be used to prevent TR339 infection of the midgut. In addition, dsRNA homologous to *Ae. aegypti argonaute-2* injected into the mosquito prior to an infectious bloodfeed resulted in increased virus replication in the midgut, as well as increased dissemination from midgut tissues.

This work, in conjunction with previous work based on RNA-mediated interference o virus replication in mosquitoes, provides a basis for additional studies that could lead to generating RNAi-based arbovirus resistance in mosquitoes.

Kimberly Marie Keene
Department of Microbiology, Immunology, and Pathology
Colorado State University
Fort Collins, CO 80523
Spring 2006

Acknowledgements

First, I would like to thank my advisors, Drs. Carol Blair and Ken Olson for accepting me into their lab and also for their guidance and support. I would like to thank Dr. Barry Beaty for his many lessons in bio-politics. I would also like to thank Drs. Jon Carlson, Paul Layborne, and Ann Powers for their time and commitment while serving on my graduate committee.

I would also like to thank the many people that provided assistance with my studies. Erik Powers and Cindy Meredith helped raise mosquitoes for my experiments. Drue Webb and Sarah VanOtterloo provided much assistance with any task that was asked of them, and I am very grateful for their help. Dr. Brian Foy contributed many hours of discussion for developing the design of my projects, and also help me to interpret the results in the "big picture". I am thankful for his insights and patience. I would like to thank Dr. Irma Sanchez-Vargas for her technical assistance.

I would especially like to thank my fellow graduate students and friends for making my time at AIDL so enjoyable, especially Isabel Salazar-Sanchez, Chris Cirimotich, Bethany Bolling, Eric Beck, Jason Richardson, and Doug Brackney.

Dedication

This work is dedicated to my loving parents, who taught me that anything is possible.

TABLE OF CONTENTS

Chapter 1

Literature Review

Introduction

Mosquito-borne diseases represent a serious threat to global health. Many of these diseases are considered emerging or re-emerging. Diseases such as malaria, filariasis, and dengue fever cause significant morbidity and mortality each year. It is estimated that over 300 million people are infected with malaria each year worldwide, resulting in over one million deaths (WHO 2004). Arboviruses (arthropod-borne viruses) account for a large percentage of the agents transmitted by mosquitoes, and are responsible for severe morbidity and mortality in human and animal populations around the world. The last 30 years have seen a dramatic increase in the emergence and re-emergence of arboviruses. Increases in urbanization, global travel, and a lack of effective vaccines and control measures have allowed for an unchecked geographic expansion of the vectors and their associated agents.

Alphaviruses

The *Alphavirus* genus (family *Togaviridae*) is composed of 29 viruses, all of which are mosquito-borne (International Committee on Taxonomy of Viruses- www.ncbi.nlm.nih.gov/ICTVdb/ICTV/index.htm). Alphaviruses are worldwide in distribution; however, the geographical distribution of each virus is limited, following the distribution of the vector (Calisher and Karabatsos, 1988). Mosquitoes are the primary arthropod vectors for alphaviruses, although a few viruses have been isolated from mites and ticks. Alphaviruses are generally maintained in a mosquito-bird/small mammal transmission cycle. In general, larger mammals such as horses or humans are incidental or dead-end hosts and are not involved in the maintenance of the virus. Some viruses,

however, such as Venezuelan equine encephalitis (VEE) can be amplified in horses and subsequently transmitted to mosquitoes during epidemics (Weaver et al., 2004).

Alphaviruses represent a serious threat to human health. Those that cause human disease are generally divided into two groups: New World and Old World. New World alphaviruses such as eastern equine encephalitis virus (EEE), western equine encephalitis virus (WEE), and Venezuelan equine encephalitis virus (VEE) are capable of causing fatal encephalitis in humans. Other members, including the Old World alphaviruses chikungunya (CHIK) and o'nyong-nyong viruses (ONN), cause febrile illnesses although little mortality is associated with infection. Ross River virus (RR) causes epidemic polyarthritis and the symptoms of this disease may persist for years (Strauss and Strauss, 1994). The distinction between the groups is based on geographic location and genetic and antigenic determinants (Calisher et al., 1980). Alphaviruses are categorized based on serologic cross-reactivity (Strauss and Strauss, 1994; Calisher and Karabatsos, 1988).

Antigenic groups of alphavirus are based on the reactivity of viruses to antibodies generated against a specific virus. Eight antigenic complexes exist for grouping alphaviruses (Calisher et al., 1980; Powers et al., 2001). Genetic evidence for relationships between alphaviruses is based on nucleotide sequence homologies. RNA-RNA hybridization and RNase T1 mapping techniques, and sequence analysis of viral genomes have been employed for determining these relationships (Olson and Trent, 1985; Weaver et al., 1992; Wengler and Filipe, 1977). Nucleotide sequence information is available for a number of alphavirus genomes, including VEE, RR, Semliki Forest virus (SF), Sindbis virus (SIN), ONN, EEE, and WEE (Faragher et al., 1988; Garoff et

al., 1980; Kinney et al., 1989; Levinson et al., 1990; Strauss et al., 1984; Takkinen, 1986).

Alphavirus Biology:

Alphaviruses are single-stranded, positive sense RNA viruses with genomes of ~11.8 kB in length. The viral RNA is packaged in an icosahedral capsid and surrounded by a lipid bilayer envelope that contains two prominent viral glycoproteins. The 5′ two-thirds of the genome encodes the non-structural proteins nsP1-4, and the 3′ one-third of the genome encodes the structural proteins: the capsid, two envelope glycoproteins (E1 and E2), E3 and 6K. Viral RNAs contain a 5′ 7-methylguanosine cap and the 3′ end is polyadenylated. The structural proteins are expressed from a subgenomic mRNA (26S RNA) that is produced from the full-length negative-sense RNA template in infected cells during replication.

Alphavirus capsids have an icosahedral symmetry of T=4 (Choi et al., 1991; Coombs and Brown, 1987). The interactions of the glycoproteins E1 and E2 with the capsid protein allows for the formation of a stable heterodimer. Three E1-E2 heterodimers interact to form the glycoprotein spike on the viral envelope (Rice and Strauss, 1982). The E2 glycoprotein is formed from the precursor PE2. The precursor is cleaved into E2, E3 and a small hydrophobic protein 6k (Strauss and Strauss, 1994). The specific function of E3 is currently unknown. E3 remains associated with the virion in some instances, as with SFV, but not with others (SINV) (Heidner et al., 1996; Strauss and Strauss, 1994). PE2 and E1 are glycosylated and pass through the endoplasmic reticulum (ER) and the Golgi to the plasma membrane, where virus assembly occurs. Cleavage of PE2 occurs in the Golgi complex (de Curtis and Simons, 1988; Naim and

Koblet, 1990). The capsid protein contains a serine protease activity that autocatalytically cleaves itself from the polyprotein (Strauss and Strauss, 1994). The capsid binds specifically to both newly made viral RNA as well as to the large subunit of ribosomes (Weiss et al., 1989).

Alphavirus replication occurs in the cytoplasm of infected cells. The virus gains entry into the cell by receptor-mediated endocytosis (RME) following attachment to the cellular receptor. The virus is then trafficked to cellular compartments in endosomal vesicles. The endosomes containing the viral particles become acidified, leading to conformational changes in E1 and E2, and fusion of the viral and cell membranes, facilitating the release of the viral capsids into the cytoplasm. Once the genomic RNA is released into the cytoplasm, it serves as a mRNA and the non-structural proteins are directly translated as a polyprotein. Non-structural proteins are translated from genomic RNA to form two polyproteins, nsP123 and nsP1234. Some alphaviruses, such as SIN, contain a leaky opal termination codon that allows for both polyproteins to be produced (Feng et al., 1990; Shirako and Strauss, 1994). ONN does not contain an opal codon (Levinson et al., 1990; Takkinen, 1986). Both 5' and 3' untranslated regions (UTRs) are important for replication and the regulation of viral RNA synthesis (Strauss and Strauss, 1994).

The alphavirus nsP1 protein has 3 known functions. nsP1 is required for the initiation of synthesis of the negative strand RNA after infection (Hahn et al., 1989; Wang et al., 1991). nsP1 also contains the enzyme activity required for capping the viral RNAs for translation (Mi et al., 1989). Both the genomic and subgenomic RNAs can be

capped by nsP1. Finally, nsP1 plays a role in modulating the proteinase activity of nsP2 (de Groot et al., 1990).

The N-terminus of the nsP2 protein has RNA helicase activity and is functional during RNA replication and transcription, regulating minus-strand RNA synthesis (Strauss and Strauss, 1994). nsP2 is also required at the initiation step for the synthesis of 26S RNA (Suopanki et al., 1998). The role of nsP3 during alphavirus infection is not well understood. Experiments performed with temperature sensitive mutants of nsP3 have shown that the protein is required for RNA syntheses (LaStarza et al., 1994). The viral polymerase is encoded by nsP4. The concentration of nsP4 is highly regulated in infected cells (Strauss and Strauss, 1994).

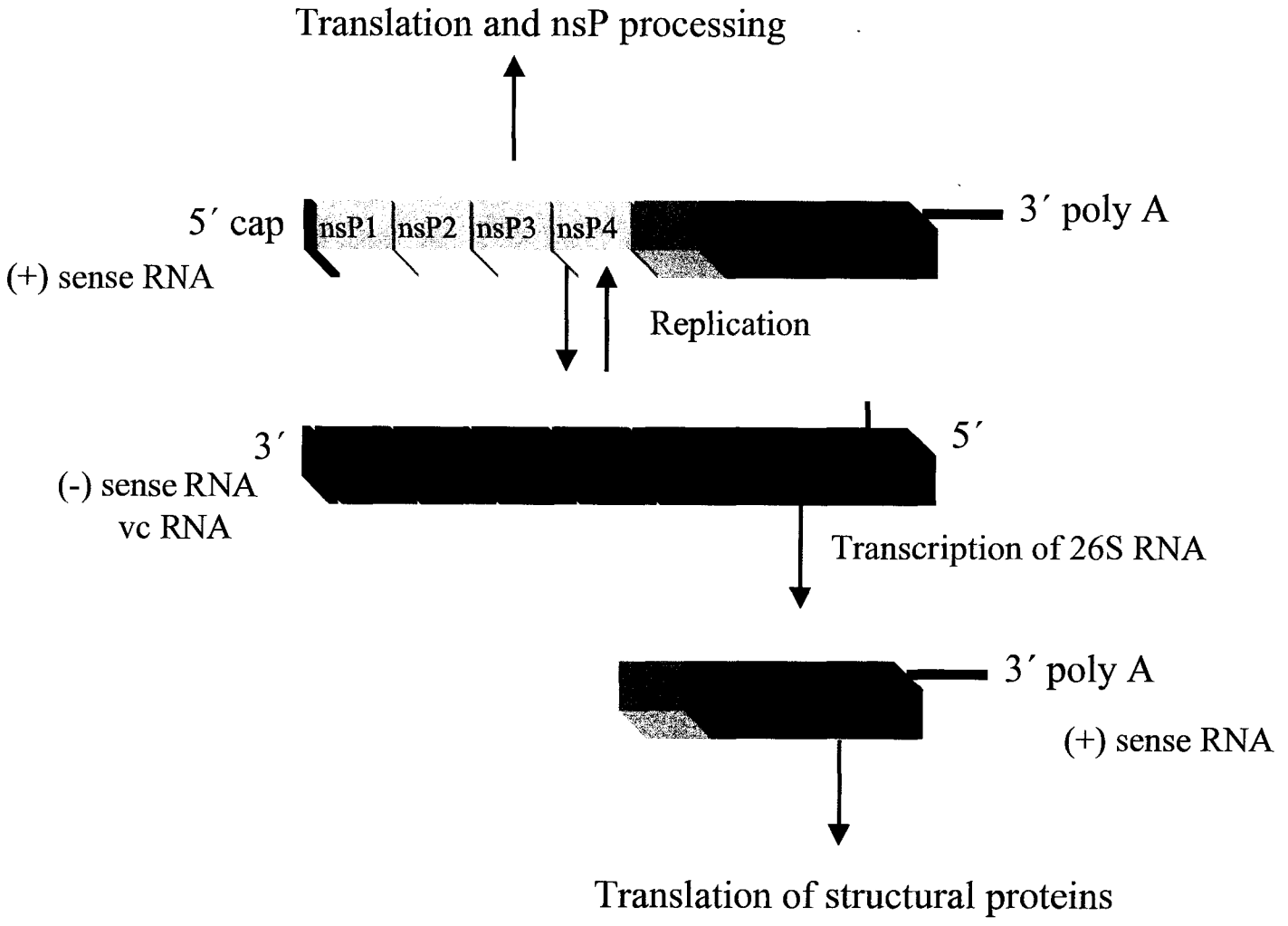
Once the non-structural proteins are produced, synthesis of the minus-strand RNA begins utilizing the minus-strand replicase, which consists of the polyprotein P123. The minus-strand RNA is a template for the synthesis of more genomic positive strand RNAs, formed by the plus-strand replicase, which includes nsP1-4. The negative strand RNA is also a template for the transcription of the 26S subgenomic mRNA. Structural proteins are translated from this subgenomic RNA as a polyprotein. The polyprotein including the structural proteins is proteolytically processed in the endoplasmic reticulum (ER) and the Golgi complex to form the mature proteins. In vertebrate cells, the E1 and E2 glycoproteins remain associated after translocation as a heterodimer and are inserted into the plasma membrane to be incorporated into the virus as it buds from the cell. Alphaviruses also bud into cytoplasmic vacuoles, and intracellular budding may be the primary mode of virus maturation in mosquito cells (Miller and Brown, 1992; Raghow et al., 1973a; Raghow et al., 1973b).

After alphavirus infection, RNA replicase complexes assemble on membranes within the cell (Barton et al., 1991). Cellular proteins are thought to be involved with virus replication and one such protein, the La autoantigen, has been identified. La protein binds to the 3′ end of RNA transcripts and can facilitate multiple rounds of transcription reinitiation (Gottlieb and Steitz, 1989). Human La protein interacts with the 3′ end of other viral RNAs, including adenovirus and HIV mRNAs. Purified La protein has been demonstrated to bind to RNA sequences contained in the 3′ end of Sindbis virus minus-strand RNA, which are believed to contain a promoter for virus RNA replication (Pardigon and Strauss, 1996). These findings indicate La protein may play an important role in Sindbis virus replication.

Alphaviruses regulate negative-strand RNA synthesis (Sawicki et al., 1981). Negative-strand RNAs are produced until 3 to 6 hours post infection, when positive-strand RNA synthesis begins. In order for negative-strand synthesis to proceed, the P123 complex must remain intact and interact with the polymerase. Cleavage of P123 into nsP1, nsP2, and nsP3 by the protease activity in nsP2 prevents negative strand synthesis from occurring and is necessary for the commencement of positive strand RNA synthesis (Strauss and Strauss, 1994).

Alphavirus genomic RNA is encapsidated into virions using the encapsidation signal found in the 5′ region of the genome (Geigenmuller-Gnirke et al., 1991). The newly formed capsids interact with the vertebrate host-cell plasma membrane in regions where E1 and E2 have been inserted to bud from the cell, acquiring both the envelope and its embedded glycoproteins (Cadd et al., 1997; Frolov et al., 1996).

Figure 1.1 Replication and Transcription of Alphavirus RNA



O'nyong nyong virus (*Alphavirus*; *Togaviridae*)

O'nyong-nyong virus (ONNV) (family *Togaviridae*, genus *Alphavirus*) is a member of the Semliki Forest antigenic complex (Karabatsos, 1975). This virus is unique among alphaviruses in that it appears to be transmitted by anopheline mosquitoes (*Anopheles gambiae* and *An. funestus*). No vertebrate reservoir has yet been identified for this virus, leading to the hypothesis that this virus may be maintained in a mosquito-human transmission cycle (Lanciotti et al., 1998). Few other arboviruses are known to infect *An. gambiae* (Karabatsos, 1985). ONNV is closely related to another alphavirus, chikungunya virus (CHIKV), which is transmitted by *Aedes aegypti*. Serologic studies have shown that CHIKV antibodies react strongly to CHIK and ONN viruses, while antibodies to ONNV react strongly with ONNV and weakly with CHIKV (Blackburn et al., 1995; Lee et al., 1997). Furthermore, phylogenetic analysis of the E1 gene sequence from CHIKV and ONNV isolates shows that the two viruses are genetically distinct with approximately 28% divergence at the nucleotide level (Powers et al., 2000).

ONNV has been responsible for three major documented outbreaks of disease in Central Africa. The first, during 1959-1962, caused an epidemic that involved over two million people (Williams et al., 1965). Serologic evidence showed that the virus continued to be transmitted to humans at a low level until the late 1960's (Lutwama et al., 1999). A second outbreak, recorded from 1996-1997 in Uganda, was also caused by ONNV but was not as extensive as the earlier outbreak. A third outbreak during the fall of 2003 in Cote d'Ivoire was recently reported (Posey et al., 2005). The ONNV strain isolated in northern Uganda 1959 is known as Gulu. The Igbo Ora strain, isolated in Uganda in 1969 was originally thought to be a distinct alphavirus, but sequence analysis

has shown it is a strain of ONNV (Lanciotti et al., 1998). The strain isolated from the 1996-97 outbreak has been designated SG650.

The clinical criteria for ONNV infection during the 1996-1997 outbreak were an acute febrile illness with polyarthralgia, rash and lymphadenopathy, along with serological evidence of infection or direct virus isolation from human sera (Kiwanuka et al., 1999). While the infection can result in an acute febrile disease, no deaths have been associated with ONNV infection (Haddow et al., 1960).

Sindbis virus

Sindbis virus (SINV) is the prototype for the Alphavirus genus. The first SINV, the AR339 strain, was isolated from a pool of *Culex univittatus* mosquitoes in the village of Sindbis, Egypt, in 1952 (Taylor et al., 1955). SINV is maintained in a transmission cycle that involves *Culex* and *Culiseta spp.* mosquitoes and wild birds (Mackenzie et al., 1994). SINV has one of the largest geographical distributions among alphaviruses, and has been isolated from both mosquitoes and vertebrates on several continents, including Europe, Africa, Asia, and Australia. SINV are grouped based on antigenic and genetic analyses (Olson and Trent, 1985) and fall into two distinct groups: the Paleoarctic/Ethiopian and the Oriental/Australian. The majority of human infections with SINV result in subclinical or very mild diseases. The clinical features of SINV infection include fever, joint pain and rash (Skogh and Espmark, 1982).

Infectious cDNA clones

The first cDNA clone from an RNA virus was produced by Taniguchi et al (Taniguchi et al., 1978). Bacteriophage Qβ RNA was reverse-transcribed and the resulting DNA was cloned into a plasmid. The plasmid was introduced into *E. coli*, where the phage RNA was transcribed and infectious bacteriophage was produced. The first cDNA clone of an animal RNA virus was generated from poliovirus RNA in 1981 (Racaniello and Baltimore, 1981). Since then, infectious cDNA clones have been generated for many plant and animal viruses. These clones benefit the study of RNA viruses by allowing site-directed mutagenesis for the study host-virus interactions, identifying potential translation products, and identifying *cis*-acting sequences involved with replication and transcription (Rice et al., 1987).

Infectious clone technology is available for many arboviruses with positive-sense RNA genomes, including SINV and ONNV. Generating DNA clones of viruses with RNA genomes allows for easy manipulation of the genetic sequence of the virus. The virus is self-replicating and self-packaging, and foreign sequences can be introduced and expressed as part of the viral genome. Prior to the development of infectious clone technology for SINV, most of the genetic analysis was performed using temperature sensitive mutants (Rice et al., 1987).

The first infectious clone of an alphavirus was generated by Rice et al (1987). Since the original construction, additional constructs and adaptations to the constructs have been made to produce viruses that are better suited for mosquito-virus interaction studies. The use of alphaviruses for transient expression systems has a number of benefits, including the ability to infect many different species, including birds, mammals

and insects, and high levels of RNA and protein expression in infected cells (Frolov et al., 1996). The two predominant strategies for employing alphavirus expression are through infectious clones and replicon systems (Frolov et al., 1996). Infectious clones have now been constructed for a number of alphaviruses, including SINV (Rice et al., 1987), ONNV (Brault et al., 2004), SFV (Liljestrom et al., 1991), VEE (Davis et al., 1989), RRV (Kuhn et al., 1991), and CHIKV (Vanlandingham et al., 2005).

Sindbis viruses derived from double-subgenomic infectious clones have greatly facilitated the study of virus-vector interactions. Double-subgenomic SINV (dsSINV) contain a duplicated subgenomic mRNA promoter that can express foreign genetic material (Kamrud et al., 1997). This material is expressed at an abundant level and persists in the mosquito cell for many days (Kamrud et al., 1997). Viruses have been engineered to express reporter genes, which make visualization of the expression patterns in the mosquito possible (Davis et al., 1989; Kuhn et al., 1991; Liljestrom et al., 1991; Myles et al., 2003; Olson et al., 2000; Pierro et al., 2003; Rice, 1996; Rice et al., 1987; Seabaugh et al., 1998; Strauss and Strauss, 1994). The first double subgenomic SIN (dsSIN) viruses developed for use in mosquitoes expressed chloramphenicol acetyltransferase (CAT) as a marker of infection visualized by immunofluorescence assay (Kamrud et al., 1997; Olson et al., 1994). A second generation of dsSIN viruses was engineered to express green fluorescent protein (GFP), which allows real time profiling of virus infection when viewed under UV light (Olson et al., 2000; Pierro et al., 2003).

SINV expression systems are excellent alternatives for expression of exogenous gene material in mosquitoes without the time and technical difficulty of developing a transgenic mosquito. Sindbis viruses have been engineered to express proteins, anti-

pathogen molecules, such as antisense and hairpin RNAs, as well as single chain antibodies (Cheng et al., 2001; de Lara Capurro et al., 2000; Gaines et al., 1996; Powers et al., 1996; Powers et al., 1994). Cheng and colleagues (Cheng et al., 2001) developed dsSIN viruses that contained the coding sequence for *Ae. aegypti* defensin, and showed that the full protein was expressed in *Ae. aegypti* upon infection. Based on immunofluorescence, the defensin protein co-localized to the same cells that were positive for E1 glycoprotein expression.

Antisense RNA to La Crosse vRNA, when expressed by dsSINV (TE/3'2J) in infected mosquitoes and cultured mosquito cells, was able to significantly reduce infection when the cells were challenged with the homologous virus (Powers et al., 1996; Powers et al., 1994). Gaines *et al* (Gaines et al., 1996) demonstrated that C6/36 (*A. albopictus*) mosquito cells infected with a dsSIN virus expressing a portion of the pre-membrane (prM) coding region of DEN-2 genome in either positive or negative sense were resistant to DEN-2 infection. Cells infected with the empty dsSIN vector were susceptible to infection. In addition, there was no protection given to the cells when challenged with DEN-3 or DEN-4 viruses, showing that the mechanism of protection was sequence dependent (Gaines et al., 1996).

Adelman and colleagues (Adelman et al., 2001) showed that *Ae. aegypti* mosquitoes failed to replicate DEN-2 when dsSIN expressing a portion of the DEN-2 genome was inoculated into mosquitoes prior to DEN-2 infection. Infection of C6/36 cells by the dsSIN-DEN2as virus prevented accumulation of DEN-2 RNA as well DEN-2 envelope antigen as measured by immunofluorescence assay (IFA) (Adelman et al., 2001). The dsSIN system has been shown to effectively express a single chain antibody

against the *Plasmodium* circumsporozoite protein and protect against *Plasmodium* infection of mosquito salivary glands (de Lara Capurro et al., 2000). This system also proved to be effective in ticks; dsSIN expressing a single chain antibody fragment to louping ill virus prevented replication (Jiang et al., 1995).

In addition to anti-pathogen molecules, dsSIN can also be engineered to knock down endogenous gene expression via RNA silencing. Transgenic mosquitoes expressing luciferase in salivary glands were intrathoracically inoculated with a dsSIN that transcribed a portion of the luciferase gene in antisense orientation. Salivary glands dissected from these mosquitoes were shown to have a 90% reduction in luciferase expression (Johnson et al., 1999). dsSIN viruses have since been used to silence a number of endogenous genes in several mosquito tissues including GATA transcription factor (*Ae. aegypti*, fat body) (Attardo et al., 2003), prophenoloxidase (*Armigeres subalbatus*-hemocytes) (Shiao et al., 2001; Tamang et al., 2004), and dopadecarboxylase (Huang et al., 2005). dsSIN expression systems also infect and knock down endogenous genes in *Bombyx mori* (Foy et al., 2004; Uhlirova et al., 2003).

Innate Immunity to Viruses in Insects

Despite the large impact of arboviral diseases on human and animal populations, little research has been aimed at studying the anti-viral defense mechanisms in insects. Only a few papers have been published that describe any proteins with potential anti-viral activity in insects. Ponnuvel and others (Ponnuvel et al., 2003) described a protein in *Bombyx mori* that had antiviral lipase activity against nucleopolyhedrovirus. This antiviral activity was limited to the anterior midgut during larval stages and was not

inducible after viral infection. Treatment of virus with purified lipase prevented infection of midgut tissue. The mechanism by which the lipase acts is unknown. However, Ponnuvel and colleagues hypothesized the enzyme destroys the viral envelope and prevents viral attachment to the receptor (Ponnuvel et al., 2003). Nakazawa and others (Nakazawa et al., 2004) described another *B. mori* protein with antiviral activity against nucleopolyhedrovirus. This protein is a serine protease, is present only in the midgut and is hormonally regulated. The mechanism by which this enzyme inactivates the virus is also unknown, but it is hypothesized to inactivate virus in the alimentary canal prior to infection.

Recently, a paper by Robalino and others (Robalino et al., 2004) described the induction of an antiviral state in shrimp following non-specific dsRNA (generated from duck immunoglobulin v heavy chain) injection into shrimp. They found that injection of a dsRNA into *Litopenaeus vannamei* resulted in an overall survival increase after subsequent challenge with taura syndrome virus (TSV; *Picornaviridae*) and white spot syndrome virus (WSSV; Baculovirus). The authors suggested the increased survivorship after infection with unrelated viruses was due to an undefined, general antiviral mechanism unrelated to RNAi. Subsequent studies showed that dsRNA cognate to a virus genome injected into shrimp provided protection in a sequence-specific manner (Robalino et al., 2005).

Innate Immunity in Mosquitoes

Mosquito-borne viruses are important pathogens that impact human and animal health. The study of the innate immune system in mosquitoes could possibly lead to

insights that may aid in the development of new strategies for combating arboviral diseases. Currently, little is known about the antiviral capabilities of the mosquito immune system. Most studies concerning the mosquito innate immune system have investigated responses and interactions between vectors and bacterial and protozoal pathogens. Mosquitoes possess no adaptive immune response that has been identified; however, the innate immune response of mosquitoes is composed of both humoral and cellular components (Dimopoulos, 2003).

The first lines of the defense in the mosquito are the physical barriers, including the insect cuticle, which covers the outer surfaces of the mosquito and the peritrophic matrix that lines the midgut lumen (Levashina, 2004). If pathogens are able to breach these barriers and enter the hemolymph, there are a number of defenses that the pathogens must overcome to establish a successful infection. Humoral arms of the innate immune system of mosquitoes include prophenoloxidase, non-self recognition molecules, and coagulation factors (Levashina, 2004). The cell-mediated components of the immune system include phagocytosis and encapsulation and are active against bacterial and protozoal pathogens (Meister and Tuschl, 2004). The activation of the recognition molecules leads to the production of antimicrobial peptides (AMPs). AMPs attack the cell membranes of the bacteria, which are subsequently destroyed by cell lysis (Levashina, 2004). AMPs are produced by cells from a number of tissues, including fat body, hemocytes, the anterior midgut, and the Malphigian tubules (Lehane et al., 1997; Richman et al., 1997; Tzou et al., 2000).

The production of AMPs is the result of signaling cascades that are triggered by pathogen-associated molecular patterns (PAMPs) on the surface of bacteria interacting

with pattern recognition receptors (PRRs) on mosquito cells (Levashina, 2004). PRRs are either small soluble proteins or transmembrane proteins on the surface of hemocytes. There are three families of AMPs reported in mosquitoes: the defensins, cecropins, and gambicins. Defensins provide protection from Gram-positive bacteria. Cecropins show activity against most Gram-positive, Gram-negative bacteria, and some fungi. Gambicins have been identified only in mosquitoes and have activity against bacteria and fungi. In addition to their anti-bacterial activity, some of the peptides have also are active against *Plasmodium* (Vizioli et al., 2001).

The cellular responses of the mosquito defense system include phagocytosis and encapsulation. Most phagocytosis of small bacteria is carried out by the hemocytes. Melanotic encapsulation is a two-step process. In the first, hemocytes recognize the foreign pathogen and bind to it, forming a capsule and isolating it from the hemolymph. In the second step, the pathogen is killed by melanization, oxidation or asphyxiation (Hoffmann et al., 1999).

DNA microarrays have been utilized in the study of mosquito innate immunity. Investigators in several research labs have performed microarray analysis to determine the transcriptional changes that result in gene expression after immune challenge (Christophides et al., 2004; Dimopoulos et al., 2002; Dimopoulos et al., 1997; Dimopoulos et al., 1998; Meister et al., 2005; Richman et al., 1997; Tahar et al., 2002). These studies demonstrated a number of mosquito genes with putative immune functions are altered upon challenge with bacteria and malaria parasites.

Few studies have been undertaken to understand antiviral immunity in mosquitoes, so relatively little is known in this area. The α/β interferon pathway, which

is normally activated upon the production of dsRNA during viral replication in vertebrate cells, is not present in mosquitoes (Kumar and Carmichael, 1998). Since the publication of the *Anopheles gambiae* genome sequence, searches have been performed to look for genes that may be involved in the antiviral response. To date, no homologues have been described to genes that are important for the antiviral response in vertebrates, such as PKR and Mx (Robalino et al., 2004; Hetru et al., 2003).

Several studies have shown the production of antiviral factors released from alphavirus infected mosquito cell cultures. Riedel and Brown (Riedel and Brown, 1979) first detected this activity in cell cultures that were infected with Sindbis virus. The antiviral activity was sensitive to proteases and was unaffected by nuclease treatment. Incubation of uninfected cell culture with the factor prevented infection of these cells with Sindbis virus. The protection was determined to be virus specific, as the treated cells were still susceptible to Semliki Forest and West Nile viruses. The factor has yet to be identified, although it has been purified (Luo and Brown, 1993). Similar activities have been reported in mosquito cell cultures infected with Semliki Forest virus and Banzi virus (Hommel and Schloemer, 1985; Newton and Dalgarno, 1983). The antiviral activity was specific for each virus, and did not provide protection against other alphaviruses and flaviviruses (Hommel and Schloemer, 1985; Newton and Dalgarno, 1983).

RNA interference

RNA interference (RNAi), RNA silencing, and post-transcriptional gene silencing (PTGS) are related pathways that detect the presence of dsRNA and subsequently

degrade dsRNA and mRNA with cognate sequence. These pathways are highly conserved evolutionarily and exist in many organisms including plants, fungi, and animals (Baulcombe, 2005). RNAi and related pathways have been shown to be involved with many functions, including regulation of development, silencing and regulation of gene expression, and defense against viruses and transposable elements (Voinnet et al., 1999) (Figure 1.2).

Post-transcriptional gene silencing (PTGS) was a phenomenon first noticed in plants by Napoli *et al* (Napoli et al., 1990). Experiments were being performed to increase expression of pigment in petunias using a transgene, however, phenotype suppression rather than enhancement occurred. Antisense RNA had long been used before the discovery of RNA interference to inhibit gene function. Antisense RNA is thought to form partial double-stranded hybrids with the endogenous mRNA and inhibit its translation. Fire *et al* (Fire et al., 1991) demonstrated specific inhibition of gene expression after injection of antisense RNA into *C. elegans*. Several years later, Guo and Kemphues, demonstrated that RNA injected into *C. elegans*, regardless of the polarity of the RNA, gave the same phenotype in the worms (Guo and Kemphues, 1995). Fire *et al,* discovered that the injection of dsRNA into *C. elegans* produced a more specific and potent effect than single stranded RNA of either polarity (Fire et al., 1998). In addition, they were able to determine that only a few molecules of dsRNA per cell were required for interference to occur. The effects were seen in the progeny of infected worms as well, due to the amplification of the siRNA signal by RdRP.

The RNAi pathway is triggered by the presence of intracellular double-stranded RNA (dsRNA) and is divided into two phases: the initiator phase and the effector phase.

The initiator phase consists of the processing of long dsRNA molecules by the RNaseIII enzyme Dicer into small interfering RNAs (siRNAs) of 21-25 base pairs. These siRNAs are considered the hallmark of the RNAi response. siRNAs are duplexes of RNA with 3′ overhangs of 2 nucleotides. Each strand has a 5′ phosphate and 3′ hydroxyl (Myers et al., 2003; Provost et al., 2002a). The siRNAs are unwound and incorporated into the RNA-induced silencing complex, or RISC, with the assistance of Dicer-2 and R2D2 during the effector phase of the pathway (Tijsterman and Plasterk, 2004), and act as guide sequences to lead the RISC to the target mRNAs and initiate degradation of the mRNA (Elbashir et al., 2001). The RISC is known to include the products of the following genes: Argonaute-2 (Ago2), Vasa intronic gene (VIG), fragile X mental retardation gene (FXR), and Tudor-Staphylococcal nuclease (TSN) (Caudy et al., 2003; Caudy et al., 2002; Hammond et al., 2001; Ishizuka et al., 2002).

Hammond *et al,* demonstrated that the reduction of levels of mRNA after dsRNA treatment of *Drosophila* cells could be correlated with the production of 21-25 nt siRNAs and were the first to suggest that the small RNAs are incorporated into a nuclease complex that uses them to direct the degradation of the mRNA in a sequence-specific manner (Hammond et al., 2000).

The initiator enzymes in the RNAi pathway, termed Dicer, were first described from *Drosophila* cell and embryo lysates (Bernstein et al., 2001). Candidate genes encoding RNaseIII enzymes were expressed in cells and tested for their ability for transform long dsRNA molecules into small RNAs of ~21-25 nt. These enzymes were also shown to be inactive against single stranded RNAs. The enzyme with this activity was termed Dicer. Antibodies directed against Dicer were successful in

immunoprecipitating a protein with enzymatic activity capable of producing the small RNAs *in vitro*. *In vivo* Dicer depletion studies were also performed and resulted in decreased Dicer activity based on the detection of reduced siRNA production (Bernstein et al., 2001). The majority of information related to Dicer enzymes has been elucidated from work with *D. melanogaster,* including the specific functions of both Dicer-1 and Dicer-2 (Pham et al., 2004; Lee et al., 2004).

Dicer is a large enzyme that was first described in *D. melanogaster* and contains an RNA helicase domain, PAZ (PIWI/Argonaute/Zwille) domain, a domain of unknown function (DUF283), 2 RNase III domains and an RNA binding domain (Provost et al., 2002a). In its active form, Dicer forms a homodimer (Zhang et al., 2004). The production of siRNAs by Dicer is ATP-dependent (Nykanen et al., 2001). Dicer has been shown to be evolutionarily conserved, and is found in many organisms in addition to *Drosophila*, such as *C. elegans*, humans (Hutvagner et al., 2001), mice (Nicholson and Nicholson, 2002), yeast (*Schizosaccharomyces pombe*) (Provost et al., 2002b), trypanosomes (Ngo et al., 1998), zebrafish (Wienholds and Plasterk, 2004), the fungi *Magnaporthe oryzae* (Kadotani et al., 2004), and *Neurospora crassa* (Catalanotto et al., 2004). The number of dicer genes in different organisms varies. In some organisms including *D. melanogaster* there are two dicer genes. In *Arabidopsis* there are 4 *dicer* genes, each of which processes siRNAs from different dsRNA sources (Schauer et al., 2002). However, in some organisms, such as humans and *C. elegans* there is only one *dicer* gene.

Recently, the RNAi pathway has been shown to have two branches: the siRNA branch and the micro-RNA (miRNA) branch. In organisms that encode two Dicer

proteins, Dicer-2 is the enzyme that produces siRNAs from long dsRNAs and initiates the RNAi pathway (Pham et al., 2004). Dicer-1 is the enzyme that produces miRNAs from endogenous transcripts, which are important for regulating developmental pathways (Brigneti et al., 1998; Lee et al., 1993; Lee et al., 2004; Lin et al., 2003; Llave et al., 2002; Palatnik et al., 2003; Reinhart et al., 2000; Wienholds and Plasterk, 2004). miRNAs are produced from pre-miRNAs that are stem-loop RNA precursors encoded in genomes (Lee et al., 2004). Silencing of genes by the miRNA pathway occurs not by degradation of the mRNA, but rather by translational arrest during protein synthesis (Ambros et al., 2003). Also, unlike silencing by the siRNA pathway, miRNA silencing does not require complete base pairing between the miRNA and the target sequence to be silenced. There are two distinct RISCs that distinguish between the silencing mechanisms (Lee et al., 2004). It is not known how the distinction is made between siRNA production and miRNA production with a single Dicer enzyme in *C. elegans*. In advanced vertebrates such as humans, interferon is the innate immune response to long dsRNA, so the function of Dicer-2 may not be required in these animals.

Another family of genes that is important in RNAi is the *argonaute* gene family. In *Drosophila*, there are five *argonaute* genes (Williams and Rubin, 2002). Two of the genes, *argonaute-1* and *argonaute-2* have been implicated in RNA interference. Argonaute-2 is an important component of the RISC complex (Liu et al., 2004; Meister and Tuschl, 2004). Argonaute-2 has been termed 'Slicer' and is the only component of human RISC that is required for the degradation of mRNA molecules (Rand et al., 2004). Rand et al (2004) suggested that the endonuclease activity resides in the PIWI domain, where it is predicted the active site for the enzyme may reside. Okamura et al (Okamura

et al., 2004) showed that *Drosophila* embryo mutants lacking Ago2 were unable to perform siRNA- directed mRNA cleavage, although miRNA directed cleavage was still possible. The embryos also lacked the capacity to load siRNAs into RISC.

D. melanogaster Argonaute-1 (DmAGO1) is not involved in siRNA-directed cleavage as is *DmAgo2*. DmAGO1is believed to function downstream of the production of the siRNAs and not be a component of RISC (Williams and Rubin, 2002). These studies also demonstrated that *Drosophila* mutants lacking AGO1 are embryonic lethals, implicating AGO1 in *Drosophila* development. AGO1 is required for miRNA-directed cleavage and dispensable for siRNA directed cleavage, showing divergent roles for different Argonaute proteins (Okamura et al., 2004). This functional differentiation of the Argonaute proteins has also been noted in plants (Vaucheret et al., 2004).

Argonaute proteins contain two domains, the PAZ domain and the PIWI domain. The crystal structures of the PAZ domain of Argonaute-2 from *Drosophila* and the thermophile *Pyrococcus furiosus* were determined and shown to have structural properties similar to proteins that bind single stranded nucleic acid (Lingel et al., 2003; Lingel et al., 2004; Song et al., 2003). The PAZ domain recognizes the 3' overhangs of the siRNA duplexes (Lingel et al., 2003; Lingel et al., 2004; Song et al., 2003; Yan et al., 2003). Interestingly, this is the same region that is recognized by certain viral suppressors of RNAi (Lakatos et al., 2004). The PIWI domain is involved with protein-protein interactions between Argonaute and Dicer and may play a role in siRNA loading into RISC (Pham et al., 2004; Tahbaz et al., 2004). Recently, Sen and Blau (Sen and Blau, 2005) showed that both human Argonaute-1 and -2 proteins reside in intracellular

structures known as 'cytoplasmic bodies'. These areas of the cell are believed to be where cellular mRNA turnover is regulated (Sen and Blau, 2005).

R2D2 is a protein with two dsRNA binding domains that bridges the initiator and effector stages of the RNAi pathway. R2D2 was co-purified with Dicer-2 in a siRNA-generating extract from *Drosophila* S2 cells. The R2D2-Dicer-2 complex is required to produce siRNAs from dsRNA and also to load the newly formed siRNAs into RISC (Liu et al., 2003).

In some organisms, such as *C. elegans* and plants, an amplification of the RNAi response occurs. In these systems, the original siRNAs act as primers, and along with endogenous RNA-dependent RNA polymerase (RdRP), synthesize new dsRNA molecules in the cells. These newly produced dsRNAs are processed by Dicer and generate an additional pool of siRNAs (Sijen et al., 2001; Tijsterman et al., 2002). This phenomenon is termed "transitive RNAi" and allows for the degradation of a full mRNA even when the initial trigger sequence represents only a portion of the gene (Waterhouse et al., 2001). Transitivity of PTGS increases the pool of siRNAs outside the original target area (Sijen et al., 2001). Amplification of the siRNA signal is bi-directional along the transcript. This same phenomenon also exists in *C. elegans*, however the signal can only spread 3′-5′ along the transcript (Voinnet et al., 1998). Transitive RNAi does not occur in species such as *D. melanogaster*, where RdRP genes are absent from the genome. Gene knockout experiments in these organisms require design of dsRNA for specific disruption of gene expression.

One unique aspect of PTGS in plants is the ability of the silencing signal to spread throughout the organism. The siRNAs, complexed with other host proteins, are part of a

silencing complex that can move to other tissues in the plant using phloem tissues (Palauqui et al., 1997; Voinnet and Baulcombe, 1997). The distance that the siRNAs travel is dependent on their exact length and from what Dicer enzyme they were generated. The exact mechanism for long-distance movement has yet to be elucidated (Himber et al., 2003).

Viruses are both inducers and targets of RNAi:

Since the discovery of RNA silencing it has been hypothesized to play a role in anti-viral defense (Goldbach et al., 2003; Vance and Vaucheret, 2001). Studies involving pathogen-derived resistance (PDR) in plants were among the first to show that viruses could be targeted by RNA silencing (Lindbo et al., 1993).

RNAi is involved with the silencing of repeat sequences in the *C. elegans* genome, which involves the methylation of DNA. In addition, several RNAi genes regulate gene silencing by methylation of DNA during meiosis (Grishok, 2005). RNAi is known to silence transposons within the nematode, preventing new insertion events from occurring (Sijen and Plasterk, 2003; Tabara et al., 1999). While the genetics of *C. elegans* have been investigated in detail, including the genetics of RNAi, there had been no virus infection model of the nematode until very recently (Lu et al., 2005; Wilkins et al., 2005). Two infection models have been recently described; Flockhouse virus (FHV, *Nodaviridae*) in *C. elegans* and vesicular stomatitis virus (VSV; *Rhaboviridae*) in *C. elegans* derived cultured cells; these should allow more studies of RNAi as an antiviral defense mechanism in invertebrates.

Lindbo *et al* (Lindbo et al., 1993) were the first to describe pathogen-derived resistance (PDR) as an anti-viral defense in plants based on RNA silencing. Transgenic plants were engineered to express a portion of the tobacco etch virus (TEV) coat protein. After challenge with TEV, the transgenic plants were found to be resistant to the virus. No overt symptoms were seen, and no virus could be recovered from the leaf tissue. The mechanism for this resistance was termed PDR. PDR was shown to be virus specific, as the plants could be infected with a non-related virus. Any RNA that was introduced into plant cells that shared homology with the RNA in the transgene was degraded.

Soon after the discovery that plants had a defense mechanism against viruses, it was shown that many plant viruses encoded protein suppressors of RNA silencing (Anandalakshmi et al., 1998; Beclin et al., 1998; Brigneti et al., 1998; Kasschau and Carrington, 1998; Voinnet et al., 2000; Voinnet et al., 1999). Several suppressors have been found that interrupt the pathway at different steps, indicating that evasion of RNA silencing has evolved more than once. One protein, helper component-proteinase, or HC-PRO, from potyviruses is thought to be the most potent suppressor of RNA silencing found to date. HC-PRO functions at the step that prevents the accumulation of siRNAs by interacting with the RNase III enzyme Dicer (Llave et al., 2000; Mallory et al., 2001). HC-PRO also has the capability to reverse established silencing of a transgene, suggesting that the protein inhibits a mechanism required for the maintenance of silencing (Llave et al., 2000).

The 19 kD protein (p19) from tombusviruses acts as a suppressor of PTGS in plants in a different manner from HC-Pro. p19 does not block production of the 21-25 nt siRNAs, rather it binds to them via the 2-nt 3′ overhangs (Silhavy et al., 2002). Notably,

the protein will bind only double stranded 21nt sequences. Single-stranded RNAs of the same length are not recognized by p19. By binding the siRNAs, p19 forms a homodimer and sequesters the guide sequences that are required for RISC incorporation and targeting mRNA degradation (Ye et al., 2003).

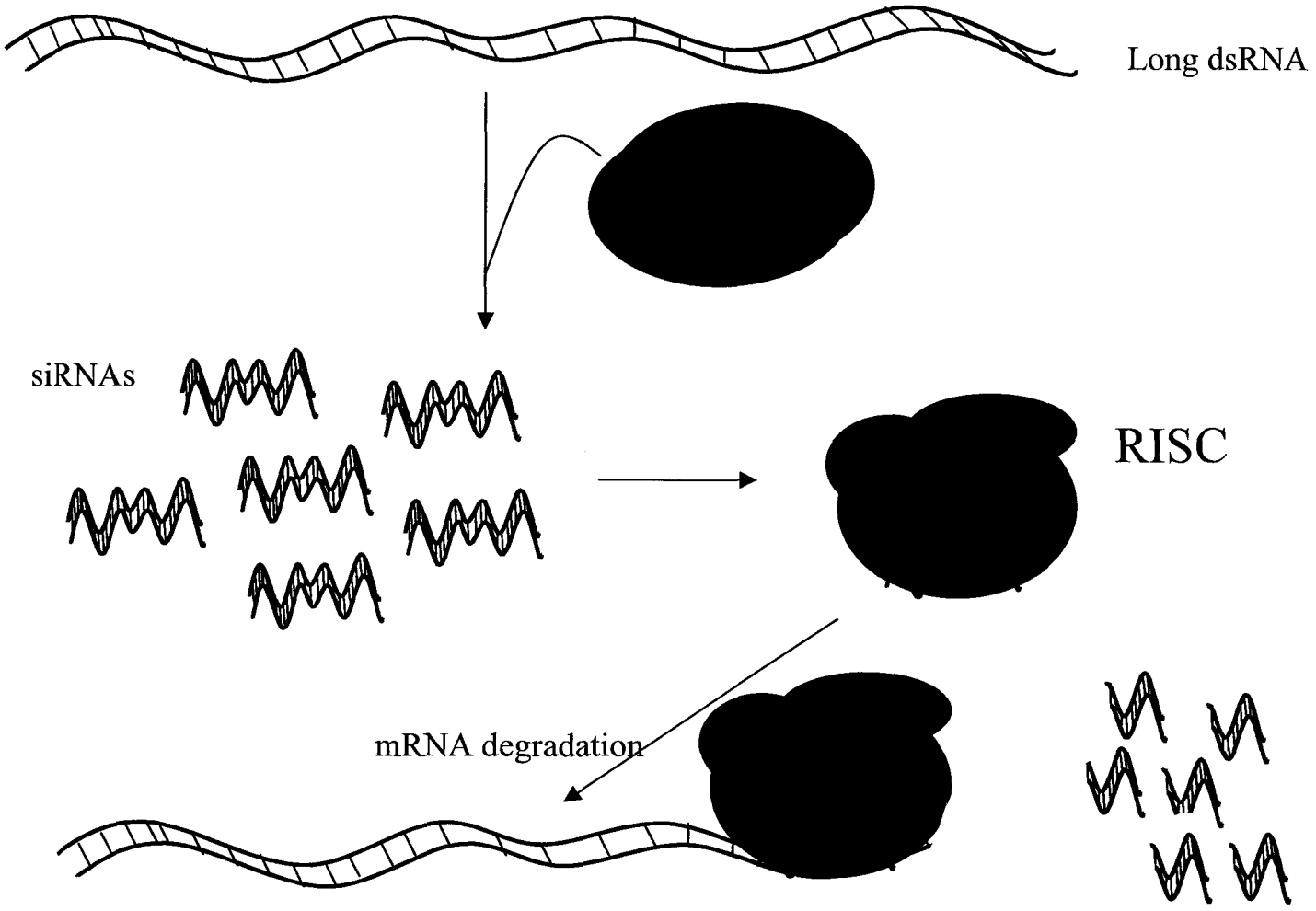
In addition to the PTGS suppressors encoded by plant viruses, the insect virus, Flockhouse virus (family *Nodaviridae*) encodes a protein, B2, which can suppress RNAi activity in both plant and *Drosophila* S2 cells (Li et al., 2002). The expression of this protein in transgenic plants expressing GFP, and transiently expressing siRNAs targeting GFP, reversed silencing of GFP. siRNAs were still detected in the tissues, indicating that the mode of action did not occur until after the production of siRNAs (Li et al., 2002). It is likely that this protein sequesters the siRNAs away from the RNAi machinery.

Targeting of replicating viruses using RNAi has prompted discussion about whether RNAi can be used as an antiviral therapy. Numerous studies in cell culture have shown that HIV replication can be halted when cells are treated with siRNAs that target the viral genome (Jacque et al., 2002; Lee et al., 2002). West Nile virus replication can be targeted and reduced in cultured cells using siRNAs targeting the virus RNA (Geiss et al., 2005). These studies showed a significant reduction in WNV RNA in cells that were pre-treated with siRNAs. However the cells that were treated subsequent to the establishment of viral replication did not have the same reduction in viral siRNA, suggesting that the RNA may be sequestered from the RNAi machinery in the cell (Geiss et al., 2005). Preliminary studies investigating RNAi as a therapy for hepatitis C virus replication have also begun. siRNAs have been used to effectively target HCV replicon

RNAs in cultured human cells as well as in a mouse model (De Francesco and Migliaccio, 2005; McCaffrey et al., 2002; Randall et al., 2003; Wang et al., 2005a).

RNAi has become an important reverse genetics tool for studying gene function. Systematic knockdown of all genes has been accomplished in both *D. melanogaster* and *C. elegans*, with great success, identifying previously unknown functions of genes in various biological pathways (Boutros et al., 2004; Kamath and Ahringer, 2003). RNA silencing has been used in many other invertebrates, including *Anopheles gambiae*. Several studies have shown that RNAi can be used to knockdown endogenous and exogenous gene expression in anopheline adult mosquitoes and cultured cells (Blandin et al., 2002; Blandin et al., 2004; Brown et al., 2003; Levashina et al., 2001), as well as *Aedes* mosquitoes (Adelman et al., 2001; Adelman et al., 2002; Attardo et al., 2003). In addition, studies have shown that it is possible to silence RNAi complex genes using RNAi (Bernstein et al., 2001; Dudley et al., 2002; Grishok et al., 2001; Hammond et al., 2001).

Fig 1.2: Simplified model of RNAi in *D. melanogaster*



RNA interference and mosquitoes:

Recently, several groups have investigated the possibility that RNA interference is a functional pathway in mosquitoes. RNA interference has been employed to down-regulate genes in both mosquito cell culture and in adult mosquitoes, which has been a useful tool in studying gene function in mosquitoes. In addition, experiments have been performed to show that arbovirus replication can be modulated in mosquitoes and mosquito cells. Several lines of evidence have now shown that the RNAi response is active in mosquitoes.

Levashina *et al* (2001) reported the first knockdown of an endogenous gene in mosquito cell culture. *Anopheles gambiae* cultured cells were transfected with a 640 bp dsRNA fragment from the TEP-1 (thioester containing protein 1) gene or a non-specific dsRNA control (GFP). The level of TEP-1 in the treated cells was determined by western blot and protein level in the dsTEP-1 cells was significantly decreased when compared to the dsGFP treated cells, indicating that transfection of dsRNA into mosquito cells can reduce the amount of protein made from a specific mRNA.

Blandin and colleagues reported the first use of RNAi to knockdown an endogenous gene in adult mosquitoes (Blandin et al., 2002). dsRNA targeting the anti-microbial peptide defensin was intrathoracically inoculated into adult female *An. gambiae* mosquitoes. When compared to a non-specific dsRNA control, those mosquitoes treated with defensin dsRNA had significantly lower levels of Defensin protein as determined by western blot, MALDI-TOF MS (Matrix assisted laser desorption/ionization- Time-of-Flight mass spectrometry) and mRNA levels as determined by RT-PCR. Mosquitoes that were treated with the dsRNA to defensin were

also more susceptible to death after bacterial challenge. dsRNA injection to disrupt gene expression has since been used to target other endogenous mosquito genes in metabolic and immune pathways (Arrighi et al., 2005; Blandin et al., 2004; Infanger et al., 2004; Moita et al., 2005; Osta et al., 2004; Sanchez-Vargas et al., 2004; Wang et al., 2005b; Zhu et al., 2003).

In addition to knocking down endogenous gene expression, RNA interference has been utilized in an anti-viral strategy to render mosquitoes resistant to arbovirus infection. There are several lines of evidence that support the idea that RNA interference can target replicating viruses both *in vitro* and *in vivo*. First, double subgenomic SINV were engineered to express a portion of the DEN-2 RNA sequence in antisense orientation. Infection of C6/36 *Ae. albopictus* cells and adult female *Ae. aegypti* mosquitoes with this virus conferred resistance to DEN-2 but not DEN-3 upon virus challenge, indicating that the interference was sequence specific (Adelman et al., 2001; Gaines et al., 1996; Olson et al., 2002).

Second, C6/36 mosquito cells were stably transformed with a plasmid that constitutively expressed an inverted repeat RNA (irRNA) derived from the prM region of the DEN-2 genome. The plasmid also contained a hygromycin resistance gene for selection (Adelman et al., 2002). This irRNA formed dsRNA when transcribed in the cell line and prevented the accumulation of DEN-2 RNA and viral antigen up to 14 days post infection with DEN-2 (Adelman et al., 2002). The cells expressing the irRNA were subjected to northern blot analysis and were shown to contain siRNAs, the hallmark of RNAi, indicating that the interference mechanism was indeed RNAi (Sanchez-Vargas et al., 2004). Interestingly, these cells were also resistant to DEN-3 upon challenge. This

may be due to a similar RNA sequence that is present in the two virus genomes that was incorporated into the irRNA sequence. The silencing phenomenon did not extend to a distantly related flavivirus, West Nile virus (EA Travanty dissertation 2005; Sanchez-Vargas et al., 2004). C6/36 mosquito cell cultures have also been treated with DEN-2 specific dsRNAs and siRNAs and subsequently challenged with DEN-2 virus. Cells that were treated with the DEN-2 specific RNA and not the control siRNA did not accumulate viral RNA or express viral antigen up to 14 dpi (Sanchez-Vargas et al., 2004, EA Travanty dissertation 2005).

The use of dsSIN systems has also been employed to silence genes in mosquitoes via an RNAi mechanism. Sanchez-Vargas *et al* (2004) used this method to knockdown expression of early trypsin in *Ae. aegypti* mosquitoes. The dsSIN virus expressed a 500 bp fragment of the early trypsin gene. Infection with this virus led to silencing of early trypsin expression in mosquito midguts as detected by quantitative RT-PCR. The silencing of the gene was accompanied by the simultaneous appearance of siRNAs with early trypsin sequence.

Other studies performed by Caplen and others (2002) showed that treatment of C6/36 cells with dsRNA to Semliki Forest virus reduced the amount of SFV RNA in the cells after virus challenge. Also, DEN-1 specific dsRNA modulated the replication kinetics of the virus in cell culture, although the effect was not as great as with SFV (Caplen et al., 2002).

RNAi based strategies are currently being employed for the generation of transgenic mosquitoes that would render them resistant to parasite infection (Travanty et al., 2004).

Summary and Goals:

RNA interference and related pathways are active in many organisms and they are quite clearly evolutionarily conserved pathways. Several lines of evidence now suggest that the pathway is active in mosquito species as well. Mosquitoes and mosquito derived cell lines can be treated with dsRNA or siRNAs to reduce expression of target genes and prevent viral infection (Adelman et al., 2002; Blandin et al., 2002; Sanchez-Vargas et al., 2004). Studies examining PTGS in plants have shown that several plant viruses encode suppressors of PTGS, suggesting that these viruses are naturally targeted by the PTGS pathway and have evolved a counterdefense strategy. This has also been shown to be true for one insect virus (Li et al., 2002). We hypothesize that RNAi can target replicating arboviruses in mosquitoes will provide insights for developing new control strategies to prevent arbovirus transmission.

The work described in the following chapters was designed to identify and characterize the mosquito genes that potentially are involved in the RNAi pathway. We are particularly interested in identifying those genes in the *dicer* and *argonaute* gene families, which have been implicated in playing important roles in RNAi in other species such as *D. melanogaster* and *C. elegans*. The complete sequencing of a number of vertebrate and invertebrate genomes has greatly aided in the search for mosquito genes. Using the information gathered from these searches, we determined which genes are potentially important in the antiviral response to alphavirus infection, testing the hypothesis that the RNAi pathway detects and modulates replicating viruses in the mosquito. The results of these studies showed the RNAi pathway is a component of the antiviral pathway in mosquitoes and provide the first evidence for an active mechanism

against virus replication in the mosquito. More research is required to determine how viruses are able to evade this response and establish a persistent infection in mosquitoes.

This work required the use of alphavirus infectious clones and recently developed reverse genetic techniques to evaluate virus replication and the role of RNA interference as an anti-viral defense in the mosquito. Using infectious clones that express reporter genes provides real-time information on how the virus infection progresses in the mosquito. Ultimately, this work will lead to a better understanding of how the mosquito innate immune system functions to protect against viral infection and provide useful information for generating new control strategies to prevent arboviral transmission.

Chapter 2: Characterization of RNA interference in an *Anopheles gambiae* cell line

The dsRNA treatment of the Sua1B cell line and luciferase assay were performed by Ngo Hoa as part of a collaboration.

Hoa, N. T., Keene, K. M., Olson, K. E., and Zheng, L. (2003). Characterization of RNA interference in an Anopheles gambiae cell line. Insect Biochem Mol Biol *33*, 949-957.

Introduction

RNA interference (RNAi) and related pathways such as post-transcriptional gene silencing (PTGS) have been observed in many eukaryotes, from single cell organisms such as *Schizosaccharomyces pombe* to more complex invertebrate animal and plant species and to higher organisms such as humans (Anderson, 2005). These pathways are evolutionarily conserved even among highly divergent species although some differences exist. RNAi is a functional pathway in the insect *Drosophila melanogaster,* which has served as one of the model organisms for determining the genetic components of RNAi as well for biochemically dissecting the pathway in detail. Early studies of RNAi in invertebrates identified two gene families as having critical roles in the pathway, the *dicer* family and the *argonaute* family.

As the threat of mosquito-borne virus diseases continues to grow worldwide, novel methods to control both the vectors and the pathogens they transmit are needed in order to reduce the burden imposed by these diseases. Recent advances in the field of transgenic insects have offered a promising new way to potentially prevent arbovirus transmission. Several studies have shown that mosquito cell lines can be engineered to be resistant to dengue virus replication (Adelman et al., 2001; Caplen et al., 2002). In addition, it is clear that endogenous genes can be silenced in mosquito cells and adult mosquitoes using a mechanism consistent with RNAi (Blandin et al., 2002; Levashina et al., 2001). Combining RNAi and transgenic technology may provide a promising way for controlling arbovirus transmission in the future.

The studies described here were conducted to test the hypothesis that the components of the RNAi pathway are present in *Anopheles gambiae* and constitute a

functional pathway in cultured mosquito cells. Gene mining studies were performed using the recently published *An. gambiae* genome sequence and known sequences from *D. melanogaster* in BLAST searches to identify putative RNAi genes from *An. gambiae*, specifically those in the *dicer* and *argonaute* gene families. Cultured *An. gambiae* mosquito cells engineered to express luciferase were tested for their ability to reduce luciferase expression via an RNAi mechanism. *An. gambiae* cell cultures were also treated with dsRNA derived from each of the *dicer* and *argonaute* genes to examine the effect on RNA mediated knockdown of transient expression of luciferase.

Materials and Methods

Identification of dicer/argonaute genes in the An. gambiae genome:

The completion and publication of the *An. gambiae* genome sequence allowed for an *in silico* search for potential dicer and argonaute genes (Holt et al., 2002). Using the *An. gambiae* database search tools at www.ncbi.nlm.gov and www.ensembl.org, BLAST searches were performed to identify homologues of the known *Drosophila dicer* (*dcr*) and *argonaute* (*ago*) genes. The Flybase accession numbers for the *Drosophila* genes are as follows: *dicer-1* (CG4792, FBgn0039016); *dicer-2* (CG6493, FBgn0034246); *ago-1* (FBgn0026611); *ago-2* (CG13452/CG7439, FBgn0046812); *ago-3* (AE003107.2); *piwi* (FBgn0004872); *aubergine* (FBgn0000146).

Preparation of templates for dsRNA production:

Total RNA was extracted from *An. gambiae* Sua1B cells (established from triturated newly hatched larvae (Dimopoulos et al., 1997) using the RNeasy Kit (Qiagen)

according to the manufacturer's instructions and reverse transcribed using an oligo d(T) primer to produce single-stranded cDNA. This cDNA was utilized as template for PCR amplification of *dicer* and *argonaute* genes. The primer pairs utilized for each gene in the study are listed in Table 2.1. Each primer pair amplified a region of the gene that was approximately 500 bp. T7 promoter sequences were incorporated into the 5' ends of primer sequences to facilitate *in vitro* transcription. In addition to the endogenous genes, dsRNA templates of approximately 500 bp were also produced for the firefly luciferase gene and the β-galactosidase gene from *E. coli*. All PCR products were cloned into pGEM-T and sequenced. Sequences were verified against database sequences by using the bl2seq tool at www.ncbi.nlm.nih.gov.

dsRNA production:

PCR products made by using the primers listed in Table 2.1 and containing 5' T7 promoter extensions were used as templates for the *in vitro* transcription reactions. After agarose gel extraction and purification of DNA products using the Gel Extraction Kit (Qiagen), the templates were transcribed using the T7 MEGAscript kit (Ambion) following the manufacturer's instructions. The transcription products were annealed by heating to 70°C briefly and cooling to room temperature. Template DNA was digested by incubating with DNaseI for 15 minutes at 37°C. The resulting dsRNA was purified using phenol-chloroform extraction and ethanol precipitation. dsRNA quality was analyzed by agarose gel electrophoresis and quantified by UV spectrometry (Amersham Pharmacia, Piscataway, NJ).

Cell culture:

The cells used in these experiments were the Sua1B cell line, which were derived from *Anopheles gambiae* larvae (Dimopoulos et al., 1997). Cells were maintained in Schneider medium (Sigma), supplemented with 10% FBS and L-glutamine and incubated at 25°C.

Treatment of cells with dsRNA to identify genes required for RNAi:

Sua1B cells were grown in 24-well plates until they reached 70% confluency. They were then transfected with 4.5 μg of dsRNA per well with sequence homology to *AgDcr1-2* or *AgAgo1-5* genes using the Effectine kit (Qiagen) following manufacturer's protocols. Twelve hours later, cells were transfected with 9 μg of dsRNA with sequence corresponding to the firefly luciferase gene. An additional 4.5 μg of the *AgDcr1-2* and *AgAgo1-5* dsRNA was also added to replace what was removed during the medium change. After 6 hours, the cells were transfected with two luciferase expressing plasmids, pCecA-luc and pRL-Act5C (Zheng and Zheng, 2002). pCecA-luc expresses firefly luciferase under the control of the Cecropin promoter, and pRL-Act5C expresses *Renilla* luciferase controlled by a *D. melanogaster* actin promoter. The cells were incubated for 24 hours at 25°C and then collected by centrifugation at 2500 x g for 10 min. Cells were rinsed once with PBS, and centrifuged again to pellet the cells. The cells were then subjected to a dual luciferase assay (Zheng and Zheng, 2002). The pRL-Act5c was included as a transfection efficiency control. Luciferase activity resulting from expression from the pCecA-luc plasmid was normalized to expression from the pRL-Act5c plasmid to minimize variation in the experiment.

RNA extractions for sequencing of AgDcr1:

RNA was isolated from a pool of 20 *A. gambiae* G3 larvae (L3) using a modified guanidinium thiocynate (GITC) extraction method (Chomczynski and Sacchi, 1987). Larvae were homogenized in 400 µl RNA lysis buffer (4M GITC, 50 mM Tris-HCl pH 8.0, 25 mM sodium citrate pH 7.5, 0.5% N-lauryl-sarcosine, 0.1M β-mercaptoethanol) using a mortar and pestle. The sample was subjected to a phenol-chloroform extraction. RNA in the aqueous layer was precipitated using an equal volume of isopropanol at –70°C overnight. The RNA was pelleted by centrifugation at 4°C for 1 hour. The RNA was resuspended in nuclease-free water and purified by EtOH precipitation overnight at –70°C. The RNA was resuspended in nuclease-free water and quantitated using UV spectrometry.

First Strand cDNA synthesis:

Total RNA isolated from *An. gambiae* larvae was used for generating first strand cDNA. Total RNA was mixed with 15 pmol of an oligo dT reverse primer and diluted to 12 µl with nuclease-free water. The RNA samples were heated to 70°C for 10 minutes and then cooled to room temperature. To each sample, 4 µl of first strand buffer (Gibco), 1 µl of 0.1 mM DTT (dithiothreitol) (Invitrogen), and 2 µl of 10 mM dNTP mix (Ameresco) and 1 µl of Superscript II-RNase H were added. The reaction was allowed to proceed for 1 hour at 42°C, followed by a 10-minute denaturing step at 75°C. Samples were incubated with RNase-1 at 37°C for 15 minutes to remove the RNA template.

PCR amplification of An. gambiae dicer-1 sequences:

Genomic sequence data from *D. melanogaster* (www.flybase.org) and *An. gambiae* (www.ensembl.org) were used to design primers to amplify the *dicer-1* gene from *An. gambiae*. Eleven sets of overlapping primers were used for PCR amplification. The primer sets and the expected product sizes are listed in Table 2.2. The positions of the primers on the gene are shown in Appendix Figure A.1. Fifty pmol of both the forward and reverse primers were added to 1 ml of PCR master mix, which contains the following: 1.5 mM $MgCl_2$, 1X PCR buffer (Promega), and 200 μM of each dNTP. One-half μl of cDNA resulting from the first-strand synthesis reaction was added to each 50 μl PCR reaction. Two units of *Taq* polymerase were added prior to the start of each reaction. For initial testing of the primer pairs, a gradient thermocycler was used to determine the optimum annealing temperature (Eppendorf). The PCR program was as follows: 30 seconds 95°C, followed by 35 cycles of 30 seconds at 95°C (denaturing), 30 seconds at the annealing temperature (ranged from 50-65°C on thermocycler), and 1 minute at 72°C (elongation). This was followed by a 10-minute extension step at 72°C, and the products were cooled to 4°C. The PCR products were analyzed by electrophoresis on 1.0% agarose gels containing ethidium bromide for DNA staining and visualization under UV light.

Bands of expected size were extracted from the gel using the Qiagen Gel Extraction Kit (Qiagen), according to manufacturer's instructions. The extracted DNA was then cloned into the pCR2.1-TOPO vector using the TOPO TA cloning kit (Invitrogen) following manufacturer's instructions and the resulting plasmids transformed into OneShot TOP10 cells. The cells were plated onto LB agar containing 50 μg/ml

ampicillin, and 40 mg/ml of X-gal and incubated overnight at 37°C. White colonies were selected for further growth overnight on new plates of the same selective agar.

After the overnight incubation, colonies were picked and added to PCR master mix (see above) containing M13 primers to select clones that contained a PCR insert. The PCR program used was as follows: 30 seconds 95°C, followed by 35 cycles of 30 seconds at 95°C, 30 seconds at 55°C, and 1 minute at 72°C. This was followed by a 10 minute extension step at 72°C, and the products were cooled to 4°C. The PCR products were analyzed by agarose gel electrophoresis on 1.0% agarose gels containing ethidium bromide for DNA staining and visualization under UV light.

Selection and purification of positive clones for sequencing:

The colonies yielding products with correct size bands determined by gel electrophoresis were chosen for DNA sequencing of the inserts. Colonies were grown overnight in 10 ml cultures of LB medium containing 50 mg/ml ampicillin in a shaking incubator at 37°C. Cells were pelleted by centrifuging at 2000 rpm. The supernatant was removed and the plasmids were prepared using the Qiagen Miniprep kit (Qiagen) following manufacturers' instructions. DNA was eluted into sterile water and the concentration was measured by UV spectrometry. Plasmid inserts were sequenced using M13 primers by Davis Sequencing (Davis, California) and at AIDL using the Big Dye v. 2.1 sequencing kit and the ABI 310 automated sequencer per manufacturers' instructions.

Table 2.1: RNAi primer pairs gene silencing in *Anopheles gambiae* cell culture

Target gene	Forward primer (5′-3′)	Reverse primer (5′-3′)	Product size
AgDcr1	cgcccaaggttcgccgcctgctcga	ctcgtacagcagcacgtaaagaat	496 bp
AgDcr1 RNase III	cgaacgccaacgatggcatc	ccaagccatgccataaagag	682 bp
AgDcr2	aggtgctgaaccaaatccac	gtacaccgagacggcaaact	479 bp
AgAgo1	tgccggccacaccgccggcaccgcc	gaccggtgtgtaggtcatgctgggc	500 bp
AgAgo2	ggttcgcgcccatacctaaacgtg	tgggacgccaggtcgaggatcttg	500 bp
AgAgo3	aacaatccgatgctggagat	cctcgtggtacgtgtcaatg	463 bp
AgAgo4	cgacttcctcaactgcatga	ttgtccttcgaatcgtgaca	408 bp
AgAgo5	ccatcaacgagctgatgaac	ccgaactgctcctggtactt	463 bp
β-galactosidase	ggtcgccagcggcacgcgcctttc	gccggtagccagcgcggatcatcgg	521 bp
Luciferase	agaactgcctgcgtgagatt	atccagatccacaaccttcg	480 bp

** All primers have a common 5' extension (5'-taatacgactcactatagg-3').
*** Locations of primer pairs are shown on the gene diagrams in Figures 2.1 and 2.2.

Table 2.2: Primers for *AgDcr-1* amplification and sequence confirmation

Primer pair name	Forward primer (5'-3')	Reverse primer (5'-3')	Estimated product size
AgDcr1 1F/1R	ATGTCGTTGTTCCACTGG	GTTTGCTCGGCTCTCGTGACCG	1600 bp
AgDcr1 1.2F/1.2R	CACTGGACGGATGGCAACA	TTGCTCGGCTCTCGTGAC	1537 bp
AgDcr1 2F/2R	GAACCTAAAGCACAAAGTCGTCG	ACGATACACTTCTAGGCAGTGG	830 bp
AgDcr1 3.2F/3R	TGCCAAGTGCCGCAATGGAGAG	GCGTTGATGCGGTACAGTACG	1280 bp
AgDcr1 4F/4R	TGTGCACGATACATCCGTTCCC	GTGGCTTACGGACAGCGGTTCG	1085 bp
AgDcr1 5F/5R	AGCAACGGTTTGAGCGGCAG	GTTTCGATGCCACCGAACGTAATG	1510 bp
AgDcr1 6.3F/6R	CTGGCGTAACGATTGCTGAGTA	TATCGGCGCTTGTGAT	950 bp
S2/Anop 1F/1R	CTGCCACCCTGCTACTACGTGCC	CATCGCCCAGRAACTCCAGMCGC	826 bp
S2/Anop 2F/1R	AACGATGGCATGAATCTGGAGCG	CATCGCCCAGRAACTCCAGMCGC	919 bp
S2/Anop 3F/3R	CTCTTTATGGCMTGGCTKGGC	GCCGCCGTACACTTGGCGATGCGATAG	1034 bp
S2/Anop 4F/4R	GTTCACCAAWGTRYTGCGYTTGCARAAG	CGCTTKGCCCGCTTYGTCTCCTCCG	551 bp

**The location of the primer pair sequences are shown in Appendix Figure A.1.

Results

Putative dcr and ago genes identified in An. gambiae: The sequences that showed the highest scores on the BLAST searches of the *An. gambiae* genome using *D. melanogaster* sequences as queries were further investigated to determine if they were indeed members of the *dcr* and *ago* gene families in *An. gambiae* (Altschul et al., 1997). The results from the BLAST searches are shown in Table 2.3. The sequences selected in BLAST searches were subjected to domain analysis using the Simple Modular Architecture Research Tool (SMART) from the Expert Protein Analysis System website (www.expasy.ch), which predicts functional domains encoded by the amino acid sequences.

The searches indicated that two full-length dicer genes were present in the *An. gambiae* genome, each having strong homology to one of the *Drosophila* dicer (*DmDcr*) genes. A clear homologue of *DmDcr1* was discovered and is now referred to as *AgDcr1* (ENSANGG00000014308). The translated sequence has 43% identity and 54% similarity to *DmDcr1* and is located in Division 13E on chromosome 2L. A second Dicer-like gene, *AgDcr2*, also was discovered and is a homologue to *DmDcr*2 (32.8% identity, 61.8% similarity). *AgDcr2* is designated as ENSANGG00000014054, and is located on chromosome 3L at position 43D. Both *Anopheles* Dicer proteins contain the same domains, and when aligned, the two amino acid sequences have 21.7% identity. As seen in *Drosophila*, the length of the Dicer-2 protein is shorter than the Dicer-1 protein.

There were also two sequences that had homology to *DmDcr1/2* but were not considered further to be *dicer* homologues because the sequences were not full length and did not encode all of the domains required to be functional Dicer-like proteins. These

sequences were found in the genomic scaffolds AAAB01008964_284 and AAAB0100846_337. AAAB01008964_284 was located on chromosome 3R and the translated sequence contained the two RNase III domains and the double-stranded RNA binding domain. AAAB0100846_337 was located on the X chromosome and encodes the helicase domains. These sequences are likely to be the result of genomic rearrangements or duplications of the sequence, and are not likely to produce functional gene products.

The searches also suggested that five *argonaute*-like genes exist in *An. gambiae*. The nucleotide sequences were translated and aligned with each of the 5 *Drosophila* Argonaute protein sequences using the SIM tool at www.expasy.ch. The results of the alignments as well as other *Anopheles* Argonaute characteristics are shown in Table 2.4.

Table 2.3: Blast searches for *Anopheles gambiae dicer* and *argonaute* genes using *Drosophila* genes as queries

	TBLASTX Highest homology		
Drosophila gene (accession number)	ENSEMBL (*Anopheles gambiae* database)	NCBI Scaffold (genome sequence)	e-value
DmDcr1 (CG4792, FBgn0039016)	ENSANGG00000014308 AY239359	AAAB01008859_134	0.0
DmDcr2 (CG6943, FBgn0034246)	ENSANGG00000014054	AAAB01008986_87	1e-197
DmAgo1 (FBgn0026611)	ENSANGG00000006700	AAAB01008986	1.1e-296
DmAgo2 (CG7439, FBgn0046812)	ENSANGG00000021748	AAAB01008834_94	1.2e-125
DmAgo3 (AE003107)	ENSANGG0000004858	AAAB01008984_66	8.5e-103
piwi (FBgn0004872)	ENSANGG0000008598	AAAB01008839_32	1.8e-179
aubergine (FBgn0000146)	ENSANGG0000006265	AAAB01008816_102	1.4e-151

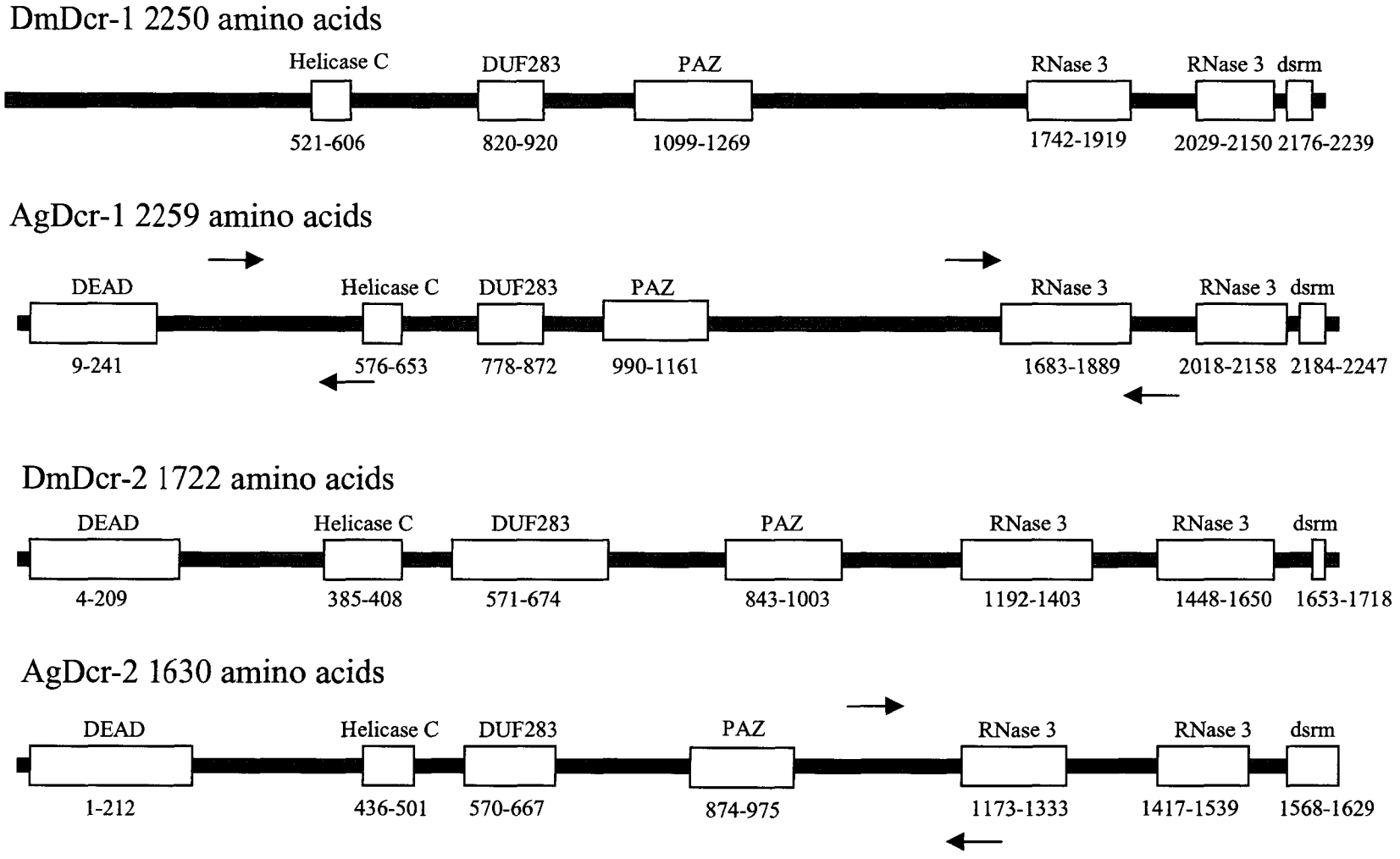


Figure 2.1: Comparison of translated dicer gene products from *D. melanogaster* and *An. gambiae*. The *AgDcr* genes are shown in schematics with translated products and the domains encoded by the genes (SMART, www.expasy.ch). The *DmDcr* homologues are shown for comparison. The approximate regions of genes amplified for dsRNA templates (from Table 2.1) are shown by inverted arrows.

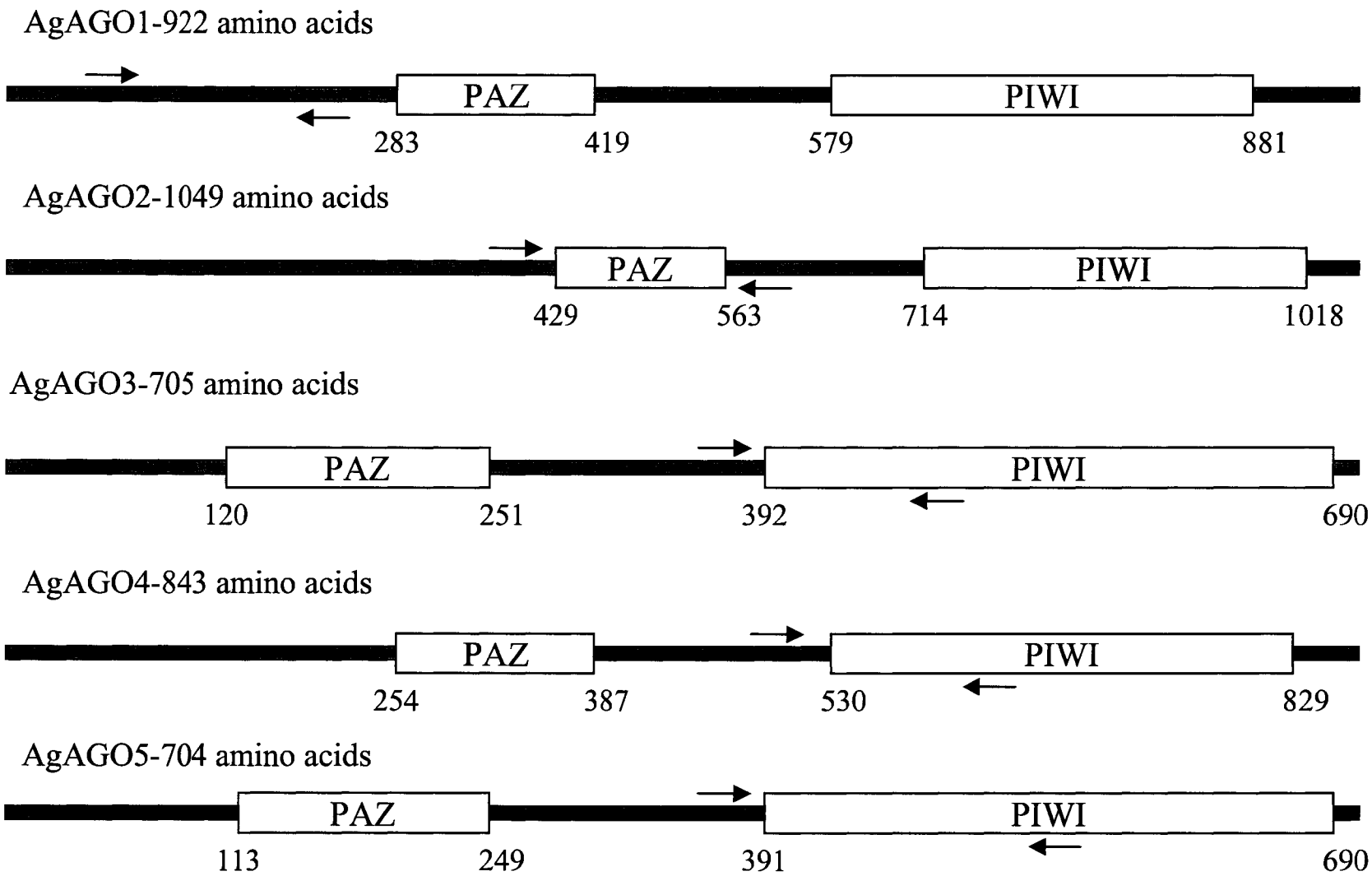


Figure 2.2: Translated products from *AgAgo* genes. The gene products of *AgAgo1-5* are shown. The approximate regions of genes amplified for dsRNA templates (from Table 2.1) are shown by inverted arrows.

Table 2.4 -*Argonaute* BLAST results and protein alignments. The BLAST results from the *An. gambiae* genome using *DmAgo* genes as queries are shown, along with their www.expasy.ch accession number, and homology with their *DmAgo* homologues.

Anopheles Argonaute	Accession number	AA length	Chromosome locations	*Drosophila* homologue	% identity translated sequence
AgAgo1	ENSANGG00000006700	992	3L/43D	DmAgo1	90.9%
AgAgo2	ENSANGG00000021748	1049	3L/43B	DmAgo2	39.7%
AgAgo3	ENSANGG00000004858	705	3R/34B	DmAgo3	38.0%
AgAgo4	ENSANGG00000008598	843	3R/42B	piwi	48.1%
AgAgo5	ENSANGG00000006265	704	3L/42B	aub	42.6%

Sequence confirmation of AgDcr1 from An. gambiae larvae:

Clones were sequenced in triplicate and aligned with the *An. gambiae* Dicer-1-like sequence found in the Ensembl database using Clustal analysis. The Clustal tools used were online at www.expasy.ch. The aligned coding sequence for *AgDcr1* along with the translated amino acid sequence are shown in Appendix Figure A.1. The translated *AgDcr1* sequence has a 43.7 % amino acid identity with *DmDcr1*. The aligned protein sequences of *AgDcr1* and *DmDcr1* are shown in Appendix Figure A.2. *AgDcr1* is 6780 nucleotides in length and codes for a protein 2259 amino acids in length. *AgDcr1* shares many of the protein domains with its homologue, *DmDcr1*. *AgDcr1* encodes two N-terminal helicases, DEAD box helicase and helicase C; a domain of unknown function 283 (DUF 283), PAZ domain, two RNase III domains and a C-terminal dsRNA binding domain (Figure 2.1).

Sequencing results indicated a region of 129 bp, beginning at nt 5458, that was shown on the Ensembl website as genomic (intron) sequence that we determined to be expressed in mRNA in the late larval stage. The region was amplified by two sets of primers, AgDcr5F/5R, and S2/Anop 1F/1R. The full-length consensus sequence resulting from these experiments was submitted to GenBank at www.ncbi.nlm.nih.gov and was assigned the accession number AY239359.

dicer-2, argonaute-2 and argonaute-3 are involved with RNAi in An. gambiae cells:

Using the primers listed in Table 2.1, ~ 500 bp regions from all of the putative dicer and argonaute genes were amplified by RT-PCR from RNA extracted from the Sua1B cell line. These products were sequenced to verify their identities and then used as templates for the *in vitro* transcription reactions to generate dsRNA. In order to determine which of the genes are involved with RNA interference, we hypothesized that knocking down the genes involved in the pathway would allow for a recovery of luciferase expression in a transformed cell line that had luciferase expression silenced.

Sua1B cells were engineered to constitutively express a luciferase reporter construct driven by the Cecropin-A (Cec-A) promoter (Zheng and Zheng, 2002). The effectiveness of RNAi in the cells was tested by transfecting the cells with luciferase dsRNA or a non-specific *β-galactosidase* control dsRNA, followed by transfection of the CecA promoter luciferase-driven construct. Treatment of this cell line with 5 μg of dsRNA derived from the luciferase sequence resulted in down-regulation of the expression of luciferase up to 4000-fold (Hoa NT: data not shown).

In order to establish a baseline of luciferase recovery, cells were pre-treated with a non-specific dsRNA, dsβgal, 12 hours prior to transfecting cells with dsLuc. Pre-treatment of the cells with dsβgal resulted in a 10-15 fold recovery of luciferase expression. Pre-treatment of the cells with dsRNA with homologous sequence to some genes in the *dicer* and *argonaute* gene families resulted in a recovery of luciferase expression. Pre-treatment of the cells with dsRNA derived from *AgDcr2*, *AgAgo2*, and *AgAgo3* resulted in greater than 80% recovery of luciferase expression. Pre-treatment of cells with dsRNA from *AgDcr1*, *AgAgo1*, *AgAgo4*, and *AgAgo5* did not result in luciferase expression recovery that was equal to or greater than the dsβgal control. A second region of *AgDcr1* was also targeted for RNAi knockdown, and that dsRNA also did not result in recovery of luciferase expression (Fig 2.3).

Figure 2.3

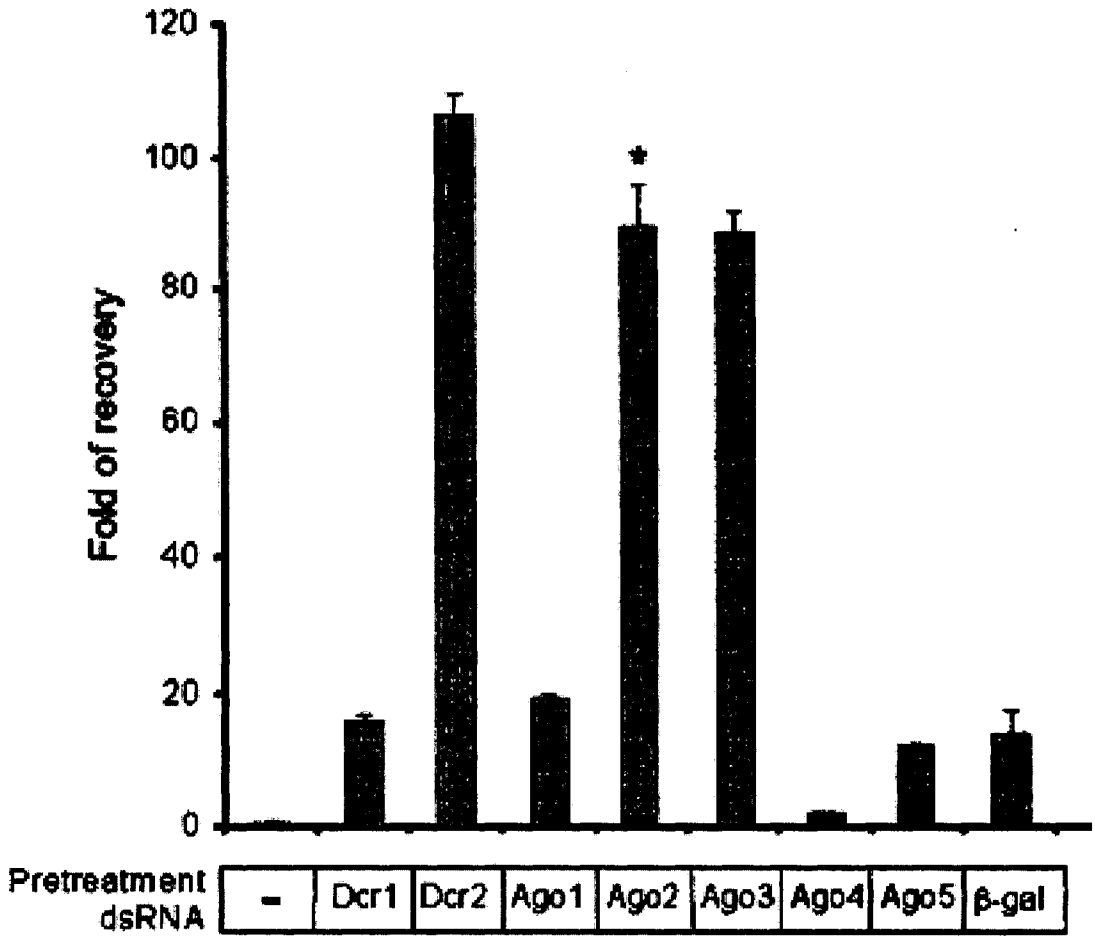


Figure 2.3: *Dcr2*, *Ago2*, and *Ago3* are involved in RNAi in an *An. gambiae* cell line. Luciferase activities were expressed as fold of recovery after pretreatment with dsRNA from each putative RNAi machinery component. dsLuc knocked-down luciferase expression approximately 4000-fold (data not shown).

Discussion

Previous studies have led us to the hypothesis that RNAi is a functional phenomenon in mosquitoes and mosquito cells (Blandin et al., 2002; Blandin et al., 2004; Levashina et al., 2001). These studies confirm that the RNAi pathway is present and functional in *An. gambiae* mosquito cells. Treatment of luciferase-expressing *An. gambiae* Sua1B cells with dsRNA cognate to the luciferase gene resulted in approximately a 4000-fold decrease in luciferase expression when compared to cells treated with a non-specific (βgal) dsRNA (dsβgal) (Hoa et al, 2002). As discussed in Chapter 1, studies concurrent with this work also demonstrated that RNA-dependent RNA polymerase (RdRP) is not involved with RNAi in *An. gambiae*, as has been reported for other organisms, including plants and *C. elegans* (Hoa et al., 2003). This finding indicates that the RNAi pathway in *An. gambiae* may be more closely related to the pathway in *D. melanogaster* than *C. elegans*. This provides more support for use of *Drosophila* as a model system for mosquitoes.

As in *D. melanogaster*, two *dicer* genes and 5 *argonaute* genes are present in the *An. gambiae* genome. The conservation of the number of genes in these families in these two organisms may also allow use of findings in *Drosophila* to predict the functions of the *An. gambiae* homologues. Although the overall mechanism of RNAi is conserved, the roles of the specific genes are known to differ between organisms. The number of dicer-encoding genes is different between organisms, ranging from one in humans and *C. elegans* to four in some plant species. In plants, it is now known that the specificity of Dicer function is dependent on the source of the dsRNA (Schauer et al., 2002). In organisms encoding two *dcr* genes, such as *D. melanogaster*, the functions of each of the

Dicer proteins have been elucidated. Dicer-1 processes endogenous dsRNA for the miRNA pathway, while Dicer-2 processes exogenous long dsRNA for the siRNA pathway (Lee et al., 2004; Pham et al., 2004).

Sequencing of *AgDcr1* cDNA confirms the presence of the gene in *An. gambiae* and the expression of a full length mRNA. This gene was sequenced because at the time of the study the gene product of *AgDcr1* was believed to be responsible for the processing of long dsRNAs into siRNAs. However, we now understand the *Drosophila* homologue of this gene to be involved primarily with the miRNA processing branch of the pathway (Lee et al., 2004). Further experiments were originally planned to express the full-length protein from this gene for use in *in vitro* RNAi studies. *AgDcr1* closely resembles the *DmDcr1,* in length and structure. The similarities between this gene and its *Drosophila* homolog, along with the knowledge that the *AgDcr2* and *DmDcr2* have high homology, indicate that the functions of the *An. gambiae dicer* genes may be very similar to those in *Drosophila.* The results from the dicer knockdown-luciferase assay experiments support this hypothesis. While the transfection of dsAgDcr2 had a dramatic effect on the recovery of luciferase expression, the transfection of dsRNA to two different regions of *AgDcr1* had no effect on luciferase expression, indicating that *AgDcr2* and not *AgDcr1* is involved with this branch of the RNAi pathway.

A recent study by Lee *et al* stated that the *Drosophila* Dicer-2 protein lacks a PAZ domain (Lee et al., 2004). Our analysis of the amino acid sequences using several sequence analysis tools and protein domain prediction programs (www.expasy.ch) indicates that both the *Drosophila* and *Anopheles* Dicer-2 proteins encode a PAZ domain, as is depicted in Figures 2.1 and 2.2. Several *Drosophila* genome databases, including

Flybase (www.flybase.org) and Ensembl (www.ensembl.org) also show the PAZ domain encoded by *Drosophila* Dicer-2. Based on the resolved crystal structure of AGO2, the PAZ domain was shown to recognize and bind the 3' OH overhangs of both siRNA duplexes and single-stranded siRNAs (Lee et al., 2004; Lingel et al., 2003; Pham et al., 2004; Yan et al., 2003).

Reduction of luciferase expression by approximately 4000-fold was achieved by transfection of the transformed luciferase-expressing cells with dsLuc. Pre-treatment of the same cells with the non-specific dsβgal twelve hours prior to dsLuc treatment resulted in about a 10-fold recovery of luciferase expression. The sequence of dsβgal did not share any homology with the genes that were tested in the experiment, and so the result was interpreted that the level of recovery was due to competition for the RNAi machinery in the cells. This 10-fold recovery of expression was set as a background limit and any experimental result not exceeding this threshold was regarded as non-significant. Reasons for this may include a non-specific interaction, or the dsRNA may be competing for the RNAi machinery. Transfection with dsRNA generated from sequences from *AgDcr2*, *AgAgo2*, and *AgAgo3* resulted in recovery of luciferase expression above 80-fold, indicating that these genes may play an important role in RNA interference in the cell line.

Interestingly, dsRNA generated from sequences from *AgDcr1* and *AgAgo1* resulted in a level of recovery that was similar to that of dsβgal and therefore these genes were first regarded as not being important for RNAi in the cells. Early studies in *D. melanogaster* implicated Dicer-1 as the important enzyme in the RNAi pathway, cleaving long dsRNA molecules into siRNAs (Bernstein, 2001). However, more recent work has

shown that it is Dicer-2, and not Dicer-1, is the enzyme important for generating the siRNAs that are ultimately used for mRNA degradation (Pham, 2004). Our *AgDcr1/2* results from this experiment support this finding.

DmAgo1 is also known to be important in the RNAi pathway in *D. melanogaster,* and surprisingly knockdown of *AgAgo1* did not result in significant luciferase recovery in mosquito cell culture. *DmAgo1* is required for the targeted degradation of mRNA during RNA silencing, but is not required for the production of siRNAs in *Drosophila* embryos (Williams and Rubin, 2002). Interestingly, *AgAgo3* was found to be involved in RNAi in *A. gambiae*. Although the gene sequence for Ago3 has been identified in *Drosophila*, no function has been assigned. It is possible that *AgAgo3* may play a similar role to *DmAgo1* in mosquito cells, however this has yet to be determined.

The results from these experiments show that RNAi is a functional pathway in mosquito cells. The experiments in the following chapter were designed to determine if the RNAi pathway could be used to target arbovirus replication in the mosquito and whether the pathway naturally modulates arboviral infection in the mosquito vector.

Chapter 3

RNA Interference is a natural antiviral response to o'nyong-nyong virus (*Alphavirus*; *Togaviridae*) infection of *Anopheles gambiae*

Some of the work from this chapter has been published:

Keene, K. M., Foy, B. D., Sanchez-Vargas, I., Beaty, B. J., Blair, C. D., and Olson, K. E. (2004). RNA Interference acts as a natural antiviral response to O'nyong-nyong virus (*Alphavirus*; *Togaviridae*) infection of *Anopheles gambiae*. PNAS *101*, 17240-12245.

Introduction

Arthropod-borne viruses (arboviruses) continue to impact human and animal health worldwide. Mosquito-borne arboviruses replicate and disseminate within susceptible vector tissues prior to transmission to vertebrate hosts. We know little about how mosquitoes cope with arbovirus infections. However, we expect that mosquitoes have defense mechanisms to counteract or modulate arbovirus infections that could impair host functions. RNA interference (RNAi) may be an important pathway that mosquitoes use to modulate arbovirus replication (Sanchez-Vargas et al., 2004). RNAi is a potent intracellular response activated by dsRNA and results in a reduced steady-state level of specific RNA molecules with sequence similarity to the dsRNA (Cogoni and Macino, 1997; Vaucheret et al., 1998). The mechanism of RNAi has been studied in some detail in *Drosophila melanogaster*. In the fruitfly, the RNase III enzyme Dicer is responsible for digesting dsRNA into 21-23 bp small interfering RNAs (siRNAs). The siRNAs are then unwound into single-stranded siRNAs in an ATP-dependent step and incorporated into an enzyme complex termed the RNA-induced silencing complex (RISC). The single-stranded siRNAs guide RISC to the target mRNA and the complex cleaves the message or inhibits its translation (Schwarz et al., 2002). An essential component of RISC is Argonaute2 (AGO2), a member of the Argonaute family of proteins. AGO2 has been co-immunoprecipitated with Dicer from *Drosophila* S2 cells (Hammond et al., 2001). A proposed interaction between AGO2 and Dicer2 facilitates the incorporation of siRNAs into RISC, which can target cognate mRNAs for destruction (Hammond et al., 2001; Pham et al., 2004).

There are several reasons to suspect that RNAi is an antagonist of arbovirus replication in mosquitoes. First, the RNAi-like pathway in plants, post-transcriptional gene silencing (PTGS), is a potent antiviral response triggered by dsRNA generated during replication by most plant viruses (Kasschau and Carrington, 1998). Second, many plant RNA viruses encode suppressors of PTGS, supporting the observations that PTGS acts as a viral defense system (Roth et al., 2004). For example, the tombusvirus p19 protein suppresses PTGS in plants by binding siRNAs produced after virus infection (Silhavy et al., 2002). Third, Li and colleagues demonstrated that the *B2* gene of the insect nodavirus, flock house virus (FHV), can suppress PTGS activity in plants and RNAi in *Drosophila* S2 cells, emphasizing that an evolutionarily conserved RNAi pathway plays a natural antiviral role (Li et al., 2002). This research group also demonstrated that vaccinia and human influenza A, B, and C viruses each encode a protein that suppresses RNAi in mammalian and insect cells (Li et al., 2004).

A number of studies have now shown that RNAi is active in anopheline and culicine mosquitoes (Adelman et al., 2001; Adelman et al., 2002; Attardo et al., 2002; Blandin et al., 2002; Blandin et al., 2004; Brown et al., 2003; Levashina et al., 2001; Osta et al., 2004; Sanchez-Vargas et al., 2004). Recently, Hoa and colleagues demonstrated that expression of Dicer2, AGO2 and AGO3 proteins are essential for RNAi activity in *A. gambiae* Sua1B cells (Hoa et al., 2003). Transient expression of luciferase in the Sua1B cell line was silenced ~4000-fold following transfection with dsRNA derived from the luciferase reporter gene. Pretreatment of the cells with dsRNA derived from cDNA sequence of *An. gambiae dcr2 (AgDcr2), AgAgo2,* or *AgAgo3* consistently yielded recovery of luciferase activity, demonstrating that RNAi can be used to silence genes

involved in its own pathway and implying that these genes have important roles in RNAi within mosquito cells. The recovery phenomenon was not observed when the cells were treated with dsRNA derived from *β-gal* or *AgAgo1, AgAgo4* and *AgAgo5* genes (Hoa et al., 2003).

A logical progression from these studies is to determine whether RNAi can act as an antagonist to arbovirus replication in *An. gambiae*. However, anopheline mosquitoes transmit few arboviruses. An exception is o'nyong-nyong virus (*Togaviridae*; *Alphavirus*; ONNV), the etiological agent of a large outbreak of human disease in East Africa from 1959-1962 (Corbet et al., 1961; Williams and Woodall, 1961; Williams and Woodall, 1965) and again in 1996 (Lanciotti et al., 1998). In epidemics, *Anopheles spp.* are almost certainly the vectors. Wild-caught *Anopheles funestus* and *An. gambiae* mosquitoes held alive for up to 20 days after capture have been found to be infected with ONNV (Corbet et al., 1961).

ONNV, like other alphaviruses, is a small, enveloped, RNA virus that replicates exclusively in the cytoplasm of infected cells (Strauss and Strauss, 1994). The genome is a positive-sense, single-stranded, non-segmented RNA of about 11.7 kb (Strauss et al., 1984; Strauss and Strauss, 1994). The 5′ two-thirds of the alphavirus genome is translated to form polyproteins that are posttranslationally processed into nsP1-nsP4 to form replicase complexes that synthesize positive or negative RNAs (Sawicki and Sawicki, 1994; Strauss and Strauss, 1994). Replication of alphavirus RNA occurs at intracellular membranes in infected cells and leads to formation of double stranded RNA (dsRNA) forms called replicative intermediates (Barton et al., 1991). The subgenomic (26S) mRNA, colinear with the 3′ one-third of the genome, is translated into a structural

polyprotein from which capsid, the envelope glycoproteins (E1 and E2), and two smaller polypeptides (E3 and 6K) are produced as cleavage products during glycoprotein processing. Alphavirus RNA genomes are readily manipulated as full-length cDNA infectious clones and recombinant double subgenomic ONNV has been generated that expresses eGFP as a marker of infection (Brault et al., 2004).

The experiments described here were performed to test the hypothesis that the RNAi pathway can naturally target a replicating arbovirus in mosquitoes. In this study, we describe the replication of a recombinant ONNV-eGFP following intrathoracic injection into *An. gambiae*. When this virus was co-injected with dsRNA derived from the viral genome, virus replication in the mosquito was significantly compromised. In contrast, mosquitoes co-injected with ONN-eGFP and dsRNA derived from *AgAgo2* were more permissive to virus replication and dissemination. These experiments demonstrate that RNAi acts as an antiviral response to ONNV-eGFP infections in *An. gambiae*.

Materials and Methods

Generation of ONNV-eGFP. The infectious cDNA clone, pONN.30a, derived from the Uganda SG-650 strain of ONNV was provided by Dr. Ann Powers (DVBID-CDC). The infectious clone was further modified by insertion of a second subgenomic promoter 5' to the structural proteins that drove expression of eGFP (p5'dsONNic-Foy/eGFP, but for simplicity is termed ONNV-eGFP in this paper), by methods described earlier (Foy et al., 2004). It was linearized to form a template for *in vitro* transcription of infectious RNA using T7 polymerase. RNA from the transcription reaction was electroporated into BHK-

21 cells. Thirty six hours later, the supernatant containing virus was collected and the virus titer was determined by plaque assay. C6/36 cells (*Aedes albopictus*) were then infected at an MOI of 0.01 and 60 hours later, supernatant containing the amplified virus was collected and titrated. The virus stock contained 2.1 x 10^7 PFU/ml.

Mosquitoes and intrathoracic inoculation of virus and/or dsRNAs. An. gambiae (G3 strain) larvae were reared on an artificial diet of ground fish food at 30°C with a 14h light/10h dark photoperiod. Adult female *An. gambiae* (2-4 days post emergence) were anesthetized by placing on ice and then intrathoracically injected with 0.5 µl of inoculum. For viral characterization studies, viral stock was diluted with DMEM to 1x10^7 PFU/ml prior to injection, thus each mosquito received approximately 5 x 10^3 PFU of virus. For co-injections of virus and dsRNA, dsRNA was diluted to a concentration of 1 µg/µl in DMEM and mixed 1:1 (vol:vol) with the undiluted stock (2.1 x 10^7 PFU/ml). Thus, each mosquito was inoculated with approximately 5.3 x 10^3 PFU of virus and 250 ng of dsRNA.

Characterization of ONNV infection in An. gambiae mosquitoes. Following inoculation, *An. gambiae* adult female mosquitoes were assayed at 3, 6, and 9 dpi. Mosquitoes were killed by brief submersion in 70% ethanol, washed in saline, and were initially observed and photographed for eGFP expression under blue light. The heads were then removed, squashed on glass slides, fixed in acetone, and assayed for viral antigen by IFA using a monoclonal antibody (30.11a) developed against Sindbis virus E2 protein but which cross-reacts with ONNV E2 protein. The primary antibody (30.11a) (1:200) was

incubated with the tissues at 37°C in a humid chamber for one hour. The slides were washed three times with sterile PBS, and then incubated with biotinylated sheep anti-mouse antibody (1:200) and a 1% Evan's blue counterstain at 37°C for 1 hour. The slides were washed again with PBS, and then incubated with fluoroscein-streptavidin (1:200) for 15 minutes at 37°C. The slides were washed again with PBS, and covered with glycerol and glass coverslips. The stained tissues were then viewed using a fluroescent microscope for analysis. The thorax and abdomen of each mosquito were individually frozen at –70°C and later titrated by plaque assay (see below).

Generation of dsRNA. cDNAs were generated from total mosquito RNA (extracted from 4th instar larvae using the Qiagen RNeasy kit) by reverse transcriptase-polymerase chain reaction (RT-PCR) using an oligo d(T) primer. Oligonucleotide primers were designed to amplify approximately 500 bp regions of *An. gambiae Dcr1-2* and *Ago1-5* cDNA. Primer sequences incorporated T7 promoter sequences at the 5' ends for facilitating dsRNA production (Table 1). Control ~500 bp cDNA templates were generated by PCR using primers specific for portions of the *E. coli β-galactosidase* cDNA clone and for the *nsP3* gene from ONNV-eGFP. To generate dsRNA, PCR products were purified by gel electrophoresis and extraction and used as templates for *in vitro* transcription using the MegaScript kit (Ambion, Austin, TX). One μg of DNA template was mixed with 2 μl of each of the following: 10X T7 polymerase Reaction Buffer, 75 mM NTP solutions (ATP, CTP, GTP, UTP). Nuclease-free water was added to bring the total reaction volume to 20 μl, and then 2 μl of the T7 enzyme mix was added. The reaction was mixed by vortexing, and briefly spun down to collect the liquid at the bottom of the tube. The reaction was

incubated at 37°C for at least 4 hours. Following the transcription reaction, 1 µl of DNase 1 was added and the reactions incubated at 37°C for 15 minutes to degrade the template DNA. The dsRNAs were annealed by heating the reactions at 70°C for 5 minutes and then allowing to cool slowly to room temperature. dsRNAs were purified by phenol-chloroform extraction and ethanol precipitation according to manufacturer's instructions. The quality of dsRNA was checked by agarose gel electrophoresis and was quantified using a spectrophotometer.

Determination of virus titer. Mosquitoes were triturated in 270 µl of Dulbecco's Modification of Eagle's Medium (DMEM) using a mortar and pestle, and large particulates were pelleted by centrifugation. The supernatant was passed through a 0.22 µm syringe-tip filter and then titrated by standard plaque assay using Vero cells (Myles et al., 2003). Plaques were counted and the data were expressed as $\log_{10}$ PFU/mL. Differences in viral titers were first analyzed by ANOVA, and the titers in the treatment groups were significantly different if $p < 0.01$. Pairwise t-tests were then performed. Highly significant differences ($p < 0.01$) between treatment groups and controls are marked on graphs with two asterisks, significant differences ($p < 0.05$) are marked with a single asterisk.

Northern blot analysis. Total RNA was isolated 2-3 dpi from 50-60 mosquitoes and mRNA was purified from total RNA using the MicroPoly(A) Purist Kit (Ambion). Ten µg of mRNA was separated on a 1% agarose-formaldehyde gel and transferred to a BrightStar Plus nylon membrane (Ambion). The blots were hybridized with ^{32}P-labeled

probe complementary to the ONNV *E2* gene (Fig. 1B) or *AgAgo2* (Fig. 4). The blots were then washed at 68°C and radioactivity was detected using a Storm Phosphorimager (Molecular Dynamics Inc.; Amersham Pharmacia Biotech Ltd.).

Table 3.1: Primer pairs for production of dsRNA from *An. gambiae* cDNA

Target gene	Forward Primer (5'-3')	Reverse Primer (5'-3')
Dcr1	cgaacgccaacgatggcatc	ccaagccatgccataaagag
Dcr2	gaggtgctgaaccaaatccac	ggtacaccgagacggcaaact
Ago1	gcaggtgtccctgttcaacct	ggtttggccgttctctagctg
Ago2	gcatgagcacgctcaacaac	gttcgagtcgtcgtacagca
Ago3	gtgtggcattgacacgtacc	gctcagctgctgcagaatgtc
Ago4	gcgacttcctcaactgcatga	gtgttgagcggcagatagttg
Ago5	gacaagtcgctctcgtacggt	gtctcgtcgaagatcacgttg
ONNV nsp3	catgtggccaaaacaaactg	cgaatttgcgtacattggtg
β-galactosidase	ggtcgccagcggcaccgcgccttc	gccggtagccagcgcggatcatcgg

**All primers had a common 5' extension (5'-taatacgactcactatagg-3').

Quantitative Reverse Transcription – Polymerase Chain Reaction (QRT-PCR):

Following dsRNA inoculation, adult female *An. gambiae* were assayed at 3 and 6 dpi for mRNA levels. Each time point consisted of 3 experimental replicates. Total RNA was isolated from a pool of 5 mosquitoes using the GITC-phenol-chloroform extraction method as previously described (Chapter 2, Material and Methods). Total RNA was quantified by absorbance at 260 nm and 500 ng of RNA was used per reaction. Primers for *AgDcr1* and *AgDcr2* QRT-PCR were designed to amplify 150-200 bp regions of the genes outside the dsRNA target region, as to amplify only mRNA. Primers were also designed to amplify a 460 bp region of the S7 ribosomal protein gene as a normalization

control (Richman et al., 1996). The primer pairs used for amplification are listed in Table 3.2.

QRT-PCR was performed in an Opticon Monitor 2 thermocycler (MJ Research) using the Quantitect SYBR Green RT-PCR kit (Qiagen). The comparative threshold cycle (Ct) method was used to determine differences in quantities of transcripts present in the different treatment groups (Applied Biosystems). Statistical analyses were performed using GraphPad software. One sample T-tests were used to determine statistical significance between the groups.

Table 3.2: Primer pairs for QRT-PCR in *An. gambiae*

Target gene	Forward Primer (5'-3')	Reverse Primer (5'-3')
AgDcr1	CGAACCACCGCTCCAATCA	ACAACGGCGTCCCGAAAT
AgDcr2	CGACGAGATTGCGGACTTTACG	ACACCACCTCCTCCTCCACGAA
S7	GGCGATCATCATCTACGTGC	GTAGCTGCTGCAAACTTCGG

Results

Construction and in vivo characterization of intrathoracically injected ONNV-eGFP.

The pONNV.30a infectious cDNA clone was engineered to contain a duplicated subgenomic promoter, 3' to the original subgenomic promoter (Fig. 1A). The eGFP gene was inserted immediately downstream of the 5' subgenomic promoter into a multiple cloning site, generating ONNV-eGFP plasmid. Recombinant virus was produced from this plasmid and injected into *An. gambiae* mosquitoes. The three RNA species to be

formed during virus replication were not detectable on northern blots of mosquito RNA after 48 hours but were readily apparent after 72 hours (Fig. 1B), indicating a relatively slow rate of viral replication in *A. gambiae* tissue. Viral titers in the mosquito 3 days post-injection (3dpi) were lower than the input titer by >1.5 log PFU (Table 3.3), but then slowly increased over a 9 day incubation period. Virus was detected in the head tissues of individual mosquitoes both by eGFP expression and by immunofluorescence analysis of the E2 glycoprotein (Table 3.3). The data indicate that these methods of detection are equivalent and that eGFP serves as a readily detectable marker of infection during this time period. eGFP expression in injected mosquitoes at 9 dpi revealed temporal and spatial infection patterns of the virus (Fig. 1C). Virus typically infected nervous, muscle, and fat body tissues. When ONNV-eGFP dissemination to the head occurred, ommatidia and cells in the maxillary palps, antennae, and in the mouthparts expressed eGFP, but the marker was not apparent in salivary glands (data not shown). In the abdomen, eGFP was most often sporadically associated with circular muscle fibers wrapping the alimentary canal and fat body.

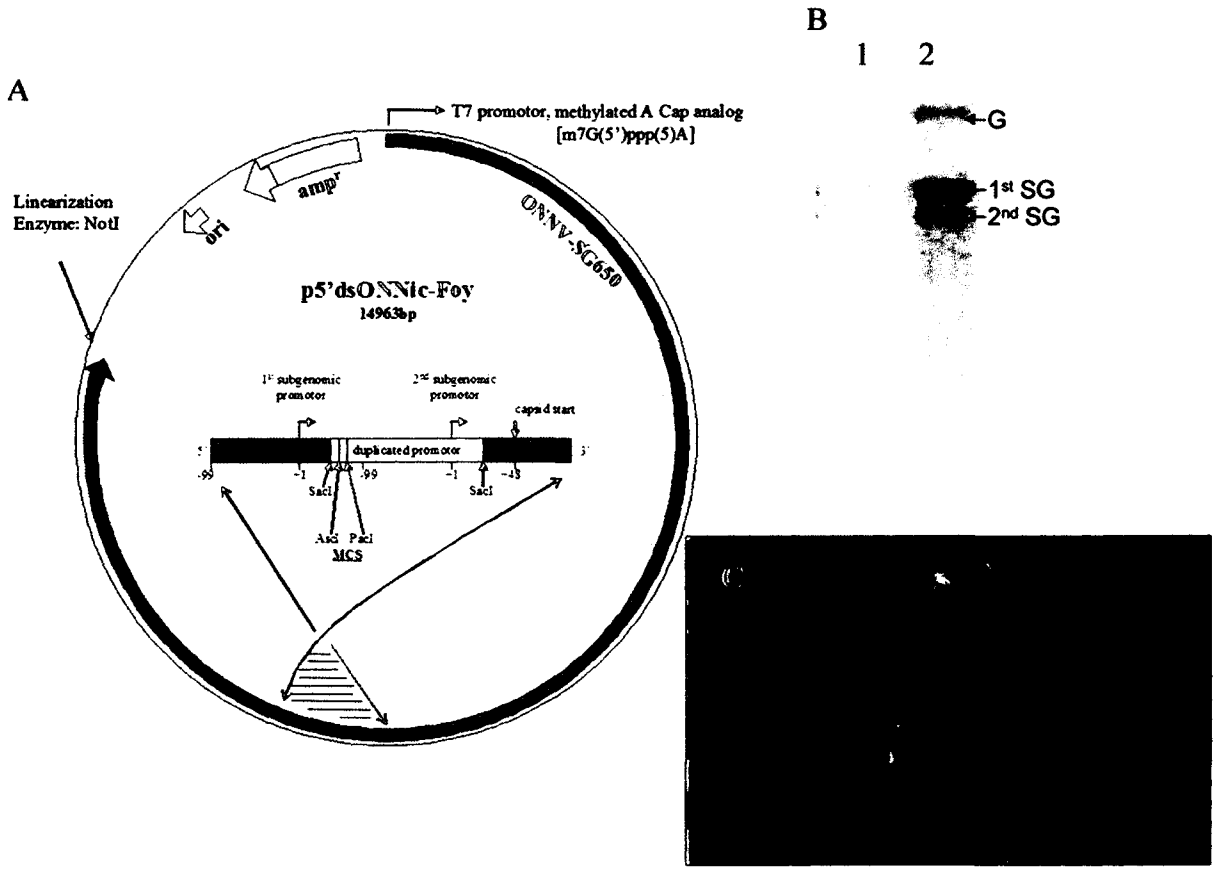


Fig. 3.1. Design and characterization of recombinant ONNV-eGFP. (A) p5'dsONNVic/Foy contains a full length cDNA of the ONNV genome with a second subgenomic promoter inserted 3' of the original subgenomic promoter. eGFP was cloned into the MCS and is transcribed under control of the first subgenomic promoter, generating ONNV-eGFP. (B) Characterization of transcripts of ONNV RNA after injection into adult female *A. gambiae*. Lane 1 and 2 show ONNV-eGFP transcript profile 48 and 72 h after infection. The blot was hybridized using a radiolabeled ONNV E2 gene as probe, and shows the production of the full length genomic and two subgenomic transcripts. (C) eGFP expression at 9 dpi occurs throughout the body.

Table 3.3. Dissemination of 5'dsONNic-Foy/EGFP following injection in *An. gambiae*

	head		thorax + abdomen
	GFP	IFA	PFU/mosq (log_{10})
3 dpi	24% (12/50)	24% (12/50)	2.1
6 dpi	33% (16/49)	30% (15/50)	1.7
9 dpi	52% (26/50)	68% (34/50)	2.3

Co- injection of ONNV and dsRNA targeting ONNV nsP3 gene sequence inhibits virus replication. We initially tested whether the RNAi pathway in *An. gambiae* could inhibit virus replication and dissemination when triggered by introduction of dsRNA derived from the *nsP3* gene (dsnsP3) to target the virus genome. Mosquitoes were co-injected with $5.3x10^3$ PFU of virus and 250ng of ONNV dsnsP3. Mosquitoes were also co-injected with the same dose of ONNV-eGFP and *βgal* dsRNA (dsβgal). Primers for generating dsRNAs are listed in Table 1. In mosquitoes receiving dsnsP3, eGFP was usually restricted to thoracic tissue surrounding the site of injection (Fig. 2A). Mosquitoes co-injected with virus and dsβgal control usually had more extensive expression of eGFP in the thorax (Fig. 2B). At 3 dpi, 86% (n=58) of mosquitoes receiving dsβgal expressed eGFP in thoracic tissues but only 38% (n=52) expressed eGFP in thoracic tissues when dsnsP3 was co-injected (Table 3.4). At 3 dpi, 23% (n=52) of mosquitoes receiving dsnsP3 showed eGFP expression in head tissues and none of the same mosquitoes had eGFP in abdominal tissue (Table 3.4). In contrast, 34% (n=58) of mosquitoes receiving dsβgal had eGFP expression in their head and abdominal tissues (Table 3.4). At 6 dpi, 57% (n=53) of mosquitoes injected with dsβgal had eGFP in their

head and abdominal tissues; however eGFP could only be seen in 36% of the head tissue and 16% of the abdominal tissue (n=52) of mosquitoes co-injected with virus and dsnsP3 (Table 3.4).

To quantitate virus abundance, mosquitoes were triturated and virus titer determined by plaque assay from these same treatment groups. No significant difference in the number of plaques was observed between the mosquitoes injected only with ONNV-eGFP and mosquitoes co-injected with dsβgal ($p > 0.08$). However, mosquitoes co-injected with ONNV-eGFP and dsnsP3 had significantly fewer plaques than those mosquitoes co-injected with ONNV-eGFP and dsβgal ($p < 0.0001$) at both 3 and 6 dpi (Fig. 3.3).

Fig. 2. eGFP expression 3 dpi in mosquitoes co-injected with dsRNAs and ONNV-eGFP. (A) Co-injection of ONNV-eGFP and dsnsP3, (B) co-injection of ONNV-eGFP and dsβgal, and (C) co-injection of ONNV-eGFP and dsAgAgo2. Mosquitoes injected with dsnsP3 show a dramatic reduction in eGFP expression when compared to dsβgal-injected controls, while mosquitoes injected with dsAgAgo2 show an increase in eGFP expression over controls.

Table 3.4: Percent of injected mosquitoes displaying eGFP expression in body tissues after co-injection of ONNV-eGFP and dsnsP3 or dsAgAgo2.

	3 dpi			6 dpi		
ONNV-eGFP	head	thorax	abdomen	head	thorax	abdomen
No dsRNA	10% (4/40)	85% (34/40)	0% (0/40)	36% (14/44)	93% (41/44)	20% (9/44)
dsβgal	34% (20/58)	86% (50/58)	34% (20/58)	57% (30/53)	89% (47/53)	57% (30/53)
dsnsP3	23% (12/52)	38% (20/52)	0% (0/52)	36% (20/55)	71% (39/55)	16% (9/55)
dsAgAgo2	98% (62/63)	100% (63/63)	97% (61/63)	100% (36/36)	100% (36/36)	100% (36/36)

Figure 3:

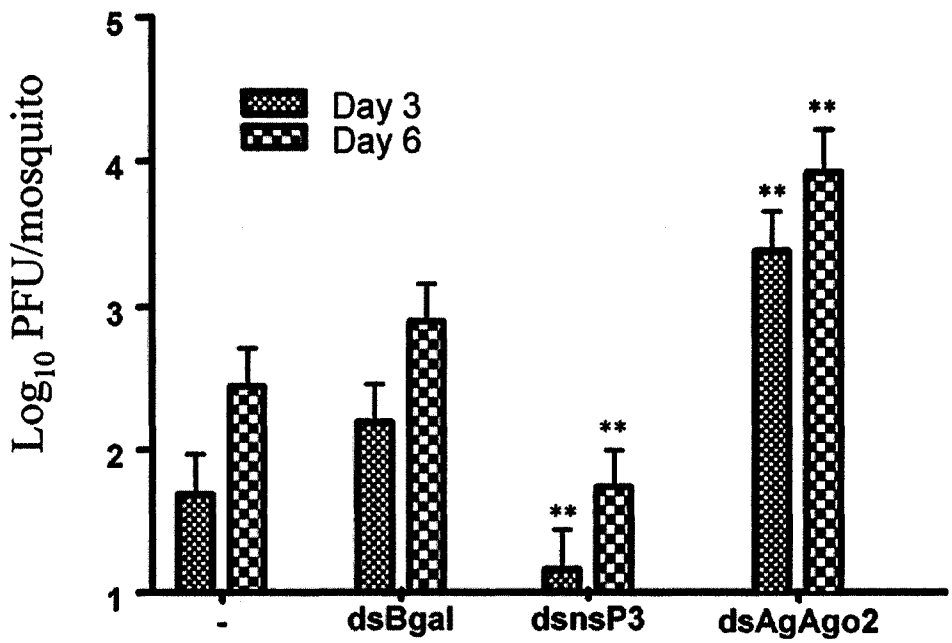


Figure 3.3. Titers of ONNV-eGFP in mosquitoes co-injected with dsRNA homologous to ONNV *nsP3* and *AgAgo2*. Compared with the non-specific *β-gal* dsRNA, mosquitoes co-injected with virus and dsnsP3 had statistically significant decreases in infection at both 3 and 6 dpi ($p < 0.01$). Viral titers of ONNV-eGFP increased significantly in mosquitoes at 3 and 6 dpi after co-injection with dsRNAs homologous with *AgAgo2* ($p < 0.0001$ and $p = 0.0006$ at 3 and 6 dpi, respectively).

Co-injection of ONNV and dsRNAs homologous to AgAgo2 increases virus replication. We then co-injected mosquitoes with dsRNAs derived from the *AgAgo2* (dsAgAgo2) to observe whether silencing of *AgAgo2* expression would make mosquitoes more permissive to ONNV. At 3 dpi these mosquitoes usually displayed dramatic increases in eGFP expression with eGFP detected in tissues in all body segments (Fig. 3.2C). At 3 dpi, only 34% (n=58) of mosquitoes treated with ONNV-eGFP and dsβgal showed eGFP in their heads and abdomens, but ≥97% (n=63) of mosquitoes treated with dsAgAgo2 showed eGFP in the same tissues (Table 3.4). At 6 dpi, 57% (n=53) of mosquitoes injected with virus and dsβgal displayed eGFP in head and abdominal tissues; 100% (n=36) injected with virus and dsAgAgo2 displayed eGFP in those tissues (Fig. 3.3).

We examined accumulation of *AgAgo2* mRNA in mosquitoes following injection with dsAgAgo2 to determine whether the dsRNA specifically silenced *AgAgo2* mRNA. Injection of non-specific dsβgal failed to silence *AgAgo2* mRNA, but injection of dsAgAgo2 silenced accumulation of *Ago2* mRNA in mosquitoes (Fig. 3.4, lane 3). We detected partial recovery of *AgAgo2* mRNA when ONNV was co-injected with dsAgAgo2 (Fig 3.4, lane 4). *AgAgo2* mRNA levels in mosquitoes injected with ONNV alone were similar to non-injected and dsβgal-injected controls.

Finally, virus titers were determined on the same mosquitoes used for determining the eGFP expression profiles. Mosquitoes treated with dsAgAgo2 had significantly more infectious virus per mosquito at 3 dpi than mosquitoes treated with *dsβgal* (16 fold increase; $p < 0.0001$; Fig. 3.3). At 6 dpi, viral titers increased in all mosquitoes tested, but dsAgAgo2 treated mosquitoes still had significantly more virus per mosquito than *dsβgal*-treated controls ($p = 0.0006$; Fig 3.3).

Figure 4:

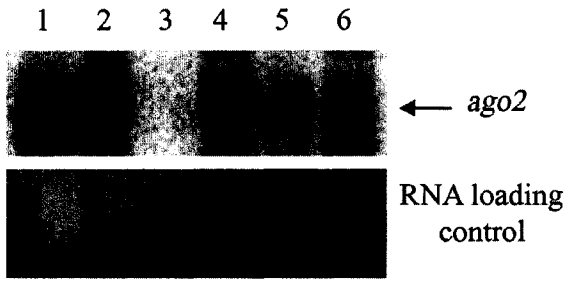


Fig. 3.4. Northern blot analysis of *AgAgo2* mRNA after injection of mosquitoes with ONNV-eGFP, dsRNA or dsRNA and ONNV-eGFP. Injection of dsAgAgo2 results in the reduction of *Ago2* transcript levels. Mosquitoes injected with dsAgAgo2 and virus showed partial recovery of the *Ago2* mRNA accumulation. Top row: Northern blot hybridized with *ago2* probe. Lane 1: mock injected, Lane 2: dsβgal 3dpi, Lane 3: dsAgAgo2 3dpi, Lane 4: dsAgAgo2 + ONNV-eGFP 3dpi, Lane 5: ONNV-eGFP 2dpi, Lane 6: ONNV-eGFP 3dpi. Bottom row. Ethidium bromide stain of northern blot showing ribosomal RNA in each lane and verifying that equivalent amounts of total RNA were added to each lane.

Co-injection of ONNV and dsRNAs homologous to AgAgo1, 3, 4 and 5. Mosquitoes were injected with virus and dsRNAs derived from *AgAgo 1, 3, 4,* and *5* (Hoa NT, et al 2003). These studies were performed to observe whether silencing of other *AgAgo* genes also could increase mosquito permissiveness to ONNV replication, possibly implicating them in RNAi modulation of ONNV-eGFP replication. At 3 and 6 dpi, mosquitoes co-injected with virus and *AgAgo*-derived dsRNAs (dsAgAgo1, 3, 4, and 5) usually displayed similar eGFP expression patterns to that seen with the dsβgal control mosquitoes (Table 3.5). The only exception was that mosquitoes injected with dsAgAgo3 consistently had greater dissemination of virus in all tissues than those injected with either *dsβgal* or dsAgAgo 1, 4, and 5 (Table 3.5). These observations were confirmed by virus titration. Virus titers in mosquitoes injected with dsAgAgo3 differed significantly from those mosquitoes injected with *dsβgal* at 3 dpi ($p = 0.0067$) and at 6 dpi ($p = 0.0141$) (Fig. 3.5). Virus titers in mosquitoes injected with dsAgAgo 1, 4, and 5 did not differ statistically from those mosquitoes injected with *dsβgal* at 3 and 6 dpi.

Co-injection of ONNV and dsRNAs homologous to AgDcr1 and AgDcr2. A. gambiae mosquitoes were injected with ONNV-eGFP and dsRNAs derived from *AgDcr1 and AgDcr2* (Hoa et al 2003). These studies were performed to see whether silencing of *AgDcr* genes would result in an increased permissiveness to ONNV infection. *AgDcr2* had previously been implicated as being involved in RNAi in *An. gambiae*, and may also be involved with modulating ONNV replication in *An. gambiae*. At 3 dpi, mosquitoes co-injected with virus and dsRNA derived from *AgDcr* genes showed greater dissemination of ONNV-eGFP to abdominal tissues than the *dsβgal* injected controls. Levels of dissemination to the head were consistent between the groups. At 6 dpi, the

AgDcr injected mosquitoes had greater levels of dissemination to all tissues when compared to the *dsβgal* control mosquitoes. Plaque titrations performed on these mosquitoes revealed that the mosquitoes injected with *AgDcr1* and *AgDcr2* had significantly more infectious virus than those mosquitoes injected with *dsβgal* at 3 dpi ($p = 0.0007$ and 0.0005 respectively). There was no statistical difference in titers at 6 dpi (Figure 3.6).

Northern blot analysis was first used to analyze changes in *AgDcr1* and *AgDcr2* gene expression after dsRNA injection and virus infection. The probes used for hybridization were unsuccessful in detecting these two mRNAs on the blots (data not shown). As a result, quantitative-reverse transcription-PCR (qRT-PCR) was used to determine changes in mRNA levels.

Levels of *AgDcr1* and *AgDcr2* transcripts were normalized to the S7 ribosomal protein transcript as an internal control (Blandin et al., 1997, Richman et al., 1996, and Levashina et al., 2001). Levels of *AgDcr1* and *AgDcr2* expression were analyzed in mosquitoes that were co-injected with ONNV-eGFP and dsRNA to either dicer transcript and were compared directly to levels in mosquitoes that were co-injected with ONNV-eGFP and *dsβgal*. Mosquitoes that were co-injected with ONNV-eGFP and *dsDcr2* demonstrated changes of -1.2 fold at 1 dpi, -2.4 fold at 3 dpi, and -1.6 fold at 6 dpi in the *AgDcr2* transcript, although none of these changes were determined to be significant according to a one-sample t-test ($p > 0.05$) (Figure 3.7). Mosquitoes that were co-injected with ONNV-eGFP and *dsDcr1* demonstrated a statistically significant change of 1.6 fold at 1dpi ($p = 0.0218$) and non-significant changes of 1.3 fold at 3 dpi, and –2.1 fold at 6 dpi ($p > 0.05$) (Figure 3.8).

Figure 5:

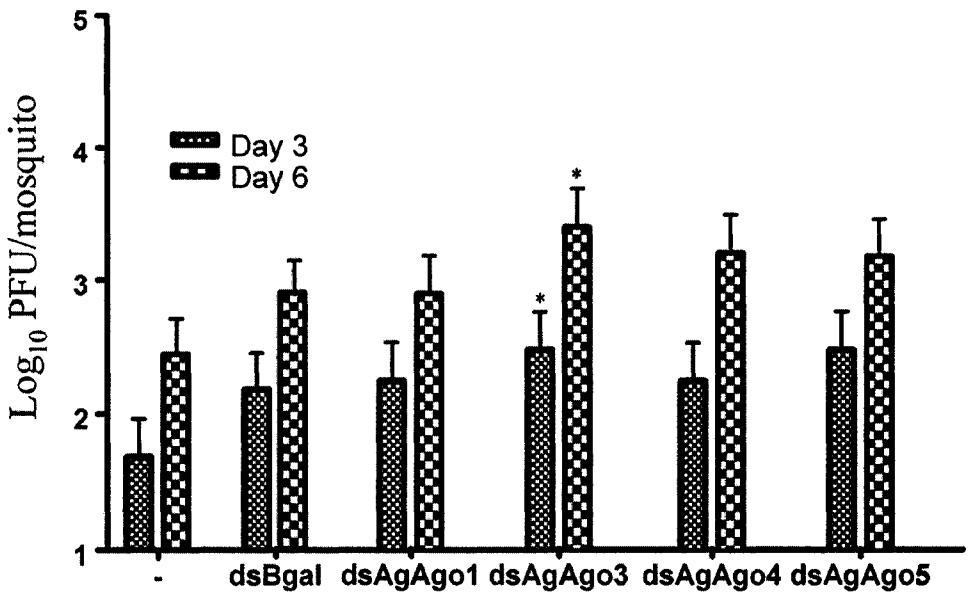


Fig. 3.5. Titers of ONNV-eGFP in mosquitoes at 3 and 6 dpi following co-injection of dsRNA homologous to *AgAgo1, 3, 4,* and *5*. Compared with the non-specific *β-gal* dsRNA, only mosquitoes co-injected with virus and dsAgAgo3 had statistically significant increases at 3 and 6 dpi ($p \leq 0.0141$) (*).

Figure 6:

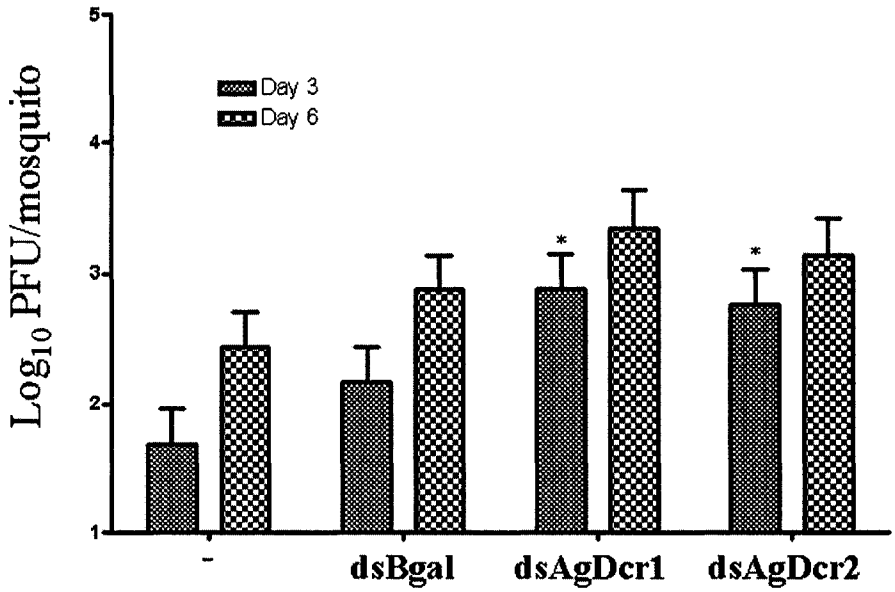


Fig 3.6: Viral titers of ONNV-eGFP in mosquitoes co-injected with dsRNA homologous to *AgDcr1 and AgDcr2*. Compared with the non-specific *β-gal* dsRNA, mosquitoes co-injected with virus and dsdcr1 and dsdcr2 had statistically significant increases in infection at 3 dpi ($p = 0.0007$ and 0.0005 respectively). Viral titers of ONNV-eGFP were not significantly different in these same groups at 6 dpi.

Table 3.5: Percent of injected mosquitoes displaying eGFP expression in body tissues after co-injection of ONNV-eGFP and dsRNAs from AgAgo1, 3, 4 and 5.

	3 dpi			6 dpi		
ONNV-eGFP	head	thorax	abdomen	head	thorax	abdomen
No dsRNA	10% (4/40)	85% (34/40)	0% (0/40)	36% (14/44)	93% (41/44)	20% (9/44)
dsβgal	34% (20/58)	86% (50/58)	34% (20/58)	57% (30/53)	89% (47/53)	57% (30/53)
dsAgAgo1	16% (9/55)	98% (39/40)	32% (18/55)	81% (26/32)	97% (31/32)	94% (30/32)
dsAgAgo3	55% (22/40)	98% (39/40)	78% (31/40)	75% (29/39)	100% (39/39)	80% (31/39)
dsAgAgo4	38% (2052)	94% (49/52)	56% (29/52)	73% (25/34)	100% (34/34)	82% (28/34)
dsAgAgo5	46% (24/52)	98% (51/52)	48% (25/52)	62% (22/36)	94% (34/36)	42% (15/36)

Table 3.6: Percent of injected mosquitoes displaying eGFP expression in body tissues after co-injection of ONNV-eGFP and dsnsP3 or dsAgDcr1/2.

	3 dpi			6 dpi		
ONNV-eGFP	head	thorax	abdomen	head	thorax	abdomen
No dsRNA	10% (4/40)	85% (34/40)	0% (0/40)	36% (14/44)	93% (41/44)	20% (9/44)
dsβgal	34% (20/58)	86% (50/58)	34% (20/58)	57% (30/53)	89% (47/53)	57% (30/53)
dsAgDcr1	56% (34/60)	98% (59/60)	73% (44/60)	91% (31/34)	100% (34/34)	91% (31/34)
dsAgDcr2	46% (24/52)	100% (52/52)	75% (39/52)	96% (45/47)	100% (47/47)	98% (46/47)

Figure 7

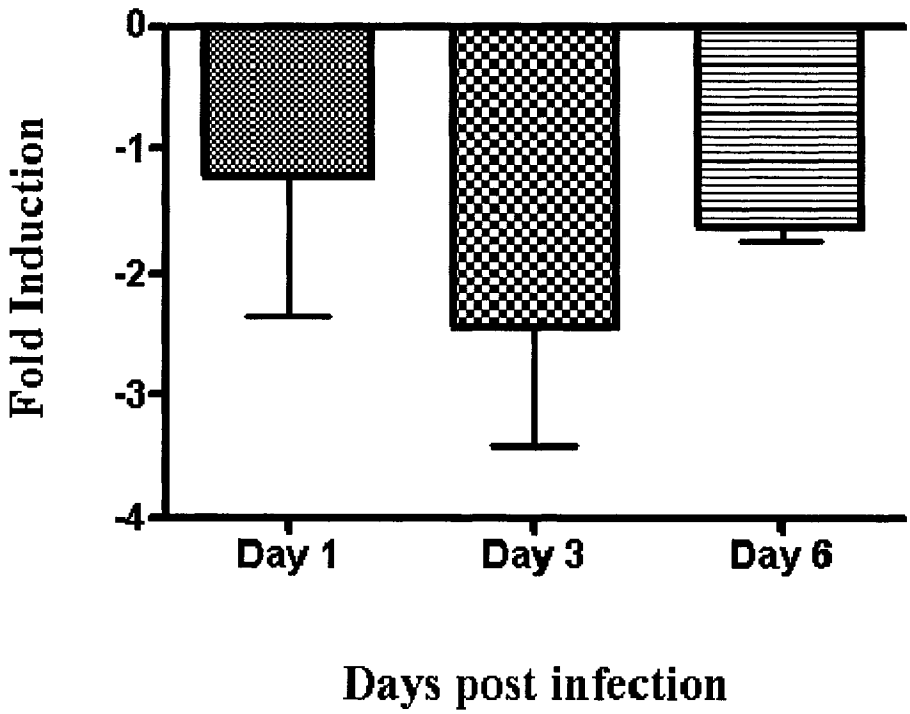


Fig. 3.7: qRT-PCR results for *AgDcr2* expression after ONNV-eGFP and dsdcr2 injection. Injection of dsAgDcr2 resulted in the reduction of *Dcr2* transcript levels at 1 dpi (-1.217 fold), 3 dpi (-2.443 fold), and 6 dpi (-1.635 fold). Each time point consisted of three replicates. No timepoint showed statistical significance of $p < 0.05$. Transcripts of *AgDcr2* in this experiment were relative to those in mosquitoes injected with ONN-eGFP and *dsβgal*.

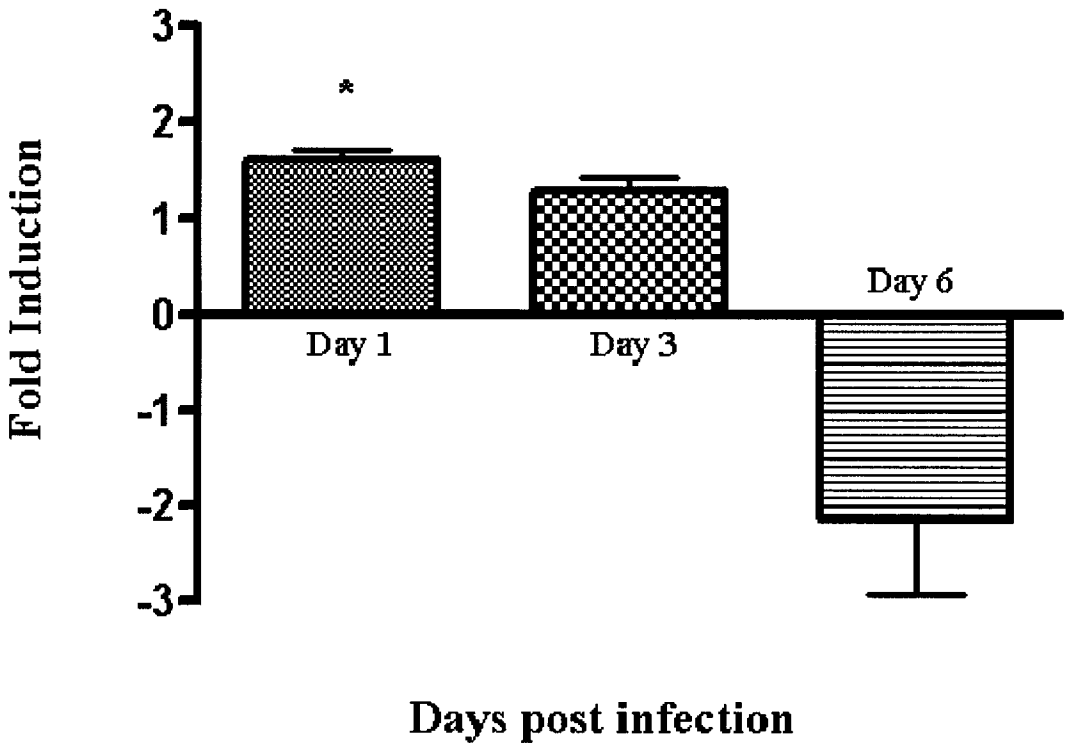


Fig. 3.8: qRT-PCR results for *AgDcr1* expression after ONNV-eGFP and dsdcr1 injection. Injection of dsAgDcr1 resulted in changes in *Dcr1* transcript levels at 1 dpi (1.607 fold), 3 dpi (1.300 fold), and 6 dpi (-2.145 fold). Each time point consisted of three replicates. Differences in transcript levels were statistically significant at 1 dpi (p = 0.0218). Transcripts of *AgDcr1* in this experiment were compared to those in mosquitoes injected with ONNV-eGFP and *dsβgal*.

Discussion

In this study, we demonstrated for the first time that RNAi can repress arbovirus replication in mosquitoes. We described an alphavirus transducing system based on ONNV and followed the course of infection after injection of recombinant virus into the *An. gambiae* hemocoel. After injection into the hemocoel, ONNV-eGFP replicated slowly in *An. gambiae.* Injection of dsRNA cognate to a portion of the ONNV genome and stimulation of RNAi further slowed ONNV replication in this mosquito species. Injection of dsRNA cognate to *AgAgo2*, a gene known to function in the *An. gambiae* RNAi pathway, silenced RNAi, thereby permitting ONNV-eGFP to more quickly replicate and disseminate in mosquitoes (Hoa et al., 2003). In addition, injection of dsRNA derived from *AgAgo3* made mosquitoes more permissive to ONNV-eGFP replication at 3 and 6 dpi. Injection of dsRNA derived from *AgDcr1* and *AgDcr2* also allowed ONNV-eGFP to replicate to higher titers at 3 dpi. In contrast, dsRNAs derived from *AgAgo 1, 4,* and *5* did not significantly alter virus replication at either 3 or 6 dpi. These data suggest a regulatory role for RNAi in controlling arbovirus infections in mosquitoes and define some of the components of the antiviral RNAi pathway.

The pathogenesis of ONNV in *An. gambiae* is unusual when compared to typical alphavirus/vector models. When injected into culicine mosquitoes, ONNV, like other alphaviruses (Sindbis virus, Venezuelan equine encephalitis virus), replicates efficiently and rapidly spreads throughout the mosquito (Bowers et al., 1995; Weaver et al., 1984). However, ONNV replication in *An. gambiae* is relatively slow following intrathoracic injection. ONNV also infects *An. gambiae* midgut tissues, but has an unusual tropism for the anterior midgut epithelium and is limited in its ability to escape from the alimentary

canal following *per os* infection (Brault et al., 2004). *An. gambiae* may not be the ideal vector for ONNV. *An. funestus* has been implicated as a potentially better vector of ONNV in Africa (Corbet et al., 1961; Williams and Woodall, 1965). It would be interesting to observe whether *An. funestus* can modulate ONNV-eGFP infection as readily as *An. gambiae*.

Our results point to *AgAgo2*, and possibly *AgAgo3*, *AgDcr1*, and *AgDcr2* proteins as critical components of a mosquito RNAi pathway involved in the inhibition of alphavirus replication. ONNV replication was most affected by silencing of *Ago2*, which has been shown to be an important RISC component in *Drosophila* (Hammond et al., 2001). Northern blot analysis showed that the presence of dsAgAgo2 leads to reduced amounts of *AgAgo2* mRNA *in vivo*, but the analysis also suggested that infection with ONNV stimulates recovery of *AgAgo2* mRNA accumulation. A possible explanation is that ONNV infection induced transcription of *AgAgo2*, although there is no supporting evidence for this hypothesis. Another possible explanation is that ONNV encodes a suppressor of RNAi in the mosquito, which has been shown with other RNA viruses (Roth et al., 2004). The ONNV suppressor might counteract the silencing resulting from dsAgAgo2 injection. Transcription of ONNV-eGFP was first detected at 72 h post-infection (figure 1B), which is when an ONNV suppressor protein may be translated in sufficient quantities to have an effect on RNAi. Even if one or both of these possibilities is correct, it is clear that the injection of dsAgAgo2 had a strong biological effect that lasted at least 6 days after treatment and resulted in increased ONNV titers in the mosquito.

Our results also show that injection of dsAgDcr1 and dsAgDcr2 results in wider dissemination patterns of ONNV-eGFP and significantly higher viral titers at 3 dpi. Initial northern blot analyses performed did not detect *AgDcr1/2* transcripts, so qRT-PCR was used for more sensitive detection of *dcr* transcripts. The injection of dsAgDcr2 resulted in a trend toward down-regulation of *AgDcr2* expression; however, the results were not statistically significant. One possible explanation for this result is that the injection of any dsRNA induces increased expression of *AgDcr2*. However, our QRT-PCR analysis did not support this, as there was no significant increase in *AgDcr2* transcript levels after injection of a non-specific dsRNA when compared to a mock control (Appendix Figure A.3). We also hypothesized that Dcr2 levels may be influenced by the presence of the replicating virus. Infection of ONNV-eGFP did not increase *Dcr2* mRNA levels when compared to mock-injected mosquitoes (Appendix Figure A.4).

After co-injection of ONN and dsAgDcr1, *AgDcr1* transcript levels increased slightly at 1 and 3 dpi. These results are not consistent with the expected down-regulation due to RNAi. Dicer-1 is now understood to be involved with the production of miRNAs, and not the processing of exogenous dsRNA into siRNAs (Lee et al., 2004). *AgDcr1* levels do not change with the injection of a non-specific dsRNA (Appendix Figure A.5). Based on these results, the injection of dsAgDcr1 may cause an increase in *AgDcr1* transcript levels to ensure the proper processing of miRNAs in mosquitoes. A detailed time course examining the *AgDcr1* transcript by qRT-PCR along with monitoring of miRNA processing in mosquitoes may provide evidence to support this hypothesis.

The reason for these conflicting results may also lie within the design of the experiment. For the QRT-PCR testing of mRNA levels, mosquitoes were pooled into groups of 5 for the analysis. Testing mosquitoes individually may be required for accurate QRT-PCR analysis of RNAi components. The results from our plaque assays on individual dsnsP3-injected mosquitoes have shown that reduction of virus does not occur in 100% of the mosquitoes.

The Argonaute family comprises a group of proteins some of which are required for RNAi with others having roles in regulating development. We have demonstrated that *AgAgo2* and *AgAgo3* are involved in RNAi both in anopheline cell culture (Hoa et al., 2003) and in *An. gambiae*. While the *Drosophila* ortholog of *Ago2* has been characterized, *Ago3* has not (Carmell et al., 2002). *DmAgo1* is required for efficient RNAi in *Drosophila*, functioning in the pathway after the production of siRNA (Williams and Rubin, 2002). We have previously hypothesized that the function of *AgAgo3* may be analogous to that of *DmAgo1* (Hoa et al., 2003). The *Drosophila* paralogues of *AgAgo4* and *AgAgo5*; *piwi* and *aubergine,* have known functions in development (Carmell et al., 2002).

Several innate immune pathways in mosquitoes have been elucidated for defense against bacterial and macroparasite infections (Blandin et al., 2004; Christophides et al., 2002; Gorman and Paskewitz, 2001; Kumar et al., 2003; Lowenberger, 2001); however, no antiviral mechanisms or pathways have been described. In vertebrate species, there are innate immune mechanisms that recognize and mount responses to dsRNA, including the interferon and PKR pathways, but neither has been detected in the mosquito. The data presented in this paper support the idea that RNAi is a mechanism to protect

mosquitoes from viral infection. We hypothesize that vector competence for an alphavirus is in part due to the balance struck between the opposing forces of vector and arbovirus evolution. Some alphaviruses may replicate and disseminate so quickly as to avoid induction of an RNAi defense or they may quickly disseminate from tissues with a strong RNAi response and into mosquito tissues with a weak RNAi response (Foy et al., 2004; Scott and Burrage, 1984; Weaver, 1986). For example, the *C. elegans* nervous system has been shown to be refractory to silencing of mRNA by RNAi (Kamath and Ahringer, 2003). In addition, many plant viruses encode suppressors of RNAi that may also be present in alphaviruses (Roth et al., 2004). Different vector species are also likely to show differences in their RNAi responses. Some mosquitoes may preferentially express negative regulators of RNAi such as the ERI-1 protein that has been identified in *C. elegans* (Kennedy et al., 2004). Data demonstrating dsRNA effectiveness in silencing a myriad of targets in *An. gambiae* (Blandin et al., 2002; Blandin et al., 2004; Osta et al., 2004) may indicate that these mosquitoes generally have a robust RNAi response, which could partly explain why anopheline mosquitoes are poor vectors of arboviruses. Mosquitoes such as *Aedes aegypti* readily transmit both alphaviruses (Chikungunya, Sindbis), and flaviviruses (yellow fever, dengue), possibly indicating that these mosquitoes have a weaker RNAi response. In support of this hypothesis, ONNV readily disseminates in *Aedes aegypti* tissues following injection, but the Sindbis virus MRE16 strain (Foy et al., 2004), which disseminates very efficiently in all culicines we have injected, could not replicate in *An. gambiae* tissues (Appendix Table A.2).

This study is the first step in understanding how RNAi naturally modulates arboviral infection in the mosquito. Investigations are underway to determine whether

the RNAi response can modulate arboviral infection and dissemination in other vector mosquitoes and by other families of arboviruses. These studies will provide a better understanding of how mosquitoes respond to virus infection and may provide us with information on how to design strategies that enhance mosquito refractoriness to arboviruses.

Chapter 4

Alphavirus infection of *Aedes aegypti* mosquitoes is modulated by RNA interference

Introduction

RNAi can be induced by dsRNA injections, double-subgenomic SINV expression systems, and transgenic techniques to silence gene expression in *Anopheles gambiae* and *Aedes aegypti* mosquitoes (Arrighi et al., 2005; Attardo et al., 2003; Bian et al., 2005; Hansen et al., 2004; Johnson et al., 1999; Michel et al., 2005; Michel and Kafatos, 2005; Vlachou et al., 2005). In the previous chapters, injection of dsRNA was used to interrupt the RNAi pathway and increase virus replication in *An. gambiae*. Given the evolutionary conservation of the RNAi pathway among eukaryotic organisms, and since *Drosophila*, anophelines and culicines are in the same order (Diptera), the RNAi pathway also would be predicted to affect virus replication in *Ae. aegypti*.

Ae. aegypti mosquitoes are vectors for a number of medically important arboviruses worldwide, including dengue and yellow fever. *Ae. aegypti* is a good model system for studying virus-vector interactions in the laboratory, and the studies performed with this species would provide a better understanding of how RNAi modulates arbovirus infection. This could lead to better technologies to exploit the system in this species. Recent genome sequence information from *Ae. aegypti* was used to find potential *dicer* and *argonaute* genes from this mosquito.

Two SIN viruses, MRE16 and AR339, have been described that differ in infection potentials of *Ae. aegypti* midguts (Myles et al., 2003). After *per os* introduction AR339, the prototype SINV, infects the mosquito midgut with poor efficiency and also has limited dissemination potential to other tissues (Myles et al., 2003; Seabaugh et al., 1998). In contrast, the MRE16 strain produces a prolific infection of the midgut and readily disseminates to other tissues following *per os* introduction (Seabaugh et al.,

1998). One hypothesis for the differences in dissemination is that RNAi modulates virus replication and dissemination from the midgut into the hemocoel (KM Myles, dissertation).

Studies were conducted using the infectious clone based on the Sindbis AR339 genome sequence, designated TR339, to examine the hypothesis that RNA interference is a mechanism by which the mosquito restricts virus escape from the midgut into the hemocoel (Myles et al., 2003). dsRNA was injected in *Ae. aegypti* to knockdown gene expression and to test the hypothesis that the RNAi pathway plays a role in modulating viral replication and controlling viral escape from the midgut.

Materials and Methods

Cell lines and virus production/Generation of TR339-eGFP: C6/36 (*Ae. albopictus*) mosquito cells were grown in L-15 medium containing 10% FBS (fetal bovine serum), 100U/ml penicillin and 100 μg/ml streptomycin at 28°C. The TR339 infectious clone (TR339ic) was provided by Dr. Robert Johnston (Division of Infectious Diseases, Department of Medicine, School of Medicine, University of North Carolina, Chapel Hill) and was further modified by Dennis Pierro by insertion of a second subgenomic promoter and the eGFP gene (p5'dsTR339ic-eGFP) (unpublished data). The infectious clone was linearized to form a template for *in vitro* transcription of infectious RNA using SP6 polymerase. RNA resulting from this reaction was electroporated into C6/36 mosquito cells. Forty-eight hours later, the medium containing the virus was collected and the virus titer was determined by plaque assay. The virus was passaged twice more in C6/36

cells for 48h at a moi of 0.01. The resulting stock virus titer was determined by plaque assay in Vero cells as previously described. The virus stock titer was 2.2 X 10^8 PFU/ml.

Mosquitoes: The *Ae. aegypti* Rexville D- Higgs' White Eye (HWE) mosquito strain was reared at 28°C, 80% relative humidity, with a photoperiod of 14 hours light and 10 hours dark. Adults were provided with a sugar source and water for nutrition.

Identification of dicer/argonaute genes in the Ae. aegypti genome: The *dcr* and *ago* sequences identified from *An. gambiae* in Chapter 2 and *Drosophila* sequences were used for an *in silico* search of the *Ae. aegypti* genome database (see Chapter 2 for accession numbers). At the current time, the *Ae. aegypti* genome project has not been completed (current release v5.0 06/2005). Sequences are deposited in a searchable database maintained by The Institute of Genomic Research at www.tigr.org. After identification as potential *dicer* genes, sequences were used by Heather Sanders (University of California, Riverside) to identify full-length cDNA clones from an *Ae. aegypti* midgut specific cDNA library.

Preparation of templates/ generation of dsRNA: cDNA templates for generation of dsRNA were amplified from total mosquito RNA extracted from 4th instar larvae as described in Chapter 3 (Generation of dsRNA). Oligonucleotide primers were designed to amplify approximately 500 bp regions of *Ae. aegypti Ago2* cDNA. Primers incorporated T7 promoter sequences at the 5' ends to facilitate dsRNA production (Table 4.1). Control ~500 bp cDNA templates were generated by PCR using primers specific

for portions of the *E. coli β-galactosidase* cDNA clone and for the *nsP3* gene from TR339-eGFP. Generation of dsRNA was performed as described in Chapter 3 (Generation of dsRNA).

Table 4.1: Primer pairs for generating dsRNA template			
Target gene	Forward Primer (5'-3')	Reverse Primer (5'-3')	Product size
β-galactosidase	ggtcgccagcggcaccgcgccttc	gccggtagccagcgcggatcatcgg	521 bp
AaAgo2	tgatgtagacgcgtcctctg	cagttcaagcagacgaacca	499 bp
TR339-nsP3	aggactactttcgtgcactg	acagaactgacgcacgt	485 bp

*All primers have a common 5' T7 promoter extension (5'-taatacgactcactatagg-3').

Table 4.2: Primer pairs for QRT-PCR in *Ae. aegypti*			
Target gene	Forward Primer (5'-3')	Reverse Primer (5'-3')	Product size
actin	cgctcgttgtcgacaatgg	cataccgaccatcacaccc	132 bp
AaAgo2	tggcacagatagtaagtcagc	agtatcaattcttcctggtttc	200 bp
TR339-nsP3	tcggaaagacggtcagaaa	gcgacggctaagatggtg	116 bp

Intrathoracic injection of dsRNA and per os infection of mosquitoes: dsRNA was diluted to a concentration of 0.5 μg/μl in DMEM prior to inoculation. Adult female *Ae. aegypti* (2-3 days post emergence) were anesthetized on ice and then intrathoracically injected with 0.5 μl of dsRNA inoculum. Each mosquito was injected with approximately 250 ng of dsRNA.

To prepare virus for mosquito infection, monolayers of C6/36 cells were infected at an MOI of 0.01. Cells were maintained in L-15 with 2% FBS for 72 hours. The virus was harvested by scraping all cells into the medium. The resulting suspension of infected cells was frozen at – 70°C in 1 mL aliquots. The virus titer was determined to be 2.2 X

10^8 PFU/ml by plaque titration in Vero cells. Three days after injection of dsRNA,virus was diluted with defibrinated sheep blood (Colorado Serum Company, Boulder, CO) to a final titer of 1.0 X 10^8 PFU/mL and placed in a water-jacketed glass membrane feeder incubated at 37°C for the duration of the bloodfeed. Mosquitoes were allowed to feed for approximately one hour through a human scented paraffin membrane. The mosquitoes were sugar-starved for 24 hours and water was removed from the mosquito cages approximately 4 hours prior to blood feeding. After blood feeding, mosquitoes were cold anesthetized and sorted. Only mosquitoes that had ingested a full blood meal were kept for further analysis.

Visualization of TR339 infection of Ae. aegypti mosquitoes/midgut dissections: *Ae. aegypti* mosquitoes were observed for virus infection and dissemination at 4 and 7 days post blood feed. Mosquitoes were killed by brief submersion in 70% ethanol and then washed in saline. Mosquito midguts were dissected in PBS and immediately examined by blue light microscopy for eGFP expression. Mosquito carcasses were examined in conjunction with midguts and scored for eGFP expression.

Determination of viral titer: At 4 and 7 dpi, five mosquitoes from each group were frozen at –70°C. The mosquitoes were triturated and virus titrated as described in Chapter 3 (Determination of virus titer). Each time point consisted of three replicates.

QRT-PCR analysis: The protocol used for QRT-PCR analysis was presented in Chapter 3 (Quantitative Reverse Transcription – Polymerase Chain Reaction (QRT-PCR)). Five mosquitoes were used at each time point, per group.

Northern Blot Analysis: Total RNA was isolated from individual mosquitoes at 3 days post injection with dsRNA using the GITC method of RNA purification as previously described (Chapter 2 M&M). Two μg of total RNA was separated on a 1% agarose-formaldehyde gel and transferred to a BrightStar Plus nylon membrane (Ambion). The blots were hybridized with ^{32}P-labeled probe complementary to *AaAgo2*. The blots were washed at 68°C and radioactivity was detected using a Storm Phosphorimager (Molecular Dynamics Inc.; Amersham Pharmacia Biotech Ltd.).

Results

Potential Dcr and Ago genes identified in Ae. aegypti: Results of BLAST searches with *Drosophila* and *Anopheles dcr* and *ago* genes are recorded in Tables 4.3 and 4.4 and the sequences are listed in Appendix figures A.6, A.8, and A.10. *Ae. aegypti* sequences identified by BLAST search were further investigated by aligning sequences with *A. gambiae dcr* and *ago* genes to determine the closest matches. The searches revealed that two *dicer*-like sequences exist in the *Ae. aegypti* genome that were similar to both *Drosophila* and *Anopheles dcr* sequences. TC64262 (982 bp) and TC59742 (2073 bp) were identified as partial *AaDcr1* and *AaDcr2* sequences, respectively. Both fragments were subjected to domain analysis using the Simple Modular Architecture Research Tool

(SMART, www.expasy.ch) as described in Chapter 2. The sequences were then aligned with the *An. gambiae* sequences.

When translated, TC64262 contains an RNase III domain followed by a dsRNA-binding domain. The translated amino acid sequence has a 94% identity with *AgDcr1*. The TC59742 translated sequence also contains an RNase III domain and a dsRNA-binding domain. TC59742 and *AgDcr2* have 96% identity when aligned. Protein alignments were performed using the bl2seq tool at www.ncbi.nlm.nih.gov. Schematics for the translated TIGR sequence fragments and their alignments to *Anopheles* dicer genes are shown in Figure 4.1.

Primers were designed for PCR amplification of the two dicer-like sequences by H. Sanders at UCR. These were used to identify full-length clones in an *Ae. aegypti* midgut specific cDNA library. The partial sequences found in TIGR are aligned with the full-length sequences of *AeDcr1* and *AeDcr2* in the schematics in Figure 4.2 and in Appendix Figures A.7 and A.9, respectively. The sequences for the full-length clones were deposited into Genbank and were assigned the accession numbers AY713295 and AY713296. The TC64262 sequence from the TIGR database aligns with the *AaDcr1* gene from nucleotides 5909 to 6579, the 3'end of the gene. This sequence encodes an RNase III domain as well as a dsRNA binding domain. The TC59742 sequence from the TIGR database aligns with the *AaDcr2* full-length clone from nucleotides 3909 to 4977, the 3' end of the gene.

Using the BLAST alignment tool at NCBI (www.ncbi.nlm.nih.gov/blast), the amino acid sequences for *Anopheles* and *Aedes* dicer sequences were aligned to determine their level of similarity. When the gene products are translated and aligned,

AaDcr1 and *AgDcr1* have 50.3% identity over the length of the gene (Figure A.11). The products of *AaDcr2* and *AgDcr2* have 51.7% amino acid identity (Figure A.12).

The BLAST searches also yielded twelve argonaute-like sequences in the *Ae. aegypti* genome. Nucleotide alignments were performed between each of these sequences to determine if any of the clones were duplications. Clones TC65831, TC65832 and TC65529, which showed highest homology to *AgAgo4,* were determined to have overlapping sequences. Clones NABWG64TR and TC63572, which most closely matched *AgAgo2,* also had overlapping sequences. Clones that showed highest homology to *AgAgo* sequences were translated and aligned with the *An. gambiae argonaute* amino acid sequences. Those that had the highest identities were bolded and are listed in Table 4.4.

Several sequence fragments were shown to have high homology to known *Drosophila* and *Anopheles* argonaute genes, but subsequent nucleotide and protein alignments showed these sequences aligned with less homology to the same regions as other fragments. These sequences are listed in appendix figure A.10, and may represent additional argonaute genes in *Ae. aegypti*. The *Ae. aegypti* genome fragments, designated AEMTBA03, AI648322, NABWG64TR, and TC57115, when translated and scanned for PFAM (protein family) domains (EMBL, www.expasy.ch) show evidence of being argonaute-like genes. AEMTBA03 (amino acids 1-98), AI648322 (aa 89-204), and NABWG64TR (aa5-109) contain PAZ domains. TC57115 (aa 1-157) contains a PIWI domain. These domains have only been found in dicer and argonaute genes.

Table 4.3: BLAST searches for *dicer* and *argonaute* genes in the *Ae. aegypti* genome using *Drosophila* sequence queries (TIGR database; www.tigr.org)

TBLASTX Highest homology		
Drosophila gene	TIGR results	e – values
DmDcr1	TC64262	6.5 e –115
DmDcr2	TC59742	6.3 e –87
DmAgo1	NACBN64TR	7.5 e -115
DmAgo2	TC63572	5.8 e –79
DmAgo3	AI648332	1.1 e –52
	TC54141	8.1 e -22
piwi	TC65831	4.7 e –94
	TC62198	2.6 e –80
	TC63456	1.4 e –78
	TC65529	9.5 e –73
	TC65832	4.8 e –66
	TC54141	8.1 e –54
	TC57115	6.5 e –47
aubergine	TC65831	6.9 e –106
	TC62198	3.0 e –88
	TC63456	9.0 e –83
	TC65529	4.9 e –82
	TC65832	5.8 e –74
	TC54141	7.0 e –57
	TC57115	2.0 e -49

Table 4.4: BLAST searches for *dicer* and *argonaute* genes in the *Ae. aegypti* genome using *Anopheles* sequence queries (TIGR database; www.tigr.org)

TBLASTX Highest homology			
Anopheles gene	TIGR results	e – values	% identity between translated sequences
AgDcr1	**TC64262**	2.2 e –111	71%
AgDcr2	**TC59742**	2.2 e –107	55%
AgAgo1	**NACBN64TR**	6.9 e –119	98%
AgAgo2	**TC63572**	1.2 e –119	67%
	NACBWG64TR	3.0 e -47	
AgAgo3	**TC54141**	2.4 e –142	63%
	TC65831	1.0 e –64	
	TC62198	4.0 e –60	
	TC63456	2.0 e –55	
	TC65529	9.0 e –50	
	AI648322	3.0 e –47	
	TC65832	6.0 e –41	
AgAgo4	**TC65831**	4.0 e –181	78%
	TC65529	9.0 e –132	
	TC62198	9.0 e –125	
	TC65832	1.0 e –115	83%
	TC63456	5.0 e –114	
	TC54141	2.0 e –72	
	TC57115	1.0 e –55	
	AEMTBA03	1.0 e -49	
AgAgo5	**TC62198**	2.0 e –124	62%
	TC65831	3.0 e-119	
	TC63456	8.0 e-118	
	TC65529	1.1 e-87	
	TC65832	3.0 e-77	
	TC54141	4.9 e-73	
	TC57115	4.4 e-61	

**Notes: Bolded selections are most likely to be the *Aedes aegypti* homologues.
Clones TC65831, TC65832 and TC65529 have overlapping sequence.
NABWG64TR and TC63572 have overlapping sequence.

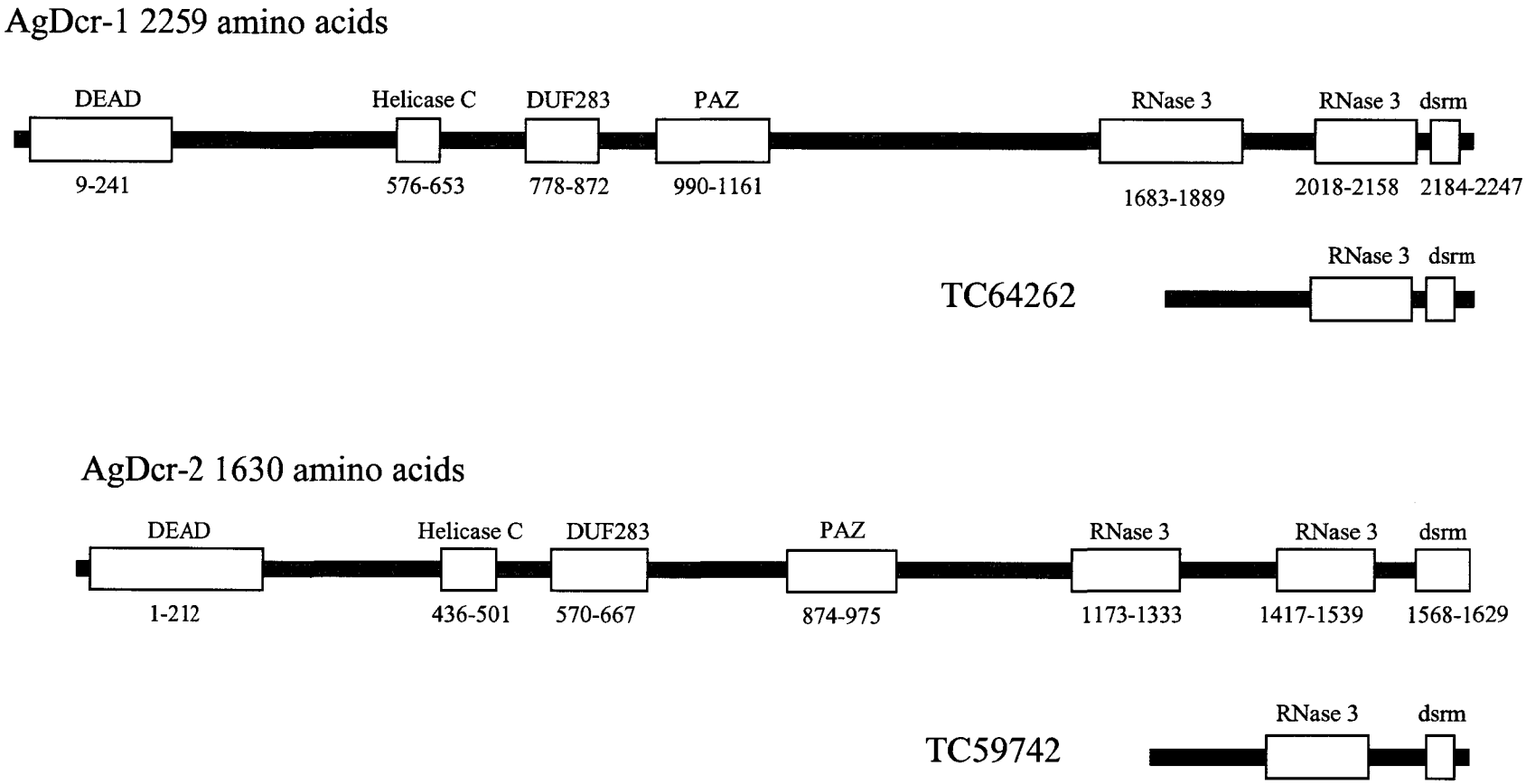


Figure 4.1: ***Ae. aegypti*** **TIGR results aligned with** ***An. gambiae dicer*** **genes.** The translated products of the sequence fragments from the putative *AeDcr1* and *AeDcr2* genes (TC64262 and TC59742, respectively) from the TIGR database are schematically shown aligned with the translated products of *AgDcr1* and *AgDcr2*. Each pair is drawn to scale.

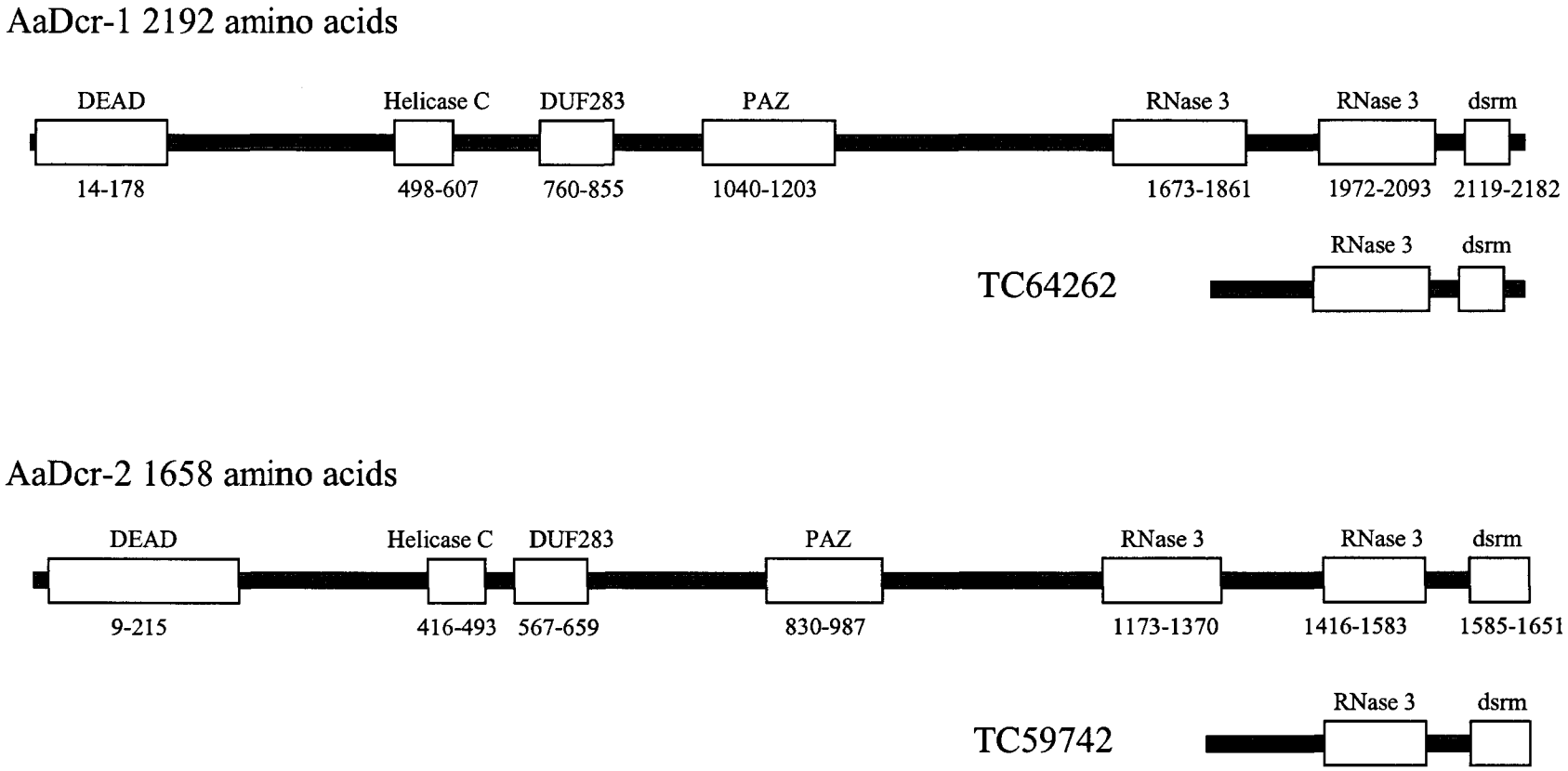


Figure 4.2: *Ae. aegypti* TIGR results aligned with *Ae. aegypti* dicer genes from cDNA library. The translated products of the sequence fragments from the putative *AeDcr1* and *AeDcr2* (TC64262 and TC59742, respectively) from the TIGR database are schematically shown aligned with the translated products of *AeDcr1* and *AeDcr2*. Each pair is drawn to scale.

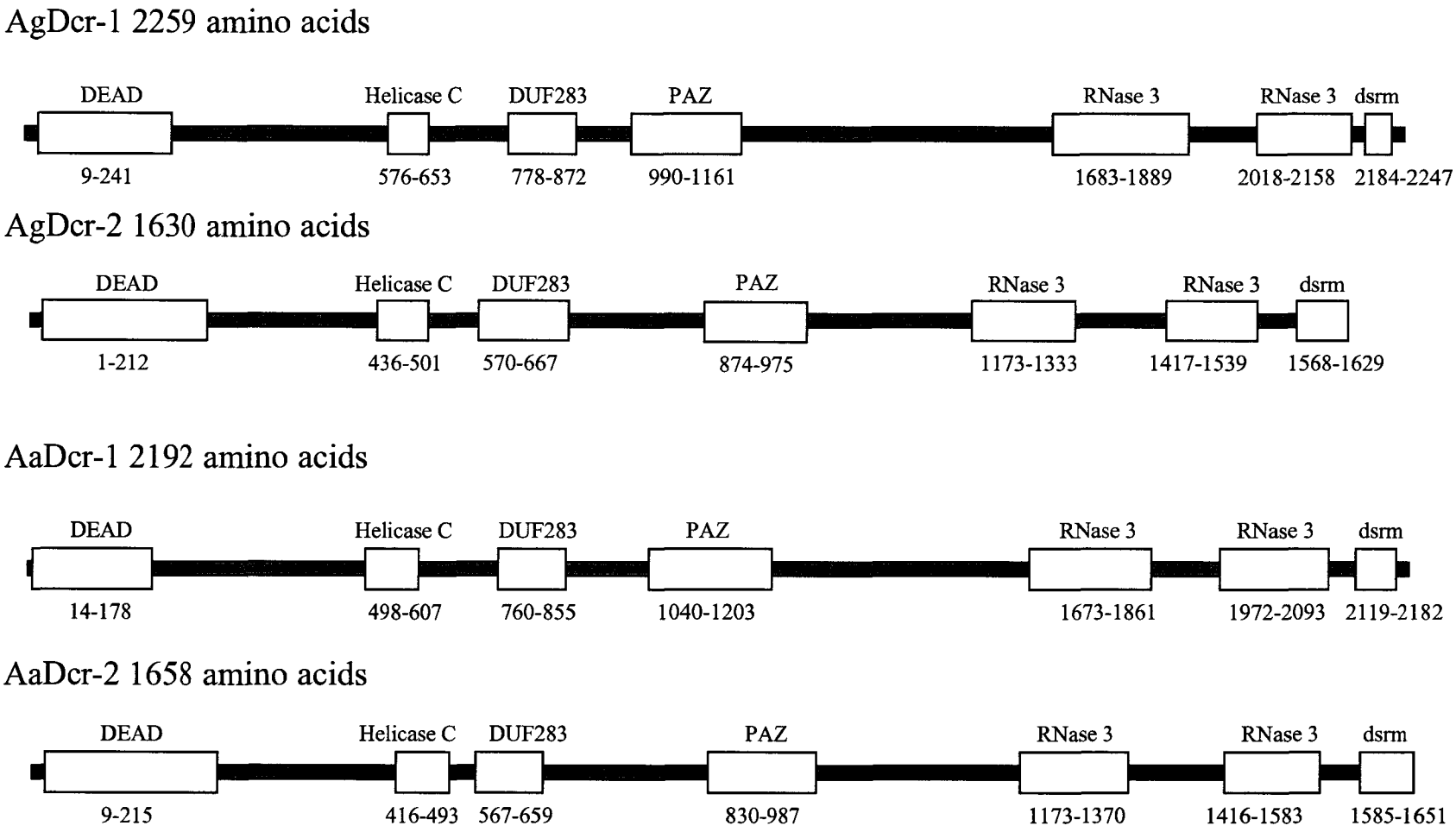


Figure 4.3: Full-length alignments of *Aedes* and *Anopheles dicer* genes. Full-length schematics of *Anopheles* and *Aedes* dicer proteins are drawn and aligned for comparison. Drawings are to scale.

Injection of dsRNA derived from TR339-nsP3 gene sequence prior to per os infection inhibits TR339-eGFP replication in the mosquito midgut: We previously showed that intrathoracic co-injection of virus and virus-derived dsRNA resulted in the silencing of virus expression in the mosquito. In these experiments, we tested whether the injection of dsRNA prior to an infectious bloodmeal could prevent virus expression in the midgut and subsequent dissemination.

The amount of eGFP expression in each midgut was given a qualitative score based on the estimated percentage of cells in the midgut expressing eGFP. If there was no detectable eGFP expression, then the midgut was scored as negative (Figure 4.4a). When 1-33% of cells in the midgut expressed eGFP, then the infection was considered "light", as seen in Figure 4.4b and c. If between 33 and 66% of cells in the midgut expressed eGFP, then the infection was scored as "moderate", as shown in Figure 4.4d. A "heavy" infection was denoted if greater that 66% of cells in the midgut were expressing eGFP (Figure 4.4e and f). The qualitative amount of eGFP expression is also shown in Table 4.5.

At 4 days after oral infection with TR339-eGFP, 83% (n=30) of mosquitoes that had been injected with dsβgal expressed eGFP in midgut tissues, but only 3% (n=32) expressed GFP when dsnsP3 was injected (Table 4.5). The single mosquito injected with dsnsP3 and scored as positive showed eGFP expression in a small focus of the midgut. At the same time point, 20% of mosquitoes that had been injected with dsβgal exhibited a disseminated infection to abdomen and head tissues, whereas no mosquitoes injected with dsnsP3 had a disseminated infection. At 7 dpi, 61% of mosquitoes injected with dsβgal (n=46) expressed eGFP in the midgut, compared with 3% of mosquitoes (n=28) that

received dsnsP3, and no dsnsP3-injected mosquitoes demonstrated eGFP expression in abdominal or head tissues. Again, mosquitoes injected with dsnsP3 only expressed eGFP in a small number of cells in the midgut (Fig.4.4b). In contrast, 2% (n=46) of mosquitoes injected with dsβgal showed eGFP expression in head and abdominal tissues.

The amount of virus present in mosquitoes from each treatment group was quantified by plaque assay. The difference in viral titers between the two treatment groups (dsβgal and dsnsP3) was statistically different at 7 dpi ($p=0.0238$) (Figure 4.5b) but not 4 dpi ($p>0.05$) (Figure 4.5a). Even though the viral titers did not differ significantly at 4 dpi, only 1/15 mosquitoes injected with dsnsP3 was positive for virus by plaque assay.

Table 4.5: Levels of eGFP expression in *Ae. aegypti* midguts from TR339-eGFP orally infected mosquitoes after injection with dsRNA, at 4 and 7 dpi.

Percentage of mosquito midguts infected

	Day 4		Day 7	
	Midgut Tissues	Disseminated Infection	Midgut Tissues	Disseminated Infection
Bgal	Light: 21/30 (70%) Moderate: 2/30 (6.7%) Heavy: 2/30 (6.7%) Total: 25/30 (83.3%)	6/30 (20%)	Light: 23/46 (50%) Moderate: 4/46 (8.7%) Heavy: 1/46 (2.2%) Total: 28/46 (60.9%)	1/46 (2.2%)
Ago2	Light: 7/25 (28%) Moderate: 6/25 (24%) Heavy: 12/25 (48%) Total: 25/25 (100%)	12/25 (48%)	Light: 20/53 (37.7%) Moderate: 14/53 (26.4%) Heavy: 5/53 (9.4%) Total: 41/53 (77.4%)	5/53 (9.4%)
nsP3	Light: 1/32 (3.1%) Moderate: 0/32 (0.0%) Heavy: 0/32 (0.0%) Total: 1/32 (3.1%)	0/32 (0.0%)	Light: 1/28 (3.6%) Moderate: 0/28 (0.0%) Heavy: 0/28 (0.0%) Total: 1/28 (3.6%)	0/28 (0.0%)

Key for infection intensities:

Light: 1-33% of midgut cells infected
Moderate: 34-66% of midgut cells infected
Heavy: 67-100% of midgut cells infected

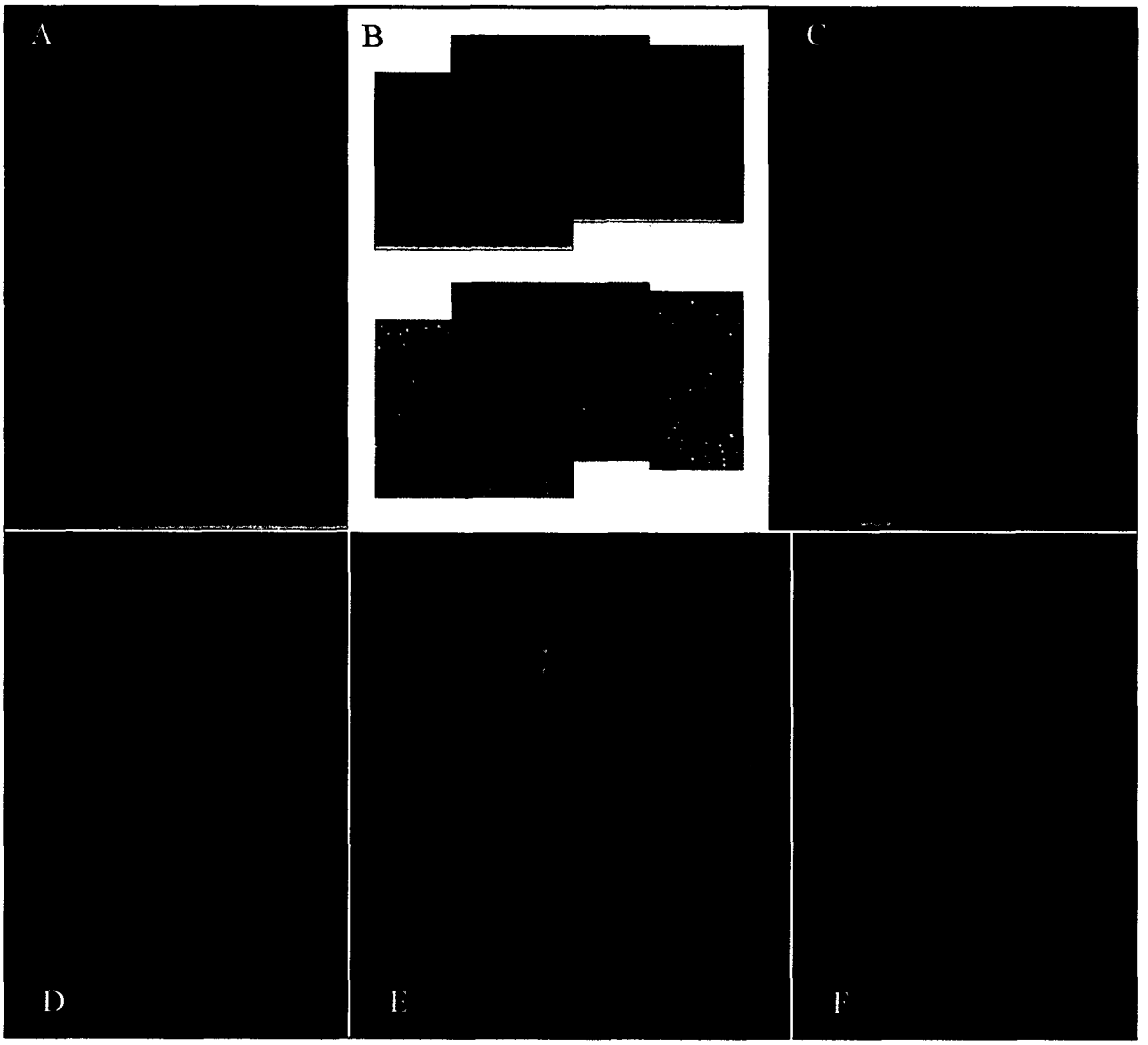


Figure 4.4: Whole mosquito midguts assayed for the presence of TR339-eGFP by blue light microscopy after blood feeding. Composite images of HWE *Ae. aegypti* midguts showing of eGFP expression at 4 dpi. (A) An uninfected midgut, displaying no eGFP expression. (B) A single midgut, photographed under both blue light (top) and white light (bottom) showing a single focus of infection. (C) This shows eGFP expression in approximately one-third of the midgut, and is considered a light infection. (D) About one-half of the midgut expressing eGFP, considered a moderate infection. (E, F) Mosquito midguts that have eGFP expression in two-thirds of the midgut or more, considered a heavy infection.

Injection of dsRNA homologous to AaAgo2 followed by per os infection of TR339-eGFP: Mosquitoes were injected with dsAaAgo2 to determine whether silencing of *AaAgo2* would increase permissiveness of the mosquitoes to oral infection by TR339-eGFP and increase dissemination of the virus from the midgut. At 4 dpi, 100% (n=25) of mosquitoes that were injected with dsAaAgo2 expressed GFP in the midgut after an infectious blood meal, compared to 83% (n=30) of mosquitoes injected with dsβgal. In addition to the number of midguts expressing eGFP, the intensity of expression also increased when dsAaAgo2 was injected. Forty eight percent of midguts injected with dsAaAgo2 had high eGFP expression (Figure 4.4e,f), while only 6.7% of the dsβgal injected mosquitoes had high expression of eGFP. The majority (70%) of the mosquitoes receiving dsβgal had light expression in the midgut (Figure 4.4b,c). The number of disseminated infections in the group injected with dsAaAgo2 also increased to 48%, compared to 20% of dsβgal controls. At 7 dpi, 77% (n=53) of mosquito midguts injected with dsAaAgo2 expressed eGFP compared with 61% (n=46) of mosquitoes from the dsβgal control group. The majority of the mosquitoes, 77.4%, receiving dsAaAgo2 also showed high eGFP expression in midgut tissues. In contrast, only 2.2% of mosquitoes receiving dsβgal had similar eGFP expression in midguts. Fifty percent of the mosquitoes receiving dsβgal were considered to have a light infection. The number of disseminated infections was also increased in the dsAaAgo2 injected group when compared to controls (9% vs. 2 %) (Table 4.5).

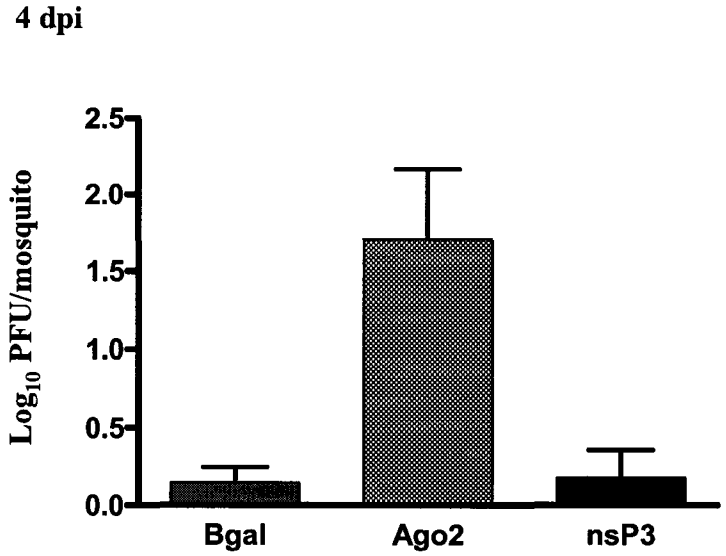
Virus titers were determined by plaque assay in 15 mosquitoes per group. At 4 dpi, viral titers in mosquitoes that were injected with dsAaAgo2 were significantly higher than titers in control mosquitoes (p=0.0042) as determined by paired t-test. At 7 dpi, the

difference in titers between the groups was not statistically significant (p=0.0664), although the average titer in dsβgal injected mosquitoes was higher than in the dsAaAgo2 injected mosquitoes (Figure 4.5 A and B).

Quantitative reverse-transcriptase PCR analysis was performed on the *AaAgo2* transcript after injection of dsRNA and infection with virus. Five mosquitoes were pooled for each RNA extraction and analysis. Quantitative analysis of *AaAgo2* mRNA did not show a reduction at 4 or 7 dpi with TR339 (p=0.1059 and 0.1007 respectively) when compared to dsβgal-injected mosquitoes (Figure 4.6).

At the time of blood feeding 5 mosquitoes each were selected from the dsβgal and dsAaAgo2 injected mosquitoes that did not ingest a blood meal. Five mosquitoes were also chosen from a group of the same age that had not been injected with dsRNA. RNA was extracted and northern blot analysis was performed on these individual mosquitoes to determine the level of *AaAgo2* transcript in the mosquitoes at the time the infectious blood meal was ingested. The analysis revealed that the injection of dsβgal did not alter the level of *AaAgo2* transcript (Figure 4.7, compare lanes 6-10 with lanes 1-5). Mosquitoes injected with dsAaAgo2 showed a reduction in the level of *AaAgo2* mRNA (Figure 4.7, lanes 11-14).

A.



B.

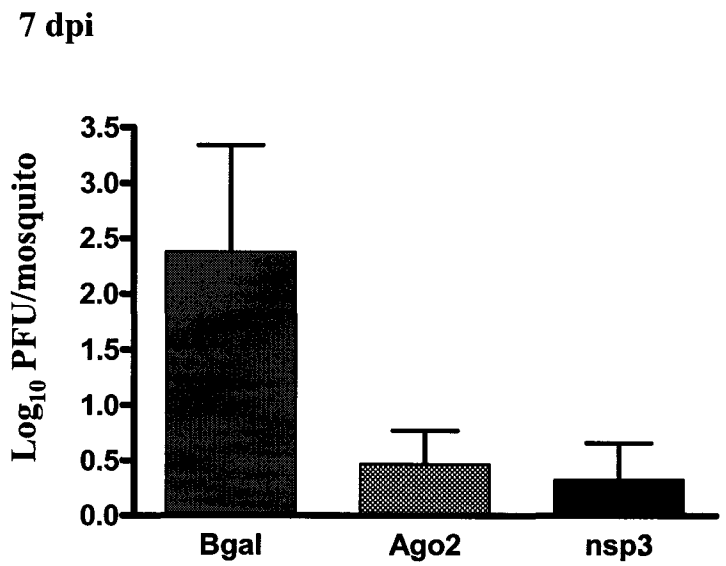


Figure 4.5: Viral titers of TR339-eGFP in mosquitoes co-injected with dsRNA homologous to TR339-*nsP3* and *AaAgo2*. Compared with the non-specific *β-gal* dsRNA, mosquitoes injected with dsnsP3 and then fed with virus had significantly lower viral titers at 7 dpi ((B), p=0.0238), but not at 4 dpi ((A), p>0.05). Viral titers of mosquitoes injected with *AaAgo2* dsRNA and then fed TR339-eGFP increased significantly at 4 dpi ((A), p=0.0042), but not at 7 dpi ((B), p=0.0664).

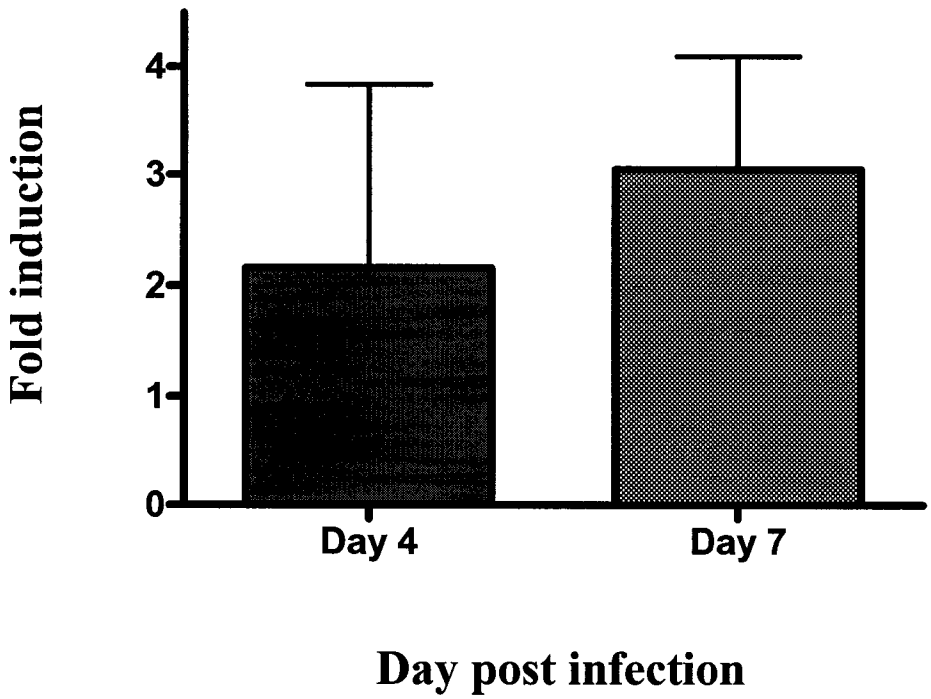


Figure 4.6: Changes in *AaAgo2* transcript determined by qRT-PCR at 4 and 7 dpi. Analysis of the *AaAgo2* transcript following injection of dsAaAgo2 and TR339-eGFP bloodfeed shows no significant change in the AaAgo2 transcript at 4 dpi (2.16 fold, p=0.1059) and 7 dpi (3.05 fold, p=.0.1007) when compared to dsβgal injected controls.

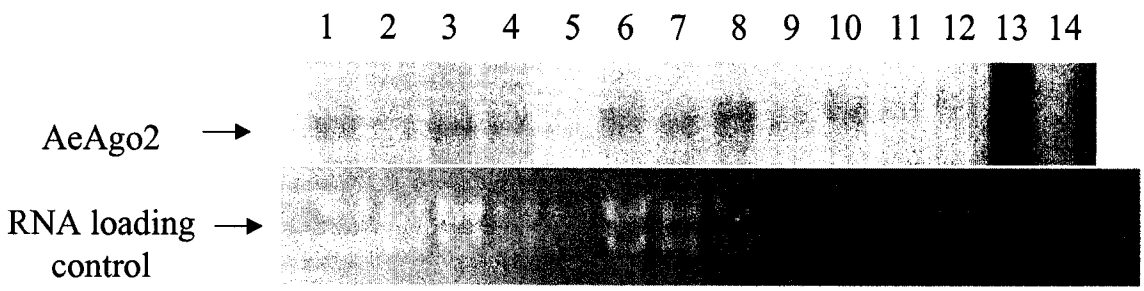


Fig. 4.7. Northern blot analysis of *AaAgo2* mRNA after injection of mosquitoes with *βgal* or *AaAgo2* dsRNA. Injection of dsAaAgo2 results in the reduction of *Ago2* transcript levels. Mosquitoes injected with dsβgal (top rows 6-10) did not have altered transcript levels from non-injected mosquitoes (top rows 1-5). Mosquitoes that were injected with dsAaAgo2 without virus infection showed decreased levels of transcript (top rows 11-14). Bottom row: Ethidium bromide stain of agarose gel showing ribosomal RNA in each lane and verifying that equivalent amounts of total RNA were added to each lane.

Discussion

In this study, we examined the ability of RNAi to modulate alphavirus infection in *Ae. aegypti* mosquitoes. In addition, we investigated the hypothesis that RNAi is the mechanism involved in the midgut escape barrier preventing the dissemination of TR339 from the mosquito midgut. Myles, *et al,* (2004) showed that a midgut escape barrier is responsible for the differences in dissemination from the midgut of *Ae. aegypti* between MRE16 and TR339. One hypothesis presented by Myles, *et al* (2004) was that RNA interference may be responsible for alphavirus modulation in the mosquito midgut.

We have previously shown that co-injection of an alphavirus and dsRNA derived from the genome of that virus results in decreased viral expression. Here, we demonstrate that injection of dsnsP3 cognate to TR339 RNA prior to an infectious bloodmeal prevents eGFP expression from the viral genome in the midgut and dissemination to secondary tissues after infection. We also show that injection of dsRNA with cognate sequence to genes in the RNAi pathway leads to increased viral gene expression in the midgut, increased viral dissemination from the midgut, and increased viral titers in the mosquito.

As seen in *Drosophila* and *Anopheles*, the genome of *Ae. aegypti* contains two dicer genes. The conservation of the number of dicer genes and domain organization may allow us to predict to the function of these genes in *Ae. aegypti*. The number of argonaute genes in *Ae. aegypti* may be greater than found in *An. gambiae* and *D. melanogaster*. BLAST analysis using *An. gambiae argonaute* genes revealed nine unique sequences that had high homology to the *Anopheles* sequences and contained PAZ and PIWI domains. The only known genes that contain these domains are in the

argonaute gene families. It is possible that this gene family has been expanded in *Ae. aegypti*. The completion of the *Ae. aegypti* genome sequence and annotation will confirm the number of *argonaute* genes encoded by the genome. More research will be needed to determine the function of any additional *ago* genes.

Injection of a dsRNA with cognate sequence to TR339 prior to an infectious blood feed reduced infection of the mosquito midgut as monitored by eGFP expression. Only one mosquito of totals of 32 and 28 assayed at 4 and 7 dpi respectively, showed a single focus of a few cells that were expressing eGFP. Viral titers from nsP3 dsRNA-injected mosquitoes were significantly lower than the dsβgal injected controls at 7 dpi, but not at 4 dpi. Small sample sizes were likely the reason for this result. Reduced expression of eGFP from the virus genome indicated that there was little virus infection of these mosquitoes at 4 dpi. In addition, only one mosquito out of 15 assayed by plaque titration contained a detectable level of infectious virus. Whole intact mosquitoes were utilized for plaque assays, and therefore, these mosquitoes were not dissected and examined for midgut infection and/or dissemination.

Injection of dsAaAgo2 prior to viral infection intensified the infection in the midgut at 4 dpi based on percentage of midgut cells expressing eGFP. The number of mosquitoes displaying a disseminated infection also increased when compared to the control dsβgal injected mosquitoes (Table 4.5). At 4 dpi, viral titers in mosquitoes injected with dsAaAgo2 prior to infectious bloodfeed increased significantly compared to controls (Figure 4.5A). This correlated well with more eGFP expression in the mosquito midgut and increased levels of dissemination demonstrated by eGFP expression (Table

4.5). These results provide evidence that the RNAi pathway can naturally limit TR339 infection in the *Ae. aegypti* mosquito, including the midgut.

Injecting dsRNA prior to feeding the virus allows the RNAi machinery to be primed for degradation of the incoming virus RNA in the case of dsnsP3, and in the case of dsAaAgo2, allows the RNAi machinery to be "disabled" before the virus is introduced. Northern blot analysis performed at the time of bloodfeeding showed that the amount of *AaAgo2* mRNA was reduced (Figure 4.7). Transient knockdown would allow the virus to establish an infection without genome degradation by RNAi, and lead to an increase in the amount of virus in the mosquito if the RNAi pathway controls virus infection. Specifically, it should allow for an increase of virus replication in the midgut, which is the first tissue the virus encounters. If the RNAi pathway were a factor in the midgut escape barrier, then down regulation of the pathway would increase dissemination from the midgut. An increase in dissemination from the midgut was seen in dsAaAgo2-injected mosquitoes. The exact reason for increased dissemination is unclear, and two possible explanations exist. The first is that the RNAi pathway controls the midgut escape barrier (MEB) and silencing of *AaAgo2* results in a breakdown of the barrier that allows for dissemination. The increase in dissemination rates could also be correlated with the amount of virus present in the midgut. *AaAgo2* silenced mosquitoes had more eGFP expression in the midgut and therefore a higher titer of virus. The increased virus load may be sufficient to increase dissemination rates from the midgut.

The knockdown of *AaAgo2* appeared to have a greater impact on viral escape from the midgut at earlier time points. This may be due to the design of the experiment. In order to test how initial infection and early dissemination changes in silencing certain

genes, it was necessary to inject dsRNA and allow mosquitoes to recover prior to bloodfeeding. As a result, observations at early time points of viral infection occurred 7 days after dsRNA injection. The effect of the dsRNA may be diminished by this point, or if a viral suppressor of RNAi is present, then the results may be altered as well. Evidence for the reduced effect of dsRNA over time was exhibited in the ONNV/*An. gambiae* experiment. Mosquitoes that were injected with dsRNA targeting the dicer genes showed significantly increased viral titers at 3, but not 6 dpi (Figure 3.6).

qRT-PCR analysis revealed that there was no significant difference in *AaAgo2* mRNA levels after injection of dsAaAgo2 followed by a bloodfeed with virus. These results were similar to what was seen with co-injection of dsAgAgo2 and ONNV-eGFP in *An. gambiae*. Although the levels of *AaAgo2* appear to increase after bloodfeed, the fold increases in the individual replicates vary and the statistical analysis shows the apparent increase is not significant. These replicates represent total RNA from a pool of mosquitoes. To limit the amount of variation between replicates, individual mosquitoes should be assayed to properly assess the variation of transcript levels between individual, rather than a pool of mosquitoes. After virus infection, the mosquitoes were assayed for *AaAgo2* mRNA at 4 and 7 dpi. These time points assay the mosquitoes at least 7 days after dsRNA injection, and may be too long to have a sustained decrease in *AaAgo2* mRNA levels. Alternatively, the virus might encode a suppressor of RNAi, which could prevent the dsAaAgo2 from being diced and incorporated into RISC, in turn preventing further degradation of the targeted mRNA. However, it is apparent the knockdown of *AaAgo2* message prior to bloodfeeding was sufficient to change the dynamics of virus infection in the mosquito.

This study shows that RNAi can naturally modulate arboviral infection in *Ae. aegypti* after an oral bloodfeed. The results of this study and our previous work in *An. gambiae* indicate that RNAi is a common antagonist of viral replication in the mosquito. Overall, knockdown of *AaAgo2* increased viral replication in midgut tissues and increased viral dissemination from the midgut. However, the phenotype of TR339 oral infection did not completely change to that observed with MRE16 oral infection (Myles et al., 2004), indicating that other factors may be involved with restricting viral movement out of the midgut. This may be due to other mosquito factors, or to genetic difference between the two virus strains. This work also demonstrated that gene expression can be effectively down-regulated by RNAi in the mosquito midgut, which may aid in developing transgenic mosquitoes and other novel strategies for preventing mosquito infection and transmission.

Chapter 5:

Summary and Conclusions

Arthropod-borne viruses remain a significant threat to human and animal populations worldwide. It is essential to understand the interactions between arboviruses and their arthropod vectors in order to develop new strategies to reduce transmission of these viruses, thereby reducing disease incidence and burden. Many approaches have been used for controlling vector-borne disease, with different levels of success. The study of the vector immune system should provide insights on how to possibly render vectors incapable of transmitting disease agents, in conjunction with other techniques such as transgenic technologies. In this dissertation, we investigated the role of RNAi as a component of the mosquito innate immune system in the response to viral infection.

RNAi has been used successfully to study gene function in mosquitoes (Blandin et al., 2002; Levashina et al., 2001). These studies demonstrated the silencing of genes in mosquitoes and mosquito cell culture can be accomplished by an RNAi-like mechanism. Included in these studies were experiments that targeted viral replication in mosquitoes and in cell culture (Adelman et al., 2001; Adelman et al., 2002; Sanchez-Vargas et al., 2004; Travanty et al., 2004). The use of dsSIN viruses has been employed to express portions of DEN viral genomes to interfere with DEN replication (Adelman et al., 2001; Gaines et al., 1996; Sanchez-Vargas et al., 2004). A mosquito cell line engineered to constitutively express an inverted repeat RNA derived from the DEN-2 genome prevented viral replication in these cells (Adelman et al., 2002). siRNAs derived from the DEN-2 genome and transfected into mosquito cells prevented the accumulation of viral RNA and viral antigen in these cells after challenge with DEN-2 (EA Travanty, Ph.D. dissertation 2005). This protection was determined to be sequence-specific, as the

siRNA did not provide protection when the cells were challenged with West Nile virus (EA Travanty, Ph.D. dissertation 2005).

Mosquitoes are responsible for the transmission of many arboviruses. Culicine mosquitoes show tremendous potential and ability to transmit many arboviruses, whereas anopheline mosquitoes do not show the same ability even though their geographic distribution is as diverse as culicines. Anopheline mosquitoes are the main vectors for *Plasmodium spp,* the protozoan agents of malaria. One hypothesis for this difference is the abilities in the RNAi responses of the two groups to recognize and eliminate viral pathogens. The complete sequencing of the *Aedes aegypti* genome and other mosquito genomes will provide more information about the RNAi genes in culicines. Additionally, studies to determine whether differences in transcript or protein levels of RNAi pathway components exist could provide insight on the robustness of the response in different mosquito species.

ONNV is an unusual virus based on its ability to be transmitted by anopheline mosquitoes (Brault et al., 2004). ONNV replicates slowly in the tissues of *An. gambiae*, but replicates well in *Ae. aegypti* (Appendix Table A.2). Based on our results, we presented the hypothesis that ONNV encodes a suppressor of RNA interference. The presence of a suppressor may allow for the virus to overcome the RNAi response and allow for replication to occur. Currently, scientists in our lab our constructing a system that will readily allow us to screen for RNAi suppressors encoded by arboviruses (B. Geiss, personal communication).

Another important analysis that will further our knowledge and understanding of the RNAi response in mosquitoes is to determine the baseline expression levels of

important RNAi genes in different tissues of the mosquito, and identify changes in their transcript and protein levels after bloodfeeding and infection with an arbovirus. For example, microarray analysis of changes in transcript levels in the *Ae. aegypti* midgut after bloodfeeding showed an increase in the level of *AaDcr2* mRNA levels (Sanders et al., personal communication). Could this be an activation of the RNAi pathway that would be a pre-emptive strike against a viral pathogen, or is the RNAi response reactive only to viral infection? It will be important to determine baseline RNAi gene expression levels and responses to viral infection to compare the robustness of the RNAi responses in different mosquito species.

These experiments used an RNAi strategy to identify genes in mosquitoes in the RNAi pathway involved with defense against alphavirus infection. Other studies have shown the applicability of downregulating genes in the RNAi pathway using an RNAi approach to study their involvement with RNAi (Dudley et al., 2002). The early studies were aimed to determine whether mosquitoes encode the genetic components of the RNAi machinery. The best-described genes known to be involved in RNAi in other organisms, such as *Drosophila*, are in the *dicer* and *argonaute* gene families. Examination of the *An. gambiae* and *Ae. aegypti* genomes indicate that genes in these families are expressed in the two mosquito species and are similar in structure to their *Drosophila* counterparts. Similar experiments using arboviruses from other virus families should determine whether RNAi is a mechanism that mosquitoes use to target other viruses as well.

It is important to note that the experimental design from Chapter 4 required injecting the mosquitoes with dsRNA 3 days prior to blood feeding. This was done in

order to allow the mosquitoes recover and ensure a high percentage of mosquitoes would bloodfeed. In addition, this allowed *AaAgo2* to be silenced, and the pathway primed to target the incoming virus using dsnsP3, prior to infection with TR339-eGFP. This experimental design may have meant that at our chosen time points for analysis, the RNAi response may have not been at optimal strength. Analysis of the ONNV titer data show that the differences in viral titers in mosquitoes injected with dsAgAgo2, dsAgDcr1 and dsAgDcr2 were not as significant at 6 dpi compared to 3 dpi. However, in the *Ae. aegypti* experiment, it was important to inject dsRNA before virus was introduced in order to observe the effects at the early time point after infection. If the effects of dsRNA do begin to wane, it will be difficult to assess the effects of gene knockouts using this method at later time points of infection. In order to study arboviruses that require a longer time to replicate in the mosquito, such as dengue virus, multiple rounds of dsRNA injection may be required. An alternative approach to this method would be to generate a transgenic mosquito that has a component of the RNAi pathway, such as *dcr2,* silenced. This approach should alleviate any issues with the waning effects of dsRNA and would eliminate the requirement for multiple rounds of injection.

The experiments in this dissertation have showed the first pathway to be directly involved with an immune response to arboviral infection. Multiple pathways have been determined to be involved with defending the mosquito from bacterial, protozoal, and fungal infections (Levashina, 2004; Meister et al., 2005). Apoptosis is another defense, which may be involved with protection of the mosquito from arbovirus infection. Although this has yet to be shown, apoptosis has been shown to be involved with clearing densonucleosis viruses from mosquito cells, as well as baculovirus from other insects

(Clem, 2005; Paterson et al., 2005). More research needs to be done to determine if this is another mechanism the mosquito employs to clear arbovirus infection. Recent advances in microarray and genome technologies, as well as reverse genetic techniques should allow for the identification of other antiviral pathways.

The results reported here clearly show that mosquitoes are able to recognize and target invading viruses. Used in conjunction with other technologies, the knowledge gained here may help reduce transmission of mosquito-borne viruses.

Literature Cited

Adelman, Z. N., Blair, C. D., Carlson, J. O., Beaty, B. J., and Olson, K. E. (2001). Sindbis virus-induced silencing of dengue viruses in mosquitoes. *Insect Mol Biol* **10,** 265-73.

Adelman, Z. N., Sanchez-Vargas, I., Travanty, E. A., Carlson, J. O., Beaty, B. J., Blair, C. D., and Olson, K. E. (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* **76,** 12925-33.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25,** 3389-402.

Ambros, V., Lee, R. C., Lavanway, A., Williams, P. T., and Jewell, D. (2003). MicroRNAs and other tiny endogenous RNAs in C. elegans. *Curr Biol* **13,** 807-18.

Anandalakshmi, R., Pruss, G. J., Ge, X., Marathe, R., Mallory, A. C., Smith, T. H., and Vance, V. B. (1998). A viral suppressor of gene silencing in plants. *Proc Natl Acad Sci U S A* **95,** 13079-84.

Anderson, P. (2005). A Place for RNAi. *Dev Cell* **9,** 311-2.

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

Attardo, G. M., Higgs, S., Klingler, K. A., Vanlandingham, D. L., and Raikhel, A. S. (2003). RNA interference-mediated knockdown of a GATA factor reveals a link to anautogeny in the mosquito *Aedes aegypti*. *Proc Natl Acad Sci U S A* **100,** 13374-9.

Barton, D. J., Sawicki, S. G., and Sawicki, D. L. (1991). Solubilization and immunoprecipitation of alphavirus replication complexes. *J Virol* **65,** 1496-506.

Baulcombe, D. (2005). RNA silencing. *Trends Biochem Sci* **30,** 290-3.

Beclin, C., Berthome, R., Palauqui, J. C., Tepfer, M., and Vaucheret, H. (1998). Infection of tobacco or *Arabidopsis* plants by CMV counteracts systemic post-transcriptional silencing of nonviral (trans)genes. *Virology* **252,** 313-7.

Bernstein, E., Caudy, A. A., Hammond, S. M., and Hannon, G. J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409,** 363-6.

Bian, G., Shin, S. W., Cheon, H. M., Kokoza, V., and Raikhel, A. S. (2005). Transgenic alteration of Toll immune pathway in the female mosquito *Aedes aegypti*. *Proc Natl Acad Sci U S A* **102,** 13568-73.

Blackburn, N. K., Besselaar, T. G., and Gibson, G. (1995). Antigenic relationship between chikungunya virus strains and o'nyong nyong virus using monoclonal antibodies. *Res Virol* **146,** 69-73.

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

Blandin, S., Moita, L. F., Kocher, T., Wilm, M., Kafatos, F. C., and Levashina, E. A. (2002). Reverse genetics in the mosquito *Anopheles gambiae*: targeted disruption of the Defensin gene. *EMBO Rep* **3,** 852-6.

Blandin, S., Shiao, S. H., Moita, L. F., Janse, C. J., Waters, A. P., Kafatos, F. C., and Levashina, E. A. (2004). Complement-like protein TEP1 is a determinant of vectorial capacity in the malaria vector *Anopheles gambiae*. *Cell* **116,** 661-70.

Boutros, M., Kiger, A. A., Armknecht, S., Kerr, K., Hild, M., Koch, B., Haas, S. A., Consortium, H. F., Paro, R., and Perrimon, N. (2004). Genome-wide RNAi analysis of growth and viability in *Drosophila* cells. *Science* **303,** 832-5.

Bowers, D. F., Abell, B. A., and Brown, D. T. (1995). Replication and tissue tropism of the alphavirus Sindbis in the mosquito *Aedes albopictus*. *Virology* **212,** 1-12.

Brault, A. C., Foy, B. D., Myles, K. M., Higgs, S., Weaver, S. C., Olson, K. E., Miller, B. R., and Powers, A. M. (2004a). Development of O'nyong-nyong virus vectors for infection and expression studies in the malaria-transmitting mosquito *Anopheles gambiae*. *Submitted.*

Brault, A. C., Foy, B. D., Myles, K. M., Kelly, C. L., Higgs, S., Weaver, S. C., Olson, K. E., Miller, B. R., and Powers, A. M. (2004b). Infection patterns of o'nyong nyong virus in the malaria-transmitting mosquito, *Anopheles gambiae*. *Insect Mol Biol* **13,** 625-35.

Brigneti, G., Voinnet, O., Li, W. X., Ji, L. H., Ding, S. W., and Baulcombe, D. C. (1998). Viral pathogenicity determinants are suppressors of transgene silencing in *Nicotiana benthamiana*. *Embo J* **17,** 6739-46.

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

Brown, A. E., Crisanti, A., and Catteruccia, F. (2003b). Comparative analysis of DNA vectors at mediating RNAi in *Anopheles* mosquito cells and larvae. *J Exp Biol* **206,** 1817-23.

Cadd, T. L., Skoging, U., and Liljestrom, P. (1997). Budding of enveloped viruses from the plasma membrane. *Bioessays* **19,** 993-1000.

Calisher, C. H., Shope, R. E., Brandt, W., Casals, J., Karabatsos, N., Murphy, F. A., Tesh, R. B., and Wiebe, M. E. (1980). Proposed antigenic classification of registered arboviruses I. *Togaviridae, Alphavirus. Intervirology* **14,** 229-32.

Calisher, C. H., and Karabatsos, N. (1988). Arbovirus serogroups: definition and geographic distribution. *In* "The Arboviruses: epidemiology and ecology" (T. P. Monath, Ed.) Vol.1, p.19. CRC Press, Boca Raton, Fl.

Caplen, N. J., Zheng, Z., Falgout, B., and Morgan, R. A. (2002). Inhibition of viral gene expression and replication in mosquito cells by dsRNA-triggered RNA interference. *Mol Ther* **6,** 243-51.

Carmell, M. A., Xuan, Z., Zhang, M. Q., and Hannon, G. J. (2002). The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. *Genes Dev* **16,** 2733-42.

Catalanotto, C., Pallotta, M., ReFalo, P., Sachs, M. S., Vayssie, L., Macino, G., and Cogoni, C. (2004). Redundancy of the two dicer genes in transgene-induced posttranscriptional gene silencing in *Neurospora crassa. Mol Cell Biol* **24,** 2536-45.

Caudy, A. A., Ketting, R. F., Hammond, S. M., Denli, A. M., Bathoorn, A. M., Tops, B. B., Silva, J. M., Myers, M. M., Hannon, G. J., and Plasterk, R. H. (2003). A micrococcal nuclease homologue in RNAi effector complexes. *Nature* **425,** 411-4.

Caudy, A. A., Myers, M., Hannon, G. J., and Hammond, S. M. (2002). Fragile X-related protein and VIG associate with the RNA interference machinery. *Genes Dev* **16,** 2491-6.

Chang, G. J., and Trent, D. W. (1987). Nucleotide sequence of the genome region encoding the 26S mRNA of eastern equine encephalomyelitis virus and the deduced amino acid sequence of the viral structural proteins. *J Gen Virol* **68 (Pt 8),** 2129-42.

Cheng, L. L., Bartholomay, L. C., Olson, K. E., Lowenberger, C., Vizioli, J., Higgs, S., Beaty, B. J., and Christensen, B. M. (2001). Characterization of an endogenous gene expressed in *Aedes aegypti* using an orally infectious recombinant Sindbis virus. *J Insect Sci* **1,** 10.

Choi, H. K., Tong, L., Minor, W., Dumas, P., Boege, U., Rossmann, M. G., and Wengler, G. (1991). Structure of Sindbis virus core protein reveals a chymotrypsin-like serine proteinase and the organization of the virion. *Nature* **354,** 37-43.

Chomczynski, P., and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* **162,** 156-9.

Christophides, G. K., Vlachou, D., and Kafatos, F. C. (2004). Comparative and functional genomics of the innate immune system in the malaria vector *Anopheles gambiae*. *Immunol Rev* **198,** 127-48.

Christophides, G. K., Zdobnov, E., Barillas-Mury, C., Birney, E., Blandin, S., Blass, C., Brey, P. T., Collins, F. H., Danielli, A., Dimopoulos, G., et al. (2002). Immunity-related genes and gene families in Anopheles gambiae. *Science* **298,** 159-65.

Clem, R. J. (2005). The role of apoptosis in defense against baculovirus infection in insects. *Curr Top Microbiol Immunol* **289,** 113-29.

Cogoni, C., and Macino, G. (1997). Isolation of quelling-defective (qde) mutants impaired in posttranscriptional transgene-induced gene silencing in *Neurospora crassa*. *Proc Natl Acad Sci U S A* **94,** 10233-8.

Coombs, K., and Brown, D. T. (1987). Topological organization of Sindbis virus capsid protein in isolated nucleocapsids. *Virus Res* **7,** 131-49.

Corbet, P. S., Williams, M. C., and Gillett, J. D. (1961). O'nyong-nyong fever: An epidemic virus disease in East Africa: vector studies at epidemic sites. *Trans R Soc Trop Med Hyg* **55,** 463-80.

Davis, N. L., Willis, L. V., Smith, J. F., and Johnston, R. E. (1989). *In vitro* synthesis of infectious venezuelan equine encephalitis virus RNA from a cDNA clone: analysis of a viable deletion mutant. *Virology* **171,** 189-204.

de Curtis, I., and Simons, K. (1988). Dissection of Semliki Forest virus glycoprotein delivery from the trans-Golgi network to the cell surface in permeabilized BHK cells. *Proc Natl Acad Sci U S A* **85,** 8052-6.

De Francesco, R., and Migliaccio, G. (2005). Challenges and successes in developing new therapies for hepatitis C. *Nature* **436,** 953-60.

de Groot, R. J., Hardy, W. R., Shirako, Y., and Strauss, J. H. (1990). Cleavage-site preferences of Sindbis virus polyproteins containing the non-structural proteinase. Evidence for temporal regulation of polyprotein processing in vivo. *Embo J* **9,** 2631-8.

de Lara Capurro, M., Coleman, J., Beerntsen, B. T., Myles, K. M., Olson, K. E., Rocha, E., Krettli, A. U., and James, A. A. (2000). Virus-expressed, recombinant single-chain antibody blocks sporozoite infection of salivary glands in *Plasmodium gallinaceum*-infected *Aedes aegypti*. *Am J Trop Med Hyg* **62,** 427-33.

Dimopoulos, G. (2003). Insect immunity and its implication in mosquito-malaria interactions. *Cell Microbiol* **5,** 3-14.

Dimopoulos, G., Christophides, G. K., Meister, S., Schultz, J., White, K. P., Barillas-Mury, C., and Kafatos, F. C. (2002). Genome expression analysis of *Anopheles gambiae*: responses to injury, bacterial challenge, and malaria infection. *Proc Natl Acad Sci U S A* **99,** 8814-9.

Dimopoulos, G., Richman, A., Muller, H. M., and Kafatos, F. C. (1997). Molecular immune responses of the mosquito *Anopheles gambiae* to bacteria and malaria parasites. *Proc Natl Acad Sci U S A* **94,** 11508-13.

Dimopoulos, G., Seeley, D., Wolf, A., and Kafatos, F. C. (1998). Malaria infection of the mosquito *Anopheles gambiae* activates immune-responsive genes during critical transition stages of the parasite life cycle. *Embo J* **17,** 6115-23.

Dudley, N. R., Labbe, J. C., and Goldstein, B. (2002). Using RNA interference to identify genes required for RNA interference. *Proc Natl Acad Sci U S A* **99,** 4191-6.

Elbashir, S. M., Lendeckel, W., and Tuschl, T. (2001). RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev* **15,** 188-200.

Faragher, S. G., Meek, A. D., Rice, C. M., and Dalgarno, L. (1988). Genome sequences of a mouse-avirulent and a mouse-virulent strain of Ross River virus. *Virology* **163,** 509-26.

Feng, Y. X., Copeland, T. D., Oroszlan, S., Rein, A., and Levin, J. G. (1990). Identification of amino acids inserted during suppression of UAA and UGA termination codons at the gag-pol junction of Moloney murine leukemia virus. *Proc Natl Acad Sci U S A* **87,** 8860-3.

Fire, A., Albertson, D., Harrison, S. W., and Moerman, D. G. (1991). Production of antisense RNA leads to effective and specific inhibition of gene expression in *C. elegans* muscle. *Development* **113,** 503-14.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391,** 806-11.

Foy, B. D., Myles, K. M., Pierro, D. J., Sanchez-Vargas, I., Uhlirova, M., Jindra, M., Beaty, B. J., and Olson, K. E. (2004). Development of a new Sindbis virus transducing system and its characterization in three Culicine mosquitoes and two Lepidopteran species. *Insect Mol Biol* **13,** 89-100.

Frolov, I., Hoffman, T. A., Pragai, B. M., Dryga, S. A., Huang, H. V., Schlesinger, S., and Rice, C. M. (1996). Alphavirus-based expression vectors: strategies and applications. *Proc Natl Acad Sci U S A* **93,** 11371-7.

Gaines, P. J., Olson, K. E., Higgs, S., Powers, A. M., Beaty, B. J., and Blair, C. D. (1996). Pathogen-derived resistance to dengue type 2 virus in mosquito cells by expression of the premembrane coding region of the viral genome. *J Virol* **70,** 2132-7.

Garoff, H., Frischauf, A. M., Simons, K., Lehrach, H., and Delius, H. (1980). Nucleotide sequence of cDNA coding for Semliki Forest virus membrane glycoproteins. *Nature* **288,** 236-41.

Geigenmuller-Gnirke, U., Weiss, B., Wright, R., and Schlesinger, S. (1991). Complementation between Sindbis viral RNAs produces infectious particles with a bipartite genome. *Proc Natl Acad Sci U S A* **88,** 3253-7.

Geiss, B. J., Pierson, T. C., and Diamond, M. S. (2005). Actively replicating West Nile virus is resistant to cytoplasmic delivery of siRNA. *Virol J* **2,** 53.

Goldbach, R., Bucher, E., and Prins, M. (2003). Resistance mechanisms to plant viruses: an overview. *Virus Res* **92,** 207-12.

Gorman, M. J., and Paskewitz, S. M. (2001). Serine proteases as mediators of mosquito immune responses. *Insect Biochem Mol Biol* **31,** 257-62.

Gottlieb, E., and Steitz, J. A. (1989). Function of the mammalian La protein: evidence for its action in transcription termination by RNA polymerase III. *Embo J* **8,** 851-61.

Grishok, A. (2005). RNAi mechanisms in *Caenorhabditis elegans*. *FEBS Lett* **579,** 5932-9.

Grishok, A., Pasquinelli, A. E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D. L., Fire, A., Ruvkun, G., and Mello, C. C. (2001). Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. *Cell* **106,** 23-34.

Guo, S., and Kemphues, K. J. (1995). par-1, a gene required for establishing polarity in C. elegans embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed. *Cell* **81,** 611-20.

Hahn, Y. S., Strauss, E. G., and Strauss, J. H. (1989). Mapping of RNA- temperature-sensitive mutants of Sindbis virus: assignment of complementation groups A, B, and G to nonstructural proteins. *J Virol* **63,** 3142-50.

Hammond, S. M., Bernstein, E., Beach, D., and Hannon, G. J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* **404,** 293-6.

Hammond, S. M., Boettcher, S., Caudy, A. A., Kobayashi, R., and Hannon, G. J. (2001). Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science* **293,** 1146-50.

Hansen, I. A., Attardo, G. M., Park, J. H., Peng, Q., and Raikhel, A. S. (2004). Target of rapamycin-mediated amino acid signaling in mosquito anautogeny. *Proc Natl Acad Sci U S A* **101,** 10626-31.

Heidner, H. W., Knott, T. A., and Johnston, R. E. (1996). Differential processing of sindbis virus glycoprotein PE2 in cultured vertebrate and arthropod cells. J Virol *70*, 2069-2073.

Hetru, C., Troxler, L., and Hoffmann, J. A. (2003). Drosophila melanogaster antimicrobial defense. *J Infect Dis* **187 Suppl 2,** S327-34.

Himber, C., Dunoyer, P., Moissiard, G., Ritzenthaler, C., and Voinnet, O. (2003). Transitivity-dependent and -independent cell-to-cell movement of RNA silencing. *Embo J* **22,** 4523-33.

Hoa, N. T., Keene, K. M., Olson, K. E., and Zheng, L. (2003). Characterization of RNA interference in an *Anopheles gambiae* cell line. *Insect Biochem Mol Biol* **33,** 949-57.

Hoffmann, J. A., Kafatos, F. C., Janeway, C. A., and Ezekowitz, R. A. (1999). Phylogenetic perspectives in innate immunity. *Science* **284,** 1313-8.

Holt, R. A., Subramanian, G. M., Halpern, A., Sutton, G. G., Charlab, R., Nusskern, D. R., Wincker, P., Clark, A. G., Ribeiro, J. M., Wides, R., et al. (2002). The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science* **298,** 129-49.

Hommel, G. A., and Schloemer, R. H. (1985). Site of suppression of Banzi viral replication by an antiviral factor released from *Aedes albopictus* cells persistently infected with Banzi virus. Virus Res *4*, 37-51.

Huang, C. Y., Chou, S. Y., Bartholomay, L. C., Christensen, B. M., and Chen, C. C. (2005). The use of gene silencing to study the role of dopa decarboxylase in mosquito melanization reactions. *Insect Mol Biol* **14,** 237-44.

Hutvagner, G., McLachlan, J., Pasquinelli, A. E., Balint, E., Tuschl, T., and Zamore, P. D. (2001). A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* **293,** 834-8.

Infanger, L. C., Rocheleau, T. A., Bartholomay, L. C., Johnson, J. K., Fuchs, J., Higgs, S., Chen, C. C., and Christensen, B. M. (2004). The role of phenylalanine hydroxylase in melanotic encapsulation of filarial worms in two species of mosquitoes. *Insect Biochem Mol Biol* **34,** 1329-38.

Ishizuka, A., Siomi, M. C., and Siomi, H. (2002). A *Drosophila* fragile X protein interacts with components of RNAi and ribosomal proteins. *Genes Dev* **16,** 2497-508.

Jacque, J. M., Triques, K., and Stevenson, M. (2002). Modulation of HIV-1 replication by RNA interference. *Nature* **418,** 435-8.

Jiang, W., Venugopal, K., and Gould, E. A. (1995). Intracellular interference of tick-borne flavivirus infection by using a single-chain antibody fragment delivered by recombinant Sindbis virus. *J Virol* **69,** 1044-9.

Johnson, B. W., Olson, K. E., Allen-Miura, T., Rayms-Keller, A., Carlson, J. O., Coates, C. J., Jasinskiene, N., James, A. A., Beaty, B. J., and Higgs, S. (1999). Inhibition of luciferase expression in transgenic *Aedes aegypti* mosquitoes by Sindbis virus expression of antisense luciferase RNA. *Proc Natl Acad Sci U S A* **96,** 13399-403.

Kadotani, N., Nakayashiki, H., Tosa, Y., and Mayama, S. (2004). One of the two Dicer-like proteins in the filamentous fungi *Magnaporthe oryzae* genome is responsible for hairpin RNA-triggered RNA silencing and related small interfering RNA accumulation. *J Biol Chem* **279,** 44467-74.

Kamath, R. S., and Ahringer, J. (2003). Genome-wide RNAi screening in *Caenorhabditis elegans*. *Methods* **30,** 313-21.

Kamrud, K. I., Olson, K. E., Higgs, S., Powers, A. M., Carlson, J. O., and Beaty, B. J. (1997). Detection of expressed chloramphenicol acetyltransferase in the saliva of *Culex pipiens* mosquitoes. *Insect Biochem Mol Biol* **27,** 423-9.

Karabatsos, N. (1975). Antigenic relationships of group A arboviruses by plaque reduction neutralization testing. *Am J Trop Med Hyg* **24,** 527-32.

Karbatsos, N., Ed. (1985). International Catalogue of Arboviruses, including certain other viruses of vertebrates. 3rd ed. San Antonio, TX: American Society of Tropical Medicine and Hygiene.

Kasschau, K. D., and Carrington, J. C. (1998). A counterdefensive strategy of plant viruses: suppression of posttranscriptional gene silencing. *Cell* **95,** 461-70.

Kennedy, S., Wang, D., and Ruvkun, G. (2004). A conserved siRNA-degrading RNase negatively regulates RNA interference in *C. elegans*. *Nature* **427,** 645-9.

Kinney, R. M., Johnson, B. J., Welch, J. B., Tsuchiya, K. R., and Trent, D. W. (1989). The full-length nucleotide sequences of the virulent Trinidad donkey strain of Venezuelan equine encephalitis virus and its attenuated vaccine derivative, strain TC-83. *Virology* **170,** 19-30.

Kiwanuka, N., Sanders, E. J., Rwaguma, E. B., Kawamata, J., Ssengooba, F. P., Najjemba, R., Were, W. A., Lamunu, M., Bagambisa, G., Burkot, T. R., Dunster, L., Lutwama, J. J., Martin, D. A., Cropp, C. B., Karabatsos, N., Lanciotti, R. S., Tsai, T. F., and Campbell, G. L. (1999). O'nyong-nyong fever in south-central Uganda, 1996-1997: clinical features and validation of a clinical case definition for surveillance purposes. *Clin Infect Dis* **29,** 1243-50.

Kuhn, R. J., Niesters, H. G., Hong, Z., and Strauss, J. H. (1991). Infectious RNA transcripts from Ross River virus cDNA clones and the construction and characterization of defined chimeras with Sindbis virus. *Virology* **182,** 430-41.

Kumar, M., and Carmichael, G. G. (1998). Antisense RNA: function and fate of duplex RNA in cells of higher eukaryotes. *Microbiol Mol Biol Rev* **62,** 1415-34.

Kumar, S., Christophides, G. K., Cantera, R., Charles, B., Han, Y. S., Meister, S., Dimopoulos, G., Kafatos, F. C., and Barillas-Mury, C. (2003). The role of reactive oxygen species on *Plasmodium* melanotic encapsulation in *Anopheles gambiae*. *Proc Natl Acad Sci U S A* **100,** 14139-44.

Lakatos, L., Szittya, G., Silhavy, D., and Burgyan, J. (2004). Molecular mechanism of RNA silencing suppression mediated by p19 protein of tombusviruses. *Embo J* **23,** 876-84.

Lanciotti, R. S., Ludwig, M. L., Rwaguma, E. B., Lutwama, J. J., Kram, T. M., Karabatsos, N., Cropp, B. C., and Miller, B. R. (1998). Emergence of epidemic O'nyong-nyong fever in Uganda after a 35-year absence: genetic characterization of the virus. *Virology* **252,** 258-68.

LaStarza, M. W., Lemm, J. A., and Rice, C. M. (1994). Genetic analysis of the nsP3 region of Sindbis virus: evidence for roles in minus-strand and subgenomic RNA synthesis. *J Virol* **68,** 5781-91.

Lee, E., Stocks, C., Lobigs, P., Hislop, A., Straub, J., Marshall, I., Weir, R., and Dalgarno, L. (1997). Nucleotide sequence of the Barmah Forest virus genome. *Virology* **227,** 509-14.

Lee, N. S., Dohjima, T., Bauer, G., Li, H., Li, M. J., Ehsani, A., Salvaterra, P., and Rossi, J. (2002). Expression of small interfering RNAs targeted against HIV-1 rev transcripts in human cells. *Nat Biotechnol* **20,** 500-5.

Lee, R. C., Feinbaum, R. L., and Ambros, V. (1993). The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75,** 843-54.

Lee, Y. S., Nakahara, K., Pham, J. W., Kim, K., He, Z., Sontheimer, E. J., and Carthew, R. W. (2004). Distinct roles for *Drosophila* Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. *Cell* **117,** 69-81.

Lehane, M. J., Wu, D., and Lehane, S. M. (1997). Midgut-specific immune molecules are produced by the blood-sucking insect *Stomoxys calcitrans*. *Proc Natl Acad Sci U S A* **94,** 11502-7.

Levashina, E. A. (2004). Immune responses in *Anopheles gambiae*. *Insect Biochem Mol Biol* **34,** 673-8.

Levashina, E. A., Moita, L. F., Blandin, S., Vriend, G., Lagueux, M., and Kafatos, F. C. (2001). Conserved role of a complement-like protein in phagocytosis revealed by dsRNA knockout in cultured cells of the mosquito, *Anopheles gambiae*. *Cell* **104,** 709-18.

Levinson, R. S., Strauss, J. H., and Strauss, E. G. (1990). Complete sequence of the genomic RNA of O'nyong-nyong virus and its use in the construction of alphavirus phylogenetic trees. *Virology* **175,** 110-23.

Li, H., Li, W. X., and Ding, S. W. (2002). Induction and suppression of RNA silencing by an animal virus. *Science* **296,** 1319-21.

Li, W. X., Li, H., Lu, R., Li, F., Dus, M., Atkinson, P., Brydon, E. W., Johnson, K. L., Garcia-Sastre, A., Ball, L. A., Palese, P., and Ding, S. W. (2004). Interferon antagonist proteins of influenza and vaccinia viruses are suppressors of RNA silencing. *Proc Natl Acad Sci U S A* **101,** 1350-5.

Liljestrom, P., Lusa, S., Huylebroeck, D., and Garoff, H. (1991). In vitro mutagenesis of a full-length cDNA clone of Semliki Forest virus: the small 6,000-molecular-weight membrane protein modulates virus release. *J Virol* **65,** 4107-13.

Lin, S. Y., Johnson, S. M., Abraham, M., Vella, M. C., Pasquinelli, A., Gamberi, C., Gottlieb, E., and Slack, F. J. (2003). The *C. elegans* hunchback homolog, hbl-1, controls temporal patterning and is a probable microRNA target. *Dev Cell* **4,** 639-50.

Lindbo, J. A., Silva-Rosales, L., Proebsting, W. M., and Dougherty, W. G. (1993). Induction of a Highly Specific Antiviral State in Transgenic Plants: Implications for Regulation of Gene Expression and Virus Resistance. *Plant Cell* **5,** 1749-1759.

Lingel, A., Simon, B., Izaurralde, E., and Sattler, M. (2003). Structure and nucleic-acid binding of the *Drosophila* Argonaute 2 PAZ domain. *Nature* **426,** 465-9.

Lingel, A., Simon, B., Izaurralde, E., and Sattler, M. (2004). Nucleic acid 3'-end recognition by the Argonaute2 PAZ domain. *Nat Struct Mol Biol* **11,** 576-7.

Liu, J., Carmell, M. A., Rivas, F. V., Marsden, C. G., Thomson, J. M., Song, J. J., Hammond, S. M., Joshua-Tor, L., and Hannon, G. J. (2004). Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **305,** 1437-41.

Liu, Q., Rand, T. A., Kalidas, S., Du, F., Kim, H. E., Smith, D. P., and Wang, X. (2003). R2D2, a bridge between the initiation and effector steps of the *Drosophila* RNAi pathway. *Science* **301,** 1921-5.

Llave, C., Kasschau, K. D., and Carrington, J. C. (2000). Virus-encoded suppressor of posttranscriptional gene silencing targets a maintenance step in the silencing pathway. *Proc Natl Acad Sci U S A* **97,** 13401-6.

Llave, C., Xie, Z., Kasschau, K. D., and Carrington, J. C. (2002). Cleavage of Scarecrow-like mRNA targets directed by a class of *Arabidopsis* miRNA. *Science* **297,** 2053-6.

Lowenberger, C. (2001). Innate immune response of *Aedes aegypti*. *Insect Biochem Mol Biol* **31,** 219-29.

Lu, R., Maduro, M., Li, F., Li, H. W., Broitman-Maduro, G., Li, W. X., and Ding, S. W. (2005). Animal virus replication and RNAi-mediated antiviral silencing in *Caenorhabditis elegans*. *Nature* **436,** 1040-3.

Luo, T., and Brown, D. T. (1993). Purification and characterization of a Sindbis virus-induced peptide which stimulates its own production and blocks virus RNA synthesis. Virology *194*, 44-49.

Lutwama, J. J., Kayondo, J., Savage, H. M., Burkot, T. R., and Miller, B. R. (1999). Epidemic O'Nyong-Nyong fever in southcentral Uganda, 1996-1997: entomologic studies in Bbaale village, Rakai District. *Am J Trop Med Hyg* **61,** 158-62.

Ma, J. B., Ye, K., and Patel, D. J. (2004). Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature* **429,** 318-22.

Mackenzie, J. S., Lindsay, M. D., Coelen, R. J., Broom, A. K., Hall, R. A., and Smith, D. W. (1994). Arboviruses causing human disease in the Australasian zoogeographic region. *Arch Virol* **136,** 447-67.

Mallory, A. C., Ely, L., Smith, T. H., Marathe, R., Anandalakshmi, R., Fagard, M., Vaucheret, H., Pruss, G., Bowman, L., and Vance, V. B. (2001). HC-Pro suppression of transgene silencing eliminates the small RNAs but not transgene methylation or the mobile signal. *Plant Cell* **13,** 571-83.

McCaffrey, A. P., Meuse, L., Pham, T. T., Conklin, D. S., Hannon, G. J., and Kay, M. A. (2002). RNA interference in adult mice. *Nature* **418,** 38-9.

Meister, G., and Tuschl, T. (2004). Mechanisms of gene silencing by double-stranded RNA. *Nature* **431,** 343-9.

Meister, S., Kanzok, S. M., Zheng, X. L., Luna, C., Li, T. R., Hoa, N. T., Clayton, J. R., White, K. P., Kafatos, F. C., Christophides, G. K., and Zheng, L. (2005). Immune signaling pathways regulating bacterial and malaria parasite infection of the mosquito *Anopheles gambiae*. *Proc Natl Acad Sci U S A* **102,** 11420-5.

Mi, S., Durbin, R., Huang, H. V., Rice, C. M., and Stollar, V. (1989). Association of the Sindbis virus RNA methyltransferase activity with the nonstructural protein nsP1. *Virology* **170,** 385-91.

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

Michel, K., and Kafatos, F. C. (2005). Mosquito immunity against *Plasmodium*. *Insect Biochem Mol Biol* **35,** 677-89.

Miller, M. L., and Brown, D. T. (1992). Morphogenesis of Sindbis virus in three subclones of *Aedes albopictus* (mosquito) cells. *J Virol* **66,** 4180-90.

Moita, L. F., Wang-Sattler, R., Michel, K., Zimmermann, T., Blandin, S., Levashina, E. A., and Kafatos, F. C. (2005). In vivo identification of novel regulators and conserved pathways of phagocytosis in *A. gambiae*. *Immunity* **23,** 65-73.

Myers, J. W., Jones, J. T., Meyer, T., and Ferrell, J. E., Jr. (2003). Recombinant Dicer efficiently converts large dsRNAs into siRNAs suitable for gene silencing. *Nat Biotechnol* **21,** 324-8.

Myles, K. M., Pierro, D. J., and Olson, K. E. (2003). Deletions in the putative cell receptor-binding domain of Sindbis virus strain MRE16 E2 glycoprotein reduce midgut infectivity in *Aedes aegypti*. *J Virol* **77,** 8872-81.

Myles, K. M., Pierro, D. J., and Olson, K. E. (2004). Comparison of the transmission potential of two genetically distinct Sindbis viruses after oral infection of *Aedes aegypti* (Diptera: Culicidae). *J Med Entomol* **41,** 95-106.

Myles, K. M. (2003). Dissertation. Colorado State University, Fort Collins, CO.

Naim, H. Y., and Koblet, H. (1990). The cleavage of p62, the precursor of E2 and E3, is an early and continuous event in Semliki Forest virus-infected *Aedes albopictus* cells. *Arch Virol* **110,** 221-37.

Nakazawa, H., Tsuneishi, E., Ponnuvel, K. M., Furukawa, S., Asaoka, A., Tanaka, H., Ishibashi, J., and Yamakawa, M. (2004). Antiviral activity of a serine protease from the digestive juice of *Bombyx mori* larvae against nucleopolyhedrovirus. *Virology* **321,** 154-62.

Napoli, C., Lemieux, C., and Jorgensen, R. (1990). Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. *Plant Cell* **2,** 279-289.

Newton, S. E., and Dalgarno, L. (1983). Antiviral activity released from *Aedes albopictus* cells persistently infected with Semliki forest virus. J Virol *47*, 652-655.

Ngo, H., Tschudi, C., Gull, K., and Ullu, E. (1998). Double-stranded RNA induces mRNA degradation in *Trypanosoma brucei*. *Proc Natl Acad Sci U S A* **95,** 14687-92.

Nicholson, R. H., and Nicholson, A. W. (2002). Molecular characterization of a mouse cDNA encoding Dicer, a ribonuclease III ortholog involved in RNA interference. *Mamm Genome* **13,** 67-73.

Nykanen, A., Haley, B., and Zamore, P. D. (2001). ATP requirements and small interfering RNA structure in the RNA interference pathway. *Cell* **107,** 309-21.

Okamura, K., Ishizuka, A., Siomi, H., and Siomi, M. C. (2004). Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. *Genes Dev* **18,** 1655-66.

Olson, K., and Trent, D. W. (1985). Genetic and antigenic variations among geographical isolates of Sindbis virus. *J Gen Virol* **66 (Pt 4),** 797-810.

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

Olson, K. E., Higgs, S., Hahn, C. S., Rice, C. M., Carlson, J. O., and Beaty, B. J. (1994). The expression of chloramphenicol acetyltransferase in *Aedes albopictus* (C6/36) cells and Aedes triseriatus mosquitoes using a double subgenomic recombinant Sindbis virus. *Insect Biochem Mol Biol* **24,** 39-48.

Olson, K. E., Myles, K. M., Seabaugh, R. C., Higgs, S., Carlson, J. O., and Beaty, B. J. (2000). Development of a Sindbis virus expression system that efficiently expresses green fluorescent protein in midguts of *Aedes aegypti* following per os infection. *Insect Mol Biol* **9,** 57-65.

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

Palatnik, J. F., Allen, E., Wu, X., Schommer, C., Schwab, R., Carrington, J. C., and Weigel, D. (2003). Control of leaf morphogenesis by microRNAs. *Nature* **425,** 257-63.

Palauqui, J. C., Elmayan, T., Pollien, J. M., and Vaucheret, H. (1997). Systemic acquired silencing: transgene-specific post-transcriptional silencing is transmitted by grafting from silenced stocks to non-silenced scions. *Embo J* **16,** 4738-45.

Pardigon, N., and Strauss, J. H. (1996). Mosquito homolog of the La autoantigen binds to Sindbis virus RNA. *J Virol* **70,** 1173-81.

Paterson, A., Robinson, E., Suchman, E., Afanasiev, B., and Carlson, J. (2005). Mosquito densonucleosis viruses cause dramatically different infection phenotypes in the C6/36 *Aedes albopictus* cell line. *Virology* **337,** 253-61.

Pham, J. W., Pellino, J. L., Lee, Y. S., Carthew, R. W., and Sontheimer, E. J. (2004). A Dicer-2-dependent 80s complex cleaves targeted mRNAs during RNAi in *Drosophila*. *Cell* **117,** 83-94.

Pierro, D. J., Myles, K. M., Foy, B. D., Beaty, B. J., and Olson, K. E. (2003). Development of an orally infectious Sindbis virus transducing system that efficiently disseminates and expresses green fluorescent protein in *Aedes aegypti*. *Insect Mol Biol* **12,** 107-16.

Ponnuvel, K. M., Nakazawa, H., Furukawa, S., Asaoka, A., Ishibashi, J., Tanaka, H., and Yamakawa, M. (2003). A lipase isolated from the silkworm *Bombyx mori* shows antiviral activity against nucleopolyhedrovirus. *J Virol* **77,** 10725-9.

Posey, D. L., O'Rourke, T., Roehrig, J. T., Lanciotti, R. S., Weinberg, M., and Maloney, S. (2005). O'Nyong-nyong fever in West Africa. *Am J Trop Med Hyg* **73,** 32.

Powers, A. M., Brault, A. C., Shirako, Y., Strauss, E. G., Kang, W., Strauss, J. H., and Weaver, S. C. (2001). Evolutionary relationships and systematics of the alphaviruses. J Virol *75*, 10118-10131.

Powers, A. M., Brault, A. C., Tesh, R. B., and Weaver, S. C. (2000). Re-emergence of Chikungunya and O'nyong-nyong viruses: evidence for distinct geographical lineages and distant evolutionary relationships. *J Gen Virol* **81,** 471-9.

Powers, A. M., Kamrud, K. I., Olson, K. E., Higgs, S., Carlson, J. O., and Beaty, B. J. (1996). Molecularly engineered resistance to California serogroup virus replication in mosquito cells and mosquitoes. *Proc Natl Acad Sci U S A* **93,** 4187-91.

Powers, A. M., Olson, K. E., Higgs, S., Carlson, J. O., and Beaty, B. J. (1994). Intracellular immunization of mosquito cells to LaCrosse virus using a recombinant Sindbis virus vector. *Virus Res* **32,** 57-67.

Provost, P., Dishart, D., Doucet, J., Frendewey, D., Samuelsson, B., and Radmark, O. (2002a). Ribonuclease activity and RNA binding of recombinant human Dicer. *Embo J* **21,** 5864-74.

Provost, P., Silverstein, R. A., Dishart, D., Walfridsson, J., Djupedal, I., Kniola, B., Wright, A., Samuelsson, B., Radmark, O., and Ekwall, K. (2002b). Dicer is required for chromosome segregation and gene silencing in fission yeast cells. *Proc Natl Acad Sci U S A* **99,** 16648-53.

Racaniello, V. R., and Baltimore, D. (1981). Cloned poliovirus complementary DNA is infectious in mammalian cells. *Science* **214,** 916-9.

Raghow, R. S., Davey, M. W., and Dalgarno, L. (1973a). The growth of Semliki Forest virus in cultured mosquito cells: ultrastructural observations. *Arch Gesamte Virusforsch* **43,** 165-8.

Raghow, R. S., Grace, T. D., Filshie, B. K., Bartley, W., and Dalgarno, L. (1973b). Ross River virus replication in cultured mosquito and mammalian cells: virus growth and correlated ultrastructural changes. *J Gen Virol* **21,** 109-22.

Rand, T. A., Ginalski, K., Grishin, N. V., and Wang, X. (2004). Biochemical identification of Argonaute 2 as the sole protein required for RNA-induced silencing complex activity. *Proc Natl Acad Sci U S A* **101,** 14385-9.

Randall, G., Grakoui, A., and Rice, C. M. (2003). Clearance of replicating hepatitis C virus replicon RNAs in cell culture by small interfering RNAs. *Proc Natl Acad Sci U S A* **100,** 235-40.

Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., Horvitz, H. R., and Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature* **403,** 901-6.

Rice, C. M. (1996). Alphavirus-based expression systems. *Adv Exp Med Biol* **397,** 31-40.

Rice, C. M., Levis, R., Strauss, J. H., and Huang, H. V. (1987). Production of infectious RNA transcripts from Sindbis virus cDNA clones: mapping of lethal mutations, rescue of a temperature-sensitive marker, and in vitro mutagenesis to generate defined mutants. *J Virol* **61,** 3809-19.

Rice, C. M., and Strauss, J. H. (1982). Association of sindbis virion glycoproteins and their precursors. *J Mol Biol* **154,** 325-48.

Richman, A. M., Bulet, P., Hetru, C., Barillas-Mury, C., Hoffmann, J. A., and Kafalos, F. C. (1996). Inducible immune factors of the vector mosquito *Anopheles gambiae*: biochemical purification of a defensin antibacterial peptide and molecular cloning of preprodefensin cDNA. *Insect Mol Biol* **5,** 203-10.

Richman, A. M., Dimopoulos, G., Seeley, D., and Kafatos, F. C. (1997). *Plasmodium* activates the innate immune response of *Anopheles gambiae* mosquitoes. *Embo J* **16,** 6114-9.

Riedel, B., and Brown, D. T. (1979). Novel antiviral activity found in the media of Sindbis virus-persistently infected mosquito (Aedes albopictus) cell cultures. J Virol *29*, 51-60.

Robalino, J., Bartlett, T., Shepard, E., Prior, S., Jaramillo, G., Scura, E., Chapman, R. W., Gross, P. S., Browdy, C. L., and Warr, G. W. (2005). Double-stranded RNA induces sequence-specific antiviral silencing in addition to nonspecific immunity in a marine shrimp: convergence of RNA interference and innate immunity in the invertebrate antiviral response? *J Virol* **79,** 13561-71.

Robalino, J., Browdy, C. L., Prior, S., Metz, A., Parnell, P., Gross, P., and Warr, G. (2004). Induction of antiviral immunity by double-stranded RNA in a marine invertebrate. *J Virol* **78,** 10442-8.

Roth, B. M., Pruss, G. J., and Vance, V. B. (2004). Plant viral suppressors of RNA silencing. *Virus Res* **102,** 97-108.

Sanchez-Vargas, I., Travanty, E. A., Keene, K. M., Franz, A. W., Beaty, B. J., Blair, C. D., and Olson, K. E. (2004). RNA interference, arthropod-borne viruses, and mosquitoes. *Virus Res* **102,** 65-74.

Sawicki, D. L., and Sawicki, S. G. (1994). Alphavirus positive and negative strand RNA synthesis and the role of polyproteins in formation of viral replication complexes. *Arch Virol Suppl* **9,** 393-405.

Sawicki, S. G., Sawicki, D. L., Kaariainen, L., and Keranen, S. (1981). A Sindbis virus mutant temperature-sensitive in the regulation of minus-strand RNA synthesis. *Virology* **115,** 161-72.

Schauer, S. E., Jacobsen, S. E., Meinke, D. W., and Ray, A. (2002). DICER-LIKE1: blind men and elephants in *Arabidopsis* development. *Trends Plant Sci* **7,** 487-91.

Schwarz, D. S., Hutvagner, G., Haley, B., and Zamore, P. D. (2002). Evidence that siRNAs function as guides, not primers, in the *Drosophila* and human RNAi pathways. *Mol Cell* **10,** 537-48.

Scott, T. W., and Burrage, T. G. (1984). Rapid infection of salivary glands in *Culiseta melanura* with eastern equine encephalitis virus: an electron microscopic study. *Am J Trop Med Hyg* **33,** 961-4.

Seabaugh, R. C., Olson, K. E., Higgs, S., Carlson, J. O., and Beaty, B. J. (1998). Development of a chimeric sindbis virus with enhanced per os infection of *Aedes aegypti*. *Virology* **243,** 99-112.

Sen, G. L., and Blau, H. M. (2005). Argonaute 2/RISC resides in sites of mammalian mRNA decay known as cytoplasmic bodies. *Nat Cell Biol* **7,** 633-6.

Shiao, S. H., Higgs, S., Adelman, Z., Christensen, B. M., Liu, S. H., and Chen, C. C. (2001). Effect of prophenoloxidase expression knockout on the melanization of microfilariae in the mosquito *Armigeres subalbatus*. *Insect Mol Biol* **10,** 315-21.

Shirako, Y., and Strauss, J. H. (1994). Regulation of Sindbis virus RNA replication: uncleaved P123 and nsP4 function in minus-strand RNA synthesis, whereas cleaved products from P123 are required for efficient plus-strand RNA synthesis. *J Virol* **68,** 1874-85.

Sijen, T., Fleenor, J., Simmer, F., Thijssen, K. L., Parrish, S., Timmons, L., Plasterk, R. H., and Fire, A. (2001). On the role of RNA amplification in dsRNA-triggered gene silencing. *Cell* **107,** 465-76.

Sijen, T., and Plasterk, R. H. (2003). Transposon silencing in the *Caenorhabditis elegans* germ line by natural RNAi. *Nature* **426,** 310-4.

Silhavy, D., Molnar, A., Lucioli, A., Szittya, G., Hornyik, C., Tavazza, M., and Burgyan, J. (2002). A viral protein suppresses RNA silencing and binds silencing-generated, 21- to 25-nucleotide double-stranded RNAs. *Embo J* **21,** 3070-80.

Skogh, M., and Espmark, A. (1982). Ockelbo disease: epidemic arthritis-exanthema syndrome in Sweden caused by Sindbis-virus like agent. *Lancet* **1,** 795-6.

Song, J. J., Liu, J., Tolia, N. H., Schneiderman, J., Smith, S. K., Martienssen, R. A., Hannon, G. J., and Joshua-Tor, L. (2003). The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. *Nat Struct Biol* **10,** 1026-32.

Strauss, E. G., Rice, C. M., and Strauss, J. H. (1984). Complete nucleotide sequence of the genomic RNA of Sindbis virus. *Virology* **133,** 92-110.

Strauss, J. H., and Strauss, E. G. (1994). The alphaviruses: gene expression, replication, and evolution. *Microbiol Rev* **58,** 491-562.

Suopanki, J., Sawicki, D. L., Sawicki, S. G., and Kaariainen, L. (1998). Regulation of alphavirus 26S mRNA transcription by replicase component nsP2. *J Gen Virol* **79 (Pt 2),** 309-19.

Tabara, H., Sarkissian, M., Kelly, W. G., Fleenor, J., Grishok, A., Timmons, L., Fire, A., and Mello, C. C. (1999). The rde-1 gene, RNA interference, and transposon silencing in C. elegans. *Cell* **99,** 123-32.

Tahar, R., Boudin, C., Thiery, I., and Bourgouin, C. (2002). Immune response of *Anopheles gambiae* to the early sporogonic stages of the human malaria parasite *Plasmodium falciparum*. *Embo J* **21,** 6673-80.

Tahbaz, N., Kolb, F. A., Zhang, H., Jaronczyk, K., Filipowicz, W., and Hobman, T. C. (2004). Characterization of the interactions between mammalian PAZ PIWI domain proteins and Dicer. *EMBO Rep* **5,** 189-94.

Takkinen, K. (1986). Complete nucleotide sequence of the nonstructural protein genes of Semliki Forest virus. *Nucleic Acids Res* **14,** 5667-82.

Tamang, D., Tseng, S. M., Huang, C. Y., Tsao, I. Y., Chou, S. Z., Higgs, S., Christensen, B. M., and Chen, C. C. (2004). The use of a double subgenomic Sindbis virus expression system to study mosquito gene function: effects of antisense nucleotide number and duration of viral infection on gene silencing efficiency. *Insect Mol Biol* **13,** 595-602.

Taniguchi, T., Palmieri, M., and Weissmann, C. (1978). QB DNA-containing hybrid plasmids giving rise to QB phage formation in the bacterial host. *Nature* **274,** 223-8.

Taylor, R. M., Hurlbut, H. S., Work, T. H., Kingston, J. R., and Frothingham, T. E. (1955). Sindbis virus: a newly recognized arthropodtransmitted virus. *Am J Trop Med Hyg* **4,** 844-62.

Tijsterman, M., Ketting, R. F., Okihara, K. L., Sijen, T., and Plasterk, R. H. (2002). RNA helicase MUT-14-dependent gene silencing triggered in *C. elegans* by short antisense RNAs. *Science* **295,** 694-7.

Tijsterman, M., and Plasterk, R. H. (2004). Dicers at RISC; the mechanism of RNAi. *Cell* **117,** 1-3.

Travanty, E. A., Adelman, Z. N., Franz, A. W., Keene, K. M., Beaty, B. J., Blair, C. D., James, A. A., and Olson, K. E. (2004). Using RNA interference to develop dengue virus resistance in genetically modified *Aedes aegypti. Insect Biochem Mol Biol* **34,** 607-13.

Travanty, E. A. (2005). Dissertation. Colorado State University, Fort Collins, CO.

Tzou, P., Ohresser, S., Ferrandon, D., Capovilla, M., Reichhart, J. M., Lemaitre, B., Hoffmann, J. A., and Imler, J. L. (2000). Tissue-specific inducible expression of antimicrobial peptide genes in *Drosophila* surface epithelia. *Immunity* **13,** 737-48.

Uhlirova, M., Foy, B. D., Beaty, B. J., Olson, K. E., Riddiford, L. M., and Jindra, M. (2003). Use of Sindbis virus-mediated RNA interference to demonstrate a conserved role of Broad-Complex in insect metamorphosis. *Proc Natl Acad Sci U S A* **100,** 15607-12.

Vance, V., and Vaucheret, H. (2001). RNA silencing in plants--defense and counterdefense. *Science* **292,** 2277-80.

Vanlandingham, D. L., Hong, C., Klingler, K., Tsetsarkin, K., McElroy, K. L., Powers, A. M., Lehane, M. J., and Higgs, S. (2005). Differential infectivities of o'nyong-nyong and chikungunya virus isolates in *Anopheles gambiae* and *Aedes aegypti* mosquitoes. *Am J Trop Med Hyg* **72,** 616-21.

Vaucheret, H., Beclin, C., Elmayan, T., Feuerbach, F., Godon, C., Morel, J. B., Mourrain, P., Palauqui, J. C., and Vernhettes, S. (1998). Transgene-induced gene silencing in plants. *Plant J* **16,** 651-9.

Vaucheret, H., Vazquez, F., Crete, P., and Bartel, D. P. (2004). The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. *Genes Dev* **18,** 1187-97.

Vizioli, J., Bulet, P., Hoffmann, J. A., Kafatos, F. C., Muller, H. M., and Dimopoulos, G. (2001). Gambicin: a novel immune responsive antimicrobial peptide from the malaria vector Anopheles gambiae. *Proc Natl Acad Sci U S A* **98,** 12630-5.

Vlachou, D., Schlegelmilch, T., Christophides, G. K., and Kafatos, F. C. (2005). Functional genomic analysis of midgut epithelial responses in *Anopheles* during Plasmodium invasion. *Curr Biol* **15,** 1185-95.

Voinnet, O., and Baulcombe, D. C. (1997). Systemic signalling in gene silencing. *Nature* **389,** 553.

Voinnet, O., Lederer, C., and Baulcombe, D. C. (2000). A viral movement protein prevents spread of the gene silencing signal in *Nicotiana benthamiana*. *Cell* **103,** 157-67.

Voinnet, O., Pinto, Y. M., and Baulcombe, D. C. (1999). Suppression of gene silencing: a general strategy used by diverse DNA and RNA viruses of plants. *Proc Natl Acad Sci U S A* **96,** 14147-52.

Voinnet, O., Vain, P., Angell, S., and Baulcombe, D. C. (1998). Systemic spread of sequence-specific transgene RNA degradation in plants is initiated by localized introduction of ectopic promoterless DNA. *Cell* **95,** 177-87.

Wang, Q., Contag, C. H., Ilves, H., Johnston, B. H., and Kaspar, R. L. (2005a). Small hairpin RNAs efficiently inhibit hepatitis C IRES-mediated gene expression in human tissue culture cells and a mouse model. *Mol Ther* **12,** 562-8.

Wang, X., Fuchs, J. F., Infanger, L. C., Rocheleau, T. A., Hillyer, J. F., Chen, C. C., and Christensen, B. M. (2005b). Mosquito innate immunity: involvement of beta 1,3-glucan recognition protein in melanotic encapsulation immune responses in *Armigeres subalbatus*. *Mol Biochem Parasitol* **139,** 65-73.

Wang, Y. F., Sawicki, S. G., and Sawicki, D. L. (1991). Sindbis virus nsP1 functions in negative-strand RNA synthesis. *J Virol* **65,** 985-8.

Waterhouse, P. M., Wang, M. B., and Lough, T. (2001). Gene silencing as an adaptive defence against viruses. *Nature* **411,** 834-42.

Weaver, S. C. (1986). Electron microscopic analysis of infection patterns for Venezuelan equine encephalomyelitis virus in the vector mosquito, *Culex* (*Melanoconion*) *taeniopus*. *Am J Trop Med Hyg* **35,** 624-31.

Weaver, S. C., Bellew, L. A., and Rico-Hesse, R. (1992). Phylogenetic analysis of alphaviruses in the Venezuelan equine encephalitis complex and identification of the source of epizootic viruses. *Virology* **191,** 282-90.

Weaver, S. C., Scherer, W. F., Cupp, E. W., and Castello, D. A. (1984). Barriers to dissemination of Venezuelan encephalitis viruses in the Middle American enzootic vector mosquito, *Culex* (*Melanoconion*) *taeniopus*. *Am J Trop Med Hyg* **33,** 953-60.

Weaver, S. C., Anishchenko, M., Bowen, R., Brault, A. C., Estrada-Franco, J. G., Fernandez, Z., Greene, I., Ortiz, D., Paessler, S., and Powers, A. M. (2004). Genetic determinants of Venezuelan equine encephalitis emergence. Arch Virol Suppl, 43-64.

Weiss, B., Nitschko, H., Ghattas, I., Wright, R., and Schlesinger, S. (1989). Evidence for specificity in the encapsidation of Sindbis virus RNAs. *J Virol* **63,** 5310-8.

Wengler, G., and Filipe, A. R. (1977). A study of neucleotide sequence homology between the nucleic acids of different alphaviruses. *Virology* **78,** 124-34.

WHO (1996). The World Health Report 1996: fighting disease, fostering development. Geneva: World Health Organization.

Wienholds, E., and Plasterk, R. H. (2004). Target-selected gene inactivation in zebrafish. *Methods Cell Biol* **77,** 69-90.

Wilkins, C., Dishongh, R., Moore, S. C., Whitt, M. A., Chow, M., and Machaca, K. (2005). RNA interference is an antiviral defence mechanism in *Caenorhabditis elegans*. *Nature* **436,** 1044-7.

Williams, M. C., and Woodall, J. P. (1961). O'nyong nyong fever: An epidemic virus disease in East Africa: Isolation and some properties of the virus. *Trans R Soc Trop Med Hyg* **55,** 135-41.

Williams, M. C., and Woodall, J. P. (1965). O'nyong nyong fever: An epidemic virus disease in East Africa: Virus isolations from anopheles mosquitoes. *Trans R Soc Trop Med Hyg* **59,** 300-6.

Williams, M. C., Woodall, J. P., Corbet, P. S., and Gillett, J. D. (1965). O'nyong-Nyong Fever: An Epidemic Virus Disease in East Africa. 8. Virus Isolations from Anopheles Mosquitoes. *Trans R Soc Trop Med Hyg* **59,** 300-6.

Williams, R. W., and Rubin, G. M. (2002). ARGONAUTE1 is required for efficient RNA interference in Drosophila embryos. *Proc Natl Acad Sci U S A* **99,** 6889-94.

Yan, K. S., Yan, S., Farooq, A., Han, A., Zeng, L., and Zhou, M. M. (2003). Structure and conserved RNA binding of the PAZ domain. *Nature* **426,** 468-74.

Ye, K., Malinina, L., and Patel, D. J. (2003). Recognition of small interfering RNA by a viral suppressor of RNA silencing. *Nature* **426,** 874-8.

Zhang, H., Kolb, F. A., Jaskiewicz, L., Westhof, E., and Filipowicz, W. (2004). Single processing center models for human Dicer and bacterial RNase III. *Cell* **118,** 57-68.

Zheng, X. L., and Zheng, A. L. (2002). Genomic organization and regulation of three cecropin genes in *Anopheles gambiae*. *Insect Mol Biol* **11,** 517-25.

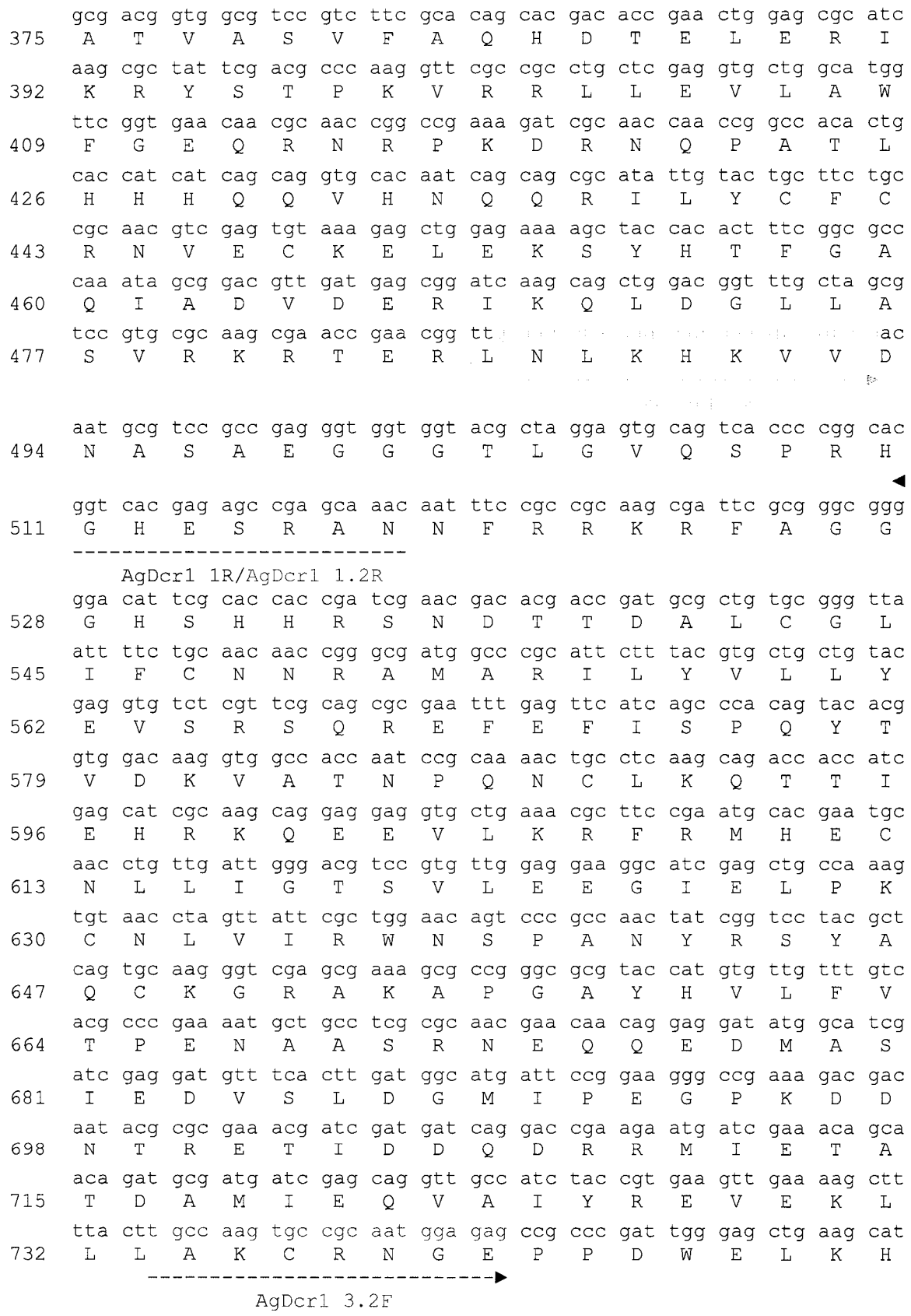
Zhu, J., Miura, K., Chen, L., and Raikhel, A. S. (2003). Cyclicity of mosquito vitellogenic ecdysteroid-mediated signaling is modulated by alternative dimerization of the RXR homologue Ultraspiracle. *Proc Natl Acad Sci U S A* **100,** 544-9.

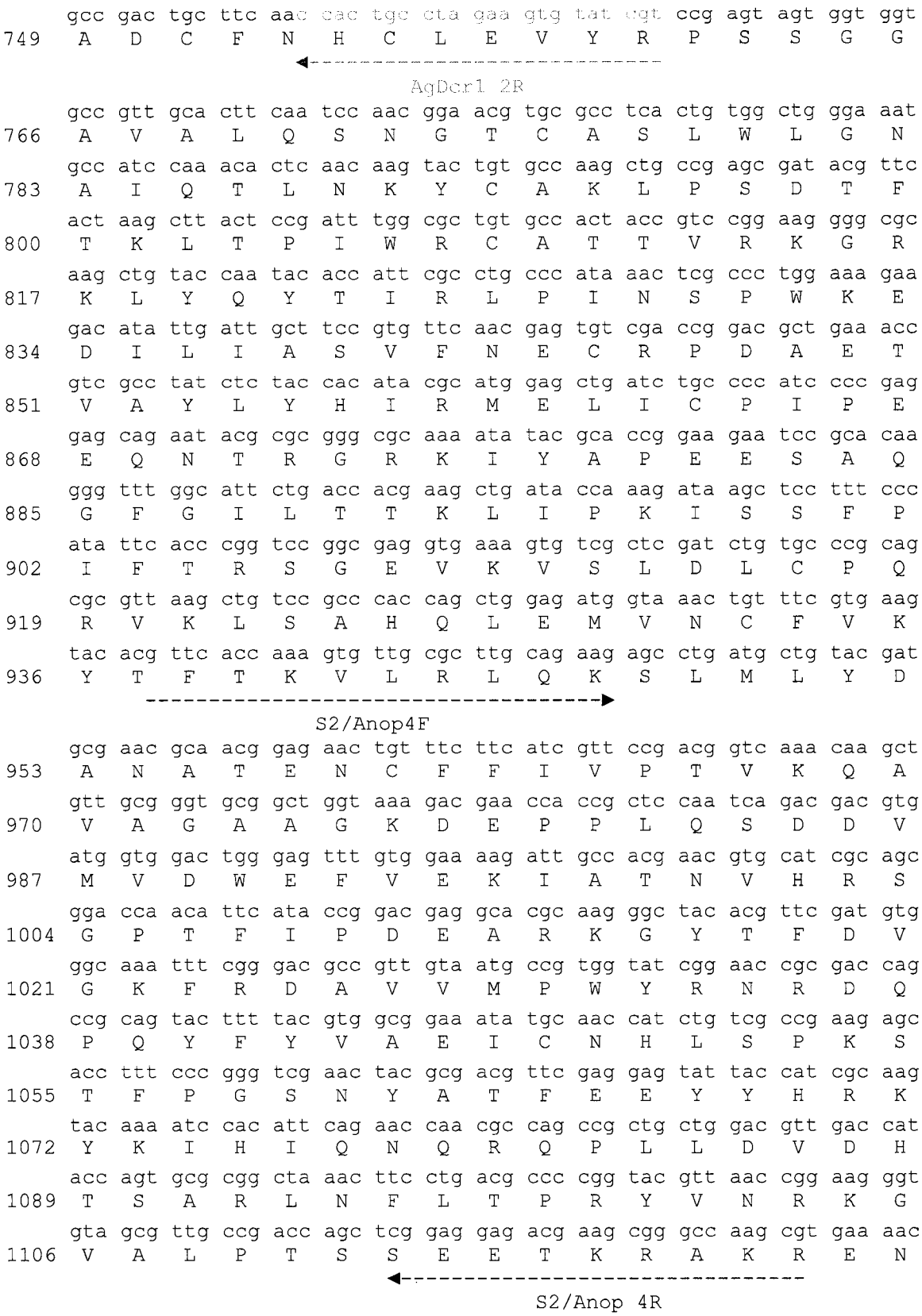
Appendices

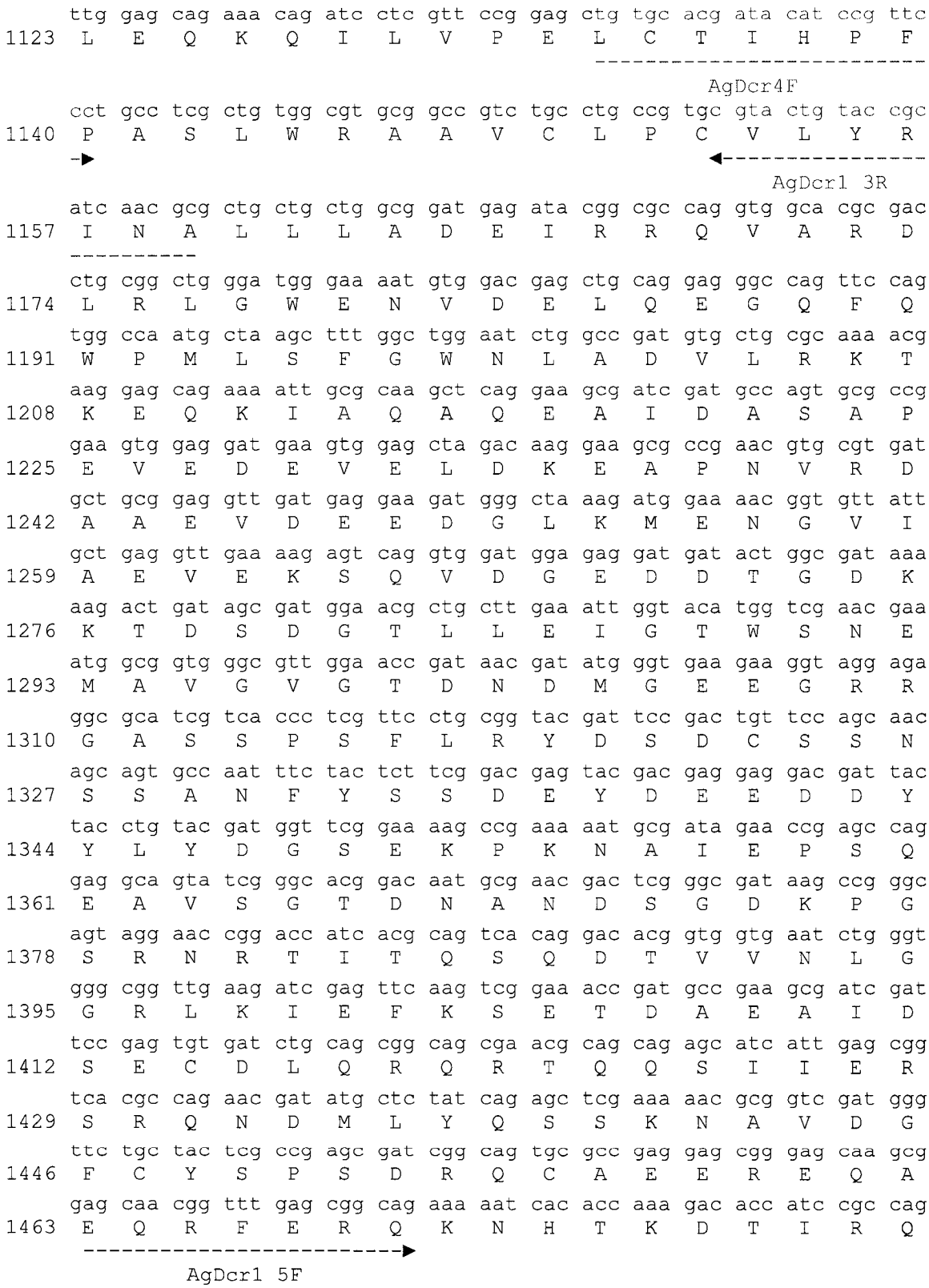
Additional Figures and Tables

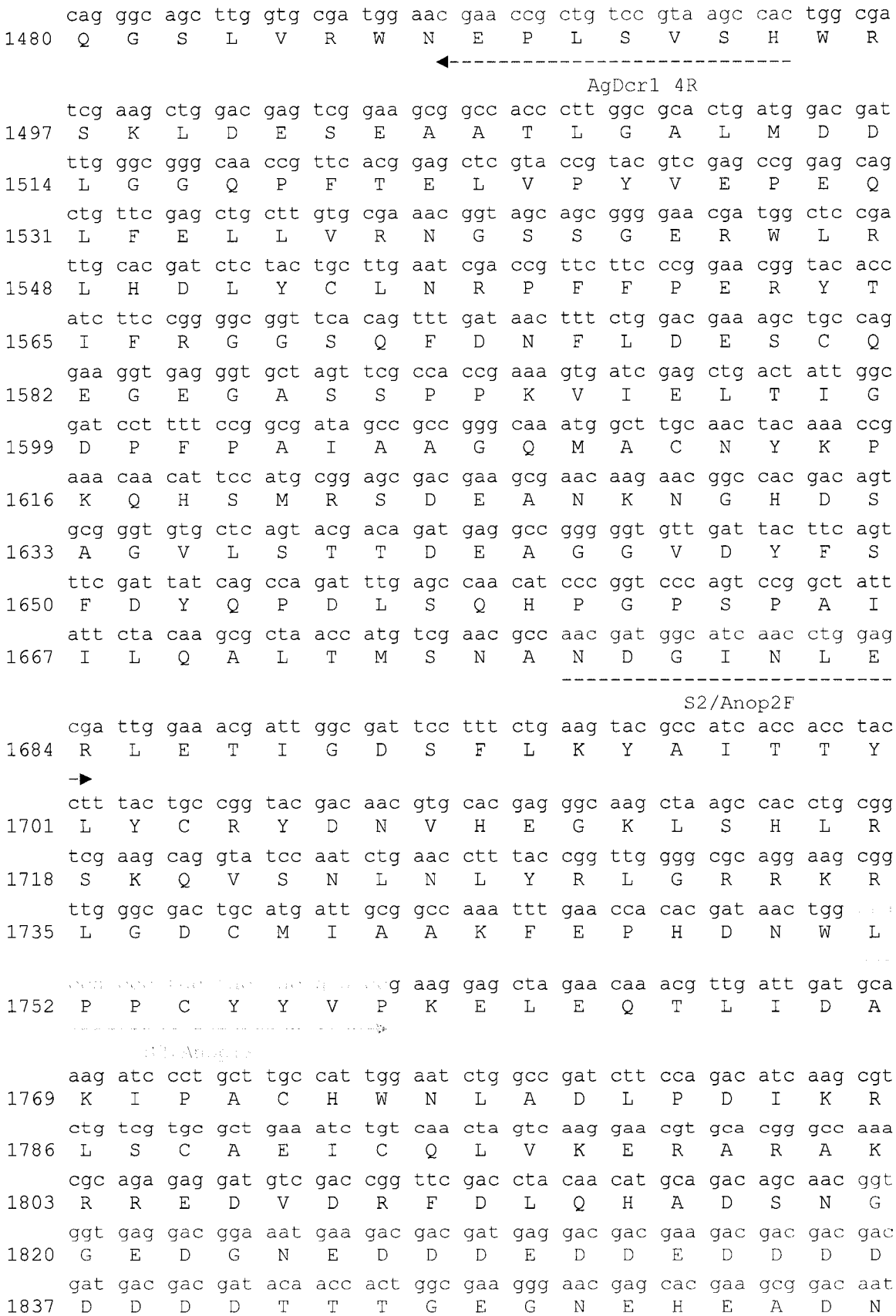
Figure A.1: Sequence of *AgDcr-1* (AY239359)

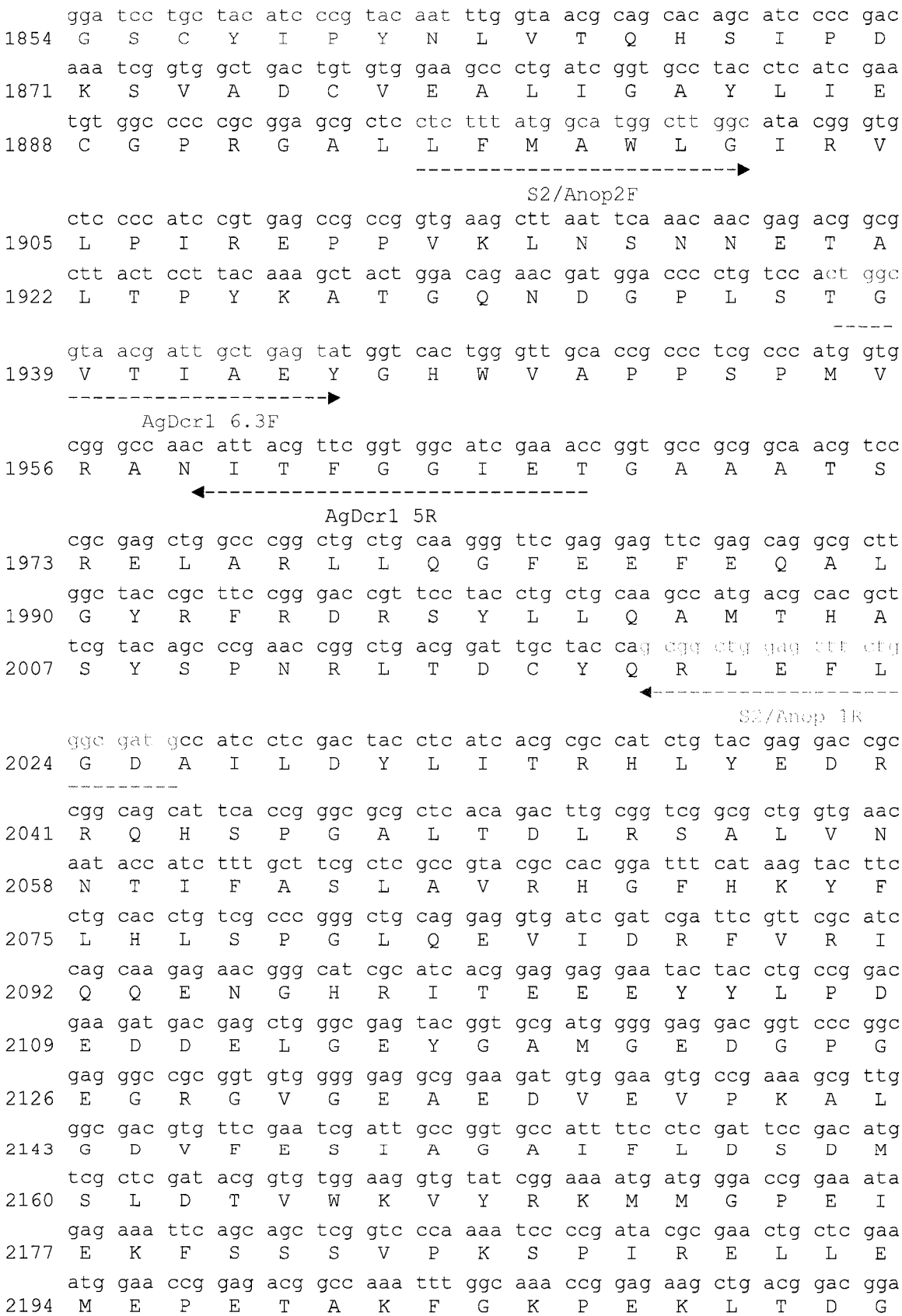
```
     atg tcg ttg ttc cac tgg acg gat ggc aac atc cac acc aca gcc ctc acg
1     M   S   L   F   H   W   T   D   G   N   I   H   T   T   A   L   T
     ---------------------->
           AgDcr1F   ----------------------->
                                AgDcr1.2F
     ccc cgg gac tat cag acc gag ctg ctg gct acg gct cgg gaa gaa aat ctt
18    P   R   D   Y   Q   T   E   L   L   A   T   A   R   E   E   N   L
     atc gtc tgt ata gcg cac aat tcc gcc aaa gag ttc ctg gcc gta aag ttg
35    I   V   C   I   A   H   N   S   A   K   E   F   L   A   V   K   L
     atc cag tcg atg cgc acc aat cgc tgg agc agt cac cct gaa gca ccg gga
52    I   Q   S   M   R   T   N   R   W   S   S   H   P   E   A   P   G
     aaa gcc atc tac ctc acg cga atg gat cga agc ttg ctg tcg tcg atg gtt
69    K   A   I   Y   L   T   R   M   D   R   S   L   L   S   S   M   V
     tcg aat ttg acc gac cta cag gtc gcc aac gtc gat gat gtg gag gac agt
86    S   N   L   T   D   L   Q   V   A   N   V   D   D   V   E   D   S
     gaa ggg tca cac gaa ccg gac ggt gct tcc aac acg ccg gtg acg gat gtt
103   E   G   S   H   E   P   D   G   A   S   N   T   P   V   T   D   V
     gct tcg gcc gat gtg ctg ttc ttc ggc agt gaa aca acg ttg ctg cag tac
120   A   S   A   D   V   L   F   F   G   S   E   T   T   L   L   Q   Y
     atc gag caa ggt acg gta cgg gtg caa gac ata agc ctg ctg att gtg gac
137   I   E   Q   G   T   V   R   V   Q   D   I   S   L   L   I   V   D
     gaa tgc cac aaa aac tac ggt cga cag gaa ctg tgg gag atc tgt gct cga
154   E   C   H   K   N   Y   G   R   Q   E   L   W   E   I   C   A   R
     ctg acg cat cag gca ccg tcg tcc gat agg ccc gct caa cgg acg aga atc
171   L   T   H   Q   A   P   S   S   D   R   P   A   Q   R   T   R   I
     ttg gga tta gct ggt cca ttg cac ggt gcg ggc tgt aca ccg gaa cgg ttg
188   L   G   L   A   G   P   L   H   G   A   G   C   T   P   E   R   L
     tgc tgg gag ctg cat tac ctt gaa cga tgc ttg cgt gca cgt atc gaa act
205   C   W   E   L   H   Y   L   E   R   C   L   R   A   R   I   E   T
     gcg agc gac atc aca agc gtt ctc aga ttc agc aca aaa cct acg gaa cta
222   A   S   D   I   T   S   V   L   R   F   S   T   K   P   T   E   L
     atc ctg gaa tgc atc cca ccc aag cca tcc aat ttg acc cag ctt ctt agg
239   I   L   E   C   I   P   P   K   P   S   N   L   T   Q   L   L   R
     atg tta ata cag cgt cag atc gct ttc ctg aag cag cat cgc tat gaa ccg
256   M   L   I   Q   R   Q   I   A   F   L   K   Q   H   R   Y   E   P
     cta gcc gtg tac gga ctg gac ggt aac gat agc gct agt gat aag ccg gac
273   L   A   V   Y   G   L   D   G   N   D   S   A   S   D   K   P   D
     aca gac ggg gaa ccg gag aag gat gaa gaa aac gat gac ctg cgc cgt gag
290   T   D   G   E   P   E   K   D   E   E   N   D   D   L   R   R   E
     cta aaa tcc atc ccc gat ccc acg gtc ggt ccg ttg agc tac cta aag cag
307   L   K   S   I   P   D   P   T   V   G   P   L   S   Y   L   K   Q
     tat ctc gag ctg ctg gac gag ttc gga ccc tgg ggc gct gat cgc ggt gca
324   Y   L   E   L   L   D   E   F   G   P   W   G   A   D   R   G   A
     ctc gaa ctt cta acg acc atc gac cag gag aag gtg aaa gtg cct tac gac
341   L   E   L   L   T   T   I   D   Q   E   K   V   K   V   P   Y   D
     cgg cac ttt ctt ctg ttc tgc atg gtg tac acc acg ctg ctg cag gca aga
358   R   H   F   L   L   F   C   M   V   Y   T   T   L   L   Q   A   R
```











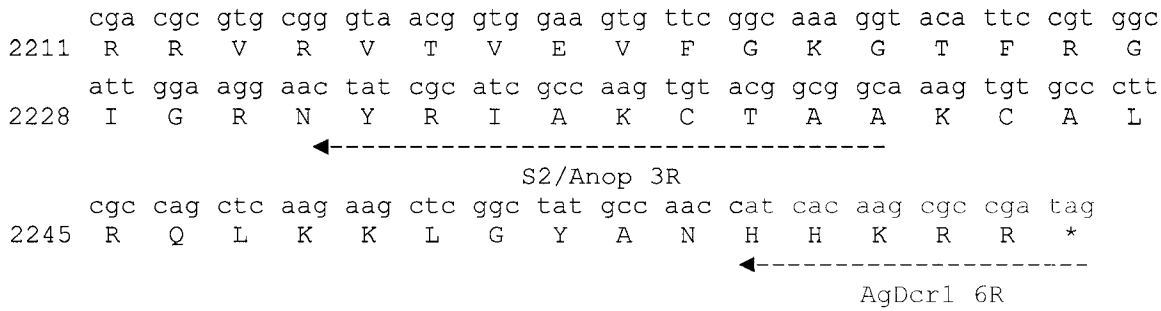


Figure A.1: *AgDcr1* sequence. The nucleotide and amino acid sequences of *AgDcr1* from *A. gambiae* larvae are shown. The sequence in red indicates the 129 bp sequence (beginning at a.a. 1819) that was determined from direct sequencing and was not present in the Ensembl annotation.

Figure A.2: Amino acid sequence alignment of dicer-1 gene products from *D. melanogaster* and *A. gambiae*

```
DmDcr1     3 FHWCDNNLHTTVFTPRDFQVELLATAYERNTIICLGHRSSKEFIALKLLQELSRRARRHG
AgDcr1     4 FHWTDGNIHTTALTPRDYQTELLATAREENLIVCIAHNSAKEFLAVKLIQSM-RTNRWSS
               *** * * ***  **** * ****** * * * *  * * *** * ** *   *  *

DmDcr1    63 RVSVYLSCEVGTSTEPCSIYTMLTHLTDLRV----------WQEQPD--MQIPFDHCWTD
AgDcr1    63 HPEAPGKAIYLTRMDRSLLSSMVSNLTDLQVANVDDVEDSEGSHEPDGASNTPVTDVASA
                        *         *   **** *              **     *

DmDcr1   111 YHVSILRPEGFLYLLETRELLLSSVELIVLEDCHDSAVYQRIRPLFENHIMPAPPADRP-
AgDcr1   123 DVLFFGSETTLLQYIEQGTVRVQDISLLIVDECHKNYGRQELWEICARLTHQAPSSDRPA
                      *   *          *     **     *            **  ***

DmDcr1   170 ---RILGLAGPLHSAGCELQQLSAMLATLEQSVLCQIETASDIVTVLRYCSRPHEYIVQC
AgDcr1   183 QRTRILGLAGPLHGAGCTPERLCWELHYLERCLRARIETASDITSVLRFSTKPTELILEC
                ********** ***    *   *  **      *******  ***    * * *  *

DmDcr1   227 APFEMDELSLVLADVLNTHKSFLLDHRYDPYEIYG----------------------TDQ
AgDcr1   243 IPPKPSNLTQLLRMLIQRQIAFLKQHRYEPLAVYGLDGNDSASDKPDTDGEPEKDEENDD
              *     *   *          **  *** *   **                       *

DmDcr1   265 FMDELKDIPDPKVDPLNVINSLLVVLHEMGPWCTQRAAHHFYQCNEKLKVKTPHERHYLL
AgDcr1   303 LRRELKSIPDPTVGPLSYLKQYLELLDEFGPWGADRGALELLTTIDQEKVKVPYDRHFLL
                *** **** * **      *  * * ***   * *          *** *  ** **

DmDcr1   325 YCLVSTALIQLYSLCEHAFHRHLGSGSDSRQTIERYSSPKVRRLLQTLRCF-----KPEE
AgDcr1   363 FCMVYTTLLQARATVASVFAQH----DTELERIKRYSTPKVRRLLEVLAWFGEQRNRPKD
              * * * * *        *  *          * *** *******  *  *      *

DmDcr1   380 V--------HTQADGLRRMRHQVDQ-ADFNRLSHTLESKCRMVDQMDQPPTETRALVATL
AgDcr1   419 RNQPATLHHHQQVHNQQRILYCFCRNVECKELEKSYHTFGAQIADVDERIKQLDGLLASV
                    * *     *             *              *        * *

DmDcr1   431 EQILHTTEDRQTNRSAARVTPTPTPAHAKPKPSSGANTAQPRTRRRVYTRRHHRDHNDGS
AgDcr1   479 RKRTERLNLKHKVVDNASAEGGGTLGVQSPRHGHESRANNFRRKRFAGGGHSHHRSNDTT
                           *      *     *           *  *       *   **

DmDcr1   491 DTLCALIYCNQNHTARVLFELLAEISRRDPDLKFLRCQYTTDRVA-DPTTEPKEAELEHR
AgDcr1   539 DALCGLIFCNNRAMARILYVLLYEVSRSQREFEFISPQYTVDKVATNPQNCLKQTTIEHR
              * ** ** **    ** *  ** * **      *   *** * **  *    *    ***

DmDcr1   550 RQEEVLKRFRMHDCNVLIGTSVLEEGIDVPKCNLVVRWDPPTTYRSYVQCKGRARAAPAY
AgDcr1   599 KQEEVLKRFRMHECNLLIGTSVLEEGIELPKCNLVIRWNSPANYRSYAQCKGRAKAPGAY
              ********** ** **********  ****** **  *  **** ****** *  **

DmDcr1   610 HVILVAPSYKSPTVGSVQLTDRSHRYICATGDTTEADSDSDDSAMPNSSGSDPYTFGTAR
AgDcr1   659 HVLFVTP---------------------------------ENAASRNEQQEDM-------
              **  * *                                    *  *    *

DmDcr1   670 GTVKILNPEVFSKQPPTACDIKLQEIQDELPAAAQLDTSNSSDEAVSMSNTSPSESSTEQ
AgDcr1   679 ----------------------------------------ASIEDVSLDGMIPEGPK---
                                                      * * **     *
```

```
DmDcr1  730 KSRRFQCELSSLTEPEDTSDTTAEIDTAHSLASTTKDLVHQMAQYREIEQMLLSKCANTE
AgDcr1  696 ---------------DDNTRETIDDQDRRMIETATDAMIEQVAIYREVEKLLLAKCRNGE
                          *    *           *     * * *** *  ** ** * *

DmDcr1  790 PPEQEQSEAERFSACLAAYRPKPHLL-------TGASVDLGSAIALVNKYCARLPSDTFT
AgDcr1  741 PPDWELKHADCFNHCLEVYRPSSGGAVALQSNGTCASLWLGNAIQTLNKYCAKLPSDTFT
            **  *   *  *  **  ***            * **  ** **   ***** *******

DmDcr1  843 KLTALWRCTRNERAGVTLFQYTLRLPINSPLKHDIVGLPMPTQTLARRLAALQACVELHR
AgDcr1  801 KLTPIWRCATTVRKGRKLYQYTIRLPINSPWKEDIL-----------------------
            ***  ***    * *  * *** ******* * **

DmDcr1  903 IGELDDQLQPIGKEGFRALEPDWECFELEPEDEQIVQLSDEPRPGTTKRRQYYYKRIASE
AgDcr1  837 ------------------------------------------------------IASV
                                                                   ***

DmDcr1  963 FCDCRPVAGAPCYLYFIQLTLQCPIPEEQNTRGRKIYPPEDAQQGFGILTTKRIPKLSAF
AgDcr1  841 FNECRPDAETVAYLYHIRMELICPIPEEQNTRGRKIYAPEESAQGFGILTTKLIPKISSF
            *  *** *    *** *   * ************** **   ********* *** * *

DmDcr1 1023 SIFTRSGEVKVSLELAKERVILTSEQIVCINGFLNYTFTNVLRLQKFLMLFDPDSTENCV
AgDcr1  901 PIFTRSGEVKVSLDLCPQRVKLSAHQLEMVNCFVKYTFTKVLRLQKSLMLYDANATENCF
             ************ *   ** *   *    * *  **** ****** *** *   ****

DmDcr1 1083 FIVPTVK---APAGGKH----------IDWQFLELIQANGNTM-PRAVPDEERQAQPFDP
AgDcr1  961 FIVPTVKQAVAGAAGKDEPPLQSDDVMVDWEFVEKIATNVHRSGPTFIPDEARKGYTFDV
            *******   * * **           ** * * *  *     *   *** *    **

DmDcr1 1129 QRFQDAVVMPWYRNQDQPQYFYVAEICPHLSPLSCFPGDNYRTFKHYYLVKYGLTIQNTS
AgDcr1 1021 GKFRDAVVMPWYRNRDQPQYFYVAEICNHLSPKSTFPGSNYATFEEYYHRKYKIHIQNQR
              * ********** ************ **** * *** ** **  **  **   ***

DmDcr1 1189 QPLLDVDHTSARLNFLTPRYVNRKGVALPTSSEETKRAKRENLEQKQILVPELCTVHPFP
AgDcr1 1081 QPLLDVDHTSARLNFLTPRYVNRKGVALPTSSEETKRAKRENLEQKQILVPELCTIHPFP
            ************************************************** ****

DmDcr1 1249 ASLWRTAVCLPCILYRINGLLLADDIRKQVSADLGLGRQ---QIEDEDFEWPMLDFGWSL
AgDcr1 1141 ASLWRAAVCLPCVLYRINALLLADEIRRQVARDLRLGWENVDELQEGQFQWPMLSFGWNL
            ***** ****** ***** ***** ** **  ** **           * **** *** *

DmDcr1 1306 SEVLKKSRESK---QKESLKDDTINGKDLADVEKKPTSEETQLDKDSKDD---------K
AgDcr1 1201 ADVLRKTKEQKIAQAQEAIDASAPEVEDEVELDKEAPNVRDAAEVDEEDGLKMENGVIAE
              ** *  * *     *          *     *           *  *

DmDcr1 1354 VEKSAIELIIE-GEEKLQEADDFIEIGTWSNDMADDIASFNQEDDDEDDAFHLPVLPANV
AgDcr1 1261 VEKSQVDGEDDTGDKKTDSDGTLLEIGTWSNEMAVGVGTDNDMGEEGRRGASSPSF----
            ****        *  *        ******* **      *            *

DmDcr1 1413 KFCDQQTRYGSPTFWDVSNGESGFKGPKSSQNKQGGKGKAKGPAKPTFNYYDSDNSLGSS
AgDcr1 1317 ------LRYDS----DCSSNSSANFYSSDEYDEEDDYYLYDGSEKPKNAIEPSQEAVSGT
                  ** *    * *   *                    *  **      *

DmDcr1 1473 YDDDDNAGPLNYMHHNYSSDDDDVADDIDAGRIAFTSKNEAETIETAQEVEKRQKQLSII
AgDcr1 1367 DNANDSGDKPGSRNRTITQSQDTVVNLGGRLKIEFKSETDAEAIDSECDLQRQRTQQSII
               *               * *        * * *   ** *           * ***
```

```
DmDcr1 1533 QATNANERQYQQTKNLLIGFNFKHED-----QKEPATIRYEESIAKLKTEIESGGMLVPH
AgDcr1 1427 ERSRQNDMLYQSSKNAVDGFCYSPSDRQCAEEREQAEQRFERQKNHTKDTIRQQGSLVRW
                 *   **  **   **     *       * *  * *      *  *   * **

DmDcr1 1588 DQQLVLK--RSDAAEAQVAKVSMM------ELLKQLLPYVNEDVLAKKL-----GDRREL
AgDcr1 1487 NEPLSVSHWRSKLDESEAATLGALMDDLGGQPFTELVPYVEPEQLFELLVRNGSSGERWL
               *     **   *   *                * ***    *   *        * *

DmDcr1 1635 LLSDLVELNADWVARHEQETYNVMGCGDSFDNYNDHH----------------------
AgDcr1 1547 RLHDLYCLNRPFF----PERYTIFRGGSQFDNFLDESCQEGEGASSPPKVIELTIGDPFP
             * **  **         * *     *  ***  *

DmDcr1 1672 -------RLNLDEKQLKLQYERIEIEPPTSTKAITSAILPAG---FSFDRQPDLVGHPGP
AgDcr1 1603 AIAAGQMACNYKPKQHSMRSDEANKNGHDSAGVLSTTDEAGGVDYFSFDYQPDLSQHPGP
                   *   **              *           *   **** **** ****

DmDcr1 1722 SPSIILQALTMSNANDGINLERLETIGDSFLKYAITTYLYITYENVHEGKLSHLRSKQVA
AgDcr1 1663 SPAIILQALTMSNANDGINLERLETIGDSFLKYAITTYLYCRYDNVHEGKLSHLRSKQVS
            ** ********************************** * ***************

DmDcr1 1782 NLNLYRLGRRKRLGEYMIATKFEPHDNWLPPCYYVPKELEKALIEAKIPTHHWKLADLLD
AgDcr1 1723 NLNLYRLGRRKRLGDCMIAAKFEPHDNWLPPCYYVPKELEQTLIDAKIPACHWNLADLPD
            **************  *** ********************  ** ****  ** **** *

DmDcr1 1842 IKNLSSVQICEMVREKA-----------DALGLEQNGGAQNGQLDDSNDSCNDF------
AgDcr1 1783 IKRLSCAEICQLVKERARAKRREDVDRFDLQHADSNGGEDGNEDDDEDDEDDDDDDDDTT
            ** **   **  * * *           *      ***      **  *   *

DmDcr1 1885 ------------SCFIPYNLVSQHSIPDKSIADCVEALIGAYLIECGPRGALLFMAWLGV
AgDcr1 1843 TGEGNEHEADNGSCYIPYNLVTQHSIPDKSVADCVEALIGAYLIECGPRGALLFMAWLGI
                     ** ****** ******** ****************************

DmDcr1 1933 RVLPITR---QLDGGNQEQRIP----GSTKPNAENVVTV--YGAWPTPRSPLLHF-----
AgDcr1 1903 RVLPIREPPVKLNSNNETALTPYKATGQNDGPLSTGVTIAEYGHWVAPPSPMVRANITFG
            *****      *   *     *    *         **   ** *  * **

DmDcr1 1979 -----APNATEELDQLLSGFEEFEESLGYKFRDRSYLLQAMTHASYTPNRLTDCYQRLEF
AgDcr1 1963 GIETGAAATSRELARLLQGFEEFEQALGYRFRDRSYLLQAMTHASYSPNRLTDCYQRLEF
                 *     **  ** ******  *** *************** ************

DmDcr1 2034 LGDAVLDYLITRHLYEDPRQHSPGALTDLRSALVNNTIFASLAVRHGFHKFFRHLSPGLN
AgDcr1 2023 LGDAILDYLITRHLYEDRRQHSPGALTDLRSALVNNTIFASLAVRHGFHKYFLHLSPGLQ
            **** ************ ******************************** * ******

DmDcr1 2094 DVIDRFVRIQQENGHCISEEYYLLSEEECD-------------------DAEDVEVPKAL
AgDcr1 2083 EVIDRFVRIQQENGHRITEEEYYLPDEDDELGEYGAMGEDGPGEGRGVGEAEDVEVPKAL
             ************** * ** * *  *                       **********

DmDcr1 2135 GDVFESIAGAIFLDSNMSLDVVWHVYSNMMSPEIEQFSNSVPKSPIRELLELEPETAKFG
AgDcr1 2143 GDVFESIAGAIFLDSDMSLDTVWKVYRKMMGPEIEKFSSSVPKSPIRELLEMEPETAKFG
            *************** **** ** **  ** **** ** ************ ********

DmDcr1 2195 KPEKLADGRRVRVTVDVFCKGTFRGIGRNYRIAKCTAAKCALRQLKKQG
AgDcr1 2203 KPEKLTDGRRVRVTVEVFGKGTFRGIGRNYRIAKCTAAKCALRQLKKLG
            ***** ********* ** ***************************** *
```

Table A.2: Viral titers of ONNV-eGFP and MRE-16-eGFP after injection into *A. gambiae* and *A. aegypti*, at 3dpi		
	PFU	
	A. gambiae	*A. aegypti*
ONNV-eGFP	3.0×10^2	7.3×10^3
MRE16-eGFP	15	9.5×10^4

* 0.5 µl of ONNV-eGFP or MRE-16-eGFP was injected into each species, 15 mosquitoes / group. Each virus was diluted to the same titer prior to injection.

Figure A.3

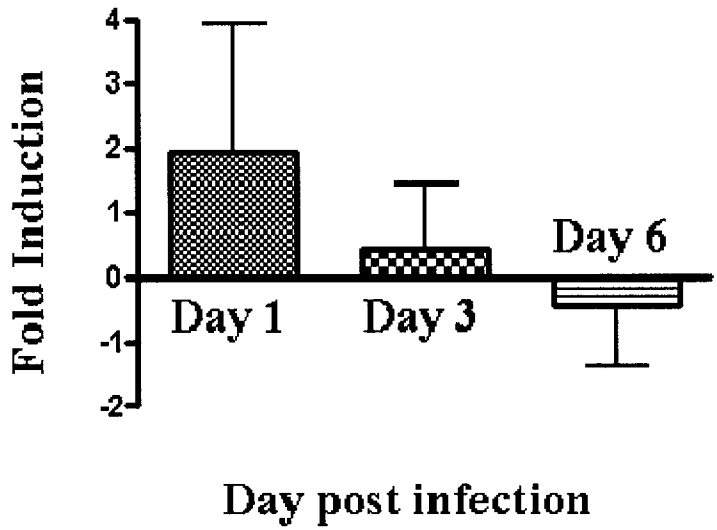


Fig. A.3: qRT-PCR results for *AgDcr2* expression comparing dsβgal and mock injected mosquitoes. Injection of dsβgal results in the induction of *AgDcr2* transcript levels at 1 dpi (1.933 fold), however, the result was not statistically significant ($p > 0.05$). At 3 and 6 dpi, the fold change is close to zero.

Figure A.4

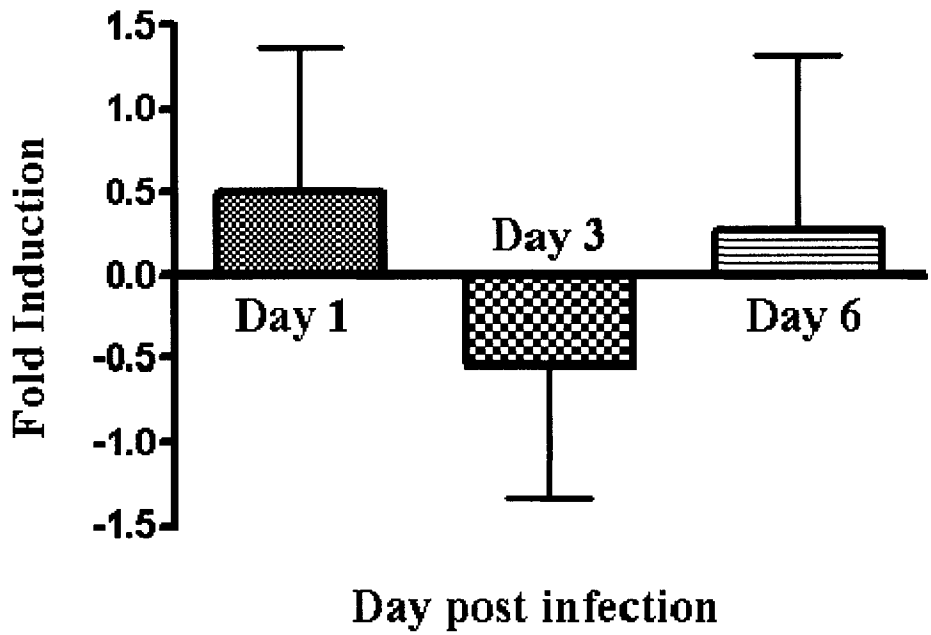


Fig. A.4: qRT-PCR results for *AgDcr2* expression comparing ONNV-eGFP and mock injected mosquitoes. Injection of ONNV does not result in a significant change of *AgDcr2* transcript levels at any time point sampled.

Figure A.5

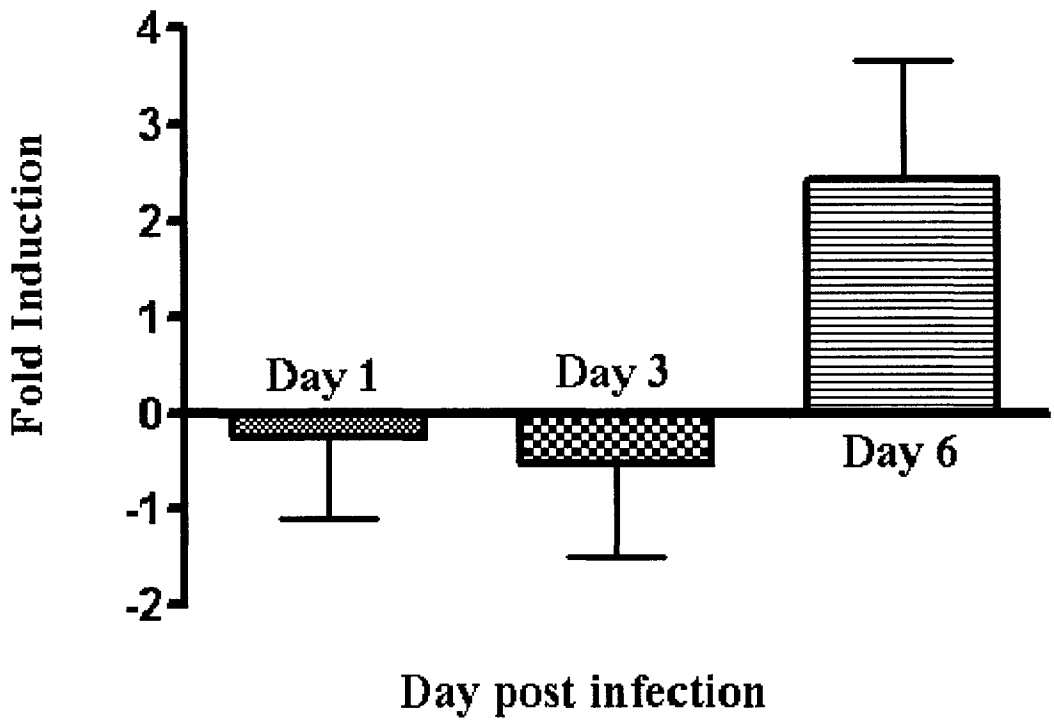


Fig. A.5: qRT-PCR results for *AgDcr1* expression comparing dsβgal and mock injected mosquitoes. Injection of dsβgal does not significantly induce expression of *AgDcr1* at any time point.

Figure A.6: TC64262 (TIGR Blast result *AaDcr1*)

GTAGCTCCCAACGTTACGAAAAAGATGGCATTGCTATGCAATCTGTGTACGGATACTGGGTACCGCCGAAATC
TCCACTCTTGATGTACGCCCCAAATCCACAGCAAACTCTGGAACATCTCTTGGATGGTTACAGTGCCTTCGAA
GACTCGTTAGGATATCAATTCACAGATCGCTCTTATCTGCTCCAAGCCATGACTCACGCATCATACTGCCCGA
ATCGATTAACAGATTGTTATCAACGTTTAGAGTTCCTCGGTGATGCAGTCCTGGACTACCTCATAACTAGGCA
TCTCTACGAAGATCCAAGACAACATTCTCCCGGCGCATTGACGGATTTGCGATCAGCCCTAGTCAATAACACG
ATCTTTGCCTCGCTAGCAGTACGGCATGATTTCCACAAGTATTTCAGGCACCTTTCGCCGGGTCTAAACGATG
TGATAGACCGGTTTGTTCGGATTCAACAGGACAATGGACACATCATCAGCGAGGAGTACTATCTAATATCCGA
GGATGAATGCGATGAGGCGGAAGATGTCGAGGTGCCGAAGGCCTTAGGCGACGTTTTTGAATCCGTAGCGGGA
GCTATCTTCTTAGATTCAAATATGTCGTTGGACGCTGTGTGGAAGGTATACCGAAGTATGATGGGACCTGAAA
TCGAACAGTTCAGCAATTCTGTACCAAAATCACCAATTAGAGAACTTCTTGAACTAGAACCAGAAACAGCTAA
GTTTGGCAAACCGGAGAAGTTGGCCGATGGACGACGCGTTCGAGTGACGGTTGAAGTATTCGGAAAAGGAACA
TTCCGTGGCATTGGGCGCAATTATCGCATTGCGAAATGTACGGCGGCAAAATGTGCCCTCCGGCAACTGAAGA
AAATGGGGCTAATCAACAAACGTCGTTAAAATATCACCAACCAGGTAGGCAAAATCAGAAAATAGCAGTGCAG
TGAAGCAATCTTATTTAGAACTCAAAGTTTACC

Figure A.7: *Aedes aegypti* dicer-1 (AY713295)
Red region indicates TIGR database sequence

```
      atg gct tac cat tgg acg gac aat att cac act acc gct ctg act ccg cgg
1      M   A   Y   H   W   T   D   N   I   H   T   T   A   L   T   P   R
      gac tat caa gtg gaa ctg ctt gcc tcg gcc aag gag cgt aat tta atc ctt
18     D   Y   Q   V   E   L   L   A   S   A   K   E   R   N   L   I   L
      tgt ctg gcg cac aac tcg tcg aaa gaa ttc att gcc cta aag ttg att cac
35     C   L   A   H   N   S   S   K   E   F   I   A   L   K   L   I   H
      gag ttg ggt tct caa cta cgg aag ccg gct gga aag cga aag cgc acg att
52     E   L   G   S   Q   L   R   K   P   A   G   K   R   K   R   T   I
      tac att tcg caa aac gat tcc gtg ttc aat ttg ata cga gac cta acg gat
69     Y   I   S   Q   N   D   S   V   F   N   L   I   R   D   L   T   D
      ctg aag gta atc aac gtg aac gat ctg gaa gat gcg gac gaa gat tac gat
86     L   K   V   I   N   V   N   D   L   E   D   A   D   E   D   Y   D
      tgg gag caa ata gtg gac gat tac caa gtg att ata acg gac gaa agg aaa
103    W   E   Q   I   V   D   D   Y   Q   V   I   I   T   D   E   R   K
      tgt ctg gat gcc ata att tgt ggc tat ttg gat ctg aac gaa gta aat ctg
120    C   L   D   A   I   I   C   G   Y   L   D   L   N   E   V   N   L
      ctg gtg atc gac gat tgc cac aaa gtg tac ggc aat gaa gag ata tcg aag
137    L   V   I   D   D   C   H   K   V   Y   G   N   E   E   I   S   K
      ctg ttt atc gac tat tac aat gta tgc cgg gag aag ccc aaa ata ttg gga
154    L   F   I   D   Y   Y   N   V   C   R   E   K   P   K   I   L   G
      ctg gca ggc cca ctg cac aat gcc gga tgt ata ccg gga agg ctg agc gca
171    L   A   G   P   L   H   N   A   G   C   I   P   G   R   L   S   A
      gaa ctg gaa caa ctc gaa tat tgt ctg cag gcg aag gcg gaa acg gcc agc
188    E   L   E   Q   L   E   Y   C   L   Q   A   K   A   E   T   A   S
      gat atc gtg acc gta ttg aga tac tgc act aaa cca aag gaa att cta ctg
205    D   I   V   T   V   L   R   Y   C   T   K   P   K   E   I   L   L
      caa tgc gct ccg cct gca aat agt aat tta gct tca tat cta aaa gaa ata
222    Q   C   A   P   P   A   N   S   N   L   A   S   Y   L   K   E   I
      gtg tta tca caa atc aca ttc ctc gaa gac cac cgg tac gat ccg tcg gaa
239    V   L   S   Q   I   T   F   L   E   D   H   R   Y   D   P   S   E
      atc tac gaa gat gac gaa ttc ctc gaa gag ttg aaa aac att cca gat cca
256    I   Y   E   D   D   E   F   L   E   E   L   K   N   I   P   D   P
      cga gcg gat cca ctt aag ttc ctc cac gag ttc ctg gcc gtg ctc gaa gaa
273    R   A   D   P   L   K   F   L   H   E   F   L   A   V   L   E   E
      atg ggc ccc tgg tgc gct gat agg gca gcg tta gca atg att gtc cag atc
290    M   G   P   W   C   A   D   R   A   A   L   A   M   I   V   Q   I
      gaa aaa caa aag att aaa act cct tac gag aga cac ttt ctg ctg ctg tgt
307    E   K   Q   K   I   K   T   P   Y   E   R   H   F   L   L   L   C
      cta gtt tcg acg gta ttc gta cag att cga tcg cac tgt gat caa att ttc
324    L   V   S   T   V   F   V   Q   I   R   S   H   C   D   Q   I   F
      caa cag tat tcc aac gaa aag gaa aag att cag gca cat tcc acg cct aaa
341    Q   Q   Y   S   N   E   K   E   K   I   Q   A   H   S   T   P   K
      gtg ctg cga tta ctg gag ata ttt cga ctg ttt agt cca gaa gga acg cga
358    V   L   R   L   L   E   I   F   R   L   F   S   P   E   G   T   R
      aat aaa gag aaa gct ata agt gaa ata ttc cca gga gag cca cct gat acg
375    N   K   E   K   A   I   S   E   I   F   P   G   E   P   P   D   T
      tgc caa tca gaa tct tgc caa aat cta ctg gaa gag ata aag ctg ata gac
392    C   Q   S   E   S   C   Q   N   L   L   E   E   I   K   L   I   D
      ttt aaa aaa ttg att gcg gat gtg gat gta gtt acc aaa agc atc gat agc
409    F   K   K   L   I   A   D   V   D   V   V   T   K   S   I   D   S
      att act gaa agt ata gcc aac ctt aaa act tct ctc aaa cag att tca gaa
426    I   T   E   S   I   A   N   L   K   T   S   L   K   Q   I   S   E
      ccg agc ata gca gag gca gat gtt gca act cag tcg aaa aaa cta cga tct
443    P   S   I   A   E   A   D   V   A   T   Q   S   K   K   L   R   S
```

```
    cca aca aaa aat ggt ata cga caa ttg aga caa cgt aga aaa cct gga atg
460  P   T   K   N   G   I   R   Q   L   R   Q   R   R   K   P   G   M
    ccc aat cgt ccc tat aga tca aac tat tac aac cag aat gat cca gat gct
477  P   N   R   P   Y   R   S   N   Y   Y   N   Q   N   D   P   D   A
    tta tgt gcg tta atc ttt tgt aac tca aag ttc att gcc aaa att ttg tac
494  L   C   A   L   I   F   C   N   S   K   F   I   A   K   I   L   Y
    agc tta ttc tac gaa agc att cgt tcc gat ccg caa cta gct tat atc aat
511  S   L   F   Y   E   S   I   R   S   D   P   Q   L   A   Y   I   N
    gtt caa tac acg gtg gac aaa acg gcc gat ccc atc aaa gat ccg aaa gag
528  V   Q   Y   T   V   D   K   T   A   D   P   I   K   D   P   K   E
    gcc gaa att gaa cac agg aaa cag gag gaa gtg ctc aaa cga ttc cga atg
545  A   E   I   E   H   R   K   Q   E   E   V   L   K   R   F   R   M
    cac gaa tgc aat ctg ctg ata ggc aca tcg gtt ctg gag gaa gga atc gac
562  H   E   C   N   L   L   I   G   T   S   V   L   E   E   G   I   D
    tta cca aag tgt aat ctc gtg att cgt tgg aac gaa ccg gtt agc tat cga
579  L   P   K   C   N   L   V   I   R   W   N   E   P   V   S   Y   R
    tcg tac gtg cag tgc aaa ggc cga gcc cga gct tcc acc gca tat cat att
596  S   Y   V   Q   C   K   G   R   A   R   A   S   T   A   Y   H   I
    ctc ttt gta act ccc aaa tct gat gct caa aat tgt gca aat ctg tgt gag
613  L   F   V   T   P   K   S   D   A   Q   N   C   A   N   L   C   E
    gat ctg agc gaa cgc atg aat cat tat att tgc gaa aag agt ata gcg aag
630  D   L   S   E   R   M   N   H   Y   I   C   E   K   S   I   A   K
    cta aat act gtt aat gaa gat tgt gaa ata tta gac gac aaa gta gaa gaa
647  L   N   T   V   N   E   D   C   E   I   L   D   D   K   V   E   E
    acc aat ccc aat tca aag caa acg tgt tgt gtt gaa atg cat tca aat agt
664  T   N   P   N   S   K   Q   T   C   C   V   E   M   H   S   N   S
    aat gaa ttc gaa tct tcg cat aag cta aac atg ctg gtc atg gaa aat tgt
681  N   E   F   E   S   S   H   K   L   N   M   L   V   M   E   N   C
    acc aac gag atg att gaa aag atc gct atc tat atg gag atc gag aag ttg
698  T   N   E   M   I   E   K   I   A   I   Y   M   E   I   E   K   L
    ctg ctt aaa aaa tgt gaa aac atg gaa cca ccg gaa agc aat caa cag tac
715  L   L   K   K   C   E   N   M   E   P   P   E   S   N   Q   Q   Y
    gcg gat tgc ttc agc cac tat ctg gag ccc tat cga cct ctg ccg gaa acg
732  A   D   C   F   S   H   Y   L   E   P   Y   R   P   L   P   E   T
    gat gaa aac agc aac agt gtt cga ctg tca aac gcc att caa atg gtc aat
749  D   E   N   S   N   S   V   R   L   S   N   A   I   Q   M   V   N
    aag tac tgc gcc aag ctt ccg agt gac acc ttc acc aaa ctt acg ccg ctg
766  K   Y   C   A   K   L   P   S   D   T   F   T   K   L   T   P   L
    tgg cga ggt gcc aaa acg gtg cga aat gga caa gag ctg tat cag tgc acc
783  W   R   G   A   K   T   V   R   N   G   Q   E   L   Y   Q   C   T
    ata cgg atg ccg atc aat tca ccc ctc aag gaa gac att gtg ggc ttg ccc
800  I   R   M   P   I   N   S   P   L   K   E   D   I   V   G   L   P
    atg cct acg gag atc ttg gcg cga cga atg gca gct tat aca gcg tgt cgt
817  M   P   T   E   I   L   A   R   R   M   A   A   Y   T   A   C   R
    gtt ctg cat aag gct gga gaa ctg gac gac tcg ttg cag ccg atc gga aag
834  V   L   H   K   A   G   E   L   D   D   S   L   Q   P   I   G   K
    gag tcg ttc cgg gca ttc gag gcc gat tgg gaa gat ttc gag ctc gag agt
851  E   S   F   R   A   F   E   A   D   W   E   D   F   E   L   E   S
    aac gat gca cag att gtg atc gaa aac tcg gaa ccg cga ccg gga act act
868  N   D   A   Q   I   V   I   E   N   S   E   P   R   P   G   T   T
    aaa aga aga cag tac tac tat aaa cgg atc gca tcg gct ttc tct gaa tgt
885  K   R   R   Q   Y   Y   Y   K   R   I   A   S   A   F   S   E   C
    aga cca aca att ggt tcc aag gta tat cta tat cac ata aat atg aca ctg
902  R   P   T   I   G   S   K   V   Y   L   Y   H   I   N   M   T   L
    cag tgt cca ata cca gaa gag cag aac acg cgt ggt cgc aag ata tat ccc
919  Q   C   P   I   P   E   E   Q   N   T   R   G   R   K   I   Y   P
```

```
     ccg gaa gat tct ccc caa ggt ttc ggt att ctg act acg aaa ctg ata ccc
936   P   E   D   S   P   Q   G   F   G   I   L   T   T   K   L   I   P
     aaa gtt agc gca ttc ccg att ttc acg cgt tct ggc gaa gtg aag gta gct
953   K   V   S   A   F   P   I   F   T   R   S   G   E   V   K   V   A
     ttg aat ctg tgc gac gag cgg cct att cta aac caa gaa caa cta gaa agg
970   L   N   L   C   D   E   R   P   I   L   N   Q   E   Q   L   E   R
     att cat atg ttc atc aac tat acg ttc acc aag gtg ctg agg ctc caa aag
987   I   H   M   F   I   N   Y   T   F   T   K   V   L   R   L   Q   K
     tat ctc atg tta tac gac cca gat gca atg gaa aat tgc ttc ttc atc gtt
1004  Y   L   M   L   Y   D   P   D   A   M   E   N   C   F   F   I   V
     cca acc atc aca cga gaa agt gtt gtt acc gtt gat tgg aat ttc att gac
1021  P   T   I   T   R   E   S   V   V   T   V   D   W   N   F   I   D
     acg ata gca aag aac gtg gac aag atg ccc act ttc ata ccg gat gaa aat
1038  T   I   A   K   N   V   D   K   M   P   T   F   I   P   D   E   N
     cgt aaa aac tat acc ttt gat agc aat cag ttc aaa gat gca gtg gtt atg
1055  R   K   N   Y   T   F   D   S   N   Q   F   K   D   A   V   V   M
     ccg tgg tat cga aac caa gat cag cct caa tat ttt tac gta gct gaa att
1072  P   W   Y   R   N   Q   D   Q   P   Q   Y   F   Y   V   A   E   I
     tgt tat cac ttg tct ccg aaa agt tca ttc cct ggt ttg aac tac agc acg
1089  C   Y   H   L   S   P   K   S   S   F   P   G   L   N   Y   S   T
     ttc gaa gaa tac tat tac aag aag tac agt att caa ata caa aat tgc aag
1106  F   E   E   Y   Y   Y   K   K   Y   S   I   Q   I   Q   N   C   K
     caa cct ctg ttg gat gtg gat cat acc agc gct aga cta aat ttc ctc acg
1123  Q   P   L   L   D   V   D   H   T   S   A   R   L   N   F   L   T
     cca agg tat gtg aat cgg aaa ggg gtg gct ctt cct aca agc tct gag gaa
1140  P   R   Y   V   N   R   K   G   V   A   L   P   T   S   S   E   E
     acc aaa cgg gcc aaa agg gaa aat ctg gag cag aag caa atc ctc ata cct
1157  T   K   R   A   K   R   E   N   L   E   Q   K   Q   I   L   I   P
     gag ctc tgt acg att cat ccg ttt cca gca tcg ttg tgg cgc gcc gct gtc
1174  E   L   C   T   I   H   P   F   P   A   S   L   W   R   A   A   V
     tgt ttg cca tgt att ctc tac aga atc aat gcc ttg ctg tta gcc gat gag
1191  C   L   P   C   I   L   Y   R   I   N   A   L   L   L   A   D   E
     ata cgc aaa gag gtg gcc cat gat ctc ggt ttg gga acc acg gac ata gcc
1208  I   R   K   E   V   A   H   D   L   G   L   G   T   T   D   I   A
     gac gaa agc ttc gaa tgg cca att tta aat ttt ggg tgg agt ctg gcc gat
1225  D   E   S   F   E   W   P   I   L   N   F   G   W   S   L   A   D
     gta ttg aaa aag tcc aga gag gct aag ctg aat gcc caa tca gca gtt gaa
1242  V   L   K   K   S   R   E   A   K   L   N   A   Q   S   A   V   E
     atg cca ccc agt gag acc aat aaa att gaa gta ttg gaa gaa gag att act
1259  M   P   P   S   E   T   N   K   I   E   V   L   E   E   E   I   T
     gaa aaa gaa agt gac aaa tct gct gcg att gag gaa gaa aag aaa gaa aac
1276  E   K   E   S   D   K   S   A   A   I   E   E   E   K   K   E   N
     ggc gaa atc gag aaa gaa gaa aag act gcc aat gat tta ctg gaa gaa gcg
1293  G   E   I   E   K   E   E   K   T   A   N   D   L   L   E   E   A
     gac aaa aaa ctt aaa gcc gag ggt ttc caa att gga acg tgg tcc aac gaa
1310  D   K   K   L   K   A   E   G   F   Q   I   G   T   W   S   N   E
     atg gca gaa act ttt gag gat gaa atg atg ttg gat ggt gat aat tgg gga
1327  M   A   E   T   F   E   D   E   M   M   L   D   G   D   N   W   G
     ctt ccg gct aac gtg gat ctg tgt tcc aga aac aga act aac atc cgt tat
1344  L   P   A   N   V   D   L   C   S   R   N   R   T   N   I   R   Y
     gga tca ccc act tcc tgg gaa gtc gga acc aag gag caa tct tcc gac ctt
1361  G   S   P   T   S   W   E   V   G   T   K   E   Q   S   S   D   L
     cgc tat tat acg gat tcc gat ggc tcg tac gtt tcg gac gat gac ttc gat
1378  R   Y   Y   T   D   S   D   G   S   Y   V   S   D   D   D   F   D
     agc ggc agt ggt gac ggt gac tcc gat aaa aac tat ggg ctg aaa atc gag
1395  S   G   S   G   D   G   D   S   D   K   N   Y   G   L   K   I   E
```

```
     ttc aaa tca gac aat gtt gcg gaa gcg atc gaa tca gcg gat gat atc ctc
1412  F   K   S   D   N   V   A   E   A   I   E   S   A   D   D   I   L
     aaa cgg caa aaa caa cgt gac att atc aat agt acg aac att aac gaa caa
1429  K   R   Q   K   Q   R   D   I   I   N   S   T   N   I   N   E   Q
     cta tac caa aac tcc aag aac ata agc agc gga ttt gat tgt acc aaa gga
1446  L   Y   Q   N   S   K   N   I   S   S   G   F   D   C   T   K   G
     agt gat ctg gag caa aat tta aaa ttg cat gaa gaa caa ttt gaa cgc tct
1463  S   D   L   E   Q   N   L   K   L   H   E   E   Q   F   E   R   S
     aca aat gaa ttc aag gaa cgc att agg agt tgt ggg acg ttg gtg aaa ttc
1480  T   N   E   F   K   E   R   I   R   S   C   G   T   L   V   K   F
     gat gag cca cta acc tta gaa aat agt tta caa aaa tca aat ggt gct aaa
1497  D   E   P   L   T   L   E   N   S   L   Q   K   S   N   G   A   K
     aag gaa agc aat aat caa ccg aat att tac gaa tta ttt cca ttt gct tca
1514  K   E   S   N   N   Q   P   N   I   Y   E   L   F   P   F   A   S
     aag gaa aca ata gat cag tac att ata tca aat acc gat gga act aaa tac
1531  K   E   T   I   D   Q   Y   I   I   S   N   T   D   G   T   K   Y
     gtg aat ctc gat ggg ttg tac gaa atg aat aaa ccg tac ttt ccg gaa aca
1548  V   N   L   D   G   L   Y   E   M   N   K   P   Y   F   P   E   T
     tat acc gta cta gga act ggt gat ata ttc gat aga ttt aat gat gaa aac
1565  Y   T   V   L   G   T   G   D   I   F   D   R   F   N   D   E   N
     ctg tta act ata gac aaa aat ggc caa aaa aca att gta ctg aat atc aaa
1582  L   L   T   I   D   K   N   G   Q   K   T   I   V   L   N   I   K
     aac ccc ttc gcc gaa ggc gaa agc tct ccc act cct gtt agt aac agt gaa
1599  N   P   F   A   E   G   E   S   S   P   T   P   V   S   N   S   E
     gaa aaa aca ata gat ctt acc gta tca act gca aat ctc aac aat ctg acc
1616  E   K   T   I   D   L   T   V   S   T   A   N   L   N   N   L   T
     gac aac cct gat gaa ttt agt ttc gac tat cag cca aat ttg aac aat cat
1633  D   N   P   D   E   F   S   F   D   Y   Q   P   N   L   N   N   H
     ccc gga cca agt cct agt att atc ctt caa gca cta aca atg tcc aat gcg
1650  P   G   P   S   P   S   I   I   L   Q   A   L   T   M   S   N   A
     aac gat ggg atc aat ttg gag cga ctg gaa acg att ggt gat tcg ttc ctg
1667  N   D   G   I   N   L   E   R   L   E   T   I   G   D   S   F   L
     aaa tac gcc ata acc acc tac ctc tat tgc acg tac gag aac gtc cac gag
1684  K   Y   A   I   T   T   Y   L   Y   C   T   Y   E   N   V   H   E
     ggc aaa ctt agc cat ctg cga tcg aaa cag gta tct aac ctg aac ctc tac
1701  G   K   L   S   H   L   R   S   K   Q   V   S   N   L   N   L   Y
     cga ttg ggc cgg cga aaa gtg ctg gga gaa agt atg ata gcc aca aag ttc
1718  R   L   G   R   R   K   V   L   G   E   S   M   I   A   T   K   F
     gag ccc cac gac aat tgg ctt cca ccg tgc tat tac gtg ccg aaa gag ttg
1735  E   P   H   D   N   W   L   P   P   C   Y   Y   V   P   K   E   L
     gaa cag gct ctc atc gat gcc aag ata ccc gct tgc cac tgg aat ttg gcc
1752  E   Q   A   L   I   D   A   K   I   P   A   C   H   W   N   L   A
     gac ctg cca gac atc aag cag ctt tcc agc gaa gaa att tgt agg cta gtt
1769  D   L   P   D   I   K   Q   L   S   S   E   E   I   C   R   L   V
     aag gag cgt gct cta gct ctc gga ctg ttg gat gat tct gac aat gaa gag
1786  K   E   R   A   L   A   L   G   L   L   D   D   S   D   N   E   E
     ggc gaa acg atg aaa gaa ata cag caa ttg gaa acg agg aac gat gag ttt
1803  G   E   T   M   K   E   I   Q   Q   L   E   T   R   N   D   E   F
     atc aag aat gct gag tat ttt gcc tgt ttt atc ccg tac aac ctg gtg act
1820  I   K   N   A   E   Y   F   A   C   F   I   P   Y   N   L   V   T
     cag cac agc atc ccg gac aag tcg gtc gct gac tgc gtg gaa gcg cta ata
1837  Q   H   S   I   P   D   K   S   V   A   D   C   V   E   A   L   I
     ggg gcc tac ttg att gaa tgc ggt cct cga ggg gca ctc cta ctg atg gca
1854  G   A   Y   L   I   E   C   G   P   R   G   A   L   L   L   M   A
     tgg tta gga atc cgg gtt ttg cca gta tat gaa gtt cct tac gac gaa aac
1871  W   L   G   I   R   V   L   P   V   Y   E   V   P   Y   D   E   N
```

```
     aat ccg aag gta cca ggt agc tcc caa cgt tac gaa aaa gat ggc att gct
1888  N   P   K   V   P   G   S   S   Q   R   Y   E   K   D   G   I   A
     atg caa tct gtg tac gga tac tgg gta ccg ccg aaa tct cca ctc ttg atg
1905  M   Q   S   V   Y   G   Y   W   V   P   P   K   S   P   L   L   M
     tac gcc cca aat cca cag caa act ctg gaa cat ctc ttg gat ggt tac agt
1922  Y   A   P   N   P   Q   Q   T   L   E   H   L   L   D   G   Y   S
     gcc ttc gaa gac tcg tta gga tat cat cac agt cgc tct tat ctg ctc caa
1939  A   F   E   D   S   L   G   Y   H   H   S   R   S   Y   L   L   Q
     gcc atg act cac gca tca tac tgc ccg aat cga tta aca gat tgt tat caa
1956  A   M   T   H   A   S   Y   C   P   N   R   L   T   D   C   Y   Q
     cgt tta gag ttc ctc ggt gat gca gtc ctg gac tac ctc ata act agg cat
1973  R   L   E   F   L   G   D   A   V   L   D   Y   L   I   T   R   H
     ctc tac gaa gat cca aga caa cat tct ccc ggc gca ttg acg gat ttg cga
1990  L   Y   E   D   P   R   Q   H   S   P   G   A   L   T   D   L   R
     tca gcc cta gtc aat aac acg atc ttt gcc tcg cta gca gta cgg cat gat
2007  S   A   L   V   N   N   T   I   F   A   S   L   A   V   R   H   D
     ttc cac aag tat ttc agg cac ctt tcg ccg ggt cta aac gat gtg ata gac
2024  F   H   K   Y   F   R   H   L   S   P   G   L   N   D   V   I   D
     cgg ttt gtt cgg att caa cag gac aat gga cac atc atc agc gag gag tac
2041  R   F   V   R   I   Q   Q   D   N   G   H   I   I   S   E   E   Y
     tat cta ata tcc gag gat gaa tgc gat gag gcg gaa gat gtc gag gtg ccg
2058  Y   L   I   S   E   D   E   C   D   E   A   E   D   V   E   V   P
     aag gcc tta ggc gac gtt ttt gaa tcc gta gcg gga gct atc ttc tta gat
2075  K   A   L   G   D   V   F   E   S   V   A   G   A   I   F   L   D
     tca aat atg tcg ttg gac gct gtg tgg aag gta tac cga agt atg atg gga
2092  S   N   M   S   L   D   A   V   W   K   V   Y   R   S   M   M   G
     cct gaa atc gaa cag ttc agc aat tct gta cca aaa tca cca att aga gaa
2109  P   E   I   E   Q   F   S   N   S   V   P   K   S   P   I   R   E
     ctt ctt gaa cta gaa cca gaa aca gct aag ttt ggc aaa ccg gag aag ttg
2126  L   L   E   L   E   P   E   T   A   K   F   G   K   P   E   K   L
     gcc gat gga cga cgc gtt cga gtg acg gtt gaa gta ttc gga aaa gga aca
2143  A   D   G   R   R   V   R   V   T   V   E   V   F   G   K   G   T
     ttc cgt ggc att ggg cgc aat tat cgc att gcg aaa tgt acg gcg gca aaa
2160  F   R   G   I   G   R   N   Y   R   I   A   K   C   T   A   A   K
     tgt gcc ctc cgg caa ctg aag aaa atg ggg cta atc aac aaa cgt cgt taa
2177  C   A   L   R   Q   L   K   K   M   G   L   I   N   K   R   R   *
```

Figure A.8: Nucleotide sequence of TC59742 (*AaDcr2*). The nucleodtide sequence with highest homology to *AgDcr2* from the TIGR *A. aegypti* genome database.

TCCTTTGCTGAATGAAAGGAATTTGATTAAATTCAATATTTATCAGATTATCCCAAACGACTGTATGCCAAAT
TGTAAGCAATCGCAACCTGGTCTATTGTGCGATAAAGTACGGACTACCTGGAATGCTAAAAATTCACAAATTC
GATCCCAAAAACGACTGGCAACCACCTCTAGCAACGGTTCCGAAGAACATCAAACGAACGATGCAATCTGTGA
ATCATTCCGCCCGAGTGCTATATCGGTTGACCTTAACAGAGGAGGAAATCAAAACCGGTGTGGTGACAGCAAA
GAACATTGATGATTTTATTGCTCAGCTTGAGCTGCATGGTAACATGCCTGATCCCTCGCCAATGGCAAATTAT
CTCTCACAGCAGACCATGGGGGACAAGACACCGGCGGATGCCATGGAAGCATTGCTAGGCGTCTGTGTGCAAT
CAGTTGGCATCGAGCGTTCCTTCAAACTGTTGCCACATTTTGGAATCCTGCCGAAAACGCACAATGTCCTAAG
ACTCCTTGCTGACAAGATCGAAAACCAACGACTGAAAACTCATATTGATATTCGTGAAGTAGATGCATTCCTC
AAGAACTACAGAAGAATTGAAGGTATTCTGGGATATAAGTTCAAGGATAGAACTTATCTCCTACAAGCACTCA
CCCATGCGTCATACCCTACCAACAGAATTACGGGAAGCTACCAACAGTTGGAATTCCTCGGCGATGCGGTATT
AGATTTCCTCATCTCGATGTACATCTTTGAGCAGAACCCTACCATGAGTCCCGGGCAGTTAACAGATCTGCGC
TCGGCTTTGGTGAATAACGTCACTTTGGCGTGCATACTTGTTCGCCATGGTCTCCACTTGTACATCCTGGCAG
AGTCGGCATCTTTCACCGACACAGTTAGCAAGTTTGTGTTGTTCCAAGAACAACACAAACACGAAATTACCGA
TCAGGTGAACTTGCTCGTCGAAGAATCAGACCGGAAGATGGCCGAATTCGTAGACGTTCCGAAGGCGCTGGGC
GATGTGTTCGAGAGTCTGGTGGCTGCGGTCTTTTTGGACTCTGGAAACGACTTTGCCGCCACCTGGCAAGTCA
TTTACGGCATGATGGGCAATGAAATATTGACCTTCACCGAGAACACACCGATCCAGATCGTGCGACAGTTGTA
TGAGTTCAAGCCGTCGTGCAAGCCAACCTTCAGCAGGGCCATCCCGGACGAGGACACGGTGCTTGTTAAGCTG
CGGTACGAGATAAGGAACCAACAACACGAAGCGTATGGTTTTGGTCAAAACAAGGACGACGCAAAGAGGGCAG
CGGCCAAAGCGGCACTGCAAGTGTTGCGCAAGCACTACCGCAGTGCTAAGTAATGTTTCGTTTTGTTTAATTT
TCTTTGTACTAGTGTCGTTTACAGATGGGAGTTGTTGAACTGTTTTCGCATGTGAATTAATATTAAATGGCGA
TAAAATGTTTTATAGAAAGCGTTTGAACGATTTATCGAGTATTAAATCAAACATAATTGTATGGAGGCAATGT
CGCAAAATGTCGACATTTTTCAAGTGTTTGTAAAACTCTAATCGTCCTACAATACTCTAATGTAAAAAAACTG
TGGAAGTGCTCATCAAATGTGTACTAAGTCGTTTGAGTCGAATCAGAACATAGTTTTAATAAAATTCCAATAG
AAATACCTATACAAAAAAAAATAGAATGAAACTCCAAATCTGACTGTTTACATGAAACCATTGACTCTCTTTT
GTGATAGAAGAGAATTGAAGACTTGGCAGTAGGTTATCTCAGATTGGCTTCCATGCTTTGTTATTAGCTTCTA
AGAGGGCGGGTTTGGAGAAGATCATTATCTAACAATTTGAATATCGGATCAAGCCAACGGTAAATCGCTGTCA
AAGCAAATACTTGTCTACGTGTGCTGGACGTTTATGATCTTCGGAACTGCGAATATGCATTTAGGCAAAGTGT
GCTTAAGATTGTCAAAGAAATGTTTTCCTTCTCAGAAGGAAAACATGAACAATAATAAGAAAAGCTATTTGAT
TGAATATATGCACGTTTTATATTACCTTT

Figure A.9: ***Aedes aegypti dicer-2*** **(AY713296), full-length sequence.**

```
      atg gat atg att atg cca cag caa gac gat ttc atc ccg cgg gac tac cag
1      M   D   M   I   M   P   Q   Q   D   D   F   I   P   R   D   Y   Q
      cgg acg atg aag aca atc tgc atg cag aaa aat aca atc atc tac ttg ccg
18     R   T   M   K   T   I   C   M   Q   K   N   T   I   I   Y   L   P
      acc gga gcc gga aag acc cat atc gcc ctg atg gtc atc aag gaa atg ggc
35     T   G   A   G   K   T   H   I   A   L   M   V   I   K   E   M   G
      aaa gat ctg gat aag cca tta act gaa ggt ggt aaa agg acg ttt ttt gtc
52     K   D   L   D   K   P   L   T   E   G   G   K   R   T   F   F   V
      gtg aat acc gtt gca tta gct aag cag cag gcc gaa ttc ctg agt cat aat
69     V   N   T   V   A   L   A   K   Q   Q   A   E   F   L   S   H   N
      ctt acc tac gac aca tcg atc tat acc agt gat cga aat gtg gat gct tgg
86     L   T   Y   D   T   S   I   Y   T   S   D   R   N   V   D   A   W
      aaa cag gac aaa tgg ctg gaa gag ttc gcg aag tac cag gtc att gtt tgc
103    K   Q   D   K   W   L   E   E   F   A   K   Y   Q   V   I   V   C
      aca tgt caa ata ttg ttg gac gtg ctg aag cat ggt tat ctc tcg gtt aaa
120    T   C   Q   I   L   L   D   V   L   K   H   G   Y   L   S   V   K
      cac atc aac ttg ctt ata ttt gat gaa tgc cac cat ggt gta gga gaa cat
137    H   I   N   L   L   I   F   D   E   C   H   H   G   V   G   E   H
      cct atg cac ggg att atg gaa caa ttt ctg agg gtt cct aag tct gac cac
154    P   M   H   G   I   M   E   Q   F   L   R   V   P   K   S   D   H
      cca cgt gta atc ggt ctt tct gga atg tta ctc tat aaa cag att aaa agt
171    P   R   V   I   G   L   S   G   M   L   L   Y   K   Q   I   K   S
      gtc gcc cta gta tct cca gaa ttg gaa cgc ttg gaa aat aca ttt aac gcg
188    V   A   L   V   S   P   E   L   E   R   L   E   N   T   F   N   A
      aca att gct acc gtt ggg agt tat gat gct ttc acc gag gtc tgc aaa ttt
205    T   I   A   T   V   G   S   Y   D   A   F   T   E   V   C   K   F
      tcg aca gat ccc aat gaa ctt ttg gtg tct tat tca acc ctt cga ctg tca
222    S   T   D   P   N   E   L   L   V   S   Y   S   T   L   R   L   S
      cct gtg atg gct gac atc gtg aac aac atc aac gct ttc agt caa acg att
239    P   V   M   A   D   I   V   N   N   I   N   A   F   S   Q   T   I
      gaa gaa ttt cac ctt cca aaa tat tta aat caa aat aaa gcc ctc cta aaa
256    E   E   F   H   L   P   K   Y   L   N   Q   N   K   A   L   L   K
      gac aga cca aag cca ctg aaa gag atc cga aaa ctc ttc aca gag ttg att
273    D   R   P   K   P   L   K   E   I   R   K   L   F   T   E   L   I
      tat cag ctt ggc gat act ggt ctt ttt ggg gga tca ata gct tta ctt ggt
290    Y   Q   L   G   D   T   G   L   F   G   G   S   I   A   L   L   G
      ttg ata gtc caa ttt gag ctg gac aaa agg caa tcg gat agt tca atg cta
307    L   I   V   Q   F   E   L   D   K   R   Q   S   D   S   S   M   L
      aga ctg gcg ttg agg tcg tgt att act ttt tgc gaa agc ctg agg cat cag
324    R   L   A   L   R   S   C   I   T   F   C   E   S   L   R   H   Q
      atc gaa aaa tta atg agt ggt ttg gac atg aag acc aaa ttg act aag ttc
341    I   E   K   L   M   S   G   L   D   M   K   T   K   L   T   K   F
      agc tct ctg aag gtc cgc caa ttg ata gat cag ctg gaa aag ttg tac gaa
358    S   S   L   K   V   R   Q   L   I   D   Q   L   E   K   L   Y   E
      gag aat cgt gac aaa aaa gcg aaa acg ctg att ttc gtc cag cga cgg ttc
375    E   N   R   D   K   K   A   K   T   L   I   F   V   Q   R   R   F
      tcg gcg aaa gtt ttg tac cat tta ctg aaa att tac ttc gct gag acg gaa
392    S   A   K   V   L   Y   H   L   L   K   I   Y   F   A   E   T   E
      gac gcc aac ctt att gtt cca gat ttc atg gtg ggc aac aat ggt tcc atg
409    D   A   N   L   I   V   P   D   F   M   V   G   N   N   G   S   M
      ccc gaa tcg att gaa caa att ttg agt gct aaa aag gat aga agg gtt ctc
426    P   E   S   I   E   Q   I   L   S   A   K   K   D   R   R   V   L
```

```
     gaa cgg ttt aaa aag aat gaa aca aac gtt att gta aca acc aac gtg tta
443   E   R   F   K   K   N   E   T   N   V   I   V   T   T   N   V   L
     gag gag ggg att gat ctt caa atg tgc aac act gtc gtc aag tat gac cat
460   E   E   G   I   D   L   Q   M   C   N   T   V   V   K   Y   D   H
     ccg caa aca ttt gcg tcg tat cag cag tca aaa gga aga gcc cgt atg aag
477   P   Q   T   F   A   S   Y   Q   Q   S   K   G   R   A   R   M   K
     aac agt cag tat atg gta atg ctg gat aac gaa aat cgt cat atc ttc ctg
494   N   S   Q   Y   M   V   M   L   D   N   E   N   R   H   I   F   L
     gag aaa tac aga ctc tat aag agt att gaa gag gag ttg cga agg tgt ctt
511   E   K   Y   R   L   Y   K   S   I   E   E   E   L   R   R   C   L
     atc ggc aaa aca ata aac cgg cca gac ccc ctc gat gcg gat gtc cac aaa
528   I   G   K   T   I   N   R   P   D   P   L   D   A   D   V   H   K
     gag cta tac aat gag atc att ccg ccg ttt ttc acc gcc aag ggt gcc aag
545   E   L   Y   N   E   I   I   P   P   F   F   T   A   K   G   A   K
     ttg gat gct ctg tcg gcc ata cag ctc ctg aac cgc tac tgc atg gga atg
562   L   D   A   L   S   A   I   Q   L   L   N   R   Y   C   M   G   M
     ccg aga gat gca ttt acc aat act aac gtc act tgg gaa cgc atc gat ctg
579   P   R   D   A   F   T   N   T   N   V   T   W   E   R   I   D   L
     aag gac ggc aga ata ata gtg gaa gtt ttg ttg ccg ctt cag tcc acc gtg
596   K   D   G   R   I   I   V   E   V   L   L   P   L   Q   S   T   V
     cga gaa aaa atc tcc ggc aat ccc atg cgc aat atc aag ttg gca aaa cga
613   R   E   K   I   S   G   N   P   M   R   N   I   K   L   A   K   R
     tca gcg gcg ttc aac gcg tgc cga aaa ttg tat gag aat aag gag ctg aat
630   S   A   A   F   N   A   C   R   K   L   Y   E   N   K   E   L   N
     gaa cat cta ata ccc ata gac tgc aaa tac cag ctt aac aat ttg aaa gat
647   E   H   L   I   P   I   D   C   K   Y   Q   L   N   N   L   K   D
     gtg tat ttc cgt cat tgg aag gat ttt gat gca gac ctc ggc aaa cta gca
664   V   Y   F   R   H   W   K   D   F   D   A   D   L   G   K   L   A
     ggc acc caa aag tgc ata cga acg cac gct atc cag tat cca aag caa aca
681   G   T   Q   K   C   I   R   T   H   A   I   Q   Y   P   K   Q   T
     aca gag tgt ttc cca cag ccg ggc aaa ccc tgt tac atc tac gtt ctg cga
698   T   E   C   F   P   Q   P   G   K   P   C   Y   I   Y   V   L   R
     att gca gct ggg ttt gcg cag gat cca aca aac gac aat gtc aac att ttc
715   I   A   A   G   F   A   Q   D   P   T   N   D   N   V   N   I   F
     cac tct ttg tat agc tcg gag aat aat ttt gga tta atg act acg aaa cct
732   H   S   L   Y   S   S   E   N   N   F   G   L   M   T   T   K   P
     ctc ccg gcc cta gca aag atg aag ttc ttt gtg act ttg gga cta atc aac
749   L   P   A   L   A   K   M   K   F   F   V   T   L   G   L   I   N
     gta cac ata gag gag act ccc atc gtg ctg cct aac ggg gga tcc gaa ata
766   V   H   I   E   E   T   P   I   V   L   P   N   G   G   S   E   I
     gaa ctt gct cta ctg aga caa ttc cat gtt acg gtc ttc cgc gac gtg cta
783   E   L   A   L   L   R   Q   F   H   V   T   V   F   R   D   V   L
     aaa ctg tgg aag gaa ttt ctc tgc tgc gat tac gac aac gag gaa aac agt
800   K   L   W   K   E   F   L   C   C   D   Y   D   N   E   E   N   S
     ttt ctg gtg gtg cct ttg aaa aat tct act cac cta gac tgg aaa ctc atc
817   F   L   V   V   P   L   K   N   S   T   H   L   D   W   K   L   I
     cga gaa ttt caa aat ttg agc gaa cca cct tcg gaa att tca acg ata gct
817   R   E   F   Q   N   L   S   E   P   P   S   E   I   S   T   I   A
     cgt aac aag atg gaa ttc gaa gcc gat aag tac cga cac aag gtc att tta
834   R   N   K   M   E   F   E   A   D   K   Y   R   H   K   V   I   L
     ccg tgg tac aag aac aac aag gaa cag cca tac gtt gtc act atg gtg cat
851   P   W   Y   K   N   N   K   E   Q   P   Y   V   V   T   M   V   H
     gaa cat ctg act ccg gag agt cct ttc cca aat ccg gag tac ggt tcc tac
868   E   H   L   T   P   E   S   P   F   P   N   P   E   Y   G   S   Y
     gcg aat tat ttc agc caa gcg tac cat ttg gca gtg gtt aag cca gat cag
885   A   N   Y   F   S   Q   A   Y   H   L   A   V   V   K   P   D   Q
```

```
      ttt ctc att gaa gtg aaa ggc atc act agt tac ctc aac cgg ttg aac ccc
902    F   L   I   E   V   K   G   I   T   S   Y   L   N   R   L   N   P
      gga gtg gaa gac gac gga aag agc acc cga agc aaa cat tgg cgt ttt aac
919    G   V   E   D   D   G   K   S   T   R   S   K   H   W   R   F   N
      gaa att ctg att ccc gaa ctg tgt cac aac tac caa ttc cct gct gac tac
936    E   I   L   I   P   E   L   C   H   N   Y   Q   F   P   A   D   Y
      tgg ttg aag gcc acc ctc ctg ccc agt gca ctt cat cga ttg cac tac ctt
953    W   L   K   A   T   L   L   P   S   A   L   H   R   L   H   Y   L
      ctg ttg gcg gaa aac att cgc gtg gat ctg gca acg ggt gca aat gtt ggc
970    L   L   A   E   N   I   R   V   D   L   A   T   G   A   N   V   G
      tgt ttg gag aat cac acg atc gaa gac gtg gac gtt gag tac aaa gag cgg
987    C   L   E   N   H   T   I   E   D   V   D   V   E   Y   K   E   R
      aaa gga aag cag ttg gag gaa cta cag ctg atg gaa ttt gaa gaa gat gaa
1004   K   G   K   Q   L   E   E   L   Q   L   M   E   F   E   E   D   E
      gac gaa gac gac gaa ttt gat ttg gaa gaa gcg aag agg tca ttg gtg gct
1021   D   E   D   D   E   F   D   L   E   E   A   K   R   S   L   V   A
      cca gag aat ttg agc gag ttg gca cga aac caa atg tgc tcc att acc ggc
1038   P   E   N   L   S   E   L   A   R   N   Q   M   C   S   I   T   G
      gat ata cct ctg ccg tgg cag gag gac gag gaa ccg gtg gat att gag cga
1055   D   I   P   L   P   W   Q   E   D   E   E   P   V   D   I   E   R
      aac tgg gat caa gtg tcg aaa ctt gat ctg gac tac tac aat gtc ttt gtg
1072   N   W   D   Q   V   S   K   L   D   L   D   Y   Y   N   V   F   V
      aat aag ttt tcc gat ctg tcg atg cgt gaa aag gcc gca gag cgt att agt
1089   N   K   F   S   D   L   S   M   R   E   K   A   A   E   R   I   S
      aca gca tac acg tca gcc gta tat aga cgt gct gct ggt agt ccg aag cga
1106   T   A   Y   T   S   A   V   Y   R   R   A   A   G   S   P   K   R
      gaa cct atg gca atc ctg gat gta ccc gta gac cag aaa ttt gcc att aaa
1123   E   P   M   A   I   L   D   V   P   V   D   Q   K   F   A   I   K
      ctg ctc cag ttg acc cca gca aat acg gtc aat gtc aat ctt caa cag aag
1140   L   L   Q   L   T   P   A   N   T   V   N   V   N   L   Q   Q   K
      aac ata atc aag gca ttg acg acg aaa tca tcg tcc gat gtt ttt gac ctg
1157   N   I   I   K   A   L   T   T   K   S   S   S   D   V   F   D   L
      gaa cgc tac gaa ttg ttg ggc gat gct ttt ttg aag ttt tcc atc tct ctc
1174   E   R   Y   E   L   L   G   D   A   F   L   K   F   S   I   S   L
      tac ctt gtt aaa tat cac aag gaa tgg cac gaa ggc ttc ctc act gca gtt
1191   Y   L   V   K   Y   H   K   E   W   H   E   G   F   L   T   A   V
      aag ggc caa att gta agc aat cgc aac ctg gtc tat tgt gcg ata aag tac
1208   K   G   Q   I   V   S   N   R   N   L   V   Y   C   A   I   K   Y
      gga cta cct gga atg cta aaa att cac aaa ttc gat ccc aaa aac gat tgg
1225   G   L   P   G   M   L   K   I   H   K   F   D   P   K   N   D   W
      caa cca cct cta gca acg gtt ccg aag aat atc aaa cga acg atg caa tct
1242   Q   P   P   L   A   T   V   P   K   N   I   K   R   T   M   Q   S
      gtg aat cat tcc gcc cga gtg cta tat cgg ttg acc tta aca gag gag gaa
1259   V   N   H   S   A   R   V   L   Y   R   L   T   L   T   E   E   E
      atc aaa acc ggt gtg gtg acc gca aag aac agt gat gat ttt att gct cag
1276   I   K   T   G   V   V   T   A   K   N   S   D   D   F   I   A   Q
      ctt gag ctg cat ggt aac atg cct gat ccc tcg cca atg gca aat tat ctc
1293   L   E   L   H   G   N   M   P   D   P   S   P   M   A   N   Y   L
      tca cag cag acc atg ggg gac aag aca ccg gcg gat gcc atg gaa gca ttg
1310   S   Q   Q   T   M   G   D   K   T   P   A   D   A   M   E   A   L
      cta ggc gtc tgt gtg caa tca gtt ggc atc gag cgt tcc ttc aaa ctg ttg
1327   L   G   V   C   V   Q   S   V   G   I   E   R   S   F   K   L   L
      cca cat ttt gga atc ctg ccg aaa acg cac aat gtc cta aga ctc ctt gct
1344   P   H   F   G   I   L   P   K   T   H   N   V   L   R   L   L   A
      gac aag atc gaa aac caa cga ctg aaa act cat att gat att cgt gaa gta
1361   D   K   I   E   N   Q   R   L   K   T   H   I   D   I   R   E   V
```

```
     gat gca ttc ctc aag aac tac aga aga att gaa ggt att ctg gga tat aag
1378  D   A   F   L   K   N   Y   R   R   I   E   G   I   L   G   Y   K
     ttc aag gat aga act tat ctc cta caa gca ctc acc cat gcg tca tac cct
1395  F   K   D   R   T   Y   L   L   Q   A   L   T   H   A   S   Y   P
     acc aac aga att acg gga agc tac caa cag ttg gaa ttc ctc ggc gat gcg
1412  T   N   R   I   T   G   S   Y   Q   Q   L   E   F   L   G   D   A
     gta tta gat ttc ctc atc tcg atg tac atc ttt gag cag aac cct acc atg
1429  V   L   D   F   L   I   S   M   Y   I   F   E   Q   N   P   T   M
     agt ccc ggg cag tta aca gat ctg cgc tcg gct ttg gtg aat aac gtc act
1446  S   P   G   Q   L   T   D   L   R   S   A   L   V   N   N   V   T
     ttg gcg tgc ata ctt gtt cgc cat ggt ctc cac ttg tac atc ctg gca gag
1463  L   A   C   I   L   V   R   H   G   L   H   L   Y   I   L   A   E
     tcg gca tct ttc acc gac aca gtt agc aag ttt gtg ttg ttc caa gaa caa
1480  S   A   S   F   T   D   T   V   S   K   F   V   L   F   Q   E   Q
     cac aaa cac gaa att acc gat cag gtg aac ttg ctc gtc gaa gaa tca gac
1497  H   K   H   E   I   T   D   Q   V   N   L   L   V   E   E   S   D
     cgg aag atg gcc gaa ttc gta gac gtt ccg aag gcg ctg ggc gat gtg ttc
1514  R   K   M   A   E   F   V   D   V   P   K   A   L   G   D   V   F
     gag agt ctg gtg gct gcg gtc ttt ttg gac tct gga aac gac ttt gcc gcc
1531  E   S   L   V   A   A   V   F   L   D   S   G   N   D   F   A   A
     acc tgg caa gtc att tac ggc atg atg ggc aat gaa ata ttg acc ttc acc
1548  T   W   Q   V   I   Y   G   M   M   G   N   E   I   L   T   F   T
     gag aac aca ccg atc cag atc gtg cga cag ttg tat gag ttc aag ccg tcg
1565  E   N   T   P   I   Q   I   V   R   Q   L   Y   E   F   K   P   S
     tgc aag cca acc ttc agc agg gcc atc ccg gac gag gac acg gtg ctt gtt
1582  C   K   P   T   F   S   R   A   I   P   D   E   D   T   V   L   V
     aag ctg cgg tac gag ata agg aac caa caa cac gaa gcg tat ggt ttt ggt
1599  K   L   R   Y   E   I   R   N   Q   Q   H   E   A   Y   G   F   G
     caa aac aag gac gac gca aag agg gca gcg gcc aaa gcg gca ctg caa gtg
1616  Q   N   K   D   D   A   K   R   A   A   A   K   A   A   L   Q   V
     ttg cgc aag cac tac cgc agt gct aag taa
1633  L   R   K   H   Y   R   S   A   K   *
```

Figure A.10: *Aedes aegypti argonaute* sequences from the TIGR database

Highest match to *AgAgo1*

>NACBN64TR
GACCGAAAATGACGACTCACTATAGGGCATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGAGGTATTCGAC
GAACCGGTTATCTTCCTGGGTGCGGACGTAACTCATCCACCGGCCGGAGACAACAAGAAGCCCTCGATTGCGG
CCGTGGTCGGTTCGATGGACGCCCATCCATCGCGTTACGCTGCTACGGTCCGGGTGCAGCAGCACCGCCAGGA
GATCATCCAGGAGCTGAGCAGTATGGTTCGGGAGTTGTTGATCATGTTCTACAAATCGACGGGCGGTTTCAAG
CCCCATCGGATCATCCTGTACCGGGATGGCGTATCGGAGGGTCAGTTCCCGCATGTGCTGCAGCACGAGCTGA
CGGCCATTCGAGAGGCGTGCATCAAGCTGGAAGCCGACTACAAGCCGGGTATTACGTTCATTGTGGTGCAGAA
GCGTCATCACACGAGACTGTTCTGCGCAGACAAGAAGGAACAGAGCGGCAAGTCGGGCAACATTCCGGCCGGT
ACGACCGTCGACGTGGGCATCACCCATCCGACCGAGTTCGATTTCTATCTCTGCAGTCACCAGGGCATTCAGG
GCACCAGCCGCCCGTCGCATTATCACGTCCTGTGGGACGACAATCACTTCGAATCCGATGAGCTCCAGTGCCT
GACGTATCAACTGTGCCACACGTATGTCCGGTGCACCCGTTCCGTTTCGATTCCGGCACCGGCGTACTACGCC
CATTTGGTGGCATTCAGGGCCAGGTACCACTTGGTCGAGAAAGAACACGATTCGGGCGAAAGATCTCACCAGA
GTGGTTGCTCCGAGGACAGAACGCCGGGTGCGATGGCTCGTGCAATCACCGTACATGCCGATACCAAAAAAGT
TATGTACTTCGCTTAAGAAAGTATTTTTTTATACAGTTTGTTCGATATATTAGTTACTTTTATAATTTTTACA
AAATTTTACATATTTATGCAATTGAAAACTTGTTTCGTATACAAAGAAATCGCGTGAATTTCAAACTTTAGTA
GCATGCTTACTGATAACAACCAGTAAAAGGTACTATTCGTCGCAAACGCCAGCTGAATAAAACCTAGAACTCA
ATTTGAAACACGCAAACACTAACCAACAAACGATATAACAGAACCGATCGATATTATATGGCT

Highest match to *AgAgo2*

>TC63572
GGATTTGGTGTTTGTGGTAATTCCATCCCCGGGAAGAGACGGCGACGTGTACGCCAAAGTTAAGCAAAAGGCA
GAACTTTGCGTTGGGCTGCTCACCCAATGTATCAAGAGCTTCACATTGGATAAGAAACGAGGCGATATGAGCA
CCATCAGCAATATTTGGCTCAAGATCAACGCCAAAACGAACGGTTCCAATCACGTGTTGGCGAAAAATTTCAA
ACCGCCGATTGCACGCAAAACGGTGATGTACGTCGGAGCCGATGTAACCCACCCTTCGCCGGAGCAAACAAAT
ATCCCAAGTGTGGTGGGATTGGCCGCTTCATACGATCTCGAGGGTTTCCGTTACAATTGCTGCTACCGTCTGC
AAGGGCCCAAAGACGAGATGATTAGAGATCTGCAAAATATCGTTATCAAACAGCTGAGACAGTTCAAGCAGAC
GAACCAAAGTCTTCCGGAACTCATCATGTACTACCGCGACGGAGTCTCCGAGGGTCAATTCCAAGAGGTACTG
ACCATAGAACTTCGTGCCATGCAAGCGGCGGCTGCTAGTGTCCAACAGGGTTACAAACCGAATATCACCTTCA
TTGTGGTCCAAAAGCGACACCATGCCCGTTTCTTCCCGACTGCCAACTGCCCGACTGAGGGAAGAAACAACAA
CGTTCAGCCGGGAACTATCGTTGACCGTTACATCACTGCTCCGAACCAGTATCAATTCTTCCTGGTTTCACAT
GCTGCTGTGCAAGGTGTGGCCAAACCCACCAAATATTGTGTGCTTTACGATGATGAAAACTGTAACCCCGATC
AGTTGCAAGCGCTGACTTACTATCTGTGCCACATGTTCACGCGCTGCAATCGAGCTGTGTCCTACCCGGCTCC
AACCTATTATGCACATCTGGCAGCCTACAGAGGACGCGTCTACATCAAAGATCGTCCTCTGAACATGAACAAC
CTGACCAAGGAGTACGAGAGAATGCAGATCAGGACCGAGATACAGGACGGTCACCCGATGTTCTTTGTTTGAA
TTCCCACGAGAAAGGGAAAATTCATTTATATTTTTTAAACTACCCATAACAAATATTTTTTCTGTGCAGTATA
AATCGCTATTTGCGACTACTACAACTAATCTGATATTTAAAAATATATTATATAATCGAAAGGAATCACATCA
TCCATCAAAACAAAATTTATAAATTTTATGAGATACATGAATGAGCTAAAATATCATCTTTACACATTACTTA
AATAACATTGGACAAAAAATGACTGACGTTTAATGTGATTCCGCTATACCCAC

Highest match to *AgAgo3*

>TC54141
TCGAATTTATTTACCATACATTTTAGATTGATTCGCGATTTGTGCGCACCTGTTCTGATCGTTGGTGAATGTG
GGTGTTGGAAAGAGTGGAGATTTCAACCGGGCCGTCACCAGTAACAACGTTCTGCAAGCGGTCAACATTCGCA
ACTGGCTCTTGGTGCACACGGCGAAGGACACCAGAATTGCCAAATCGTTTATGGATTGCGTTGAACGCAGCTG
TCGTCCCATGGGTATTCAAATTGGACCGCCAGCAATTGAGGTTCTGCAGGCCGACAAAACCGAGCTGTACGTA
CAGCTGCTGCGTACGAAAATTCGCCAGGAAACGCAAATCGTTGTTATAATCTGCCCAACGTCAAGAGATGACC
GATACGCTGCTATCAAACGTATTTGCTGTTCGGAAATCCCAGTCCCATCGCAGGTCATCAACGCTCGCACTCT
GAGTAACGAGGCCAAAAACCGAGCCATTGTGCAGAAGATCATCCTCCAGATGAACTGCAAACTGGGTGGAACG
CTGTGGAGCATTCGTATCCCGTTCGACAACGTGATGATCGCCGGCATCGATACCTATCACGATCCGAAGCAGA
AGAGCAACTCCGTTTCGGCATTCGTAGCTTCGTTGAATGGTGACTACACCAGATGGTATTCACGTGCTTGCAT
TCAGAGCAAGAAGGAGGAGTTCATCAACGGATTGTGTGCATCGATGGAGAAGTCTTTGAAGGCTTACCAGAAG
GCAAACTGTCAGCTGCCGAAGAAGATTATCATTTTCAGAGATGGCGTTGGAGACGGTCAGCTGCGAATGTGCT
CCGAGTACGAAATTCCACAACTGCAGGAGTCGTGTAAGCTAGTGGAACCGGATTACAATCCGGAGATAACATT
CATTGTCGTCCAGAAGCGTATCAACACCCGTATGTTCCGTATCGACGGTAACAACTTGGAAAACCCGAACCCT
GGCACCGTTCTGGATCACACCATCACCCGACGCAACCACTTCGATTACTTCCTGGTGCCGCAGTCGGTCCGGC
AGGGCAGCGTCTCGCCCACTCACTACATTGTGGTCCACAACCAGTCGAACCACTCGCCGGACATCCTGCAACG
GTTGAGCTACAAACTGTGCTATCTGTACTACAACTGGCCCGGAAGTGTTCGAGTTCCTGCTTGCTGTCAGTAC
GCCCACAAAATGGCCTATCTCATTGGACAGTCGGTGAAACGGAATCCGGACGAGGCGCTGAACGACAAACTGT
TCTACCTGTGAATGGCGTCATCGGTTTACCGCCAGCTGGGAGTTTTGTTTTATCCAAAGGGACTGAGTTGTTT
TCGATTTATGTTTTCTTCTCGAAATGAGATGTGTTTTATCAAGTTTCTATAGAGAGAGAGATACCTAAAATGA
GTGCCTTGTTAATCAGCGACATGATTTTCCCGCTGAGATAACCTATAAAATGTTTTTGTTACAACATATTCGG
TAACTAAAGACTGAATTCCATCTATTTTTGAGTTACCTCATCAATGACGCGTTTTGAGTGAACTAGAAGAGAA
AAATCTAACGTCTTTGAGCTCCTCGCTCGCCCTGGTTTTCAATTTCAGGCACTCGATACATTTTAAAACATTT
CAACACATCTGAACACGCCAAATACGTCGTTGGCTGTTTGTACCAGAGTCGTACAAATTACATTCTCCACAAT
ATGTTAACTTCTAACGTTGGACTGAGCAATGATAACGTTCTACTAACCTACATTTATTAGAATGAAATAACAG
TTAGTGCATAATCAAGGAAGTTTTTCGTTATTTTTTTATAAACTCCAATCCTATCGGAAGGTACCTACAGGAT
TTAAGTGCCCATCCGATGGATCGTTAGGGAGAAATCTCATATTTTATCTAATAAAATAAACATAAATTCAT

Highest matches to *AgAgo4*

>TC65831
GCATTTCCGCAACAATCCGATGCTGTCAACCGTTCGGTTAGACCATTGGTACATTATTGTACCGAGTCGGGCC
CAGCGAGAAGCCAACGAATTCCTAGGCTGCTTGATGCAGGCTGCCCAGGGAATGCGATTCGATGTCAAACGAT
GTGAGTTTGTTACCATCCCGGACGACAGTCCCGGAACCTATGTTCGGATGTTGGACAATGTTGTGAACAAGGA
TCCCCAACTGATCATGTGCGTTGTGACGAATCAGAAGGCCGATCGGTACACGGCCATCAAAAAGAAGTGCTGC
GTTGATCGCGCCGTTCCAACGCAGGTCATTTGCCAGAAGACCATCACACCGAAGGTAGGCAATGTGCGGACAC
TGATGTCCGTTGCCACCAAGGTTGTCATCCAGATGAATTGCAAACCGGGCGGAGTACCGTGGAAGGTGAAGAT
CCCGCTCAGCGGCATGATGACCATCGGTTTCTATGTGTGTCACGACACGAACGATAAGTCGAAATCTTACGGA
GCGATGGTGGCCACGTTCGATCACGATAACCGAGGCACTCCCAAGTTCTTTTCAACCGTGAGTCAGCACAGAC
ACGGGGAAGAAATATGCAACTATCTTCCTCTGAACACGATCAAAGCTCTGAACGAGTATCGCAAGGAGTACAA
CGAACTGCCGAAGCGTATCTTCTTCTACCGGGATGGTGTCGGAGAGGGTCAACTGCACTACGTGTACGAACAC
GAGGTCAAATCGATCGTGGACAAGCTGAACGAGATCTACAAATCCGCCGGAGCCGAGCAGGACGTCATGTTCA
CCTTCATCATCGTGAGCAAGCGCATCAATACGAGATTCTTCGATCGCAAACAGAATCCAAGGCCAGGAACGGT
GGTCGATGATGTGGTGACAAACCCGGAACGTACCGATTTCTACATTGTTTCGCAGTCGGTCCGCCAGGGAACG
GTTTCCCCCACGGCGTACAACGTTCTCTACGACACGTCCGGGCTGAAAGTGGACCACCTGCAGATGTTGTCCT
ACAAGCAGTGCCATCTGTACTACAACTGGTCCGGAACAACGCGGGTACCAGCGGTTTGCCAGTACGCGCACAA
ACTGTCCTTCCTGATCGGCCAGTTTATCCACCAGGCGCCGAGTAATCTGCTCGAGAAGAAGCTCTACTTCCTG
TAGGGGAAGTAATGCATCGTACTCTGGTTTTGGTTGGTTTTACAGTCTTTTACAATATTTAACTACGTTGACT
TCAAGCATTGCATGTTGACCAAGCAATTAGCGCATGTATTTAAGCAGATACCATTGCCGTAGATTTATTTTAC
TTCTTTTCGTTTCATAATATTGAATTGGCTAAGCATAATATAATAATACTTATATTATTTTTGTTGGAATTGA
CTAGTTTGTTGATACATTATTTTTATGTAAAGTAAAATTATTGCTTTAATAAATATTCGCTATTCCAGTATGG
CAC

>TC65529
CGGCACGAGGCAAAGGTTGTCATCCAAATGAACTGCAAACTGGGCGGTGTTCCATGGAAGGTCAAGATCCCAC
TCAACGGTTTGATGACCATTGGGTTCGACGTTTGCCATGATGGAAAGGACAAGTCCAAATCCTTCGGAGCCAT
GGTTGCCACGTTAGATCACGATAACAAGGGTACTCCGAAATTCTTCTCGACTGTGAGTCAGCACACGCACGGT
GAAGAAATTTCCAATTATCTGCCCATCAATACGGTCAAGGCATTGAACGAATATCGTAAGGAATTCGGAGAGC
TGCCCAAACGTATCATCTTCTACCGAGACGGTGTCGGCGAGGGTCAACTTCACTACGTGTACGAGCATGAAGT
TAAATCCATCGTTGAGAAGCTGAACCAGATCTACAAATCCGCTGGAATCGATCAGGATGTGTTGCTCACCTTC
TTCATTGTGAACAAACGAATCAACACACGCTTCTTCGATCATCGTCAGAATCCCAGACCCGGAACAGTGGTCG
ATGATGTCGTCACCCTACCAGAAAGAACTGATTTCTACCTTGTGTCGCAATCGGTCCGTCAAGGTACGGTTTC
GCCTACGGCGTACAACGTTATCTACGATACATCAGGCCTGAAGTCGATCACCTGCAGATGCTGTCGTACAAGC
AGTGCCACTTGTACTACAACTGGTCCGGAACGACTCGTGTCCCCGCGGTATGCCAGTATGCACACAAGCTGTC
CTTCCTGGTGGCGCAGTTCCTTCACCAGGCACCGAGCAATCTGCTTGAAAAGAAGCTGTACTTCTTGTAAACA
GCTGGAAAAAGATTTGCATGCCGAAAATGATCGCGGTTCTAGCTCATAAGTATTGCAGTTACTTTTTATTAGT
TAGGTATACACGTTCTTTCCAACCAGAGAGATAATTTTTACATTTTGACACGTTCCCGCAGCTGTGAATTTTT
TAAACAGTTGTCAACGACCTCAGATCAAGATTGAAAAAATGAAAAAAGATAGTAATTTGTATAACTTGCGCAT
CTTCTTAATCGATGTTAAAGAAATCCCCGCCCAGTAATTTCTAATTTTCCGAAAAAAAATGTGTTACAACTGC
AGAACGTATATATCACAAATAACTAATATAAACATTACCTATAGCATGCTTACTCACAAAGTAATTTTAATCG
TGTGTCACTAAAAATTGAATAAACGTTCGATGACAAATAAATGAAGC

>TC65832
GAGATTCCTCTCAACGGACTGATGACGGTAGGATTCGATGTGTGTCACGATACGAACGATAAGTCGAAATCCT
ACGGAGCAATGGTGGCCACGTTTGACCACGACAATCGAGGCACCCCCAAGTTCTTCTCCACCGTGAGTCAGCA
TGGCCATGGCGAAGAAATTTGCAACTATTTGCCCCTGAACACGGTCAAGGCTCTGAACGAGTATCGCAAAGAG
TACAACGAACTGCCAAAGCGTATCTTCTTCTACCGGGATGGTGTCGGAGAGGGTCAACTGCACTACGTGTACG
AACACGAGGTCAAGGCCATTGTGGACAAGCTGAACGAGATCTACAAATCCGCCGGTGCCGAGCAGGACGTCAT
GTTCACCTTTTTCATTGTGAACAAGCGCATCAATACGAGATTTTTCGATCGCAAACAGAATCCAAGGCCAGGA
ACTGTAGTCGATGACGTGGTTACGCTCCCGGAACGTACCGATTTCTACATCGTATCGCAATCGGTCCGCCAGG
GAACGGTCTCCCCCACGGCGTACAACGTCATCTACGACACGTCCGGACTGAAAGTGGACCATCTGCAGATGTT
GTCCTACAAGCAGTGCCATCTGTACTACAATTGGTCCGGAACGACGAGGGTGCCAGCGGTTTGCCAGTACGCC
CACAAACTGTCTTTCCTGGTCGGGCAGTTTATCCATCAGGCTCCGAGCAACCTGCTCGAGAAGAAACTTTACT
TCCTGTAGCGAAGTATCCCATCTGTTTGGAACCGTTCCCCCTGGAGCAGCGCTTGCGGGGAACGATATTTATT
ATTTTGTTTTCTGACTGGTTAACAGCCTGAGAGGAATTATTATTGAGTATTTGTGCGTTTTGGATTTACCAAT
TAATCCAATTGCCAAGGTTTTACGTTTTTATATATCTGACTTATGATATGTTCCTACCGCGTCGTGTGAATTA
TATATACATTGGAACAACATATTTTTACGTTTCGCAATGGTTATGATAATGTAAATAGAGACAATAAAATTAT
CA

Highest matches to *AgAgo5*

>TC62198
CGGCACGAGGGCAATGCCTTCGACAAGCGGCACGTGGTATGCGTTTCCAGATTGAGGATCCCCAACTCATCGT
CATTCCCAACGATTCACCGGCGGTCTACATCGACAGTCTGAATTCGATCGTTCAGCGGGACCCACAAATGATC
ATGTGCGTGGTTACGAACGATAAGGCTGATCGATACGCTGCAATCAAGAAAAAGTGCTGCGTTGATCGCGCTG
TTGCAACACAGGTCATCAAAACGCGTACGATCACACCCAAAGGTGGGAATGTACGAACGTTGATGTCCGTCGC
AACGAAAGTTGCAATCCAAGTGAACTGCAAGCTGGGCGGTATTCCTTGGATCCTGAAGAACCCTCTGAGCTCT
ATCATGGTCGTTGGATTTGACGTTTGCCATGATACCCGAGATAAGTCCAAATCATACGGAGCGCTGGTGGCGT
CGATGTATGGGGCGGGATGTCGACATCCGAAGTATTTCTCCACGGTTAACCACCACTCAAATGGGGAGGAACT
TTCGAACTTTATGGCTCAAAACATAATCAAAGCTGTGCATTCGTATCGCGCGGATTTCGGAAATGCGCTACCC
GAACGCATCATTGTGTATCGCGACGGAGTCGGCGAAGGACAACTCAAATACGTCTACGATCATGAAGTGAATG
CGATCAAGGAGAAACTCCTTCTAGCATGTAAGGAAAGGCGAGAAGCGGCAAGATTAACATTTTTCGTAGTTAA
CAAACGAATCAACACTAGACTGTTCCATCAGAAGCGGAATCCCGTGCCGGGTACGATCGTGGATGACGTTATA
ACCCTACCCGAAAGGAACGACTTCTACCTCGTTTCACAAAGCGTCCGGCAAGGTACGGTTTCGCCCACAAGCT
ACAACATACTCAAGGACGAGTCCGGTTTGAATGCCGACAAGCTACAGTTGTATACCTTCAAGCAGACGCACAT
GTACTACAACTGGTCCGGAACGGTTGGCGTACCTGCGGTTTGCCAATACGCACACAAGTTGGCAGCATTGGCC
GGGCAGCACTTGCATCAAGCTCCCAATAGTTTGTTGGAGAAAAAATTGTATTATTTGTAAATTTTGCATATTA
TCACATGTATTCACACGAAT

>TC63456
GAGTCTTATCGTGCAACGAGACCCCCAGCTGATCATGTGTTTGGTAACGAATGACAAGGCCGATCGTTATTCG
GCCATCAAGAAGAAATGCTGCGTGGACCGAGCCGTACCAACACAGGTCTTGAAGACCCGAACCATTACACCCA
AAGGAGGAAATGTACGTACGTTGATGTCCGTGGCCACCAAAGTAGCGATCCAGATGAACTGCAAACTGGGTGG
AATCCCATGGGTTATCAAAAGCCCTCTGGCTTCCGTGATGGTCATTGGGTACGACGTATGTAAGGATTCTAAG
GATAGATCGAAAGGATACGGAGCATTGGTTGCCTCCATGTACGGAGGTGGCGTCAAGCATCCAAAGTACTATT
CCACTGTTAACCAGCATGCCTATGGAGAGGAACTGTCCAATTATTTAGCCTTGAACGTTATCAAGGCGATCCG
TGCATATCAGTCAAGCTTCGGAAATATCCTTCCGCAACGTATCGTCATCTATAGGGATGGCGTCGGCGAAGGA
GAGTTGGGCTATGTGCATGAGCATGAAGTCGGGGCCGTGAAGGAGAAACTGGAGGCCGCTTACAAGGCACACG
AATTTCAGTCGAAACTGACCTTTTTCGTAGTTAATAAACGCATCAATACGCGATTATTCCATGACAGGAGAAA
TCCCACGCCCGGAACTATAGTAGATGACGTCATTACACTACCAGAAAGGAACGATTTCTACCTGGTTTCCCAG
AGCGTCCGGCAAGGAACGGTGTCCCCAACGTCGTACAACATTCTCAAAGACGAATCCGGACTTAGTGCCGATC
GGTTGCAGTTGTACACATTCAAGCAGACGCACATGTATTACAATTGGTCCGGTACCGTCGGAGTGCCGGCCGT
CTGTCAGTACGCCCACAAATTGGCCGCCCTGGCCGGTCAGTATCTCCACCAAGCGCCCAGCACGTGGCTGGAG
AAGAAACTTTACTTTTTGTAGACCTAAGCAGATATCGACTATGGTTTGTACGTTTATGGATGTTTTAGATGTT
TCGAAGAATCTTAATAAATTGACGGTAATGAAACG

Matches with *Argonaute* genes but unknown *A. gambiae* homologues

>AEMTBA03
AATACGTACACAATCGCCGATGTGGAGTTCAACACAAGTCCAGAGAGCGCATTCGATGCAGCTGGTACACGGG
TTACGTTTATGCAGTACTATAAGGATCGGTACAACATTACAATTCGCGATCCCCGTCAGCCTATGCTGGTGTC
TCGTGCTAAGCAGAGGGATATTCGAGCTGGAAAGTCGGAGCTAATCTATCTTGTTCCGGAACTTGTTCGAGCT
ACTGGCATTACCGACGAAATGCGGAAGAACTTCAATCTGATGCGAACGCTTGCTGATCATACTCGATTGACGC
CGGATAAACGTATACAGCGCTTGGAGCATTTTAACCTTCGATTCAACAGTCGAAGGAAAGCTCTGAAATCTTC
CAGTTTTGGAAGACAGAGTTGGATCGACGCCTAGTCGAGGTGCCGGCAAGAGTGCTTAGACCAGAAGAAATCT
TTTTCCATCCTTCGCAAGAAAATTCAAAGTGACGGCAGGAGACGTAGCCGATTGGCAAATGGCCTTCAGAAAC
AATCCAATGTATATCTCGATTCCGTTGGTCAACTGGTACTTTATTGTGCCGGCAGGATCGGAGAAGCTTATGG
TTGATTTTATGCAGTGTCTGAAACAGGCTTGCCAAAAGGATGCGTTTCCAAATTG

>AI648322
AMTADAADCGAAGTTACGATCAAGATCATTTACAAGCGCAAACAGCGAATGAGCGAGAACATTCAATTCTACA
ACATTCTCTTCCAACGCATTATGAAAGTCCTGAAGATGGTCGAGATGGGTCGTAAGAACTTTGACCCCTCGGC
CCCGAAACTGATTCCGCAGCATCGTCTGGAGATCTGGCCGGGATACGTTAGCGCTGTGGACGAATATGAGGGC
GGCTTGATGTTGAATCTCGACGTGTCCCATCGTGTGCTGCTCCAGACGACGGTTTTGGATCACATCCGAACGC
TGGCAAGGGCCAATCCCCAAGATTACAAAAACATGGCCACCAAATCGCTGCTGGGAGCCGTTATATTGACCCG
ATACAACAACAAAACGTACCGTATCGACGACATTCTGTTCGATCAGAATCCGACCATGACTTTCGAAGCCAAC
GGGCAGCCAATTTCGTACCTGCAGTATTACAAACAGCAGTACAACATCGACATCCACGATCTGAAGCAGCCGC
TGCTGATTAATCGCAAGGAGCGTCGTGTATCTGGCCAAGACAAACCGATGGAAATGATCATGTGCTTGATCCC
GGAGATTTGCTACCTTACGGATTGACCGACC

>NABWG64TR
AGAACTAAATATCATAGTTCGATGCTACTCCTATAGGGATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGA
AATATAATTCTATTAAATGCCCGGCGAATCAGCAAAAATTCAAGCTTGAAAATGGCTTTGAAATGACCATTGA
TCAATACTTCCGGTCGAAGAACAAACAACTTCGGTATCCTTCTCTGCCGGTTCTACATGTTGGAAGCCTCGTG
CGCAACGTAATGCTGCCGATTGAACTGTGCAGTATTCCGCCAGGACAGGCCCTCAACAAGAAACACCCGGATC
AATGCACTCAGTTCATAATTAGGAAGTCGGCCACCGATACCGCAACCCGAAAGCGCAAAATCATGGATTTGTT
CAACCAGATCGGCTACAACAATGCTCCAACCATCAAAGAATTTGGAGTGAGCGTCGGAAATAACTTCGAAACG
GTTGACGGCTGTATATTGGATCCTCCTGAGCTGTCCTACCGCAATGACCGTAGGGTAAAACCAATGAGGGGAG
TTTGGCGGGCTGACAACATGAACTTCATAATCCCGAGTACGGAGATTACAAGACGGGAATTGAGTTGGACAAT
ATTGAACCTGGACGGCAGAACTCGTCCAGATGCCATCGACGAGTTTGGACGTAACATTTATCAGATGTCCCTA
AAGCAGGGTGTCCAACTACAGCAGTTCAGCATGAAGAACAATTTTCTATGAACCTAGAGATATGAGATTCGCC
GTGAAGGATCTGGACACATTTTTGATGAGCTCAAAAAAGCGCAAAATTGATTTGGTGTTTGTGGTATTCCATC
CCCGGGAAGAGACGGCGACGTGTACGCCAAAATTTAAGCAAAAGGCACAACTTTGCGTTGGGCTGCTCACCCA
ATGTATCAAGAGCTTCCACATTGGATAAAAAACGAGGCTATTATGAGCACCATCAGCAATATTTGGGCTCAGG
ATTCACGCCAAAACGGACGGGTTCCAAACACTTGTTGGGCAAAAATTTTCAAACCGCCCAATGGCCCGCCAAA
AGGGTGA

>TC57115
GAGTTTGCTGGAACGCTACCTCAGCGTATCATCGTGTACCGTGATGGCGTTGGCGAAGGTCAACTACAATACG
TGTACCAACACGAAATCACTGCTATGAAGGAAAAATTGAATATTGCCTTCAAAGATCAACCAAATGCTTCCAG
ACTGACGTTCTGTGTGGTGAGCAAAAGAATTAACACCCGTTTGTTCCAAGGCGGCCAAAACCCACTGCCCGGA
ACAATTGTCGATGATGTCATAACTCTGCCAGAAAGGAATGATTTCTTCCTTGTTTCTCAAAGCGTTCGCCAGG
GAACTGTGTCCCCGACCAGTTACAATATTCTTAGAGATGAATCCGGTCTGACTGCTGACCAGCTATAACTCTA
CACCTACAAGCAGACCCATCTTTACTACAACTGGTCCGGTACCGTGGGAGTTCCAGCAGTTTGCCAATACGCC
CATAAGTTAGCCTTCTTGGCTGGTCAACATCTGCATCAGGCTCCCAACAATTTGCTGGAAAAGAAACTCTATT
ATCTGTAAACGTAATGTTGCTGTTTCGAATGTAAGCTCAAGAGATCAAACGCGAATGTTTGATGTAAAGGAGC
TTTTCTTAGAATTTGGCTGTGCTTTCTAATGAGCTTGGCAAGTGTTGTGAAGAAAAACTTGAGATTTATTGAC
ATATTTCATAAAGAATAAAAAATAACATAGTACAAAATGGATGTGCTGCTTTACCCCGAATGTTAGTTTACAA
ACGTCAGTTTCTCGATTTTATCTTCCCAGAAGTGTCCATTGTTCCAATTTTTTTTTTCAATTTTATTAGAGAG
GTTTTAATAATCTCGTGATCATTCACCTCTAATTGTTCCAAATCTTTTCGTAATAGAAATTTCCTTTACTTTC
CTGGGTATAGAGTTTCATAGTACCTGTCACATGATATACGAATGCAAAAATGGCAACTTAGGCAAAACAAGCT
CTCAGTAAATAGCTGTGGAAGTGCTTTAGAACACTCAGCAGAGAAATAGGCTCCGTCCCAGTTGGAACGTAAT
ATCGATAAGAGAAAATCTTGTAGTGTTTCACGCACCATATTTTAAGTTCAGTCAGCACTGTCTAGTTTTCCTT
CCTAGACATCACCCCATTTAACATTGCGTGATGCACTTCTAAGAGTCAAGAAATATGAATTGAAAAAAAAAAT
CGAGTACATTTTCTTATTCTATATTCTGAAACTTGCTCAAAATGTCGAGAAAAGGGAAAGTTCACTTTGGCAA
AATGTATGAAAATAATTCTACAATCACGCTTTAATTCAACCGGAGGTCTCACATCCAGTAGAAAACAAATAAC
CTGTTCAGGTTCATTAGATTGCATTGCCGAGTTTATCATTTCCATTTCAGAAGGTAGAAGTATGTCGAACAAA
AATCAAAGCTTTTGTTTTACGAATAGATTTGTGATTAGCTATAAAAATGTTCGTAATATTTTCAAAGATTCAA
TAAAAAATTTATATTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGCCAGCGTTGCGGCCGCTCC
AACATCAGAATGGTC

Figure A.11: Amino acid alignment of AgDcr1 and AaDcr1

50.3% identity in 2370 residues overlap

```
AgDcr1     4 FHWTDGNIHTTALTPRDYQTELLATAREENLIVCIAHNSAKEFLAVKLIQSMRTNRWSSH
AaDcr1     3 YHWTD-NIHTTALTPRDYQVELLASAKERNLILCLAHNSSKEFIALKLIHELGSQLRKPA
               **** ************* **** * * *** * **** *** * ***

AgDcr1    64 PEAPGKAIYLTRMDRSLLSSMVSNLTDLQVANVDDVEDSEGSHEPDGASNTPVTDVASAD
AaDcr1    62 GKRK-RTIYISQNDS--VFNLIRDLTDLKVINVNDLEDADEDYDWE--------QIVDDY
                    **    *          **** * ** * **

AgDcr1   124 VLFFGSETTLLQYIEQGTVRVQDISLLIVDECHKNYGRQELWEICARLTHQAPSSDRPAQ
AaDcr1   111 QVIITDERKCLDAIICGYLDLNEVNLLVIDDCHKVYGNEEISKLF--IDYYNVCREKP--
                  *   *  *  *         **  * *** **  *                  *

AgDcr1   184 RTRILGLAGPLHGAGCTPERLCWELHYLERCLRARIETASDITSVLRFSTKPTELILECI
AaDcr1   167 --KILGLAGPLHNAGCIPGRLSAELEQLEYCLQAKAETASDIVTVLRYCTKPKEILLQCA
               ********* *** * **  **  ** ** *  ******  ***  *** *  * *

AgDcr1   244 PPKPSNLTQLLRMLIQRQIAFLKQHRYEPLAVYGLDGNDSASDKPDTDGEPEKDEENDDL
AaDcr1   225 PPANSNLASYLKEIVLSQITFLEDHRYDPSEIY----------------------EDDEF
               **  ***   *      ** **  *** *   *                      * *

AgDcr1   304 RRELKSIPDPTVGPLSYLKQYLELLDEFGPWGADRGALELLTTIDQEKVKVPYDRHFLLF
AaDcr1   263 LEELKNIPDPRADPLKFLHEFLAVLEEMGPWCADRAALAMIVQIEKQKIKTPYERHFLLL
               *** ****   **  *   *  * * *** *** **     *   * * ** *****

AgDcr1   364 CMVYTTLLQARATVASVFAQHDTELERIKRYSTPKVRRLLEVLAWFGEQRNRPKDRNQPA
AaDcr1   323 CLVSTVFVQIRSHCDQIFQQYSNEKEKIQAHSTPKVLRLLEIFRLFSPEGTRNKEKAISE
               * * *   * *      * *   * * *   ***** ****    *     * *

AgDcr1   424 TLHHHQQVHNQQRILYCFCRNVECKELEKSYHTFGAQIADVDERIKQLDGLLASVRKRTE
AaDcr1   383 IFPGEPPDTCQSESCQNLLEEIKLIDFKKL-------IADVDVVTKSIDSITESIANLKT
                        *                 *        *****   *  *    *

AgDcr1   484 RLNLKHKVVDNASAEGGGTLGVQSPRHGHESRANNFRRKRFAGGGHSHHRSN----DTTD
AaDcr1   436 SLK---QISEPSIAEADVATQSKKLRSPTKNGIRQLRQRRKPGMPNRPYRSNYYNQNDPD
               *           **          *          *  *  *      ***       *

AgDcr1   540 ALCGLIFCNNRAMARILYVLLYEVSRSQREFEFISPQYTVDKVATNPQNCLKQTTIEHRK
AaDcr1   493 ALCALIFCNSKFIAKILYSLFYESIRSDPQLAYINVQYTVDKTA-DPIKDPKEAEIEHRK
               *** *****    * *** * **  **      *  ****** *  *    *   *****

AgDcr1   600 QEEVLKRFRMHECNLLIGTSVLEEGIELPKCNLVIRWNSPANYRSYAQCKGRAKAPGAYH
AaDcr1   552 QEEVLKRFRMHECNLLIGTSVLEEGIDLPKCNLVIRWNEPVSYRSYVQCKGRARASTAYH
               ************************** ********** *  **** ****** *  ***

AgDcr1   660 VLFVTPENAASR----------------------NEQQEDMASIEDV-------SLDG
AaDcr1   612 ILFVTPKSDAQNCANLCEDLSERMNHYICEKSIAKLNTVNEDCEILDDKVEETNPNSKQT
               *****   *                          *   **     *        *

AgDcr1   689 MIPEGPKDDNTRETIDDQDRRMIETATDAMIEQVAIYREVEKLLLAKCRNGEPPDWELKH
AaDcr1   672 CCVEMHSNSNEFESSHKLNMLVMENCTNEMIEKIAIYMEIEKLLLKKCENMEPPESNQQY
               *     *  *          *  *  ***  *** * ***** ** * ***
```

```
AgDcr1   749 ADCFNHCLEVYRPSSGGAVALQSNGTCASLWLGNAIQTLNKYCAKLPSDTFTKLTPIWRC
AaDcr1   732 ADCFSHYLEPYRPLP------ETDENSNSVRLSNAIQMVNKYCAKLPSDTFTKLTPLWRG
             **** * ** ***              *  * ****  ***************** **

AgDcr1   809 ATTVRKGRKLYQYTIRLPINSPWKEDIL--------------------------------
AaDcr1   786 AKTVRNGQELYQCTIRMPINSPLKEDIVGLPMPTEILARRMAAYTACRVLHKAGELDDSL
             * *** *  *** *** ***** ****

AgDcr1   837 ------------------------------------------------IASVFNECRPDA
AaDcr1   846 QPIGKESFRAFEADWEDFELESNDAQIVIENSEPRPGTTKRRQYYYKRIASAFSECRPTI
                                                           *** * ****

agdcr1   849 ETVAYLYHIRMELICPIPEEQNTRGRKIYAPEESAQGFGILTTKLIPKISSFPIFTRSGE
AaDcr1   906 GSKVYLYHINMTLQCPIPEEQNTRGRKIYPPEDSPQGFGILTTKLIPKVSAFPIFTRSGE
                 ***** * * *************** ** * ************ * *********

AgDcr1   909 VKVSLDLCPQRVKLSAHQLEMVNCFVKYTFTKVLRLQKSLMLYDANATENCFFIVPTVKQ
AaDcr1   966 VKVALNLCDERPILNQEQLERIHMFINYTFTKVLRLQKYLMLYDPDAMENCFFIVPTITR
             *** * **  *  *   ***    *  ********** *****  * *********

AgDcr1   969 AVAGAAGKDEPPLQSDDVMVDWEFVEKIATNVHRSGPTFIPDEARKGYTFDVGKFRDAVV
AaDcr1  1026 ESV--------------VTVDWNFIDTIAKNVDKM-PTFIPDENRKNYTFDSNQFKDAVV
                           * *** *   ** **    ******* ** ****   * ****

AgDcr1  1029 MPWYRNRDQPQYFYVAEICNHLSPKSTFPGSNYATFEEYYHRKYKIHIQNQRQPLLDVDH
AaDcr1  1071 MPWYRNQDQPQYFYVAEICYHLSPKSSFPGLNYSTFEEYYYKKYSIQIQNCKQPLLDVDH
             ****** ************ ****** *** ** ******  ** * ***  ********

AgDcr1  1089 TSARLNFLTPRYVNRKGVALPTSSEETKRAKRENLEQKQILVPELCTIHPFPASLWRAAV
AaDcr1  1131 TSARLNFLTPRYVNRKGVALPTSSEETKRAKRENLEQKQILIPELCTIHPFPASLWRAAV
             ************************************** ******************

AgDcr1  1149 CLPCVLYRINALLLADEIRRQVARDLRLGWENVDELQEGQFQWPMLSFGWNLADVLRKTK
AaDcr1  1191 CLPCILYRINALLLADEIRKEVAHDLGLG---TTDIADESFEWPILNFGWSLADVLKKSR
             **** **************  ** ** **           * ** * *** ***** *

AgDcr1  1209 EQKIAQAQEAIDASAPEVEDEVELDKEAPNVRDAAEVDEEDGLKMENGVIAEVEKSQVDG
AaDcr1  1248 EAKL-NAQSAVEMP-PSETNKIEVLEEEITEKESDKSAAIEEEKKENGEIEKEEKTANDL
             * *   ** *     *      *   *                * *** *   **   *

AgDcr1  1269 EDDTGDKKTDSDGTLLEIGTWSNEMAVGVGTDNDMGEEGRRGASSPSFLRYDSDCSSNSS
AaDcr1  1306 LEEA-DKKLKAEG--FQIGTWSNEMAE-TFEDEMMLDGDNWGLPANVDL-----CSRNRT
                  ***    *    *********     *  *      *      *     ** *

AgDcr1  1329 ANFYSSDEYDEEDDYYLYDGSEKPKNAIEPSQEAVSGTDNANDSGDKPGSRNRTITQSQD
AaDcr1  1357 NIRYGSPTSWEVG---TKEQSSDLRYYTDSDGSYVSDDDFDSGSGDGDSDKNYG------
                * *    *         *             **  *    ***     *

AgDcr1  1389 TVVNLGGRLKIEFKSETDAEAIDSECDLQRQRTQQSIIERSRQNDMLYQSSKNAVDGF-C
AaDcr1  1408 --------LKIEFKSDNVAEAIESADDILKRQKQRDIINSTNINEQLYQNSKNISSGFDC
                     *******   **** *  *      *  **     *  *** ***   ** *

AgDcr1  1448 YSPSDRQCAEEREQAEQRFERQKNHTKDTIRQQGSLVRWNEPLSVSHWRSKLDESEAATL
AaDcr1  1460 TKGSDLE--QNLKLHEEQFERSTNEFKERIRSCGTLVKFDEPLT-------LENSLQKSN
                ** *          *  ***  *  *  **  * **   ***        *  *
```

```
AgDcr1 1508 GALMDDLGGQPFTELVPYVEPEQLFELLVRNGSSGERWLRLHDLYCLNRPFFPERYTIFR
AaDcr1 1511 GAKKESNNQPNIYELFPFASKETIDQYIISN-TDGTKYVNLDGLYEMNKPYFPETYTVLG
            **           ** *    *        *   *     *  **  * * *** **

AgDcr1 1568 GGSQFDNFLDESCQEGEGASSPPKVIELTIGDPFPAIAAGQMACNYKPKQHSMRSDEANK
AaDcr1 1570 TGDIFDRFNDENLLTID--KNGQKTIVLNIKNPF---AEGESS----PTPVSNSEEKTID
             *  ** * **            * * * *  **   * *       *   *

AgDcr1 1628 NGHDSAGVLSTTDEAGGVDYFSFDYQPDLSQHPGPSPAIILQALTMSNANDGINLERLET
AaDcr1 1621 LTVSTANLNNLTDNP---DEFSFDYQPNLNNHPGPSPSIILQALTMSNANDGINLERLET
                 *     **     * ******* *  ****** **********************

AgDcr1 1688 IGDSFLKYAITTYLYCRYDNVHEGKLSHLRSKQVSNLNLYRLGRRKRLGDCMIAAKFEPH
AaDcr1 1678 IGDSFLKYAITTYLYCTYENVHEGKLSHLRSKQVSNLNLYRLGRRKVLGESMIATKFEPH
            **************** * ************************** **  *** *****

AgDcr1 1748 DNWLPPCYYVPKELEQTLIDAKIPACHWNLADLPDIKRLSCAEICQLVKERARAKRREDV
AaDcr1 1738 DNWLPPCYYVPKELEQALIDAKIPACHWNLADLPDIKQLSSEEICRLVKERALALGLLD-
            **************** ******************** **  *** ****** *    *

AgDcr1 1808 DRFDLQHADSNGGEDGNEDDDEDDEDDDDDDDDTTTGEGNEHEADNGSCYIPYNLVTQHS
AaDcr1 1797 -----------------DSDNEEGETMKEIQQLETRNDEFIKNAEYFACFIPYNLVTQHS
                           * *  *         *        *    * **********

AgDcr1 1868 IPDKSVADCVEALIGAYLIECGPRGALLFMAWLGIRVLPIREPPVKLNSNNETALTPYKA
AaDcr1 1840 IPDKSVADCVEALIGAYLIECGPRGALLLMAWLGIRVLPVYEVPY---DENNPKVPGSSQ
            **************************** **********  * *      *

AgDcr1 1928 TGQNDGPLSTGVTIAEYGHWVAPPSPMVRANITFGGIETGAAATSRELARLLQGFEEFEQ
AaDcr1 1897 RYEKDGIAMQSV----YGYWVPPKSPLLMY----------APNPQQTLEHLLDGYSAFED
                **     *    ** ** * **              *      *  ** *   **

AgDcr1 1988 ALGYRFRDRSYLLQAMTHASYSPNRLTDCYQRLEFLGDAILDYLITRHLYEDRRQHSPGA
AaDcr1 1943 SLGYH-HSRSYLLQAMTHASYCPNRLTDCYQRLEFLGDAVLDYLITRHLYEDPRQHSPGA
             ***    ************* ***************** ************ *******

AgDcr1 2048 LTDLRSALVNNTIFASLAVRHGFHKYFLHLSPGLQEVIDRFVRIQQENGHRITEEEYYLP
AaDcr1 2002 LTDLRSALVNNTIFASLAVRHDFHKYFRHLSPGLNDVIDRFVRIQQDNGH-IISEEYYLI
            ********************* ***** ******  ********** *** *  *****

AgDcr1 2108 DEDDELGEYGAMGEDGPGEGRGVGEAEDVEVPKALGDVFESIAGAIFLDSDMSLDTVWKV
AaDcr1 2061 SEDE------------------CDEAEDVEVPKALGDVFESVAGAIFLDSNMSLDAVWKV
             **                   ***************** ******** **** ****

AgDcr1 2168 YRKMMGPEIEKFSSSVPKSPIRELLEMEPETAKFGKPEKLTDGRRVRVTVEVFGKGTFRG
AaDcr1 2103 YRSMMGPEIEQFSNSVPKSPIRELLELEPETAKFGKPEKLADGRRVRVTVEVFGKGTFRG
            ** ******* ** ************ ************* ******************

AgDcr1 2228 IGRNYRIAKCTAAKCALRQLKKLGYANHHK
AaDcr1 2163 IGRNYRIAKCTAAKCALRQLKKMGLINKRR
            ********************* *  *
```

Figure A.12: AgDcr2 and AaDcr2 amino acid alignment (51.7% identity)

```
AgDcr2     6 EDFAPRNYQVQMKEICLAKNTIIFLPTGSGKTYIALMVMKEISHQLRNTVHEGGKRTFFL
AaDcr2     9 DDFIPRDYQRTMKTICMQKNTIIYLPTGAGKTHIALMVIKEMGKDLDKPLTEGGKRTFFV
               ** ** **  ** **  ***** **** *** ***** **    *     ********

AgDcr2    66 ANTVALAKQQAQFFARHMPFNVRLYTSEVNVDAWKSDRWHEEFSEGQVIICTAQILLDVL
AaDcr2    69 VNTVALAKQQAEFLSHNLTYDTSIYTSDRNVDAWKQDKWLEEFAKYQVIVCTCQILLDVL
               ********** *           ***  ****** * * ***   *** ** *******

AgDcr2   126 RHGYMSPANINLIVFDECHRAVGQHPMHAIMKEIVAAPASERPRVLGLSGTLLFKELKMA
AaDcr2   129 KHGYLSVKHINLLIFDECHHGVGEHPMHGIMEQFLRVPKSDHPRVIGLSGMLLYKQIKSV
               *** *   ***  *****  ** **** **      * *  *** **** ** *  *

AgDcr2   186 SQVPDELERLENTFSSTIATVANYDDYATVASFSTNPNEVLVTYSKPAVHLMPLVKEMWP
AaDcr2   189 ALVSPELERLENTFNATIATVGSYDAFTEVCKFSTDPNELLVSYS--TLRLSPVMADIVN
               *  *********  *****  **    *  *** *** ** **     * *

AgDcr2   246 RIDAFVEWLLQVYLPEYSTQSTRTLQKSFCKPLKEVKRALTEFKHQLEEHGMYAGSLAIL
AaDcr2   247 NINAFSQTIEEFHLPKYLNQN-KALLKDRPKPLKEIRKLFTELIYQLGDTGLFGGSIALL
               * **        ** *  *    * *   *****     **   **   *   ** * *

AgDcr2   306 AVIVQLEVSKRQSPCDKARQVYRSAISFCESIRHELVEAMSGLRGTHQILSFSSDQARKL
AaDcr2   306 GLIVQFELDKRQSDSSMLRLALRSCITFCESLRHQIEKLMSGLDMKTKLTKFSSLKVRQL
               *** *  ****     *   ** * **** **     ****        ***   * *

AgDcr2   366 LKYLEDSYRTAEDKNKQALVFVKRRFTAKVLYHLIRIYFHCLSKRDNEDELVEPIVKPDF
AaDcr2   366 IDQLEKLYEENRDKKAKTLIFVQRRFSAKVLYHLLKIYF-----AETEDA---NLIVPDF
               **  *    **    * ** *** *******  ***         **        ***

AgDcr2   426 IVGANAALEESIDAILVVREDRRVIENFRKRKINVLCATNVLEEGIDLQMCNMVIMYDAP
AaDcr2   418 MVGNNGSMPESIEQILSAKKDRRVLERFKKNETNVIVTTNVLEEGIDLQMCNTVVKYDHP
               ** *    ***  **    **** * * *   **   ************** *  ** *

AgDcr2   486 LSYASFMQSKGRARMKTSTYLMMTPAADLQQFAKRMKLYRDIENRLKEELVGKTINRPEP
AaDcr2   478 QTFASYQQSKGRARMKNSQYMVMLDNENRHIFLEKYRLYKSIEEELRRCLIGKTINRPDP
               **  ********* * *  *        *     **  **  *   * ******* *

AgDcr2   546 LENDVRKELLDDLIPPFYTPFKAKLDALSAIQLLNRYCMSMPRDLFTGSNVTWERIDRSP
AaDcr2   538 LDADVHKELYNEIIPPFFTAKGAKLDALSAIQLLNRYCMGMPRDAFTNTNVTWERIDLKD
               *  ** ***    **** *   ***************** **** **  ********

AgDcr2   606 TEIIVTVKLPLQSTVREVIHGQTMKNLKLAKQSAAFNACKRLFEVGELNMYLLPIATKDK
AaDcr2   598 GRIIVEVLLPLQSTVREKISGNPMRNIKLAKRSAAFNACRKLYENKELNEHLIPIDCKYQ
               *** * ********* * *  * * **** *******  * *  ***  * **  *

AgDcr2   666 VEELSEQYFKLWRKMSDEPNPKQAGTMKYVRGHKIVYPEETVGCTPQADGEQCYVYIVRM
AaDcr2   658 LNNLKDVYFRHWKDF-DADLGKLAGTQKCIRTHAIQYPKQTTECFPQP-GKPCYIYVLRI
               *   **  *    *    * *** *  * * * **  *  * **  *  ** *  *

AgDcr2   726 RAHFDANTHLENVRIFQELYSSANNFGILTRKRLPRLARMKLFVTLGAIGVEIVPEPVCI
AaDcr2   716 AAGFAQDPTNDNVNIFHSLYSSENNFGLMTTKPLPALAKMKFFVTLGLINVHIEETPIVL
               * *       ** **  **** ****  * * ** ** ** ***** * * *   *

AgDcr2   786 TLAPDSGELQRLKRFHLLLFRDLLKVWKPFTVLDALPEENGFLIVPMLRSQSIDWELMGK
AaDcr2   776 PNGGSEIELALLRQFHVTVFRDVLKLWKEFLCCDYDNEENSFLVVPLKNSTHLDWKLIRE
                     **  *  **   *** ** ** *   *   *** ** **   *   ** *
```

```
AgDcr2  846 FPYL-RPAAETSTRARQHLRFETEQYLLRVVHPWYKNDPDQNYVVVRVRDDLRPTSPFPN
AaDcr2  836 FQNLSEPPSEISTIARNKMEFEADKYRHKVILPWYKNNKEQPYVVTMVHEHLTPESPFPN
            *  *  *  * ** **    **   *   *  *****   * ***  *   * * *****

AgDcr2  905 AKYDSYEQYFAQEHHQVVVRHEDQFLIEVKGITTSLNRLHPGAEADGGASTRSRYWEFQE
AaDcr2  896 PEYGSYANYFSQAYHLAVVK-PDQFLIEVKGITSYLNRLNPGVE-DDGKSTRSKHWRFNE
              * **  ** *  *  **   ***********  **** ** * * * ****  * * *

AgDcr2  965 ILIPELVHNFEFPADYWLKATLLPSALHRVHYLLLAEGIRVDLARNAQVACR-------T
AaDcr2  954 ILIPELCHNYQFPADYWLKATLLPSALHRLHYLLLAENIRVDLATGANVGCLENHTIEDV
            ****** **  ***************** ******* ******  * * *

AgDcr2 1018 AVDPKQPSGKQLLQV-LNQIHSDS----------------------------ISLADRV
AaDcr2 1014 DVEYKERKGKQLEELQLMEFEEDEDEDDEFDLEEAKRSLVAPENLSELARNQMCSITGDI
             *  *   ****    *     *                                 *

AgDcr2 1048 QYPWEENEQPKDLERNWDTVSKIDIDYYASFVKKYEQQAVTIVTDVNRSVTETNIDRMLA
AaDcr2 1074 PLPWQEDEEPVDIERNWDQVSKLDLDYYNVFVNKFSDLSMREKAAERISTAYTSA-----
              ** * * * * ***** *** * ***  ** *              *   *

AgDcr2 1108 NLNMNKNEAEASPVKALPALEDGHGPVAQIAMLRLTMDNTANVALQQSDLLQALTTKSSA
AaDcr2 1129 ---VYRRAAGSPKREPMAILDVPVDQKFAIKLLQLTPANTVNVNLQQKNIIKALTTKSSS
                   *          *         *  * **  ** ** ***     *******

AgDcr2 1168 DVFNLERFEVLGDAFLKFAVSVYILFRHTSWHEGYLTTCKGRMVSNRNLLYCAMGYGLPG
AaDcr2 1186 DVFDLERYELLGDAFLKFSISLYLVKYHKEWHEGFLTAVKGQIVSNRNLVYCAIKYGLPG
            *** *** * ********  * *    *  **** **  **  ****** ***  *****

AgDcr2 1228 KIKAHPFDPKNDWVPPLSTVPGAVRRAMVDANESPALLYNLKLTEEEIQSGVVAQATVDK
AaDcr2 1246 MLKIHKFDPKNDWQPPLATVPKNIKRTMQSVNHSARVLYRLTLTEEEIKTGVVTAKNSDD
              * * ******* *** ***    * *   * *   ** * ******  ***     *

AgDcr2 1288 FLPLIE-QAPAPSQSTLHTVLQQVQIRDKVVADVTEALLGVCVKTVGYERSFRFLSHLGI
AaDcr2 1306 FIAQLELHGNMPDPSPMANYLSQQTMGDKTPADAMEALLGVCVQSVGIERSFKLLPHFGI
            *    *     *  *     * *    **  **  ********  ** ****  * * **

AgDcr2 1347 IPKGADVPMLL--RTTTFPIGDYFPVRHKVDQLLCNPERIEATLGYRFRNRTYLLQAFTH
AaDcr2 1366 LPKTHNVLRLLADKIENQRLKTHIDIR-EVDAFLKNYRRIEGILGYKFKDRTYLLQALTH
             **   *  **               *  **  * *  ***  *** *  ******* **

AgDcr2 1405 TSYTSNSLTGSYQQLEFLGDAVLDFLVSMYIYERNPSMSPGQLTDLRSALVNNVNLACVL
AaDcr2 1425 ASYPTNRITGSYQQLEFLGDAVLDFLISMYIFEQNPTMSPGQLTDLRSALVNNVTLACIL
             **  *  ***************** **** * ** ***************** *** *

AgDcr2 1465 VRNELHRHILSQSPMLTDAIGKFVAVHRCHRNQGSNWVRLLTEESDTPMAEYVDVPKVLG
AaDcr2 1485 VRHGLHLYILAESASFTDTVSKFVLFQEQHKHEITDQVNLLVEESDRKMAEFVDVPKALG
            **  **  **  *   **   ***     *       * ** ****  *** ***** **

AgDcr2 1525 DVLEALIGAIYLDSGNDLAATWEVCFRLLRDEIADFTRKTPIQVVRQLYEH-PEASPHFS
AaDcr2 1545 DVFESLVAAVFLDSGNDFAATWQVIYGMMGNEILTFTENTPIQIVRQLYEFKPSCKPTFS
            ** * *  *  ****** **** *       **  **  **** ******  *   * **

AgDcr2 1584 APFVEEEVVYVKLSYTHRSQRQTVYGFGKNKDDAKRAAAKIAL
AaDcr2 1605 RAIPDEDTVLVKLRYEIRNQQHEAYGFGQNKDDAKRAAAKAAL
                 *  * *** *  * *    **** ********** **
```