Discovery, Characterization, and Functional Analysis of microRNAs in Culicidae

Edward Andrew Mead

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Biochemistry

Dr. Zhijian Jake Tu, Chair
Dr. Glenda Gillaspy
Dr. John Jelesko
Dr. Kevin Myles
Dr. Thomas O. Sitz

May 13, 2009
Blacksburg, VA

Keywords: microRNA, RNAi, Anopheles, Aedes, malaria, dengue, mosquito, small RNA

Discovery, Characterization, and Functional Analysis of microRNAs in Culicidae

Edward Andrew Mead

Abstract

MicroRNAs (miRNAs) are non-coding RNAs that often play a fundamental role in gene regulation. Currently, hundreds to over a thousand miRNAs are predicted to be present in many eukaryote species, with many to be discovered; the functions of most are unknown. While much attention has gone towards model organisms, a much greater depth of understanding remains to be gained for the miRNAs of many organisms directly important to humans. There are few verified miRNAs for any mosquito species, despite the role of mosquitoes in many of humanity's worst diseases. *Anopheles gambiae* and *Aedes aegypti*, carriers of malaria and dengue, respectively, are responsible for over a million deaths a year. To date, there are sixty-six microRNAs in *An. gambiae* in miRBase, a central repository for miRNA sequences. Many of these are based on homology to primarily *Drosophila* miRNAs. While sequence conservation suggests an important function for these miRNAs, expression has not been experimentally verified for most mosquito miRNAs.

Using small RNA cloning and northern blots, I discovered and analyzed 27 different microRNAs in aged female *An. stephensi* mosquitoes, the age group responsible for transmission of malarial parasites. Three of these miRNAs are only found in mosquitoes (miR-1889, -1890, and –1891). Cloning and northern analysis revealed an abundance of a miRNA that is linked to longevity in flies, miR-14, across different life stages of mosquitoes. It was also shown that miR-989 was expressed almost exclusively in the adult ovary and its expression fluctuated in response to bloodfeeding, suggesting a possible role in reproduction, an area of great importance to controlling mosquito populations.

Building upon the above cloning experiment, a later high-throughput sequencing effort uncovered 98 miRNA precursors from *Ae. aegypti*. There are a total of 13 novel miRNAs that have not been found in other organisms by bioinformatic predictions or experiments. These "mosquito-specific" miRNAs may play a role in processes such as blood-feeding or vector-host interactions. A detailed examination of the expression of eight of these miRNAs was conducted in *An. gambiae*, *An. stephensi*, *Ae. aegypti*, and *T. amboinensis* to determine their expression profile, conservation, and provide hints to their function. My work revealed conserved and sometime stage-specific expression profiles of some of the mosquito-specific miRNAs. I also provided evidence for three lineage-specific miRNAs that may shed light on the divergence of different mosquito lineages.

Extending the finding that miR-989 may be involved in mosquito reproduction, we conducted a detailed analysis of its evolution, expression, possible targets and regulation. miR-989 is conserved in holometabolous insects. miR-989 expression in female *An. stephensi* and *Ae. aegypti* dramatically rises following pupal emergence until strong signal is observed, until a blood meal is taken. Expression remains quite strong then begins a steep decline in expression at 32-40 hours post blood meal (PBM), and even by 96 hours PBM, remains weak. Bioinformatic predictions of miR-989 targets coupled with a PCR-based approach uncovered three potential target leads, though preliminary results were artifacts. Although the miR-989 post-emergence expression profile correlates with the expression of Juvenile Hormone, a key reproductive hormone in mosquitoes, no observable induction occurred when abdominal ligation samples were administered methoprene, a JH analog. However, methoprene impacted a number of other miRNAs, with up to a 3.87 fold induction (miR-1891), and a 3.15 fold suppression (miR-9a) of signal. Subsequent northern analysis provided visual confirmation of observable fold changes for miR-1891 and miR-9a, but not for miRNAs that showed changes below two fold. This analysis provides a foundation to study Juvenile Hormone regulation of miRNAs in mosquitoes. In summary, we have expanded the understanding of microRNAs in mosquitoes. An improved understanding of mosquito physiology can assist in efforts to control mosquito-borne infectious diseases.

Acknowledgements

I would like to thank my advisor Zhijian Jake Tu for his direction, faithful support, and willingness to discuss projects. I also would like to thank my committee members Glenda Gillaspy, John Jelesko, Kevin Myles, and Tom Sitz for their suggestions, guidance and patience, and friendly conversations during my Ph.D. process.

I would like to thank all of my labmates through the years. Jim Biedler and Yumin Qi went out of their way to help me and give me advice on numerous occasions, and offer a friendly ear. Thanks to Fan Yang for all of his help with formatting. I would also like to thank the Adelman, Li, Myles and Zhu labs for all their friendship, assistance, and keeping me on my toes.

Lily Zhang has been a source of inspiration to me through the dissertation process, and has always been there with cheerful words and friendship. Thanks to Brian Jordan, and Chang-su Lim for their friendship and advice. I would like to say “xie xie nimen” to Wen Dou, Haiyang Fu, Ying Jiang, Fangning Liu, Shaohui Wu, Ai Xia, and Xing Zhang for their friendship, assistance, and fun memories.

Lastly, I would like to thank someone I didn’t know named Federica for letting me use her retired computer in the lab. To anyone else I may have forgotten to thank in my old age or from running too many acrylamide gels during the last few years, my sincerest apologies and I thank you as well.

Table of Contents

List of Figures

Chapter 1.

Chapter 2.

Chapter 3.

Chapter 4.

List of Tables

Chapter 2.

Chapter 3.

Chapter 4.

List of Abbreviations

Ae. aegypti	*Aedes aegypti*
Ago	Argonaute
An. gambiae	*Anopheles gambiae*
An. stephensi	*Anopheles stephensi*
BF	blood-fed
bp	base pair(s)
CA	corpora allata
C. elegans	*Caenorhabditis elegans*
CSP	circumsporozoite protein
Dcr	Dicer
DGCR-8	DiGeorge syndrome critical region 8
DHF	dengue hemorrhagic fever
DIG	digoxygenin
D. melanogaster	*Drosophila melanogaster*
DNA	deoxyribonucleic acid
d.o.	day(s) old
ds	double-stranded
ET	Early Trypsin
Exp-5	Exportin-5
JH	juvenile hormone
kb	kilobase(s)

LB	Luria broth
LNA	locked nucleic acid
Luc	luciferase
miRNA	microRNA
miRNP	microribonucleoprotein
MPSS	Massive Parallel Signature Sequencing
mya	million years ago
NBF	non-bloodfed
NS	no score
nt	nucleotide(s)
oligo	oligonucleotide
PCR	polymerase chain reaction
P. falciparum	*Plasmodium falciparum*
piRNA	piwi-interacting RNA
piwi	P-element-induced wimpy testis
pol	polymerase
pre-miRNA	precursor microRNA
pri-miRNA	primary microRNA
qRT-PCR	quantitative real-time PCR
RAKE	RNA-primed array-based Klenow enzyme
RBC	red blood cell
rasiRNA	repeat-associated small interfering RNA
RITS	RNA-induced initiation of transcriptional gene silencing

RNA	ribonucleic acid
RNAa	RNA activation
RNAi	RNA interference
rRNA	ribosomal RNA
SF	sugar-fed
SFM	serum free media
S. pombe	*Schizosaccharomyces pombe*
siRNA	small interfering RNA
sm	small
ss	single-stranded
T. amboinensis	*Toxorhynchites amboinensis*
TE	transposable element
tRNA	transfer RNA
UTR	untranslated region
UV	ultraviolet
WT	wildtype

Chapter 1

Introduction

1.1. Mosquito biology and vector-borne disease

1.1.1. A mosquito primer

Mosquitoes are "true flies", members of the order Diptera (Eldridge, 2005). Among the approximately 4500 species of the mosquito family, Culicidae (Besansky, 2006), there are two major subfamilies: Anophelinae, and Culicinae (Harbach and Kitching, 1998; Mitchell et al., 2002; Harbach and Kitching, 2005; Harbach, 2007). Toxorhynchitinae may be part of Culicinae, though there is insufficient evidence to verify whether they are monophyletic (Harbach and Kitching, 1998; Harbach, 2007), and it is often still categorized in the literature as a third subfamily. The mosquito family Culicidae contains vectors for some of the most deadly diseases known to mankind, which impact the lives of hundreds of millions of people each year (Gubler, 1998).

Mosquito adults are dimorphic, with readily visible differences between males and females. Adult males tend to emerge from egg cases earlier than females. Males have robust antennae and maxillary palps, and use an organ at the base of the antenna, the Johnston's organ, to detect females by the frequency of the wing beat (Gibson and Russell, 2006). Females are typically larger than males. Most females will only mate

once with a male in the wild, and store sperm in one to three organs known as the spermethecae (Gwadz and Collins, 1996; Eldridge, 2005). While both males and females have a slender proboscis for nectar feeding, in the females of some species there are additional stylets encased in the proboscis for hematophagy (Eldridge, 2005; Black and Kondratieff, 2005). These females also produce anticoagulants and vasodilators in their saliva to enhance bloodflow during feedings (Black and Kondratieff, 2005).

Hematophagous mosquitoes are drawn towards animals by environmental cues including temperature, CO_2, and sweat, key indicators of prey (Enserink, 2002). For anthropophilic bloodfeeders such as *An. gambiae* and *Ae. aegypti*, fine detection of odors such as ammonia and lactic acid inform them they have found a human (Enserink, 2002). Bloodfeeding females will periodically take a bloodmeal (in some cases, up to 4 times their body weight) to acquire a sufficient quantity of amino acids to generate the yolk protein vitellogenin (Holt et al., 2002). Approximately 2-3 days later, the females of most species lay a clutch of approximately 50-500 embryos in stagnant water (Gwadz and Collins, 1996). Aedine females prefer to lay on damp surfaces where flooding may not occur for months. As such, their eggs can resist drying for up to several years (Clements, 1992; Eldridge, 2005).

Once leaving their egg cases, mosquitoes progressively increase in size through 4 aquatic larval instars followed by an aquatic pupal stage before eclosion (emerging as a winged adult) (Eldridge, 2005). The larval stages are characterized by vigorous filter-feeding behavior for most species, though larval members of the subfamily Toxorhynchitinae

feed upon the larvae of other mosquitoes (Eldridge, 2005). Like other Dipteran insects, mosquitoes are holometabolous, undergoing a dramatic metamorphosis between juvenile and adult morphology (Clements, 1992). Emergence typically occurs within two days of the pupal transition (Eldridge, 2005). Juvenile Hormone (JH) plays an important role in development, and must be absent in many insect pupae to complete transformation into adults (Hagedorn, 1996). Following emergence, adult females undergo a dramatic increase in the expression of JH, and this hormone has been found to play a major role in ovarian development (Hagedorn, 1996). In all, from egg to adult takes about one to two weeks depending on environment and mosquito species (Gwadz and Collins, 1996).

1.1.2. *Anopheline* mosquitoes

Anopheles is a genus in the subfamily Anophelinae which serve as a vector for malaria. There are approximately 500 species of *Anopheles,* in six known subgenera (Gwadz and Collins, 1996; Besansky et al., 2005). Evidence suggests that *Anopheles* split away from *Drosophila* approximately 250 million years ago (mya), and from *Aedes* approximately 145-200 mya (Gaunt and Miles, 2002; Krzywinski et al., 2006; Nene et al., 2007). Representatives of Anopheles may be found on every continent except Antarctica, and species diversity is densely concentrated in the tropics (Gwadz and Collins, 1996).

Though *Anopheline* mosquitoes can transmit the causative agents of several different diseases, they have been most studied for their role as vectors of *Plasmodium* parasites,

the causative agent of malaria (Gwadz and Collins, 1996). The anthropophilic *An. gambiae* is a major vector of malaria in sub-saharan Africa (Besansky et al., 2005). In the Middle East and Southeast Asia, the more zoophilic *An. stephensi* is a major vector of malaria (Brown et al., 2003).

1.1.3. Malaria overview and *Plasmodium* biology

Malaria takes a terrible toll upon mankind. Currently an estimated 300-500 million individuals, primarily in tropical areas, are infected each year (Hartl, 2004). Among these, one to three million individuals, primarily children under five years of age, die annually (Hartl, 2004).

Malaria is a disease caused by protists of the genus *Plasmodium*. More than 200 members of *Plasmodium* are known (Perkins, and Austin, 2008), and at least four have been implicated in human malaria; all are carried by *Anopheline* mosquitoes (Winzeler, 2008). Of chief concern amongst the *Plasmodium* species is *P. falciparum*, responsible for the most severe form of malaria, which results in the majority of malaria-related deaths in humans (Winzeler, 2008).

The transmission of malaria occurs almost exclusively by mosquito vectors during bloodfeeding (http://www.who.int). Beginning with a bloodmeal from an infected individual, parasitic gametocytes are taken up by the mosquito, where they pass through

several succesive stages, ultimately emerging from midgut oocysts as sporozoites which migrate to the salivary glands of the mosquito (Doering et al., 2008). Humans are infected during a subsequent bloodfeeding (Doering et al., 2008). In all, the transition from gametocyte to infective sporozoite occurs over the course of 10-21 days, depending on environmental factors, the species of parasite, and the species of mosquito involved (http://www.cdc.gov/malaria/biology).

After introduction to a human host by a mosquito, the sporozoites rapidly migrate to the liver where they infect hepatocytes and proliferate, changing into merozoites (http://www.cdc.gov/malaria/biology). Merozoites are periodically released from the liver to infect red blood cells (http://www.cdc.gov/malaria/biology). The parasite lives intracellularly, feeding on hemoglobin (http://www.cdc.gov/malaria/biology). Some merozoites will differentiate into gametocytes which may be taken up by a mosquito, beginning the cycle anew (http://www.cdc.gov/malaria/biology).

1.1.4. The struggle against malaria

The severity of malaria has led to a variety of attempts to control or erradicate either the parasite or the vector, with incomplete success (http://www.cdc.gov/malaria/biology).

Plasmodium infection may be fought with antimalarial compounds, though resistance to many of the older quinine derivative drugs has been developing (Gould et al., 2006). Artemisinin and related compounds have proven highly effective, particularly when

combined with other antimalarials (http://www.who.int). Other drugs that appear promising, including "G25", are still undergoing trials (Wengelnik et al., 2002).

Developing vaccines against the parasite also faces difficulties. Many of the parasite's antigens remain hidden inside host cells (reviewed in Chen et al., 2000). The parasites display antigenic switching, allowing rapid evolution of antigens on the surface of infected red blood cells (reviewed in Chen et al., 2000). Thus, there is tremendous variation among the displayed surface antigens of a population of parasites making a vaccine difficult to design (Gould et al., 2006). In particular, a key surface antigen involved in red blood cell (RBC) agglutination is *P. falciparum* erythrocyte membrane protein 1 (PfEMP1) (Chen et al., 2000). PfEMP1 has been found to have a variety of variants (reviewed in Chen et al., 2000). Currently, some vaccines directed against circumsporozoite protein, CSP, an antigen displayed at the sporozoite stage, have shown promise (Matuschewski, 2006). However, no vaccines developed thus far have completely immunized against malaria (Matuschewski, 2006).

The battle against malaria has also been fought against its mosquito vector. Current methods range from simple physical implementations to complex biological control (Spielman, 1994). Mosquito control methods can be separated into individual protective methods and regional protective methods (Eldridge, 2005). Very effective individual methods include mosquito netting on individuals and in homes, and using DEET (Eldridge, 2005). Draining sources of stagnant water and educating the populace about mosquito control methods can be effective regional protective methods (Spielman, 1994,

Eldridge, 2005). Both individual and regional approaches can be prohibitively costly for rural third-world communities to implement (Spielman, 1994).

Organochlorine insecticides such as DDT proved useful in virtually erradicating malaria from the Western Hemisphere and in limiting it in several tropical countries; however, cost as well as concerns over safety, resistance of the insect, and environmental issues have restricted their use (Morel et al., 2002, Eldridge, 2005). Though more costly, new developments in biological insecticides, referred to as "third-generation insecticides", have proven to be more environmentally safe (Eldridge, 2005). These include microbial insecticides such as *Bacillus thuringiensis israelensis*, and growth regulators such as altosid, among others (Eldridge, 2005). Sterile Insect Technique, whereby large colonies of sterilized males are released to overwhelm local populations of females can lead to localized erradication, but would prove extremely difficult to implement over an entire continent such as Africa (Coleman et al., 2004). Further, local communities may resist the idea of being innundated with more mosquitoes (Aultman et al., 2000). Other possibilities include natural biological control strategies like mosquito-eating fish, mosquitoes which eat larval stages of vector mosquitoes (*Toxorhynchites*), and protozoa or fungal diseases (Scholte et al., 2004).

Replacement of existing populations of mosquitoes with genetically engineered mosquitoes that are refractory to malaria is another approach (Sinkins et al., 2006). Mosquito transformation has proven difficult, but can be done, and there have been recent improvements in site-directed genetic engineering using a phiC31-based system in

Ae. aegypti (Nimmo et al., 2005). A phiC31 bacteriophage integrase inserts a transgene into phiC31 integration sites in the genome (Chalberg et al., 2006). Recently it was found that transgenic mosquitoes could be generated that express an anti-parasitic peptide (SM1). These mosquitoes demonstrated increased fitness when feeding on *Plasmodium*-infected blood compared to wildtype (WT) mosquitoes (Marrelli et al., 2007). However, successfully transformed genotypes would need to be driven to fixation in a population by a gene drive mechanism (Nimmo et al., 2005).

While genetic modifications can be made to yield a mosquito which is refractory to malaria or dengue, evolutionary time frames would be required if the genetic construct relied upon strict Mendelian heredity to spread through a population. It would be particularly difficult to spread a transgene if there was an associated fitness cost to the organism, which is common with genetic engineering (Cha et al., 2006). This has made gene drive elements particularly important to vector borne disease research. Gene drive elements provide a means to force a gene of interest into a population to fixation at rates faster than would be allowed by standard Mendelian heredity, potentially even with associated fitness costs (Cha et al., 2006).

Current drive mechanisms, such as transposable elements, homing endonuclease genes (HEGs), or infection with symbiotic *Wolbachia*, still have significant hurdles to overcome if they are to succeed (Coleman et al., 2004). Transposable elements, DNA capable of moving to different sites within a genome, have been used to genetically engineer insects for a quarter century (Moreira et al., 2002). Natural genomic parasites,

TEs are excellent transgenic vectors as they have evolved to invade hosts and stably integrate into their genomes. Transposons can multiply through a genome mobilizing a gene of interest, and thereby offer non-Mendelian patterns of inheritance bypassing evolutionary timeframes for achieving gene fixation (Gould et al., 2006). This is best demonstrated by the rapid spread of P element in *D. melanogaster* (Engels, 1997). Unfortunately, transposition imposes fitness costs in moving to different locations; there is also the possibility of losing the transgene, or silencing of the element (Coleman et al., 2004; Gould et al., 2006). Many TEs have degenerated over time and are unable to transpose. Further, transposons are often derived from *Drosophila* and other insects and do not remobilize well in mosquitoes, limiting their use as a gene drive mechanism (Moreira et al., 2002; Coates, 2005). Low rates of transformation also are the norm. One of the best TE-based vectors available for *Ae. aegypti*, piggyBac, has an 8% success rate for transformation efficiency (Ninmo et al., 2006). HEGs are another potential gene drive mechanism (Sinkins et al., 2006). In yeast they are copied only to a similar site in the homologous chromosome, limiting the likelihood of damage (Sinkins et al., 2006). *Wolbachia* offer a potential gene drive mechanism through cytoplasmic incompatibility, through selective pressure for infected organisms.

One exciting development in recent years has been the discovery of the maternal-effect dominant embryonic arrest (MEDEA) element in the flour beetle, *T. castaneum*. MEDEA is a naturally occurring maternal selfish genetic element that selectively leads to the death of progeny that do not carry the maternally-inherited MEDEA construct (Chen et al.,

2007). Studies of a synthetic MEDEA construct in *Drosophila* have demonstrated successful gene drive (Chen et al., 2007).

In the *Drosophila* design, MEDEA relies upon the maternal expression of a "toxin", a miRNA which targets a critical zygotically-expressed mRNA for knockdown and an "antidote" which is an early zygotically expressed mRNA coding for the same protein, lacking miRNA target sites (Chen et al., 2007). In Chen et al., the toxin gene codes for artificial miRNAs to knockdown expression of Myd88, crucial to embryonic development (Chen et al., 2007). The antidote is a Myd88 gene that codes for a transcript lacking 3'-UTR sites for the miRNAs, thus restoring expression of Myd88 (Chen et al., 2007). In *Drosophila* cage studies, when this element entered into a population at an allele frequency of approximately 25%, selection for MEDEA led to fixation in a population in approximately 20 generations across 7 cage trials (Chen et al., 2007).

There is a push in the vector-borne disease field to use combinations of the above methods to attack the problem (Sinkins et al., 2006). While any one of these efforts by itself may prove insufficient to rid an area of malaria, the multiplicative effect should be devastating to the disease (Sinkins et al., 2006). While costly and labor-intensive, such combinations may be our best hope at fighting malaria.

1.1.5. *Aedine mosquitoes*

Aedes is a genus in the subfamily Culicinae and is an important vector for multiple arboviruses including Chikungunya virus, yellow fever virus, La Crosse virus, and the dengue viruses (Nasci et al., 1996). *Aedes* mosquitoes are found throughout the world (Nasci et al., 1996) Recent systematic studies have suggested a nomenclature change, elevating *Stegomyia* to the genus level, resulting in *Aedine* mosquitoes being renamed as *Stegomyia* (Reinert et al., 2004).Several other species of significance for vector-borne disease research include *Ae. albopictus*, and *Ae. vexans*, and the anthropophilic *Ae. aegypti* (Enserink, 2002).

1.1.6. Dengue viruses, yellow fever virus, and chikungunya virus

Dengue fever is a serious disease through the tropics and subtropics. Estimates suggest that dengue affects approximately 50 million people annually, with roughly 1% of dengue cases developing into the deadlier dengue hemorrhagic fever (DHF) (Mackenzie, 2004; Nene, 2007). Approximately 1-5% of individuals contracting DHF die from the disease (Senior, 2007). Dengue is a disease caused by four Flavivirus serotypes of dengue virus (Nasci et al., 1996). Interestingly, prior exposure to one of the four serotypes increases sensitivity to the remaining three, increasing the likelihood of developing the more serious DHF (Nasci et al., 1996).

Dengue is transmitted through *Aedine* mosquitoes and impacts Asia, Africa, and the Americas (Nasci et al., 1996). Recent concerns have been mounting over the spread of the Asian tiger mosquito, *Ae. albopictus*, across temperate regions including much of the

United States and parts of Europe (Enserink, 2008). There is evidence that dengue viruses can replicate in and can be transmitted by *Aedes albopictus*, opening the possibility for a major dengue outbreak within the United States (Enserink, 2008).

Yellow fever virus is another Flavivirus transmitted by Aedine mosquitoes (Nasci et al., 1996). Yellow fever has a significantly higher mortality rate than malaria and dengue. Estimates suggest that approximately 200,000 people contract yellow fever annually resulting in approximately 30,000 deaths per year (Tomori, 2004; Nene, 2007). Yellow fever is preventable as there is an effective vaccine readily available (Nasci et al., 1996).

Chikungunya is an alphavirus transmitted by *Ae. aegypti* (Enserink, 2007). It causes fever, rashes, and pain in the joints. No vaccine currently exists to treat chikungunya.

1.2. Non-coding RNAs and their classification

Non-coding small RNA pathways have been identified across a diverse array of eukaryotes and even a species of single-celled algae, *Chlamydomonas reinhardtii* (Zhao et al., 2007). RNA interference (RNAi) utilizes small RNAs known as siRNAs (short interfering RNAs) to silence gene expression through the recognition and degradation of mRNA sharing a complementary sequence (Du and Zamore, 2005). RNAi is thought to be an ancient pathway that may have first arisen early in eukaryotes as a way of protecting cells from viral invaders (Zamore, 2002). Post-transcriptional gene silencing (PTGS) can occur through the canonical RNAi pathway involving RISC-mediated

transcript degradation, as well as by suppression of translation involving the miRNA pathway in metazoans (Kim, 2005). Gene silencing can occur by rasiRNAs (repeat-associated short interfering RNAs) which permit histone or DNA methylation of recognized sequences, resulting in heterochromatin formation (Kim, 2005).

A common thread among these diverse pathways is the use of double-stranded regions of RNA processed to smaller, single-stranded RNA which is selectively paired with a complementary target, providing specificity in the silencing (Du and Zamore, 2005). Depending on the particular pathway, small RNAs may be generated specifically or at random from precursor dsRNA molecules (Du and Zamore, 2005).

MiRNAs function in stem cell differentiation, in development, in the regulation of transcription factors, even in the regulation of miRNA expression (Du and Zamore, 2005; Kim, 2005). Given the range of biological roles of miRNAs in eukaryotes, we suspect that in-depth studies of miRNAs in mosquitoes may provide a better understanding of gene regulation and the physiology of mosquitoes, assisting current efforts of vector-disease control.

1.2.1. MicroRNAs and their biogenesis

MicroRNAs (miRNAs) are non-coding RNAs that have been found across Eukaryota and in some viruses that infect them. Genes coding for miRNAs can occur in intergenic regions, though in mammals miRNAs were found to occur primarily within introns of

protein-coding genes (Miska, 2005; Kim et al., 2007; Liu et al., 2007). MiRNAs in close proximity to one another, known as a "cluster", are thought to often be derived from the same primary transcript as they are frequently observed to have a similar expression pattern (Ruby et al., 2007). Clustered miRNA are typically no farther than 10kb apart (Ruby et al., 2007). MiRNA biogenesis typically begins by RNA polymerase II-mediated transcription (Bartel, 2004; Miska, 2005; Pasquinelli et al., 2005); though in humans it was discovered that RNA pol III also transcribes some miRNAs which are associated with Alu repeats (Borchert et al., 2006). Transcription produces a primary miRNA (pri-miRNA) anywhere from hundreds of bases to more than 10,000 bases in length (Bartel, 2004; Miska, 2005; Pasquinelli et al., 2005; Lee et al., 2004; Cai et al., 2004). The transcript contains a poly(A) tail, and a 7-methylguanosine cap typical of pol II transcription (Kim and Nam J-W, 2006). Primary miRNAs (pri-miRNAs) may contain up to several mature miRNAs within a single transcript (Miska 2005; Lau et al., 2001). RNA editing of pri-miRNA has been observed in at least one case, possibly increasing functional diversity (Luciano et al., 2004). In pri-miR-22 in humans and mice, RNA editing (A to I) was observed in several tissues (Luciano et al., 2004).

Cleavage of pri-miRNAs generates precursor-miRNAs (pre-miRNAs) that are further processed to yield mature miRNAs (Saito et al., 2005). Initial processing of the pri-miRNA occurs at the transcription site (Pawlicki and Steitz, 2008). In *Drosophila*, the pri-miRNA is processed by a Drosha-Pasha complex (Saito et al., 2005). Recent analyses suggested that 90% of known miRNA precursors in *Drosophila* are between 73-102 nucleotides (nt) in length (Stark et al., 2007). Drosha is an RNAase III enzyme, whereas

Pasha (known as DGCR-8 in mammals) is a dsRNA protein (Lee et al., 2003; Han et al., 2004). Together, this complex is known as the microprocessor (Tomari and Zamore, 2005). Drosha measures approximately two RNA helical turns from the loop (~22nt) when deciding where to cleave the pri-miRNA (Tomari and Zamore, 2005). It is thought that the bulges in the stem may be important for microprocessor cleavage and processing (Stark et al., 2007). Drosha cleavage leaves 5'-monophosphate and 2nt 3'-overhangs (Lee et al., 2003; Zhang et al., 2004). Recently, a number of accessory proteins including hnRNP A1, and the RNA helicases p68 and p72 have been discovered to be necessary for *in vivo* processing of pri-miRNAs into pre-miRNAs (Guil et al., 2007; Fukuda et al., 2007; reviewed in Liu et al., 2008).

Pre-miRNAs are exported from the nucleus to the cytoplasm by Ran-GTP and Exportin-5 in *Drosophila* (Bohnsack et al., 2004). Exp-5 is a nuclear transport receptor (Kim and Nam 2006) and evidence suggests that it may protect pre-miRNAs from nuclear degradation (Yi et al., 2003; Liu et al., 2008). Following export, the double-stranded pre-miRNA are recognized by Dicer-1 in flies (Saito et al., 2005). Dicer is a roughly 200 kDa protein (Kim and Nam 2006) that contains a c-terminal double-stranded-RNA binding domain (dsRBD) and a PAZ (Piwi, Argonaute, and Zwille) domain that recognizes 3'overhangs (Zhang et al., 2004). This enzyme complexes with Loquacious (TRBP in humans) and cleaves an approximately 22 bp sequence with 2nt 3' overhangs from the pre-miRNA (Saito et al., 2005). Purified Dicer alone is a functional cleaver of pre-miRNAs, and its cofactors may act to enhance stability and function in the miRNP (micro-ribonucleoprotein) effector complex (Kim and Nam 2006).

The product of Dicer cleavage is referred to as an miRNA:miRNA* duplex, where the antisense strand is referred to as the miRNA* (Lau et al., 2001; Bartel, 2004; Miska et al., 2005). In flies, most miRNA/miRNA* duplexes are sorted into Ago-1-containing complexes (Tomari et al., 2007). The duplex molecules are separated by a helicase and the single strand with the lowest 5' stability is chosen as the mature miRNA (Bartel, 2004). Pyrophosphate sequencing, a method of "deep" sequencing, revealed that approximately 100 times as many miRNAs as miRNA*s were observed for a given miRNA in *C. elegans*, except in cases where both strands could generate functional miRNAs (Ruby et al., 2006; Tyler et al., 2008; Liu et al., 2008). MiRNA cloning studies suggest that the miRNA* strand is rapidly degraded (Bartel, 2004). The biogenesis of miRNAs is summarized in Figure 1.1.

1.2.2. Modes of miRNA action

Mature miRNAs are paired with complementary targets with the assistance of Argonaute in a micro-ribonucleoprotein (miRNP) complex (Mourelatos et al., 2002; Bartel, 2004). In plants, miRNAs tend to have perfect or near perfect complementarity to their targets and bind along the entire length of the transcript (Brodersen et al., 2008). In animals, miRNAs typically target the 3'-untranslated region (UTR) and have less perfect complementarity to their targets (Pasquinelli et al., 2005; Miska, 2005). In the event of perfect complementarity between a miRNA and its target, observable gene silencing is 3-10 fold greater than for partial complementarity (Zeng et al., 2002; Doench et al., 2003;

Wu et al., 2008). The target is either degraded by the RNase III activity of Argonaute if there is perfect (or nearly perfect) base pairing or translation of the target transcript is typically repressed if there is imperfect base pairing (Pasquinelli et al., 2005; Miska, 2005); although in some cases imperfect base pairing may lead to degradation of the transcript (Wu et al., 2006; Giraldez et al., 2006; Wu et al., 2008). A controversial finding in *Arabidopsis* suggests that, similar to animals, the primary method of gene silencing in plants by miRNAs is translational repression (Brodersen et al., 2008).

Translational repression and mRNA turnover by miRNA activity can occur through several possible routes (reviewed in Wu and Belasco, 2008). Initiation of translation is typically targeted, as this is a rate-limiting step (Richter and Sonenberg, 2005; Liu et al., 2008). Translation initiation may be blocked through competition with elF4E for the cap or interfering with 80s ribosomal interaction with the mRNA (Kiriakidou et al., 2007; Mathonnet et al., 2007; Thermann and Hentze, 2007; Wakiyama et al., 2007; Wu et al., 2008). However, some sources find that miRNA may impact initiation in a cap-independent manner (Petersen et al., 2006; Lytle et al., 2007; Wu et al., 2008). Other routes of miRNA translational repression include blocking translational elongation or slowing the ribosomes (Wu and Belasco, 2008). In some instances, imperfect pairing of miRNAs and their 3'-UTR targets causes ribosomes to release prematurely (Petersen et al., 2006; Erson and Petty, 2008). Though repressed transcripts are found in cytoplasmic foci known as P-bodies, evidence suggests that P-bodies likely do not generate translational repression (Bushati and Cohen, 2007; Chu and Rana, 2006; Jackson and Standart, 2007) despite a localized abundance of enzymes that degrade RNA including

Xm1, Dcp1 and 2, and Lsm 1-7 (Ingelfinger et al., 2002; van Dijk et al., 2002; Liu et al., 2005; Sen and Blau, 2005; Wu et al., 2008). Degradation may occur through cleavage of the target by the endonuclease Argonaute when perfect or near perfect complementarity occurs (Zamore et al., 2000). Deadenylation may also lead to degradation even when imperfect pairing occurs (Wu et al., 2006; Giraldez et al., 2006; Wu et al., 2008).

1.2.3. Mirtrons

A new category of pre-miRNAs known as mirtrons have been found to be derived from stem loops in introns and are present in mammals, flies, and worms (Ruby et al., 2007; Okamura et al., 2007; Berezikov et al., 2007). Mirtrons undergo Drosha-independent processing, and are spliced as introns but enter into the miRNA pathway to function as as miRNAs (Ruby et al., 2007; Okamura et al., 2007), summarized in figure 1.1 (Kim et al., 2009). Their presence indicates that the source of an RNA may be less relevant for determining which RNAs become miRNAs, than their structural characteristics (Liu et al., 2008).

1.2.4. SiRNAs

SiRNAs are another category of small RNA involved in gene regulation. MiRNAs and small interfering RNAs (siRNAs) have many similarities both in biogenesis and in function, though there are some already defined key differences. Unlike miRNAs, which are cleaved from single-stranded hairpins, siRNAs tend to be derived from dsRNA

generated by inverted repeats, antisense transcripts, and bidirectional generation of transcripts (Watanabe et al., 2008). SiRNAs often have perfect or near perfect complementarity to their targets (Watanabe et al., 2008). In addition, different Dicer and Argonaute proteins are generally responsible for processing and utilizing miRNAs and siRNAs (Okamura et al., 2004). In *D. melanogaster*, the Dcr-2/R2D2 complex in the cytosol plays a key role in determining the downstream pathway for both miRNAs and siRNAs (Tomari et al., 2007; Forstemann et al., 2007; Liu et al., 2008). Products of both Dcr-1 (miRNA/miRNA* duplexes) and Dcr-2 (siRNA duplexes) interact with the Dcr-2/R2D2 complex (Lee et al., 2004; Tomari et al., 2007; Forstemann et al., 2007). Considerable mismatch in a duplex, uncommon in siRNA duplexes, results in poor association with R2D2 and leads to a downstream association with Ago-1 (Tomari et al., 2007; Forstemann et al., 2007).

It has been found that both miRNAs and siRNAs can function in either repression or degradation of transcripts often depending upon the degree of complementarity to the target (Pilai, 2007), though this is not always the case (Eulalio et al., 2007). In metazoans miRNAs typically function in translational repression and siRNAs typically function in the degradation of transcripts (Wu et al., 2008). In plants, miRNAs are often highly complementary to their targets like siRNAs (Brodersen et al., 2008). In animals and plants, mRNAs have been found with siRNA targets along the entire length of the transcript (Brodersen et al., 2008) and usually are highly complementary to their targets (Wu et al., 2008). Interestingly, one study found that a length of approximately 21

nucleotides, the most frequent size of siRNAs generated *in vivo*, optimizes the balance between target specificity and minimizing off-target effects (Qiu et al., 2005).

SiRNAs are involved in RNA interference, which is important for cellular suppression of the activity of "exogenous" viruses and transposable elements as well as gene regulation (Watanabe et al., 2008; Ghildiyal et al., 2008). Several labs have discovered endogenously-encoded siRNAs in plants, animals and fungi (Sontheimer et al., 2005). As it stands now miRNAs function mostly as regulators of cellular genes, while siRNAs function to control viruses, and transposons, as well as control genes. Recent studies have found endogenous siRNAs in flies and mice that silence mRNA transcription. (Watanabe et al., 2008; Ghildiyal et al., 2008). In an analysis of *Drosophila* head samples, 41% of endogenous siRNAs mapped to mRNAs, whereas in Drosophila S2 cells, about 10% mapped to transposable elements (Ghildiyal et al., 2008). In plants and fission yeast, siRNAs have been found to play a role in chromatin remodeling (Grobhans and Filipowitz, 2008; Swiezewski et al., 2007). Artificial siRNAs have recently been designed to target specific promoters in human genes, leading to their activation in a process known as RNAa (Li et al., 2008).

1.2.5. RasiRNAs and piRNAs

Repeat-associated siRNAs (rasiRNAs) and piwi-interacting RNAs (piRNAs) are 24-27 nt RNAs (in *Drosophila*) generated from long dsRNA and play a role in silencing transposable elements (TEs) (Sontheimer et al., 2005; Pal-Bhadra et al., 2004; Kim, 2005; Amaral et al., 2008). In flies, rasiRNAs and piRNAs are interchangeable terms (Zamore,

2007), though some suggest that piRNAs are a subset of rasiRNAs (Saito et al., 2006). RasiRNAs, including piRNAs, are found in mammals as well (Lau et al., 2006; Aravin et al., 2007). In *Drosophila*, though not yet in mammals, piRNAs have been found to repress transposable element activity in the germline of both males and females (Brennecke et al., 2007; Gunawardane et al., 2007; Bushati and Cohen, 2007).

One of the first systems to be examined in detail was *S. pombe*, where silencing of transposable elements and repeat sequences utilizes rasiRNAs that guide heterochromatin formation at specific sites through the RITS (RNA-induced Initiation of Transcriptional gene Silencing) pathway (Kim, 2005). TE transcript generation may result in dsRNAs by read-through transcription from a nearby promoter on the opposite strand, or hairpin formation by inverted repeats in transcripts (Vastenhouw et al., 2004). In *Drosophila* ovaries, rasiRNAs were found to be derived primarily from the antisense strands of double-stranded RNAs, and their processing and function do not require crucial enzymes for RNAi (Ago-2, Dcr-2) nor for miRNA transcript silencing (Ago-1 and Dcr-1), suggesting the involvement of a separate pathway (Vagin et al., 2006). It has also been observed that the biogenesis of piRNAs may not involve Ago-1 or -2, but instead relies upon PIWI proteins (Vagin et al., 2006; Brennecke et al., 2007; Gunawardane et al., 2007; Bushati and Cohen, 2007). Interestingly, different Ago proteins are involved in the use of piRNAs compared to anti-viral siRNAs in flies; while Ago-2 plays a role in cellular defense against viruses, Ago-3, and two other Ago proteins, Piwi, and Aubergine are the proteins involved in silencing selfish DNA in flies (Zamore, 2007). Antisense rasiRNAs are first generated from large genomic loci, and utilize the three different ago

proteins to repeat a cycle of recognition and cleavage of nascent transposable element transcripts to amplify the number of rasiRNA present in the event of TE activity (Zamore, 2007). The key enzymes involved in the fly pathway are those utilized in heterochromatic gene silencing through H3K9 methylation (Pal-Bhadra et al., 2004), and cytosine methylation (Amaral et al., 2008).

That being said, there is some evidence from *Drosophila* germline studies and from *A. thaliana* for interrelationships between the RNAi and heterochromatic gene silencing pathways because of co-localization of key enzymes of both pathways, and because mutations in some heterochromatic gene silencing enzymes can impact RNAi (Buchon et al., 2006; Pal-Bhadra et al., 2004; Findley et al., 2003; Lippman et al., 2003). In sum, the literature seems to support that rasiRNA production and function in the *Drosophila* germline do not require RNAi enzymes (Vagin et al., 2006), though there could be redundancy/ compensation between the pathways (Deshpande et al., 2005). It has been demonstrated that relieving key enzymes involved in rasiRNA processing and function leads to an elevated rate of retrotransposition in *Drosophila* germline tissues, which suggests that rasiRNA could provide useful information about potentially active transposable elements (Aravin et al., 2001; Findley et al., 2003; Kogan et al., 2003; Vagin et al., 2006).

1.3. Discovery and identification of miRNAs

While novel miRNA discovery in model organisms may be reaching its limits, significant miRNA discovery in many organisms still remains (Berezikov et al., 2006; Ruby et al., 2006; Stark et al., 2007). MiRNAs have been discovered by several routes including cloning populations of small RNAs present in a sample of interest, microarray-based detection of miRNAs, and bioinformatics predictions using the genome of an organism of interest, based on similarity known miRNAs in other organisms. Acceptable verification of miRNAs requires experimental evidence of their expression and demonstrating the presence of a valid hairpin precursor in the genome of the organism (Ambros et al., 2003; Berezikov et al., 2006).

Small RNA cloning has been the most widely used method for discovering miRNAs (Stark et al., 2007), and can provide a detailed snapshot of small RNAs, including miRNAs present in a specific sample at a specific time (Aravin et al., 2003; Lim et al., 2003; Chen et al., 2005). Cloning was used to detect 62 different miRNAs across multiple stages of *D. melanogaster* (Aravin et al., 2003), and has been used to detect 35 predicted and new miRNAs in *Anopheline* mosquitoes (Winter et al., 2007; Mead and Tu, 2008).

Early small RNA cloning methods such as Lau et al. (2001) provided the skeleton for current cloning methodologies, though relied upon radioactivity, which is often not necessary in current cloning methods. Early cloning also relied upon large amounts of total RNA (500 μg; Lau et al., 2001) whereas current cloning methods such as the Illumina small RNA cloning kit (Illumina, San Diego, CA) require only 10 μg total RNA.

Briefly, total RNA samples are isolated in a method such as Ambion's mirVana kit or by trizol, to retain small RNAs often lost in conventional total RNA extractions (Ambion, Austin TX). The RNA is then size fractionated on an acrylamide gel, and small RNA is selectively isolated based on size (Lau et al., 2001). Linkers are sequentially added to the ends of the small RNA, to allow RT-PCR amplification of all small RNA present, which can then be readily cloned (Lau et al., 2001).

Although early cloning methods provided information about abundant miRNAs, new high throughput methods such as pyrophosphate 454 sequencing and Illumina sequencing can provide "deep sequencing" by generating millions of sequences in a typical run allowing even low abundance miRNAs to be detected (Ahmadian et al., 2006). This also provides a more accurate quantitative picture of the prevalence of a given miRNA in the sample, compared to lower scale sequencing efforts, though concerns about potential PCR biases for quantative analyses remain. Pyrosequencing relies upon "sequencing by synthesis" where the sequential addition of a nucleotide to the free end of a growing strand leads to the generation of light which can be quantitated to provide the identity of a given nucleotide at each addition (Ahmadian et al., 2006). Illumina reversible terminator sequencing is an alternate method of "sequencing by synthesis" that involves adding nucleotides one at a time to a growing strand, but from a pool of the four possible nucleotides, each with a different fluorescent tag (Bentley, 2006). Addition is halted after a single nucleotide addition by adding a terminator, after which a laser can interact with the tag to generate fluorescence that is then read by a photodetector (Bentley, 2006).

Following each reading, the terminator is removed and a new nucleotide is added (Bentley, 2006).

Cloning provides experimental verification of the expression of a given miRNA, and has the added benefit of revealing novel and organism-specific miRNAs and other small RNAs, which may be overlooked in bioinformatics predictions. A drawback to this method is that it provides a limited picture of the miRNAs present in an organism. It neglects those present in unsampled stages or tissues (Brennecke and Cohen, 2003; Nam et al., 2005).

Microarrays have been used for discovering miRNAs, providing the ability to examine large numbers of miRNAs at once (Kim and Nam, 2006). However, they are limited to analyzing the presence of known or previously predicted miRNAs; they cannot identify novel miRNAs in a population. Microarrays may not be ideally suited for miRNA analyses due to the fact that the short size of a miRNA makes it difficult to have an optimal hybridization temperature for all miRNAs under examination (Kim and Nam, 2006). Precursor miRNAs may provide longer sequences for generating antisense matches, but may not account for regulation prior to the mature miRNAs (Kim and Nam, 2006). A newer approach known as the RAKE (RNA-primed array-based Klenow enzyme) assay utilizes the miRNA as a primer to incorporate labeled nucleotides, and cannot allow for extension in the event of mismatches allowing better discrimination among closely related mature miRNA sequences (Nelson et al., 2004; Kim and Nam, 2006). Though allowing higher throughput, this and other microarray-based methods may

ultimately still be unable to match the northern for dynamic quantitative range (Kim and Nam, 2006).

Bioinformatic predictions of miRNAs rely upon similarity to known miRNAs. Predicted and known mature sequences from previously studied organisms are often highly conserved, allowing one to find genomic matches in a different organism by BLAST. Predictions are then analyzed more closely for similiarities to valid miRNAs. Flanking genomic regions may be readily obtained from ENSEMBL (Flicek et al., 2008) and analyzed for their ability to fold into a hairpin stem loop structure, using programs such as RNAfold (Hofacker et al., 1994) with the putative mature miRNA in one arm of the stem. Bulges and loops in pairing to the antisense arm should tend to be symmetrical (Stark et al., 2007). MiRscan is a program which compares precursor sequences of hairpins derived from ~100 mers from the genomes of two organisms suspected to be homologs (Lim et al., 2003). The program looks for varying degrees of conservation. Valid miRNA tend to have very high conservation of the mature miRNA sequence, less conservation in the corresponding miRNA* antisense sequence, and low conservation in the loop and flanking stem regions (Lim et al., 2003). From these results, a score is given which indicates the degree to which a given miRNA has characteristics of known miRNA, based upon analyses from *C. elegans* and *C. briggsae*; the higher the score, the more the putative precursors demonstrate these characteristics (Lim et al., 2003). This program has a reasonably high rate of successfully identifying known miRNAs in other organisms (Lim et al., 2003). Roughly three-fourths of known vertebrate miRNAs in test runs were recognized (Lim et al., 2003), however, personal examination has found that in

at least some cases, stem loops from highly divergent organisms are unable to accurately be predicted to be pre-miRNAs. Also, given its heavy reliance upon conservation, species-specific miRNAs may be unable to be verified using this program. A revised version of the original program can be downloaded (Ruby et al., 2007).

There have been additional methods of detecting miRNAs, which, while less commonly used, have been successful. Collectively, these tend to suffer from low-throughput problems due to individual analysis of miRNAs, and require knowledge of the sequence beforehand for primer or probe designs (Kim and Nam, 2006). RT-PCR has been used to detect precursor miRNAs in cancer cells (Kim and Nam, 2006). This method is capable of detecting extremely low copy numbers of miRNAs and can provide accurate information on expression levels (Kim and Nam, 2006). Of course, miRNAs of high expression may also be detected by Northern, without relying upon PCR-based approaches. Flow cytometry has also been successfully used for detecting miRNAs (Kim and Nam, 2006). This method utilized fluorescent probes with unique miRNA signatures coupled to beads, which were then detected by flow cytometry to give unique profiles of miRNA expression in cancer samples (Kim and Nam, 2006).

1.4. MiRNA target analysis

Target analysis is an informative route to predicting the function of a given miRNA. MiRNA targets are readily predicted by software such as miRanda, a program which looks for putative targets of a given miRNA from databases of RNA sequences based

upon conserved signature characteristics of miRNA:target interactions. These include critical base pairing at the 5' end of the mature miRNA as well as sequences with imperfect seed matches but compensatory 3' end pairing (John et al., 2004).

Seed pairing, pairing between nucleotides two through eight from the 5' end of the miRNA (the "seed" region) to the target, has been found to be the most important region for determining predicted targets, based upon known target sites (Lai, 2002; Rajewsky and Socci, 2004; Grün and Rajewsky, 2008). MiRNAs are frequently grouped into "families" based upon their seed regions as family members often have overlapping targets (Hwang et al., 2007) To the extent that seed pairing contains mismatches, compensatory pairing in the 3' end of the miRNA must be stronger to make up for it. G:U wobbles are poorly tolerated (Grün and Rajewsky, 2008). While often thought of as of low significance, there is increasing evidence that mispairings, wobbles, and the 3'-end of the miRNA also play an important role in miRNA specificity (Slack, 2008). It is known that strong pairing at the 3' end can improve silencing (Kiriakidou et al., 2004; Brennecke et al., 2005; Liu et al., 2008). Multiple miRNAs may bind a target to increase the impact of silencing or to add redundancy, and spacing of the sites within the 3'-UTR of animals is important to function additively; too little space between sites reduces transcript repression (Doench and Sharp, 2004; Krek et al., 2005; Saestrom et al., 2007; Liu et al., 2008). A site's surroundings are also important as secondary structure such as hairpins as well as the surrounding sequence itself, may affect the accessibility of a potential target by miRNPs (Ameres et al., 2007; Long et al., 2007; Liu et al., 2008).

MiRNAs often have hundreds of targets based upon bioinformatic predictions and a given mRNA may be targeted by multiple miRNAs, allowing for redundancy and precise regulation (Lewis et al., 2005). Bioinformatic predictions of miRNA targets suggest that miRNAs may regulate a third of human transcripts (Lewis et al., 2005); However a proteomic analysis of targets in the *Drosophila* oocyte suggested only 4% of genes in the oocyte were regulated by miRNAs (Nakahara et al., 2005). Few targets have been empirically examined, compared to the number of miRNAs and their numbers of predicted targets (Grün and Rajewsky, 2008). Thus, most miRNA target prediction algorithms rely upon a very small list of known miRNA:target interactions to train predictive software which can lead to learning too many parameters from too few examples (Grün and Rajewsky, 2008). As many predictive programs concentrate upon binding at the seed region, many false positives are predicted to occur (Grün and Rajewsky, 2008). If a program weighted targets solely upon a seven nucleotide seed region, a false positive would occur every 16,000 bases of genomic sequence (Grün and Rajewsky, 2008). However, as prediction methods become increasingly savvy, incorporating new findings, the false positive rate should decline dramatically. Further, the user can reduce the false positive rate by limiting searches to the 3'-UTR regions.

Conservation of targets is another way to reduce the false positive rate, as evolutionary constraints should retain pairing specificity to conserved miRNAs (Grün and Rajewsky, 2008). This has been independently confirmed by a number of labs (Grün et al., 2005; Krek et al., 2005; Lewis et al., 2005; Xie et al., 2005; Lall et al., 2006; Grün and Rajewsky, 2008). Mature miRNAs are highly conserved across divergent organisms. A

recent publication found that more than half of worm miRNAs have sequence homology to *H. sapiens* and *D. melanogaster* miRNAs (Ibáñez-Ventoso et al., 2008). A drawback to focusing upon conservation is that it is limited to analyses using conserved miRNAs, and misses species-specific miRNAs (Grün and Rajewsky, 2008).

While A-to-I RNA editing is common in the human transcriptome, and at least one instance of A-to-I editing of a miRNA has been observed (Luciano et al., 2004), computational predictions suggest that RNA editing of 3'-UTRs, while common in primate transcripts, is uncommon specifically in miRNA targets within the transcripts, suggesting that RNA editing may have only a minor impact on predicting targets in humans (Liang and Landweber, 2007).

1.5. MiRNA Functions

The miRNA pathway typically silences expression of endogenous genes in eukaryotes (Vasudevan et al., 2007). MiRNAs function in stem cell differentiation, longevity, apoptosis, development, in the regulation of transcription factors, even in the regulation of miRNA expression (Du and Zamore, 2005; Kim, 2005). MiRNA have also been found to be important in maintaining cellular identity after differentiation (Kloosterman, 2006). MiRNAs are intimately involved in reproduction, with a unique miRNA profile signature for murine ovaries and testes (Mishima et al., 2008). Fifty-five different miRNAs showed high or unique expression changes comparing testes and ovaries (Mishima et al., 2008).

Many miRNAs are suspected to function in the "fine-tuning" of expression (Lai et al., 2005) though some miRNAs such as let-7 are known to operate as major developmental switches (Bushati and Cohen, 2007). Some miRNAs have been proposed to act at multiple levels, including at the global level of gene expression in a given cell to substantially expand the expression diversity possible with a limited gene repertoire (Makeyev and Maniatis, 2008).

Though miRNAs usually supress translation, it was observed in human cells undergoing cell cycle arrest that some miRNAs can actually upregulate mRNA expression (Vasudevan et al., 2007). During the cell cycle, it is thought that certain miRNAs act as repressors during proliferation, and act to upregulate expression during cell starvation (Vasudevan et al., 2007; Wu et al., 2008). However, direct application to humans should be considered with caution, as both cell lines examined (HEK293 and HeLa) are transformed lines. In another study, miR-373 was found to activate gene expression by targeting E-cadherin and CSDC2 promoters in human PC-3 cell culture (Place et al., 2008).

Recent studies in miRNA function suggest that some miRNAs may not act in the cytosol, nor even in the cell they were generated in (Valadi et al., 2007; Hwang et al., 2007). Six different miRNAs were identified in exosomes from human and mouse bone cells which had been transferred to other cells and functioned in the new cells, suggesting miRNAs might be involved in cell-to-cell communication using miRNAs (Valadi et al., 2007). In

another study, a 6nt sequence was discovered in the 3'-end of a human miRNA, miR-29b, which directs localization of the miRNA to the nucleus (Hwang et al., 2007).

1.5.1. MiRNA functional analysis

Understanding the functions of miRNAs has led to exciting new discoveries in gene regulation. Yet most known miRNAs do not yet have experimentally-validated functions in most organisms. Functional analysis of miRNAs is being undertaken by three major routes: forward genetic studies, reverse genetics, and target analysis (Weinholds and Plasterk, 2005).

Forward genetics has traditionally been a cornerstone of functional analyses of genes. Forward genetics studies often focus upon studies of mutations to determine genetic underpinnings of the observed phenotype. Typically, loss-of-function or gain-of-function studies are used in miRNA forward genetic analyses (Weinholds and Plasterk, 2005). Gain-of-function analysis led to the discoveries of some of the first miRNA genes in the 1990's, prior to their recognition as miRNA genes a decade later (Ruby et al., 2007). Hormonal regulation of miRNAs is another route to determining their function through forward genetics. The researcher hopes to establish a link between a key hormone in an organism, and the expression patterns of miRNAs. In a study on hormonal regulation in Drosophila S2 cells, Sempere found that the Juvenile Hormone analog methoprene could regulate miRNA expression of several miRNAs, in opposition to the effects observed by ecdysone (Sempere et al., 2003).

Reverse genetic studies of miRNAs focus upon targeting a known individual miRNA or group of miRNAs, to determine its function. Gene knockout and rescue is a powerful way to determine function. Though deletion mutation libraries of *D. melanogaster* exist (Oh et al., 2003), studies in other insect species must rely upon alternate technology. Currently, there is limited site-specific genetic engineering technology available for studies in mosquitoes, though new developments such as success with the phi C31 in *Ae. aegypti* are encouraging (Ninmo et al., 2006). Further, many insect cell lines are poorly characterized and not as hardy as mammalian cell lines. Gene knockdown through RNAi has been successfully used to study the functions of individual miRNA and miRNA in aggregate (Weinholds and Plasterk, 2005). Similarly, knockdown may be accomplished through the use of antisense oligos, which deplete a cell of a complementary miRNA, Knockdown innovations include using chemically-modified oligos such as morpholinos, antagomirs, or LNAs, which provide unusually high affinity for complementary sequences, critical for studies of short RNA sequences.

In some studies in plants and mammals, mass knockdown by compromising the miRNA pathway, such as deleting a dicer gene, has been used to study the global impact of miRNAs in development. Such studies have demonstrated a profound role for miRNAs collectively, but do little in establishing the function of individual miRNAs. Other methods for examining function through reverse genetics include overexpression of the miRNA in question, and developing expression profiles of miRNAs (Weinholds and Plasterk, 2005).

Target analysis involves a bioinformatic route to examine miRNA function from predicting what messenger RNA may be impacted by an miRNA. Target studies have been carried out for many miRNAs, and the miRNA repository miRBase catalogues bioinformatics predictions of miRNA targets using miRanda (Griffiths-Jones et al., 2006). Knowledge of related miRNAs may also give clues to the function of a specific miRNA, as miRNAs in the same family, those sharing the same seed region, often have similar or overlapping targets (Hwang et al., 2007). Predicted targets are often experimentally examined through reporter assays utilizing recombinant plasmids with targets located in the 3'-UTR to test for their interaction with endogenous miRNAs, (Ambion). These methods of target analysis provide useful directions to function-related research, but do not provide an empirical demonstration of miRNA function. A method published in 2008 by Andachi utilizes the miRNA itself as a primer in its miRNP-associated interaction with mRNA targets, to provide experimental evidence for miRNA targets (Andachi, 2008)

1.6. MiRNAs in mosquitoes

Methods for predicting miRNAs in *An. gambiae* have relied upon on homology to known miRNAs in distant species such as *D. melanogaster* and *C. elegans,* and predictions of novel miRNAs based upon characteristics of known miRNAs. (Lai et al., 2003; Wang et al., 2005; Chatterjee and Chaudhuri, 2006; Ruby et al., 2007; Xu et al., 2008). Considerable effort has gone into characterization of the small RNAs found in the

primary model organism for insects, *Drosophila melanogaster* (Aravin et al., 2003), and 160 miRNAs are known in *D. melanogaster* (Ruby et al., 2007), 152 of these formally listed in miRBase (Griffiths-Jones, 2004; Griffiths-Jones et al., 2006).

MiRBase (Griffiths-Jones et al., 2006) is a formal repository for miRNAs, listing details of 9539 entries from many well-known species. As of May 2009; miRBase 13.0 lists 67 mature miRNAs for *An. gambiae*, and 11 miRNAs have been predicted for *Ae. aegypti* (Flicek et al., 2008). Of these, there is only experimental evidence of expression for 35 different miRNAs in *Anopheles* adults and of these, only five were found to be mosquito-specific: miR-1174, -1175, -1889, -1890, -1891 (Winter et al., 2007; Mead and Tu, 2008). In the literature, recent estimates of expected miRNAs in *An. gambiae* run well over 100, similar to *D. melanogaster* (Ruby et al., 2007; Xu et al., 2008).

While most miRNAs in mosquitoes do not yet have published profiles, many of those that have published profiles, such as miR-9a, and the ubiquitous let-7, typically have matched the expression profiles of homologs observed in *D. melanogaster* (Mead and Tu, 2008).

Several additional miRNAs that were found to be expressed in mosquitoes may prove especially interesting to the vector-borne disease community. MiR-14 plays an important role in longevity and apoptosis in the fly, and if a similar link could be established in mosquitoes, it would be of paramount importance, as longevity is crucial to the vectorial capacity of mosquitoes in the transmission of dengue and malaria (Xu et al., 2003; Mead

and Tu, 2008; Spielman, 1994). MiR-989 was observed to change in expression in response to bloodfeeding and is expressed nearly exclusively in the female ovary, implying a role in reproduction (Mead and Tu, 2008). In addition, miR-989, as well as miR-34, miR-1174, and miR-1175 all exhibited changes in expression in response to *Plasmodium* infection (Winter et al., 2007). Host-pathogen interactions may provide a key target for interrupting vector-borne diseases (Winter et al., 2007).

1.7. Research significance and objectives

Small RNA cloning offers a means to verify predicted miRNAs, and discover novel miRNAs, opening the door to functional studies. MiRNAs have been found to play a role in many important biological functions through post-transcriptional regulation, yet to date there has been little work in the study of miRNAs in mosquitoes despite the fact that mosquitoes are vectors for diseases that kill over a million people annually (WHO). This project provides a foundation for understanding gene regulation by miRNAs in mosquitoes. Given the importance of miRNAs in eukaryotic biology, a better understanding of miRNAs in mosquitoes will undoubtedly inform strategies to control mosquito-borne infectious diseases. MiRNAs may also be directly involved in mosquito-pathogen interaction as discussed below.

In the rapidly expanding field of small RNA-based gene regulation strategies, it seems likely that the mosquito would be the best target for small RNA strategies to control malaria. Studies of human infection with *P. falciparum* have failed to show any

correlation of human miRNA expression changes with plasmodium infection in a small RNA cloning study; a similar study with murine cells infected with *P. berghei* yielded the same results (Xue et al., 2008). MiRNAs have not been detected in *P. falciparum*, and the parasite lacks genes for Argonaute and Dicer, suggesting that knockdown strategies against *P. falciparum* transcripts should not work (Xue et al., 2008).

I used small RNA cloning to identify moderate to highly-expressed miRNAs in a population of *Anopheles stephensi* females (Mead and Tu, 2008). Expression analysis in greater detail was conducted upon eight selected miRNAs, providing a foundation for analysis of their roles in mosquitoes (Mead and Tu, 2008). Cloning in *Anopheles stephensi* and *Aedes aegypti* revealed miRNAs that could not be found outside of Culicidae. Representatives were selected for further expression analysis to derive hints at conservation and function. MiR-989 was a miRNA uncovered in cloning which was expressed primarily in the adult female ovary and its expression changed in response to bloodfeeding, hinting that miR-989 may play a role in reproduction and that it would be a promising candidate for further analyses (Mead and Tu, 2008).

Specific Aims of this research were:

Specific Aim 1. Clone and characterize miRNAs from the female Asian malaria mosquito, *Anopheles stephensi* and determine the expression profiles of selected miRNAs. Small RNA cloning from an aged population of female *An. stephensi* and

subsequent expression analysis provided one of the first looks at an area of gene regulation in *Anopheline* mosquitoes, detailing moderate to highly expressed miRNAs.

Specific Aim 2. Determine expression profiles across developmental stages for known and novel mosquito-specific miRNAs in *An. gambiae*, *An. stephensi, Ae. aegypti* and *T. amboinensis*. This was part of a larger effort that uncovered 98 miRNA precursors from mosquitoes, including 13 mosquito-specific miRNAs. My work focused on the expression analysis of eight mosquito-specific miRNAs and revealed conserved and sometimes stage-specific expression profiles of some of the mosquito-specific miRNAs. I also provided evidence for three narrowly conserved miRNAs that may shed light on the divergence of different mosquito lineages.

Specific Aim 3. Examine the evolution, expression, regulation, and function of miR-989. Detailed bioinformatic analyses were conducted of miR-989 in insects to examine the evolutionary relationships of miR-989 among and within insect orders. The expression profile of miR-989 in adult female tissues in *An. stephensi* and *Ae. aegypti* was determined. The expression profile of miR-989 was also analyzed during a time course across the adult female lifespan with and without bloodfeeding. MiR-989 targets were predicted using the miRanda method and experimentally examined by a PCR-based method. Modified abdominal ligation assays in *Ae. aegypti* did not support JH regulation of miR-989. However, high throughput small RNA sequencing of JH treated and control samples uncovered strong candidate miRNAs that are under JH regulation.

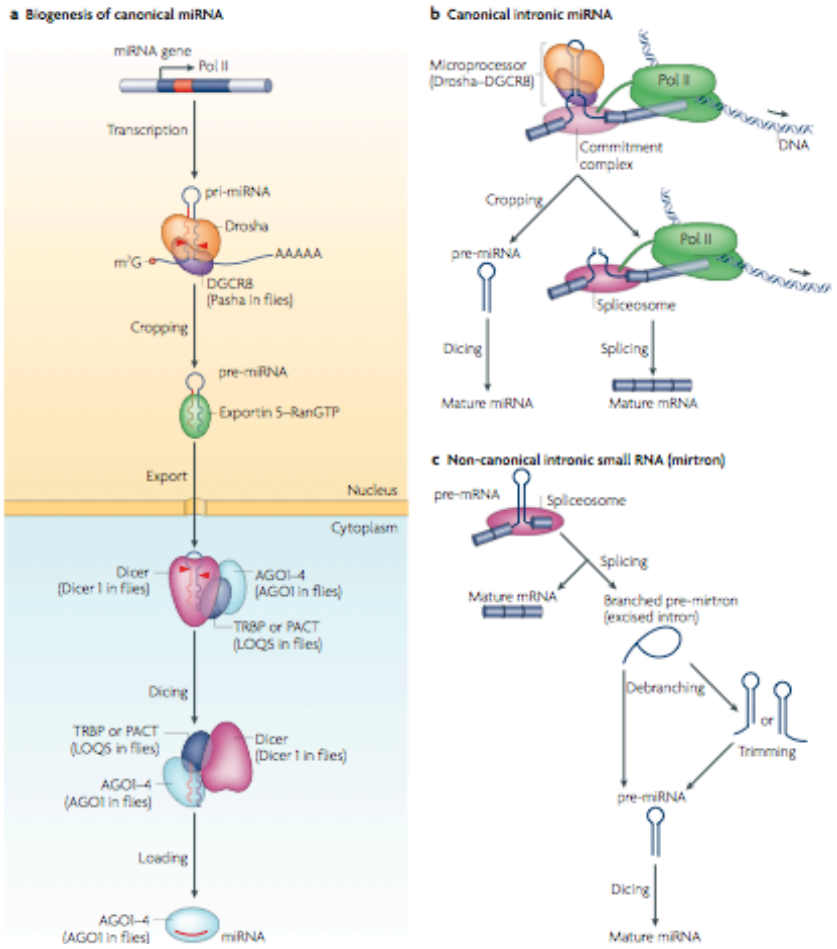


Figure 1.1. MiRNA biogenesis pathway. From Kim et al., 2009[1]. Accessible online at: http://www.nature.com/nrm/journal/v10/n2/abs/nrm2632.html

Figure 2 | **miRNA biogenesis pathway. a** | Canonical microRNA (miRNA) genes are transcribed by RNA polymerase II (Pol II) to generate the primary transcripts (pri-miRNAs). The initiation step (cropping) is mediated by the Drosha–DiGeorge syndrome critical region gene 8 (DGCR8; Pasha in *Drosophila melanogaster* and *Caenorhabditis elegans*) complex (also known as the Microprocessor complex) that generates ~65 nucleotide (nt) pre-miRNAs. Pre-miRNA has a short stem plus a ~2-nt 3′ overhang, which is recognized by the nuclear export factor exportin 5 (EXP5). On export from the nucleus, the cytoplasmic RNase III Dicer catalyses the second processing (dicing) step to produce miRNA duplexes. Dicer, TRBP (TAR RNA-binding protein; also known as TARBP2) or PACT (also known as PRKRA), and Argonaute (AGO)1–4 (also known as EIF2C1–4) mediate the processing of pre-miRNA and the assembly of the RISC (RNA-induced silencing complex) in humans. One strand of the duplex remains on the Ago protein as the mature miRNA, whereas the other strand is degraded. Ago is thought to be associated with Dicer in the dicing step as well as in the RISC assembly step. In *D. melanogaster*, Dicer 1, Loquacious (LOQS; also known as R3D1) and AGO1 are responsible for the same process. In flies, most miRNAs are loaded onto AGO1, whereas miRNAs from highly base-paired precursors are sorted into AGO2. The figure shows the mammalian processing pathways with fly components in brackets. **b** | Canonical intronic miRNAs are processed co-transcriptionally before splicing. The miRNA-containing introns are spliced more slowly than the adjacent introns for unknown reasons. The splicing commitment complex is thought to tether the introns while Drosha cleaves the miRNA hairpin. The pre-miRNA enters the miRNA pathway, whereas the rest of the transcript undergoes pre-mRNA splicing and produces mature mRNA for protein synthesis. **c** | Non-canonical intronic small RNAs are produced from spliced introns and debranching. Because such small RNAs (called mirtrons) can derive from small introns that resemble pre-miRNAs, they bypass the Drosha-processing step. Some introns have tails at either the 5′ end or 3′ end, so they need to be trimmed before pre-miRNA export. m7G, 7-methylguanosine.

1. Reprinted by permission from VN Kim and Macmillan Publishers Ltd: Nature Reviews Molecular Cell Biology from Kim, V.N., Han, J., and Siomi, M.C. (2009) Biogenesis of small RNAs in animals. *Nat Rev Mol Cell Biol* **10**(2):126-139. Copyright 2009.

Chapter 2.

Cloning, Characterization, and Expression of microRNAs from the Asian Malaria Mosquito, *Anopheles stephensi*

Published in *BMC Genomics* **9**: 244.

Edward Andrew Mead and Zhijian Tu

2.1. Abstract

MicroRNAs (miRNAs) are non-coding RNAs that are now recognized as a major class of gene-regulating molecules widely distributed in metozoans and plants. MiRNAs have been found to play important roles in apoptosis, cancer, development, differentiation, inflammation, longevity, and viral infection. There are a few reports describing miRNAs in the African malaria mosquito, *Anopheles gambiae*, on the basis of similarity to known miRNAs from other species. *An. stephensi* is the most important malaria vector in Asia and it is becoming a model *Anopheline* species for physiological and genetics studies.

We report the cloning and characterization of 27 distinct miRNAs from 17-day old *An. stephensi* female mosquitoes. Seventeen of the 27 miRNAs matched previously predicted *An. gambiae* miRNAs, offering the first experimental verification of miRNAs from mosquito species. Ten of the 27 are miRNAs previously unknown to mosquitoes, four of which did not match any known miRNAs in any organism. Twenty-five of the 27 *Anopheles* miRNAs had conserved sequences in the genome of a divergent relative, the yellow fever mosquito *Aedes aegypti*. Two clusters of miRNAs were found within introns of orthologous genes in *An. gambiae, Ae. aegypti,* and *Drosophila melanogaster*. Mature miRNAs were detected in *An. stephensi* for all of the nine selected miRNAs, including the four novel miRNAs (miR-x1- miR-x4), either by northern blot or by Ribonuclease Protection Assay. Expression profile analysis of eight of these miRNAs revealed distinct expression patterns from early embryo to adult stages in *An. stephensi*. In both *An. stephensi* and *Ae. aegypti*, the expression of miR-x2 was restricted to adult females and predominantly in the ovaries. A significant reduction of miR-x2 level was observed 72 hrs after a blood meal. Thus miR-x2 is likely involved in female reproduction and its function may be conserved among divergent mosquitoes. A mosquito homolog of miR-14, a regulator of longevity and apoptosis in *D. melanogaster*, represented 25% of all sequenced miRNA clones from 17-day old *An. stephensi* female mosquitoes. *An. stephensi* miR-14 displayed a relatively strong signal from late embryonic to adult stages. MiR-14 expression is consistent during the adult lifespan regardless of age, sex, and blood feeding status. Thus miR-14 is likely important across all mosquito life stages.

This study provides experimental evidence for 23 conserved and four new microRNAs in *An. stephensi* mosquitoes. Comparisons between miRNA gene clusters in *Anopheles* and *Aedes* mosquitoes, and in *D. melanogaster* suggest the loss or significant change of two miRNA genes in *Ae. aegypti*. Expression profile analysis of eight miRNAs, including the four new miRNAs, revealed distinct patterns from early embryo to adult stages in *An. stephensi*. Further analysis showed that miR-x2 is likely involved in female reproduction and its function may be conserved among divergent mosquitoes. Consistent expression of miR-14 suggests that it is likely important across all mosquito life stages from embryos to aged adults. Understanding the functions of mosquito miRNAs will undoubtedly contribute to a better understanding of mosquito biology including longevity, reproduction, and mosquito-pathogen interactions, which are important to disease transmission.

2.2. Introduction

MicroRNAs (miRNAs) are non-coding RNAs that are now recognized as a major class of gene-regulating molecules widely distributed in metozoans and plants (Miska, 2005; Du and Zamore, 2007). Many miRNA genes are transcribed by RNA polymerase II, yielding primary miRNAs (pri-miRNAs) of hundreds to thousands of bases in length (Miska, 2005; Bartel, 2004; Pasquinelli et al., 2005; Zhou et al., 2007). A given pri-miRNA can be either monocistronic, containing a sequence for one mature miRNA, or polycistronic, containing a sequence for multiple mature miRNA products (Miska, 2005; Bartel, 2004; Lau et al., 2001). In *Drosophila*, the pri-miRNA is processed by a Drosha-Pasha complex to yield pre-miRNA, small stem-loop

structures that are approximately 70 nucleotides (nt) in length (Denli et al., 2004; Bushati and Cohen, 2007). These stem-loops, or hairpins, are then exported to the cytoplasm and processed by Dicer-1 to form an miRNA:miRNA* duplex (Miska, 2005; Bartel, 2004; Lau et al., 2001). The duplex molecules are separated by a helicase, and based upon the strength of 5' end pairing, one single strand is chosen as the mature miRNA (Bartel, 2004). The opposing strand is referred to as the miRNA* strand, and believed to rapidly degrade following separation (Bartel, 2004). Mature miRNAs associate with an Argonaute protein and bind their mRNA targets, which are often in the 3' untranslated region (UTR), resulting in inhibition of translation or possibly target mRNA degradation in animals (Du and Zamore, 2007). The "seed region" (bases 2-8 from the 5' end) contributes significantly to miRNA-target recognition (Brennecke et. al., 2005; Grimson et al., 2007).

MiRNAs have been found to play important roles in apoptosis, cancer, development, differentiation, inflammation, longevity, and viral infection (Bushati and Cohen, 2007; Lai, 2005; Du and Zamore, 2005; Kim, 2005; Ambros and Chen, 2007; O'Connell et al., 2007; le Sage et al., 2007; Cullen, 2006). Estimates of the extent of miRNA gene regulation vary from 4% of transcripts in the *Drosophila* ovary (Nakahara et al., 2005) to a third of human genes (Lewis et al., 2005). It is estimated that approximately 110 different miRNAs are expressed across the different life stages of *D. melanogaster* (Lai et al., 2003). In flies, the adult stage is characterized by significant miRNA expression (Aravin et al., 2003). In a study to uncover *Drosophila* miRNAs, Lai *et al.* (2003) reported 38 putative miRNAs in the African malaria mosquito, *Anopheles gambiae*, that are conserved with *Drosophila* miRNAs (Lai et al., 2003). There are two additional reports describing *An. gambiae* miRNAs on the basis of similarity to

known miRNAs (Wang et al., 2005; Chatterjee and Chaudhuri, 2006). However, there is no direct experimental evidence for any of these miRNAs in mosquitoes. We are interested in identifying conserved as well as mosquito-specific microRNAs and exploring their potential functions in mosquito biology and mosquito-pathogen interactions. We carried out our cloning work on 17-day old adult female mosquitoes, which are highly relevant to disease transmission because it takes approximately two weeks for *Plasmodium* parasites to mature and become infective within a female mosquito (http://www.cdc.gov/malaria/biology). We used *An. stephensi* because this species is an important malaria vector in Asia and it is becoming a model *Anopheline* species for physiological and genetics studies. Here we report direct cloning and characterization of 23 conserved and four new miRNAs from the *An. stephensi* adult female. Comparative analysis uncovered the loss or significant change of two miRNA genes in a divergent mosquito *Ae. aegypti*. We also determined the expression profile of several selected miRNAs including the four new miRNAs across all life stages of *An. stephensi*. We performed further expression analysis on two miRNAs that are implicated in mosquito reproduction and longevity.

2.3. Materials and methods

Mosquitoes

An. stephensi (Indian wild-type strain) and *Ae. aegypti* (Liverpool strain) mosquitoes were maintained in humidified incubators at 27°C on a 12 hour light:dark cycle.

Small RNA cloning and sequencing

For small RNA cloning, approximately 1000 female *An. stephensi* adults were collected at 17 days post-emergence. Mosquitoes were fed blood and allowed to oviposit prior to collection. Aged females are the most relevant mosquito life stage for malaria transmission as it takes approximately two weeks for *Plasmodium* parasites to mature within the mosquito to an infective stage (http://www.cdc.gov/malaria/biology). RNA was isolated using a mirVana small RNA Isolation Kit (Ambion, Austin, TX). Subsequent steps were performed following published protocols (Lau et al., 2001) from the David Bartel laboratory (Whitehead Institute, MIT, Cambridge, MA). RNA oligo markers (18-mer, 24-mer; sequence as detailed from (Lau et al., 2001)) were purchased from Dharmacon (Lafayette, CO) and 2' ends were deprotected prior to use as indicated by Dharmacon. These oligos were end-labeled with ^{32}P (^{32}P-γATP, Perkin-Elmer, Waltham, MA) by T4 polynucleotide kinase (Promega, Madison, WI) and added to the sample to isolate RNAs between 18 and 24 nt in length. Next, 5' and 3' linkers were sequentially ligated to the isolated small RNA as well as the RNA markers. cDNA were generated by RT-PCR using primers derived from the two linker sequences. cDNA were ligated into a 2.1 TOPO TA vector (Invitrogen, Carlsbad, CA). Ligated plasmids were transformed in One-Shot Mach 1-T1^{R} Competent cells (Invitrogen). We did not concatemerize cDNA prior to cloning. Sequencing was performed by Virginia Bioinformatics Institute core facilities and the University of Washington High Throughput Genomics Unit.

Small RNA Sequence Analyses

Sequences obtained from cloning were analyzed with ClustalW (Thompson et al., 1994) to identify inserts and orientation. After removing low quality sequences determined by inspection of chromatographs, insert sequences were analyzed using BLAST (Altschul et al.,1997) against different databases as described below to identify matches to known miRNAs, mRNAs, tRNAs, rRNA, and other small RNAs (Table 2.1). Sequences that were identical to previously predicted *An. gambiae* miRNAs were identified by comparing with the miRBase miRNA registry (http://microrna.sanger.ac.uk) and *An. gambiae* miRNA predictions listed in (Chatterjee and Chaudhuri, 2006) using BLAST (Altschul et al.,1997). Comparisons to miRBase also revealed sequences that match miRNAs from other organisms but were not reported in mosquitoes. The remaining *An. stephensi* small RNA sequences that matched the *An. gambiae* genome assembly were further analyzed to uncover new miRNAs from mosquitoes (Table 2.2). *An. gambiae* genome assembly was used to retrieve miRNA precursors for secondary structural analysis because there was little sequence information available for the *An. stephensi* genome and the two species are in the same subgenus *Celia*. The matching *An. gambiae* sequences plus 100 nt flanking sequences were obtained through Ensembl (http://www.ensembl.org) and lowest energy conformations were generated by Vienna RNAfold (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi). The above mentioned *An. gambiae* precursor sequences were used to search for conserved sequences in *Ae. aegypti*, and in some cases, in *D. melanogaster*. Conserved sequence pairs were then examined via miRscan (http://genes.mit.edu/mirscan). One *An. stephensi* miRNA (ast-miR-304) was unscorable by miRscan (Table 2.2) and the precursor of this miRNA was examined using the biogenesis criteria described in the literature: 1) the miRNA appears in the stem of a hairpin structure; 2) there is considerable similarity between the miRNA

and miRNA*; and 3) few bulges if any are present within the miRNA:miRNA* pairing in the stem (Ambros et al., 2003; Gu et al., 2007).

Northern blot

Northern hybridizations shown in Figure 2.1 were conducted using miRCURY-LNA probes from Exiqon (Vedbaek, Denmark) for eight different miRNAs. Total RNA was isolated from *An. stephensi* of the same age or developmental stage and used for each blot. Ten micrograms of total RNA was loaded for each sample in a 15% denaturing polyacrylamide gel. After initial size confirmation using let-7 RNA, 19 nt and 23 nt single-stranded DNA oligos were used as size markers in subsequent experiments. Gels were stained with ethidium bromide for verification of equal loading and RNA quality before transferring to a BrightStar-Plus membrane (Ambion). Subsequent steps were based upon Wienholds et al. (2005a). Following UV crosslinking, membranes were prehybridized in a rotating oven for 30 minutes at 42°C using ULTRAhyb-Oligo Hybridization Buffer (Ambion), followed by overnight hybridization in the same buffer at 42°C with a final concentration of 0.1 nM digoxigenin (DIG)-labeled antisense miRCURY LNA probe (Exiqon). All probes used in northern blots were designed to hybridize with *An. stephensi* miRNAs. The membranes were washed twice for 35 minutes in an Ambion-recommended wash buffer (2xSSC, 0.5% SDS) at 42°C, and then once in low stringency buffer for 5 minutes at room temperature. The membranes were incubated for 30 minutes in blocking buffer followed by a 1 hr incubation with Anti-DIG-alkaline phosphatase fAb (Roche, Basel, Switzerland) in blocking buffer. The membranes were washed three times for 15 minutes in low stringency wash buffer followed by twice with alkaline phosphatase buffer. The membranes were then immersed in

CDP-Star solution (Roche) for 5 minutes, and placed inside saran wrap for exposure to X-ray film for 30 minutes.

For the northern blots shown in Figure 2.5, a similar procedure was followed. The samples were 3, 5, 10, 17, and 24 day old *An. stephensi* adult females as well as 3, 5, 10, and 17 day old *An. stephensi* adult males that were maintained on sugar water. The same cohort of adult females that were fed on blood on day 5 and allowed to oviposit two days later were collected at day 10, 17, and 24 after emergence. RNA isolation and northerns were carried out as detailed above. For tissue sample northern blots (Figure 2.6), five day old adult *An. stephensi* females were split into two groups, one maintained on sugar water, the other provided a blood meal then maintained on sugar water. Approximately twenty-four hrs post bloodfeeding, heads, ovaries, midguts and remainders (everything left behind) were collected from 60 bloodfed and 60 non-bloodfed six day old mosquitoes. This procedure was repeated again at 72 hrs after blood feeding with eight day old mosquitoes. The mosquitoes were not permitted to oviposit. All tissues were stored in RNA*later* (Ambion) during collection, then vortexed and stored in –80°C until RNA isolation as above using a *mir*Vana miRNA Isolation kit (Ambion). Northern hybridization was performed as described above except that 5 ug of total RNA were used. The same procedure was followed for *Ae. aegypti* mosquito tissue analysis. In separate experiments, 10 ug of total RNA isolated from whole body were used to compare the expression of miR-x2 in males and non-bloodfed females of both species. We have performed replicates for all expression profile analyses and obtained consistent expression patterns between replicates.

Ribonuclease Protection Assay (RPA)

RPA was used to examine ast-miR-76. Double-stranded oligos with a T7 promoter 5' of an antisense sequence of the miRNA were produced by annealing two single-stranded oligos. The annealed oligos were used to synthesize an RNA probe that was 9 nt longer than the miRNA, as seen below:

ast-miR-76
5'TTCGTTGTCGACGAAACCTGTTTTCTCCCTATAGTGAGTCGTATTA 3'
3'AAGCAACAGCTGCTTTGGACAAAAGAGGGATATCACTCAGCATAAT 5'

Our design also considered the possibility for future multiplexing by adding extra 4 adenosines immediately following +1 which were indigestible according to the instruction manual of the "mirVana miRNA Probe Construction Kit" (Ambion). This permitted the undigested probe to run at ~29 nt, and digested/protected probe to run at ~24 nt. We utilized a "MEGAscript RNAi kit" (Ambion) to synthesize radiolabeled (^{32}P-UTP, Perkin-Elmer) antisense transcripts. Templates were synthesized for 4 hours, and purified from acrylamide gels by elution and isopropanol precipitation (1x volume) with 1/10 volume 3M NaOAc and glycogen. 50,000 cpm of probe was added to 5ug total RNA and denatured at 96°C for 3 minutes followed by 42°C hybridization overnight. Next, the hybridized samples were digested with RNase A/T1 for 45 minutes at 37°C to remove single-stranded RNA, and to trim unhybridized regions of the probes. Afterwards, the samples were ethanol precipitated and resuspended in 5ul 2x loading buffer. Resuspended samples were denatured at 96°C for 5 minutes followed by a quick chill on ice before loading. RNA were separated using a 15% denaturing polyacrylamide gel at 150V. Locations of bands were examined by X-ray film exposure, using an intensifying screen. Size markers were as described for northern blot analysis.

2.4. Results

One hundred and forty-eight *An. stephensi* small RNA sequences showed 100% match to the *An. gambiae* genome assembly and were identified as probable miRNA sequences (Table 2.1). These small RNAs were represented by 27 distinct sequences (Table 2.2). Thirteen additional small RNA sequences had 1 mismatch to the *An. gambiae* genome assembly (see Table 2.3). To be conservative, we only considered these 13 sequences as possible miRNA candidates (Table 2.1 and Table 2.3) and we did not include them in our list of *Anopheles* miRNAs.

Confirmation of previously predicted Anopheline miRNAs by direct cloning and northern blot.

Seventeen of the 27 *An. stephensi* sequences shown in Table 2.2 match predicted *An. gambiae* miRNA hairpins (Table 2.2, category I). Fifteen of the 17 matches coincide with the predicted mature miRNAs described either at miRBase (http://microrna.sanger.ac.uk) (Table 2.2, category Ia) or in Chatterjee and Chaudhuri (2006) (Table 2, category Ib). Two sequences appear to be miRNAs*, the passenger strand of the miRNA:miRNA* duplex. The copy numbers of ast-miR-8*, and ast-mir-277* are less than those of ast-miR-8, and ast-miR-277, respectively. Intriguingly, miR-14, which is involved in the regulation of apoptosis and longevity in *D. melanogaster* (Xu et al., 2003), represents 25% of all the identified miRNAs. Northern analysis using total RNA from 17-day old females with antisense Locked Nucleic Acid (LNA) probes against 4 selected miRNAs (miR-9a, -14, -210, and let-7) all showed bands of the correct size, confirming cloning results (Figure 2.1, the last lane). The LNA oligos contain a mixture of DNA nucleotides and LNA nucleotides with 2'-4' methylene linkage providing high binding affinity and enhanced specificity to targets as compared to ordinary DNA oligos (Valoczi et al., 2004).

Six new miRNAs in Anopheles that are similar to known miRNAs from other organisms.

Three *An. stephensi* small RNAs display perfect or nearly perfect (only 1 mismatch) match to published miRNAs from organisms other than mosquitoes (Table 2.2, category IIa) and three more show high similarity (>84% identity over 19 or more nt) to miRNAs from other organisms (Table 2.2, category IIb). These small RNAs are named according to their closest miRBase matches, which are listed in Table 2.3. All precursor sequences for each of the six miRNAs obtained from the *An. gambiae* genome assembly formed good hairpins (see Figure 2.2B for an example). Four of the six *Anopheline* miRNAs (miR-12, miR-375, miR-2a, and miR-76) have conserved sequences in *Ae. aegypti* and comparisons between the *Anopheles* and *Aedes* hairpins produced miRscan scores ranging from 8.46 to 15.99, supporting their miRNA status (see Table 2.2 for the range of expected miRscan scores). MiRscan looks for hallmarks of miRNAs within a pair of conserved precursor stem-loop sequences by calculating a score based on seven criteria, the most important of which is the conservation of the base pairing between a miRNA and its antisense (http://genes.mit.edu/mirscan) (Lim et al., 2003a; Lim et al., 2003b). Two of the miRNAs, ast-miR-304 and ast-miR-306, do not have conserved sequences (either mature miRNA or hairpin) in *Ae. aegypti* and are described in the next section. We selected ast-miR-76, a miRNA that showed the lowest similarity to known miRNAs in category IIa and IIb, for further verification using Ribonuclease Protection Assay (Figure 2.2), which is theoretically more sensitive than northern blot (Figure 2.1). A product of expected size was detected thus supporting the expression of ast-miR-76.

Two of the above miRNAs are worth noting. ast-miR-2a matches dme-miR-2a perfectly thus is named as miR-2a. There is an aga-miR-2 in miRBase (derived from two precursors aga-miR-2-1 and aga-miR-2-2), which is reverse complementary to ast-miR-2a. Sequence comparison showed that ast-miR-2a is not derived from aga-miR-2* because of the existence of multiple indels/mismatches between the alignment of aga-miR-2 and aga-miR-2*. The orientation of our cloned ast-mir-2a is consistent with the miRscan prediction based on *An. gambiae* and *Ae. aegypti* hairpins (Table 2.2, score 15.99). In addition, ast-mir-304 is reverse complementary to its top match (dme-miR-304) with 86% identity, which is described in the next section.

Anopheles miR-304 and miR-306 are in two separate miRNA clusters located in introns of protein-coding genes.

As mentioned above, *Anopheles* miR-304 and miR-306 do not have conserved sequences in *Ae. aegypti*. However, they both have homologs in *D. melanogaster* (Figure 2.3). Comparisons between miR-306 sequences in *An. gambiae* and *D. melanogaster* produced a miRscan score of 7.53, which is consistent with its miRNA status. Comparisons between miR-304 sequences in *An. gambiae* and *D. melanogaster* produced no score (NS) during miRscan analysis. The failure to produce a positive score by miRscan does not automatically indicate that miR-304 is not a true miRNA because nine out of the 88 known *C. elegans*/*C. briggsae* miRNAs produced no scores and two even gave negative scores (Lim et al., 2003a). A closer examination of the miR-304 hairpin from *An. gambiae* suggested that it met all of the previously described criteria for miRNA structures (Ambros et al., 2003a; Gu et al., 2007).

Interestingly, miR-304 is closely flanked by miR-12 and miR-283 while miR-306 is in a different cluster with miR-9b and miR-79 (Figure 2.3). Clustering of *An. gambiae* miR-9b and miR-79 is noted on miRBase, but not miR-306; clustering of *An. gambiae* miR–304, -12, and –283 is not predicted in miRBase (Griffiths-Jones, 2004; Griffiths-Jones et al., 2006). Both clusters are within introns of protein coding genes. The miR-12, -304, -283 cluster occurs within a conserved gene of unknown function, while the miR-9b, -79, -306 cluster occurs within an ortholog of a gene coding for a *Drosophila* serine-threonine kinase group protein. The exons flanking each of the miRNA clusters are conserved between *An. gambiae, Ae. aegypti*, and *D. melanogaster*, which indicates that the clusters are orthologous. The order of miRNAs within the introns is conserved between *An. gambiae* and *D. melanogaster*, again supporting the miRNA status of the *Anopheles* miR-304 and miR-306. All miRNAs in the two clusters are in the same orientation as the flanking genes, indicating that these miRNAs may be transcribed from the promoters of their respective flanking genes, which is consistent with previous reports (Smalheiser, 2003; Kim and Kim, 2007). In this regard, it is suggestive that, albeit reverse complementary to each other, the mature miR-304 in *Anopheles* and *Drosophila* are both correctly annotated because the orientation of transcription of miR-304 is consistent with the flanking gene in both species.

Discovery of four novel mosquito miRNAs that have no apparent homolog outside of mosquitoes.

As shown in category IIc of Table 2.2, there are four *An. stephensi* small RNAs that produced no hits to any known miRNAs from any species based upon searches of the miRBase with an e-value cutoff of 10. These sequences are temporarily named ast-miR-x1 through ast-miR-x4. All

four sequences match perfectly to unique locations in the *An. gambiae* genome assembly. Putative precursor sequences flanking the four miRNAs in *An. gambiae* showed strong hairpin structures (Figure 2.4). The precursors of all four miRNAs showed high similarity to *Ae. aegypti* genomic sequences and these hairpin pairs gave miRscan scores between 14.4-17.45, indicating that they strongly resemble the structure and conservation pattern of known miRNAs. In particular, the mature miRNA sequences were 100% conserved between *An. gambiae, An. stephensi,* and *Ae. aegypti*. Northern blot analysis using 17-day old female samples provided further confirmation for these four miRNAs (Figure 2.1, last lane).

Expression profile of mosquito miRNAs across different developmental stages.

We decided to expand the expression analysis beyond the adult stage and investigate the expression profiles of eight miRNAs across different developmental stages, from early embryo to female adult. Shown in Figure 2.1 are northern blot results for four known miRNAs as well as all four new miRNAs. Each miRNA showed a distinct pattern. MiR-9a is expressed in all life stages examined and its expression is reduced in adults, which is consistent with what was observed in *D. melanogaster* (Aravin et al., 2003). The level of miR-210 appears to be higher in late embryo and adult females than in other stages. Let-7 expression begins in late larvae in mosquitoes and continues into adult, again similar to what was observed in *D. melanogaster*. We also determined the expression profile of miR-14, which represents 25% of the miRNA sequences during our cloning experiment (Table 2.2). As shown in Figure 2.1, miR-14 displays a strong signal starting from late embryonic to adult stages. MiR-14 is observed in all life stages in *D. melanogaster* as well (37). Further examination of miR-14 expression across adult lifespan

showed a relatively consistent expression regardless of age, sex, and hematophagy (Figure 2.5). The four novel miRNAs (miR-x1, -2, -3, -4) have unique expression patterns as well, which will likely provide useful clues to their function for future research. For example, miR-x2 showed adult-specific expression while miR-x3 showed predominantly pre-adult expression (Figure 2.1). MiR-x2 was not detected in adult males (Figure 2.6). MiR-x3 was at best weakly expressed in adult males while miR-x1 and x4 were clearly expressed in adult males (data not shown).

MiR-x2 in both An. stephensi and Ae. aegypti is predominantly found in the ovaries and its level is significantly reduced 72 hrs after blood feeding.

We decided to carry out a detailed expression analysis of miR-x2 in both *An. stephensi* and *Ae. aegypti*. When different *An. stephensi* tissues were analyzed (Figure 2.6), miR-x2 showed strong signals in ovaries while no expression was detected in the midgut samples. The expression of miR-x2 was weak in the heads and the "remainders" (thorax plus abdomen without midgut and ovary). Following blood feeding, miR-x2 expression in the ovaries remained high at 24 hours post bloodmeal, but declined sharply by 72 hours post bloodmeal. MiR-x2 was hardly detectable among the other tissues in either of the post bloodmeal time points. The same pattern of expression was observed in the distantly related *Ae. aegypti* (Figure 2.6). We also confirmed the lack of miR-x2 expression in adult males in both *An. stephensi* and *Ae. aegypti* (Figure 2.6). Thus miR-x2 expression profile is conserved between the two divergent mosquitoes in all samples tested.

2.5. Discussion

Two criteria are critical for demonstrating a valid miRNA (31,38). First, expression of an approximately 22 nt RNA should be detected by small RNA cloning or by RNA hybridization methods such as northern blot (Ambros et al., 2003; Berezikov et al., 2006). Second, the miRNA should be traceable back to a precursor with a hairpin structure (Ambros et al., 2003; Berezikov et al., 2006). All 27 miRNAs described in this report, including two miRNA* sequences, met the above criteria. The expression was indicated by direct cloning of small RNAs from 17-day old female *An. stephensi* samples and the precursor hairpin was identified using the genome assembly of a related *Anopheles* species of the same subgenus, *An. gambiae*. Additional evidence to validate miRNA status that was proposed in the literature includes proof of processing to a mature form via Dicer, and conservation of the mature sequence and its precursor (Ambros et al., 2003; Berezikov et al., 2006). For all 27 *Anopheles* miRNAs, we were able to identify conserved sequences encompassing the entire hairpin structure either from the genome assembly of a divergent mosquito species, *Ae. aegypti* (25 out of 27) or from *D. melanogaster* (see Table 2.3). Furthermore, miRscan analysis based on conserved hairpin alignment provided strong support for 26 of the 27 Anopheline miRNAs. In the case of ast-miR-304, which has no conserved sequence in *Ae. aegypti* and no miRscan score, a closer examination of the miR-304 hairpin from *An. gambiae* suggests that it met all of previously described criteria for miRNA structures (Ambros et al., 2003; Gu et al., 2007). Furthermore, the *Anopheles* miR-304 and the *D. melanogaster* miR-304 are both in a conserved miRNA cluster and both are flanked by miR-12 and miR-283, thus lending additional support for the validity of *Anopheles* miR-304. Finally miRNA expression was detected for all nine selected ast-miRNAs during either northern blot or

RPA, complementing our cloning results. Thus we feel that the overall support for the presence of these 27 miRNAs in *Anopheles* mosquitoes is strong. One miRNA, ast-miR-2a, is worth noting. Cloning results as well as comparisons to dme-miR-2a and miRscan predictions are all consistent. However, ast-miR-2 is reverse complementary to aga-miR-2 in miRBase.

As mentioned earlier, miR-304 and miR-306 do not have obvious homologs in *Ae. aegypti*. In both *An. gambiae* and *D. melanogaster*, miR-304 is closely flanked by miR-12 and miR-283 while miR-306 is in a different cluster with miR-9b and miR-79 (Figure 2.3). Both clusters are in the intron of orthologous genes in the two species. Comparisons of the two miRNAs clusters in these three species (Figure 2.3) would suggest that miR-304 and miR-306 may have either been lost in *Ae. aegypti* or evolved to significantly different sequences. As the mature miRNA sequences of the *Anopheles* and *Drosophila* miR-304 are reverse complementary to each other, it is possible that the miR-304 precursor sequence may have undergone an inversion after the two dipteran lineages separated. Thus comparisons of clear orthologous miRNA clusters will likely shed new light on the evolution (expansion, loss, rearrangement) of miRNAs, which has not been well studied.

The expression profiles of the eight mosquito miRNAs are informative. When *D. melanogaster* data are available for comparison as in the cases of let-7 and miR-9a, similar expression profiles were found between *An. stephensi* and *D. melanogaster*. This is not surprising as these conserved miRNAs are likely to have similar functions in these Dipteran insects. MiR-14 was expressed in the same stages in *An. stephensi* as in *D. melanogaster*. We provided extended expression analysis on miR-14, the miRNA that represents 25% of the sequenced miRNAs during the

cloning of the 17-day old *An. stephensi* samples. We showed that the miR-14 level increased slightly during embryonic development and remained relatively high through larvae, pupae and adult stages and we did not observe significant changes in adults regardless of age, sex, and blood feeding status. These results do not necessarily imply that mosquito miR-14 is important to longevity, a function of miR-14 demonstrated in *D. melanogaster*. Nonetheless, it appears that miR-14 is important across different mosquito life stages from embryos to aged adults. Further research on the targets and function of miR-14 in mosquitoes will help determine whether it is important to mosquito longevity. The expression profile of miR-x1 through miR-x4 confirmed our cloning results for these new miRNAs, and provided useful information for future research into their functions. The expression of miR-x2 in *An. stephensi* was adult specific as well as female specific. MiR-x2 was predominantly expressed in the ovary and its level was reduced 72 hrs after blood feeding. These results indicate that miR-x2 is likely involved in *An. stephensi* female reproduction. The same pattern of expression was shown for miR-x2 in *Ae. aegypti*, a mosquito that is highly divergent from *An. stephensi*. Thus the function of miR-x2 may be conserved among divergent mosquitoes.

We have identified 10 miRNAs previously unknown to mosquitoes, four of which did not match any known miRNAs in any organism. These four miRNAs are conserved between *An. stephensi*, *An. gambiae*, and *Ae. aegypti*, suggesting important functions possibly common to all mosquitoes as the *Anopheles* and *Aedes* genera are two of the most divergent among all mosquitoes, separated approximately 145-200 million years ago (Krzywinski et al., 2006). Considering the modest number of small RNA clones we sequenced during this study and the specific developmental stage of our total RNA source, we suspect that a number of novel

miRNAs still await to be discovered in mosquitoes. This study demonstrates the importance of direct cloning and expression profile analysis to the identification and characterization of conserved as well as mosquito-specific microRNAs, some of which will likely regulate genes that significantly affect mosquito biology and perhaps mosquito-pathogen interactions.

Three research articles became available when this manuscript was being reviewed. Two of these papers describe additional miRNAs from *Drosophila* (Ruby et al., 2007; Stark et al., 2007) and a third paper reports cloning of 10 distinct miRNAs and primer extension analysis of an additional 8 miRNAs from *An. gambiae* midgut samples infected with *Plasmodium bergei* (Winter et al., 2007). Comparisons between the 27 miRNAs reported in this manuscript and those published in the above three papers showed that *An. stephensi* miR-x1 and miR-x2 are nearly identical to *An. gambiae* miR-996 and miR-989, respectively (Winter et al., 2007) and homologs of these two miRNAs are found in the expanded list of *Drosophila* miRNAs (Ruby et al., 2007; Stark et al., 2007). Thus miR-x1 and miR-x2 are not unique to mosquitoes. Ast-miR-76 showed a perfect match to the newly identified miR-981 in *Drosophila* (Ruby et al., 2007; Stark et al., 2007), thus ast-miR-76 should be renamed as ast-miR-981. Additional five of the cloned miRNAs reported in Winter et al. (2007) match five of the 27 miRNAs reported in this manuscript. They are let-7, miR-281, miR-34, miR-12, and miR-306. *An. stephensi* miR-x3 and miR-x4 showed no similarity to any miRNAs in the miRbase which included updates from the above three papers. Thus miR-x3 and miR-x4 remain novel and potentially specific to mosquitoes. We also compared miRNA expression profiles obtained in our study using northern analysis with the profiles of overlapping miRNAs reported in Winter et al. (2007), which were obtained using primer extension. Winter and colleagues state that miR-989 (miR-x2) is expressed only in

midguts of *An. gambiae* (page 6958). However, we have shown that miR-x2 (miR-989) is predominantly expressed in ovaries in both *An. stephensi* and *Ae. aegypti* and miR-x2 was not detected in midguts under our condition. We note that the Figure 2.2 of Winter et al. (2007) does show a much stronger expression in "leftovers", which include ovaries, than in midguts.

Conclusions

This study provides experimental evidence for 23 conserved and four new microRNAs in *An. stephensi* mosquitoes. Comparisons between miRNA gene clusters in *An. gambiae*, *Ae. aegypti,* and *D. melanogaster* suggest the loss or significant change of two miRNA genes in *Ae. aegypti*. Expression profile analysis of eight miRNAs, including the four new miRNAs, revealed distinct patterns from early embryo to adult stages in *An. stephensi*. Further analysis showed that miR-x2 is likely involved in female reproduction and its function may be conserved among divergent mosquitoes. Consistent expression of miR-14 suggests that miR-14 is likely important across all mosquito life stages from embryos to aged adults. Understanding the functions of mosquito miRNAs will undoubtedly contribute to a better understanding of mosquito biology including longevity, reproduction, and mosquito-pathogen interactions, which are important to disease transmission.

2.6. Acknowledgements

We would like to thank Randy Saunders for maintaining mosquito colonies, and Ray Miller for improving the resolution of the hairpin images. We thank Jim Biedler and Zach Adelman for

reviewing a draft of the manuscript. This work is supported by NIH grants (AI070854 and AI063252) to Z.T.

Authors' Contributions

EAM assisted in experimental design. EAM conducted the experimental work, analyzed data, and wrote a rough draft of the manuscript. ZT prepared the experimental design, helped with data analysis, and revised the manuscript; ZT is the principal investigator who oversaw this project.

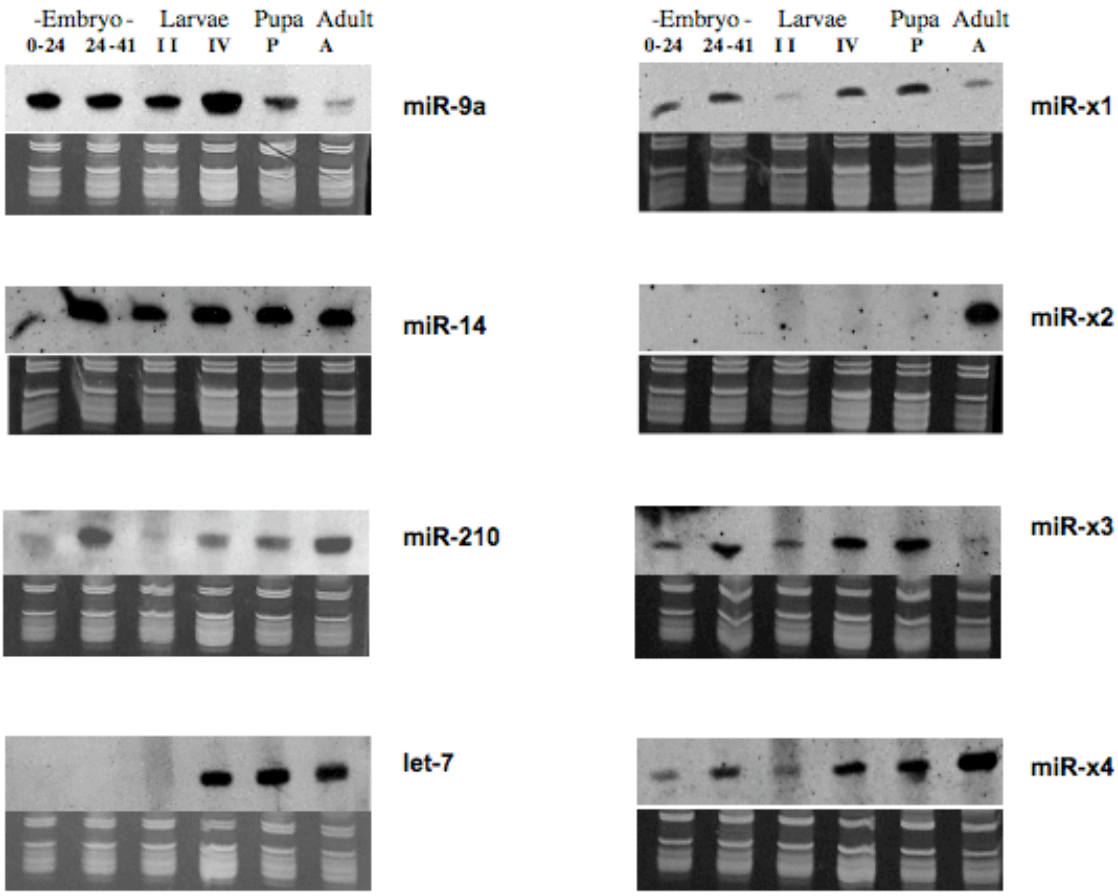


Figure 2.1. Northern analysis of eight miRNAs across different developmental stages in *An. stephensi*. Shown here are eight northern blots performed using Dig-labeled miRCURY LNA probes designed for hybridization to either miR-14, let-7, miR-9a, miR-210, or to one of the four novel miRNAs (miR-x1– x4). The top panels are northern results and the bottom panels are RNA gels for verification of small ribosomal and tRNA integrity and equal loading of total RNA. ssDNA size markers (19 and 23 nts, not shown) were also visualized on the RNA gel for size estimation. Ten micrograms of total RNA for each sample were used. Developmental stages examined were early embryo (Embryo 0-24: 0-24 hrs after egg deposition), late embryo (Embryo 24-41: 24-41 hrs after egg deposition), intermediate and late larval stages (II and IV,

respectively), Pupa (P), and Adult (A). To be consistent with our cloning experiment, 17-day old adult females were used in these northern experiments.

Figure 2.2.A

```
An. gambiae   5′ GCCTTATCTGGATCGGGTTTCGCTGGCTAACGTGTTGTGCGAGCGAGCAAAAAAAACTTCAATTCGTTGTCGACGAAACCTGCACGGAGTAAGGC 3′
Ae. aegypti   5′ -CCTTGCCTGGATCGGGTTTCGC-GGCGAACGGGCTT------CCACCATGATGAATAGCAGTTCGTTGTCGACGAAACCTGCACATTGCAAGGC 3′
                   ++++  ++++++++++++++++ +++ ++++ + +       + + ++  +  ++   ++ ++++++++++++++++++++++   + +++++
```

Figure 2.2.B

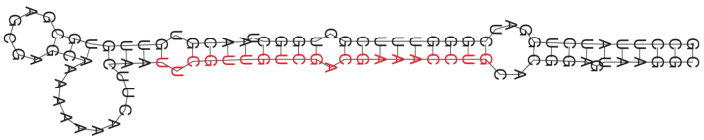


Figure 2.2.C

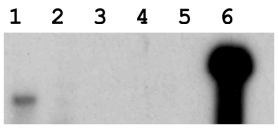


Figure 2.2. Analysis of ast-mir-76, a miRNA that was previously unknown in mosquitoes. The mature miRNA sequence was cloned from *An. stephensi*, and the hairpin precursor sequence was obtained from the *An. gambiae* genome assembly. **A)** Alignment of the pre-miRNA hairpins found in *An. gambiae* and *Ae. aegypti*. The mature miRNA is marked in red while miRNA* is marked in blue. Conserved nucleotides are indicated by a "+". **B)** Hairpin structure of *An. gambiae* mir-76. The mature miRNA predicted by miRscan (Table 2) is shown in red. **C)** RPA analysis of ast-mir-76. Lane 1, *An. stephensi* RNA with probe and digested; Lane 2, yeast RNA with probe and digested; Lane 3, yeast RNA without probe and digested; Lanes 4 and 5, empty lanes; Lane 6, undigested probe. A band of the correct size was only observed in *An. stephensi* total RNA samples (Lane 1). The size of the protected RNA product in lane 1 was estimated to be 24 nucleotides using markers as described in Figure 2.1. This size is as expected (the protected 20-nt long ast-miR-76 plus 4 undigested adenosines, see Methods).

Figure 2.3.A

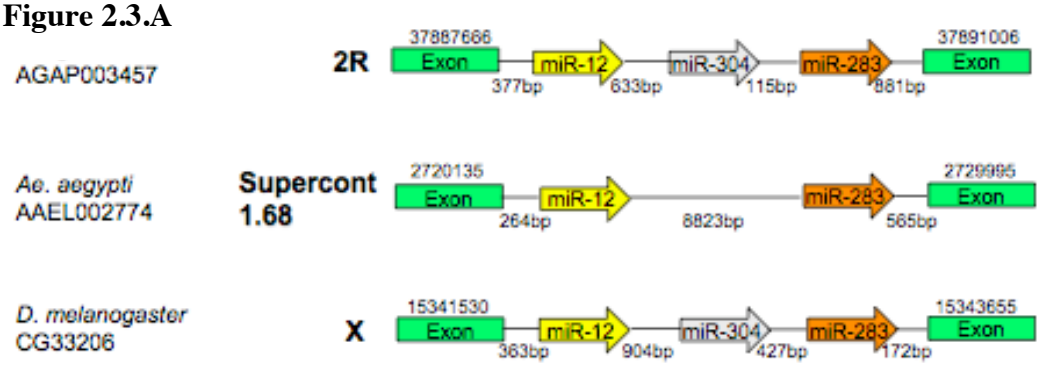


Figure 2.3.B

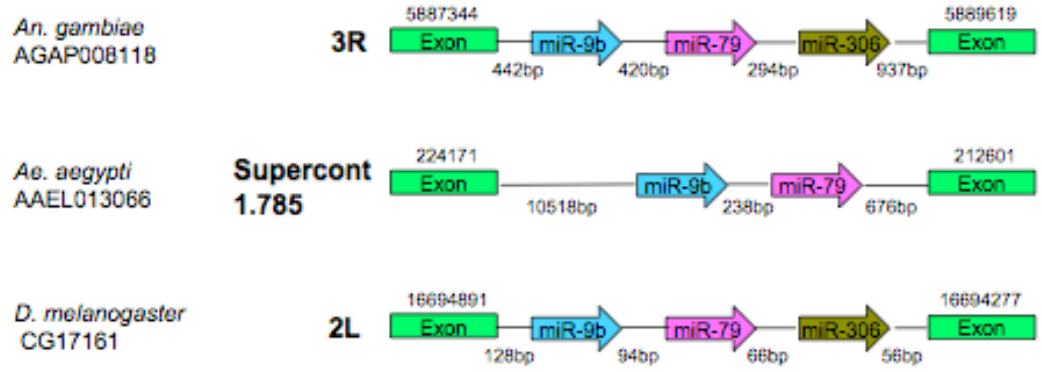


Figure 2.3. Clustering of miRNAs genes. A) A miRNA gene cluster within an intron of a conserved gene of unknown function. The miRNA gene cluster contains miR-12, -283, and -304. B) A miRNA gene cluster within an intron of a gene coding for a serine-threonine kinase group protein. The miRNA gene cluster contains miR-9b, -79, and -306. Note that one miRNA was not found in the genome of *Ae. aegypti* in both panels. Species name and gene identification are provided at the left side of the figure. Chromosome or supercontig numbers are indicated right next to diagram depicting the miRNA gene clusters. Chromosomal or supercontig positions of the regions depicted are above the boxes showing the exons. MiRNA genes are shown as open arrows. The distance between the miRNA genes and neighboring exons are indicated below the diagram. The figure is not drawn to scale. The exons shown in both panels are orthologous as indicated by conserved amino acid sequences.

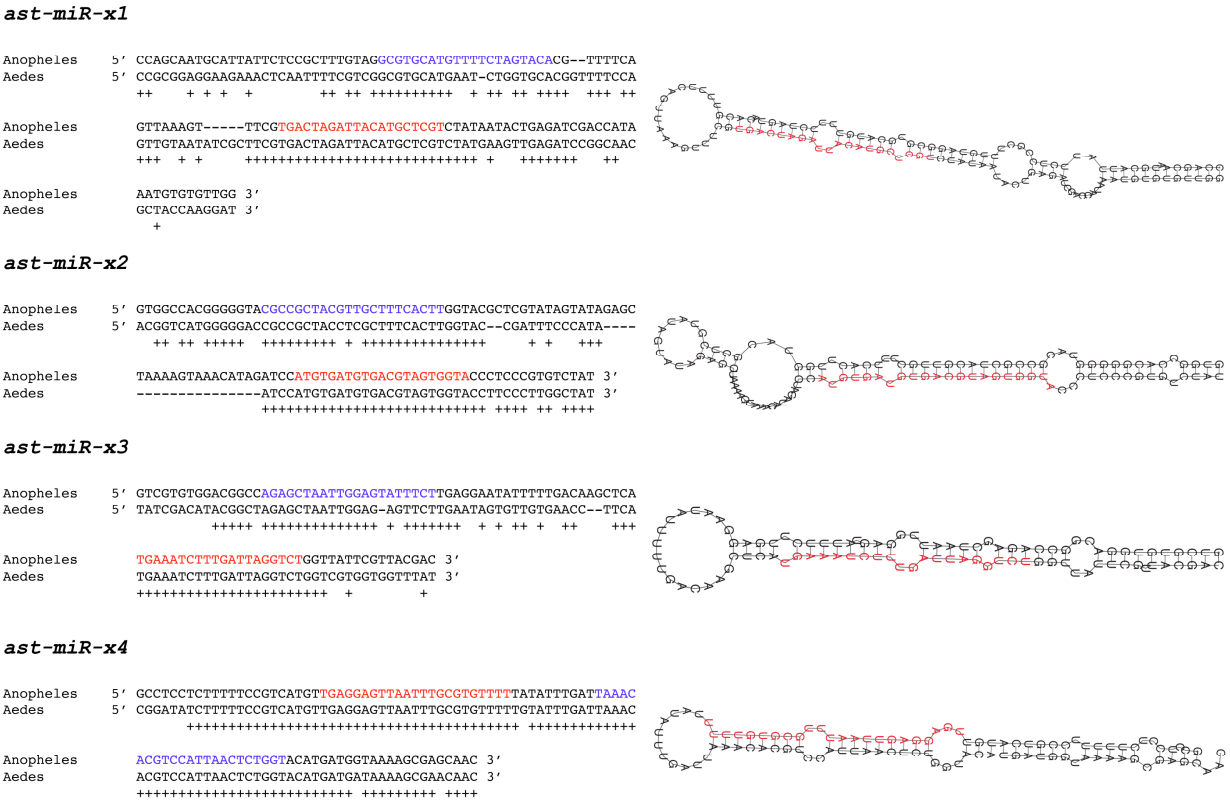


Figure 2.4. Sequence alignment and predicted secondary structure of four novel miRNAs. Shown on the left are the sequence alignments between *An. gambiae* and *Ae. aegypti* miRNA precursor hairpins. Plus signs indicate conservation. The mature miRNA is marked in red while miRNA* is marked in blue. Note the perfect conservation of the mature miRNA (red), high conservation of the miRNA* sequence (blue), and lower conservation of the surrounding stem and loop structure, a hallmark conservation pattern of pre-miRNAs. Shown on the right are the predicted secondary structures of corresponding *An. gambiae* miRNA hairpins. The mature miRNA is marked in red on the hairpin.

Figure 2.5.A

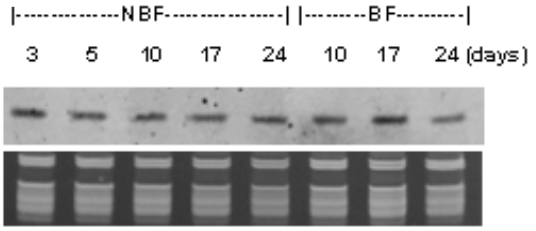


Figure 2.5.B

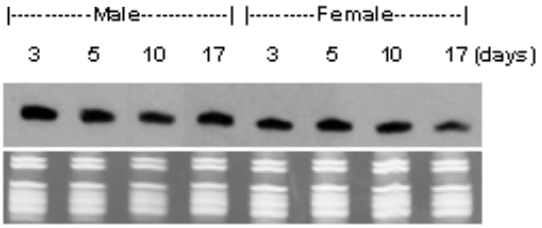


Figure 2.5. MiR-14 expression across *An. stephensi* adult lifespan. Shown here are northern blots performed using Dig-labeled miRCURY LNA probes designed for hybridization to miR-14. The top panel is the northern result and the bottom panel is a corresponding RNA gel for verification of small ribosomal and tRNA integrity and equal loading of total RNA. ssDNA size markers (19 and 23 nts, not shown) were also visualized on the RNA gel for size estimation. Ten micrograms of total RNA for each sample were used. **A)** miR-14 expression in *An. stephensi* adult females fed with either sugar water (NBF, non-bloodfed) or blood meal (BF, bloodfed). The samples were 3, 5, 10, 17, and 24 day old adult females that were maintained on sugar water as well as adult females that were fed on blood on day 5 and collected at day 10, 17, and 24 after emergence. Bloodfed females were allowed to oviposit two days after the blood meal**. B)** miR-14 expression in *An. stephensi* males and NBF females between 3-17 days of age. We did not extend the comparative analysis to 24 days because the majority of males do not survive that long.

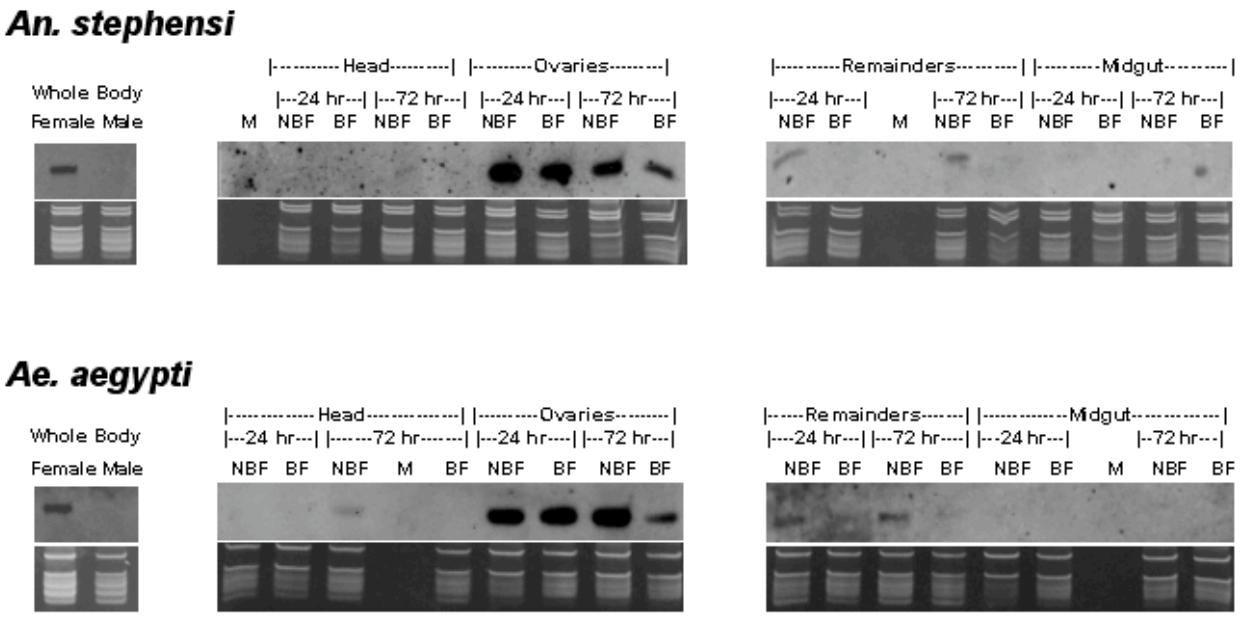


Figure 2.6. Expression of miR-x2 in *An. stephensi* and *Ae. aegypti*: sex-specificity, tissue distribution and the impact of blood feeding. Shown here are northern blots performed using Dig-labeled miRCURY LNA probes designed for hybridization to miR-x2. The top panels are northern results and the bottom panels are RNA gels for verification of small ribosomal and tRNA integrity and equal loading of total RNA. ssDNA size markers (19 and 23 nts, not shown) were also visualized on the RNA gel for size estimation. On the left for each species, a comparison between 5-day old adult male and 5-day old non-bloodfed female is shown. Five micrograms of total RNA isolated from the whole mosquitoes were used. The middle and right panels are comparisons between adult female tissues or body parts in each species. Tissues used were Heads, Ovaries, Midguts, and Remainders. There were four samples for each tissue: BF, tissue sample from bloodfed females at 24 and 72 hrs post-bloodfeeding; NBF, tissue sample from non-bloodfed (sugar-fed) females at equivalent time points compared to the blood-fed samples. Five micrograms of total RNA for each sample were used. The markers lane is

designated with an ‘M’ although the markers are not within the gel image panel because they are below the size of the ribosomal and tRNA.

RNA Species	Number Present	% of Total Clones
MiRNAs shown in Table 2.2 [1]	148	40.22%
Possible miRNAs listed in Table 2.3 [1]	13	3.53%
rRNA	3	0.82%
tRNA	16	4.35%
Unidentified [2]	51	13.86%
Low Quality or Short Sequences [3]	137	37.23%
TOTAL	**368**	**100%**

Table 2.1. Classification of Cloned Small RNAs in *An. stephensi*.

Notes:

1. We decided to include only miRNA candidates that match 100% to a locus in the *An. gambiae* genome as true miRNA candidates shown in Table 2.2. There are several additional miRNA candidates that have one nucleotide mismatch to loci in the *An. gambiae* genome, which could either result from real differences between *An. stephensi* and *An. gambiae*, or errors introduced during cloning or sequencing. These miRNA candidates are not included in Table 2.2 but are provided in Table 2.3 as we feel further investigation is necessary to ascertain their identities.

2. Sequences do not match any known miRNAs or any other small RNAs or mRNAs.

3. Low quality sequences and sequences less than 17 nucleotides were not analyzed further.

miRNA [1]	miRBase Name [3]	Sequence [4]	Occur [5]	Location [6]	Score [7]
ast-let-7 (Ia)	ast-let-7	UGAGGUAGUUGGUUGUAUAGU	12	3R, 10270763 (-)	14.8
ast-miR-124 (Ia)	ast-miR-124	UAAGGCACGCGGUGAAUGC	1	3R, 29002032 (+)	11.32
ast-miR-14 (Ia)	ast-miR-14	UCAGUCUUUUUCUCUCUCCUA	38	3R, 24898098 (+)	16.9
ast-miR-210 (Ia)	ast-miR-210	UUGUGCGUGUGACAACGGCUA	7	X, 21450383 (+)	12.72
ast-miR-276 (Ia)	ast-miR-276-5p	AGCGAGGUAUAGAGUUCCUA	4	2L, 18991766 (+)	15.96
ast-miR-277 (Ia)	ast-miR-277	UAAAUGCACUAUCUGGUACGA	4	2R, 28234532 (-)	15.78
ast-miR-277* (Ia) [2]		CGUGUCAGAGGUGCAUUUA	1	2R, 28234583 (-)	15.78
ast-miR-281 (Ia)	ast-miR-281	UGUCAUGGAAUUGCUCUCUUUA	24	2L, 17362444 (-)	16.14
ast-miR-283 (Ia)	ast-miR-283	AAAUAUCAGCUGGUAAUUCU	1	2R, 37890092 (-)	14.92
ast-miR-317 (Ia)	ast-miR-317	UGAACACAUCUGGUGGUAUCU	10	2R, 28252007 (-)	10.57
ast-miR-8 (Ia)	ast-miR-8	UAAUACUGUCAGGUAAAGAUGU	13	3L, 38943098 (+)	14.73
ast-miR-8* (Ia) [2]		CAUCUUACCGGGCAGCAUUA	1	3L, 38943058 (+)	14.73
ast-miR-9a (Ia)	ast-miR-9a	UCUUUGGUUAUCUAGCUGUAU	3	2L, 15089338 (-)	13.6
ast-miR-11 (Ib)	ast-miR-11	CAUCACAGUCUGAGUUCUUGCU	1	2R, 13042084 (-)	14.97
ast-miR-276a (Ib))	ast-miR-276-3p	UAGGAACUUCAUACCGUGCUCU	2	2L, 18991809 (+)	14.72
ast-miR-34 (Ib)	ast-miR-34	UGGCAGUGUGGUUAGCUGGUU	5	2R, 28232720 (-)	17.77 [8]
ast-miR-87 (Ib)	ast-miR-87	GGUGAGCAAAUAUUCAGGUGU	1	X, 261196 (-)	12.12
ast-miR-12 (IIa)	ast-miR-12	UGAGUAUUACAUCAGGUACUGGU	2	2R, 37888125 (-)	8.45
ast-miR-375 (IIa)	ast-miR-375	UUUGUUCGUUUGGCUCGAGUUA	1	3R, 51640581 (-)	10.44
ast-miR-2a (IIa)	ast-miR-2-1	UAUCACAGCCAGCUUUGAUGAG	2	2L, 37757111 (-)	15.99
ast-miR-304 (IIb)	ast-miR-1889	ACACAUUACAGAUUGGGAUUA	2	2R, 37888805 (-)	NS [9]
ast-miR-306 (IIb)	ast-miR-306	UCAGGUACUGGAUGACUCU	1	3R, 5888649 (-)	7.53 [9]
ast-miR-76 (IIb)	ast-miR-981	UUCGUUGUCGACGAAACCUG	2	X, 1228349 (+)	12.73
ast-miR-x1 (IIc)	ast-miR-996	UGACUAGAUUACAUGCUCGU	1	2R, 55572846 (-)	16.19
ast-miR-x2 (IIc)	ast-miR-989	AUGUGAUGUGACGUAGUGGUA	6	3L, 2905484 (+)	15.15
ast-miR-x3 (IIc)	ast-miR-1890	UGAAAUCUUUGAUUAGGUCU	1	3R, 21181098 (-)	17.45
ast-miR-x4 (IIc))	ast-miR-1891	UGAGGAGUUAAUUUGCGUGUUUU	2	3R, 5819094 (-)	14.40

Table 2.2. Sequence and Characteristics of Cloned miRNAs in *An. stephensi*.

Notes:

1. The names in this column were temporarily assigned according to similarity to known miRNAs. ast stands for *An. stephensi*. *An. stephensi* miRNAs shown here are divided into two main categories as indicated in bracket. Category I includes miRNAs that match *An. gambiae*

miRNA predictions that were previously reported in miRBase (category Ia) or Chatterjee and Chaudhuri, 2006 (category Ib). Category II are miRNAs that have not been reported in *An. gambiae* or any other mosquito species. This category includes miRNAs that displayed perfect or near-perfect (1 nucleotide mismatch only) match to miRNAs from non-mosquito species in miRBase (category IIa), miRNAs that displayed 80% or higher overall similarity to miRNAs from non-mosquito species in the miRBase (category IIb), as well as miRNAs that showed no match to any miRNAs in miRBase at the default e-value cutoff of 10 (category IIc). miRNAs in category IIc are temporarily labeled with an "x" in front of an Arabic numeral.

2. A "*" delineates that the sequence matches the miRNA* strand of miRNA:miRNA* heteroduplex.

3. These are formal names assigned by miRBase, which were received during the proofing stage. See the last paragraph of the Discussion section for details.

4. The longest sequence of each miRNA is shown. Variants with different ends are shown in Table 2.3. The observation of end variants has literary precedence (see (Gu et al., 2007)).

5. Occurrence refers to the number of times a sequence appeared during our cloning and sequencing.

6. Location refers to the location of match in the *An. gambiae* genome as there is no *An. stephensi* genome assembly available. The location is indicated by chromosome name, start position of the mature miRNA, and strand orientation. All matches are 100%.

7. Score refers to the result of miRscan analysis. Unless otherwise noted, the pre-miRNA sequences used for miRscan are pairs from *An. gambiae* and *Ae. aegypti*. In the original miRscan paper (Lim et al., 2003a), most of the validated *C. elegans* miRNAs received scores of 9 or

above although a small number of them received scores significantly less and some even received negative scores.

8. The precursor sequence of *Ae. aegypti* miR-34 contains a large segment in the loop region, which may be the cause for no score or "NS" by miRscan for the *An. gambiae* and *Ae. aegypti* pre-miR-34 alignment. Thus *An. gambiae* and *D. melanogaster* pre-miR-34 alignment was used for miRscan, which produced a score of 17.99.

9. We could not find homologs for miR-304 and miR-306 in *Ae. aegypti*. Homologs were found in *D. melanogaster* (see Figure 2.3), which were used for miRscan analysis shown here.

Proposed Name	Closest Match	Non-Redundant List of Raw Sequences Observed (5' TO 3')	Times Observed	*An. gambiae* Precursor Sequence (5' to 3')
ast-let-7	aga-let-7	TGAGGTAGTTGGTTGTATAGTA	10	See miRBase Registry at http://microrna.sanger.ac.uk/
		TGAGGTAGTTGGTTGTATA	2	
ast-miR-2a	dme-miR-2a	TCACAGCCAGCTTTGATGAGTA	1	GCATGTCCTGTCTCTCAAAGTGGCTGTGAAATGGTGCACTTTCGATCC
		TATCACAGCCAGCTTTGATA	1	GATATCTCATATCACAGCCAGCTTTGATGAGCTAGGCCGTGT
ast-miR-8	aga-miR-8	TAATACTGTCAGGTAAAGATGTTA	4	See miRBase Registry at http://microrna.sanger.ac.uk/
		TAATACTGTCAGGTAAAGATGTA	7	
		TAATACTGTCAGGTAAAGATA	2	
ast-miR-8*	aga-miR-8*	CATCTTACCGGGCAGCATTA	1	ACATCTTACCGGGCAGCATTAGATATGTTATCGGATATTTCTAATACT GTCAGGTAAAGATGT
ast-miR-9a	aga-miR-9a	GAAATCTTTGGTTATCTAGCTGTA	1	See miRBase Registry at http://microrna.sanger.ac.uk/
		TCTTTGGTTATCTAGCTGTAT	1	
		TCTTTGGTTATCTAGCTGTA	1	
ast-miR-11	aga-miR-11	CATCACAGTCTGAGTTCTTGCTA	1	GCATCCGGCAAGAACTTGGAACTGTGATCTGTGTGGTTATTAGCCGC GTTCGAAGCTCTTCACATCACAGTCTGAGTTCTTGCTCGATGG
ast-miR-12	dme-miR-12	CTGAGTATTACATCAGGTACTGGTA	1	CCGCCGGGGTGAGTATTACATCAGGTACTGGTGTGTAATTTAAACGAC
		TGAGTATTACATCAGGTACTGGTA	1	CATCGGACATGGGGGCTGCCCAACTCCTCTTCTGTCCCCGGTCAAGCT ATCAGTACTTGTGTTATA
ast-miR-14	aga-miR-14	TCAGTCTTTTTCTCTCTCCTA	37	See miRBase Registry at http://microrna.sanger.ac.uk/
		CAGTCTTTTTCTCTCTCCTA	1	
ast-miR-34	aga-miR-34	TGGCAGTGTGGTTAGCTGGTTGT	1	AGGAGGCAATATACGCTCTGGCAGTGTGGTTAGCTGGTTGTGTGGTT
		TGGCAGTGTGGTTAGCTGGTTA	2	TTCCCATCTTCACAGCCACTATCCGCCCTGCCGTCGCGCTAATGC
		TGGCAGTGTGGTTAGCTGGTT	1	
		TGGCAGTGTGGTTAGCT	1	

ast-miR-76	cel-miR-76 (16bases)	TTCGTTGTCGACGAAACCTG	2	GCCTTATCTGGATCGGGTTTCGCTGGCTAACGTGTTGTGCGAGCGAG CAAAAAAAACTTCAATTCGTTGTCGACGAAACCTGCACGGAGTAAGG C
ast-miR-87	aga-miR-87	GGTGAGCAAATATTCAGGTGTA	1	GATTTCTCTCGGCCCAGCCTGAAATTTGCTAAACCTGATACGTGTTCG AAACCAAGGTGAGCAAATATTCAGGTGTGTCGACGAGCGGTCACATT CGTCG
ast-miR-124	aga-miR-124	TAAGGCACGCGGTGAATGCTA	1	See miRBase Registry at http://microrna.sanger.ac.uk/
ast-miR-210	aga-miR-210	CTTGTGCGTGTGACAACGGCTATC CTTGTGCGTGTGACAACGGCTA TTGTGCGTGTGACAACGGCTA	1 2 4	See miRBase Registry at http://microrna.sanger.ac.uk/
ast-miR-276	aga-miR-276	AGCGAGGTATAGAGTTCCTA GCGAGGTATAGAGTTCCTA	2 2	See miRBase Registry at http://microrna.sanger.ac.uk/
ast-miR-276a	aga-miR-276a	AGGAACTTCATACCGTGCTCTA TAGGAACTTCATACCGTGCTA	1 1	CCATCAGCGAGGTATAGAGTTCCTACGGTAATCGATTGAAACTTTGTA GGAACTTCATACCGTGCTCTTGG
ast-miR-277	aga-miR-277	TAAATGCACTATCTGGTACGATA TAAATGCACTATCTGGTA	3 1	See miRBase Registry at http://microrna.sanger.ac.uk/
ast-miR-277*	aga-miR-277*	CGTGTCAGAGGTGCATTTA	1	GTTTTGGGGTACGTGTCAGAGGTGCATTTACATCGAACTATTCCAGTT GAGGTATTTGTAAATGCACTATCTGGTACGACATTCCAGAAT
ast-miR-281	aga-miR-281	TGTCATGGAATTGCTCTCTTTA	24	See miRBase Registry at http://microrna.sanger.ac.uk/
ast-miR-283	aga-miR-283	AAATATCAGCTGGTAATTCTA	1	See miRBase Registry at http://microrna.sanger.ac.uk/

ast-miR-304	dme-miR-304	ACACATTACAGATTGGGATTA	2	TGTGGCGTGGATGGGCGGCAATCTCAAATTGTAATAGTGTGTCACCG TTAACGGCTGGACCATTGTGCTGCCATCGTGATCCACCGAGAAAAAC ACACATTACAGATTGGGATTACCCCGTCCATGCTCCG
ast-miR-306	dme-miR-306	TCAGGTACTGGATGACTCTTA	1	GTCCTTTTCACATGGTCAGGTACTGGATGACTCTCAGTTGTGTGTAAA AGATGCTGAGGGCCTTCTGGTACCTACCCCAGTGAGGACGCCA
ast-miR-317	aga-miR-317	TGAACACATCTGGTGGTATCTTA TGAACACATCTGGTGGTA TGAACACATCTGGTGGT	4 5 1	See miRBase Registry at http://microrna.sanger.ac.uk/
ast-miR-375	ame-miR-375	TTTGTTCGTTTGGCTCGAGTTA	1	TGTTGCGATGAGACAGAATTTGGATTACTTAAGCCACGCGTACAGAA ACTTTTTAGAATGAAAGAGTTTGTTCGTTTGGCTCGAGTTATGCCGGT TCTAAATTGCAGCA
ast-miR-x1	N/A	TGACTAGATTACATGCTCGT	1	CCAGCAATGCATTATTCTCCGCTTTGTAGGCGTGCATGTTTTCTAGTA CACGTTTTCAGTTAAAGTTTCGTGACTAGATTACATGCTCGTCTATAA TACTGAGATCGACCATAAATGTGTGTTGG
ast-miR-x2	N/A	ATGTGATGTGACGTAGTGGTA	6	GTGGCCACGGGGGTACGCCGCTACGTTGCTTTCACTTGGTACGCTCG TATAGTATAGAGCTAAAAGTAAACATAGATCCATGTGATGTGACGTAG TGGTACCCTCCCGTGTCTAT
ast-miR-x3	N/A	TGAAATCTTTGATTAGGTCT	1	GTCGTGTGGACGGCCAGAGCTAATTGGAGTATTTCTTGAGGAATATT TTTGACAAGCTCATGAAATCTTTGATTAGGTCTGGTTATTCGTTACGA C
ast-miR-x4	N/A	TGAGGAGTTAATTTGCGTGTTTT	2	GCCTCCTCTTTTTCCGTCATGTTGAGGAGTTAATTTGCGTGTTTTTATA TTTGATTAAACACGTCCATTAACTCTGGTACATGATGGTAAAAGCGAG CAAC
Possible ast-miR-8	aga-miR-8	TAATACTGTCAGGTAAGGATGTA	1	GGGTGTCTGTTCACATCTTACCGGGCAGCATTAGATATGTTATCGGAT ATTTCTAATACTGTCAGGTAAAGATGTCGTCCGAGCCC

Possible ast-miR-12	dme-miR-12	TGAGTATTACGTCAGGTACTGGTA	2	CCGCCGGGGTGAGTATTACATCAGGTACTGGTGTGTAATTTAAACGA CCATCGGACATGGGGGCTGCCCAACTCCTCTTCTGTCCCCGGTCAAG CTATCAGTACTTGTGTTATACTCTCCTCGTCGG
Possible ast-miR-14	aga-miR-14	TCAGTCTTTCTCTCTCTCCTA	1	GCCCGATAAGCCTGTGGGAGCGAGATTAAGGCTTGCTGGTTATCAC GTTAAACGTAGTCAGTCTTTTTCTCTCTCCTATCGGTACTAACGGTG C
		TCAGTCTTTTTCTCTCCCCTA	1	GCCCGATAAGCCTGTGGGAGCGAGATTAAGGCTTGCTGGTTATCAC GTTAAACGTAGTCAGTCTTTTTCTCTCTCCTATCGGTACTAACGGTG C
Possible ast-miR-279	aga-miR-279	TGACTAGATCCACACTCCTTA	1	TTATGTTTAGAAAAAAACTTTTCCTATCAGTGTAAATGGGTGTGAATC TAGTGATTTCACATGAATTTTCGATTGTGACTAGATCCACACTCATTA ATGTTGTTTGGAACCATTCAGCCGACATTCAAGAGAAGGAATGAGTA TTTTTACCTTTGCTTAA
Possible ast-miR-281	aga-miR-281	TGTCATGGAATTGCTCTCTTCA	1	See information on ast-miR-281
		TGTCATGGAATTGCTCTCTCTA	1	
Possible ast-miR-72	cel-miR-72	TGTCAAGATGTTGGCATAGCTA	1	AAATTGGCAACTTGGCAAGATGTTGGCATAGCTAAAAGATTTGGATTT AGCTATTCAACTTCTTGTTTATTTTTCCAATTT
Possible ast-miR-9b	aga-miR-9b	ACTTTGGTGATTTAGCTGTA	4	TGCCACTTATTGGGACTTTGGTGATTTTAGCTGTATGTTGAATGCATT CCTAACCACATATAGCTTTATCACCAAAAACCTAATGTGTGTGTA

Table 2.3. MiRNA variants observed in *An. stephensi* females. A non-redundant list of all variants of each miRNA, some with different ends (mostly 3' truncation or end addition of T or A) are shown. Such end variants are common (Gu et al., 2007) and we also observed the presence of several "linker only" sequences which usually contained an additional TA, validating that these end variations are introduced during cloning. Sequence end variants were resolved by comparisons to the genome and to previously published miRNA. Only sequences that are longer than 17 nucleotides were counted. Within the listed precursor sequences are mature miRNAs, highlighted in red. *Aedine* precursors receiving favorable mirScan scores are also listed.

Chapter 3.

Direct Sequencing and Expression Analysis of miRNAs in *Aedes aegypti*

Manuscript in preparation for submission to *Genome Biology*

Song Li*, Edward A Mead*, Shaohui Liang*, and Zhijian Tu

3.1. Abstract

MicroRNAs (miRNAs) are a novel class of gene regulators whose biogenesis involves hairpin structures called precursor miRNAs, or pre-miRNAs. A pre-miRNA is processed to make a miRNA:miRNA* duplex, which is separated to generate a mature miRNA and a miRNA*. The miRNA* strand is often rapidly degraded except in cases where both the miRNA and miRNA* strands may generate functional miRNAs. MiRNAs play key regulatory roles during embryonic development as well as other cellular processes. They are also implicated in control of viral infection as well as innate immunity. Direct experimental evidence for mosquito miRNAs has been recently reported in *Anopheline*

---

* These authors contributed equally and are listed in alphabetical order.

mosquitoes based on small-scale cloning efforts.

In this study, we obtained approximately 120,000 small RNA sequences by 454 pyro-sequencing of samples that were isolated from mixed-age embryos, sugar-fed midguts, and blood-fed midguts from the yellow fever mosquito, *Aedes aegypti*. We also performed bioinformatic analysis on the *Ae. aegypti* genome assembly to identify evidence for additional miRNAs. The combination of these approaches allowed us to uncover 98 different pre-miRNAs in *Ae. aegypti* which produced 86 distinct miRNAs and 20 distinct miRNA* sequences. The majority of the miRNAs have homologs in *D. melanogaster* or other related species. Thirteen miRNAs, including eight novel miRNAs identified in this study, are only found in mosquitoes so far. Eighty-nine of the *Ae. aegypti* miRNA and miRNA* sequences showed a perfect match to small RNA sequences from at least one of the three samples.

We identified two cases in which the previously reported miRNA* sequences are at least dozens fold more abundant than the miRNA sequences and six cases in which revision of previously annotated miRNAs may be necessary. There are 14 clusters of miRNAs and there are 17 cases where more than one pre-miRNA hairpin produces the same or highly similar mature miRNAs in *Ae. aegypti*. Expression profiles of 17 of the 86 miRNAs were analyzed by northern blot in *Ae. aegypti*, confirming the results of miRNA discovery. A number of miRNAs showed higher levels in midguts from blood-fed females than that from sugar-fed females, which was confirmed by northern blots on two of these miRNAs. Northern blots also revealed several miRNAs that showed stage-specific expression.

Detailed expression analysis of eight of the 13 mosquito-specific miRNAs in four divergent genera identified cases of clearly conserved expression patterns and obvious differences. Four of the 13 miRNAs are lineage-specific within mosquitoes. Taken together, this study provides the first systematic analysis of miRNAs in *Ae. aegypti* and offers a substantially expanded list of miRNAs for all mosquitoes. New insights were gained on the evolution of conserved and lineage-specific miRNAs in mosquitoes. The expression profiles of a few miRNAs suggest stage-specific functions or functions related to blood feeding.

3.2. Introduction

MicroRNAs (miRNAs) are approximately 22 nucleotide long non-coding RNAs that regulate the expression of cellular genes by binding to target mRNAs for cleavage or translational repression (Bushati and Cohen, 2007). Thousands of miRNAs have been reported in animals and plants (Griffiths-Jones et al., 2006). Many miRNAs are highly conserved across divergent species while others are specific to a particular evolutionary lineage (miRBase: http://microrna.sanger.ac.uk/). Lineage-specific miRNAs can arise from non-miRNA sequences, which has been observed in *Drosophila*, or by modifications in existing miRNAs present in an organism (Zhang et al, 2008). MiRNA genes can occur in intergenic regions and within introns (Miska, 2005; Kim et al., 2007; Liu et al, 2007). For some miRNAs, their biogenesis starts with transcription of miRNA genes, mostly by RNA polymerase II, which produces a primary miRNA (e.g., Bartel, 2004). A given primary miRNA can be either monocistronic, containing one mature miRNA, or polycistronic, containing multiple mature miRNA. In *Drosophila*, the primary miRNA is processed by a Drosha-Pasha complex to yield small stem-loop structures that are approximately 70 nucleotides in length called pre-miRNAs (Saito et al., 2005). Drosha cleavage leaves a 5' monophosphate and a 2nt 3' overhang at the end of the hairpin (Lee et al., 2003; Zhang et al, 2004). Following export to the cytosol, the double-stranded pre-miRNA is recognized by Dicer-1 (Saito et al, 2005), which complexes with Loqs and selectively cleaves a ~22 bp sequence with 2nt 3' overhangs from the pre-miRNA (Saito et al., 2005). The product is referred to as an miRNA:miRNA* duplex (Lau et al., 2001; Bartel, 2004; Miska et al., 2005). The duplex molecules are separated

by a helicase. The miRNA* strand, which is opposite to the miRNA strand, is rapidly degraded (Bartel, 2004). Recent "deep" sequencing methods revealed that approximately 100 fold as many miRNAs as miRNA*s were observed for a given miRNA except in cases where both strands could generate functional miRNAs (Ruby et al., 2006; Tyler et al., 2008; Liu et al, 2008). A new category of pre-miRNAs known as mirtrons are derived from introns and undergo Drosha-independent processing (Ruby JG et al, 2007; Okamura et al., 2007; Berezikov et al., 2007). Some researchers suggest that their presence indicates that RNA sources may be less relevant for determining which RNAs become miRNAs than their structural characteristics (Liu et al, 2008). Mature miRNAs pair with target mRNAs with the assistance of Argonaute to achieve translational repression and/or mRNA turnover (reviewed in Wu and Belasco, 2008).

MicroRNAs play key regulatory roles during embryonic development, stem cell division, cancer development, neurogenesis, heart development, haematopoietic cell differentiation, and cell death (reviewed in Bushati and Cohen, 2007). They are also implicated in control of viral infection in vertebrates and in one recent report miRNAs were linked to malaria infection in mosquitoes (Caldas and Brenton, 2005; Lecellier et al., 2005; Sullivan et al., 2005; Winter et al., 2007). Similarly, miRNAs are also linked to innate immunity (reviewed in Glinsky, 2008; Carissimi et al., 2009).

There are reports describing *Anopheles gambiae* miRNAs on the basis of similarity to known miRNAs from other species or computational prediction (Lai et al., 2003; Wang et al., 2005; Chatterjee and Chaudhury, 2006; Xu et al., 2008). Direct experimental

evidence for mosquito miRNAs has been recently reported only in *Anopheline* mosquitoes using small-scale sequencing (Winter et al., 2007; Mead and Tu, 2008).

In this report, we describe the discovery and characterization of 98 pre-miRNAs that produce 86 distinct miRNAs and 20 distinct miRNA* sequences in *Ae. aegypti* by using high throughput sequencing and bioinformatics approaches. The majority of the *Ae. aegypti* miRNAs have homologs in *D. melanogaster* or other related species. Eighty-nine of the sequences showed perfect matches to small RNA sequences. Expression profiles of 17 of the miRNAs were analyzed by northern blot in *Ae. aegypti*, confirming the results of miRNA discovery, and displaying stage-specific expression patterns. Thirteen miRNAs, including five previously reported in *Anopheline* mosquitoes, are only found in mosquitoes so far. Four of such miRNAs are lineage-specific within mosquitoes. Therefore this study provides the first systematic analysis of miRNAs in *Ae. aegypti* and offers a substantially expanded list of miRNAs for all mosquitoes. New insights are gained on the evolution of conserved and lineage-specific miRNAs in mosquitoes. The expression profiles of a few miRNAs suggest stage-specific functions or functions related to blood feeding.

3.3. Materials and methods

Insects

Ae. aegypti (Liverpool strain), *An. gambiae* (G3 strain), *An. stephensi* (Indian wild type strain), and *T. amboinensis* (CDC strain originally from San Juan, PR) mosquitoes were

reared in a humidified insectary at 27°C on a 12 hour light:dark cycle. *Culex quinquefasciatus* embryos at 0-24 hrs post-egg-deposition were kindly provided by Drs Aaron Brault and David Clark at the University of California, Davis. *Drosophila melanogaster* wildtype (Catalina strain, stock number 14021-0231.47) samples were provided by the Tucson Drosophila Stock Center (Tucson, AZ). *D. melanogaster* W1118 eggs at 0-24 hrs post-egg-deposition were provided by Duke University Model Systems (Duke University, Durham, NC).

Ae. aegypti sample preparation for 454 pyrosequencing

Three samples were prepared for 454 sequencing of small RNAs. Female *Ae. aegypti* were either fed on mice or kept on sugar water for three days post-emergence. Midguts were dissected at 24 hours post blood meal and midguts from sugar-fed females were also dissected at the same time interval. A third sample was mixed age embryos. All eggs were laid on filter paper during one-hour intervals. The filter paper was then kept in an incubator (27°C, ~70% Relative Humidity) before collection at appropriate hours and stored in RNAlater (Ambion, Austin, Texas). The collection design ensured that embryos between 0-48 hrs post-egg-deposition were collected in the same quantity to allow broad coverage of all embryonic stages. All eggs were then mixed. These samples were sent to vertis Biotechnologie AG (Freising-Weihenstephan, Germany) for small RNA cloning. All samples were ground in liquid nitrogen and RNA smaller than 200 bp were enriched with the mirVana miRNA isolation kit (Ambion). The population of miRNAs with a length of 15-30 bp was passively eluted from polyacrylamide gels. The RNA was then

precipitated with ethanol and dissolved in water. The small RNAs collected had a poly(A)-tail added to their 3'-OH by poly-(A) polymerase. The 5'-phosphate of the small RNAs were ligated to an RNA adapter. First-strand cDNA synthesis was then performed using an oligo(dT)-linker primer and MMLV-RNase H reverse transcriptase. The resulting cDNAs were PCR-amplified to about 20 μg/μl. Primers used for PCR amplification were designed for amplicon sequencing according to the instructions of 454 Life Sciences (Branford, CT). The PCR-amplified cDNAs were size-selected using electroelution to obtain product of 119-134 bp. These cDNAs were then sequenced by 454 Life Sciences. Approximately 40,000 cDNA sequences were obtained from each sample.

Identification of pre-miRNA sequences in Ae. aegypti

Sequences that match known mosquito and *D. melanogaster* pre-miRNAs (miRBase v.12.0, September 2008) were used to identify miRNAs in *Ae. aegypti*. Potential homologues were identified in the genome assemblies of *Ae. aegypti*, *An. gambiae* and *Cx. quinquefasciatus*. The homologous sequences plus 200 bp flanking sequences were retrieved from three mosquito genomes and aligned using Clustalx (Thompson et al., 1994) with a gap open penalty of five and a gap extension penalty of 0.05. The alignments were manually inspected and pre-miRNAs were identified and confirmed by RNAfold (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi) using default settings. Eighty-nine pre-miRNAs were identified in *Ae. aegypti* including pre-miRNAs for five previously reported miRNAs that have only been found in mosquitoes so far. Expanding

this homologous search approach by using all miRNA and miRNA* sequences in miRbase (September 2008 version) as queries did not yield additional pre-miRNAs in mosquitoes.

To identify novel miRNAs (or pre-miRNAs), we searched the *Ae. aegypti* 454 small RNA sequence library for sequences that matched both *Ae. aegypti* and *Cx. quinquefasciatus* genome assemblies by BLAST (e-value cut-off is 0.01). We focused on the comparison between *Ae. aegypti* and *Cx. quinquefasciatus* because these two species are more closely related to each other than to *An. gambiae*. Briefly, we first generated a new 454 small RNA sequence library (version 2) by removing sequences that match previously characterized miRNAs, *Ae. aegypti* transposable elements (TEfam: http://tefam.biochem.vt.edu), and known mosquito non-coding RNAs. We also masked all cDNA transcripts from *Cx. quinquefasciatus* genome assembly using RepeatMasker (http://www.repeatmasker.org/) on a Linux server. In this case, we were using cDNA sequences instead of repetitive sequences as the "repeat" library. We then identified sequences in the masked *Cx. quinquefasciatus* genome that matched version 2 of the 454 small RNA library by BLAST (e-value cut-off is 0.01). The *Cx. quinquefasciatus* genome sequence was then used to identify homologous sequences in the *Ae. aegypti* assembly. The homologous sequences were aligned and folded as described above. Nine novel pre-miRNAs were identified in *Ae. aegypti*. Efforts to identify novel miRNAs simply based on whole-genome comparison between *Ae. aegypti* and *Cx. quinquefasciatus* did not yield additional miRNAs, nor did efforts based on 454 small RNA sequences alone.

454 sequence count

Approximately 40,000 cDNA sequences were obtained from each sample. To determine the number of hits per sample, 98 *Ae. aegypti* pre-miRNAs identified above were used as query for BLAST analysis. We require 100% match in at least 18 bp for a sequence to be counted. This approach does not distinguish between paralogous pre-miRNAs, nor does it distinguish between miRNAs that share at least 18 bp perfect match.

Sample collection for northern blots

For *Ae. aegypti* midgut samples, the sample collection was the same as described for 454 sequencing. Sample collections for different developmental stages of *Ae. aegypti*, *An. stephensi*, and *An. gambiae* are briefly described below. Embryo collections were made at 0-12, 12-24, and 24-36 hours after placing a damp collection cup within a cage. To generate points after 12 hours, egg containers were set aside and allowed to incubate at 27°C in a damp collection cup. 0-36 hour samples represent equal mixed pools of 0-12, 12-24 and 24-36 hour samples. Larval samples were collected at each instar, and pooled to generate early (I and II instars) and late (III and IV instars) larvae as listed in each figure. Pupal samples were collected from a pool of varied ages. Adults one to five days following eclosion were collected. In some cases, we did not separate early and late larval samples and used one mixed larval sample instead. These variations are specified in the figure legends of the northern blots.

For *T. amboinensis*, embryos were collected at 0-24 hours post-oviposition. Samples were collected for 1st and 2nd instar larvae and pooled to generate an early larvae sample, and a separate collection of 4th instar larvae was collected for late larvae. Pupae were collected from a pool of varied ages to ensure a representative sample, and male and female adults were collected at two to five days post emergence.

Northern blot

All samples were either directly processed for RNA isolation or flash frozen on liquid nitrogen immediately following collection, then stored at –80°C. Total RNA isolation was carried out upon the samples using a mirVana miRNA isolation kit (Ambion, Austin, TX). The amount of total RNA used for each sample is specified in the relevant figure legend. Northern blots were carried out based upon Mead and Tu (2008). Briefly, total RNA were loaded onto 15% denaturing polyacrylamide gels, and run beside 19 and 23 nucleotide long ssDNA markers. The RNA gels were transferred to BrightStar-Plus nylon membranes (Ambion), crosslinked using a UV crosslinker, and prehybridized, then hybridized overnight in the ULTRAhyb-Oligo Hybridization Buffer (Ambion) with the appropriate DIG-labeled probe at 42°C. Wash conditions were the same as described in Mead and Tu (2008). Antisense 5' digoxigenin-labeled miRCURY LNA probes were purchased from Exiqon (Vedbaek, Denmark). Probe sequences were reverse complementary to sequences shown in Table 1. The probe for miR-1174 is derived from aga-miR-1174 (miRBase) and it had one base difference to aae-miR-1174.

3.4. Results

Discovery of 98 distinct pre-miRNA sequences in Ae. aegypti

As shown in Table 1, we have uncovered 98 different pre-miRNAs in *Ae. aegypti* which produce 86 distinct miRNAs and 20 distinct miRNA* sequences. Some of the 98 pre-miRNA sequences produce identical miRNAs and miRNA*s. Eighty-nine of the *Ae. aegypti* miRNA and miRNA* sequences showed a perfect match to small RNA sequences from at least one of the three samples. Twenty-nine of the 98 pre-miRNAs do not have small RNA sequences in the embryo and midgut samples. However, these 29 pre-miRNAs all form hairpins and are conserved in other mosquito species.

Possible modifications of known miRNAs, cases of abundant miRNA sequences, miRNA gene clusters and duplications*

As shown in Table 1, there are a few cases where the *Ae. aegypti* miRNA sequences, as indicated by direct sequencing, start or end with one or a few extra nucleotides compared to the known miRNAs reported from *D. melanogaster* or *Anopheline* mosquitoes (miRBase). To minimize the contribution of sequencing error, we only consider cases where there are at least six such sequences and these differences are the majority of the 454 sequences. Such cases either suggest a difference between species or could provide leads to further studies to investigate whether these previously reported miRNA sequences may need to be revised. These miRNAs include aae-miR-2a, aae-miR-210, aae-miR263b, aae-miR-281, and aae-miR-283. In vast majority of the cases, mature miRNAs are much more abundant than miRNA*. However, in the case of miR-281 and miR-1175, their miRNA* are dozens or more than a hundred fold more than the miRNA

sequences. There are a few other cases in which the miRNA* is more abundant than the miRNA sequences. However, the numbers of hits are low in these cases and it is difficult to assess how significant the differences may be. There are 14 clusters of miRNAs that are defined as more than one miRNA hairpin within 10 kb (Griffiths-Jones et al., 2006). Twelve of these clusters have members that are separated by less than one kb. There are 17 cases where one of the pre-miRNAs is duplicated and thus than one pre-miRNA hairpin produces the same or highly similar mature miRNAs.

Novel miRNAs that are only found in mosquitoes so far

Nine of the 98 pre-miRNA hairpins are novel and have only been found in mosquitoes so far. These nine pre-miRNAs produce seven mature miRNAs (miR-M1, -M2, -M3, -M4a and -M4b; miR-N1 and miR-N2). All seven mature miRNAs have multiple hits from small RNA sequencing, confirming their status as miRNAs. A few of these also have hits in the miRNA* strand. Shown in Figure 1 are the sequence alignments of the pre-miRNA sequences discovered in this study and the hairpins they form. Two physically linked pre-miRNA hairpins (miR-M1-5p and miR-M1-3p) produce the same miR-M1 in *Ae. aegypti*. Two physically linked pre-miRNA hairpins (miR-M4a and miR-M4b) produce similar but not identical miR-M4a and miR-M4b. MiR-M1, -M2, -M3, -M4a and -M4b are found in all three available mosquito genome assemblies.

Two physically linked pre-miRNA hairpins (miR-N1-5p and miR-N1-3p) produce the same miR-N1 and they are also in close proximity to the miR-N2 hairpin in *Ae. aegypti*.

The two miR-N1 and one miR-N2 hairpins are in the first intron of a gene in *Ae. aegypti* (Vectorbase Gene ID AAEL009263) encoding a putative transcription factor with a basic leucine zipper domain. Sequence analysis suggests that miR-N1 is found in the orthologous gene in *C. quinquefasciatus* but not found in *An. gambiae*. MiR-N2 is only found in *Ae. aegypti*. MiR-N1 exists in two hairpins in the intron of the homologous gene in *Cx. quinquefasciatus* and there is a third hairpin that has a predicted miRNA with a similar 5' sequence as miR-N1. We name this miRNA miR-N3 and it is only found in *Cx. quinquefasciatus* (Vectorbase Gene ID CPIJ000468). Cqu-miR-N3 is not listed in Table 1 that only shows miRNAs from *Ae. aegypti*.

Thus, we have identified eight novel mosquito-specific miRNAs in this study. We define "mosquito-specific" miRNAs here as those that are only found in mosquitoes so far. It is possible that species closely-related to mosquitoes may have some of these miRNAs. However, BLAST searches using low stringent parameters (word size at seven, e-value cut-off at 10 or above) failed to identify any reliable homologues from miRBase or non-redundant GenBank sequences. Taken together, the evidence strongly indicates that what we are reporting in this study are novel miRNAs. Thus this study increased the number of novel miRNAs that are only found in mosquitoes from five (Winter et al., 2007; Mead and Tu, 2008) to 13.

Expression profiles of conserved miRNAs during different developmental stages in Ae. aegypti

We chose nine conserved miRNAs and eight mosquito-specific miRNAs for further analysis using northern blot to confirm the small RNA sequencing results and to determine the expression profiles of these miRNAs in different developmental stages. Expression analysis of the eight mosquito-specific miRNAs is described in the context of a multi-species survey in a later section. The nine conserved miRNAs include let-7, miR-1, -133, -14, -184, -210, -9a, -970, and -998. All nine miRNAs showed signals at ~20 nt by northern during at least one of the developmental stages. Shown in Figure 2 are the expression profiles of five of the nine conserved miRNAs. The expression profiles of these miRNAs are similar to the profiles of homologous miRNAs in *D. melanogaster* (Semprere et al., 2003) and *An. stephensi* (Mead and Tu, 2008). One exception is that the level of miR-14 appears to be lower in the *Ae. aegypti* embryos than in those of *D. melanogaster* and *An. stephensi*.

Elevated levels of transcripts of two miRNAs after blood feeding in the midgut of Ae. aegypti

The numbers of small RNA sequences in the midgut samples from sugar-fed and blood-fed female *Ae. aegypti* may not be high enough for quantitative comparison. Nonetheless, we decided to compare the relative mature miRNA levels for miRNAs that showed more than 25 hits in at least one of the midgut samples using the total number of all miRNA hits to normalize the data (Ruby et al., 2007). As shown in the last column of Table 2, except for miR-989 and miR-281*, all miRNAs showed an increase after blood feeding. We then performed northern blots using miR-184 and miR-998 and both results

confirmed higher miRNA levels in bloodfed samples than in sugar-fed samples (Figure 3). We have previously analyzed the level of miR-989 in the midgut before and after blood feeding (Mead and Tu, 2008). The signal was too weak to confirm or rule out reduction of miR-989 after blood feeding.

Multi-species survey of eight mosquito-specific miRNAs revealed conserved and lineage-specific miRNAs

Previously five miRNAs were reported to be only found in mosquitoes. These are miR-1174, miR-1175 (Winter et al, 2007), miR-1889, -1890, and -1891 (Mead and Tu, 2008). MiR-1889 has some similarity to miR-304 but the difference is significant enough for miRbase to assign a unique name for it. As described above, we uncovered eight additional mosquito miRNAs, bringing the number of total "mosquito-specific" miRNAs to 13. We conducted a detailed multi-species expression analysis of eight of the 13 mosquito-specific miRNAs using northern blot. When appropriate, we examined the expression of these miRNAs across the lifestages of four mosquito species, *An. gambiae*, *An. stephensi*, *Ae. aegypti*, *and T. amboinensis*.

Four of the eight miRNAs (miR-M1, -1175, -1890, and -1891) are detected in all of the above four species (Figure 4). Furthermore, the expression patterns of these miRNAs are similar in the four species and expression is detected in multiple developmental stages in three of the four miRNAs. This is consistent with the observation that conserved miRNAs tend to be widely expressed (Ruby et al., 2007; Stark et al., 2007). Although it is difficult

to compare between northern blots due to the difference in relative signal intensity, there appears to be some variations in the expression patterns between species. For example, miR-1891 is not detected in the embryo of *Ae. aegypti* and *T. amboinensis* while it is found in the embryo of the two *Anopheline* species. MiR-M1 is expressed only in the embryos in all four species.

Four other miRNAs (miR-1174, miR-N1, miR-N2, and miR-N3) are only detected in a subset of the four mosquitoes. As shown in Figure 5, miR-1174 is not found in *T. amboinensis* but strong signals were detected in the other three species. It is interesting to point out that miR-1174 and miR-1175 are in the same contig separated by only ~200 bp. The expression profile of miR-1174 and miR-1175 are similar in all three blood-feeding mosquitoes, suggesting that they may be under the same transcriptional control. MiR-1174 and miR-1175 share some sequence similarity at the 5' end. Thus it is possible that miR-1174 and miR-1175 resulted from gene duplication and miR-1174 may either have been lost in *T. amboinensis* or evolved beyond recognition by the miR-1174 probe. *Ae. aegypti* miR-1174 and *An. gambiae* miR-1174 differ by one nt. It is also possible that miR-1174 simply was not duplicated in *T. amboinensis*. As described earlier, miR-N1, N2, and N3 are from the same intronic cluster. As shown in Figure 6, miR-N1 was abundant in both *Ae. aegypti* and *Cx. quinquefasciatus* embryos. It was undetectable in *An. stephensi*. MiR-N2 was abundant in *Ae. aegypti* embryos, but undetectable in the embryos of *Cx. quinquefasciatus*. MiR-N2 was also undetectable in *An. stephensi*. MiR-N3 was found in *Cx. quinquefasciatus* embryos, but not in *Ae. aegypti*. MiR-N3 was also undetectable in *An. stephensi*.

We performed northern blots using all of the above eight miRNA probes to see if any signal was detected in *D. melanogaster*. We used 5 μg of total RNA from different developmental stages or a specific stage expected for a particular miRNA. None of the eight probes produced any miRNA signal while the positive control (*Ae. aegypti* sample) showed intense signals (data not shown). An artifact band was observed at ~40nt even under stringent conditions by northern for miR-N3 in Drosophila, though this was too large to be a valid miRNA. This is consistent with these miRNAs being only found in mosquitoes.

Zygotic transcription and expression patterns of miR-M1 in mosquito embryos

We further examined the expression of miR-M1 in the three hematophagous mosquito species in different embryonic stages and in bloodfed females. The lack of expression in 24 and 48 hr post-blood-meal females suggests that miR-M1 in the embryos is transcribed zygotically. Expression fluctuates over time within the developing embryo, and differences are observed between *Aedine* embryos and *Anopheline* embryos. In both *Anopheline* species, miR-M1 expression is strong at 0-12 hrs, and increases slightly by 12-24 hrs. It drops significantly by 24-36 hrs, particularly in *An. stephensi*. For *Aedes*, expression was weak in 0-12 hrs, but very strong at 12-24, and 24-36 hours. *Aedine* egg development proceeds at a slower rate, so it is possible that an *Aedine* and *Anopheline* egg at the same level of embryonic development may have comparable expression of miR-M1. As miR-M1 does not extend into the larval stage, it likely tapers off by 42 hours

in *Anopheles,* when the larvae emerge from the egg cases. For *Aedes*, emergence does not occur for at least several more hours.

3.5. Discussion

This study provides the first systematic analysis of miRNAs in *Ae. aegypti* and offers a substantially expanded list of miRNAs for all mosquitoes. We also provide experimental evidence for 89 of the miRNA and miRNA* sequences and revealed six cases in which revision of previously annotated miRNAs may be necessary. We also identified two cases (miR-281 and miR-1175) in which the previously reported miRNA* sequences are at least dozens fold more abundant than the miRNA sequences. In both cases, the miRNA and the miRNA* sequences are 100% identical among *Ae. aegypti*, *Cx. quinquiefasciatus*, and *An. gambiae*. It is therefore possible that miR-281* and miR-1175* are functional. There are 14 clusters of miRNAs and there are 17 cases where more than one pre-miRNA hairpin produces the same or highly similar mature miRNAs in *Ae. aegypti*. These miRNAs are a rich source for future comparative analysis to uncover the evolutionary patterns of miRNA duplication and biogenesis in mosquitoes. Perhaps most importantly, we identified eight novel miRNAs that are only found in mosquitoes so far. This discovery expanded the list of mosquito-specific miRNAs from five (Winter et al., 2007; Mead and Tu, 2008) to 13. A better understanding of the functions of these mosquito-specific miRNAs will undoubtedly offer novel insights in mosquito biology.

A number of miRNAs showed higher levels in midgut from blood-fed females than that from sugar-fed females, which was confirmed by northern blots on two of these miRNAs, aae-miR-184 and aae-miR-998. There is only one case of lower expression levels in blood-fed samples compared to sugar-fed samples, which is aae-miR-989. The predominantly elevated expression in the midgut after blood feeding is interesting and understanding the function and regulation of these miRNAs may help illuminate mosquito blood feeding physiology. As the annotation of the 3'-UTRs of mosquito genes improves, miRNA target analysis will become more fruitful and may suggest the functions of these miRNAs in the midgut. The expression profiles of a few miRNAs suggest stage-specific functions. One such example is miR-M1, which is only expressed in embryos. MiR-M1 is zygotically transcribed and it is present in most of the embryonic stages. The embryonic expression pattern of miR-M1 in *Aedine and Anopheline* mosquitoes is similar if we take into account the difference in the development time of the two genera. Thus the conserved expression pattern and sequence conservation across all major branches of Culicidae suggest that miR-M1 is important during mosquito embryogenesis.

Another interesting group of miRNAs are the miR-N1, -N2, and -N3 cluster. Two miR-N1 and one miR-N2 hairpins are in the first intron of a gene in *Ae. aegypti* encoding a putative transcription factor. Two miR-N1 hairpins and a miR-N3 hairpin are found in the orthologous gene in *Cx. quinquefasciatus*. None of these miRNAs are found in *An. gambiae*. In addition, miR-N2 is only found in *Ae. aegypti* while miR-N3 exists only in *Cx. quinquefasciatus*. These miRNAs share the same seven to eight bp 5' sequences in

the seed regions important for target recognition. MiR-N1, N2, and N3 are all expressed in the embryos. Thus it is possible that these miRNAs derive from duplication events and the duplicated miRNAs may evolve into new sequences that acquire new functions. We postulate that, given their abundance and their displayed lineage specificity, the N1/N2/N3 cluster may play a role in determining important lineage specific traits between Culicinae and other mosquitoes, and even divergence within Culicinae itself. It will be interesting to determine what the targets of these miRNAs are and how their expression may be coordinated with the transcription factor in which they reside.

Taken together, this study provides the first systematic analysis of miRNAs in *Ae. aegypti* and offers a substantially expanded list of miRNAs for all mosquitoes, including eight novel mosquito-specific miRNAs. New insights are gained on the evolution of conserved and lineage-specific miRNAs in mosquitoes. The expression profiles of a few miRNAs suggest stage-specific functions or functions related to blood feeding. A better understanding of the functions of these miRNAs will offer novel insights in mosquito biology and may lead to novel approaches to combat mosquito-borne infectious diseases.

3.6. Acknowledgements

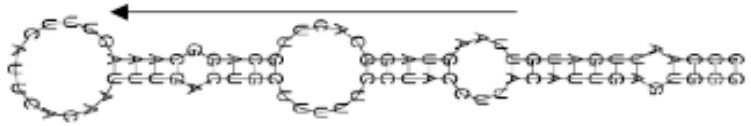
We thank Randy Saunders for mosquito-rearing and for help with sample preparations. This work is supported by NIH grant AI070854, FNIH grant GC7 #316, and the Virginia Experimental Station.

Author contributions

Song Li initiated the work, prepared samples for 454 sequencing, conducted the first northern blot using DIG-labeled probes, performed bioinformatics, provided relevant tables, and wrote a draft of part of the materials and methods section. Shaohui Liang performed northern analysis on all conserved miRNAs, performed a comparison of miRNA expression in midgut samples, and provided relevant figures. E. A. Mead performed northern analysis on all mosquito-specific miRNAs, developed a lab protocol for northern blots, worked on the initial bioinformatics analysis on the miR-N1, N2, and N3 cluster, provided relevant figures, assisted in sample dissections, wrote a draft of several parts of this manuscript and assisted in revisions. Zhijian Tu designed and oversaw the project, performed part of the bioinformatics analysis, and wrote most of the manuscript.

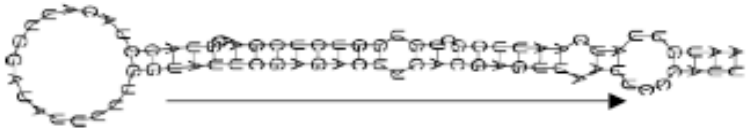
Figure 3.1.A.

aae-miR-M1-5p



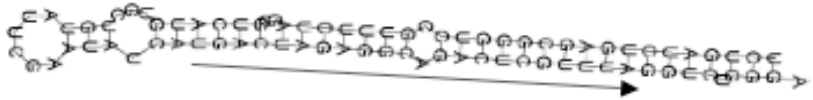
```
aae-miR-M1-5p   CCCAATTGATGTTAAGTAGGCACTTGCAGGCAAAG----TTTCATTCACAAT
cqu-miR-M1a     CCTAAGCGATGTTAAGTAGGCACTTGCAGGCAAAGTTTCTTCAGTTT-CGCT
aae-miR-M1-3p   CCTGATCGATGTTAAGTAGGCACTTGCAGGCAAA---TTTGTAGCTT-TAAT
cqu-miR-M1b     CCTGATCGATGTTAAGTAGGCATTTGCAGGCAAAA----CGTTGCTTATAAT
                **     ********************** **********          *     *

aae-miR-M1-5p   ----TTGACTGCTTTTTCCTACCTTACATTGAGTG-GG
cqu-miR-M1a     ACCGTTGACTGCTCTTTCCTACTTTACATTAAAA--AG
aae-miR-M1-3p   ATGATTGACTGCTTTTTCCTACTTTACATTAAAA--AG
cqu-miR-M1b     TTGGTTGACTGCTCTTTCCTACTTTACATTGAGA--AG
                   ********* ******** ******* *     *
```

aae-miR-M2



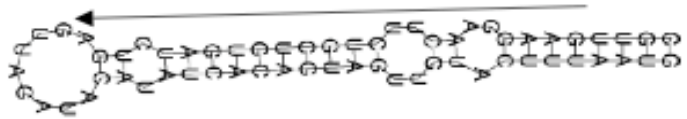
```
aae-miR-M2    AATGTTATCAATTCGCTGTGGTCTCGAACGTACCTACATTGGATATTTT
cqu-miR-M2    AACTTTATCAATTCGCTGTGGTCTCGAACGTACCTACGTTGGATATTTT
              **  ********************************** ***********

aae-miR-M2    TGGTATTCGAGACTTCACGAGTTAATTCCCATT
cqu-miR-M2    TGGTATTCGAGACTTCACGAGTTAATTCCCGTT
              ******************************** **
```

aae-miR-M3



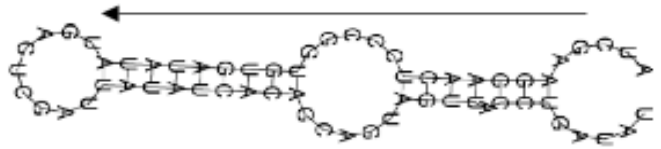
```
aae-miR-M3    TCTGATCTGAGCGGGTCCGTTTCTAGTGTCATGTGCTGTA-TTCGAA
cqu-miR-M3    CC-GACCCGAGCGGGTCCGTTTCTAGTATCATGTGC-GTCCTTCGAA
               * ** * ******************** ******** **   ******

aae-miR-M3    TATCATGACTAGAGGCAGACTCGTTTAGGTCTGGGA
Cqu-miR-M3    AGTCATGACTAGAGGCAGACTCGTTTAGGTCTTTAA
               ********************************   *
```

aae-miR-M4a



```
aae-miR-M4a   CGTTGAAGGAACTTCTGCTGTGATCTGAGTTAGATACATATCACAGTAGTTG
cqu-miR-M4a   TATTGAAGGAACTTCTGCTGTGATCTGAGTTGGAACCATATCACAGTAGTTG
               ****************************** **  ****************

aae-miR-M4a   TACTTTAATG
cqu-miR-M4a   TACTTTAATG
              **********
```

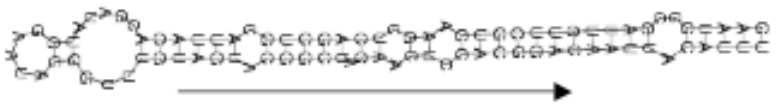
aae-miR-M4b



```
aae-miR-M4b   ATCGAAGGAACTCCCGGTGTGATATATGACTCGATTATATCACAGCAGTAG
cqu-miR-M4b   GTAGAAGGAACTCCCGGTGTGATATATGATTCGACTATATCACAGCAGTAG
               * ************************** **** ****************

aae-miR-M4b   TTACCTGATAT
cqu-miR-M4b   TTACCTGATAC
              **********
```

Figure 3.1.B.

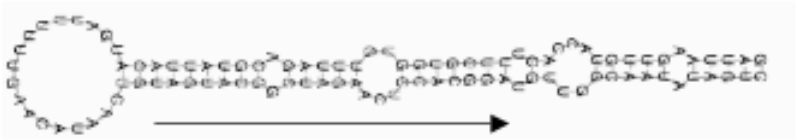
aae-miR-N1-5p



aae-miR-N1-5p

```
Aedes    GAATGGCGATTGTTCG-TGAAGCTCAGCTGCATTACA----CGATATGG
Culex    AGATGTTGAAT-TTCGCTGGAGTTCTGCCGGATTGCATTGACGAT-TGG
           ***  ** * **** ** ** ** ** * *** **    **** ***

Aedes    AATACGGTTTGTAGTACGGCTAGAACTCCACGGACAATT-ACATTT
Culex    AA----ATTTGTAGTACGGCTAGAACTCCACGGACATTCGACATTT
         **     ******************************* *  ******
```

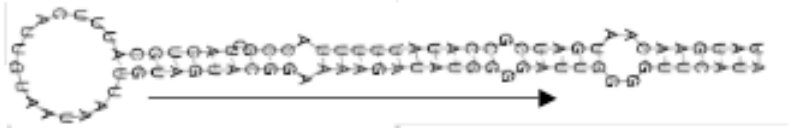
aae-miR-N1-3p



aae-miR-N1-3p

```
Aedes    GATT-AAGTTGTACCACTTTCGTGGTGTTTAGACGTATTACATGATTTTTTGAACA
Culex    CTTTCGAGGTG-ATCGTACCCGTGGTGTTTAGTCGTAGTGCATTGGTTTTTAGA--
           **  ** ** * *     ************ **** * ***   *****  *

Aedes    TAACTGTAGTACGGCTAGAACTCCACGGATGTTGGCAA-TATAGTC
Culex    --ACTGTAGTACGGCTAGAACTCCACGGATATTGGCGCCTAATGCT
           ****************************** *****   **  *
```

aae-miR-N2



cqu-miR-N3

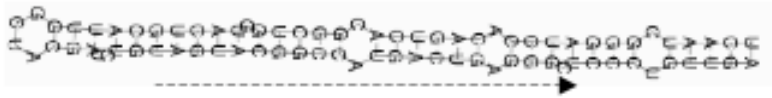


Figure 3.1. Alignments and structures of several novel mosquito pre-miRNAs. See Table 1 for naming and sequence locations of these miRNAs. A). Five novel pre-miRNA hairpins. Left side panels are the hairpin structures. Right panel are the sequence alignments between aae-miRNAs and cqu-miRNAs. In the case of miR-M1, there are two physically linked copies in both *Ae. aegypti* and *Cu. quiquefasciatus*, as shown in the alignment and in Table 1. The two copies in *Ae. aegypti* produce the same mature miRNAs and the hairpins are named -5p and -3p, reflecting their relative positions. The two copies in *Cu. quiquefasciatus* are named -1a and -1b because their mature miRNAs differ by one nucleotide. Only the hairpin for aae-miR-M1-5p is shown. All five miRNAs shown in panel A have homologs in *An. gambiae*. B). Pre-miRNA hairpins within the intron of a gene encoding a transcription factor. There are two hairpins for the same miR-N1 (-5p and -3p) in both *Ae. aegypti* and *Cu. quinquefasciatus*. Only the *Ae. aegypti* hairpin structures are shown. Aae-miR-N2 and cqu-miR-N3 are only found in *Ae. aegypti* and *Cu. quinquefasciatus*, respectively. Arrows point to the mature miRNA sequences from 5' to 3'. Dashed arrow for cqu-miR-N3 reflects the fact that we do not yet have the direct sequence for this miRNA. The mature cqu-miR-N3 sequence was predicted according to the conserved seed sequence shared with miR-N1 and miR-N2 and was confirmed by northern blots using anti cqu-miR-N3 as probe.

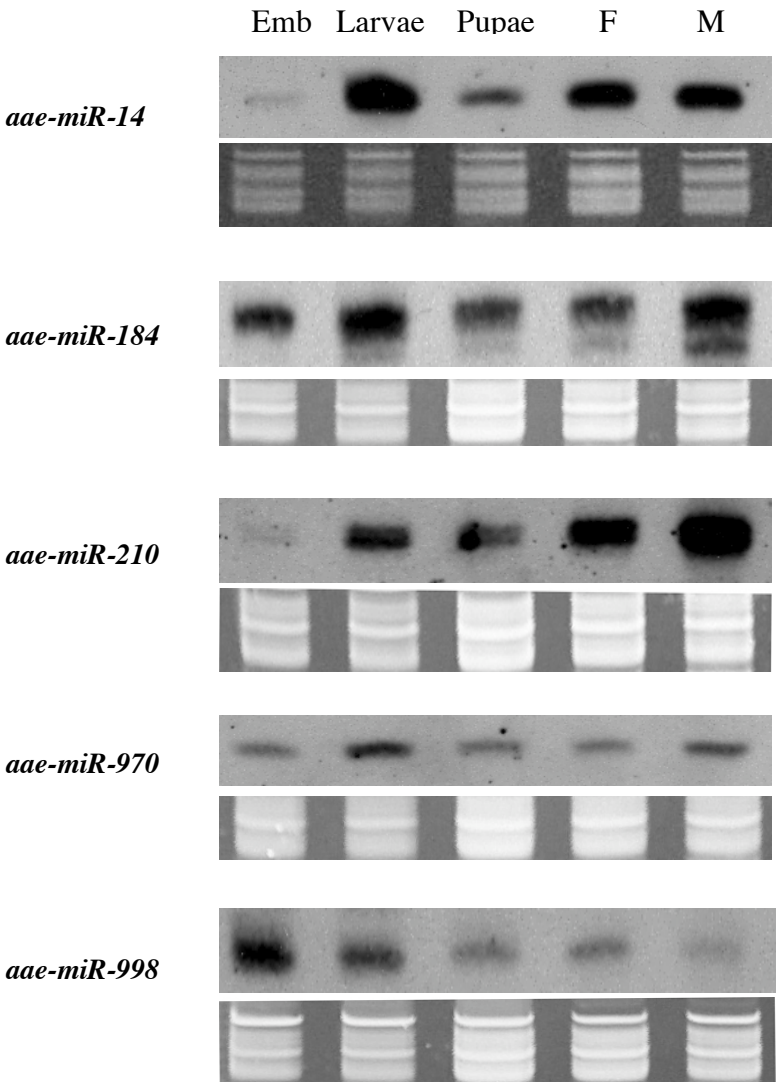


Figure 3.2. Expression profiles of *Ae. aegypti* homologs of previously known miRNAs. Only *Ae. aegypti* RNA samples were examined. Emb, pooled embryos between 0-36 hr after egg deposition; Larvae; mixed instar larvae; Pupae, mixed puape; F, adult females one to five days after emergence; M, adult males one to five days after emergence. 10 μg of total RNA per sample were used.

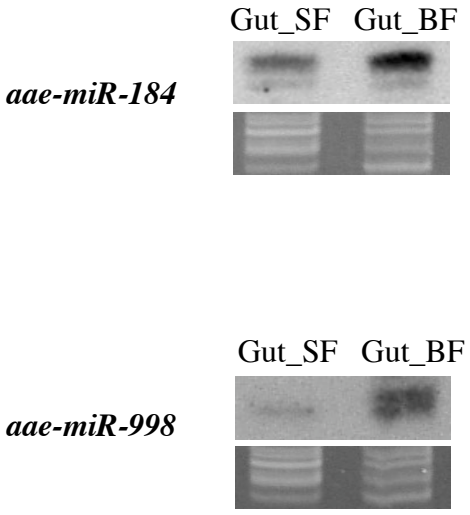


Figure 3.3. Higher levels of miRNAs are observed in the female *Ae. aegypti* midgut 24 hrs after blood feeding (Gut_BF) compared to sugar feeding (Gut_SF). Three-day old females were either fed on blood or sugar and dissected 24 hrs later. 10 μg of total RNA per sample were used.

Figure 3.4.A.

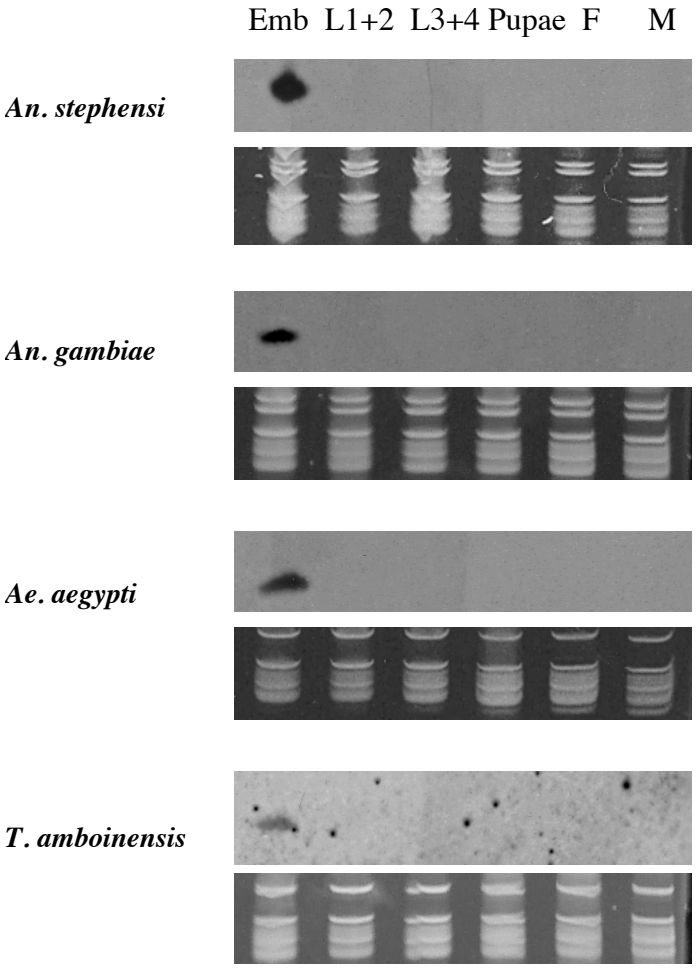

Emb L1+2 L3+4 Pupae F M
An. stephensi
An. gambiae
Ae. aegypti
T. amboinensis

Figure 3.4.B.

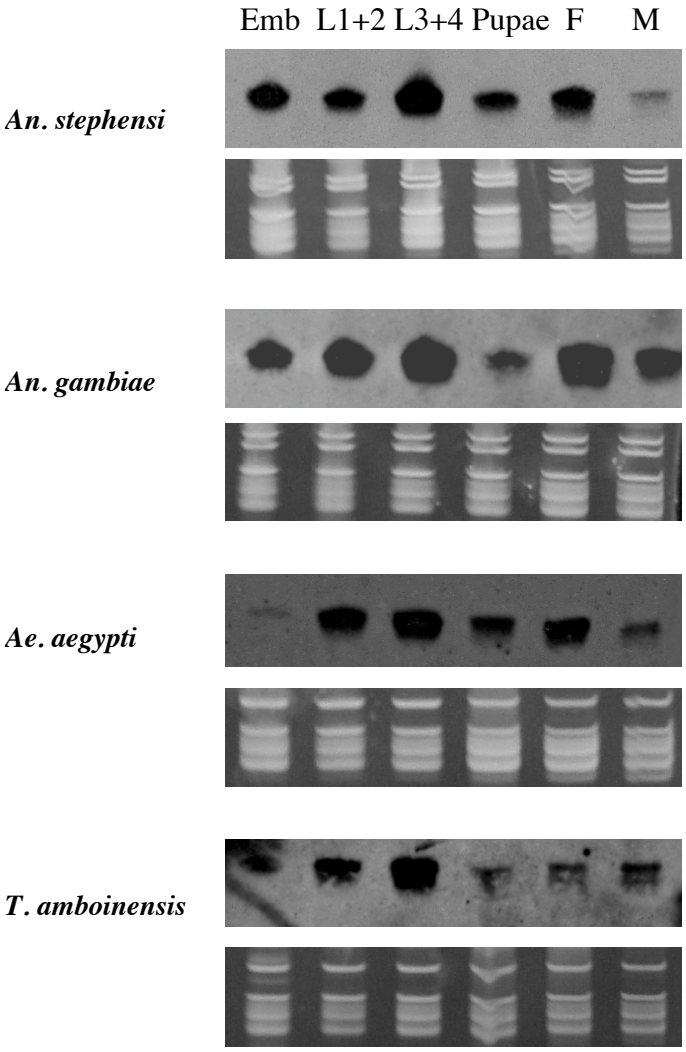

Emb L1+2 L3+4 Pupae F M
An. stephensi
An. gambiae
Ae. aegypti
T. amboinensis

Figure 3.4.C.

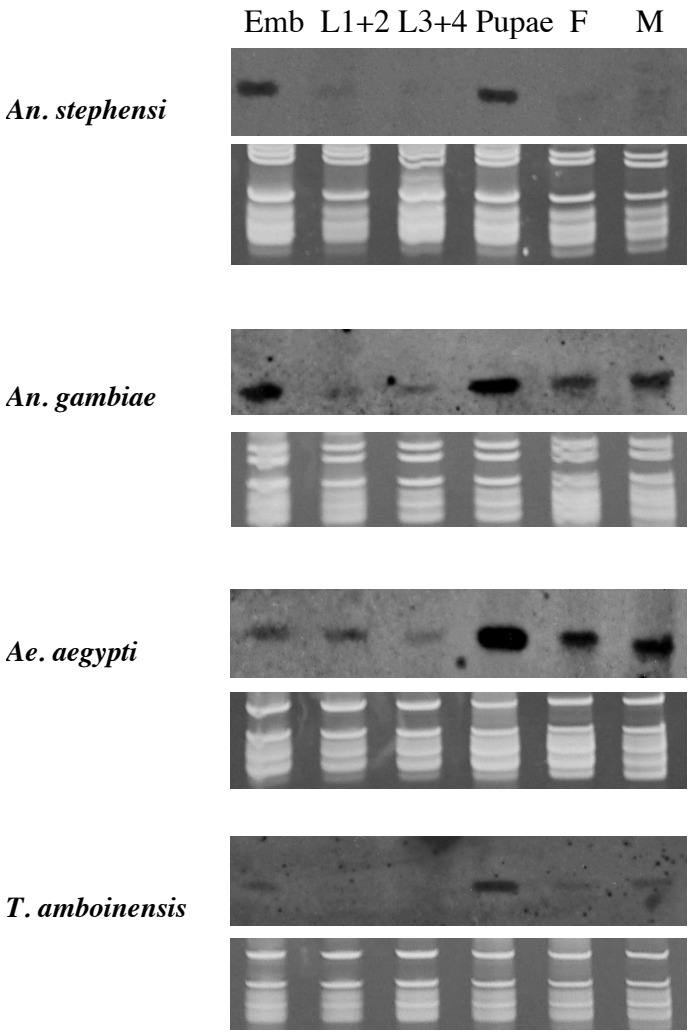

Emb L1+2 L3+4 Pupae F M
An. stephensi
An. gambiae
Ae. aegypti
T. amboinensis

Figure 3.4.D.

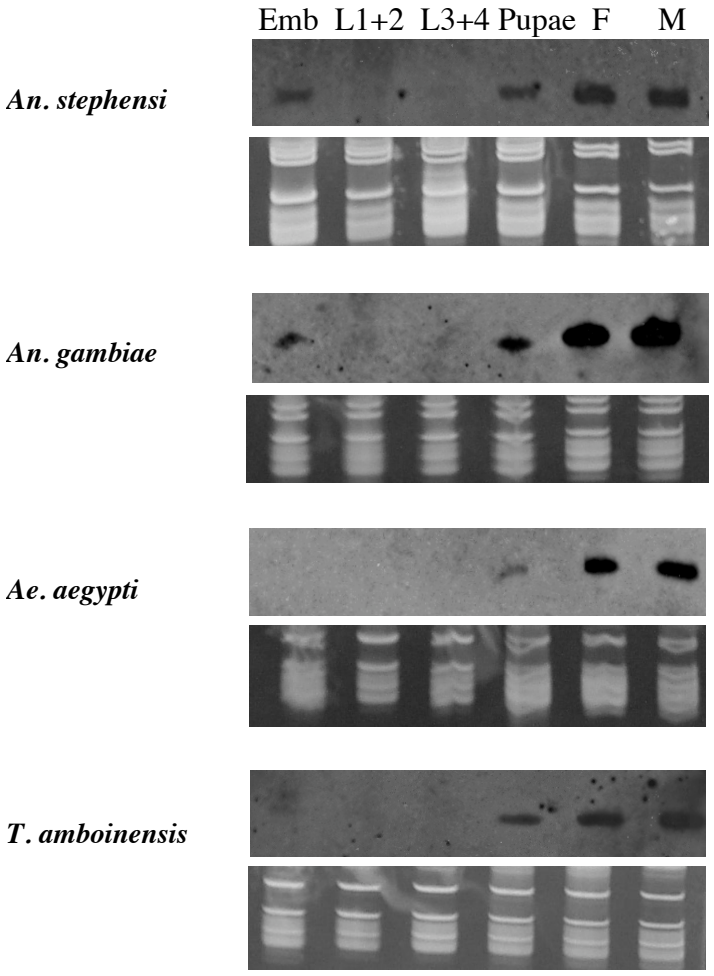


Figure 3.4. Four mosquito-specific miRNAs that are expressed in all four species of three highly divergent genera. MiRNAs examined include miR-M1 (A), miR-1175 (B), miR-1890 (C), and miR-1891 (D). Expression was examined across the developmental stages of *An. stephensi, An. gambiae, Ae. aegypti,* and *T. amboinensis*. Emb, pooled embryos between 0-36 hr after egg deposition; L1+2, pooled 1st and 2nd instar larvae; L3+4, pooled 3rd and 4th instar larvae; Pupae, mixed pupae; F, adult females one to five

days after emergence; M, adult males 1-5 days after emergence. 15 μg of total RNA per sample for *An. stephensi, An. gambiae*, and *Ae. aegypti* were used. 10 μg of *T. amboinensis* total RNA per sample were used. For *T. amboinensis* northerns, 3rd instar larvae were not included. For the *T. amboinensis* miR-M1 northern, hybridization and washes were carried out at 49°C instead of 42°C to reduce background across the membrane.

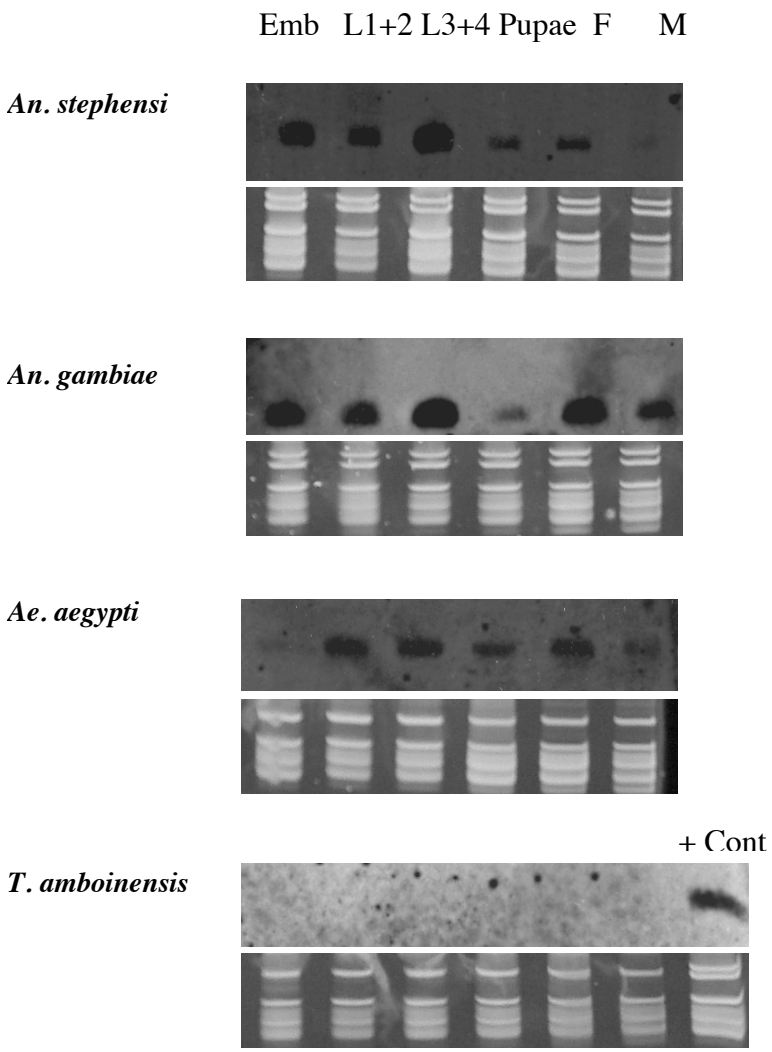


Figure 3.5. MiR-1174 is expressed in *An. stephensi, An. gambiae,* and *Ae. aegypti,* but not in *T. amboinensis*. Emb, pooled embryos between 0-36 hr after egg deposition; L1+2, pooled 1st and 2nd instar larvae; L3+4, pooled 3rd and 4th instar larvae; Pupae, mixed puape; F, adult females one to five days after emergence; M, adult males one to five days after emergence. 15 μg of total RNA per sample for *An. stephensi, An. gambiae*, and *Ae. aegypti* were used. 10 μg of *T. amboinensis* total RNA per sample were used. For *T. amboinensis*, 3rd instar larvae were not included. "+ Cont" indicates a positive control for the *T. amboinensis* northern, *An. stephensi* embryos (12-24 hr).

Figure 3.6.A.

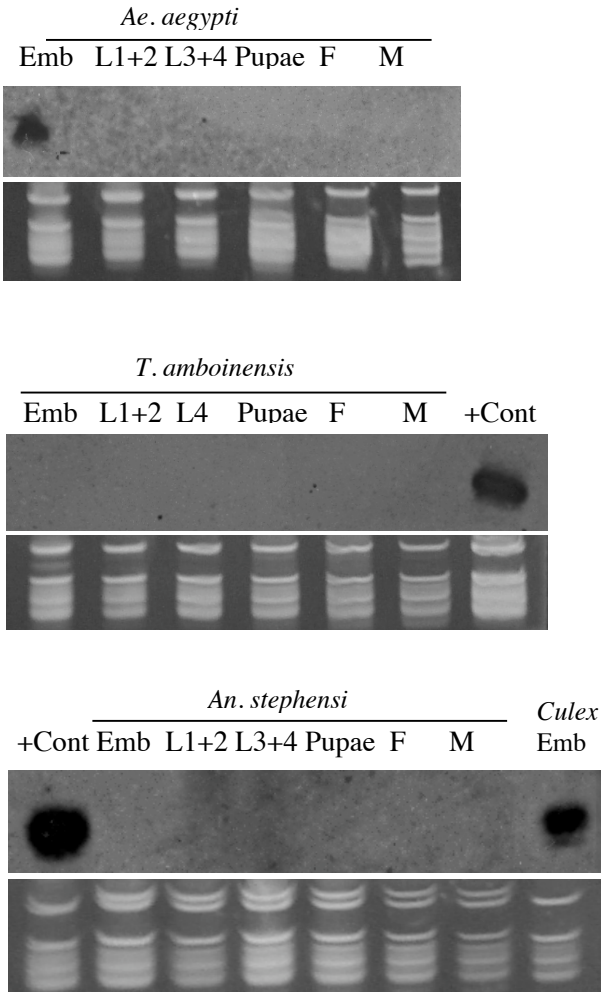

Ae. aegypti
Emb L1+2 L3+4 Pupae F M
T. amboinensis
Emb L1+2 L4 Pupae F M +Cont
An. stephensi
+Cont Emb L1+2 L3+4 Pupae F M
Culex
Emb

Figure 3.6.B. **Figure 3.6.C.**

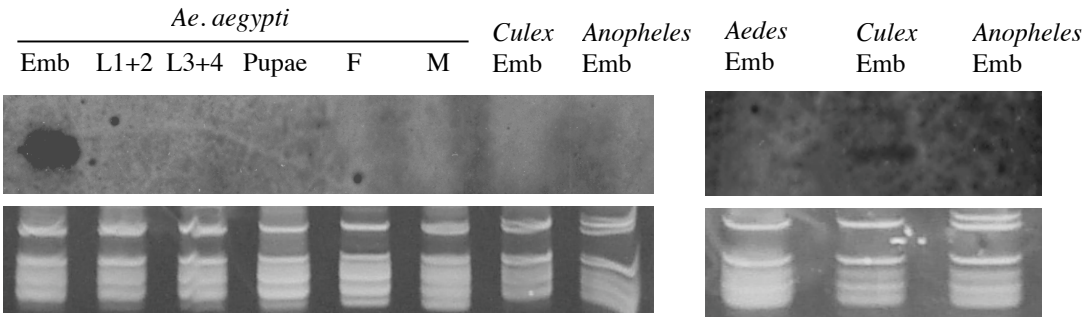


Figure 3.6. MiR-N1, miR-N2, and miR-N3 expression is restricted in particular lineages in mosquitoes. A) miR-N1 are expressed in *Culicinae* mosquitoes (*Ae. aegypti* and *Cx. quinquefasciatus*), but not in *An. stephensi* nor *T. amboinensis*. Emb, pooled embryos between 0-36 hr after egg deposition; L1+2, pooled 1st and 2nd instar larvae; L3+4, pooled 3rd and 4th instar larvae; Pupae, mixed pupae; F, adult females one to five days after emergence; M, adult males one to five days after emergence. *Culex* Emb, *Cx. quinquefasciatus* embryos 0-24 hrs after egg deposition. "+ Cont", positive control, *Ae. aegypti* embryos (12-24 hr). 15 μg of total RNA per sample for *An. stephensi* and *Ae. aegypti* were used. 10 μg of *T. amboinensis* total RNA per sample were used. For *T. amboinensis*, 3rd instar larvae were not included. B) miR-N2 is expressed in *Ae. aegypti* but not detected in *Cx. quinquefasciatus* and *An. stephensi* embryos. Symbols are as in A. *Anopheles* Emb, pooled *An. stephensi* embryos between 0-36 hr after egg deposition. C) miR-N3 is expressed in *Cx. quinquefasciatus,* but not detected in *Ae. aegypti* and *An. stephensi* embryos. Symbols are as in A and B. *Aedes* Emb, pooled *Aedes aegypti* embryos between 0-36 hr after egg deposition.

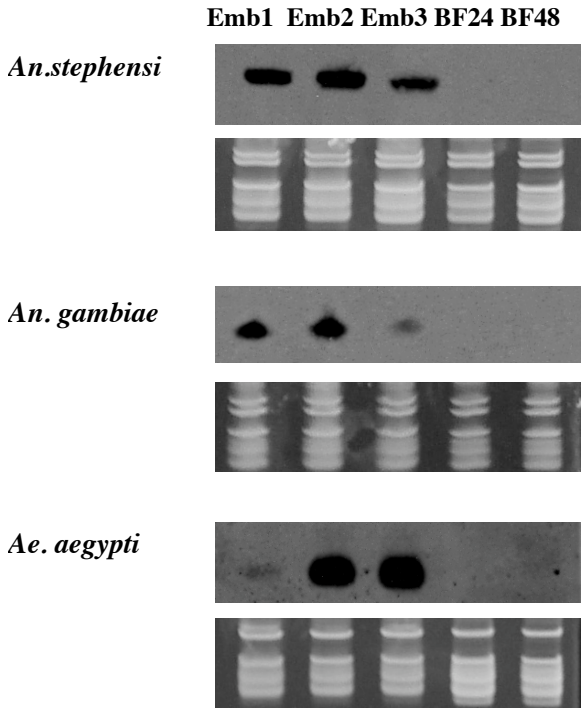


Figure 3.7. MiR-M1 is zygotically transcribed and abundant in the embryos. Emb1, pooled embryos between 0-12 hr after egg deposition; Emb2, pooled embryos between 12-24 hr after egg deposition; Emb3, pooled embryos between 24-36 hr after egg deposition; BF24, bloodfed females 24 hrs after a blood meal; BF48, bloodfed females 48 hrs after a blood meal. Females were fed on blood at five days old, and collections were made at 24 and 48 hours after the blood meal. 15 μg of total RNA per sample were used. Lack of miR-M1 in bloodfed females at 48 hrs indicates zygotic transcription of miR-M1.

Name [1,2]	Sequence [3]	Contig [5]	Start [5]	End [5]	Strand	Emb [6]	SF [6]	BF [6]
aae-miR-M1-5p	TTAAGTAGGCACTTGCAGGCAAA	CONTIG_15246	25749	25829	pos	28	0	0
aae-miR-M1-3p	TTAAGTAGGCACTTGCAGGCAAA	CONTIG_15246	25911	25981	pos	28	0	0
aae-miR-M2	TATTCGAGACTTCACGAGTTAAT	CONTIG_11944	151669	151750	pos	0	0	7
aae-miR-M3	TGACTAGAGGCAGACTCGTTTA	CONTIG_2929	123975	124056	pos	33	0	6
aae-miR-M3_star	AGCGGGTCCGTTTCTAGTGTCATG	CONTIG_2929	123975	124056	pos	3	0	0
aae-miR-M4a	GAAGGAACTTCTGCTGTGATCTGA	CONTIG_18252	93868	93929	pos	91	2	0
aae-miR-M4a_star	TATCACAGTAGTTGTACTTTAA	CONTIG_18252	93868	93929	pos	3	0	0
aae-miR-M4b	GAAGGAACTCCCGGTGTGATATA	CONTIG_18252	93731	93792	pos	62	0	0
aae-miR-M4b_star	TATCACAGCAGTAGTTACCTGA	CONTIG_18252	93731	93792	pos	24	0	0
aae-miR-N1-5p	TAGTACGGCTAGAACTCCACGGA	CONTIG_16241	318005	318105	neg	288	1	1
aae-miR-N1-3p	TAGTACGGCTAGAACTCCACGGA	CONTIG_16241	317700	317799	neg	288	1	1
aae-miR-N2	TAGTACGGAAAAGATATGGGGA	CONTIG_16241	318135	318222	neg	5	0	0
aae-miR-1174	TCAGATCTACTTAATACCCAT	CONTIG_7232	17531	17665	pos			
aae-miR-1175	TGAGATTCTACTTCTCCGACTTAA	CONTIG_7232	17737	17816	pos	0	0	1
aae-miR-1175_star	TAAGTGGAGTAGTGGTCTCATCGCT	CONTIG_7232	17737	17816	pos	0	2	59
aae-miR-1890	TGAAATCTTTGATTAGGTCTGG	CONTIG_10435	60786	60931	pos	1	0	0
aae-miR-1891-1	TGAGGAGTTAATTTGCGTGTTT	CONTIG_10244	3984	4069	neg	2	0	0
aae-miR-1891-2	TGAGGAGTTAATTTGCGTGTTT	CONTIG_18287	4656	4741	pos	2	0	0
aae-miR-1889	CACGTTACAGATTGGGGTTTCC	CONTIG_4323	23786	23921	neg	0	0	3
aae-bantam	TGAGATCATTTTGAAAGCTGATT	CONTIG_3301	208288	208426	neg	1	2	2
aae-let-7	TGAGGTAGTTGGTTGTATAGT	CONTIG_2929	321614	321684	pos	0	0	5
aae-miR-1	TGGAATGTAAAGAAGTATGGAG	CONTIG_25061	66256	66338	pos	1	0	1
aae-miR-10	ACCCTGTAGATCCGAATTTGTT	CONTIG_17639	8749	8833	pos	1	0	0
aae-miR-100	AACCCGTAGATCCGAACTTGTG	CONTIG_2929	307441	307568	pos			
aae-miR-1000-1	ATATTGTCCTGTCACAGCAGT	CONTIG_35246	172	265	neg			
aae-miR-1000-2	ATATTGTCCTGTCACAGCAGT	CONTIG_9774	12985	13078	pos			
aae-miR-11	CATCACAGTCTGAGTTCTTGCTT	CONTIG_23911	46399	46488	pos	1861	42	436
aae-miR-11_star	CGAGAACTCCGGCTGTGACCTGTG	CONTIG_23911	46399	46488	pos	4	0	2
aae-miR-12	TGAGTATTACATCAGGTACTGGT	CONTIG_4323	23386	23503	neg			
aae-miR-124	TAAGGCACGCGGTGAATGCCAAG	CONTIG_531	21343	21423	neg	3	0	0
aae-miR-125	TCCCTGAGACCCTAACTTGTGAC	CONTIG_2929	321885	321975	pos	0	0	3
aae-miR-133	TTGGTCCCCTTCAACCAGCTGT	CONTIG_24522	50243	50347	neg			
aae-miR-137-1	TTATTGCTTGAGAATACACGTA	CONTIG_7773	15414	15486	neg			
aae-miR-137-2	TTATTGCTTGAGAATACACGTA	CONTIG_29705	92457	92529	pos			
aae-miR-13b	TATCACAGCCATTTTGACGAGTT	CONTIG_12793	17619	17710	neg	168	2	42
aae-miR-14	TCAGTCTTTTTCTCTCTCCTAT	CONTIG_12112	693	785	neg	1	1	6
aae-miR-184	TGGACGGAGAACTGATAAGGGC	CONTIG_19030	3452	3535	neg	957	76	1307
aae-miR-190	AGATATGTTTGATATTCTTGGTTG	CONTIG_10108	2092	2180	neg	83	2	43
aae-miR-193	TACTGGCCTACTAAGTCCCAAC	CONTIG_19568	51035	51158	pos			
aae-miR-2a	TATCACAGCCAGCTTTGATGAGCT	CONTIG_12793	16130	16216	neg	1217	80	249
aae-miR-210	TTGTGCGTGTGACAACGGCTAT	CONTIG_19443	85221	85292	neg	3	12	12
aae-miR-219	TGATTGTCCAAACGCAATTCTTG	CONTIG_15776	56993	57078	pos			
aae-miR-2b	TATCACAGCCAGCTTTGAAGAGCG	CONTIG_12793	17471	17562	neg	758	74	147
aae-miR-2c	TATCACAGCCAGCTTTGATGAGC	CONTIG_12793	17971	18048	neg	1218	80	249
aae-miR-252	CTAAGTACTAGTGCCGCAGGAGA	CONTIG_3685	79002	79177	neg	38	0	1
aae-miR-252_star	CCTGCTGCCCAAGTGCTTATCGAA	CONTIG_3685	79002	79177	neg	5	0	0
aae-miR-263	AATGGCACTGGAAGAATTCACGGG	CONTIG_27430	4746	4821	neg	53	1	1

aae-miR-263_star	CGTGTTCTGGCAGTGGCATCCC	CONTIG_27430	4746	4821	neg	6	0	0
aae-miR-263b	CTTGGCACTGGGAGAATTCACAG	CONTIG_1908	33267	33357	pos	6	0	0
aae-miR-263b_star	TGGATCTTTTCGTGCCATCGT	CONTIG_1908	33267	33357	pos	1	0	0
aae-miR-275	TCAGGTACCTGAAGTAGCGCGCG	CONTIG_1651	96625	96720	pos	0	0	5
aae-miR-275_star	CGCGCTAAGCAGGAACCGAGACT	CONTIG_1651	96625	96720	pos	1	0	6
aae-miR-276-1	TAGGAACTTCATACCGTGCTCT	CONTIG_417	10988	11177	neg			
aae-miR-276-2	TAGGAACTTCATACCGTGCTCT	CONTIG_7617	38953	39087	pos			
aae-miR-277	TAAATGCACTATCTGGTACGACA	CONTIG_12646	3556	3644	pos			
aae-miR-278	TCGGTGGGACTTTCGTCCGTTT	CONTIG_1172	33037	33128	pos	6	0	2
aae-miR-279	TGACTAGATCCACACTCATTAA	CONTIG_17556	116934	117006	neg	26	0	2
aae-miR-281	CTGTCATGGAATTGCTCTCTTTA **4**	CONTIG_27100	34036	34204	pos	1	16	50
aae-miR-281_star	AAAGAGAGCTATCCGTCGACAGTA	CONTIG_27100	34036	34204	pos	164	5686	4806
aae-miR-282-1	AATCTAGCCTCTCCTAGGCTTTGTCTG	CONTIG_15805	19414	19548	pos			
aae-miR-282-1star	ACATAGCCTGACAGAGGTTAGG	CONTIG_15805	19414	19548	pos	2	0	0
aae-miR-282-2	AATCTAGCCTCTCCTAGGCTTTGTCTG	CONTIG_15982	5479	5613	neg			
aae-miR-282-2star	ACATAGCCTGACAGAGGTTAGG	CONTIG_15982	5479	5613	neg	2	0	0
aae-miR-283	CAATATCAGCTGGTAATTCTGGG	CONTIG_4323	32334	32426	neg	6	1	131
aae-miR-285	TAGCACCATTCGAAATCAGT	CONTIG_1634	212427	212492	neg	0	1	0
aae-miR-286a-1	TGACTAGACCGAACACTCGTATCC	CONTIG_21241	18407	18503	pos	8	0	0
aae-miR-286a-2	TGACTAGACCGAACACTCGTATCC	CONTIG_8291	17656	17752	pos	8	0	0
aae-miR-286b	TGACTAGACCGAACACTCGCGTCCT	CONTIG_18252	93413	93509	pos	10	0	0
aae-miR-305	ATTGTACTTCATCAGGTGCTCTGG	CONTIG_1651	105677	105766	pos	0	0	2
aae-miR-305_star	CGGCACATGTTGGAGTACACTTAA	CONTIG_1651	105677	105766	pos	0	0	7
aae-miR-306	TCAGGTACTGAGTGACTCTCAG	CONTIG_24640	55004	55128	pos	119	0	27
aae-miR-307	TCACAACCTCCTTGAGTGAGCGA	CONTIG_1157	147943	148043	neg	1	0	0
aae-miR-308	AATCACAGGAGTATACTGTGAG	CONTIG_6324	6462	6528	pos			
aae-miR-308_star	CGCGGTATATTCTTGTGGCTTGA	CONTIG_6324	6462	6528	pos	2	0	0
aae-miR-31	TGGCAAGATGTTGGCATAGCTGAAA	CONTIG_21960	98153	98301	neg	1	0	3
aae-miR-315	TTTTGATTGTTGCTCAGAAAGCC	CONTIG_21501	12498	12562	pos	27	1	0
aae-miR-315_star	CTTTCGAGCAGTAATCAAAGTc	CONTIG_21501	12498	12562	pos	5	0	0
aae-miR-316	TGTCTTTTTCCGCTTACTGCCG	CONTIG_13460	10738	10828	neg	0	0	1
aae-miR-317-1	TGAACACAGCTGGTGGTATCTCAGT	CONTIG_12640	88437	88524	pos	36	38	737
aae-miR-317-2	TGAACACAGCTGGTGGTATCTCAGT	CONTIG_8451	9550	9637	neg	36	38	737
aae-miR-3a-1	TCACTGGGCAAAGTTTGTCGCA	CONTIG_21241	18967	19043	pos	6	0	0
aae-miR-3a-2	TCACTGGGCAAAGTTTGTCGCA	CONTIG_8291	18216	18292	pos	6	0	0
aae-miR-3b	TCACTGGGCATAGTTTGTCGCA	CONTIG_18252	94047	94117	pos	3	0	0
aae-miR-3b_star	CGTCAAACTCCGTTCAGTTGGTG	CONTIG_18252	94047	94117	pos	1	0	0
aae-miR-33	GTGCATTGTAGTTGCATTGCA	CONTIG_18815	108787	108866	pos			
aae-miR-34	TGGCAGTGTGGTTAGCTGGTTGTG	CONTIG_12646	4278	4400	pos	44	55	639
aae-miR-375	TTTGTTCGTTTGGCTCGAGTTA	CONTIG_14081	238834	238944	neg	1	0	0
aae-miR-7	TGGAAGACTAGTGATTTTGTTGTT	CONTIG_31115	46028	46112	pos	14	0	0
aae-miR-71	AGAAAGACATGGGTAGTGAGATA	CONTIG_12793	18215	18396	neg	39	2	7
aae-miR-71_star	TCTCACTACCTTGTCTTTCATG	CONTIG_12793	18215	18396	neg	6	0	0
aae-miR-79	ATAAAGCTAGATTACCAAAGCAT	CONTIG_24640	55215	55284	pos	2	0	0
aae-miR-8	TAATACTGTCAGGTAAAGATGTC	CONTIG_16942	34594	34660	pos	32	1	38
aae-miR-8_star	CATCTTACCGGGCAGCATTAGA	CONTIG_16942	34594	34660	pos	7	2	4
aae-miR-87	GGTGAGCAAATTTTCAGGTGT	CONTIG_15587	30157	30251	pos			
aae-miR-927	TTTAGAATTCCTACGCTTTACC	CONTIG_1795	170460	170534	pos			
aae-miR-929-1	ATTGACTCTAGTAGGGAGTCC	CONTIG_12461	76647	76734	neg			
aae-miR-929-2	ATTGACTCTAGTAGGGAGTCC	CONTIG_34315	3294	3381	pos			
aae-miR-92a	TATTGCACTTGTCCCGGCCTAT	CONTIG_6821	67755	67832	pos	74	0	5

aae-miR-92a_star	CGGTACGGACAGGGGCAACATT	CONTIG_6821	67755	67832	pos	6	0	0
aae-miR-92b	AATTGCACTTGTCCCGGCCTGC	CONTIG_6824	4049	4131	pos	74	0	5
aae-miR-92b_star	AGGTCGTGACTTGTGCCCGTTTG	CONTIG_6824	4049	4131	pos	12	0	0
aae-miR-932	TCAATTCCGTAGTGCATTGCAG	CONTIG_28378	72232	72319	neg			
aae-miR-957	TGAAACCGTCCAAAACTGAGGC	CONTIG_591	147619	147787	pos			
aae-miR-965	TAAGCGTATAGCTTTTCCC	CONTIG_3410	87378	87524	pos			
aae-miR-970	TCATAAGACACACGCGGCTAT	CONTIG_11324	29252	29338	pos	1	0	3
aae-miR-981	TTCGTTGTCGACGAAACCTGCA	CONTIG_7302	184768	184853	neg			
aae-miR-988	CCCCTTGTTGCAAACCTCACGC	CONTIG_17684	77966	78086	neg	3	0	1
aae-miR-988_star	GTGTGCTTTGTGACAATGAGA	CONTIG_17684	77966	78086	neg	2	0	2
aae-miR-989	TGTGATGTGACGTAGTGGTAC	CONTIG_6774	51048	51137	neg	2	33	3
aae-miR-993	GAAGCTCGTTTCTATAGAGGTATCT	CONTIG_28305	82579	82660	pos			
aae-miR-996	TGACTAGATTACATGCTCGTCT	CONTIG_17556	112174	112267	neg			
aae-miR-998	TAGCACCATGAGATTCAGCTC	CONTIG_23911	46676	46764	pos	164	6	42
aae-miR-999	TGTTAACTGTAAGACTGTGTCT	CONTIG_6027	163128	163269	pos			
aae-miR-9a-1	TCTTTGGTTATCTAGCTGTATGA	CONTIG_19539	116871	116951	pos	16	1	2
aae-miR-9a-2	TCTTTGGTTATCTAGCTGTATGA	CONTIG_19541	4061	4141	pos	16	1	2
aae-miR-9b	TCTTTGGTGATTTTAGCTGTATGC	CONTIG_24640	55513	55604	pos	134	0	69
aae-miR-9c	TCTAAAGCTTTAGTACCAGAGGTC	CONTIG_24640	28160	28257	pos	5	0	0
aae-miR-iab-4-1	CGGTATACCTTCAGTATACGTAAC	CONTIG_17255	44653	44726	neg			
aae-miR-iab-4-2	CGGTATACCTTCAGTATACGTAAC	CONTIG_23219	23581	23654	pos			

Table 3.1. Sequence, location, and expression of miRNAs in *Aedes aegypti*. MiRNAs discovered in the *Ae. aegypti* embryo (Emb), sugar fed midguts (SF) and blood fed midguts (BF). Shown is the number of hits present among 454 sequenced samples for each identified miRNA. Also shown is the genomic location in *Ae. aegypti*, the strand where the sequence occurs, and the sequence of the mature miRNA.

Notes:

1. The first block (miR-M1 through miR-N2) are miRNAs that are newly reported from mosquitoes. We have not detected homology to any known miRNAs or genomic sequences outside of mosquito species (BLASTN, word size seven, and an e-value cutoff at 1). The second block (miR-1174 through miR-1889) contains miRNAs that are homologous to previously reported "mosquito-specific" miRNAs. The third block,

comprised of the remaining miRNAs, contains miRNAs that have homologs outside of mosquitoes.

2. Naming for the novel miRNAs is temporary. "-1", and "-2" suffixes refer to different hairpins that produce the same mature miRNA. "-a", and "-b" suffixes refer to different hairpins that produce similar but not identical miRNAs.

3. Underlined sequences are cases where there are extra bases at the 5' or 3' ends and they are the majority in at least six small RNA sequences. Italicized sequences are cases where such sequences are detected but they are either not the majority or there are less than six 454 sequence hits. As poly-As are added to the small RNAs, As at the 3' end cannot be confirmed by this sequencing approach.

4. There appears to be a one base shift in this sequence compared to the miRBase entry.

5. The contig, start, and end positions refer to the locations of the pre-miRNA hairpins.

6. The last three columns are the number of sequence hits in small RNA libraries obtained by 454 sequencing. The total hits for all Ae. aegypti miRNAs are 8369, 6260, and 9922 in embryos (Emb), sugar-fed midguts (SF), and blood-fed midgets (BF), respectively. There are a few cases where we did not distinguish the hits from nearly identical miRNAs.

Name	Gut_SF	Gut_BF	Gut_SF (hits per 100)	Gut_BF (hits per 100)	Fold Change
aae-miR-11 [2]	42	436	0.671	4.394	6.55
aae-miR-1175_star	2	59	0.032	0.595	18.61
aae-miR-13b [3]	2	42	0.032	0.423	13.25
aae-miR-184	76	1307	1.214	13.173	10.85
aae-miR-190	2	43	0.032	0.433	13.56
aae-miR-281	16	50	0.256	0.504	1.97
aae-miR-281_star	5686	4806	90.831	48.438	0.53
aae-miR-283	1	131	0.016	1.320	82.65
aae-miR-2a/2b/2c [3, 4]	80	249	1.278	2.510	1.96
aae-miR-306 [5]	0	27	0.000	0.272	N/A
aae-miR-317-1	38	737	0.607	7.428	12.24
aae-miR-317-2	38	737	0.607	7.428	12.24
aae-miR-34	55	639	0.879	6.440	7.33
aae-miR-8	1	38	0.016	0.383	23.98
aae-miR-989	33	3	0.527	0.030	0.06
aae-miR-998 [2]	6	42	0.096	0.423	4.42
aae-miR-9b [5]	0	69	0.000	0.695	N/A

Table 3.2. Comparison of the number of miRNA sequences in sugar-fed and blood-fed midgut samples. Only miRNAs that showed 25 or more sequences in one of the gut samples are shown. The second and third columns are raw number of sequence reads in sugar-fed (Gut_SF) and blood-fed (Gut_BF) samples, respectively. Columns four and five contain numbers numbers normalized according to the total miRNA hits according to literary precedence (Ruby et al., 2007). Although the number of small RNA sequences may not be sufficient for this approach to be quantitative, overall trends and suggest leads for expression analysis by northern blot.

Notes:

1. MiR-306 and miR-9b are physically linked, less than 500 bp apart.

2. MiR-11 and miR-998 are physically linked, less than 300 bp apart.

3. MiR-13b and miR-2a/2b/2c are in a physically linked cluster. Mir-13b is less than 200 bp apart from miR-2b.

4. It is difficult to distinguish between hits that match miR-2a/2b/2c.

Chapter 4.

Characterization and Function of miR-989 in mosquitoes

4.1. Abstract

MicroRNAs are important regulators of gene expression across eukaryotes. In mosquitoes, miR-989 was observed to be highly expressed in the adult female ovary suggesting a potential role in reproduction. In this chapter, I describe our efforts to investigate the evolution, expression, target and regulation of miR-989.

In genomic analyses of 14 different invertebrates, miR-989 was broadly conserved among holometabolous insects, but was not found outside this superorder. A detailed expression analysis of miR-989 in *Anopheline* and *Aedine* mosquitoes showed that miR-989 expression increased sharply following eclosion, reaching a peak approximately two days after emergence. It remained at a high level through 32 hours post blood feeding, then dropped dramatically, staying at a low expression level through at least 96 hours post-bloodmeal. The post-eclosion expression pattern of miR-989 in female mosquitoes resembled the profile of Juvenile Hormone (JH), a key reproductive hormone. MiRanda, a widely-used target prediction algorithm revealed 20 putative targets in *Ae. aegypti* and 16 in *An. gambiae* with scores above a previously established cutoff of 110. We used a

recently reported experimental approach to verify these predicted targets. However, the results are not yet conclusive.

By applying methoprene, a JH analog on ligated abdomens that are deprived of endogenous source of JH, we evaluated the possible induction of miR-989 by this key hormone in *Ae. aegypti*. We did not observe any obvious induction compared to the controls. In addition, high-throughput sequencing of small RNA clones using methoprene-treated samples indicated minimal, if any, induction of miR-989 by the JH analog, methoprene. However, nine known miRNAs were affected by methoprene, with up to 3.87 fold induction, and 3.15 fold suppression. Subsequent northern analysis verified changes for the miRNAs that showed 3.87 and 3.15 fold differences but not for miRNAs that showed changes below two fold. This analysis provides the foundation to study JH regulation of miRNAs in mosquitoes.

4.2. Introduction

miR-989 is a microRNA that was recently discovered in *Drosophila* and *Anopheline* mosquitoes (Winter et al., 2007; Ruby et al., 2007; Mead and Tu, 2008). Direct sequencing of *Ae. aegypti* miR-989 is described in Chapter 3. The previously reported *An. stephensi* mature miR-989 (Mead and Tu, 2008) has a single base difference compared to miR-989 found in *D. melanogaster, Cx. pipiens, An. gambiae* and *Ae. aegypti*. It is possible that the difference reported in Mead and Tu, 2008 is an artifact, as

the mature miR-989 sequence of Winter et al., 2007, and of Ruby et al., 2007 is conserved across all other dipterans examined.

In *An. gambiae,* miR-989 was found to be present in Chromosome 3L in the positive orientation at location 2905395-2905525 (Winter et al., 2007; Mead and Tu, 2008; Griffiths-Jones et al., 2006). In *D. melanogaster*, it was observed to be present in Chromosome 2R in the reverse orientation at location 10032856-10033029 (Ruby et al., 2007; Griffiths-Jones et al., 2006). MiR-989 is intergenic (Ruby et al., 2007; Griffiths-Jones et al., 2006), whereas most miRNAs are intragenic (Miska, 2005; Kim et al., 2007; Liu et al., 2007). MiR-989 is also not clustered with other miRNAs in *Drosophila* (Ruby et al., 2007; Griffiths-Jones et al., 2006), and the closest known miRNA, miR-308, in the same orientation on the same chromosome in *Drosophila* is roughly 70 kb away (Ruby et al., 2007). MiR-989 is therefore a monocistronic miRNA, based on the current literature.

Ruby et al. report a large hairpin precursor, 174 nt long, with an unusually large sequence intervening between the miRNA and miRNA* in *D. melanogaster* (Ruby et al., 2007). They found that the intervening sequence varied considerably amongst fly species (Ruby et al., 2007). While the observed intervening sequence was 99 nt in *D. melanogaster*, the intervening sequence in *D. pseudoobscura* was nearly half as long, at 52 nt in length, for a final hairpin more similar in length to the 131 nt hairpin reported in *An. gambiae* (Ruby et al., 2007; Winter et al., 2007; Griffiths-Jones et al., 2006).

MiR-989 was present in all seven *Drosophila* species which were been examined in Ruby et al., representing both the Sophophora and Drosophila subgenera (Ruby et al., 2007). Among *Drosophila*, expression was strongly observed in adult bodies with little or no expression observed in other life stages (Ruby et al., 2007). This agrees with the expression pattern observed in *An. stephensi* lifestages for miR-989 where expression was nearly exclusively in adult females (Mead and Tu, 2008). MiR-989 was expressed almost exclusively in the mosquito ovaries, and its expression decreased in response to bloodfeeding (Mead and Tu, 2008). Dme-miR-989 differs from aga-miR-989 by a single nucleotide in the 3' end of the mature miRNA (Griffiths-Jones et al., 2006).

Little is known about miR-989, particularly in mosquitoes, though the limited expression information available suggested this miRNA might play an important role in mosquito reproduction. Thus, we sought to discover the breadth of its conservation in insects to uncover miR-989 targets and investigate the regulation of miR-989.

4.3. Materials and Methods

Mosquitoes

Anopheles stephensi (Indian strain) and *Ae. aegypti* (Liverpool strain) mosquitoes were maintained as in Mead and Tu, 2008. For small RNA cloning of JH induction samples, Rockefeller strain *Ae. aegypti* mosquitoes were used, and also were maintained as in Mead and Tu, 2008. *T. amboinensis* (Puerto Rico) mosquitoes were maintained in the

same insectary under the same conditions, but did not require bloodfeeding. *T. amboinenesis* larvae were maintained on *Ae. aegypti* larvae.

Bioinformatic Analysis of miR-989

Bioinformatic analyses were conducted using Ensembl, Vectorbase, and NCBI genome browsers as well as the silkworm Database and the Nasonia Genome Project to examine conservation of miR-989 (Flicek et al., 2008; Lawson et al., 2007; Werren et al., 2004; Wang et al., 2007). Using dme-miR-989 and aga-miR-989 sequences (Griffiths-Jones et al., 2006) as references, BLASTs were conducted against 13 insect species including three mosquitoes (*Ae. aegypti, An. gambiae, Culex pipiens*), four *Drosophila* species (*D. pseudoobscura, D. melanogaster, D. simulans, D. yakuba*), and six additional insects including *Bombyx mori* (silkworm), *Tribolium castaneum* (flour beetle), *Apis melifera* (European honeybee), *Acyrthosiphon pisum* (pea aphid), *Nasonia vitripennis* (jewel wasp), and *Pediculus humanus corporis* (body louse). Additionally, an outgroup invertebrate *Caenorhabditis elegans* (nematode) was also examined. Alignments were prepared using clustalw at the Kyoto University Bioinformatics Center website (http://align.genome.jp/). Default settings were used (Gap open penalty: 15, and gap extension penalty of 6.66).

Folding of hairpins was conducted using RNAfold (Hofacker et al., 1994). Published precursor sequences (*D. melanogaster*, *An. gambiae*) were retrieved from miRBase (Griffiths-Jones et al., 2006). Precursor sequences were examined by RNAfold for

miRNAs demonstrating conservation of the mature miR-989 from the list of species above. This included *D. melanogaster* and *An. gambiae*, *Ae. aegypti, Cx. pipiens, A. mellifera*, *B. mori,* and *T. castaneum*. For brevity and clarity, we report only the RNAfold result for the *D. melanogaster* hairpin and not for the other Drosophilid flies.

Detailed northern analysis of miR-989 expression in female mosquitoes

As post-bloodmeal expression of miR-989 fluctuated (Mead and Tu, 2008), the expression of miR-989 in *An. stephensi*, and *Ae. aegypti* was examined in depth across a detailed time course. Samples were taken every two hours post-emergence from a colony of mosquitoes, through 8 hours, then at 12 hours, 24 hours, and at each 24 hour point through 5 days.

Following bloodfeeding, samples were taken every hour for the first 6 hours, followed by an eight hour and a 16 hour timepoint. Subsequent timepoints were taken every eight hours through 96 hours post bloodfeeding. Previous analyses showed a slight increase in expression of miR-989 at 24 hours post-bloodfeeding, followed by significantly lower expression at 48 and lower expression still by 72 hours post-bloodfeeding. Recovery of signal to pre-bloodfeeding levels took several more days. A cohort of non-bloodfed mosquitoes were also collected to provide a reference control for miR-989 expression.

Northerns were conducted using the procedure detailed in Mead and Tu, (2008). However, for the miR-989 time course I used 15 well gels, and loaded five μg of total

RNA per well. To verify that the observed results are not due to procedural artifacts, RNA northerns were repeated, running key samples, those which demonstrated a shift in expression, adjacent to one another on the same gel.

Target analysis of miR-989

Analysis of targets in *Cx. pipiens*, *Ae. aegypti*, and *An. stephensi* was carried out using miRanda, a program which looks for putative targets of a given miRNA from databases of RNA sequences. This relies upon conserved signature characteristics of miRNA:target interactions including critical base pairing at the 5' end of the mature miRNA as well as sequences with imperfect seed matches but compensatory 3' end pairing (John et al., 2005).

I downloaded databases from Ensembl for *An. gambiae*, and *Ae. aegypti* 3'-UTRs of the transcriptomes. These databases were analyzed by miRanda, v1.0b (John et al., 2005). Settings used were a Gap Open Penalty: -8.000000, Gap Extend: -2.000000, Score Threshold: 90.000000, Energy Threshold: -20.000000 kcal/mol. After ranking the transcriptome results, I looked up the top 20 scores based upon individual targets for each category, and the top twenty for each category having one target or more (total score). There was little difference whether I looked for top score from a single target, or top total score, so most hits found were single target results. Scores of 110 and above were considered "acceptable" based upon examination of known miRNAs and targets by John et al. (2005). Transcripts in the top 20 were referenced in Ensembl, Vectorbase, and

Flybase to attempt to uncover an identity for each; for many, functions were predicted based upon homology to known transcripts in other organisms, primarily *C. elegans* and *D. melanogaster*. Targets related to ovarian functions were particularly of interest.

In addition, the top 20 miRNA targets for miR-989 reported at TargetScan Fly (Friedman et al., 2009), were also examined. TargetScan Fly examines 8-mer and 7-mer conserved seed sites in the 3'-UTR of the *D. melanogaster* sequences (Friedman et al., 2009). It examines conservation of the site across 12 Drosophilid species, and uses information about target location and base composition in predicting a score (Friedman et al., 2009).

PCR-based Target Analysis

Abdomens were obtained from 10 *Aedes aegypti* non-bloodfed females through dissection and placed into 20 µl buffer (5 mM KCl, 0.1mM NaH_2PO_4, pH 6.8). The remaining protocol was heavily based upon the methods listed in Andachi, 2008. Buffers used are according to formulations described in Andachi (2008), unless otherwise indicated. Briefly, samples were homogenized in cell fractionation buffer (PARIS kit, Ambion) to retain both protein and nucleic acids intact, thus leaving units of the RISC complex associated with the miRNA:mRNA duplex. Centrifugation was carried out at 20,000xg for five minutes to precipitate cellular debris and nuclei. The supernatant was treated with SDS on ice for one minute to destabilize proteins present. From here, an aliquot was drawn and added to RT reaction buffer to dilute out the SDS prior to an RT-PCR reaction. This reaction was placed at 4° C in a thermocycler and at one minute

intervals the temperature was increased by approximately 4.1° C, to a final temperature of 37° C. After 37° C was reached, the reaction was maintained at this temperature for one hour. This step utilizes the miRNA as a primer for generating a hybrid cDNA that retains the miRNA as its 5' end. Next, phenol:chloroform extraction was conducted to remove proteins, and other impurities.

An ethanol precipitation followed, and the pellets were resuspended in reaction buffer and incubated at 25° C for 10 minutes to synthesize short strands, followed by incubation at 70° C for 30 minutes to synthesize full-length cDNA. In this reaction, moderately degraded residual mRNA fragments function as primers to generate the cDNA second strand. Samples were cleaned up using a Qiagen PCR purification kit. At this point, the samples were ready for use in PCR. The twenty top candidates were chosen for PCR based upon the miRanda results discussed above.

Forward primers were selected from the terminal protein-coding region, and were designed to be close to 57° C, the Melting Temperature (Tm) of the miR-989 DNA primer. Primer design was carried out using Primer 3 in the SDSC workbench (Subramaniam, 1998). Primers are listed in Table 4.1. Nested PCR was carried out using Taq polymerase and buffers (Takara, Tokyo, Japan) according to instructions provided with the polymerase. Forward primer1 was used first, for 30 cycles, followed by a repeated PCR using 1μl of the first reaction using nested primer Forward Primer two, downsteam of the first forward primer. Reverse primers were the miR-989 primer. Template for the first PCR round was the cDNA generated above. Thermocycler

conditions used were 94º C for three minutes followed by 30 cycles of 94º C for 30 seconds, 55º C for 30 seconds, and 72º C for one minute. After 30 cycles, an additional five minutes at 72º C were allowed to fill in ends of products.

Five µl of each sample were separated on a 1% agarose gel beside a 100 bp ladder (Promega) for size reference. Following electrophoresis, samples exhibiting bands were cloned into pGEM T-Easy plasmids via TA cloning according to manufacturer's instructions (Promega). Recombinant plasmids were used to transform competent *E. coli* (Mach1-T1^R cells, Invitrogen) according to manufacturer's instructions, using LB plates and media prepared to 100 µg/mL ampicillin. Clones were sequenced at Virginia Bioinformatics Institute (Blacksburg, VA) to identify inserts by BLAST against the *Ae. aegypti* genome (Ensembl).

Antisense-mediated knockdown of miR-989 in Ae. aegypti

Antisense-mediated knockdown of miR-x2 was carried out by intrathoracic injection of ~0.5 µl anti-miR-x2 LNA into 40 recently emerged (0-3 hrs old) females. A cohort was injected with anti-GFP (not predicted to affect mosquitoes), while another set was subjected only to chilling and handling and remained uninjected. Mortality rate was examined on a daily basis. Additionally, representative random samples were examined at 24 and 72 hours post injection, and at 24 and 72 hours post-bloodfeeding for gross morphological changes in the ovaries.

Abdominal Ligation

Abdominal ligation was conducted based upon a variant protocol developed by Jinsong Zhu (Virginia Tech). Liverpool strain *Ae. aegypti* mosquitoes were collected between 0-1 hour post-emergence, and the head and approximately one-third of the anterior thorax were removed by excision, effectively removing the corpora allata (CA). The remaining thorax was glued with Vetbond (3M) to a glass slide, abdomen up, to seal the wound and maintain the tissues in a viable state. From here, the abdomen was treated in one of three ways: 1µl acetone, 1 µl methoprene (500 ng/µl) in acetone, or no chemical treatment (Figure 4.2). A cohort of "undecapitated" mosquitoes was also maintained for the duration, before decapitation (no chemical treatment) immediately prior to freezing in liquid nitrogen. Sixteen mosquitoes were prepared for each methoprene sample, and acetone only sample, whereas eight mosquitoes were prepared for both untreated samples. This was to ensure that sufficient material would be available for small RNA cloning.

Incubations were carried out for four and 12 hours, at which time, the samples were frozen in liquid nitrogen. At this point, cohort undecapitated samples were decapitated and immediately frozen simultaneously with the decapitated cohort. RNA isolation was carried out upon the samples to generate total RNA, using a miRVana kit as detailed in Mead and Tu, 2008. This experiment has been repeated several times, examining expression of miR-x2 after four, six, 12, 14, 24 and 36 hours as well.

Sample collection and preparation for small RNA analysis samples was as described above for abdominal ligation, with the following modifications to treatment. Rockefeller strain *Ae. aegypti* female adults were anaesthetized by cold treatment on ice for 40-50 minutes following collection. Six to eight females were collected for each of two sets of samples, methoprene-treated, and acetone-treated. For methoprene treatment, 0.5 ng methoprene was delivered in 0.5 µl volume of acetone, and for acetone-treatment, 0.5 µl acetone was delivered as a control. Following 12 hours incubation, the samples were stored in 150 µl TRIzol (Invitrogen, Carlsbad, CA) at –80º C until RNA isolation was conducted. A replicate sample set was also prepared in the same manner and stored as above.

Northern analysis of abdominal ligation samples was conducted using 10 µg of total RNA as in Mead and Tu, 2008. Membranes were probed with a dig-labeled LNA anti-miR-989 probe, as detailed in Mead and Tu, 2008.

Small RNA Analysis of JH Induction

Total RNA isolation was conducted using TRIzol according to the manufacturer (Invitrogen). Quantitative Real Time PCR (qRT-PCR) was used to examine early trypsin induction. First strand synthesis was performed with one µg total RNA using an Omniscript RT kit (Qiagen, Valencia, CA) as instructed by the manufacturer, using the Protector RNase Inhibitor (Roche, Indianapolis, IN), and an oligo dT primer (Roche). Prior to the reverse transcription reaction, the samples were treated with DNase I,

Amplification grade (Invitrogen) to remove potential genomic DNA contamination. cDNA from the RT step was then used for PCR amplification with SYBR Green dye (Invitrogen) for quantitative fluorescent visualization of amplification in qRT-PCR using an ABI 7300 RT-PCR system (Applied Biosystems, Austin, TX). Samples were analyzed in triplicate and compared to an internal control, ribosomal protein S7 to determine the induction of early trypsin by methoprene.

Small RNA cloning was conducted using the Illumina DGE-Small RNA Sample Prep Kit according to manufacturers instructions (Illumina, San Diego, CA). Briefly, 10 μg of each total RNA sample was used for gel purification of small RNA. Subsequent rounds of addition of linkers first to the 3' end then to the 5' end of the small RNA with gel purfication at each step were carried out. RT-PCR was used to generate double-stranded cDNA libraries.

To test the quality of the libraries, a 1μl aliquot of each library was used to clone into pCR II Blunt TOPO plasmid according to the Zero Blunt TOPO PCR Cloning Kit (Invitrogen) manual. The plasmids were used to transform One-Shot Chemically competent *E. coli* cells (Invitrogen) and grow colonies according to the protocol. Plasmid isolation was carried out using a Wizard Plus Miniprep DNA purification system (Promega, Madison, WI), isolating plasmids for 10 colonies in each library. Aliquots of each set of 10 plasmids were sent to Virginia Bioinformatics Institute for sequencing to verify the presence of inserts which could be mapped back to the genome for each group.

Following verification of valid inserts, Illumina reversible terminator "deep" sequencing was conducted by the lab of Dr. Gao (VCU).

Results were tabulated by BLAST and examined for changes in expression. Small RNA fold changes were normalized based upon the total number of hits for each sample. Two miRNAs, miR-9a and –1891 which demonstrated a three to four fold change were examined by northern blot to verify results. Additionally, the expression of miR-305, and –989, two miRNAs of interest to the lab, were analyzed as well by northern.

Northerns were conducted as above, to examine the expression of miR-9a, -305, -989, and –1891, but using 15-well gels to allow more samples to be analyzed at once. For the smaller well sizes, five μg total RNA was loaded for acetone and methoprene samples. Northern blots were probed using antisense miRCURY LNA Dig-labeled oligos (Exiqon, Vedbaek, Denmark) for each miRNA of interest. A replicate set of samples was also examined to verify the northern results of the first sample set. Both sets were observed to have strong induction of early trypsin (26.6 fold for the first set (used in cloning) and 91 fold for the second set) by methoprene.

4.4. Results

Conservation of miR-989

MiR-989 was examined by BLAST across 13 insects and a nematode, with special emphasis upon dipterans. MiR-989 was conserved in all four Drosophilds examined (*D. pseudoobscura, D. melanogaster, D. simulans,* and *D. yakuba*) and in all three mosquitoes (*Ae. aegypti, An. gambiae*, and *Cx. pipiens*) examined as well (Figure 4.3.A). MiR-989 was also present in the genomes of *B. mori, A. mellifera*, and *T. castaneum* (figure 4.3.B). The *B. mori* genome showed two hairpins with identical sequences, in two different scaffolds, though the flanking region surrounding the hairpin was sparsely conserved between them. MiR-989 was not found in the body louse (*P. humanus*), the nematode *(C. elegans),* the wasp, *N. vitripenis*, and *A. pisum* (aphid), all having less than 17 nucleotides in common with the mature sequence. The mature miR-989 sequence was found to have at least 18 nt in common across all the species where it was present and in particular the seed region, nucleotides 2-8 from the 5' end, shows perfect conservation. MiR-989 is intergenic in all dipterans examined, according to the current annotations.

Temporal expression of miR-989 in female mosquitoes

Results of northerns have expanded upon previous analyses of the expression of miR-989 in adult female mosquitoes. There was a sharp rise in the expression of miR-989 beginning approximately 12 hours following emergence from the pupal case, and this increased until approximately 48 hours post-emergence (Figure 4.4.A). A strong signal was detected until a drop in expression was observed at approximately 32-40 hours post-bloodmeal. In *Ae. aegypti*, the drop was more rapid and more pronounced than in *An. stephensi*. There was also a slight drop in expression observed between 5- 8 hours post-

bloodfeeding. These results are consistent across species, showing up in both *An. stephensi* and *Ae. aegypti*. Re-running key samples together on the same blot also showed the same expression pattern suggesting that the findings were not procedural artifacts. The expression of miR-989 was examined in *T. amboinensis* lifestages, and found to occur only in adult females (Figure 4.4.B).

Target analysis of miR-989

MiRanda yielded 20 target transcripts for miR-989 above a score of 110 for *Ae. aegypti* and 16 target transcripts which met the score criteria for *An. gambiae*. Results for *Ae. aegypti* targets are reported below in Table 4.1. Although targeted mRNA themselves have orthologs, target site conservation was not detectable by BLAST. Due to limited annotation of the *Ae. aegypti* genome, 11 out of 20 targets were hypothetical proteins, though DNA binding and protein transport targets were well represented.

PCR-based detection of miR-989 targets

Target prediction by miRanda revealed 20 potential targets for miR-989 in *Ae. aegypti*. These 20 were examined through a PCR-based strategy taking advantage of the association of miRNA with their targets to use the miRNA as a primer for cDNA synthesis. Of the 20 targets examined, three yielded products detectable by gel electrophoresis (Figure 4.5). These correspond to amplification of products by primers for AAEL002261-RA, the Aedine ortholog of GTP cyclohydrolase I, AAEL-000246, a

putative cadherin, and AAEL-013816, a hypothetical protein with DNA binding domains. However subsequent sequence analysis indicated the presence of retrotransposon artifacts.

Analysis of JH regulation through abdominal ligation and small RNA cloning

Abdominal ligation experiments revealed that miR-989 showed no observable change in expression following introduction of the JH analog, methoprene. This occurred regardless of the length of time the samples were allowed to incubate following treatment, and regardless of whether the samples were treated or remained untreated. To further analyze the role of JH on miR-989, I conducted small RNA cloning on a methoprene treated and an acetone treated abdominal ligation sample which had been incubated for 12 hours, and determined to have a 26 fold induction of early trypsin, a marker for JH induction.

BLAST results of small RNA cloning are displayed in Table 4.2, showing fold changes in methoprene treated samples. I observed 95 different known miRNAs, and miRNA*s. Among these, fold inductions of miRNA by methoprene ranged from eight-fold increase through 4-fold suppression of miRNA expression. However, counts below 30 reads for a given miRNA in a sample may not give sufficient representation to accurately examine fold changes without their values being swamped by random sampling error. Barring these, fold changes ranged from 3.87 fold increase to a 3.15 fold suppression of miRNA expression for 9 miRNAs.

To provide additional support and to verify the high-throughput sequencing data, northern blots were carried out to determine if miRNAs were indeed changing in expression. I examined the expression of miR-1891, and miR-9a, both of which had >30 reads per sample and which covered the range of fold changes observed. MiR-1891 and –9a display fold changes by northern that correlate with their small RNA cloning results. However, the lesser changes observed for miR-989 and miR-305, according to small RNA cloning, at 1.45 and 1.92 fold, respectively were undetectable by northern.

4.5. Discussion

miR-989 is broadly conserved across holometabolous insects

Bioinformatics can provide clues to the orthology and function of miR-989. The conservation of miR-989 was examined broadly across insects, despite limitations in genome availability, assembly and annotation. MiR-989 was absent in *C. elegans* and other invertebrates represented in miRBase, which suggests but does not prove, that miR-989 is not present outside of insects. Within Insecta, we examined representatives of the infraclass Neoptera, containing the majority of insects.

There are multiple competing classification schemes for phylogeny within Neoptera. Savard et al., 2006, use 18S ribosomal sequences to assign a classification tree of representatives from diverse orders within Neoptera (Figure 4.6). Conservation of miR-989 was analyzed in at least one representative of each branch, leading to the discovery

that miR-989 could be found among all branches of holometabolous insects surveyed, but not in hemimetabolous insects examined (*A. pisum,* and *P. humanus* (not shown in Figure 4.6))

MiR-989 was found within all seven dipterans examined (four Drosophilids, and three mosquitoes). Though the miR-989 analysis here examined conservation of miR-989 within the *Drosophila* subgenus Sophophora, Ruby et al., 2007 report the presence of miR-989 in representatives of both Sophophora, and Drosophila subgenera (Ruby et al., 2007). The genomic analysis presented here is an examination of the presence of miR-989 among only two of the three subfamilies within mosquitoes (*Anophelinae,* and *Culicinae*) due to a lack of sequence information about *Toxorhyncitinae*. However, I was still able to confirm the presence of miR-989 in adult females in this subfamily as well by northern analysis (Figure 4.4.B) Thus, miR-989 is present in all subfamilies of mosquitoes and both subgenera of *Drosophila*.

Hairpin conservation is expected to be low for dme-miR-989 due to the wide variation in hairpins observed even within the same genus for this miRNA (Ruby et al., 2007). Nevertheless, among the hairpins examined, *D. simulans* and *D. yakuba* show a very high conservation of the majority of the stem loop. *D. pseudoobscura* shows a much lower conservation. Very little conservation of the large dme-miR-989 stem loop occurs outside of the mature miRNA.

No ame-miR-989 exists in the literature (Griffiths-Jones et al., 2006; Weaver et al., 2007). Although the stem loop for ame-miR-989 is unorthodox, there is considerable variation among stem loops for miR-989 even within the same genus, *Drosophila* (Ruby et al., 2007). Additionally the RNAfold prediction for dme-miR-989 (not completely shown due to space limitations) also suggests the presence of two loop regions away from a central loop, similar to the structure found for ame-miR-989. Lastly, the high conservation of a mature miR-989 strongly suggests it has been retained as a functional miRNA in *A. mellifera*. Hymenoptera is a basal order among holometabolous insects (Savard et al., 2006). Within Hymenoptera, the mature miR-989 was found in *A. mellifera*, but not in *N. vitripennis*. A loss of miR-989 may have occurred within Hymenoptera.

It was observed that two copies of miR-989 were present in *B. mori*, in scaffold 4254, and nscaf1898 (http://silkworm.swu.edu.cn). Though the genome is not fully assembled, it is unlikely that these are simply the same genomic sequence, as the genomic region surrounding the hairpins differs. Thus these are likely to be paralogous miRNAs. miR-989 matched two small non-coding RNA sequences found in a *B. mori* ovarian sample (*Bombyx mori* non-coding RNA, ovarian small RNA-35893, and –9097, Kawoka et al., 2008).There may be temporal differences from mosquito expression, as the samples were isolated in later pupal ovaries (Kawoka et al., 2008), a time where minimal miR-989 is observed in mosquitoes (Mead and Tu, 2008). The expression of miR-989 in *B. mori* provides empirical evidence for the presence of this miRNA across divergent groups of holometabolous insects.

Temporal expression suggests that miR-989 may play a role in reproduction

miR-989 was observed to increase in expression following pupal emergence and drop in expression following bloodfeeding in a pattern reminiscent of JH expression (Figure 4. 7). There is a correlation between the expression of JH and the expression of miR-989, particularly during the first few days following adult emergence. However, correlation does not equal causation. It is also possible that both JH and miR-989 are regulated by a third agent, such as allatostatin or allatotropin, neurosecretory peptides which have been found to regulate production of the corpora allata (Hagedorn, 1996). Allostatin type A precursor transcripts have been found in both the brain and the ovaries in crickets, and the mature peptide inhibits both CA production of JH and ovarian production of vitellogenin and ecdysteroid, respectively (Meyering-Vos, and Hoffmann, 2003), though it is unknown whether a similar process occurs in mosquitoes. However, immunostaining suggests that allatostatin-C and allatotropin in *Ae. aegypti* and *An. albimanius* play regulatory roles in both the corpora allata, and the ovaries (Martinez et al., 2005). Whether these processes are related to the function of miR-989 remains to be determined, and there is currently no empirical support for either allatostatins or allatotropin regulation of miR-989. Removal of the heads did not impact miR-989 expression in our studies.

miR-989 targeted mRNA are conserved

Targets for miR-989 are listed in Table 4.1, below. Orthologs for many mRNA targets could be found among Dipteran insects underscoring the validity of the target search and complementing the genomic conservation of miR-989 in Diptera. Though some mRNA functions could not be clearly identified due to limited annotation and functional studies in mosquitoes, in some cases functions were hinted at by motifs. Among top scoring candidates, DNA binding and protein transport functions were well represented, suggesting that miR-989 may play a role in fine regulation of mRNAs in ovarian development.

The authors of the original miRanda paper (John et al., 2005) find that miRanda results can be considered reliable based upon testing with known miRNA targets and with scrambled miRNAs. With *lin*-41, a well-studied target of let-7, they found two targets with scores of 115, and 110 (John et al., 2005). However, they recommend that the results should be screened for:

1. Conservation of the target site across different species.
2. A target score of >110 if the gene does not have more than one target.

In addition, any results need ultimately to be verified experimentally, as false-positive rates for miRanda have been estimated to be between 24% (Watanabe et al., 2007) to 35% (Grün and Rajewsky, 2008).

The targeted transcripts were conserved across *An. gambiae* and *D. melanogaster*, though predicted target sites could not be detected by BLAST to determine whether conservation occurs, as there is simply too little homology present. This suggests that the nucleotides

are not all conserved, though it is likely that some bases which do not pair with the miRNA will be subjected to "semi-random" mutation over time, and will diverge, provided that they do not affect secondary structure or complementarity with the miRNA.

It is unlikely that all targets will be valid. After all, the transcriptome reflects all transcripts from the organism, so a given mRNA may not ever be colocalized with the miRNA. In our case, miR-989 is expressed almost exclusively in the adult ovary with little expression elsewhere. Ovarian transcriptome libraries may show that a given mRNA is expressed in the ovaries, though may not provide sufficient localization information to verify the colocalization of miR-989 and the mRNA in question. Therefore, all predicted target results and implied functions should be taken with a high degree of caution, and merely suggestive of further avenues of inquiry.

PCR-based Target Analysis may lead to empirical support for miR-989 functions

Abdomens of female mosquitoes were chosen for sampling to enrich for the ovaries while still allowing very rapid sample collection in the absence of RNase inhibitor solutions that could interfere with subsequent enzymatic steps. Experimental examination of predicted targets through the RT-PCR-based method of Andachi followed by PCR of putative targets for miR-989 revealed three products that gave significant bands when analyzed by gel electrophoresis. Products were observed using primers for GTP cyclohydrolase I (AAEL002261-RA), a cadherin (AAEL000246-RA), and a conserved

hypothetical protein (AAEL013816-RA). However, retrotransposon artifacts were observed.

Complementary target sequences in the 3'-UTRs to miR-989 were not well conserved by BLAST analysis in the *An. gambiae* genome for any of the targets analyzed through PCR, but it is expected that some variation can occur as there can be mutiple non-complementary bases, particularly outside of the seed region, making BLAST analysis difficult.

Target prediction scores by miRanda ranged from 120 for GTP cyclohydrolase I, (the highest scoring individual target among the 3'-UTRs of the *Ae. aegypti* transcriptome), to the 110 cut-off that an individual target should have for accuracy as suggested by the miRanda authors (John et al., 2005). MiRanda focuses upon complementarity using a given miRNA, especially in the critical seed region and the thermodynamics involved (John et al., 2005). That being said, even with a score over 110, approximately a quarter to a third of predicted miRanda targets may be false positives. If the prediction was wrong, it is possible that a predicted target could be amplified, though in the event of miRNA remaining intact with their real targets as suggested by Andachi, false positives would not occur as the miRNA:target interaction is required to provide priming in cDNA synthesis.

It is possible that not all miRNAs even under the best conditions would remain in association with their real targets. If disruption were too strong, dislodged miR-989 could

happen to find that specific mRNA, and it could be detected based upon PCR amplification. If a dislodged miR-989 happened to find an unreal mRNA target or contaminating genomic DNA target in solution, it could yield the retrotransposons observed. These may already be present in abundance and dwarf a real target in PCR. The steps were processed quickly and efficiently, and maintained in cold as discussed above, limiting the extent to which miRNA:target dissociation might occur in the absence of Argonaute, or due to chemical/physical dislodging during protein removal. Cold may have been a double-edged sword, leading to permitting errant PCR to occur once dislodging occurred. A small quantity of sample (10 mosquito abdomens) was used, and the sample was highly diluted several times during processing, spatially favoring reassociation of a miRNA with its valid target in the event of dissociation.

Non-specific mispriming interactions were detectable as we intentionally left off two nt on the 3' end of the DNA miR-989 PCR primer to recognize whether the targets were the same as predicted. The size determined by electrophoresis, and sequencing the products, are strong ways to recognize false positives. Nested PCR also helps to get rid of potential artifacts. If small RNA cloning from the samples is carried out we could potentially estimate a false positive rate based upon the rate at which products have been generated from regions other than the 3'-UTR. Targets in regions outside of the 3'-UTR are far less common than within the 3'-UTR (Gu et al., 2009). While this preliminary experiment was fruitless, it is worthwhile for this lab to continue to troubleshoot the procedure as this method offers experimental validation of targets, a solid way to determine the function of miRNAs.

Microinjection of Ae. aegypti with anti-miR-989 did not lead to observable phenotypic changes

This study was a preliminary study with antisense knockdown in adult mosquitoes. Knockdown attempts with anti-miR-989 failed to demonstrate that knockdown of miR-989 had an observable phenotype. Several potential confounds exist. First, it is unknown whether there will be penetrance of the LNA into ovarian cells. The primary site of LNA concentration in mammals is the kidneys, suggesting that the LNA is removed from circulation over time (Roberts et al., 2006). Another potential confound is that antisense-mediated miRNA knockdown has at times led to significantly different results than that of a mutational deficiency (Bushati and Cohen, 2007). This suggests that the results of knockdown should be interpreted with caution, as they do not indicate that miR-989 does not play an important role, but rather, that it's role could not be evaluated, as knockdown may not have occurred. It is also important to note that most depletions of individual miRNAs do not yield a seriously abnormal phenotype in animals (Bushati and Cohen, 2007), so it is very possible that knockdown of miR-989 will not result in a visible phenotype. Therefore, we probably would have to rely upon other means of assessing the function of miR-989.

Abdominal ligation to study the impact of JH on miR-989 expression in Ae. aegypti

Our northern analyses in both *An. stephensi* and *Ae. aegypti* adult females have shown that miR-989 expression correlates with JH levels in adult females, at least before blood feeding (Hagedorn, 1996). It is possible that the transcription of miR-989 may be regulated by JH. A relatively simple way to examine the role of JH on miR-989 expression is by knockdown through abdominal ligation (Hagedorn et al., 1997). By removing the CA shortly after emergence, JH levels are kept at a minimum as the synthesis of JH halts.

With the CA removed, JH levels should remain minimal, and if JH is responsible for the synthesis of miR-989, expression should remain low even through 36 hours post-excision. However, with the addition of methoprene, a JH analog, expression of miR-989 should be rescued if JH does regulate miR-989. The control samples of acetone only, and of no treatment, should show little miR-989 whereas the methoprene-treated sample should show a strong miR-989 signal. This is a good test of JH regulation as it examines whether restoring one component (JH, in this case an analog of JH) results in restoration of miR-989 expression.

My preliminary analyses of miR-989 in samples subjected to abdominal ligation showed little difference between the different samples at the same timepoint. The expression of miR-989 was present in samples at 4 hours post-treatment, regardless of treatment. This was unexpected as northern analyses in both *An. stephensi* and *Ae. aegypti* without decapitation show a very low expression at four hours (see figure 4.4.A). Further, these

samples had been verified for early trypsin expression, indicating that JH induction had occurred.

The preliminary results led us to believe that removal of the head and ~1/3 of the thorax either removed an inhibition of miR-989 expression, or else signalled the expression of miR-989. I conducted another abdominal ligation experiment, where newly-emerged *Ae. aegypti* samples were prepared as above. Decapitation of 0-1 hr female mosquitoes was followed by treatment of either methoprene, acetone, or no treatment, then incubation of all samples for four or 12 hours. An additional group was left non-decapitated in the incubator for four or 12 hours, and at the end was decapitated, and untreated (Figure 4.8). The results were analyzed by northern, and no difference was seen between the samples, suggesting that decapitation alone was not creating a change in miR-989 expression. Further examination is warranted as JH signaling is a major regulator from the head for ovarian development (Hernandez-Martinez et al., 2007).

Small RNA Analysis after JH Induction

JH is a key regulator of ovarian development and pre-vitellogenesis in adult female mosquitoes (Noriega et al., 1997). The corpora allata (CA) located near the brain is the site of JH production. Though several JH variants exist in nature, JH III is the principal JH variant in adult female mosquitoes (Hagedorn, 1996). JH expression begins to rise from two pg/female following eclosion, and within 48 hours, reaches a peak titer of 20 pg/female (Noriega et al., 1997). After bloodfeeding, a decline in JH is observed, with a

pronounced and rapid drop in JH occurring approximately one day post bloodmeal, to a level slightly lower than at eclosion (Hagedorn, 1996). After another day, the levels begin to climb again back towards the peak (Hagedorn, 1996). These results are summarized in Figure 4.7.

I decided to evaluate the effect of JH through small RNA cloning. Analysis of early trypsin (ET) expression, known to be induced by JH, suggested that methoprene induction in our abdominal ligation experiments has been weak, so I used 12 hr samples prepared by the Zhu lab that had previously exhibited ET strong induction, for small RNA cloning. Analysis of ET expression suggested a 26.6 fold induction, comparable with strong induction by methoprene at 12 hours post-induction observed in Noriega et al., 1997.

BLAST results of small RNA cloning were compiled in a table (Table 4.2) and examined for the fold change in miRNA expression for known miRNAs. To confirm expression changes, northerns were run, showing that higher expression changes in the three to four fold range were detectable for miR-1891 and miR-9a whereas lower fold changes were not. It may be that, like a microarray, fold changes in sequencing reads of less than two fold may not be accurate. A replicate sample was examined and observed to demonstrate similar changes in expression by northern though not as dramatically.

MiR-9a regulates sensory organ and wing development in flies, with upregulation of miR-9a resulting in upregulation of *apterous*, necessary for wing development

(Biryukova et al., 2009), so it is logical that methoprene administration would decrease the response in adult females, as these functions may not need to be carried out in adult females. Indeed a drop of miR-9a is observed in adult females compared to other timepoints when examining lifestages (Mead and Tu, 2008). In the case of miR-1891, the opposite pattern is observed, an increase in expression in adult females (Mead and Tu, 2008), when JH levels should be increasing, a pattern which agrees with the results of small RNA cloning for methoprene administration. Unfortunately, I do not have functional information on this important miRNA to examine its role. Cloning revealed seven additional miRNAs with greater than two fold expression change and more than 30 hits per miRNA in each condition. These are summarized in Table 4.2. These results suggest that JH induces a cascade effect upon gene regulation through miRNAs, as nine individual miRNAs were observed to be affected, which in turn will likely regulate hundreds to thousands of transcripts in the mosquito.

Despite the limitations of quantitative reliability, it is interesting to note that the expression of the four miRNAs observed by Sempere et al. to change expression in *Drosophila* cells in response to methoprene administration were all observed to change in expression in response to methoprene in female adult mosquitoes, and expression changed in the same direction, either increasing or decreasing in expression as in flies. In flies, miR-34 was observed to increase in expression (Sempere et al., 2003), and in mosquitoes, small RNA cloning indicated a 2.27 fold increase in miR-34 expression in response to methoprene. MiR-100, -125 and let-7 were all observed to decrease in expression in flies. In mosquitoes, decreases in expression were observed as well (1.28,

1.75, and 1.20 fold decreases respectively). However, these results should be taken with a high degree of caution as they are below the threshold of quantitative reliability that was established by northern analysis.

Summary and Future Directions

In summary, miR-989 is a miRNA that is conserved broadly across the holometabolous insects but does not appear to be present in other organisms. The expression of miR-989 in mosquitoes clearly followed a pattern of rising and falling in response to key events in the female mosquito life such as adult emergence and blood feeding, both of which are tied to changes in ovarian development, which agrees with previous expression analyses in the literature (Mead and Tu, 2008). However, a link to a key reproductive hormone, JH, could not be established beyond a weak fold-change according to small RNA cloning. As miR-989 is also expressed in non-bloodfeeding insects, it would be interesting to examine whether miR-989 also follows a cyclical pattern of rising and falling during ovarian development in these insects. *In situ* hybridization would be a good method to determine the spatiotemporal profile of miR-989 in the ovaries, to further narrow down the regulation and function of this miRNA. Nine known miRNAs appear to be affected by methoprene treatment, with up to 3.87 fold induction, and 3.15 fold suppression, providing the first glimpse at the cascade effect set off by JH through miRNA expression in mosquitoes. This analysis provides the foundation to study JH regulation of miRNAs in mosquitoes.

PCR-based target verification is an attractive method to provide experimental evidence for miRNA targets. Direct high-throughput sequencing of a cDNA sample prepared for PCR-based target detection would potentially provide an extensive catalog of all the targets of miR-989 as well as targets for other miRNAs present. This would provide experimental evidence for the function of a large category of the miRome, the collective miRNAs in the sample.

4.6. Acknowledgements

Except where noted below, Edward Andrew Mead conducted the experiments and analyses, and wrote the manuscript.

Jeff Busche carried out abdominal ligation for the samples used in small RNA cloning experiment, conducted TRIzol-based total RNA isolation upon these abdominal ligation samples, and upon a second set of samples for use in replicate northern analysis. JB provided Zhu lab protocols for abdominal ligation, qRT-PCR, and trizol isolation.

Xing Zhang conducted qRT-PCR analyses for the expression of early trypsin in the abdominal ligation samples examined.

Jinsong Zhu provided direction regarding abdominal ligation and JH induction and provided use of his laboratory facilities and supplies. JZ also provided an initial set of

abdominal ligation samples for preliminary examination of miR-989 expression by northern blot (not shown).

Zhijian Tu provided project direction, funding and assisted in revision of the manuscript. In addition, ZT provided an initial bioinformatic analysis of the small RNA cloning results for the abdominal ligation experiment.

We would like to acknowledge Kevin Myles for his insightful comments on miRNA expression in abdominal ligation samples, Randy Saunders for maintaining mosquito colonies, and Jim Biedler for discussions about bioinformatic analyses. Yumin Qi provided advice regarding qRT-PCR.

AAEL002261-RA, expected product: ~950 bp primer 1; ~900 bp primer 2

GTCGATCGGCTATCTGCCCTGCAAGAAGATTCTCGGACTCAGCAAGCTGGCCAGAATTGT
GGAAATCTTCTCCCGCCGGCTGCAGGTCCAGGAACGCCTCACCAAACAGATCGCCGTGGC
CGTTACGCAAGCCGTTCAACCCGCCGGTGTGGCCGTCATCATTGAGGGAGTGCACATGTG
CATGGTGATGCGAGGCGTGCAGAAGATCAACAGTAAAACCGTCACCTCGACGATGTTGGG
CGTCTTCCGGGACGATCCGAAGACCCGCGAGGAATTCCTTACGCTGTGTAACAACAAGTA
GGAAAAGTGCGAGCAACTTAATCGAATCCCGCATGAATGGTGCCACGAGGGCAGCGATAA
TTCTGCTGGCATACCTTTTATCCGTTTAACTCTCTTCTAGTGTAATGTGTATAACCCAGC
TATTAGGTGTCTAAGTCGAAAGCAACTGGAAACTCTGTTATCGGAAGTCCAAATATGTGT
GCGTGTGTAAAAACTGTTATTTCCTGGAAAATATTAATTGTTGATGTTGTCAAAAGAAAG
TAAAATTAATCAAATGAGATAACGAGAGATTGAGAACGAGTGAGTAAAGCATACAAAAAT
TTACAAAATATATGAAACACATGATAGAACCAGTTTTCCGAGGGGCCAAGCTCTTACCAA
GTAGCGATCGCGGAAACAATCATACTCTCAACGAAAATTGCCAATAATTTTAAATACCTT
AGGCGTGTAAGCATACAAGGGACATTTTAACTCTAGGCGAACGAGAAATACGATTTTAGC
GAAACGCAGCAATAGCATCACGTATTTTAGAGGAAGTATTTTACACACACTAGCCCAGTT
GTACTTGTTGTTACTCACAAAAAGATTAAGCGAAAAAATTTGAATACTTTGTTAAGGTTG
ATATTGTTGATCGGAGATACGTTTTACAAAATAAAAAAAATATGTTTGTTACATTCTGAT
TTTATCATTGAAATGTCATATCACGACAAATCGTGAATCGTTGATTATCGGTTTTGAATC

AAEL000246-RA, expected product: ~300bp primer 1; ~250 bp primer 2

GCTGAATCACCGTACATGTCAAAGGAGCGTATAGCTCCAGACTTGTACAGTCCTCGGACT
GATCCACATGGATCATTCAAGTCCAGTGTACAATCACTATTGAAAAATGATTATCAACGT
CACAAACAGGGGGCTGATTCTGACCGTATGTCAGAAGGGTCGGATAAAAATCCATACAAT
TTCCCGTATACAGCGGAAGAAGACCATTCAAGTCACACAACAATGAAATCCAACGACGGA
GCTGGTCAAATAATGCATCGGTTTAACAACCCTACTTCAATACAGCCAATGGCGCCTTTA
ACGAATTTGAACGATAATAACGAAGGACCTAACGATGACGACGAGACAACTGTATGAGAT
CCAACGGTTTGGCGTAGCATTCCTCACATCACAAGTGATTTCAAAAGCACAAATTCTATT

AAEL013816-RA, expected product: ~850 bp primer 1; ~650bp primer2

CGAAAAACGTCGGGGTCACTAACTCTCCGGTGCGGAAGAATTCCTGCGAACAGATGAACT
CCCCGAGGGCCACAGCCGGAGGACAAGGTCCACTACCGCTCAACTACGTGACCAGCGGGT
ACGGTTCATCATCTAGTGCTACTCAATTGATTCATAACATGTCCATAAGTCCACAAACTG
CGCATCTCTCCGGTCGATCGAACTATACCCAGCAATCGTCGCCGCAGCATTCGACGGGCT
CATCCAGCTGGGGACAACAATCGTCGCAATACTACCGATAGGTTCGTAGCAGCCCTAGGA
AATCGCCGTTTCTTTCGAAAGACGAACGTCACAGTAGAAGTAACAATCATTTCTCTGATT
ATTTATACACTAAGAATAATGACATTGCAGCAAATACATGTTTTAGATAAATCTAATCTA
AACTAACTGTCCTACTCTCTGAAATCTAATACTATTTGTTCATAAGAAAGCCGCCACCCG
AAATCCGCAACCGCGAGAGATTCACCTTCCTAAATTCCATATGTAAAGGAGTGTTTCCAC
ATCAAGTTGGTTAATCTTGCGTTTAGTATCATAAGGGCGTAACTGCAATGATCTATCTCT
CTTCAGGTAACGATCGCTGCATATACGCCTGTATGATGCTAAACGCGAGAAATGATGTAT
CGAATTACAAACACTCTTTGTCAATAATGGAAGTCATAGATCTAATCTATTATTCGACAA
TCATAAACTGTTTCCATTTACTTTCATCTATATGCGTACAGTGCCGCATACGAATTTGAG
TCACTGTACCAACGCTATAATCCTCGCCAGAATATTTTCCACATCGGCTCCACTATCGTT
CTGAACCAACAGATCCCATTCCGTTACATCGTCCAGATCGCACTCCGACTGAACGTCATC

Figure 4.1. Location of primers for PCR-based target analysis. cDNA sequences from Ensembl correspond to the terminal protein coding region and portion of the 3'-UTR of the transcripts listed. Protein coding regions are designated with blue letters. The reverse primer corresponds to miR-989, which interacts with the shown target (aquamarine). The miR-989 primer is based upon the sequence of miR-989 from the miRBase registry. However, two nt are deleted from the 3' end, to allow for recognition of predicted target pairings as detailed in Andachi et al, 2008. Forward primer one (red) and forward primer two (green) locations are detailed for each targeted transcript. Forward primer one is used in the first round of nested PCR whereas forward primer two is used in the second round.

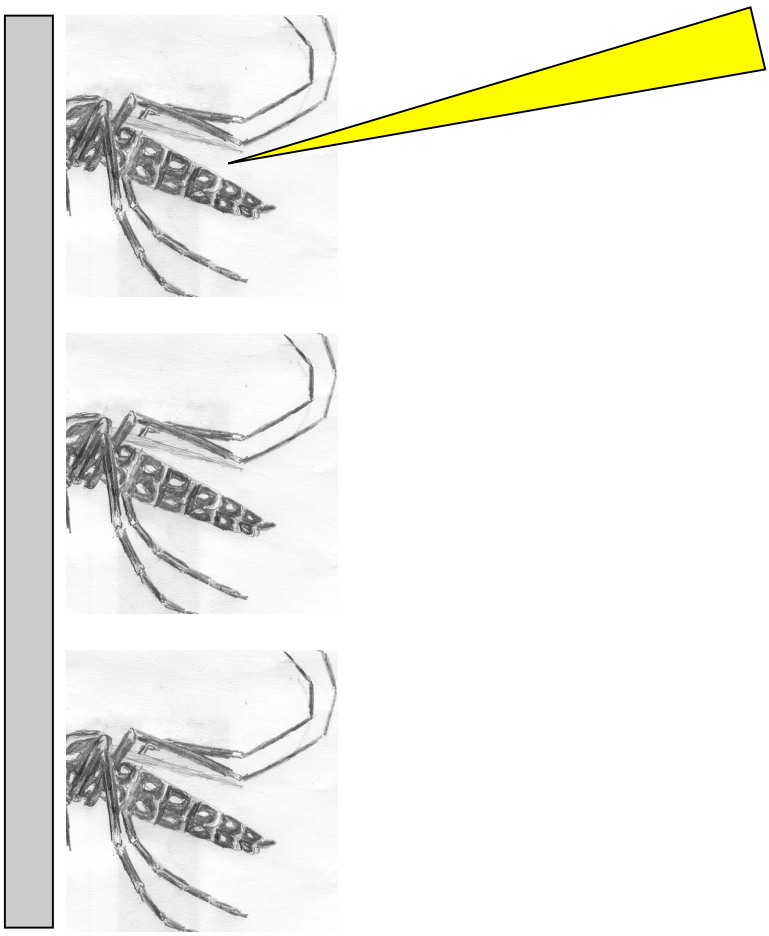
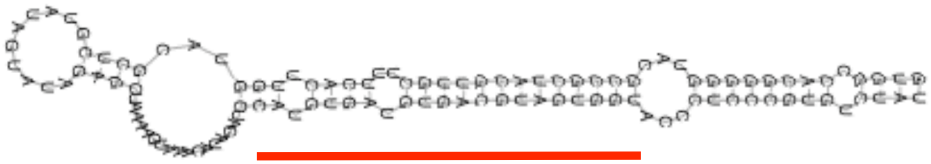


Figure 4.2. Chemical treatment following abdominal ligation. Female *Ae. aegypti* (Liverpool) were decapitated at 0-1 hours following pupal emergence, and the remaining thorax was affixed to a slide using Vetbond (3M, St. Paul, MN, USA). A set was comprised of 8-16 mosquitoes on a slide. Each set was treated with either 500 ng of methoprene, a JH analog, diluted in one µl acetone, or else an equal volume one µl of acetone. For treatment, the solvent was applied directly to the abdomen. The sets of samples were stored under humid conditions at 27°C after which total RNA isolation by

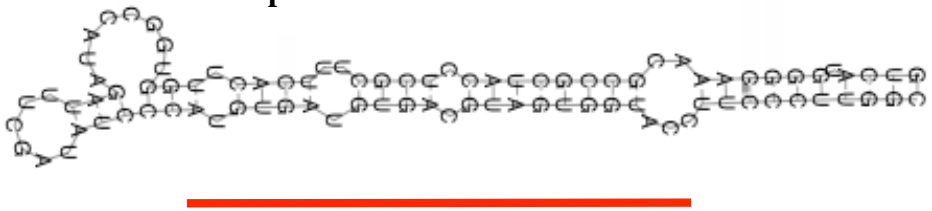
miRVana kit was conducted. RNA was examined for the expression of early trypsin, a known marker of JH induction. Mosquito images were drawn by EAM based upon an image at http://neeladri.files.wordpress.com/2006/10/aedes-aegypti.gif

Figure 4.3.A

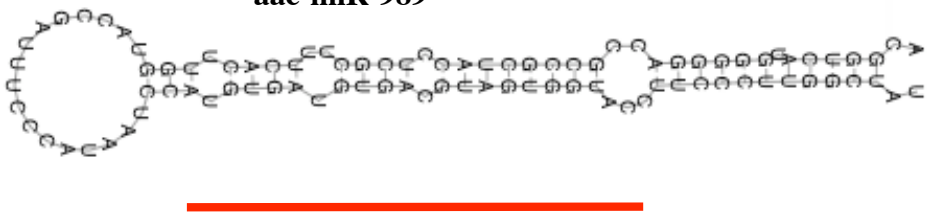
aga-miR-989



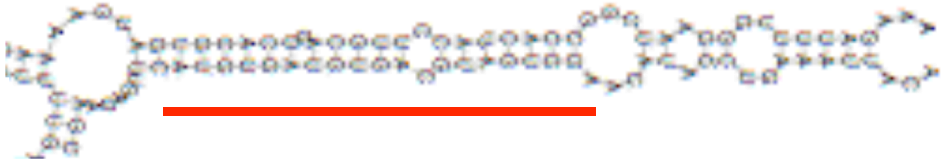
cpi-miR-989



aae-miR-989



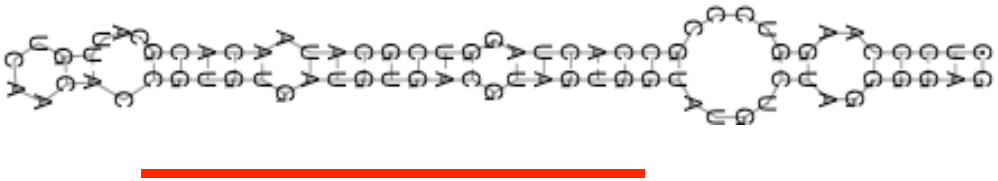
dme-miR-989



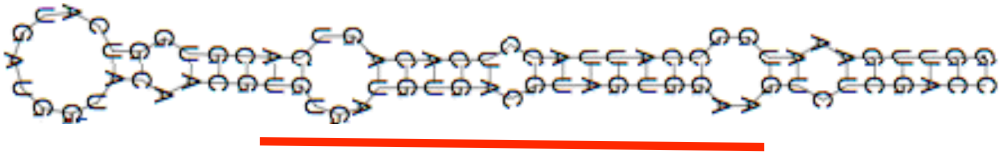
```
An.gambiae      -GTGGCCA-CGGGGGTA-CG-CCGCTACGT
Ae.aegypti      ACGGGTCA-TGGGGG-ACCG-CCGCTACCT
Cx.pipiens      -----TCA-T-GGGGAAACG-CCGCTACCT
D.melanogaster A---AAGATTTTGGGAATCGGCCACTACCT
                       *    **  * ** ** ******

An.gambiae      TGCTTTCA--------------------
Ae.aegypti      CGCTTTCA--------------------
Cx.pipiens      CGCTTTCA--------------------
D.melanogaster TGCAGTCACGTGATGAAAAGACACAGGTGA
                ** ***

An.gambiae      ----------------------------
Ae.aegypti      ----------------------------
Cx.pipiens      ----------------------------
D.melanogaster CACTGATCCGGATTTGGTAGTTGACAAATC


An.gambiae      ---CTTGGTA-CGCTCGTATAGTATAGAGC
Ae.aegypti      ---CTTGGTA-C--CGATTTCCCATA----
Cx.pipiens      ---CTTGGTGGCCATAGATTTCGATA----
D.melanogaster CTCCATGCCGAGATTAGTTTCATTTTGCGT
                  * **            *    *

An.gambiae      TAAAAGTAAACATA-GATCCATGTGATGTG
Ae.aegypti      ----------------ATCCATGTGATGTG
Cx.pipiens      ----------------TCCCATGTGATGTG
D.melanogaster CTTTTGAATTCGAATAGTTCATGTGATGTG
                                ***********

An.gambiae      ACGTAGTGGTACCCTCCCGTGTCTAT-
Ae.aegypti      ACGTAGTGGTACCTTCCCTTGGCTAT-
Cx. pipiens     ACGTAGTGGTACCTTCCCTTGGC----
D.melanogaster ACGTAGTGGAACATACCTGAAATTACA
                ********* **   **
```

Figure 4.3.B

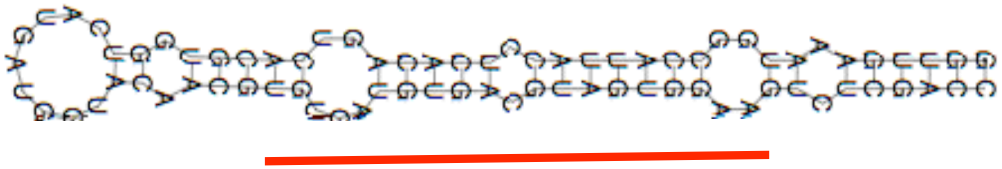
tca-miR-989



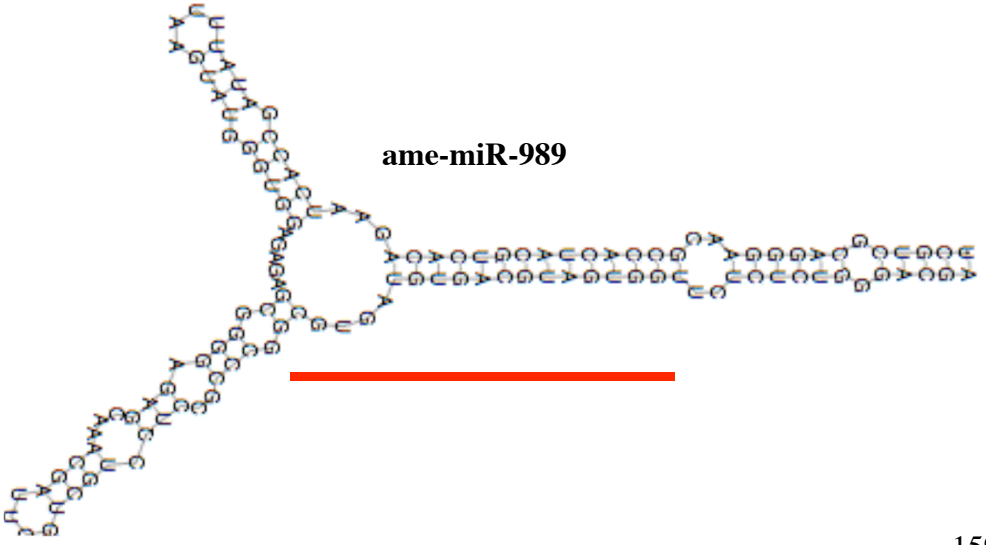
bmo-miR-989-1



bmo-miR-989-2



ame-miR-989



```
A.mellifera     TCGTCGCAGGGAA-----CGCCACTACGTC
T.castaneum     ------CTCCCAAGGTCCCGCCACTAGGTC
B.mori_1        ------GGTTGAAA---TGGCCATTACCTC
B.mori_2        ------GGTTGAAA---TGGCCATTACCTC
                            **      **** **  **

A.mellifera     CACAGAATCACCGATATTTAAGTATGGGTG
T.castaneum     -GCATAA-CACGCATTGGTCA---------
B.mori_1        -ACA--GTCACGTGGTCATGATGGT-----
B.mori_2        -ACA--GTCACGTGGTCATGATGGT-----
                  **    ***       * *

A.mellifera     GAGAGAGCGGGGAGAGCAAACGATTCGTCG
T.castaneum     -------------------------------
B.mori_1        -------------------------------
B.mori_2        -------------------------------

A.mellifera     TTCCTCCGCCCGGCGTGATGTGACGTAGTG
T.castaneum     -------ACACCGTGTGATGTGACGTAGTG
B.mori_1        -------ACAACGTGTGATGTGACGTAGTG
B.mori_2        -------ACAACGTGTGATGTGACGTAGTG
                        *   * ****************

A.mellifera     GTTCTCTCTGGGACGA
T.castaneum     GTATGTCTAGGGGAG-
B.mori_1        GAAGTCTCGACC----
B.mori_2        GAAGTCTCGACC----
                *
```

Figure 4.3. Conservation of pre-miR-989 among holometabolous insects. Figure 4.3.A shows conservation of hairpins among holometabolous species within the order Diptera. Figure 4.3.B shows conservation of hairpins among holometabolous species outside of the order Diptera. Hairpin structures for miR-989 hairpins were generated using RNAfold at http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi. Hairpins for miR-989 in *An. gambiae* and *Ae. aegypti* were from Mead and Tu, 2008. The sequence for *D. melanogaster* pre-miR-989 was found on miRBase (http://microrna.sanger.ac.uk/sequences/) and re-folded for the figure. A partial hairpin is shown for dme-miR-989 to focus on conserved regions as it is an unusually lengthy stem loop structure. The entire hairpin sequence is shown in the table. The sequences for *T. castaneum, A. mellifera*, and *B. mori* miR-989 were discovered by blast using aga-miR-989, and surrounding genomic sequences were examined using Ensembl. Clustal W was used for alignment (http://align.genome.jp) of sequences and the alignment was manually corrected where needed. *B. mori* showed two identical hairpin structures from different scaffolds though flanking regions exhibited only partial conservation (not shown). The miR-989 is indicated by a red line, and conserved nucleotides are indicated in the table by a star (*). Nucleotides which differ from aga-miR-989 or dme-miR-989 are indicated in red in figure 4.3.B.

Figure 4.4.A

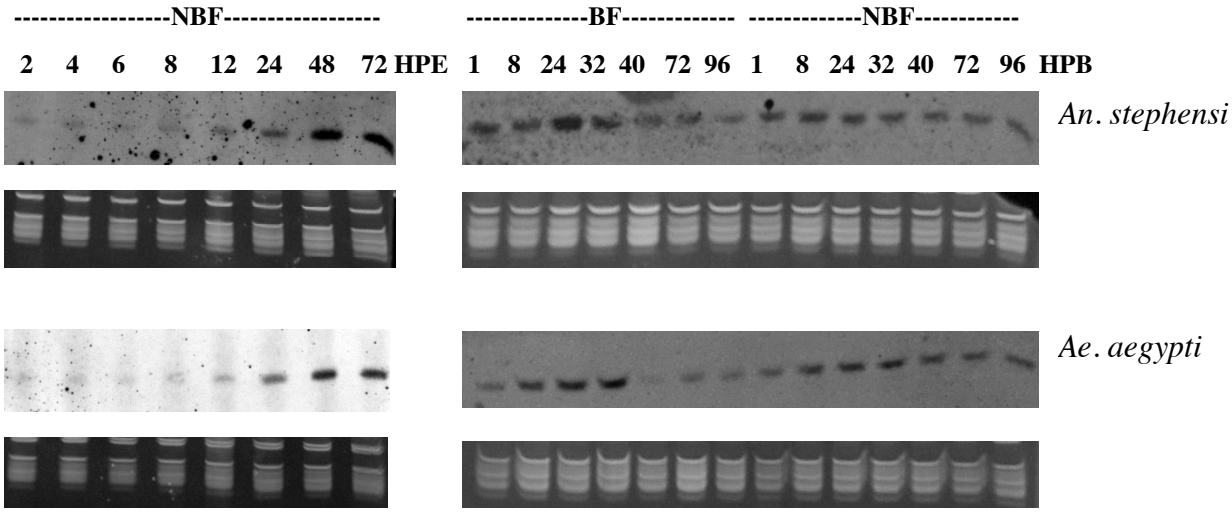


Figure 4.4.B

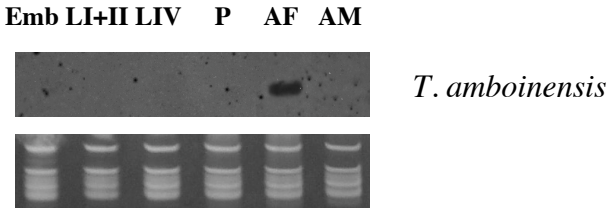


Figure 4.4. Expression of miR-989 in the three subfamilies of mosquitoes. The presence of miR-989 was examined in a time course in *An. stephensi* and *Ae. aegypti* females following pupal emergence, and following bloodfeeding (BF), top panels (Figure 4.4.A). Numbers below the headings indicate the time in hours post-eclosion (HPE) or hours post-bloodfeeding (HPB), as appropriate. Expression of miR-989 increased following eclosion to a peak at 48 HPE. Expression remained strong following bloodfeeding,

though by 32-40 HPB, a drop in expression of miR-989 was observed. By 96 HPB, the expression of miR-989 had begun to return. These results can be compared to NBF (non-bloodfed) control mosquitoes from the same cohort, which do not exhibit a decrease in miR-989 expression over the same period. In Figure 4.4.B the expression of miR-989 was examined across the lifestages of *T. amboinensis* (top panel). As with other mosquito species (Mead and Tu, 2008), miR-989 was detected almost exclusively in the adult female. In both figures, ribosomal and tRNA of each sample is shown for the PAGE gels prior to transfer to nylon membranes showing that equal loading occurred.

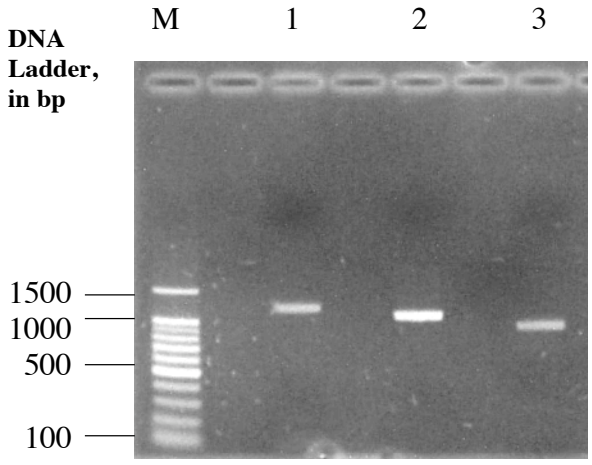


Figure 4.5. PCR-based miR-989 target analysis. Examination of 20 targets for miR-989 was conducted based upon a protocol in Andachi, 2008. I relied upon endogenous miRNAs as primers for cDNA synthesis from a female *Ae. aegypti* abdominal extract. After cDNA was obtained, I conducted two rounds of nested PCR to examine the presence of individual targets predicted to have high scores by miRanda for aga-miR-989. Shown above are the samples showing a prominent PCR product from nested PCR reactions run adjacent to 100 bp markers (Promega) on a 1% agarose TAE gel. Marker (M) sizes increase in 100 bp increments to 1000, with a 1500 bp band shown at the top. From left to right, the lanes are labeled 1-3, and correspond to the primers used in nested PCR to exons in the following transcripts 1. AAEL002261-RA, 2. AAEL000246-RA, 3. AAEL013816-RA. Predicted sizes are 900 bp, 250 bp, and 650 bp respectively. Observed products are 1.2 kb, 1.1 kb, and 1 kb respectively.

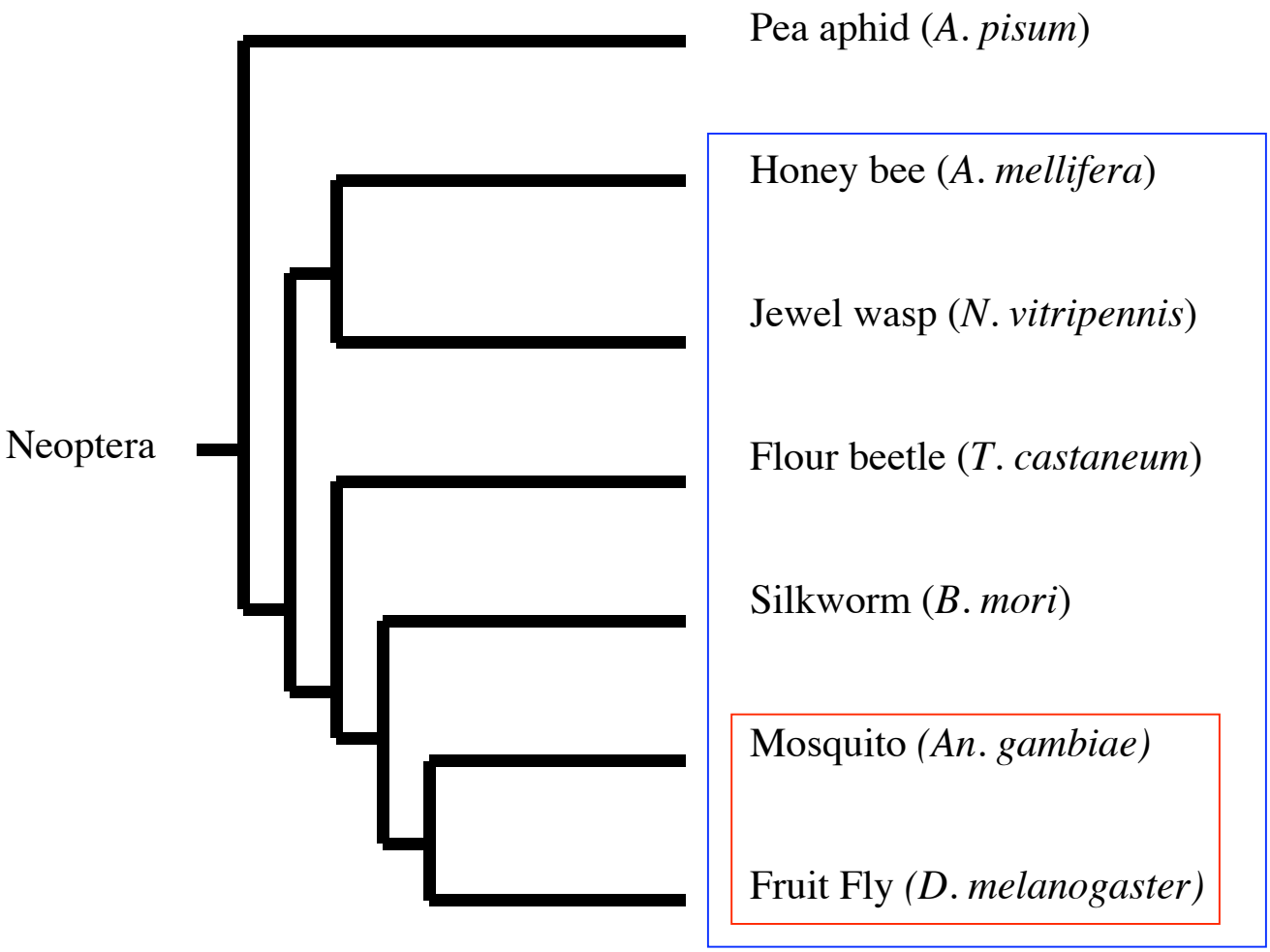


Figure 4.6. Phylogenetic relationships among neopteran insects. Based upon Savard et al., 2006. Branch lengths above do not reflect divergence times or distance. Relationships were determined by Savard et al., through 18s RNA sequence homology. MiR-989 was found in all branches of holometabolous insects (blue box) but not in hemimetabolous insects, suggesting it is unique to holometabolans. It was not found within wasps, suggesting that at least some members of the order Hymenoptera have lost this miRNA.

It was present among all seven members of Diptera examined (four flies, three mosquitoes; red box).

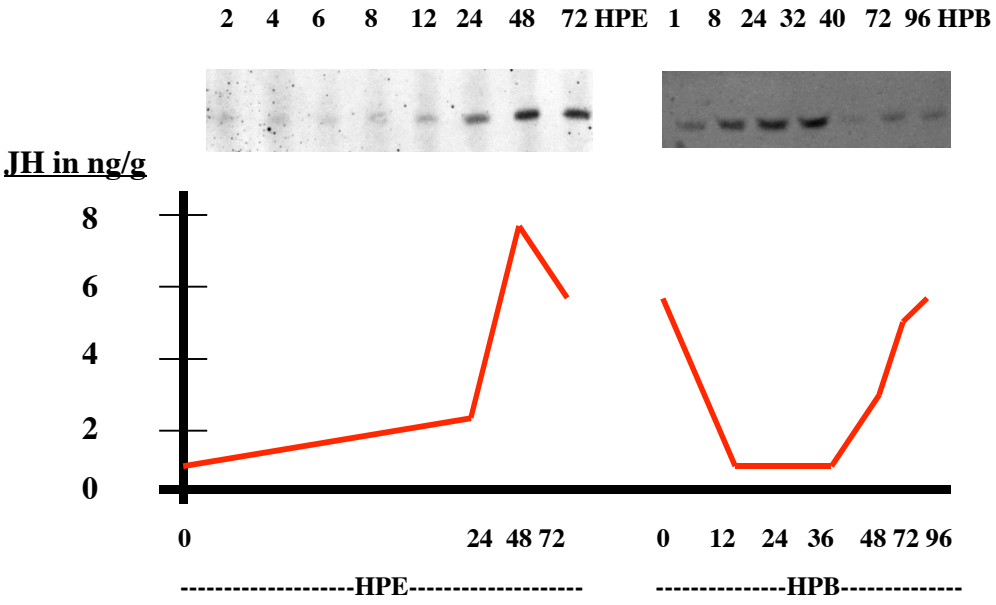


Figure 4.7. Comparison of JH and miR-989 expression in *Ae. aegypti* females. Graph of JH expression is based upon Hagedorn, 1996. The vertical axis of the graph is listed in ng/g mosquito tissue. The horizontal axis displays Hours Post Emergence (HPE), and Hours Post Bloodmeal (HPB). Contrast with the observed expression of miR-989 in *Ae. aegypti* females (top panel). Times listed above the panel correspond to hours post emergence and hours post bloodmeal. For HPE (left panel), non-bloodfed females were used, for the right panel (HPB), bloodfed mosquitoes were used. Expression post-emergence tracks very closely with the expression of JH, whereas expression post-blood meal shows a delayed expression change compared to the expression of JH at the same time, though the changes are in the same direction.

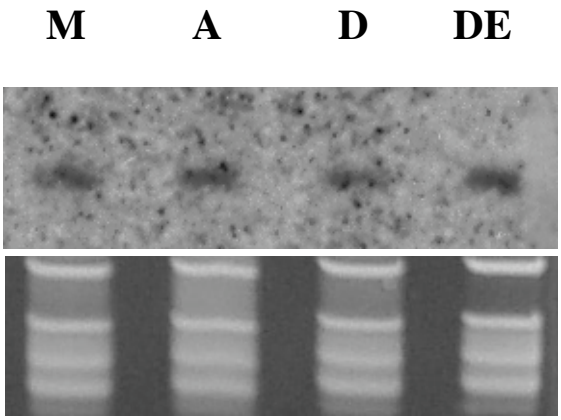


Figure 4.8. Expression of miR-989 does not respond to the JH analog, methoprene. A northern was conducted using *Ae. aegypti* females 0-1 hours post-eclosion in an abdominal ligation experiment examining treatment with the 0.5 ng of the JH analog, methoprene (M) in one µl acetone applied to the abdomen as detailed in materials and methods. All mosquitoes were from the same cohort. As a control, a cohort was treated with one µl acetone (A). Further controls included mosquitoes that were decapitated at the 0-1 hours post-emergence but not chemically-treated (D), and mosquitoes which remained untreated in any way until immediately before freezing (four to five hours post-eclosion), at which time they were immediately decapitated and frozen (DE). All samples were quick frozen in liquid nitrogen at four to five hours post-eclosion and stored at –80°C until processed for total RNA. As shown, miR-989 expression remained the same, regardless of treatment.

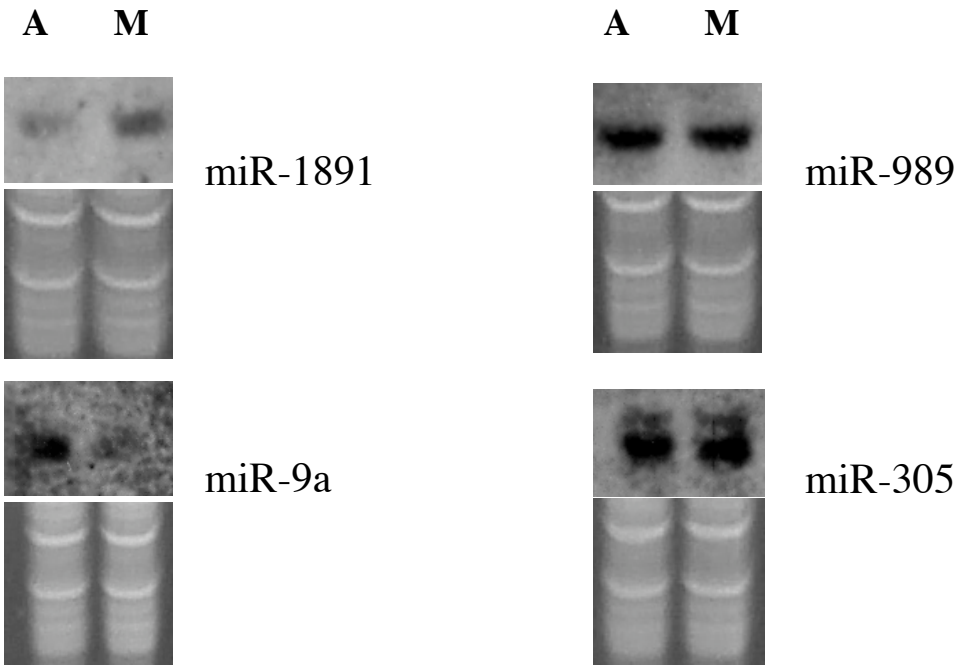


Figure 4.9. Northern blot analysis of four miRNAs exhibiting changes in expression according to small RNA cloning. Four miRNAs of interest, miR-1891, -9a, -989, -305 from small RNA cloning of abdominal ligation samples were examined by northern analysis. In the figures above, A and M refer to the Acetone only sample, and the Methoprene treated sample, respectively. At the top of each panel, the results from northern using the listed antisense miRCURY dig-labeled probe to the indicated miRNAare shown. Beneath, are shown the tRNA and small ribosomal RNA of the gel that was transferred to the northern membrane, exhibiting equal loading, and intact bands. Analysis was carried out using five μg total RNA per lane. Expression of miR-1891 and

miR-9a are shown to exhibit upregulation and downregulation, respectively, going from A to M treatment, correlating with the results from small RNA cloning, where three to four fold changes were observed for these two miRNAs. MiR-989 and –305 were also examined and do not clearly demonstrate a change in expression. These two samples were mildly upregulated among the small RNA clones. This experiment was repeated, and though the northerns demonstrated the same trends, the observed change was a little less strong for miR-1891.

Target Identity	Protein Ortholog	Putative Function	Score	Binding Energy (kcal/mol)	Query Match	Primer F1 5'→3'	Primer F2 5'→3'
AAEL002261-RA	GTP cyclohydrolase I	Pteridin biosynthesis, metabolism	120	-23.64	2 21	ATCTGCCCTGCAAGAAG	CGCCTCACCAAACAGAT
AAEL007316-RA	conserved hypothetical protein	LIM, zinc-binding Transglutaminase-like	120	-22.82	1 21	GGACTTTATGGCAACCCTA	CCTCAAGAGGGCCAGTA
AAEL012819-RA	vacuolar ATP synthase subunit g Q1HQI4_AEDAE	vacuolar ATP synthase subunit g Necrosis	117	-25.50	1 21	ACCGAAGAGATCGAGAAG	TCAAGGATGTGCTCGAG
AAEL000246-RA	cadherin	Transmembrane adhesion	115	-20.95	1 21	CTCGGACTGATCCACAT	CGTATGTCAGAAGGGTCG
AAEL013968-RA	conserved hypothetical protein	tRNA methyltransferase	115	-21.26	1 21	GGAAGATGTAACGCAGG	CAGTGAAAGAAGAGGCG
AAEL003059-RA	conserved hypothetical protein	Rab GTPase activator	114	-22.46	4 21	CAAACTGCGAGATCAGG	ATTTCCACTGCCAGACG
AAEL010506-RA	GTP-binding protein alpha subunit, gna	Protein signaling Programmed cell death	114	-22.00	1 21	ACTCGTCGGTAATCCTGT	ACCACAAAGAGACGCAAT
AAEL011054-RA	hypothetical protein	Unknown	113	-23.67	1 21	ATGACACAGCAACCAAAC	GAACATCTCGCAGATGC
AAEL012211-RA	microfibrillar-associated protein, putative	Transcription factor	113	-24.46	1 20	GCGCCATTACTAGGCTT	GAAGACAGAGGTAACCCG
AAEL013816-RA	conserved hypothetical protein	Zinc finger ring domain AT-hook DNA binding domain Postsynaptic domain	113	-22.27	2 21	TCACTAACTCTCCGGTGC	CCGGTCGATCGAACTAT
AAEL004610-RA	hypothetical protein	(Low Complexity)	112	-21.66	2 21	ATCTTTCGGGATTGCTTC	ACATTAACAGACGGAGGC
AAEL009375-RA	plekhh1	Development- axon guidance	112	-23.80	4 21	GCTCCTCATAGCGGACT	ACGCTCACTTCCCACTC
AAEL003789-RA	exportin, putative	Protein transport	111	-20.48	1 21	GTTTCTGGCTGTGCTGT	CAACCTTTGCTCACCGT
AAEL012166-RA	conserved hypothetical protein	Development Protein transport	111	-22.32	2 20	GATGCCATTACGGCTAAC	GCGAAAAGGTATAAATGTGATAA
AAEL012166-RB	conserved hypothetical protein	Development Protein transport	111	-22.32	2 20	GATGCCATTACGGCTAAC	CTATAATTTCCGGGAACG
AAEL012166-RC	conserved hypothetical protein	Development Protein transport	111	-22.32	2 20	GATGCCATTACGGCTAAC	CTATAATTTCCGGGAACG

AAEL002809-RA	down syndrome critical region protein	Vacuolar transport	110	-22.83	1 21	CTGCTGAAAGCCGAGAT	ACCGAGATCCAGAACATT
AAEL002809-RB	down syndrome critical region protein	Vacuolar transport	110	-22.83	1 21	CTGCTGAAAGCCGAGAT	ACCGAGATCCAGAACATT
AAEL005677-RA	hypothetical protein	Zinc Finger, C2H2-like	110	-22.94	3 21	GACCAACCCTCAATCAA	TAACTCCGGTGACGACC
AAEL014377-RA	conserved hypothetical protein	BRCT domain- cell cycle FHA domain- phosphopeptide recognition	110	-20.39	1 21	GTGCGAGTCAAGTGCAT	AACTACAAACCGCAGCA

Table 4.1. Top targets in *Ae. aegypti* for miR-989 as determined by miRanda. Target identities are the Ensembl ID for a given target transcript. Protein orthology was predicted based upon Ensembl, Vectorbase, and Flybase. Putative Functions were also determined at the same websites. Query match indicates the nucleotides in common between the target and miR-989, beginning at the 5'end. The seed region is nt 2-8. MiRanda scores and binding energy give an estimate of the degree of complementarity, with higher scores and lower binding energy indicating more favorable pairing. Based upon single hit scores of 110 or above, twenty targeted mRNA were selected for PCR-based target analysis. Initial primer design was carried out using Primer 3, v0.6 on the San Diego Supercomputer Center Biology workbench (http://workbench.sdsc.edu) at UC San Diego and ordered from Sigma (St. Louis, MO). Forward primers are shown above. As nested PCR was carried out, first round forward primers are designated as Forward Primer one primers, and second round nested primers are designated as Forward Primer two primers. Reverse primers were miR-989 for the initial cDNA synthesis, followed by a truncated miR-989 oligo, with the terminal 3' 2 nt removed to demonstrate that amplified products arose from miR-989.

miRNA	Meth	Acet	Fold Change
aae-miR-219	3	0	N/A
aae-miR-1000-1	8	1	8.240696627
aae-miR-1000-2	8	1	8.240696627
aae-miR-252_star	6	1	6.18052247
aae-miR-304	4	1	4.120348314
aae-miR-M2	4	1	4.120348314
aae-miR-1891-2	252	67	3.874357369
aae-miR-1891-1	252	67	3.874357369
aae-miR-210	11	3	3.776985954
aae-miR-87	51	15	3.502296066
aae-miR-10	173	56	3.182233296
aae-miR-282-1_star	3	1	3.090261235
aae-miR-282-2_star	3	1	3.090261235
aae-miR-92a_star	26	10	2.678226404
aae-miR-305_star	682	287	2.447802744
aae-miR-988	97	44	2.270873786
aae-miR-M3	5159	2341	2.270063749
aae-miR-13b	187	91	2.116772348
aae-miR-981	4	2	2.060174157
aae-miR-133	4	2	2.060174157
aae-miR-965	4	2	2.060174157
aae-miR-305	1043	560	1.918537183
aae-miR-11	384	208	1.901699222
aae-miR-996	1149	635	1.863889847
aae-miR-79	120	68	1.817800727
aae-miR-281	181	103	1.810153021
aae-miR-8_star	3855	2510	1.582066011
aae-miR-308	211	143	1.519918696
aae-miR-2-2	930	638	1.501537591
aae-miR-2-1	933	642	1.496995707
aae-miR-2-3	933	642	1.496995707
aae-miR-316	57	40	1.467874087
aae-miR-989	5871	4182	1.446112204
aae-miR-286a-2	7	5	1.44212191
aae-miR-286a-1	7	5	1.44212191
aae-miR-7	168	123	1.406948205
aae-miR-9c	23	17	1.393647224
aae-miR-285	8	6	1.373449438
aae-miR-375	132	100	1.359714943
aae-miR-184	28565	22021	1.336198964
aae-miR-281_star	20906	16476	1.307052711
aae-miR-14	60	49	1.261331116
aae-miR-252	276	226	1.25798245
aae-miR-306	49	41	1.231079679
aae-miR-279	613	514	1.228489064
aae-miR-275_star	182	162	1.157258323
aae-miR-N2	123	119	1.064711854
aae-miR-1175	86	84	1.054612961
aae-miR-3-2_star	6	6	1.030087078
aae-miR-12	392	399	1.012015375
aae-miR-283	233	239	1.004227152
aae-miR-998	95	98	0.9985538

aae-miR-317-1	459	474	0.997489386
aae-miR-317-2	459	474	0.997489386
aae-miR-193	17	18	0.972860018
aae-miR-92a	152	163	0.960571999
aae-miR-276-2	8259	8890	0.956972911
aae-miR-276-1	8259	8890	0.956972911
aae-miR-92b	153	165	0.955171654
aae-miR-1175_star	653	716	0.939450925
aae-miR-N1-2	255	284	0.92490213
aae-miR-N1-1	255	284	0.92490213
aae-miR-190	412	461	0.92059843
aae-miR-124	211	237	0.917081745
aae-miR-275	35123	40550	0.892225609
aae-miR-137-2	12	14	0.882931781
aae-miR-137-1	12	14	0.882931781
aae-miR-8	6780	8069	0.865533572
aae-bantam	1097	1350	0.83704113
aae-let-7	13655	16959	0.829402621
aae-miR-iab-4-1	31	40	0.798317486
aae-miR-iab-4-2	31	40	0.798317486
aae-miR-970	538	702	0.789439955
aae-miR-315	16	21	0.78482825
aae-miR-100	859	1131	0.782356145
aae-miR-M4b_star	3	4	0.772565309
aae-miR-M4b-x1	24	32	0.772565309
aae-miR-34	349	470	0.764894448
aae-miR-263	3282	4565	0.740579582
aae-miR-1	1126145	1588166	0.730419498
aae-miR-33	183	259	0.727822144
aae-miR-308_star	520	742	0.721893909
aae-miR-11_star	38	56	0.69898766
aae-miR-957	49	73	0.691428313
aae-miR-71	1307	1956	0.688304607
aae-miR-71_star	4	6	0.686724719
aae-miR-31	4707	7279	0.666110713
aae-miR-993	47	74	0.654244496
aae-miR-277	1178	1860	0.652388483
aae-miR-263_star	46	73	0.649095967
aae-miR-999	92	146	0.649095967
aae-miR-278	19	31	0.631343693
aae-miR-927	11	18	0.629497659
aae-miR-2-2_star	3	5	0.618052247
aae-miR-125	198	356	0.5729136
aae-miR-1174	59	113	0.537833076
aae-miR-286b	1	2	0.515043539
aae-miR-1890	65	135	0.495967853
aae-miR-9b	26	55	0.486950255
aae-miR-282-1	11	25	0.453238314
aae-miR-282-2	11	25	0.453238314
aae-miR-263b	55	139	0.407588412
aae-miR-2929_star	1	3	0.343362359
aae-miR-9a-2	1186	3873	0.315435909
aae-miR-9a-1	1186	3873	0.315435909
aae-miR-932	1	4	0.25752177
aae-miR-988_star	0	1	0
aae-miR-M4a_star	0	1	0

Total No. miRNAs	**1293044**	**1758421**
Total No. Reads	**8096499**	**8340099**

Table 4.2. Expression changes in miRNAs in response to methoprene. Newly-emerged adult female *Ae. aegypti* were decapitated and subjected to treatment with either methoprene "Meth" or solvent alone (acetone "Acet") and allowed to incubate for 12 hours to allow induction of pathways affected by JH. Following RNA isolation, small RNA cloning was carried out upon the samples, followed by by Illumina sequencing, which uncovered changes in miRNA expression up to four fold up and down regulation of miRNAs. Shown in this figure is the number of hits corresponding to known miRNAs of each given identity for each of two samples. A ratio "Meth/Acet" was determined by dividing the results for the acetone only sample by the number of reads for the methoprene-treated sample. To correct for differences in sequencing, the Meth/Acet ratio was multiplied by the ratio of the total number of reads for each sample, yielding the "Fold Change" number, a corrected value for miRNA fold change. In total, nine miRNAs (highlighted in orange) had >30 hits per sample and exhibited two fold or higher change. Two of the nine miRNAs (miR-9a, and miR-1891) were found to correlate with small RNA cloning results. 2 additional miRNAs (in green, miR-305 and miR-989; both less than two fold change) were examined by northern blot, but failed to correlate with small RNA cloning.

Chapter 5.

Summary and Future Directions

This study presents the discovery, characterization, and exploration of function and regulation of microRNAs (miRNAs) in *Anopheline* and *Aedine* mosquitoes, the vectors of malaria and dengue. Initial studies led to the discovery of 27 different miRNAs in aged female *Anopheles stephensi*, the age group responsible for transmission of malarial parasites. Comparative bioinformatics uncovered two clusters of miRNA genes in *Drosophila, Aedes,* and *Anopheles* where two miRNA genes were absent in *Ae. aegypti*. Among the miRNA discovered, three were mosquito-specific (miR-1889, -1890, and –1891). One miRNA discovered, miR-989, was primarily expressed in the ovaries of divergent mosquitoes, and its level fluctuated in response to a bloodmeal, suggesting that this miRNA plays an important role in mosquito reproductive processes. In addition, cloning and northern analysis in *An. stephensi* revealed an abundance of miR-14 across the lifestages of mosquitoes. In flies, this miRNA has been linked to longevity. These studies shed light on an understudied area of gene regulation in mosquitoes that may eventually help facilitate vector-disease control efforts, given the prominent roles that reproduction and longevity play in mosquito vectorial capacity. Given new developments in sequencing technology, "deep" sequencing approaches may be useful in future research to more thoroughly characterize the miRNAs present in *Anopheline* mosquitoes. Higher throughput expression profiles of miRNAs of interest could be achieved through qRT-PCR.

Results from the initial cloning studies of *An. stephensi* were expanded upon to study the expression of eight mosquito-specific miRNAs. Northerns were conducted to examine expression in the lifestages of the three subfamilies of mosquitoes, which revealed varying degrees of conservation within Culicidae, the mosquitoes. MiR-M1, -1175, -1890, and –1891 were expressed across all subfamilies of mosquitoes, and exhibited similar expression patterns in different species, implying roles in functions required for all mosquitoes. MiR-1174 was not observed to be expressed in *T. amboinensis*, a mosquito which does not take bloodmeals. A cluster of three miRNAs, the N1/N2/N3 cluster, was highly abundant in embryos and displayed high lineage specificity, with miR-N1 only present among Culicinae examined, miR-N2 only within *Ae. aegypti*, and miR-N3 present only within *Culex quinquefasciatus*. These miRNAs likely target similar transcripts as they share a common seed region. Given sequence similarities between them, these miRNAs may have arisen from duplication events in ancestral Culicinae. The high degree of lineage specificity invites further studies to determine if these miRNAs could play roles in lineage-specific traits or if they were involved in Culicine divergence. The N1/N2/N3 cluster and miR-M1 exhibited embryonic expression, and miR-M1 is likely to be zygotically-expressed as it does not appear in bloodfed ovary samples. Function-based studies to answer the roles of these miRNAs would be a natural progression of the research.

miR-989 was found broadly across holometabolous insects and not outside, suggesting it may regulate functions specific to holometabolous insects. miR-989 expression in female

An. stephensi and *Ae. aegypti* dramatically rises following pupal emergence and remains strong until about 32 hours post-blood meal at which point it rapidly begins to drop and remains minimally expressed through at least 96 hours. While similar in profile to the expression of Juvenile Hormone (JH), there was only minimal regulation at best of miR-989 by JH based on small RNA cloning of an abdominal ligation study. Among the targets predicted for miR-989, twenty possible candidates were examined in a PCR-based experimental approach, though sequencing suggested that artifacts were amplified. Small RNA cloning of an *Ae. aegypti* female abdominal ligation samples revealed many miRNAs which are likely regulated by the JH, one of the few instances of hormonal regulation of miRNAs with empirical support. As further developments occur in genomics, new databases become available to provide a more thorough analysis of the targets of miR-989. We have opened up many new avenues for research for many miRNAs that are likely regulated by JH. Where are they located? What do they target?

In summary, we have greatly expanded the understanding of miRNAs in mosquitoes, which are important regulators of transcription in some of the world's worst disease vectors. An improved understanding of mosquito physiology can assist in efforts to control mosquito-borne infectious diseases.

References

Ahmadian, A., Ehn, M., and Hober, S. (2006) Pyrosequencing: History, biochemistry and future. *Clinica Chimica Acta* **363**: 83-94.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. and Walker, P. (2002) Cell Junctions, Cell Adhesion, and the Extracellular Matrix, Chapter 19, pp. 1065-1125 *in* Molecular Biology of the Cell, 4th Edition. Garland Science, New York.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403-410.

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. *Nucl Acids Res* **25**(17)**:** 3389-3402.

Amaral, P.P., Dinger, M.E., Mercer, T.R., and Mattick, J.S. (2008) The Eukaryotic Genome as an RNA Machine. *Science* **319**(5871): 1787 – 1789.

Ambros, V., Bartel, B., Bartel, D.P., Burge, C.B., Carrington, J.C., Chen, X., Dreyfuss, G., Eddy, S.R., Griffiths-Jones, S., Marshall, M., Matzke, M., Ruvkun, G., and Tuschl, T. (2003) A uniform system for microRNA annotation. *RNA* **9**: 277-279.

Ambros, V., and Chen, X. (2007) The regulation of genes and genomes by small RNAs. *Development* **134**(9): 1635-1641.

Ameres, S.L., Martinez, J., and Schroeder, R. (2007) Molecular basis for target RNA recognition and cleavage by human RISC. *Cell* **130**: 101–112.

Andachi, Y. (2008) A novel biochemical method to identify target genes of individual microRNAs: identification of a new *Caenorhabditis elegans* let-7 target. *RNA* **14**(11): 2440-2451.

Aravin, A.A., Lagos-Quintana, M., Yalcin, A., Zavolan, M., Marks, D., Snyder, B., Gaasterland, T., Meyer, J., and Tuschl, T. (2003) The small RNA profile during *Drosophila melanogaster* development. *Dev Cell* **5**(2): 337-350.

Aravin, A.A., Naumova, N.M., Tulin, A.V., Vagin, V.V., Rozovsky, Y.M., and Gvozdev, V.A. (2001) Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the *D. melanogaster* germline. *Curr Biol* **11**(13):101-27.

Aravin, A.A., Sachidanandam, R., Girard, A., Fejes-Toth, K., and Hannon, G.J. (2007) Developmentally regulated piRNA clusters implicate MILI in transposon control. *Science* **316**(5825): 744-747.

Aultman, K.S., Walker, E.D., Gifford, F., Severson, D.W., Beard, C.B., and Scott, T.W. (2000) Research ethics. Managing risks of arthropod vector research. *Science* **288**(5475):2321-2.

Bartel, B. (2005) MicroRNAs directing siRNA biogenesis. *Nat Struct Mol Biol* **12**(7): 569-571.

Bartel, D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**(2): 281-297.

Bashirullah, A., Pasquinelli, A.E., Kiger, A.A., Perrimon, N., Ruvkun, G., and Thummel, C.S. (2003) Coordinate regulation of small temporal RNAs at the onset of *Drosophila* metamorphosis. *Dev Biol* **259**(1): 1-8.

Baskerville, S., and Bartel, D.P. (2005) Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA* **11**(3): 241-247.

Beerntsen, B.T., Champagne, D.E., Coleman, J.L., Campos, Y.A., and James, A.A. (1999) Characterization of the *Sialokinin* I gene encoding the salivary vasodilator of the yellow fever mosquito, *Aedes aegypti*. *Insect Mol Biol* **8**(4): 459-467.

Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P., and Izaurralde, E. (2006) mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. *Genes Dev* **20**: 1885–1898.

Bentley, D.R. (2006) Whole-genome re-sequencing. *Curr Opin Gen & Dev* **16**(6): 545-552.

Bentwich, I., Avniel, A., Karov, Y., Aharonov, R., Gilad, S., Barad, O., Barzilai, A., Einat, P., Einav, U., Meiri, E., Sharon, E., Spector, Y., and Bentwich, Z. (2005) Identification of hundreds of conserved and nonconserved human microRNAs. *Nat Genet* **37**(7): 766-770.

Berezikov, E., Chung, W.J., Willis, J., Cuppen, E., and Lai, E.C. (2007) Mammalian mirtron genes. *Mol Cell* **28**: 328–336.

Berezikov, E., Cuppen, E., and Plasterk, R.H.A. (2006) Approaches to microRNA discovery**.** *Nat Genet* **38:** 52-57.

Berezikov, E., Thuemmler, F., Van Laake, L.W., Kondova, I., Bontrop, R., Cuppen, E., and Plasterk, R.H. (2006) Diversity of microRNAs in human and chimpanzee brain. *Nat Genet* **38**: 1375–1377.

Besansky, N.J. on behalf of the *Anopheles* Genomes Cluster Committee. (2006) Proposal for a Twelve Genomes Cluster for Genus *Anopheles:* Genome sequencing white paper. 10 printed pages.

Besansky NJ, Adams J, Ashburner M, Benedict M, Carlton J, Maureen Coetzee M, Collins FH, della Torre A, Hemingway J, Roos DS (2005) Eight Genomes Cluster for Genus *Anopheles*: Genome sequencing white paper. 17 printed pages.

Biedler, J., Qi, Y., Holligan, D., della Torre, A., Wessler, S., and Tu, Z. (2003) Transposable element (TE) display and rapid detection of TE insertion polymorphism in the *Anopheles gambiae* species complex. *Insect Mol Biol* **12**(3): 211-216.

Biedler, J., and Tu, Z. (2003) Non-LTR retrotransposons in the African malaria mosquito, *Anopheles gambiae*: unprecedented diversity and evidence of recent activity. *Mol Biol Evol* **20**(11): 1811-1825.

Biotage. (2004) The principle of Pyrosequencing. Uppsala, Sweden.

Biryukova, I., Asmar. J., Abdesselem, H., and Heitzler, P. (2009) *Drosophila* mir-9a regulates wing development via fine-tuning expression of the LIM only factor, dLMO. *Dev Biol* **327**(2): 487-496.

Black, W.C. IV, and Kondratieff, B. C. (2005) Evolution of Arthropod Disease Vectors, Chapter 2, *in*: The Biology of Disease Vectors, Volume 2. (W. C. Marquardt, *ed*.), pp. 9-23. Elsevier Academic Press, Burlington.

Bohnsack, M.T., Czaplinski, K., and Gorlich, D. (2004) Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA* **10**(2): 185-191.

Borchert, G.M., Lanier, W., and Davidson, B.L. (2006) RNA polymerase III transcribes human microRNAs. *Nat Struct Mol Biol* **13**(12):1097-101.

Buchon, N, and Vaury, C. (2006) RNAi: a defensive RNA-silencing against viruses and transposable elements. *Heredity* **96**:195–202.

Bray, N., Dubchak, I., and Pachter, L. (2003) AVID: A global alignment program. *Genome Res* **13**(1): 97-102.

Brennecke, J., Aravin, A.A., Stark, A., Dus, M., Kelis, M., Sachidanandam, R., and Hannon, G.J. (2007) Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* **128**: 1089–1103.

Brennecke, J., and Cohen, S.M. (2003) Towards a complete description of the microRNA complement of animal genomes. *Genome Biol* **4**(9): 228.

Brennecke, J., Stark, A., Russell, R.B., and Cohen, S.M. (2005) Principles of microRNA-target recognition. *PLoS Biol* **3**(3): e85.

Brenner, S., Johnson, M., Bridgham, J., Golda, G., Lloyd, D.H., Johnson, D., Luo S., McCurdy, S., Foy, M., Ewan, M., Roth, R., George, D., Eletr, S., Albrecht, G., Vermaas, E., Williams, S.R., Moon, K., Burcham, T., Pallas, M., DuBridge, R.B., Kirchner, J., Fearon, K., Mao. J., and Corcoran, K. (2000) Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nat Biotechnol* **18**: 630–634.

Brodersen, P., Sakvarelidze-Achard, L., Bruun-Rasmussen, M., Dunoyer, P., Yamamoto, Y.Y., Sieburth, L., and Voinnet, O. (2008) Widespread translational inhibition by plant miRNAs and siRNAs. *Science* **320**(5880): 1185 – 1190.

Brown, A.E., Bugeon, L., Crisanti, A., Catteruccia, F. (2003) Stable and heritable gene silencing in the malaria vector *Anopheles stephensi*. *Nuc Acids Res* 31(15): e85.

Bushati, N., and Cohen, S.M. (2007) microRNA Functions. *Ann Rev Cell Dev Biol* **23**(1): 175-205.

Cai, X., Hagedorn, C.H., and Cullen, B.R. (2004) Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA* **10**: 1957–1966.

Caldas, C., and Brenton, J.D. (2005) Sizing up miRNAs as cancer genes. *Nat Med* **11**(7): 712-714.

Calin, G.A., Liu, C.G., Sevignani, C., Ferracin, M., Felli, N., Dumitru, C.D., Shimizu, M., Cimmino, A., Zupo, S., Dono, M., Dell'Aquila, M.L., Alder, H., Rassenti, L., Kipps, T.J., Bullrich, F., Negrini, M., and Croce, C.M. (2004) MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias. *Proc Natl Acad Sci U S A* **101**(32): 11755-11760.

Carissimi, C., Fulci, V., and Macino, G. (2009) MicroRNAs: novel regulators of immunity. *Autoimmun Rev* **8**(6):520-4.

Cha, S.-J., Mori, A., Chadee, D.D., Severson, D.W. (2006) Cage trials using an endogenous meiotic drive gene in the mosquito *Aedes aegypti* to promote population replacement. *Am J Trop Med Hyg* **74**(1): 62–68.

Chang, S., Johnston, R.J. Jr., Frokjaer-Jensen, C., Lockery, S., and Hobert, O. (2004) MicroRNAs act sequentially and asymmetrically to control chemosensory laterality in the nematode. *Nature* **430**(7001): 785-789.

Chang, S., Johnston, R.J. Jr., and Hobert, O. (2003) A transcriptional regulatory cascade that controls left/right asymmetry in chemosensory neurons of *C. elegans*. *Genes Dev* **17**(17): 2123-2137.

Chalberg, T.W., Portlock, J.L., Olivares, E.C., Thyagarajan, B., Kirby, P.J., Hillman, R.T., Hoelters, J., Calos, M.P. (2006) Integration specificity of phage phiC31 integrase in the human genome. *J Mol Biol* **357**(1): 28-48.

Chatterjee, R., and Chaudhuri, K. (2006) An approach for the identification of microRNA with an application to *Anopheles gambiae*. *Acta Biochim Pol* **53**: 303-309.

Chen, C.H., Huang, H., Ward, C.M., Su, J.T., Schaeffer, L.V., Guo, M., and Hay, B.A. (2007) A synthetic maternal-effect selfish genetic element drives population replacement in *Drosophila*. *Science* **316**(5824):597-600.

Chen, C.Z., Li, L., Lodish, H.F., and Bartel, D.P. (2004) MicroRNAs modulate hematopoietic lineage differentiation. *Science* **303**(5654): 83-86.

Chen, C.Z., and Lodish, H.F. (2005) MicroRNAs as regulators of mammalian hematopoiesis. *Semin Immunol* **17**(2): 155-165.

Chen, P.Y., Manninga, H., Slanchev, K., Chien, M., Russo, J.J., Ju, J., Sheridan, R., John, B., Marks, D.S., Gaidatzis, D., Sander, C., Zavolan, M., and Tuschl, T. (2005) The developmental miRNA profiles of zebrafish as determined by small RNA cloning. *Genes Dev* **19**(11): 1288-1293.

Chen, Q., Schlichtherle, M., and Wahlgren, M. (2000) Molecular Aspects of Severe Malaria. *Clin Microbio Rev* **13**(3): 439-450.

Chu, C.Y., and Rana, T.M. (2006) Translation repression in human cells by microRNA-induced gene silencing requires RCK/p54. *PLoS Biol* **4**: e210.

Clements, A.N. (1992) The Biology of Mosquitoes, Vol.1: Development, Nutrition, and Reproduction. Chapman & Hall, New York.

Coates, C.J., Kaminski, J.M., Summers, J.B., Segal, D.J., Miller, A.D., and Kolb, A.F. (2005) Site-directed genome modification: derivatives of DNA-modifying enzymes as targeting tools. *Trends Biotechnol* **23**(8):407-19.

Cohen, E.D., Mariol, M.C., Wallace, R.M.H., Weyers, J., Kamberov, Y.G., Pradel, J., and Wilder, E.L. (2002). DWnt4 regulates cell movement and focal adhesion kinase during *Drosophila* ovarian morphogenesis. *Dev Cell* **2**(4): 437-448.

Coleman, P.G., and Alphey, L. (2004) Genetic control of vector populations: an imminent prospect. *Trop Med Int Health* **9**(4):433-7.

Cullen B.R. (2006) Viruses and microRNAs. *Nat Genet* **38**: S25-S30.

Curtis C.F. (1994) The case for Malaria Control by Genetic Manipulation of its Vectors. *Parasitol Today* **10**(10): 371-374.

Denli, A., Tops, B.B.J., Plasterk, R.H.A., Ketting, R.F., and Hannon, G.J. (2004) Processing of primary microRNAs by the Microprocessor complex. *Nature* **432**: 231-235.

Dennis, C. (2002) Small RNAs: the genome's guiding hand? *Nature* **420**(6917): 732.

Deshpande, G., Calhoun, G., and Schedl, P. (2005) *Drosophila* argonaute-2 is required early in embryogenesis for the assembly of centric/centromeric heterochromatin, nuclear division, nuclear migration, and germ-cell formation. *Genes Dev* **19**(14):1680-5.

Dissanayake, S., Marinotti, O., Ribeiro, J.M.C., and James, A.A. (2006) angaGEDUCI: *Anopheles gambiae* gene expression database with integrated comparative algorithms for identifying conserved DNA motifs in promoter sequences. *BMC Genomics* **7**: 116.

Doench, J.G., Petersen, C.P., and Sharp, P.A. (2003) siRNAs can function as miRNAs. *Genes Dev* **17**: 438–442.

Doench, J.G., and Sharp, P.A. (2004) Specificity of microRNA target selection in translational repression. *Genes Dev* **18**: 504–511.

Doerig, C., Billker, O., Haystead, T., Sharma, P., Tobin, A.B., and Waters, N.C. (2008) Protein kinases of malaria parasites: an update. *Trends Parasitol*. **24**(12): 570-577.

Du, T., and Zamore, P.D. (2007) Beginning to understand microRNA function. *Cell Res* **17**(8): 661-663.

Du, T., and Zamore, P.D. (2005) microPrimer: the biogenesis and function of microRNA. *Development* **132**(21): 4645-4652.

Edwards, M.J., Lemos, F.J., Donnelly-Doman, M., and Jacobs-Lorena, M. (1997) Rapid induction by a blood meal of a carboxypeptidase gene in the gut of the mosquito *Anopheles gambiae*. *Insect Biochem Mol Biol* **27**(12): 1063-1072.

Edwards, M.J., Moskalyk, L.A., Donelly-Doman, M., Vlaskova, M., Noriega, F.G., Walker, V.K., and Jacobs-Lorena, M. (2000) Characterization of a carboxypeptidase A gene from the mosquito, *Aedes aegypti*. *Insect Mol Biol* **9**(1): 33-38.

Edwards, M.J., Severson, D.W., and Hagedorn, H.H. (1998) Vitelline envelope genes of the yellow fever mosquito, *Aedes aegypti*. *Insect Biochem Mol Biol* **28**(12): 915-925.

Eldridge, B.F. (2005) Mosquitoes, the Culicidae, Chapter 9, *in*: The Biology of Disease Vectors, Volume 2. (W. C. Marquardt, *ed*.), pp. 95-111. Elsevier Academic Press, Burlington.

Engels, W.R. (1997) Invasions of P elements. *Genetics* **145**:11-15.

Enright, A.J., John, B., Gaul, U., Tuschl, T., Sander, C., and Marks, D.S. (2003) MicroRNA targets in *Drosophila*. *Genome Biol* **5**(1): R1.

Enserink, M. (2002) What mosquitoes want: secrets of host attraction. *Science* **298**(5591): 90 – 92.

Enserink, M. (2007) Chikungunya: No Longer a Third World Disease. *Science* **318**(5858): 1860-1861.

Enserink, M. (2008) Entomology. A mosquito goes global. *Science* **320**(5878):864-6.

Erson, A.E., and Petty, E.M. (2008) MicroRNAs in development and disease. *Clin Genet* **74**: 296–306.

Eulalio, A., Rehwinkel, J., Stricker, M., Huntzinger, E., Yang, S.F., Doerks, T., Dorner, S., Bork, P., Boutros, M., and Izaurralde, E. (2007) Target-specific requirements for enhancers of decapping in miRNA-mediated gene silencing. *Genes Dev* **21**: 2558–2570.

Findley, S.D., Tamanaha, M., Clegg, N.J., and Ruohola-Baker, H. (2003) Maelstrom, a *Drosophila* spindle-class gene, encodes a protein that colocalizes with Vasa and RDE1/AGO1 homolog, Aubergine, in nuage. *Development* **130**(5):859-71.

Flangan, T.R., and Hagedorn, H.H. (2008) Vitellogenin synthesis in the mosquito: the role of juvenile hormone in the development of responsiveness to ecdysone. *Physiol Ent* **2**(3): 173 –178.

Flicek, P., Aken, B.L., Beal, K., Ballester, B., Caccamo, M., Chen, Y., Clarke, L., Coates, G., Cunningham, F., Cutts, T., Down, T., Dyer, S.C., Eyre, T., Fitzgerald, S., Fernandez-Banet, J., Gräf, S., Haider, S., Hammond, M., Holland, R., Howe, K.L., Howe, K., Johnson, N., Jenkinson, A., Kähäri, A., Keefe, D., Kokocinski, F., Kulesha, E., Lawson, D., Longden, I., Megy, K., Meidl, P., Overduin, B., Parker, A., Pritchard, B., Prlic, A., Rice, S., Rios, D., Schuster, M., Sealy, I., Slater, G., Smedley, D., Spudich, G., Trevanion, S., Vilella, A.J., Vogel, J., White, S., Wood, M., Birney, E., Cox, T., Curwen, V., Durbin, R., Fernandez-Suarez, X.M., Herrero, J., Hubbard, T.J.P., Kasprzyk, A., Proctor, G., Smith, J., Ureta-Vidal, A., and Searle, S. (2008) Ensembl 2008. *Nucl Acids Res*. **36**: D707-D714.

Forstemann, K., Horwich, M.D., Wee, L., Tomari, Y., and Zamore, P.D. (2007) *Drosophila* microRNAs are sorted into functionally distinct argonaute complexes after production by dicer-1. *Cell* **130**:287–297.

Frazer, K.A., Pachter, L., Poliakov, A., Rubin, E.M., and Dubchak, I. (2004) VISTA: computational tools for comparative genomics. *Nuc Acids Res* Jul 1;32 (Web Server issue):W273-9.

Friedman, R.C., Farh, K.K., Burge, C.B., and Bartel, D. (2009) Most Mammalian mRNAs are Conserved Targets of microRNAs. *Genome Res* **19**(1): 92-105.

Fukuda, T., Yamagata, K., Fujiyama, S., Matsumoto, T., Koshida, I., Yoshimura, K., Mihara, M., Naitou, M., Endoh, H., Nakamura, T., Akimoto, C., Yamamoto, Y., Katagiri, T., Foulds, C., Takezawa, S., Kitagawa, H., Takeyama, K., O'Malley, B.W., and Kato, S. (2007) DEAD-box RNA helicase subunits of the Drosha complex are required for processing of rRNA and a subset of microRNAs. *Nat Cell Biol* **9**: 604–611.

Gaunt, M.W., and Miles, M.A. (2002) An insect molecular clock dates the origin of the insects and accords with palaeontological and biogeographic landmarks. *Mol Biol Evol* **19**: 748-761.

Ghildiyal, M., Seitz, H., Horwich, M.D., Li, C., Du, T., Lee, S., Xu, J., Kittler, E.L., Zapp, M.L., Weng, Z., and Zamore, P.D. (2008) Endogenous siRNAs derived from transposons and mRNAs in *Drosophila* somatic cells. *Science* **320**(5879): 1077-1081.

Gibson, G., and Russell, I. (2006) Flying in Tune: Sexual Recognition in Mosquitoes. *Curr Biol* **16**(13): Pages 1311-1316

Giraldez, A.J., Cinalli, R.M., Glasner, M.E., Enright, A.J., Thomson, J.M., Baskerville, S., Hammond, S.M., Bartel, D.P., and Schier, A.F. (2005) MicroRNAs regulate brain morphogenesis in zebrafish. *Science* **308**(5723): 833-838.

Giraldez, A.J., Mishima, Y., Rihel, J., Grocock, R.J., Van Dongen, S., Inoue, K., Enright, A.J., and Schier, A.F. (2006) Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. *Science* **312**: 75–79.

Glazov, E.A., McWilliam, S., Barris, W.C., and Dalrymple, B.P. (2008) Origin, evolution and biological role of miRNA cluster in DLK-DIO3 genomic region in placental mammals. *Mol Biol Evol* **25**: 939–948.

Glinsky. G.V. (2008) An SNP-guided microRNA map of fifteen common human disorders identifies a consensus disease phenocode aiming at principal components of the nuclear import pathway. *Cell Cycle* **7**(16):2570-83.

Gonzalez, S., Pisano, D.G., and Serrano, M. (2008) Mechanistic principles of chromatin remodeling guided by siRNAs and miRNAs. *Cell Cycle* **7**(16): 2601-2608.

Gould, F., Magori, K., and Huang, Y.X. (2006) Genetic strategies for controlling mosquito-borne diseases. *American Scientist*. **94**(3): 238-246.

Goymer P. (2007) MicroRNA: Introducing the mirtron. *Nat Rev Mol Cell Biol* **8**, 597.

Grad, Y., Aach, J., Hayes, G.D., Reinhart, B.J., Church, G.M., Ruvkun, G., and Kim, J. (2003) Computational and experimental identification of C. elegans microRNAs. *Mol Cell* **11**(5): 1253-1263.

Griffiths-Jones, S. (2004) The microRNA Registry. *Nucl Acids Res* **32**: D109-D111.

Griffiths-Jones, S., Grocock, R.J., van Dongen, S., Bateman, A., and Enright, A.J. (2006) miRBase: microRNA sequences, targets and gene nomenclature. *Nucl Acids Res* **34**: D140-D144.

Grimson, A., Farh, K.K., Johnston, W.K., Garrett-Engele, P., Lim, L.P., and Bartel, D.P. (2007) MicroRNA Targeting Specificity in Mammals: Determinants beyond Seed Pairing. *Mol Cell* **27**: 91-105.

Grosshans, H., and Filipowicz, W. (2008) Molecular biology: the expanding world of small RNAs. *Nature* **451**(7177):414-6.

Grün D., and Rajewsky, N. (2008) Computational prediction of microRNA targets in vertebrates, fruitflies and nematodes. In: MicroRNAs: From Basic Science to Disease Biology. (A.K. Cambridge, *eds*.), pp. 172-186. Cambridge University Press, Cambridge.

Grün, D., Wang, Y.L., Langenberger, D., Gunsalus, K.C., and Rajewsky, N. (2005) microRNA target predictions across seven *Drosophila* species and comparison to mammalian targets. *PLoS Comp Biol* **1**: e13.

Gubler, D.J. (1998) Resurgent Vector-Borne Diseases as a Global Health Problem. *Emerg Infect Dis* **4**(3): 442-450.

Guil, S., and Caceres, J.F. (2007) The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. *Nat Struct Mol Biol* **14**: 591–596.

Gunawardane, L.S., Saito, K., Nishida, K.M., Miyoshi, K., Kawamura, Y., Nagami, T., Siomi, H., and Siomi, M.C. (2007) A slicer-mediated mechanism for repeat-associated siRNA 5 end formation in *Drosophila*. *Science* **315**:1587–1590.

Gu, S., Jin, L., Zhang, F., Sarnow, P., and Kay, M.A. (2009) Biological basis for restriction of microRNA targets to the 3' untranslated region in mammalian mRNAs. *Nat Struct Mol Biol* **16**: 144 - 150.

Gu, Z., Eleswarapu, S., and Jiang, H. (2007) Identification and characterization of microRNAs from the bovine adipose tissue and mammary gland. *FEBS Let* **581**: 981-988.

Gwadz, R., and Collins, F.H. (1996) Anopheline Mosquitoes and the Agents they Transmit pg.73-84. In: The Biology of Disease Vectors 1996 Barry J Beaty and William C. Marquardt University Press, Colorado.

Hagedorn, H.H. (1996) Physiology of Mosquitoes, Chapter 17, pp. 273-297, *in*: The Biology of Disease Vectors. (B. Beaty and W.C. Marquardt, *eds*.), Elsevier Academic Press, Burlington.

Hagedorn, H.H. (2005) Mosquito Endocrinology. *in*: The Biology of Disease Vectors, Volume 2. (W. C. Marquardt, *ed*.), pp. 317-327. Elsevier Academic Press, Burlington.

Hagedorn, H.H., Turner, S., Hagedorn, E.A., Pontecorvo, D., Greenbaum, P., Pfeiffer, D., Wheelock, G., and Flanagan, T.R. (1997). Post-emergence growth of the ovarian follicles of *Aedes aegypti*. *J Insect Physiol* **23**: 203-206.

Han, J., Lee, Y., Yeom, K.H., Kim, Y.K., Jin, H., and Kim, V.N. (2004) The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev* **18**: 3016–3027.

Harbach, R.E. (2007) The Culicidae (Diptera): a review of taxonomy, classification and phylogeny. *Zootaxa* 1668:591–638.

Harbach, R.E., and Kitching, I.J. (2005) Reconsideration of anopheline mosquito phylogeny (Diptera: Culicidae: Anophelinae) based on morphological data. *Syst Biodiversity* **3**(4): 345-374.

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

Hartl, D.L. (2004) The origin of malaria: mixed messages from genetic diversity. *Nat Rev Microbiol* **2**(1): 15-22.

Hatfield, S.D., Shcherbata, H.R., Fischer, K.A., Nakahara, K., Carthew, R.W., and Ruohola-Baker, H. (2005) Stem cell division is regulated by the microRNA pathway. *Nature* **435**(7044): 974-978.

Hernández -Martinez, S, Li,Y., Lanz-Mendoza,H., Rodríguez, M.H., and. Noriega, F.G. (2009) Immunostaining for allatotropin and allatostatin-A and -C in the mosquitoes Aedes aegypti and Anopheles albimanus. *Cell Tissue Res* **321**(1): 105–113.

Hernández-Martínez, S., Li, Y., Rodriguez, M.H., Lanz-Mendoza, H., and Noriega, F.G. (2005) Allatotropin and PISCF- and YXFGL-amide-allatostatins distribution in Aedes aegypti and Anopheles albimanus mosquitoes. *Cell and Tissue Research* **321**:105–113.

Hernández -Martinez, S., Mayoral, J.G.,Li, Y.,and Noriega, F.G. (2007) Role of juvenile hormone and allatotropin on nutrient allocation, ovarian development and survivorship in mosquitoes. *J Insect Physiol* **53**: 230–234.

Hipfner, D.R., Weigmann, K., and Cohen, S.M. (2002) The bantam gene regulates *Drosophila* growth. *Genetics* **161**(4): 1527-1537.

Hofacker, I.L., Fontana, W., Stadler, P.F., Bonhoeffer, S., Tacker, M., and Schuster, P. (1994) Fast folding and comparison of RNA secondary structures. *Monatshefte f Chemie* **125:** 167–188.

Holt, R.A., Subramanian, G.M., Halpern, A., Sutton, G.G., Charlab, R., Nusskern, D.R., Wincker, P., Clark, A.G., Ribeiro, J.M., Wides, R., Salzberg, S.L., Loftus, B., Yandell, M., Majoros, W.H., Rusch, D.B., Lai, Z., Kraft, C.L., Abril, J.F., Anthouard, V., Arensburger, P., Atkinson, P.W., Baden, H., de Berardinis, V., Baldwin, D., Benes, V., Biedler, J., Blass, C., Bolanos, R., Boscus, D., Barnstead, M., Cai, S., Center, A., Chaturverdi, K., Christophides, G.K., Chrystal, M.A., Clamp, M., Cravchik, A., Curwen, V., Dana, A., Delcher, A., Dew, I., Evans, C.A., Flanigan, M., Grundschober-Freimoser, A., Friedli, L., Gu, Z., Guan, P., Guigo, R., Hillenmeyer, M.E., Hladun, S.L., Hogan, J.R., Hong, Y.S., Hoover, J., Jaillon, O., Ke, Z., Kodira, C., Kokoza, E., Koutsos, A., Letunic, I., Levitsky, A., Liang, Y., Lin, J.J., Lobo, N.F., Lopez, J.R., Malek, J.A., McIntosh, T.C., Meister, S., Miller,

J., Mobarry, C., Mongin, E., Murphy, S.D., O'Brochta, D.A., Pfannkoch, C., Qi, R., Regier, M.A., Remington, K., Shao, H., Sharakhova, M.V., Sitter, C.D., Shetty, J., Smith, T.J., Strong, R., Sun, J., Thomasova, D., Ton, L.Q., Topalis, P., Tu, Z., Unger, M.F., Walenz, B., Wang, A., Wang, J., Wang, M., Wang, X., Woodford, K.J., Wortman, J.R., Wu, M., Yao, A., Zdobnov, E.M., Zhang, H., Zhao, Q., Zhao, S., Zhu, S.C., Zhimulev, I., Coluzzi, M., della Torre, A., Roth, C.W., Louis, C., Kalush, F., Mural, R.J., Myers, E.W., Adams, M.D., Smith, H.O., Broder, S., Gardner, M.J., Fraser, C.M., Birney, E., Bork, P., Brey, P.T., Venter, J.C., Weissenbach, J., Kafatos, F.C., Collins, F.H., and Hoffman, S.L. (2002) The Genome Sequence of the Malaria Mosquito *Anopheles gambiae*. *Science* **298**(5591): 129-149.

Hutvagner, G., Simard, M.J., Mello, C.C., and Zamore, P.D. (2004) Sequence-specific inhibition of small RNA function. *PLoS Biol* **2**(4): E98.

Hwang, H.W., Wentzel, E.A., and Mendell, J.T. (2007) A hexanucleotide element directs microRNA nuclear import. *Science* **315**: 97–100.

Ibáñez-Ventoso, C., Vora, M., and Driscoll, M. (2008) Sequence Relationships among *C. elegans, D. melanogaster* and Human microRNAs Highlight the Extensive Conservation of microRNAs in Biology. *PLoS ONE* **3**(7): e2818.

Illumina, Inc. (2009a) Specification Sheet: Illumina sequencing. Genome Analyzer II System. Illumina, Inc. San Diego, CA.

Ingelfinger, D., Arndt-Jovin, D.J., Luhrmann, R., and Achsel, T. (2002) The human LSm1–7 proteins colocalize with the mRNA-degrading enzymes Dcp1/2 and Xrnl in distinct cytoplasmic foci. *RNA* **8**: 1489–1501.

Jackson, R.J., and Standart, N. (2007) How do microRNAs regulate gene expression? *Sci STKE* **367**: re1.

John, B., Enright, A.J., Aravin, A., Tuschl, T., Sander, C., and Marks, D.S. (2004) Human MicroRNA Targets. PLoS Biol 2(11): e363.

Johnston, R.J., and Hobert, O. (2003) A microRNA controlling left/right neuronal asymmetry in *Caenorhabditis elegans*. *Nature* **426**(6968): 845-849.

Kawashima, T., Shoguchi, E., Satou, Y., and Satoh, N. (2008) Comparative Genomics of Invertebrates, Chapter 6. *in*: Comparative Genomics: Basic and Applied Research (J. Brown, *ed*.), CRC Press, Boca Raton.

Kawaoka, S., Hayashi, N., Katsuma, S., Kishino, H., Kohara, Y., Mita, K., and Shimada, T. (2008) Bombyx small RNAs: genomic defense system against transposons in the silkworm, Bombyx mori. *Insect Biochem Mol Biol* **38**(12): 1058-1065.

Kim, Y.K., and Kim, V.N. (2007) Processing of intronic microRNAs. *EMBO J*. **26**: 775–783.

Kim, V.N. (2005) Small RNAs: Classification, Biogenesis, and Function. *Mol Cells* **19**(1): 1-15.

Kim, V.N. and Nam, J.W. (2006) Genomics of microRNA. *TRENDS Genet* **22**(3)**:** 165-173.

Kim, V.N., Han, J., and Siomi, M.C. (2009) Biogenesis of small RNAs in animals. *Nat Rev Mol Cell Biol* **10**(2):126-39.

Kloosterman, W.P., Wienholds, E., de Bruijn, E., Kauppinen, S., and Plasterk, R.H. (2006) *In situ* detection of miRNAs in animal embryos using LNA-modified oligonucleotide probes. *Nat Methods* **3**(1):27-9.

Kogan, G.L., Tulin, A.V., Aravin, A.A., Abramov, Y.A., Kalmykova, A.I., Maisonhaute, C., and Gvozdev, V.A. (2003) The GATE retrotransposon in *Drosophila melanogaster*: mobility in heterochromatin and aspects of its expression in germline tissues. *Mol Genet Genomics* **269**(2):234-42.

Kiriakidou, M., Nelson, P.T., Kouranov, A., Fitziev, P., Bouyioukos, C., Mourelatos, Z., and Hatzigeorgiou, A. (2004) A combined computational-experimental approach predicts human microRNA targets. *Genes Dev* **18**: 1165–1178.

Kiriakidou, M., Tan, G.S., Lamprinaki, S., De Planell-Saguer, M., Nelson, P.T., and Mourelatos, Z. (2007) An mRNA m7G cap binding-like motif within human Ago2 represses translation. *Cell* **129**: 1141–1151.

Krek, A., Grün, D., Poy, M.N., Wolf, R., Rosenberg, L., Epstein, E.J., MacMenamin, P., da Piedade, I., Gunsalus, K.C., Stoffel, M., and Rajewsky, N. (2005) Combinatorial microRNA target predictions. *Nat Genet* **37**: 495–500.

Krzywinski, J., Grushko, O.G., and Besansky, N.J. (2006) Analysis of the complete mitochondrial DNA from *Anopheles funestus*: An improved dipteran mitochondrial genome annotation and a temporal dimension of mosquito evolution. *Mol Phylogenet Evol* **39**(2): 417-423.

Lagos-Quintana, M., Rauhut, R., Meyer, J., Borkhardt, A., and Tuschl, T. (2003) New microRNAs from mouse and human. *RNA* **9**(2): 175-179.

Lai, E.C. (2002) MicroRNAs are complementary to the 3'UTR sequence motifs that mediate negative post-transcriptional regulation. *Nat Genet* **30**: 363-364.

Lai, E.C. (2005) miRNAs: Whys and Wherefores of miRNA-Mediated Regulation. *Curr Biol* **15**(12): R458-R460.

Lai, E.C., Tam, B. and Rubin, G.M. (2005) Pervasive regulation of *Drosophila* Notch target genes by GY-box-, Brd-box-, and K-box-class microRNAs. *Genes Dev* **19**(9): 1067-1080.

Lai, E.C., Tomancak, P., Williams, R.W. and Rubin, G.M. (2003) Computational identification of *Drosophila* microRNA genes. *Genome Biol* **4**(7): R42.

Lall, S., Grün, D., Krek, A., Chen, K., Wang, Y.L., Dewey, C.N., Sood, P., Colombo, T., Bray, N., Macmenamin, P., Kao, H.L., Gunsalus, K.C., Pachter, L., Piano, F., and Rajewsky, N. (2006) A genome wide map of conserved microRNA targets in *C. elegans*. *Curr Biol* **16**: 460-471.

Lau, N.C., Lim, L.P., Weinstein, E.G., and Bartel, D.P. (2001) An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* **294**(5543): 858-862.

Lau, N.C., Seto, A.G., Kim, J., Kuramochi-Miyagawa, S., Nakano, T., Bartel, D.P., and Kingston, R.E. (2006) Characterization of the piRNA complex from rat testes. *Science* **313**(5785): 363-367.

Lawson, D., Arensburger, P., Atkinson, P., Besansky, N.J., Bruggner, R.V., Butler, R., Campbell, K.S., Christophides, G.K., Christley, S., Dialynas, E., Emmert, D., Hammond, M., Hill, C.A., Kennedy, R.C., Lobo, N.F., MacCallum, M.R., Madey, G., Megy, K., Redmond, S., Russo, S., Severson, D.W., Stinson, E.O., Topalis, P., Zdobnov, E.M., Birney, E., Gelbart, W.M., Kafatos, F.C., Louis, C., and Collins, F.H. (2007) VectorBase: a home for invertebrate vectors of human pathogens. *Nucl Acids Res* **35**: D503-D505.

Leaman, D., Chen, P.Y., Fak, J., Yalcin, A., Pearce, M., Unnerstall, U., Marks, D.S., Sander, C., Tuschl, T., and Gaul, U. (2005) Antisense-mediated depletion reveals essential and specific functions of microRNAs in *Drosophila* development. *Cell* **121**(7): 1097-1108.

le Sage, C., Nagel, R., Egan, D.A., Schrier, M., Mesman, E., Mangiola, A., Anile, C., Maira, G., Mercatelli, N., Ciafrè, S.A., Farace, M.G., and Agami, R. (2007) Regulation of the p27Kip1 tumor suppressor by miR-221 and miR-222 promotes cancer cell proliferation. *EMBO J* **26**: 3699–3708.

Lecellier, C.H., Dunoyer, P., Arar, K., Lehmann-Che, J., Eyquem, S., Himber, C., Saib, A., and Voinnet, O. (2005) A cellular microRNA mediates antiviral defense in human cells. *Science* **308**(5721): 557-560.

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**(5): 843-854.

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Rådmark, O., Kim, S., and Kim, V.N. (2003) The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**: 415–419.

Lee, Y., Kim, M., Han, J., Yeom, K.H., Lee, S., Baek, S.H., and Kim, V.N. (2004) MicroRNA genes are transcribed by RNA polymerase II. *EMBO J* **23**: 4051–4060.

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.

Lewis, B.P., Burge, C.B., and Bartel, D.P. (2005) Conserved Seed Pairing, Often Flanked by Adenosines, Indicates that Thousands of Human Genes are MicroRNA Targets. *Cell* **120**(1): 15-20.

Lewis, B.P., Shih, I.H., Jones-Rhoades, M.W., Bartel, D.P., and Burge, C.B. (2003) Prediction of mammalian microRNA targets. *Cell* **115**(7): 787-798.

Li, L.C., Okino, S.T., Zhao, H., Pookot, D., Place, R.F., Urakami, S., Enokida, H., and Dahiya, R. (2006) Small dsRNAs induce transcriptional activation in human cells. *Proc Natl Acad Sci U S A* **103**(46): 17337-17342.

Liang, H., and Landweber, L.F. (2007) Hypothesis: RNA editing of microRNA target sites in humans? *RNA* **13**: 463-467.

Lim, L.P., Glasner, M.E., Yekta, S., Burge, C.B., and Bartel, D.P. (2003a) Vertebrate MicroRNA Genes. *Science* **299**:1540.

Lim, L.P., Lau, N.C., Weinstein, E.G., Abdelhakim, A., Yekta, S., Rhoades, M.W., Burge, C.B., and Bartel, D.P. (2003b) The microRNAs of *Caenorhabditis elegans*. *Genes Dev* **17**(8): 991-1008.

Lindow, M., and Krogh, A. (2005) Computational evidence for hundreds of non-conserved plant microRNAs. *BMC Genomics* **6**(1): 119.

Lippman, Z., May, B., Yordan, C., Singer, T., and Martienssen, R. (2003) Distinct mechanisms determine transposon inheritance and methylation via small interfering RNA and histone modification. *PLoS Biol* **1**(3):E67.

Liu, J., Valencia-Sanchez, M.A., Hannon, G.J., and Parker, R. (2005) MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. *Nat Cell Biol* **7**: 719–723.

Liu, X., Fortin, K., and Mourelatos, Z. (2007) MicroRNAs: Biogenesis and Molecular Functions. *Brain Pathol* **18**: 113-121.

Long, D., Lee, R., Williams, P., Chan, C.Y., Ambros, V., and Ding, Y. (2007) Potent effect of target structure on microRNA function. *Nat Struct Mol Biol* **14**: 287–294.

Lu, C., Tej, S.S., Luo, S., Haudenschild, C.D., Meyers, B.C., and Green, P.J. (2005) Elucidation of the small RNA component of the transcriptome. *Science* **309**(5740): 1567-1569.

Lu, J., Getz, G., Miska, E.A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B.L., Mak, R.H., Ferrando, A.A., Downing, J.R., Jacks, T., Horvitz, H.R., and Golub, T.R. (2005) MicroRNA expression profiles classify human cancers. *Nature* **435**(7043): 834-838.

Luciano, D.J., Mirsky, H., Vendetti, N.J., and Maas, S. (2004) RNA editing of a miRNA precursor. *RNA* **10**(8): 1174–1177.

Lytle, J.R., Yario, T.A., and Steitz, J.A. (2007) Target mRNAs are repressed as efficiently by microRNA-binding sites in the 5' UTR as in the 3' UTR. *PNAS* **104**: 9667–9672.

Maciel-de-Freitas, R., Eiras, A.E., and Lourenço-de-Oliveira, R. (2008) Calculating the survival rate and estimated population density of gravid *Aedes aegypti* (Diptera, Culicidae) in Rio de Janeiro, Brazil. *Cad Saude Publica* **24**(12): 2747-2754.

Maciel-de-Freitas, R., Peres, R.C., Alves, F., and Brandolini, M.B. (2008) Mosquito traps designed to capture *Aedes aegypti* (Diptera: Culicidae) females: preliminary comparison of Adultrap, MosquiTRAP and backpack aspirator efficiency in a dengue-endemic area of Brazil. *Mem Inst Oswaldo Cruz*. **103**(6): 602-605.

Mackenzie, J.S., Gubler, D.J., and Petersen, L.R. (2004) Emerging flaviviruses: the spread and resurgence of Japanese encephalitis, West Nile and dengue viruses. *Nat Med* **10**: S98–S109.

Makeyev, E.V., and Maniatis, T. (2008) Multilevel Regulation of Gene Expression by MicroRNAs. *Science* **319**: 1789-1790.

Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.J., Chen, Z., Dewell, S.B., Du, L., Fierro, J.M., Gomes, X.V., Godwin, B.C., He, W., Helgesen, S., Ho, C.H., Irzyk, G.P., Jando, S.C., Alenquer, M.L., Jarvie, T.P., Jirage, K.B., Kim, J.B., Knight, J.R., Lanza, J.R., Leamon, J.H., Lefkowitz, S.M., Lei, M., Li, J., Lohman, K.L., Lu, H., Makhijani, V.B., McDade, K.E., McKenna, M.P., Myers, E.W., Nickerson, E., Nobile, J.R., Plant, R., Puc, B.P., Ronan, M.T., Roth, G.T., Sarkis, G.J., Simons, J.F., Simpson, J.W., Srinivasan, M., Tartaro, K.R., Tomasz, A., Vogt, K.A., Volkmer, G.A., Wang, S.H., Wang, Y., Weiner, M.P., Yu, P., Begley, R.F., and Rothberg, J.M. (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**(7057): 376-380.

Marinotti, O., Calvo, E., and Nguyen, Q.K, Dissanayake S, Ribeiro JM, James AA. (2006) Genome-wide analysis of gene expression in adult *Anopheles gambiae*. *Insect Mol Biol* **15**(1): 1-12.

Marrelli, M.T., Li, C., Rasgon, J.L., Jacobs-Lorena, M. (2007) Transgenic malaria-resistant mosquitoes have a fitness advantage when feeding on Plasmodium-infected blood. *Proc Natl Acad Sci* USA **104**:5580–5583.

Mathonnet, G., Fabian, M.R., Svitkin, Y.V., Parsyan, A., Huck, L., Murata, T., Biffo, S., Merrick, W.C., Darzynkiewicz, E., Pillai, R.S., Filipowicz, W., Duchaine, T.F. and Sonenberg, N. (2007) MicroRNA inhibition of translation initiation *in vitro* by targeting the cap-binding complex eIF4F. *Science* **317**: 1764–1767.

Matranga, C., Tomari, Y., Shin, C., Bartel, D.P., and Zamore, P.D. (2005) Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. *Cell* **123**(4): 607-620.

Matuschewski, K. (2006) Vaccine development against malaria. *Curr Opin Immun* **18**(4): 449-457.

Mayor, C., Brudno, M., Schwartz, J.R., Poliakov, A., Rubin, E.M., Frazer, K.A., Pachter, L.S., and Dubchak, I. (2000) VISTA: visualizing global DNA sequence alignments of arbitrary length. *Bioinformatics* **16**(11): 1046-1047.

Mead, E.A., and Tu, Z. (2008) Cloning, Characterization, and Expression of microRNAs from the Asian Malaria Mosquito, *Anopheles stephensi*. *BMC Genomics* **9**: 244.

Meister, G., and Tuschl, T. (2004) Mechanisms of gene silencing by double-stranded RNA. *Nature* **431**(7006): 343-349.

Meyering-Vos, M., and Hoffman, K.H. (2003) Expression of allatostatins in the Mediterranean field cricket, Gryllus bimaculatus de Geer (Ensifera, Gryllidae). *Comp Biochem Phys Part B: Biochem Mol Biol* **136** (2): 207-215

Miranda, K.C., Huynh, T., Tay, Y., Ang, Y.S., Tam, W.L., Thomson, A.M., Lim, B., and Rigoutso, I. (2006) A pattern-based method for the identification of microRNA-target sites and their corresponding RNA/RNA complexes. *Cell* **126**: 1203-1217.

Mishima, T., Takizawa, T., Luo, S.S., Ishibashi, O., Kawahigashi, Y., Mizuguchi, Y., Ishikawa, T., Mori, M., Kanda, T., Goto, T., and Takizawa, T. (2008) MicroRNA (miRNA) cloning analysis reveals sex differences in miRNA expression profiles between adult mouse testis and ovary. *Reproduction* **136**(6):811-22.

Miska, E.A. (2005) How microRNAs control cell division, differentiation and death. *Curr Opin Genet Dev* **15**(5): 563-568.

Miska, E.A., Alvarez-Saavedra, E., Townsend, M., Yoshii, A., Sestan, N., Rakic, P., Constantine-Paton, M., and Horvitz, H.R. (2004) Microarray analysis of microRNA expression in the developing mammalian brain. *Genome Biol* **5**(9): R68.

Mitchell, A., Sperling, F.A.H., and Hickey, D.A. (2002) Higher level phylogeny of mosquitoes (Diptera: Culicidae): mtDNA data support a derived placement for *Toxorhynchites*. *Insect Syst Evol* **33**: 163–174.

Moita, C., Simoes, S., Moita, L.F., Jacinto, A., and Fernandes, P. (2005) The cadherin superfamily in *Anopheles gambiae*: a comparative study with *Drosophila melanogaster*. *Comp Funct Genom* **6**(4): 204–216.

Monticelli, S., Ansel, K.M., Xiao, C., Socci, N.D., Krichevsky, A.M., Thai, T.H., Rajewsky, N., Marks, D.S., Sander, C., Rajewsky, K., Rao, A., and Kosik, K.S. (2005) MicroRNA profiling of the murine hematopoietic system. *Genome Biol* **6**(8): R71.

Moreira, L.A., Ito, J., Ghosh, A., Devenport, M., Zieler, H., Abraham, E.G., Crisanti, A., Nolan, T., Catteruccia, F., and Jacobs-Lorena, M. (2002) Bee venom phospholipase inhibits malaria parasite development in transgenic mosquitoes *J Biol Chem* **277**(43):40839-43.

Morel, C.M., Toure, Y.T., Dobrokhotov, B., and Oduola, A.M. (2002) The mosquito genome--a breakthrough for public health. *Science* **298**(5591): 79.

Mourelatos, Z., Dostie, J., Paushkin, S., Sharma, A., Charroux, B., Abel, L., Rappsilber, J., Mann, M., and Dreyfuss, G. (2002) miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. *Genes Dev* **16**:720–728.

Nakahara, K., Kim, K., Sciulli, C., Dowd, S.R., Minden, J.S., and Carthew, R.W. (2005) Targets of microRNA regulation in the *Drosophila* oocyte proteome. *Proc Natl Acad Sci U S A* **102**(34): 12023-12028.

Nam, J.W., Shin, K.R., Han, J., Lee, Y., Kim, V.N., and Zhang, B.T. (2005) Human microRNA prediction through a probabilistic co-learning model of sequence and structure. *Nucl Acids Res* **33**(11): 3570-3581.

Nasci, R.S., and Miller, B.R. (1996) Culicine Mosquitoes and the Agents they Transmit. In: The Biology of Disease Vectors. (B.J. Beaty and C.William, *eds*.) Marquardt University Press, Colorado.

Nelson, P.T., Baldwin, D.A., Scearce, L.M., Oberholtzer, J.C., Tobias, J.W., and Mourelatos, Z. (2004) Microarray-based, high-throughput gene expression profiling of microRNAs. *Nat Methods* **1**(2): 155–161.

Nene, V., Wortman, J.R., Lawson, D., Haas, B., Kodira, C., Tu, Z.J., Loftus, B., Xi, Z., Megy, K., Grabherr, M., Ren, Q., Zdobnov, E.M., Lobo, N.F., Campbell, K.S., Brown, S.E., Bonaldo, M.F., Zhu, J., Sinkins, S.P., Hogenkamp, D.G., Amedeo, P., Arensburger, P., Atkinson, P.W., Bidwell, S., Biedler, J., Birney, E., Bruggner, R.V., Costas, J., Coy, M.R., Crabtree, J., Crawford, M., Debruyn, B., Decaprio, D., Eiglmeier, K., Eisenstadt, E., El-Dorry, H., Gelbart, W.M., Gomes, S.L., Hammond, M., Hannick, L.I., Hogan, J.R., Holmes, M.H., Jaffe, D., Johnston, J.S., Kennedy, R.C., Koo, H., Kravitz, S., Kriventseva, E.V., Kulp, D., Labutti, K., Lee, E., Li, S., Lovin, D.D., Mao, C., Mauceli, E., Menck, C.F., Miller, J.R., Montgomery, P., Mori, A,, Nascimento, A.L., Naveira, H.F., Nusbaum, C., O'leary, S., Orvis, J., Pertea, M., Quesneville, H., Reidenbach, K.R., Rogers, Y.H., Roth, C.W., Schneider, J.R., Schatz, M., Shumway, M., Stanke, M., Stinson, E.O., Tubio, J.M., Vanzee, J.P., Verjovski-Almeida, S., Werner, D., White, O., Wyder, S., Zeng, Q., Zhao, Q., Zhao, Y., Hill, C.A., Raikhel, A.S., Soares, M.B., Knudson, D.L., Lee, N.H., Galagan, J., Salzberg, S.L., Paulsen, I.T., Dimopoulos, G., Collins, F.H., Birren, B., Fraser-Liggett, C.M., and Severson, D.W. (2007) Genome Sequence of *Aedes aegypti*, a Major Arbovirus Vector *Science* **316**(5832): 1718 – 1723.

Nimmo, D.D., Alphey, L., Meredith, J.M., and Eggleston, P. (2006) High efficiency site-specific genetic engineering of the mosquito genome. *Insect Mol Biol* **15**(2): 129-136.

Noriega, F.G., Shah, D.K., and Wells, M.A. (1997) Juvenile hormone controls early trypsin gene transcription in the midgut of *Aedes aegypti*. *Insect Mol Biol* **6**(1): 63-66.

O'Connell, R.M., Taganov, K.D., Boldin, M.P., Cheng, G., and Baltimore, D. (2007) MicroRNA-155 is induced during the macrophage inflammatory response. *PNAS* **104**(5): 1604-1609.

Oh, S.-W., Kingsley, T., Shin, H.-H., Zheng, Z., Chen, H.-W., Chen, X., Wang, H., Ruan, P., Michelle Moody, M., and Hou, S.X. (2003) A P-Element Insertion Screen

Identified Mutations in 455 Novel Essential Genes in Drosophila. *Genetics* **163**: 195-201

Ohler, U., Yekta, S., Lim, L.P., Bartel, D.P., and Burge, C.B. (2004) Patterns of flanking sequence conservation and a characteristic upstream motif for microRNA gene identification. *RNA* **10**(9): 1309-1322.

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**(14): 1655-1666.

Okamura, K., Hagen, J.W., Duan, H., Tyler, D.M., and Lai, E.C. (2007) The mirtron pathway generates microRNA-class regulatory RNAs in *Drosophila*. *Cell* **130**:89–100.

Orban, T.I., and Izaurralde, E. (2005) Decay of mRNAs targeted by RISC requires XRN1, the Ski complex, and the exosome. *RNA* **11**: 459–469.

Pal-Bhadra, M., Leibovitch, B.A., Gandhi, S.G., Rao, M., Bhadra, U., Birchler, J.A., and Elgin, S.C. (2004) Heterochromatic silencing and HP1 localization in *Drosophila* are dependent on the RNAi machinery. *Science* **303**(5658):669-72.

Pasquinelli, A.E., Hunter, S., and Bracht, J. (2005) MicroRNAs: a developing story. *Curr Opin Genet Dev* **15**(2): 200-205.

Pawlicki, J.M., and Steitz, J.A. (2008) Primary microRNA transcript retention at sites of transcription leads to enhanced microRNA production. *J Cell Biol* **182**(1):61-76.

Pélisson, A., Sarot, E., Payen-Groschêne, G., and Bucheton, A. (2007) A novel repeat-associated small interfering RNA-mediated silencing pathway downregulates complementary sense gypsy transcripts in somatic cells of the *Drosophila* ovary. *J Virol* **81**(4): 1951-1960.

Perkins, S.L., and Austin, C. (2008) Four New Species of Plasmodium from New Guinea Lizards: Integrating Morphology and Molecules. *J Parasitol* **1**: [Epub ahead of print] doi:10.1645/GE-1750R.

Petersen, C.P., Bordeleau, M.E., Pelletier, J., and Sharp, P.A. (2006) Short RNAs repress translation after initiation in mammalian cells. *Mol Cell* **21**: 533–542.

Pfeffer, S., Lagos-Quintana, M., and Tuschl, T. (2003) Cloning of small RNA molecules, Unit 26.4. Pp. 1-18 in: Current protocols in molecular biology. (F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl *eds*.), John Wiley and Sons, New York.

Pillai, R.S., Bhattacharyya, S.N., Artus, C.G., Zoller, T., Cougot, N., Basyuk, E., Bertrand, E., and Filipowicz, W. (2005) Inhibition of translational initiation by Let-7 microRNA in human cells. *Science* **309**(5740):1573-6.

Place, R.F., Li, L.C., Pookot, D., Noonan, E.J., and Dahiya, R. (2008) MicroRNA-373 induces expression of genes with complementary promoter sequences. *PNAS* **105**(5): 1608-1613.

Poy, M.N., Eliasson, L., Krutzfeldt, J., Kuwajima, S., Ma, X., Macdonald, P.E., Pfeffer, S., Tuschl, T., Rajewsky, N., Rorsman, P., and Stoffel, M. (2004) A pancreatic islet-specific microRNA regulates insulin secretion. *Nature* **432**(7014): 226-230.

Prestridge, D.S. (1991) SIGNAL SCAN: A computer program that scans DNA sequences for eukaryotic transcriptional elements. *CABIOS* 7: 203-206.

Qiu, S., Adema, C.M., and Lane, T. (2005) A computational study of off-target effects of RNA interference. *Nucl Acids Res* **33**(6): 1834-1847.

Raikhel, A.S. (2005) Vitellogenesis of Disease Vectors, From Physiology to Genes. In: Biology of disease vectors. (W.C. Marquardt, *ed*), pp. 329-346. Elsevier academic press, Berlington.

Rajewsky, N., and Socci, N.D. (2004) Computational identification of microRNA targets. *Dev Biol* **267**(2): 529-535.

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**(6772): 901-906.

Reinert, J.F., Harbach, R.E., and Kitching, I.J. (2004) Phylogeny and classification of Aedini (Diptera: Culicidae), based on morphological characters of all life stages. *Zool J Linn Soc* **142**: 289–368.

Richter, J.D., and Sonenberg, N. (2005) Regulation of cap-dependent translation by eIF4E inhibitory proteins. *Nature* **433**: 477–480.

Roberts, J., Palma, E., Sazani, P., Orum, H., Cho, M., and Kole, R. (2006) Efficient and Persistent Splice Switching by Systemically Delivered LNA Oligonucleotides in Mice. *Mol Ther* **14** (4):471-475.

Ruby, J.G., Jan, C., Player, C., Axtell, M.J., Lee, W., Nusbaum, C., Ge, H., and Bartel, D.P. (2006) Large-scale sequencing reveals 21URNAs and additional microRNAs and endogenous siRNAs in *C. elegans*. *Cell* **127**: 1193–1207.

Ruby, J.G., Jan, C.H., and Bartel, D.P. (2007) Intronic microRNA precursors that bypass Drosha processing. *Nature* **448**: 83–86.

Ruby, J.G., Stark, A., Johnston, W.K., Kellis, M., Bartel, D.P., and Lai, E.C. (2007) Evolution, biogenesis, expression, and target predictions of a substantially expanded set of *Drosophila* microRNAs. *Genome Res* **17**(12): 1850-1864.

Ruvkun, G., and Slack, F. (1998) Suppressors of the retarded heterochronic gene let-7 identify multiple alleles of lin-41, which encodes a novel RING finger containing protein. East Coast Worm Meeting.

Saetrom, P., Heale, B.S., Snove. Jr., O Aagaard, L., Alluin, J., and Rossi, J.J. (2007) Distance constraints between microRNA target sites dictate efficacy and cooperativity. *Nuc Acids Res* **35**: 2333–2342.

Saito, K., Ishizuka, A., Siomi, H., and Siomi, M.C. (2005) Processing of pre-microRNAs by the Dicer-1-Loquacious complex in *Drosophila* cells. *PLoS Biol* **3**(7): e235.

Saito, K., Nishida, K.M., Mori, T., Kawamura, Y., Miyoshi, K., Nagami, T., Siomi, H., and Siomi, M.C. (2006). Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the *Drosophila* genome. *Genes Dev* **20**: 2214–2222.

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**(1): 65-74.

Savard, J., Tautz, D., Richards, S., Weinstock, G.M., Gibbs, R.A., Werren, J.H., Tettelin, and H., Lercher, M.J. (2006) Phylogenomic analysis reveals bees and wasps (*Hymenoptera*) at the base of the radiation of holometabolous insects. *Genome Res*. **16**:1334–1338.

Schmittgen, T.D., Jiang, J., Liu, Q., and Yang, L. (2004) A high-throughput method to monitor the expression of microRNA precursors. *Nucleic Acids Res* **32**(4): e43.

Scholte, E.J., Knols, B.G., Samson, R.A., and Takken, W.(2004) Entomopathogenic fungi for mosquito control: a review. *J Insect Sci*. **4**:19.

Sempere, L.F., Sokol, N.S., Dubrovsky, E.B., Berger, E.M., and Ambros, V. (2003) Temporal regulation of microRNA expression in *Drosophila melanogaster* mediated by hormonal signals and broad-Complex gene activity. *Dev Biol* **259**(1):9-18.

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**: 633–636.

Senior, K. (2007) Dengue fever: what hope for control? *Lancet Infect Dis* **7**(10): 636.

Sinkins, S.P., and Gould, F. (2006) Gene drive systems for insect disease vectors. *Nat Rev Genet* **7**(6):427-35.

Slack, F. J. (2008) MicroRNAs in Development and Cancer. *Molecular Cell Biology & Biotechnology Seminar Series*, Spring, 2008.

Slack, F., and Grosshans, H. (2002) Micro-RNAs: small is plentiful. *J Cell Biol* **156**:17-21.

Slack, F., and Ruvkun G. (1999) The heterochronic gene lin-41 encodes a temporally regulated RING finger protein that controls the timing of appearance of LIN-29 protein. *International C. elegans Meeting* **62**.

Smalheiser, N.A. (2003) EST analyses predict the existence of a population of chimeric microRNA precursor-mRNA transcripts expressed in normal human and mouse tissues. *Genome Biol* **4**(7): 403.

Smalheiser, N.R., and Torvik, V.I. (2005) Mammalian microRNAs derived from genomic repeats. *Trends Genet* **21**(6): 322-326.

Smith, A.D., Xuan, Z., and Zhang, M.Q. (2008) Using quality scores and longer reads improves accuracy of Solexa read mapping. *BMC Bioinformatics* **9**: 128.

Smith, C. (2005) Genomics: big tasks for small molecules. *Nature* **435**(7044): 991.

Sontheimer, E.J., and Carthew, R.W. (2005) Silence from within: endogenous siRNAs and miRNAs. *Cell* **122**(1): 9-12.

Spielman, A. (1994) Why Entomological Antimalaria Research Should not Focus on Transgenic Mosquitoes. *Parasitol Today* **10**: 374-376.

Stark, A., Brennecke, J., Russell, R.B., and Cohen, S.M. (2003) Identification of *Drosophila* MicroRNA targets. *PLoS Biol* **1**(3): E60.

Stark, A., Kheradpour, P., Parts, L., Brennecke, J., Hodges, E., Hannon, G.J., and Kellis, M. (2007) Systematic discovery and characterization of fly microRNAs using 12 *Drosophila* genomes**.** *Genome Res* **17**(12)**:** 1865-1879.

Subramaniam, S. (1998) The Biology Workbench--a seamless database and analysis environment for the biologist. *Proteins* **32**: 1-2.

Sullivan, C.S., Gründhoff, A.T., Tevethia, S., Pipas, J.M., and Ganem, D. (2005) SV40-encoded microRNAs regulate viral gene expression and reduce susceptibility to cytotoxic T cells. *Nature* **435**(7042): 682-686.

Swiezewski, S., Crevillen, P., Liu, F., Ecker, J.R., Jerzmanowski, A., and Dean, C. (2007) Small RNA-mediated chromatin silencing directed to the 32 region of the *Arabidopsis* gene encoding the developmental regulator, FLC. *PNAS* **104**(9): 3633-3638.

Takken, W., Dekker, T., and Wijnholds, Y.G. (1997) Odor-mediated flight behavior of *Anopheles gambiae* giles Sensu stricto and *An. stephensi* liston in response to CO2, acetone, and 1-octen-3-ol (Diptera: Culicidae). *J Insect Behav* **10**: 395-407.

Thermann, R., and Hentze, M.W. (2007) *Drosophila* miR2 induces pseudo-polysomes and inhibits translation initiation. *Nature* **447**: 875–878.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice**.** *Nucl Acids Res* **22**(22)**:** 4673-4680.

Thomson, J.M., Parker, J., Perou, C.M., and Hammond, S.M. (2004) A custom microarray platform for analysis of microRNA gene expression. *Nat Methods* **1**(1): 47-53.

Tomari, Y., Du, T., and Zamore, P.D. (2007) Sorting of *Drosophila* small silencing RNAs. *Cell* **130**: 299–308.

Tomari, Y., and Zamore, P.D. (2005) MicroRNA biogenesis: Drosha can't cut it without a partner. *Curr Biol* **15**: R61–R64.

Tomori O. (2004) Yellow fever: the recurring plague. *Crit Rev Clin Lab Sci* **41**: 391–427.

Tu, Z.J. (2001) Eight novel families of miniature inverted repeat transposable elements in the African malaria mosquito, *Anopheles gambiae*. *PNAS* **98**(4): 1699-1704.

Tyler, D.M., Okamura, K., Chung, W-J., Hagen, J.W., Berezikov, E., Hannon, G.J., and Lai, E.C. (2008) Functionally distinct regulatory RNAs generated by bidirectional transcription and processing of microRNA loci. *Genes Dev* **22**: 26-36.

Vagin, V.V., Sigova, A., Li, C., Seitz, H., Gvozdev, V., and Zamore, P.D. (2006) A distinct small RNA pathway silences selfish genetic elements in the germline. *Science* **313**: 320–324.

Valadi, H., Ekström, K., Bossios, A., Sjöstrand, M., Lee, J.J., and Lötvall, J.O. (2007) Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nat Cell Biol* **9**(6): 654-659.

Valoczi, A., Hornyik, C., Varga, N., Burgyan, J., Kauppinen, S., and Havelda, Z. (2004) Sensitive and specific detection of microRNAs by northern blot analysis using LNA-modified oligonucleotide probes. *Nucl Acids Res* **32**(22): e175.

van Dijk, E., Cougot, N., Meyer, S., Babajko, S., Wahle, E., and Seraphin, B. (2002) Human Dcp2: a catalytically active mRNA decapping enzyme located in specific cytoplasmic structures. *EMBO J* **21**: 6915–6924.

Vastenhouw, N.L., and Plasterk, R.H. (2004). RNAi protects the Caenorhabditis elegans germline against transposition. *Trends Genet* **20**:314–319.

Vasudevan, S., Tong, Y., and Steitz, J.A. (2007) Switching from repression to activation: microRNAs can up-regulate translation. *Science* **318**: 1931–1934.

Vella, C.M., Choi, E-Y., Lin, S-Y., Reinert, K., and Slack F.J. (2004) The C. elegans microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3'UTR. *Genes Dev* 18: 132-137.

Wakiyama, M., Takimoto, K., Ohara, O., and Yokoyama, S. (2007) Let-7 microRNA-mediated mRNA deadenylation and translational repression in a mammalian cell-free system. *Genes Dev* **21**: 1857–1862.

Wang, J., Xia, Q., He, X., Dai, M., Ruan, J., Chen, J., Yu, G., Yuan, H., Hu, Y., Li, R., Feng, T., Ye, C., Lu, C., Wang, J., Li, S., Wong, G.K., Yang, H., Wang, J., Xiang,

Z., Zhou, Z., and Yu, J. (2007) SilkDB: A knowledgebase for silkworm biology and genomics. *Nucl Acids Res* **33**: D399-D402.

Wang, X., Zhang, J., Li, F., Gu, J., He, T., Zhang, X., and Li, Y. (2005) MicroRNA identification based on sequence and structure alignment. *Bioinformatics* **21**(18): 3610-3614.

Watanabe, T., Totoki, Y., Toyoda, A., Kaneda, M., Kuramochi-Miyagawa, S., Obata, Y., Chiba, H., Kohara, Y., Kono, T., Nakano, T., Surani, M.A., Sakaki, Y., and Sasaki, H. (2008) Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* **453**(7194):539-543.

Watanabe, Y., Tomita, M., and Kanai, A. (2007) Computational Methods for MicroRNA Target Prediction. In: Methods in Enzymology, Vol 427: MicroRNA Methods, (J.J. Rossi, G.J. Hannon, *eds*.), pp. 65-86. Elsevier, New York.

Weaver, D.B., Anzola, J.M., Evans, J.D., Reid, J.G., Reese, J.T., Childs, K.L., Zdobnov, E.M., Samanta, M.P., Miller, J., Elsik, C.G. (2007) Computational and transcriptional evidence for microRNAs in the honey bee genome. *Genome Biol* **8**(6): R97.

Wengelnik, K., Vidal, V., Ancelin, M.L., Cathiard, A.M., Morgat, J.L., Kocken, C.H., Calas, M., Herrera, S., Thomas, A.W., and Vial, H.J. (2008) A class of potent antimalarials and their specific accumulation in infected erythrocytes. *Science* **295**(5558): 1311-4.

Werren, J.H., Gadau, J., Beukeboom, C., Desplan, J., Lynch, R., Rivers, S., and van de Zande, L. (2004) Proposal to sequence the *Nasonia* Genome (white paper).

Wheeler, W.C.Whiting, M., Wheeler, Q.D., and Carpenter, J.M. (2001) The phylogeny of the extant hexapod orders. *Cladistics* **17**:113-169.

Wienholds, E., Kloosterman, W.P., Miska, E., Alvarez-Saavedra, E., Berezikov, E., de Bruijn, E., Horvitz, H.R., Kauppinen, S., and Plasterk, R.H.A. (2005a) MicroRNA Expression in Zebrafish Embryonic Development. *Science* **309**: 310-311.

Wienholds, E., and Plasterk, R.H. (2005b) MicroRNA function in animal development. *FEBS Lett* **579**(26): 5911-5922.

Winter, F., Edaye, S., Huttenhofer, A., and Brunel, C. (2007) *Anopheles gambiae* miRNAs as actors of defence reaction against *Plasmodium* invasion. *Nucl Acids Res* **35**(20): 6953-6962.

Winzeler, E.A., (2008) Malaria research in the post-genomic era. *Nature* **455**(7214): 751-756.

Wu, L., Fan, J., and Belasco, J.G. (2006) MicroRNAs direct rapid deadenylation of mRNA. *PNAS* **103**: 4034–4039.

Wu, L., and Belasco, J.G. (2008) Let me count the ways: mechanisms of gene regulation by miRNAs and siRNAs. *Mol Cell* **29**(1):1-7.

Xiao, F., Zuo, Z., Cai, G., Kang, S., Gao, X., and Li, T. (2009) miRecords: an integrated resource for microRNA–target interactions. *NAR* **37:** D105-D110.

Xie, X., Lu, J., Kulbokas, E.J., Golub, T.R., Mootha, V., Lindblad-Toh, K., Lander, E.S., and Kellis, M. (2005) Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. *Nature* **434**(7031): 338-345.

Xu, P., Vernooy, S.Y., Guo, M., and Hay, B.A. (2003) The *Drosophila* microRNA Mir-14 suppresses cell death and is required for normal fat metabolism. *Curr Biol* **13**(9): 790-795.

Xu, Y., Zhou, X., and Zhang, W. (2008) MicroRNA prediction with a novel ranking algorithm based on random walks. *Bioinformatics* **24**(13):i50-8.

Xue, X., Zhang, Q., Huang, Y., Feng, L., and Pan, W. (2008) No miRNA were found in Plasmodium and the ones identified in erythrocytes could not be correlated with infection. *Malaria J* 7:47 (6 printed pages).

Yi, R., Qin, Y., Macara, I.G., and Cullen, B.R. (2003) Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev* **17**: 3011–3016.

Tomari, Y., Du, T., and Zamore, P.D. (2007) Sorting of Drosophila Small Silencing RNAs. *Cell* **130**(2): 299-308.

Zamore, P.D. (2007) RNA silencing: Genomic defence with a slice of pi. *Nature* **446**: 864-865.

Zamore, P.D., Tuschl, T., Sharp, P.A., and Bartel, D.P. (2000) RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* **101**: 25–33.

Zeng, Y., Wagner, E.J., and Cullen, B.R. (2002) Both natural and designed microRNAs can inhibit the expression of cognate mRNAs when expressed in human cells. *Mol Cell* **9**: 1327–1333.

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.

Zhang, R., Wang, Y.Q., and Su, B. (2008). Molecular evolution of a primate-specific microRNA family. *Mol Biol Evol* **25**(7): 1493-1502.

Zhao, T., Li, G., Mi, S., Li, S., Hannon, G., Wang, X.J., and Qi, Y. (2007) A complex system of small RNAs in the unicellular green alga *Chlamydomonas reinhardtii*. *Genes Dev* **21**: 1190-1203.

Zhao, Y., Samal, E., and Srivastava, D. (2005) Serum response factor regulates a muscle-specific microRNA that targets Hand2 during cardiogenesis. *Nature* **436**(7048): 214-220.

Zhou, X., Ruan, J., Wang, G., and Zhang, W. (2007) Characterization and identification of microRNA core promoters in four model species**.** *PLoS Comput Biol* **3**(3)**:** e37.

Zuker, M., Mathews, D.H., and Turner, D.H. (1999) Algorithms and Thermodynamics for RNA Secondary Structure Prediction: A Practical Guide in RNA Biochemistry and Biotechnology. In: NATO ASI Series, (J. Barciszewski, and B.F.C. Clark, *eds*.) Kluwer Academic, New York.