DISSERTATION

SMALL RNAS IN *AEDES AEGYPTI*: ONE GIANT STEP FOR VIRUS CONTROL IN MOSQUITOES

Submitted by

Adeline E. Williams

Department of Microbiology, Immunology, and Pathology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2022

Doctoral Committee:

Advisor: Ken Olson

Mike Antolin
Eric Calvo
Alexander Franz
Jeffrey Wilusz

ABSTRACT

SMALL RNAS IN *AEDES AEGYPTI*: ONE GIANT STEP FOR VIRUS CONTROL IN MOSQUITOES

Aedes aegypti mosquitoes are key vectors of medically relevant arthropod-borne (arbo) viruses such as Zika (ZIKV), dengue (DENV1-4), and yellow fever (YFV). When *Ae. aegypti* become infected with arboviruses, RNA interference (RNAi) is a critical antiviral immune mechanism that is a key determinant for successful virus transmission. The major antiviral pathway is the RNAi small-interfering RNA (siRNA) pathway, although evidence shows that the Piwi-interacting RNA (piRNA) pathway also acts as an important RNAi mechanism for controlling persistently infective viruses. The overarching goals of this work were twofold: (1) to determine the potency of the *Ae. aegypti* siRNA pathway against Zika virus and (2) to understand molecular mechanisms underlying piRNA-mediated antiviral immunity and its implications on mosquito vector competence. To achieve these goals, we (1) engineered transgenic *Ae. aegypti* mosquitoes that synthetically triggered the endogenous siRNA pathway against ZIKV and then quantified virus resistance in these mosquitoes, (2) sequenced small RNAs (sRNAs) of the mosquito virome that may impact vector competence and virus persistence, and (3) characterized structural features of Piwi4, an antiviral protein, involved in sRNA binding and subcellular localization to gain insights on its role in the piRNA and siRNA pathways.

A major challenge in the fight against arboviruses is the lack of effective vaccines and limited therapeutic options. Vector control remains the primary method of preventing disease,

and integrated vector management (IVM), including the genetic control of mosquitoes, is imperative to prevent emerging arboviral diseases. To this end, we designed an antiviral effector gene – a ZIKV-specific double stranded (ds) RNA –that synthetically triggered the mosquito's siRNA pathway after a bloodmeal in transgenic *Ae. aegypti*. Small RNA analyses in transgenic midguts revealed ZIKV-specific 21 nucleotide (nt) siRNAs 24 hours after a non-infectious bloodmeal. Nearly complete (90%) inhibition of ZIKV replication was found 7-to-14 days post-infection (dpi); furthermore, significantly fewer transgenic mosquitoes contained ZIKV in their salivary glands ($p = 0.001$), which led to a reduction in the number of ZIKV-containing saliva samples as measured by transmission assay. Our work shows that the siRNA pathway can be synthetically exploited to generate ZIKV-resistant *Ae. aegypti* mosquitoes. In the context of gene drive, antiviral effectors expressed in transgenic *Ae. aegypti* will be an invaluable tool for a population replacement vector control approach.

piRNA-mediated antiviral immunity involves an endogenous viral element (EVE) – viral derived cDNA (vDNA) integrated into host genomes – as well as infection with a cognate virus, which together trigger piRNA amplification and lead to virus silencing. EVEs are from viruses that infected a population in previous generations, and most are derived from insect-specific viruses (ISVs) that persistently infect *Ae. aegypti*. We hypothesized that ISVs and ISV-derived piRNA populations, like EVEs, have geographic structure and impact vector competence to arboviruses. To test this hypothesis, we sequenced sRNAs from geographically distinct *Ae. aegypti* and characterized virus-derived sRNAs (vsRNAs). Overall, the distribution of total sRNAs was highly variable. Small RNAs derived from ISVs were diverse and dependent on geographic origin. We next infected *Ae. aegypti* from Poza Rica, Mexico with DENV2 and analyzed changes in the sRNA virome. DENV2 intrathoracic inoculation resulted in DENV2-

specific siRNAs and piRNAs. We also found increased loads of sRNAs against the ISVs verdadero (*Partitiviridae*: unclassified), Aedes anphevirus (*Xinmoviridae*: *Anphevirus*), and chaq-like virus (*Partitiviridae*: unclassified) after DENV2 infection compared to ISV-derived sRNAs in controls. Overall, our study highlights the diversity of infective ISVs and the complexity of the sRNA virome across *Ae. aegypti* populations, which likely has consequences on sRNA crosstalk, virus replication, and vector competence.

To gain insights on how Piwis, piRNA-binding proteins, are involved in virus control, we characterized structural features of an antiviral Piwi, Piwi4, involved in RNA binding and subcellular localization. We found that Piwi4 PAZ (Piwi/Argonaute/Zwille), the domain that binds the 3'-terminal ends of piRNAs, bound to mature (3'-terminal 2'-O-methylated) and 3'-terminal unmethylated RNAs with similar micromolar affinities (K_D = 1.7 ± 0.8 μM and K_D of 5.0 ± 2.2 μM, respectively) in a sequence independent manner. Through site-directed mutagenesis studies, we identified highly conserved residues involved in RNA binding and found that subtle changes in the amino acids flanking the binding pocket across PAZ proteins had significant impacts on binding behaviors, likely by impacting protein secondary structure. We also found that Piwi4 was both cytoplasmic and nuclear in mosquito tissues, and we identified a Piwi4 nuclear localization signal in the N-terminal region of the protein. These studies provide insights on the dynamic role of Piwi4 in RNAi and pave the way for future studies aimed at understanding Piwi4 interactions with diverse RNA populations.

ACKNOWLEDGEMENTS

I am grateful for many mentors and colleagues with whom I have made lifelong connections and friendships during my academic journey. This work would not have been possible without the support of my PI and mentor, Dr. Ken Olson, who provided me the expertise, resources, and flexibility to grow as an independent scientist. Dr. Eric Calvo, who gave me the opportunity to work in his lab for a year at the NIH, challenged me in new sectors of biology while guiding me to pursue my own ideas. Dr. Alexander Franz hosted me in his lab at the University of Missouri in the Summer of 2019 where I learned techniques in transgenesis, and who I could always count on for thorough manuscript editing and advice. Drs. Jeff Wilusz and Mike Antolin provided their expertise in RNA biology and population genetics, respectively, and whose insights and experimental suggestions benefited this work enormously.

I would not be here without the many staff scientists, post-docs, and fellow PhD students with whom I learned from and worked with throughout my journey. Dr. Irma Sanchez-Vargas worked by my side every day at CSU, and she trained me and shared her many years of expertise in arbovirology with me in the BSL3, as well as her coffee. I could always rely on my good friend Dr. Bill Reid for thorough advice and stimulating discussions, and who taught me everything transgenics and gene drive. Our lab manager Susi Bennett who helps with anything and everything, be it an experiment, an administrative question, or a laugh. My Calvo lab family – Drs. Ines Martin-Martin, Paola Carolina Valenzuela Leon, Gaurav Shrivastava, and Leticia Smith – who taught me the ins and outs of many new lab techniques and brainstormed and troubleshot all things my projects. Marylee Layton, Dr. Claudia Ruckert, and Dr. Mark Stenglein trained and advised me on many small RNA sequencing projects. Drs. Karla Saavedra Rodriguez

and Teca Magalhaes provided us mosquitoes from Mexico and Brazil, respectively, that were essential for our studies. My colleagues who encouraged and supported my PhD adventure before it began: Drs. Serap Aksoy, Brian Weiss, Geoff Attardo, and Aurelien Vigneron at Yale, and Drs. Carolina Barillas-Mury and Alvaro Molina-Cruz at NIH. Finally, the too-many-to-list CSU CVID/AIDLers, as well as my friends at NIH/NIAID/LMVR, who were always eager to lend a helping hand or ear.

This work is the culmination of multi-institutional collaborations of which I am proud to be a part. In addition to those already listed, I have worked closely with Dr. Nelson Lau at the University of Boston, Dr. Mac Fraser at the University of Notre Dame, Dr. Dawn Wesson and Brendan Carter at Tulane University, and Drs. Apo Gittis, Sundar Ganesan and Glenn Nardone in the Research Technology Branch at the NIH. Space does not allow me to elaborate on how much these people taught and supported me in various aspects of these studies.

I also thank those that supported my professional growth outside the lab. Heidi Runge and Stephie Zellner, our graduate student coordinators, Dr. Mark Zabel, who recruited me, the Supporting Healthy Graduate Students committee, and the CSU Graduate School and MIP Department as a whole, all who supported me financially and emotionally.

Lastly, I must acknowledge my family and friends who have stuck by my side throughout this entire process. My partner in life and science, G, discussed every experiment, celebrated every success no matter how small, and pep-talked every failure. And moved me across the country no less than four times. My fortuitous East-meets-West FoCo family – Aine Lehane, Katie and Joe Bonner, and Caroline Benham. My gal Alaina Perkins. My mom and brother who shouldered every complaint. And my dad, who will always be our inspiration for the endless pursuit of knowledge.

DEDICATION

In memory of my dad, Dr. Mark Lane Williams.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

As you are now so once were we

- James Joyce, *Ulysses*

CHAPTER 1: INTRODUCTION

1.1 The global burden of mosquito-borne arboviruses

1.1.1 Vector-borne diseases: a public health priority

A key driver of global health progress over the past 30 years has been the large reduction in the burden of major infectious diseases, particularly those that impact children [1]. Despite these advances, emerging infectious diseases – particularly viral zoonoses – continue to be serious public health threats [2]. Zoonoses, or diseases that originate in non-human animals but can infect people, have caused most pandemics in recent history [2]. Arthropod-borne (arbo) viral zoonoses are of particular concern because these viruses often infect a wide range of mammalian hosts, which increases the likelihood of spillover [2,3]. The risk that emerging infectious diseases will become global pandemics will continue to increase because it is a significant negative consequence of rapidly increasing globalization, urbanization, and economic growth [4–7]. Because socially disadvantaged populations tend to be disproportionately affected, existing health inequalities also exacerbate infectious disease outbreaks [4,8].

Collectively, vector-borne diseases account for almost a quarter of all infectious diseases [9,10]. Vector-borne diseases cause more than 700,000 deaths annually, and approximately 80% of the world's population is at risk of contracting one. The burden disproportionately impacts poor communities in tropical and sub-tropical regions [9,10]. Vector-borne disease life cycles are influenced by complex interactions between pathogens, vectors, and hosts (Figure 1.1). As such, the distribution of these diseases is determined by many biological, environmental, and societal factors, each of which must be addressed to break transmission cycles.

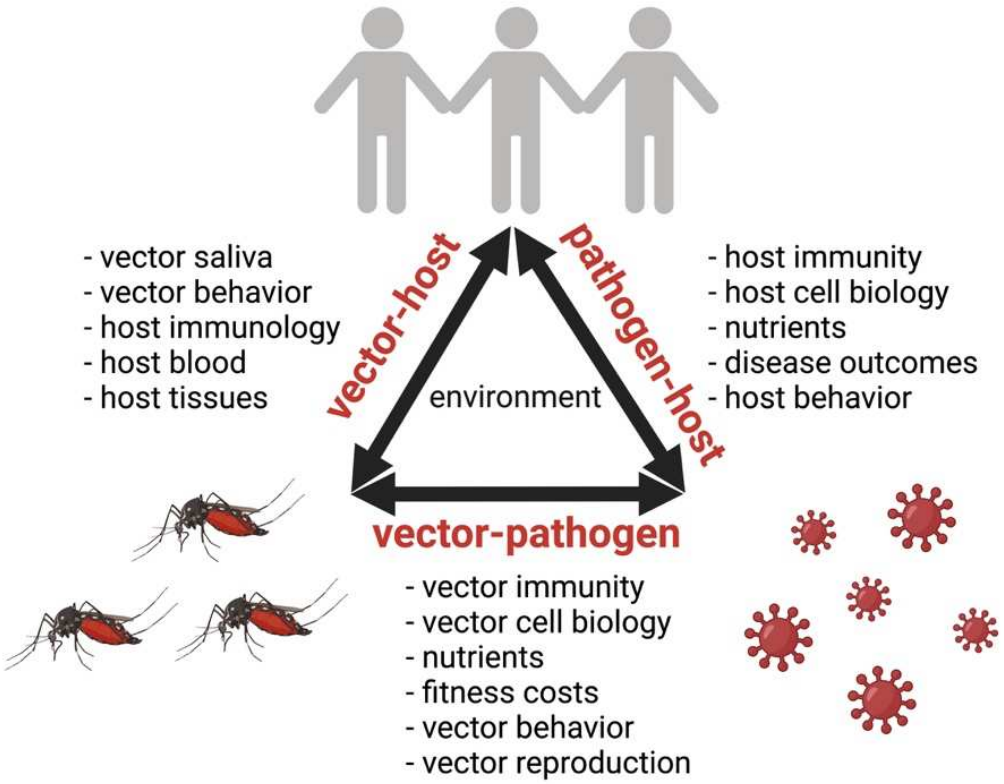


Figure 1.1. Vector-borne disease life cycles are impacted by unique intersections between the vectors, hosts, and pathogens. Vector-pathogen, vector-host, and pathogen-host interactions (highlighted in red) are influenced by a myriad of variables, outlined to the side, and all interactions are dictated by the natural and built environment. Figure made with BioRender.com

Vector-borne diseases, particularly arboviruses, are increasingly prevalent, and the subsequent risk that they re-emerge and metamorphosize into an epidemic is naturally up as well [11,12]. Examples abound. West Nile virus, introduced into the United States in 1999, has become one the most broadly distributed arboviruses in the world and has caused more than 50,000 clinical cases in the United States alone [13,14]. There have been significant yellow fever outbreaks in the Americas (Brazil) and Africa (Angola, the Democratic Republic of the Congo, and Nigeria) in the past five years [15]. Dengue viruses, which already caused approximately 40,000 cases every year [9], resurged in Brazil in 2019 and caused more than two million cases [16]. The Zika virus outbreak in 2015-2016 impacted an estimated one hundred million people in

the Americas alone [17]. These diseases, along with other emerging arboviruses such as Rift valley fever, Mayaro, chikungunya, and tick-borne encephalitis viruses [18], are serious public health threats. The sheer diversity of arboviruses and the wide-reaching effects they have on human health underscore the continued importance of monitoring and addressing them.

The major contributors to vector-borne disease re-emergence are shifts in public health policies, biological changes in the pathogens and vectors, and societal and demographic changes in the affected hosts [13,19]. From the 17th to 20th centuries, vector-borne diseases were the leading cause of human illness [19]. Emphasis on vector control consequently dominated public health approaches [19]. Unfortunately, funding decreases have led to decaying infrastructure, eroded social support, and an overreliance on traditional methods. Overuse of insecticides and chemotherapeutic drugs has selected for resistance in vectors and pathogens, respectively. Climate change also heightens the burden of vector-borne disease because vector abundance, reproduction, survival rates, feeding activities, and the rate of pathogen development within the vector all increase with increasing temperatures [10,13,20–22].

1.1.2 Arbovirus transmission by their vectors

There are more than 500 arboviruses in the Centers for Disease Control and Prevention (CDC) Arbovirus catalog [23], most of which are zoonoses [24,25]. Arboviruses infect and multiply in their respective arthropod vectors at high enough titers so that they can be transmitted to new susceptible mammalian hosts through multiple routes of transmission, including during bloodmeal probing. The vectors themselves can be infected by ingesting a viremic bloodmeal or, to a lesser extent, through transovarial, venereal, or vertical transmission [26]. While arboviruses may have significant morbidity or mortality on host populations, they typically do not have deleterious effects on their vectors [26].

Arboviruses lie within seven virus families: *Togaviridae*, *Flaviviridae*, *Phenuiviridae*, *Nairoviridae*, *Peribunyaviridae, Reoviridae*, *Rhabdoviridae*, and *Asfarviridae*. With the notable exception of African swine fever virus, they are RNA viruses that often have envelopes. Generally, during a productive infection, the arbovirus glycoprotein will first attach to a susceptible host cell receptor, penetrate the plasma membrane, and enter the cell. The virus capsid disassembles, and virus genomic RNA is released into the cell cytoplasm. At the endoplasm reticulum, host ribosomes translate viral proteins, and the virus RNA-dependent RNA polymerase then replicates the virus genome. Viral nucleocapsids bud from the ER lumen intracellularly, and once they are fully assembled, they are released out of the cell [26] (Figure 1.2).

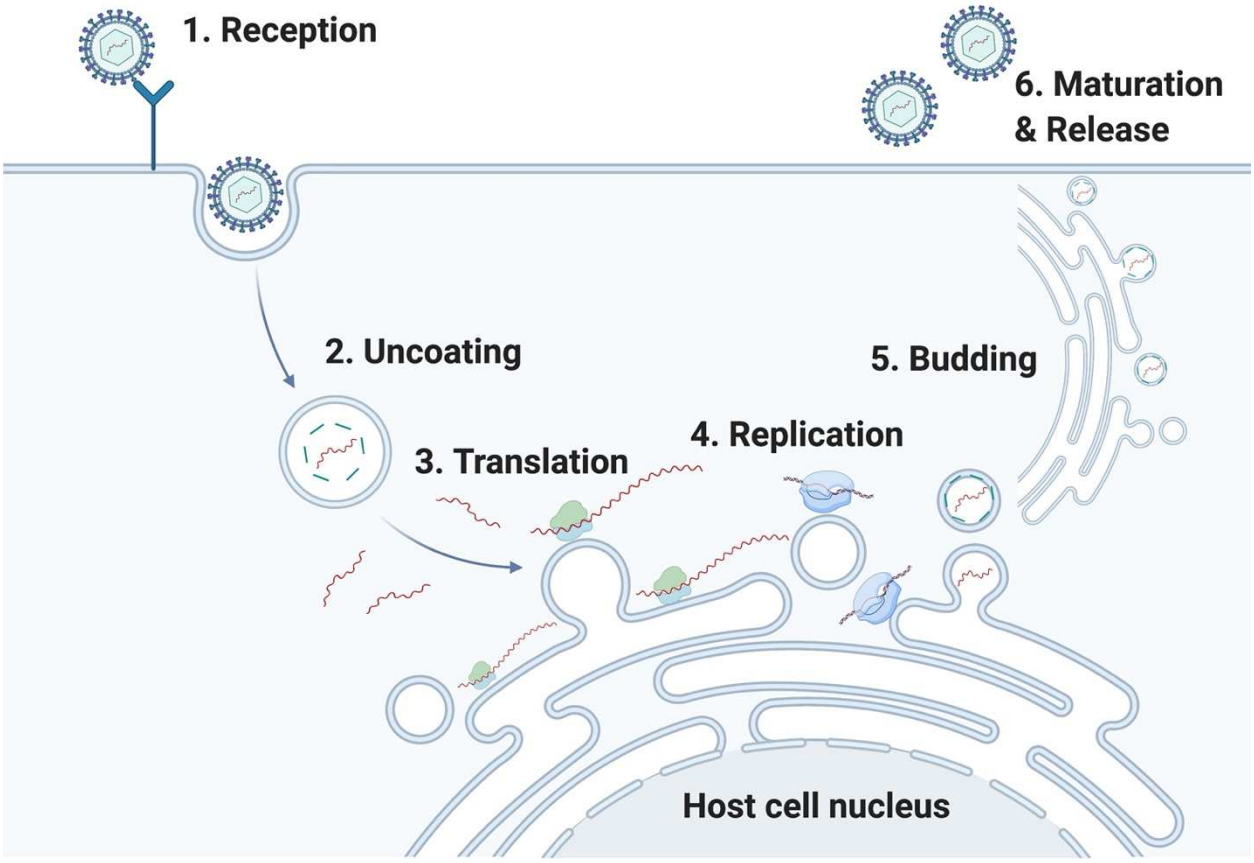


Figure 1.2. Generalized schematic illustrating positive-stranded RNA arbovirus replication in host cells. (1) Reception: the virion glycoprotein attaches to a host cell receptor, **(2) Uncoating:** virus enters cells by pinocytosis or by direct fusion with the plasma membrane, and the nucleocapsid enters the cytoplasm from the plasma membrane or from a endosome and releases its RNA, **(3) Translation:** viral proteins are produced on host ribosomes, **(4) Replication:** viral RNA is transcribed and replicated, **(5) Budding:** virus proteins and nucleic acids coalesce and **(6) Maturation and Release:** mature through host cell membranes and are released from the cell. **3-4** occur simultaneously during the "eclipse" stage, when infectious virions cannot be detected. Figure made with BioRender.com. Adapted from Biology of Disease Vectors, Chapter 14 [26].

The overall ability of a vector to transmit a pathogen at a given time is dependent on numerous extrinsic factors, including vector population size, longevity, length and number of gonadotrophic cycles, feeding behaviors, and diel activities. Together, this is known as vectorial capacity. It can be mathematically quantified by Macdonald's Equation [27]:

$$V = \frac{m \times a^2 \times p^n \times b}{-\ln(p)}$$

where m = vector density in relation to the host, a = probability a vector feeds on a host in one day, p = probability a vector will survive one day, n = duration of the extrinsic incubation period (EIP), and b = vector competence [27]. Based on this equation, a, p, and n most strongly affect vectorial capacity because a is squared (this is because a vector must feed twice to transmit a pathogen), n is an exponent of p and has a nonlinear effect, and p is in both the numerator and denominator [27].

For transmission to occur, a vector must also be biologically susceptible to the infection and permit pathogen reproduction and development. The intrinsic ability of a vector to transmit a pathogen is known as vector competence [26]. In the laboratory, it is often quantified by the percentage of females with virus in the saliva after the appropriate EIP [28]. The EIP is the period between virus uptake after a vector feeds on a viremic host until the vector can transmit virus, which for a mosquito is in her saliva during probing. While small changes in feeding rates

or mosquito population size may have significant impacts on vectorial capacity, based on Macdonald's Equation, vector competence (*b*) must essentially be driven to zero to have significant impacts on vectorial capacity. These are important considerations when designing vector control strategies.

The arthropoda phylum is highly diverse, and most arthropods are unable to transmit arboviruses. As such, the arthropod-arbovirus relationship is specific. For example, mostly only culicine mosquitoes transmit mosquito-borne arboviruses. The primary exception is O'nyong-nyong virus, which is transmitted by anopheline mosquitoes; however, there are reports of anophelines transmitting other arboviruses, for example, Mayaro virus (*Togaviridae*: *Alphavirus*) [29]. Of the culicines, mosquitoes in the genera *Aedes* and *Culex* are the main incriminated vectors of arboviruses. Most *Aedes*-transmitted arboviruses of the *Flaviviridae* and *Togaviridae* families have either primate reservoirs or circulate between humans and mosquitoes, and major arboviruses of the former family such as dengue (DENV) and yellow fever (YFV) are associated with hemorrhagic disease; on the other hand, most *Culex*-transmitted arboviruses of the *Flaviviridae* and *Togaviridae* families have avian reservoirs and are associated with encephalopathies [30]. The major arthropod species capable of transmitting their respective arboviruses of medical importance are summarized in Table 1.1 at the end of this chapter.

1.1.3 Challenges in arbovirus control

A major challenge in the fight against arboviruses is the lack of effective vaccines and a dearth of therapeutic options. Indeed, there are few licensed vaccines against arboviruses available. The most successful include the live-attenuated YFV 17D vaccine [31] (although YFV remains endemic in many parts of the world due to limited access [11]), live attenuated and live recombinant Japanese encephalitis vaccines [32], and an inactivated tick-borne encephalitis

vaccine [33] that was recently approved for use in the USA [34]. It is challenging to develop effective vaccines because many arboviruses are of similar structure and therefore induce antibodies that may cross react. This makes it harder to assess vaccine efficacy of vaccines and develop effective diagnostics; in the case of DENV, cross-reactive antibodies against different serotypes may exacerbate disease outcomes. This means there is a risk that a DENV vaccine could enhance pathogenesis for DENV-naïve individuals that live in DENV-endemic areas, as was reported for the Dengvaxia vaccine [11]. Furthermore, because arbovirus outbreaks are unpredictably seasonal, the timing required to develop an effective vaccine and test it in clinical trials often outlasts the period of the outbreak [11].

Given the challenges of developing safe and effective therapeutics against arboviruses, the key strategy for disease prevention is still reducing pathogen transmission by targeting the vector [35]. Since the sustained use of insecticides selects for insecticide resistance, the incorporation of multiple interventions that interrupt human-vector contact, known as integrated vector management (IVM), is imperative in the fight against arboviruses [35,36].

Table 1.1. Major arthropod vectors that transmit medically relevant arboviruses that infect humans. * genus/species is incriminated but is less efficient vector than that previously listed. Table adapted & expanded upon from those in Agarwal et al., 2017 [37] & Braack et al., 2018 [38].

Arthropod vector		**Medically relevant arbovirus**		**Geographic distribution**	**Selected References**
genus	**major species**	**genus**	**species**		
Aedes *Haemagogus** *Sabethes**	*Ae. aegypti*	*flavivirus*	yellow fever virus (YFV)	Africa Central & South America	Jentes et al., 2011 [39]
Aedes	*Ae. aegypti* *Ae. albopictus*	*alphavirus*	chikungunya virus (CHIKV)	Africa Asia Pacific Americas Southern Europe	Leparc-Goffart et al., 2014 [40] Vega-Rúa et al., 2014 [41]
Aedes	*Ae. aegypti* *Ae. albopictus*	*flavivirus*	dengue virus (DENV)	Africa Asia Pacific Americas	Simmons et al., 2012 [42] Grunnil & Boots, 2016 [43]
Aedes *Culex** *Anopheles** *Eretmapodites** *Mansonia**	*Ae. aegypti* *Ae. albopictus*	*flavivirus*	Zika virus (ZIKV)	Africa Asia Pacific Americas	Boyer et al., 2018 [44] Gutiérrez-Bugallo et al., 2019 [45]
Aedes *Culex*	*Ae. camptorhynchus* *Ae. vigilax* *Cx. annulirostris*	*alphavirus*	Ross River virus (RRV)	South Pacific	Russel, 2002 [46]
Aedes *Culex*	*Ae. aegypti* *Ae. albopictus* *Ae. vexans* *Cx. pipiens* *Cx. antennatus*	*phlebovirus*	Rift valley fever virus (RVFV)	Africa Middle East	Nielsen et al., 2020 [47] Tantely et al., 2015 [48]
Aedes *Culex*	*Ae. triseriatus* *Ae. albopictus* *Cx. pipiens* *Cx. restuans*	*orthobunyavirus*	La Crosse virus (LACV)	United States	Harris et al., 2015 [49]
Aedes *Culex*	*Ae. triseriatus* *Cx. taeniorhynchus*	*alphavirus*	Venezuelan equine encephalitis virus (VEEV)	Americas	Brault et al., 2004 [50]

Anopheles	*An. funestus* *An. gambiae*	*alphavirus*	o'nyong nyong virus (ONNV)	sub-Saharan Africa	Rezza et al., 2017 [51]
Culiseta	*Cs. melanura* *Cs. morsitans*	*alphavirus*	eastern equine encephalitis (EEEV)	Americas	Armstrong & Andreadis, 2010 [52]
Culicoides *Aedes** *Culex**	*Culicoides paraensis* *Ae. triseriatus** *Cx. quinquefasciatus**	*orthobunyavirus*	Oropouche orthobunyavirus (OROV)	Central & South America	Sakkas et al., 2018 [53]
Culex	*Cx tritaeniorhynchus* *Cx. gelidus* *Cx. quinquefasciatus*	*flavivirus*	Japanese encephalitis virus (JEV)	Asia	Lindahl et al., 2012 [54]
Culex	*Cx. pipiens* *Cx. tarsalis* *Cx. quinquefasciatus*	*flavivirus*	West Nile virus (WNV)	Africa Europe Asia United States	Nash et al., 2001 [55]
Culex *Coquillettidia** *Culiseta**	*Cx. univittatus* *Cx. pipiens**	*alphavirus*	Sindbis virus (SINV)	Europe Asia Africa Oceania	Adouchief et al., 2016 [56]
Haemagogus *Aedes** *Anopheles**	*Hag. janthinomys*	*alphavirus*	Mayaro virus (MAYV)	Central & South America	Ali et al., 2019 [57] Brustolin et al., 2018 [29]
Hyalomma	*H. marginatum* *H. rufipes* *H. anatolicum* *H. asiaticum*	*orthonairovirus*	Crimean-Congo hemorrhagic fever orthonairovirus (CCHFV)	Africa Asia Middle East	Papa et al., 2017 [58]
Ixodes *Dermacentor* *Haemaphysalis*	*Ix. ricinus* *Ix. persulcatus* *Ix. ovatus* *Dermacentor reticulatus* *Haemaphysalis concinna*	*flavivirus*	tick-borne encephalitis virus (TBEV)	Europe Asia	Chitimia-Dobler et al., 2019 [59]
Ixodes	*Ix. scapularis* *Ix. cookei* *Ix. marxi*	*flavivirus*	Powassan (POWV)	United States Canada Russia	Hermance & Thangamani, 2017 [60]

1.2 *Aedes aegypti*-transmitted arboviruses and mosquito antiviral immunity

1.2.1 Aedes aegypti *biology and epidemiology*

Aedes aegypti (Diptera: Culicidae), or the yellow fever mosquito, is the major vector of medically relevant arboviruses, including YFV, DENV1-4, Zika (ZIKV), and chikungunya (CHIKV). Native to West Africa, it was introduced to the New World about 500 years ago through the slave trade [61]. Ancestral populations in sub-Saharan Africa, known as *Ae. aegypti formosus*, still exist in tree holes within the forests, and their preferred bloodmeals are nonhuman mammals. They are generally not considered efficient vectors of yellow fever or dengue viruses [62]. However, the domesticated form, *Ae. aegypti aegypti*, breeds in urban settings with artificial containers and prefers to feed on humans. This adaptation has enabled their global spread. Where and when domestication occurred is controversial, but genetic analyses suggest it was around 400-550 years ago when human settlements began to form near forests in West Africa; there are mosquito populations with mixed *Ae. aegypti formosus* and *Ae. aegypti aegypti* genetic backgrounds in Senegal and Angola, which may represent the origins of the domestic lineage [61]. Since then, *Ae. aegypti* have become highly invasive and have colonized tropical and subtropical regions, even highlands and semi-arid regions. They are currently found on every continent except Antarctica [63].

Like all mosquitoes, *Ae. aegypti* undergo complete metamorphosis and exhibit an egg stage, four aquatic larval stages, an aquatic pupal stage, and a non-aquatic adult stage (Figure 1.3). Aedine mosquitoes lay individual eggs in moist environments, which are resistant to drying and can survive for several months. The eggs hatch in wet conditions into first instar larvae that are barely visible to the naked eye.

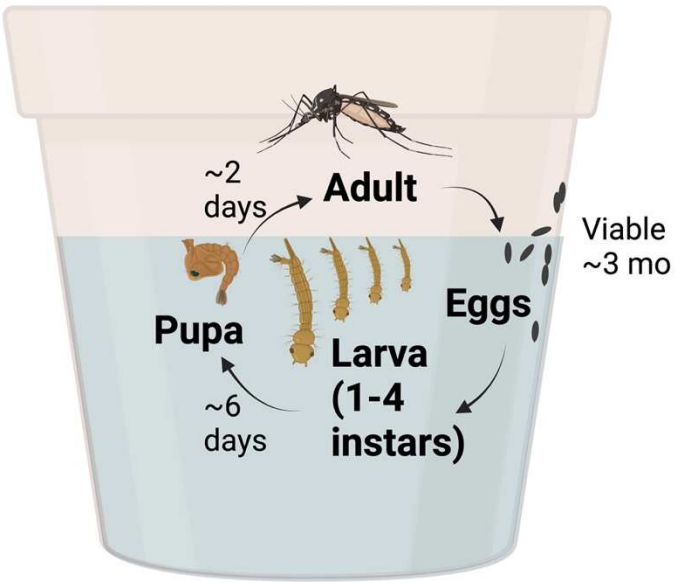


Figure 1.3. *Aedes* life cycle. *Ae. aegypti* undergo complete metamorphosis and exhibit an egg stage, four aquatic larval stages (1st - 4th instars), and an aquatic pupal stage, followed by the non-aquatic adult stage. Eggs are viable for ~ 3 months. In moist, favorable conditions, eggs hatch into 1st instar larvae that, over ~ 6 days, develop into 2nd, 3rd, and 4th instar larvae. Pupa require two days for development into adults. Figure made with BioRender.com.

Larvae use an elongated hind air tube to breathe and feed on organic matter in the environment. Larval habit strongly impacts mosquito fitness across development stages. Studies have shown that the type and location of the breeding container is significantly associated with the body and wing size of adult females in the field [64]. In controlled settings, the interactions between water temperature, diet, and initial larval density have been shown to significantly impact hatching rates, developmental rates, time to adult emergence [65,66]. Couret and colleagues argued that the nutrient requirements needed to trigger hormonal developmental cascades – and thus progress to subsequent life stages – are dependent on both water temperature and larval density [65]. Many years ago, it was observed that high *Ae. aegypti* larval densities were associated with high larval mortality, long larval development times, and decreased adult sizes [67]. These effects were independent of the amount of available nutrients [67]. Wada attributed these effects to intraspecific physical interference that force larva to competitively

sequester resources from the environment [67,68]. Chemical interference is also a factor: larva were shown to produce growth retardants that stalled the development of other larvae at times of high competition [68]. It has since become clear that density-dependent effects, and compensatory or overcompensatory mortality, are more common in container-breeding mosquitoes such as *Ae. aegypti* [69]. Due to intraspecific competition, the larvae stratify into populations that develop at different speeds, depending on resource availability. Larval biology therefore has important implications for vector control. Targeting the larval stages with biological (*e.g. Bacillus thuringiensis israelensis* [BTI] application) or chemical (*e.g.* oils or film) applications is a key strategy for inhibiting adult mosquito fitness and emergence. However, if this control strategy is performed incompletely, it can counterproductively promote delayed larval development [69]. Larval health also impacts adult vector competence [70]. Herd and colleagues found that starved *Ae. aegypti* larva had significantly thinner midgut basal laminas as adults and exhibited significantly higher rates of ZIKV dissemination compared to larva reared with ample nutritional sources [70]. Collectively, these studies underscore the importance of larval health on the mosquito's entire lifespan as well as its role as an arbovirus vector.

After larvae go through all four instars, they develop into pupa. In favorable conditions, pupation occurs approximately six days after hatching. Mosquito pupae are easily distinguishable from larvae; they are C-shaped and have two superior respiratory trumpets. Approximately two days after pupation, pupa metamorphosize into adult mosquitoes and leave the water. Adults have three parts – a head, thorax, and abdomen – with slender wings and a proboscis. *Ae. aegypti* adults have characteristic white scales on brown thoraxes in the shape of a lyre, which makes them morphologically easy to identify. Only adult females bloodfeed, and as such, their

mouthparts and proboscises are well-adapted for piercing and sucking [71]. While salivary pumps inject saliva containing proteins that facilitate bloodfeeding, suction pumps in the head create the pressure necessary for blood uptake [71,72].

The female proboscis is made up of a variety of parts called stylets. The four types of stylets on *Ae. aegypti* females are the labrum, mandibles, laciniae, and hypopharynx, which all lie at their base in a groove of the labium (Figure 1.4). When the mosquito is feeding, the labella presses against the host and bends backward, while the stylets penetrate and the laciniae cut the skin [71]. When the mosquito is resting, the piercing structures are enveloped by the labium, and the proboscis is folded under the body [71]. Adult males are generally smaller than females and have distinguishably hairy maxillary palps; because they feed on sugar, their proboscises are adapted for sucking plant juices and have smaller piercing structures than females [71,73].

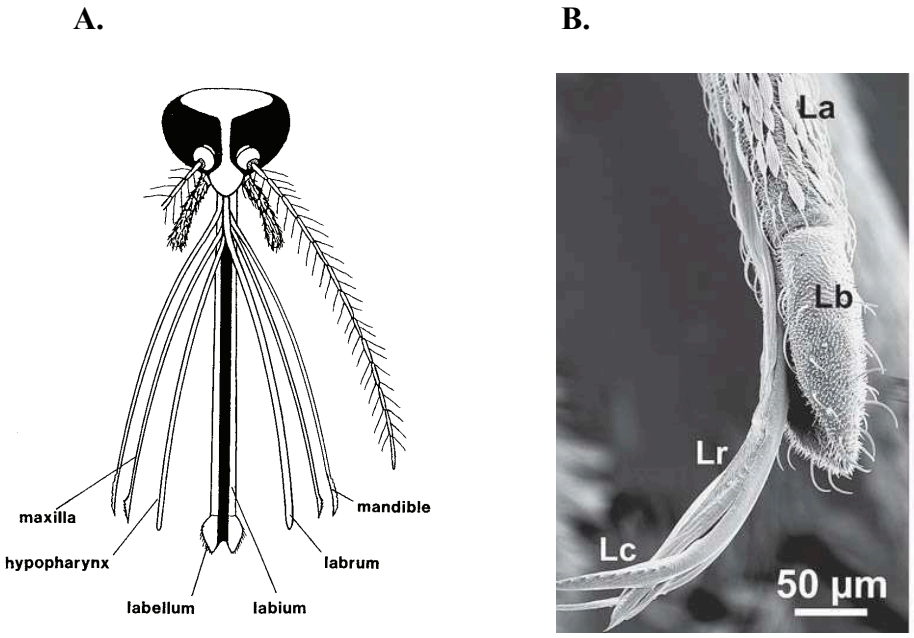


Figure 1.4. Female culicine mouthparts. A. The head of a female culicine, indicating mouthparts. The labium serves as the proboscis sheath; the labrum and hypopharynx serve as the food canal; the hypopoharynx serves as the salivary canal. Image from Baranitharan et al., 2018, "Introduction of Vector Mosquitoes" [74]. **B.** Electron micrograph of *Anopheles stephensi* (Diptera: Culicidae) piercing structures. The labrum, mandible, and laciniae serve as the puncturing organ during feeding; the lacinial teeth provide anchorage, while the lacinia is the

piecing organ. La = labium; Lb = labella; Lr = labrum; Lc = laciniae. Scale bar = 50 μM. Image from Krenn & Aspock, 2012 [71].

Adult mosquitoes first mate between 36 to 48 hours after emergence [75]. Reproduction is sexual and occurs in three stages: coupling, copulation, and insemination [75]. Egg fertilization does not occur during mating; even though the male injects seminal fluids into the female during insemination, the fluids are stored in the spermathecae. *Ae. aegypti* have three spermathecae – one large medial one and two smaller lateral ones. Most female mosquitoes store sperm in at least two of the lobes [76]. Copulation and insemination last between six to ten seconds, and enough sperm is injected to last the entire lifetime of the mosquito [75,77]. Females are generally considered monogamous, although reports of polyandry have been reported when the initial mating event was interrupted [75]. Males are polygamous. Male *Ae. aegypti* can effectively inseminate at least five females and can mate at least three times without it impacting fertility [78,79] . Males recover fertility every two days [79]. After they are injected, seminal fluids impact female behavior by driving her to host seek, among other effects [75]. *Ae. aegypti* typically mate near their hosts [78,80].

Both female and male mosquitoes require carbohydrates from sugars, but only females are hematophagous because they need blood to reproduce. Sugars are stored in the ventral diverticulum (known also as “the crop”), an elastic sac that branches from the alimentary canal in the thorax. Over time, sugars are re-routed from the crop to the midgut [81]. Blood, which supports ovarian development, goes directly into the midgut. To accommodate bloodfeeding, the mosquito exoskeleton is not rigid. It is made up of hard plates called tergites and sternites. These are connected by less sclerotized membranes, known as the pleura, that become elasticized after a bloodmeal [82]. The products of blood digestion accumulate in the hemolymph – the

mosquito's "blood"– which are rapidly sequestered into the fat body. This triggers the release of juvenile hormone (JH), which stimulates the fat body to begin vitellogenesis, the process of yolk and egg formation. As vitellogenesis proceeds, the oocyte grows and activates stretch receptors that signal the brain to release allatostatins, which blocks JH production and slows down vitellogenesis [82]. As JH wanes, the steroid 20-hydroxyecdysone (20E) increases, which directs the synthesis of yolk protein precursors (YPPs) in the fat body [83]. Vitellogenins (Vg) and YPPs pass between ovarian follicular cells and are absorbed by oocysts through receptor-mediated endocytosis.

Bloodmeal host seeking is a complex process that is influenced by hormones and life cycle stage. For example, recently emerged or blood engorged females do not host seek; in the latter case, stretch receptors in the abdomen suppress bloodfeeding [84]. Because mosquitoes only associate with hosts when they need a bloodmeal, complex physiological signals coordinate bloodfeeding. Mosquitoes use visual and chemical cues to locate viable hosts. For example, carbon dioxide and water vapor contained in breath, as well as lactic acid from sweat, attract host-seeking mosquitoes by interacting with receptors on the antennae [82]. *Aedes* mosquitoes are crepuscular feeders, meaning they host seek in the mornings and evenings ("at dawn and dusk."); however, *Ae. aegypti* specifically are aggressive feeders. They often partially bloodfeed multiple times a day, including during midday [82]. They can also bloodfeed multiple times during their lifetime and can undergo multiple gonotrophic cycles [26]. They tend to feed on moving objects closer to the ground because their attractants, like carbon dioxide, are heavier than air [82]. Mosquitoes salivate while they bloodfeed. Mosquito saliva is a complex mixture of antiplatelets, anticoagulants, vasodilators, and immunomodulators. To facilitate bloodfeeding, these molecules reduce pain and inflammation and counteract host hemostasis [85]. Many of the

unique proteins that are overexpressed in female salivary glands are secreted, and affect not only host physiology, but pathogen transmission as well [86–89].

1.2.2 The route of a virus in a mosquito and virus infection barriers

After infectious blood is ingested into the mosquito midgut, virus begins replicating there. For successful arbovirus transmission to occur, the virus must invade the midgut epithelium, disseminate from the midgut, and replicate in the mosquito's secondary tissues. After virus infects the salivary glands, it is released into the salivary ducts. At this point, infected saliva can be egested during probing, which enters the host through the bite site [90] (Figure 1.5).

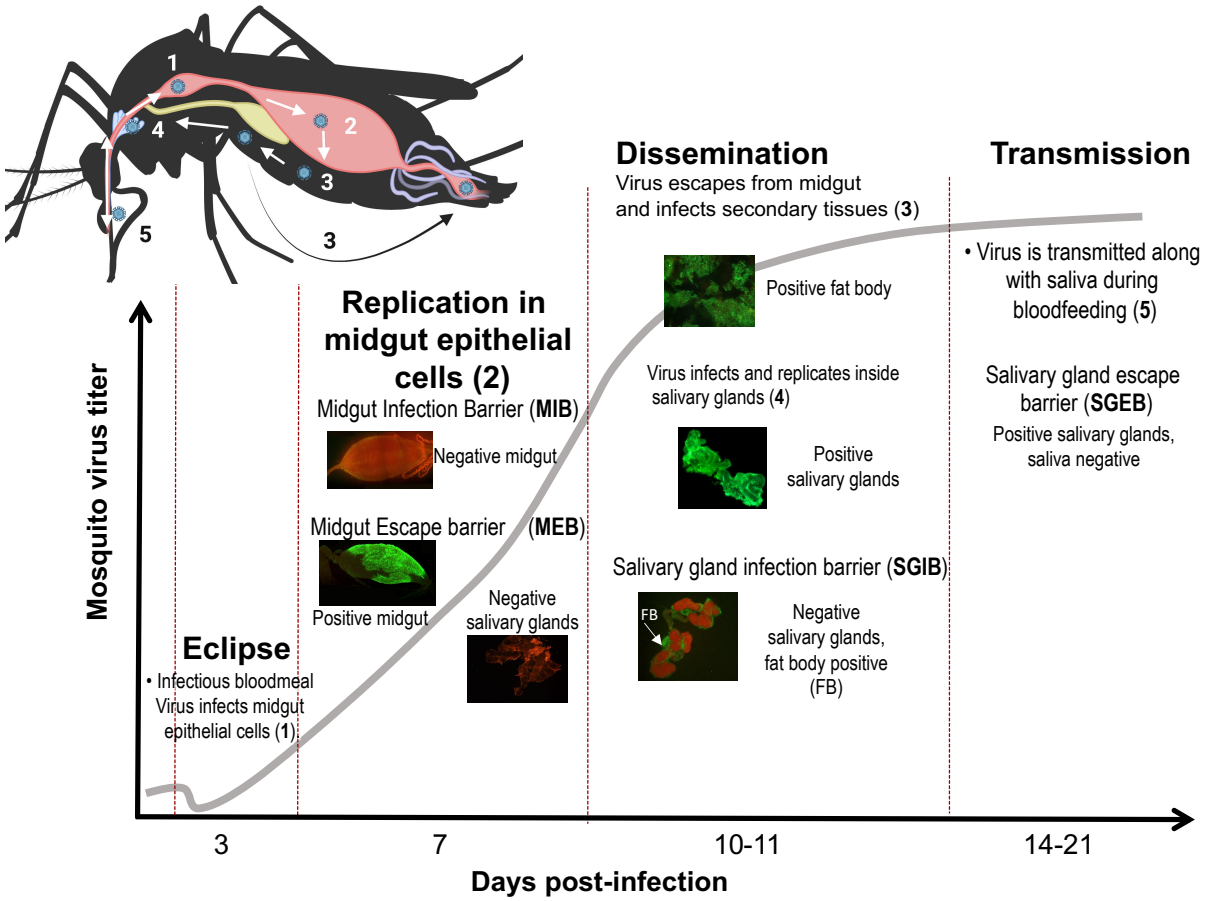


Figure 1.5. Arbovirus infection of and transmission by a mosquito. (1) A mosquito first ingests a viremic bloodmeal into the midgut; during the eclipse phase, while virus is replicating, no infectious virus is detectable. **(2)** Virus replicates in midgut epithelial cells. If a mosquito exhibits a midgut infection barrier (MIB), virus replication is blocked; if a mosquito exhibits a midgut escape barrier (MEB), virus infects the midgut but is unable to disseminate and infect the salivary glands. **(3)** Virus then escapes the midgut epithelium, disseminates into secondary tissues such as the ovaries or fat body (FB), and **(4)** infects the salivary glands. If a mosquito exhibits a salivary gland infection barrier (SGIB), secondary tissues are infected while salivary glands are uninfected. Virus is then **(5)** secreted into the saliva and egested during a subsequent bloodmeal. If a mosquito exhibits a salivary gland escape barrier (SGEB), salivary glands are infected while saliva remains uninfected. Tissue IFAs stained with a pan-flaviviral antibody in green. Days post-infection (dpi) reflects DENV and ZIKV replication dynamics in *Ae. aegypti* [91,92] and differs between virus families. Figure gratefully reproduced and adapted from Dr. Irma Sanchez-Vargas. Mosquito figure made with BioRender.com.

Vector competence is determined by a mosquito's natural barriers against infection. Mosquitoes can exhibit four different types of infection barriers: midgut infection (MIB), midgut escape (MEB), salivary gland infection (SGIB) and salivary gland escape (SGEB) barriers (Figure 1.5). MIBs block initial infection of the midgut, while MEBs block virus dissemination from an infected midgut. SGIBs prevent viral infection of the salivary gland basal lamina. Thus, mosquitoes with SGIBs will have disseminated infections, but will lack infected salivary glands. Finally, mosquitoes with SGEBs have infected salivary glands but their saliva is uninfected [93].

Of the four barriers, the molecular mechanisms underlying MIBs are the best characterized across vector-arbovirus pairings. MIBs occur when a virus is unable to infect midgut epithelial cells or is unable to spread to other cells [30]. Examples include mosquito innate immune responses and the absence of suitable epithelial cell receptors in the midgut to which the virus can bind [30]. For example, *Ae. aegypti* use RNA interference (RNAi) as an antiviral immune response when virus enters midgut cells and forms long double stranded RNA (dsRNA) intermediates. Several midgut receptors necessary for arbovirus infection have also been identified for mosquitoes, including C-type lectin family pattern recognition receptors (for

SINV [94] and DENV [95]), divalent metal ion transporters (for SINV [96]), laminin-binding proteins (for DENV2-4 [97]), an Hsp 90-like protein (for DENV4 [98,99]), and enolase (for DENV2 [100]). Interestingly, only a small number of cells in the epithelium are susceptible to virus infection, and different viruses display different cell infection patterns [30]. For example, DENV2 does not infect the anterior of the midgut, while Eastern equine encephalitis (EEEV) infects only 30% of midgut cells [30].

Although the mechanisms underlying MEBs are less understood than MIBs, physical blocking of virus by mosquito tissues such as the basal lamina or the peritrophic matrix is thought to be a factor. The basal lamina is a proteinaceous matrix in the extracellular matrix that is secreted by epithelial cells and continuously renewed [30]. The basal lamina is porous and densely packed with collagen type IV, laminin, perlecan, and entactin [30]. During bloodfeeding, the mosquito gut expands and overstretches; tears and gaps caused by degradation and re-synthesis of the basal lamina have been shown to be an escapeway for arboviruses [30,101–103]. For example, CHIKV dissemination in *Ae. aegypti* was coincident with increased collagenase activity, diminished collagen IV abundance, and basal lamina shedding in the midgut approximately 24 hours post-bloodmeal [101]. Overexpressing an inhibitor of matrix metalloproteinases – proteins that exhibit collagenase activity in the midgut epithelium – significantly increased CHIKV midgut dissemination in *Ae. aegypti* [101]. The authors attributed this phenotype to basal lamina degradation and/or restoration [101]. Subsequent electron microscopy studies showed chikungunya virions accumulated at strands of the basal lamina 24-32 hours post-bloodmeal, and that a second bloodmeal caused virions to again congregate in the midgut basal lamina [102]. Increased ZIKV and DENV2 dissemination was also reported in *Ae.*

aegypti that received two bloodmeals [91,103]. Bloodmeal-induced microperforations in the basal lamina were attributed to increased ZIKV transmission [103].

Mechanisms underlying salivary gland infection and escape barriers remain elusive. It has been suggested that the salivary gland basal lamina is an SGIB for Rift Valley fever virus in *Aedes* [104], much like the midgut basal lamina that acts as an MEB. The lateral and median salivary gland lobes display different surface proteins that could be necessary for entryway for specific viruses [105,106]. For example, heparan-sulfate proteoglycans – which occur in the internal ducts of the lateral lobes but are absent in the median lobes – are necessary for salivary gland infection by certain strains of SINV [107]. *Ae. aegypti* exhibit low rates of SGIBs against DENV2, but higher rates against ZIKV or CHIKV [93]. On the other hand, SGEBs were found for multiple strains of ZIKV, DENV2, and CHIKV [93]. Studying the genetics underlying the infection barriers in *Ae. aegypti* discussed in this section so far has laid the foundation for the generation of transgenic mosquitoes that inhibit virus infection and replication.

1.2.3 RNA interference and antiviral immunity[1]

As early as 1991, it was observed that certain *Ae. aegypti* strains naturally harbor arbovirus resistance genes that could be transferred to susceptible mosquito populations through selective breeding [108,109]. Quantitative trait loci (QTL) experiments later identified the regions of the *Ae. aegypti* genome that are associated with arbovirus resistance [110]. Specifically, the *dcr2* gene was found to be associated with *Ae. aegypti* resistance to DENV2, suggesting RNA interference was playing a role in mosquito antiviral immunity [111].

[1] This section includes parts of the manuscript cited as "Williams, A.E., Franz, A.W.E., Reid, W.R., Olson, K.E. Antiviral effectors and gene drive strategies for mosquito population suppression or replacement to mitigate arbovirus transmission by *Aedes aegypti*. *Insects*. 2020; *11*(1): 52." The article is reproduced with permission and minor modifications have been made.

RNAi, or post-transcriptional gene silencing, is a conserved biological response to double stranded RNA that results in sequence-specific gene silencing [112]. It was first formally discovered in *Caenorhabditis elegans* when sense and antisense RNA molecules were injected into the nematodes to form dsRNA. The sense and antisense strands were substantially more effective at gene silencing when injected together rather than with either strand alone [113]. This discovery led to the reassessment of homology-dependent gene silencing events that had previously been described in several other systems [112]. For example, before RNAi was formally described, virus-specific dsRNAs had been found to silence virus replication in *Ae. aegypti* mosquitoes and *Aedes*-derived cell lines (C6/36 [*Ae. albopictus*] and Aag2 [*Ae. aegypti*] cells) [114,115]. Targeting a region within the pre-membrane (prM) gene of DENV2, in the sense or antisense orientation from a double subgenomic Sindbis (dsSIN) virus vector, led to significantly lower DENV2 titers in mosquitoes [114] and cells [115]. In retrospect, as the sequence was transcribed from the dsSIN plasmid, dsRNA intermediates probably triggered RNAi. Resistance was sequence-specific because the cells that expressed DENV2-specific prM sequences were resistant to DENV2, but not to DENV3 or DENV4 [115]. Once RNAi was described in 1998, the same team found that expressing a DENV2 dsRNA in Aag2 cells led to significantly less DENV2 titers, as well as an accumulation of 21-25 nt DENV2-specific RNAs, *in vivo* [116].

The best characterized antiviral immune mechanism in *Ae. aegypti* is the RNAi small-interfering RNA (siRNA) pathway [117]. In this pathway, viral dsRNA intermediates are recognized by mosquito Dicer-2 and processed into 21 bp siRNA duplexes. These siRNA duplexes interact with Dicer-2 and R2D2 and are incorporated into the RNA-induced silencing complex (RISC) that contains the endonuclease Argonaute-2. One of the siRNA strands, the

guide strand, directs RISC to complementary (viral) RNA sequences in the cell cytoplasm. Argonaute-2 of RISC then cleaves the targeted RNA (Figure 1.6).

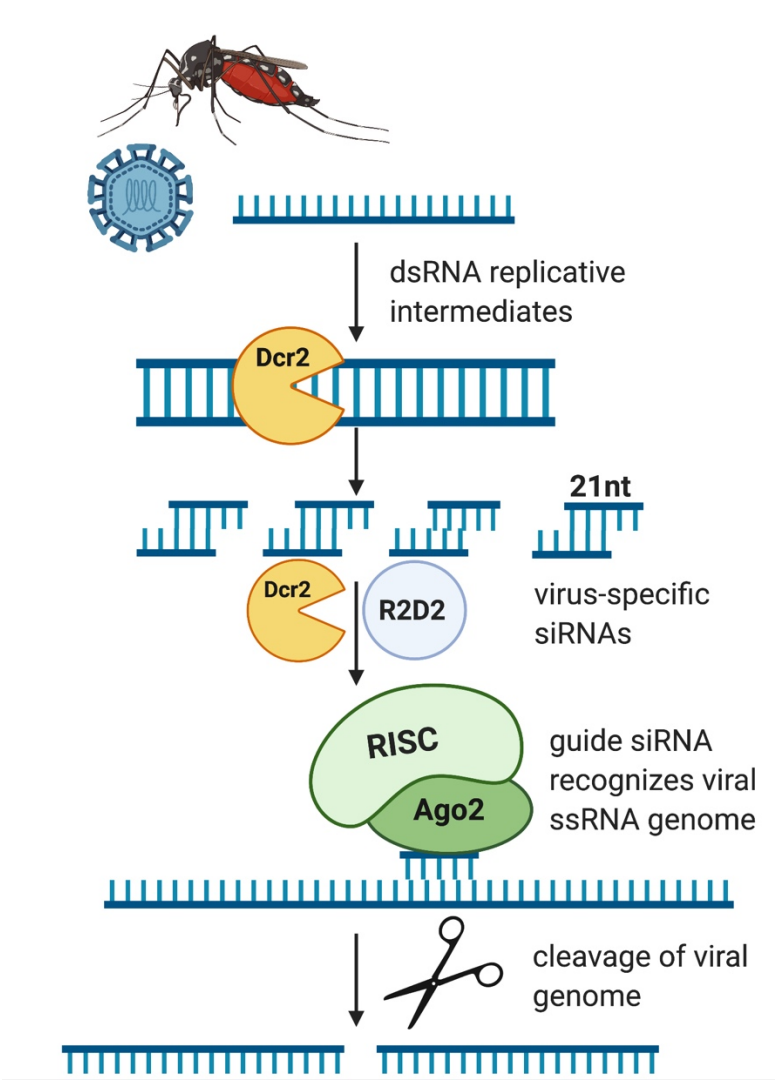


Figure 1.6. The small interfering RNA pathway as an antiviral mechanism in *Ae. aegypti*. Virus replication forms dsRNA intermediates that trigger the endogenous siRNA antiviral pathway. Long dsRNAs are processed by the RNase III enzyme Dicer-2 (Dcr2) into 21 nucleotide (nt) virus-specific siRNAs. siRNAs interact with Dicer-2 and the RNA-binding protein R2D2 and are escorted to the RNAi silencing complex. A single strand derived from the siRNA guides RISC to complementary sequences in the cell, where the endonuclease Argonaute-2 (Ago2) then performs site-specific cleavage on single-stranded RNAs (ssRNAs). Figure made with BioRender.com.

In *Ae. aegypti*, the P-element induced wimpy testes (PIWI)-interacting RNA (piRNA) pathway is another RNAi mechanism involved in antiviral immunity [118–130], as well as embryonic development [131,132] and gene regulation [133,134]. From *Drosophila* studies, the piRNA pathway is known to silence transposable elements (TEs) that integrate into the germline genome and threaten its integrity [135–138]. piRNAs, ~ 24-32 nt small RNAs (sRNA), bind Piwi proteins. piRNA-bound Piwis assemble into piRNA-induced RNA silencing complexes (piRISCs), and effector Piwis are targeted to complementary RNA substrates [139–141]. *Drosophila* express three Piwis: Piwi, Aubergine (Aub), and Argonaute-3 (Ago3). Aub and Ago3 are expressed exclusively in the germline, while Piwi is expressed in the germline but also in neighboring somatic cells [136,142,143]. Aub and Ago3 silence their targets post-transcriptionally in the cytoplasm; Piwi, however, translocates into the nucleus, forms a nuclear effector complex, and silences its targets co-transcriptionally [144–146]. Depletion of the piRNA pathway in *Drosophila* leads to TE insertion accumulation in the genome and consequently DNA damage, defects in embryonic development, and female sterility [147–152].

There are more piRNA pathway functions than initially thought [137,153–156]. For example, across organisms, Piwis display differential expression patterns in the germline, soma, cytoplasm, or nucleus [154]. This means their roles and functions are likely to be broad. In arthropods, for instance, both somatic and germline piRNAs are common [157]. In *Ae. aegypti*, there are not many TE-derived piRNAs relative to the TE content of the genome; instead, piRNAs are also derived from viruses (vpiRNAs) or *Ae. aegypti* genes [133]. Interestingly, although vpiRNAs are relatively abundant in *Aedes*, they are generally not observed in *Drosophila* [123]. The Piwi protein repertoire has expanded in culicines, further suggesting functional divergence [118]. *Aedes* express seven Piwis. Notably, four of these Piwis, Piwi4-6

and Ago3, are abundantly expressed in both the soma and germline [118,158]. *Culex* mosquitoes also exhibit an expanded family of Piwi proteins [123], and recent work indicated that *Culex quinquefasciatus* cells (Hsu) produce vpiRNAs against the insect-specific virus (ISV) Merida virus (*Rhabdoviridae*: *Merhavirus*) [159].

Over a decade ago, it was observed that arbovirus infections in *Ae. aegypti* cells and mosquitoes led to the production of 24-32 nt vpiRNAs [118,121,122,160–162]. Mature vpiRNAs exhibit the hallmarks of "ping-pong" amplification [120–122]. In the ping-pong cycle, primary piRNAs are processed from precursor single stranded (ss) RNAs (in the case of virus-derived vpiRNA, precursor ssRNAs are viral RNA). Primary piRNAs guide the cleavage of complementary sequences, producing secondary piRNAs. Secondary piRNAs then perpetuate the loop by binding to complementary precursor ssRNAs that are processed into primary piRNAs. In *Ae. aegypti,* Piwi5 and Ago3 are responsible for vpiRNA biogenesis [118]. Mature vpiRNAs therefore exhibit a uridine bias on the antisense primary piRNA (1U bias) and an adenine bias on the complementary sense secondary piRNA (10A bias) because these nucleotides serve as binding sites for Piwi5 and Ago3, respectively, during ping-pong amplification [118]. Mature piRNAs also exhibit 5' phosphates and 2'-O-methylated 3'-terminal ends that protect the mature piRNA from degradation and serve as a binding site for the PAZ (Piwi/Argonaute/Zwille) domain of Piwi proteins [163].

Ping-pong amplification relies on complex protein interactions. For example, the Tudor proteins Veneno and Atari-PB associate with Ago3, along with the co-factors Vasa and Yb, which in turn interact with Piwi5, and form a complex in the cytoplasm that is necessary for ping-pong amplification and secondary piRNA formation [129,158] (Figure 1.7). The cytoplasmic form of the slicer Pasilla is also required for piRNA biogenesis and interacts with

Piwi5 [158]. Zucchini and Nibbler orthologs are required for piRNA 3'-terminal end formation and trimming [164]. Zucchini and Nibbler produce "trailer piRNAs," which are diverse downstream cleavage products that result from phased piRNA precursor processing [164]. This suggests that piRNA "pools" may be more adaptable than previously thought. A recent study reported that SINV infection did not significantly alter the interactomes of the somatic *Ae. aegypti* Piwis, suggesting that vpiRNA processing does not require specific proteins but rather relies on the proteins involved in processing of other piRNA populations [158]. A current model of the mechanisms known underlying vpiRNA biogenesis and piRNA-mediated antiviral immunity is presented in Figure 1.7.

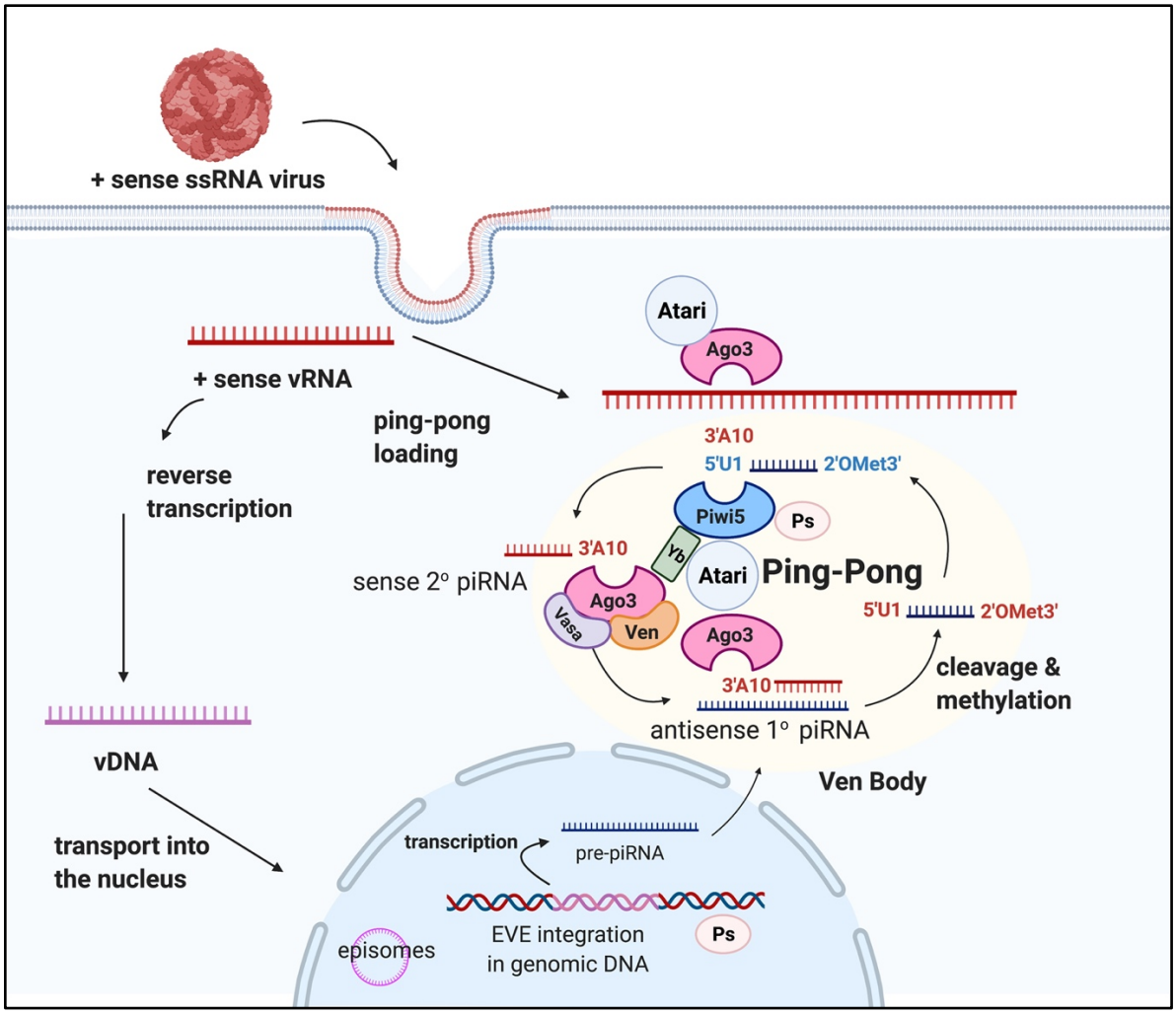


Figure 1.7. The current model of piRNA-mediated antiviral immunity. Non-retroviral positive-sense (+) single-stranded (ss) RNA viruses are reverse transcribed by endogenous retrotransposons into viral DNA (vDNA) in the cytoplasm. vDNA is then transported into the nucleus and may form episomes or may be integrated into the mosquito genome as an endogenous viral element (EVE). EVEs are transcribed into pre-piRNAs that are transported into the cytoplasm. Upon infection with a cognate virus, positive-sense ssRNA, derived from the virus genome, is bound by Argonaute-3 (Ago3) at an adenosine nucleotide. Ago-3-bound ssRNA hybridizes with complementary antisense primary piRNAs, which are cleaved into 24-32 nt piRNAs and methylated at the 3'-terminal end. Mature antisense piRNAs, bound by Piwi5 at a uridine nt, then perpetuate the cycle by interacting with complementary positive sense ssRNAs. piRNA amplification proceeds by a mechanism known as ping-pong in granular Ven bodies. The Tudor proteins Veneno (Ven) and Atari-PB (Atari) associate with Ago3, along with the co-factors Vasa and Yb, which in turn interact with Piwi5, and form a complex in the cytoplasm that is necessary for ping-pong amplification and secondary piRNA formation. The cytoplasmic form of the slicer Pasilla (Ps) is required for piRNA biogenesis and interacts with Piwi5. Figure made with Biorender.com.

1.3 Endogenous viral elements and Piwis

1.3.1 EVEs and ISVs in Ae. aegypti

Endogenous viral elements, or EVEs, are viral fragments integrated into host genomes that are inherited as host alleles [165]. EVEs are generated when a double-stranded DNA (dsDNA) copy of the viral genome becomes fully or partially integrated into germline or somatic cell genomes [165]. The viral origins of EVEs are diverse and may be derived from single-stranded DNA (ssDNA) viruses [166], retroviruses [167,168], or non-retroviral RNA viruses [169]. EVEs are commonly found across eukaryotes. For example, EVEs derived from retroviruses of at least thirty virus families compose 5-8% of the human genome [170] and play important roles in host physiology and antiviral immunity [171].

Non-retroviral RNA arboviruses, as opposed to retroviruses, neither enter the nucleus nor express a reverse transcriptase and therefore do not make a DNA intermediate; yet they counterintuitively present as non-retroviral integrated RNA virus sequences (NIRVS) in eukaryote genomes [165,169,172]. NIRVS are often found in TE-rich regions, clustered with

long-terminal repeat (LTR) retrotransposons, suggesting that retrotransposon TEs are involved in endogenization of non-retroviral RNA viruses [173–176]. Specifically, endogenous retrotransposons reverse transcribe viral RNA into viral cDNA (vDNA), which either forms episomes in the nucleus or is integrated into the host genome [162,177,178]. Treating SINV-infected *Ae. aegypti* cells with reverse transcriptase inhibitors prevents vDNA production [127].

NIRVS are EVEs derived from non-retroviral RNA viruses, and they occur much less frequently than EVEs derived from retroviruses [165]. They are often partial genome fragments and are the result of sporadic evolutionary events. As such, NIRVS are often distinct than the presently circulating virus from which they are derived because RNA viruses are susceptible to mutations; NIRVS can therefore be thought of as "blueprints" of ancient viral pseudogenes that have historically infected an organism [128,172,179,180]. *Ae. aegypti* NIRVS tend to originate from insect-specific viruses (ISVs), although NIRVS genetically related to arboviruses have been reported as well [128,174,176]. ISVs exclusively infect insects [172]. They are closely associated with their vector hosts and are transmitted vertically throughout mosquito populations when they infect the germline [181], although mechanical ISV transmission is also possible [182]. Most NIRVS are of ISV origin because many ISVs persistently infect mosquito germlines, which increases the probability of acquiring an inheritable NIRVS [128]. For example, the first NIRVS discovered in *Ae. aegypti* and *Ae. albopictus* were from four regions similar to RNAs of the ISVs cell fusing agent (CFAV; *Flaviviridae*: *Flavivirus*) and Kamiti River (KRV; *Flaviviridae*: *Flavivirus*) [169,172]. Since their initial discovery, more than 300 NIRVS have been annotated in *Ae. aegypti* alone, mostly of *Flaviviridae* or *Rhabdoviridae* origins [174,183,184], although there are probably many more. *Aedes* mosquitoes harbor approximately 10 times more EVEs

than other mosquito species [127]; this is again perhaps a consequence of the highly abundant and persistently infective nature of ISVs in *Aedes*, particularly in the germline [128].

NIRVS are enriched in piRNA clusters, regions of the genome that generate piRNAs [174,176]. This suggests NIRVS function with vpiRNAs and are involved in antiviral immunity. The most direct evidence supporting this hypothesis relied on a strain of *Ae. aegypti* that expressed a CFAV NIRVS that was approximately 96% similar to a previously isolated strain of CFAV [130]. These mosquitoes, when infected with the cognate CFAV strain, controlled CFAV infection by producing CFAV-specific vpiRNAs [130]. Using CRISPR/Cas9, the authors of that study then knocked out the CFAV NIRVS in mosquitoes and infected them with the cognate CFAV strain [130]. CFAV NIRVS-KO mosquitoes lacked CFAV-specific piRNAs and harbored significantly greater CFAV loads in somatic and germline tissues as compared to controls [130]. Based on these results, the authors proposed that both a NIRVS and exogenous viral infection are required to produce antisense and sense primary piRNAs, which induces secondary piRNA production, virus silencing, and ping-pong amplification [130] (Figure 1.7).

The impact of piRNA-mediated antiviral immunity against arboviruses of medical importance remains controversial. As discussed previously, NIRVS tend to be of ISV origin, although NIRVS genetically related to arboviruses have been reported in *Ae. aegypti* as well [174]. However, sequencing *Ae. aegypti* EVEs is challenging because they tend to be in highly repetitive regions of the genome, so it is likely NIRVS are poorly annotated in the current *Ae. aegypti* genome assembly. Efforts to generate and assemble long reads across these repetitive regions are underway to improve genome annotations [127]. Obstacles, however, remain. Long-read sequencing machines display high error rates (~13-15% [185]). Furthermore, NIRVS display geographic structure [180], so EVEs sequenced from colony mosquitoes will not

necessarily reflect those in field mosquitoes that become infected with diverse ISVs. Further investigation on the mechanisms underlying EVE and piRNA-mediated antiviral immunity, and its relevance in the field, is warranted.

1.3.2 Piwi-piRNA structural biology

The Argonaute superfamily of proteins is divided into two subfamilies – Ago and Piwi. Agos bind siRNAs and microRNAs (miRNAs) and are ubiquitously expressed. Piwis, on the other hand, bind piRNAs and are traditionally thought to be expressed mostly in the germline. Argonautes are evolutionarily conserved and are found in all higher eukaroyes [139]. They have two characteristic domains: PAZ and PIWI. The PAZ domain in Piwis preferentially binds 3'-terminal 2'-O-methylated piRNAs because of a hydrophobic binding pocket that is flexible enough to accommodate the methyl group [163]. This contrasts with Ago PAZ that have more restrictive RNA binding pockets, which drives protein preferential binding to 3'-terminal 2'-non-methylated sRNAs [163]. The PIWI domain is largely homologous to RNase H and harbors a catalytic tetrad DEDH (two aspartate, one histidine, and one glutamic acid amino acids) that exhibits slicing activity [186]. Other conserved domains include the N-terminal (N) domain, required for full catalytic activity and guide-target positioning, and the MID domain, which binds the 5' phosphate of guide RNAs [186]. Argonautes are bilobed, where one lobe is comprised of the N-PAZ domains, and the other is comprised of the MID-PIWI domains; the lobes are connected to each other by 2 linker regions [187]. A "specificity" loop in the MID domain binds to the 5'-terminal uracil, while a conserved lysine and a metal ion bind the 5' phosphate group of the guide RNA strand [188,189].

Comparing structural features of Agos and Piwis has revealed important mechanistic insights on their contrasting functions. For example, the PIWI domain of one of the three

Drosophila Piwis (Piwi) harbors a DVDH tetrad and lacks the slicing activity displayed by Agos [189]. A Piwi mutant that harbored the canonical DEDH catalytic tetrad did display slicing activity and disassociated from partially complementary RNA targets [189]. These observations suggest that slicing activity would likely compromise Piwi-mediated co-transcriptional silencing [189]. That study is one of the few to crystalize and solve the structure of a Piwi, which was isolated by immunoprecipitation from fly ovarian somatic cells bound to a piRNA [189]. That same team used similar techniques to crystalize silkworm Piwi (Siwi) as well [190].

Structural information on Argonautes, particularly Piwis, remains limited. Recombinant Argonaute proteins are difficult to express and purify, especially at high concentrations, and tend to be unstable without their RNA binding partners. Immunoprecipitating endogenous Argonautes, on the other hand, either requires specific antibodies raised against the protein or expressing the protein with a tag that can bind to tag-specific antibodies. Generating specific antibodies against Argonautes, however, is often difficult. Antibodies are often collected from animals who have been immunized with a recombinant protein and generating recombinant Argonautes is challenging. Additionally, antibodies produced against small peptides may cross react to other similar Argonaute proteins because they tend to be highly conserved.

1.4 Antiviral effectors and transgenic technologies in *Aedes aegypti*[2]

1.4.1 Antiviral effectors in transgenic Ae. aegypti

The *Ae. aegypti* antiviral immune mechanisms and natural barriers against infection described thus far have provided avenues to engineer virus-resistant mosquitoes. For example, as

[2] This section includes parts of the manuscript cited as "Williams, A.E., Franz, A.W.E., Reid, W.R., Olson, K.E. Antiviral effectors and gene drive strategies for mosquito population suppression or replacement to mitigate arbovirus transmission by *Aedes aegypti*. *Insects*. 2020; *11*(1): 52." The article is reproduced with permission and minor modifications have been made.

early as 2006, the siRNA pathway was exploited to engineer transgenic DENV2-resistant mosquitoes [191]. These mosquitoes expressed from the *Ae. aegypti carboxypeptidase A (CpA)* bloodmeal-inducible promoter an inverted-repeat (IR) RNA derived from the DENV2 prM encoding RNA sequence. The transcribed IR, mimicking a fragment of the viral dsRNA intermediate, was processed by the mosquito's RNAi machinery into DENV2-specific siRNAs, causing destruction of the viral RNA in infected midgut cells. Fourteen days post-infection (dpi), 96% of the transgenic mosquitoes did not exhibit any level of DENV2 infection and consequently did not transmit the virus in *in vitro* transmission assays [191]. This same strategy has been used in transgenic *Ae. aegypti* expressing the same IR from the salivary gland-specific *Ae30K b* promoter [192], and in an improved homozygous line ("Carb109M"), where the IR, again under control of the *CpA* promoter, was expressed from a more stable genetic locus ("Carb109;" Chr3: 409699138) [193]. These mosquitoes continue to be nearly 100% resistant to DENV2 through > 55 generations as compared to their wild-type (WT) counterparts and display minimal fitness costs [193].

Another antiviral effector strategy that has been explored makes use of ribozymes, which are short RNA molecules that have enzyme-like capabilities to cleave RNA molecules [194]. Ribozymes are found in all domains of life and have the capacity to self-cleave (*e.g.*, group I and II introns) or to cleave RNA at specific sites (*e.g.*, hammerhead and hairpin ribozymes) [194]. Ribozymes can be engineered to function *in trans* to cleave a target RNA, and several types have been investigated as antiviral mediators in *Aedes* spp. cell lines and transgenic mosquitoes [195–197]. For example, the *Ae. albopictus* cell line C6/36 was transformed to express "dual" trans-splicing group 1 introns that targeted both DENV and CHIKV [196]. These group 1 ribozymes targeted the highly conserved 5′ and 3′ cyclization sequences of DENVs and a highly conserved

region in the CHIKV *NS1* gene. Additionally, these ribozymes were designed to express a ΔN-Bax 3′ exon upon splicing of the viral genome, which caused the induction of cell apoptosis. Thus, the ribozyme led to the destruction of the arbovirus *per se* as well as the infected cell. All cell lines transfected with these ribozymes completely suppressed both DENV1-4 and CHIKV replication *in vitro* [196]. Because the ribozymes target highly conserved sequences, this strategy could be a promising method to develop transgenic mosquitoes resistant to multiple arboviruses; furthermore, targeting highly conserved viral sequences could limit the viruses' abilities to develop resistance against the effector transgene.

Transgenic *Ae. aegypti* expressing a hammerhead ribozyme (hRZ) were recently developed targeting the attenuated CHIKV strain 181/25 [197]. In principle, hammerhead ribozymes target small (15–16 nt) sequences, are active without relying on the host-cell machinery and are stable at a wide variety of temperatures [197]. Seven hRZs were designed to target the CHIKV structural polyprotein-encoding region of the viral RNA genome, and each individual hRZ completely inhibited CHIKV 181/25 in cell culture. Six out of seven transgenic *Ae. aegypti* lines engineered to express a CHIKV 181/25 hRZ completely inhibited CHIKV replication in the salivary glands at 7 dpi [197].

Transgenic *Ae. aegypti* have also been engineered to overexpress genes involved in the conserved, antiviral immune Janus kinase (JAK)-signal transducer and activator of transcription (STAT) pathway [198]. A cytokine signaling pathway in mammals, the Dipteran JAK/STAT pathway is activated by ligand binding to and dimerization of the transmembrane receptor Domeless (Dome) (or its species-specific orthologs) [199]. Associated JAKs, such as the Hop kinase, self-phosphorylate, then phosphorylate the Dome receptor, and form docking sites for other STAT proteins; following their phosphorylation, the STATs are translocated into the

nucleus where they serve as transcription factors for several antiviral restriction factors [199]. Transgenic mosquitoes were engineered to overexpress either the Dome receptor or the Hop kinase, as well as to overexpress both elements simultaneously [198]. Reduction of DENV2 and DENV4 infections clearly varied among individual transgenic mosquitoes, and they were not able to effectively antagonize ZIKV or CHIKV [198]. Therefore, this transgenic approach may not be sufficient to completely block virus transmission.

Ae. aegypti expressing a cluster of small synthetic DENV3 and CHIKV-targeting miRNAs were engineered to suppress both viruses by way of the microRNA pathway [200]. Similar to the siRNA pathway, the miRNA pathway is highly conserved and uses miRNAs as guides for Argonaute-1 mediated sequence-specific degradation. However, pri-miRNAs (instead of long dsRNA, as in the siRNA pathway) are processed by Drosha in the nucleus, exported by Exportin5 into the cytoplasm, and processed by Dicer-1 into 21–25 nt miRNAs, which are then loaded into miRNA-induced silencing complexes (miRISCs). Eventually, this then results in a sequence-specific degradation of mRNAs or translational gene silencing [200]. Exploiting this pathway, four effector constructs were tested in transgenic mosquitoes: two lines expressing four or six miRNAs complementary to DENV3 or CHIKV RNAs, respectively, both under control of the constitutively expressing *polyubiquitin* (*PUb*) promoter, and two lines expressing ten and three miRNAs targeting both DENV3 and CHIKV RNAs, respectively, either from the *PUb* or the *CpA* promoter [200]. Transgenic mosquitoes engineered to target either DENV3 or CHIKV (but not both) showed elevated levels of resistance to their viral target, while those mosquitoes engineered to target both viruses simultaneously were significantly less susceptible to DENV3 (~10% infection prevalence) but were unable to affect CHIKV replication [200]. These results suggested that DENV3 may be more vulnerable to miRNA-mediated silencing than CHIKV.

This same miRNA-based strategy has been used to engineer *Ae. aegypti* targeting multiple ZIKV strains (Cambodian FSS13025 and Puerto Rican PRVABC59) [201] as well as all four serotypes of DENV [202]. In the case of the former, the transgenic mosquitoes expressed a polycistronic cluster of eight synthetic ZIKV-targeting miRNAs, although only five of them were properly processed. Nonetheless, heterozygote mosquitoes of the transgenic line significantly reduced ZIKV titers by > 2 logs plaque-forming units per milliliter (PFU/mL) whereas homozygote mosquitoes of the line completely inhibited ZIKV midgut infection and dissemination at 4 and 14 dpi. Further, these transgenic mosquitoes completely lacked virus in saliva as shown by transmission assays [201]. On the other hand, the miRNA-targeting approach against four serotypes of DENV blocked DENV2 and DENV4, but was less effective against DENV1 or DENV3 [202]. In that study, transgenic mosquitoes expressed eight miRNAs targeting conserved regions within the 5' and 3'UTR regions of the DENV genome, either under the *PUb* or the *CpA* promoters [202]. Transgenic mosquitoes expressing the miRNA construct under the *CpA* promoter (*CpA-8miR*) were more effective at preventing DENV replication than those expressing the construct under the *PUb* promoter [202]. Significantly less *CpA-8miR* mosquitoes became infected with DENV2-4 7 dpi as compared to controls, but there was no difference in infection prevalence when both groups were infected with DENV1 [202]. The transgenic *CpA-8miR* mosquitoes that did become infected with DENV2-4 strain were infected a significantly lower titers as compared to controls 7 dpi [202]. Furthermore, significantly less *CpA-8miR* mosquitoes exhibited DENV2 or DENV4-infected saliva or salivary glands compared to controls 7 and 14 dpi [202]. There was no difference in transmission efficiency for transgenic mosquitoes when infected with DENV1 or DENV3 compared to controls [202].

A convincing strategy for targeting all four serotypes of DENV relied on transgenic mosquitoes engineered to express single-chain antibodies [203]. Based on a monoclonal antibody (1C19) derived from human patient IgGs, Buchman and colleagues (2020) developed a homozygous transgenic line of *Ae. aegypti* that expressed a broadly neutralizing, single-chain variable fragment targeting DENV1-4 [203]. This single-chain antibody recognized the BC loop of domain II in the envelope protein and was optimized for expression in mosquitoes [204]. Remarkably, these mosquitoes have shown to be resistant to all four DENV serotypes, displaying a complete lack of virus infection in midguts, carcasses, and saliva at 14 dpi as confirmed by plaque assays or quantitative RT-PCR [203]. Table 1.2 lists all transgenic *Ae. aegypti* lines that have been engineered to date to be arbovirus resistant.

Table 1.2. Transgenic *Ae. aegypti* engineered (to date) to be resistant to arbovirus infections. IR = inverted-repeat construct; *CpA = Ae. aegypti Carboxypeptidase A* promoter; *Ae30K b = Ae. aegypti Aegyptin* promoter; *AetRNA^{val}* Pol III = *Ae. aegypti* RNA polymerase III valine promoter; *AeVg1 = Ae. aegypti vitellogenin 1* promoter; *PUb = Ae. aegypti Polyubiquitin* promoter; dpi = days post-infection.

Study	Transgenic Strategy	Virus Targeted	Method of transgene insertion	Promoter for antiviral effector	Prevalence of transgenics with disseminated infection
Franz et al., 2006 [191] Franz et al., 2009 [219] Franz et al., 2014 [193]	IR triggering siRNA anti-viral pathway	DENV2	*mariner/mosI*	*CpA*	0% (14 dpi)
Mathur et al., 2010 [192]	IR triggering siRNA anti-viral pathway	DENV2	*mariner/mosI*	*Ae30k b*	0% (saliva, 14 dpi)
Mishra et al., 2016 [197]	antiviral hammerhead ribozymes	CHIKV 181/25	*piggyBac*	*AetRNAval* Pol III	0% (7 dpi)
Jupatanakul et al., 2017 [198]	overexpression of JAK/STAT players	DENV2/4	*piggyBac*	*AeVg1*	$\geq$43% (14 dpi)

Yen et al., 2018 [200]	synthetic miRNAs targeting virus genome	DENV3/ CHIKV	*mariner/mosI*	*PuB/CpA*	~10% (DENV3, 21 dpi) ~10-50% (CHIKV, 6 dpi)
Buchman et al., 2019 [201]	synthetic RNAs targeting virus genome, triggering miRNA antiviral immunity	ZIKV	*piggyBac*	*CpA*	0% (14 dpi)
Buchman et al., 2019 [203]	broadly neutralizing single chain antibody	DENV1-4	*piggyBac*	*CpA*	0% (14 dpi)
Williams et al., 2020 [Chapter 2] [205]	IR triggering siRNA anti-viral pathway	ZIKV	CRISPR/Cas9	*CpA*	~10% (14 dpi)
Liu et al., 2021 [202]	synthetic miRNAs targeting virus genome	DENV1-4	*mariner/mosI*	*CpA or PuB*	~50% (saliva, 14 dpi DENV1) ~8-35% (saliva, 14 dpi DENV2) ~50% (saliva, 14 dpi DENV3) ~8-50% (saliva, 14 dpi, DENV4)

1.4.2 Technologies to generate transgenic Ae. aegypti

Transposable elements have been used to introduce exogenous genetic material into *Ae. aegypti* as early as 1989 [206]. Autonomous TEs are selfish genetic elements (SGEs) that can remobilize following their integration into the host genome and increase their copy number during remobilization; non-autonomous transposable elements, on the other hand, require an externally provided transposase and, consequently, cannot re-mobilize on their own. Several Class II DNA transposable elements have since been characterized and used in *Ae. aegypti* as non-autonomous gene insertion vectors, including *piggyBac* [207] (derived from a *Trichoplusia ni* cell line [208]), *Hermes* [209] (derived from *Musca domestica* [210]), and *mariner Mos1* [211] (derived from *D. mauritiana* [212]). These elements have been extensively described [213]. *PiggyBac* and *mariner Mos1* are the most commonly used non-autonomous transposons to

generate transgenic *Ae. aegypti* [214]. A DNA plasmid-based transgene containing a gene of interest is flanked by the short inverted terminal repeats (ITR) of the transposon, serving as binding sites for the homologous transposase, which is supplied by a separate helper plasmid (Figure 1.8). Once co-expressed in the preblastoderm mosquito embryo, the transposase binds to both ITRs of the transgenic construct and cleaves the construct before integrating it into the *Ae. aegypti* genome in a cut-and-paste manner.

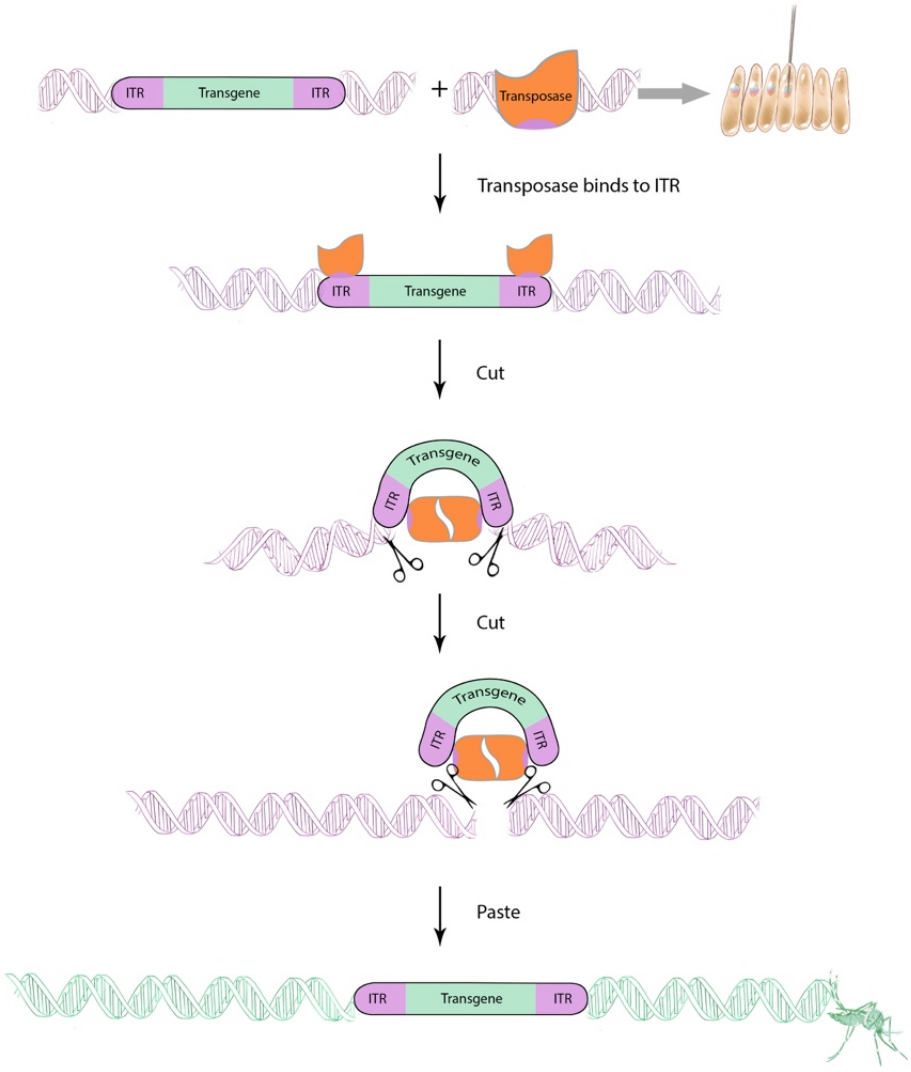


Figure 1.8. Non-autonomous class II transposable elements can be used to generate transgenic *Aedes aegypti*. Two DNA plasmids are co-injected into pre-blastoderm embryos: (1) a plasmid encoding the transgene, which is flanked by inverted terminal repeat (ITR) sequences that serve as binding sites for the transposase; (2) another "helper" plasmid encoding the

transposase. Once the transposase is transcribed and expressed, it binds to the ITR region of the transgene-bearing construct. Two transposase units dimerize, induce a double-stranded break (DSB), and cleave the transgene (including the ITRs) out of the plasmid. The transgene-transposase complex then binds and inserts the transgenic cargo at defined short recognition sequence motifs into the mosquito genome. Image from Williams et al., 2020 [215].

The CRISPR/Cas9 system was first used in *Ae. aegypti* in 2015 [216], employing the Cas9 protein from *Streptococcus pyogenes*. Cas9 is part of the bacterial immune response to phages and can introduce site-specific double-stranded breaks (DSBs) in the genomes of diverse species when programmed with a cognate guide (g)RNA containing crRNA/tracrRNA molecule(s). The Cas9 effector complex identifies a specific sequence immediately downstream of the target, termed the protospacer adjacent motif (PAM), in the host genome. A 17–20 bp stretch of the sgRNA binds to the PAM site, thereby stabilizing the Cas9/gDNA complex, which results in a DSB cleavage three base pairs upstream of the PAM. Owing to the simplicity of the required PAM sequence, it is predicted that, on average, a *S. pyogenes*-Cas9 PAM is present in the *Ae. aegypti* genome once every 17 base pairs [217]. Following a DSB, the cell must repair the genomic DNA molecule, which largely occurs through non-homologous end joining (NHEJ) or through homology directed repair (HDR). In somatic tissues and in the absence of a homologous DNA template, NHEJ is the likely means of DNA repair, while in the germline, HDR is the more likely DNA repair mechanism helping to conserve genome integrity. Methods and applications for CRISPR/Cas9 editing in *Ae. aegypti* are outlined in [217,218].

1.5 Goals and hypotheses of this dissertation

The overarching goals of the work described herein were twofold: (1) to determine the potency of the *Ae. aegypti* siRNA pathway against Zika virus and (2) to understand molecular mechanisms underlying piRNA-mediated antiviral immunity and its implications on mosquito

vector competence. To address the first goal, in Chapter 2, we hypothesized that the siRNA pathway could be co-opted to render transgenic mosquitoes resistant to ZIKV. We therefore engineered transgenic *Ae. aegypti* mosquitoes that synthetically triggered the endogenous siRNA pathway against ZIKV and then quantified virus resistance in these mosquitoes. We approached the second goal from two perspectives – through an *in vivo* approach, by sequencing sRNAs in mosquitoes from distinct origins, and through an *in vitro* approach, by studying RNA binding dynamics of a recombinant Piwi protein. First, in Chapter 3, we hypothesized that *Ae. aegypti* from distinct geographic origins were persistently infected with unique ISVs and had diverse sRNA profiles that may impact vector competence and virus persistence. We therefore sequenced and compared sRNAs from mosquitoes across the Americas and studied how ISV-specific sRNAs were impacted by arboviral (DENV) infection in mosquitoes. Next, in Chapter 4, we hypothesized that characterizing the structural features of an antiviral Piwi, Piwi4, involved in RNA binding and subcellular localization could provide insights on its role in the siRNA and piRNA pathways. We therefore expressed a recombinant domain of Piwi4 and measured its binding affinities for different sRNAs as well as generated Piwi4-specific polyclonal antibodies to visualize Piwi4 expression in mosquito tissues.

CHAPTER 2: THE ANTIVIRAL SMALL-INTERFERING RNA PATHWAY INDUCES ZIKA VIRUS RESISTANCE IN TRANSGENIC *AEDES AEGYPTI*[3]

2.1 Introduction

We previously found that synthetic resistance to arboviruses in *Ae. aegypti* could be achieved by co-opting the innate immune responses of the mosquito by utilizing its RNAi pathway [191]. Transgenic *Ae. aegypti* expressing an inverted repeat (IR) RNA targeting the pre-membrane region of the DENV2 genome triggered the siRNA pathway, resulting in ~100% DENV2 resistance [191–193]. However, assessment of several DENV2-targeting transgenic *Ae. aegypti* lines that differed from each other only by their genomic transgene insertion sites displayed a range in both effectiveness of DENV2 suppression and transgene stability [193,219]. Most notably, while multiple transgenic lines demonstrated high levels of DENV2 suppression, one of these lines lost its resistance phenotype after only 17 generations in laboratory colony [193,219]. Meanwhile, another transgenic line has remained refractory to DENV2 for > 55 generations [193]. Given the high impact of transgene genomic insertion site on both expression levels and long-term stability of antiviral effectors, this aspect should always be taken into consideration when engineering transgenic resistance to pathogens in *Ae. aegypti*.

To generate transgenic DENV2-refractory *Ae. aegypti*, the *mariner Mos1* transposon was used. However, the use of Class II DNA transposons for transposon-mediated genome insertion is quasi-random and prone to insertional position effects [193,220]. For our study, we therefore chose a genomic locus that was previously identified in Dong et al. (2017) to have robust and

[3] This section includes the complete manuscript cited as "Williams, A.E., Sanchez-Vargas, I., Reid, W.R., Lin, J., Franz, A.W.E., Olson, K.E., 2020. The antiviral small-interfering pathway induces Zika virus resistance in transgenic *Aedes aegypti*. *Viruses*. 2020; *12*(11): 1231." The article is reproduced with permission and minor modifications have been made.

stable transgene expression in the midgut [101,221]. In that study, Dong et al. (2017) used *mariner Mos1* to overexpress a tissue inhibitor of metalloproteinases (TIMP) transgene under control of the *carboxypeptidase A* promoter in the female midgut following bloodmeal ingestion [101]. Because transgene insertion by way of the *mariner Mos1* transposon is quasi-random, this method allows for the discovery of novel genomic loci that are permissive for strong and stable transgene expression. Indeed, Dong et al. (2017) identified a total of seven unique insertion sites that resulted in three viable transgenic lines, among which one line displayed strong and robust transgene expression in the midgut [101]. Functional testing of the transgenic mosquito line demonstrated a significant increase in the dissemination of chikungunya virus, which proved that the inserted transgene had the expected biological activity in the mosquito midgut [101]. Further, the authors mapped the inserted transgene to a single intergenic locus on chromosome 2 and performed genetic outcrosses to confirm single locus insertion [101]. Recently, work in the Franz laboratory has shown that the transgenic line from Dong et al. (2017) retains stable transgene expression for more than 10 generations. Given that the chromosome 2 locus resulted in strong and stable transgene expression in the midgut, we used the CRISPR/Cas9 system to introduce our anti-ZIKV effector proximal to the identified *mariner Mos1* insertion site.

For construction of the IR effector, we selected a cDNA sequence derived from the NS3/4A region of the ZIKV genome, which had been identified to be a highly effective target at suppressing the virus when transiently tested in *Ae. aegypti* [222]. In addition, while the approach to insert a transgene into a specific genome locus allows for robust transgene expression, it also allows for the direct side-by-side comparison and optimization of other antiviral effectors since any insertional position effects are cancelled out between them.

2.2 Results

2.2.1 Generation of transgenic Ae. aegypti *expressing a ZIKV-specific IR effector*

Based on our previous success in generating DENV2 resistance [191–193], we aimed to engineer ZIKV-resistant *Ae. aegypti* by expressing a ZIKV-specific IR RNA sequence *in vivo*, intended to trigger the mosquito's antiviral RNAi pathway (Figure 2.1 A). The IR sequence (Figure 2.2) was chosen based on previously published data that had identified the NS3/4A region (nt residues 6309-6846 of the PRVABC59 strain; 538 bp in each direction) of the viral genome to be highly conserved amongst different ZIKV strains and a robust RNAi target [222]. To identify the optimal site for Cas9-mediated transgene insertion around locus Chr2:321382225, we first injected three replicates of ~ 100 *Ae. aegypti* embryos (Higgs' white eye [HWE] strain) with injection mixes containing 300 ng/µL Cas9 protein complexed with 80 ng/µL of sgRNA. Injected embryos were allowed to develop for 16–24 h, after which we extracted their genomic DNA and amplified it across the target locus to assess for indels using Sanger sequencing and the Synthego ICE tool. Overall, two sgRNAs (#5 and #6, Table 2.1) demonstrated activity (Table 2.2), while the two other sgRNAs demonstrated no activity. The two sgRNAs that were active overlapped each other on the sense and antisense strands of the genomic DNA, respectively, and were located approximately 600 bp downstream of the *mariner Mos1* insertion site in the previously generated transgenic line "Timp-P4" [101]. A common donor was constructed for both sgRNAs that contained a 7 bp gap from the predicted cut site of sgRNA 5 and a 12 bp gap from the cut site of sgRNA 6 (Figure 2.3 A, B). This plasmid sequence has been deposited into GenBank under the accession number MT926371.

A. **B.**

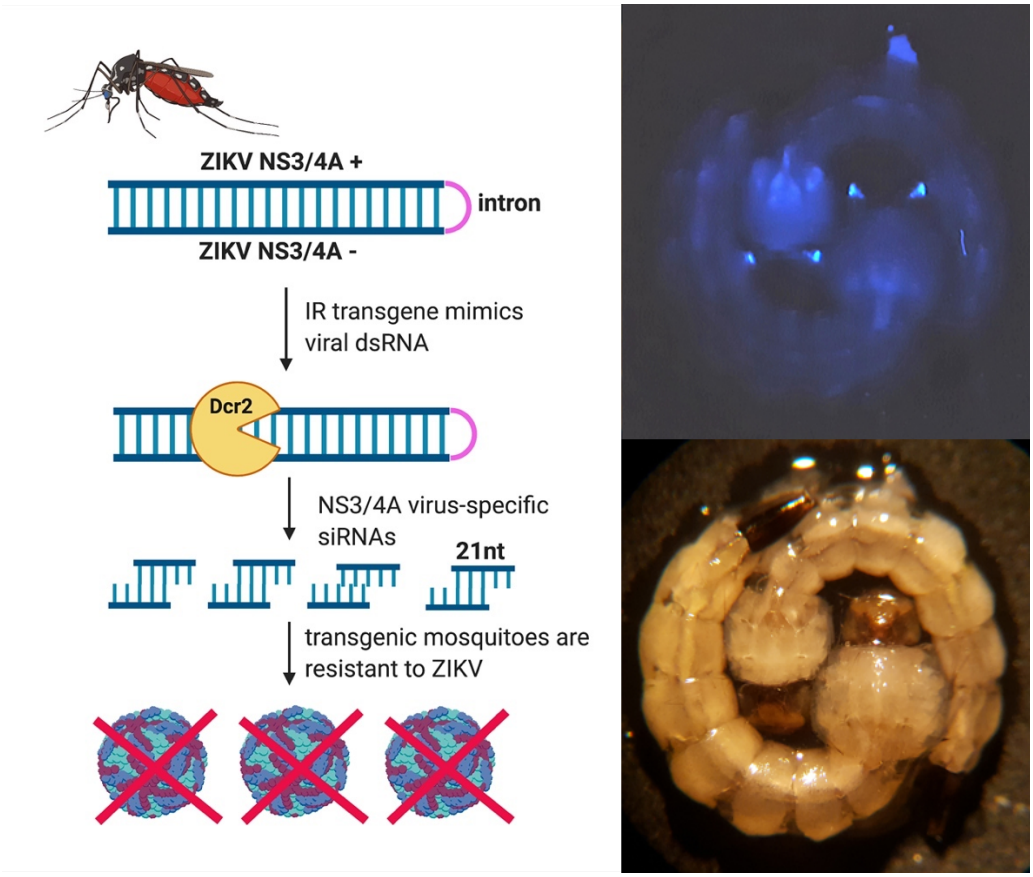


Figure 2.1. Transgene-mediated resistance to ZIKV in *Ae. aegypti*. (**A**) Schematic outline depicting the transgenic strategy for inducing Zika virus (ZIKV) resistance. Transgenic mosquitoes express an inverted repeat (IR) complementary to the NS3/4A regions of the ZIKV genome in their midguts after a bloodmeal. The IR mimics a viral double-stranded RNA (dsRNA), which triggers the endogenous siRNA antiviral immune pathway. Dcr2 = dicer-2. Figure created with BioRender.com. (**B**) (**top**) Transgenic mosquito larvae expressing eCFP from the photoreceptor-specific promoter 3xP3, enabling marker-based identification by fluorescent microscopy and selection of transgenic individuals. (**bottom**) Transgenic larvae under white light.

CGACCAACAACACCATAATGGAAGACAGTGTGCCGGCAGAGGTGTGGACCAGACAC
GGAGAGAAAAGAGTGCTCAAACCGAGGTGGATGGACGCCAGAGTTTGTTCAGATCA
TGCGGCCCTGAAGTCATTCAAGGAGTTTGCCGCTGGGAAAAGAGGAGCGGCTTTTG
GAGTGATGGAAGCCCTGGGAACACTGCCAGGACACATGACAGAGAGATTCCAGGAA
GCCATTGACAACCTCGCTGTGCTCATGCGGGCAGAGACTGGAAGCAGGCCTTACAA
AGCCGCGGCGGCCCAATTGCCGGAGACCCTAGAGACCATAATGCTTTTGGGGTTGCT
GGGAACAGTCTCGCTGGGAATCTTCTTCGTCTTGATGAGGAACAAGGGCATAGGGA
AGATGGGCTTTGGAATGGTGACTCTTGGGGCCAGCGCATGGCTCATGTGGCTCTCGG
AAATTGAGCCAGCCAGAATTGCATGTGTCCTCATTGTTGTGTTCCTATTGCTGGTGGT

GCTCATACCTGAGCCAGAAAAGCAAAGATCAGCTGCCTAATGATATATTTTTAATCA
CTAACAGAACTTTGAACAAAATCTGATGATTGGGTGCTCACCTCCGATCTTTGCTTTT
CTGGCTCAGGTATGAGCACCACCAGCAATAGGAACACAACAATGAGGACACATGCA
ATTCTGGCTGGCTCAATTTCCGAGAGCCACATGAGCCATGCGCTGGCCCCAAGAGTC
ACCATTCCAAAGCCCATCTTCCCTATGCCCTTGTTCCTCATCAAGACGAAGAAGATT
CCCAGCGAGACTGTTCCCAGCAACCCCAAAAGCATTATGGTCTCTAGGGTCTCCGGC
AATTGGGCCGCCGCGGCTTTGTAAGGCCTGCTTCCAGTCTCTGCCCGCATGAGCACA
GCGAGGTTGTCAATGGCTTCCTGGAATCTCTCTGTCATGTGTCCTGGCAGTGTTCCCA
GGGCTTCCATCACTCCAAAAGCCGCTCCTCTTTTCCCAGCGGCAAACTCCTTGAATG
ACTTCAGGGCCGCATGATCTGAACAAACTCTGGCGTCCATCCACCTCGGTTTGAGCA
CTCTTTTCTCTCCGTGTCTGGTCCACACCTCTGCCGGCACACTGTCTTCCATTATGGT
GTTGTTGGTCGC

Figure 2.2. ZIKV-specific NS3/4A IR sequence. Underline indicates NS3/4A inverted repeat (IR) gene, while red is the small intron of the *Ae. aegypti sialokinin1* gene

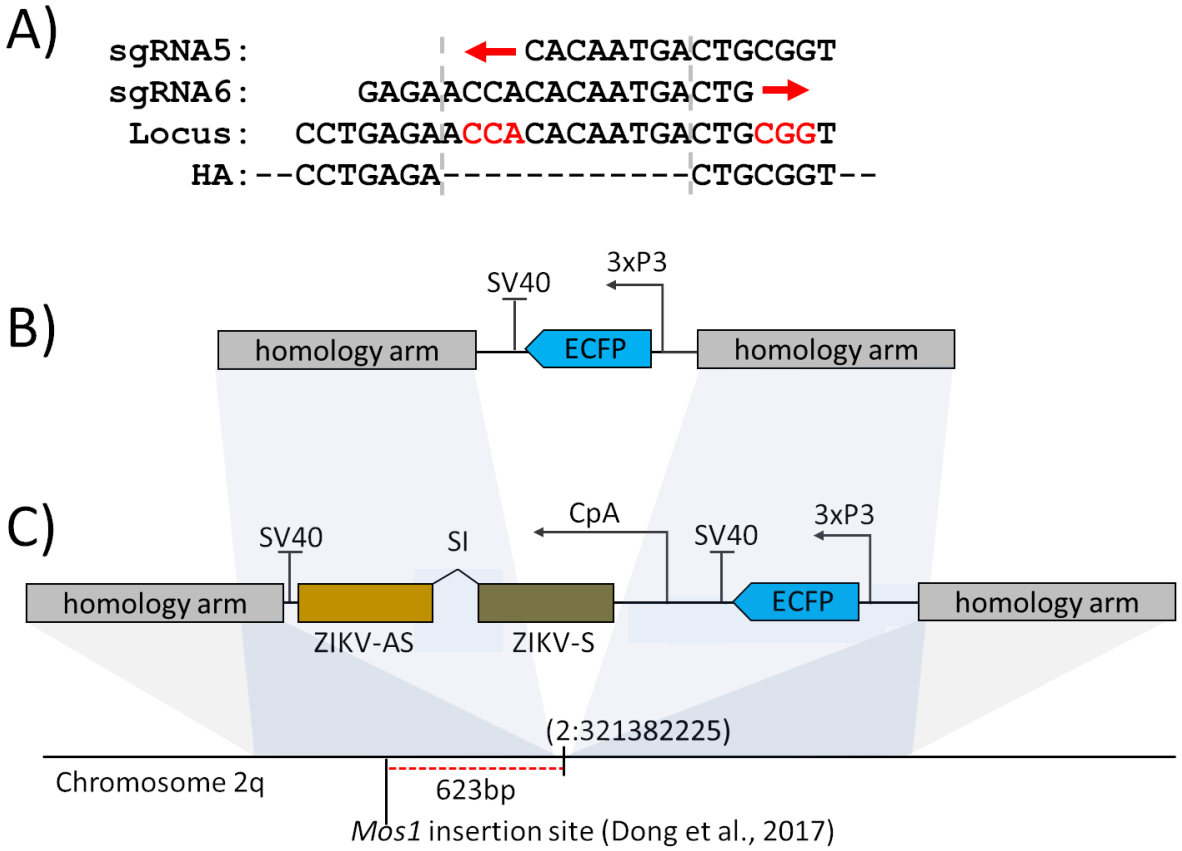


Figure 2.3. Schematic illustrations of sgRNAs and plasmids used in this study for CRISPR/Cas9 site-specific insertion. (**A**) Nucleotide sequences of sgRNAs (sgRNA5 and sgRNA6), their genomic target sequence, and the junctions of the flanking homology arms (HA) used for transgene integration into the Chr2:321382225 site. The red arrows indicate the directionality of the protospacers, while the respective PAM sequences are displayed in red text. (**B**) 3xP3-ECFP-SV40 donor used to test CRISPR/Cas9 mediated insertion efficiency. (**C**)

ZIKV-targeting IR construct. CpA = *carboxypeptidase A* promoter, ECFP = enhanced cyan fluorescent protein, 3xP3 = eye-specific promoter, SV40 = large T antigen terminator from simian virus 40 used as a polyadenylation signal, ZIKV-AS = antisense cDNA of the anti-ZIKV IR effector, ZIKV-S = sense cDNA of the anti-ZIKV IR effector, SI = small intron of the *Ae. aegypti sialokinin1* gene.

Table 2.1. Genomic location of the desired target insertion site for the inverted repeat transgene and the locations of the closest sgRNAs that did not contain predicted off-target sites in the *Ae. aegypti* genome.

Location	Embryos injected (total N)	*Ae. aegypti* Liverpool strain equivalent cut site[1]	Δ distance (bp)	Synthego score (average indel% +/ SEM)
Dong et al., 2017 [101]	n/a	2:321382845-321382846	0	n/a
sgRNA3	424	2:321382201	644	0
sgRNA4	282	2:321382958	112	0
sgRNA5	282	2:321382222	623	17.0 ± 2.5
sgRNA6	324	2:321382228	629	43.3 ± 3.4

[1]*Ae. aegypti* genome version AGWG AaegL5, vectorbase.org

Table 2.2. Testing of sgRNAs for genome editing rates in the Chr2:321382225 locus of *Ae. aegypti*.

Injection mix	Embryos injected	Male	Female	Survival (%)
sgRNA 5	735	19	22	5.6*
sgRNA 6	928	20	23	4.6
sgRNAs 5+6	882	38	36	8.4
Anti ZIKV IR	1510	84	100	12.1*

*Indicates that the experiment resulted in an established transgenic line.

Following the construction of the donor plasmid (Figure 2.3 C), we injected *Ae. aegypti* embryos with the same Cas9/sgRNA concentration used for the sgRNA efficacy assessment along with 80 fmol/μL donor plasmid and 100 ng/μL of ku70 dsRNA [218] to suppress NHEJ events. Surviving G0 males and females were individually outcrossed to HWE,

provided three bloodmeals, and assessed for transformation rates (Table 2.2). One pool was positive for site-specific insertion when sgRNA 5 was used, which was then selected for future experiments utilizing the IR effector construct.

Following successful insertion of an eCFP marker construct into the *Ae. aegypti* Chr2:321382225 locus, we used the same methodology to site-specifically insert the anti-ZIKV IR construct. A total of 1510 preblastoderm embryos were injected with injection mixes containing 300 ng/µL Cas9 protein, 80 ng/µL of sgRNA, 80 fmol/µL donor plasmid, and 100 ng/µL of ku70 dsRNA (to silence NHEJ events [218]), yielding 184 G0 survivors (12%), which were outcrossed to HWE to obtain the G1 population. In the G1 generation, 37 individuals displayed eye-specific eCFP expression within one of the male pools (consisting of ~ 200 virgin female *Ae. aegypti* that had been allowed to individually mate with ~ 20 of the surviving G0 males). Transgenic individuals, hereafter termed "anti-ZIKV-NS3/4A," were again outcrossed to female HWE to obtain the G2 generation. The transgene insertion locus and effector sequence were confirmed via Sanger sequencing, and the population was genetically balanced with the help of a 3xP3-mCherry marker inserted into the Chr2:321382225 site (unpublished) to obtain a homozygous population (Materials and Methods, Section 2.4.4). The anti-ZIKV IR construct sequence has been deposited in GenBank under the accession number MT926370.

2.2.2 The ZIKV-specific IR effector is processed by the endogenous siRNA machinery of the mosquito

To confirm that the anti-ZIKV-NS3/4A IR was successfully processed by the mosquito's siRNA machinery, we performed sRNA sequencing on midguts at 24 h post-non-infectious bloodmeal. ZIKV-specific 21 nt siRNAs were abundantly detected in midguts of the transgenic

mosquitoes at 24 h post-bloodfeed, but not in the midguts of the HWE control (Figure 2.4 A). The exact size of 21 nt indicates that the NS3/4A IR was processed by the endogenous RNAi siRNA pathway. Further positional analysis revealed that the ZIKV-specific siRNAs aligned exclusively to the NS3/4A region, which was specific for the IR effector (Figure 2.4 B). These data indicate that the ZIKV IR was successfully recognized and processed by the RNAi machinery in the midgut of females at 24 h post-bloodmeal.

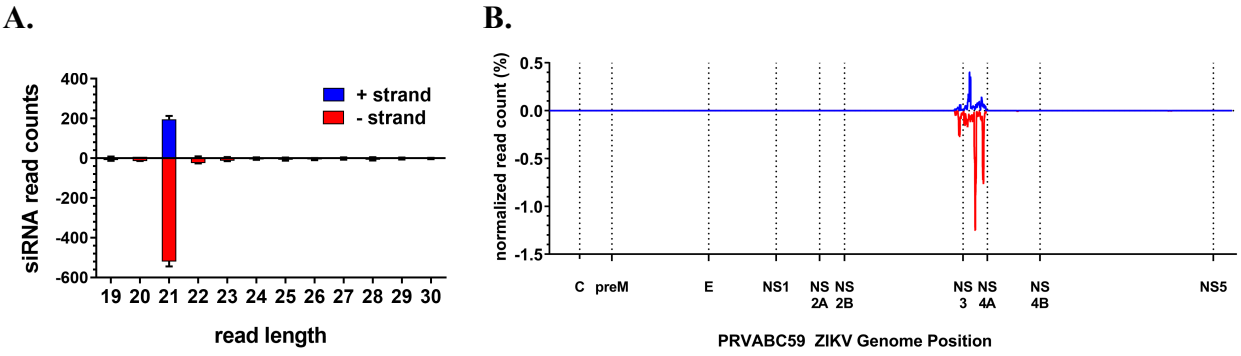
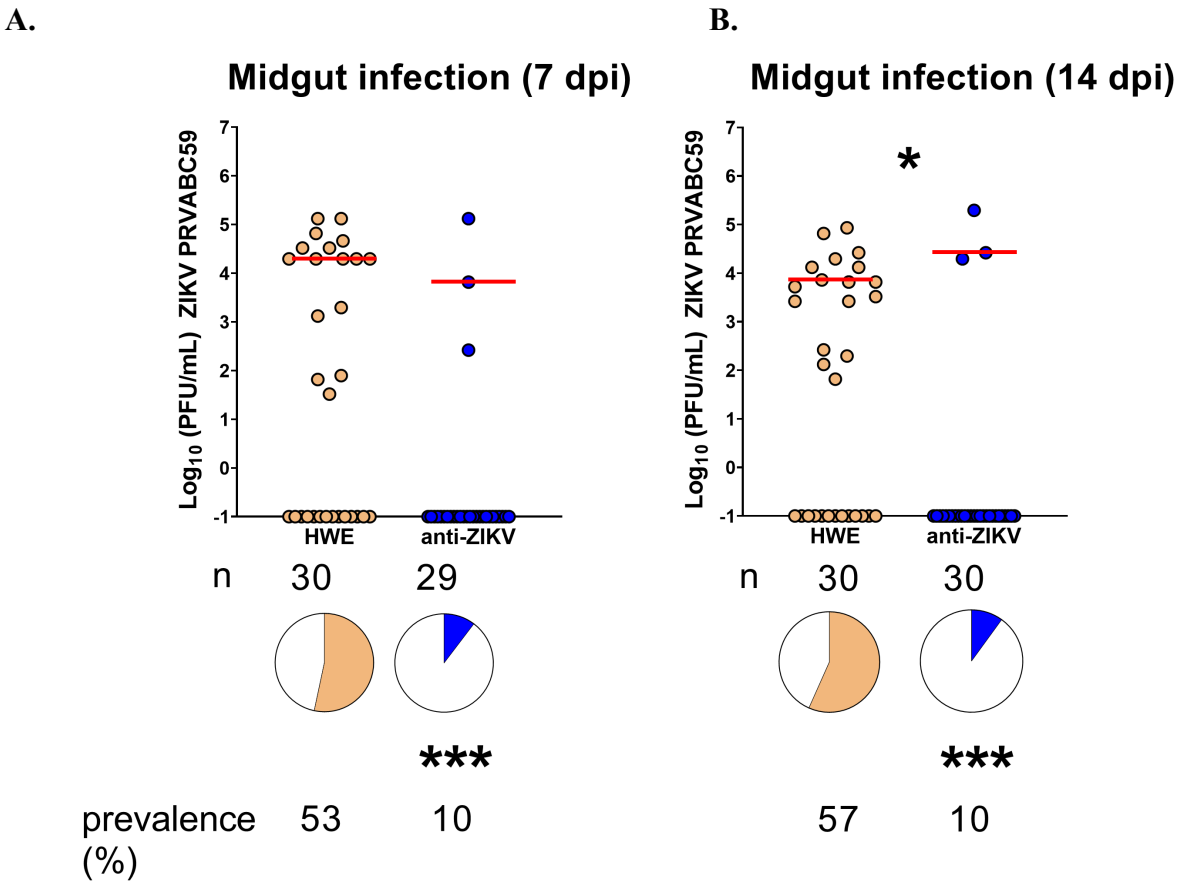


Figure 2.4. The anti-ZIKV-NS3/4A IR effector is processed by the midgut's RNAi machinery. (**A**) Anti-ZIKV-NS3/4A mosquitoes process the IR effector into 21 nt ZIKV-specific siRNAs. Black bars indicate mean with range error bars. (**B**) Positional analysis of ZIKV-specific siRNAs reveals that they are complementary to the NS3/4A region of the ZIKV (PRVABC59) genome, which is the region that is targeted by the transgene.

2.2.3 Ae. aegypti *expressing the anti-ZIKV IR effector are resistant to ZIKV*

To test if the anti-ZIKV-NS3/4A mosquitoes were resistant to ZIKV, we challenged them, along with the parental non-transgenic HWE control, with 3×10^5 plaque-forming units (PFU)/mL ZIKV PRVABC59. We dissected midguts at 7 and 14 dpi and performed plaque assays on these tissues to determine virus titers present in the transgenic versus the HWE control mosquitoes. We chose these time points based on peak ZIKV midgut (7 dpi) or disseminated infection (14 dpi) [92]. At 7 dpi, three (10%) transgenic midguts were infected with the virus as compared to 53% of HWE control midguts ($p = 0.0006$, Fisher's exact test) (Figure 2.5 A). The three transgenic mosquito midguts that were ZIKV infected displayed similar median titers as the

HWE midguts that were infected (p = 0.9329, Mann–Whitney U-test). At 14 dpi, again, three (10%) of the transgenic midguts were infected as compared to 57% of the HWE control midguts (p = 0.0003, Fisher's exact test) (Figure 2.5 B), exhibiting similar, albeit slightly higher, median titers (p = 0.0377, Mann–Whitney U-test).



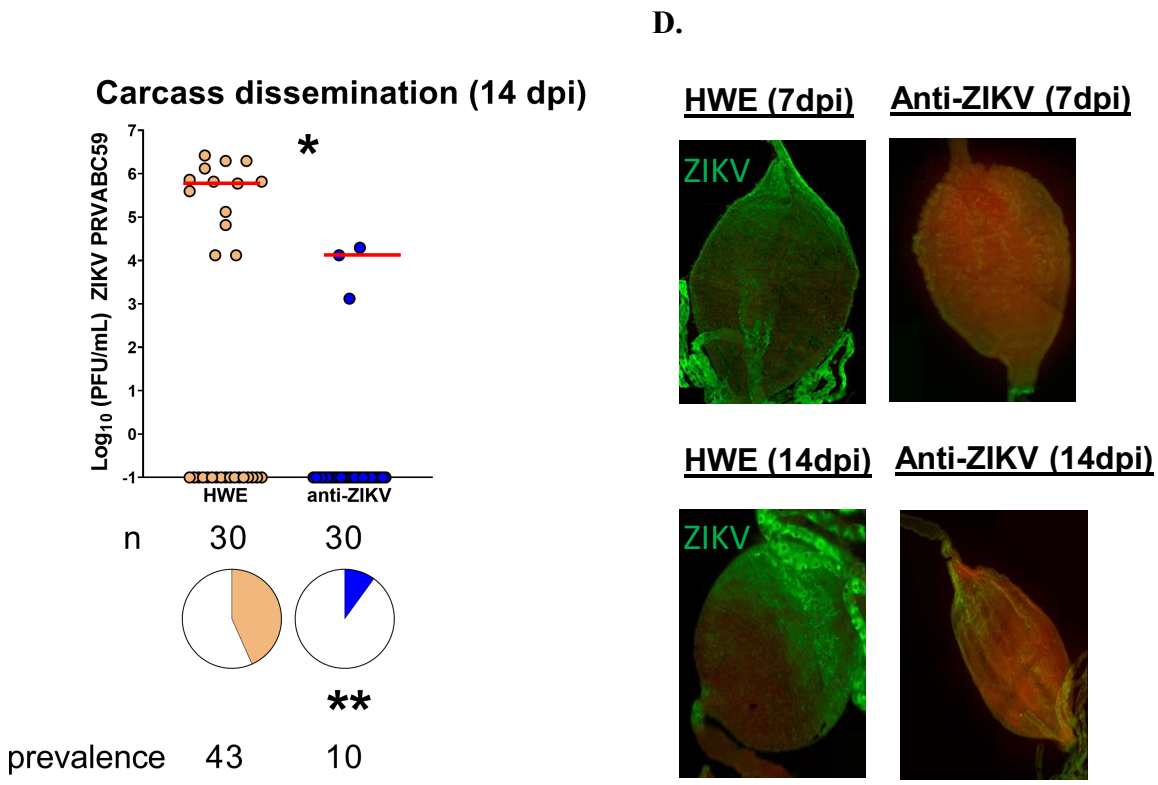


Figure 2.5. ZIKV IR effector expressing mosquitoes (anti-ZIKV-NS3/4A) are resistant to ZIKV. Infected midguts at (**A**) 7 dpi or (**B**) 14 dpi. (**C**) Infected carcasses showing disseminated infection at 14 dpi. (**D**) Representative immunofluorescence assay images of midguts obtained from HWE and anti-ZIKV-NS3/4A mosquitoes at 7 dpi (top) or 14 dpi (bottom). Primary antibodies recognized the ZIKV E and NS1 proteins. Red bars indicate median virus titers. Stars above graph compare virus titers between infected groups. Stars below pie charts compare infection prevalence; n = number of mosquitoes tested; * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. "Anti-ZIKV" = anti-ZIKV-NS3/4A (transgenic) mosquitoes.

To measure virus dissemination, we also performed plaque assays on matching carcasses at 14 dpi. We found that the transgenic carcasses that matched the infected transgenic midguts were also infected but showed significantly lower median titers when compared to the HWE carcasses ($p = 0.0232$, Mann–Whitney U-test) (Figure 2.5 C). Overall, 10% of the anti-ZIKV-NS3/4A carcasses were infected as compared to 43% of the HWE control carcasses ($p = 0.0074$, Fisher's exact test) (Figure 2.5 C). We also performed immunofluorescence assays on midguts dissected at both time points from both the HWE and anti-ZIKV-NS3/4A groups using

monoclonal antibodies that bind to the E and NS1 ZIKV proteins. At both time points, all HWE midguts that were imaged showed the presence of viral antigen, while none of the transgenic midguts imaged did so (Figure 2.5 D).

To determine whether the anti-ZIKV-NS3/4A mosquitoes could block virus transmission, we performed transmission assays by titrating mosquito salivary glands and saliva at 14 dpi. We challenged HWE and anti-ZIKV-NS3/4A mosquitoes with bloodmeals containing 7×10^5 PFU/mL of ZIKV PRVABC59. We found that five transgenic mosquitoes (17%) harbored infectious virus in their salivary glands, three of which also displayed saliva containing virus. In comparison, 59% of controls had infected salivary glands ($p = 0.001$, Fisher's exact test) and 33% released saliva containing virus (Figure 2.6). When only considering infected mosquitoes of both the transgenic and control groups, there were no significant differences between viral titers in the salivary glands ($p = 0.1338$, Mann–Whitney U-test) or in the saliva ($p > 0.9999$, Mann–Whitney U-test). These results show that the anti-ZIKV-NS3/4A mosquitoes significantly block virus replication in their salivary glands, which leads to a decrease in virus prevalence in saliva.

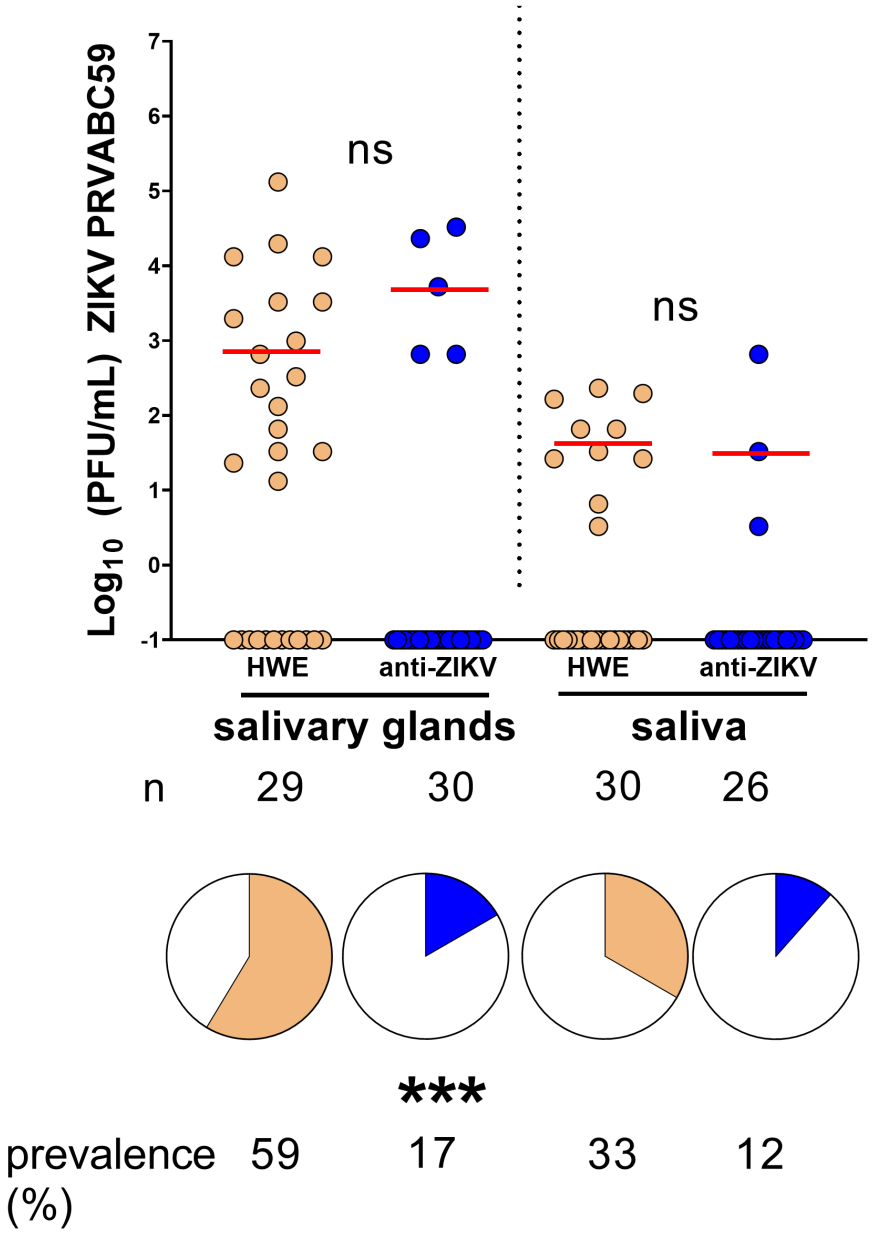


Figure 2.6. Transgenic anti-ZIKV-NS3/4A mosquitoes block ZIKV transmission. Infected salivary glands (**left**) or saliva (**right**) at 14 dpi. Red bars indicate median virus titers. "ns" above graph compares virus titers between infected groups. Stars below pie charts compare infection prevalence; n = number of mosquitoes tested; n = number of mosquitoes tested; ns = not significant; *** = $p < 0.001$.

2.2.4 Ae. aegypti *expressing the anti-ZIKV IR effector lose their resistance to the virus when their midgut infection barriers are bypassed*

We next sought to confirm whether the ZIKV resistance observed in the anti-ZIKV-NS3/4A transgenic mosquitoes was due to transgenic expression of the IR effector in the midgut, the initial site of infection after a mosquito ingests a viremic bloodmeal. We therefore hypothesized that by bypassing the IR-mediated midgut infection barrier, the transgenic mosquitoes would lose virus resistance. To bypass this barrier, we intrathoracically injected both HWE and anti-ZIKV-NS3/4A mosquitoes with 100 PFU ZIKV (PRVABC59). By infecting mosquitoes through artificial means — intrathoracic inoculation — as opposed to a more natural route of infection — bloodfeeding — we could assess the efficacy of the transgene-induced midgut infection barrier. After intrathoracic inoculation, we then separated both mosquito types into two groups: one that would receive a sugarmeal and another group that would receive a non-infectious bloodmeal (to trigger expression of the anti-ZIKV-NS3/4A IR). We then performed plaque assays on whole mosquitoes at 8 days post-virus injection. Regardless of bloodfeeding status, all mosquitoes became infected with ZIKV at high titers (Figure 2.7). This result indicates that the ZIKV resistance phenotype displayed by the anti-ZIKV-NS3/4A mosquitoes is caused by the transgenically imposed midgut infection barrier.

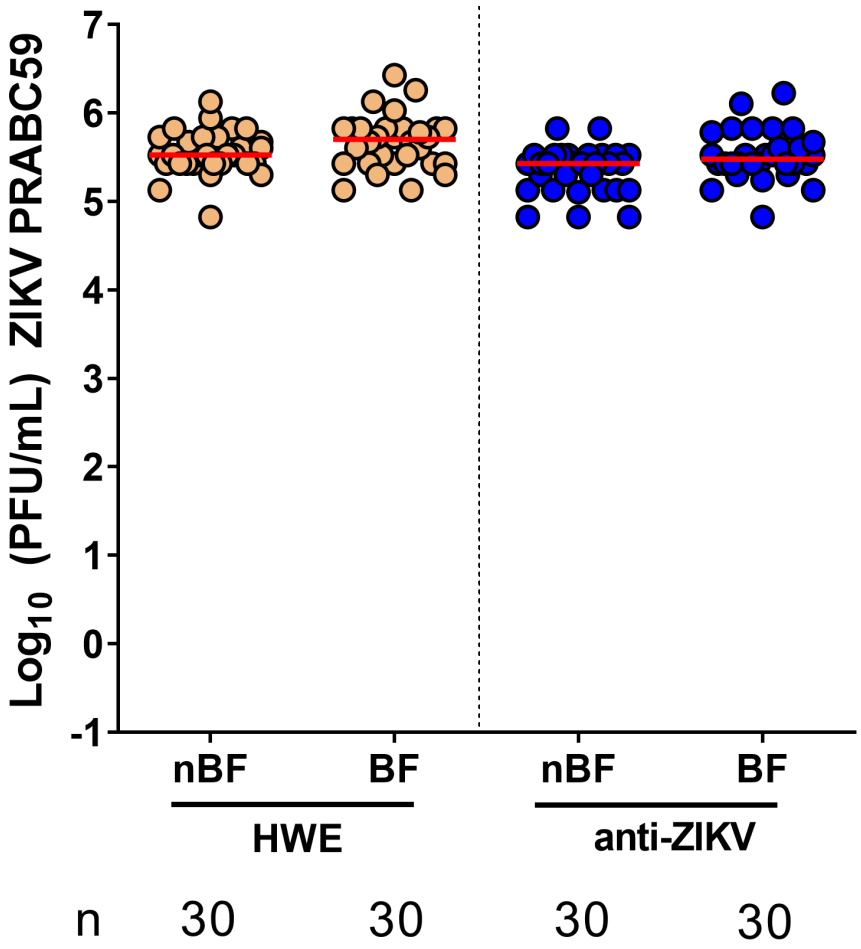


Figure 2.7. Anti-ZIKV-NS3/4A transgenics lose virus resistance when their midgut infection barrier is bypassed. After intrathoracic virus injection, both HWE mosquitoes (**left**) and anti-ZIKV-NS3/4A ("anti-ZIKV") transgenic mosquitoes (**right**) either received a non-infectious bloodmeal (to induce expression of the transgene) or a sugarmeal. Red bars indicate medians. nBF = non-bloodfed (sugarmeal), BF = bloodfed, n = number of mosquitoes tested.

2.2.5 Resistance of anti-ZIKV-NS3/4A transgenics to ZIKV shows a tendency to be virus strain-specific

The endogenous endonuclease Argonaute-2 slices RNAs that are complementary to the guide siRNA in a sequence-specific manner. We therefore asked whether the anti-ZIKV-NS3/4A transgenics would be resistant to any ZIKV strains that differed in their nucleotide sequences encoding the NS3/4A region of the viral genome. We aligned the anti-ZIKV-NS3/4A IR (derived from the PRVABC59 strain of the Asian lineage) to the homologous region of a different ZIKV

strain (Dakar 41525 of the African lineage) and found that the Dakar 41525 strain was 89% identical to the transgenic IR effector. We next challenged HWE and anti-ZIKV-NS3/4A mosquitoes with 6×10^4 PFU/mL ZIKV Dakar 41525 and collected midguts and carcasses at 14 dpi for plaque assays. We found that significantly fewer anti-ZIKV-NS3/4A midguts were infected ($p = 0.047$, Fisher's exact test) as compared to the HWE controls, and all infected transgenic midguts led to virus dissemination in the corresponding carcasses (Figure 2.8). Similar to ZIKV PRVABC59, the ZIKV Dakar 41525-infected anti-ZIKV-NS3/4A mosquitoes showed viral titers that were similar to the viral titers of the HWE control, in both the midguts ($p = 0.5191$, Mann–Whitney U-test) and the carcasses ($p = 0.7778$). Although we observed evidence of resistance against ZIKV Dakar 41525, the anti-ZIKV-NS3/4A mosquitoes appeared to be slightly more susceptible to a ZIKV strain that was not identical in sequence to the effector cargo. These results imply that the transgenic IR effector may be less protective against ZIKV strains that differ in their viral genome sequence by more than 10% when compared to the virus-derived sequence of the transgene.

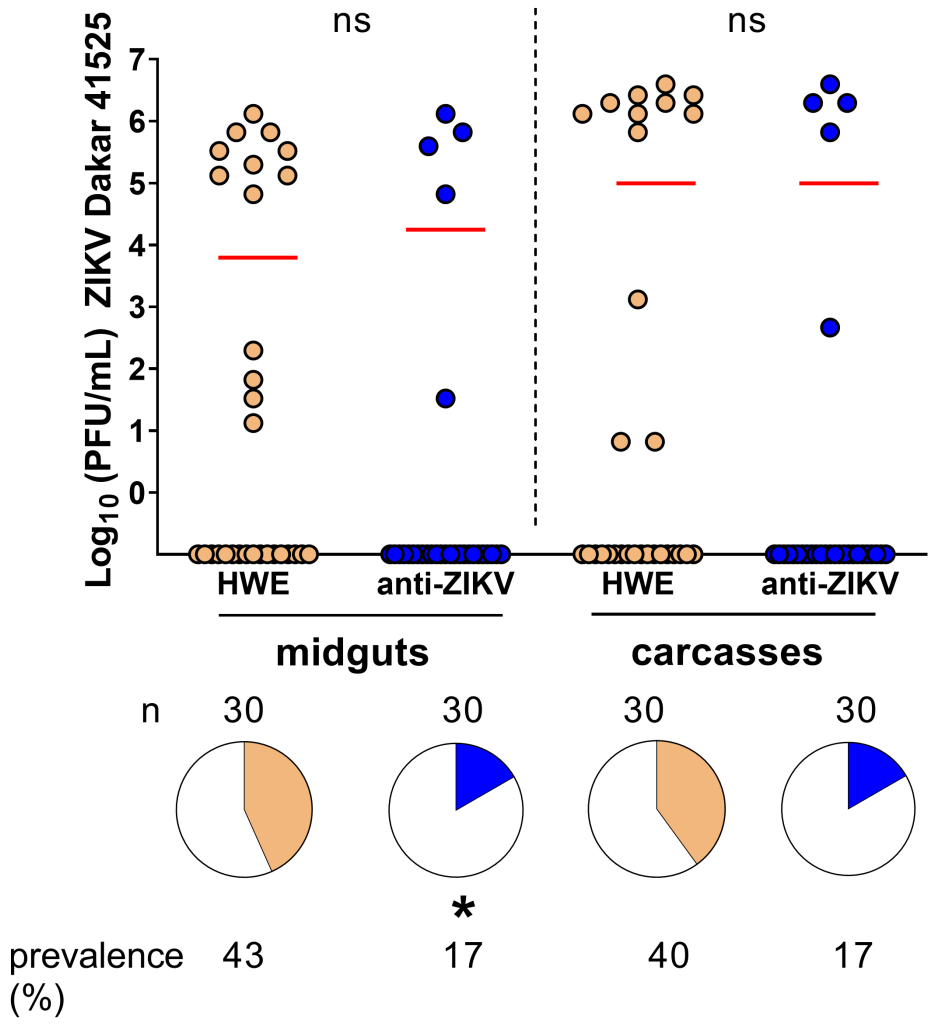


Figure 2.8. Anti-ZIKV-NS3/4A mosquitoes are significantly resistant to ZIKV Dakar 41525 in their midguts. Infected midguts (**left**) or carcasses (**right**) of anti-ZIKV-NS3/4A ("anti-ZIKV") and HWE mosquitoes at 14 dpi after infection with ZIKV Dakar. Red bars indicate median virus titers. "ns" above graph refers to virus titers between infected groups. Stars below pie charts compare infection prevalence; n = number of mosquitoes tested, ns = not significant; * = $p < 0.05$.

2.3 Discussion

Building from our previous success in generating DENV2-resistant *Ae. aegypti*, in this study, we sought to engineer ZIKV-resistant *Ae. aegypti* that would express a ZIKV-derived long dsRNA. We inserted the IR effector-containing transgene into a specific locus on chromosome 2q using CRISPR/Cas9-mediated site-specific insertion. We showed that the Chr2:321382225

locus can be used to reliably insert transgenes via CRISPR/Cas9 technology with as few as ~ 700 embryos needed to be injected to obtain a transgenic line. Site-specific transgene integration allows for the efficacy of various anti-ZIKV effectors to be compared side-by-side without being confronted with varying position effects. In addition, the ability to balance the genomic insertion locus with dominant markers expressing different fluorescent proteins allowed us to obtain homozygous populations within two genetic crosses.

Once we established the transgenic anti-ZIKV-NS3/4A line of mosquitoes, we challenged them with two ZIKV strains to test the efficacy of the IR effector to silence the viruses. We found that 90% of the transgenic mosquitoes blocked ZIKV (PRVABC59) infection in their midguts at 7 and 14 dpi, which also prevented disseminated infections. More than 80% of these mosquitoes released no virus in their saliva and showed no infections in their salivary glands as measured by transmission assays. We confirmed that the observed resistance was due to the midgut infection barrier induced by the transgene, which was expected since the *CpA* promoter is tissue-specific for the midgut [221]. Our RNAi-based strategy, however, appears to be relatively sequence-specific, or at least dependent on ZIKV replication dynamics; significantly fewer transgenic mosquito midguts (the tissue in which the transgene is expressed) became infected with a heterologous ZIKV strain as compared to controls, while marginally fewer transgenic mosquito carcasses were infected as compared to controls (Figure 2.8).

The anti-ZIKV-NS3/4A sequence expressed in our transgenic mosquito line was selected based on previous findings in which mosquitoes were highly protected against ZIKV (PRVABC59) when intrathoracically injected (prior to virus infection) with dsRNAs derived from the viral sequence [222]. The anti-NS3/4A sequence (in [222], this same sequence is termed “dsZIKV5”) overlaps with the last 183 bp of the NS3 region and 355 bp of the NS4A

region (nearly the entire NS4A sequence) of ZIKV PRVABC59. In Magalhaes et al. (2019), mosquitoes were intrathoracically inoculated with 250 ng of five different long dsRNAs targeting ZIKV, including the anti-ZIKV-NS3/4A dsRNA [222]. Three days later, all mosquito groups, including three control groups, received an artificial virus-containing bloodmeal to assess virus resistance [222]. Only one mosquito injected with the anti-ZIKV-NS3/4A dsRNA was infected at 7 dpi, and no mosquitoes were infected at 14 dpi [222]. These results contrasted with those from our transgenic mosquitoes, where 10% of the mosquitoes consistently became ZIKV infected at 7 or 14 dpi. These differences may have been because the mosquitoes in [222] received a higher dose of the anti-ZIKV/NS3/4A dsRNA by intrathoracic injection as compared to the number of anti-ZIKV-NS3/4A dsRNA molecules expressed in our transgenic mosquitoes. Timing may also be a factor influencing virus resistance—the mosquitoes that received the anti-ZIKV-NS3/4A dsRNA by injection were challenged with ZIKV (PRVABC59) three days later, whereas the anti-ZIKV-NS3/4A transgene is expressed in the mosquito midgut at the time of infection. As a point of similarity, mosquitoes that were injected with the anti-ZIKV-NS3/4A dsRNA were less resistant to divergent ZIKV strains, as we observed here for the transgenic anti-ZIKV-NS3/4A mosquitoes.

It is also worth noting that although Magalhaes et al. (2019) challenged mosquitoes with a higher ZIKV titer (8.7×10^6 PFU/mL) than we did (~10^5 PFU/mL), the authors used frozen ZIKV stocks, which have been shown to exhibit decreased infectivity when compared to fresh virus cultures [222,223]. In this study, we consistently challenged both the HWE control and the anti-ZIKV-NS3/4A mosquitoes with freshly cultured ZIKV at ~10^5 PFU/mL across the different experiments. We aimed to feed similar ZIKV concentrations across replicate experiments because it has been shown that mosquito infection prevalence is highly dependent on infectious

virus titers, as well as virus and mosquito strains [224]. It is therefore important that studies like ours and others report the infectious bloodmeal titers, as well as the genetic background of the mosquitoes tested, as these variables may confound the results.

This is the first transgenic line of *Ae. aegypti* engineered to trigger the endogenous siRNA pathway targeting ZIKV. Previously, Buchman et al. (2019) generated ZIKV-resistant transgenic *Ae. aegypti* that expressed synthetic sRNAs, which induced the miRNA pathway [201]. Five of eight synthetic miRNAs were processed in homozygous mosquitoes, which led to a 100% reduction in ZIKV genome equivalents in midguts, carcasses, and saliva [201]. These miRNAs targeted the capsid (C), pre-membrane (prM), NS1, and NS5 encoding regions of the ZIKV genome [201]. Given the fact that our results consistently showed that 10% of the transgenic mosquitoes were infected after siRNA targeting of the virus, it might be possible that ZIKV is less susceptible to the siRNA pathway. Indeed, it has been demonstrated that ZIKV subgenomic flaviviral RNA (sfRNA) interacts with specific mosquito proteins and suppresses RNAi in *Ae. aegypti* [225]. This viral immune evasion strategy may be a limiting factor when attempting to enhance RNAi-based immune responses *in vivo* to engineer ZIKV-resistance in mosquitoes. Additionally, targeting multiple regions of the ZIKV genome may be more efficacious in viral blocking than just targeting a single long region. The observed discrepancies between the viral silencing efficiencies of siRNAs and synthetic miRNAs could also be caused by different expression patterns of Dicer-1 and Dicer-2. The siRNA pathway uses Dicer-2 to process long dsRNAs into siRNAs; it is possible that Dicer-2 is less strongly expressed than Dicer-1 and therefore produces proportionally fewer siRNAs as compared to Dicer-1-mediated miRNAs.

Although in their respective transgenic *Ae. aegypti* lines the anti-ZIKV IR effector of this work was processed into siRNAs at a similar rate as an anti-DENV2 IR effector in a previous study [193], we observed lower rates of resistance against ZIKV as compared to DENV2. We consistently observed that 93–100% of transgenic anti-DENV2 IR *Ae. aegypti* were resistant to DENV2 (but not to other serotypes) at 14 dpi regardless of the DENV2 genotype used in the challenge experiment; we tested four different DENV2 strains that were 90-96% similar to the Jamaican 1409 strain which the transgene IR is based, and transgenic anti-DENV2 mosquitoes were resistant to all of them [193]. However, in this study, we found that only 83–90% of the anti-ZIKV-NS3/4A mosquitoes were resistant to ZIKV, depending on virus strain used in the challenge experiment, at 14 dpi. Furthermore, the anti-ZIKV-NS3/4A mosquitoes that did become infected with ZIKV exhibited similar titers as the controls. Because the anti-ZIKV-NS3/4A mosquitoes were homozygous for the transgene, and because we screened all transgenics for fluorescent eye coloration before experiments, we do not believe that the transgenic mosquitoes that remained susceptible to the virus lacked the transgene. Furthermore, given that we have engineered 93–100% DENV2 resistance by this same strategy in *Ae. aegypti* of the same genetic background (HWE), we doubt that this susceptibly is due to genetic differences between mosquitoes in siRNA processing. However, it is possible that even with robust effector transgene expression, the siRNA pathway may become saturated. ZIKV may also be less susceptible to siRNA targeting as compared to DENV2, which could suggest that viruses of the same family (in this case, *Flaviviridae*) are not equally susceptible to mosquito innate immune mechanisms or various types of antiviral effectors.

Future work aimed at the development of antiviral effectors that target multiple arboviruses synchronously is needed since several arboviruses may co-circulate in a given

region. This study was proof-of-concept, but further optimization may improve rates of resistance. For example, engineering mosquitoes to express several long dsRNAs targeting multiple regions of the ZIKV genome, multiple ZIKV strains, or even multiple arboviruses could broaden the transgenic approach explored in this chapter. Along these lines, in our lab, we crossed the anti-DENV2 transgenic *Ae. aegypti* that express a DENV2-specific IR at the Carb109 locus [191,193] with the anti-ZIKV mosquitoes discussed herein that express the ZIKV-specific IR at the Timp-P4 locus. We are in the process of evaluating DENV2- and ZIKV-specific sRNA production in this transgenic line as well as evaluating resistance against both viruses. Using complementary antiviral effector strategies in tandem may also provide an avenue toward achieving resistance against multiple arboviruses. For example, in addition to generating ZIKV or DENV resistance by triggering *Ae. aegypti* innate immunity, DENV1-4 resistance has been achieved through transgenic expression of a single chain antibody [203] and CHIKV resistance has been achieved through transgenic expression of an antiviral ribozyme [197]. Using multiple strategies will minimize the risk of viruses developing resistance against these effectors. Ultimately, novel vector control techniques targeting *Ae. aegypti* populations will rely on potent antiviral effectors that block virus replication *in vivo* and, in this way, virus transmission.

2.4 Future work aimed at coupling antiviral effectors with gene drive technology[4]

The characterization of antiviral effectors in transgenic *Ae. aegypti* as discussed herein and elsewhere has fueled interest in the genetic manipulation of mosquitoes as additional tools to

[4] This section includes parts of the manuscript cited as "Williams, A.E., Franz, A.W.E., Reid, W.R., Olson, K.E. Antiviral effectors and gene drive strategies for mosquito population suppression or replacement to mitigate arbovirus transmission by *Aedes aegypti*. *Insects*. 2020; *11*(1): 52." The article is reproduced with permission and minor modifications have been made.

combat disease transmission. Population replacement is based on replacing wild-type populations with transgenic insects that have been engineered to be resistant to pathogen infection/transmission. However, the use of transgenes to confer arboviral resistance may carry unintended fitness costs, ultimately leading to their loss in wild-type populations. Coupling the effector to a gene drive system can overcome fitness costs by pushing the transgene through the population at levels higher than expected from typical Mendelian inheritance. Gene drive, or meiotic drive, is the super-Mendelian inheritance pattern of a selfish genetic element (SGE) that allows it to rapidly spread through populations, even if it does not improve the survival or reproduction of its host [226] (Figure 2.9). Gene elements that are inherited from a parent with an allele frequency of > 50% are considered to exhibit gene drive [226].

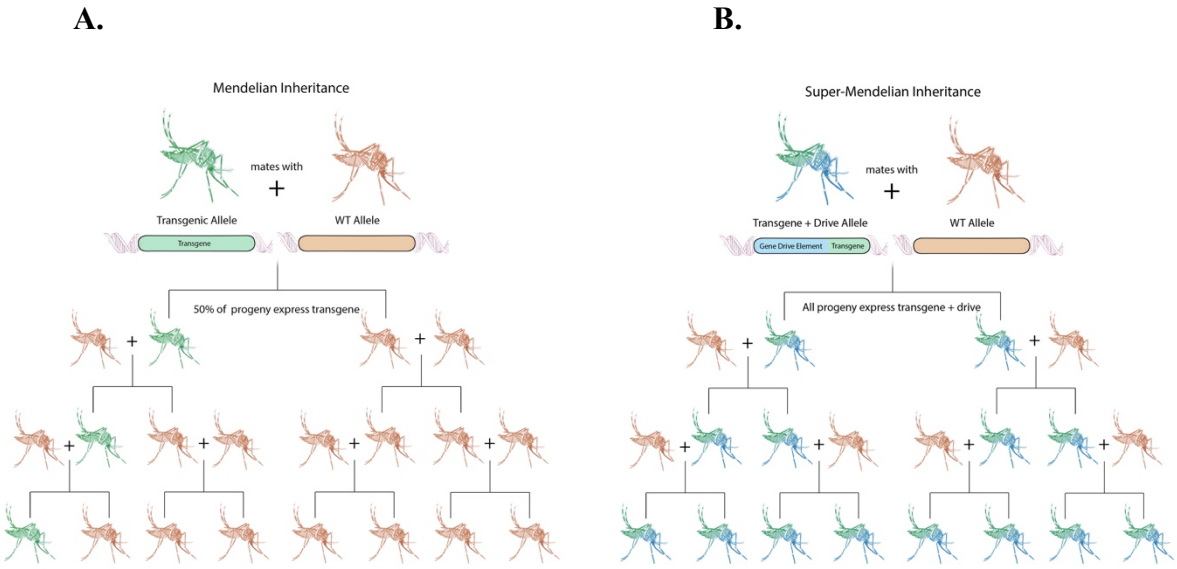


Figure 2.9. Mendelian versus "Super" Mendelian inheritance. (**A**) Mendel's law of independent assortment predicts an inheritance rate of 50% for a transgene when it is not sex-linked. Without repeated introduction, loss of the transgene is expected because of multiple factors, including genetic drift and fitness cost of the transgene. (**B**) Homing endonuclease-based gene drives supersede Mendel's law of independent assortment by converting wild-type alleles into gene drive-bearing alleles in the germline. This then leads to fixation of the gene drive in the target population.

Homing endonucleases such as CRISPR/Cas9 are SGEs that induce a double-stranded break (DSB) at a specific target site of the genomic DNA to insert a copy of the genetic element into the target site [227–229]. SGEs can be genetically linked to exogenous transgenes such as antiviral effectors to allow for super-Mendelian inheritance of the effector gene even if it carries a fitness load for the organism. However, the design and application of antiviral effectors in combination with SGEs to convert pathogen-susceptible wild-type *Ae. aegypti* populations into pathogen-resistant populations requires several important aspects for consideration. These systems need to be heritable and robustly transmitted to the subsequent generations of the target population. Effective gene drive systems must have the following attributes [230]:

(1) Compensate for any loss of fitness associated with the effector gene

(2) Link tightly to complex effector genes that are associated with a fluorescent marker

(3) Drive the effector gene relatively quickly to fixation within the target population

(4) Adapt to genetically diverse strains of mosquitoes

(5) Remain confined to the targeted species irrespective of population structure and mating dynamics between species

(6) Resist mutations that diminish or block drive to be sustained in nature

(7) Be socially accepted by those communities who might benefit

The CRISPR/Cas9 machinery along with the appropriate sgRNA can be designed as a gene drive system to be allele-specific and inherited by subsequent generations (Figure 2.10). A CRISPR/Cas9-based gene drive construct must generally encode, at a minimum: (1) flanking homology arms serving as a DNA template complementary to the mosquito genome to facilitate HDR-mediated knock-in of the transgenic cargo, (2) the Cas9 enzyme, under control of a germline-specific promoter, and (3) an endogenously expressed sgRNA [231]. Recently, a study

examined the efficiency of 12 *Ae. aegypti* RNA polymerase III U6 promoters, four of which, including the U6 promoter initially described by Konet and colleagues [231], were found to facilitate efficient CRISPR/Cas9 mediated genome editing [232,233].

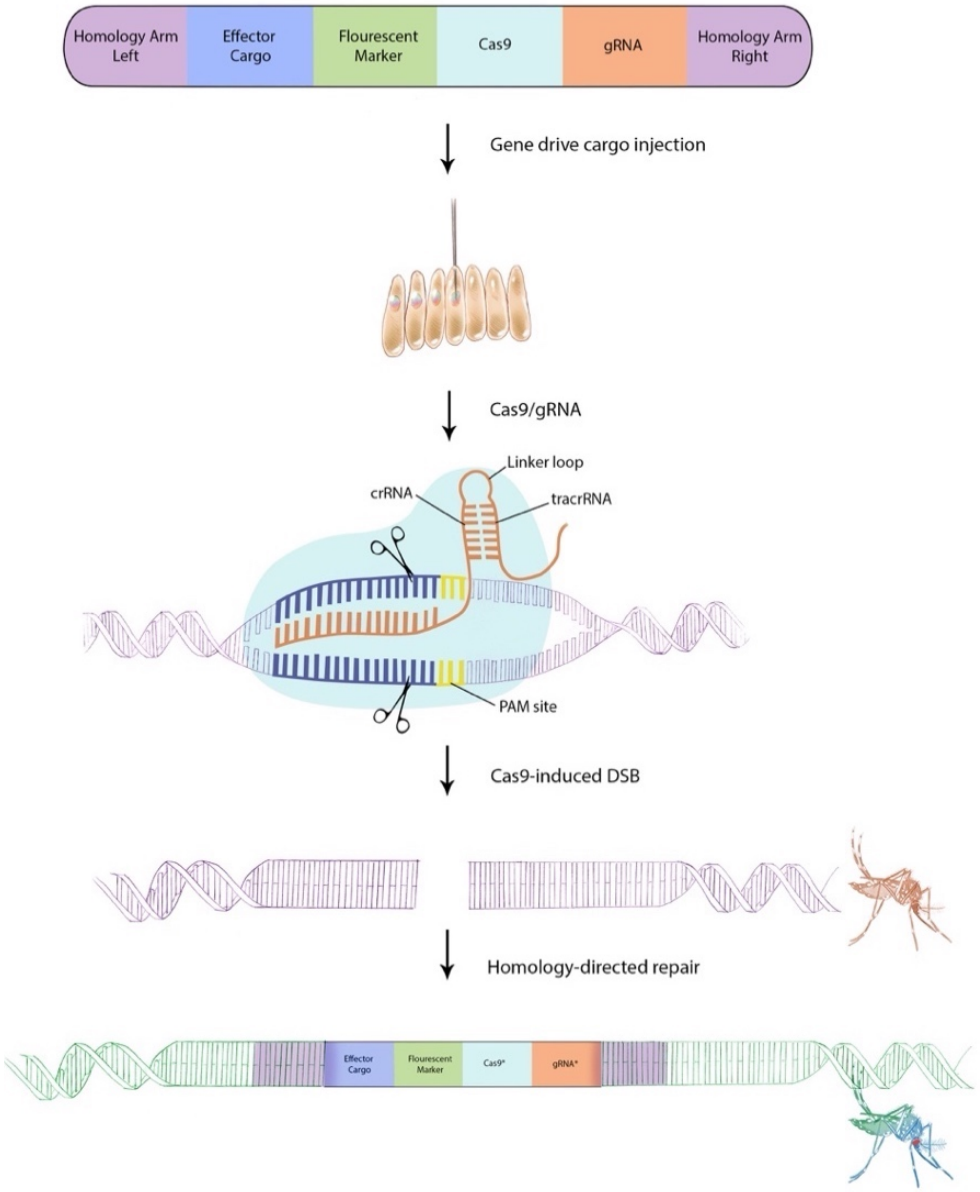


Figure 2.10. Generalized flow chart for the establishment of a one-component CRISPR/Cas9-based gene drive system in *Ae. aegypti*. Antiviral effector cargo, when expressed alongside the necessary components of the CRISPR/Cas9 system, can be driven into mosquito populations. Such a gene drive construct would contain (1) flanking homology arms that are complementary to the CRISPR/Cas9 target site in the mosquito genome, (2) the antiviral effector under control of a tissue-specific promoter, (3) a discernable marker (such as a fluorescent protein under control of a photoreceptor-specific promoter), (4) the Cas9 enzyme

under control of a germline-specific promoter, and (5) a sgRNA under control of an RNA polymerase III (U6) promoter. The construct is injected into pre-blastoderm embryos for site-specific germline integration. Once expressed, the sgRNA forms a complex with the Cas9 enzyme and guides it to complementary sequences upstream of a PAM site. The Cas9 enzyme then induces a DSB 3–5 bp upstream of the PAM site. If the cell uses HDR to repair the DSB, the homologous sequence in the donor plasmid will be used as template to repair the DSB. The G1 offspring then represent transgenic mosquitoes, which are capable of gene driving by targeting the wild-type allele with CRISPR/Cas9 and repairing the cleavage though HDR, thereby inserting the entire gene drive system including the antiviral effector cargo.

There are several challenges regarding the design of such a gene drive construct. An optimal genome locus must be identified to allow stable, site-specific insertion of the gene drive system. This locus should be highly conserved among the diverse populations of *Ae. aegypti* mosquitoes to facilitate robust spread of an antiviral effector gene. The efficiency of various sgRNAs to target a specific locus must also be evaluated. Because Cas9 efficiency has been reported to vary significantly between different target sites, the optimal sgRNA sequence should be identified in a comparative assay. If the targeted locus is not conserved among individuals of the target population or is prone to mutations or indel formation, the selected sgRNA will no longer be complementary to the target sequence. Consequently, the CRISPR drive (along with the antiviral effector) can no longer be inserted into the homologous allele; the gene drive system stalls, meaning that the antiviral effector will no longer be passed on to subsequent generations. In addition to naturally occurring sequence variation that could prevent a sequence homology dependent gene drive from spreading, resistance to the gene drive system could also develop during the repair process of the genomic DNA following its cleavage by Cas9. If the DNA repair undergoes homologous recombination with the gene drive-containing DNA template, the drive system will continue spreading through the population. If, however, the DNA repair undergoes NHEJ, the CRISPR/Cas9 target site becomes modified and is no longer recognized by the Cas9/sgRNA complex. In addition, maternal effect, when the phenotype of the

offspring is influenced by the mother's genotype or phenotype as dictated by her environment, can result in drive-resistant allele formation [234]. Generally, sequence homology dependent gene drive systems are sensitive to the development of drive-resistant alleles [235].

Several strategies have been developed to address these problems. For example, multiplexing the CRISPR/Cas9 construct by using more than one sgRNA has been shown to efficiently result in multiple simultaneous gene disruptions [236], and experiments with *D. melanogaster* have shown that sgRNA multiplexing significantly reduces resistance allele formation rates [237]. Since the efficiency of the CRISPR/Cas9-mediated gene drive heavily relies on the timing and the level of expression of its components in specific tissues of the insect, the choice of optimal promoters for Cas9 and sgRNA expression is regarded as one of the most critical aspects of the overall gene drive design. Restricting Cas9 expression to the germline has been shown to improve HDR rates [238,239]. Li et al., (2017) generated several *Ae. aegypti* lines that stably expressed Cas9 in the germline [236]. Using an optimal polymerase III promoter for sgRNA expression, and the *exuperentia* or *ubiquitin L40* promoter for Cas9 expression, that same group developed the first two-component CRISPR/Cas9-based gene drive (split drive) system for *Ae. aegypti* based on the "Copy Cat" approach. This split drive system was inherited with a rate exceeding 90% over several generations [233,236].

In conclusion, genetic control strategies targeting *Ae. aegypti* represent a novel set of tools for vector control programs. Population modification techniques in *Ae. aegypti* are becoming especially attractive now that a range of tissue-specific promoters have been identified, antiviral effector genes have been optimized and tested under laboratory conditions, and efficient gene drive systems are being developed. Despite the challenges discussed in this section, coupling an antiviral effector, such as the ZIKV-specific IR discussed herein, to a gene

drive element in transgenic mosquitoes is the next step in the development of a population replacement vector control strategy.

2.5 Materials and Methods

2.5.1 Mosquito rearing and maintenance

All *Ae. aegypti* mosquito colonies were maintained at 28 °C with 75–80% relative humidity and a 12 h light/12 h dark cycle. Routine maintenance regimens are described in [193]. Briefly, mated females were fed artificial bloodmeals consisting of defibrinated sheep blood (Colorado Serum Co., Denver, CO, USA) and 10 mM ATP approximately 4 days post emergence. Females were encaged with oviposition cups (consisting of paper towel strips and small water-filled plastic cups) for 5 days, and the eggs were then retrieved and dried. Stored eggs were viable for up to 3 months. Eggs were hatched in sterile water, and larvae were fed with ground TetraMin (Melle, Germany) fish food.

2.5.2 Identification of active sgRNA target sites

CHOPCHOP [240,241] was used to design four sgRNAs as close to the original *mariner Mos1* insertion site [101] as possible while avoiding any predicted off-target sequences. The genomic DNA from a pool of 10 female and 10 male *Ae. aegypti* (HWE strain [242]) was sequenced across the locus containing the sgRNA target sites to confirm their presence and integrity. Each of the sgRNAs was then tested for DNA cleavage activity in the mosquito embryo by injection of three sets of ~100 *Ae. aegypti* embryos for each sgRNA. The injection mixes contained 300 ng/μL Cas9-NLS protein (PNABio, Thousand Oaks, CA, USA) complexed with 80 ng/μL sgRNAs synthesized using the ENGen sgRNA kit (NEB, Ipswitch, MA, USA). Embryos were collected from hypergravid females over a 15 min period and then manually

aligned using a fine spotting paint brush, transferred to double-face Scotch tape (Scotch Brand, St. Paul, MN, USA), and covered with Halocarbon 27 oil (Millipore Sigma, St. Louis, MO, USA). No later than 30 min after collection, preblastoderm embryos were then injected using a Femtojet microinjector (Eppendorf, Hamburg, Germany) set to a constant injection pressure of 600 hPa and a backpressure of 250 hPa. The Halocarbon 27 oil was then immediately washed from the embryos with deionized water, and the embryos were allowed to develop for 16–24 h in a humid Petri dish prior to genomic DNA extraction using DNAzol (Thermo Fisher Scientific, Waltham, MA, USA). PCR products spanning the sgRNA target sites were then amplified using primers BR-20 and BR-23 (Supplemental Table 2.1), gel purified using the Zymo gel purification kit (Zymo Research, Irvine, CA, USA), Sanger sequenced at the University of Missouri DNA Core (Columbia, MO, USA), and assessed for trace sequence decay using the Inference of CRISPR Edits (ICE) tool from Synthego (Synthego Performance Analysis, ICE Analysis. 2019. v2.0. Synthego).

2.5.3 Construction of donor plasmid DNAs

To initially test whether the Chr2:32138225 locus could be successfully targeted via homology-directed DNA repair using CRISPR/Cas9, we designed a donor plasmid containing the enhanced cyan fluorescent protein (eCFP) coding sequence under control of the photoreceptor-specific 3xP3 promoter [243,244]. The transcription terminator originated from the large T-antigen encoding gene of SV40. This expression cassette was then flanked by an upstream homology arm amplified from the HWE strain of *Ae. aegypti* containing the genomic sequence upstream of the CRISPR/Cas9 target site (1238 bp; Chr2:321380975-321382213) and a downstream homology arm containing the genomic sequence downstream of the CRISPR/Cas9 target site (1743 bp; Chr2:321382225-321383968).

The anti-ZIKV effector DNA construct was based on this eye marker construct into which the IR effector expression cassette was inserted. The IR molecule consisted of 538 bp cDNA sequences derived from the ZIKV NS3/4A encoding region in sense and antisense orientations [222], which were separated by the small *sialokinin1* intron [245]. The IR molecule was placed under control of the bloodmeal inducible, midgut-specific *CpA* promoter [191,246]. The same transcription terminator and homology arms as described above were used for this construct, now containing eye marker and IR effector.

Plasmid construction was performed based on a combination of conventional restriction enzyme-mediated cloning and Gibson assembly-based cloning. The annotated sequences for both DNA constructs, the eye marker-based reporter and the complete anti-ZIKV effector construct, are available at NCBI under accessions MT926371 and MT926370, respectively. All primer sequences and the gBlock sequence for the anti-ZIKV effector and small *sialokinin1* intron are provided in Supplemental Table 2.1.

2.5.4 Establishment of a transgenic line of Ae. aegypti *containing an anti-ZIKV IR effector*

Donor plasmids were isolated using the Zymo plasmid midiprep kit (Zymo Research, Irvine, CA, USA) and added at a final concentration of 80 fmol/µL to an injection mix containing 300 ng/µL Cas9-NLS (PNABio), 80 ng/µL sgRNA, and 100 ng/µL ku70 dsRNA. Preblastoderm embryos were collected from hypergravid *Ae. aegypti* females (HWE strain) over a 15-min period, aligned using a fine spotter paint brush for an additional 20–30 min, then transferred to double-face Scotch brand tape and covered with Halocarbon 27 oil (Sigma-Aldrich, St. Louis, MI, USA). The posterior ends of the embryos were then injected following the same methodology as used for the sgRNA activity testing. The oil was rinsed off and the embryos were transferred to moistened Kimwipe tissue and allowed to develop for 7 days prior

to hatching. Surviving G0 males were individually outcrossed to 7–10 virgin HWE females, allowed to mate for 5 days, then pooled. Surviving G0 females were mass-crossed to an equal number of HWE males and allowed to mate for 5 days. Outcrossed pools were provided with three subsequent bloodmeals (defibrinated sheep blood, Colorado Serum Co.) and allowed to lay the eggs of the G1 generation. The G1 generation was subsequently hatched and screened for the presence of the eCFP marker and survivors were individually outcrossed a second time to HWE *Ae. aegypti* and then screened by PCR to confirm transgene integration into the Chr2:321382225 locus. Genomic DNA was extracted from whole mosquitoes using DNAzol (Thermo Fisher Scientific, Waltham, MA, USA). Transgene integrity and integration were confirmed by generating three different PCR amplicons — 1. spanning the left homology arm (BR-20) and the *sialokinin1* intron (BR-347); 2. spanning the *CpA* promoter (BR-348) and the *sialokinin1* intron (BR-345), and 3. spanning the *sialokinin1* intron (BR-223) and eCFP (BR-348)—using Thermo Fisher Superscript II DNA polymerase under the following cycling conditions: 98 °C for 3 min followed by 35 cycles of 98 °C for 20 s, 62 °C for 20 s, 72 °C for 1 min, and a final extension step of 72 °C for 5 min. Critically, the ramping rate for annealing was increased to 3 °C per second, which allowed for PCR product formation to proceed without impairment by the hairpin. DNA sequences of the amplicons were confirmed by Sanger sequencing. All primers sequences are listed in Supplemental Table 2.1.

Finally, the transgenic anti-ZIKV effector harboring mosquitoes (eCFP marker) were outcrossed to another transgenic line of *Ae. aegypti* (HWE strain) containing a 3xP3-mCherry (red) eye marker also integrated into the Chr2:321382225 locus. The F1 progeny were then screened to obtain individuals containing both red and blue eye markers; a subset of the heterozygous individuals was used for sRNA profiling, while remaining heterozygotes were

reciprocally crossed and screened to obtain true single-locus homozygotes (blue-eye marker only). Homozygosity was confirmed in a sample of the resulting line by outcrossing transgenic males to wild-type virgin HWE *Ae. aegypti*.

2.5.5 Small RNA sequencing

Adult female mosquitoes were fed a non-infectious artificial bloodmeal approximately 4 days post-emergence. Midguts were dissected 24 h later, cleaned of blood, and placed in TRIzol (Thermo Fisher Scientific, Waltham, MA, USA) for RNA extraction following the manufacturer's instructions. Each sample was a pool of 15 cleaned midguts/sequencing library. Total RNA samples were then sent to the University of Missouri DNA Core, where TruSeq sRNA libraries were prepared and subjected to Illumina NextSeq Mid Output SE75 deep sequencing. Small RNA sequencing analyses were performed using a pipeline developed in-house by Dr. Greg Ebel's lab [159]. The sRNA data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus [247] and are accessible through GEO Series accession number GSE156825. The HWE sRNA dataset is presented under accession number GSM4745099 and the anti-ZIKV-NS3/4A sRNA dataset is presented under accession number GSM4745100.

2.5.6 Virus challenge experiments

ZIKV isolates used in this study were PRVABC59 of the Asian lineage (accession number KU501215) and Dakar 41525 of the African lineage (accession number KU955591). ZIKV was propagated in Vero cells at a 0.01 multiplicity of infection (moi) for 72 h using Dulbecco's modified Eagle medium (DMEM) supplemented with inactivated 3% fetal-bovine serum (FBS). Infected cells were then pelleted, resuspended in a small volume (~3 ml) of infected cell culture supernatant, and added to defibrinated sheep blood (Colorado Serum Co.) at

a 1:1 (vol/vol) ratio. Mosquitoes were fed for ~1 h using an artificial membrane feeder that maintained ~1–2 mL blood-virus mixture at 37 °C for each carton. Engorged females were visually selected after feeding and were maintained in 64 oz. cartons supplied with sucrose and water until further analysis.

2.5.7 Mosquito tissue plaque assays for ZIKV detection

Tissue samples were homogenized in 500 µL (midguts or salivary glands) or 1000 µL (carcasses) DMEM (7% inactivated FBS, 1% penicillin/streptomycin, 1% glutamine, 1% non-essential amino acids). Each sample was then passed through a 0.2 µm Acrodisc Syringe Filter fitted with Supor Membrane (Pall Life Sciences, East Hills, NY, USA). Vero cells were seeded in 24-well plates and were left for three days to achieve confluence. Cells were infected with 10-fold serial dilutions of the homogenates (up to $1/10^5$ PFU/mL) for 1 h at 37 °C. After infection, 1 mL of a sterilized 1% agarose solution containing a nutrient supplement (10% 1× Medium 199 (Sigma-Aldrich), 5% inactivated FBS, 4% sodium bicarbonate, 2% diethylaminoethyl (DEAE)-dextran, 0.5% MEM amino acids (Mediatech Inc., Manassas, VA, USA), 0.5% MEM vitamins) was overlaid on each well. Plates were left to solidify for ~1 h and were then moved to the 37 °C incubator for 6 days. To visualize plaques, 150 µL of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT, 3 mg/mL in 1× phosphate-buffered saline [PBS]) was added to each well followed by ~24 h incubation. Plaques were visually quantified the next day. Viral titers of each sample were calculated as plaque-forming units per milliliter (PFU/mL).

2.5.8 Intrathoracic Inoculation of ZIKV

Mosquitoes were intrathoracically inoculated with ZIKV as described previously [248,249]. Three-day-old females were anesthetized at 4 °C and inoculated with 100 PFU of virus suspended in a 69 nL volume of growth medium. Two days later, a subset of inoculated

mosquitoes was exposed to artificial bloodmeals consisting of defibrinated sheep blood (Colorado Serum Co.) and 10 mM ATP, while the other group of mosquitoes was maintained on a sugar diet. Eight days post-virus injection, whole mosquitoes were processed for plaque assay as described above.

2.5.9 ZIKV transmission assays

Saliva was collected from female mosquitoes at 14 dpi as previously described [250,251]. Legs and wings were removed from the mosquitoes, and the proboscises were inserted into a 1 µL capillary (microcaps, Drummond Scientific Company, Broomall, PA, USA) filled with immersion oil type B. Mosquitoes were allowed to salivate into the oil at room temperature for 1 h. The oil containing the saliva was expelled under pressure into 1.5 mL Eppendorf tubes containing 300 µL DMEM medium (20% FBS, 1% penicillin/streptomycin, 1% glutamine, 1% non-essential amino acids) and flash frozen on dry ice. Capillaries were visually analyzed for the presence of saliva, and capillaries that did not contain trace amounts of saliva were discarded. Following salivation, salivary glands were dissected from the same mosquitoes and placed into 1.5 mL Eppendorf tubes containing 500 µL DMEM medium (20% FBS, 1% penicillin/streptomycin, 1% glutamine, 1% non-essential amino acids). Corresponding carcasses were also collected. Saliva, salivary glands, and carcasses were frozen at −80 °C and were processed for plaque assay as described above. Saliva samples were not filtered before cell infection.

2.5.10 Immunofluorescence assays to detect ZIKV antigen

Dissected tissues were fixed in 4% paraformaldehyde and permeabilized with 0.2% Triton X-100. Immunofluorescence assays were performed using the monoclonal antibody 4G2 (1:200 in PBS) targeting a conserved epitope of the flavivirus E protein, as well as using a ZIKV

NS1-specific mouse monoclonal antibody (1E11, Immune Technology Corporation, 1:200 in PBS). Anti-mouse IgG, biotinylated species-specific whole antibody from sheep (Amersham BioSciences, Cat. # RPN1001V1), was used as secondary antibody (1:200 in PBS, supplemented with 1% Evan's blue counterstain). Detection was achieved by addition of Streptavidin-Fluorescein conjugate (Amersham Biosciences, Cat. # RPB1232V1; 1:200 in PBS). Slides were mounted using Mowiol (10%) supplemented with DABCO (1,4-diazobicyclo-[1.2.2]-octane) and visualized with an Olympus BH2 microscope.

2.5.11 Statistical analyses

All statistical analyses were performed with GraphPad Prism (version 8, LaJolla, CA, USA). Comparisons of virus titers were performed using the non-parametric Mann–Whitney U-test, excluding uninfected mosquitoes. A two-tailed Fisher's exact test was used to compare infection prevalence. Significance was defined as $p < 0.05$.

CHAPTER 3: SMALL RNAS IN GEOGRAPHICALLY DISTINCT *AEDES AEGYPTI*: IMPLICATIONS ON VECTOR COMPETENCE

3.1 Introduction

Ae. aegypti vary in their abilities to transmit arboviruses due to complex interactions between extrinsic environmental factors and intrinsic genetic factors [252,253]. One example is infection with insect-specific viruses (ISVs) that may impact mosquito antiviral immunity and interfere or enhance replication of medically relevant arboviruses. This phenomenon has been reported for many ISV-vector pairings. For *Culex pipiens,* the Culex Y Virus (*Birnaviridae*: unclassified) suppressed RNAi when its VP3 protein bound long dsRNA and blocked Dcr2-mediated siRNA production [254]. For *Coquillettidia xanthogaster*, Palm Creek virus (*Flaviviridae*: *Flavivirus*) suppressed West Nile virus replication in cell culture [255] and mosquitoes [256]. In *Ae. aegypti,* CFAV infection led to increased DENV replication, likely by promoting heightened expression of ribonuclease kappa, known to promote infection of viruses that enter cells by endocytosis [257]. Dual infection of *Ae. albopictus* cells (Aa23) with Phasi Charoen-like virus (PCLV; *Phenuiviridae*: *Phasivirus*) and CFAV significantly inhibited ZIKV, DENV, and LACV [258]. Finally, Aedes anphevirus (AeAV; *Xinmoviridae*: *Anphevirus*) modestly reduced DENV replication in Aa20 cells [259]. Furthermore, the endosymbiont *Wolbachia*, known to limit DENV replication in *Ae. aegypti* [260], enhanced AeAV replication [259].

Recent efforts to characterize the virome of culicines have led to the identification of many ISVs that are broadly distributed across *Ae. aegypti* populations, including CFAV [261,262], PCLV [263–265], AeAV [259], Dezidougou virus (unclassified: *Negevirus*) [266], Aedes aegypti densovirus (*Paroviridae*: *Brevihamaparvovirus*) [267], Humaita-Tubiacanga virus

(HTV; unclassified) [268], and verdadero virus (*Partitiviridae*: unclassified) [269]. Some of these viruses, like verdadero virus, were once thought to only infect plants or fungi [269]. However, recent additions to the sequencing data banks and improved annotations have revealed the host range of many viruses are broader than once thought. Because fungi are common in the mosquito microbiome and because they feed on nectar from plants, it is possible that viruses have "jumped hosts" during horizontal virus transfer [182]. Indeed, it is likely that many lesser known or unclassified ISVs infect *Ae. aegypti*, with some estimates of 27 ISVs infecting a single mosquito at a given time [270]. ISVs are closely associated with their vector hosts and are mostly transmitted vertically throughout mosquito populations, possibly through transovarial or venereal transmission in nature [181]. Given that ISVs have been shown to impact arbovirus replication, the ecology, distribution, and diversity of these viruses in the field will have important implications on vector competence for arboviruses in *Ae. aegypti* around the world.

Arbovirus replication may be impacted by both persistently infective ISVs as well as ISV non-retroviral integrated RNA virus sequences (NIRVS) that have been shown to have antiviral functions. Because ISVs persistently infect the germline at high prevalence across mosquito populations, most NIRVS are of ISV origin [271], although NIRVS genetically related to arboviruses have been reported in *Ae. aegypti* as well [174]. Next generation sequencing has revealed that NIRVS are almost exclusively processed into vpiRNAs of the opposite orientation of the virus genome, probably because endogenous retrotransposons reverse transcribe viral RNA into vDNA [127]; however, upon infection with a cognate virus, ping-pong amplification produces an abundance of both sense and antisense vpiRNAs that associate with Piwi proteins and silence viral RNA [130] (Figure 1.7). Some studies have suggested that there is potential cross talk between NIRVS-derived sRNAs and flaviviral infections [272]. For example, one

study showed that *Ae. aegypti* were persistently infected with CFAV because they harbored CFAV-specific siRNAs, and to a lesser extent, CFAV-specific vpiRNAs [272]. When these mosquitoes were infected with ZIKV, however, they expressed heightened levels of CFAV-specific vpiRNAs that aligned to different regions of the CFAV genome than the CFAV-specific vpiRNAs expressed by ZIKV-uninfected mosquitoes [272]. The CFAV-specific vpiRNAs did not have homology with ZIKV RNA but did have plus and minus strand ping-pong signatures [272], suggesting the presence of a CFAV NIRVS. These observations beg the question: does the presence of certain vsRNAs stimulate the expression of other vsRNAs? How ISV sRNA populations impact replication of other viruses, as well as their impacts on vector competence in the mosquito vector, remains poorly understood.

In field *Ae. aegypti*, NIRVS likely vary by ecotype as the result of distinctly circulating viruses infecting different mosquito populations. To investigate how sRNAs derived from both NIRVS and persistently infective ISVs compare in geographically distinct *Ae. aegypti*, we sequenced sRNAs derived from populations recently collected from Mexico, Brazil, and the southern USA. We chose these locations based on the diversity of viruses historically and presently circulating there, as well as differences in vector competences reported for mosquito populations across these areas [273]. For example, for *Ae. aegypti* in Mexico, the intersection of a neovolcanic axis (NVA) with the Gulf of Mexico was shown to be a barrier to gene flow, where mosquitoes north of this barrier were significantly more competent for DENV2 and displayed distinct mitochondrial haplotypes from mosquitoes south of this barrier [273] (Figure 3.1). Along these same lines, we hypothesized that *Ae. aegypti* from distinct geographic origins are persistently infected with unique ISVs and therefore have distinct sRNA profiles that may impact vector competence.

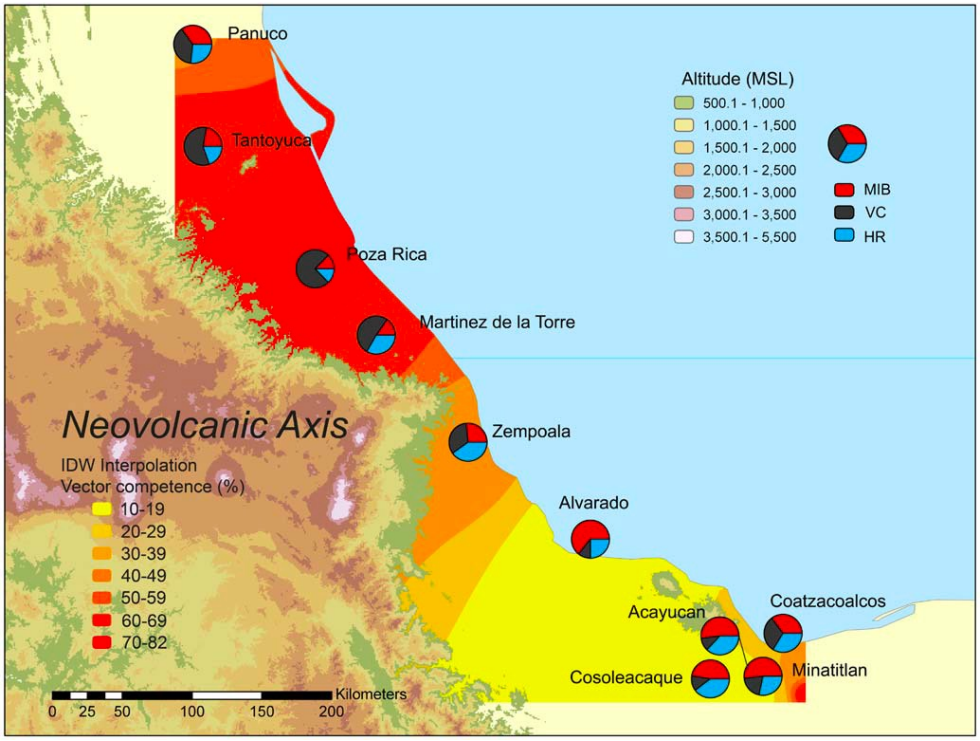


Figure 3.1. Map of the state of Veracruz, Mexico indicating *Aedes aegypti* DENV2 vector competence levels relative to the Neovolcanic Axis. *Ae. aegypti* were collected from 10 locations across the Veracruz coastal plan and their vector competences to DENV2 were compared. Pie charts indicate the proportion of mosquitoes that were vector competent (VC, black), exhibited midgut infection barriers (MIB, red), or blocked virus infection of the salivary glands (HR = head resistant, blue). Vector competency rates were interpolated by Inverse Distance Weighting (IDW), and geographic areas are colored from yellow to red according to predicted vector competence rates. Figure and accompanying legend from Lozano-Fuentes et al., 2009 [273].

Across mosquito strains sampled herein, we found that the overall distribution of sRNAs were variable, and that vsiRNAs and vpiRNAs were diverse and dependent on geographic origin. The most abundant ISV sRNAs were derived from Phasi Charoen-like virus, verdadero virus, chaq-like virus, Aedes anphevirus, and two Aedes aegypti totiviruses (*Totiviridae*: unclassified) first discovered in Guadeloupe (GuAaTV) and Ghana (GhAaTV). To gain insights on how ISV- and arbovirus-specific sRNAs interact, we infected mosquitoes from Poza Rica that exhibited diverse ISV-derived sRNA profiles with DENV2 (strain: Jamaica 1409) and repeated sRNA sequencing 7 dpi. DENV2-infected mosquitoes displayed heightened levels of total vsiRNAs, as

well as increased loads of specific ISV-derived sRNAs. Overall, our studies underscore the complexity of the mosquito virome, which may – along with other genetic factors – have significant impacts of vector competence in the field. Future studies comparing virus-derived sRNAs in mosquito strains with diverse viromes infected with arboviruses of medical importance are underway to compare RNAi cross talk and its impacts on virus replication and transmission.

3.2 Results

3.2.1 Small RNA distributions of mosquitoes uninfected with arboviruses are highly variable across Ae. aegypti *strains*

To gain an overall sense of the diversity of sRNA populations in WT mosquitoes from different geographic origins, we sequenced sRNAs from midgut and abdomen tissues of *Ae. aegypti* from across the Americas (Figure 3.2). Adult mosquitoes were reared from eggs that had been recently collected from Poza Rica, Veracruz state, Mexico in 2012, from cities throughout Chiapas, Mexico (Supplemental Figure 3.1) and interbred as a genetically diverse laboratory strain (GDLS) [274] in 2015, from Tapachula, Chiapas state, Mexico in a separate collection performed in 2019, and from Recife, Brazil in 2018. Notably, the NVA that acts as a barrier to gene flow for *Ae. aegypti* populations [273] separated Poza Rica in the state of Veracruz (northeast of the NVA) and the state of Chiapas (southwest of the NVA). Since their time of collection, eggs were maintained in colony in the laboratory and were renewed approximately every three months as described in the Materials and Methods section. We also included abdomen tissues from adult *Ae. aegypti* collected directly from the field in New Orleans, USA as a comparison to the samples that had been maintained in colony for at least one generation. In the case of the mosquitoes reared in the laboratory, for this study, we maintained one set on

sugar and offered another set a non-infectious bloodmeal and collected samples either 24 or 48 hours later. We were interested in how sRNAs differed between these populations in the midgut or after a bloodmeal because this is when and where mosquitoes may first be exposed to viruses of medical importance. We chose this timepoint because it is peak expression of the antiviral Piwi Piwi4 that may be involved in vsRNA processing [127] (see Chapter 4 for more information). We limited this preliminary analysis to somatic tissues because we were interested in detecting sRNAs that may crosstalk with arboviruses that replicate in the midgut and disseminate to secondary tissues as opposed to those that preferentially replicate in the germline. Pools of fifteen mosquitoes per sample, performed in triplicate, were collected for Illumina sRNA sequencing. We considered reads of 18-22 nt as siRNAs and reads of 24-32 nt as piRNAs, and all data shown is normalized by reads per million (RPM). 21-25 nt sRNAs that aligned to known miRNAs, allowing for 2 mismatches, were considered miRNAs, while 18-22 nt or 24-32 nt sRNAs that aligned to viral sequences were considered vsiRNAs or vpiRNAs, respectively.

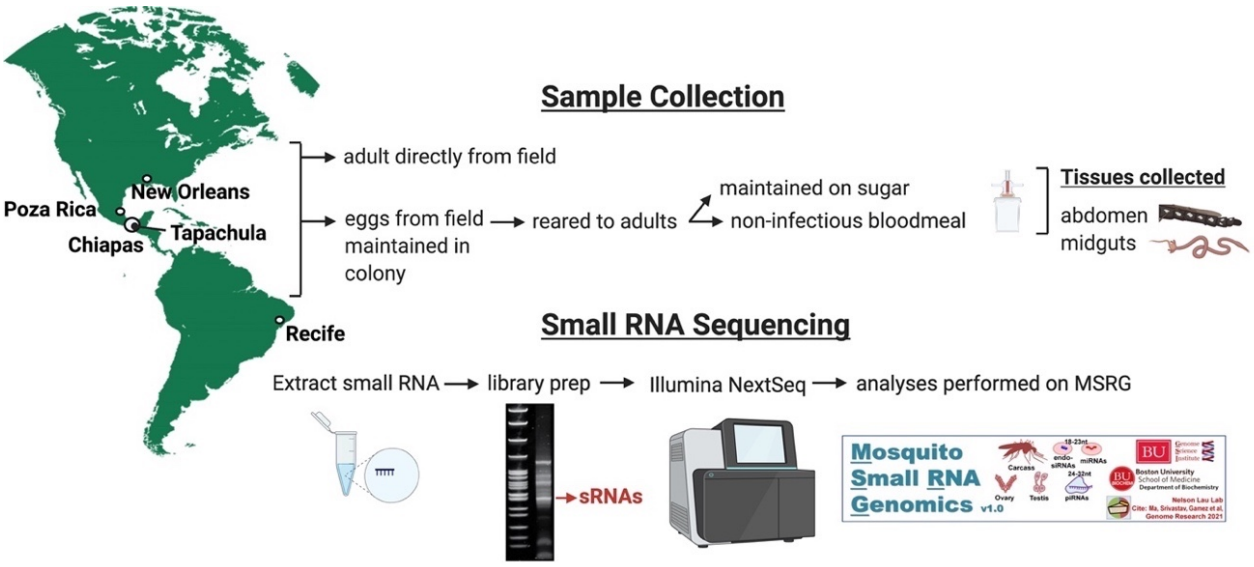


Figure 3.2. Mosquito collection sites and sample preparation. Mosquitoes were collected across the Americas from New Orleans, LA, USA (as adults), Poza Rica, Mexico (as eggs), Tapachula, Mexico (as eggs), throughout Chiapas, Mexico (as eggs) and Recife, Brazil (as eggs). Abdomen tissues (ovaries removed) from adults collected directly from the field in New Orleans,

USA were processed for sequencing. Eggs from the field were maintained in colony at CSU until the time of sample preparation. Adults reared in the laboratory were either maintained on sugar or provided a non-infectious bloodmeal. Somatic tissues (midguts or abdomens) were collected 24 or 48 hours later and processed for sRNA sequencing. Data was analyzed with the mosquito small RNA genomics (MSRG) pipeline.

The overall sRNA distributions differed by somatic tissue type in samples where midguts vs. abdomens were available. Midgut tissues exhibited a greater proportion of 18-22 nt siRNAs (42-48% of total RPM on combined average; Figure 3.3 A, B, E, F) than did abdomens (23-31% of total RPM on combined average; Figure 3.3 C, D, G, H). Abdomens, on the other hand, exhibited a greater proportion of 24-32 nt piRNAs (62-72% of total RPM on combined average; Figure 3.3 A, B, E, F) than did midguts (36-40% of total RPM on combined average; Figure 3.3 C, D, G, H).

When comparing abdomen tissues across mosquitoes, the overall distributions of total sRNAs were highly variable. For example, abdomens from Recife, Brazil harbored more piRNAs (69-72% of total RPM on combined average) than siRNAs (23-25% of total RPM on combined average; Figure 3.3 C, D), which was like abdomens from Poza Rica, Mexico that harbored 62-69% piRNAs and 25-31% siRNAs (Figure 3.3 G, H). On the other hand, abdomens from New Orleans, USA had roughly equal proportions of siRNAs (45-46% of total RPM on combined average) and piRNAs (44-47% of total RPM on combined average) (Figure 3.3 I, J). We also compared sRNAs in abdomens from a GDLS collected and interbred from mosquitoes collected in cities (including Tapachula) throughout the state of Chiapas, Mexico in 2015 (Supplemental Figure 3.1) to abdomens more recently collected in 2019 from a single city, Tapachula, in Chiapas, Mexico. Even though the GDLS and Tapachula mosquitoes have similar geographic origins, the sRNA distributions were different. GDLS abdomens from mosquitoes collected throughout Chiapas had similar percentages of siRNAs (44-46% on combined average)

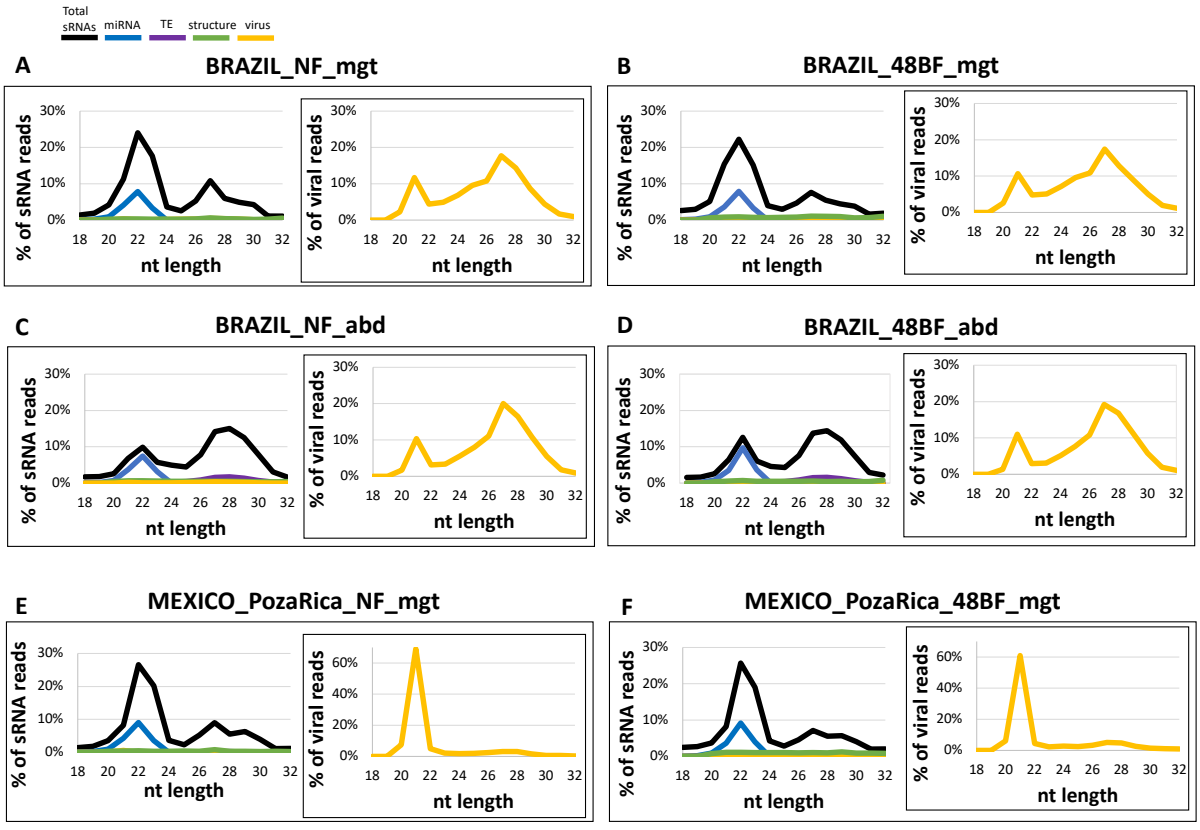
and piRNAs (40-43% on combined average), while those from the city of Tapachula had more piRNAs (51-56% on combined average) than siRNAs (34-36% on combined average) (Supplemental Figure 3.2).

When considering virus-derived sRNAs only (as opposed to *all* sRNAs), proportions of 18-22 nt vsiRNAs and 24-32 nt vpiRNAs were also highly variable across mosquitoes (Figure 3.3 inset graphs with yellow lines). For example, 12-15% and 65-81% of viral genome-derived sRNAs were vsiRNAs in the mosquitoes from Recife, Brazil (Figure 3.3 A-D insets) or Poza Rica, Mexico (Figure 3.3 E-H insets), respectively. The high variability in virus-derived sRNA distributions was true across all samples (Figure 3.3 insets; Supplemental Figure 3.2). A striking example are the mosquitoes from two streets in New Orleans: 33% or 60% of vsRNAs in mosquitoes from Burdette Street were vsiRNAs or vpiRNAs, respectively, while 52% or 37% of vsRNAs in mosquitoes from Freret Street were vsiRNAs or vpiRNAs, respectively (Figure 3.3 I, J inset graphs).

For the mosquitoes reared in the laboratory and offered an artificial bloodmeal, the overall size distribution of sRNAs did not differ by bloodfeeding status when 24 or 48 hour-post bloodfed samples were compared to their respective non-fed samples (Figure 3.3 A-H). This was true regardless of tissue type (abdomen vs. midgut) or geographic origin. These observations suggest that both the machinery required to process sRNAs and the RNA substrates that are processed are expressed independently of the acquisition or presence of an artificial bloodmeal. ISVs infect the germline and are transmitted transovarially (as opposed to during bloodfeeding) [275,276], so it is perhaps unsurprising that ISV-derived sRNAs would not be impacted by a bloodmeal. However, the bloodfeeding process is complex and often triggers the upregulation of many pathways necessary for metabolism, reproduction, and immunity [277]. We were therefore

surprised not to find any differences in sRNA distributions in the mosquitoes 24 or 48 hours post-bloodfed compared to those that had never bloodfed.

Most of the total sRNAs in this study were of unknown origin because overall, low proportions mapped to miRNAs (blue lines), transposable elements (purple lines), structural RNAs such as ribosomal or transfer RNAs (green lines), and viruses (yellow lines). The high variability in sRNA distributions may be due to differences in persistently infective viruses, microbiomes, nutritional or fitness parameters, or other unknown factors. Taken together, these results highlight the overall variability in sRNA production across mosquitoes and suggest that the extent by which mosquitoes are persistently infected with viruses differs across specimens.



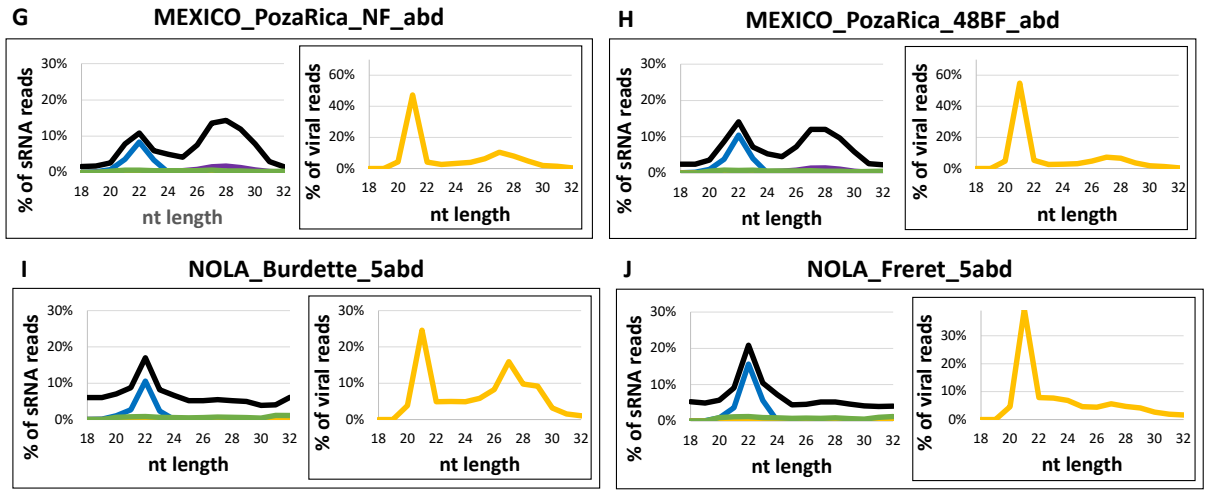


Figure 3.3. Small RNA size distributions in selected samples of this study. Proportions of sRNA reads per million (RPM) for total sRNAs (sRNAs, black), microRNAs (miRNA, blue), transposable elements (TE, purple), structural RNAs (*e.g.* tRNAs or rRNAs, green), and virus-derived RNAs (yellow), relative to total RPM, are shown by nucleotide (nt) length. Inset graphs show virus-derived sRNA RPM relative to total virus-specific RPM in yellow by nucleotide length. Samples are from **A-D)** Brazil = Recife, Brazil; **E-H)** PozaRica = Poza Rica, Mexico; **I-J** Burdette/Freret = streets in New Orleans (NOLA), USA from which samples were collected. NF = non-fed and maintained on sugar; = 48 hours post non-infectious bloodfeed; mgt = midgut tissue; abd = abdomen tissue, ovaries removed. All samples are uninfected with arboviruses.

3.2.2 Insect-specific viruses display unique patterns of infection dependent on ecotype

The mosquito small RNA genomics resource (MSRG) aligns sRNA reads to more than 220 mosquito arboviruses manually curated from NCBI GenBank and the Virus Pathogen Resource (VIPR) as of 2022 [272]. The pipeline also aligns reads to all EVEs collectively annotated from the *Ae. aegypti* genome. Using this resource, we compared virus-derived sRNAs in our sample set. All vsRNAs were of ISV origins, and the most abundant were specifically derived from Phasi Charoen-like virus, verdadero virus, chaq-like virus, Aedes anphevirus, and Aedes aegypti totiviruses originally identified in either Guadeloupe or Ghana (Figure 3.4). Mosquitoes from the state of Chiapas, Mexico (including those recently collected from the city of Tapachula) and Recife, Brazil overwhelmingly harbored more PCLV vsRNAs (76-99% of all vsRNA reads) compared to any other ISV surveyed. The only library within the PCLV sRNA-

rich subset that had any other notable ISV-specific vsRNA were GDLS mosquitoes 24 hours post-bloodfeeding that displayed 21% HTV-derived vsRNAs.

Mosquitoes from Poza Rica, Mexico and New Orleans, USA mostly lacked PCLV and displayed more diverse vsRNA profiles. For those from Poza Rica, Mexico, verdadero and chaq-like virus vsRNAs dominated the dataset; for those from New Orleans, USA, two totiviruses – Guadaloupe and Ghana Aedes aegypti totiviruses – dominated the dataset. Mosquitoes from both Poza Rica and New Orleans also displayed high AeAV-derived vsRNA RPMs, ranging between 14-37% of all vsRNAs captured. To a lesser extent, we also captured sRNAs that aligned with the collective endogenous viral element data set and Aedes aegypti vigra-like virus (AaVV) (Figure 3.4). Profiles of the ISVs from which sRNAs with at least 10 RPMs/library are derived are summarized in Table 3.1

Table 3.1. Insect-specific virus profiles from which most abundant virus small RNAs are derived across samples analyzed in this study. *Highest order of classification is provided.

virus	abbreviation	order; family; genera*	genome	geographic origin	references
Phasi Charoen-like virus	PCLV	Bunyavirales; *Phenuiviridae; Phasivirus*	(-) sense ssRNA	Recife, Brazil Chiapas state, Mexico New Orleans, USA Poza Rica, Mexico	Chandler et al., 2014 Zakrzewski et al., 2018 Zhang et al., 2018 Shi et al., 2019 Ramos-Nino et al., 2020 Cunha et al., 2020 Munivenkatappa et al., 2021 Olmo et al., 2021
verdadero virus		Durnavirales; *Partitiviridae*	dsRNA	Poza Rica, Mexico	Cross et al., 2020 Parry et al., 2021
chaq-like virus		Durnavirales*; Partitiviridae*	dsRNA	Poza Rica, Mexico	Cross et al., 2020 Parry et al., 2021
Aedes anphevirus	AeAV	Mononegavirales*; Xinmoviridae; Anphevirus*	(-) sense ssRNA	Poza Rica, Mexico New Orleans, USA	Parry & Asgari, 2018 Di Giallonardo et al., 2018 Thongsripong et al., 2021 Olmo et al., 2021
Ghana Aedes aegypti totivirus	GhAaTV	*Totiviridae*	dsRNA	New Orleans, USA	Amoa-Bosompem et al., 2020 Parry et al., 2021 Agboli et al., 2021
Guadeloupe Aedes aegypti totivirus	GuAaTV	*Totiviridae*	dsRNA	New Orleans, USA	Shi et al., 2019 Parry et al., 2021 Olmo et al., 2021
Aedes aegypti virga-like virus	AaVV	unclassified	(+) sense ssRNA	Recife, Brazil	Amoa-Bosompem et al., 2020 Agboli et al., 2021
Humaita-Tubiacanga virus	HTV	unclassified	(+) sense ssRNA	Chiapas state, Mexico	Aguiar et al., 2015 Parry et al., 2021 Olmo et al., 2021

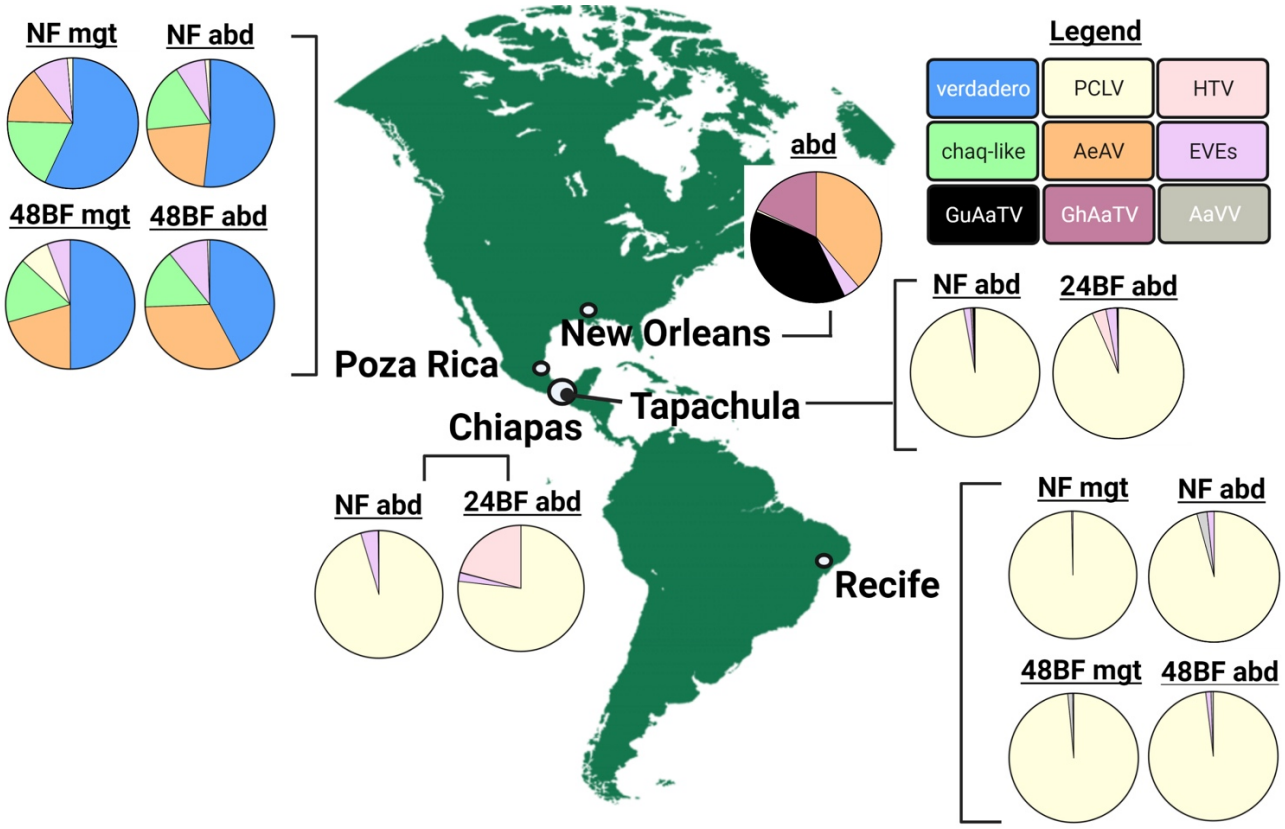


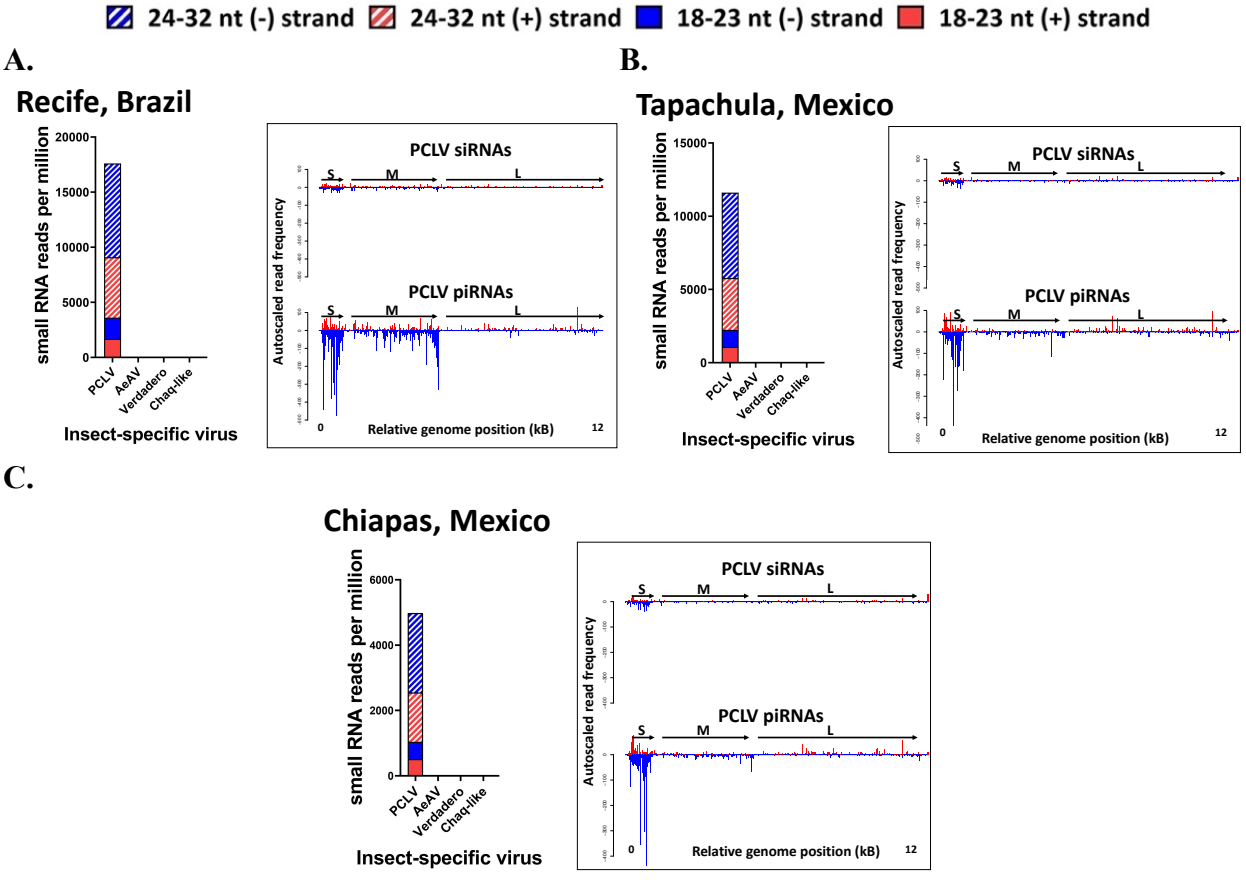
Figure 3.4. Virus-derived small RNA diversity in *Ae. aegypti* samples across the Americas. Percentage of virus-derived sRNA reads per million (RPM) relative to total virus-derived sRNA RPMs by sample type and geographic origin. Sample types: NF = non-fed; mgt = midgut; abd = abdomen; 24BF = 24 hours post-bloodfeed; 48BF = 48 hours post-bloodfeed. Viruses: verdadero = verdadero virus; chaq-like = chaq-like virus; PCLV = Phasi Charoen-like virus; AeAV = Aedes anphevirus; EVEs = collection of annotated endogenous viral elements in the *Ae. aegypti* genome; GuAaTV = Guadaloupe Aedes aegypti totivirus; GhAaTV = Ghana Aedes aegypti totivirus; HTV = Humaita Tubiacanga virus; AaVV = Aedes aegypti virga-like virus.

The mosquitoes largely harboring PCLV-specific vsRNAs had greater total RPM counts of PCLV-specific vsRNAs compared to the total number of verdadero, chaq-like, or AeAV-specific vsRNA RPMs in the mosquitoes largely lacking PCLV (Figure 3.5 bar graphs). The PCLV-specific vsRNAs were mostly vpiRNAs (striped bars) as opposed to vsiRNAs (solid color bars) (Figure 3.5 A-C bar graphs). Furthermore, vpiRNAs of PCLV origin, which is a negative sense RNA virus from the *Bunyavirales* order (Table 3.1), were nearly equal levels of the negative (blue) or positive (red) sense orientation (Figure 3.5 A-C), suggesting ping pong

amplification. Indeed, the PCLV-specific positive and negative-sense vpiRNAs mostly aligned to the S and M segments of the PCLV genome (Figure 3.5 A-C coverage plots), suggesting a NIRVS signature derived from these regions of the PCLV genome, consistent with other reports [127,176,272].

Mosquitoes that largely lacked PCLV sRNAs displayed lower, but more diverse, ISV-specific sRNA RPMs, which revealed patterns of active persistently infective viruses and unique NIRVS that differed across geographic origin. For mosquitoes from Poza Rica, verdadero sRNAs derived from the RNA-dependent RNA polymerase (RdRP) were the most abundant ISV-derived sRNA population present (Figure 3.5 D bar graph). Verdadero is a segmented dsRNA partitivirus, where single RNA segments are packaged individually in non-enveloped particles [269]. Verdadero-derived RdRP sRNAs were roughly equal proportions of vsiRNAs and vpiRNAs, where vsiRNAs were present in both the negative and positive sense orientations, while vpiRNAs tended to be positive sense (Figure 3.5 D bar graph). The corresponding coverage plot further revealed that verdadero-specific vsRNAs spanned the entire RdRP RNA segment, which would be expected from a veradero-specific NIRVS derived from the entire RdRP gene (Figure 3.5 D coverage plot). Because each partitivirus RNA segment is packaged individually [269], it may be that entire partitivirus RNA segments become endogenized into the mosquito genome. Taken together, these results suggest that verdadero virus actively infects the Poza Rica colony, and these mosquitoes may express a verdadero-specific NIRVS derived from the RdRP region. These observations were also true for the chaq-like RNA segment, where chaq-like virus sRNAs were mostly vsiRNAs compared to vpiRNAs, and chaq-like piRNAs were almost exclusively of the positive sense orientation (Figure 3.5 D). Chaq-like virus sequences are found together with parent partitiviruses (in this case, verdadero virus), and it is not fully

understood whether they are an "optional" partitivirus RNA segment or a satellite virus [269]. Like verdadero virus, chaq-like virus is a single component RNA virus, so it may not be surprising that an entire chaq-like NIRVS is present in the mosquito genome. In contrast to verdadero and chaq-like RNA viruses, on the other hand, Aedes anphevirus sRNAs in both Poza Rica and New Orleans mosquitoes were mostly positive sense vpiRNAs that aligned toward the 5' end of the genome (Figure 3.5 D-E). AeAV is a negative sense RNA virus from the *Mononegavirales* order [259]. Given that most sRNAs are in the opposite orientation of the AeAV genome and align to a specific region of the genome, these results suggest the mosquitoes harbor an AeAV-specific NIRVS as well.



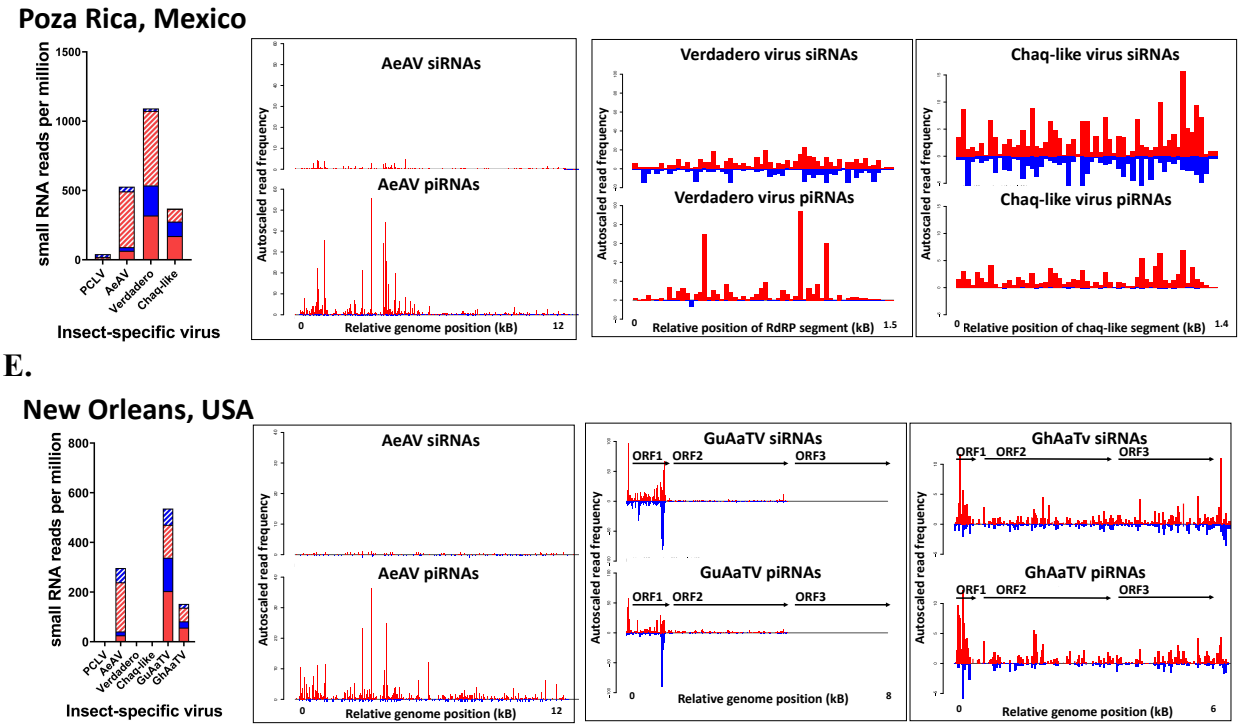


Figure 3.5. Mosquitoes from distinct geographic origins harbor persistently infective insect-specific viruses and display unique NIRVS signatures. Bar graphs: Total number of negative (blue) or positive (red) sense vsiRNAs (solid colors) or vpiRNAs (checkered colors) in reads per million (RPM) against the insect-specific viruses Phasi Charoen-like virus (PCLV), Aedes anphevirus (AeAV), verdadero virus, chaq-like virus, Guadeloupe Aedes aegypti totivirus (GuAaTV) or Ghana Aedes aegypti totivirus (GhAaTV) in combined mosquito samples from (**A)** Recife, Brazil, (**B**) Tapachula, Mexico, (**C**) Chiapas, Mexico, (**D**) Poza Rica, Mexico, or (**E)** New Orleans, USA. **Coverage plots:** Corresponding representative coverage plots of (**A-C**) PCLV-specific (**D-E**) AeAV-specific (**D**) verdadero- and chaq-like-specific or (**E)** GuAaTV- and GhAaTV-specific 18-23 nucleotide (nt) vsiRNAs (top) or 24-32 nt vpiRNAs (bottom). The X-axis shows the coordinates of the virus genome in kilobases (kB) while the Y-axis is autoscaled read frequency. The small (S), medium (M), and large (L) segments of the PCLV genome are marked as are the open reading frames (ORF) of GuAaTV and GhAaTV. The verdadero genome region is derived from the RNA-dependent RNA polymerase RNA segment (NCBI Accession: MT742175) while the chaq-like RNA segment is NCBI Accession: MT742176.

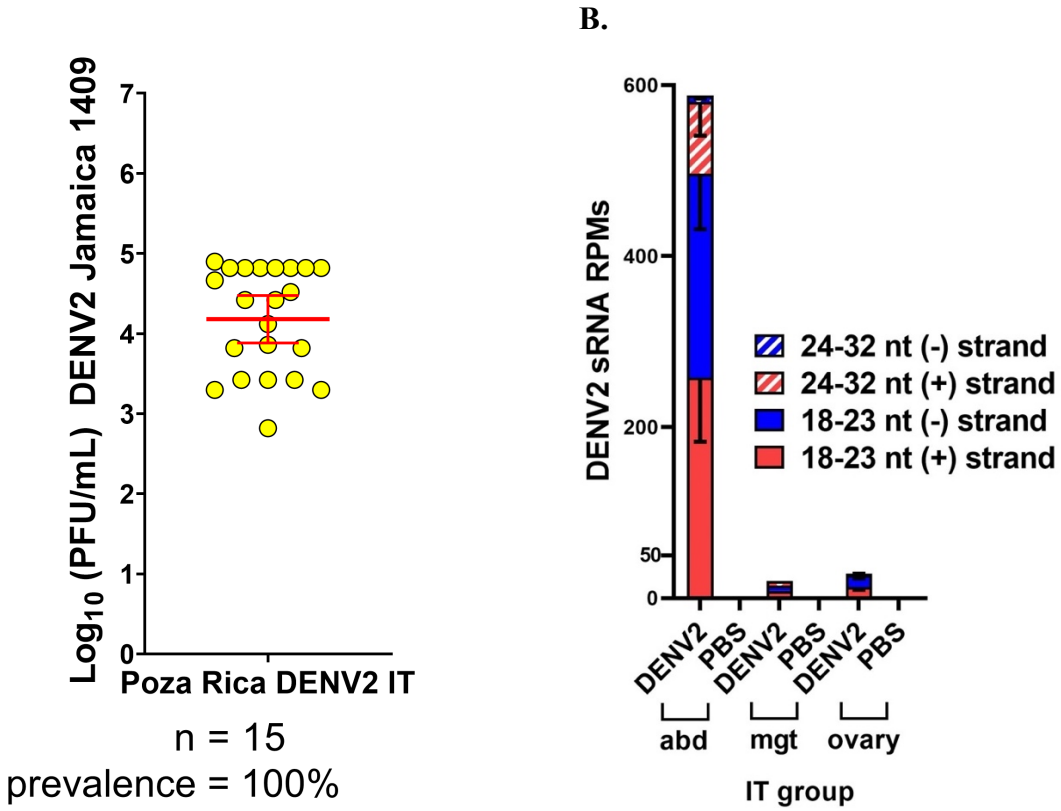
The Guadeloupe and Ghana Aedes aegypti totiviruses, GuAaTV, and GhAaTV respectively, unique to the New Orleans mosquitoes, also revealed interesting patterns of ISV persistent infection and ISV-specific NIRVS. GuAaTV-specific sRNA RPMs were most abundant in the dataset and presented in equal proportion of vsiRNAs and vpiRNAs (Figure 3.5 E bar graph). The corresponding GuAaTV coverage plot revealed that almost all vsRNAs

aligned to ORF1 (the capsid) of the GuAaTV genome in roughly equal proportions of the positive and negative sense orientation, indicative of ping-pong amplification (Figure 3.5 E). Together, these results suggest the mosquitoes from New Orleans harbor a GuAaTV-specific NIRVS and are actively infected with this virus. On the other hand, GhAaTV-specific vsRNAs tended to span the entire virus genome, but both vsiRNAs and vpiRNAs were present in both orientations (Figure 3.5 coverage plot). These mosquitoes are therefore likely actively infected with GhAaTV. It is also possible, given the GhAaTV vpiRNA profile, that these mosquitoes harbor a GhAaTV-specific NIRVS that spans the entire genome.

In summary, we used sRNA profiles to discover actively infecting ISVs across mosquito species and to show evidence of vpiRNA ping pong amplification from novel NIRVS. We highlighted NIRVS signatures against PCLV, AeAV, GuAaTV, and possibly verdadero or chaq-like viruses, dependent on mosquito geographic origin. Those mosquitoes harboring putative PCLV (from Chiapas, Mexico and Recife, Brazil) or GuAaTV (from New Orleans, USA) NIRVS appear to be actively infected with these viruses because of (1) the presence of PCLV- or GuAaTV-specific vsiRNAs that span the viral genomes in both the positive and negative orientations and (2) PCLV- or GuAaTV-specific vpiRNAs in the negative and positive sense orientations that align to specific regions of the viral genome, indicative of ping-pong amplification. Mosquitoes from New Orleans were also actively infected with GhAaTV but did not appear to express a GhAaTV-specific NIRVS. Mosquitoes from Poza Rica appear to be actively infected with verdadero and chaq-like viruses and may possess NIRVS derived from the entire verdadero and chaq-like RNA segments. Finally, mosquitoes from both Poza Rica and New Orleans appear to harbor an AeAV-specific NIRVS, but the extent to which they were infected with AeAV remains unclear.

3.2.3 DENV2 infection results in DENV2-specific vsiRNAs and, to a lesser extent, vpiRNAs

To gain insights on how ISV sRNAs may cross talk with sRNAs from arboviruses, we intrathoracically inoculated mosquitoes from Poza Rica, Mexico (that harbor verdadero, chaq-like, and AeAV sRNAs, Figure 3.4, 3.5, Table 3.1) with 500 PFU of DENV2 and repeated sRNA sequencing 7 dpi on midguts, abdomens, and ovaries. As a negative control, we also intrathoracically inoculated a group of mosquitoes with PBS. We chose to inject the mosquitoes with DENV2 (strain: Jamaica 1409) so that they would all become infected with a constant dose of virus [249]. At 7 dpi, we confirmed that 100% of intrathoracically inoculated mosquitoes were infected by plaque assay and found that the average DENV2 titer was 10^4 PFU / ml in whole bodies (Figure 3.6 A). Small RNA sequencing revealed, as expected, mosquitoes that were injected with DENV2 displayed DENV2-specific vsiRNAs and vpiRNAs not seen in the group injected with PBS (Figure 3.6 B). IT treatment bypasses the midgut, and we observed a much greater abundance of DENV2-specific sRNAs in the abdomen than in the midgut, likely because this is where virus was introduced into the mosquito. We also detected DENV2 sRNAs in the ovaries, albeit at low levels, most likely because DENV2 does not efficiently infect the ovaries (~1-4% of DENV2-infected mosquitoes display DENV2 vertical transmission) [91]. DENV2 siRNAs spanned the entire viral genome (Figure 3.6 C), which is expected during an active infection. DENV2-specific vpiRNAs were substantially less abundant than DENV2 siRNAs, and they were mostly positive sense, most likely because they were directly processed from the DENV2 positive sense RNA genome (Figure 3.6 B). We did observe a notable peak of DENV2-specific piRNAs that aligned to the 3'UTR of the virus for reasons unknown (Figure 3.6 C). However, we observed no evidence of ping pong amplification, and these mosquitoes therefore likely lack a DENV2-specific NIRVS.



C.

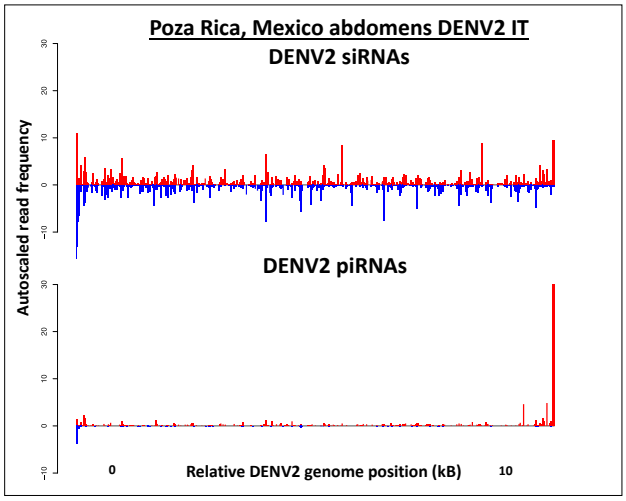


Figure 3.6. Intrathoracic inoculation of DENV2 results in DENV2-derived sRNAs in all tissues. A. DENV2 plaque-forming units (PFU) / mL in whole Poza Rica, Mexico mosquitoes 7 days post intrathoracic inoculation (IT) with DENV2. Red bars indicate mean and 95% confidence interval. **B.** DENV2-specific siRNAs (solid colors) and piRNAs (checkered colors) by negative (blue) or positive (red) strand in abdomen (abd), midgut (mgt), or ovary tissues from

DENV2-injected mosquitoes compared to those intrathoracically inoculated with PBS as a control. Y axis shows DENV2-specific small RNA (sRNA) reads per million (RPM) for each IT group (by tissue type). Bars indicate average RPM with standard errors of the mean (SEM). **C.** Representative coverage plot of DENV2-specific 18-23 nucleotide (nt) vsiRNAs (top) or 24-32 nt vpiRNAs (bottom) in DENV2 IT abdomens. The X-axis shows the coordinates of the virus genome in kilobases (kB) while the Y axis represents the autoscaled read frequency.

When we compared overall sRNA distributions in DENV2- and PBS-injected mosquitoes, we found that the majority of sRNAs in midguts and abdomens were 18-23 nt in length and that 24-32 nt piRNAs had diminished compared to the non-injected Poza Rica samples in Figure 3.3 E-H. Because this was true regardless of whether the mosquitoes received PBS or DENV2, it may be that the siRNA and/or miRNA pathways act as an injury response to intrathoracic inoculation. Abdomens from DENV2-injected mosquitoes displayed slightly elevated levels of vsiRNAs (50% of all vsRNA reads) compared to PBS controls (35% of all vsRNA reads) (Supplemental Figure 3.3), but vpiRNA proportions were not affected. vsiRNAs levels in midguts or ovaries were similar in DENV2- or PBS-injected mosquitoes (Supplemental Figure 3.3), most likely because of the low levels of DENV2 in these tissues 7 days post DENV2 IT.

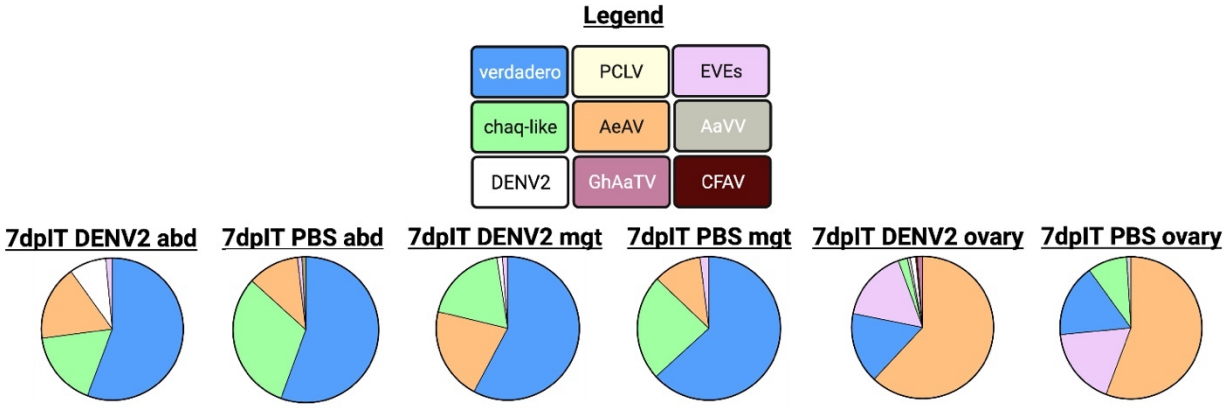
3.2.4 DENV2 infection is associated with increased abundances of small RNAs against insect-specific viruses

We next analyzed how DENV2 infection impacted vsRNA abundances against native ISV-derived sRNAs observed in the same population of uninfected mosquitoes from Poza Rica. We once again found that the most prevalent vsRNAs were of verdadero, chaq-like, and AeAV origins, regardless of injection (IT) group (Figure 3.7 A). In fact, DENV2-infected mosquitoes harbored greater proportions of ISV-derived sRNAs compared to DENV2-specific sRNAs (Figure 3.7 A). vsRNA diversity was similar in these samples compared to their uninfected

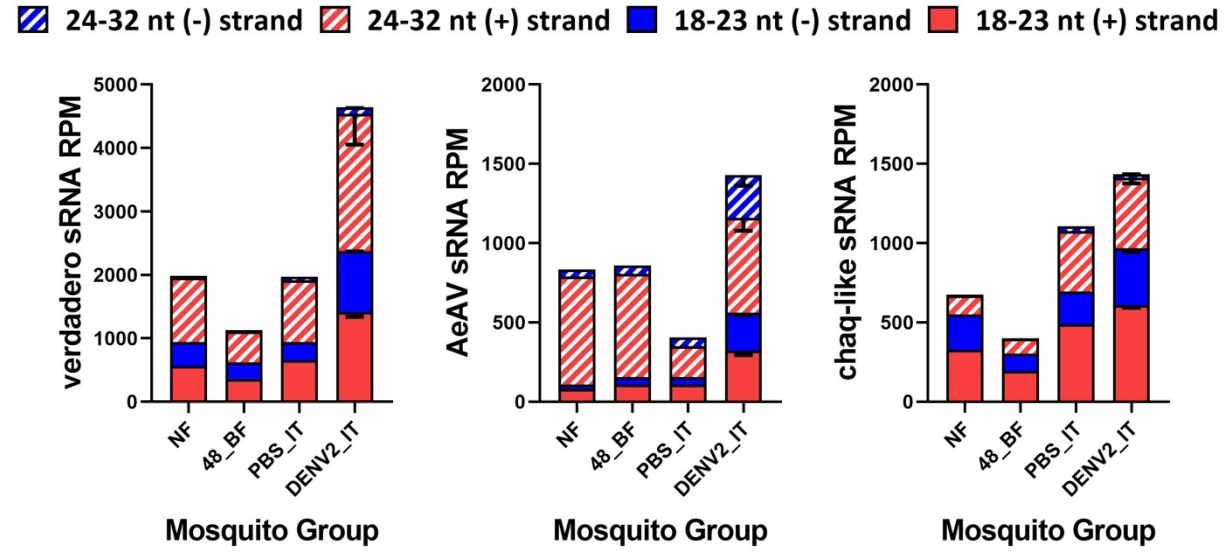
counterparts (Figure 3.4). Interestingly, ovaries, not previously analyzed, displayed a higher proportion of AeAV sRNAs compared to verdadero or chaq-like sRNAs (Figure 3.7 A) or compared to the proportion of AeAV sRNAs seen in somatic tissues (Figure 3.4). They also harbored more EVE-derived sRNAs than somatic tissues (Figure 3.7). For reference, a table of annotated EVEs from the *Ae. aegypti* genome, found in Palatini et al., 2017 [174], is provided in Supplemental Table 3.1.

We next focused our analyses on abdomens since this was the tissue that exhibited the greatest abundance of DENV2-specific sRNAs 7 days post-IT. DENV2-infected abdomens displayed heightened ISV-derived sRNA RPMs compared to their DENV2 uninfected counterparts (Figure 3.7 B). This observation was most apparent for verdadero- and AeAV-derived sRNAs, although the trend was consistent for chaq-like-derived sRNAs as well. Corresponding coverage plots revealed that the ISV vsiRNAs spanned the entirety of the genomic regions surveyed (Figure 3.7 C, top plots). Small RNA coverage plots of verdadero and chaq-like viruses appeared similar in DENV2-infected and uninfected mosquitoes, but AeAV coverage plots appeared different by infection status (Figure 3.7 C). Specifically, DENV2-infected mosquitoes displayed a strong peak of AeAV-specific positive sense piRNAs toward the center of the AeAV genomic RNA as well as negative sense AeAV-specific piRNA peaks surrounding this positive sense peak. The AeAV coverage plot in PBS-treated mosquitoes looked similar to those from mosquitoes maintained on sugar or 48 hours post-non infectious bloodmeal (Figure 3.5 D). These mosquitoes displayed positive sense AeAV-derived vpiRNAs that tended to align more toward the 5’ end of the AeAV genome, and they mostly lacked negative sense AeAV-specific vpiRNAs. Given that vpiRNAs differed by infection status, these results may suggest cross talk of AeAV and DENV2 sRNAs.

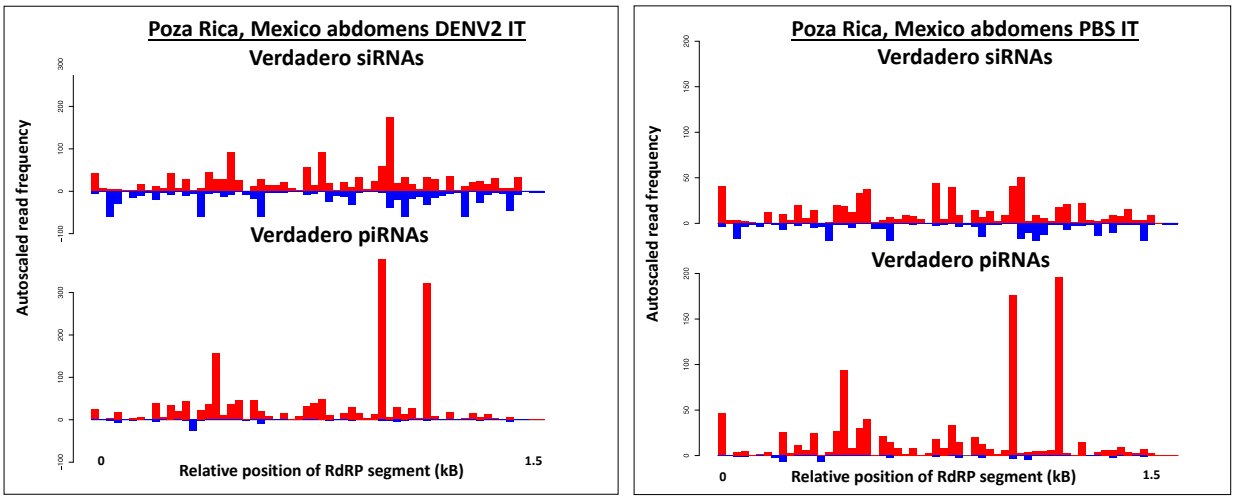
A.


Legend
verdadero
PCLV
EVEs
chaq-like
AeAV
AaVV
DENV2
GhAaTV
CFAV
7dpIT DENV2 abd
7dpIT PBS abd
7dpIT DENV2 mgt
7dpIT PBS mgt
7dpIT DENV2 ovary
7dpIT PBS ovary

B.


24-32 nt (-) strand
24-32 nt (+) strand
18-23 nt (-) strand
18-23 nt (+) strand
verdadero sRNA RPM
AeAV sRNA RPM
chaq-like sRNA RPM
NF
48_BF
PBS_IT
DENV2_IT
Mosquito Group

C.


Poza Rica, Mexico abdomens DENV2 IT
Verdadero siRNAs
Verdadero piRNAs
Autoscaled read frequency
Relative position of RdRP segment (kB)
Poza Rica, Mexico abdomens PBS IT
Verdadero siRNAs
Verdadero piRNAs
Autoscaled read frequency
Relative position of RdRP segment (kB)

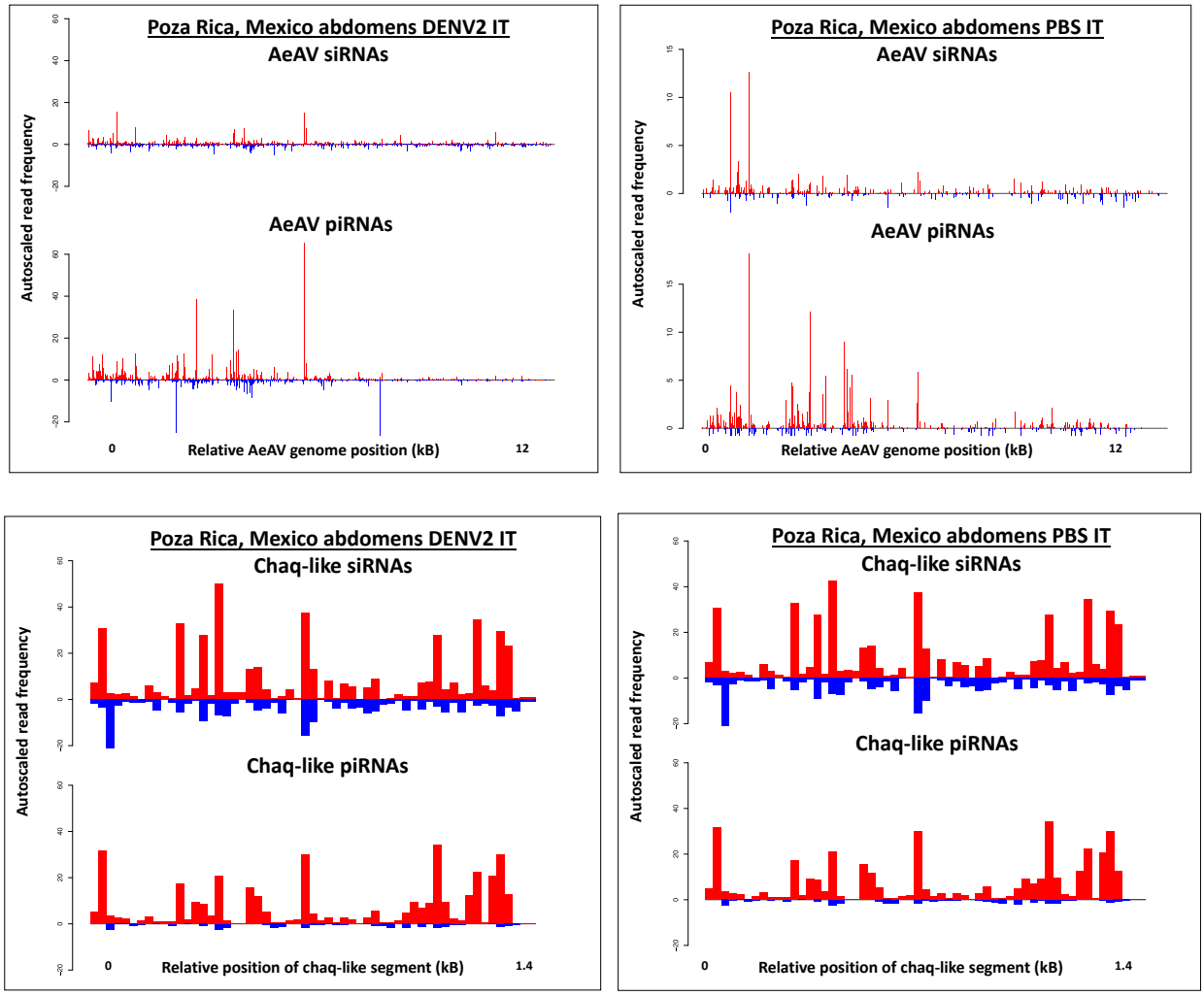


Figure 3.7. Virus-specific sRNAs in DENV2-infected and non-infected *Ae. aegypti* from Poza Rica, Mexico. A. Virus-derived sRNA diversity by tissue type and DENV2 infection status in midguts (mgt), abdomens (abd), and ovaries of Poza Rica, Mexico *Ae. aegypti* after intrathoracic inoculation (IT). verdadero = verdadero virus; chaq-like = chaq-like virus; PCLV = Phasi Charoen-like virus; AeAV = Aedes anphevirus; EVEs = collection of annotated endogenous viral elements in the *Ae. aegypti* genome; DENV2 = dengue-2; GhAaTV = Ghana Aedes aegypti totivirus; CFAV = cell fusing agent virus; AaVV = Aedes aegypti virga-like virus. **B.** Verdadero virus, chaq-like virus, and AeAV specific siRNAs (solid colors) or piRNAs (checkered colors) in abdomens of Poza Rica, Mexico *Ae. aegypti* that were maintained on sugar (NF = non-fed), 48 hours post-non-infectious bloodmeal (48_BF), intrathoracically inoculated (IT) with PBS, or IT with DENV2. **C.** Representative coverage plots of verdadero, chaq-like viruses, and AeAV-specific 18-23 nucleotide (nt) vsiRNAs (top) or 24-32 nt vpiRNAs (bottom). The X-axis shows the coordinates of the virus genome in kilobases (kB) while the Y-axis is autoscaled read frequency. The verdadero genome region is derived from the RNA-dependent RNA polymerase-encoding RNA (NCBI Accession: MT742175) while the chaq-like RNA segment is available at NCBI Accession: MT742176.

Aedes anphevirus was first identified in an embryonic *Ae. aegypti* cell line transfected with *Wolbachia* (*w*MelPop-CLA strain), and *Wolbachia* was found to enhance AeAV replication in cell culture [259]. It remains elusive whether *Ae. aegypti* are natural hosts for certain strains of *Wolbachia* [278], although they are generally considered not to be [270]. However, our Poza Rica line exhibited low RPM levels of negative sense *Wolbachia*-specific sRNAs derived from a strain of *Wolbachia pipientis* (wAlbB HN2016 origin) originally found in wild caught *Aedes albopictus* mosquitoes [279] (Supplemental Figure 3.4). Given that these RPM levels were low and were mostly in a single (negative sense) orientation, it is unclear whether *Wolbachia pipientis* HN2016 is actively infecting these mosquitoes. Nonetheless we were curious whether heightened AeAV-specific vsRNAs in the DENV2-infected mosquitoes were also associated with changes in *Wolbachia* sRNAs. We observed a pronounced spike in *Wolbachia pipientis*-specific sRNAs in the DENV2 IT abdomen group compared to the uninfected groups (Supplemental Figure 3.4). The sRNAs were mostly negative sense siRNAs and piRNAs. These results may suggest that DENV2 infection enhances or triggers *Wolbachia* replication/sRNA production in mosquitoes as well, and that acute infection with DENV2 is associated with complex interactions with not just the virome, but the microbiome as well.

3.3 Discussion

In this study, we compared virus-derived sRNA populations in geographically distinct *Ae. aegypti* populations to gain insights on how insect-specific viruses, endogenous sRNAs, and arbovirus sRNAs interact and whether such interactions would impact mosquito vector competence. We found that total sRNA distributions were highly variable across mosquito samples, and that ISV sRNAs displayed a geographic structure. Mosquitoes from the state of

Chiapas, Mexico or Recife, Brazil mostly harbored PCLV sRNAs and displayed less sRNA diversity than mosquitoes from New Orleans, USA or Poza Rica, Mexico that mostly lacked PCLV. On the other hand, mosquitoes from Poza Rica mostly harbored verdadero virus, chaq-like virus, and AeAV sRNAs, whereas those from New Orleans also harbored AeAV but exhibited Guadualope and Ghana totivirus sRNAs as well. After infecting mosquitoes from Poza Rica with DENV2, we observed an increase in ISV vsRNA RPM abundances as well as the generation of unique AeAV-specific vpiRNAs not seen in DENV2-noninfected counterparts.

The overall variability in sRNA distributions across samples is likely attributed to differences in persistently infecting viruses and the mosquito virome but may also be caused by other sources such as endogenous sRNA loci that are not yet characterized. We emphasized herein that total sRNA distributions were highly variable across mosquitoes (Figure 3.2, Supplemental Figure 3.2), yet most sRNAs in our samples were not of miRNA, TE, or structural RNA origin. Of course, we are limited to hitherto known sequences deposited into publicly available databases, and this observation may be due to incomplete miRNA, TE, or virus annotations present in *Ae. aegypti*. Interestingly, siRNA clusters in the abdomen tended to overlap with peaks in miRNA abundances, whereas siRNA clusters in the midguts did not overlap with known miRNAs, TEs, structural RNA, or viral RNA profiles (Figure 3.3 A-H; Supplemental Figure 3.3 A-D). Because the midgut is the organ into which a bloodmeal is first ingested, it represents the initial tissue to become exposed to pathogens present in the blood. Therefore, the midgut may harbor more anti-pathogen siRNAs that are not yet known compared to the abdomen. In *Drosophila*, sRNAs derived from endogenous siRNA and piRNA loci are involved in mRNA or TE silencing [280–283], but little is known of endogenous siRNA loci in *Ae. aegypti*. Unexpectedly, 22 nt siRNAs that did not map to *Ae. aegypti* structural RNAs were

most abundant in all samples, particularly in midguts, compared to 18-21 nt siRNAs; however, 21 nt vsiRNAs were overwhelmingly the most abundant vsiRNA in all samples compared to those that were 18-20 or 22 nt in length. This discrepancy suggests most siRNAs of unknown origin are not necessarily of viral origin. Furthermore, piRNAs were also mostly from unknown origins. A recent extensive study found that piRNA-containing loci in *Ae. aegypti* appear to flank TE-rich regions and be mostly intergenic [272], which implies many piRNAs are not of TE origin, as we saw here. Further research is warranted investigating endogenous sRNA loci in *Ae. aegypti*, including where in the genome they reside, their functions, and their origin.

Although sRNA distributions were variable, we consistently observed more siRNAs in midguts than abdomens from the same mosquito type independent of bloodfeeding status. This result suggests that the siRNA machinery is more active in the midgut compared to the abdomen, at least at the time points tested. Perhaps the proteins involved in the siRNA pathway are more expressed in this tissue, or dsRNA substrates processed into siRNAs are of greater prevalence here compared to other tissues. However, the act of bloodfeeding itself does not appear to significantly alter sRNA distributions in the midgut. It should be noted here that our artificial feeding system relies on sterile animal blood and a membrane as a model for non-sterile blood and a skin barrier, which undoubtably would also impact sRNAs and immunity. Nonetheless, given that the midgut appears to be more effective at processing siRNAs, transgenic mosquitoes that synthetically trigger the siRNA pathways against arboviruses (as in Chapter 2) will likely be more successful if the transgene is expressed in this tissue as opposed to outside this tissue. Therefore, using a midgut-specific promoter to express synthetic dsRNAs, such as *CpA*, may be more effective at triggering the endogenous siRNA pathway than promoters aimed at expression in other tissues.

We show herein that all mosquitoes in our sample set appear to be persistently infected with ISVs that are processed into significant abundances of sRNAs, particularly vpiRNAs. Except for mosquitoes from New Orleans, all samples used in this study had been maintained in colony in the same conditions for at least three generations. Therefore, it appears ISVs acquired in the field can be vertically transmitted over long periods of time. This is consistent with a recent report that found that five of the most abundant viral taxa sequenced from field and colony reared mosquitoes from Florida were the same [182]. In that study, the authors maintained mosquitoes from the field for seven generations and found that a totivirus similar to GuAaTV was vertically transmitted [182]. In our study, the ISV-derived vsRNA profiles of the adult field mosquitoes from New Orleans did have similarities with colony mosquitoes originating from Poza Rica – most notably, both populations harbored AeAV sRNAs – but the mosquitoes from New Orleans also harbored totivirus sRNAs not seen in our other samples. Colony rearing may select for certain ISV populations in the absence of fitness pressures of the field, but from our study, we cannot say whether mosquitoes maintained in colony harbored significantly different ISVs compared to mosquitoes in the field.

We found that mosquitoes with significant PCLV vsRNA loads displayed less diverse sRNA viromes compared to mosquitoes that mostly lacked PCLV. PCLV comprised 77-99% of all virus-derived sRNAs when it was the main ISV present, although we did detect vsRNAs against other ISVs at low abundances in every sample. PCLV is a phlebovirus of the *Phenuiviridae* family that is globally distributed in *Ae. aegypti* cells and mosquitoes from Thailand [263], Australia [270], China [264], the West Indies [265], and India [284]. A recent preprint reported that PCLV was found mosquitoes throughout Brazil, Singapore, France, and some, but not all, countries they surveyed in Africa [285]. Similar to our findings – we detected

more PCLV-derived sRNAs in RPM compared to any other vsRNA – the authors of that study found PCLV to be one of the most abundant viruses in their sample set as well [285]. However, they found that PCLV more often coinfected mosquitoes with HTV as opposed to infection with either virus alone [285]. This contrasts with our findings, where HTV was not very prevalent in our samples and was only found in significant abundance in the *Ae. aegypti* GDLS. The mosquitoes analyzed by Olmo et al., 2021 were all directly from the field, so it is possible that PCLV can outcompete other ISVs such as HTV in colony. However, we did find mosquitoes that mostly, but not completely, lacked PCLV, including those from the Poza Rica colony and from the field in New Orleans. Notably, the Poza Rica and Chiapas *Ae. aegypti* colonies have been maintained in the same insectary for many generations, so it does not seem that PCLV has cross-contaminated mosquito lines. Similar to our findings, Olmo and colleagues also found mosquitoes that mostly, but not completely, lacked PCLV; they were from Dakar and Senegal [285]. Additionally, the authors identified a mosquito population from Gabon, Africa that completely lacked PCLV [285], whereas all mosquitoes analyzed herein harbored at least a few PCLV sRNAs.

The ISV geographic structure observed herein correlates with the reported geographic structure of *Ae. aegypti* populations in Mexico, which was associated with differences in vector competence for DENV [273]. In that study, Lozano-Fuentes and colleagues found distinct mitochondrial haplotypes north or south of a NVA that appears to act as a barrier against gene flow [273]. This population structure was correlated with DENV vector competence, where mosquitoes north of the NVA were significantly more competent to DENV2 than those south of the NVA [273]. Although it is unclear how the NVA limits mosquito gene flow, the authors noted that enhanced human movement and maritime and cruise ship activity on the west Pacific

coast of Mexico could traffic *Ae. aegypti* more so than on the coast of the Gulf of Mexico, which could be the principal barrier [273]. In our study, mosquitoes from Poza Rica, Veracruz, Mexico (northeast of the NVA) harbored substantially different ISVs than those in Tapachula, Chiapas, Mexico (southwest of the NVA). Perhaps these barriers also impact ISV circulation among *Ae. aegypti* populations. It remains unclear how mosquito population structure and genetics impact ISVs, but it would be interesting to correlate mosquito haplotypes with ISV profiles. Future investigations comparing vector competence between mosquito populations harboring different ISV profiles are also underway.

How ISVs interact with arboviruses remains elusive, particularly *in vivo.* Cell culture studies have shown that both PCLV [258] and AeAV [259] impact DENV2 replication, so virome diversity in field mosquitoes likely has significant consequences on arbovirus transmission potentials. It is possible that there is competition between ISVs and active replication of certain viruses would suppress the replication of other ISVs. Or perhaps antiviral immunity is more targeted and/or effective against some viruses than others, especially during complex infections. In the recent preprint referenced previously, Olmo and colleagues found that there were several cellular processes downregulated during DENV infection and upregulated by the presence of HTV and PCLV, the majority of which involved genes related to histones [285]. Furthermore, histone H4 was significantly downregulated in mosquitoes infected with ZIKV that lacked PCLV and HTV, yet levels of histone H4 were significantly higher in the presence of HTV and PCLV compared to mosquitoes lacking these viruses [285]. The authors concluded that PCLV and HTV prevented downregulation of histone H4 that was induced by DENV and ZIKV infection [285]. In our study, however, mosquitoes that mostly lacked PCLV and HTV from Poza Rica have been previously shown to be highly competent for DENV [273] (Figure 3.1), so

other ISVs and/or mechanisms are likely also at play. Further investigations are warranted investigating the molecular mechanisms underlying ISV-arbovirus interactions in the mosquitoes themselves. We, and others, have shown that *Ae. aegypti* are infected by a plethora of ISVs, so future studies analyzing ISVs should consider the complexity of the system, specifically how ISVs interact with each other and interact with arboviruses in the presence of many ISVs at once.

We found that sRNA analyses were good resource to detect not only actively infecting viruses, but also NIRVS that may otherwise be difficult to sequence. For example, our sRNA data suggest mosquitoes harbor PCLV, AeAV, and GuAaTV NIRVS, and possibly verdadero and chaq-like NIRVS, depending on geographic origin. Mosquitoes harboring PCLV-derived sRNAs displayed evidence of ping-pong amplification because vpiRNAs aligning exclusively to the small (S) or medium (M) segments of the PCLV genome were present in both sense and antisense orientations. These mosquitoes most likely (1) express two PCLV NIRVS derived from these segments, (2) are actively infected with PCLV, and (3) display effective vpiRNA-mediated silencing. These observations were largely similar for Guadalupe Aedes totivirus from New Orleans field mosquitoes that appeared to exhibit GuAaTV ping pong signatures against ORF1 (the capsid). In the case of Aedes anphevirus, however, AeAV-specific vpiRNAs aligning to the 5' end of the genome were mostly present in the sense orientation. Because AeAV has a negative sense RNA genome, our results suggest that the mosquitoes express an AeAV NIRVS, but either they are not actively infected with AeAV (or that AeAV is replicating at low levels) or that vpiRNA-mediated silencing is not effectively occurring.

Overall, we found that vsRNAs were diverse across *Ae. aegypti*. In addition to the most abundant ISV sRNAs highlighted in Figure 3.5 (and when considering sRNAs with $\geq$ 10 RPM in replicates), we detected vsRNAs of HTV, AaVV, and CFAV origins. All vsRNAs were from

ISVs that had been previously identified in field mosquitoes. HTV is an unclassified positive sense ssRNA virus that has been identified in *Ae. aegypti* from Brazil [268,285], Australia, and Thailand [270]. AaVV was first identified in *Ae. aegypti* from the greater Accra region of Ghana that had been maintained in colony for two generations [286]. It is closely related to the *Virgaviridae* family of positive sense RNA plant viruses and was not effectively vertically transmitted in colony after three generations [286]. Interestingly, however, we detected AaVV sRNAs in the Recife colony that had been maintained in colony for at least six generations, albeit at low RPM abundances. CFAV is a well-known flavivirus that replicates in C6/36 [160] and Aag2 [257] cells and was isolated from Puerto Rican *Ae. aegypti* populations in 2006 [262]. It has subsequently been identified in Thai [287], American [288], Australian [270], Mexican [289], and Indonesian [290] *Ae. aegypti*. CFAV-specific NIRVS have been identified in Aag2 cells [177] as well as in the *Ae. aegypti* mosquito genome assembly [174] (Supplemental Table 3.1). We only detected CFAV sRNAs in ovaries, which again highlights that tissues are not equally infected with ISVs and that they do not process sRNAs in equal manners.

We found that DENV2-infected Poza Rica *Ae. aegypti* expressed DENV2-specific siRNAs (Figure 3.6 B, C), which led to a modest increase of total virus-derived siRNAs (Supplemental Figure 3.3). It is well established that DENV2 infection triggers *Ae. aegypti* siRNA immunity [116] and leads to the production of DENV2-specific piRNAs [123,128], so these results were expected. Interestingly, DENV2-infected mosquitoes expressed > five-fold increased ISV-derived siRNAs compared to DENV2-specific siRNAs (Figure 3.6 B, Figure 3.7 B). Considering these mosquitoes harbored, on average, 10^4 PFU/ml of DENV2, it is surprising that there were relatively few DENV2-specific vsRNAs. It could be that the mosquitoes were infected with much greater loads of ISVs compared to DENV2 or that ISVs trigger RNAi

antiviral immunity to a greater extent than does DENV2. We also observed that all mosquitoes that underwent intrathoracic inoculation (with either PBS or DENV2) exhibited a spike in total siRNAs and a decrease in total piRNAs. It is possible that the act of injection causes an injury response in the mosquitoes, which triggers innate immune responses. We are in the process of infecting mosquitoes from Poza Rica, Mexico and Recife, Brazil with DENV2 and ZIKV through a more natural route of infection – bloodfeeding – to see how arbovirus-specific sRNAs behave in mosquitoes that have different ISV profiles.

The increase in ISV sRNAs after DENV2 infection has never been reported in mosquitoes and warrants further investigation. The increase in siRNAs specifically (Figure 3.7 B) suggests arbovirus infection promotes ISV replication. This did not appear to be limited to a single ISV because we observed increased vsRNAs for the three most prevalent ISVs in the mosquitoes we sequenced. Perhaps arbovirus infection compromises RNAi immunity, which allows other persistently infective viruses to replicate more. We also noted that AeAV-specific vpiRNAs in the DENV2-infected mosquitoes were derived from different regions of the virus genome than those AeAV vpiRNAs from uninfected mosquitoes (Figure 3.7), suggesting cross talk of ISV and arbovirus vsRNAs. A similar phenomenon has been reported for ZIKV-infected *Ae. aegypti* that expressed unique CFAV vpiRNAs not seen in ZIKV-uninfected mosquitoes [272]. AeAV was first discovered in *Wolbachia*-transfected *Ae. aegypti* cells and was shown to reduce DENV2 replication [259]. Cells with *Wolbachia* supported enhanced AeAV replication compared to those lacking *Wolbachia*, suggesting the endosymbiont enhanced AeAV replication [259]. In line with that study, we also found elevated *Wolbachia pipientis* sRNA loads in the DENV2-infected mosquitoes harboring higher counts of AeAV sRNAs, perhaps suggesting increased replication of both AeAV and *Wolbachia*. However, it should be noted that

Wolbachia-specific sRNAs were low and exhibited high variability in replicates. Because *Ae. aegypti* is not generally considered a natural host for *Wolbachia,* future studies confirming active *Wolbachia* infection in this strain of mosquitoes is warranted. Nonetheless, taken together, our results suggest that sRNAs against ISVs, arboviruses, and perhaps even other endosymbionts likely have complex interactions that impact each other's replication dynamics.

3.4 Methods

3.4.1 Mosquito sample collection and rearing

New Orleans, LA collection

Aedes aegypti from New Orleans, LA, USA were collected through a collaboration with Mr. Brendan Cater and Dr. Dawn Wesson of the Tulane University Department of Tropical Medicine. Adult mosquitoes were collected via a handheld aspirator from vegetation and BG Sentinel 2 traps (Biogents, Regensburg, Germany) with scent lure in New Orleans, Louisiana in September and October of 2020. After collection, mosquitoes were stored in a -20°C freezer and were identified to species and processed the same day as collection. Midguts and ovaries of the mosquitoes were removed using forceps and scalpel under a microscope (Leica EZ4HD, Leica Microsystems, Wetzlar, Germany). Midguts and mosquito bodies without heads were placed into separate tubes of 500 µL of TRIzol (Thermo Fisher Scientific, Waltham, MA) respectively and were stored at -80°C until further processing.

GDLS collection

A genetically diverse laboratory strain (GDLS) was established through a series of crosses using *Ae. aegypti* derived from 10 geographically distinct populations in Chiapas,

Mexico as described in de Valdez et al., 2011 [274] (Supplemental Figure 3.1). This line has been maintained at CSU in colony since its re-establishment in 2015.

Recife, Brazil collection

Aedes aegypti from Brazil were collected through a collaboration between Dr. Tereza Magalhaes, by then a Research Scientist at Colorado State University, and Dr. Danilo de Carvalho Leandro, an Associate Professor at *Colégio de Aplicação* of the *Universidade Federal de Pernambuco*. In February 2018, ovitraps containing wood paddles and water with *Bacillus thuringiensis israelensis* (Bti serves as attractant for gravid females and kills larvae if they emerge in the traps) were installed in peri-domestic environments for one week in four neighborhoods of Recife, Pernambuco State, Brazil. The wood paddles containing *Ae. aegypti* eggs were then collected, allowed to dry, and shipped to Colorado, USA. On March 26th, 2018, ten wood paddles representative from the four neighborhoods were immersed in water in the insectary at CSU. After a few hours, some of the eggs had hatched and the L1 larvae were immediately transferred to another water container to avoid Bti ingestion by the same (the wood paddles were embedded with Bti). After a few days, the pupae were transferred to a mosquito cage for adult emergence. This consisted of the F_0 generation.

Poza Rica, Mexico collection

Larval *Ae. aegypti* from Poza Rica, Mexico were collected in 2015 as described in [291]. Larva from at least 30 different containers were returned to the laboratory and reared to adulthood. Adult *Ae. aegypti* females were offered a bloodmeal and eggs were collected and stored as the Poza Rica colony.

Tapachula, Mexico collection

Eggs from *Ae. aegypti* from Tapachula, Mexico were collected in 2019 through a collaboration with Dr. Karla Saavedra Rodriguez of the Colorado State University Center for Vector-borne Infectious Diseases as described in [292]. Briefly, ovitraps made of 1 L black polypropylene cups lined with Whatman filter paper were filled with water to 75% capacity and were placed in Pobres Unidos. Egg papers were hatched at the Regional Center for Research in Public Health/National Institute of Public Health (CRISP/INSP). *Ae. aegypti* mosquitoes were morphologically identified, placed in cages, and bloodfed from rabbit to obtain the F_1 generation.

Colony rearing

Mosquitoes reared in colony were maintained at 28 °C with 75–80% relative humidity and a 12 h light/12 h dark cycle. Mated females were fed defibrinated sheep blood (Colorado Serum Co., Denver, CO, USA) and 10 mM ATP approximately 4 days post emergence in an artificial feeding system. Females were encaged with oviposition cups (consisting of paper towel strips and small water-filled plastic cups) for 5 d, and the eggs were then retrieved and dried. Stored eggs were viable for up to 3 months. Eggs were hatched in sterile water, and larvae were fed with ground TetraMin (Melle, Germany) fish food.

3.4.2 Intrathoracic inoculation of DENV2

Mosquitoes were intrathoracically inoculated with DENV2 as previously described [249]. Three-day-old females were anesthetized at 4 °C and inoculated with 500 PFU DENV2 (Jamaica 1409 strain, accession number M20558) suspended in 50 nL of growth medium. Mosquitoes were maintained on a sugar diet for 7 d until further analysis.

3.4.3 Plaque assays

A subset of mosquitoes infected with DENV2 by intrathoracic inoculation were collected for plaque assay to assess infection prevalence and titer. Mosquitoes were homogenized in 1 mL

DMEM (7% inactivated FBS, 1% penicillin/streptomycin, 1% glutamine, 1% non-essential amino acids). Each sample was then passed through a 0.2 μm Acrodisc Syringe Filter fitted with Supor Membrane (Pall Life Sciences, East Hills, NY, USA). LLC-MK2 cells were grown to confluent monolayers in 24-well plates and infected with 10-fold serial dilutions of the homogenates (up to $1/10^5$ PFU/mL) for 1 h at 37 °C. After infection, 1 mL of a sterilized 1% agarose solution containing a nutrient supplement (10% 1× Medium 199 (Sigma-Aldrich), 5% inactivated FBS, 4% sodium bicarbonate, 2% diethylaminoethyl (DEAE)-dextran, 0.5% MEM amino acids (Mediatech Inc., Manassas, VA, USA), 0.5% MEM vitamins (Mediatech Inc., Manassas, VA, USA) was overlaid on each well. Plates were left to solidify for ~1 h and were then moved to the 37 °C incubator for 6 days. To visualize plaques, 150 μL of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT, 3 mg/mL in 1× PBS) was added to each well followed by ~24 h incubation. Plaques were visually quantified the next day. Viral titers of each sample were calculated as PFU/mL.

3.4.4 Small RNA library preparation and deep sequencing

Small RNAs were extracted from mosquitoes using the mirVana miRNA isolation kit (Invitrogen) following the manufacturer's instructions. RNA was quantified using a NanoDrop Spectrophotometer (ThermoFisher). 100 ng – 1 μg RNA was used for sRNA library preparation using the NEBNext small RNA library prep set for Illumina (New England Biolabs). When indicated in the manufacturer's instructions, reagents were diluted 1:2 in nuclease-free water. To increase ligation efficiency of piRNAs, the ligation reaction was performed O/N at 16 °C. PCR amplification of the reaction using the SR Primer for Illumina and multiplex index primers was performed following the manufacturer's instructions for 15 cycles. Libraries were then loaded into a 2% agarose gel and separated by size for 2-3 h at 100 V. Small RNAs, ~ 140 – 150 bp

including adaptors, were excised from the gel and purified using the NucleoSpin Gel & PCR clean up kit (TakaRa). Libraries were eluted from the column using 20 µl pre-heated nuclease-free water. Small RNA libraries were quantified using the NEBNext library quant kit for Illumina (New England Biolabs) following the manufacturer's instructions and pooled at 2-5 nM. Pooled libraries were again quantified using Qubit Fluorometric Quantification (ThermoFisher) and analyzed on a TapeStation (Agilent) before sequencing. Single end sequencing was performed on an Illumina 500/550 v2.5 NextSeq 75 cycle kit (Illumina) following the manufacturer's instructions.

3.4.5 Bioinformatic analyses

Small RNA datasets were analyzed using the MSRG bioinformatic analysis pipeline developed in Ma et al., 2021 [272] and available at https://laulab.bu.edu/msrg/. To generate sRNA graphs by length and type of sRNA, the genic/intergenic and transposon/virus count sRNA pipeline was used as described in Ma et al., 2021 [272]. Briefly, sRNA libraries were first trimmed by cutadapt [293] and mapped to virus sequences, allowing for two mismatches, using Bowtie [294]. Remaining sequences were mapped to miRNAs and structure RNAs using Bowtie, allowing for two mismatches, and removed. All remaining reads were mapped to the *Ae. aegypti* genome using Bowtie, allowing for two mismatches. In this ultimate subset of the data, reads were also aligned to TE families that had been previously processed to reduce redundancy, as described in the Supplementary Data in Ma et al., 2021 [272]. TEs with 55% similarity were grouped and annotated as the same family using the MeShClust program, which left 1,242 TE families for alignment and preserved ~60% genomic coverage.

3.4.6 Data access

Deep sequencing datasets generated in this study have been submitted to the NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) under series accession number GSE154531.

CHAPTER 4: *AEDES AEGYPTI* PIWI4 STRUCTURAL FEATURES ARE NECESSARY FOR RNA BINDING AND NUCLEAR LOCALIZATION[5]

4.1 Introduction

One of the *Ae. aegypti* Piwis, Piwi4 (termed "AePiwi4" throughout this chapter), has been associated with antiviral immunity, but its function remains unknown. Although unnecessary for sRNA production, AePiwi4 associated with Semliki Forest vsiRNAs and vpiRNAs in infected cells, as well as with several protein players involved in both the siRNA (Ago2 and Dcr2) and piRNA (Ago3, Piwi5, Piwi6, Yb, vreteno, Tejas, and minotaur) pathways [125,158]. Furthermore, silencing *AePiwi4* depleted 3′ 2′ O-methylated (mature) SINV-specific vsiRNAs and vpiRNAs and increased SINV, DENV2, and CHIKV replication in infected cells [127]. This phenotype was recapitulated in DENV2-infected *Ae. aegypti* mosquitoes, where silencing *AePiwi4* increased infectious virus titers between 5–10 dpi [127]. AePiwi4 also associated with highly conserved satellite repeat-derived piRNAs (tapiR1 and tapiR2) that were 3′ 2′ O-methylated [131]. Knocking down *AePiwi4* reduced tapiR1 and tapiR2 transcripts, and depleting tapiR1 in embryos arrested their development and prevented the degradation of maternally deposited transcripts [131]. Taken together, the role(s) of AePiwi4 appear to be diverse and span across several different RNAi pathways.

AePiwi4 has been consistently associated with long (28–30 nt), mature 3′ 2′ O-methylated (henceforth termed "3′m") piRNAs, and it was found in both the cytoplasmic and

[5] This section includes the complete manuscript cited as "Williams, A.E., Shrivastava. G., Gittis, A.G., Ganesan, S., Martin-Martin, I., Leon, P. C. V., Olson, K.E., Calvo, E., 2021. *Aedes aegypti* Piwi4 structural features are necessary for RNA binding and nuclear localization. *International Journal of Molecular Sciences*. 2021; *22*(23): 12733." The article is reproduced with permission and minor modifications have been made.

nuclear fractions in an embryonic *Ae. aegypti* cell line (Aag2) [127,131,158]. We therefore set out to characterize AePiwi4 structural motifs involved in piRNA binding and nuclear localization to gain further insights on AePiwi4 function. In human Piwis (Hiwi1, Hiwi2, and Hili), the PAZ (Piwi/Argonaute/Zwille) domain preferentially binds 3′m piRNAs because of a hydrophobic-binding pocket that is flexible enough to accommodate the methyl group [163]. This contrasts with the human Argonaute-1 (Ago1) PAZ domain where its more restrictive RNA-binding pocket exhibits preferential binding to 3′ 2′-OH (henceforth termed "3′nm") groups present on miRNAs [163]. We hypothesized that a flexible AePiwi4 PAZ domain would also determine AePiwi4 preferential binding to mature, long piRNA populations.

We first compared PAZ sequences across previously crystalized Piwis to determine AePiwi4 PAZ structural features and binding pockets involved in 3′ end piRNA recognition. We then characterized recombinant AePiwi4 PAZ-binding dynamics with the 3′-terminal ends of mature and non-methylated piRNAs by surface plasmon resonance (SPR). We found that mutating putative RNA-binding residues depleted or significantly impacted binding to both 3′m and 3′nm sRNAs, while a T41R change, present in *A. aegypti* Ago3, significantly improved binding. Finally, we characterized a functional nuclear localization signal (NLS) in the N-terminal region of the AePiwi4 protein. We found that subtle structural differences across Piwi proteins may have important impacts on preferential RNA-binding behaviors and subcellular localization.

4.2 Results

4.2.1 Biophysical properties of AePiwi4 by structural modeling and alignment

Using I-TASSER and Chimera, we first modeled AePiwi4 against the recently crystalized *Drosophila melanogaster* Piwi protein [189] (PDB: 6KR6) (Figure 4.1 A). The

quality of the predicted model was assessed by its C-score = −1.50. C-scores fall between −5 and 2, and more than 90% of the quality predictions are correct for models that have C-scores of −1.5 or higher [295]. Furthermore, the average template modeling (TM) score (0.53 ± 0.15) was > 0.5, which indicates a model of correct topology [295]. We then superimposed the AePiwi4 model against crystalized human (Hiwi1; PDB: 3O7V) [163] and mouse (Miwi; PDB: 2XFM) [296] PAZ to determine the amino acid residues of AePiwi4 PAZ to be M270 –T380 (Figure 4.1 B). A summary of predicted biophysical properties of AePiwi4 PAZ is provided in Table 4.1. An electrostatic density map of Piwi4 (Figure 4.1 C) revealed an inner pocket that was highly positively charged, analogous to the *Drosophila* Piwi linker regions that bind RNA nucleotides. AePiwi4 PAZ also displayed long stretches of flexibility with neighboring hydrophobic regions (Figure 4.1 D). The AePiwi4 PAZ model suggests that the protein contains hydrophobic regions buried within a flexible protein structure, allowing AePiwi4 to bind long 3′m piRNAs.

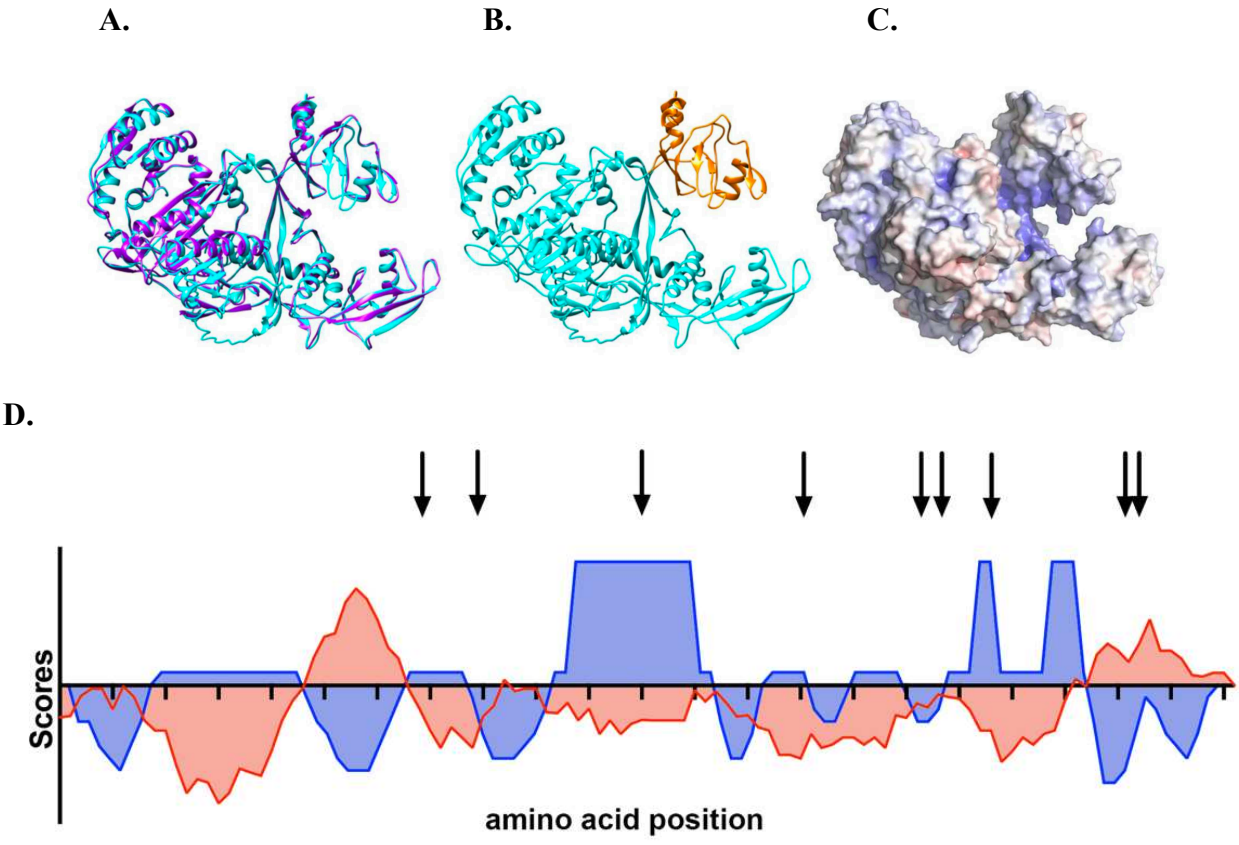


Figure 4.1. Predicted *Ae. aegypti* Piwi4 RNA-binding properties. (**A**) Predicted *Ae. aegypti* Piwi4 model (blue) superimposed with the crystalized *Drosophila* Piwi structure (purple); (**B**) predicted *Ae. aegypti* Piwi4 structure (blue) with PAZ domain highlighted in orange; (**C**) electrostatic density of *Ae. aegypti* Piwi4 where red = negatively charged and blue = positively charged. Structure is rotated on the right to reveal inner positively charged pocket; (**D**) predicted AePiwi4 PAZ flexibility scores (Karplus-Schulz prediction, blue), hydrophobicity (Hopp-Woods prediction, red), and RNA binding sites (black arrows) based on other crystalized Piwi structures. Graphs superimposed to each other in relation to their threshold scores.

Table 4.1. Biophysical properties of predicted AePiwi4 PAZ. Biophysical parameters and associated values for AePiwi4 PAZ. pI = isoelectric point. Average hydropathy = the Kyte-Doolittle hydropathy value. Abs280 = the absorbance at 280 nm in an oxidized environment. Molar Extinction Coefficient = the amount of light the protein absorbs, used to calculate protein concentration.

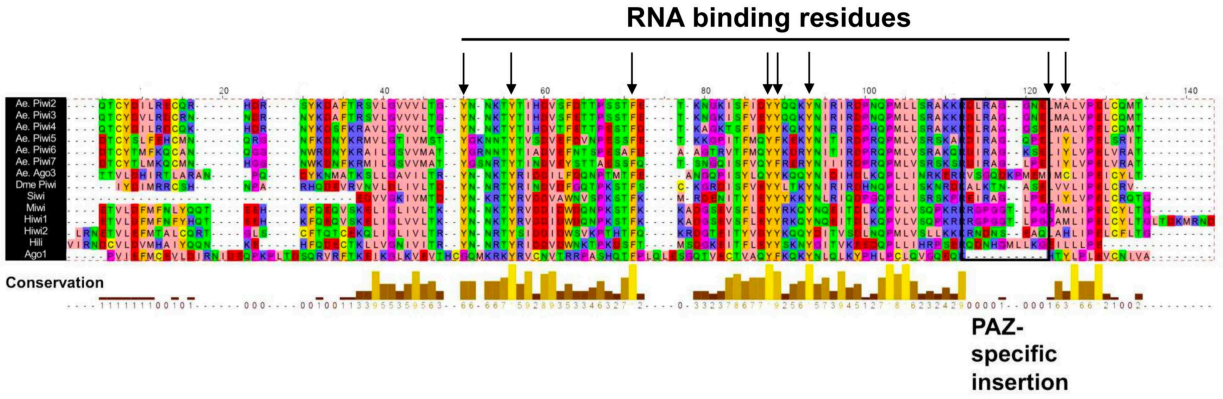
Parameter	Value
Molecular weight (g/mol)	12,945
Net charge at neutral pH 7	5.08
pI	9.17
Average hydropathy (GRAVY)	-0.66
Aliphatic index	69
Abs280	0.82
Molar Extinction Coefficient ($M^{-1}cm^{-1}$)	10,430

To determine putative AePiwi4 PAZ amino acids involved in RNA binding, we aligned Piwi PAZ sequences derived from proteins whose structures had been crystalized, including *Drosophila* Piwi [189] (PDB: 6KR6), silkworm Piwi (Siwi [190]) (PDB: 5GUH), mouse Piwi (Miwi [296]) (PDB: 2XFM), and human Hili (PDB: 3O7X) and Hiwi1 (PDB: 3O6E) [163] (Figure 4.2 A). We also included human Hiwi2 PAZ, as its binding properties to the 3′ ends of piRNAs were characterized by isothermal calorimetry (ITC) in Tian et al., 2011 [163]. As a comparison, we included the outgroup Argonaute protein human Ago1 (PDB: 4KXT), whose more restricted PAZ domain dictates its RNA-binding preference for 3′nm miRNAs [163,297]. We then compared known RNA-binding residues from the crystal structures with the residues of *Ae. aegypti* PAZ domains (black arrows, Figure 4.1 D; Figure 4.2 A). We found that

the residues involved in RNA binding tended to be highly conserved across the different organisms, and many were tyrosine and phenylalanine aromatic residues, whose hydroxyl groups form hydrogen bonds with phosphate oxygens of RNA nucleotides [163]. We also noted that all Piwi PAZ domains analyzed herein displayed the Piwi PAZ specific insertion element (black box, Figure 4.2 A) that provides the flexibility necessary for accommodating 3′m piRNA ends [163]. This insertion site lies between two beta barrels, which, when absent (as in Ago1), results in a sharp turn between the barrels and a narrower binding pocket [163]. Although the amino acids within the Piwi PAZ-specific insertion site are not conserved across organisms, or even across subfamily Piwi proteins in the same organism [163], we observed that the first five amino acids of the Piwi PAZ-specific insertion sites were highly conserved across the *Ae. aegypti* Piwis. The only exception to this observation was *Ae. aegypti* Ago3, whose residues shared no similarity to its AePiwi PAZ counterparts. Further, Ago3 displayed two amino acids within the insertion site not seen in the other PAZ sites, perhaps suggesting a more flexible binding pocket than the other *Ae. aegypti* Piwis.

We also generated a phylogenetic tree to compare evolutionary relatedness between Piwi PAZ sequences from the various organisms (Figure 4.2 B). We included human Ago1 again as an outgroup. We also added the two additional *Drosophila* Piwis, Aub, and Ago3. We found that all *Ae. aegypti* Piwi PAZ, except for Ago3, clustered with *Drosophila* Aub. On the other hand, *Ae. aegypti* Ago3 was most closely related to *Drosophila* Ago3. We also observed that AePiwi4 clustered with the germline AePiwis Piwi2-3 as opposed to the somatic AePiwis Piwi5, Piwi6, and Ago3.

A.



B.

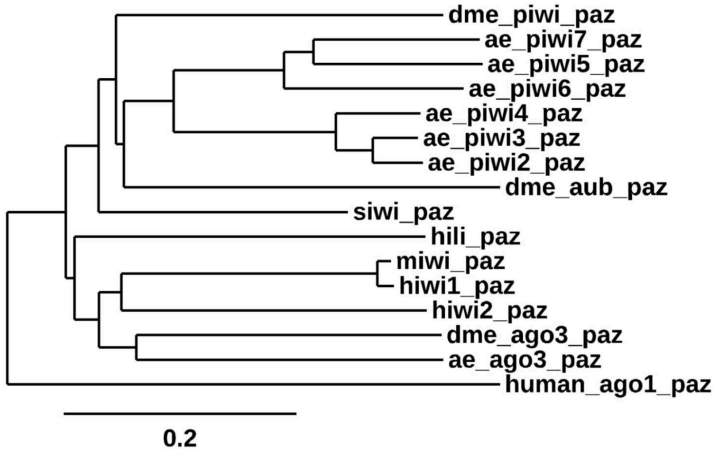


Figure 4.2. Predicted RNA binding residues and phylogenic tree of Piwi PAZ across organisms. (**A**) Alignment of Piwi PAZ domains, including all *Ae. aegypti* Piwis (Ae. Piwi2-7 and Ae. Ago3) and crystalized or characterized Piwi PAZ (*Drosophila* (Dme) Piwi, silkworm Piwi (Siwi), mouse Piwi (Miwi), and human Piwis Hiwi1, Hiwi2, and Hili). The human Argonaute protein Ago1 was also included as an outgroup. Black arrows indicate known RNA-binding residues by crystal structures, which in this alignment include amino acid numbers 50, 56, 71, 88, 89, 93, 123, and 125. Black box indicates the Piwi PAZ-specific insertion site. Colors (Zappo color scheme) indicate biochemical properties where peach = aliphatic/hydrophobic, aromatic = orange, blue = positively charged, red = negatively charged, green = hydrophilic, pink = conformationally special, and yellow = cysteine; (**E**) phylogenetic tree of all Piwi PAZ included in the alignment shown in 1D, with the addition of *Drosophila* Aubergine and Ago3 PAZs. Scale bar indicates number of substitutions per site.

4.2.2 Ae. aegypti *Piwi4 PAZ binds 3'-terminal 2'-O-methylated and 3'-terminal non-methylated piRNAs in a sequence-independent manner*

To determine whether *Ae. aegypti* Piwi4 PAZ bound to both 3′ 2′ O-methylated and 3′ 2′ OH piRNAs, we cloned *AePiwi4 PAZ* with a histidine 6xHis-tag into pET-17b for bacterial expression (Figure 4.3 A, B), purified it, and characterized RNA-binding dynamics by SPR. We confirmed AePiwi4 PAZ expressed in BL21(DE3) pLysS *E. coli* by Western blot using antibodies against its 6xHis-tag (Figure 4.3 C). We purified AePiwi4 PAZ from the soluble fraction by nickel chromatography followed by size exclusion chromatography. Soluble AePiwi4 PAZ was stable in 20 mM Tris-HCl pH 7.4 and 150 mM NaCl and ran at the expected size (14 kDa) by SDS/PAGE gel electrophoresis (Figure 4.3 D). Protein identity was confirmed by Edman degradation.

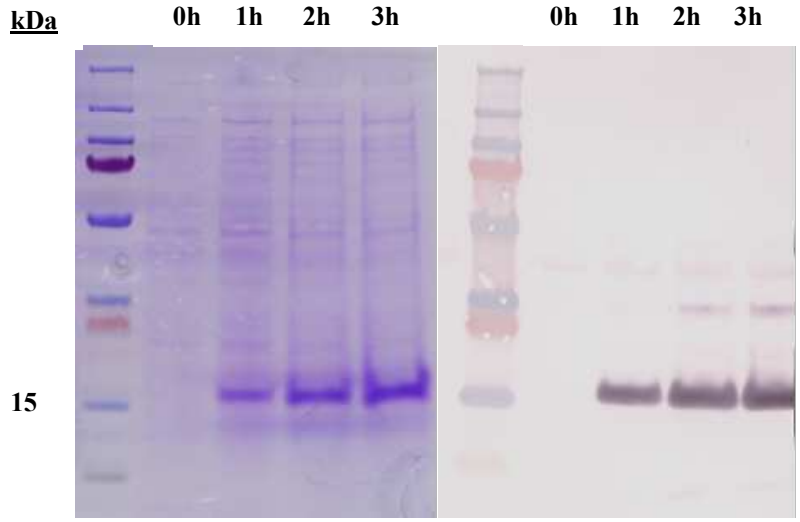
A.

>AePiwi4_PAZ_DNA
catatgCAGACATGCTACGACATCTTGCGCGATTGTCAGAAGCACGATCGTAACTACAA
GGATTCGTTCAAACGTGCCGTACTTGGTGTCGTTGTACTGACCGGTTACAACAATAA
AACCTATACCATTCACGACGTCACGTTTGAAACCACTCCGGAGAGTACGTTCGATAC
CAAGGCCGGTAAAACGTCCTTCATTGAGTATTACAAACAGAAGTACAACATTCGCAT
TCGTGATCCACATCAGCCTATGTTGCTGTCGCGAGCCAAGAAACGCGATCTGCGCGC
TGGAGGCAGCGAACTCATGGCCCTTGTTCCAGAACTGTGCCAGATGACGCATCATCA
CCATCACCATTGActcgag

B.

>AePiwi4_PAZ_protein
MQTCYDILRDCQKHDRNYKDSFKRAVLGVVVLTGYNNKTYTIHDVTFETTPESTFDTK
AGKTSFIEYYKQKYNIRIRDPHQPMLLSRAKKRDLRAGGSELMALVPELCQMTHHHHH
H

C.



D.

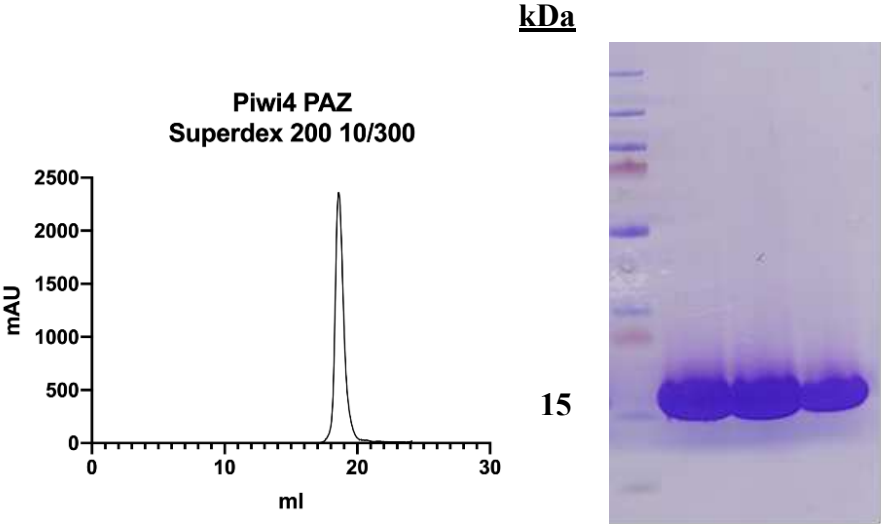


Figure 4.3. AePiwi4 PAZ expression and purification. A. *AePiwi4 PAZ* nucleotide sequence, including restriction enzyme sites (lower case letters) used for cloning and a 6xHis-tag (underlined). **B.** AePiwi4 PAZ amino acid sequence, including 6xHis-tag. **C.** AePiwi4 PAZ expression trial in *E. coli* BL21(DE3) pLYsS cells assessed by Coomassie-stained (left) and anti-6xHis-tag western blot. Lanes indicate hours post-induction. **D.** AePiwi4 PAZ protein purification by size exclusion chromatography (left) where peak elutions were confirmed by Coomassie-stained SDS/PAGE gel (right). Expected AePiwi4 PAZ protein size = 14 kDa.

To test RNA-binding dynamics, we performed SPR with AePiwi4 PAZ and three different RNAs: a 3′m 28 nt PCLV-specific vpiRNA, a 3′nm vpiRNA of the same sequence, and a 3′nm 28 nt scrambled sequence. We chose this piRNA sequence because we had found it was present across tissues in several *Ae. aegypti* strains (Chapter 3), and because a recent publication noted that PCLV piRNAs are broadly distributed across culicine mosquito cell lines, perhaps due to a PCLV-specific endogenous viral element in the genome [272]. Using synthetic RNA sequences biotinylated on the 5′ end, we immobilized the RNA on a CM5 Biacore surface that had been pre-coated with neutravidin. Immobilizing RNA by the 5′ end allowed us to test binding affinities to moieties at the 3′ end. Using increasing concentrations of AePiwi4 PAZ analytes flowed over the surface of the chip with immobilized ligand, we were able to determine the dissociation constants (K_D) from steady-state binding levels (R_{eq}) against the analyte concentration (C, in molar concentration) once binding reached equilibrium. Experiments were performed in four replicates.

We found that AePiwi4 PAZ bound the 3′m 28 nt piRNA with a K_D of 1.7 ± 0.8 μM (Figure 4.4 A), the 3′nm 28 nt piRNA with a K_D of 5.0 ± 2.2 μM (Figure 4.4 B), and the scrambled 28 nt 3′nm piRNA with a K_D of 2.5 ± 0.1 μM (Figure 4.4 C). AePiwi4 PAZ bound to 3′ 2′ O-methylated piRNAs with marginally greater affinity than it did to 3′ 2′ unmethylated piRNAs ($p = 0.05$), and there was no significant difference in binding affinities for known or scrambled RNA sequences ($p = 0.25$).

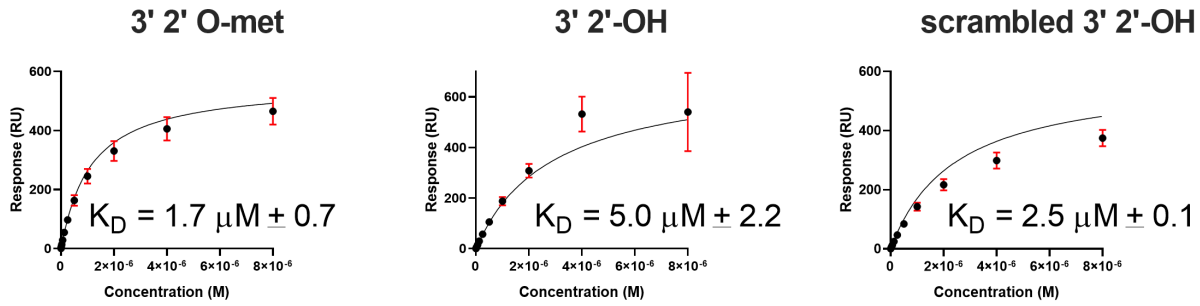


Figure 4.4. Affinity-binding equilibrium curves for AePiwi4 PAZ to piRNAs. Fitted affinity-binding equilibrium curves for AePiwi4 PAZ analyte to a 28 nt (**A**) 3′ 2′ O-methylated, (**B**) non-methylated, or (**C**) scrambled 28 nt non-methylated piRNA. Equilibrium K_D was calculated from steady-state binding levels $R_{eq} = (CR_{max})/(K_D + C)$ + offset, where C = concentration, R_{max} = analyte-binding capacity of the surface in response units (RU), K_D = dissociation constant, and offset = response at zero analyte concentration. M = molar concentration. Red bars indicate mean and standard deviation for R_{max} values.

4.2.3 Ae. aegypti *Piwi4 PAZ mutants reveal the amino acids necessary for piRNA binding*

For further insights into how the AePiwi4 PAZ structure dictates its RNA-binding preferences, we generated AePiwi4 PAZ mutants that displayed amino acid changes within predicted RNA-binding pockets. We focused our efforts around two highly conserved residues shown to form hydrogen bonds with RNA—Y40 and F55—as well as two residues that flank Y40 and appeared to be moderately conserved across the *Ae. aegypti* Piwi PAZ—T39 and T41. Through site-directed mutagenesis, we generated five mutants that we then expressed in bacteria and purified: T39A, Y40A, T41A, F55A, and T41R (Figure 4.5, 4.6). Y40A and F55A displayed alanine substitutions for the highly conserved tyrosine or phenylalanine amino acids, respectively, while T39A and T41A displayed alanine substitutions for the threonines that flank Y40. The T41R mutation reflected the arginine present in only one *Ae. aegypti* Piwi, Ago3, but also in most other organisms' Piwi PAZ domains analyzed herein (Figure 4.2 A). Inserts were confirmed by Sanger sequencing, and mutations in protein sequence were confirmed by mass spectrometry. Binding behaviors were assessed by SPR, as described previously.

A.

>AePiwi4_PAZ_T39A_DNA
catatgCAGACATGCTACGACATCTTGCGCGATTGTCAGAAGCACGATCGTAACTACAA
GGATTCGTTCAAACGTGCCGTACTTGGTGTCGTTGTACTGACCGGTTACAACAATAA
A**G**CCTATACCATTCACGACGTCACGTTTGAAACCACTCCGGAGAGTACGTTCGATAC
CAAGGCCGGTAAAACGTCCTTCATTGAGTATTACAAACAGAAGTACAACATTCGCAT
TCGTGATCCACATCAGCCTATGTTGCTGTCGCGAGCCAAGAAACGCGATCTGCGCGC
TGGAGGCAGCGAACTCATGGCCCTTGTTCCAGAACTGTGCCAGATGACGCATCATCA
CCATCACCATTGActcgag

B.

>AePiwi4_PAZ_T39A_protein
MQTCYDILRDCQKHDRNYKDSFKRAVLGVVVLTGYNNK**A**YTIHDVTFETTPESTFDTK
AGKTSFIEYYKQKYNIRIRDPHQPMLLSRAKKRDLRAGGSELMALVPELCQMTHHHHH
H

C.

>AePiwi4_PAZ_Y40A_DNA
catatgCAGACATGCTACGACATCTTGCGCGATTGTCAGAAGCACGATCGTAACTACAA
GGATTCGTTCAAACGTGCCGTACTTGGTGTCGTTGTACTGACCGGTTACAACAATAA
AACC**GC**TACCATTCACGACGTCACGTTTGAAACCACTCCGGAGAGTACGTTCGATAC
CAAGGCCGGTAAAACGTCCTTCATTGAGTATTACAAACAGAAGTACAACATTCGCAT
TCGTGATCCACATCAGCCTATGTTGCTGTCGCGAGCCAAGAAACGCGATCTGCGCGC
TGGAGGCAGCGAACTCATGGCCCTTGTTCCAGAACTGTGCCAGATGACGCATCATCA
CCATCACCATTGActcgag

D.

>AePiwi4_PAZ_Y40A_protein
MQTCYDILRDCQKHDRNYKDSFKRAVLGVVVLTGYNNKT**A**TIHDVTFETTPESTFDTK
AGKTSFIEYYKQKYNIRIRDPHQPMLLSRAKKRDLRAGGSELMALVPELCQMTHHHHH
H

E.

>AePiwi4_PAZ_T41A_DNA
catatgCAGACATGCTACGACATCTTGCGCGATTGTCAGAAGCACGATCGTAACTACAA
GGATTCGTTCAAACGTGCCGTACTTGGTGTCGTTGTACTGACCGGTTACAACAATAA
AACCTAT**G**CCATTCACGACGTCACGTTTGAAACCACTCCGGAGAGTACGTTCGATAC
CAAGGCCGGTAAAACGTCCTTCATTGAGTATTACAAACAGAAGTACAACATTCGCAT
TCGTGATCCACATCAGCCTATGTTGCTGTCGCGAGCCAAGAAACGCGATCTGCGCGC
TGGAGGCAGCGAACTCATGGCCCTTGTTCCAGAACTGTGCCAGATGACGCATCATCA
CCATCACCATTGActcgag

F.

>AePiwi4_PAZ_T41A_protein
MQTCYDILRDCQKHDRNYKDSFKRAVLGVVVLTGYNNKTY**A**IHDVTFETTPESTFDTK
AGKTSFIEYYKQKYNIRIRDPHQPMLLSRAKKRDLRAGGSELMALVPELCQMTHHHHH
H

G.

>AePiwi4_PAZ_T41R_DNA
catatgCAGACATGCTACGACATCTTGCGCGATTGTCAGAAGCACGATCGTAACTACAA
GGATTCGTTCAAACGTGCCGTACTTGGTGTCGTTGTACTGACCGGTTACAACAATAA
AACCTAT**CG**CATTCACGACGTCACGTTTGAAACCACTCCGGAGAGTACGTTCGATAC
CAAGGCCGGTAAAACGTCCTTCATTGAGTATTACAAACAGAAGTACAACATTCGCAT
TCGTGATCCACATCAGCCTATGTTGCTGTCGCGAGCCAAGAAACGCGATCTGCGCGC
TGGAGGCAGCGAACTCATGGCCCTTGTTCCAGAACTGTGCCAGATGACGCATCATCA
CCATCACCATTGActcgag

H.

>AePiwi4_PAZ_T41R_protein
MQTCYDILRDCQKHDRNYKDSFKRAVLGVVVLTGYNNKTY**R**IHDVTFETTPESTFDTK
AGKTSFIEYYKQKYNIRIRDPHQPMLLSRAKKRDLRAGGSELMALVPELCQMTHHHHH
H

I.

>AePiwi4_PAZ_F55A_DNA
catatgCAGACATGCTACGACATCTTGCGCGATTGTCAGAAGCACGATCGTAACTACAA
GGATTCGTTCAAACGTGCCGTACTTGGTGTCGTTGTACTGACCGGTTACAACAATAA
AACCTATACCATTCACGACGTCACGTTTGAAACCACTCCGGAGAGTACG**GC**CGATA
CCAAGGCCGGTAAAACGTCCTTCATTGAGTATTACAAACAGAAGTACAACATTCGC
ATTCGTGATCCACATCAGCCTATGTTGCTGTCGCGAGCCAAGAAACGCGATCTGCGC
GCTGGAGGCAGCGAACTCATGGCCCTTGTTCCAGAACTGTGCCAGATGACGCATCAT
CACCATCACCATTGActcgag

J.

>AePiwi4_PAZ_F55A_protein
MQTCYDILRDCQKHDRNYKDSFKRAVLGVVVLTGYNNKTYTIHDVTFETTPEST**A**DTK
AGKTSFIEYYKQKYNIRIRDPHQPMLLSRAKKRDLRAGGSELMALVPELCQMTHHHHH
H

Figure 4.5. AePiwi4 PAZ mutant nucleotide and protein sequences. Mutant *AePiwi4 PAZ* (**A**) T39A, (**C**) Y40A, (**E**) T41A, (**G**) T41R, (**I**) F55A nucleotide sequences, including restriction

enzyme sites (lower case letters) used for cloning and a 6xhis tag (underlined). Mutant site is bolded and underlined. Mutant AePiwi4 PAZ (**B**) T39A, (**D**) Y40A, (**F**) T41A, (**H**) T41R, (**J**) F55A amino acid sequences, including 6xHis-tag. Mutation site is bolded and underlined.

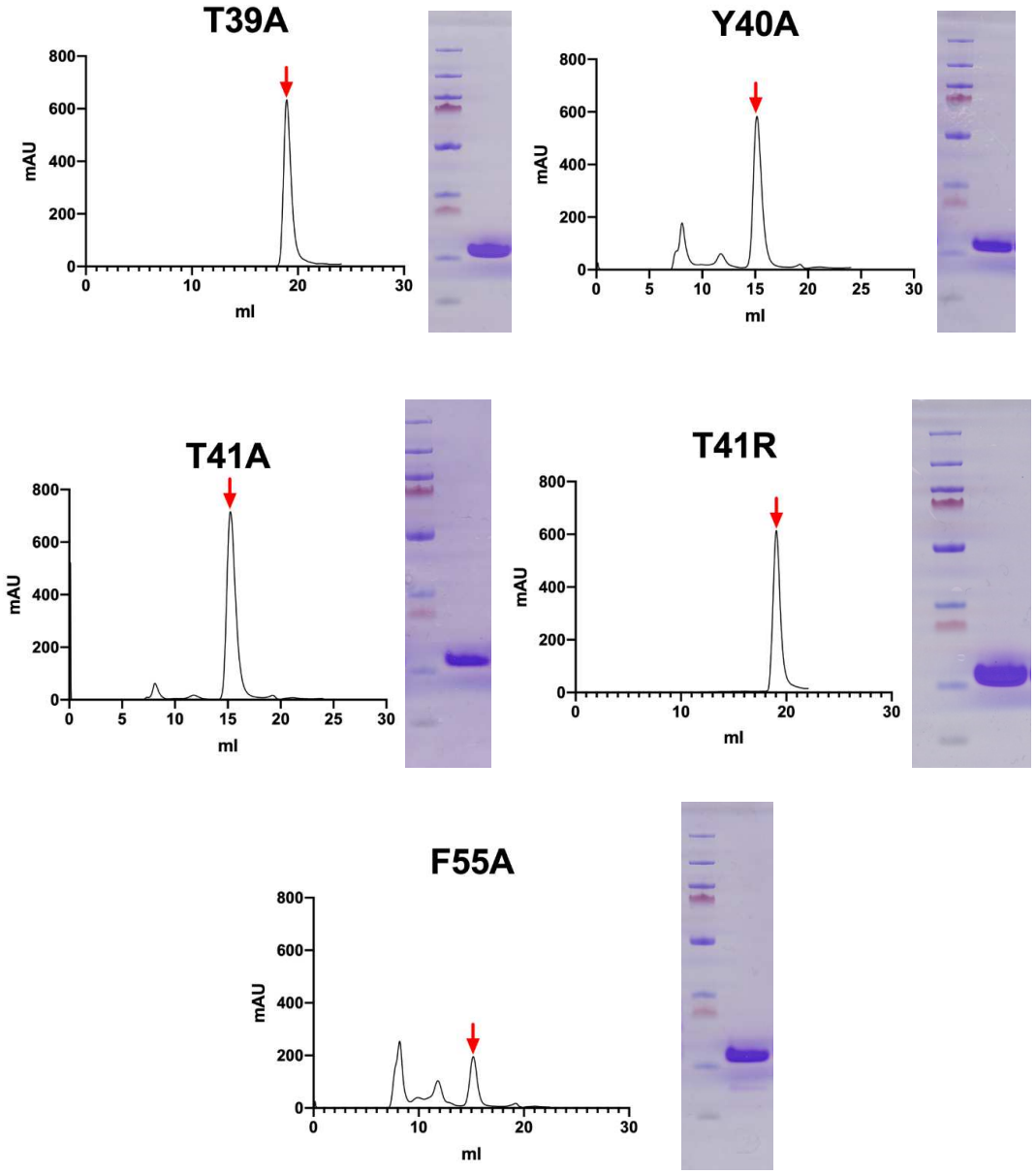
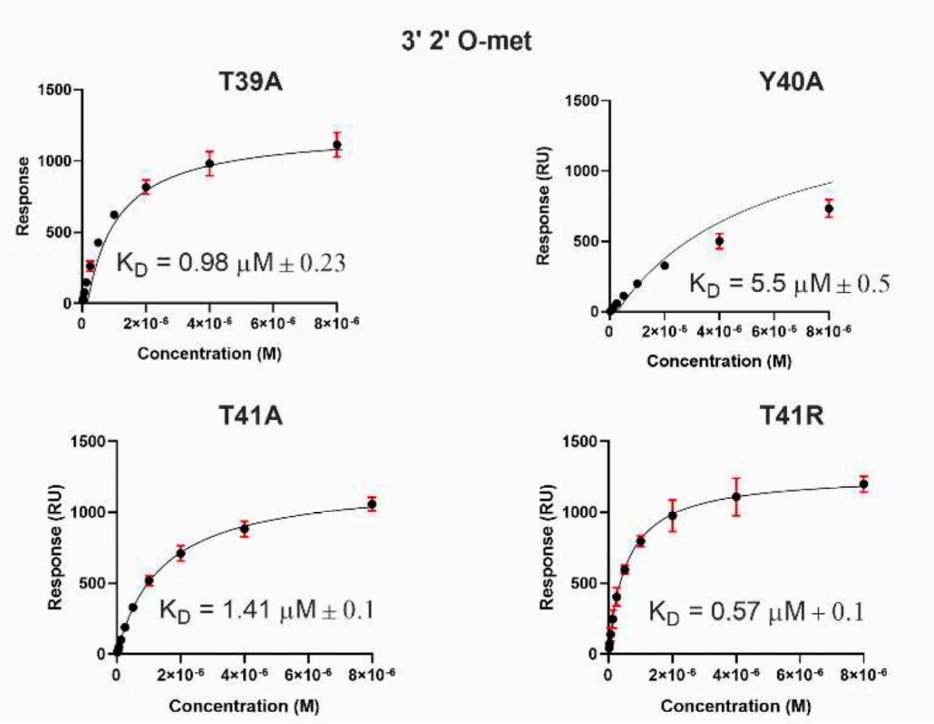
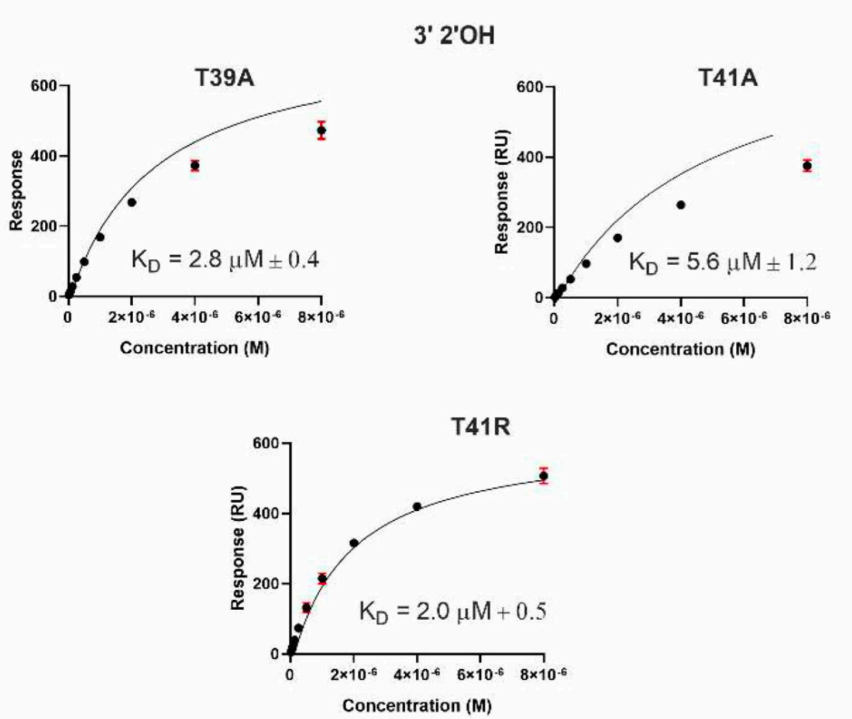


Figure 4.6. Expression and purification of AePiwi4 PAZ proteins displaying respective mutations in predicted RNA binding pockets. Size Exclusion Chromatograms and corresponding Coomassie-stained SDS/PAGE gels of purified T39A, Y40A, T41A, T41R, or F55A AePiwi4 PAZ mutants. Predicted size of all proteins are ~ 14 kDa. Red arrows indicate protein peaks of interest.

Dissociation constants for all mutant proteins binding both the 3′m and 3′nm 28 nt piRNA are displayed in Figure 4.7 and summarized in Table 4.2. We found that F55 was essential for both 3′m and 3′nm binding because when mutated, no binding occurred for either ligand. Y40A also depleted 3′nm binding and significantly inhibited 3′m piRNA binding (K_D = 5.5 ± 0.5 μM; p = 0.04). We found that disrupting the amino acids flanking Y40 with alanine mutations had no significant impact on binding 3′m piRNAs as compared to WT PAZ binding to this RNA (T39A: p = 0.2; T41A: p = 0.3). However, we did observe a significantly increased affinity of T39A for the 3′nm piRNA (K_D = 2.8 ± 0.4 μM; p = 0.02), suggesting that this residue does have an impact on 3′nm binding. Furthermore, we observed that mutating T41 to match the amino acid present in *Ae. aegypti* Ago3 PAZ tended to improve 3′m binding (K_D = 0.57 ± 0.1 μM) and significantly improved 3′nm binding (K_D = 2.0 ± 0.5 μM; p = 0.02).



(A)



(**B**)

Figure 4.7. Affinity-binding equilibrium curves for AePiwi4 PAZ mutants to piRNAs. Fitted affinity-binding equilibrium curves for AePiwi4 PAZ mutant analytes to a 28 nt (**A**) 3′ 2′ O-methylated or (**B**) non-methylated piRNA. Equilibrium K_D was calculated from steady-state binding levels $R_{eq} = (CR_{max})/(K_D + C)$ + offset, where C = concentration, R_{max} = analyte-binding capacity of the surface in response units (RU), K_D = dissociation constant, and offset = response at zero analyte concentration. M = molar concentration. Red bars indicate mean and standard deviation for R_{max} values.

Table 4.2. Summary of disassociation constants for AePiwi4 PAZ mutants binding 3′ 2′ O-methylated (met) or non-methylated (nmet) piRNA. Equilibrium K_D was calculated from steady-state binding levels $R_{eq} = (CR_{max})/(K_D + C)$ + offset, where C = concentration, R_{max} = analyte-binding capacity of the surface in response units (RU), K_D = dissociation constant, and offset = response at zero analyte concentration. * = $p \leq 0.05$ by unpaired t-test with WT as comparison group.

Immobilized ligand	Binding AePiwi4 PAZ	K_D values (μM)	R_{max} (RU)
3' 2'met piRNA	WT	1.7 ± 0.7	1364 ± 3
	T39A	1.0 ± 0.2	1343 ± 2
	Y40A	5.5 ± 0.5 *	1103 ± 62
	T41A	1.4 ± 0.07	1272 ± 14
	T41R	0.57 ± 0.1	1388 ± 18

	F55A	no binding	
3' 2'nmet piRNA	WT	5 ± 2.2	585 ± 3
	T39A	2.8 ± 0.4 *	596 ± 5
	Y40A	no binding	
	T41A	5.6 ± 1.2	580 ± 8
	T41R	2.0 ± 0.5 *	590 ± 3
	F55A	no binding	

In Hiwi1 PAZ, preferential binding of 3′m RNA over 3′nm RNA is mostly dictated by backbone confirmation of the protein rather than the amino acid composition of the binding pocket [163]. To investigate whether the mutations that impacted the RNA binding impacted the AePiwi4 PAZ secondary structure, we performed circular dichroism (CD) spectroscopy analysis with the T39A, Y40A, T41A, and T41R mutants and compared the CD curves to that of the WT AePiwi4 PAZ (Figure 4.8). We analyzed the data using CAPITO [298]. WT AePiwi4 PAZ displayed a CD curve most like proteins that had a mostly irregular structure but also that had between 30% and 49% beta strands and 6–16% alpha helices. All mutants maintained a mostly irregular secondary structure; however, they displayed different CD curves and percentages of alpha helices and beta-sheets compared to the WT protein (Figure 4.8 insets). T41R was most like proteins that were made up of between 16% and 26% beta strands and 28–46% alpha helices. Y40A also displayed a spectrum that aligned more with proteins that had a greater abundance of alpha helices—26–40% alpha helices but only 14–22% beta-strands. The T39A CD curve clustered with proteins that were made up of 9–25% alpha helices and 30–41% beta-strands, while the T41A curve clustered with proteins that were made up of 31–50% alpha helices and 4–21% beta-strands. Taken together, these data indicate that single amino acid changes in the AePiwi4 PAZ backbone can alter secondary structure, which likely impacted RNA-binding behaviors.

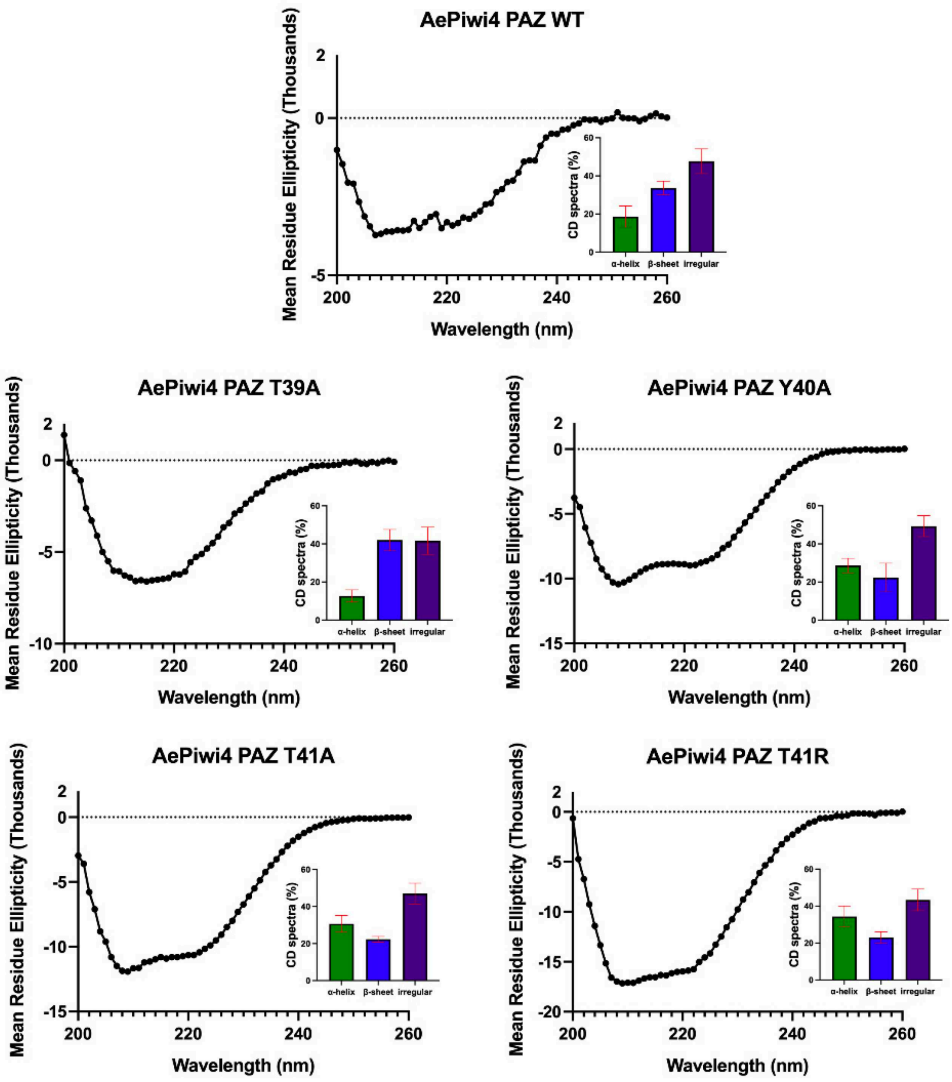


Figure 4.8. Circular dichroism (CD) spectra analyses of AePiwi4 PAZ mutants. CD spectra curves, by mean residue ellipticity for AePiwi4 PAZ WT and mutant proteins, recorded over 200–260 nm. Insets show the calculated percentages of secondary structures determined by CD analysis using CAPITO. Red bars indicate mean and standard deviation of three similarity hits based on lowest area differences under the curve.

4.2.4 AePiwi4 co-localizes in the cytoplasm and nucleus in Ae. aegypti tissues

For further insights into the function of *Ae. aegypti* Piwi4, we characterized the sub-cellular localization of the native protein in *Ae. aegypti* mosquitoes. While this manuscript was under preparation, Joosten and colleagues (2021) reported that Piwi4, Piwi5, and Piwi6 were in both the nucleus and the cytoplasm in an *Ae. aegypti* embryonic cell line infected or uninfected

with SINV [158]. To determine the sub-cellular localization of native AePiwi4 in both somatic and germline tissues in the mosquito, we generated polyclonal antibodies against AePiwi4 by immunizing mice with the AePiwi4 PAZ recombinant protein. We confirmed that these antibodies recognized both recombinant AePiwi4 PAZ and full-length proteins by Western blot (Figure 4.9). To confirm that the antibodies recognized AePiwi4 from mosquito tissues, we prepared whole mosquito lysates from three *Ae. aegypti* females 48 h post-bloodmeal (time of peak AePiwi4 expression [127]) for Western blot. Anti-AePiwi4 mouse serum reacted to whole mosquito lysate at the expected size of AePiwi4 (100 kDa) (Figure 4.9). Mass spectrometry analyses further confirmed that AePiwi4-specific peptides were present at the same location on a corresponding SDS/PAGE gel slice.

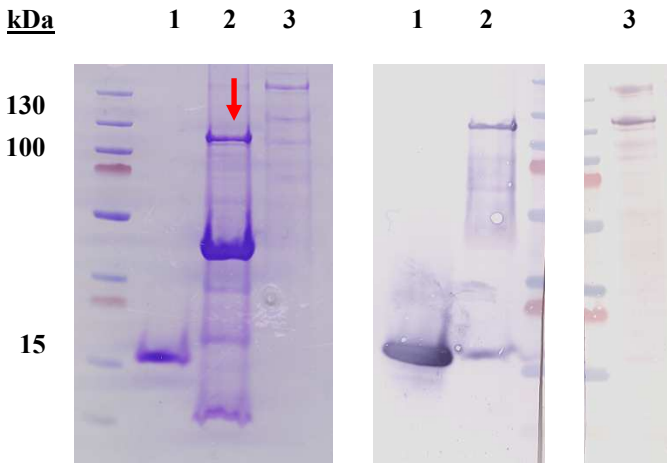
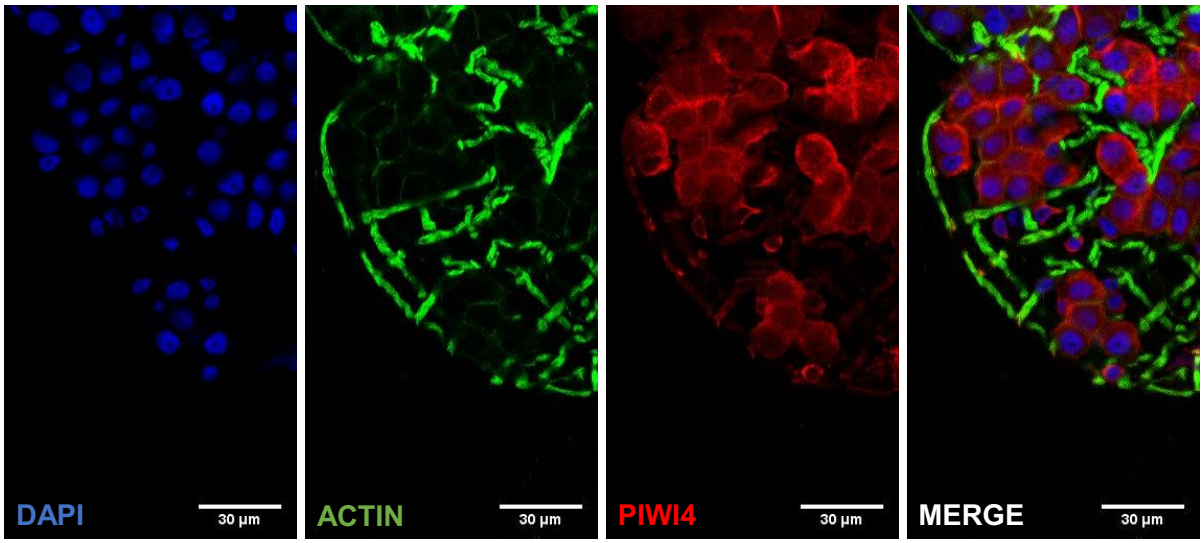


Figure 4.9. Anti-AePiwi4 PAZ reacts to AePiwi4 by western blot with recombinant proteins and whole mosquito lysate. Left. Coomassie-stained SDS/PAGE gel with recombinant AePiwi4 PAZ (lane 1), recombinant AePiwi4 full length (lane 2, indicated by red arrow), and whole mosquito lysate (lane 3). **Right.** Corresponding western blot of recombinant AePiwi4 PAZ (lane 1), recombinant AePiwi4 full length (lane 2), and whole mosquito lysate (lane 3) reacts against mouse serum after immunization with recombinant AePiwi4 PAZ.

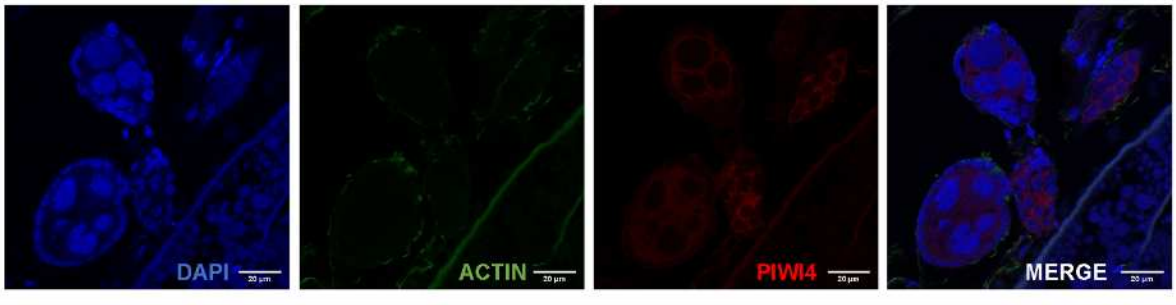
To determine the sub-cellular localization of AePiwi4, we next performed immunofluorescence assays and Western blots using both somatic and germline-derived

mosquito tissues. AePiwi4 tended to stain cytoplasmically in both midguts (Figure 4.10 A, C) and unfertilized embryos from ovary tissues (Figure 4.10 B, D). However, when we fractionated ovaries from *A. aegypti* mosquitoes 48 h post-bloodmeal into cytoplasmic and nuclear fractions for Western blot, we found that AePiwi4 was present in both fractions (Figure 4.10 E). Antibodies targeting H3 histone were used as a marker for the nuclear fraction (Figure 4.10 F). These results suggested AePiwi4 may be trafficked in and out of the nucleus in mosquito tissues.

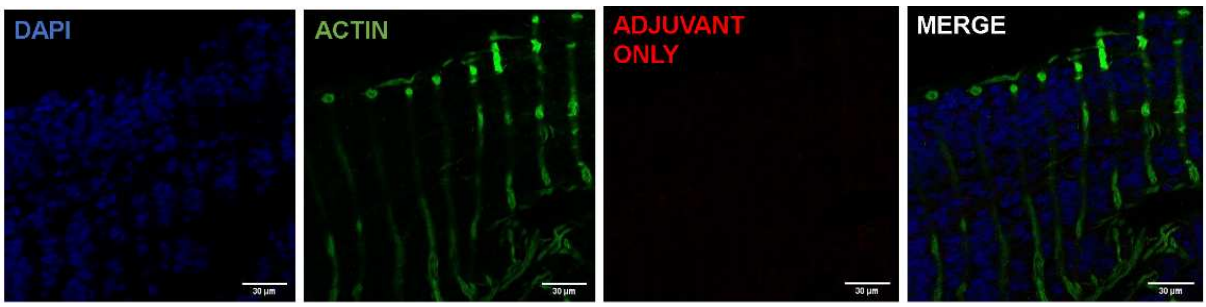
A.



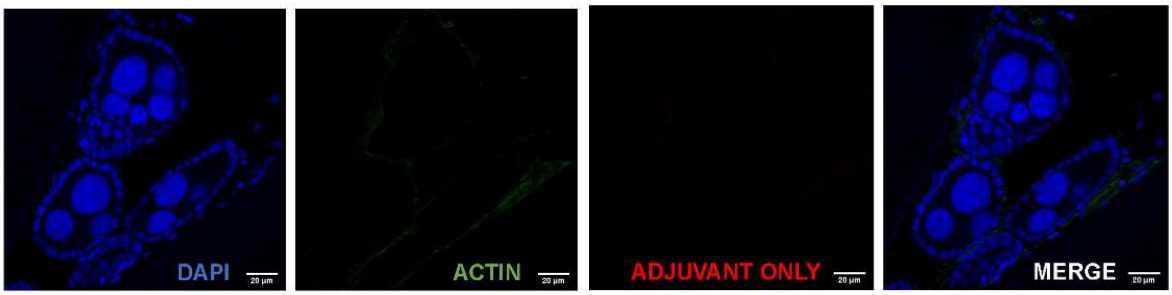
B.



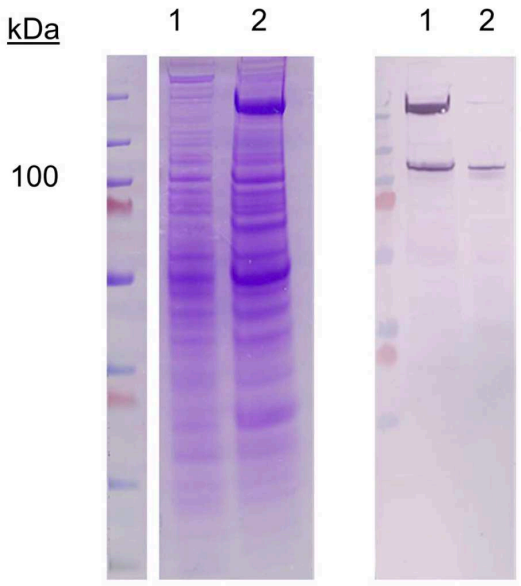
C.


DAPI
ACTIN
ADJUVANT ONLY
MERGE

D.


DAPI
ACTIN
ADJUVANT ONLY
MERGE

E.


kDa
1
2
1
2
100

F.

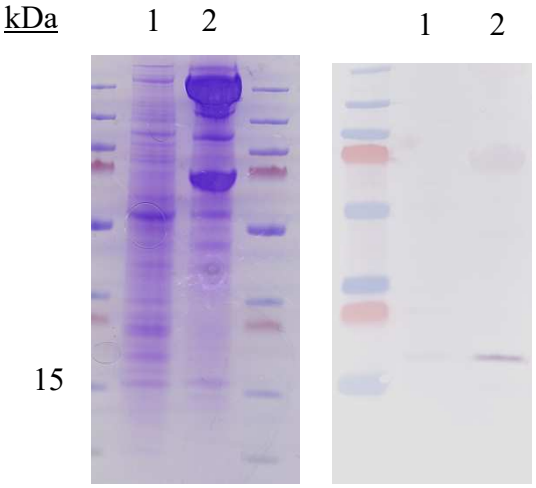


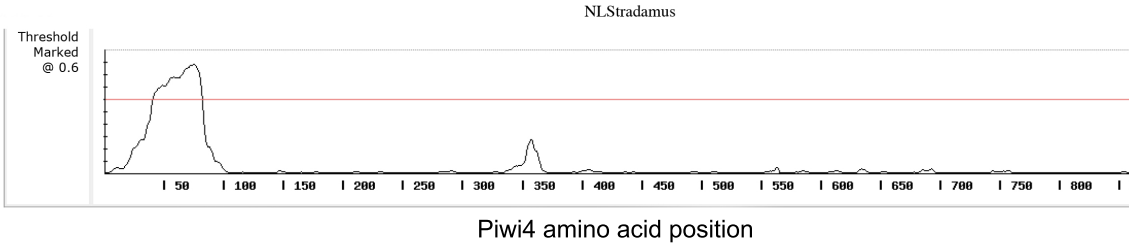
Figure 4.10. AePiwi4 expression in mosquito tissues. IFA of *Ae. aegypti* midgut (A) or ovaries with unfertilized embryos (B) stained with anti-AePiwi4 (red), phalloidin (green), or DAPI (blue). Negative control IFA images of *Ae. aegypti* (C) midguts and (D) ovaries containing unfertilized eggs stained with mouse serum immunized with Magic Mouse adjuvant alone (red), phalloidin (green), and DAPI (blue) as negative controls for IFA analyses. Scale bars are 30 µM and 20 µM for midguts and embryos, respectively; (E) Coomassie-stained SDS/PAGE gel (left) and corresponding Western blot of cytoplasmic (lane 1) or nuclear (lane 2) fractions of *Ae. aegypti* mosquito ovary tissue; (F) Nuclear marker against *Ae. aegypti* ovary tissue cytoplasmic and nuclear fractions. Left. Coomassie-stained SDS/PAGE gel of ovary cytoplasmic (lane 1) or nuclear (lane 2) fractions display different banding patterns. Right. anti-H3 histone reacts against the nuclear fraction (lane 2) and not the cytoplasmic fraction (lane 1). Expected H3 size = 17 kDa.

4.2.5 Ae. aegypti *Piwi4 expresses a nuclear localization signal in the N-terminal region of the protein*

To further explore AePiwi4 nuclear localization, we identified a putative NLS in the N-terminal region of the protein (Figure 4.11 A). In *Drosophila melanogaster* Piwi, the NLS is expressed in the intrinsically disordered domain in a similar region of the N-terminal. We therefore generated a phylogenetic tree of the intrinsically disordered regions of the *Ae. aegypti* and *D. melanogaster* Piwi proteins to see if Piwis with known or putative NLS signals would cluster together (Figure 4.11 B). We found that the intrinsically disordered domains of *Ae. aegypti* Ago3 and *D. melanogaster* Ago3 clustered together, that *Drosophila* Aub and *Ae.*

aegypti Piwi7 clustered together, and that *Ae. aegypti* Piwis2-6 clustered with *Drosophila* Piwi, the only *Drosophila* Piwi with an NLS. These results suggested that *Ae. aegypti* Piwi2-6 may also harbor nuclear localization signals in their intrinsically disordered domains in the N-terminal regions.

A.



B.

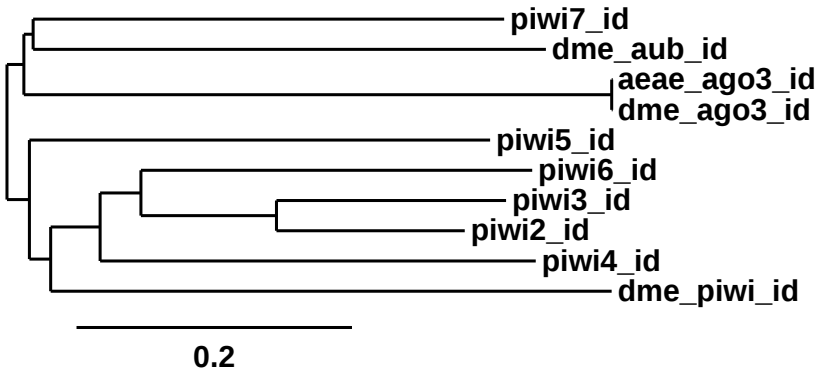
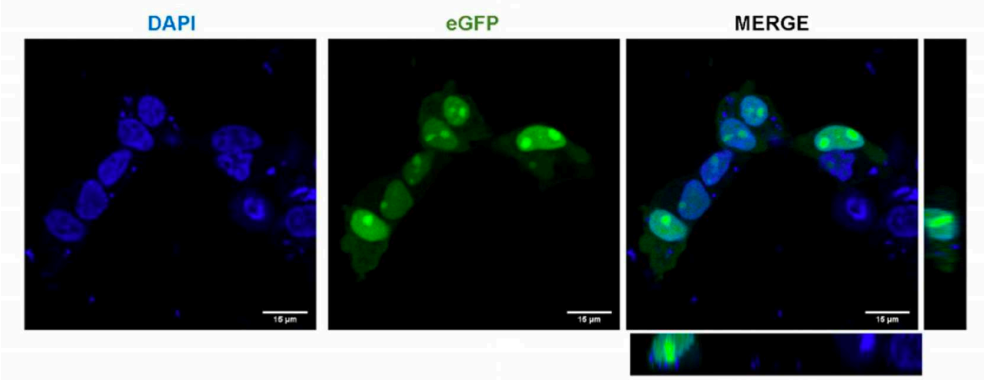


Figure 4.11. Putative N-terminal NLS present in AePiwi4. A. Putative AePiwi4 NLS relative to a threshold predicted by NLStradamus (bottom) [299]. **B.** Phylogenetic tree comparing the intrinsically disordered (id) domains of *Drosophila melanogaster* (dme) and *Aedes aegypti* (aeae) Piwi proteins. Scale bar indicates number of substitutions per site.

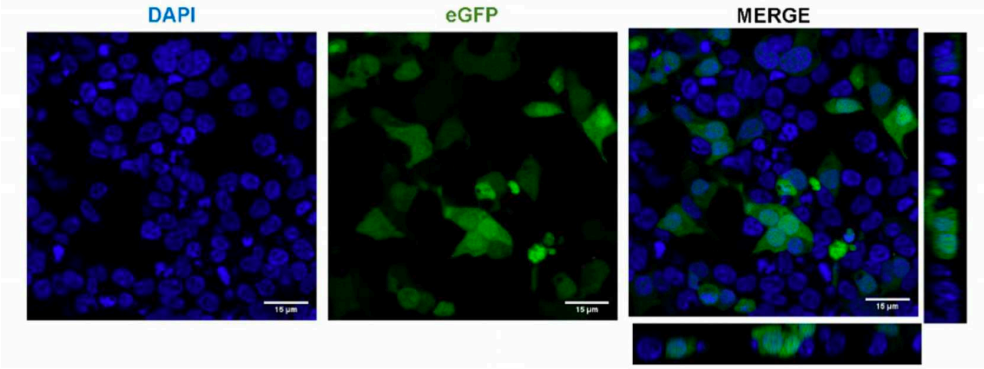
To confirm that the putative AePiwi4 NLS was responsible for protein nuclear localization, we cloned the putative AePiwi4 NLS (amino acid residues 42–83, Figure 4.11 A), as well as the entire N-terminal region containing the NLS (amino acid residues 1–83), fused to an eGFP; we henceforth named these constructs AePiwi4NLS-eGFP and AePiwi4Nterminal-eGFP, respectively. We used the same backbone containing either a known SV40 NLS fused to

the eGFP [244] as a positive control or an eGFP alone as a negative control; we henceforth named these constructs SV40NLS-eGFP and eGFP, respectively. We then transfected HEK293 cells with these constructs and visualized eGFP and DAPI colocalization 24 h post-transfection. As expected, we found that the known SV40NLS-eGFP localized in the nucleus while the eGFP alone appeared diffused throughout the cells (Figure 4.12 A, B). Plasmids harboring either AePiwi4NLS-eGFP or AePiwi4Nterminal-eGFP migrated into the nucleus, as evidenced by eGFP expression colocalized with DAPI staining (Figure 4.12 C, D). The nuclear staining appeared punctated, perhaps indicative of nucleolar staining. We also observed that both AePiwi4NLS-eGFP and AePiwi4Nterminal-eGFP displayed cytoplasmic eGFP expression as well, which was not observed in cells transfected with the SV40 NLS plasmid (Figure 4.12 C, D).

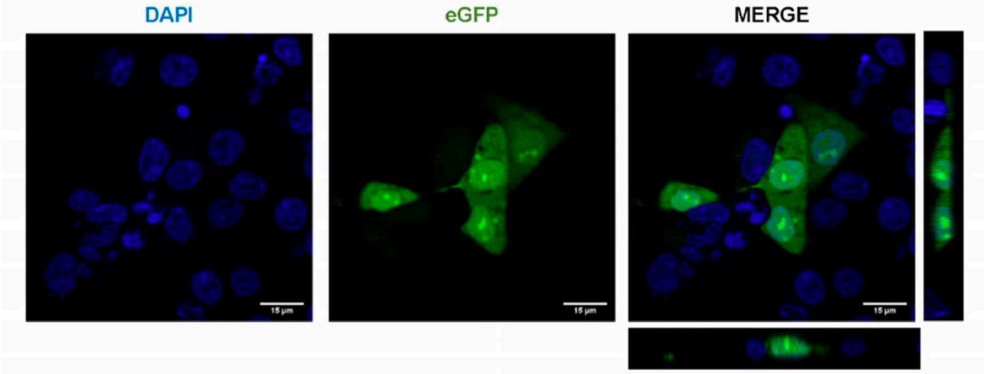
To quantitively compare eGFP fluorescent intensities across sample types, we subtracted total eGFP fluorescent intensity sums from eGFP fluorescent intensity sums in nuclear surfaces, normalized by number of cells, in three independent views across slides (Figure 4.12 E). We found that the resulting eGFP nuclear intensity sums outside of the nuclear surfaces was significantly higher for cells transfected with the eGFP construct as compared to cells transfected with SV40NLS-eGFP ($p = 0.006$), AePiwi4NLS-eGFP ($p = 0.01$), or AePiwi4Nterminal-eGFP ($p = 0.05$). There were no significant differences in eGFP fluorescent intensity sums outside of nuclear surfaces between cells transfected with AePiwi4NLS-eGFP ($p = 0.34$) or AePiwi4Nterminal-eGFP ($p = 0.23$) compared to those transfected with the SV40NLS-eGFP positive control. Taken together, these results suggested that *Ae. aegypti* Piwi4 expresses an NLS in the intrinsically disordered domain in the N-terminal region of the protein.


DAPI
eGFP
MERGE

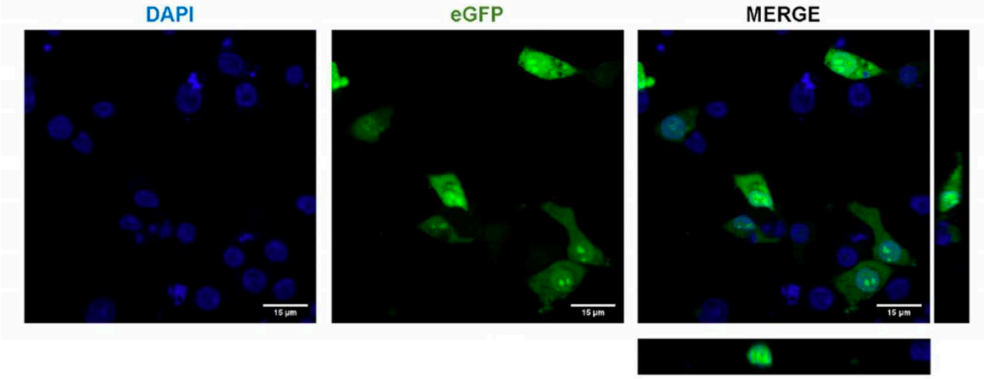
(**A**)


DAPI
eGFP
MERGE

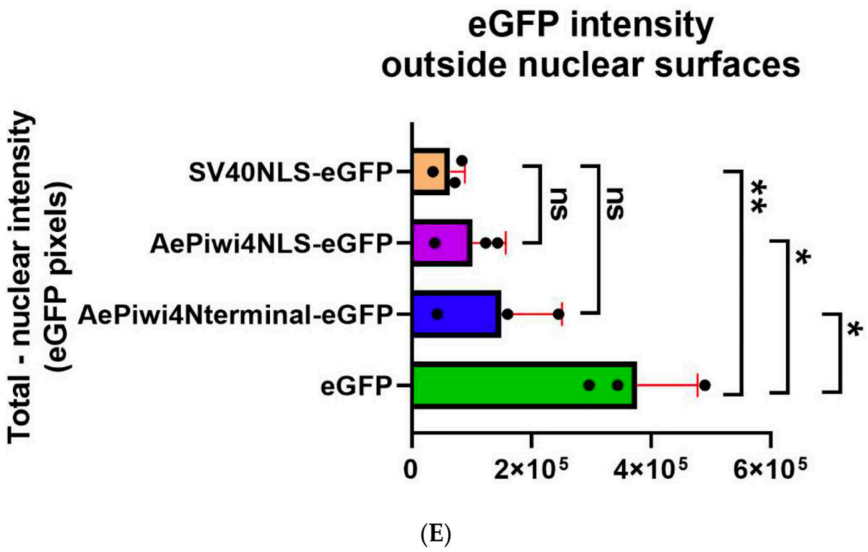
(**B**)


DAPI
eGFP
MERGE

(**C**)


DAPI
eGFP
MERGE

(**D**)



(E)

Figure 4.12. AePiwi4 harbors an NLS in the N-terminal region of the protein. Representative slices from image stacks of HEK293 cells transfected with (**A**) SV40NLS-eGFP, (**B**) eGFP, (**C**) AePiwi4NLS-eGFP, or (**D**) AePiwi4Nterminal-eGFP. Cells were stained with DAPI (blue) 24 h post-transfection. DAPI (left), eGFP (middle), and merged (right) channels are shown separately. Orthogonal views presented in merged channel. Scale bar = 15 μM. (**E**) Quantification of total eGFP fluorescence intensity sums subtracted from eGFP intensity sums in nuclear surfaces for each sample type, normalized by number of cells. Each black dot represents an individual picture. Red bars indicate SEM. ** = $p \leq 0.01$, * = $p \leq 0.05$, ns = not significant.

4.3 Discussion

In this study, we characterized *Ae. aegypti* Piwi4 structural features involved in RNA binding and nuclear localization to gain insights into the protein's function. AePiwi4 had previously been associated with various 28–30 nt 3′ 2′ O-methylated piRNAs, so we focused our efforts on the PAZ domain that binds the 3′ ends of piRNAs. We assessed AePiwi4 PAZ RNA-binding dynamics by SPR and found that AePiwi4 PAZ bound to both mature and unmethylated piRNAs with micromolar affinities in a sequence independent manner. We identified key residues in AePiwi4 PAZ involved in RNA binding and found that they were highly conserved across organisms. We also highlighted a unique arginine amino acid flanking a tyrosine residue necessary for 3′nm RNA binding that was present in most other organisms' Piwi PAZ but was

only present in a single *Ae. aegypti* Piwi PAZ (Ago3). Mutating this residue in AePiwi4 PAZ to match that of Ago3 improved both 3′m and 3′nm RNA binding. Through circular dichroism, we showed that single amino acid changes in Piwi PAZ changes the secondary structure of the protein. Finally, we found that AePiwi4 was both cytoplasmic and nuclear in mosquito tissues, and that signals in the intrinsically disordered region drove nuclear localization.

We report herein Piwi-RNA-binding affinities for a Piwi protein of an arthropod vector that harbors an expanded repertoire of Piwis, which complements studies performed with canonical human and *Drosophila melanogaster* Piwi PAZ. We found that *Ae. aegypti* Piwi4 PAZ bound 3′ 2′ O-methylated and non-methylated piRNAs with K_Ds of 1.7 ± 0.8 μM and 5.0 ± 2.2 μM, respectively. The preference of AePiwi4 PAZ for 3′m piRNAs over 3′nm piRNAs was less pronounced than what has been reported for other Piwi PAZ (Table 4.3). For example, Hiwi1, Hiwi2, and Hili bound 3′m piRNAs with K_Ds of 6.5 μM, 2 μM, or 10 μM, respectively, but they bound non-methylated piRNAs with weaker affinities—K_Ds of 16 μM, 12 μM, or 34 μM, respectively [163].

Table 4.3. Summary of known Piwi PAZ binding affinities for 3'-terminal 2'-O-methylated (3' 2' O-met) or non-methylated 3'-terminal (3' 2' OH) piRNAs.

		Ligand K_D binding affinity (μM)		
Organism	**Piwi PAZ binding partner**	**3' 2' O-met**	**3' 2' OH**	**Reference**
Drosophila	Piwi	4	unreported	Yamaguchi et al., 2020 [189]
Human	Hiwi1	6.5	16	Tian et al., 2011 [163]
Human	Hiwi2	2	12	Tian et al., 2011 [163]
Human	Hili	10	34	Tian et al., 2011 [163]

Immunoprecipitations of AePiwi4 from uninfected or infected Aag2 cells followed by sRNA sequencing of associated RNAs have revealed the protein associates with *bona fide* piRNAs resistant to beta-elimination, a method that selects for 3′m piRNAs and depletes 3′nm

miRNAs [127,131]. Our results suggest, however, that AePiwi4 can bind to both 3′m and 3′nm sRNAs with only a marginally higher affinity for the former over the latter. Further investigations on AePiwi4-associated sRNAs from different cellular compartments may provide new insights on protein-RNA trafficking and the range of sRNAs with which AePiwi4 interacts. For example, the role(s) of *Ae. aegypti* Piwis may function with both pre-processed non-methylated RNAs and mature piRNAs across cellular compartments or with sRNA populations outside the piRNA pathway. Halbach et al. (2020) compared AePiwi4-mediated silencing of a satellite repeat-derived target by way of a piRNA to that of miRNA silencing [131]. In that study, the authors found that the 3′ end of a satellite repeat-derived piRNA (tapR1) was not absolutely required for silencing, while the seed region was not sufficient for silencing, a pattern they compared to miRNA-mediated silencing [131]. In another study, Tassetto and colleagues found that silencing *AePiwi4* impacted both 3′m piRNA and siRNA production and argued that AePiwi4 links the siRNA and piRNA pathways [127]. Our AePiwi4 RNA-binding studies indicate that the PAZ domain of AePiwi4 is indeed able to interact with diverse populations of sRNAs with similar affinities, perhaps suggesting AePiwi4 has broad functions or unique roles in RNA binding that may differ from model Piwis. Future studies comparing RNA-binding dynamics across the *Ae. aegypti* Piwis will elucidate the roles they play in RNAi.

In this study, we compared 3′m and 3′nm RNA binding with protein partners by SPR. Other studies have characterized PAZ RNA binding by ITC using small eight nucleotide RNAs, and we note that caution should be taken when comparing hard dissociation constant values across these different techniques. While ITC provides valuable information on number of binding sites and heat released from a binding reaction, we found that immobilizing the 5′ end of longer, more physiologically relevant RNAs by SPR enabled us to efficiently calculate

dissociation constants for many protein-binding partners against stabilized ligands in a single experiment. This method may be useful for other studies aimed at understanding protein-RNA binding at specific motifs.

Our data suggest that small differences in Piwi PAZ amino acid composition across Piwi proteins alter protein secondary structure, which thereby impact the protein's affinity for certain RNAs. Given that preferential binding of 3′m RNA over 3′nm RNA was mostly dictated by backbone confirmation of the protein in Hiwi1, it is likely that subtle differences in Piwi protein structure may have profound impacts on preferential RNA-binding behaviors that are important for defining the functions of different Piwis. Although the amino acids that directly form hydrogen bonds with RNA were highly conserved, residues that flank these sites tended to be more variable. Perhaps those residues that impact the stability of the PAZ structure, rather than those that directly bind the RNA itself, drive Piwi functional divergence.

The number of Piwi proteins has expanded in culicine mosquitoes as compared to anophelines and drosophilids [301], and understanding their evolutionary relationships with other Piwis, their shared or unique structural features, and their interactions with diverse RNA populations may provide insights into how their functions have diverged. We found that the PAZ domains of *Ae. aegypti* Piwis 2-7 are more evolutionarily related to that of *D. melanogaster* Aubergine than to *D. melanogaster* Piwi or Ago3 PAZ. Aubergine is a cytoplasmic, germline-specific protein that participates in ping-pong amplification by binding antisense primary piRNAs and producing secondary sense piRNAs that fuel the cycle [302], a role similar to that of *Ae. aegypti* Piwi5 [118]. A recent study showed that piRNA binding to Aub PAZ induces a protein confirmational change that triggers symmetric dimethylarginine (sDMA) methylation of the Aub intrinsically disordered domain in the N-terminal region [303].

The sDMA modification then serves as a binding site for Krimper, which simultaneously binds unmethylated Ago3 to bring the proteins in close proximity for RNA transfer during ping-pong amplification [303]. Indeed, Joosten and colleagues recently characterized a Krimper ortholog in *Aedes*, Atari, which bound Ago3 without sDMA modifications [158]. Perhaps a similar mode of RNA binding-dependent autoregulation and sDMA signaling also govern the *Aedes aegypti* Piwis Piwi2-6.

Drosophila Piwi, the only nuclear Piwi in the fly, expresses a bipartite NLS in its intrinsically disordered domain [304]. *Drosophila* Piwi nuclear localization is autoregulated by confirmational changes that occur once the protein binds piRNAs; the NLS remains buried within the protein structure until RNA binding triggers a conformational change and exposes the NLS [304]. Once the protein is imported into the nucleus and releases the piRNA, the protein is trafficked back into the cytoplasm. We found that *Ae. aegypti* Piwi4 also expresses signals in the intrinsically disordered region that drive proteins to the nucleus (Figure 4.12 C, D), which, if similar to *Drosophila* Piwi, could autoregulate protein trafficking based on protein confirmational changes that dictate signal exposure. We observed that the AePiwi4 NLS did not drive complete expression of eGFP into the nucleus, as evidenced by diffused cytoplasmic fluorescence in addition to punctated nuclear staining (Figure 4.12 C, D). It is possible that the intrinsically disordered region of AePiwi4 contains both an NLS and nuclear export signals (NES) that drive protein trafficking in and out of the nucleus. Investigations on how AePiwi4 regulates its trafficking into different cellular compartments require future studies and could be useful in understanding its role in different RNAi-mediated processes. It is also possible that several *Ae. aegypti* Piwis autoregulate their subcellular localization in similar manners as AePiwi4. Our phylogenetic analyses revealed that the regions of the AePiwi4 and *Drosophila*

melanogaster Piwi proteins that harbored nuclear localization signals, the intrinsically disordered domains, also clustered with *A. aegypti* Piwi2, Piwi3, Piwi5, and Piwi6 (Figure 4.11 B). Different Piwis likely have sophisticated and diverse regulation mechanisms that control their expression patterns in different compartments of the cell.

Growing evidence reveals that the piRNA pathway is involved in gene regulation in somatic tissues and contributes to diverse human diseases including cancer [156,305] and neurodegenerative disorders [306]. Because somatic piRNAs and Piwi expression are common in arthropods [157], they could be valuable models for understanding the molecular mechanisms underlying the lesser understood Piwi or piRNA functions. Future studies aimed at understanding how Piwi-RNA binding impacts protein structure and function will be useful for learning more about how this pathway is involved in immunity, gene regulation, and disease in arthropod vectors as well as in other organisms.

4.4 Materials and Methods

4.4.1 Ae. aegypti *Piwi4 structure model prediction*

A model of the *Ae. aegypti* Piwi4 structure was generated using the I-TASSER software (version 5.1 Zhang Lab, University of Michigan, Ann Arbor, MI, USA [295,307,308] and visualized using Chimera (University of California, San Francisco, CA, USA). This software predicts secondary and tertiary structures based on the similarity of other proteins whose structures have been solved. The AePiwi4 amino acid sequence was queried against the *Drosophila* Piwi structure that was crystalized in [189], which allowed us to determine the predicted *Ae. aegypti* Piwi4 PAZ domain. AePiwi4 PAZ was then superimposed to other

crystalized PAZ proteins, including Hili PAZ [163], Hiwi1 PAZ [163], Miwi PAZ [296], and Siwi PAZ [190].

4.4.2 Cloning

Aedes aegypti Piwi4 (AAEL007698), including a 6xHis-tag, was synthesized by BioBasic Inc. (Markham, ON, Canada). Both *AePiwi4* full length (FL) and PAZ domain (residues 270–380) nucleotide sequences were sub-cloned into pCR-Blunt II-TOPO vector using the Zero Blunt TOPO PCR Cloning Kit (Thermo Fisher Scientific, Waltham, MA, USA) following the manufacturer's instructions. pCR-Blunt II-TOPO vectors containing either *AePiwi4* FL or *AePiwi4* PAZ were transformed in OneShot Top10 chemically competent *E. coli* (Invitrogen, Waltham, MA, USA). Using standard restriction enzyme-mediated cloning and the pCR-Blunt II-TOPO vectors described above as PCR templates, *AePiwi4* FL or *AePiwi4* PAZ were then cloned into pET-17b vectors. pET-17b vectors containing either *AePiwi4* FL or *AePiwi4* PAZ were transformed in OneShot Top10 chemically competent *E. coli* (Invitrogen, Waltham, MA, USA). Inserts were confirmed by Sanger sequencing. Primers used in this study are displayed in Supplemental Table 4.1.

The putative AePiwi4 NLS, as well as the entire N-terminal region of AePiwi4 containing the putative NLS, were cloned from the *AePiwi4* FL-containing pCR-Blunt II-TOPO vector into a backbone containing an eGFP by In-Fusion cloning (TakaRa, San Jose, CA, USA) following the manufacturer's instructions. The parent SV40NLS-eGFP backbone was a gift from Rob Parton (Addgene plasmid # 67652; **http://n2t.net/addgene:67652** (accessed on 28 April 2021); RRID: Addgene_67652) [300]. Briefly, an eGFP-alone plasmid was generated by NcoI digestion and religation with the T4 DNA ligation Mighty Mix (TakaRa, San Jose, CA, USA) of the SV40NLS-eGFP plasmid. AePiwi4NLS-eGFP or AePiwi4Nterminal-eGFP were then cloned

into the eGFP alone plasmid using In-Fusion primers listed in Supplemental Table 4.1. All constructs were transformed into Stellar Competent Cells (TakaRa, San Jose, CA, USA), and inserts were confirmed by Sanger sequencing.

4.4.3 Recombinant protein expression

pET-17b vectors containing either *AePiwi4* FL or PAZ were transformed into BL21(DE3) pLysS chemically competent *E. coli* cells (Thermo Fisher Scientific, Waltham, MA, USA). Transformed *E. coli* were plated onto Luria–Bertani (LB) agar plates with 100 µg/mL ampicillin and 34 µg/mL chloramphenicol that were left O/N at 37 °C. Individual colonies were picked into starter cultures of 4 mL LB broth (supplemented with 100 µg/mL ampicillin and 34 µg/mL chloramphenicol) that were left shaking at 220 RPM O/N at 37 °C. Starter cultures were then added to 150 mL LB broth (supplemented with 100 µg/mL ampicillin and 34 µg/mL chloramphenicol) that were left shaking at 220 RPM 37 °C until OD_{600} = 0.4 (~1 h). Protein expression was then induced with 0.1 mM isopropyl β- d-1-thiogalactopyranoside (IPTG) for 4 h shaking at 160 RPM at 25 °C. Bacteria was pelleted and stored at −30 °C until protein purification.

For larger scale expression, 150 mL LB broth (supplemented with 100 µg/mL ampicillin and 34 µg/mL chloramphenicol) starter cultures that had been inoculated with glycerol scrapings of BL21(DE3) pLysS *E. coli* containing either pET-17b-*AePiwi4* FL or pET-17b-*AePiwi4* PAZ were left shaking at 220 RPM O/N at 37 °C. Starter cultures were then added to 1 L LB broth (supplemented with 100 µg/mL ampicillin and 34 µg/mL chloramphenicol) and expression was induced following the above protocol.

Expression was confirmed by SDS-PAGE separation and anti-6xHis-tag Western blot in both the soluble and inclusion body fractions for both proteins.

4.4.4 Recombinant protein purification

The soluble AePiwi4 PAZ protein was purified by affinity chromatography followed by size-exclusion chromatography using Nickel-charged HiTrap Chelating HP (GE Healthcare, Chicago, IL, USA) and Superdex 200 10/300 GL columns (GE Healthcare, Chicago, IL, USA), respectively. Frozen *E. coli* pellets were resuspended with Buffer A (10 mM Tris, 500 mM NaCl, 5 mM imidazole, pH 8), left on ice for 10–15 min, and pulse sonicated 4 x for 30 s—2 min. The lysates were then spun at 15,000× *g* for 30 min at 4 °C. The resulting supernatants were filtered with a 0.8 µM filter (MilliporeSigma, Burlington, MA, USA) and loaded onto a pre-equilibrated Nickel-charged HiTrap Chelating HP column using a peristaltic pump. The column was then pre-washed with 3 column volumes (CV) of Buffer A, followed by a 3 CV wash with Wash Buffer 1 (10 mM Tris, 500 mM NaCl, 20 mM imidazole, pH 8) and 3 CV wash with Wash Buffer 2 (10 mM Tris, 500 mM NaCl, 100 mM imidazole, pH 8). The protein was eluted from the column using 3 CV of Elution Buffer (10 mM Tris, 500 mM NaCl, 300 mM imidazole, and pH 8) and visualized by SDS-PAGE gel electrophoresis.

Eluted protein was concentrated down to ~500 µL using an Amicon stirred cell with a cellulose membrane of 3 kDa nominal molecular weight (MilliporeSigma, Burlington, MA, USA). The resulting concentrated protein was spun down at 4000× *g* for 10 min to remove large debris and loaded onto a Superdex 200 10/300 GL column that had been equilibrated with 20 mM Tris-HCl, 150 mM NaCl, pH 7.4. Peak elutions that corresponded to the correct size of Piwi4 PAZ (14 kDa) were confirmed by SDS-PAGE gel electrophoresis. N-terminal protein sequence was also confirmed by Edman degradation.

4.4.5 SDS-PAGE

All proteins were heated to 95 °C for 5 min under reducing conditions in 1X LDS (Thermo Fisher Scientific, Waltham, MA, USA) and were separated using 4–12% Bis-Tris protein gels (Thermo Fisher Scientific, Waltham, MA, USA). Gels were stained with Coomassie Brilliant Blue (GenScript, Piscataway, NJ, USA). Protein concentrations were determined using the Nanodrop ND-1000 spectrophotometer adjusted by the molar extinction coefficient.

4.4.6 Western blot

Aedes aegypti mosquito midguts and ovaries, as well as recombinant proteins, were separated by SDS-PAGE gel electrophoresis for Western blots. *Ae. aegypti* that had been fed defibrinated sheep blood (Denver Serum Company, Denver, CO, USA) were collected 48 h post-bloodmeal, and their midguts (cleaned of blood in 1X PBS) and ovaries were dissected and flash frozen on dry ice. 15 midguts or ovaries / tube were resuspended in 100 µL cold hypotonic lysis buffer (10 mM Hepes pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, 0.2 mM PMSF) and left on ice for 15 min. Samples were vortexed vigorously for 30 s and then pelleted at 1000× *g* for 15 min. The supernatant was collected as the cytoplasmic fraction. The remaining pellets were then resuspended in 100 µL solubilization buffer (15 mM Tris, 150 mM NaCl, 5 mM EDTA, 0.5% Triton X-100, 10% glycerol, 0.2 mM PMSF) and spun down at 100,000× *g*. The supernatant was collected as the nuclear fraction.

30 µg of protein was processed for SDS/PAGE separation, as described previously, and run alongside 10 µM AePiwi4 PAZ or AePiwi4 FL inclusion bodies. Proteins were transferred to a PVDF membrane (iBlot, Invitrogen, Waltham, MA, USA) that was blocked for 2 h at RT in blocking buffer (5% powdered milk (Carnation), 50 mM Tris-HCl pH 7.4, 150 mM NaCl, 1% Tween 20 (TBST)). Membranes were incubated O/N at 4 °C with anti-Piwi4 PAZ mouse serum (1:500 in blocking buffer), a 6xHis-tag monoclonal antibody (ThermoFisher Scientific,

Waltham, MA, USA, diluted 1:5000 in blocking buffer), or anti-Histone H3 as a nuclear marker (Novus Biologicals, Littleton, CO, USA; generated in rabbit, diluted 1:1000 in blocking buffer). Membranes were washed with TBST (3× for 10 min) and with TBS (1× 10 min) and incubated at RT for 1–2 h with goat anti-mouse or anti-rabbit antibodies conjugated to alkaline phosphatase (1 mg/mL, diluted 1:10,000). Membranes were again washed with TBST and TBS, and proteins were detected for 5–10 min using Western Blue Stabilized alkaline phosphatase substrate (Promega, Madison, WI, USA).

4.4.7 Mosquito rearing

Ae. aegypti mosquitoes (Liverpool (LVP) strain) were reared in standard insectary conditions at the Laboratory of Malaria and Vector Research, NIAID, NIH (28 °C, 60–70% humidity, 14:10 h light/dark cycle) under the expert supervision of Karina Sewell, Andre Laughinghouse, Kevin Lee, and Yonas Gebremicale. Mosquitoes had a solution of 10% sucrose *ad libitum* and were offered defibrinated sheep blood (Denver Serum Company, Denver, CO, USA) in an artificial feeding system. Larva were fed Tetramin.

4.4.8 Sequence alignment

Nucleotide and amino acid sequences were retrieved from the NCBI databases. Multiple alignments and phylogenetic trees were obtained by Clustal Omega [309] and visualized on Jalview [310].

4.4.9 Site-directed mutagenesis

AePiwi4 PAZ protein mutants were generated using the QuikChange II Site-Directed mutagenesis kit (Agilent, Santa, Clara, CA, USA) following the manufacturer's instructions. Primers were designed using PrimerX (**https://www.bioinformatics.org/primerx/cgi-bin/DNA_1.cgi**, [accessed on 14 March 2021]) and are displayed in Supplemental Table 4.1.

The pET-17b vector containing *AePiwi4* FL was used as template for the reactions, and XL1-Blue supercompetent cells (Agilent, Santa, Clara, CA, USA) were transformed with the mutant plasmids. Mutation nucleotide sequences were confirmed by Sanger sequencing, and protein mutant sequences were confirmed by mass spectrometry.

The pET-17b vectors containing the AePiwi4 PAZ mutations were transformed into BL21(DE3) pLysS chemically competent *E. coli* cells (Thermo Fisher Scientific, Waltham, MA, USA), and all proteins were expressed and purified, as described previously.

4.4.10 Surface plasmon resonance (SPR)

All SPR experiments were carried out in a T100 instrument (GE Healthcare, Chicago, IL, USA) following the manufacturer's instructions. Sensor CM5, amine coupling reagents, and HBS-P buffers were also purchased from GE Healthcare (Chicago, IL, USA). HBS-P was supplemented with EDTA (HBS-PE, 10 mM Hepes pH 7.4, 150 mM NaCl, 3 mM EDTA, and 0.005% (*v*/*v*) P20 surfactant) and was used as the running buffer while Conditioning Solution 2 (50 mM NaOH, 1 M NaCl) was used as the regeneration and conditioning solution for all experiments. Briefly, the CM5 sensor was coated 40 µg/mL neutravidin and pre-conditioned with 3 × 60 s injections of Conditioning Solution 2. ~500–1000 RUs of biotinylated RNAs were then captured to flow cells 2 or 4, which were then conditioned with 3 × 60 s injections of Conditioning Solution 2. Protein analyte was introduced unto the surface with 180 s injections (30 µL/s). Results were analyzed using the Biacore T200 Evaluation software v2.0.3 provided by GE Healthcare (Chicago, IL, USA). Equilibrium dissociation constants were calculated from steady-state binding levels (R_{eq}) against molar concentration of the analyte (C). The fitted equation was $R_{eq} = ((CR_{max})/(K_D + C)) + \text{offset}$, where R_{max} = analyte-binding capacity of the surface in response units (RU) and offset = response at zero analyte concentration, which

accounts for buffer-mediated effects on the refractive index. SPR experiments were carried out 2–4 x.

4.4.11 Circular dichroism

0.1 mg/mL of purified AePiwi4 PAZ WT, T39A, Y40A, T41A, or T41R in 20 mM Tris 75 mM, NaCl pH 7.4 were used for CD analyses. Continuous measurements with a pitch of 0.2 nm were recorded from 200–260 nm wavelengths with a bandwidth of 1 nm. Mean residue ellipticity was calculated with the following equation: (molecular weight of each protein in daltons/((number of amino acids − 1) × θ_λ))/(10 × pathlength in cm × protein concentration in g/mL). All readings were normalized by subtracting with blank (buffer) mean residue ellipticity. Data were analyzed using CAPITO [298].

4.4.12 RNA synthesis

The 3′ 2′ O-CH_3 and 3′ 2′ OH 28 nt RNAs were synthesized by Eurofins Genomics (Louisville, KY, USA). Sequences are listed in Supplemental Table 4.2. RNA was resuspended at 1–2 mM in DEPC-treated water and stored at −80 °C.

4.4.13 Mouse polyclonal antibody production

Polyclonal antibodies against *A. aegypti* Piwi4 PAZ were raised in mice. Mice (Balb/c; Charles River, Frederick, MD, USA) were IM immunized with 10 µg of AePiwi4 PAZ in combination with Magic Mouse Adjuvant (CD Creative Diagnostics, Shirley, NY, USA). Negative control mice were immunized with Magic Mouse Adjuvant alone. At 21 d post-immunization, mice received a 2nd booster immunization with 10 µg of AePiwi4 PAZ in combination with Magic Mouse Adjuvant (or adjuvant alone for negative control group). Blood was collected 35 d post-immunization. The antibody levels were confirmed by ELISA.

4.4.14 Mass spectrometry

Mosquito tissue samples and recombinant proteins were prepared and separated by an SDS-PAGE gel as previously described, which was then stained with Coomassie blue. Bands of interest were excised from the gel and submitted for liquid chromatography coupled with mass spectrometry at the Research and Technology Branch (NIAID, NIH, Rockville, MD, USA). Briefly, the gel slices from the SDS-PAGE gel were cut into small pieces and subjected to in-gel trypsin digestion. The gel slices were destained to remove Coomassie blue staining and were then reduced and alkylated. After dehydration with acetonitrile and air-drying, a sequencing grade trypsin (Promega, Madison, WI, USA) solution was added onto the gel slices and was allowed to be absorbed into the gel slice. The gel slices were then incubated overnight at 30 °C for in-gel digestion. The peptides released from in-gel digestion were extracted by acetonitrile and then applied for LC-MS/MS analysis. Proteomic analyses were performed, as previously described [311].

4.4.15 IFA

Mosquito midguts or ovaries were dissected 48 h post-bloodmeal, flash fixed for 30 s in cold 4% paraformaldehyde (PFA) in PBS and cleaned of blood in cold PBS. The midguts or ovaries were then left shaking in 4% PFA in PBS O/N at 4 °C. The next day, midguts or ovaries were washed 3X in PBS and blocked O/N in blocking buffer (2% BSA, 0.5% Triton-X-100, PBS). The midguts or ovaries were then incubated with either serum from mice immunized with AePiwi4 PAZ or with Magic Mouse adjuvant alone (1:500, diluted in blocking buffer) O/N at 4 °C. The midguts or ovaries were washed with blocking buffer a minimum of 3× for 30 min and were then incubated with secondary goat anti-mouse antibodies conjugated to Alexa Fluor 594 (Thermo Fisher Scientific, Waltham, MA, USA; diluted 1:1000 in blocking buffer) for 1 h at RT. The midguts or ovaries were again washed with blocking buffer a minimum of 3× for 30 min,

followed by incubation with 1 μg/mL DAPI (diluted 1:1000 in blocking buffer) and phalloidin conjugated to Alexa Fluor 488 (Thermo Fisher Scientific, Waltham, MA, USA; diluted 1:250 in blocking buffer) for 20 min at RT. The midguts or ovaries were washed 2× for 20 min with blocking buffer and 1× with 0.5% Triton-X-100 in PBS and were then mounted onto slides with ProLong Gold antifade mountant with DAPI (Thermo Fisher Scientific, Waltham, MA, USA).

4.4.16 HEK293 cell culture and transfection

HEK293 cells were cultured in 35 mm dishes with a No. 15 coverslip pre-coated with Poly-D-Lysine (MatTek Life Sciences, Ashland, MA, USA). Briefly 300,000 cells were plated on individual dishes in DMEM media. The next day, the cells were transfected with transfection complex containing 500 ng of (1) SV40NLS-eGFP, (2) eGFP alone, (3) AePiwi4NLS-eGFP, or (4) AePiwi4Nterminal-eGFP in 0.5 μL Lipofectamine 3000 (Invitrogen, Waltham, MA, USA) in serum free Opti-MEM media, according to the manufacturer's protocol. Twenty-four hours post-transfection, the cells were washed 3X with PBS and fixed with 4% PFA in PBS for 30 min at RT. The cells were then washed 3X with PBS and permeabilized with 0.5% Triton-X-100 in PBS for 30 min at RT. The cells were then stained with DAPI (1 μg/mL in 2% BSA, 0.5% Triton-X-100, PBS). The cells were visualized using a Leica Confocal SP8 microscope. Images were processed with Imaris software version 9.2.1 and post-processing was carried out in Fiji ImageJ version 1.52n for representative purposes.

4.4.17 Statistics

Surface Plasmon Resonance equilibrium curves were fitted with a non-linear regression generated by the Biacore Evaluation software v2.0.3 provided by GE Healthcare (described in "Surface Plasmon Resonance" Methods section), which were then visualized with GraphPad Prism. The equilibrium dissociation constants, calculated based on steady state, were generated

by that same software. Differences between dissociation constants were compared using an unpaired two-tailed t-test with GraphPad Prism.

Quantifications of eGFP fluorescent intensities were calculated by subtracting eGFP pixel total intensity sums by average nuclear intensity sums, normalized by the number of cells, in three independent views across a slide. Nuclear surfaces were determined by DAPI display and eGFP pixel intensity values were extracted using Imaris software version 9.2.1. Differences between eGFP intensities outside of nuclear surfaces were compared using an unpaired two-tailed T-test with GraphPad Prism.

CHAPTER 5: SUMMARY AND FUTURE DIRECTIONS

5.1 Summary of major findings

Collectively, the projects described herein explored molecular mechanisms underlying two major antiviral RNAi pathways in *Ae. aegypti*: the siRNA and piRNA pathways. In our first study, we synthetically triggered the siRNA pathway against ZIKV in transgenic mosquitoes that expressed a ZIKV-specific dsRNA in the midgut after a bloodmeal. The transgenic mosquitoes were 83-90% resistant to the virus [205]. Through a similar strategy, our lab had previously generated DENV2-resistant *Ae. aegypti* that have been nearly 100% resistant to DENV2 for more than 50 generations [191,193]. Together, these works showed that the siRNA pathway is a potent antiviral mechanism, and that it can be synthetically triggered to render mosquitoes largely resistant to arboviruses.

In our next studies, we shifted focus to the piRNA pathway, which controls insect-specific viruses but can also target arboviruses. ISVs from geographically distinct *Ae. aegypti* displayed ecotypes and overall sRNA distributions were highly variable. Infecting mosquitoes with DENV2 resulted in an increase of virus-derived siRNAs and was associated with increased sRNAs against several insect-specific viruses. Infective ISVs were diverse across mosquitoes, and their dynamics with arboviruses and RNAi appear to be incredibly complex. These observations leave unanswered questions such as (1) do virus-derived sRNAs cross talk with each other (*i.e.* does the infection of one virus impact sRNAs against another?) (2) is there a maximum siRNA/piRNA threshold (*i.e.* can a mosquito only produce a certain number of sRNAs until the RNAi system is overwhelmed?) (3) to what extent is vector competence impacted by the mosquito virome (*i.e.* do ISVs have a significant impact on arbovirus replication

dynamics in the field?) and (4) are there proteins that bridge RNAi pathways (*i.e.* do the siRNA and piRNA pathways work together to control viruses?).

Along these lines, we studied RNA binding dynamics of an antiviral Piwi protein, Piwi4, that has been postulated to link the siRNA and piRNA pathways [127]. We found that Piwi4 PAZ bound 3'-terminal 2'-O-methylated and non-methylated 3'-terminal sRNAs with similar micromolar affinities, suggesting Piwi4 has the capacity to interact with diverse RNA substrates. Native Piwi4 was both cytoplasmic and nuclear in mosquito somatic and germline tissues, and we identified a nuclear localization signal in the N-terminal region of the protein. Our *in vitro* work is further evidence that the functions of non-canonical Piwis like Piwi4 are dynamic and may cross RNAi pathways.

5.2 Ongoing and future work: genetic control strategies to mitigate arbovirus transmission[6,7]

In Chapter 2, we characterized an antiviral effector against ZIKV in transgenic mosquitoes at a previously identified locus ("Timp-P4") known to drive transgene expression. We also discussed future considerations for coupling antiviral effectors to gene drive elements for a population replacement vector control strategy. Two major challenges in building an antiviral single component gene drive system will be (1) to identify loci that allow for optimal

[6] This section includes parts of the manuscript cited as "Williams, A.E., Franz, A.W.E., Reid, W.R., Olson, K.E. Antiviral effectors and gene drive strategies for mosquito population suppression or replacement to mitigate arbovirus transmission by *Aedes aegypti*. *Insects*. 2020; *11*(1): 52." The article is reproduced with permission and minor modifications have been made.

[7] This section includes parts of the preprint cited as "Reid, W.R., Lin, J., Williams, A.E., Juncu, R., Olson, K.E., Franz, A.W.E. Genomic insertion locus and Cas9 expression in the germline affect CRISPR/Cas9-based gene drive performance in the yellow fever mosquito *Aedes aegypti*. *bioRxiv*. 2021. Sections of this article are reproduced/summarized with permission and minor modifications have been made.

expression of both the antiviral effector and the gene drive element and (2) to express Cas9 under an optimum germline promoter. Along these lines, Reid and colleagues – in collaboration with our group at CSU – generated transgenic *Ae. aegypti* that express gene drive cassettes targeting either the Carb109 locus, previously shown to allow strong expression of an anti-DENV2 IR [191,193] or the Timp-P4 locus, tested in Chapter 2 for expression of the anti-ZIKV IR transgene [205,312]. The gene drive element is ideally active in the germline during early gametogenesis, so we also tested Cas9 expression under three promoters, *nanos* [313,314], *zpg* [315,316], and *β2-tubulin* [317], previously shown to be restricted to the germline. Transgenic mosquitoes expressing Cas9 under *β2-tubulin* did not exhibit drive (Figure 5.1 A). This was also true for mosquitoes targeting the gene drive element to the Timp-P4 site when Cas9 was expressed under *nanos* (Figure 5.1 B). Lack of drive was evident because when these lines were outcrossed, ~50% of their progeny inherited the transgene cassette, which would be expected under Mendelian inheritance in the absence of drive. Transgenic mosquitoes expressing the Cas9 under either the *nanos* (Figure 5.1 C) or *zpg* (Figure 5.1 D) promoters, both targeted at the Carb109 site, did exhibit significant rates of drive. However, mosquitoes expressing Cas9 under the *nanos* promoter that showed the highest rates of drive also exhibited higher rates of gene drive blocking indel formation in larvae lacking the gene drive cassette as compared to mosquitoes that expressed the Cas9 under *zpg*. In families exhibiting rates of 86-90% gene drive inheritance, there were 7 larvae lacking the gene drive cassette that also harbored gene drive blocking indels out of 12 or 27 total larvae tested (26-58%). On the other hand, mosquitoes expressing Cas9 under *zpg* exhibited low rates of indel formation (0-7.4%). These data show that genomic target site and Cas9 promoter selection strongly influence drive efficiency, and that not all target sites that allow for strong antiviral effector expression (in the case of Timp-P4) allow for gene drive.

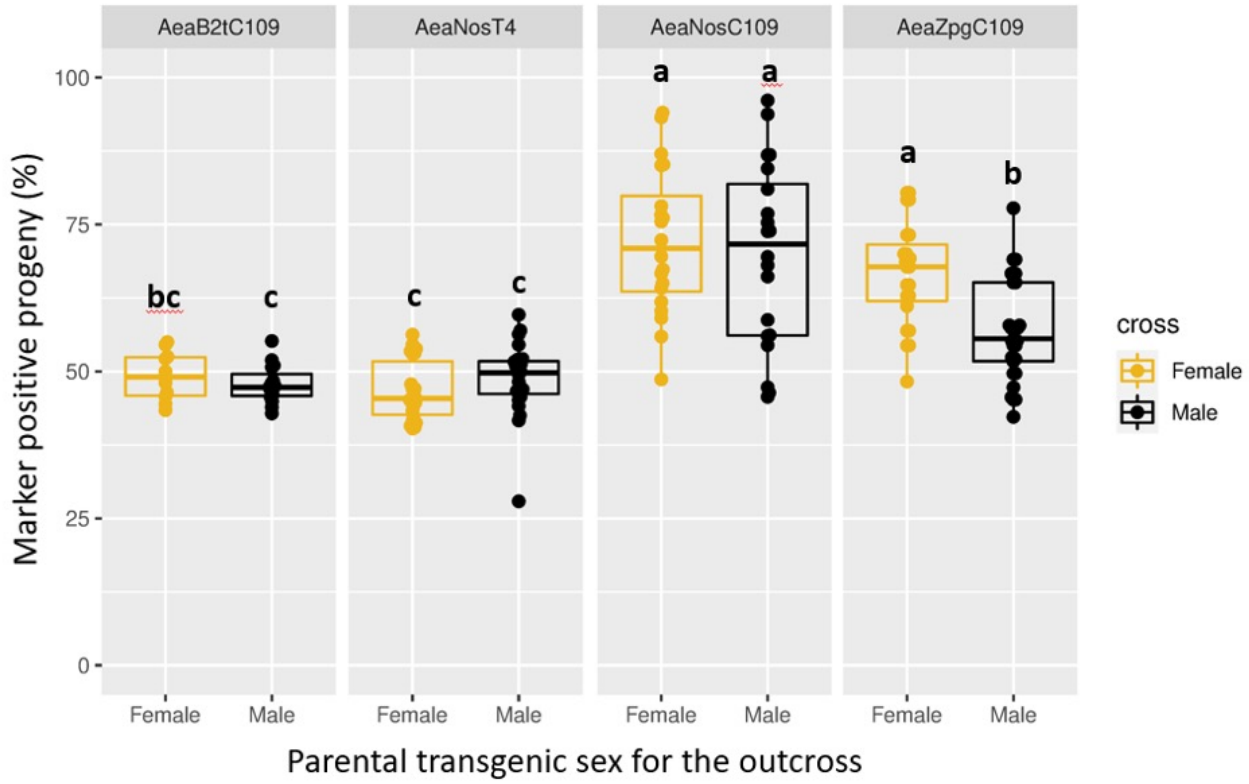


Figure 5.1. Frequency of fluorescent marker inheritance after outcrossing transgenic *Ae. aegypti* female or male parental lines expressing gene drive cassettes under different promoters targeting either Carb109 (C109) or Timp-P4 (T4). Frequency of marker inheritance after individuals from four gene drive lines were outcrossed (OX) to Higgs' white eye (HWE) strain of *Ae. aegypti*. Gene drive lines are (first panel) Aeaβ2tC109, (second panel) AeaNosT4, (third panel) AeaNosC109, and (fourth panel) AeaZpgC109 where Aea = *Aedes aegypti*; β2t = *β2-tubulin*, Nos = *nanos*, zpg = *zero population growth*, which are promoters for Cas9 expression; C109 = carb109, T4 = Timp-P4, which are the targeted loci. Each data point represents the percentage of transgene inheritance resulting from the offspring of the parental crosses where each transgenic female parental was allowed to mate to two HWE males, and each transgenic male parental was allowed to mate to two HWE females. A minimum of 20 larvae was set for each group to be scored. For female parental crosses, the population sizes of each data point ranged from 20 to 142 larvae (59 ± 3). For male parental crosses, the population sizes of each data point ranged from 21 to 415 larvae (91 ± 5 larvae). Box and whisker plots show medians and upper and lower quartiles. Yellow graphs = OX was with female HWE; black graphs = OX was with male HWE. Groups superseded with the same letter are not statistically different ($p > 0.05$).

Cage studies will be necessary to validate gene drive systems linked to anti-pathogen effectors in transgenic *Ae. aegypti* under more realistic field conditions. Such experiments will be especially useful when assessing the likelihood of resistance that may evolve against the gene

drive. Gene drive models such as MGDrive (Mosquito Gene Drive Explorer) [318,319] are also powerful tools to estimate required mosquito release numbers and time periods to achieve transgene fixation. MGDrive is versatile and can be used to estimate the efficacy of a variety of gene drive systems given a wide variety of parameters. The key components of the model include genetic inheritance (sex-specific homing efficiency and homing resistance rates), mosquito life history (fitness cost parameters measured at each life stage, including egg production, mortality risk, and longevity), and landscape (target population size and size of geographic target region). From these user defined variables, the model can estimate required release numbers with respect to sex ratios and time periods required for transgene fixation.

Before transitioning transgenic *Ae. aegypti* to field conditions outside the laboratory, there is already growing concern that such a gene drive could invade non-target insect populations or self-propagate out of control. Therefore, "confinable" gene drives have been investigated and modeled using the MGDrivE platform [233,318,319]. These systems include split drives, whereby a two-component system is established with the trans-acting Cas9 and the sgRNA segregating independently from one generation to the next. Mosquitoes expressing either (but not both) of these constructs do not exhibit gene drive; however, crossing mosquitoes that harbor the sgRNA component with those that harbor the Cas9 expression cassette would result in progeny that exhibit gene drive activity with varying inheritance patterns (estimated to range from 71-100%), depending on the efficiencies of the sgRNA and Cas9 promoters [233]. MGDrivE modeling estimates that 10 weekly releases of 10,000 homozygous males expressing the split gene drive into a population of 10,000 mosquitoes could drive the transgene to fixation within ~ 8 months and could last > 4 years [233]. The split drive is expected to be confinable: the sgRNA/anti-pathogen allele is estimated to reach 15% allele frequency in untargeted neighboring

populations (compared to 50% in the absence of a split gene drive) after 10 releases, before being gradually eliminated because of fitness costs. Importantly, the release of wild-type males can reverse the action of the split gene drive system if necessary [233]. Other "exhaustible" gene drive systems that have been explored in *Ae. aegypti* include daisy chain – a split drive system that encodes an allele resistant to the gene drive element [320] – and self-elimination technologies – where the gene drive cassette is targeted for excision [321].

Transparency, safety, and standardized quality control methods must be implemented once transgenic mosquitoes exhibiting gene drive are ready for field applications. Several publications outline standard operating procedures and containment guidelines for genetically modified organisms in the context of gene drive [322–324]. Community engagement and safety regulations will be as crucial as scientific rigor before gene drive technologies can be applied to mitigate the prevalence of arboviruses in the field.

5.3 Concluding remarks

Arboviruses must share harmonious relationships with their vectors so that they can persist in nature. Arboviruses therefore cause little-to-no fitness costs in the vector so they can be transmitted to new hosts. In the case of those transmitted by bloodfeeding, arboviruses must replicate at high enough titers so they can disseminate to the salivary glands and be egested in saliva. Rapid virus replication triggers acute siRNA innate immunity, which the virus must evade efficiently enough to continue to replicate. On the other hand, insect-specific viruses persistently infect the mosquito germline, which triggers piRNA-mediated immunity with which the virus must co-exist at sufficient levels that allows for vertical transmission. In both cases, the virus-vector relationship has evolved over thousands of years to achieve a seemingly healthy balance in viral RNA and RNAi response that allows for virus transmission.

Given the intimate relationship of mosquito vectors and RNA viruses, *Ae. aegypti* are a prime model for studying RNAi. Two years before the formal discovery of RNAi [113], *Ae. aegypti* were known to use dsRNA substrates to silence virus replication by sequence-specific homology [114]. Since then, the mechanisms underlying exogenous siRNA-mediated antiviral immunity have been well studied in *Ae aegypti* and elsewhere and have paved the way for the generation of transgenic virus-resistant mosquitoes. More elusive are the functions of and mechanisms underlying the piRNA pathway, where the current model is that the Piwis are expressed in the germline and associate with piRNAs to target transposable elements that threaten germline integrity. The paradigm is shifting. piRNAs in the soma have been associated with not only antiviral immunity, but genome rearrangement, epigenetic programming, cell regeneration, memory, and cancer [325]. Piwis are increasingly being identified in somatic tissues [325,326], and their sub-cellular localization is more complex than initially thought [304,327]. What better a system to study the nuances of lesser understood piRNA and Piwi functions than *Ae. aegypti,* an organism that is becoming increasingly easier to rear, that exhibits an expanded Piwi repertoire, and that relies on RNAi-mediated silencing in diverse functions? The complexity of RNA interference as an antiviral mechanism is overwhelming, but the field is rapidly evolving with *Ae. aegypti* at the forefront.

REFERENCES

1. Vos, T. *et al.* Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* **396**, 1204–1222 (2020).
2. Olival, K. J. *et al.* Host and viral traits predict zoonotic spillover from mammals. *Nature* **546**, 646–650 (2017).
3. R., P. C. *et al.* Cross-Species Virus Transmission and the Emergence of New Epidemic Diseases. *Microbiol. Mol. Biol. Rev.* **72**, 457–470 (2008).
4. Quinn, S. C. & Kumar, S. Health inequalities and infectious disease epidemics: a challenge for global health security. *Biosecur. Bioterror.* **12**, 263–273 (2014).
5. Kreuder Johnson, C. *et al.* Spillover and pandemic properties of zoonotic viruses with high host plasticity. *Sci. Rep.* **5**, 14830 (2015).
6. Wu, T. *et al.* Economic growth, urbanization, globalization, and the risks of emerging infectious diseases in China: A review. *Ambio* **46**, 18–29 (2017).
7. Farzanegan, M. R., Feizi, M. & Gholipour, H. F. Globalization and the Outbreak of COVID-19: An Empirical Analysis. *Journal of Risk and Financial Management* vol. 14 (2021).
8. Roser, M. & Ritchie, H. Burden of Disease. *Our World Data* (2016).
9. WHO. Vector-borne diseases. *WHO* https://www.who.int/news-room/fact-sheets/detail/vector-borne-diseases (2020).
10. Torto, B. & Tchouassi, D. P. Grand Challenges in Vector-Borne Disease Control Targeting Vectors. *Frontiers in Tropical Diseases* vol. 1 (2021).
11. Wilder-Smith, A. *et al.* Epidemic arboviral diseases: priorities for research and public health. *Lancet Infect. Dis.* **17**, e101–e106 (2017).
12. Pierson, T. C. & Diamond, M. S. The continued threat of emerging flaviviruses. *Nat. Microbiol.* **5**, 796–812 (2020).
13. Weaver, S. C. & Reisen, W. K. Present and future arboviral threats. *Antiviral Res.* **85**, 328–345 (2010).
14. Ronca, S. E., Ruff, J. C. & Murray, K. O. A 20-year historical review of West Nile virus since its initial emergence in North America: Has West Nile virus become a neglected tropical disease? *PLoS Negl. Trop. Dis.* **15**, e0009190–e0009190 (2021).
15. Sacchetto, L. *et al.* Neighbor danger: Yellow fever virus epizootics in urban and urban-rural transition areas of Minas Gerais state, during 2017-2018 yellow fever outbreaks in Brazil. *PLoS Negl. Trop. Dis.* **14**, e0008658–e0008658 (2020).
16. Brito, A. F. *et al.* Lying in wait: the resurgence of dengue virus after the Zika epidemic in Brazil. *Nat. Commun.* **12**, 2619 (2021).
17. Moore, S. M. *et al.* Leveraging multiple data types to estimate the size of the Zika epidemic in the Americas. *PLoS Negl. Trop. Dis.* **14**, e0008640–e0008640 (2020).
18. Kading, R. C., Brault, A. C. & Beckham, J. D. Global Perspectives on Arbovirus Outbreaks: A 2020 Snapshot. *Trop. Med. Infect. Dis.* **5**, 142 (2020).
19. Gubler, D. J. Resurgent vector-borne diseases as a global health problem. *Emerg. Infect. Dis.* **4**, 442–450 (1998).
20. Campbell-Lendrum, D., Manga, L., Bagayoko, M. & Sommerfeld, J. Climate change and

vector-borne diseases: what are the implications for public health research and policy? *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **370**, 20130552 (2015).
21. Ryan, S. J., Carlson, C. J., Mordecai, E. A. & Johnson, L. R. Global expansion and redistribution of *Aedes*-borne virus transmission risk with climate change. *PLoS Negl. Trop. Dis.* **13**, e0007213–e0007213 (2019).
22. Rocklöv, J. & Dubrow, R. Climate change: an enduring challenge for vector-borne disease prevention and control. *Nat. Immunol.* **21**, 479–483 (2020).
23. CDC. Arbovirus Catalog. *CDC* https://wwwn.cdc.gov/arbocat/.
24. Pfeffer, M. & Dobler, G. Emergence of zoonotic arboviruses by animal trade and migration. *Parasit. Vectors* **3**, 35 (2010).
25. Go, Y. Y., Balasuriya, U. B. R. & Lee, C.-K. Zoonotic encephalitides caused by arboviruses: transmission and epidemiology of alphaviruses and flaviviruses. *Clin. Exp. Vaccine Res.* **3**, 58–77 (2014).
26. Higgs, S. & Beaty, B. J. Natural Cycles of Vector-Borne Pathogens. in *Biology of Disease Vectors* (ed. Marquardt, W. C.) 167–185 (Elsevier, 2005).
27. Black IV, W. C. & Moore, C. G. Population biology as a tool to study vector-borne diseases. in *Biology of Disease Vectors* (ed. Marquardt, W. C.) 187–206 (Elsevier, 2005).
28. Souza-Neto, J. A., Powell, J. R. & Bonizzoni, M. *Aedes aegypti* vector competence studies: A review. *Infect. Genet. Evol.* **67**, 191–209 (2019).
29. Brustolin, M., Pujhari, S., Henderson, C. A. & Rasgon, J. L. *Anopheles* mosquitoes may drive invasion and transmission of Mayaro virus across geographically diverse regions. *PLoS Negl. Trop. Dis.* **12**, e0006895 (2018).
30. Franz, A. W. E., Kantor, A. M., Passarelli, A. L. & Clem, R. J. Tissue Barriers to Arbovirus Infection in Mosquitoes. *Viruses* **7**, 3741–3767 (2015).
31. Theiler, M. & Smith, H. H. THE EFFECT OF PROLONGED CULTIVATION IN VITRO UPON THE PATHOGENICITY OF YELLOW FEVER VIRUS . *J. Exp. Med.* **65**, 767–786 (1937).
32. Vannice, K. S. *et al.* The future of Japanese encephalitis vaccination: expert recommendations for achieving and maintaining optimal JE control. *npj Vaccines* **6**, 82 (2021).
33. Kubinski, M. *et al.* Tick-Borne Encephalitis Virus: A Quest for Better Vaccines against a Virus on the Rise. *Vaccines* **8**, 451 (2020).
34. Centers for Disease Control and Prevention. Tick-borne encephalitis vaccine. *CDC* https://www.cdc.gov/tick-borne-encephalitis/vaccine/index.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fticks%2Ftbe_vaccine%2Findex.html (2022).
35. Luckhart, S., Lindsay, S. W., James, A. A. & Scott, T. W. Reframing critical needs in vector biology and management of vector-borne disease. *PLoS Negl. Trop. Dis.* **4**, e566–e566 (2010).
36. Achee, N. L. *et al.* Alternative strategies for mosquito-borne arbovirus control. *PLoS Negl. Trop. Dis.* **13**, e0006822–e0006822 (2019).
37. Agarwal, A., Parida, M. & Dash, P. K. Impact of transmission cycles and vector competence on global expansion and emergence of arboviruses. *Rev. Med. Virol.* **27**, e1941 (2017).
38. Braack, L., Gouveia de Almeida, A. P., Cornel, A. J., Swanepoel, R. & de Jager, C. Mosquito-borne arboviruses of African origin: review of key viruses and vectors. *Parasit.*

Vectors **11**, 29 (2018).
39. Jentes, E. S. *et al.* The revised global yellow fever risk map and recommendations for vaccination, 2010: consensus of the Informal WHO Working Group on Geographic Risk for Yellow Fever. *Lancet Infect. Dis.* **11**, 622–632 (2011).
40. Leparc-Goffart, I., Nougairede, A., Cassadou, S., Prat, C. & de Lamballerie, X. Chikungunya in the Americas. *Lancet* **383**, 514 (2014).
41. Anubis, V.-R. *et al.* High Level of Vector Competence of *Aedes aegypti* and *Aedes albopictus* from Ten American Countries as a Crucial Factor in the Spread of Chikungunya Virus. *J. Virol.* **88**, 6294–6306 (2014).
42. Simmons, C. P., Farrar, J. J., van Vinh Chau, N. & Wills, B. Dengue. *N. Engl. J. Med.* **366**, 1423–1432 (2012).
43. Grunnill, M. & Boots, M. How Important is Vertical Transmission of Dengue Viruses by Mosquitoes (Diptera: Culicidae)? *J. Med. Entomol.* **53**, 1–19 (2016).
44. Boyer, S., Calvez, E., Chouin-Carneiro, T., Diallo, D. & Failloux, A.-B. An overview of mosquito vectors of Zika virus. *Microbes Infect.* **20**, 646–660 (2018).
45. Gutiérrez-Bugallo, G. *et al.* Vector-borne transmission and evolution of Zika virus. *Nat. Ecol. Evol.* **3**, 561–569 (2019).
46. Russell, R. C. Ross River Virus: Ecology and Distribution. *Annu. Rev. Entomol.* **47**, 1–31 (2002).
47. Nielsen, S. S. *et al.* Rift Valley Fever – epidemiological update and risk of introduction into Europe. *EFSA J.* **18**, e06041 (2020).
48. Tantely, L. M., Boyer, S. & Fontenille, D. A review of mosquitoes associated with Rift Valley fever virus in Madagascar. *Am. J. Trop. Med. Hyg.* **92**, 722–729 (2015).
49. Harris, M. C. *et al.* La Crosse Virus in *Aedes japonicus japonicus* mosquitoes in the Appalachian Region, United States. *Emerg. Infect. Dis.* **21**, 646–649 (2015).
50. C., B. A. *et al.* Venezuelan equine encephalitis emergence: Enhanced vector infection from a single amino acid substitution in the envelope glycoprotein. *Proc. Natl. Acad. Sci.* **101**, 11344–11349 (2004).
51. Rezza, G., Chen, R. & Weaver, S. C. O'nyong-nyong fever: a neglected mosquito-borne viral disease. *Pathog. Glob. Health* **111**, 271–275 (2017).
52. Armstrong, P. M. & Andreadis, T. G. Eastern equine encephalitis virus in mosquitoes and their role as bridge vectors. *Emerg. Infect. Dis.* **16**, 1869–1874 (2010).
53. Sakkas, H., Bozidis, P., Franks, A. & Papadopoulou, C. Oropouche Fever: A Review. *Viruses* **10**, 175 (2018).
54. Lindahl, J., Chirico, J., Boqvist, S., Thu, H. T. V. & Magnusson, U. Occurrence of Japanese encephalitis virus mosquito vectors in relation to urban pig holdings. *Am. J. Trop. Med. Hyg.* **87**, 1076–1082 (2012).
55. Nash, D. *et al.* The Outbreak of West Nile Virus Infection in the New York City Area in 1999. *N. Engl. J. Med.* **344**, 1807–1814 (2001).
56. Adouchief, S., Smura, T., Sane, J., Vapalahti, O. & Kurkela, S. Sindbis virus as a human pathogen—epidemiology, clinical picture and pathogenesis. *Rev. Med. Virol.* **26**, 221–241 (2016).
57. Ali, R. *et al.* Changing patterns in the distribution of the Mayaro virus vector *Haemagogus* species in Trinidad, West Indies. *Acta Trop.* **199**, 105108 (2019).
58. Papa, A., Tsergouli, K., Tsioka, K. & Mirazimi, A. Crimean-Congo Hemorrhagic Fever: Tick-Host-Virus Interactions. *Front. Cell. Infect. Microbiol.* **7**, 213 (2017).

59. Chitimia-Dobler, L. *et al.* Repeated isolation of tick-borne encephalitis virus from adult *Dermacentor reticulatus* ticks in an endemic area in Germany. *Parasit. Vectors* **12**, 90 (2019).
60. Hermance, M. E. & Thangamani, S. Powassan Virus: An Emerging Arbovirus of Public Health Concern in North America. *Vector Borne Zoonotic Dis.* **17**, 453–462 (2017).
61. Powell, J. R., Gloria-Soria, A. & Kotsakiozi, P. Recent History of *Aedes aegypti*: Vector Genomics and Epidemiology Records. *Bioscience* **68**, 854–860 (2018).
62. Dickson, L. B., Sanchez-Vargas, I., Sylla, M., Fleming, K. & Black IV, W. C. Vector Competence in West African *Aedes aegypti* is Flavivirus Species and Genotype Dependent. *PLoS Negl. Trop. Dis.* **8**, e3153 (2014).
63. Lwande, O. W. *et al.* Globe-Trotting *Aedes aegypti* and *Aedes albopictus*: Risk Factors for Arbovirus Pandemics. *Vector Borne Zoonotic Dis.* **20**, 71–81 (2020).
64. Schneider, J. R., Morrison, A. C., Astete, H., Scott, T. W. & Wilson, M. L. Adult Size and Distribution of *Aedes aegypti* (Diptera: Culicidae) Associated with Larval Habitats in Iquitos, Peru. *J. Med. Entomol.* **41**, 634–642 (2004).
65. Couret, J., Dotson, E. & Benedict, M. Q. Temperature, larval diet, and density effects on development rate and survival of *Aedes aegypti* (Diptera: Culicidae). *PLoS One* **9**, e87468–e87468 (2014).
66. Steinwascher, K. Competition among *Aedes aegypti* larvae. *PLoS One* **13**, e0202455–e0202455 (2018).
67. Wada, Y. Effect of larval density on the development of *Aedes aegypti* (L.) and the size of adults. *Quaest. Entomol.* **1**, 223–249 (1965).
68. Dye, C. Models for the population dynamics of the yellow fever mosquito, *Aedes aegypti*. *J. Anim. Ecol.* **53**, 247–268 (1984).
69. Juliano, S. A. POPULATION DYNAMICS. *J. Am. Mosq. Control Assoc.* **23**, 265–275 (2007).
70. Herd, C. S., Grant, D. G., Lin, J. & Franz, A. W. E. Starvation at the larval stage increases the vector competence of *Aedes aegypti* females for Zika virus. *PLoS Negl. Trop. Dis.* **15**, e0010003 (2021).
71. Krenn, H. W. & Aspöck, H. Form, function and evolution of the mouthparts of blood-feeding Arthropoda. *Arthropod Struct. Dev.* **41**, 101–118 (2012).
72. Lehane, M. J. *The Biology of Blood-Sucking in Insects*. (Cambridge University Press, 2005). doi:10.1017/CBO9780511610493.
73. Wahid, I., Sunahara, T. & Mogi, M. Maxillae and Mandibles of Male Mosquitoes and Female Autogenous Mosquitoes (Diptera: Culicidae). *J. Med. Entomol.* **40**, 150–158 (2003).
74. Baranitharan, M., Gokulakrishnan, J. & Sridhar, N. *Introduction of Vector Mosquitoes*. (Lambert Academic Publishing, 2018).
75. Carvalho, D. O. *et al.* Mosquito pornoscopy: Observation and interruption of *Aedes aegypti* copulation to determine female polyandric event and mixed progeny. *PLoS One* **13**, e0193164–e0193164 (2018).
76. Carrasquilla, M. C., Lounibos, L. P., Honorio, N. A. & Murr, S. Spermathecal Filling in *Aedes aegypti* and *Aedes albopictus*: Effects of Female and Male Body Sizes and Species. *J. Med. Entomol.* **56**, 334–340 (2019).
77. Roth, L. M. A study of mosquito behavior: an experimental laboratory study of the sexual behavior of *Aedes aegypti* (Linnaeus). *Am. Midl. Nat.* **40**, (1948).

78. Helinski, M. E. H. & Harrington, L. C. Male mating history and body size influence female fecundity and longevity of the dengue vector *Aedes aegypti*. *J. Med. Entomol.* **48**, 202–211 (2011).
79. Felipe Ramírez-Sánchez, L., Camargo, C. & Avila, F. W. Male sexual history influences female fertility and re-mating incidence in the mosquito vector *Aedes aegypti* (Diptera: Culicidae). *J. Insect Physiol.* **121**, 104019 (2020).
80. Yuval, B. MATING SYSTEMS OF BLOOD-FEEDING FLIES. *Annu. Rev. Entomol.* **51**, 413–440 (2005).
81. Eldridge, B. F. Mosquitoes, the Culicidae. in *Biology of Disease Vectors* (ed. Marquardt, W. C.) 95–111 (Elsevier, 2005).
82. Black 4th, W. C. & Kondratieff, B. C. Evolution of Arthropod Disease Vectors. in *Biology of Disease Vectors* (ed. Marquardt, W. C.) 9–23 (Elsevier, 2005).
83. Liu, S. *et al.* Antagonistic actions of juvenile hormone and 20-hydroxyecdysone within the ring gland determine developmental transitions in *Drosophila*. *Proc. Natl. Acad. Sci.* **115**, 139 LP – 144 (2018).
84. Gwadz, R. W. Regulation of blood meal size in the mosquito. *J. Insect Physiol.* **15**, 2039–2044 (1969).
85. Calvo, E., Mans, B. J., Andersen, J. F. & Ribeiro, J. M. C. Function and Evolution of a Mosquito Salivary Protein Family. *J. Biol. Chem.* **281**, 1935–1942 (2006).
86. Schneider, B. S. & Higgs, S. The enhancement of arbovirus transmission and disease by mosquito saliva is associated with modulation of the host immune response. *Trans. R. Soc. Trop. Med. Hyg.* **102**, 400–408 (2008).
87. Ribeiro, J. M. C., Martin-Martin, I., Arcà, B. & Calvo, E. A Deep Insight into the Sialome of Male and Female *Aedes aegypti* Mosquitoes. *PLoS One* **11**, e0151400–e0151400 (2016).
88. Manning, J. E., Morens, D. M., Kamhawi, S., Valenzuela, J. G. & Memoli, M. Mosquito Saliva: The Hope for a Universal Arbovirus Vaccine? *J. Infect. Dis.* **218**, 7–15 (2018).
89. Tirloni, L., Calvo, E., Konnai, S. & da Silva Vaz Jr, I. Editorial: The Role of Saliva in Arthropod-Host-Pathogen Relationships. *Front. Cell. Infect. Microbiol.* **10**, 630626 (2021).
90. Hardy, J. L., Houk, E. J., Kramer, L. D. & Reeves, W. C. Intrinsic Factors Affecting Vector Competence of Mosquitoes for Arboviruses. *Annu. Rev. Entomol.* **28**, 229–262 (1983).
91. Sánchez-Vargas, I., Harrington, L. C., Doty, J. B., Black 4th, W. C. & Olson, K. E. Demonstration of efficient vertical and venereal transmission of dengue virus type-2 in a genetically diverse laboratory strain of *Aedes aegypti*. *PLoS Negl. Trop. Dis.* **12**, e0006754 (2018).
92. Robison, A., Young, M. C., Byas, A. D., Rückert, C. & Ebel, G. D. Comparison of Chikungunya Virus and Zika Virus Replication and Transmission Dynamics in *Aedes aegypti* Mosquitoes. *Am. J. Trop. Med. Hyg.* **103**, 869–875 (2020).
93. Sanchez-Vargas, I., Olson, K. E. & Black 4th, W. C. The Genetic Basis for Salivary Gland Barriers to Arboviral Transmission. *Insects* **12**, 73 (2021).
94. B., K. W., M., N. E., Shane, S. M., D., Y. A. & D., R. K. DC-SIGN and L-SIGN Can Act as Attachment Receptors for Alphaviruses and Distinguish between Mosquito Cell- and Mammalian Cell-Derived Viruses. *J. Virol.* **77**, 12022–12032 (2003).
95. Liu, Y. *et al.* Transmission-blocking antibodies against mosquito C-type lectins for

dengue prevention. *PLoS Pathog.* **10**, e1003931–e1003931 (2014).
96. Rose, P. P. *et al.* Natural resistance-associated macrophage protein is a cellular receptor for Sindbis virus in both insect and mammalian hosts. *Cell Host Microbe* **10**, 97–104 (2011).
97. Sakoonwatanyoo, P., Boonsanay, V. & Smith, D. R. Growth and Production of the Dengue Virus in C6/36 Cells and Identification of a Laminin-Binding Protein as a Candidate Serotype 3 and 4 Receptor Protein. *Intervirology* **49**, 161–172 (2006).
98. Salas-Benito, J. S. & del Angel, R. M. Identification of two surface proteins from C6/36 cells that bind dengue type 4 virus. *J. Virol.* **71**, 7246–7252 (1997).
99. Salas-Benito, J. *et al.* Evidence that the 45-kD glycoprotein, part of a putative dengue virus receptor complex in the mosquito cell line C6/36, is a heat-shock related protein. *Am. J. Trop. Med. Hyg.* **77**, 283–290 (2007).
100. Muñoz, M. de L. *et al.* Proteomic Identification of Dengue Virus Binding Proteins in *Aedes aegypti* Mosquitoes and *Aedes albopictus* Cells. *Biomed Res. Int.* **2013**, 875958 (2013).
101. Dong, S. *et al.* Chikungunya virus dissemination from the midgut of *Aedes aegypti* is associated with temporal basal lamina degradation during bloodmeal digestion. *PLoS Negl. Trop. Dis.* **11**, e0005976 (2017).
102. Kantor, A. M., Grant, D. G., Balaraman, V., White, T. A. & Franz, A. W. E. Ultrastructural Analysis of Chikungunya Virus Dissemination from the Midgut of the Yellow Fever Mosquito, *Aedes aegypti*. *Viruses* vol. 10 (2018).
103. Armstrong, P. M. *et al.* Successive blood meals enhance virus dissemination within mosquitoes and increase transmission potential. *Nat. Microbiol.* **5**, 239–247 (2020).
104. Romoser, W. S. *et al.* Pathogenesis of Rift Valley fever virus in mosquitoes -- tracheal conduits & the basal lamina as an extra-cellular barrier. *Arch. Virol. Suppl.* 89–100 (2005) doi:10.1007/3-211-29981-5_8.
105. Barreau, C., Conrad, J., Fischer, E., Lujan, H. D. & Vernick, K. D. Identification of surface molecules on salivary glands of the mosquito, *Aedes aegypti*, by a panel of monoclonal antibodies. *Insect Biochem. Mol. Biol.* **29**, 515–526 (1999).
106. Juhn, J. *et al.* Spatial mapping of gene expression in the salivary glands of the dengue vector mosquito, *Aedes aegypti*. *Parasit. Vectors* **4**, 1 (2011).
107. Ciano, K. A., Saredy, J. J. & Bowers, D. F. Heparan sulfate proteoglycan: an arbovirus attachment factor integral to mosquito salivary gland ducts. *Viruses* **6**, 5182–5197 (2014).
108. Miller, B. R. & Mitchell, C. J. Genetic Selection of a Flavivirus-Refractory Strain of the Yellow Fever Mosquito *Aedes aegypti*. *Am. J. Trop. Med. Hyg.* **45**, 399–407 (1991).
109. Bennett, K. E. *et al.* Variation in vector competence for dengue 2 virus among 24 collections of *Aedes aegypti* from Mexico and the United States. *Am. J. Trop. Med. Hyg. Am J Trop Med Hyg Am. J. Trop. Med. Hyg.* **67**, 85–92 (2002).
110. Bosio, C. F., Fulton, R. E., Salasek, M. L., Beaty, B. J. & Black 4th, W. C. Quantitative trait loci that control vector competence for dengue-2 virus in the mosquito *Aedes aegypti*. *Genetics* **156**, 687–698 (2000).
111. Lambrechts, L. *et al.* Specificity of resistance to dengue virus isolates is associated with genotypes of the mosquito antiviral gene Dicer-2. *Proceedings. Biol. Sci.* **280**, 20122437 (2013).
112. Hannon, G. J. RNA interference. *Nature* **418**, 244–251 (2002).
113. Fire, A. *et al.* Potent and specific genetic interference by double-stranded RNA in

Caenorhabditis elegans. *Nature* **391**, 806–811 (1998).
114. E., O. K. *et al.* Genetically Engineered Resistance to Dengue-2 Virus Transmission in Mosquitoes. *Science (80-.).* **272**, 884–886 (1996).
115. Gaines, P. J. *et al.* Pathogen-derived resistance to dengue type 2 virus in mosquito cells by expression of the premembrane coding region of the viral genome. *J. Virol.* **70**, 2132–2137 (1996).
116. Sánchez-Vargas, I. *et al.* Dengue virus type 2 infections of *Aedes aegypti* are modulated by the mosquito's RNA interference pathway. *PLoS Pathog.* **5**, e1000299–e1000299 (2009).
117. Blair, D. C. & Olson, E. K. The Role of RNA Interference (RNAi) in Arbovirus-Vector Interactions. *Viruses* vol. 7 (2015).
118. Miesen, P., Girardi, E. & van Rij, R. P. Distinct sets of PIWI proteins produce arbovirus and transposon-derived piRNAs in *Aedes aegypti* mosquito cells. *Nucleic Acids Res.* **43**, 6545–6556 (2015).
119. Schnettler, E. *et al.* Knockdown of piRNA pathway proteins results in enhanced semliki forest virus production in mosquito cells. *J. Gen. Virol.* **94**, 1680–1689 (2013).
120. Petit, M. *et al.* piRNA pathway is not required for antiviral defense in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci.* **113**, E4218–E4227 (2016).
121. Vodovar, N. *et al.* Arbovirus-derived piRNAs exhibit a ping-pong signature in mosquito cells. *PLoS One* **7**, (2012).
122. Morazzani, E. M., Wiley, M. R., Murreddu, M. G., Adelman, Z. N. & Myles, K. M. Production of virus-derived ping-pong-dependent piRNA-like small RNAs in the mosquito soma. *PLoS Pathog.* **8**, (2012).
123. Miesen, P., Joosten, J. & van Rij, R. P. PIWIs Go Viral: Arbovirus-Derived piRNAs in Vector Mosquitoes. *PLOS Pathog.* **12**, e1006017 (2016).
124. Varjak, M. *et al.* Spindle-E acts antivirally against alphaviruses in mosquito cells. *Viruses* **10**, (2018).
125. Varjak, M. *et al. Aedes aegypti* Piwi4 Is a Noncanonical PIWI Protein Involved in Antiviral Responses. *mSphere* **2**, (2017).
126. Lee, W.-S., Webster, J. A., Madzokere, E. T., Stephenson, E. B. & Herrero, L. J. Mosquito antiviral defense mechanisms: a delicate balance between innate immunity and persistent viral infection. *Parasit. Vectors* **12**, 165 (2019).
127. Tassetto, M. *et al.* Control of RNA viruses in mosquito cells through the acquisition of vDNA and endogenous viral elements. *Elife* **8**, e41244 (2019).
128. Blair, C. D. Deducing the Role of Virus Genome-Derived PIWI-Associated RNAs in the Mosquito–Arbovirus Arms Race. *Frontiers in Genetics* vol. 10 (2019).
129. Joosten, J. *et al.* The Tudor protein Veneno assembles the ping-pong amplification complex that produces viral piRNAs in *Aedes* mosquitoes. *Nucleic Acids Res.* **47**, 2546–2559 (2019).
130. Suzuki, Y. *et al.* Non-retroviral Endogenous Viral Element Limits Cognate Virus Replication in *Aedes aegypti* Ovaries. *Curr. Biol.* **30**, 3495-3506.e6 (2020).
131. Halbach, R. *et al.* A satellite repeat-derived piRNA controls embryonic development of *Aedes*. *Nature* **580**, 274–277 (2020).
132. Betting, V. *et al.* A piRNA-lncRNA regulatory network initiates responder and trailer piRNA formation during mosquito embryonic development. *RNA* **27**, 1155–1172 (2021).
133. Arensburger, P., Hice, R. H., Wright, J. A., Craig, N. L. & Atkinson, P. W. The mosquito

Aedes aegypti has a large genome size and high transposable element load but contains a low proportion of transposon-specific piRNAs. *BMC Genomics* **12**, (2011).
134. Girardi, E. *et al.* Histone-derived piRNA biogenesis depends on the ping-pong partners Piwi5 and Ago3 in *Aedes aegypti*. *Nucleic Acids Res.* **45**, 4881–4892 (2017).
135. Aravin, A. A., Hannon, G. J. & Brennecke, J. *</i>The Piwi-piRNA Pathway Provides an Adaptive Defense in the Transposon Arms Race*. http://science.sciencemag.org/.
136. O'Donnell, K. A. & Boeke, J. D. Mighty Piwis Defend the Germline against Genome Intruders. *Cell* vol. 129 37–44 (2007).
137. Senti, K. A. & Brennecke, J. The piRNA pathway: A fly's perspective on the guardian of the genome. *Trends in Genetics* vol. 26 499–509 (2010).
138. Tóth, K. F., Pezic, D., Stuwe, E. & Webster, A. The pirna pathway guards the germline genome against transposable elements. in *Advances in Experimental Medicine and Biology* vol. 886 51–77 (Springer New York LLC, 2016).
139. Höck, J. & Meister, G. The Argonaute protein family. *Genome Biology* vol. 9 (2008).
140. Czech, B. *et al.* piRNA-Guided Genome Defense: From Biogenesis to Silencing. (2018) doi:10.1146/annurev-genet-120417.
141. Ozata, D. M., Gainetdinov, I., Zoch, A., O'Carroll, D. & Zamore, P. D. PIWI-interacting RNAs: small RNAs with big functions. *Nature Reviews Genetics* vol. 20 89–108 (2019).
142. Malone, C. D. *et al.* Specialized piRNA Pathways Act in Germline and Somatic Tissues of the *Drosophila* Ovary. *Cell* **137**, 522–535 (2009).
143. Théron, E. *et al.* The interplay between the argonaute proteins piwi and aub within *Drosophila* germarium is critical for oogenesis, piRNA biogenesis and TE silencing. *Nucleic Acids Res.* **46**, 10052–10065 (2018).
144. Le Thomas, A. *et al.* Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. *Genes Dev.* **27**, 390–399 (2013).
145. Senti, K.-A., Jurczak, D., Sachidanandam, R. & Brennecke, J. piRNA-guided slicing of transposon transcripts enforces their transcriptional silencing via specifying the nuclear piRNA repertoire. (2015) doi:10.1101/gad.267252.
146. Fabry, M. H. *et al.* piRNA-guided co-transcriptional silencing coopts nuclear export factors. *Elife* **8**, (2019).
147. Cox, D. N. *et al.* A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. *Genes Dev.* **12**, 3715–3727 (1998).
148. Aravin, A. A. *et al.* Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the *D. melanogaster* germline. *Curr. Biol.* **11**, 1017–1027 (2001).
149. Klattenhoff, C. *et al. Drosophila* rasiRNA Pathway Mutations Disrupt Embryonic Axis Specification through Activation of an ATR/Chk2 DNA Damage Response. *Dev. Cell* **12**, 45–55 (2007).
150. Li, C. *et al.* Collapse of Germline piRNAs in the Absence of Argonaute3 Reveals Somatic piRNAs in Flies. *Cell* **137**, 509–521 (2009).
151. Muerdter, F. *et al.* A genome-wide RNAi screen draws a genetic framework for transposon control and primary piRNA biogenesis in *Drosophila*. *Mol. Cell* **50**, 736–748 (2013).
152. Handler, D. *et al.* The genetic makeup of the *Drosophila* piRNA pathway. *Mol. Cell* **50**, 762–777 (2013).
153. Théron, E., Dennis, C., Brasset, E. & Vaury, C. Distinct features of the piRNA pathway in

somatic and germ cells: From piRNA cluster transcription to piRNA processing and amplification. *Mobile DNA* vol. 5 (2014).
154. Mani, S. R. & Juliano, C. E. Untangling the web: The diverse functions of the PIWI/piRNA pathway. *Mol. Reprod. Dev.* **80**, 632–664 (2013).
155. Weick, E. M. & Miska, E. A. piRNAs: From biogenesis to function. *Development (Cambridge)* vol. 141 3458–3471 (2014).
156. Han, Y. N. *et al.* PIWI Proteins and PIWI-Interacting RNA: Emerging Roles in Cancer. *Cellular Physiology and Biochemistry* vol. 44 1–20 (2017).
157. Lewis, S. H. *et al.* Pan-arthropod analysis reveals somatic piRNAs as an ancestral defence against transposable elements. *Nat. Ecol. Evol.* **2**, 174–181 (2018).
158. Joosten, J. *et al.* PIWI proteomics identifies Atari and Pasilla as piRNA biogenesis factors in *Aedes* mosquitoes. *Cell Rep.* **35**, (2021).
159. Rückert, C. *et al.* Small RNA responses of *Culex* mosquitoes and cell lines during acute and persistent virus infection. *Insect Biochem. Mol. Biol.* **109**, 13–23 (2019).
160. Scott, J. C. *et al.* Comparison of dengue virus type 2-specific small RNAs from RNA interference - competent and - incompetent mosquito cells. *PLoS Negl. Trop. Dis.* **4**, (2010).
161. Hess, A. M. *et al.* Small RNA profiling of Dengue virus-mosquito interactions implicates the PIWI RNA pathway in anti-viral defense. *BMC Microbiol.* **11**, (2011).
162. Goic, B. *et al.* Virus-derived DNA drives mosquito vector tolerance to arboviral infection. *Nat. Commun.* **7**, (2016).
163. Tian, Y., Simanshu, D., Ma, J. & Patel, D. Structural Basis for piRNA 2-O-methylated 3-end Recognition by Piwi PAZ (Piwi/Argonaute/Awille) Domains. (2011) doi:10.1073/pnas.1017762108.
164. Joosten, J., Van Rij, R. & Miesen, P. Slicing of viral RNA guided by endogenous piRNAs triggers the production of responder and trailer piRNAs in *Aedes* mosquitoes. (2020) doi:10.1101/2020.07.08.193029.
165. Holmes, E. C. The evolution of endogenous viral elements. *Cell Host Microbe* **10**, 368–377 (2011).
166. Bejarano, E. R., Khashoggi, A., Witty, M. & Lichtenstein, C. Integration of multiple repeats of geminiviral DNA into the nuclear genome of tobacco during evolution. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 759–764 (1996).
167. Benveniste, R. E. & Todaro, G. J. Evolution of C-type viral genes: inheritance of exogenously acquired viral genes. *Nature* **252**, 456–459 (1974).
168. Weiss, R. A., Mason, W. S. & Vogt, P. K. Genetic recombinants and heterozygotes derived from endogenous and exogenous avian RNA tumor viruses. *Virology* **52**, 535–552 (1973).
169. Crochu, S. *et al.* Sequences of flavivirus-related RNA viruses persist in DNA form integrated in the genome of *Aedes* spp. mosquitoes. *J. Gen. Virol.* **85**, 1971–1980 (2004).
170. Belshaw, R., Katzourakis, A., Pačes, J., Burt, A. & Tristem, M. High Copy Number in Human Endogenous Retrovirus Families is Associated with Copying Mechanisms in Addition to Reinfection. *Mol. Biol. Evol.* **22**, 814–817 (2005).
171. Frank, J. A. & Feschotte, C. Co-option of endogenous viral sequences for host cell function. *Curr. Opin. Virol.* **25**, 81–89 (2017).
172. Blair, C. D., Olson, K. E. & Bonizzoni, M. The Widespread Occurrence and Potential Biological Roles of Endogenous Viral Elements in Insect Genomes. *Curr. Issues Mol.*

Biol. **34**, 13–30 (2020).
173. Horie, M. *et al.* Endogenous non-retroviral RNA virus elements in mammalian genomes. *Nature* **463**, 84–87 (2010).
174. Palatini, U. *et al.* Comparative genomics shows that viral integrations are abundant and express piRNAs in the arboviral vectors *Aedes aegypti* and *Aedes albopictus*. *BMC Genomics* **18**, 512 (2017).
175. Suzuki, Y. *et al.* Uncovering the Repertoire of Endogenous Flaviviral Elements in *Aedes* Mosquito Genomes. *J. Virol.* **91**, e00571-17 (2017).
176. Whitfield, Z. J. *et al.* The Diversity, Structure, and Function of Heritable Adaptive Immunity Sequences in the *Aedes aegypti* Genome. *Curr. Biol.* **27**, 3511-3519.e7 (2017).
177. Silverman, N. *et al.* Control of RNA viruses in mosquito cells through the acquisition of vDNA and endogenous viral elements. (2019) doi:10.7554/eLife.41244.001.
178. Goic, B. *et al.* RNA-mediated interference and reverse transcription control the persistence of RNA viruses in the insect model *Drosophila*. *Nat. Immunol.* **14**, 396–403 (2013).
179. Parrish, N. F. *et al.* piRNAs derived from ancient viral processed pseudogenes as transgenerational sequence-specific immune memory in mammals. *RNA* **21**, 1691–1703 (2015).
180. Pischedda, E. *et al.* Insights Into an Unexplored Component of the Mosquito Repeatome: Distribution and Variability of Viral Sequences Integrated Into the Genome of the Arboviral Vector *Aedes albopictus*. *Frontiers in Genetics* vol. 10 (2019).
181. Vasilakis, N. & Tesh, R. B. Insect-specific viruses and their potential impact on arbovirus transmission. *Curr. Opin. Virol.* **15**, 69–74 (2015).
182. Coatsworth, H. *et al.* Intrinsic variation in the vertically transmitted core virome of the mosquito *Aedes aegypti*. *Mol. Ecol.* **n/a**, (2022).
183. Crava, C. M. *et al.* Population genomics in the arboviral vector Aedes aegypti reveals the genomic architecture and evolution of endogenous viral elements. *Mol. Ecol.* **30**, 1594–1611 (2021).
184. Russo, A. G., Kelly, A. G., Enosi Tuipulotu, D., Tanaka, M. M. & White, P. A. Novel insights into endogenous RNA viral elements in Ixodes scapularis and other arbovirus vector genomes. *Virus Evol.* **5**, vez010–vez010 (2019).
185. Dohm, J. C., Peters, P., Stralis-Pavese, N. & Himmelbauer, H. Benchmarking of long-read correction methods. *NAR Genomics Bioinforma.* **2**, lqaa037 (2020).
186. Müller, M., Fazi, F. & Ciaudo, C. Argonaute Proteins: From Structure to Function in Development and Pathological Cell Fate Determination. *Frontiers in Cell and Developmental Biology* vol. 7 (2020).
187. Wu, J., Yang, J., Cho, W. C. & Zheng, Y. Argonaute proteins: Structural features, functions and emerging roles. *J. Adv. Res.* **24**, 317–324 (2020).
188. Miyoshi, T., Ito, K., Murakami, R. & Uchiumi, T. Structural basis for the recognition of guide RNA and target DNA heteroduplex by Argonaute. *Nat. Commun.* **7**, 11846 (2016).
189. Yamaguchi, S. *et al.* Crystal structure of *Drosophila* Piwi. *Nat. Commun.* **11**, 858 (2020).
190. Matsumoto, N. *et al.* Crystal Structure of Silkworm PIWI-Clade Argonaute Siwi Bound to piRNA. *Cell* **167**, 484-497.e9 (2016).
191. Franz, A. W. E. *et al.* Engineering RNA interference-based resistance to dengue virus type 2 in genetically modified *Aedes aegypti*. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 4198–4203 (2006).

192. Mathur, G. *et al.* Transgene-mediated suppression of dengue viruses in the salivary glands of the yellow fever mosquito, *Aedes aegypti*. *Insect Mol. Biol.* **19**, 753–763 (2010).
193. Franz, A. W. E. *et al.* Fitness impact and stability of a transgene conferring resistance to dengue-2 virus following introgression into a genetically diverse *Aedes aegypti* strain. *PLoS Negl. Trop. Dis.* **8**, e2833–e2833 (2014).
194. Lilley, D. M. J. Structure, folding and mechanisms of ribozymes. *Curr. Opin. Struct. Biol.* **15**, 313–323 (2005).
195. Nawtaisong, P. *et al.* Effective suppression of Dengue fever virus in mosquito cell cultures using retroviral transduction of hammerhead ribozymes targeting the viral genome. *Virol. J.* **6**, 73 (2009).
196. Carter, J. R. *et al.* Suppression of the Arboviruses Dengue and Chikungunya Using a Dual-Acting Group-I Intron Coupled with Conditional Expression of the Bax C-Terminal Domain. *PLoS One* **10**, e0139899–e0139899 (2015).
197. Mishra, P., Furey, C., Balaraman, V. & Fraser, M. J. Antiviral Hammerhead Ribozymes Are Effective for Developing Transgenic Suppression of Chikungunya Virus in *Aedes aegypti* Mosquitoes. *Viruses* **8**, 163 (2016).
198. Jupatanakul, N. *et al.* Engineered *Aedes aegypti* JAK/STAT Pathway-Mediated Immunity to Dengue Virus. *PLoS Negl. Trop. Dis.* **11**, e0005187–e0005187 (2017).
199. Souza-Neto, J. A., Sim, S. & Dimopoulos, G. An evolutionary conserved function of the JAK-STAT pathway in anti-dengue defense. *Proc. Natl. Acad. Sci.* **106**, 17841 LP – 17846 (2009).
200. Yen, P.-S., James, A., Li, J.-C., Chen, C.-H. & Failloux, A.-B. Synthetic miRNAs induce dual arboviral-resistance phenotypes in the vector mosquito *Aedes aegypti*. *Commun. Biol.* **1**, 11 (2018).
201. Buchman, A. *et al.* Engineered resistance to Zika virus in transgenic *Aedes aegypti* expressing a polycistronic cluster of synthetic small RNAs. *Proc. Natl. Acad. Sci.* **116**, 3656 LP – 3661 (2019).
202. Liu, W.-L. *et al.* Transgenic refractory *Aedes aegypti* lines are resistant to multiple serotypes of dengue virus. *Sci. Rep.* **11**, 23865 (2021).
203. Buchman, A. *et al.* Broad dengue neutralization in mosquitoes expressing an engineered antibody. *PLOS Pathog.* **16**, e1008103 (2020).
204. A., S. S. *et al.* The Potent and Broadly Neutralizing Human Dengue Virus-Specific Monoclonal Antibody 1C19 Reveals a Unique Cross-Reactive Epitope on the bc Loop of Domain II of the Envelope Protein. *MBio* **4**, e00873-13 (2022).
205. Williams, A. E. *et al.* The Antiviral Small-Interfering RNA Pathway Induces Zika Virus Resistance in Transgenic *Aedes aegypti*. *Viruses* **12**, 1231 (2020).
206. Morris, A. C., Eggleston, P. & Crampton, J. M. Genetic transformation of the mosquito *Aedes aegypti* by micro-injection of DNA. *Med. Vet. Entomol.* **3**, 1–7 (1989).
207. Kokoza, V., Ahmed, A., Wimmer, E. A. & Raikhel, A. S. Efficient transformation of the yellow fever mosquito *Aedes aegypti* using the piggyBac transposable element vector pBac[3xP3-EGFP afm]. *Insect Biochem. Mol. Biol.* **31**, 1137–1143 (2001).
208. Cary, L. C. *et al.* Transposon mutagenesis of baculoviruses: analysis of *Trichoplusia ni* transposon IFP2 insertions within the FP-locus of nuclear polyhedrosis viruses. *Virology* **172**, 156–169 (1989).
209. Jasinskiene, N. *et al.* Stable transformation of the yellow fever mosquito, *Aedes aegypti*, with the *Hermes* element from the housefly. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 3743–3747

(1998).
210. Warren, W. D., Atkinson, P. W. & O'Brochta, D. A. The *Hermes* transposable element from the house fly, *Musca domestica*, is a short inverted repeat-type element of the hobo, Ac, and Tam3 (hAT) element family. *Genet. Res.* **64**, 87–97 (1994).
211. Coates, C. J., Jasinskiene, N., Miyashiro, L. & James, A. A. *Mariner* transposition and transformation of the yellow fever mosquito, *Aedes aegypti*. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 3748–3751 (1998).
212. Medhora, M., Maruyama, K. & Hartl, D. L. Molecular and functional analysis of the mariner mutator element *Mos1* in *Drosophila*. *Genetics* **128**, 311–318 (1991).
213. Adelman, Z. N., Jasinskiene, N. & James, A. A. Development and applications of transgenesis in the yellow fever mosquito, *Aedes aegypti*. *Mol. Biochem. Parasitol.* **121**, 1–10 (2002).
214. Criscione, F., O'Brochta, D. A. & Reid, W. Genetic technologies for disease vectors. *Curr. Opin. insect Sci.* **10**, 90–97 (2015).
215. Williams, A. E., Franz, A. W. E., Reid, W. R. & Olson, K. E. Antiviral Effectors and Gene Drive Strategies for Mosquito Population Suppression or Replacement to Mitigate Arbovirus Transmission by *Aedes aegypti*. *Insects* **11**, 52 (2020).
216. Dong, S. *et al.* Heritable CRISPR/Cas9-Mediated Genome Editing in the Yellow Fever Mosquito, *Aedes aegypti*. *PLoS One* **10**, e0122353 (2015).
217. Kistler, K. E., Vosshall, L. B. & Matthews, B. J. Genome engineering with CRISPR-Cas9 in the mosquito *Aedes aegypti*. *Cell Rep.* **11**, 51–60 (2015).
218. Basu, S. *et al.* Silencing of end-joining repair for efficient site-specific gene insertion after TALEN/CRISPR mutagenesis in *Aedes aegypti*. *Proc. Natl. Acad. Sci.* **112**, 4038 LP – 4043 (2015).
219. Franz, A. W. E. *et al.* Stability and loss of a virus resistance phenotype over time in transgenic mosquitoes harbouring an antiviral effector gene. *Insect Mol. Biol.* **18**, 661–672 (2009).
220. Wilson, R. *et al.* Post-integration behavior of a *Mos1 mariner* gene vector in *Aedes aegypti*. *Insect Biochem. Mol. Biol.* **33**, 853–863 (2003).
221. Moreira, L. A. *et al.* Robust gut-specific gene expression in transgenic *Aedes aegypti* mosquitoes. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 10895–10898 (2000).
222. Magalhaes, T. *et al.* Induction of RNA interference to block Zika virus replication and transmission in the mosquito *Aedes aegypti*. *Insect Biochem. Mol. Biol.* **111**, 103169 (2019).
223. Weger-Lucarelli, J. *et al.* Vector Competence of American Mosquitoes for Three Strains of Zika Virus. *PLoS Negl. Trop. Dis.* **10**, e0005101 (2016).
224. Roundy, C. M. *et al.* Variation in *Aedes aegypti* Mosquito Competence for Zika Virus Transmission. *Emerg. Infect. Dis.* **23**, 625–632 (2017).
225. Göertz, G. P. *et al.* Subgenomic flavivirus RNA binds the mosquito DEAD/H-box helicase ME31B and determines Zika virus transmission by *Aedes aegypti*. *Proc. Natl. Acad. Sci.* **116**, 19136 LP – 19144 (2019).
226. Crow, J. F. Unmasking a cheating gene. *Science* **283**, 1651–1652 (1999).
227. Burt, A. Site-specific selfish genes as tools for the control and genetic engineering of natural populations. *Proceedings. Biol. Sci.* **270**, 921–928 (2003).
228. Deredec, A., Burt, A. & Godfray, H. C. J. The population genetics of using homing endonuclease genes in vector and pest management. *Genetics* **179**, 2013–2026 (2008).

229. Akbari, O. S., Papathanos, P. A., Sandler, J. E., Kennedy, K. & Hay, B. A. Identification of germline transcriptional regulatory elements in *Aedes aegypti*. *Sci. Rep.* **4**, 3954 (2014).
230. James, A. A. Gene drive systems in mosquitoes: rules of the road. *Trends Parasitol.* **21**, 64–67 (2005).
231. Konet, D. S. *et al.* Short-hairpin RNA expressed from polymerase III promoters mediates RNA interference in mosquito cells. *Insect Mol. Biol.* **16**, 199–206 (2007).
232. Dang, Y. *et al.* Optimizing sgRNA structure to improve CRISPR-Cas9 knockout efficiency. *Genome Biol.* **16**, 280 (2015).
233. Li, M. *et al.* Development of a confinable gene drive system in the human disease vector *Aedes aegypti*. *Elife* **9**, e51701 (2020).
234. Windbichler, N. *et al.* Homing endonuclease mediated gene targeting in *Anopheles gambiae* cells and embryos. *Nucleic Acids Res.* **35**, 5922–5933 (2007).
235. Drury, D. W., Dapper, A. L., Siniard, D. J., Zentner, G. E. & Wade, M. J. CRISPR/Cas9 gene drives in genetically variable and nonrandomly mating wild populations. *Sci. Adv.* **3**, e1601910 (2017).
236. Li, M. *et al.* Germline Cas9 expression yields highly efficient genome engineering in a major worldwide disease vector, *Aedes aegypti*. *Proc. Natl. Acad. Sci.* **114**, E10540 LP-E10549 (2017).
237. Champer, J. *et al.* Reducing resistance allele formation in CRISPR gene drive. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 5522–5527 (2018).
238. Windbichler, N. *et al.* A synthetic homing endonuclease-based gene drive system in the human malaria mosquito. *Nature* **473**, 212–215 (2011).
239. Bier, E., Harrison, M. M., O'Connor-Giles, K. M. & Wildonger, J. Advances in Engineering the Fly Genome with the CRISPR-Cas System. *Genetics* **208**, 1–18 (2018).
240. Labun, K., Montague, T. G., Gagnon, J. A., Thyme, S. B. & Valen, E. CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. *Nucleic Acids Res.* **44**, W272–W276 (2016).
241. Labun, K. *et al.* CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res.* **47**, W171–W174 (2019).
242. Wendell, M. D., Wilson, T. G., Higgs, S. & Black, W. C. Chemical and gamma-ray mutagenesis of the white gene in *Aedes aegypti*. *Insect Mol. Biol.* **9**, 119–125 (2000).
243. Horn, C., Jaunich, B. & Wimmer, E. A. Highly sensitive, fluorescent transformation marker for *Drosophila* transgenesis. *Dev. Genes Evol.* **210**, 623–629 (2000).
244. Franz, A. W. E. *et al.* Comparison of transgene expression in *Aedes aegypti* generated by *mariner Mos1* transposition and ΦC31 site-directed recombination. *Insect Mol. Biol.* **20**, 587–598 (2011).
245. Beerntsen, B. T., Champagne, D. E., Coleman, J. L., Campos, Y. A. & James, A. A. Characterization of the *Sialokinin I* gene encoding the salivary vasodilator of the yellow fever mosquito, *Aedes aegypti*. *Insect Mol. Biol.* **8**, 459–467 (1999).
246. Edwards, M. J. *et al.* Characterization of a *carboxypeptidase A* gene from the mosquito, *Aedes aegypti*. *Insect Mol. Biol.* **9**, 33–38 (2000).
247. Edgar, R., Domrachev, M. & Lash, A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* **30**, 207–210 (2002).
248. Rosen, L. & Gubler, D. The use of mosquitoes to detect and propagate dengue viruses. *Am. J. Trop. Med. Hyg.* **23**, 1153–1160 (1974).

249. Sanchez-Vargas, I., Williams, A. E., Franz, A. W. E. & Olson, K. E. Intrathoracic Inoculation of Zika Virus in *Aedes aegypti*. *Bio-protocol* **11**, e4165 (2021).
250. Dubrulle, M., Mousson, L., Moutailler, S., Vazeille, M. & Failloux, A.-B. Chikungunya virus and *Aedes* mosquitoes: saliva is infectious as soon as two days after oral infection. *PLoS One* **4**, e5895 (2009).
251. Sanchez-Vargas, I., Harrington, L. C., Black 4th, W. C. & Olson, K. E. Analysis of Salivary Glands and Saliva from *Aedes albopictus* and *Aedes aegypti* Infected with Chikungunya Viruses. *Insects* **10**, 39 (2019).
252. Beerntsen, B. T., James, A. A. & Christensen, B. M. Genetics of mosquito vector competence. *Microbiol. Mol. Biol. Rev.* **64**, 115–137 (2000).
253. Tabachnick, W. J. Nature, nurture and evolution of intra-species variation in mosquito arbovirus transmission competence. *Int. J. Environ. Res. Public Health* **10**, 249–277 (2013).
254. van Cleef, K. W. R. *et al.* Mosquito and *Drosophila* entomobirnaviruses suppress dsRNA- and siRNA-induced RNAi. *Nucleic Acids Res.* **42**, 8732–8744 (2014).
255. Hobson-Peters, J. *et al.* A new insect-specific flavivirus from northern Australia suppresses replication of West Nile virus and Murray Valley encephalitis virus in co-infected mosquito cells. *PLoS One* **8**, e56534–e56534 (2013).
256. Hall-Mendelin, S. *et al.* The insect-specific Palm Creek virus modulates West Nile virus infection in and transmission by Australian mosquitoes. *Parasit. Vectors* **9**, 414 (2016).
257. Zhang, G., Asad, S., Khromykh, A. A. & Asgari, S. Cell fusing agent virus and dengue virus mutually interact in *Aedes aegypti* cell lines. *Sci. Rep.* **7**, 6935 (2017).
258. Schultz, M. J., Frydman, H. M. & Connor, J. H. Dual Insect specific virus infection limits Arbovirus replication in *Aedes* mosquito cells. *Virology* **518**, 406–413 (2018).
259. Parry, R. & Asgari, S. Aedes Anphevirus: an Insect-Specific Virus Distributed Worldwide in *Aedes aegypti* Mosquitoes That Has Complex Interplays with *Wolbachia* and Dengue Virus Infection in Cells. *J. Virol.* **92**, e00224-18 (2018).
260. Frentiu, F. D. *et al.* Limited Dengue Virus Replication in Field-Collected *Aedes aegypti* Mosquitoes Infected with *Wolbachia*. *PLoS Negl. Trop. Dis.* **8**, e2688 (2014).
261. Bolling, B. G., Weaver, S. C., Tesh, R. B. & Vasilakis, N. Insect-Specific Virus Discovery: Significance for the Arbovirus Community. *Viruses* **7**, 4911–4928 (2015).
262. Cook, S. *et al.* Isolation of a new strain of the flavivirus cell fusing agent virus in a natural mosquito population from Puerto Rico. *J. Gen. Virol.* **87**, 735–748 (2006).
263. Chandler, J. A. *et al.* Metagenomic shotgun sequencing of a Bunyavirus in wild-caught *Aedes aegypti* from Thailand informs the evolutionary and genomic history of the Phleboviruses. *Virology* **464**–**465**, 312–319 (2014).
264. Zhang, X. *et al.* Discovery and high prevalence of Phasi Charoen-like virus in field-captured *Aedes aegypti* in South China. *Virology* **523**, 35–40 (2018).
265. Ramos-Nino, M. E. *et al.* High prevalence of Phasi Charoen-like virus from wild-caught *Aedes aegypti* in Grenada, W.I. as revealed by metagenomic analysis. *PLoS One* **15**, e0227998 (2020).
266. Vasilakis, N. *et al.* Negevirus: a proposed new taxon of insect-specific viruses with wide geographic distribution. *J. Virol.* **87**, 2475–2488 (2013).
267. Kittayapong, P., Baisley, K. J. & O'Neill, S. L. A mosquito densovirus infecting *Aedes aegypti* and *Aedes albopictus* from Thailand. *Am. J. Trop. Med. Hyg.* **61**, 612–617 (1999).
268. Aguiar, E. R. G. R. *et al.* Sequence-independent characterization of viruses based on the

pattern of viral small RNAs produced by the host. *Nucleic Acids Res.* **43**, 6191–6206 (2015).
269. Cross, S. T. *et al.* Partitiviruses Infecting *Drosophila melanogaster* and *Aedes aegypti* Exhibit Efficient Biparental Vertical Transmission. *J. Virol.* **94**, e01070-20 (2022).
270. Zakrzewski, M. *et al.* Mapping the virome in wild-caught *Aedes aegypti* from Cairns and Bangkok. *Sci. Rep.* **8**, 4690 (2018).
271. Houé, V., Bonizzoni, M. & Failloux, A.-B. Endogenous non-retroviral elements in genomes of *Aedes* mosquitoes and vector competence. *Emerg. Microbes Infect.* **8**, 542–555 (2019).
272. Ma, Q. *et al.* A mosquito small RNA genomics resource reveals dynamic evolution and host responses to viruses and transposons. *Genome Res.* **31**, 512–528 (2021).
273. Lozano-Fuentes, S. *et al.* The Neovolcanic Axis Is a Barrier to Gene Flow among *Aedes aegypti* Populations in Mexico That Differ in Vector Competence for Dengue 2 Virus. *PLoS Negl. Trop. Dis.* **3**, e468 (2009).
274. Wise de Valdez, M. R. *et al.* Genetic elimination of dengue vector mosquitoes. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 4772–4775 (2011).
275. Elrefaey, A. M. *et al.* Understanding the Mechanisms Underlying Host Restriction of Insect-Specific Viruses. *Viruses* **12**, 964 (2020).
276. Patterson, E. I., Villinger, J., Muthoni, J. N., Dobel-Ober, L. & Hughes, G. L. Exploiting insect-specific viruses as a novel strategy to control vector-borne disease. *Curr. Opin. insect Sci.* **39**, 50–56 (2020).
277. Talyuli, O. A. C., Bottino-Rojas, V., Polycarpo, C. R., Oliveira, P. L. & Paiva-Silva, G. O. Non-immune Traits Triggered by Blood Intake Impact Vectorial Competence. *Frontiers in Physiology* vol. 12 (2021).
278. Ross, P. A. *et al.* An elusive endosymbiont: Does *Wolbachia* occur naturally in *Aedes aegypti*? *Ecol. Evol.* **10**, 1581–1591 (2020).
279. Sinha, A., Li, Z., Sun, L. & Carlow, C. K. S. Complete Genome Sequence of the *Wolbachia wAlbB* Endosymbiont of *Aedes albopictus*. *Genome Biol. Evol.* **11**, 706–720 (2019).
280. Czech, B. *et al.* An endogenous small interfering RNA pathway in *Drosophila*. *Nature* **453**, 798–802 (2008).
281. Ghildiyal, M. *et al.* Endogenous siRNAs derived from transposons and mRNAs in *Drosophila* somatic cells. *Science* **320**, 1077–1081 (2008).
282. Mirkovic-Hösle, M. & Förstemann, K. Transposon defense by endo-siRNAs, piRNAs and somatic pilRNAs in *Drosophila*: contributions of Loqs-PD and R2D2. *PLoS One* **9**, e84994–e84994 (2014).
283. Wen, J. *et al.* Diversity of miRNAs, siRNAs, and piRNAs across 25 *Drosophila* cell lines. *Genome Res.* **24**, 1236–1250 (2014).
284. Munivenkatappa, A. *et al.* Identification of Phasi Charoen-Like Phasivirus in Field Collected *Aedes aegypti* from Karnataka State, India. *Vector Borne Zoonotic Dis.* **21**, 900–909 (2021).
285. Olmo, R. P. *et al.* Insect-specific viruses regulate vector competence in *Aedes aegypti* mosquitoes via expression of histone H4. *bioRxiv* 2021.06.05.447047 (2021) doi:10.1101/2021.06.05.447047.
286. Amoa-Bosompem, M. *et al.* Entomological Assessment of the Status and Risk of Mosquito-borne Arboviral Transmission in Ghana. *Viruses* vol. 12 (2020).

287. Yamanaka, A., Thongrungkiat, S., Ramasoota, P. & Konishi, E. Genetic and evolutionary analysis of cell-fusing agent virus based on Thai strains isolated in 2008 and 2012. *Infect. Genet. Evol.* **19**, 188–194 (2013).
288. Calzolari, M. *et al.* Insect-specific flaviviruses, a worldwide widespread group of viruses only detected in insects. *Infect. Genet. Evol.* **40**, 381–388 (2016).
289. Espinoza-Gómez, F. *et al.* Detection of sequences from a potentially novel strain of cell fusing agent virus in Mexican *Stegomyia (Aedes) aegypti* mosquitoes. *Arch. Virol.* **156**, 1263–1267 (2011).
290. Hoshino, K., Isawa, H., Tsuda, Y., Sawabe, K. & Kobayashi, M. Isolation and characterization of a new insect flavivirus from *Aedes albopictus* and *Aedes flavopictus* mosquitoes in Japan. *Virology* **391**, 119–129 (2009).
291. Vera-Maloof, F. Z., Saavedra-Rodriguez, K., Elizondo-Quiroga, A. E., Lozano-Fuentes, S. & Black IV, W. C. Coevolution of the Ile1,016 and Cys1,534 Mutations in the Voltage Gated Sodium Channel Gene of *Aedes aegypti* in Mexico. *PLoS Negl. Trop. Dis.* **9**, e0004263–e0004263 (2015).
292. Solis-Santoyo, F. *et al.* Insecticide resistance in *Aedes aegypti* from Tapachula, Mexico: Spatial variation and response to historical insecticide use. *PLoS Negl. Trop. Dis.* **15**, e0009746 (2021).
293. Didion, J. P., Martin, M. & Collins, F. S. Atropos: specific, sensitive, and speedy trimming of sequencing reads. *PeerJ* **5**, e3720–e3720 (2017).
294. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
295. Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: A unified platform for automated protein structure and function prediction. *Nat. Protoc.* **5**, 725–738 (2010).
296. Simon, B. *et al.* Recognition of 2'-O-Methylated 3′-End of piRNA by the PAZ Domain of a Piwi Protein. *Structure* **19**, 172–180 (2011).
297. Ma, J. B., Ye, K. & Patel, D. J. Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature* **429**, 318–322 (2004).
298. Wiedemann, C., Bellstedt, P. & Görlach, M. CAPITO—a web server-based analysis and plotting tool for circular dichroism data. *Bioinformatics* **29**, 1750–1757 (2013).
299. Nguyen Ba, A. N., Pogoutse, A., Provart, N. & Moses, A. M. NLStradamus: A simple Hidden Markov Model for nuclear localization signal prediction. *BMC Bioinformatics* **10**, (2009).
300. Ariotti, N. *et al.* Modular Detection of GFP-Labeled Proteins for Rapid Screening by Electron Microscopy in Cells and Organisms. *Dev. Cell* **35**, 513–525 (2015).
301. Gamez, S., Srivastav, S., Akbari, O. S. & Lau, N. C. Diverse Defenses: A Perspective Comparing Dipteran Piwi-piRNA Pathways. *Cells* vol. 9 (2020).
302. Wang, W. *et al.* Slicing and Binding by Ago3 or Aub Trigger Piwi-Bound piRNA Production by Distinct Mechanisms. *Mol. Cell* **59**, 819–830 (2015).
303. Huang, X. *et al.* Binding of guide piRNA triggers methylation of the unstructured N-terminal region of Aub leading to assembly of the piRNA amplification complex. *Nat. Commun.* **12**, 4061 (2021).
304. Yashiro, R. *et al.* Piwi Nuclear Localization and Its Regulatory Mechanism in *Drosophila* Ovarian Somatic Cells. *Cell Rep.* **23**, 3647–3657 (2018).
305. Liu, Y. *et al.* The emerging role of the piRNA/piwi complex in cancer. *Molecular Cancer* vol. 18 (2019).

306. Wu, X. *et al.* The Biogenesis and Functions of piRNAs in Human Diseases. *Molecular Therapy - Nucleic Acids* vol. 21 108–120 (2020).
307. Yang, J. *et al.* The I-TASSER Suite: protein structure and function prediction. *Nat. Methods* **12**, 7–8 (2015).
308. Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**, (2008).
309. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, (2011).
310. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version 2-A multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).
311. Martin-Martin, I. *et al.* ADP binding by the *Culex quinquefasciatus* mosquito D7 salivary protein enhances blood feeding on mammals. *Nat. Commun.* **11**, 2911 (2020).
312. Reid, W. R. *et al.* Genomic insertion locus and Cas9 expression in the germline affect CRISPR/Cas9-based gene drive performance in the yellow fever mosquito *Aedes aegypti*. *bioRxiv* 2021.12.08.471839 (2021) doi:10.1101/2021.12.08.471839.
313. Adelman, Z. N. *et al. nanos* gene control DNA mediates developmentally regulated transposition in the yellow fever mosquito *Aedes aegypti*. *Proc. Natl. Acad. Sci.* **104**, 9970–9975 (2007).
314. Carballar-Lejarazú, R. *et al.* Next-generation gene drive for population modification of the malaria vector mosquito, *Anopheles gambiae*. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 22805–22814 (2020).
315. Hammond, A. *et al.* Regulating the expression of gene drives is key to increasing their invasive potential and the mitigation of resistance. *PLoS Genet.* **17**, e1009321–e1009321 (2021).
316. Kyrou, K. *et al.* A CRISPR–Cas9 gene drive targeting doublesex causes complete population suppression in caged *Anopheles gambiae* mosquitoes. *Nat. Biotechnol.* **36**, 1062–1066 (2018).
317. Terradas, G., Hermann, A., James, A. A., McGinnis, W. & Bier, E. High-resolution in situ analysis of Cas9 germline transcript distributions in gene-drive *Anopheles* mosquitoes. *G3 Genes|Genomes|Genetics* **12**, jkab369 (2022).
318. Wu, S. L. *et al.* MGDrivE 2: A simulation framework for gene drive systems incorporating seasonality and epidemiological dynamics. *PLOS Comput. Biol.* **17**, e1009030 (2021).
319. Sánchez C., H. M., Wu, S. L., Bennett, J. B. & Marshall, J. M. MGDrivE: A modular simulation framework for the spread of gene drives through spatially explicit mosquito populations. *Methods Ecol. Evol.* **11**, 229–239 (2020).
320. Noble, C. *et al.* Daisy-chain gene drives for the alteration of local populations. *Proc. Natl. Acad. Sci.* **116**, 8275 LP – 8282 (2019).
321. Zapletal, J. *et al.* Making gene drive biodegradable. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **376**, 20190804 (2021).
322. Adelman, Z. *et al.* Rules of the road for insect gene drive research and testing. *Nat. Biotechnol.* **35**, 716–718 (2017).
323. Benedict, M. Q. *et al.* Recommendations for Laboratory Containment and Management of Gene Drive Systems in Arthropods. *Vector Borne Zoonotic Dis.* **18**, 2–13 (2018).
324. ACME. Containment Practices for Arthropods Modified with Engineered Transgenes

Capable of Gene Drive Addendum 1 to the Arthropod Containment Guidelines, Version 3.2. *Vector Borne Zoonotic Dis.* **22**, 3–17 (2022).
325. Ross, R. J., Weiner, M. M. & Lin, H. PIWI proteins and PIWI-interacting RNAs in the soma. *Nature* **505**, 353–359 (2014).
326. Drews, F. *et al.* Two Piwis with Ago-like functions silence somatic genes at the chromatin level. *RNA Biol.* **18**, 757–769 (2021).
327. Nazer, E., Gómez Acuña, L. & Kornblihtt, A. R. Seeking the truth behind the myth: Argonaute tales from 'nuclearland'. *Mol. Cell* **82**, 503–513 (2022).

APPENDIX

Chapter 2 Supplemental Information

Supplemental Table 2.1. Primers used in Chapter 2.

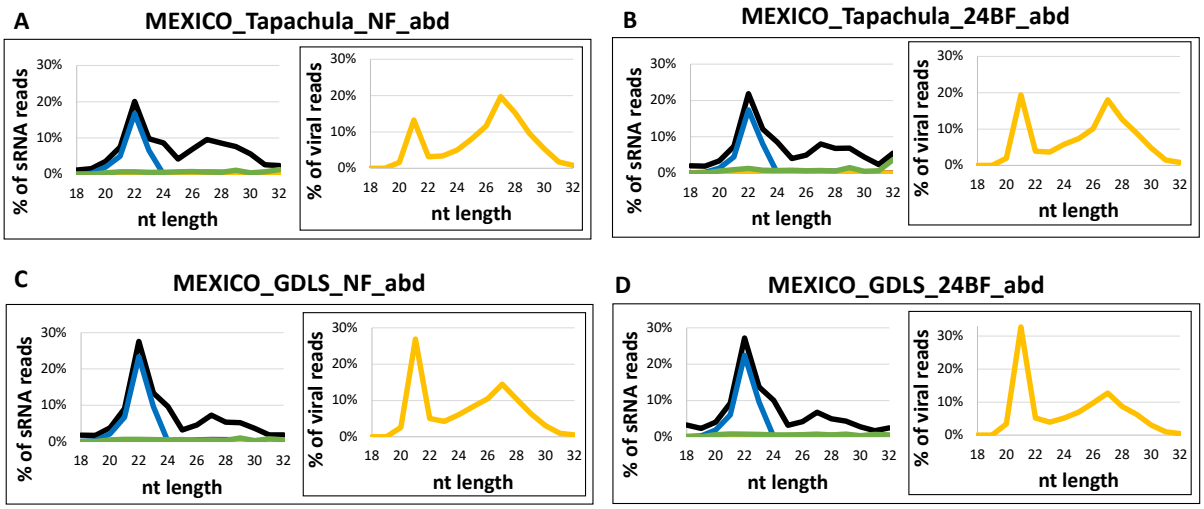
Primer	Sequence (5' to 3')	Purpose
BR-4	GAAATTAATACGACTCACTATAGGTGGTTCTCAGGCGACAGTTGTTTTAGAGCTAGAAATAGC	Forward primer for sgRNA 3
BR-5	GAAATTAATACGACTCACTATAGGTAATGGCACAGGCTATAGGTTTTAGAGCTAGAAATAGC	Forward primer for sgRNA 4
BR-6	GAAATTAATACGACTCACTATAGGACCAACCGCAGTCATTGTGGTTTTAGAGCTAGAAATAGC	Forward primer for sgRNA 5
BR-7	GAAATTAATACGACTCACTATAGGAGAACCACACAATGACTGGTTTTAGAGCTAGAAATAGC	Forward primer for sgRNA 6
BR-20	AACGAGATGCCTTCTCCTGA	Genomic primer for indel assessment
BR-23	AAAATGGCGTTCGATGAGAC	Genomic primer for indel assessment
BR-51	CATAAGCTTGCTAATCGGTGTCAATTCCATC	Upstream homology arm forward
BR-52	CATGGATCCTAACTCGAGTCTCAGGCGACAGTATGGCTC	Upstream homology arm reverse to introduce a XhoI site
BR-53	CATGGATCCATACTCGAGCTGCGGTTGGTCGTGTATTTGAAG	Downstream homology arm forward to introduce a XhoI site
BR-54	CATCTCGAGAAATTCGAGCTCGCCCGG	3xP3FOR-XhoI forward
BR-55	CATCTCGAGCCGTACGCGTATTCGATAAG	SV40REV-XhoI reverse
BR-73	GGGAAGGTTTCTGGTACAGG	Downstream homology arm reverse
BR-94	CATTATCTCGAGCTTCACGTTTTCCCAGGTCAGAAGC	attP-XhoI forward
BR-10 2	CTGCTATGTGGCGCGGTATTATC	β-lactamase breaking primer for Gibson assembly

BR-103	GATAATACCGCGCCACATAGCAG	β-lactamase breaking primer for Gibson assembly
BR-124	AATCTCGAGCTAGCACTAGTCGCGCTCGCGCG	attP-NheI reverse
BR-125	CAATATCTAGAGCCGTACGCGTATTCGATAAG	SV40-XbaI reverse
BR-126	ACAATTCTAGAGAAATTCGAGCTCGCCCGG	3xP3-XbaI forward
BR-138	TCGAGAATTCGCTAGCACTAGTCGCGCTCGCGCG	attP-NheI-EcoRI reverse
BR-139	CATTATCCGCGGCTTCACGTTTTCCCAGGTCAGAAGC	attP-SacII forward
BR-211	AGAAACCGCGGAGATCTCGGCCGCGACTCTAGATC	SV40-SacII-BglII forward
BR-212	CAATAGCTAGCTAAGATACATTGATGAGTTTGGACAAAC	SV40-NheI reverse
BR-213	AAACCTCTCGAGCGACCAAC	Anti-ZIKV dsRNA trigger + intron-XhoI
BR-214	GGTTTCTGCAGCTGCCTAATGATATATTTTTAAT	Anti-ZIKV dsRNA trigger + intron-PstI
BR-219	ATCGAGGCCTGAGCTCTTACGCGTGCTAGTTTTTG	*Carboxypeptidase A* promoter-StuI forward
BR-220	ATAGTAATCTCGAGCGGCCGCCTGCAG	*Carboxypeptidase A* promoter-XhoI reverse
BR-223	TCGGCATGGACGAGCTGTAC	eCFP anchored primer running out of the end of the color cds
BR-243	CGGCCGCGACTCTAGATC	Intermediate for final Gibson assembly
BR-245	TAGTGATTAAAAATATATCATTAGGCAGCTGATCTTTGCTTTTCTGGCTCAG	ZIKV dsRNA trigger sense strand for final Gibson assembly
BR-250	TCTCCGCGGTTTCTCTAGTGGATCCCCCGGGCGACCAACAACACCATAATGG	ZIKV dsRNA trigger sense strand for final Gibson assembly
BR-248	CCCGGGGGATCCACTAG	Intermediate for final Gibson assembly
BR-345	ATCACTAGCCCGGGCTGCA	*Carboxypeptidase A* anchored primer running into the IR
BR-347	GGAGGTGAGCACCCAATCATCAG	*sialokinin1* intron anchored primer running towards the IR promoter
BR-348	AATCACTAACAGAACTTTGAACAAAATC	*sialokinin1* intron anchored primer running towards the inverted-repeat IR terminator

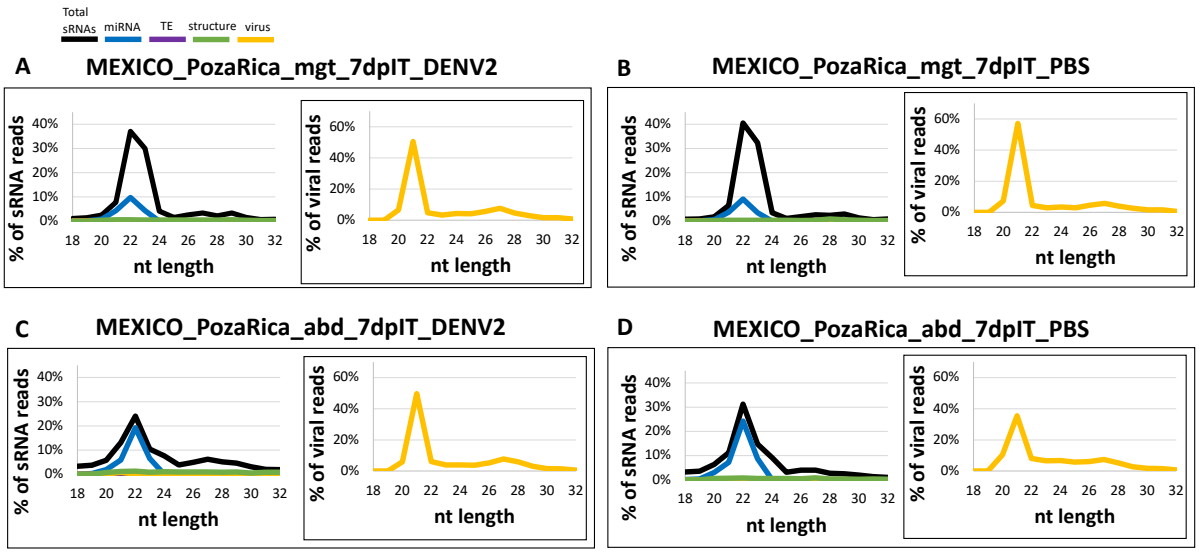
BR-240	AAACCTCTCGAGCGACCAA CAACACCATAATGGAAGAC AGTGTGCCGGCAGAGGTGT GGACCAGACACGGAGAGAA AAGAGTGCTCAAACCGAGG TGGATGGACGCCAGAGTTTG TTCAGATCATGCGGCCCTGA AGTCATTCAAGGAGTTTGCC GCTGGGAAAAGAGGAGCGG CTTTTGGAGTGATGGAAGCC CTGGGAACACTGCCAGGAC ACATGACAGAGAGATTCCA GGAAGCCATTGACAACCTC GCTGTGCTCATGCGGGCAGA GACTGGAAGCAGGCCTTAC AAAGCCGCGGCGGCCCAAT TGCCGGAGACCCTAGAGAC CATAATGCTTTTGGGGTTGC TGGGAACAGTCTCGCTGGG AATCTTCTTCGTCTTGATGA GGAACAAGGGCATAGGGAA GATGGGCTTTGGAATGGTGA CTCTTGGGGCCAGCGCATGG CTCATGTGGCTCTCGGAAAT TGAGCCAGCCAGAATTGCAT GTGTCCTCATTGTTGTGTTC CTATTGCTGGTGGTGCTCAT ACCTGAGCCAGAAAAGCAA AGATCGGAGGTGAGCACCC AATCATCAGATTTTGTTCAA AGTTCTGTTAGTGATTAAAA ATATATCATTAGGCAGCTGC AGAAACC	XhoI/PstI tagged *sialokinin1* intron and anti-ZIKV RNAi trigger sequence used to construct IR

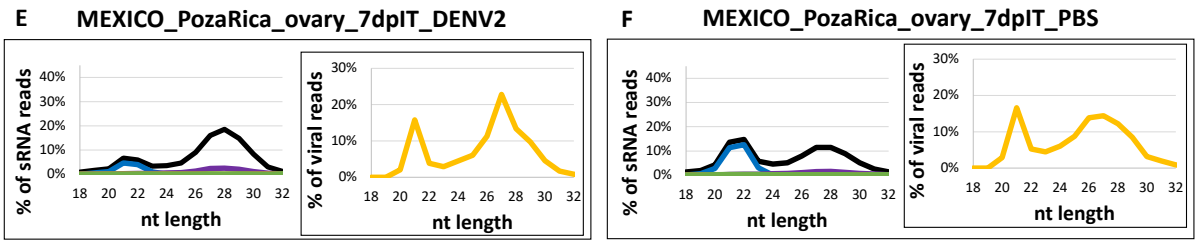
Chapter 3 Supplemental Information

Supplemental Figure 3.1. Origins of the genetically diverse laboratory strain in the state of Chiapas, Mexico. Map of Chiapas, Mexico indicating cities from which mosquito eggs were collected to establish the genetically diverse laboratory strain (GDLS) of *Ae. aegypti* sequenced in this study. Map adapted from alamy.com and reproduced with permission.

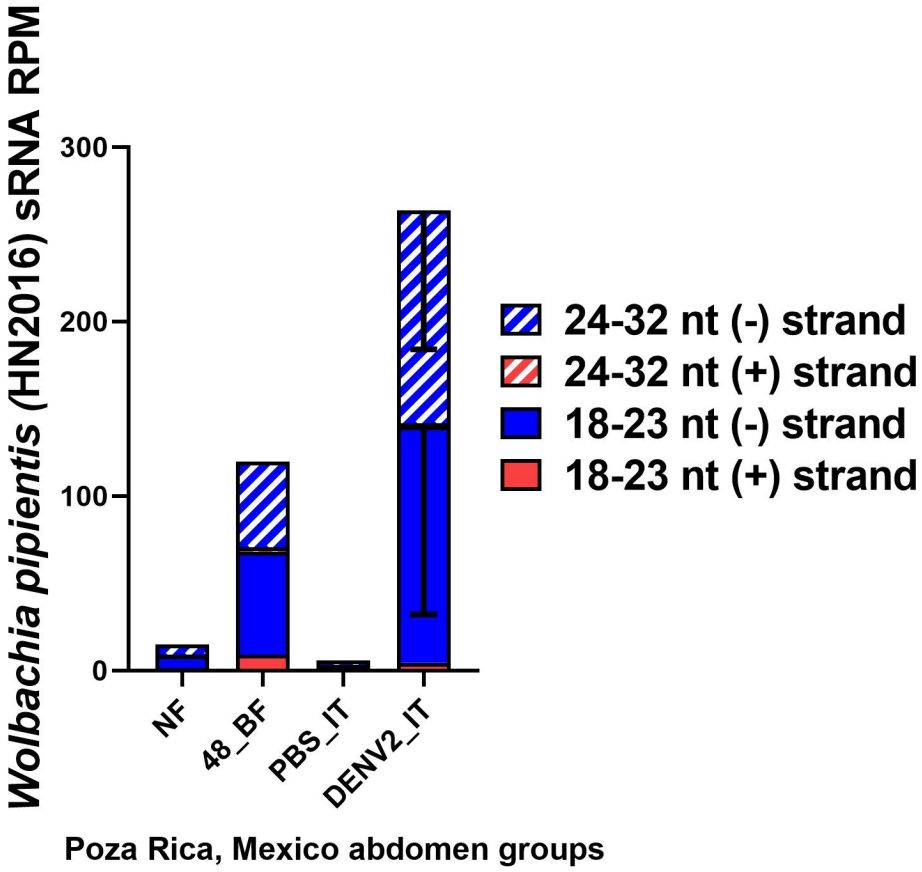


Supplemental Figure 3.2. Small RNA size distributions of DENV2-uninfected samples from the state of Chiapas, Mexico. Proportions of small RNA reads per million (RPM) for total small RNAs (sRNAs, black), microRNAs (miRNA, blue), transposable elements (TE, purple), structural RNAs (*e.g.* tRNAs or rRNAs, green), and virus-derived RNAs (yellow), relative to total RPM, are shown by nucleotide (nt) length. Inset graphs show virus-derived sRNA RPM relative to total virus-specific RPM in yellow by nucleotide length. Samples are from **A-B)** Tapachula = Tapachula, Mexico; **C-D)** GDLS = genetically diverse laboratory strain, generated from lines of *Ae. aegypti* derived from regions throughout Chiapas, Mexico as initially described in de Valdez et al., 2011 [274]; NF = non-fed and maintained on sugar; 24BF = 24 hours post non-infectious bloodfeed; mgt = midgut tissue; abd = abdomen tissue, ovaries removed.





Supplemental Figure 3.3. Small RNA size distributions of DENV2-infected and uninfected tissues from Poza Rica, Veracruz, Mexico. Total small RNAs (sRNAs, black), microRNAs (miRNA, blue), transposable elements (TE, purple), structural RNAs (*e.g.* tRNAs or rRNAs, green), and virus-derived RNAs (yellow) reads per million (RPM) normalized by total number of RPMs, are shown by nucleotide (nt) length. Inset graphs show virus-derived sRNA RPMs, normalized by total number of virus-specific RPMs, in yellow by nucleotide length. DENV2 or PBS intrathoracically inoculated (IT) midgut (**A-B**), abdomen (**C-D**), and ovary (**E-F**) tissues from Poza Rica, Mexico mosquitoes. nt = nucleotide; mgt = midgut; abdomen = abd.



Supplemental Figure 3.4. DENV2-infected *Ae. aegypti* display increased small RNAs derived from the endosymbiont *Wolbachia pipientis.* *Wolbachia pipientis* (HN2016 origin) small RNAs (sRNAs) normalized by reads per million (RPM) in Poza Rica, Mexico *Ae. aegypti*

maintained on sugar (NF = non-fed), 48 hours post-bloodmeal (48_BF), after intrathoracic inoculation with PBS (PBS_IT), or IT with dengue-2 (DENV2_IT). Bars indicate average RPM and standard errors of the mean (SEM).

Supplemental Table 3.1. List of non-retroviral integrated sequences (NIRVS) identified in the *Aedes aegypti* genome (Aag3 assembly), adapted from Palatini et al., 2017 [174]. Additional NIRVS annotations can be found in Crava et al., 2021[183] and Russo et al., 2019 [184].

NIRVS locus name	Viral family	Length	main ORFs	orientation	% identity	Genomic context
AeBun1	*Hantaviridae*	1,457	G protein (partial) [Imjin River V]	R	31%	intergenic
AeFlavi10	*Flaviviridae*	391	NS5 (partial) [CFA]	R	96%	intergenic
AeFlavi109	*Flaviviridae*	674	2 partial ORFs NS1 (partial) [Hanko V]	R	59-95%	intergenic
AeFlavi111	*Flaviviridae*	834	NS1 (partial) [Hanko V]	R	48%	intergenic
AeFlavi113	*Flaviviridae*	287	NS1 (partial) [ParV]	R	69%	intergenic
AeFlavi115	*Flaviviridae*	734	NS1 (partial) [AeFV]	R	66%	intergenic
AeFlavi31	*Flaviviridae*	1,376	NS5 (partial) [AeFV]	F	78%	exon of AAEL017001
AeFlavi32	*Flaviviridae*	446	NS5 (partial) [AeFV]	F	81%	intergenic
AeFlavi34	*Flaviviridae*	2,166	E, NS1 protein (partial) (AeFv)	R	63-100%	intergenic
AeFlavi4	*Flaviviridae*	1,673	2 ORFs for E ed M protein (partial) (KRV, AeFV)	R	40-48%	intergenic
AeFlavi41	*Flaviviridae*	143	polyprotein [AeFV]	R	44%	intergenic
AeFlavi42	*Flaviviridae*	544	NSA-NS3 (partial) [AeFV]	F	29%	intergenic
AeFlavi44	*Flaviviridae*	239	NS2A-NS3 (partial) [KRV]	F	34%	intergenic
AeFlavi45	*Flaviviridae*	2,012	NS3 (partial) [CFA]	F	55%	intergenic
AeFlavi50	*Flaviviridae*	1,584	NS3-NS4B (partial) [KRV]	F	38%	intergenic
AeFlavi53	*Flaviviridae*	1,021	NS1 (partial) [AeFV]	F	50%	intergenic
AeFlavi56	*Flaviviridae*	239	NS2A-NS3 (partial) [CFA]	F	34%	intergenic
AeFlavi57	*Flaviviridae*	788	NS3 (partial) [CFA]	F	41%	intergenic

AeFlavi59	*Flaviviridae*	572	NS3 (partial) [KRV]	F	61%	intergenic
AeFlavi60	*Flaviviridae*	2,063	NS3 (partial) [CFA]	F	55%	intergenic
AeFlavi62	*Flaviviridae*	1,148	NS3 (partial) [CFA]	F	56%	intergenic
AeFlavi66	*Flaviviridae*	221	NS3 (partial) [KRV]	F	63%	intergenic
AeFlavi73	*Flaviviridae*	1,581	NS3-NS2B (partial) [KRV]	F	27%	intergenic
AeFlavi77	*Flaviviridae*	961	NS1-NS2 (partial) [HankV]	F	49%	exon of AAEL0007866
AeFlavi80	*Flaviviridae*	224	NS2A-NS3 (partial) [CFAV]	F	32%	intergenic
AeFlavi81	*Flaviviridae*	188	NS5 (partial) [CFAV]	F	72%	intergenic
AeFlavi83	*Flaviviridae*	414	NS5 (partial) [KRV]	F	76%	intergenic
AeFlavi86	*Flaviviridae*	665	Flavi NS1 conserved (partial) [AeFv]	R	69%	intergenic
AeFlavi88	*Flaviviridae*	1,022	Flavi NS1 conserved (partial) [AeFv]	R	67%	intergenic
AeFlavi91	*Flaviviridae*	287	Flavi NS1 conserved (partial) [AeFv]	R	75%	intergenic
AeFlavi92	*Flaviviridae*	167	NS1-NS2A (partial) [KRV]	R	39%	intergenic
AeFlavi94	*Flaviviridae*	284	NS1 (partial) [AeFV]	R	76%	intergenic
AeFlavi95	*Flaviviridae*	5,400	4 partial ORFs for NS1 [KRV]	R	44-60%	intergenic
AeReo1	*Reoviridae*	707	VP5 (partial) Liao Ning Virus	R	86%	intergenic
AeRha100	*Rhabdoviridae*	686	G protein (partial) [GrassCarpRhaV]	F	29%	intergenic
AeRha102	*Rhabdoviridae*	782	2 ORFs N protein (partial) [EkpomaV]	F	38%	intergenic
AeRha104	*Rhabdoviridae*	461	N protein (partial) [CarajasV]	F	33%	intergenic
AeRha105	*Rhabdoviridae*	629	L protein (partial) [WuhanAntV]	F	41%	intergenic
AeRha108	*Rhabdoviridae*	509	L protein (partial) [WuhanAntV]	F	41%	intergenic
AeRha11	*Rhabdoviridae*	350	G protein (partial) [Wuhan House Fly V]	R	40%	intergenic
AeRha110	*Rhabdoviridae*	446	N protein (partial) [Tongilchon V1]	R	50%	intergenic
AeRha111	*Rhabdoviridae*	446	N protein (partial) [Tongilchon V1]	R	50%	intergenic

AeRha112	*Rhabdoviridae*	404	N protein (partial) [Culex triraeniorhynchus V]	F	47%	intergenic
AeRha113	*Rhabdoviridae*	1,145	N protein (partial) [Santa Barbara V]	R	30%	intergenic
AeRha116	*Rhabdoviridae*	470	N protein (partial) [Tongilchon V1]	R	50%	intergenic
AeRha117	*Rhabdoviridae*	170	N protein (partial) [Tongilchon V1]	R	38%	intergenic
AeRha118	*Rhabdoviridae*	726	2 Orfs for L protein (partial) [Wuhan Mosquito V9]	F	53-58%	intergenic
AeRha123	*Rhabdoviridae*	698	G protein (partial) [GrassCarpRhaV]	F	26%	intergenic
AeRha125	*Rhabdoviridae*	872	N protein (partial) [Culex triraeniorhynchus V]	R	38%	exon of AAEL009525
AeRha127	*Rhabdoviridae*	938	L protein (partial) [ShuanguaIV1]	F	56%	exon of AAEL001772
AeRha13	*Rhabdoviridae*	230	N protein (partial) [CarajasV]	R	33%	intergenic
AeRha130	*Rhabdoviridae*	869	N protein (partial) [Sunguru V]	F	31%	exon of AAEL009873
AeRha132	*Rhabdoviridae*	551	N protein (partial) [CocalV]	F	32%	exon of AAEL009870
AeRha133	*Rhabdoviridae*	685	2 partial ORFS for G [Rhabdo V]	F	31-76%	intergenic
AeRha136	*Rhabdoviridae*	161	G protein (partial) [Tupaia V]	F	38%	intergenic
AeRha137	*Rhabdoviridae*	380	N protein (partial) [Tongichon V1]	F	57%	exon of AAEL009940
AeRha138	*Rhabdoviridae*	446	N protein (partial) [Tongichon V1]	R	50%	intergenic
AeRha139	*Rhabdoviridae*	902	G protein (partial) [Grass carpV]	R	25%	intergenic
AeRha14	*Rhabdoviridae*	161	G protein (partial) [TupaiaV]	R	38%	intergenic
AeRha142	*Rhabdoviridae*	704	G protein (partial) [TongilchonV]	R	32%	intergenic
AeRha143	*Rhabdoviridae*	2,053	N protein (partial) [GarbaV]	F	33%	end of AAEL002517, beginning of AAEL002535
AeRha144	*Rhabdoviridae*	956	N protein (partial) [Berrimah V]	F	33%	exon of AAEL002535

AeRha146	*Rhabdoviridae*	803	L protein (partial) [IsfashanV]	F	62%	exon of AAEL017355
AeRha147	*Rhabdoviridae*	524	N protein (partial) [DurhamV]	F	71%	intergenic
AeRha148	*Rhabdoviridae*	578	N protein (partial) [GarbaV]	R	33%	intergenic
AeRha149	*Rhabdoviridae*	581	N protein (partial) [GarbaV]	R	33%	intergenic
AeRha15	*Rhabdoviridae*	251	G protein (partial) [TupaiaV]	F	41%	intergenic
AeRha150	*Rhabdoviridae*	581	N protein (partial) [Drosophila melanogaster Sigma Virus]	R	32%	intron of AAEL002051
AeRha152	*Rhabdoviridae*	185	N protein (partial) [DrosoAffinisSigma]	F	44%	intergenic
AeRha153	*Rhabdoviridae*	293	N protein (partial) [Culex tritaeniorhynchus V]	F	72%	intergenic
AeRha154	*Rhabdoviridae*	1,223	N protein (partial) [Santa BarbaraV]	F	30%	exon of AAEL012729
AeRha156	*Rhabdoviridae*	170	N protein (partial) [YataV]	F	38%	exon of AAEL003033
AeRha157	*Rhabdoviridae*	509	N protein (partial) [Culex triraeniorhynchus V]	F	41%	exon of AAEL003033
AeRha16	*Rhabdoviridae*	795	N protein (partial) [Sunguru V]	F	34%	intergenic
AeRha160	*Rhabdoviridae*	390	L protein (partial) [ChandipuraV]	F	63%	intergenic
AeRha162	*Rhabdoviridae*	719	N protein (partial) [CarajasV]	R	30%	intergenic
AeRha163	*Rhabdoviridae*	992	N protein (partial) [Coastal Plain virus]	R	29%	intergenic
AeRha168	*Rhabdoviridae*	524	N protein (partial) [GarbaV]	R	34%	intergenic
AeRha169	*Rhabdoviridae*	404	N protein (partial) [KotonkanV]	F	28%	intergenic
AeRha2	*Rhabdoviridae*	326	N protein (partial) [Bass Congo V]	R	47%	intron of AAEL000093
AeRha245	*Rhabdoviridae*	287	G protein (partial) [Xincheng Mosquito V)	F	45%	intergenic
AeRha246	*Rhabdoviridae*	167	G protein (partial) [Xincheng Mosquito V]	F	62%	intergenic
AeRha27	*Rhabdoviridae*	281	L protein (partial) [Riverside V]	R	67%	intron of AAEL014445
AeRha28	*Rhabdoviridae*	389	G protein (partial) [Tongilchon V1]	F	31%	intergenic

AeRha29	*Rhabdoviridae*	275	G protein (partial) [Tongilchon V]	F	33%	intergenic
AeRha3	*Rhabdoviridae*	386	N protein (partial) [Walkabout Creek V]	R	36%	intron of AAEL000093
AeRha30	*Rhabdoviridae*	857	2 ORFs G protein (partial) [Rhabdo V]	F	31-34%	intergenic (PIRC5)
AeRha33	*Rhabdoviridae*	905	G protein (partial) [Culex Triteeniorhynchus RV]	F	31%	intergenic
AeRha36	*Rhabdoviridae*	881	N protein (partial) [Culex triraeniorhynchus V]	F	35%	exon of AAEL004959
AeRha38	*Rhabdoviridae*	365	L protein (partial) [Tongilcon V]	R	78%	intergenic
AeRha42	*Rhabdoviridae*	879	N protein (partial) [Scophthalmus maximum V]	R	98%	exon of AAEL005456 and extending on its 5' end
AeRha45	*Rhabdoviridae*	884	N protein (partial) [Culex triraeniorhynchus V]	R	36%	exon of AAEL000808
AeRha47	*Rhabdoviridae*	516	N protein (partial) [CarajasV]	F	38%	intergenic
AeRha49	*Rhabdoviridae*	1,088	N protein (partial) [Tongilchon V1]	R	43%	exon of AAEL006218
AeRha5	*Rhabdoviridae*	473	N protein (partial) [Garba V]	R	33%	intergenic
AeRha51	*Rhabdoviridae*	1,088	N protein (partial) [Tongilchon V1]	R	43%	exon of AAEL006217
AeRha54	*Rhabdoviridae*	290	N protein (partial) [Culex triraeniorhynchus V]	R	50%	intron of AAEL000126
AeRha55	*Rhabdoviridae*	878	N protein (partial) [EkpomaV]	F	37%	exon of AAEL000130
AeRha57	*Rhabdoviridae*	524	N protein (partial) [Coastal PlainsV]	R	35%	exon of AAEL000120
AeRha58	*Rhabdoviridae*	167	N protein (partial) [Gata Virus]	F	35%	exon of AAEL000114
AeRha6	*Rhabdoviridae*	287	N protein (partial) [Tongilchon V1]	R	57%	intergenic
AeRha61	*Rhabdoviridae*	986	N protein (partial) [Coastal PlainsV]	R	31%	exon of AAEL000991
AeRha63	*Rhabdoviridae*	998	N protein (partial) [Barrimah V]	R	35%	exon of AAEL001003

AeRha66	*Rhabdoviridae*	986	N protein (partial) [Barrimah V]	R	33%	exon of AAEL000976
AeRha68	*Rhabdoviridae*	887	N protein (partial) [Barrimah V]	R	33%	exon of AAEL000997
AeRha70	*Rhabdoviridae*	302	N protein (partial) [Culex triraeniorhynchus V]	R	39%	intergenic
AeRha71	*Rhabdoviridae*	230	N protein (partial) [Vesicular stomatitis Alagoas V]	R	33%	intergenic
AeRha72	*Rhabdoviridae*	848	G protein (partial) [Tongilchon V 1]	R	31%	intergenic
AeRha75	*Rhabdoviridae*	440	2 ORFS for partial G protein (partial) [Rhabdo V]	F	41-53%	intergenic
AeRha77	*Rhabdoviridae*	679	2 ORFs for G protein (partial) [Rhabdo V]	F	27-42%	intergenic
AeRha8	*Rhabdoviridae*	239	N protein (partial) [Culex triraeniorhynchus V]	R	33%	intergenic
AeRha80	*Rhabdoviridae*	653	G protein (partial) [Tongilchon V1]	F	29%	intergenic
AeRha82	*Rhabdoviridae*	359	G protein (partial) [Yongija Tick V2]	R	26%	intergenic
AeRha83	*Rhabdoviridae*	767	G protein (partial) [Culex tritaeniorhynchus V]	R	27%	intergenic
AeRha85	*Rhabdoviridae*	881	N protein (partial) [Tongilchon V 1]	F	43%	Exon of AAEL007520
AeRha87	*Rhabdoviridae*	767	N protein (partial) [BerrimahV]	F	33%	Exon of AAEL007529
AeRha89	*Rhabdoviridae*	482	N protein (partial) [CocalV]	R	32%	Exon of AAEL001267-AAEL001259
AeRha9	*Rhabdoviridae*	422	G protein (partial) [Tongilchon V]	R	32%	intergenic
AeRha91	*Rhabdoviridae*	890	G protein (partial) [Culex triraeniorhynchus V]	F	28%	Exon of AAEL007844
AeRha93	*Rhabdoviridae*	590	G protein (partial) [GrassCarpRhaV]	F	29%	Exon of AAEL007844
AeRha96	*Rhabdoviridae*	785	N protein (partial) [EkpomaV]	F	34%	intergenic
AeRha98	*Rhabdoviridae*	557	G protein (partial) [MeridaV]	F	34%	intergenic

Chapter 4 Supplemental Information

Supplemental Table 4.1. Primer sequences used in Chapter 4. Primer names (left), sequences (middle) and reason for use (right).

primer name	sequence	purpose
piwi4Nde_F	CATATGTCTGACCGTTACTCTCAAGGG	Piwi4 restriction enzyme-mediated cloning pET-17b
piwi4Xho_R	CTCGAGTCAGTGGTGGTGGT	Piwi4 restriction enzyme-mediated cloning pET-17b
piwi4_int_R	TGGTAGACATCTCCAGTACGTTCG	Piwi4 sequencing primer
piwi4pazNDE_F	CATATGCAGACATGCTACGACATCTTG	Piwi4 PAZ restriction enzyme-mediated cloning pET-17b
piwi4pazxhohis_R	CTCGAGTCAATGGTGATGGTGATGATGCGTCATCTG	Piwi4 PAZ restriction enzyme-mediated cloning pET-17b
QC-PAZ17-T41A-F	CAACAATAAAACCTATGCCATTCACGACGTCAC	Piwi4 PAZ T41A mutation
QC-PAZ17-T41A-R	GTGACGTCGTGAATGGCATAGGTTTTATTGTTG	Piwi4 PAZ T41A mutation
QC-PAZ17-T41R-F	GTTACAACAATAAAACCTATCGCATTCACGACGTCACGTTTG	Piwi4 PAZ T41R mutation
QC-PAZ17-T41R-R	CAAACGTGACGTCGTGAATGCGATAGGTTTTATTGTTGTAAC	Piwi4 PAZ T41R mutation
QC-PAZ17-Y40A-F	GTTACAACAATAAAACCGCTACCATTCACGACGTCAC	Piwi4 PAZ Y40A mutation
QC-PAZ17-Y40A-R	GTGACGTCGTGAATGGTAGCGGTTTTATTGTTGTAAC	Piwi4 PAZ Y40A mutation
QC-PAZ17-T39A-F	CGGTTACAACAATAAAGCCTATACCATTCACGAC	Piwi4 PAZ T39A mutation
QC-PAZ17-T39A-R	GTCGTGAATGGTATAGGCTTTATTGTTGTAACCG	Piwi4 PAZ T39A mutation
QC-PAZ17-F55A-F	GAAACCACTCCGGAGAGTACGGCCGATACCAAGGCCGGTAAAAC	Piwi4 PAZ F55A mutation
QC-PAZ17-F55A-R	GTTTTACCGGCCTTGGTATCGGCCGTACTCTCCGGAGTGGTTTC	Piwi4 PAZ F55A mutation
Inf-NLS-NcoI-F	GGCTGCCGCCACCATGGGATCGAGGGAACCGAGAGAACAC	Piwi4 NLS infusion-mediated cloning

Inf-NLS-NcoI-R	CCCTTGCTCACCATGGCTCCAACGCCACGGCGACTGCG	Piwi4 NLS/N terminal infusion-mediated cloning
InfNLSlong-NcoI-F	GGCTGCCGCCACCATGGGATCTGACCGTTACTCTCAAGGG	Piwi4 N terminal infusion-mediated cloning

Supplemental Table 4.2. RNA sequences used in Chapter 4. RNA sequences used for SPR experiments. BioON = biotinylated. 2OMe = 3'-terminal 2'-O-methylated.

RNA	**sequence**
pclvpirnabio	[BioON]CGAUAAGUGAUCUUUCAGCACUGCAGAA
pclvpirnametbio	[BioON]CGAUAAGUGAUCUUUCAGCACUGCAGA[2OMeA]
scrampirnabio	[BioON]AUGAGGACAUAAGCCUACGUAGUAUUCC

LIST OF ABBREVIATIONS

AaVV: Aedes aegypti virga-like virus
abd: abdomen
AeAV: Aedes anphevirus
Ago: argonaute
arboviruses: arthropod-borne viruses
Aub: Aubergine
BF: bloodfeed
bp: base pair
BTI: *Bacillus thuringiensis israelensis*
CD: circular dichroism
CDC: Centers for Disease Control and Prevention
cDNA: complementary DNA
CFAV: cell-fusing agent virus
CHIKV: chikungunya virus
CpA: *carboxypeptidase A*
CRISPR: clustered regularly interspaced short palindromic repeats
Dcr2: Dicer-2
DENV: dengue virus
DNA: deoxyribonucleic acid
DMEM: Dulbecco's modified Eagle medium
dpi: days post-infection
dpIT: days post-intrathoracic inoculation
DSB: double-stranded break
dsDNA: double-stranded DNA
dsRNA: double-stranded RNA
dsSIN: double subgenomic Sindbis virus vector
eCFP: enhanced cyan fluorescent protein
eGFP: enhanced green fluorescent protein
EIP: extrinsic incubation period
EVE: endogenous viral element
FBS: fetal bovine serum
GDLS: genetically diverse laboratory strain
GhAaTV (also AaTV in other texts): Ghana Aedes aegypti totivirus
GPI: glycosylphosphatidylinositol
GuAaTV (also GAATV in other texts): Guadeloupe Aedes aegypti totivirus
HA: homology arms
HDR: homology directed repair
HEG: homing endonuclease genes
hRz: hammerhead ribozyme
HTV: Humaita-Tubiacanga virus
HWE: Higgs' white eye
ID: intrinsically disordered

IFA: immunofluorescence assay
IR: inverted repeat
ISV: insect-specific virus
IT: intrathoracic inoculation
ITC: isothermal calorimetry
ITR: inverted terminal repeat
IVM: integrated vector management
JH: juvenile hormone
K_D: dissociation constant
KO: knock out
LACV: La Crosse virus
LTR: long terminal repeat
met: 3'-terminal 2'-O-methylated
MEB: midgut escape barrier
mgt: midgut
MIB: midgut infection barrier
miRISC: miRNA-induced silencing complex
miRNA: microRNA
moi: multiplicity of infection
MSRG: Mosquito small RNA genomics resource
MTT: 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
NES: nuclear export signal
NF: non-blood fed
NHEJ: non-homologous end joining
NIRVS: non-retroviral integrated RNA virus sequences
NLS: nuclear localization signal
nmet: 3'-terminal 2' OH
ns: non-significant
NS: non-structural protein
nt: nucleotide
NVA: neovolcanic axis
PAM: protospacer adjacent motif
PAZ: Piwi/Argonaute/Zwille
pbm: post-bloodmeal
PBS: phosphate buffer saline
PCLV: Phasi Charoen-like virus
PCR: polymerase chain reaction
PFU: plaque-forming unit
piRISC: piRNA-induced silencing complex
piRNA: PIWI-interacting RNA
PIWI: P-element induced wimpy testis
prM: pre-membrane protein
Pub: polyubiquitin
QTL: quantitative trait loci
RdRP: RNA dependent RNA polymerase
RIDL: release of insects carrying a dominant lethal

RISC: RNA-induced silencing complex
RNA: Ribonucleic acid
RNAi: RNA interference
RPM: reads per million
RT-PCR reverse-transcription PCR
SDS-PAGE: sodium dodecyl-sulfate polyacrylamide gel electrophoresis
SGE: selfish genetic element
SGEB: salivary gland escape barrier
SGIB: salivary gland infection barrier
sfRNA: subgenomic flaviviral RNA
sgRNA: single guide RNA
SINV: Sindbis virus
siRNA: small-interfering RNA
SPR: surface plasmon resonance
sRNA: small RNA
ssDNA: single-stranded DNA
ssRNA: single-stranded RNA
TALEN: transcription activator-like effector nucleases
TE: transposable element
TIMP: tissue inhibitor of metalloproteinases
tTAV: tetracycline-repressible transcriptional transactivator
vDNA: viral derived cDNA
vpiRNA: virus-derived piRNA
vsRNA: virus-derived small RNAs
vsiRNA: virus-derived siRNA
WT: wild type
YFV: yellow fever virus
YPP: yolk protein precursors
ZIKV: Zika virus
ZFN: zinc finger nuclease
20E: 20-hydroxyecdysone