

The Pennsylvania State University

The Graduate School

ARBOVIRAL INFECTION DYNAMICS AND EVOLUTIONARY GENETICS OF MOSQUITO VECTORS

A Dissertation in
Molecular, Cellular, and Integrative Biosciences
by
Cory Henderson

© 2021 Cory Henderson

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2021

The dissertation of Cory Henderson was reviewed and approved by the following:

Jason L. Rasgon
Professor of Entomology
Center for Infectious Disease Dynamics
Huck Institutes of the Life Sciences
Dissertation Advisor

Elizabeth McGraw
Professor and Huck Scholar in Entomology
Center for Infectious Disease Dynamics
Center for Chemical Ecology
Molecular, Cellular, and Integrative Biosciences
Chair of Committee

Joyce Jose
Assistant Professor of Biochemistry & Molecular Biology
Center for Infectious Disease Dynamics
Center for Structural Biology

Isabella Cattadori
Associate Professor of Biology
Center for Infectious Disease Dynamics
Molecular, Cellular, and Integrative Biosciences

Melissa Rolls
Paul Berg Professor of Biochemistry and Molecular Biology Director of the Center for Cellular
Dynamics Chair of the Intercollege Graduate Degree Program in Molecular, Cellular and
Integrative Biosciences

ABSTRACT

Increasing globalization and climate change have contributed to increase and spread of arboviral pathogens. Mayaro virus (MAYV) is an arboviral pathogen in the genus *Alphavirus* that is circulating in South America with potential to spread to naïve regions. MAYV is also one of the few viruses with the ability to be transmitted by mosquitoes in the genus *Anopheles*, as well as the typical arboviral transmitting mosquitoes in the genus *Aedes*. *Anopheles* mosquitoes are the principal vectors for malaria and lymphatic filariasis, and they demonstrate limited evidence for arboviral transmission in laboratory and natural contexts. Vector management approaches require an understanding of the ecological, epidemiological, and biological context of the species in question. To address these aspects of vector management, I studied the interactions between co-circulating arboviral pathogens in vector mosquitoes, investigated the molecular basis for arboviral infection of vector species, and produced novel genome assemblies from understudied mosquito vector species.

First, I describe co- and super-infection of *Aedes aegypti* with MAYV and Zika virus (ZIKV) to understand the interaction dynamics of these viral pathogens. Zika virus (ZIKV) is a *Flavivirus* primarily transmitted by *Aedes* mosquitoes and has emerged as a major arboviral public health concern due to its capacity to cross the placenta effecting developing fetus in pregnant women. Our results show a significant reduction in the infection rate of ZIKV during a co-infection of ZIKV and MAYV but no variation in the infection rate of MAYV. I also demonstrate an increase in the infection rate of MAYV when the mosquito was previously infected with Zika virus. This study highlights the dynamics of co- and super-infection and its role in arboviral infections and

transmission also highlights the importance of considering these dynamics during risk assessment in epidemic areas.

To investigate the response of anopheline mosquitoes to alphaviral infection at time points integral to transmission, I detail the transcriptomic and small RNA responses of *An. stephensi* to infection with MAYV via infectious bloodmeal at 2, 7, and 14 days post infection (dpi). 487 unique transcripts were significantly regulated, 79 putative novel miRNAs were identified, and an siRNA response is observed targeting the MAYV genome. Gene ontology analysis of transcripts regulated at each timepoint suggested activation of the Toll pathway at 7 dpi and repression of pathways related to autophagy and apoptosis at 14 dpi. These findings provide a basic understanding of the infection response of *An. stephensi* to MAYV and help to identify host factors which might be useful to target to inhibit viral replication in *Anopheles* mosquitoes.

Finally, I present novel genome assemblies for *An. crucians*, *An. freeborni*, *An. albimanus*, and *An. quadrimaculatus* and examine the evolutionary relationship between these species. I create a phylogeny using the newly sequenced anopheline genomes and further identify 525 single copy orthologs demonstrating evidence for positive selection on at least one branch of the phylogeny. Gene ontology terms such as calcium ion signaling, histone binding, and protein acetylation identified as being biased in the set of selected genes. These novel genome sequences will be useful in developing our understanding of the diverse biological traits that drive vectorial capacity in anophelines.

TABLE OF CONTENTS

List of Figures.....	viii
List of Tables.....	x
List of Abbreviations.....	xii
Acknowledgments.....	xiv
Chapter 1: Introduction.....	1
Dissertation Objectives.....	2
Chapter 2: Co-Infection and Super-Infection of Mayaro and Zika Virus Demonstrates Viral Interference and Modulation of Infection and Dissemination Capacity in Mosquitos	5
Introduction.....	5
Methods	8
Mosquito Rearing	8
Virus Production	8
Vector Competence Assay	8
Statistical Analysis	10
Results	11
Discussion	16

Conclusion	18
Chapter 3: Mayaro Virus Infection Elicits an Innate Immune Response in <i>An. stephensi</i>	19
Introduction.....	19
Methods	20
<i>An. stephensi</i> Rearing	20
Viral Production and Infection via Bloodmeal	20
Transcriptomic Library Preparation and Sequencing.....	21
Small RNA Library Preparation and Sequencing.....	22
Transcriptomic RNA Sequencing Data Analysis	23
Small RNA Sequencing Data Analysis.....	23
Results and Discussion.....	24
Transcriptome	24
RNA Sequencing.....	24
Differential Expression.....	26
Gene Ontology.....	28
Small RNA	30
miRNA Identification.....	30
miRNA Differential Expression.....	34
miRNA Target Prediction	36
piRNA Identification.....	38
siRNA Identification	40

Conclusion	41
Chapter 4: Evolutionary Dynamics of North American Anophelines.....	42
Introduction.....	42
Methods	43
Species Identification.....	43
Genome Assembly.....	44
Phylogenetics and Scanning for Signatures of Selection	45
Results and Discussion.....	46
Genome Assembly	46
Phylogenetics.....	47
Evidence for Selection	51
Conclusion	52
Chapter 5: Conclusions and Future Directions	53
Appendix A: Supplement to Chapter 3	56
Appendix B: Supplement to Chapter 4	90
Bibliography.....	94

LIST OF FIGURES

Figure 2-1: Global distribution of MAYV and ZIKV.....	6
Figure 2-2: Co-infection and Super-infection in mosquitoes.....	7
Figure 2-3: Infectious virus loads of MAYV and ZIKV in co-infection and super-infection <i>Ae. aegypti</i>	12
Figure 3-1: Principal Components Analysis (PCA) on filtered read counts mapping to annotated genes from the Astel2 build of the <i>Anopheles stephensi</i> genome in Vectorbase.	25
Figure 3-2: Volcano plots visualizing differential expression of <i>Anopheles stephensi</i> transcripts in response to Mayaro infection.	27
Figure 3-3: Histograms representing the number of miRNAs shared between treatments.....	32
Figure 3- 4: Principal Components Analysis (PCA) on read counts mapping to miRNAs identified in the Astel2 build of the <i>Anopheles stephensi</i> genome in Vectorbase.	33
Figure 3-5: Volcano plots visualizing differential expression of identified <i>Anopheles stephensi</i> miRNAs in response to Mayaro infection.....	35
Figure 3-6: Histograms demonstrating read depth across the Mayaro virus genome for reads with a piRNA size profile (24 - 35 nt) and a siRNA size profile (18 – 23).	39
Figure 4-1: Maximum likelihood phylogenies with bootstrap values, gene concordance factors, and site concordance factors displayed on corresponding nodes.	48

Figure 4-2: Admixture plots created using ADMIXTURE software with inferred number of populations (K) equal to 5, 6, 7, and 8..... 50

Figure B-1: Maximum likelihood tree with *Ae. aegypti* as an outgroup..... 91

LIST OF TABLES

Table 2-1: <i>Ae. aegypti</i> orally challenged with MAYV and ZIKV simultaneously or serially to investigate the effect of CI and SI on vector competence.	11
Table 2-2: IR, DIR and TE for double positive mosquitoes in CI and SI experiments.	15
Table 4-1: Genome sizes, N50, and BUSCO scores for all samples sequenced.	46
Table A-1: Includes number of mosquitoes in each treatment and time point and associated mortality.....	56
Table A-2: Nanodrop readings for all RNA extractions collected, pooling scheme, and qPCR data using primers specific for Mayaro virus strain BeAn to confirm infection status	57
Table A-3: Differentially expressed transcripts from the <i>Anopheles stephensi</i> Astel2 genome. Only those with logFC greater than +/- 2	59
Table A-4: GO term overrepresentation for differentially regulated transcripts.....	62
Table A – 5: Read counts to known miRNAs and total library sizes	65
Table A – 6: Read counts to novel miRNAs.....	67
Table A – 7: Differentially expressed miRNAs	69
Table A-8: Overrepresented GO terms represented by targets of significantly regulated miRNAs	70
Table A-9: miRNA sequences.....	72
Table A-10: miRNA targets in the Astel2 genome assembly	74
Table B-1: Mitochondrial genome sequence alignment	90
Table B-3: Gene ontology overrepresentation results for selected for terms identified	

by aBSREL 92

Table B-4: Table B-3: Gene ontology overrepresentation results for selected for terms identified

by MEME 93

LIST OF ABBREVIATIONS

MAYV: Mayaro Virus

M: Mayaro Virus

ZIKV: Zika Virus

Z: Zika Virus

DENV: Dengue Virus

CHIKV: Chikungunya Virus

SV: Sindbis Virus

ISV: Insect Specific Virus

U: Uninfected (Control)

CI: Co-Infection

SI: Super-Infection

IR: Infection Rate

TE: Transmission Efficiency

DIR: Dissemination Rate

VC: Vector Competence

MOI: Multiplicity of Infection

dpi: Days Post Infection

FFU: Focus Forming Unit

RNAi: RNA Interference

miRNA: Micro RNA

piRNA: PIWI Interacting RNA

PCA: Principal Components Analysis

GLM: General-Linearized Model

GO: Gene Ontology

FBS: Fetal Bovine Serum

Log2FC: Log2 Fold Change

MEME: Mixed Effects Model of Evolution

aBSREL: adaptive Branch-Site Random Effects Likelihood

ACKNOWLEDGMENTS

This work was supported by NIH grants R01AI150251, R01AI128201, R01AI116636, USDA Hatch funds (Accession #1010032; Project #PEN04608), and a grant with the Pennsylvania Department of Health using Tobacco Settlement Funds to Jason L. Rasgon, BBSRC awards BB/T001240/1 and BB/V011278/1, Royal Society Wolfson Fellowship RSWF\R1\180013, NIH grants R21AI138074, URKI grants 20197 and 85336, EPSRC grant V043911/1, and NIHR grant NIHR2000907 to Grant Hughes. Cory Henderson was supported by an NSF graduate research fellowship program award (ID 2018258101). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, the Department of Health or Public Health England.

CHAPTER 1: INTRODUCTION

Vectors are living organisms which can transmit disease causing agents between vertebrate hosts. Mosquitoes transmit disease-causing pathogens during the act of bloodfeeding, which is required by many species for reproduction [1]. Vector borne diseases generally fall into the broad categories of parasites, bacteria, or viruses, and result in over 700,000 deaths per year, primarily in tropical and sub-tropical regions and disproportionately affecting the poorest populations residing in these regions [2].

Anopheles mosquitoes are the principal vectors for malaria and lymphatic filariasis, which historically impacted the Americas but now the region has the lowest malaria endemicity in the world [3]. Due to the near absence of malaria in the Americas and because so few examples of arboviral transmission by *Anopheles* mosquitoes have been observed in natural or laboratory settings, they are generally overlooked during surveys for pathogen presence and transmission in regions where they are actively circulating, potentially masking their involvement in transmission systems. *Anopheles* mosquitoes do demonstrate limited evidence for viral transmission; in natural contexts *Anopheles gambiae* and *funestus* act as the primary transmitting vectors for O'nyong'nyong virus, and in laboratory contexts they have demonstrated the capacity for transmission of Mayaro (MAYV) and Chikungunya virus (CHIKV) [4-6].

MAYV is a mosquito-borne, enveloped positive-sense single-stranded RNA virus in the genus *Alphavirus*, first isolated from the blood of five febrile workers in Mayaro county, Trinidad in 1954 [6]. Symptoms of MAYV infection are similar to other arboviral infections such as Dengue (DENV) or CHIKV, and include rash, fever, retro-orbital pain, headache, diarrhea, and

arthralgia [7]. While no epidemics or outbreaks with MAYV being the causative agent have been recorded outside of South America, there have been imported cases reported in the Netherlands, Germany, France, and Switzerland [8-11], which demonstrates a need to understand the capacity for the virus to spread into naïve regions, such as the United States.

The principal mosquitoes transmitting MAYV naturally are thought to be the canopy-dwellers of the genus *Haemogogus*, maintaining the sylvatic cycle between non-human primates as primary hosts and birds as secondary hosts [12]. Human infections are sporadic due to the rare display of anthropophilic biting behaviors by *Haemogogus* mosquitoes, with transmission due to these species primarily occurring in rural regions with close proximity to forests [13]. Vector competence studies have identified anthropophilic and urban adapted species such as *Aedes aegypti* and *Ae. albopictus*, as well as the malaria parasite transmitters *Anopheles gambiae*, *An. stephensi*, *An. freeborni*, and *An. quadrimaculatus*, as being competent vectors for MAYV under laboratory conditions [14-17].

In this dissertation I strive to advance our understanding of arboviral infection dynamics in mosquito vectors, with a focus specifically on MAYV in its canonical vector species as well as the role that anopheline mosquitoes could play in transmission in naïve regions. The specific objectives for each chapter are described in detail below.

DISSERTATION OBJECTIVES

In Chapter 2 I investigate the infection dynamics of MAYV and ZIKV in their shared mosquito vector *Aedes aegypti*. ZIKV has emerged as a major arboviral public health concern due to its capacity to cross the placenta effecting developing fetus in pregnant women [18, 19]. Overlapping distribution of MAYV and ZIKV endemicity and the presence of shared vector

species in the form of *Aedes* mosquitoes suggests the potential for co-infection of a vector mosquito by both pathogens as well as subsequent transmission to humans. The ability of mosquito vectors to carry multiple pathogens simultaneously in nature has been demonstrated in field-collected *Aedes albopictus* [20], and under laboratory conditions there is evidence to show that concurrent infection and transmission of multiple arboviruses is possible [21 – 24]. Here I present data from co-infection and super-infection of *Ae. aegypti* mosquitoes with MAYV and ZIKV to study their interactions and the consequent effects on vector competence.

Chapter 3 shifts focus more to the molecular basis for infection and transmission of MAYV by anopheline mosquitos. *Anopheles* are generally poor transmitters of arboviral pathogens, however there is increasing evidence that they can facilitate transmission of certain alphaviral pathogens [4 -6, 14]. The mechanisms by which MAYV can infect and be transmitted anopheline mosquitos have never been investigated in detail, however some work has been performed elucidating the molecular pathways involved in ONNV infection of *Anopheles* mosquitoes as this is the only arbovirus known to be naturally transmitted by *Anopheles* in nature. To address these questions, I perform a transcriptomic on MAYV infected *An. stephensi*, describing the molecular response at 2, 7, and 14 days post infection (dpi). I also detail the small RNA response elicited, describing differential expression of known and novel miRNAs as well as elucidating the siRNA mounted by the mosquito against MAYV.

In Chapter 4 I investigated the genomic context of North American anopheline vectors that potentially have the capacity to transmit MAYV. Vector management approaches require an understanding of the ecological, epidemiological, and biological context of the vector species, and with the increase in interest and application of genetic strategies, such as gene

drives utilizing CRISPR-Cas9, the availability of genomic sequences of vector species is becoming an immensely important aspect of vector control. I produced novel genome sequences for the North and Central American *Anopheles* species *An. crucians*, *An. freeborni*, *An. quadrimaculatus*, and *An. albimanus* to aid in furthering our understanding of the diverse biological traits that drive vectorial capacity in anophelines within the Americas.

In Chapter 5, the results from the studies described are summarized and for future research are explored.

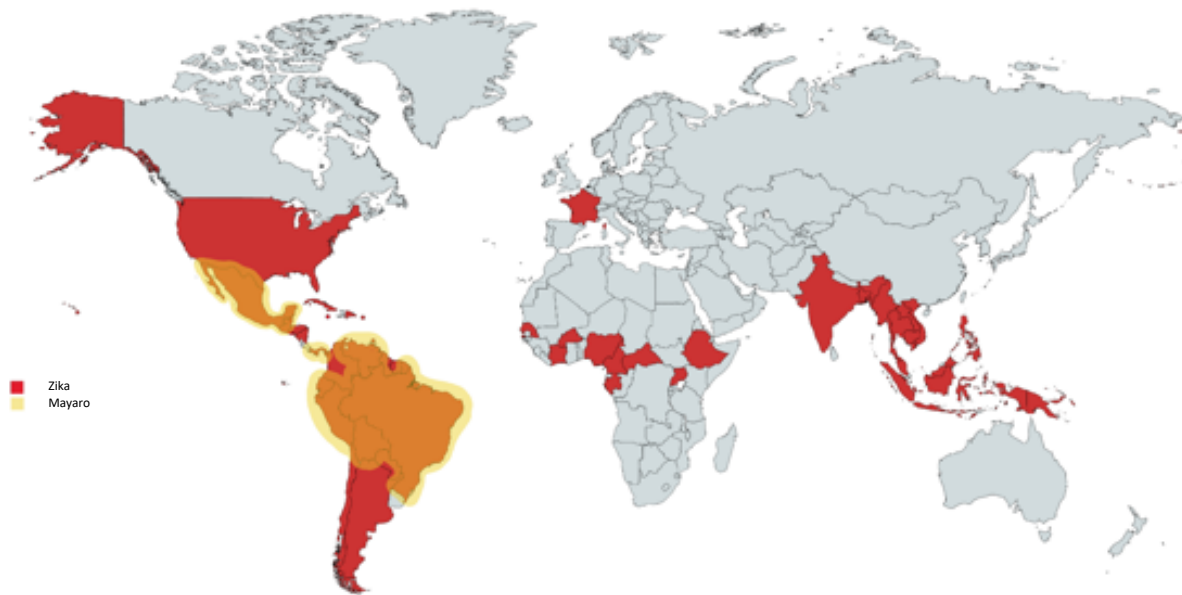
**CHAPTER 2: CO-INFECTION AND SUPER-INFECTION OF MAYARO VIRUS AND ZIKA VIRUS
DEMONSTRATE VIRAL INTERFERENCE AND MODULATION OF INFECTION AND
DISSEMINATION CAPACITY IN MOSQUITOS**

INTRODUCTON

The past several decades have seen a dramatic increase in the incidence of arboviral diseases, attributing to increased international trade and transport, climate change, and urban crowding. Increasing landscape fragmentation and anthropic activities are disturbing natural ecosystems, favoring the spill-over of new pathogens with epidemic potential to new geographic regions [25]. In the neotropical regions of the Americas, more than 150 different arboviruses have been reported and in many regions different arboviruses co-exist and co-circulate in sylvatic or urban cycles [26 – 28]. Despite this diversity and the co-circulation of multiple arboviruses in the same geographical area, some pathogens often go unnoticed in both vector and host due to inappropriate or lack of entomological surveillance strategies, overlapping clinical symptoms, and/or the absence of specific diagnostic tests [29]. At the same time, clinical studies have identified the co-circulation of multiple arboviruses such as DENV, ZIKV, CHIKV and MAYV in epidemic areas [30 – 33]. Further the presence of multiple arboviruses has been demonstrated in field collected *Aedes albopictus* [20], supporting the presence and capacity of vectors to carry multiple pathogens simultaneously in nature. Studies under laboratory conditions have also shown that concurrent infection and transmission of multiple arboviruses is possible [21 – 24]. Overlapping distribution and the presence of common vectors i.e. *Aedes* mosquitoes in South and Central America (Figure 2-1) suggest the

potential for co-infection among human patients, possibly mediated by the bites of co-infected mosquitoes.

Figure 2-1. Global distribution of MAYV and ZIKV. Map depicts countries with past and actual



detection of MAYV (yellow) and ZIKV (red). Map is based on data provided by the CDC as well as the literature, and it was generated using the free online tool

<https://mapchart.net/detworld.html>.

Simultaneous exposure of a vector to multiple pathogens through a single infection event is considered co-infection (CI), while serial infection of the vector with different pathogens during subsequent feeding events is considered super-infection (SI) (Figure 2-2).

When multiple arboviruses infect the same vector, simultaneously or subsequently, three different viral interaction can occur: a) an agonistic interaction that result in a partial or total inhibition of one or both viruses, b) a synergic interaction that enhance one or both viruses and c) a neutral co-existence of both viruses without any alteration of viral fitness [34]. Here, I co-infect and super-infect *Aedes aegypti* mosquitoes with MAYV and ZIKV to study their interactions and the consequent effects on vector competence.

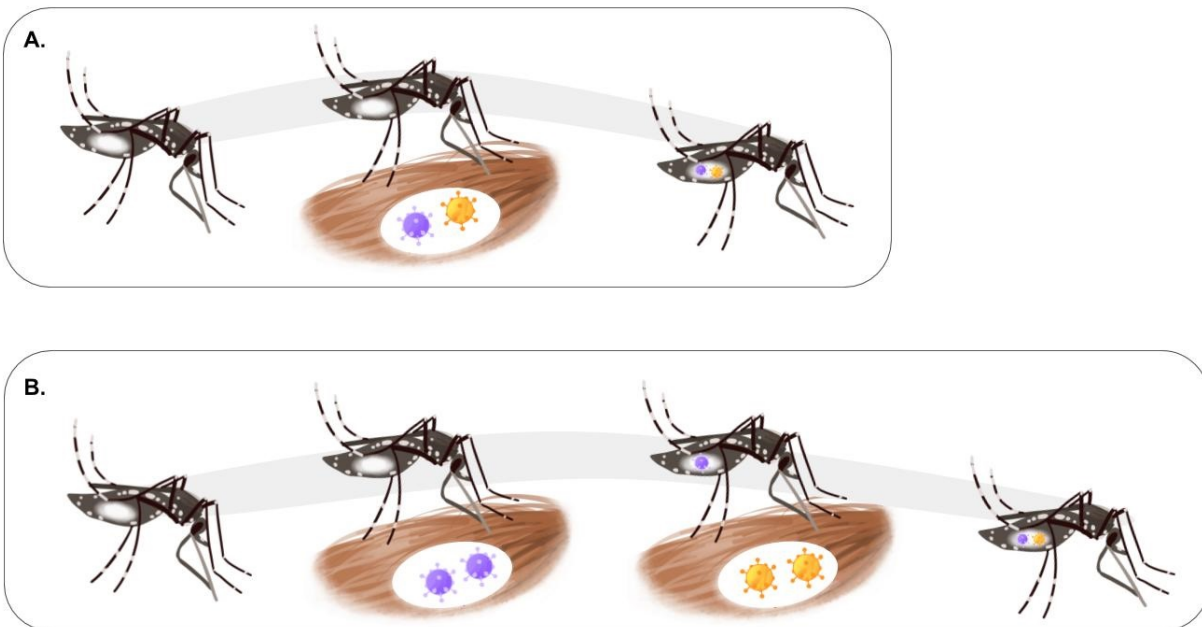


Figure 2-2. Co-infection and Super-infection in mosquitoes. Competent mosquitoes get infected with multiple arboviruses during a single feeding event (Co-infection; Figure 2-2a) or multiple sequential feeding events (Super-infection; Figure 2-2b).

METHODS

Mosquito Rearing

Ae. aegypti (Rockefeller strain) was obtained from Johns Hopkins University. Mosquito colonies were reared and maintained at the Millennium Sciences Complex insectary (The Pennsylvania State University, University Park, PA, USA) under the following environmental conditions: 27°C ± 1°C, 12:12 h light:dark diurnal cycle at 80% relative humidity. The larvae were feed with Ground fish flakes (TetraMin, Melle, Germany). Adult mosquitoes have free access to 10% sucrose solution.

Virus Production

Zika virus strain MR766 (NR-50065; BEI Resources, Manassas, VA, USA) and Mayaro virus strain BE AR 505411 (NR-49910; BEI Resources, Manassas, VA, USA) were propagated on Vero cells, stocks solutions were aliquoted and stored at -70°C until used. Viral stock titers were obtained by the focus forming unit (FFU).

Vector Competence Assay

Unfed 5-7 days old female mosquitoes were separated in groups of approximately 80 individuals each and fed with virus spiked infected human blood via a glass feeder jacketed with 37°C distilled water for 30-45 min. In the CI experiment blood containing MAYV, ZIKV, or both viruses were presented to different mosquito groups (henceforth called M, Z and CI) (day 0). A control group fed with uninfected blood (U) was included as the negative control. After feeding, mosquitoes were anesthetized at 4°C, full-engorged females were selected and transferred to a clean card box cup with access to 10% sugar solution *ad libitum*. In the SI experiment, two sequent feeding events were scheduled. During the first feeding operation (day -6) 3 groups were

exposed to uninfected blood, 2 groups to ZIKV (Z) and 2 groups to MAYV (M). During the second feeding operation (day 0) all the groups were exposed to a different bloodmeals resulting in 7 different possible combinations: UM, UZ, MU, ZU, MZ, ZM and UU. After both feedings, full-engorged females were selected and housed as previously described.

Mortality for all groups was recorded every two days. At 7- and 14-days post-infection (dpi), living mosquitoes were anesthetized with triethylamine (Sigma, St. Louis, MO) for approximately 30 seconds and subsequently dissected by separating the legs from the body of the mosquito. After leg detachment, the mosquitoes were forced to salivate in a glass capillary filled with a 1:1 solution of 50% sucrose solution and fetal bovine serum (FBS) FBS for 30 minutes. Bodies and legs were collected in separate 2mL tube containing 1 ml of mosquito diluent (20% heat-inactivated FBS in Dulbecco's phosphate-buffered saline, 50 µg/ml penicillin/streptomycin, 50 ug/ml gentamicin, and 2.5 µg/ml fungizone) with a single zinc-plated, steel, 4.5 mm BB (Daisy, Rogers, AR, USA). Tissues were then homogenized at 30 Hz for 2 min using TissueLyser II (Qiagen GmbH, Hilden, Germany) and centrifuged for 30 sec at 11000 rpm. Saliva samples were collected in a 2ml tube containing 0.1 ml of mosquito diluent. All samples were stored at -70°C until used. Samples were analyzed by Focus Forming Unit assay.

Briefly, 30 µl of 10-fold serial dilutions of each homogenized tissue samples (body or legs; saliva samples were not diluted) were used to infect a Vero cell monolayer in a 96 wells plate. After 24 hours (for MAYV) or 48 hours (for ZIKV), cells were fixed, permeabilized and labeled using the monoclonal anti-CHIKV E2 envelope glycoprotein clone CHK-48 (which reacts with MAYV) (BEI Resources, Manassas, VA, USA) or the monoclonal anti-Flavivirus group antigen (Clone D1-4G2-4-15). CI and SI samples were analyzed twice, one for each specific antibody.

Subsequentially, the primary antibody was labeled with the Alexa-488 goat anti-mouse IgG secondary antibody (Invitrogen, Life Science, Eugene OR, USA) and green fluorescent foci observed and enumerated with an Olympus BX41 microscope equipped with an UPlanFI 4× objective and a FITC filter.

Infection rates (IR) (positive bodies/total of exposed mosquitoes), dissemination rates (DIR) (positive legs/positives bodies), and transmission efficiency (TE) (positive saliva/ total of exposed mosquitoes) were calculated for all the tested groups. In the co-infection experiment I compared M vs CI and Z vs CI. In the super-infection experiment I compared MU vs MZ, ZU vs ZM, UM vs ZM, and UZ vs MZ.

The titers of MAYV and ZIKV in body, legs and saliva of each mosquito in each group were calculated and expressed in FFU/mL. Viral titers in specific tissues were compared between group following the scheme used for the comparison of IR, DIR and TE and described in the previous paragraph.

Statistical analysis

GraphPad Prism software version 8.2.1 (441) was used to analyze all the data. Differences in the IR, DIR and TE of groups challenged in both experiments were analyzed by Fisher's exact test. Two-tailed Mann-Whitney U test was used to compare viral titers in body, legs, and saliva samples of different groups. A p value of < 0.05 was considered statistically significant.

RESULTS

Ae. aegypti were orally challenged with MAYV and ZIKV simultaneously or serially to investigate the effect of CI and SI on vector competence (Table 2-1). Infection rate (IR), Dissemination rate (DIR) and Transmission efficiency (TE) were analyzed and compared by quantitative live virus titration in the body, legs, and saliva respectively (Figure 2-3).

		Coinfection		
	Groups	IR & p-value	DIR & p-value	TE & p-value
MAYV positivity				
7dpi	M vs CI	14/23 (60%) vs 9/19 (47%); ns	13/14 (92%) vs 7/9 (78%); ns	1/23 (4%) vs 0/19; ns
14dpi	M vs CI	17/23 (73%) vs 12/19 (63%); ns	17/17 (100%) vs 11/12 (92%); ns	1/22 (4%) vs 2/19; ns
ZIKV positivity				
7dpi	Z vs CI	25/25 (100%) vs 14/19 (73%); p= 0.0107	18/25 (72%) vs 4/14 (28%); p= 0.0019	0/25 vs 0/19; ns
14dpi	Z vs CI	25/26 (96%) vs 12/19 (63%); p= 0.0064	24/25 (96%) vs 11/12 (92%); ns	8/26 (31%) vs 3/19 (16%); ns
		Superinfection		
MAYV positivity				
7dpi	MU vs MZ	12/12 (100%) vs 4/5 (80%); ns	12/12 (100%) vs 3/4 (75%); ns	3/12 (25%) vs 0/5; ns
14dpi	MU vs MZ	12/12 (100%) vs 7/8 (87%); ns	10/12 (83%) vs 6/7 (85%); ns	1/12 (8%) vs 1/8 (12%); ns
7dpi	UM vs ZM	9/14 (64%) vs 20/21 (95%) p= 0.0278	8/9 (88%) vs 18/20 (90%); ns	2/14 (14%) vs 0/21; ns
14dpi	UM vs ZM	13/14 (93%) vs 18/22 (81%); ns	13/13 (100%) vs 17/18 (94%); ns	0/12 vs 1/22 (4%); ns
ZIKV positivity				
7dpi	ZU vs ZM	14/14 (100%) vs 21/21 (100%); ns	14/14 (100%) vs 21/21 (100%); ns	9/14 (64%) vs 10/21 (48%); ns
14dpi	ZU vs ZM	11/11 (100%) vs 23/23 (100%); ns	10/11 (90%) vs 20/23 (87%); ns	7/11 (63%) vs 8/23 (35%); ns
7dpi	UZ vs MZ	20/20 (100%) vs 5/0 (100%); ns	17/20 (85%) vs 4/5 (75%); ns	0/20 vs 0/5; ns
14dpi	UZ vs MZ	17/17 (100%) vs 8/8 (100%); ns	17/17 (100%) vs 8/8 (100%); ns	9/17 (53%) vs 5/8 (62%); ns

Table 2-1. CI= coinfection. M= Mayaro virus. Z= Zika virus. U= Uninfected Blood. P= p-value. ns= not significant. In Superinfection experiment two letters code represent two sequential feeding, ie. MZ= Mayaro virus in the first bloodmeal and Zika virus in the second.

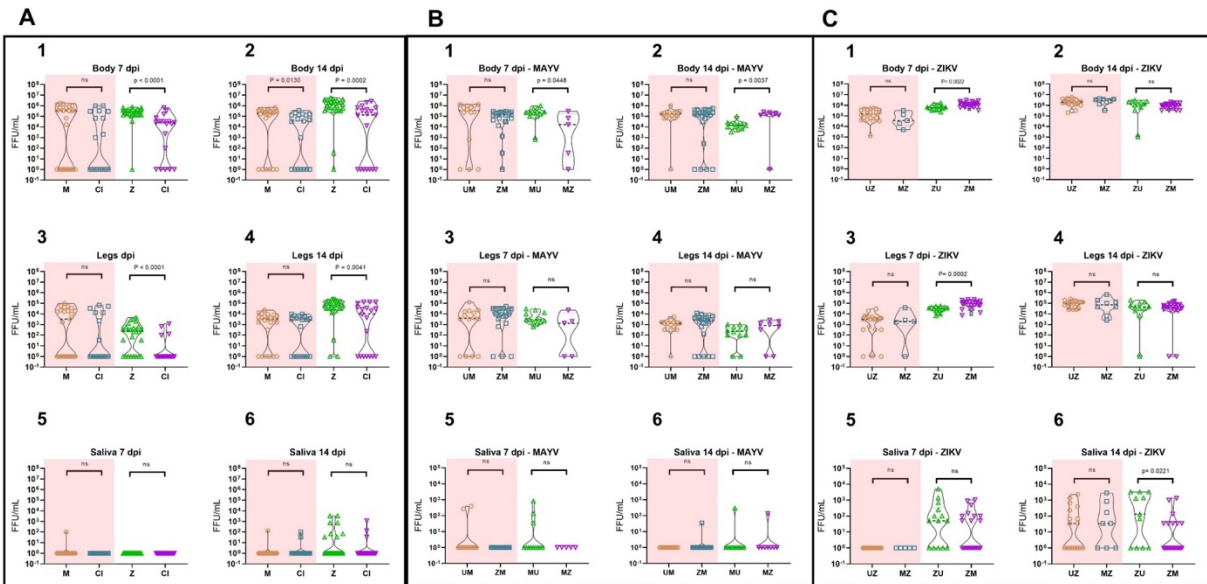


Figure 2-3. Infectious virus loads of MAYV and ZIKV in co-infection and super-infection *Ae. aegypti*. Panel A shown MAYV and ZIKV viral loads in the co-infection experiment. Titers obtained from single infected groups were compared with those from the coinfecting group. Panel B shows the viral loads of MAYV in the super-infection experiment. In this panel, two different comparisons were made: UM vs ZM and MU vs MZ. Panel C shows the viral loads of ZIKV in the super-infection experiment. In this panel two different comparisons were done: UZ vs MZ and ZU vs ZM. Titers are expressed in FFU/mL. CI= co-infection. M= Mayaro virus. Z= Zika virus. U= Uninfected Blood. p = p value. ns= not significant.

In the CI experiment, I observed a statistical reduction of IR for ZIKV in the co-infected group compared to the single infected group at 7- and 14-days post-infection (100% vs 73% p = 0.0107 and 96% vs 63% p = 0.0064 respectively). Similar reduction was also observed for the DIR at 7-days post-infection (72% vs 28% p =0.0019). I did not observe any statistically significant

variation of IR or DIR for MAYV in the co-infection group in comparison to the single infected group at any time point.

Zika virus titer in the body and legs of co-infected mosquitoes were statistically lower compared to the titer recorded in the body and legs of single infected mosquitoes at 7 dpi (body $p < 0.0001$; legs $p < 0.0001$) and 14 dpi (body $p = 0.0002$; legs $p = 0.0041$) (Figure 2-3 panel A graphs 1-2-3-4). Similarly, MAYV titers showed a significant reduction in the body of co-infected mosquitoes at 14 dpi ($p = 0.0130$) (Figure 2-3 panel A graph 2).

I did not find any statistically significant differences in the TEs of MAYV or ZIKV between single infected and co-infected groups or in their respective saliva viral titers.

These data suggest that the simultaneous intake of both viruses has a negative impact on their respective capacity to infect the midgut and the subsequent spread to the whole body of the mosquito, but not on the overall capacity of the mosquitoes to transmit both viruses in accordance to what is previously demonstrated for CHIKV/ZIKV co-infection [22, 23].

The SI experiment is designed to mimic the most common scenario in nature for mosquito species that blood feed multiple times during a single gonotrophic cycle, as the case of *Ae. aegypti* or *Ae. albopictus*. In this experiment I recorded only one statistically significant difference for IR, DIR or TE among all the combinations tested: the increase of the IR for MAYV in the ZM group compared to the UM group (95% vs 64% $p = 0.0278$) at 7 dpi. This result indicates a positive effect of a previous ZIKV infection on MAYV ability to establish a stable infection in the midgut.

I did not observe any statistically significant variation of MAYV titer in mosquitoes previously exposed to ZIKV compared to those previously exposed to uninfected blood. Conversely, our results indicate that ZIKV infection superimposed to MAYV infection can influence MAYV titer. I observed a statistically significant decrease of MAYV titer in body of MZ group at 7 dpi (group MZ; $p= 0.0448$; Figure 2-3 panel B graph 1) however, the same experimental group showed a statistically significant increase of MAYV body titer at 14 dpi (group ZM $p= 0.0037$; Figure 2-3 panel B graph 2).

The titer of ZIKV was positively affected by a subsequent MAYV infection at 7dpi. As shown in Figure 3 panel C graphs 1 and 3, I observed a statistically significant enhancement of ZIKV titer in the body and legs of ZM group compared to ZU group in ($p=0.0022$ and $p=0.0002$ respectively) at 7 dpi. These differences disappear at 14 dpi where the titers are almost identical in both body and legs groups. Saliva titers of ZIKV showed a reduction in viral titer in the saliva of the ZM group compared to the ZU group at day 14- dpi ($p = 0.0221$; Figure 2-3 panel C graph 6). Then, I calculated the IR, DIR and TE for double positive mosquitoes in CI and SI experiments (table 2-2).

		Coinfection		
	Groups	IR – Double positive	DIR - Double positive	TE - Double positive
7dpi	CI	7/19 (37%)	0/19*	0/19
14dpi	CI	7/19 (37%)	6/19* (32%)	0/19
		Superinfection		
7dpi	MZ	4/5 (80%)	3/5* (60%)	0/5
14dpi	MZ	7/8 (86%)	6/8* (75%)	1/8 (12%)
7dpi	ZM	20/21 (95%)	18/21* (86%)	0/21
14dpi	ZM	18/22 (82%)	16/22* (73%)	1/22 (5%)

Table 2-2. Infection rate (IR), dissemination rate (DIR), and transmission efficiency (TE) for double positive mosquitoes in CI and SI experiments. In the SI experiment, the rates of superinfected groups (MZ and ZM) have been compared with Fisher’s exact test at different time points. CI= coinfection. M= Mayaro virus. Z= Zika virus. U= Uninfected Blood. * = DIR was calculated as number of mosquitoes with double positive legs/ total number of exposed mosquitoes. In Superinfection experiment two letters code represent two sequential feeding, ie. MZ= Mayaro virus in the first bloodmeal and Zika virus in the second

In the CI experiment I observed the same IR for double positive mosquitoes at 7- and 14-dpi (7/19 (37%)). Conversely, mosquitoes with double disseminated infection were observed only at 14-dpi (6/19 (32%)). No transmission was observed at any time point for CI group.

In SI experiment, I observed higher IR of double positive mosquitoes compared to CI experiment as detailed in Table 2. In addition, double disseminated infection was recorded starting from 7-dpi, contrary to what was observed in the CI experiment. These data suggest that vectors are more permissive to double infection with multiple arboviruses in case of subsequent exposure rather than simultaneous exposure. Saliva samples tested positive for

both viruses only at 14-days post infection. In the ZM group, 1/22 (4.5%) saliva samples tested positive for both the viruses; in the MZ group 1/8 (12.5%) saliva samples tested positive for MAYV and ZIKV. Comparison between IR, DIR and TE of MZ and ZM groups did not show any statistically significant difference. Finally, I analyzed and compared the titers of double positive body and legs sample, however no statistically significant correlation between MAYV and ZIKV titers were found in both co-infection and super-infection experiments.

DISCUSSION

Here I show that *Ae. aegypti* can be infected with MAYV and ZIKV simultaneously after sequential or co-infection with both the arboviral pathogens and that simultaneous transmission of MAYV and ZIKV by *Ae. aegypti* is possible as a result of a super-infection event. Even though there is no data on MAYV/ZIKV co-infected mosquitoes in nature, pools of mosquitoes collected in different areas of Mato Grosso (Brazil) were double-positive for MAYV and DENV-4 [35]. In addition, reports of MAYV co-infected patients with CHIKV [32], ZIKV [36], and DENV-1 [37], indicates that MAYV is also co-circulating in urban and peri-urban areas, probably due to the high number of vector species that can contribute to its transmission cycle resulting in a high potential number of double infected mosquitos [14]. Currently, most surveillance programs pool and analyze mosquitoes to save time and resources. Unfortunately, this surveillance strategy makes it impossible to investigate the percentage of the mosquitoes carrying multiple arboviruses and their potential contribution to the transmission cycle of those viruses. The implementation of precise sampling strategies of field-collected mosquitoes would be fundamental to investigate the prevalence of multi-infected vectors in the endemic areas and correlate the data with the incidence of co-infected patients.

Our results showed a significant reduction in the IR of ZIKV during a co-infection. *Rückert et al.* has previously demonstrated a similar decrease in IR for ZIKV in a CI experiment with CHIKV [22], an *Alphavirus* belonging to the same antigenic complex of MAYV (Semliki Forest virus complex). Similarly, *Muturi et al.* demonstrated that in co-infected *Ae. albopictus*, suppressed the replication of DENV-4 [38]. I did not observe any variation in the IR of MAYV during co-infection with ZIKV. Our data corroborate those of previous studies using different *Flavivirus/Alphavirus* co-infection models. In these studies, the authors demonstrated that IR and DIR of CHIKV were not influenced by the presence of ZIKV or DENV-2 [22, 39]. These data suggest that in co-infected mosquitoes the IR of flaviviruses is negatively affected by the presence of an *Alphavirus* but not the opposite. However, additional studies including multiple *Flavivirus/Alphavirus* couples, are needed to evaluate this hypothesis and to identify the molecular basis of this phenomenon.

Interestingly, I observed an increase of MAYV IR in previously ZIKV-infected mosquitoes (ZM group). Although the molecular mechanism is out of this study's scope, I hypothesize that a prior ZIK exposure to a mosquito may prime the mosquito's immune responses and other molecular machinery, creating a suitable environment for the MAYV infection and replication process [24]. Further, a high viral titer requirement for mosquito feeding experiments could possibly explain the absence of differences in the IR for ZIKV (all groups present IR= 100%) and had limited our capacity in to evaluate if previous or subsequent MAYV infection have the same effect on IR for ZIKV.

To note, our experimental models assume that the vector obtains a blood meal and gets infected with a similar titer of both viruses - a rare scenario in the field where the possible

combination of viruses and their respective titers are incredibly variable. More studies are required to evaluate the impact of different titer combinations for MAYV and ZIKV on the VC in case of CI or SI. In a related study *Muturi et al.* showed that using different virus titer ratios between DENV-4 and SINV might result in opposite effect on the replication of one or both viruses in *in vitro* [38]. This result suggests that not only the viral titers (or MOI in case of *in vitro* studies) but also their ratio profoundly influence the VC in case of CI or SI.

CONCLUSION

In this chapter I show that *Ae. aegypti* can be infected with MAYV and ZIKV simultaneously after sequential or co-infection with both the arboviral pathogens and that simultaneous transmission of MAYV and ZIKV by *Ae. aegypti* is possible as a result of a super-infection event. More precisely, our results showed a significant reduction in the IR of ZIKV during a co-infection and I observed an increase of MAYV IR in previously ZIKV-infected mosquitoes (ZM group). These data suggest that in co-infected mosquitoes the IR of flaviviruses is negatively affected by the presence of an *Alphavirus* but not the opposite. As highlighted before, the number of variables and their possible combinations (pathogens, titers, time-lapse between the exposition and different extrinsic periods of incubations) might result in very different outputs. For this reason, co-infection and super-infection are poorly understood mechanisms, especially when applied to the field for which very few data are available. However, it is imperative to consider the interaction of multiple viruses in future VC studies.

CHAPTER 3: MAYARO VIRUS INFECTION ELICITS AN INNATE IMMUNE RESPONSE IN *AN.*

STEPHENSII

INTRODUCTION

As arboviral pathogens are transmitted between hosts primarily by arthropod vectors, transmission requires the virus to infect and disseminate from the midgut and salivary glands of the mosquito following an infectious bloodmeal [1]. The molecular underpinnings controlling why MAYV and these closely related viruses can infect *Anopheles* salivary glands is of epidemiological interest, yet remains poorly understood. A more complete understanding of this phenomenon requires investigation of the molecular pathways involved in viral infection of anopheline mosquitoes. Recent transcriptomic studies have identified a number of genes involved in classical immune pathways, RNA interference (RNAi), metabolism, energy production, and transport as being regulated in response to arboviral infection of mosquitoes [46-49]. In addition, studies focusing on small RNA identification and regulation have identified RNAi activity, such as miRNA, piRNA, and siRNA expression, in response to infection of mosquitoes by arboviruses [50-53].

The available evidence suggests that, should MAYV be introduced into a naïve region, outbreaks and epidemics of the resulting disease could be driven by anopheline vectors [3, 14, 54]. Because anopheline mosquitoes could act as the primary transmitting vectors for MAYV, this study also provides an opportunity to understand how vector competence might emerge in this system and provide insight into why *Anopheles* are generally poor viral transmitters when compared to *Aedes* mosquitoes. I used RNA sequencing to study the transcriptomic and small

RNA responses of *An. stephensi* to infection with MAYV via infectious bloodmeal at 2, 7, and 14 days post infection (dpi).

METHODS

Anopheles stephensi Rearing

Protocols pertaining to mosquito rearing and presentation of infectious bloodmeal has been described elsewhere [14]. Briefly, *An. stephensi* (Liston strain) were provided by Johns Hopkins University (Baltimore, MD, USA). Mosquito colonies were reared and maintained at the Millennium Sciences Complex insectary (The Pennsylvania State University, University Park, PA, USA) at 27°C ±1°C, 12 hour light 12 hour dark diurnal cycle at 80% relative humidity in 30×30×30-cm cages. Ground fish flakes (TetraMin, Melle, Germany) were used to feed larvae, and upon emergence adult mosquitoes were maintained with a 10% sucrose solution.

Viral Production and Infection via Bloodmeal

Mayaro virus strain BeAn 343102 (BEI Resources, Manassas, VA, USA) was utilized in this study, a genotype D strain originally isolated from a monkey in Para, Brazil, in May 1978. Virus-infected supernatant was aliquoted and stored at –80°C until used for mosquito infections. Viral stock titers were obtained by focus forming assay (FFA) technique. Adult female mosquitoes at 6 days post emergence that had not previously blood-fed were used for experimentation. Mosquitoes were allowed to feed on either human blood spiked with MAYV at 1*10⁷ FFU/mL or a control bloodmeal with no virus via a glass feeder jacketed with 37°C distilled water for 1 h. *An. stephensi* was chosen as the model vector in this study as it was the most susceptible to infection in Brustolin et. al. 2018, and as they bloodfed most readily on the membrane feeder required for infection [14].

At 2, 7, and 14 days post infection, mosquitoes were anesthetized with triethylamine (Sigma, St. Louis, MO, USA) and RNA was extracted from each individual mosquito using mirVana RNA extraction kit (Life Technologies) applying the protocol for extraction of total RNA. Infection was confirmed via qPCR using primers published by Wiggins et. al. 2018 (Forward: 5'-TGGACCTTTGGCTCTTCTTATC-3', Reverse: 5'-GACGCTCACTGCGACTAAA-3') [102], a CT value of 20 or less was used to confirm infection (Appendix A). 3 pools of total RNA were created for each time point and infection status to be used for library preparation, each consisting of 750 ng of RNA from 4 mosquitoes for a total of 3 mg per pool as confirmed via nanodrop. The protocol for mosquito rearing, viral production, and infection via bloodmeal is described in more detail in Brustolin et al. 2018 [14].

Transcriptomic Library Preparation and Sequencing

All pools were sent to University of Texas Medical Branch for library preparation where total RNA was quantified using a Qubit fluorescent assay (Thermo Scientific) and RNA quality was assessed using an RNA 6000 chip on an Agilent 2100 Bioanalyzer (Agilent Technologies). See Etebari et al. 2017 for more detail on library preparation and sequencing [47]. 1 mg of total RNA per pool was poly-A selected and fragmented using divalent cations and heat (940 C, 8 min). The NEBNext Ultra II RNA library kit (New England Biolabs) was used for RNA-Seq library construction. Fragmented poly-A RNA samples were converted to cDNA by random primed synthesis using ProtoScript II reverse transcriptase (New England Biolabs). After second strand synthesis, the double-stranded DNAs were treated with T4 DNA polymerase, 5' phosphorylated and then an adenine residue was added to the 3' ends of the DNA. Adapters were then ligated to the ends of these target template DNAs. After ligation, the template DNAs were amplified (5-

9 cycles) using primers specific to each of the non-complimentary sequences in the adapters. This created a library of DNA templates that have non-homologous 5 'and 3 'ends. A qPCR analysis was performed to determine the template concentration of each library. Reference standards cloned from a HeLa S3 RNA-Seq library were used in the qPCR analysis. Cluster formation was performed using 15.5-17 billion templates per lane using the Illumina cBot v3 system. Sequencing by synthesis, paired end 75 base reads, was performed on an Illumina NextSeq 5500 using a protocol recommended by the manufacturer.

Small RNA Library Preparation and Sequencing

Small RNA libraries were created using the New England Biolabs small RNA library protocol. See Saldaña et. al. 2017 for more information on small RNA sequencing [51]. Library construction used a two-step ligation process to create templates compatible with Illumina based next generation sequence (NGS) analysis. Where appropriate, RNA samples were quantified using a Qubit fluorometric assay. RNA quality was assessed using a pico-RNA chip on an Agilent 2100 Bioanalyzer. Library creation uses a sequential addition of first a 3 'adapter sequence followed by a 5 'adapter sequence. A cDNA copy was then synthesized using ProtoScript reverse transcriptase and a primer complimentary to a segment of the 3 'adapter. Amplification of the template population was performed in 15 cycles (94°C for 30 sec; 62°C for 30 sec; 70°C for 30 sec) and the amplified templates were PAGE (polyacrylamide gel electrophoresis) purified (147 bp DNA) prior to sequencing. All NGS libraries were indexed. The final concentration of all NGS libraries was determined using a Qubit fluorometric assay and the DNA fragment size of each library was assessed using a DNA 1000 high sensitivity chip and an

Agilent 2100 Bioanalyzer. Single end 75 base sequencing by synthesis on an Illumina NextSeq 5500.

Transcriptomic RNA Sequencing Data Analysis

Raw sequencing data was uploaded to the ICS-ACI high performance computing cluster at Pennsylvania State University to perform all computational analyses. Transcriptomic libraries had adapters trimmed and low-quality bases removed using Trimmomatic read trimming software with base settings [55]. Quality trimmed reads were aligned to the current build of the *Anopheles stephensi* Indian strain genome in Vectorbase (AsteI2) using the STAR RNA sequencing aligner [56]. Reads less than 75 bp in length and with a mapping quality of less than 20 were dropped from the analysis, and read counts were calculated in R using the rSubread package [57], following which a principal components analysis was performed and differential expression conducted using a negative binomial GLM with the EdgeR package [58]. Contrasts considered in the GLM were infected against control at 2, 7, and 14 dpi, and differences between 2 -7 dpi and 7 – 14 dpi for infected treatments corrected for the response from the control treatments between the same time points. Gene IDs that were differentially expressed with a log₂FC value of +/- 1 and P value < 0.05 were uploaded to g:Profiler to run GO term overrepresentation analysis [59].

Small RNA Sequencing Data Analysis

Small RNA libraries had adapters trimmed using Trimmomatic and were subsequently passed into the miRDeep2 pipeline to identify novel and known miRNAs in all samples and determine expression of all known and novel miRNAs at each time point and treatment status [55, 60]. Novel miRNAs with a miRDeep score of less than 3, a minimum free energy value of

less than -20 , or a non-significant Randfold p-value were considered false IDs and excluded from further analysis. miRNA targets were identified in the *Astel2* genome using miRanda software package with a score threshold of 140, free energy cutoff of -20 , gap open penalty of -9 , gap extend penalty of -4 , and strict mapping in the seed region [61]. Differential expression of miRNAs in response to infection status and time point was conducted using a negative binomial GLM with the EdgeR package and contrasts as described for the transcriptomic analysis [58]. miRNAs which were differentially expressed with $\log_2FC \pm 1$ and P value < 0.05 had their miRanda genomic targets uploaded to g:Profiler to determine if any GO terms were overrepresented by transcripts potentially regulated by differentially expressed miRNAs [62]. piRNAs and siRNAs were isolated from the small RNA libraries by selecting all 18 – 24 nt reads (siRNA) and 24 – 35 nt (piRNA) reads from the trimmed datasets and filtering out all identified mature miRNAs, and those mapping to the MAYV NC_003417.1 genome were considered potential piRNAs or siRNAs. piRNA and siRNA alignment to the *Astel2* genome was performed using Bowtie RNA sequencing aligner within the Mosquito Small RNA Genomics (MSRG) resource pipeline [63]. Observation of fastQC output for Control Day 7 Replicate 1 small RNA sequencing revealed poor sequencing results, so this replicate was omitted from all analyses in the small RNA focused portions of this study [64].

RESULTS AND DISCUSSION

Transcriptome

RNA Sequencing

I assayed genome-wide gene expression in pools of *An. stephensi* (Liston strain) experimentally infected with MAYV at 2, 7, and 14 dpi, along with blood fed uninfected

negative controls. RNAseq libraries were sequenced on the Illumina NextSeq 5500 platform, yielding 20.- 30 million paired end reads per library.. Principal components analysis (PCA) performed on read counts of each annotated gene in the *An. stephensi* (Indian strain) reference transcriptome (AsteI2) at each time point distributed infected and control samples into distinct groups (Figure 3-1).

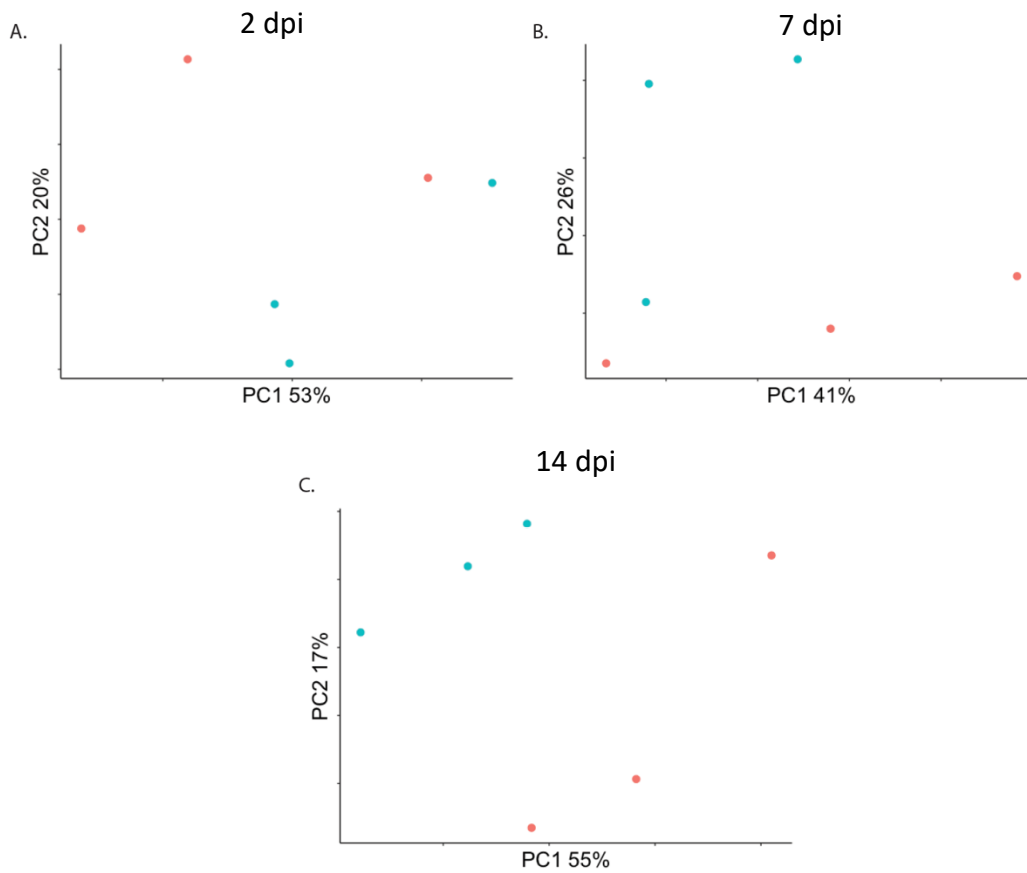


Figure 3-1: Principal Components Analysis (PCA) on filtered read counts mapping to annotated genes from the AsteI2 build of the *Anopheles stephensi* genome in Vectorbase. A., B., and C. are read counts from samples in the 2, 7, and 14 dpi groupings respectively. In all PCAs, blue is Mayaro infected, and red are control.

Differential Expression

To determine which *An. stephensi* genes are differentially expressed between Control and MAYV infection status during different time points, I applied the general-linearized model (GLM) in the EdgeR package on filtered and normalized read counts mapping to the Astel2 genome (Appendix A; Figure 3-2). Contrasts considered in the GLM were infected compared to control at 2, 7, and 14 dpi, and differences between 2 -7 dpi and 7 – 14 dpi for infected treatments correcting for results from control treatments between those time points. Genes were considered significantly regulated if they had a log fold-change (log₂FC) value of +/- 1 and P value < 0.05.

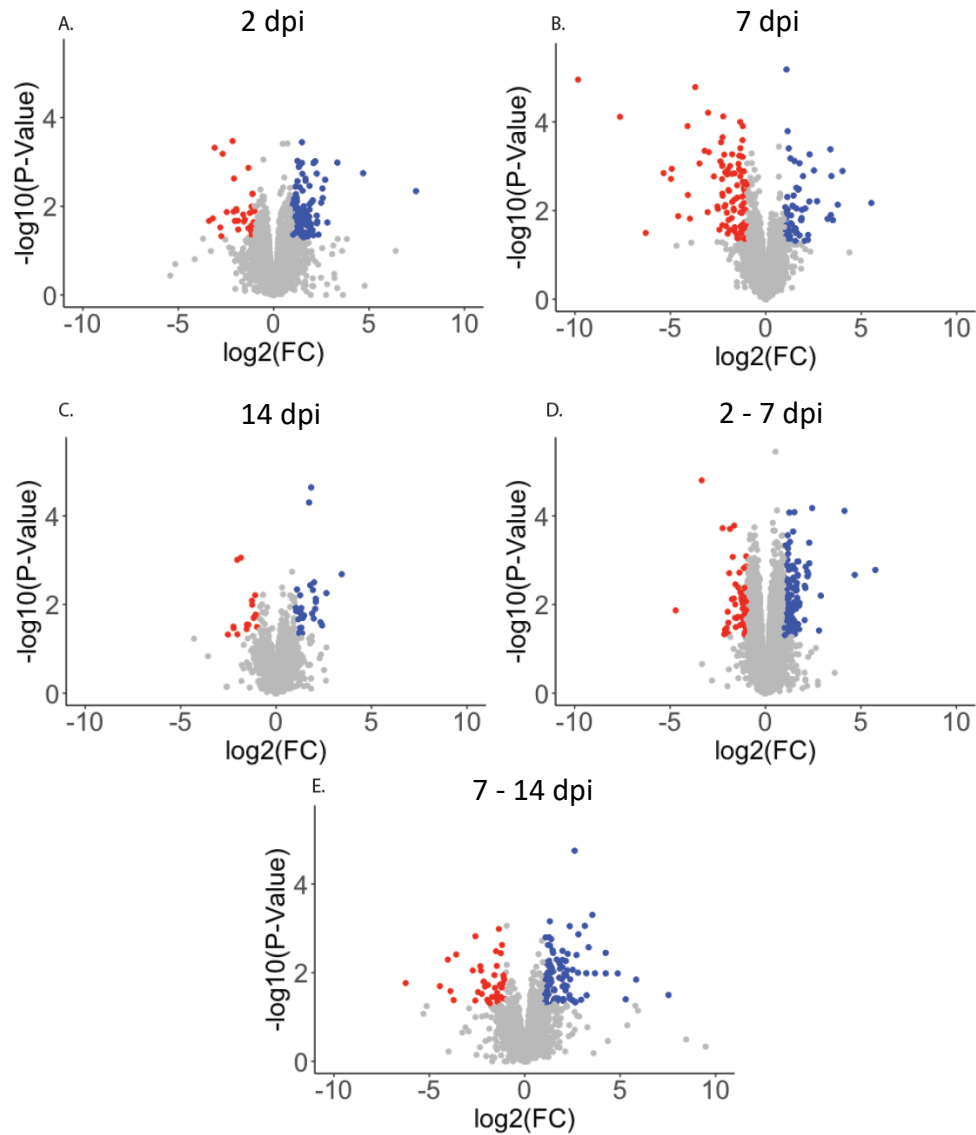


Figure 3-2: Volcano plots visualizing differential expression of *Anopheles stephensi* transcripts in response to Mayaro infection. The Y-axis shows $-\log_{10}$ transformed P-values, and the X-axis shows \log_2 transformed fold change values. Red points represent transcripts downregulated by more than $-1 \log_2\text{FC}$ in response to infection with a $\text{FDR} < 0.05$, while blue points are transcripts upregulated by more than $1 \log_2\text{FC}$ in response to infection with a $\text{P value} < 0.05$. A. - C. are transcripts regulated in the 2 dpi, 7 dpi, and 14 dpi groupings respectively, while D. and E. are transcripts regulated in the infected treatment between 2 - 7 dpi and 7 - 14 dpi respectively.

There were 161 (64 enriched, 97 depleted), 45 (29 enriched, 16 depleted), and 204 (149 enriched, 55 depleted) of 10,313 annotated genes regulated between control and infected at 2, 7, and 14 dpi respectively. 3 genes were regulated in the same direction at each time point, 2 enriched (ASTEI09037 – an ortholog of drICE caspase, and ASTEI03083 – a likely alkaline phosphatase) and 1 depleted (ASTEI04716 – ortholog of lush odorant receptor). The gene with the strongest response to infection at any time point was ASTEI04601 (no *Drosophila* ortholog but conserved in mosquitoes) at 2 dpi with a log₂FC of -9.8 and the most enriched gene was ASTEI04639 (no specified product) at 14 dpi with log₂FC of 5.7. When considering changes between time points for the infected treatment when controlling for the response from the uninfected treatments, there were 96 positively and 44 negatively regulated genes between 2 - 7 dpi, and 129 upregulated and 32 downregulated genes between 7 - 14 dpi. Regulated transcripts for 2-7 dpi ranged from -6.2 (ASTEI08168 – serine protease ortholog in *Ae. albopictus*), to 7.5 (ASTEI09252 – serine type endopeptidase ortholog in many *Aedes* and *Anopheles*) log₂FC in terms of magnitude of expression, and -3.4 (ASTEI10804 – Histone H2A) to 7.4 (ASTEI04639 – no specified product) log₂FC for 7-14 dpi. When considering a FDR threshold as a multiple testing correction very few transcripts in any contrast can be considered significantly regulated; 3 transcripts at 2 dpi (ASTEI04601, ASTEI05497, ASTEI05732) and 2 transcripts at 14 dpi (ASTEI00644, ASTEI08604) fall below a FDR < 0.1 threshold for significance.

Gene Ontology

A gene ontology (GO) over-representation analysis was performed using g:Profiler on gene IDs which were significantly enriched or depleted in any considered contrast in the GLM described above when using a P value cutoff of 0.05 and any overrepresented GO terms with an

FDR < 0.5 were considered significant (Appendix A) [59]. At 2 dpi significantly regulated molecular function terms were overrepresented by peptidase activity, specifically serine type endopeptidase activity, however this overrepresentation is not observed when just considering depleted or enriched genes it is only present when considering all regulated genes at 2 dpi together. Depleted genes at 2 dpi are overrepresented by cell-cell adhesion terms in the biological function category, and components of the membrane for cellular component, enriched genes are not significantly biased for any terms. At 7 dpi molecular function terms related to odorant binding and carboxylic acid binding were overrepresented in depleted genes, and for enriched genes however with a slightly less than significant FDR of 0.06. Molecular function terms relating to dioxygenases, transferases, and oxidoreductases are overrepresented by depleted genes at 7 dpi, and different types of amino acid binding as well as serine type endopeptidase terms are overrepresented by enriched genes. Biological process terms related to nervous system processes, sensory perception of smell, and mannose metabolism are overrepresented by depleted genes at 7 dpi. At 14 dpi the enriched transcripts were biased for molecular function GO terms related to sensory perception, specifically perception of mechanical, light, and sound stimuli, while depleted transcripts were biased for those related to MAPK/JNK signaling cascades, apoptosis, and amino acid biosynthesis for molecular functions and peroxisome and nucleosome for cellular component terms. From 2 to 7 dpi molecular function terms related to serine protease activity were represented by both upregulated and downregulated genes, but primarily by upregulated genes, and upregulated genes were biased for cell membrane proteins for cellular component. 7 to 14 dpi had upregulated genes biased for molecular functions related to ATP dependent microtubule

activity and G-protein coupled receptors while downregulated genes were biased for different types of amino acid binding. Biological process terms overrepresented by upregulated genes between 7 and 14 dpi were related to sensory perception, and downregulated genes were overrepresented for various types of metabolic and biosynthetic processes, NF- κ B signaling, and the MAPK/JNK cascade.

Endopeptidases, specifically serine proteases were biased in enriched terms at 7 dpi and from 2 – 7 dpi, suggesting an activation of the Toll pathway as part of the innate humoral response to infection once the virus has had time to establish an infection in the mosquito [62]. Activation of serine proteases is not uncommon in pathogenic infection of insects, and has been identified specifically as upregulated in *Ae. aegypti* in response to dengue and Zika virus infection, and in *An. gambiae* and *An. coluzzii* in response to O'nyong'nyong virus infection [46-48, 66]. At late stages of infection there was depletion of the autophagic and apoptotic inducing JNK and MAPK cascades in addition to repression of JAK/STAT signaling pathways through repression of MAPK signaling. Autophagy and apoptosis both have demonstrated positive impacts on alphaviral replication [62, 67], suggesting another possible molecular response from the mosquito to prevent viral replication at late stages of infection.

Small RNA

miRNA Identification

I next identified novel miRNAs in the small RNA transcriptomes of the MAYV infected samples and controls using miRDeep [60]. I searched for matches in our sequencing reads to all miRNAs in the miRBase database for the species *An. gambiae*, *Aedes aegypti*, *Culex quinquefasciatus*, *Drosophila melanogaster*, *Bombyx mori*, *Apis mellifera*, and *Acyrtosiphon*

pisum. I found matches to 73 known miRNAs, all from *An. gambiae*, and 79 novel miRNAs identified across all samples, with between 12 – 21 percent of quality filtered reads mapping to identified miRNAs per-sample (Appendix A). I found no explicit relationship between diversity of miRNA population or percentage of library belonging to miRNA mapping reads and either dpi or infection status.

Of the 153 total miRNAs identified across all samples, 83 were present in at least one replicate per treatment (Figure 3-3). PCA of read counts normalized to library size for all identified miRNAs in each sample shows a correlation between infection status and grouping along the PC1/PC2 axis at each time point (Figure 3-4).

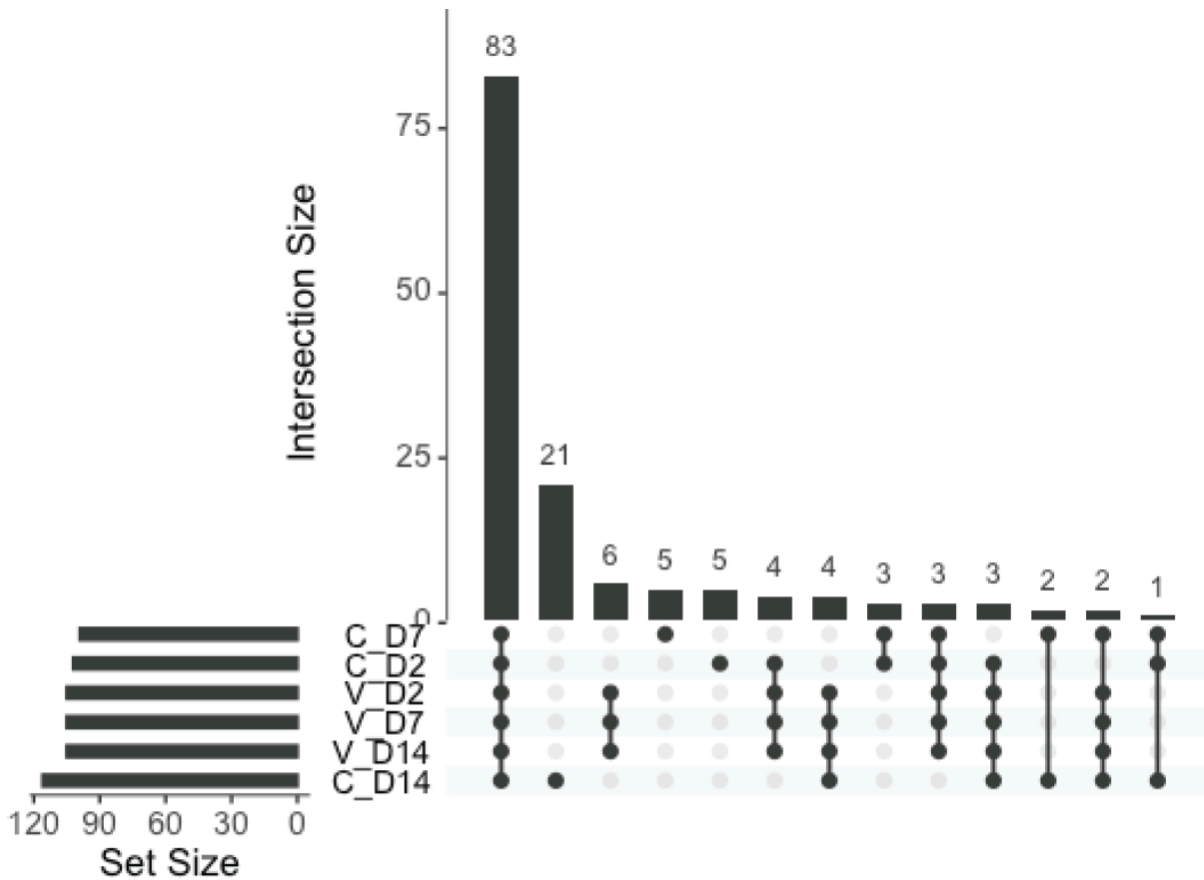


Figure 3-3: The top histogram represents the number of miRNAs shared between treatments (intersection size), and each row below the histogram represents a treatment. The lines connecting treatments below the top histogram represent treatments which share that number of miRNAs, and the histogram to the side of the treatments represents the number of miRNAs contained within each treatment.

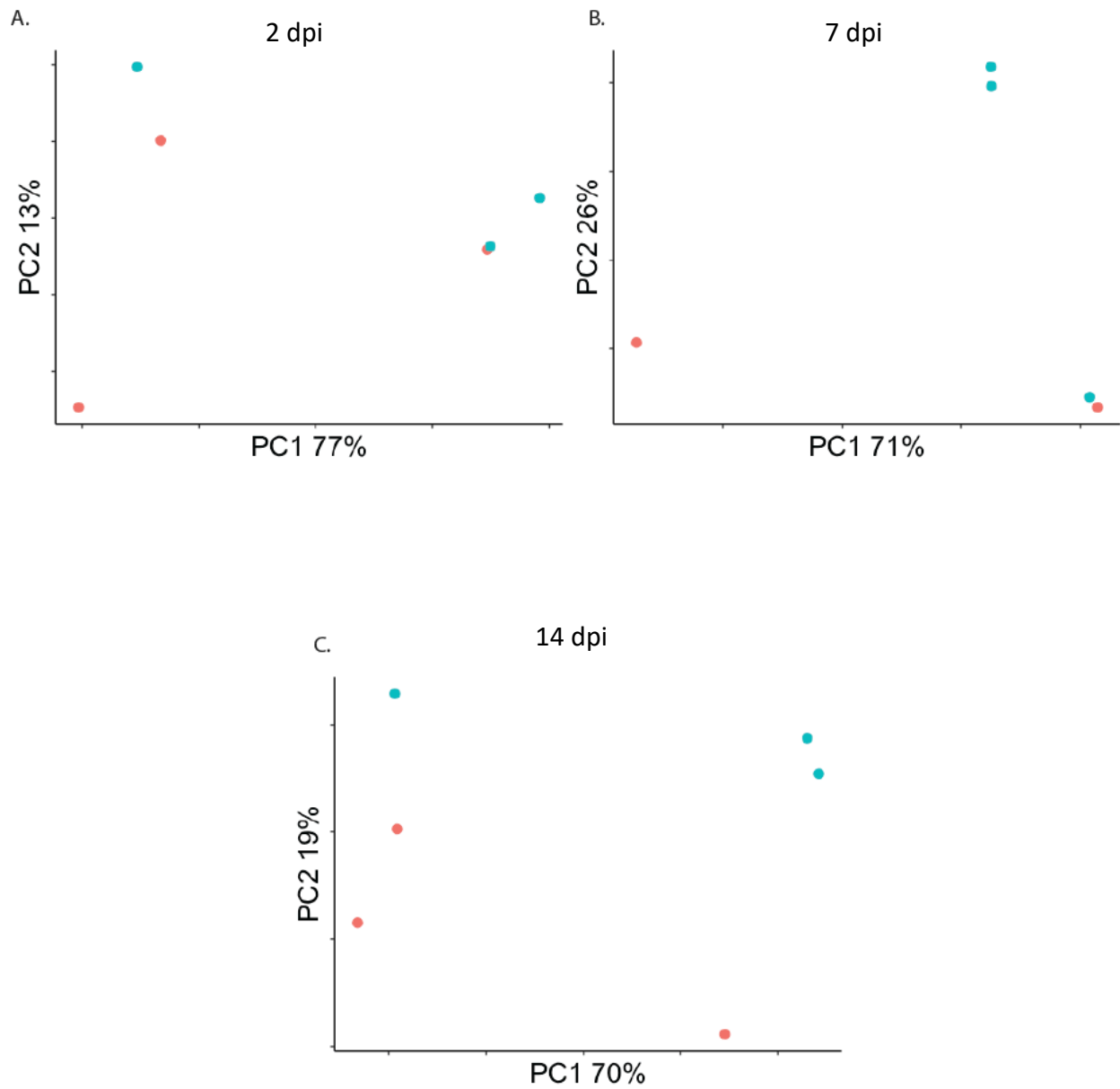


Figure 3- 4: Principal Components Analysis (PCA) on read counts mapping to miRNAs identified in the Astel2 build of the *Anopheles stephensi* genome in Vectorbase. A. - C. are the 2 dpi, 7 dpi, and 14 dpi groupings respectively. In all PCAs, blue is Mayaro infected, and red is control.

miRNA Differential Expression

I next identified known and novel miRNAs that were differentially expressed by infection status (Fig. 3-5; Appendix A). Contrasts considered in the GLM were infected against control at 2, 7, and 14 dpi, and differences between 2 -7 dpi and 7 – 14 dpi for infected treatments relative to control. miRNAs were considered differentially expressed by having a log fold-change (log₂FC) value of +/- 1 and P value < 0.05.

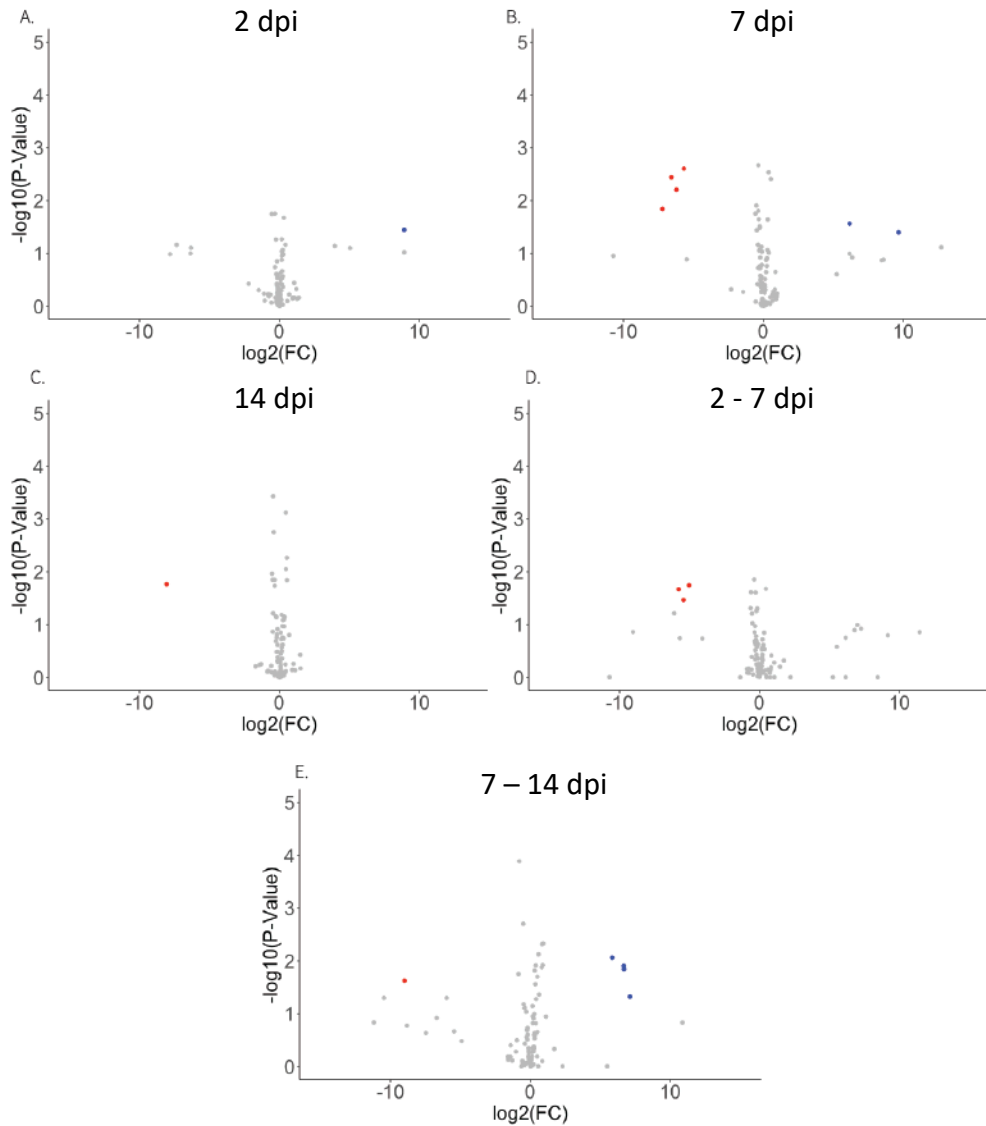


Figure 3-5: Volcano plots visualizing differential expression of identified *Anopheles stephensi* miRNAs in response to Mayaro infection. The Y-axis shows $-\log_{10}$ transformed P-values, and the X-axis shows \log_2 transformed fold change values. Red points represent transcripts downregulated by more than $-1 \log_2\text{FC}$ in response to infection with a $\text{FDR} < 0.05$, while blue points are transcripts upregulated by more than $1 \log_2\text{FC}$ in response to infection with a $\text{FDR} < 0.05$. A. - C. are the 2 dpi, 7 dpi, and 14 dpi groupings respectively, while D. and E. are miRNAs regulated in the infected treatment between 2 - 7 dpi and 7 - 14 dpi respectively.

Amongst all the comparisons, I observed a total of 8 miRNAs differentially regulated: novel miRNAs as-mirNOV10, as-mirNOV16, and as-mirNOV17 as well as known miRNAs aga-miR-286b, aga-miR-2944a, aga-miR-2944b, aga-miR-307, and aga-miR-309. as-mirNOV10 was upregulated at 2 dpi, as-mirNOV16 was upregulated at 7 dpi, and as-mirNOV 17 was downregulated at 14 dpi and between 7 – 14 dpi. The known miRNAs were downregulated as a group at 7 dpi and in the 2 - 7 dpi contrast but upregulated in the 7 – 14 dpi contrast.

The miRNA cluster miR-309/286/2944 has been found to be upregulated in *An. gambiae* in response to bloodfeeding [68 – 69], and to be associated with Argonaute proteins post-bloodmeal [68]. Experimental repression of aga- miR-309 can retard oocyte development [68], so its downregulation in response to MAYV infection may suggest that viral replication or the host immune response sequesters resources normally requires for host oocyte development, and as a result associated miRNAs are also downregulated. With *Aedes aegypti*, Chikungunya virus infection does not impact the number of eggs laid but does have a detrimental impact on the viability of the eggs produced, however this is not observed with Zika virus [70, 71]. Infection of *Culex tarsalis* with West Nile Virus demonstrates a decrease in fecundity as measured by egg raft size and number of eggs laid [72].

miRNA Target Prediction

I next identified putative targets in the *An. stephensi* genome for all known and novel miRNAs identified in all samples [61]. For each of the identified miRNAs I found an average 537 potential annotated targets within the Astel2 genome. Targets for significantly regulated miRNAs were loaded into g:Profiler and any overrepresented GO terms with an FDR < 0.5 were considered significant (Appendix A) [62].

No overrepresented GO terms were identified by the genomic targets of as-mirNOV-10 or aga-miR-307. as-mirNOV-16 was significantly upregulated in response to infection at 7 dpi and the only GO terms overrepresented by the predicted genomic targets of this miRNA are associated with protein binding. as-mirNOV17 was downregulated at 14 dpi and between 7 – 14 dpi and has GO terms related to transmembrane ion channels overrepresented by its genomic targets. aga-mir-2944a/aga-mir-2944b were both downregulated at 7 dpi and between 2 – 7 dpi but upregulated between 7 – 14 dpi and both have GO terms primarily associated with intracellular signaling and various binding function. Potential genomic targets of aga-mir-2944b also appear to be involved with lipid localization and transport. aga-mir-286b and aga-miR-309 were both also downregulated at 7 dpi and between 2 and 7 dpi, and were upregulated between 7 and 14 dpi; aga-mir-286b only had acetylglucosaminyltransferase activity overrepresented by its genomic targets, and aga-miR-309 had terms related to calcium ion transport, actin filament binding, and catalytic activity overrepresented by its genomic targets.

The novel miRNA as-mirNOV-17 has 498 predicted genomic targets, and those targets overlap with 8 upregulated and 3 downregulated genes at 14 dpi and 6 upregulated and 2 downregulated genes between 7 and 14 dpi, when as-mirNOV-17 was significantly repressed in response to MAYV infection. Neither as-mirNOV-10 or as-mirNOV-16 showed a bias for upregulated or downregulated genomic targets in the contrasts they were significantly regulated in. The known miRNAs also showed a bias for upregulated targets between 2-7 dpi and 7-14 dpi where they are repressed and activated in each contrast respectively. These patterns are consistent with the miRNAs acting as effector molecules for RNAi, except for the known miRNAs between 7 - 14 dpi where their expression is enhanced but they still have a bias

for upregulated genomic targets [67]. Recent studies have demonstrated that through targeting of promotor elements miRNAs can have a positive impact on gene transcription, so this could explain the phenomenon happening between 7 – 14 dpi where miRNA targets are upregulated when the miRNAs themselves are also upregulated [73].

piRNA Identification

I identified putative piRNAs in the trimmed small RNA datasets for each sample by isolating all 24-35 nt reads, removing those that were identified positively as miRNAs, and mapping the remaining reads to the Mayaro Virus NC_003417.1 genome using the Bowtie sequence aligner, all of which was performed using the MSRG pipeline [55, 63, 74]. There was viral piRNA expression in infected samples with a bias for the positive strand over the negative strand, and the proportion of potential piRNAs mapping to the viral genome remained consistent across time points with no particular peaks or hotspots identified across the viral genome (Figure 3-6).

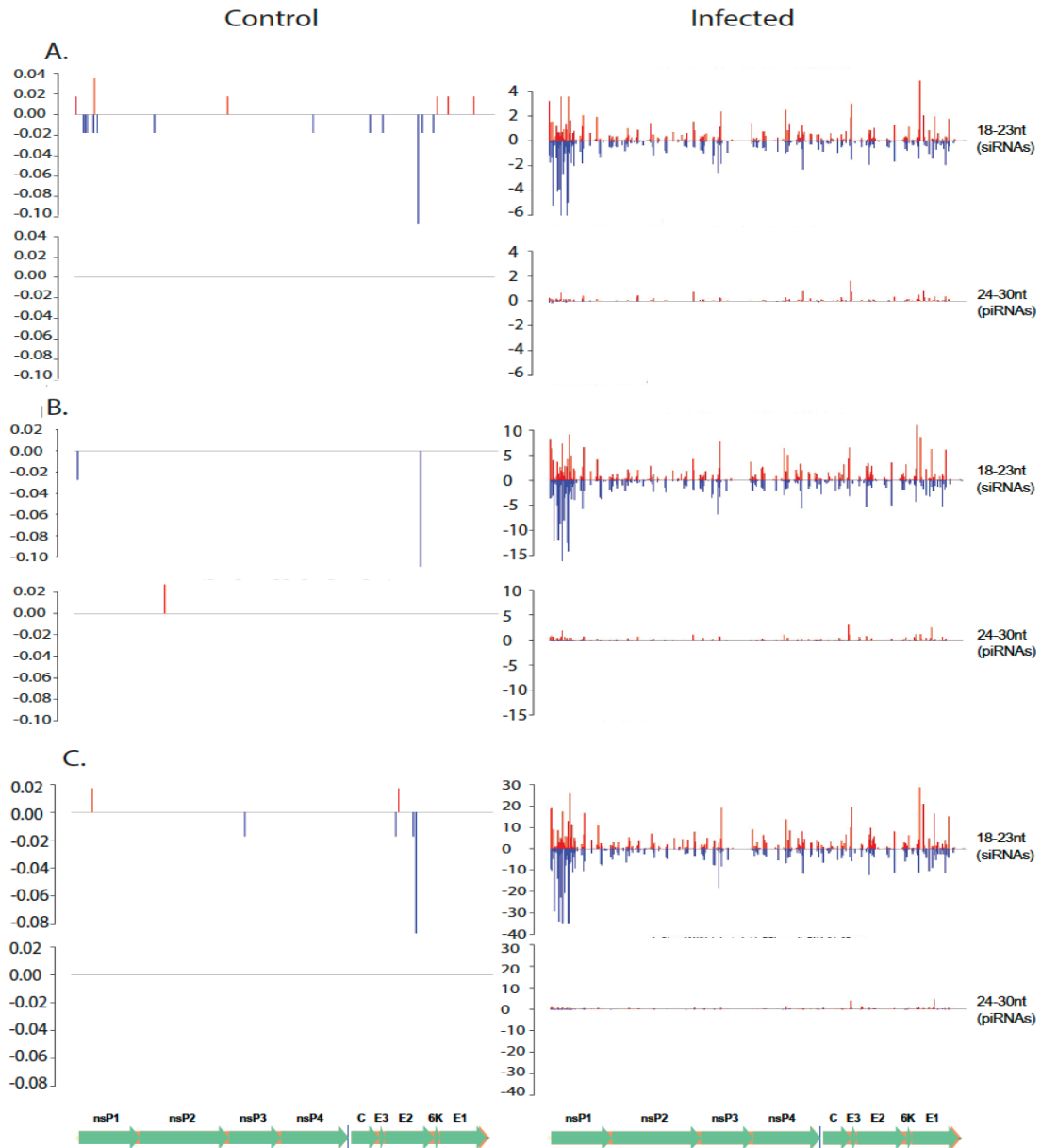


Figure 3-6: Histograms demonstrating read depth across the Mayaro virus genome for reads with a piRNA size profile (24 - 35 nt) and a siRNA size profile (18 – 23). Y-axis is read depth in reads per million, and X-axis is position in viral genome. Blue reads map to the negative strand, while red reads map to the positive strand. A. - C. 2, 7, and 14 dpi respectively, those on the left are control and on the right are infected for that time point. Underneath the y-axis is a visualization of the positions of the proteins along the viral genome.

Virus-derived piRNA-like small RNAs have been identified in insects and insect cells infected with Flaviviruses, Bunyaviruses and Alphaviruses. Knockdown of the piRNA pathway proteins leads to enhanced replication of arboviruses in mosquito cells, suggesting their potential antiviral properties in mosquitoes [75 – 80]. For example, knockdown of Piwi-4 in *Ae. aegypti* Aag2 cell line increased replication of Semliki-Forest virus, and silencing of Ago3 and Piwi-5 led to significantly reduced production of piRNAs against Sindbis virus [76, 79].

siRNA Identification

I also identified putative viral siRNAs by considering the 18-24nt reads excluded from miRNAs which mapped to the Mayaro Virus NC_003417.1 genome [55, 63, 74]. As expected, almost no siRNAs to MAYV were in the control samples, but in the infected samples there was significant MAYV siRNAs (Figure 3-6). There were abundant siRNAs on both positive and negative strands of MAYV, indicative of Dicer processing of double-stranded RNA generated during initial virus replication [81]. Interestingly, two viral genes appeared to make the most siRNAs, the nonstructural protein 1 and envelope protein 1, and these siRNA reads increased from 2 to 14 dpi.

The siRNA pathway is thought to be the main antiviral component of immunity in insects at the cellular level [53]. Functional studies in *Aedes* mosquitoes have implicated the siRNA pathway as integral to the antiviral response in the midgut stage of infection as well as at later systemic stages of infection for Dengue and Sindbis virus [53]. Studies of the siRNA response in *An. gambiae* have shown that O'nyong'nyong virus does stimulate a siRNA response at 3 dpi during the midgut stages of infection, but that its not a necessary component of antiviral immunity. They do find siRNA production is active and functional at this time point but

repression of Ago1 and Dcr2 does not significantly alter viral titers, however repression of JAK/STAT and Imd pathways did significantly increase viral titers [53]. Our siRNA results agree with the findings from *An. Gambiae* and O'nyong'nyong virus in that the siRNA response is detectable at early stages of infection and is persistent until later stages.

CONCLUSION

The transcriptomic profiles analyzed in this chapter suggest that MAYV activates the Toll pathway at early and mid-stages of infection as an innate humoral response from the host to fight infection due to the bias for serine proteases as enriched at these time points. At later stages of infection GO term overrepresentation suggests there is a repression of JNK and MAPK signaling cascades, which could potentially impact autophagic and apoptotic processes as a way to limit MAYV replication. The small RNA profiles indicate a robust response by the *An. stephensi* host to primarily generate viral siRNAs immediately after MAYV infection that increases with time of viral persistence, contrasting against a minor piRNA response primarily to the virus positive strand. *An. stephensi* miRNA expression may also respond to MAYV infection with putative roles in regulating potential transcript targets as part of the vector host response to MAYV infection. this study provides basic understanding of MAYV infection of anopheline vectors, and analyzing the molecular responses of mosquito vectors to infection with medically relevant arboviruses can lead to the discovery of important pathways necessary for the arbovirus to be transmitted by the mosquito vector thus providing the necessary context for the development of molecular strategies to block transmission in regions where the arbovirus is circulating or could be introduced.

CHAPTER 4: EVOLUTIONARY DYNAMICS OF NORTH AMERICAN ANOPHELINES

INTRODUCTION

High quality chromosome level genome assemblies exist for certain model anopheline species such as *An. gambiae* and *An. stephensi* [65, 82, 83], however, most *Anopheles* species complexes are understudied and as such comprehensive comparative genomics studies on these species are difficult to perform. Currently, while native anopheline species ranges span the entirety of the continental United States, genome assemblies of many of these species are absent from the literature. The primary anopheline species in the U.S. are *An. freeborni* in the west and *An. quadrimaculatus* in the east, and members of the *An. punctipennis* species group such as *An. crucians* are also present throughout the eastern half of the country. *An. albimanus* is one of the primary anopheline mosquitoes in Mexico, Panama and the Caribbean and is also found in Florida [3].

In this study I present novel genome assemblies for *An. crucians*, *An. freeborni*, *An. albimanus*, and *An. quadrimaculatus*. I examine the evolutionary relationship between these species, and look for examples of selection occurring in shared single copy orthologs identified within these novel genome assemblies. I identify 525 genes demonstrating evidence for positive selection on at least one branch of the phylogeny constructed using the novel genome assemblies, with gene ontology terms such as calcium ion signaling, histone binding, and protein acetylation identified as being biased in the set of selected genes.

METHODS

Species Identification

Wild *Anopheles* larvae belonging to the species *An. crucians* and *freeborni* were provided to the Rasgon Vector Genetics Laboratory by collaborators at the Brunswick County Vector Control Agency (North Carolina, *An. crucians*) and the Bennton County Mosquito Control Agency (Washington state, *An. freeborni*). Larvae were shipped from their collection localities in whirl-paks (Whirl-Pak, Madison, WI, USA) in water from their collection source, at which point they were placed into pans and reared in the Millennium Sciences Complex Insectary (The Pennsylvania State University, University Park, PA, USA) at $27^{\circ}\text{C} \pm 1^{\circ}\text{C}$, 12 hour light 12 hour dark diurnal cycle at 80% relative humidity. Ground fish flakes (TetraMin, Melle, Germany) were used to feed larvae, and upon emergence adult mosquitoes were maintained with a 10% sucrose solution. *An. quadrimaculatus* and *albimanus* were taken from colony populations maintained in the Millennium Sciences Complex Insectary initially provided by BEI Resources (Manassas, VA, USA).

Fourth instar larvae were isolated and identified using morphological features under light microscopy according to reference material for *Anopheles* larval identification in “Mosquitoes of Pennsylvania”, and species ranges of identifications were confirmed using ranges published in “Mosquitoes of New York” [84, 85]. Larvae were reared separately after identification according to date of emergence and species. Following emergence species identification was then confirmed in adults by placing them at 4°C to force them into a state of paralysis, which was then maintained for species confirmation using morphological characteristics for adults as described in “Mosquitos of Pennsylvania” by placing individuals on

a chill block under a light microscope [85]. Samples NC crucians 5 – 9 were originally identified as *An. punctipennis* by morphology, but mitochondrial assemblies place them more confidently as *An. crucians*. *An. crucians* is in the *An. punctipennis* species group and share many morphological features, making species identification from morphological based measures difficult, often requiring a molecular diagnostic [103].

Genome Assembly

DNA was extracted from individual mosquitos using the Qiagen DNEasy Blood and Tissue Kit with standard protocols. Extracted DNA underwent library preparation and genome sequencing at the New York University Langone Genome Technology Center (New York, NY, USA) using a NovaSeq 6000 with 150 bp paired end sequencing to ~30X coverage per sample. Raw sequencing reads had adapters trimmed and low-quality bases removed using Trimmomatic read trimming software with base settings [55]. Velvet was run on trimmed and paired reads for each sample with k-mer length of 85 bp, minimum coverage of 5x and expected coverage of 30x to generate the novel genome assemblies [86]. Genome completeness was measured using BUSCO with the insecta_odb10 database to scan for complete conserved single-copy orthologs present within the novel assemblies [87].

Mitochondrial genomes were assembled from interleaved raw sequencing reads for each sample using the mitoBIM pipeline [88], and were then aligned using BLAST against a database of cytochrome c oxidase subunit 1 sequences collected from NCBI to determine closest species assignment for the mitochondrial sequences (Appendix B) [89]. Augustus de-novo gene prediction software was used to identify potential proteins within the novel assemblies using a training set of 5876 full length coding sequences from the *An. gambiae* PEST

strain AgamP3 (GCA_000005575) genome assembly and functional prediction on the potential gene sets was achieved with InterProScan [90, 91].

Phylogenetics and Scanning for Signatures of Selection

To identify single copy-orthologs for use in creating a phylogeny for the novel genomes and for identifying signatures of selection, Exonerate protein2genome was run for every known protein present in the *An. gambiae* PEST strain UP000007062 proteome published on uniProt against all novel genome assemblies [92, 93]. Any Exonerate hits that had scores above 1000, were represented at least 80% of full length in the novel genomes, were not duplicated in any genome, and were present in all genomes considered were kept for further analysis. Multiple sequence alignment (MSA) for all orthologs were generated using MUSCLE and pal2nal was used to convert the MSA for each ortholog into an in-frame codon alignment [94, 95]. Concatenated in-frame codon alignments for all orthologs were used to create a consensus phylogenetic tree using iqTree, then individual gene-trees were made for each ortholog and compared with the consensus tree to determine gene concordance and site concordance factors for each node of the consensus tree [96]. HyPhy software was used to run Mixed Effects Model of Evolution (MEME) to identify signatures of episodic positive selection occurring along any branch of the consensus phylogeny for all orthologs tested, and adaptive Branch-Site Random Effects Likelihood (aBSREL) to identify if positive selection occurred along any specific proportion of branches for the orthologs in question [97, 98].

RESULTS AND DISCUSSION

Genome Assembly

I created novel genome assemblies from wild-collected *An. crucians* and *An. freeborni* and from colony populations of *An. quadrimaculatus* and *An. albimanus*. For the natural samples representing *An. crucians*, I collected whole genome sequences from 6 individuals (NC crucians 2, 3, 4, 5, 6, and 8), and two *An. freeborni* were sequenced (WA freeborni 1 and 2). Of the colony *An. quadrimaculatus* and *An. albimanus*, 2 individuals each were sent for whole genome sequencing (COL albi 1 and 2, and COL quad 3 and 4).

An. crucians had the largest genome size of those sequenced in this study with 310.8 Mb average size per assembly, *An. freeborni* had an average genome size of 270.05 Mb, *An. quadrimaculatus* had an average genome size of 229.95 Mb, and *An. albimanus* had the smallest genome size of those sampled with an average assembly size of 170.85 Mb. N50 was highest for colony sourced samples which is likely an artifact of lower genetic diversity within colony populations vs natural populations, ranging from 213.5 Kb for an *An. albimanus* sample to 25.9 Kb for an *An. crucians* sample (Table 4-1). *An. albimanus* is the only species represented in this study that already has a genome assembly published, and they report a genome size of 170.85 Mb [99], comparable with our results. BUSCO scores suggest all assemblies are relatively complete with between 98.3% and 99.5% single-copy ortholog representation from the BUSCO insecta_odb10 dataset [87].

Sample	Species	Assembly Size (MB)	N50 (KB)	Single Copy BUSCO	Duplicated BUSCO	Fragmented BUSCO	Missing BUSCO	Total BUSCO
COL albi 1	<i>An. albimanus</i>	170.3	213.5	1346	13	2	6	1367
COL albi 2	<i>An. albimanus</i>	171.4	200.1	1348	10	2	7	1367
COL quad 3	<i>An. quadrimaculatus</i>	229.2	110.3	1322	32	5	8	1367
COL quad 4	<i>An. quadrimaculatus</i>	230.7	100.1	1316	28	10	13	1367
NC crucians 2	<i>An. crucians</i>	309.4	30.0	1315	29	10	13	1367
NC crucians 3	<i>An. crucians</i>	304.2	34.0	1323	26	9	9	1367
NC crucians 4	<i>An. crucians</i>	312.8	35.9	1320	25	12	10	1367
NC crucians 5	<i>An. crucians</i>	321.6	25.9	1318	31	11	7	1367
NC crucians 6	<i>An. crucians</i>	308.1	35.6	1320	25	8	14	1367
NC crucians 8	<i>An. crucians</i>	308.8	39.3	1319	30	10	8	1367
WA freeborni 1	<i>An. freeborni</i>	268.7	53.6	1335	12	3	17	1367
WA freeborni 2	<i>An. freeborni</i>	274.1	57.0	1339	14	4	10	1367

Table 4-1 Genome sizes, N50, and BUSCO scores for all samples sequenced.

A trend for high numbers of genes was observed in the natural genomes, with as many as 40 thousand predicted for some *An. crucians* samples, and functional prediction using InterProScan determined many of these gene predictions to be classified as intrinsically disordered proteins [91]. Furthermore, many of the genes identified as intrinsically disordered protein in the novel assemblies occurred as the beginning or end of a contig, and with the shorter contig lengths in natural vs colony sourced genomes, one possible explanation for the large number of poorly categorized proteins in the natural assemblies is that they are full length proteins which have been fragmented by the lower quality of the assembly.

Phylogenetics

A total of 790 single copy orthologs from the *An. gambiae* PEST strain proteome on UniProt were identified that were shared between all novel genome assemblies [92]. A maximum likelihood phylogeny was created using a concatenated alignment for a selection of 673 orthologs that were shared with *Aedes aegypti* which acted as an outgroup. Individual trees were made for each ortholog for calculation of site and gene concordance factors using the concatenated maximum likelihood phylogeny as a reference (Figure 4-1).

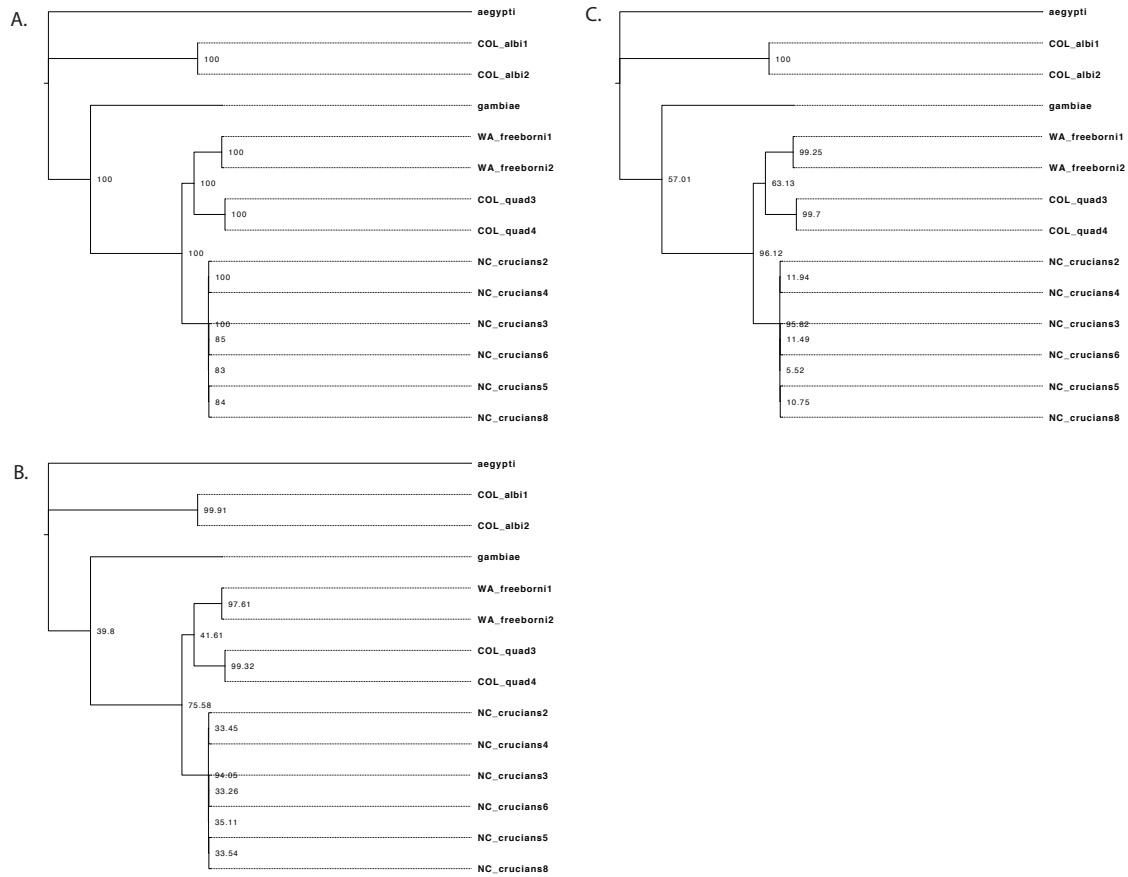


Figure 4-1: Maximum likelihood phylogenies with bootstrap values (A.), gene concordance factors (B.), and site concordance factors (C.) displayed on corresponding nodes.

An. albimanus and *An. gambiae* are placed on their own independent branches, while *An. freeborni*, *An. quadrimaculatus*, and *An. crucians* are on a branch together with 100% bootstrap support. *An. freeborni* and *An. quadrimaculatus* diverged with 100% bootstrap support into independent branches, while *An. crucians* diverges into its own branch with 100% bootstrap support, however the topology within this branch was not well supported. There are nodes within the branch that makes the *An. crucians* portion of the phylogeny, each containing 2 samples with 2 gene concordance factors of 11.94, 11.49, and 10.75.

Mitochondrial genomes were assembled from the raw sequencing data collected for each sample, and these genomes were compared to a database of cytochrome c oxidase subunit 1 sequences representing *An. gambiae*, *quadrimaculatus*, *albimanus*, *freeborni*, *crucians*, and *punctipennis* with BLAST to determine species identification (Appendix B) [89]. The potential for admixture was considered between *An. crucians* and *punctipennis* as this has been observed previously for *An. crucians* and *bradleyi* [100], but analysis of genome variants with ADMIXTURE and did not identify any evidence for hybridization between samples from the samples which could have represented the two species or confidently place them into two groups along species lines (Figure 4-2) [101]. PCA on genomic variants identified within the ortholog set demonstrates that the *An. crucians* branch separates into two loose groups, but they cluster much closer together compared with the other species represented within this study (Figure 3). The evidence suggests that instead the branch for *An. crucians* represents a single species, not 2 species or a combination of 2 species and hybridized individuals.

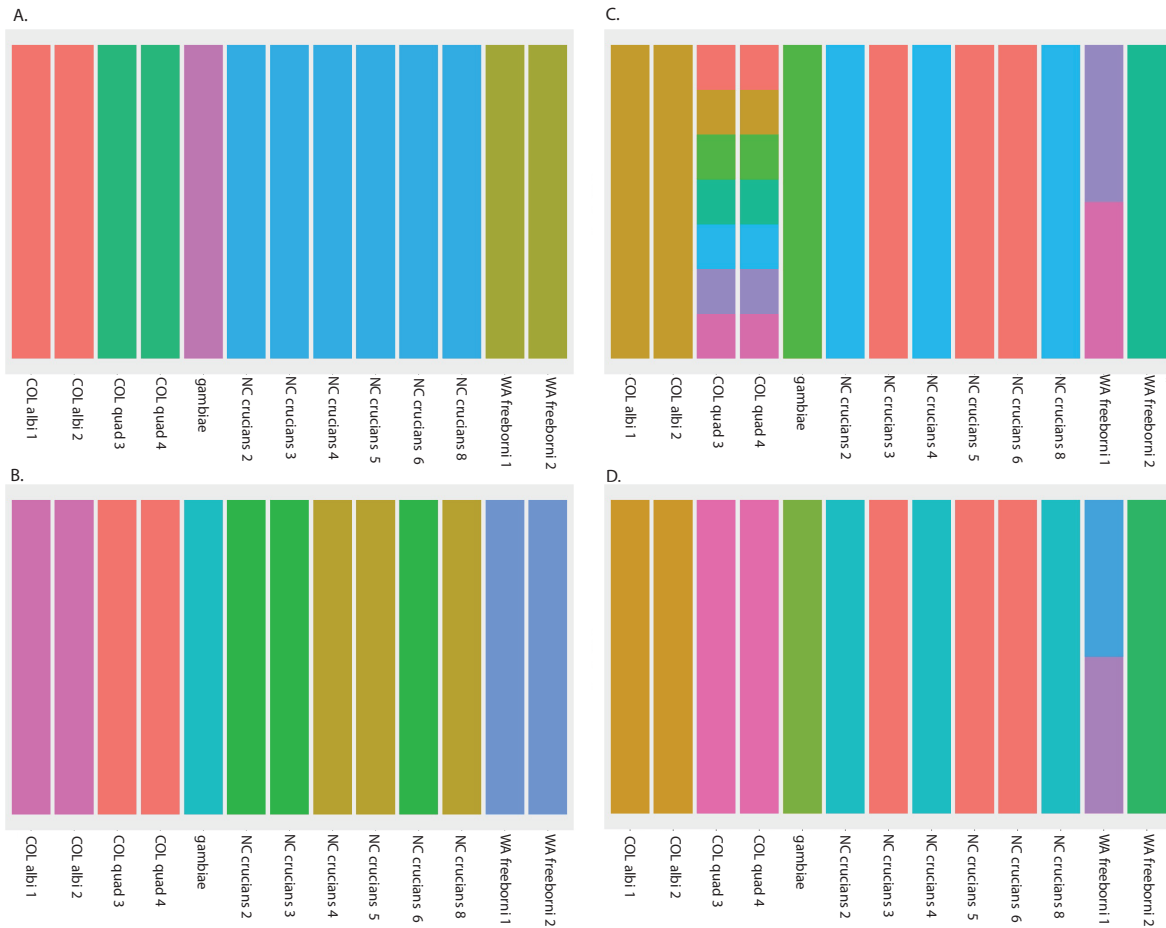


Figure 4-2: Admixture plots created using ADMIXTURE software with inferred number of populations (K) equal to 5 (A), 6 (B), 7 (C), and 8 (D). Y-axis shows the proportion of a sample’s genome from 0% to 100%, X axis has the novel genome assemblies investigated in this study and *An. gambiae* as an outgroup, and the color of the bar for each sample represents the inferred populations representing that sample’s genome. The color’s height on the Y axis represents the proportion of ancestry that population represents in the corresponding genome assembly.

Evidence for Selection

dN/dS based analyses were performed using HyPhy to analyze signatures of selection in alignments of the full set of 790 single copy orthologs identified in all newly sequenced samples with a reduced tree not containing the *Aedes aegypti* outgroup (Appendix B). Signatures of episodic positive selection at individual sites in the orthologs were identified using MEME [97], and selection of orthologs along a proportion of branches within the phylogeny using aBSREL [98]. Of the 790 orthologs considered, MEME identified at least one site under episodic positive selection in 525 orthologs in at least one branch of the phylogeny. Gene ontology analysis using g:Profiler on the genes under episodic selection identified molecular functions related to histone binding and protein acetylation, and cellular components of the cytoplasm and endomembrane complex (Appendix B) [59]. aBSREL identifies genes under selection in any branch within the phylogeny, and a more diverse set of terms is identified 289 unique orthologs regulated in at least 1 of the 19 nodes identified in the phylogeny. Gene ontology analysis using g:Profiler on terms regulated in any particular branch identified genes involved with water ion transport and glutamic acid ligation selected for nodes leading to *An. quadrimaculatus* samples (Appendix B) [59]. Nodes leading to all samples but *An. gambiae* are biased for terms related to cellular components of DNA driven RNA polymerase complex, peroxisomes, intracellular transferase complex, and protein complexes. NC *An. crucians* sample 3 had biased biological process terms related to calcium ion transport, as well as broader terms such as RNA processing and capping, gene expression, and cyclic compound metabolic processing.

Limiting the spread of vector borne disease can be improved by a deep understanding of the biology of the vector species of interest, and increasing our understanding of vector

genetics is essential for progress in eliminating vector borne pathogens. Phylogenetics of novel genome assemblies for *An. freeborni*, *crucians*, *albimanus*, and *quadrimaculatus* show that *An. crucians* groups closely with *An. quadrimaculatus* and *freeborni* separate from *An. albimanus*, the only South American species represented, which forms its own clade. Scanning for signatures of selection shows a diverse set of genes being selected for along different branches of the phylogeny, most notably calcium ion transport in *An. crucians* and *quadrimaculatus*. These novel genome sequences will be useful in developing our understanding of the diverse biological traits that drive vectorial capacity in anophelines.

CONCLUSION

Limiting the spread of vector borne disease can be improved by a deep understanding of the biology of the vector species of interest and increasing our understanding of vector genetics is essential for progress in eliminating vector borne pathogens. Phylogenetics of novel genome assemblies for *An. freeborni*, *crucians*, *albimanus*, and *quadrimaculatus* discussed in this chapter show that *An. crucians* groups closely with *An. quadrimaculatus* and *freeborni* separate from *An. albimanus*, the only South American species represented, which forms its own clade. Scanning for signatures of selection shows a diverse set of genes being selected for along different branches of the phylogeny, most notably calcium ion transport in *An. crucians* and *quadrimaculatus*.

CHAPTER 5: CONCLUSIONS AND FUTURE DIRECTIONS

The results from Chapter 2 suggest that we need to keep in mind the importance of the different arboviruses that the vector can intake during its life and all the viruses that are part of its natural virome. Some insect specific viruses (ISV) have demonstrated a capacity to modulate VC for different arboviruses: Nhumirin virus suppressed replication of West Nile virus, Japanese Encephalitis virus and Saint-Luis Encephalitis virus [40], Palm Creek virus showed the capacity to suppress West Nile virus and Eilat virus modulated CHIKV dissemination [41, 42]. Evidence for suppression of arboviruses by ISVs in the mosquito vector suggest new strategies using ISV-infected vectors to control arboviral infections at vector level [43 – 45]. In these studies the SI and SIE are at the base of the proposed strategies, however the molecular mechanisms and the biological pathways involved in these cases of viral interference are still not clear, especially for emerging or neglected arboviruses. Therefore it is essential to focus our attention on SI and CI mechanisms to understand better the natural dynamics of arboviral infection and circulation in the field and also to improve the development of innovative control strategies.

Understanding the response at the molecular level of vectors infected with pathogens can inform us on potential biological pathways to target in efforts to contain disease outbreaks. *Anopheles* mosquitos are generally ignored for their potential role in transmission of arboviruses as they are generally believed to be poor viral vectors, however there is increasing evidence to suggest they may play a role in transmission of some Alphaviruses [4, 5, 14]. Chapter 3 suggests pathways through which the *Anopheles* vector responds to infection by MAYV, one such *Alphavirus* which can be transmitted by anopheline mosquitos. Studies have investigated the ability of *An. gambiae* to transmit ONNV when repressing various immune

pathways to determine which were integral to defense from viral infection, finding that JAK/STAT and immune signaling pathways were required for the mosquito to mount a successful response to ONNV and that there was a siRNA response but that it was not integral to antiviral defense [53]. Replicating this study with MAYV and *An. stephensi* repressing transcripts and pathways involved with infection identified in this study and those identified with ONNV and *An. gambiae* would determine what aspects of the Alphaviral infection response are conserved between these two closely related viruses and vectors. Additionally, investigating both early and later stages of infection with this experimental setup would allow for investigation of the midgut and salivary gland stages of infection, while the previously described study only investigated the midgut stages. This would be important because our results from Chapter 3 suggest that different infection responses may be occurring at early and later stages of infection, however the most informative stage of the infection response from a vector control perspective is the midgut stage.

Vector control efforts are becoming more wholistic, incorporating the molecular, physiological, genetic, and ecological aspects of the vector species and pathogens in question. Increased focus on genetically modifying vector species for disease control efforts requires the production of novel genome assemblies from wild vectors to aid in developing our understanding of the diverse biological traits that drive vectorial capacity. While the genomes produced in Chapter 4 are useful because they provide the foundation for hypothesis driven studies into the biology of these vectors and a chance to investigate aspects of the evolutionary history of North American anophelines, the genome builds could be vastly improved with the application of long read sequencing to aid in characterizing high repeat

regions and create assemblies with fewer gaps allowing for a more thorough investigation of transposable element activity and copy number variation, as was achieved with Neafsey et. al. 2015 [99].

This dissertation I investigate various aspects of the potential for MAYV transmission within North America, with a particular focus on the role that anophelines might play in such a scenario where MAYV is introduced. Here I have demonstrated how MAYV interacts with ZIKV in their shared vector species, showing that MAYV appears to confer a benefit from co- and super-infection scenarios with ZIKV. I have investigated what molecular pathways are involved in the infection of *An. stephensi* with MAYV showing a potential reliance on the TOLL pathway and a signature for siRNA targeting of the viral genome. Finally, I produce novel genomes of North American anopheline vectors which could facilitate MAYV transmission, possibly eluding vector surveillance do to the ignorance of *Anopheles* as potential viral vectors. These studies provide necessary context for MAYV transmission in North America and provide resources for future vector control efforts aiming to inhibit arboviral transmission.

APPENDIX A: Supplement to Chapter 3

Table A-1: Includes number of mosquitoes in each treatment and time point and associated mortality.

Bloodfed Control							Mayaro (BeAn-8/Aug/2018): 1*10 ⁷ FFU/mL						
N in group	92						N in group	100					
N presented bloodmeal	92						N presented bloodmeal	100					
N bloodfed	65						N bloodfed	91					
% fed of those presented	70.65217391						% fed of those presented	91					
DPI	N mosquito 2 dpi	% mortality 2 dpi	N mosquito 7 dpi	% mortality 7 dpi	N mosquito 14 dpi	% mortality 14 dpi	DPI	N mosquito 2 dpi	% mortality 2 dpi	N mosquito 7 dpi	% mortality 7 dpi	N mosquito 14 dpi	% mortality 14 dpi
N subgroup	22	-	21	-	22	-	N subgroup	30	-	30	-	31	-
D0	0	0%	0	0%	0	0%	D0	0	0%	0	0%	0	0%
D1	3	12%	1	4%	0	0%	D1	4	10%	2	5%	0	0%
D2	5	20%	3	13%	1	4%	D2	7	18%	3	8%	2	5%
D3			5	22%	2	8%	D3			5	13%	6	15%
D4			5	22%	3	13%	D4			8	21%	8	20%
D5			5	22%	5	21%	D5			9	23%	9	23%
D6			6	26%	5	21%	D6			9	23%	9	23%
D7			6	26%	5	21%	D7			9	23%	9	23%
D8				0%	7	29%	D8					9	23%
D9				0%	7	29%	D9					9	23%
D10				0%	7	29%	D10					9	23%
D11				0%	7	29%	D11					10	25%
D12				0%	7	29%	D12					11	28%
D13				0%	8	33%	D13					14	35%
D14				0%	8	33%	D14					15	38%
N survive	17		15		14		N survive	23		21		16	
	11-Aug-18		16-Aug-18		23-Aug-18			11-Aug-18		16-Aug-18		23-Aug-18	

Table A-2: Nanodrop readings for all RNA extractions collected, pooling scheme, and qPCR data using primers specific for Mayaro virus strain BeAn to confirm infection status.

Sample	Nanodrop (ng/uL)	260/280	260/230	uL initial volume	ug initial RNA	uL sent for sequencing	ug sent for sequencing	Pool allocation	qPCR CT Value
BeAn-2_dpi-stephensi-1	83.59	2.07	1.41	45	3.76				13
BeAn-2_dpi-stephensi-2	130.83	2.1	1.67	45	5.89	5.73	0.75	1	13.9
BeAn-2_dpi-stephensi-3	90.33	2.17	1.63	45	4.06	8.30	0.75	1	11.8
BeAn-2_dpi-stephensi-4	96.02	2.13	1.23	45	4.32				11.1
BeAn-2_dpi-stephensi-5	71.68	2.13	1.46	45	3.23				12.3
BeAn-2_dpi-stephensi-6	113.15	2.04	1.05	45	5.09				11.4
BeAn-2_dpi-stephensi-7	119.81	2.11	1.56	45	5.39				10.6
BeAn-2_dpi-stephensi-8	71.12	2.14	1.95	45	3.20	10.55	0.75	1	11.2
BeAn-2_dpi-stephensi-9	91.53	2.05	1.38	45	4.12				13.7
BeAn-2_dpi-stephensi-10	79.16	2.13	1.9	45	3.56	9.47	0.75	1	11.1
BeAn-2_dpi-stephensi-11	65.06	2.16	1.61	45	2.93	11.53	0.75	2	8.9
BeAn-2_dpi-stephensi-12	90.2	2.16	1.85	45	4.06	8.31	0.75	2	15.1
BeAn-2_dpi-stephensi-13	122.17	2.13	1.7	45	5.50	6.14	0.75	2	10.6
BeAn-2_dpi-stephensi-14	79.47	2.15	1.72	45	3.58	9.44	0.75	2	12.3
BeAn-2_dpi-stephensi-15	77.17	2.15	1.3	45	3.47				9.6
BeAn-2_dpi-stephensi-16	96.07	2.15	1.65	45	4.32	7.81	0.75	3	9.5
BeAn-2_dpi-stephensi-17	70.73	2.15	1.88	45	3.18	10.60	0.75	3	9.1
BeAn-2_dpi-stephensi-18	112.07	2.11	1.41	45	5.04				14.3
BeAn-2_dpi-stephensi-19	138.15	2.14	1.96	45	6.22	5.43	0.75	3	11
BeAn-2_dpi-stephensi-20	64.72	2.14	1.04	45	2.91				9
BeAn-2_dpi-stephensi-21	112.32	2.09	1.66	45	5.05	6.68	0.75	3	9.2
BeAn-2_dpi-stephensi-22	79.31	2.09	1.37	45	3.57				14.8
BeAn-2_dpi-stephensi-23	86.94	2.11	1.56	45	3.91				8.1
ctrl-2_dpi-stephensi-1	143.63	2.05	1.65	45	6.46	5.22	0.75	1	28.2
ctrl-2_dpi-stephensi-2	104.92	2.12	1.79	45	4.72	7.15	0.75	1	29
ctrl-2_dpi-stephensi-3	43.06	2.15	1.43	45	1.94	17.42	0.75	2	28.6
ctrl-2_dpi-stephensi-4	100.02	2.15	1.48	45	4.50	7.50	0.75	1	29.1
ctrl-2_dpi-stephensi-5	83	2.1	1.4	45	3.74				29.3
ctrl-2_dpi-stephensi-6	43	2.16	1.96	45	1.94	17.44	0.75	1	27.9
ctrl-2_dpi-stephensi-7	89.7	2.15	1.61	45	4.04	8.36	0.75	2	28.1
ctrl-2_dpi-stephensi-8	101.38	2.15	1.33	45	4.56				28.7
ctrl-2_dpi-stephensi-9	33.33	2.19	1.77	45	1.50	22.50	0.75	2	28.3
ctrl-2_dpi-stephensi-10	59.91	2.11	1.95	45	2.70	12.52	0.75	3	28.2
ctrl-2_dpi-stephensi-11	84.65	1.94	1.5	45	3.81	8.86	0.75	2	27.8
ctrl-2_dpi-stephensi-12	63.3	2.09	1.15	45	2.85				28
ctrl-2_dpi-stephensi-13	105.31	2.02	1.3	45	4.74	7.12	0.75	3	28.7
ctrl-2_dpi-stephensi-14	40.68	2.12	1.49	45	1.83	18.44	0.75	3	29.2
ctrl-2_dpi-stephensi-15	112.09	2.13	1.75	45	5.04	6.69	0.75	3	29.7
ctrl-2_dpi-stephensi-16	97.54	2.1	1.26	45	4.39				31
ctrl-2_dpi-stephensi-17	114.77	2.12	1.02	45	5.16				29.3
BeAn-7_dpi-stephensi-1	57.04	1.97	1.11	45	2.57				10.8
BeAn-7_dpi-stephensi-2	98.76	2.03	1.6	45	4.44	7.59	0.75	1	14.7
BeAn-7_dpi-stephensi-3	65.37	2.08	1.05	45	2.94				8.9
BeAn-7_dpi-stephensi-4	47.15	2.05	1.39	45	2.12	15.91	0.75	1	9.7
BeAn-7_dpi-stephensi-5	77.02	2.01	1.53	45	3.47	9.74	0.75	1	9.7
BeAn-7_dpi-stephensi-6	57.14	2	1.44	45	2.57	13.13	0.75	1	10.1
BeAn-7_dpi-stephensi-7	45.48	1.95	0.98	45	2.05				9.1
BeAn-7_dpi-stephensi-8	68.51	2.03	1.67	45	3.08	10.95	0.75	2	15.3
BeAn-7_dpi-stephensi-9	45.91	2	1.04	45	2.07				16.1
BeAn-7_dpi-stephensi-10	63.68	2.01	1.35	45	2.87				10.3
BeAn-7_dpi-stephensi-11	73.56	2.04	1.43	45	3.31	10.20	0.75	2	15
BeAn-7_dpi-stephensi-12	43.2	2.06	1.03	45	1.94				11.1
BeAn-7_dpi-stephensi-13	49.84	2.05	1.01	45	2.24				10.6

Sample	Nanodrop (ng/uL)	260/280	260/230	uL initial volume	ug initial RNA	uL sent for sequencing	ug sent for sequencing	Pool allocation	qPCR CT Value
BeAn-7_dpi-stephensi-14	43.37	2.04	1.66	45	1.95	17.29	0.75	2	11
BeAn-7_dpi-stephensi-15	55.95	2.07	1.83	45	2.52	13.40	0.75	3	9.1
BeAn-7_dpi-stephensi-16	65.32	1.98	1.5	45	2.94	11.48	0.75	3	10.5
BeAn-7_dpi-stephensi-17	167.54	2.06	1.44	45	7.54	4.48	0.75	3	26.9
BeAn-7_dpi-stephensi-18	39.82	2.17	1.46	45	1.79	18.83	0.75	3	27.4
BeAn-7_dpi-stephensi-19	56.89	2.09	1.31	45	2.56				27.3
BeAn-7_dpi-stephensi-20	75.02	2.03	1.46	45	3.38	10.00	0.75	2	29.7
ctrl-7_dpi-stephensi-1	24.36	1.95	1.33	45	1.10	30.79	0.75	1	27.9
ctrl-7_dpi-stephensi-2	99.53	2.09	0.98	45	4.48	7.54	0.75	1	29.5
ctrl-7_dpi-stephensi-3	28.17	1.97	1.22	45	1.27	26.62	0.75	1	28.4
ctrl-7_dpi-stephensi-4	36.63	1.95	1.31	45	1.65	20.48	0.75	2	27.6
ctrl-7_dpi-stephensi-5	91.74	2.06	1.43	45	4.13	8.18	0.75	2	28
ctrl-7_dpi-stephensi-6	80.67	2.04	1.7	45	3.63	9.30	0.75	1	30.1
ctrl-7_dpi-stephensi-7	91.93	2.08	1.86	45	4.14	8.16	0.75	2	27.9
ctrl-7_dpi-stephensi-8	50.77	2.02	0.77	45	2.28				27.7
ctrl-7_dpi-stephensi-9	56.9	2.02	1.34	45	2.56	13.18	0.75	2	29.8
ctrl-7_dpi-stephensi-10	52.17	2.08	1.72	45	2.35	14.38	0.75	3	29.9
ctrl-7_dpi-stephensi-11	54.18	2.13	0.81	45	2.44	13.84	0.75	3	30
ctrl-7_dpi-stephensi-12	27.23	1.99	1.1	45	1.23	27.54	0.75	3	30.4
ctrl-7_dpi-stephensi-13	40.86	2.06	1.83	45	1.84	18.36	0.75	3	24.8
ctrl-7_dpi-stephensi-14	45.64	1.92	0.79	45	2.05				32.1
ctrl-7_dpi-stephensi-15	16.25	1.96	0.28	45	0.73				29.6
BeAn-14_dpi-stephensi-1	27.41	2.13	0.9	45	1.23	27.36	0.75	1	10.8
BeAn-14_dpi-stephensi-2	19.81	1.98	0.89	45	0.89	37.86	0.75	2	14.7
BeAn-14_dpi-stephensi-3	41.18	2.1	1.51	45	1.85	18.21	0.75	1	8.9
BeAn-14_dpi-stephensi-4	22.88	1.98	1.24	45	1.03	32.78	0.75	1	9.7
BeAn-14_dpi-stephensi-5	17.37	2.06	1.65	45	0.78				9.7
BeAn-14_dpi-stephensi-6	19.99	2.06	1.24	45	0.90	37.52	0.75	2	10.1
BeAn-14_dpi-stephensi-7	26.01	2.15	0.82	45	1.17	28.84	0.75	3	9.1
BeAn-14_dpi-stephensi-8	29.36	1.93	1.03	45	1.32	25.54	0.75	2	15.3
BeAn-14_dpi-stephensi-9	26.54	1.92	1.29	45	1.19	28.26	0.75	3	16.1
BeAn-14_dpi-stephensi-10	22.48	2	0.53	45	1.01	33.36	0.75	3	10.3
BeAn-14_dpi-stephensi-11	19.17	1.91	1.2	45	0.86	39.12	0.75	3	15
BeAn-14_dpi-stephensi-12	17.56	1.88	1.88	45	0.79	35.02	0.615	2	11.1
BeAn-14_dpi-stephensi-13	16.81	1.96	1.7	45	0.76				10.6
BeAn-14_dpi-stephensi-14	11.77	2.08	1.22	45	0.53				11
BeAn-14_dpi-stephensi-15	55.22	2.03	0.89	45	2.48	13.58	0.75	1	9.1
BeAn-14_dpi-stephensi-16	16.75	1.97	0.84	45	0.75				10.5
ctrl-14_dpi-stephensi-1	41.32	2.01	0.58	45	1.86				26.9
ctrl-14_dpi-stephensi-2	28.74	1.94	0.87	45	1.29	26.10	0.75	1	27.4
ctrl-14_dpi-stephensi-3	27.22	2.02	1.19	45	1.22	27.55	0.75	3	27.3
ctrl-14_dpi-stephensi-4	49.09	1.99	1.37	45	2.21	15.28	0.75	1	29.7
ctrl-14_dpi-stephensi-5	35.58	1.98	1.69	45	1.60	21.08	0.75	1	27.9
ctrl-14_dpi-stephensi-6	23.74	1.93	0.55	45	1.07				29.5
ctrl-14_dpi-stephensi-7	21.88	1.97	1.41	45	0.98	34.28	0.75	2	28.4
ctrl-14_dpi-stephensi-8	41.04	2.01	1.92	45	1.85	18.27	0.75	2	27.6
ctrl-14_dpi-stephensi-9	43.39	2.01	1.51	45	1.95	17.29	0.75	2	28
ctrl-14_dpi-stephensi-10	27.18	1.86	1.17	45	1.22	27.59	0.75	2	30.1
ctrl-14_dpi-stephensi-11	62.99	2	1.29	45	2.83	11.91	0.75	3	27.9
ctrl-14_dpi-stephensi-12	59.34	2.03	1.61	45	2.67	12.64	0.75	3	27.7
ctrl-14_dpi-stephensi-13	34.36	1.89	1.64	45	1.55	21.83	0.75	1	29.8
ctrl-14_dpi-stephensi-14	50.95	2	0.83	45	2.29	14.72	0.75	3	29.9

Table A-3: Differentially expressed transcripts from the *Anopheles stephensi* Astel2 genome.

Only those with logFC greater than +/- 2 are included here.

gene ID	logFC	logCPM	F	P-Value	FDR	Direction	Time Point
ASTEI08168	5.53	3.14	5.39	0.00	0.02	POSITIVE	2 dpi
ASTEI09093	4.02	2.66	15.44	0.00	0.00	POSITIVE	2 dpi
ASTEI08441	3.77	3.82	4.39	0.01	0.03	POSITIVE	2 dpi
ASTEI10616	3.52	5.66	18.32	0.00	0.00	POSITIVE	2 dpi
ASTEI02705	3.44	1.53	9.74	0.00	0.00	POSITIVE	2 dpi
ASTEI11378	3.40	6.59	19.12	0.00	0.00	POSITIVE	2 dpi
ASTEI04118	3.39	5.08	40.26	0.00	0.00	POSITIVE	2 dpi
ASTEI00531	3.24	6.23	18.28	0.00	0.00	POSITIVE	2 dpi
ASTEI07580	2.68	-0.79	8.97	0.00	0.00	POSITIVE	2 dpi
ASTEI11168	2.52	-1.09	11.46	0.00	0.00	POSITIVE	2 dpi
ASTEI00925	2.29	3.08	15.09	0.00	0.00	POSITIVE	2 dpi
ASTEI07890	2.21	0.94	11.12	0.00	0.00	POSITIVE	2 dpi
ASTEI05504	2.21	0.12	3.69	0.02	0.05	POSITIVE	2 dpi
ASTEI01707	2.06	-0.97	4.45	0.01	0.03	POSITIVE	2 dpi
ASTEI07892	2.04	2.57	19.19	0.00	0.00	POSITIVE	2 dpi
ASTEI02198	-2.02	2.91	5.57	0.00	0.02	NEGATIVE	2 dpi
ASTEI05508	-2.04	4.37	6.03	0.00	0.01	NEGATIVE	2 dpi
ASTEI09416	-2.07	-0.99	4.98	0.01	0.02	NEGATIVE	2 dpi
ASTEI04602	-2.09	4.85	29.97	0.00	0.00	NEGATIVE	2 dpi
ASTEI08621	-2.12	3.34	6.97	0.00	0.01	NEGATIVE	2 dpi
ASTEI00155	-2.13	4.78	10.51	0.00	0.00	NEGATIVE	2 dpi
ASTEI07835	-2.17	4.01	4.22	0.01	0.03	NEGATIVE	2 dpi
ASTEI10380	-2.18	0.15	4.74	0.01	0.03	NEGATIVE	2 dpi
ASTEI00113	-2.24	7.35	11.16	0.00	0.00	NEGATIVE	2 dpi
ASTEI02435	-2.26	3.63	9.79	0.00	0.00	NEGATIVE	2 dpi
ASTEI09659	-2.27	2.63	28.99	0.00	0.00	NEGATIVE	2 dpi
ASTEI07450	-2.28	1.66	8.30	0.00	0.01	NEGATIVE	2 dpi
ASTEI09280	-2.28	2.93	3.81	0.02	0.05	NEGATIVE	2 dpi
ASTEI03596	-2.29	4.13	13.46	0.00	0.00	NEGATIVE	2 dpi
ASTEI01820	-2.29	6.21	14.79	0.00	0.00	NEGATIVE	2 dpi
ASTEI03524	-2.35	1.43	8.54	0.00	0.01	NEGATIVE	2 dpi
ASTEI03523	-2.38	2.00	9.08	0.00	0.00	NEGATIVE	2 dpi
ASTEI06119	-2.51	2.96	7.17	0.00	0.01	NEGATIVE	2 dpi
ASTEI07635	-2.53	8.91	24.31	0.00	0.00	NEGATIVE	2 dpi
ASTEI03160	-2.72	1.18	13.42	0.00	0.00	NEGATIVE	2 dpi
ASTEI08671	-2.99	5.70	14.15	0.00	0.00	NEGATIVE	2 dpi
ASTEI00189	-3.02	4.21	15.40	0.00	0.00	NEGATIVE	2 dpi
ASTEI00191	-3.20	0.25	8.49	0.00	0.01	NEGATIVE	2 dpi
ASTEI02121	-3.48	3.30	13.96	0.00	0.00	NEGATIVE	2 dpi
ASTEI05732	-3.70	4.84	18.15	0.00	0.00	NEGATIVE	2 dpi
ASTEI07193	-3.98	4.83	13.32	0.00	0.00	NEGATIVE	2 dpi
ASTEI03159	-4.09	5.79	23.08	0.00	0.00	NEGATIVE	2 dpi
ASTEI02616	-4.10	0.51	11.99	0.00	0.00	NEGATIVE	2 dpi
ASTEI09342	-4.93	7.42	19.83	0.00	0.00	NEGATIVE	2 dpi
ASTEI06683	-4.97	2.30	13.07	0.00	0.00	NEGATIVE	2 dpi
ASTEI11252	-5.36	2.97	15.50	0.00	0.00	NEGATIVE	2 dpi
ASTEI02390	-7.64	5.56	31.22	0.00	0.00	NEGATIVE	2 dpi
ASTEI04601	-9.84	6.97	37.68	0.00	0.00	NEGATIVE	2 dpi

gene ID	logFC	logCPM	F	P-Value	FDR	Direction	Time Point
ASTEI09440	3.44	7.42	26.16	0.00	0.00	POSITIVE	7 dpi
ASTEI08024	2.64	5.43	23.96	0.00	0.00	POSITIVE	7 dpi
ASTEI08512	2.61	2.52	21.92	0.00	0.00	POSITIVE	7 dpi
ASTEI07892	2.51	2.57	19.19	0.00	0.00	POSITIVE	7 dpi
ASTEI10032	2.35	3.84	3.71	0.02	0.05	POSITIVE	7 dpi
ASTEI11215	2.33	10.79	237.73	0.00	0.00	POSITIVE	7 dpi
ASTEI00207	2.28	-0.70	4.24	0.01	0.03	POSITIVE	7 dpi
ASTEI04792	2.08	-0.43	5.81	0.00	0.01	POSITIVE	7 dpi
ASTEI08260	2.05	1.86	4.40	0.01	0.03	POSITIVE	7 dpi
ASTEI00542	-2.03	-0.93	6.27	0.00	0.01	NEGATIVE	7 dpi
ASTEI04716	-2.04	1.63	13.55	0.00	0.00	NEGATIVE	7 dpi
ASTEI11356	-2.23	-0.60	11.51	0.00	0.00	NEGATIVE	7 dpi
ASTEI05917	-2.24	0.66	5.44	0.00	0.02	NEGATIVE	7 dpi
ASTEI07216	-2.59	6.44	30.46	0.00	0.00	NEGATIVE	7 dpi
ASTEI04566	-2.59	7.14	25.66	0.00	0.00	NEGATIVE	7 dpi
ASTEI07348	-3.57	4.16	17.49	0.00	0.00	NEGATIVE	7 dpi
ASTEI08794	-4.30	0.71	3.68	0.02	0.05	NEGATIVE	7 dpi
ASTEI04639	5.76	-0.35	10.53	0.00	0.01	POSITIVE	14 dpi
ASTEI10115	4.68	-1.08	4.30	0.01	0.03	POSITIVE	14 dpi
ASTEI02210	4.14	1.91	18.02	0.00	0.00	POSITIVE	14 dpi
ASTEI06873	3.63	-0.28	7.08	0.01	0.02	POSITIVE	14 dpi
ASTEI02862	2.90	1.14	16.10	0.00	0.00	POSITIVE	14 dpi
ASTEI08509	2.74	4.37	25.49	0.00	0.00	POSITIVE	14 dpi
ASTEI07892	2.74	2.57	19.19	0.00	0.00	POSITIVE	14 dpi
ASTEI02105	2.44	-0.07	32.56	0.00	0.00	POSITIVE	14 dpi
ASTEI11202	2.40	-0.85	8.00	0.00	0.01	POSITIVE	14 dpi
ASTEI03994	2.29	1.18	7.02	0.00	0.01	POSITIVE	14 dpi
ASTEI03385	2.29	0.18	8.45	0.00	0.01	POSITIVE	14 dpi
ASTEI11500	2.24	0.91	5.85	0.00	0.01	POSITIVE	14 dpi
ASTEI08399	2.20	1.22	4.68	0.01	0.03	POSITIVE	14 dpi
ASTEI11170	2.19	0.01	7.07	0.00	0.01	POSITIVE	14 dpi
ASTEI04566	2.18	7.14	25.66	0.00	0.00	POSITIVE	14 dpi
ASTEI11358	2.11	1.10	5.66	0.00	0.02	POSITIVE	14 dpi
ASTEI11252	2.10	2.97	15.50	0.00	0.00	POSITIVE	14 dpi
ASTEI08794	2.09	0.71	3.68	0.02	0.05	POSITIVE	14 dpi
ASTEI05293	2.06	1.97	6.27	0.00	0.01	POSITIVE	14 dpi
ASTEI01707	2.05	-0.97	4.45	0.01	0.03	POSITIVE	14 dpi
ASTEI06683	2.05	2.30	13.07	0.00	0.00	POSITIVE	14 dpi
ASTEI05504	-2.04	0.12	3.69	0.02	0.05	NEGATIVE	14 dpi
ASTEI03026	-2.25	4.85	5.83	0.00	0.01	NEGATIVE	14 dpi
ASTEI08512	-2.81	2.52	21.92	0.00	0.00	NEGATIVE	14 dpi
ASTEI08506	-3.32	0.87	28.27	0.00	0.00	NEGATIVE	14 dpi
ASTEI08604	-3.34	3.98	10.29	0.00	0.00	NEGATIVE	14 dpi
ASTEI03019	-4.71	-0.10	3.77	0.02	0.05	NEGATIVE	14 dpi

gene ID	logFC	logCPM	F	P-Value	FDR	Direction	Time Point
ASTEI02857	2.44	-0.13	6.83	0.00	0.01	POSITIVE	2 - 7 dpi
ASTEI10322	2.06	-0.22	6.34	0.00	0.01	POSITIVE	2 - 7 dpi
ASTEI01756	2.04	4.70	4.82	0.01	0.02	POSITIVE	2 - 7 dpi
ASTEI01542	-2.07	2.19	6.88	0.00	0.01	NEGATIVE	2 - 7 dpi
ASTEI02705	-2.46	1.53	9.74	0.00	0.00	NEGATIVE	2 - 7 dpi
ASTEI01165	-2.94	0.46	12.37	0.00	0.00	NEGATIVE	2 - 7 dpi
ASTEI05917	-3.09	0.66	5.44	0.00	0.02	NEGATIVE	2 - 7 dpi
ASTEI04058	-3.15	2.79	8.97	0.00	0.00	NEGATIVE	2 - 7 dpi
ASTEI08441	-3.61	3.82	4.39	0.01	0.03	NEGATIVE	2 - 7 dpi
ASTEI08168	-4.40	3.14	5.39	0.00	0.02	NEGATIVE	2 - 7 dpi
ASTEI08794	-5.83	0.71	3.68	0.02	0.05	NEGATIVE	2 - 7 dpi
ASTEI02210	5.34	1.91	18.02	0.00	0.00	POSITIVE	7 - 14 dpi
ASTEI04639	4.78	-0.35	10.53	0.00	0.01	POSITIVE	7 - 14 dpi
ASTEI04566	4.17	7.14	25.66	0.00	0.00	POSITIVE	7 - 14 dpi
ASTEI08513	4.17	4.54	97.55	0.00	0.00	POSITIVE	7 - 14 dpi
ASTEI00542	3.65	-0.93	6.27	0.00	0.01	POSITIVE	7 - 14 dpi
ASTEI06873	3.63	-0.28	7.08	0.01	0.02	POSITIVE	7 - 14 dpi
ASTEI04601	3.61	6.97	37.68	0.00	0.00	POSITIVE	7 - 14 dpi
ASTEI02105	3.38	-0.07	32.56	0.00	0.00	POSITIVE	7 - 14 dpi
ASTEI07216	2.74	6.44	30.46	0.00	0.00	POSITIVE	7 - 14 dpi
ASTEI08509	2.74	4.37	25.49	0.00	0.00	POSITIVE	7 - 14 dpi
ASTEI05945	2.65	0.42	14.78	0.00	0.00	POSITIVE	7 - 14 dpi
ASTEI05812	2.45	5.72	7.28	0.00	0.01	POSITIVE	7 - 14 dpi
ASTEI06978	2.42	0.77	4.95	0.01	0.02	POSITIVE	7 - 14 dpi
ASTEI07348	2.41	4.16	17.49	0.00	0.00	POSITIVE	7 - 14 dpi
ASTEI05589	2.35	-0.66	5.61	0.00	0.02	POSITIVE	7 - 14 dpi
ASTEI03994	2.32	1.18	7.02	0.00	0.01	POSITIVE	7 - 14 dpi
ASTEI00171	2.24	1.35	13.33	0.00	0.00	POSITIVE	7 - 14 dpi
ASTEI06822	2.20	0.96	8.34	0.00	0.01	POSITIVE	7 - 14 dpi
ASTEI08009	2.18	1.10	10.22	0.00	0.00	POSITIVE	7 - 14 dpi
ASTEI07271	2.12	0.70	4.67	0.01	0.03	POSITIVE	7 - 14 dpi
ASTEI10115	2.11	-1.08	4.30	0.01	0.03	POSITIVE	7 - 14 dpi
ASTEI03290	2.10	-0.23	4.15	0.01	0.04	POSITIVE	7 - 14 dpi
ASTEI10400	2.02	3.97	20.72	0.00	0.00	POSITIVE	7 - 14 dpi
ASTEI03026	-2.25	4.85	5.83	0.00	0.01	NEGATIVE	7 - 14 dpi
ASTEI08512	-2.61	2.52	21.92	0.00	0.00	NEGATIVE	7 - 14 dpi
ASTEI08604	-3.21	3.98	10.29	0.00	0.00	NEGATIVE	7 - 14 dpi

Table A-4: GO term overrepresentation for differentially regulated transcripts.

GO term accession	GO category	GO term name	GO term size	Time Point	FDR - Total	FDR - Up	FDR - Down
GO:0004175	Molecular Function	endopeptidase activity	284	2 dpi	0.02	0.16	0.11
GO:0004252	Molecular Function	serine-type endopeptidase activity	190	2 dpi	0.02	0.16	0.11
GO:0008236	Molecular Function	serine-type peptidase activity	211	2 dpi	0.02	0.16	0.11
GO:0017171	Molecular Function	serine hydrolase activity	211	2 dpi	0.02	0.16	0.11
GO:0003824	Molecular Function	catalytic activity	2969	2 dpi	0.02	0.16	0.11
GO:0008233	Molecular Function	peptidase activity	448	2 dpi	0.03	0.16	0.12
GO:0098609	Biological Process	cell-cell adhesion	34	2 dpi	0.07	1.00	0.01
GO:0098742	Biological Process	cell-cell adhesion via plasma-membrane adhesion molecules	33	2 dpi	0.07	1.00	0.01
GO:0007156	Biological Process	homophilic cell adhesion via plasma membrane adhesion molecules	33	2 dpi	0.07	1.00	0.01
GO:0016020	Cellular Component	membrane	2747	2 dpi	0.00	0.70	0.00
GO:0016021	Cellular Component	integral component of membrane	2385	2 dpi	0.00	0.70	0.00
GO:0031224	Cellular Component	intrinsic component of membrane	2397	2 dpi	0.00	0.70	0.00
GO:0005549	Molecular Function	odorant binding	110	7 dpi	0.00	0.07	0.00
GO:0043177	Molecular Function	organic acid binding	9	7 dpi	0.02	0.06	0.04
GO:0031406	Molecular Function	carboxylic acid binding	10	7 dpi	0.02	0.06	0.04
GO:0031543	Molecular Function	peptidyl-proline dioxygenase activity	2	7 dpi	0.04	1.00	0.02
GO:0031545	Molecular Function	peptidyl-proline 4-dioxygenase activity	2	7 dpi	0.04	1.00	0.02
GO:0019798	Molecular Function	procollagen-proline dioxygenase activity	2	7 dpi	0.04	1.00	0.02
GO:0004656	Molecular Function	procollagen-proline 4-dioxygenase activity	2	7 dpi	0.04	1.00	0.02
GO:0004197	Molecular Function	cysteine-type endopeptidase activity	20	7 dpi	0.03	0.02	1.00
GO:1901338	Molecular Function	catecholamine binding	1	7 dpi	0.03	0.02	1.00
GO:0072545	Molecular Function	tyrosine binding	1	7 dpi	0.03	0.02	1.00
GO:0072544	Molecular Function	L-DOPA binding	1	7 dpi	0.03	0.02	1.00
GO:0035240	Molecular Function	dopamine binding	1	7 dpi	0.03	0.02	1.00
GO:0016787	Molecular Function	hydrolase activity	1345	7 dpi	0.08	0.02	0.90
GO:0004175	Molecular Function	endopeptidase activity	284	7 dpi	0.03	0.02	1.00
GO:0008233	Molecular Function	peptidase activity	448	7 dpi	0.09	0.03	1.00
GO:0003824	Molecular Function	catalytic activity	2969	7 dpi	0.09	0.03	0.90
GO:0016597	Molecular Function	amino acid binding	2	7 dpi	0.04	0.03	1.00
GO:0031418	Molecular Function	L-ascorbic acid binding	7	7 dpi	0.09	1.00	0.04
GO:0016702	Molecular Function	oxidoreductase activity, acting on single donors with incorporation of molecular oxygen, incorporation of two atoms of oxygen	9	7 dpi	0.09	1.00	0.04
GO:0016701	Molecular Function	oxidoreductase activity, acting on single donors with incorporation of molecular oxygen	10	7 dpi	0.09	1.00	0.04
GO:0015923	Molecular Function	mannosidase activity	10	7 dpi	0.09	1.00	0.04
GO:0004559	Molecular Function	alpha-mannosidase activity	10	7 dpi	0.09	1.00	0.04
GO:0016706	Molecular Function	2-oxoglutarate-dependent dioxygenase activity	9	7 dpi	0.09	1.00	0.04
GO:0048029	Molecular Function	monosaccharide binding	11	7 dpi	0.10	1.00	0.04
GO:0016409	Molecular Function	palmitoyltransferase activity	13	7 dpi	0.10	1.00	0.04
GO:0042834	Molecular Function	peptidoglycan binding	4	7 dpi	0.08	0.05	1.00
GO:0008234	Molecular Function	cysteine-type peptidase activity	60	7 dpi	0.09	0.05	1.00
GO:0004252	Molecular Function	serine-type endopeptidase activity	190	7 dpi	0.09	0.05	1.00
GO:0007608	Biological Process	sensory perception of smell	109	7 dpi	0.00	0.15	0.00
GO:0007600	Biological Process	sensory perception	178	7 dpi	0.00	0.20	0.00
GO:0003008	Biological Process	system process	191	7 dpi	0.00	0.21	0.00
GO:0007606	Biological Process	sensory perception of chemical stimulus	160	7 dpi	0.00	0.18	0.00
GO:0050877	Biological Process	nervous system process	187	7 dpi	0.00	0.21	0.00
GO:0032501	Biological Process	multicellular organismal process	379	7 dpi	0.06	0.48	0.01
GO:0019511	Biological Process	peptidyl-proline hydroxylation	2	7 dpi	0.13	1.00	0.02
GO:0018126	Biological Process	protein hydroxylation	3	7 dpi	0.13	1.00	0.03
GO:0006013	Biological Process	mannose metabolic process	6	7 dpi	0.19	1.00	0.05

GO term accession	GO category	GO term name	GO term size	Time Point	FDR - Total	FDR - Up	FDR - Down
GO:0050954	Biological Process	sensory perception of mechanical stimulus	5	14 dpi	0.01	0.00	1.00
GO:0007605	Biological Process	sensory perception of sound	5	14 dpi	0.01	0.00	1.00
GO:0051403	Biological Process	stress-activated MAPK cascade	3	14 dpi	0.03	1.00	0.00
GO:0032872	Biological Process	regulation of stress-activated MAPK cascade	3	14 dpi	0.03	1.00	0.00
GO:0070302	Biological Process	regulation of stress-activated protein kinase signaling cascade	3	14 dpi	0.03	1.00	0.00
GO:0007254	Biological Process	JNK cascade	3	14 dpi	0.03	1.00	0.00
GO:0031038	Biological Process	stress-activated protein kinase signaling cascade	3	14 dpi	0.03	1.00	0.00
GO:0046328	Biological Process	regulation of JNK cascade	3	14 dpi	0.03	1.00	0.00
GO:0080135	Biological Process	regulation of cellular response to stress	6	14 dpi	0.10	1.00	0.01
GO:0043408	Biological Process	regulation of MAPK cascade	6	14 dpi	0.10	1.00	0.01
GO:0080134	Biological Process	regulation of response to stress	8	14 dpi	0.17	1.00	0.01
GO:0006915	Biological Process	apoptotic process	40	14 dpi	0.03	0.34	0.09
GO:0012519	Biological Process	programmed cell death	41	14 dpi	0.03	0.34	0.09
GO:0008210	Biological Process	cell death	42	14 dpi	0.03	0.34	0.09
GO:0000165	Biological Process	MAPK cascade	12	14 dpi	0.24	1.00	0.03
GO:0042398	Biological Process	cellular modified amino acid biosynthetic process	15	14 dpi	0.31	1.00	0.04
GO:0050953	Biological Process	sensory perception of light stimulus	15	14 dpi	0.05	0.05	1.00
GO:0000786	Cellular Component	nucleosome	30	14 dpi	0.03	1.00	0.00
GO:0032993	Cellular Component	protein-DNA complex	32	14 dpi	0.03	1.00	0.00
GO:004815	Cellular Component	DNA packaging complex	32	14 dpi	0.03	1.00	0.00
GO:0000785	Cellular Component	chromatin	66	14 dpi	0.26	1.00	0.00
GO:0005694	Cellular Component	chromosome	123	14 dpi	0.60	1.00	0.00
GO:0042579	Cellular Component	microbody	35	14 dpi	0.60	1.00	0.01
GO:0005777	Cellular Component	peroxisome	35	14 dpi	0.60	1.00	0.01
GO:0004252	Cellular Component	serine-type endopeptidase activity	190	14 dpi	0.00	0.00	0.04
GO:0017171	Cellular Component	serine hydrolase activity	211	14 dpi	0.00	0.00	0.04
GO:0008236	Cellular Component	serine-type peptidase activity	211	14 dpi	0.00	0.00	0.04
GO:0004175	Cellular Component	endopeptidase activity	284	14 dpi	0.00	0.00	0.07
GO:0008233	Cellular Component	peptidase activity	448	14 dpi	0.00	0.00	0.09
GO:0016787	Cellular Component	hydrolase activity	1345	14 dpi	0.00	0.01	0.30
GO:0004181	Cellular Component	metallocarboxypeptidase activity	19	14 dpi	0.00	0.01	0.32
GO:0004180	Cellular Component	carboxypeptidase activity	28	14 dpi	0.01	0.04	0.38
GO:0003824	Cellular Component	catalytic activity	2969	14 dpi	0.03	0.05	0.38
GO:0140096	Cellular Component	catalytic activity, acting on a protein	942	14 dpi	0.03	0.05	0.30
GO:0008235	Cellular Component	metalloexopeptidase activity	43	14 dpi	0.04	0.05	0.41
GO:0006508	Biological Process	proteolysis	548	2 - 7 dpi	0.00	0.01	0.33
GO:0016021	Cellular Component	integral component of membrane	2385	2 - 7 dpi	0.05	0.02	0.78
GO:0031224	Cellular Component	intrinsic component of membrane	2397	2 - 7 dpi	0.05	0.02	0.78
GO:0016020	Cellular Component	membrane	2747	2 - 7 dpi	0.13	0.03	0.87
GO:1990939	Molecular Function	ATP-dependent microtubule motor activity	13	2 - 7 dpi	0.00	0.00	1.00
GO:0008569	Molecular Function	ATP-dependent microtubule motor activity, minus-end-directed	13	2 - 7 dpi	0.00	0.00	1.00
GO:0004930	Molecular Function	G protein-coupled receptor activity	123	2 - 7 dpi	0.00	0.00	0.57
GO:0004888	Molecular Function	transmembrane signaling receptor activity	278	2 - 7 dpi	0.03	0.01	0.70
GO:0060089	Molecular Function	molecular transducer activity	301	2 - 7 dpi	0.04	0.01	0.71
GO:0038023	Molecular Function	signaling receptor activity	301	2 - 7 dpi	0.04	0.01	0.71
GO:0008528	Molecular Function	G protein-coupled peptide receptor activity	20	2 - 7 dpi	0.04	0.01	1.00
GO:0001653	Molecular Function	peptide receptor activity	20	2 - 7 dpi	0.04	0.01	1.00
GO:0003777	Molecular Function	microtubule motor activity	45	2 - 7 dpi	0.05	0.01	1.00
GO:0072545	Molecular Function	tyrosine binding	1	2 - 7 dpi	0.11	1.00	0.05
GO:0072544	Molecular Function	L-DOPA binding	1	2 - 7 dpi	0.11	1.00	0.05
GO:1901338	Molecular Function	catecholamine binding	1	2 - 7 dpi	0.11	1.00	0.05
GO:0004055	Molecular Function	argininosuccinate synthase activity	1	2 - 7 dpi	0.11	1.00	0.05
GO:0004613	Molecular Function	phosphoenolpyruvate carboxykinase (GTP) activity	1	2 - 7 dpi	0.11	1.00	0.05
GO:0004611	Molecular Function	phosphoenolpyruvate carboxykinase activity	1	2 - 7 dpi	0.11	1.00	0.05
GO:0035240	Molecular Function	dopamine binding	1	2 - 7 dpi	0.11	1.00	0.05

GO term accession	GO category	GO term name	GO term size	Time Point	FDR - Total	FDR - Up	FDR - Down
GO:0050954	Biological Process	sensory perception of mechanical stimulus	5	14 dpi	0.01	0.00	1.00
GO:0007605	Biological Process	sensory perception of sound	5	7 - 14 dpi	0.00	0.00	1.00
GO:0007186	Biological Process	G protein-coupled receptor signaling pathway	155	7 - 14 dpi	0.01	0.01	0.40
GO:0050953	Biological Process	sensory perception of light stimulus	15	7 - 14 dpi	0.05	0.02	1.00
GO:0019627	Biological Process	urea metabolic process	2	7 - 14 dpi	0.24	1.00	0.05
GO:0042398	Biological Process	cellular modified amino acid biosynthetic process	15	7 - 14 dpi	0.22	1.00	0.05
GO:0038061	Biological Process	NIK/NF-kappaB signaling	1	7 - 14 dpi	0.22	1.00	0.05
GO:0006575	Biological Process	cellular modified amino acid metabolic process	50	7 - 14 dpi	0.42	1.00	0.05
GO:0031098	Biological Process	stress-activated protein kinase signaling cascade	3	7 - 14 dpi	0.24	1.00	0.05
GO:0035998	Biological Process	7,8-dihydroneopterin 3'-triphosphate biosynthetic process	3	7 - 14 dpi	0.24	1.00	0.05
GO:0032874	Biological Process	positive regulation of stress-activated MAPK cascade	1	7 - 14 dpi	0.22	1.00	0.05
GO:0032872	Biological Process	regulation of stress-activated MAPK cascade	3	7 - 14 dpi	0.24	1.00	0.05
GO:0006526	Biological Process	arginine biosynthetic process	2	7 - 14 dpi	0.24	1.00	0.05
GO:0000050	Biological Process	urea cycle	2	7 - 14 dpi	0.24	1.00	0.05
GO:0044283	Biological Process	small molecule biosynthetic process	114	7 - 14 dpi	0.29	0.92	0.05
GO:0007254	Biological Process	JNK cascade	3	7 - 14 dpi	0.24	1.00	0.05
GO:0051403	Biological Process	stress-activated MAPK cascade	3	7 - 14 dpi	0.24	1.00	0.05
GO:0046330	Biological Process	positive regulation of JNK cascade	1	7 - 14 dpi	0.22	1.00	0.05
GO:0046328	Biological Process	regulation of JNK cascade	3	7 - 14 dpi	0.24	1.00	0.05
GO:0007250	Biological Process	activation of NF-kappaB-inducing kinase activity	1	7 - 14 dpi	0.22	1.00	0.05
GO:0070302	Biological Process	regulation of stress-activated protein kinase signaling cascade	3	7 - 14 dpi	0.24	1.00	0.05
GO:0070304	Biological Process	positive regulation of stress-activated protein kinase signaling cascade	1	7 - 14 dpi	0.22	1.00	0.05
GO:0009437	Biological Process	camitine metabolic process	2	7 - 14 dpi	0.24	1.00	0.05
GO:0071941	Biological Process	nitrogen cycle metabolic process	3	7 - 14 dpi	0.24	1.00	0.05
GO:0006578	Biological Process	amino-acid betaine biosynthetic process	2	7 - 14 dpi	0.24	1.00	0.05
GO:0006577	Biological Process	amino-acid betaine metabolic process	2	7 - 14 dpi	0.24	1.00	0.05
GO:0045329	Biological Process	camitine biosynthetic process	2	7 - 14 dpi	0.24	1.00	0.05
GO:0006525	Biological Process	arginine metabolic process	3	7 - 14 dpi	0.24	1.00	0.05
GO:0051066	Biological Process	dihydrobiopterin metabolic process	3	7 - 14 dpi	0.24	1.00	0.05
GO:0030286	Cellular Component	dynein complex	29	7 - 14 dpi	0.00	0.00	1.00
GO:0005875	Cellular Component	microtubule associated complex	39	7 - 14 dpi	0.00	0.00	1.00
GO:0005929	Cellular Component	cilium	54	7 - 14 dpi	0.01	0.01	1.00
GO:0120025	Cellular Component	plasma membrane bounded cell projection	60	7 - 14 dpi	0.01	0.01	1.00
GO:0042995	Cellular Component	cell projection	70	7 - 14 dpi	0.02	0.01	1.00

Table A – 5: Read counts to known miRNAs and total library sizes.

miRNA ID	C D2_R1	C D2_R2	C D2_R3	C D7_R2	C D7_R3	C D14_R1	C D14_R2	C D14_R3
aga-bantam	119055	120421	90422	120421	113592	119649	101089	100060
aga-let-7	33039	34175	37050	34175	64398	54833	53182	67266
aga-miR-1	154772	156946	205152	156946	419484	431212	301595	512374
aga-miR-10	34165	32537	41503	32537	64179	64438	56259	75243
aga-miR-100	106935	86262	125409	86262	145538	145042	119098	169480
aga-miR-1000	204	228	219	228	250	199	157	169
aga-miR-10365	918	874	1066	874	1456	1669	1042	1481
aga-miR-10378	2141	2043	2248	2043	3314	2471	1874	2390
aga-miR-11	62235	39260	52835	39260	50386	53212	45187	58879
aga-miR-1175	0	8388	6820	8388	8592	8177	7210	7610
aga-miR-124	666	0	0	0	0	0	0	0
aga-miR-125	14373	13313	14621	13313	0	20623	0	22187
aga-miR-133	1927	2117	2053	2117	2277	1826	1507	1746
aga-miR-137	1265	1279	1516	1279	1969	1928	1347	1924
aga-miR-13b	7331	8391	7378	8391	8952	7690	6277	6177
aga-miR-14	152184	139850	158627	139850	250212	245936	214478	234768
aga-miR-184	81205	69172	78506	69172	146493	138250	124940	183786
aga-miR-1890	1499	1146	1289	1146	1527	1368	891	1243
aga-miR-1891	10762	8771	10226	8771	19089	19011	16416	18256
aga-miR-190	9326	8622	9605	8622	11951	10430	8574	10243
aga-miR-2	10150	11722	10048	11722	14451	12014	9412	11397
aga-miR-210	0	3511	4012	3511	5451	4831	4177	4140
aga-miR-219	58	58	64	58	116	89	67	114
aga-miR-263a	116791	66614	85404	66614	89631	89074	77251	110975
aga-miR-263b	8512	9895	9630	9895	28143	21987	18282	24577
aga-miR-275	62387	46861	61476	46861	31988	41427	31787	39996
aga-miR-276	409993	334423	417714	334423	461104	425690	333503	436858
aga-miR-277	98701	103556	105214	103556	126513	101192	83814	92170
aga-miR-278	4102	3476	4418	3476	4168	3969	2987	3767
aga-miR-279	30241	24218	25905	24218	47496	55154	57129	58861
aga-miR-2796	7543	6548	7832	6548	15137	12371	9920	11177
aga-miR-281	64669	76510	64464	76510	154045	173602	121568	185401
aga-miR-282	0	4	0	4	0	0	0	0
aga-miR-283	10176	11926	10376	11926	15614	21344	15538	18362
aga-miR-285	0	0	0	0	0	0	0	0
aga-miR-286	738	0	416	0	50	67	21	32
aga-miR-286b	36564	9991	18741	9991	20	0	0	12
aga-miR-2944a	66380	21029	40823	21029	192	217	115	129
aga-miR-2944b	23946	8317	13256	8317	85	75	45	70
aga-miR-2945	6726	5081	6260	5081	5226	4818	4830	6066
aga-miR-2b	0	0	0	0	0	0	0	0
aga-miR-305	39203	31218	49332	31218	23296	31614	18814	27513
aga-miR-306	68672	42489	49221	42489	77929	84137	73070	102661
aga-miR-307	95	84	96	84	98	79	58	43
aga-miR-308	0	0	52	0	83	117	101	83
aga-miR-309	19524	6579	9607	6579	37	18	0	13
aga-miR-31	4928	4250	5113	4250	6183	4948	4183	4927
aga-miR-315	1478	1602	1511	1602	2887	2740	1990	2669
aga-miR-33	30	24	24	24	35	42	26	41
aga-miR-34	145122	105612	142837	105612	311109	314005	322137	348362
aga-miR-375	523	381	547	381	616	735	648	884
aga-miR-7	61096	54783	66965	54783	100087	89224	76482	103137
aga-miR-79	90	59	46	59	85	65	46	61
aga-miR-8	171657	179904	167813	179904	251062	204789	199112	159914
aga-miR-87	2696	2963	3138	2963	2872	3078	2387	3023
aga-miR-927	5810	5330	5548	5330	10629	12385	11860	15265
aga-miR-929	176	194	183	194	324	217	154	171
aga-miR-92a	1541	1869	1425	1869	2832	3270	2619	3036
aga-miR-92b	2324	3011	2087	3011	4641	4568	3756	4047
aga-miR-932	0	0	1317	0	0	0	1201	0
aga-miR-957	3561	3734	4199	3734	6098	4986	4177	5897
aga-miR-965	855	681	767	681	1128	1912	1211	1770
aga-miR-970	5632	3837	6063	3837	5800	4805	5263	4695
aga-miR-980	54	105	64	105	0	0	81	81
aga-miR-981	3265	2940	3543	2940	4957	3937	2604	3381
aga-miR-988	541	519	485	519	672	840	705	845
aga-miR-989	28281	45229	27173	45229	104748	93392	115063	114957
aga-miR-993	295	394	330	394	531	349	389	349
aga-miR-996	3604	3254	3084	3254	4248	4249	3408	3551
aga-miR-998	5905	2593	3966	2593	2834	2581	2270	2542
aga-miR-9a	48662	39613	40667	39613	64933	74118	58358	73128
aga-miR-9c	94882	72665	58195	72665	95801	105444	105279	116812
aga-miR-iab-4	144	176	165	176	178	162	123	145
Total miRNA	2623050	2254482	2551279	2254482	3617884	3576305	3033309	3794362
Library Size	17800192	16812408	18092880	14833688	20551719	20259396	17801449	18082055
Percent of Library	0.15	0.13	0.14	0.15	0.18	0.18	0.17	0.21

miRNA ID	V_D2_R1	V_D2_R2	V_D2_R3	V_D7_R1	V_D7_R2	V_D7_R3	V_D14_R1	V_D14_R2	V_D14_R3
aga-bantam	112837	136133	102680	111462	134941	109245	95011	97239	105574
aga-let-7	35969	34222	27317	51292	71379	53993	60269	53750	57241
aga-miR-1	188647	142520	132669	274291	405670	290122	408394	297046	292864
aga-miR-10	47381	37830	30006	46901	72552	51911	58566	51374	47666
aga-miR-100	125042	98411	79701	122825	171757	134882	145167	102407	105187
aga-miR-1000	239	246	229	229	320	287	223	211	176
aga-miR-10365	1310	1204	1065	1632	2202	1700	1418	1463	1385
aga-miR-10378	2422	2552	2264	2486	3914	3068	2695	2175	2155
aga-miR-11	57237	52136	38736	48676	65689	56581	47737	34337	41780
aga-miR-1175	8719	9499	7785	9415	10531	11017	7882	6991	7666
aga-miR-124	0	866	667	0	0	0	0	0	0
aga-miR-125	15317	14938	11696	20725	0	20150	21517	0	0
aga-miR-133	2240	2247	1967	2147	2689	2328	1891	1716	1953
aga-miR-137	1710	1522	1532	1812	2854	2376	1854	1491	1526
aga-miR-13b	8740	8844	7699	7444	9873	8290	7031	7647	7794
aga-miR-14	191196	185235	144242	277848	337996	271125	268409	272506	314716
aga-miR-184	77009	63505	45769	94576	126698	111503	119924	99640	94853
aga-miR-1890	1459	1220	1040	1175	1637	1491	937	774	926
aga-miR-1891	11200	10841	7874	16488	21478	17504	12295	15256	18834
aga-miR-190	11013	10062	8327	10089	14749	11207	9835	9640	9546
aga-miR-2	10779	10879	8721	11248	15498	13138	10968	11795	11584
aga-miR-210	4098	4164	3292	4848	6503	5407	0	4360	0
aga-miR-219	51	57	56	90	108	78	101	78	73
aga-miR-263a	85635	67354	35389	52806	84221	64997	68711	67395	67240
aga-miR-263b	9508	8920	6643	15343	21894	16395	17856	21440	22214
aga-miR-275	60839	55736	39089	38971	43191	38428	39011	32315	34151
aga-miR-276	460864	428826	335593	396740	550525	462175	384639	344823	370848
aga-miR-277	110398	116427	87903	98805	131706	107412	92960	76650	83013
aga-miR-278	5168	4856	4026	4427	6356	5116	3707	3091	3418
aga-miR-279	31609	33341	26106	47931	61549	54274	44739	31603	39809
aga-miR-2796	8875	7602	5579	12222	15910	12732	13497	14911	14443
aga-miR-281	129457	84667	72944	169687	209390	202697	156124	70826	88995
aga-miR-282	0	0	9	0	0	0	0	0	0
aga-miR-283	13430	12815	13109	18768	21686	21008	18438	11464	15242
aga-miR-285	5927	0	0	0	0	0	0	0	0
aga-miR-286	461	221	27	29	39	44	83	33	29
aga-miR-286b	18682	10335	2695	21	33	52	0	9	0
aga-miR-2944a	49090	28230	8022	163	251	245	251	125	89
aga-miR-2944b	16312	8839	2604	78	106	0	103	72	59
aga-miR-2945	7774	6159	4796	4796	7208	5671	4669	3784	3752
aga-miR-2b	0	0	0	0	0	0	0	2084	0
aga-miR-305	43828	40658	39331	36683	34772	32402	29172	22077	27095
aga-miR-306	51774	52326	42073	51795	81932	75565	76366	42652	48361
aga-miR-307	90	68	66	54	88	89	74	68	76
aga-miR-308	0	0	89	102	126	0	129	0	94
aga-miR-309	11714	7350	2313	26	34	49	12	11	5
aga-miR-31	5666	5161	4188	5399	6903	5838	4736	3975	4780
aga-miR-315	1691	1464	1337	1771	2798	2149	2305	2510	2315
aga-miR-33	37	13	18	43	54	44	48	0	19
aga-miR-34	141121	126866	88333	225956	297609	238978	261961	252171	290366
aga-miR-375	627	586	496	587	909	717	717	552	476
aga-miR-7	71445	67030	49018	72308	108960	84422	71531	71432	73310
aga-miR-79	62	60	52	56	64	62	66	50	50
aga-miR-8	184363	203460	198499	190383	249885	254239	172337	147044	152942
aga-miR-87	3115	2902	2709	2984	3765	3236	2893	2422	2168
aga-miR-927	5800	5613	3863	6584	10299	7683	9662	8766	9272
aga-miR-929	255	227	156	206	270	265	199	263	229
aga-miR-92a	1503	1677	1454	2434	3187	2738	3296	2062	2839
aga-miR-92b	2162	2394	1844	3120	3743	3351	4197	3036	4044
aga-miR-932	0	1664	1220	1328	0	1624	0	1655	1679
aga-miR-957	4281	3231	2803	3977	6557	4810	4928	4357	4150
aga-miR-965	656	695	567	1138	1371	1224	1859	1219	1460
aga-miR-970	7071	5456	4161	6824	7779	6003	5043	4839	6837
aga-miR-980	70	130	56	98	159	65	128	70	106
aga-miR-981	4016	3787	3185	4697	6527	5038	4173	4109	3756
aga-miR-988	448	523	420	602	756	639	872	675	835
aga-miR-989	27779	46324	44514	59831	93585	86823	78115	56021	76117
aga-miR-993	356	397	311	345	538	415	369	440	427
aga-miR-996	3766	4076	2512	4586	4102	5244	3654	3717	3828
aga-miR-998	4639	2876	2473	2671	3568	3341	2511	1887	2191
aga-miR-9a	52625	51347	41698	56242	76862	69047	57281	38460	39244
aga-miR-9c	68109	76491	63351	63376	99393	91046	86582	53104	64958
aga-miR-iab-4	138	159	126	174	199	150	129	120	123
Total miRNA	2798105	2600968	2076112	2977219	4000416	3376672	3227809	2658496	2897260
Library Size	15851816	18043897	17499685	15721343	23780788	20775885	18174128	19188815	20525521
Percent of Library	0.18	0.14	0.12	0.19	0.17	0.16	0.18	0.14	0.14

Table A – 6: Read counts to novel miRNAs.

miRNA ID	C_D2_R1	C_D2_R2	C_D2_R3	C_D7_R2	C_D7_R3	C_D14_R1	C_D14_R2	C_D14_R3
as-mir1	65047	62538	75691	62538	95150	88284	70946	90862
as-mir10	0	0	0	0	0	115	0	105
as-mir11	0	0	0	0	0	70	0	0
as-mir12	0	18	0	18	58	66	56	90
as-mir13	0	13	28	13	44	76	41	56
as-mir14	31	12	20	12	53	38	47	68
as-mir15	0	27	0	27	28	27	13	15
as-mir16	0	0	0	0	0	35	13	40
as-mir17	0	0	0	0	17	28	38	52
as-mir18	13	0	13	0	32	31	24	31
as-mir19	28	15	10	15	26	26	51	52
as-mir2	43786	31667	31210	31667	50846	54645	46149	48181
as-mir20	0	0	0	0	0	25	0	0
as-mir21	21	9	0	9	37	16	22	47
as-mir22	0	0	0	0	0	18	0	43
as-mir23	0	0	0	0	7	5	0	0
as-mir24	0	0	0	0	0	14	0	0
as-mir25	0	0	52	0	0	12	21	22
as-mir26	0	0	0	0	0	6	0	0
as-mir27	0	18	0	18	56	65	69	63
as-mir28	0	0	0	0	0	15	0	0
as-mir29	0	0	0	0	0	10	0	0
as-mir3	20620	41312	31341	41312	16402	30683	19721	21766
as-mir30	0	0	0	0	0	0	479	589
as-mir31	0	0	0	0	157	0	304	273
as-mir32	0	0	0	0	0	0	123	0
as-mir33	0	0	0	0	0	0	108	0
as-mir34	0	0	0	0	0	0	32	0
as-mir35	0	0	0	0	0	0	21	0
as-mir36	0	0	0	0	0	0	7	0
as-mir37	0	0	0	0	0	0	14	0
as-mir38	0	0	0	0	0	0	8	0
as-mir39	0	0	0	0	0	0	9	0
as-mir4	0	15294	0	15294	34320	31846	28327	27328
as-mir40	165	0	0	0	0	0	0	506
as-mir41	0	0	0	0	0	0	0	401
as-mir42	0	0	0	0	0	0	0	165
as-mir43	0	0	28	0	29	0	0	45
as-mir44	0	0	0	0	0	0	0	35
as-mir45	0	0	0	0	0	0	0	17
as-mir46	0	0	0	0	0	0	0	10
as-mir47	0	0	0	0	0	0	0	40
as-mir48	13641	0	15274	0	0	0	0	0
as-mir49	80	0	0	0	99	0	0	0
as-mir5	5109	7500	7281	7500	14363	12806	12834	14856
as-mir50	24	0	0	0	0	0	0	0
as-mir51	13	0	0	0	0	0	0	0
as-mir52	6	0	0	0	0	0	0	0
as-mir53	9	11	10	11	0	0	0	0
as-mir54	61	0	0	0	0	0	0	0
as-mir55	0	24	0	24	0	0	0	0
as-mir56	0	11	0	11	0	0	0	0
as-mir57	0	49	0	49	42	0	0	0
as-mir58	0	0	46	0	0	0	0	0
as-mir59	0	0	9	0	0	0	0	0
as-mir6	1647	1932	1658	1932	4608	5069	4688	5215
as-mir60	0	0	6	0	0	0	0	0
as-mir61	0	0	0	0	118	0	0	0
as-mir62	0	0	0	0	6	0	0	0
as-mir63	0	0	0	0	13	0	0	0
as-mir64	0	0	0	0	10	0	0	0
as-mir65	0	0	0	0	8	0	0	0
as-mir66	0	0	0	0	0	0	0	0
as-mir67	0	0	0	0	0	0	0	0
as-mir68	0	0	0	0	0	0	0	0
as-mir69	0	0	0	0	0	0	0	0
as-mir7	0	0	0	0	0	1780	0	0
as-mir70	0	0	0	0	0	0	0	0
as-mir71	0	0	0	0	0	0	0	0
as-mir72	0	0	0	0	0	0	0	0
as-mir73	0	0	0	0	0	0	0	0
as-mir74	0	0	0	0	0	0	0	0
as-mir75	0	0	0	0	0	0	0	0
as-mir76	0	0	0	0	0	0	0	0
as-mir77	0	0	0	0	0	0	0	0
as-mir78	0	0	0	0	0	0	0	0
as-mir79	0	0	0	0	0	0	0	0
as-mir8	0	0	0	0	798	1188	0	0
as-mir9	424	405	441	405	735	615	0	0

miRNA ID	V_D2_R1	V_D2_R2	V_D2_R3	V_D7_R1	V_D7_R2	V_D7_R3	V_D14_R1	V_D14_R2	V_D14_R3
as-mir1	84045	83931	63039	81098	113000	93841	82923	72332	75750
as-mir10	85	0	72	125	134	122	114	0	0
as-mir11	0	0	0	0	110	0	0	0	0
as-mir12	0	26	18	36	45	35	58	60	65
as-mir13	21	17	9	38	27	49	42	50	29
as-mir14	25	18	13	46	51	27	0	97	65
as-mir15	7	0	0	25	0	0	17	15	0
as-mir16	0	5	0	0	23	13	33	22	31
as-mir17	12	0	0	0	24	19	0	0	0
as-mir18	10	15	0	22	19	19	28	25	30
as-mir19	16	22	15	24	51	29	47	28	34
as-mir2	38138	38632	32618	40263	59692	57694	46355	27523	36210
as-mir20	0	0	0	0	20	0	0	25	0
as-mir21	28	13	21	17	36	28	32	31	24
as-mir22	0	0	0	15	0	16	13	21	0
as-mir23	0	0	0	0	0	9	8	5	0
as-mir24	0	0	0	0	0	0	0	0	0
as-mir25	0	0	48	0	44	0	31	24	38
as-mir26	0	0	0	0	0	0	0	0	5
as-mir27	0	40	0	34	0	62	0	0	65
as-mir28	0	0	0	0	0	0	0	0	0
as-mir29	0	0	0	0	0	0	0	0	0
as-mir3	22312	34009	44472	28598	26589	24222	35272	29655	39389
as-mir30	0	0	0	0	0	0	0	0	0
as-mir31	0	0	134	0	306	0	0	151	0
as-mir32	0	0	0	0	0	0	0	0	0
as-mir33	0	0	0	0	0	0	0	0	0
as-mir34	0	0	0	0	0	0	0	0	0
as-mir35	0	0	0	0	0	0	0	0	0
as-mir36	0	8	5	0	0	0	7	10	0
as-mir37	0	0	0	0	0	0	0	0	0
as-mir38	0	0	0	6	0	0	0	0	0
as-mir39	0	0	0	0	0	0	0	0	0
as-mir4	0	16028	14912	26382	36645	29311	29524	27660	32580
as-mir40	0	108	0	0	0	175	253	342	308
as-mir41	0	0	0	0	0	0	0	0	0
as-mir42	0	0	0	62	0	92	114	0	0
as-mir43	0	0	0	0	31	35	0	31	39
as-mir44	0	0	0	0	0	0	0	0	0
as-mir45	3	0	0	0	0	0	0	0	0
as-mir46	0	0	0	0	0	0	0	0	0
as-mir47	0	0	0	0	0	0	0	0	0
as-mir48	15173	0	0	0	0	0	0	0	0
as-mir49	0	0	0	19	0	0	0	0	0
as-mir5	7963	7364	5933	11401	16638	10467	13396	13793	14856
as-mir50	0	0	0	0	0	0	0	0	0
as-mir51	0	0	0	0	13	7	0	0	11
as-mir52	0	0	4	0	0	0	0	0	0
as-mir53	12	0	0	10	0	9	0	0	0
as-mir54	38	0	0	0	0	0	0	0	0
as-mir55	0	0	0	11	25	32	0	0	0
as-mir56	0	0	12	0	0	0	0	0	0
as-mir57	0	0	0	43	0	36	0	0	0
as-mir58	0	0	0	0	0	0	0	0	0
as-mir59	0	0	0	0	0	0	0	0	0
as-mir6	1861	1836	1331	3651	4293	3834	4955	3800	5201
as-mir60	0	0	0	0	0	0	0	0	19
as-mir61	0	0	0	0	0	0	0	0	0
as-mir62	0	0	0	0	0	0	0	0	0
as-mir63	0	0	0	0	0	0	0	0	0
as-mir64	0	0	0	0	5	0	0	0	0
as-mir65	0	0	0	0	0	0	0	0	0
as-mir66	0	0	0	0	0	0	75	0	0
as-mir67	0	0	0	0	0	0	12	0	0
as-mir68	0	0	0	0	0	0	0	7	0
as-mir69	0	0	0	0	0	0	0	6	0
as-mir7	0	0	0	0	1972	0	1723	0	0
as-mir70	0	0	0	0	15	0	0	0	6
as-mir71	7	0	4	0	11	0	0	0	0
as-mir72	5	0	0	0	0	0	0	0	0
as-mir73	32	0	0	0	0	0	0	0	0
as-mir74	0	6	0	0	0	0	0	0	0
as-mir75	0	18	0	0	0	0	0	0	0
as-mir76	0	0	0	11	0	0	0	0	0
as-mir77	0	0	0	8	0	0	0	0	0
as-mir78	0	0	0	7	0	0	0	0	0
as-mir79	0	0	0	0	12	0	0	0	0
as-mir8	0	0	0	0	0	0	0	0	1104
as-mir9	461	400	348	401	688	544	520	498	448

Table A – 7: Differentially expressed miRNAs.

miRNA ID	logFC	logCPM	F	PValue	FDR	direction	intersection	intersection	time point
as-mirNOV10	8.99	4.11	6.06	0.04	0.57	POSITIVE	2	3	2 dpi
as-mir1NOV16	6.21	2.27	6.36	0.03	0.31	POSITIVE	2	2	7 dpi
aga-miR-309	-7.13	10.57	16.78	0.00	0.08	NEGATIVE	0	3	7 dpi
aga-miR-286b	-7.78	11.31	11.12	0.01	0.13	NEGATIVE	1	1	7 dpi
aga-miR-2944a	-6.23	12.45	21.04	0.00	0.07	NEGATIVE	0	3	7 dpi
aga-miR-2944b	-6.74	10.93	13.54	0.00	0.11	NEGATIVE	0	1	7 dpi
as-mirNOV17	-8.12	2.05	7.71	0.02	0.16	NEGATIVE	8	3	14 dpi
aga-miR-286b	-6.67	11.31	6.10	0.03	0.52	NEGATIVE	1	0	2 - 7 dpi
aga-miR-2944a	-5.58	12.45	11.52	0.01	0.47	NEGATIVE	5	4	2 - 7 dpi
aga-miR-2944b	-5.96	10.93	7.44	0.02	0.47	NEGATIVE	6	2	2 - 7 dpi
aga-miR-309	-6.33	10.57	9.48	0.01	0.47	NEGATIVE	9	2	2 - 7 dpi
aga-miR-286b	7.80	11.31	7.05	0.02	0.14	POSITIVE	3	1	7 - 14 dpi
aga-miR-2944a	6.46	12.45	14.94	0.00	0.09	POSITIVE	8	0	7 - 14 dpi
aga-miR-2944b	7.29	10.93	10.55	0.01	0.09	POSITIVE	13	0	7 - 14 dpi
aga-miR-307	1.09	4.84	7.68	0.02	0.13	POSITIVE	16	3	7 - 14 dpi
aga-miR-309	7.30	10.57	11.72	0.01	0.09	POSITIVE	11	0	7 - 14 dpi
as-mirNOV17	-8.76	2.05	5.42	0.05	0.21	NEGATIVE	6	2	7 - 14 dpi

Table A-8: Overrepresented GO terms represented by targets of significantly regulated miRNAs.

Term Name	GO ID	GO Category	FDR	Term Size	Intersection Size	miRNA ID
protein binding	GO:0005515	Molecular Function	0.00	1754	146	aga-mir-2944a
purine nucleotide binding	GO:0017076	Molecular Function	0.01	879	80	aga-mir-2944a
ribonucleotide binding	GO:0032553	Molecular Function	0.01	885	80	aga-mir-2944a
purine ribonucleoside triphosphate binding	GO:0035639	Molecular Function	0.01	868	79	aga-mir-2944a
purine ribonucleotide binding	GO:0032555	Molecular Function	0.01	877	80	aga-mir-2944a
ATP binding	GO:0005524	Molecular Function	0.02	714	65	aga-mir-2944a
kinase activity	GO:0016301	Molecular Function	0.02	297	33	aga-mir-2944a
adenyl nucleotide binding	GO:0030554	Molecular Function	0.02	725	66	aga-mir-2944a
adenyl ribonucleotide binding	GO:0032559	Molecular Function	0.02	723	66	aga-mir-2944a
nucleotide binding	GO:0000166	Molecular Function	0.02	1006	86	aga-mir-2944a
nucleoside phosphate binding	GO:1901265	Molecular Function	0.02	1006	86	aga-mir-2944a
calcium channel activity	GO:0005262	Molecular Function	0.02	18	6	aga-mir-2944a
calcium ion transmembrane transporter activity	GO:0015085	Molecular Function	0.02	25	7	aga-mir-2944a
guanyl-nucleotide exchange factor activity	GO:0005085	Molecular Function	0.02	50	10	aga-mir-2944a
transferase activity, transferring phosphorus-containing groups	GO:0016772	Molecular Function	0.03	384	39	aga-mir-2944a
carbohydrate derivative binding	GO:0097367	Molecular Function	0.03	1007	84	aga-mir-2944a
small molecule binding	GO:0036094	Molecular Function	0.04	1062	87	aga-mir-2944a
phosphotransferase activity, alcohol group as acceptor	GO:0016773	Molecular Function	0.04	268	29	aga-mir-2944a
anion binding	GO:0043168	Molecular Function	0.04	981	81	aga-mir-2944a
intracellular signal transduction	GO:0035556	Biological Process	0.00	199	30	aga-mir-2944a
acetylglucosaminyltransferase activity	GO:0008375	Molecular Function	0.03	10	3	aga-mir-286b
calcium ion binding	GO:0005509	Molecular Function	0.00	174	26	aga-mir-309
guanyl-nucleotide exchange factor activity	GO:0005085	Molecular Function	0.02	50	11	aga-mir-309
calcium channel activity	GO:0005262	Molecular Function	0.05	18	6	aga-mir-309
actin filament binding	GO:0051015	Molecular Function	0.05	19	6	aga-mir-309
catalytic activity, acting on a protein	GO:0140096	Molecular Function	0.05	942	78	aga-mir-309
protein binding	GO:0005515	Molecular Function	0.01	1754	150	as-mirNOV16
binding	GO:0005488	Molecular Function	0.04	4842	343	as-mirNOV16
inorganic molecular entity transmembrane transporter activity	GO:0015318	Molecular Function	0.01	275	30	as-mirNOV17
calcium ion binding	GO:0005509	Molecular Function	0.01	174	22	as-mirNOV17
channel activity	GO:0015267	Molecular Function	0.01	185	22	as-mirNOV17
passive transmembrane transporter activity	GO:0022803	Molecular Function	0.01	185	22	as-mirNOV17
ion channel activity	GO:0005216	Molecular Function	0.01	179	21	as-mirNOV17
ion transmembrane transporter activity	GO:0015075	Molecular Function	0.01	309	30	as-mirNOV17
inorganic cation transmembrane transporter activity	GO:0022890	Molecular Function	0.01	187	21	as-mirNOV17
metal ion transmembrane transporter activity	GO:0046873	Molecular Function	0.02	126	16	as-mirNOV17
cation transmembrane transporter activity	GO:0008324	Molecular Function	0.02	194	21	as-mirNOV17
potassium ion transmembrane transporter activity	GO:0015079	Molecular Function	0.03	42	8	as-mirNOV17
metal ion binding	GO:0046872	Molecular Function	0.03	1327	87	as-mirNOV17
cation binding	GO:0043169	Molecular Function	0.03	1333	87	as-mirNOV17
transporter activity	GO:0005215	Molecular Function	0.03	536	42	as-mirNOV17
transmembrane transporter activity	GO:0022857	Molecular Function	0.03	508	40	as-mirNOV17
cation channel activity	GO:0005261	Molecular Function	0.04	106	13	as-mirNOV17
potassium channel activity	GO:0005267	Molecular Function	0.04	37	7	as-mirNOV17

Term Name	GO ID	GO Category	FDR	Term Size	Intersection Size	miRNA ID
hydrolase activity, acting on acid anhydrides	GO:0016817	Molecular Function	0.01	405	55	aga-mir-2944b
hydrolase activity, acting on acid anhydrides, in phosphorus-containing anhydrides	GO:0016818	Molecular Function	0.01	405	55	aga-mir-2944b
phosphorus-oxygen lyase activity	GO:0016849	Molecular Function	0.01	24	9	aga-mir-2944b
protein binding	GO:0005515	Molecular Function	0.01	1754	183	aga-mir-2944b
pyrophosphatase activity	GO:0016462	Molecular Function	0.01	402	55	aga-mir-2944b
nucleoside-triphosphatase activity	GO:0017111	Molecular Function	0.01	389	55	aga-mir-2944b
guanylate cyclase activity	GO:0004383	Molecular Function	0.01	12	6	aga-mir-2944b
3',5'-cyclic-nucleotide phosphodiesterase activity	GO:0004114	Molecular Function	0.01	9	5	aga-mir-2944b
purine nucleotide binding	GO:0017076	Molecular Function	0.01	879	99	aga-mir-2944b
cyclase activity	GO:0009975	Molecular Function	0.01	18	7	aga-mir-2944b
calcium channel activity	GO:0005262	Molecular Function	0.01	18	7	aga-mir-2944b
cyclic-nucleotide phosphodiesterase activity	GO:0004112	Molecular Function	0.01	9	5	aga-mir-2944b
purine ribonucleotide binding	GO:0032555	Molecular Function	0.01	877	99	aga-mir-2944b
binding	GO:0005488	Molecular Function	0.02	4842	431	aga-mir-2944b
ribonucleotide binding	GO:0032553	Molecular Function	0.02	885	99	aga-mir-2944b
purine ribonucleoside triphosphate binding	GO:0035639	Molecular Function	0.02	868	97	aga-mir-2944b
guanyl-nucleotide exchange factor activity	GO:0005085	Molecular Function	0.02	50	12	aga-mir-2944b
calcium ion transmembrane transporter activity	GO:0015085	Molecular Function	0.02	25	8	aga-mir-2944b
histone-lysine N-methyltransferase activity	GO:0018024	Molecular Function	0.02	15	6	aga-mir-2944b
histone methyltransferase activity	GO:0042054	Molecular Function	0.02	15	6	aga-mir-2944b
lipid transporter activity	GO:0005319	Molecular Function	0.03	22	7	aga-mir-2944b
guanyl ribonucleotide binding	GO:0032561	Molecular Function	0.05	162	24	aga-mir-2944b
GTP binding	GO:0005525	Molecular Function	0.05	159	24	aga-mir-2944b
ribonucleoside binding	GO:0032549	Molecular Function	0.05	162	24	aga-mir-2944b
cytoskeletal protein binding	GO:0008092	Molecular Function	0.05	161	24	aga-mir-2944b
guanyl nucleotide binding	GO:0019001	Molecular Function	0.05	162	24	aga-mir-2944b
purine ribonucleoside binding	GO:0032550	Molecular Function	0.05	159	24	aga-mir-2944b
ATP-dependent microtubule motor activity	GO:1990939	Molecular Function	0.05	13	5	aga-mir-2944b
purine nucleoside binding	GO:0001883	Molecular Function	0.05	159	24	aga-mir-2944b
ATP-dependent microtubule motor activity, minus-end-directed	GO:0008569	Molecular Function	0.05	13	5	aga-mir-2944b
protein-lysine N-methyltransferase activity	GO:0016279	Molecular Function	0.05	18	6	aga-mir-2944b
lysine N-methyltransferase activity	GO:0016278	Molecular Function	0.05	18	6	aga-mir-2944b
S-adenosylmethionine-dependent methyltransferase activity	GO:0008757	Molecular Function	0.05	52	11	aga-mir-2944b
microtubule motor activity	GO:0003777	Molecular Function	0.05	45	10	aga-mir-2944b
nucleoside binding	GO:0001882	Molecular Function	0.05	163	24	aga-mir-2944b
nucleotide binding	GO:0000166	Molecular Function	0.05	1006	105	aga-mir-2944b
nucleoside phosphate binding	GO:1901265	Molecular Function	0.05	1006	105	aga-mir-2944b
GTPase activity	GO:0003924	Molecular Function	0.05	111	18	aga-mir-2944b
intracellular signal transduction	GO:0035556	Biological Process	0.00	199	42	aga-mir-2944b
Rho protein signal transduction	GO:0007266	Biological Process	0.00	30	11	aga-mir-2944b
Ras protein signal transduction	GO:0007265	Biological Process	0.00	41	13	aga-mir-2944b
cyclic nucleotide metabolic process	GO:0009187	Biological Process	0.00	24	10	aga-mir-2944b
regulation of Rho protein signal transduction	GO:0035023	Biological Process	0.01	28	10	aga-mir-2944b
cyclic nucleotide biosynthetic process	GO:0009190	Biological Process	0.01	23	9	aga-mir-2944b
regulation of Ras protein signal transduction	GO:0046578	Biological Process	0.01	35	11	aga-mir-2944b
regulation of intracellular signal transduction	GO:1902531	Biological Process	0.02	63	15	aga-mir-2944b
cGMP biosynthetic process	GO:0006182	Biological Process	0.02	12	6	aga-mir-2944b
signal transduction	GO:0007165	Biological Process	0.02	696	84	aga-mir-2944b
small GTPase mediated signal transduction	GO:0007264	Biological Process	0.02	78	17	aga-mir-2944b
response to stimulus	GO:0050896	Biological Process	0.02	993	112	aga-mir-2944b
cGMP metabolic process	GO:0046068	Biological Process	0.02	12	6	aga-mir-2944b
cyclic purine nucleotide metabolic process	GO:0052652	Biological Process	0.02	17	7	aga-mir-2944b
regulation of response to stimulus	GO:0048583	Biological Process	0.03	98	19	aga-mir-2944b
regulation of small GTPase mediated signal transduction	GO:0051056	Biological Process	0.03	42	11	aga-mir-2944b
cell communication	GO:0007154	Biological Process	0.03	715	84	aga-mir-2944b
calcium ion transmembrane transport	GO:0070588	Biological Process	0.03	24	8	aga-mir-2944b
signaling	GO:0023052	Biological Process	0.03	717	84	aga-mir-2944b
regulation of signal transduction	GO:0009966	Biological Process	0.04	95	18	aga-mir-2944b
lipid transport	GO:0006869	Biological Process	0.04	51	12	aga-mir-2944b
movement of cell or subcellular component	GO:0006928	Biological Process	0.04	96	18	aga-mir-2944b

Table with columns: miRNA ID, Target IDs. The table contains multiple sections for miRNAs including miR-1890, miR-1891, miR-190, miR-2, miR-219, and miR-263a, each listing associated target IDs.

Table with columns 'PKA ID' and 'Page ID'. The table contains multiple rows of alphanumeric IDs, organized into sections labeled 'page-mil-34', 'page-mil-37', 'page-mil-7', 'page-mil-8', 'page-mil-9', 'page-mil-17', 'page-mil-29', 'page-mil-32', and 'page-mil-35'. Each row contains a long alphanumeric string representing a unique identifier or data point.

Figure B-1: Maximum likelihood tree without *Ae. aegypti* as an outgroup. Annotating the nodes are bootstrap values (A.), gene concordance factors (B.), and site concordance factors (C.).

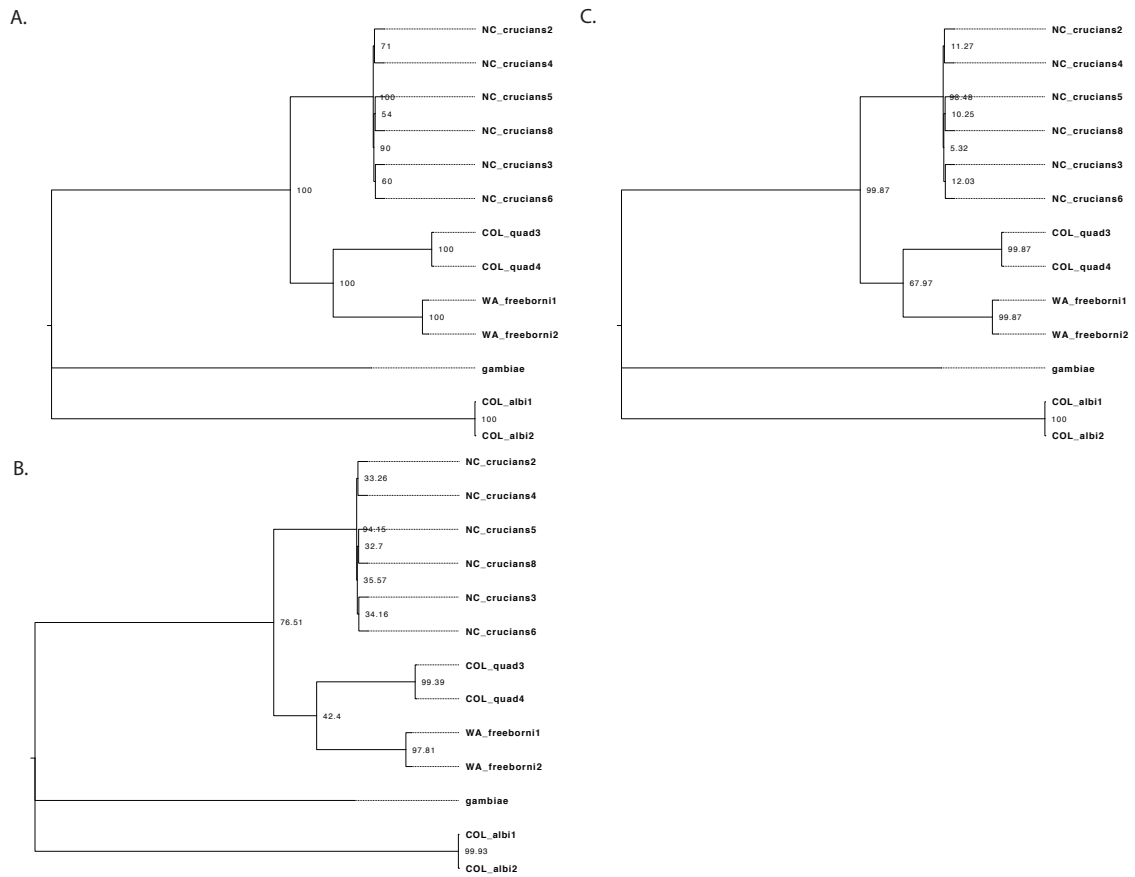


Table B-4: Table B-3: Gene ontology overrepresentation results for selected for terms identified

By MEME.

GO Category	Term Name	Term ID	P Value	Term Size	Intersection Size
Molecular Function	lysine-acetylated histone binding	GO:0070577	0.037	3	3
Molecular Function	acetylation-dependent protein binding	GO:0140033	0.037	3	3
Cellular Component	endomembrane system	GO:0012505	0.004	861	67
Cellular Component	cytoplasm	GO:0005737	0.006	3380	195

BIBLIOGRAPHY

1. Franz A, Kantor A, Passarelli A, Clem R. Tissue Barriers to Arbovirus Infection in Mosquitoes. *Viruses*. 2015;7(7):3741-3767. doi:10.3390/v7072795
2. Vector-borne diseases. <https://www.who.int/news-room/fact-sheets/detail/vector-borne-diseases>. Accessed March 2, 2021.
3. Sinka ME, Rubio-Palis Y, Manguin S, et al. The dominant Anopheles vectors of human malaria in the Americas: occurrence data, distribution maps and bionomic précis. *Parasit Vectors*. 2010;3(1):72. doi:10.1186/1756-3305-3-72
4. Rezza G, Chen R, Weaver SC. O'nyong-nyong fever: a neglected mosquito-borne viral disease. *Pathog Glob Health*. 2017;111(6):271-275. doi:10.1080/20477724.2017.1355431
5. Yadav P, Barde P, Singh D, Mishra A, Mourya D. EXPERIMENTAL TRANSMISSION OF CHIKUNGUNYA VIRUS BY ANOPHELES STEPHENSI MOSQUITOES. *Acta Virol*. 2003;47:45-47. http://www.elis.sk/download_file.php?product_id=8&session_id=jc2533t2t4ld65j13rlddb4r55. Accessed February 24, 2019.
6. Anderson CR, Wattley GH, Ahin NW, Downs WG, Reese AA. Mayaro Virus: A New Human Disease Agent. *Am J Trop Med Hyg*. 1957;6(6):1012-1016. doi:10.4269/ajtmh.1957.6.1012
7. Esposito DLA, Fonseca BAL da. Will Mayaro virus be responsible for the next outbreak of an arthropod-borne virus in Brazil? *Brazilian J Infect Dis*. 2017;21(5):540-544. doi:10.1016/j.bjid.2017.06.002
8. Hassing R-J, Leparc-Goffart I, Blank SN, et al. Imported Mayaro virus infection in the Netherlands. *J Infect*. 2010;61(4):343-345. doi:10.1016/j.jinf.2010.06.009

9. Llagonne-Barets M, Icard V, Leparç-Goffart I, et al. A case of Mayaro virus infection imported from French Guiana. *J Clin Virol*. 2016;77:66-68. doi:10.1016/j.jcv.2016.02.013
10. Receveur MC, Grandadam M, Pistone T, Malvy D. Infection with Mayaro virus in a French traveller returning from the Amazon region, Brazil, January, 2010. *Euro Surveill*. 2010;15(18). <http://www.ncbi.nlm.nih.gov/pubmed/20460093>. Accessed February 24, 2019.
11. Neumayr A, Gabriel M, Fritz J, et al. Mayaro virus infection in traveler returning from Amazon Basin, northern Peru. *Emerg Infect Dis*. 2012;18(4):695-696. doi:10.3201/eid1804.111717
12. Mota MT de O, Ribeiro MR, Vedovello D, Nogueira ML. Mayaro virus: a neglected arbovirus of the Americas. *Future Virol*. 2015;10(9):1109-1122. doi:10.2217/fvl.15.76
13. Abad-Franch F, Grimmer GH, de Paula VS, Figueiredo LTM, Braga WSM, Luz SLB. Mayaro Virus Infection in Amazonia: A Multimodel Inference Approach to Risk Factor Assessment. Weaver SC, ed. *PLoS Negl Trop Dis*. 2012;6(10):e1846. doi:10.1371/journal.pntd.0001846
14. Brustolin M, Pujhari S, Henderson CA, Rasgon JL. Anopheles mosquitoes may drive invasion and transmission of Mayaro virus across geographically diverse regions. Christofferson RC, ed. *PLoS Negl Trop Dis*. 2018;12(11):e0006895. doi:10.1371/journal.pntd.0006895
15. Wiggins K, Eastmond B, Alto BW. Transmission potential of Mayaro virus in Florida *Aedes aegypti* and *Aedes albopictus* mosquitoes. *Med Vet Entomol*. 2018;32(4):436-442. doi:10.1111/mve.12322
16. Smith GC, Francy DB. Laboratory studies of a Brazilian strain of *Aedes albopictus* as a potential vector of Mayaro and Oropouche viruses. *J Am Mosq Control Assoc*. 1991;7(1):89-93. <http://www.ncbi.nlm.nih.gov/pubmed/1646286>. Accessed February 24, 2019.

17. Tesh RB, Higgs S, Hausser NL, et al. Experimental Transmission of Mayaro Virus by *Aedes aegypti*. *Am J Trop Med Hyg.* 2011;85(4):750-757. doi:10.4269/ajtmh.2011.11-0359
18. Krauer, F. *et al.* Zika Virus Infection as a Cause of Congenital Brain Abnormalities and Guillain-Barré Syndrome: Systematic Review. *PLoS Med.* **14**, e1002203 (2017).
19. Mlakar, J. *et al.* Zika Virus Associated with Microcephaly. *N. Engl. J. Med.* **374**, 951–958 (2016).
20. Caron, M. *et al.* Recent introduction and rapid dissemination of Chikungunya virus and Dengue virus serotype 2 associated with human and mosquito coinfections in Gabon, central Africa. *Clin Infect Dis* **55**, e45-53 (2012).
21. Nuckols, J. T. *et al.* Evaluation of Simultaneous Transmission of Chikungunya Virus and Dengue Virus Type 2 in Infected *Aedes aegypti* and *Aedes albopictus* (Diptera: Culicidae). *J. Med. Entomol* **52**, 447–451 (2015).
22. Rückert, C. *et al.* Impact of simultaneous exposure to arboviruses on infection and transmission by *Aedes aegypti* mosquitoes. *Nat. Commun.* **8**, 15412 (2017).
23. Göertz, G.P. ; Vogels, C.B.F. ; Geertsema, C. ; Koenraadt, C.J.M. ; Pijlman, G. . *et al.* Mosquito co-infection with Zika and chikungunya virus allows simultaneous transmission without affecting vector competence of *Aedes aegypti*. *PLoS Negl. Trop. Dis.* **10**, e0005654 (2017).
24. Magalhaes, T. *et al.* Sequential Infection of *Aedes aegypti* Mosquitoes with Chikungunya Virus and Zika Virus Enhances Early Zika Virus Transmission. *Insects* **9**, 177 (2018).
25. Jones, R. T. *et al.* The impact of industrial activities on vector-borne disease transmission. *Acta Trop.* (2018) doi:10.1016/J.ACTATROPICA.2018.08.033.

26. Yuill, T. M. *The ecology of tropical arthropod-borne viruses. Ann. Rev. Ecol. Syst* vol. 17 www.annualreviews.org (1986).
27. Marklewitz, M. & Junglen, S. Evolutionary and ecological insights into the emergence of arthropod-borne viruses. *Acta Tropica* vol. 190 52–58 (2019).
28. Mayer, S. V., Tesh, R. B. & Vasilakis, N. The emergence of arthropod-borne viral diseases: A global prospective on dengue, chikungunya and zika fevers. *Acta Tropica* vol. 166 155–163 (2017).
29. Wilder-Smith, A. *et al.* Epidemic arboviral diseases: priorities for research and public health. *The Lancet Infectious Diseases* vol. 17 e101–e106 (2017).
30. Carrillo-Hernández, M. Y., Ruiz-Saenz, J., Villamizar, L. J., Gómez-Rangel, S. Y. & Martínez-Gutierrez, M. Co-circulation and simultaneous co-infection of dengue, chikungunya, and zika viruses in patients with febrile syndrome at the Colombian-Venezuelan border. *BMC Infect. Dis.* **18**, 61 (2018).
31. Ball, J. D. *et al.* Clinical and Epidemiologic Patterns of Chikungunya Virus Infection and Coincident Arboviral Disease in a School Cohort in Haiti, 2014–2015. *Clin. Infect. Dis.* (2018) doi:10.1093/cid/ciy582.
32. White, S. K. *et al.* Detection and phylogenetic characterization of arbovirus dual-infections among persons during a chikungunya fever outbreak, Haiti 2014. *PLoS Negl. Trop. Dis.* **12**, (2018).
33. Zambrano, H. *et al.* Case report: Zika virus and chikungunya virus coinfections: A series of three cases from a single center in Ecuador. *Am. J. Trop. Med. Hyg.* **95**, 894–896 (2016).
34. Ciota, A. T. The role of co-infection and swarm dynamics in arbovirus transmission. *Virus*

Res. **265**, 88–93 (2019).

35. Pereira Serra, O. *et al.* Mayaro virus and dengue virus 1 and 4 natural infection in culicids from Cuiabá, state of Mato Grosso, Brazil. *Mem Inst Oswaldo Cruz Rio Janeiro* **111**, 20–29 (2016).
36. Lednicky, J. *et al.* Mayaro virus in child with acute febrile illness, Haiti, 2015. *Emerg. Infect. Dis.* **22**, 2000–2002 (2016).
37. de Souza Costa, M. C. *et al.* Arbovirus investigation in patients from Mato Grosso during Zika and Chikungunya virus introduction in Brazil, 2015–2016. *Acta Trop.* **190**, 395–402 (2019).
38. Muturi, E. J. & Bara, J. Sindbis virus interferes with dengue 4 virus replication and its potential transmission by *Aedes albopictus*. *Parasit. Vectors* **8**, 65 (2015).
39. Le Coupanec, A. *et al.* Co-Infection of Mosquitoes with Chikungunya and Dengue Viruses Reveals Modulation of the Replication of Both Viruses in Midguts and Salivary Glands of *Aedes aegypti* Mosquitoes. *Int. J. Mol. Sci.* **18**, 1708 (2017).
40. Kenney, J. L., Solberg, O. D., Langevin, S. A. & Brault, A. C. Characterization of a novel insect-specific flavivirus from Brazil: Potential for inhibition of infection of arthropod cells with medically important flaviviruses. *J. Gen. Virol.* **95**, 2796–2808 (2014).
41. 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).
42. Nasar, F., Erasmus, J. H., Haddow, A. D., Tesh, R. B. & Weaver, S. C. Eilat virus induces both homologous and heterologous interference. *Virology* **484**, 51–58 (2015).
43. Vasilakis, N. & Tesh, R. B. Insect-specific viruses and their potential impact on arbovirus transmission. *Curr Opin Virol* vol. 15 69–74 (2015).

44. Blitvich, B. J. & Firth, A. E. Insect-Specific Flaviviruses: A Systematic Review of Their Discovery, Host Range, Mode of Transmission, Superinfection Exclusion Potential and Genomic Organization. *Viruses* **7**, 1927–1959 (2015).
45. Laureti, M., Paradkar, P. N., Fazakerley, J. K. & Rodriguez-Andres, J. Superinfection Exclusion in Mosquitoes and Its Potential as an Arbovirus Control Strategy. *Viruses* **12**, 1259 (2020).
46. Bonizzoni M, Dunn WA, Campbell CL, Olson KE, Marinotti O, James AA. Complex Modulation of the *Aedes aegypti* Transcriptome in Response to Dengue Virus Infection. Moreira LA, ed. *PLoS One*. 2012;7(11):e50512. doi:10.1371/journal.pone.0050512
47. Etebari K, Hegde S, Saldana MA, et al. Global transcriptome analysis of *Aedes aegypti* mosquitoes in response to Zika virus infection. *bioRxiv*. August 2017:179416. doi:10.1101/179416
48. Waldock J, Olson KE, Christophides GK. *Anopheles gambiae* Antiviral Immune Response to Systemic O'nyong-nyong Infection. Traub-Csekö YM, ed. *PLoS Negl Trop Dis*. 2012;6(3):e1565. doi:10.1371/journal.pntd.0001565
49. Sim C, Hong YS, Vanlandingham DL, et al. Modulation of *Anopheles gambiae* gene expression in response to o'nyong-nyong virus infection. *Insect Mol Biol*. 2005;14(5):475-481. doi:10.1111/j.1365-2583.2005.00578.x
50. Palatini U, Miesen P, Carballar-Lejarazu R, et al. Comparative genomics shows that viral integrations are abundant and express piRNAs in the arboviral vectors *Aedes aegypti* and *Aedes albopictus*. *BMC Genomics*. 2017. doi:10.1186/s12864-017-3903-3
51. Saldaña MA, Etebari K, Hart CE, et al. Zika virus alters the microRNA expression profile

- and elicits an RNAi response in *Aedes aegypti* mosquitoes. Armstrong PM, ed. *PLoS Negl Trop Dis*. 2017;11(7):e0005760. doi:10.1371/journal.pntd.0005760
52. Varjak M, Donald CL, Mottram TJ, et al. Characterization of the Zika virus induced small RNA response in *Aedes aegypti* cells. Olson KE, ed. *PLoS Negl Trop Dis*.
53. Carissimo G, Pondeville E, McFarlane M, et al. Antiviral immunity of *Anopheles gambiae* is highly compartmentalized, with distinct roles for RNA interference and gut microbiota. *Proc Natl Acad Sci U S A*. 2015;112(2):E176-85. doi:10.1073/pnas.1412984112
54. Hay SI, Sinka ME, Okara RM, et al. Developing Global Maps of the Dominant *Anopheles* Vectors of Human Malaria. *PLoS Med*. 2010;7(2):e1000209. doi:10.1371/journal.pmed.1000209
55. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30(15):2114-2120. doi:10.1093/bioinformatics/btu170
56. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15-21. doi:10.1093/bioinformatics/bts635
57. Liao Y, Smyth GK, Shi W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res*. 2019;47(8):e47-e47. doi:10.1093/nar/gkz114
58. Chen Y, Mccarthy D, Ritchie M, Robinson M, Smyth GK. *EdgeR User's Guide.*; 2008. <https://www.bioconductor.org/packages/release/bioc/vignettes/edgeR/inst/doc/edgeRUsersGuide.pdf>. Accessed September 24, 2018.
59. Uri Reimand J, Arak T, Adler P, et al. g:Profiler-a web server for functional interpretation of gene lists (2016 update). *Nucleic Acids Res*. 2016;(1). doi:10.1093/nar/gkw199
60. GitHub - rajewsky-lab/mirdeep2: Discovering known and novel miRNAs from small RNA

sequencing data. <https://github.com/rajewsky-lab/mirdeep2>. Accessed October 8, 2019.

61. Betel D, Wilson M, Gabow A, Marks DS, Sander C. The microRNA.org resource: targets and expression. *Nucleic Acids Res.* 2008;36(Database issue):D149-53. doi:10.1093/nar/gkm995
62. Brackney DE. Implications of autophagy on arbovirus infection of mosquitoes. *Curr Opin Insect Sci.* 2017;22:1-6. doi:10.1016/j.cois.2017.05.001
63. Ma Q, Srivastav SP, Gamez S, et al. A mosquito small RNA genomics resource reveals dynamic evolution and host responses to viruses and transposons. *Genome Res.* 2021;31(3):gr.265157.120. doi:10.1101/gr.265157.120
64. Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
65. Chakraborty M, Ramaiah A, Adolphi A, et al. Hidden genomic features of an invasive malaria vector, *Anopheles stephensi*, revealed by a chromosome-level genome assembly. *BMC Biol.* 2021;19(1):1-16. doi:10.1186/s12915-021-00963-z
66. Carissimo G, Pain A, Belda E, Vernick KD. Highly focused transcriptional response of *Anopheles coluzzii* to O'nyong nyong arbovirus during the primary midgut infection. *BMC Genomics.* 2018;19(1):526. doi:10.1186/s12864-018-4918-0
67. Samuel GH, Adelman ZN, Myles KM. Antiviral Immunity and Virus-Mediated Antagonism in Disease Vector Mosquitoes. 2017. doi:10.1016/j.tim.2017.12.005
68. Fu X, Dimopoulos G, Zhu J. Association of microRNAs with Argonaute proteins in the malaria mosquito *Anopheles gambiae* after blood ingestion. *Sci Rep.* 2017;7(1). doi:10.1038/s41598-017-07013-1
69. Biryukova I, Ye T, Levashina E. Transcriptome-wide analysis of microRNA expression in

the malaria mosquito *Anopheles gambiae*. *BMC Genomics*. 2014;15(1):557. doi:10.1186/1471-2164-15-557

70. Resck MEB, Padilha KP, Cupolillo AP, et al. Unlike Zika, Chikungunya virus interferes in the viability of *Aedes aegypti* eggs, regardless of females' age. *Sci Rep*. 2020;10(1):1-9. doi:10.1038/s41598-020-70367-6

71. da Silveira ID, Petersen MT, Sylvestre G, et al. Zika Virus Infection Produces a Reduction on *Aedes aegypti* Lifespan but No Effects on Mosquito Fecundity and Oviposition Success. *Front Microbiol*. 2018;9:3011. doi:10.3389/fmicb.2018.03011

72. Styer LM, Meola MA, Kramer LD. West Nile Virus Infection Decreases Fecundity of *Culex tarsalis* Females. *J Med Entomol*. 2007;44(6):1074-1085. doi:10.1093/jmedent/44.6.1074

73. Xiao M, Li J, Li W, et al. MicroRNAs activate gene transcription epigenetically as an enhancer trigger. *RNA Biol*. 2017;14(10):1326-1334. doi:10.1080/15476286.2015.1112487

74. Espósito DLA, da Fonseca BAL. Complete genome sequence of Mayaro virus (Togaviridae, Alphavirus) strain BeAr 20290 from Brazil. *Genome Announc*. 2015;3(6). doi:10.1128/genomeA.01372-15

75. Scott JC, Brackney DE, Campbell CL, et al. Comparison of Dengue Virus Type 2-Specific Small RNAs from RNA Interference-Competent and –Incompetent Mosquito Cells. O'Neill SL, ed. *PLoS Negl Trop Dis*. 2010;4(10):e848. doi:10.1371/journal.pntd.0000848

76. Schnettler E, Donald CL, Human S, et al. Knockdown of piRNA pathway proteins results in enhanced semliki forest virus production in mosquito cells. *J Gen Virol*. 2013;94(PART7):1680-1689. doi:10.1099/vir.0.053850-0

77. Morazzani EM, Wiley MR, Murreddu MG, Adelman ZN, Myles KM. Production of Virus-

Derived Ping-Pong-Dependent piRNA-like Small RNAs in the Mosquito Soma. Ding S-W, ed. *PLoS Pathog.* 2012;8(1):e1002470. doi:10.1371/journal.ppat.1002470

78. Vodovar N, Bronkhorst AW, van Cleef KWR, et al. Arbovirus-Derived piRNAs Exhibit a Ping-Pong Signature in Mosquito Cells. Pfeffer S, ed. *PLoS One.* 2012;7(1):e30861. doi:10.1371/journal.pone.0030861

79. Miesen P, Girardi E, Van Rij RP. Distinct sets of PIWI proteins produce arbovirus and transposon-derived piRNAs in *Aedes aegypti* mosquito cells. *Nucleic Acids Res.* 2015;43(13):6545-6556. doi:10.1093/nar/gkv590

80. Schnettler E, Ratinier M, Watson M, et al. RNA interference targets arbovirus replication in Culicoides cells. *J Virol.* 2013;87(5):2441-2454. doi:10.1128/JVI.02848-12\

81. Sabin LR, Zheng Q, Thekkat P, et al. Dicer-2 Processes Diverse Viral RNA Species. *PLoS One.* 2013;8(2):55458. doi:10.1371/journal.pone.0055458

82. Miles A, Harding NJ, Bottà G, et al. Genetic diversity of the African malaria vector *Anopheles gambiae*. *Nature.* 2017;552:96-100. doi:10.1038/nature24995

83. Holt RA, Mani Subramanian G, Halpern A, et al. The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science (80-).* 2002;298(5591):129-149. doi:10.1126/science.1076181

84. Means RG. *Mosquitoes of New York: Genera of Culicidae Other than Aedes Occurring in ...* - Robert G. Means - Google Books. New York: University of the State of New York; 1987. https://books.google.com/books/about/Mosquitoes_of_New_York_Genera_of_Culicid.html?id=yvUhAQAAAJ. Accessed March 2, 2021.

85. Darsie RF, Hutchinson ML. *THE MOSQUITOES OF PENNSYLVANIA Identification of Adult*

Females and Fourth Instar Larvae, Geographical Distribution, Biology and Public Health Importance.; 2009.

86. Zerbino DR. Using the Velvet de novo assembler for short-read sequencing technologies. *Curr Protoc Bioinforma.* 2010;CHAPTER(SUPPL. 31):Unit. doi:10.1002/0471250953.bi1105s31
87. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva E V., Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics.* 2015;31(19):3210-3212. doi:10.1093/bioinformatics/btv351
88. Hahn C, Bachmann L, Chevreur B. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads - A baiting and iterative mapping approach. *Nucleic Acids Res.* 2013;41(13):e129-e129. doi:10.1093/nar/gkt371
89. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol.* 1990;215(3):403-410. doi:10.1016/S0022-2836(05)80360-2
90. Stanke M, Diekhans M, Baertsch R, Haussler D. Using native and syntenically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics.* 2008;24(5):637-644. doi:10.1093/bioinformatics/btn013
91. Jones P, Binns D, Chang HY, et al. InterProScan 5: Genome-scale protein function classification. *Bioinformatics.* 2014;30(9):1236-1240. doi:10.1093/bioinformatics/btu031
92. Bateman A, Martin MJ, Orchard S, et al. UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Res.* 2021;49(D1):D480-D489. doi:10.1093/nar/gkaa1100
93. Slater GSC, Birney E. Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics.* 2005;6(1):31. doi:10.1186/1471-2105-6-31
94. Edgar RC. MUSCLE: Multiple sequence alignment with high accuracy and high

throughput. *Nucleic Acids Res.* 2004;32(5):1792-1797. doi:10.1093/nar/gkh340

95. Suyama M, Torrents D, Bork P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. doi:10.1093/nar/gkl315

96. Minh BQ, Schmidt HA, Chernomor O, et al. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Teeling E, ed. *Mol Biol Evol.* 2020;37(5):1530-1534. doi:10.1093/molbev/msaa015

97. Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, Kosakovsky Pond SL. Detecting individual sites subject to episodic diversifying selection. *PLoS Genet.* 2012;8(7):1002764. doi:10.1371/journal.pgen.1002764

98. Smith MD, Wertheim JO, Weaver S, Murrell B, Scheffler K, Kosakovsky Pond SL. Less Is More: An Adaptive Branch-Site Random Effects Model for Efficient Detection of Episodic Diversifying Selection. *Mol Biol Evol.* 2015;32(5):1342-1353. doi:10.1093/molbev/msv022

99. Neafsey DE, Waterhouse RM, Abai MR, et al. Highly evolvable malaria vectors: The genomes of 16 Anopheles mosquitoes. *Science (80-)*. 2015;347(6217). doi:10.1126/science.1258522

100. Kreutzer RD, Kitzmiller JB. Hybridization Between Anopheles crucians and Anopheles bradleyi. *Evolution (N Y)*. 1971;25(1):195. doi:10.2307/2406511

101. Alexander DH, Lange K. Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics.* 2011;12(1):246. doi:10.1186/1471-2105-12-246

102. Wiggins K, Eastmond B, Alto BW. Transmission potential of Mayaro virus in Florida *Aedes aegypti* and *Aedes albopictus* mosquitoes. *Med Vet Entomol.* 2018;32(4):436-442. doi:10.1111/mve.12322

103. Hodge JM, Yurchenko AA, Karagodin DA, et al. The new Internal Transcribed Spacer 2 diagnostic tool clarifies the taxonomic position and geographic distribution of the North American malaria vector *Anopheles punctipennis*. *Malar J.* 2021;20(1). doi:10.1186/s12936-021-03676-4

VITA

Cory Henderson

Education

- 2021 Ph. D. in Molecular Cellular and Integrative Biosciences; The Pennsylvania State University, State College, PA
- 2017 B.S. in Biology, B.A. in Anthropology; Minor in Chemistry; University of North Carolina at Greensboro, Greensboro, NC

Selected Research Experience

- 2017 – 21 Graduate Research Assistant; The Pennsylvania State University, State College, PA
- 2018 Field Assistant; International Centers of Excellence for Malaria Research, Thailand
- 2015 – 17 Undergraduate Research Assistant; University of North Carolina at Greensboro, Greensboro, NC
- 2016 Research Assistant; The Olduvai Palaeoecological and Paleoanthropological Project

Selected Academic Honors (Pennsylvania State University)

- 2018 Graduate Research Fellowship; National Science Foundation
- 2017 Willaman Distinguished Graduate Fellow; Pennsylvania State University
- 2017 Braddock Award; Pennsylvania State University

Selected Service (Pennsylvania State University)

- 2019 Assistant Instructor for Bioinformatics Workshop with the African Primatological Society Annual Conference
- 2018 – 19 Camp Director and Instructor of Finding Your Roots Summer Camp; Science-U
- 2017 – 18 Treasurer for Center for Infectious Disease Dynamics Graduate Student Association

Publications

CP Egeland, CM Fadem, RM Byerly, **C Henderson**, C Fitzgerald, AZP Mabulla, E Baquedano, A Gidna (2019). Geochemical and physical characterization of lithic raw materials in the Olduvai Basin, Tanzania. *Quaternary International*. DOI: 10.1016/J.QUAINT.2019.09.036.

M Brustolin, S Pujhari, **C Henderson**, JL Rasgon (2018). *Anopheles* mosquitoes may drive invasion and transmission of Mayaro virus across geographically diverse regions. *PLOS Neglected Tropical Diseases*. DOI: 10.1371/journal.pntd.0006895.

RL Anemone, CW Emerson, TW Jones, J Liu, **C Henderson** (2018). Taking virtual anthropology to the field: Developing three dimensional digital outcrop models (3D-DOMs) of fossil localities. IN: *New Geospatial Approaches Within Anthropology*, RL Anemone and GC Conroy (eds.), University of New Mexico Press, Albuquerque, NM.