FULL TITLE: Mayaro Virus Infection Elicits an Innate Immune Response and Represses Autophagy in *Anopheles stephensi*

SHORT TITLE: Mayaro virus molecular infection dynamics using an *Anopheles* model

Cory Henderson[1], Marco Brustolin[1], Shivanand Hegde[2], Grant L. Hughes[2], Christina Bergey[3], and Jason L. Rasgon[1*]

[1]Departments of Entomology and Disease Epidemiology, The Pennsylvania State University, University Park, PA, United States of America.

[2] Departments of Vector Biology and Tropical Disease Biology, Centre for Neglected Tropical Disease, Liverpool School of Tropical Medicine, Liverpool, United Kingdom

[3]Department of Genetics, Rutgers University, New Brunswick, NJ, United States of America

*Corresponding author: Jason L. Rasgon (jlr54@psu.edu)

Keywords: Toll Pathway, Autophagy, Transcriptome, Small RNA, miRNA, piRNA, *Alphavirus*, Mayaro Virus, *Anopheles stephensi*

ABSTRACT:

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* in addition to the typical arbovirus transmitting mosquitoes in the genus *Aedes*. Few studies have investigated the infection response of *Anopheles* mosquitoes to arboviruses. In this study we 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 and 79 novel miRNAs were identified. 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 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.

AUTHOR SUMMARY:

Mayaro virus (MAYV) is a mosquito-borne Alphavirus responsible for outbreaks in South America and the Caribbean. In this study we infected *Anopheles stephensi* with MAYV and sequenced mRNA and small RNA to understand how MAYV infection impacts gene transcription and the expression of small RNAs in the mosquito vector. Genes involved with innate immunity and autophagy are regulated in response to MAYV infection of *An. stephensi,* we also discover novel miRNAs and describe their expression patterns following bloodmeal ingestion. These

results suggest that MAYV does induce a molecular response to infection in its mosquito vector species and that MAYV may have mechanisms to evade the vector immune response.

INTRODUCTION

Mayaro virus (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 [1]. Symptoms of MAYV infection are similar to other arboviral infections such as dengue or Chikungunya viruses, and include rash, fever, retro-orbital pain, headache, diarrhea, and arthralgia [2]. While no epidemics or outbreaks with Mayaro virus being the causative agent have been recorded outside of South America, there have been imported cases reported in the Netherlands, Germany, France, and Switzerland [3-6], 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 Mayaro virus naturally are thought to be the canopy-dwellers in the genus *Haemogogus*, maintaining the sylvatic cycle between non-human primates as primary hosts and birds as secondary hosts [7]. 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 [8]. Vector competence studies have identified anthropophilic and urban adapted species such as *Aedes aegypti* and *Ae. albopictus,* as well as the malaria parasite vectors *Anopheles gambiae*, *An. stephensi*, *An. freeborni*, and *An. quadrimaculatus*, as being competent vectors for Mayaro virus under laboratory conditions [9-12]. Transmission of an arbovirus with an anopheline mosquito as a primary vector is rare, only having been observed occurring

regularly for O'nyong'nyong virus by *An. gambiae* and *An. funestus* in Uganda [13], with some limited evidence for Chikungunya and Semliki Forest virus [14].

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 [15]. 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 [16-19]. In addition, studies focusing on small RNA identification and regulation have identified RNAi activity, such as miRNA and piRNA expression, in response to infection of mosquitoes by arboviruses [20-23].

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 [9, 24-25]. Because anopheline, and not aedine, 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. We 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).

MATERIALS AND METHODS

Anopheles stephensi Rearing

_______Protocols pertaining to mosquito rearing and presentation of infectious bloodmeal has been described elsewhere [9]. Briefly, *An. stephensi* (Liston strain) 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 Mayaro virus at $1*10^7$ FFU/mL or a control bloodmeal with no virus via a glass feeder jacketed with 37°C distilled water for 1 h.

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: 5⃞-

TGGACCTTTGGCTCTTCTTATC-3⊠, Reverse: 5⊠-GACGCTCACTGCGACTAAA-3⊠) [10], a CT value of 20 or less was used to confirm infection (Supplementary Table 1). 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 [9].

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 all. 2017 for more detail on library preparation and sequencing [17]. 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.

<u>Small RNA Library Preparation and Sequencing</u>

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 [21]. 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.

<u>Transcriptomic RNA Sequencing Data Analysis</u>

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 were trimmed and low-quality bases removed using Trimmomatic read trimming software with base settings [26]. 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 [27]. 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 [28], following which a principal components analysis was performed and differential expression conducted using a negative binomial GLM with the EdgeR package [29]. 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 [30].

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 [31, 26]. 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 AsteI2 genome using miRanda software package [32]. 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 [29]. miRNAs which were differentially expressed

with log2FC +/- 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 [33]. piRNAs were isolated from the small RNA libraries by

selecting all 24 – 30 nt reads from the trimmed datasets and filtering out all identified mature

miRNAs, and those mapping to the AsteI2 genome were considered potential piRNAs. piRNA

alignment to the AsteI2 genome was performed using STAR RNA sequencing aligner allowing

for 3 mismatches across the length of the read [27].

RESULTS/DISCUSSION

Transcriptome

RNA Sequencing

We 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 20.6 - 28.4 million paired end reads per library. (Supplementary Table 1). 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 1).

Differential Expression

To determine which genes exhibit differential expression by infection status and

between time points a general-linearized model (GLM) was performed on filtered and

normalized read counts mapping to the Aste12 genome (Table 1, details in Supplementary Table 2; Figure 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. Genes were considered significantly regulated if they had a log fold-change (log2FC) value of +/- 1 and P value < 0.05.

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, ASTEI03083) and 1 depleted (ASTEI04716). The gene with the strongest response to infection at any time point was ASTEI04601 at 2 dpi with a log2FC of -9.8 and the most enriched gene was ASTEI04639 at 14 dpi with log2FC pf 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) to 7.5 (ASTEI09252) log2FC in terms of magnitude of expression, and -3.4 (ASTEI10804) to 7.4 (ASTEI04639) 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 (Supplementary Table 3) [33]. At 2 dpi depleted terms were overrepresented by proteins categorized as integral membrane components, but no terms were considered overrepresented by upregulated transcripts. At 7 dpi depleted transcripts were biased for odorant binding proteins, and enriched transcripts by endopeptidases. At 14 dpi the upregulated transcripts were not biased for any GO terms MAPK/JNK signaling cascades were overrepresented by downregulated genes at 14 dpi with activity being localized to peroxisomes and chromatin. Between 2 and 7 dpi upregulated transcripts were biased for serine type endopeptidases, specifically serine hydrolases. 7 to 14 dpi saw terms associated with G protein-coupled receptor signaling overrepresented by upregulated transcripts.

Endopeptidases, specifically serine proteases were upregulated 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 [33]. 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 [16-18, 34]. At late stages of infection there was depletion of the autophagic inducing JNK and MAPK cascades in addition to repression of JAK/STAT signaling pathways through repression MAPK signaling, both of which have positive impacts on alphaviral replication [33, 35].

Small RNA

miRNA Identification

We next identified novel miRNAs in the small RNA transcriptomes of the MAYV infected samples and controls using miRDeep [31]. We 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 *Acyrthosiphon pisum*. We found matches to 74 known miRNAs, all from *An. gambiae*, and 79 novel miRNAs identified across all samples, with between 2.2 – 4.0 million reads mapping to identified miRNAs per-sample (Supplementary Table 4). We found no explicit relationship between diversity of miRNA population 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). PCA of read counts for all identified miRNAs in each sample showed no obvious grouping patterns at 2 and 7 dpi, however there was a correlation between infection status and PC1/PC2 placement at 14 dpi (Figure 4).

miRNA Differential Expression

We next identified known and novel miRNAs that were differentially expressed by infection status (Figure 5; Table 2, Supplementary Table 5). 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 with results from control treatments filtered out. miRNAs were

considered differentially expressed by having a log fold-change (log2FC) value of +/- 1 and P

value < 0.05.

There were a total of 7 miRNAs differentially regulated in any considered contrast, novel

miRNAs as-mir10, as-mir16, and as-mir17 as well as known miRNAs aga-miR-286b-5p, aga-miR-

2944a-5p, aga-miR-2944b-5p, and aga-miR-309. as-mir10 was upregulated at both 2 and 7 dpi,

as-mir16 was upregulated at 7 dpi, and as-mir 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 only miRNA found to be significantly regulated

with an FDR cutoff of 0.1 was aga-miR-2944a-5p downregulated at 7 dpi.

The miR-309/286/2944 has been found to be upregulated in *An. gambiae* in response to

bloodfeeding [36 – 37], and to be associated with Argonaute proteins post-bloodmeal [36].

When experimentally repressed aga- miR-309 was found to retard oocyte development [36], so

it's downregulation in response to MAYV infection may suggest that viral replication sequesters

resources normally requires for host oocyte development, and as a result associated miRNAs

are also downregulated.

miRNA Target Prediction

We next identified putative targets in the *An. stephensi* genome for all known and novel

miRNAs identified in all samples [32]. For each of the identified miRNAs we found an average

537 potential annotated targets within the Astel2 genome (Supplementary Table 6). Targets for

significantly regulated miRNAs were loaded into g:Profiler and any overrepresented GO terms

with an FDR < 0.5 were considered significant (Table 3) [33].

as-mir-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-mir17 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-5p and aga-mir-2944b-5p 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 functions, and aga-mir-2944b-5p also appears to be involved with lipid localization and transport.

The novel miRNA as-mir-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-mir-17 was significantly repressed in response to MAYV infection. 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 novel miRNAs between 7 - 14 dpi where their expression is enhanced but they still have a bias for upregulated genomic targets [38]. 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 [39].

piRNA Identification

Virus-derived piRNA-like small RNAs (25–30 nt), 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 [40 – 45]. 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 [41, 44].

We identified putative piRNAs in the trimmed small RNA datasets for each sample by isolating all 24-30 nt reads, removing those that were identified positively as miRNAs, and mapping the remaining reads to the Mayaro Virus BeAr 20290 genome using the STAR sequence aligner [26 – 27, 46]. There was no particular bias for viral piRNA abundance in infected samples over control, and the proportion of potential piRNAs mapping to the viral genome remained consistent across time points (Table 4). There was a bias for piRNAs mapping to the negative strand over the positive strand, and a hotspot in the reads mapping to the negative strand at position 6990 overlapping with nonstructural protein 4 (nsP4) and at 7915 overlapping with structural C polyprotein (C) (Figure 6, Supplementary Figure 1) [47]. There does appear to be a bias in potential piRNAs mapped to the viral genome for U in position 1, but no bias for A at position 10 was observed in this study (Figure 7, Supplementary Figure 2) [48].

Mosquito cells infected with Alphaviruses and Bunyaviruses show clear U1 and A10 ping-pong piRNA signature [41, 43, 49]. In the current study reads in the 25-30 nt range mapping to the MAYV genome have a clear U1 bias but not A10, suggesting perhaps primary piRNA production without secondary piRNA biogenesis as a result as the A10 bias is suggested to be a product of cleavage activity (Figure 7) [48]. It is also observed that there are reads in the control groupings also matching the MAYV genome with the same hotspots being produced.

What may be observed is instead a general antiviral response as the viral ORFs nsP4 and C are hotspots [50], suggesting these reads are possibly constituently circulating, or expressed following bloodmeal ingestion (Figure 6).

Conclusion

The transcriptomic profiles suggest that MAYV activates the Toll pathway at mid-stages of infection as an innate humoral response from the host to fight infection. At later stages of infection MAYV appears to repress autophagic processes to promote replication. The small RNA profiles produced suggest a potential reliance on piRNA biogenesis as a generalized antiviral immune response, but that it was not active against MAYV in this study in the sense that a ping-pong profile of potential piRNA reads mapping to the viral genome was not observed. miRNAs were also elicited in response to infection and some overlap was observed with transcripts identified as regulated in response to infection, but not to the extent that they appear to be strongly regulating transcriptional profiles in response to infection.

Acknowledgments.

We thank the UTMB sequencing core for assistance with next generation sequencing. 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 JLR, BBSRC awards BB/T001240/1 and V011278/1, Royal Society Wolfson Fellowship RSWF\R1\180013, NIH grants R21AI138074 and R21AI129507, URKI grant 20197, and NIHR grant NIHR2000907 to GLH. CAH was supported by an NSF graduate research fellowship program award (ID 2018258101). SH was supported

the Liverpool School of Tropical Medicine Director's Catalyst Fund award. GLH is affiliated to the National Institute for Health Research Health Protection Research Unit (NIHR HPRU) in Emerging and Zoonotic Infections at University of Liverpool in partnership with Public Health England (PHE), in collaboration with Liverpool School of Tropical Medicine and the University of Oxford. GLH is based at LSTM. 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. We would also like to thank Dr. Martin Donnelly at Liverpool School of Tropical Medicine for comments on an early version of this manuscript.

References

1.   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

2.   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

3.   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

4.   Llagonne-Barets M, Icard V, Leparc-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

5.   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.

6.   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

7.   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

8.   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

9.   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

10.   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

11.   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.

12. 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

13. 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

14. 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=jc2533t2t4ld65j13rldd b4r55. Accessed February 24, 2019.

15. 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

16. 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

17. 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

18. 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

19. 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

20.   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

21.   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

22.   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*. 2017;11(10):e0006010. doi:10.1371/journal.pntd.0006010

23.   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

24.   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

25.   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

26.   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

27.   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

28. 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

29. 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/edgeR UsersGuide.pdf. Accessed September 24, 2018.

30. 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

31. GitHub - rajewsky-lab/mirdeep2: Discovering known and novel miRNAs from small RNA sequencing data. https://github.com/rajewsky-lab/mirdeep2. Accessed October 8, 2019.

32. 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

33. 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

34. 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

35. 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

36. 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

37.   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

38.   Blair CD, Olson KE. The role of RNA interference (RNAi) in arbovirus-vector interactions. *Viruses*. 2015;7(2):820-843. doi:10.3390/v7020820

39.   Xiao M, Li J, Li W, et al. MicroRNAs activate gene transcription epigenetically as an enhancer trigger. *RNA Biol*. 2017;14(10):1326-1334. 777doi:10.1080/15476286.2015.1112487

40.   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

41.   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

42.   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

43.   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

44. 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

45. 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\

46. 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

47. Lavergne A, Thoisy B de, Lacoste V, et al. Mayaro virus: Complete nucleotide sequence and phylogenetic relationships with other alphaviruses. *Virus Res*. 2006;117(2):283-290. doi:10.1016/j.virusres.2005.11.006

48. Brennecke J, Aravin AA, Stark A, et al. Discrete Small RNA-Generating Loci as Master Regulators of Transposon Activity in Drosophila. *Cell*. 2007;128(6):1089-1103. doi:10.1016/j.cell.2007.01.043

49. Dietrich I, Jansen S, Fall G, et al. RNA Interference Restricts Rift Valley Fever Virus in Multiple Insect Systems. *mSphere*. 2017;2(3). doi:10.1128/msphere.00090-17

50. Auguste AJ, Liria J, Forrester NL, et al. Evolutionary and Ecological Characterization of Mayaro Virus Strains Isolated during an Outbreak, Venezuela, 2010. *Emerg Infect Dis*. 2015;21(10):1742-1750. doi:10.3201/eid2110.141660

Figure legends:

Figure 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.

Figure 2: Volcano plots visualizing differential expression of Anopheles stephensi transcripts in response to Mayaro infection. The Y-axis shows -log10 transformed P-values, and the X-axis shows log2 transformed fold change values. Red points represent transcripts downregulated by more than -1 log2FC in response to infection with a FDR < 0.05, while blue points are transcripts upregulated by more than 1 log2FC in response to infection with a 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.

Figure 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 4: Principal Components Analysis (PCA) on read counts mapping to miRNAs identified in the AsteI2 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.

Figure 5: Volcano plots visualizing differential expression of identified *Anopheles stephensi* miRNAs in response to Mayaro infection. The Y-axis shows -log10 transformed P-values, and the X-axis shows log2 transformed fold change values. Red points represent transcripts downregulated by more than -1 log2FC in response to infection with a FDR < 0.05, while blue points are transcripts upregulated by more than 1 log2FC in response to infection with a 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.

Figure 6: Histograms demonstrating read depth across the Mayaro virus genome for reads with a piRNA size profile (24 - 30 nt). Y-axis is read depth, and X-axis is position in viral genome. Blue demonstrates reads for that sample mapping to the positive strand, while red demonstrates those mapping to the negative strand. A. - C. are control for 2, 7, and 14 dpi respectively, and D. - F. are infected for 2, 7, and 14 dpi. One replicate per treatment is chosen here to demonstrate, but all replicates are in Supplementary Figure 1.

Figure 7: Nucleotide bias at first 15 bp of potential piRNA reads mapping to the Mayaro virus genome. X-axis is position in read, and Y-axis is nucleotide bias. Size of nucleotide demonstrates relative bias at that position in the read. A. - C. are control for 2, 7, and 14 dpi respectively, and

D. - F. are infected for 2, 7, and 14 dpi. One replicate per treatment is chosen here to demonstrate, but all replicates are in Supplementary Figure 2.

Table 1: Differentially expressed *Anopheles stephensi* transcripts in response to Mayaro virus infection. A. – C. demonstrate differentially expressed transcripts between control and infected treatments at 2, 7, and 14 dpi respectively. D. and E. represent differentially expressed transcripts between 2 – 7 dpi and 7 – 14 dpi for infected treatments correcting results from control treatments. Only transcripts differentially expressed with logFC of +/- 2 are shown here, more detailed results are provided in Supplementary Table 2.

A.

gene ID	logFC	logCPM	F	P-Value	FDR	Direction
				2 dpi		

B.

gene ID	logFC	logCPM	F	P-Value	FDR	Direction
				7 dpi		

C.

gene ID	logFC	logCPM	F	P-Value	FDR	Direction
				14 dpi		

D.

gene ID	logFC	logCPM	F	P-Value	FDR	Direction
				7 - 2 dpi		

E.

gene ID	logFC	logCPM	F	P-Value	FDR	Direction
				14 - 7 dpi		

Table 2: Differentially expressed *Anopheles stephensi* miRNAs in response to Mayaro virus infection. A. and B. demonstrate differentially expressed miRNAs between control and infected treatments at 2, and 7 dpi respectively. C. and D. represent differentially expressed miRNAs between $2-7$ dpi and $7-14$ dpi for infected and control treatments.

A.

miRNA ID	logFC	logCPM	F	PValue	FDR	direction
			2 dpi			
as_mir10	8.91902202	3.99960865	6.04342539	0.03588223	0.6006979	POSITIVE

B.

miRNA ID	logFC	logCPM	F	PValue	FDR	direction
			7 dpi			
as_mir10	9.66274326	3.99960865	5.71198169	0.04016716	0.24336571	POSITIVE
as-mir16	6.1676765	2.17663742	6.87820325	0.02735664	0.21674873	POSITIVE
aga_miR 286b_5p	7.2138078	11.1822434	7.9400609	0.01440288	0.1791391	NEGATIVE
aga-miR-2944a-5p	-5.6923915	12.3223923	13.9128375	2.48E-03	0.08143048	NEGATIVE
aga-miR-2944b-5p	-6.2234331	10.8056821	10.5456059	0.00628318	0.1078613	NEGATIVE

C.

miRNA ID	logFC	logCPM	F	PValue	FDR	direction
			14 dpi			
as-mir17	-8.0636413	1.98897721	8.46359478	0.01706672	0.17196133	NEGATIVE

D.

miRNA ID	logFC	logCPM	F	PValue	FDR	direction
			2-7 dpi			
aga-miR-2944a-5p	-5.0392774	12.3223923	7.30236167	0.01797716	0.42794374	NEGATIVE
aga-miR-2944b-5p	-5.4387568	10.8056821	5.57879027	0.0342666	0.50420854	NEGATIVE
aga-miR-309	-5.7845649	10.4454892	6.79900551	2.15E-02	0.42794374	NEGATIVE

E.

miRNA ID	logFC	logCPM	F	PValue	FDR	direction
			7-14 dpi			
aga-miR-286b-5p	7.12647125	11.1822434	4.77967974	0.04747986	0.26011923	POSITIVE
aga-miR-2944a-5p	5.85356522	12.3223923	9.4915733	0.00867114	0.13109483	POSITIVE
aga-miR-2944b-5p	6.69560733	10.8056821	7.96456984	1.43E-02	0.13109483	POSITIVE
aga-miR-309	6.67824682	10.4454892	8.37514119	0.0124383	0.13109483	POSITIVE
as-mir17	-9.009352	1.98897721	7.3170676	0.02368014	0.16397693	NEGATIVE

Table 3: Overrepresented GO terms represented by targets of significantly regulated miRNAs.

A

	aga-mir-2944b-5p			
Term Name	ID	FDR	Term Size	Intersection Size
Molecular Function				
protein binding	GO:0005515	7.53E-03	1791	188
Rho guanyl nucleotide exchange factor activity	GO:0005089	7.53E-03	27	10
phosphorus oxygen lyase activity	GO:0016849	7.53E-03	23	9
Ras guanyl-nucleotide exchange factor activity	GO:0005088	7.53E-03	29	10
nucleoside-triphosphatase activity	GO:0017111	1.12E-02	313	45
binding	GO:0005488	1.12E-02	4675	420
histone methyltransferase activity	GO:0042054	1.12E-02	12	6
histone-lysine N-methyltransferase activity	GO:0018024	1.12E-02	12	6
hydrolase activity, acting on acid anhydrides	GO:0016817	1.17E-02	328	45
hydrolase activity, acting on acid anhydrides, in phosphorus containing anhydrides	GO:0016818	1.17E-02	325	45
pyrophosphatase activity	GO:0016462	1.17E-02	322	45
S-adenosylmethionine-dependent methyltransferase activity	GO:0008757	1.17E-02	40	11
cyclic nucleotide phosphodiesterase activity	GO:0004112	1.17E-02	9	5
3',5'-cyclic-nucleotide phosphodiesterase activity	GO:0004114	1.17E-02	9	5
guanyl-nucleotide exchange factor activity	GO:0005085	1.17E-02	54	13
Rho GTPase binding	GO:0017048	1.42E-02	36	10
protein-lysine N-methyltransferase activity	GO:0016279	2.18E-02	15	6
lysine N-methyltransferase activity	GO:0016278	2.18E-02	15	6
guanylate cyclase activity	GO:0004383	2.79E-02	11	5
purine nucleotide binding	GO:0017076	2.79E-02	832	92
purine ribonucleotide binding	GO:0032555	2.79E-02	830	92
lipid transporter activity	GO:0005319	3.01E-02	22	7
protein methyltransferase activity	GO:0008276	3.01E-02	22	7
ribonucleotide binding	GO:0032553	3.07E-02	838	92
purine ribonucleoside triphosphate binding	GO:0035639	3.43E-02	821	90
N-methyltransferase activity	GO:0008170	3.43E-02	23	7
calcium channel activity	GO:0005262	3.43E-02	12	5
phosphoric ester hydrolase activity	GO:0042578	4.34E-02	74	7
GTPase activity	GO:0003924	4.74E-02	105	18
nucleoside phosphate binding	GO:1901265	4.85E-02	956	101
nucleotide binding	GO:0000166	4.85E-02	956	101
Biological Process				
intracellular signal transduction	GO:0035556	3.45E-06	193	42
Rho protein signal transduction	GO:0007266	4.33E-03	29	11
Ras protein signal transduction	GO:0007265	4.33E-03	40	13
cyclic nucleotide metabolic process	GO:0009187	9.48E-03	23	9
cyclic nucleotide biosynthetic process	GO:0009190	9.48E-03	23	9
regulation of Rho protein signal transduction	GO:0035023	9.48E-03	28	10
regulation of Ras protein signal transduction	GO:0046578	1.13E-02	35	11
signal transduction	GO:0007165	1.13E-02	681	84
small GTPase mediated signal transduction	GO:0007264	1.33E-02	75	17
regulation of intracellular signal transduction	GO:1902531	1.39E-02	67	15
cell communication	GO:0007154	1.59E-02	696	84
signaling	GO:0023052	1.61E-02	696	84
regulation of response to stimulus	GO:0048583	1.82E-02	94	19
response to stimulus	GO:0050896	2.51E-02	876	100
regulation of signal transduction	GO:0009966	2.64E-02	90	18
regulation of cell communication	GO:0010646	3.11E-02	92	18
regulation of small GTPase mediated signal transduction	GO:0051056	3.11E-02	42	11
histone lysine methylation	GO:0034968	3.42E-02	14	6
lipid transport	GO:0006869	3.73E-02	50	12
regulation of signaling	GO:0023051	3.94E-02	95	18
cellular response to stimulus	GO:0051716	4.33E-02	826	93
lipid localization	GO:0010876	4.66E-02	52	12

B

	aga-mir-2944a-5p			
Term Name	ID	FDR	Term Size	Intersection Size
Molecular Function				
protein binding	GO:0005515	3.46E-04	1791	150
purine ribonucleoside triphosphate binding	GO:0035639	6.99E-03	821	75
adenyl ribonucleotide binding	GO:0032559	6.99E-03	684	64
purine ribonucleotide binding	GO:0032555	6.99E-03	830	76
ribonucleotide binding	GO:0032553	6.99E-03	838	76
adenyl nucleotide binding	GO:0030554	6.99E-03	686	64
purine nucleotide binding	GO:0017076	6.99E-03	832	76
nucleoside phosphate binding	GO:19C1265	6.99E-03	956	83
nucleotide binding	GO:0000166	6.99E-03	956	83
ATP binding	GO:0005524	6.99E-03	675	63
Rho guanyl-nucleotide exchange factor activity	GO:0005089	6.99E-03	27	8
kinase activity	GO:0016301	7.97E-03	290	33
Ras guanyl nucleotide exchange factor activity	GO:0005088	7.97E-03	29	8
binding	GO:0005488	1.28E-02	4675	309
small molecule binding	GO:0036094	2.01E-02	1011	84
transferase activity, transferring phosphorus-containing groups	GO:0016772	2.07E-02	386	39
carbohydrate derivative binding	GO:0097367	2.17E-02	959	80
phosphotransferase activity, alcohol group as acceptor	GO:0016773	2.45E-02	264	29
Rho GTPase binding	GO:0017048	2.49E-02	36	8
phosphorus oxygen lyase activity	GO:0016849	4.22E-02	23	6
drug binding	GO:0008144	4.60E-02	815	68
anion binding	GO:0043168	4.78E-02	979	79
Biological Process				
intracellular signal transduction	GO:0035556	4.13E-03	193	29

C

	aga-mir-309			
Term Name	ID	FDR	Term Size	Intersection Size
Molecular Function				
calcium ion binding	GO:0005509	1.91E-03	170	26

D

	ae-mir16			
Term Name	ID	FDR	Term Size	Intersection Size
Molecular Function				
protein binding	GO:0005515	1.15E-02	1791	153
binding	GO:0005488	1.86E-02	4675	336

E

	ae-mir17			
Term Name	ID	FDR	Term Size	Intersection Size
Molecular Function				
calcium ion binding	GO:0005509	0.00361388	172	22
inorganic molecular entity transmembrane transporter activity	GO:0015318	0.00453046	267	29
channel activity	GO:0015267	0.01989068	179	21
passive transmembrane transporter activity	GO:0022803	0.01989068	179	21
metal ion binding	GO:0046872	0.0225287	1097	78
ion transmembrane transporter activity	GO:0015075	0.02452428	293	29
cation binding	GO:0043169	0.02955644	1106	78
ion channel activity	GO:0005216	0.03515638	173	20

Table 4: Number and proportion of reads from each small RNA sample sequenced mapping to the host genome, mapping to miRNAs, mapping to the viral genome, and of those mapping to the viral genome that are potential piRNAs.

ID	Total Reads QC Passed	Mapping to Host Genome	Proportion of Total Mapping to Host	Mapping to Host miRNAs	Proportion of Host Reads Mapping to miRNAs	Mapping to Viral Genome	Proportion of Total mapping to Virus	Viral piRNAs	Proportion Mapping to Viral as Potential piRNAs
C_D2_R1	17800192	16126876	91%	3033086	19%	199006	1%	7960	4%
C_D2_R2	16812408	15056380	90%	2254482	15%	272599	2%	8569	3%
C_D2_R3	18092880	15127997	84%	2551279	17%	275676	2%	7103	3%
C_D7_R1	21189926	18142804	86%	179672	1%	797862	4%	15559	2%
C_D7_R2	14833688	13591924	92%	2254482	17%	96190	1%	6641	7%
C_D7_R3	20551719	18828285	92%	3617884	19%	187169	1%	8276	4%
C_D14_R1	20259396	18306112	90%	3575305	20%	180217	1%	10368	6%
C_D14_R2	17801450	16105271	90%	3033309	19%	140748	1%	8647	6%
C_D14_R3	18082055	16502692	91%	3794362	23%	148854	1%	10335	7%
V_D2_R1	15851816	14109550	89%	2798105	20%	214624	1%	9641	4%
V_D2_R2	18043897	15456543	86%	2600968	17%	317283	2%	13052	4%
V_D2_R3	17499685	14894859	85%	2076112	14%	255560	1%	10647	4%
V_D7_R1	15721343	14395465	92%	2977219	21%	262465	2%	11763	4%
V_D7_R2	23780788	21664761	91%	4000416	18%	277922	1%	17215	6%
V_D7_R3	20775885	18719374	90%	3376672	18%	295778	1%	16856	6%
V_D14_R1	18174128	16399105	90%	3227809	20%	332311	2%	16246	5%
V_D14_R2	19188815	17548229	91%	2658496	15%	303330	2%	12630	4%
V_D14_R3	20525522	18517727	90%	2897260	16%	309427	2%	16246	5%

Supplementary Figure 1: Histograms demonstrating read depth across the Mayaro virus genome for reads with a piRNA size profile (24 - 30 nt). Represents all samples, chosen subset demonstrated by Figure 6.

Supplementary Figure 2: Nucleotide bias at first 15 bp of potential piRNA reads mapping to the Mayaro virus genome. Represents all samples, chosen subset demonstrated by Figure 7.

Supplementary Table 1: Information related to infection of *Anopheles stephensi* with Mayaro virus. Includes number of mosquitoes in each treatment and time point and associated mortality, nanodrop readings for all RNA extractions collected, pooling scheme for sequencing of mRNA and small RNA, and qPCR data from each sample using primers specific for Mayaro virus strain BeAn to confirm infection status.

Supplementary Table 2: Differentially expressed transcripts from the *Anopheles stephensi* AsteI2 genome.

Supplementary Table 3: GO term overrepresentation for differentially regulated transcripts.

Supplementary Table 4: Read counts mapping to the identified *Anopheles stephensi* miRNAs in each small RNA sample sequenced.

Supplementary Table 5: Differential expression of *Anopheles stephensi* miRNAs.

Supplementary Table 6: Genomic targets from the *Anopheles stephensi* AsteI2 genome for all

identified miRNAs.

Figure 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 respectivey. In all PCAs, blue are Mayaro infected, and red are control.

Figure 2: Volcano plots visualizing differential expression of Anopheles stephensi transcripts in response to Mayaro infection. The Y-axis shows -log10 transformed P-values, and the X-axis shows log2 transformed fold change values. Red points represent transcripts downregulated by more than -1 logFC in response to infection with a P value < 0.05, while blue points are transcripts upregulated by more than 1 logFC in response to infection with a P value < 0.05. A. - C. are trasncripts 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.

Figure 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 4: Principal Components Analysis (PCA) on read counts mapping to miRNAs identified in the AsteI2 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.

Figure 5: Volcano plots visualizing differential expression of identified *Anopheles stephensi* miRNAs in response to Mayaro infection. The Y-axis shows -log10 transformed P-values, and the X-axis shows log2 transformed fold change values. Red points represent transcripts downregulated by more than -1 logFC in response to infection with a P Value < 0.05, while blue points are transcripts upregulated by more than 1 logFC in response to infection with a P Value < 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.

Figure 6: Histograms demonstrating read depth across the Mayaro virus genome for reads with a piRNA size profile (24 - 30 nt). Y-axis is red depth, and X-axis is position in viral genome. Blue demonstrates reads for that sample mapping to the positive strand, while red demonstrates those mapping to the negative strand. A. - C. are control for 2, 7, and 14 dpi respectively, and D. - F. are infected for 2, 7, and 14 dpi. One replicate per treatment is chosen here to demonstrate, but all replicates are in Suplementary Figure 1.

A.

D.

B.

E.

C.

F.

Figure 7: Nucleotide bias at first 15 bp of potential piRNA reads mapping to the Mayaro virus genome. X-axis is position in read, and Y-axis is nucleotide bias. Size of nucleotide demonstrates relative bias at that position in the read. A. - C. are control for 2, 7, and 14 dpi respectively, and D. - F. are infected for 2, 7, and 14 dpi. One replicate per treatment is chosen here to demonstrate, but all replicates are in Suplementary Figure 2.