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

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

Cory Henderson[1], Marco Brustolin[1], Shivanand Hegde[2], Gargi Dayama[3], Nelson Lau[3], Grant L. Hughes[2], Christina Bergey[4], 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 Biochemistry, BU School of Medicine, Boston University, Boston, MA, United States of America

[4]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*, as well as the typical arboviral transmitting mosquitoes in the genus *Aedes*. Few studies have investigated the infection response of *Anopheles* mosquitoes. 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, 78 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.

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 signaling pathways related to cell death are regulated in response to MAYV infection of *An. stephensi*, we also discover novel miRNAs and describe the expression patterns of miRNAs,

44 siRNAs, and piRNAs following bloodmeal ingestion. These results suggest that MAYV does

45 induce a molecular response to infection in its mosquito vector species.

46 **INTRODUCTION**

47 Mayaro virus (MAYV) is a mosquito-borne, enveloped positive-sense single-stranded
48 RNA virus in the genus *Alphavirus*, first isolated from the blood of five febrile workers in
49 Mayaro county, Trinidad in 1954 [1]. Symptoms of MAYV infection are similar to other arboviral
50 infections such as Dengue (DENV) or Chikungunya viruses, (CHIKV) and include rash, fever,
51 retro-orbital pain, headache, diarrhea, and arthralgia [2]. While no epidemics or outbreaks with
52 MAYV being the causative agent have been recorded outside of South America, there have
53 been imported cases reported in the Netherlands, Germany, France, and Switzerland [3-6],
54 which demonstrates a need to understand the capacity for the virus to spread into naïve
55 regions, such as the United States.

56 The principal mosquitoes transmitting MAYV naturally are thought to be the canopy-
57 dwellers of the genus *Haemogogus*, maintaining the sylvatic cycle between non-human
58 primates as primary hosts and birds as secondary hosts [7]. Human infections are sporadic due
59 to the rare display of anthropophilic biting behaviors by *Haemogogus* mosquitoes, with
60 transmission due to these species primarily occurring in rural regions with close proximity to
61 forests [8]. Vector competence studies have identified anthropophilic and urban adapted
62 species such as *Aedes aegypti* and *Ae. albopictus,* as well as the malaria parasite transmitters
63 *Anopheles gambiae*, *An. stephensi*, *An. freeborni*, and *An. quadrimaculatus*, as being competent
64 vectors for MAYV under laboratory conditions [9-12]. Transmission of an arbovirus with an
65 anopheline mosquito as a primary vector is rare, only having been observed occurring regularly
66 for O'nyong'nyong virus by *An. gambiae* and *An. funestus* in Uganda [13], with some limited
67 evidence for CHIKV and Semliki Forest virus (SFV) [14].

68 As arboviral pathogens are transmitted between hosts primarily by arthropod vectors,
69 transmission requires the virus to infect and disseminate from the midgut and salivary glands of
70 the mosquito following an infectious bloodmeal [15]. The molecular underpinnings controlling
71 why MAYV and these closely related viruses can infect *Anopheles* salivary glands is of
72 epidemiological interest, yet remains poorly understood. A more complete understanding of
73 this phenomenon requires investigation of the molecular pathways involved in viral infection of
74 anopheline mosquitoes. Recent transcriptomic studies have identified a number of genes
75 involved in classical immune pathways, RNA interference (RNAi), metabolism, energy
76 production, and transport as being regulated in response to arboviral infection of mosquitoes
77 [16-19]. In addition, studies focusing on small RNA identification and regulation have identified
78 RNAi activity, such as miRNA, piRNA, and siRNA expression, in response to infection of
79 mosquitoes by arboviruses [20-23].
80 The available evidence suggests that, should MAYV be introduced into a naïve region,
81 outbreaks and epidemics of the resulting disease could be driven by anopheline vectors [9, 24-
82 25]. Because anopheline, and not aedine, mosquitoes could act as the primary transmitting
83 vectors for MAYV, this study also provides an opportunity to understand how vector
84 competence might emerge in this system and provide insight into why *Anopheles* are generally
85 poor viral transmitters when compared to *Aedes* mosquitoes. We used RNA sequencing to
86 study the transcriptomic and small RNA responses of *An. stephensi* to infection with MAYV via
87 infectious bloodmeal at 2, 7, and 14 days post infection (dpi).

88 **MATERIALS AND METHODS**

89 *Anopheles stephensi* Rearing

90 Protocols pertaining to mosquito rearing and presentation of infectious bloodmeal has
91 been described elsewhere [9]. Briefly, *An. stephensi* (Liston strain) were provided by Johns
92 Hopkins University (Baltimore, MD, USA). Mosquito colonies were reared and maintained at the
93 Millennium Sciences Complex insectary (The Pennsylvania State University, University Park, PA,
94 USA) at 27˚C ± 1˚C, 12 hour light 12 hour dark diurnal cycle at 80% relative humidity in
95 30×30×30-cm cages. Ground fish flakes (TetraMin, Melle, Germany) were used to feed larvae,
96 and upon emergence adult mosquitoes were maintained with a 10% sucrose solution.

97 Viral Production and Infection via Bloodmeal

98 Mayaro virus strain BeAn 343102 (BEI Resources, Manassas, VA, USA) was utilized in this
99 study, a genotype D strain originally isolated from a monkey in Para, Brazil, in May 1978. Virus-
100 infected supernatant was aliquoted and stored at −80˚C until used for mosquito infections. Viral
101 stock titers were obtained by focus forming assay (FFA) technique. Adult female mosquitoes at
102 6 days post emergence that had not previously blood-fed were used for experimentation.
103 Mosquitoes were allowed to feed on either human blood spiked with MAYV at 1*107 FFU/mL
104 or a control bloodmeal with no virus via a glass feeder jacketed with 37˚C distilled water for 1 h.
105 At 2, 7, and 14 days post infection, mosquitoes were anesthetized with triethylamine
106 (Sigma, St. Louis, MO, USA) and RNA was extracted from each individual mosquito
107 using mirVana RNA extraction kit (Life Technologies) applying the protocol for extraction of
108 total RNA. Infection was confirmed via qPCR using primers published by Wiggins et. al. 2018
109 (Forward: 5′-TGGACCTTTGGCTCTTCTTATC-3′, Reverse: 5′-GACGCTCACTGCGACTAAA-3′) [10], a
110 CT value of 20 or less was used to confirm infection (Supplementary Table 1). 3 pools of total
111 RNA were created for each time point and infection status to be used for library preparation,

112 each consisting of 750 ng of RNA from 4 mosquitoes for a total of 3 mg per pool as confirmed
113 via nanodrop. The protocol for mosquito rearing, viral production, and infection via bloodmeal
114 is described in more detail in Brustolin et al. 2018 [9].

115 Transcriptomic Library Preparation and Sequencing

116 All pools were sent to University of Texas Medical Branch for library preparation where
117 total RNA was quantified using a Qubit fluorescent assay (Thermo Scientific) and RNA quality
118 was assessed using an RNA 6000 chip on an Agilent 2100 Bioanalyzer (Agilent Technologies).
119 See Etebari et all. 2017 for more detail on library preparation and sequencing [17]. 1 mg of total
120 RNA per pool was poly-A selected and fragmented using divalent cations and heat (940 C, 8
121 min). The NEBNext Ultra II RNA library kit (New England Biolabs) was used for RNA-Seq library
122 construction. Fragmented poly-A RNA samples were converted to cDNA by random primed
123 synthesis using ProtoScript II reverse transcriptase (New England Biolabs). After second strand
124 synthesis, the double-stranded DNAs were treated with T4 DNA polymerase, 5' phosphorylated
125 and then an adenine residue was added to the 3' ends of the DNA. Adapters were then ligated
126 to the ends of these target template DNAs. After ligation, the template DNAs were amplified (5-
127 9 cycles) using primers specific to each of the non-complimentary sequences in the adapters.
128 This created a library of DNA templates that have non-homologous 5' and 3' ends. A qPCR
129 analysis was performed to determine the template concentration of each library. Reference
130 standards cloned from a HeLa S3 RNA-Seq library were used in the qPCR analysis. Cluster
131 formation was performed using 15.5-17 billion templates per lane using the Illumina cBot v3
132 system. Sequencing by synthesis, paired end 75 base reads, was performed on an Illumina
133 NextSeq 5500 using a protocol recommended by the manufacturer.

134 Small RNA Library Preparation and Sequencing

135 Small RNA libraries were created using the New England Biolabs small RNA library
136 protocol. See Saldaña et. al. 2017 for more information on small RNA sequencing [21]. Library
137 construction used a two-step ligation process to create templates compatible with Illumina
138 based next generation sequence (NGS) analysis. Where appropriate, RNA samples were
139 quantified using a Qubit fluorometric assay. RNA quality was assessed using a pico-RNA chip on
140 an Agilent 2100 Bioanalyzer. Library creation uses a sequential addition of first a 3' adapter
141 sequence followed by a 5' adapter sequence. A cDNA copy was then synthesized using
142 ProtoScript reverse transcriptase and a primer complimentary to a segment of the 3' adapter.
143 Amplification of the template population was performed in 15 cycles (94˚C for 30 sec; 62˚C for
144 30 sec; 70˚C for 30 sec) and the amplified templates were PAGE (polyacrylamide gel
145 electrophoresis) purified (147 bp DNA) prior to sequencing. All NGS libraries were indexed. The
146 final concentration of all NGS libraries was determined using a Qubit fluorometric assay and the
147 DNA fragment size of each library was assessed using a DNA 1000 high sensitivity chip and an
148 Agilent 2100 Bioanalyzer. Single end 75 base sequencing by synthesis on an Illumina NextSeq
149 5500.

150 Transcriptomic RNA Sequencing Data Analysis

151 Raw sequencing data was uploaded to the ICS-ACI high performance computing cluster
152 at Pennsylvania State University to perform all computational analyses. Transcriptomic libraries
153 had adapters trimmed and low-quality bases removed using Trimmomatic read trimming
154 software with base settings [26]. Quality trimmed reads were aligned to the current build of the
155 Anopheles stephensi Indian strain genome in Vectorbase (AsteI2) using the STAR RNA

156 sequencing aligner [27]. Reads less than 75 bp in length and with a mapping quality of less than
157 20 were dropped from the analysis, and read counts were calculated in R using the rSubread
158 package [28], following which a principal components analysis was performed and differential
159 expression conducted using a negative binomial GLM with the EdgeR package [29]. Contrasts
160 considered in the GLM were infected against control at 2, 7, and 14 dpi, and differences
161 between 2 -7 dpi and 7 – 14 dpi for infected treatments corrected for the response from the
162 control treatments between the same time points. Gene IDs that were differentially expressed
163 with a log2FC value of +/- 1 and P value < 0.05 were uploaded to g:Profiler to run GO term
164 overrepresentation analysis [30].

165 Small RNA Sequencing Data Analysis

166 Small RNA libraries had adapters trimmed using Trimmomatic and were subsequently
167 passed into the miRDeep2 pipeline to identify novel and known miRNAs in all samples and
168 determine expression of all known and novel miRNAs at each time point and treatment status
169 [31, 26]. Novel miRNAs with a miRDeep score of less than 3, a minimum free energy value of
170 less than – 20, or a non-significant Randfold p-value were considered false IDs and excluded
171 from further analysis. miRNA targets were identified in the AsteI2 genome using miRanda
172 software package [32]. Differential expression of miRNAs in response to infection status and
173 time point was conducted using a negative binomial GLM with the EdgeR package and contrasts
174 as described for the transcriptomic analysis [29]. miRNAs which were differentially expressed
175 with log2FC +/- 1 and P value < 0.05 had their miRanda genomic targets uploaded to g:Profiler
176 to determine if any GO terms were overrepresented by transcripts potentially regulated by
177 differentially expressed miRNAs [33]. piRNAs and siRNAs were isolated from the small RNA

178 libraries by selecting all 18 – 24 nt reads (siRNA) and 24 – 35 nt (piRNA) reads from the trimmed
179 datasets and filtering out all identified mature miRNAs, and those mapping to the MAYV
180 NC_003417.1 genome were considered potential piRNAs or siRNAs. piRNA and siRNA alignment
181 to the AsteI2 genome was performed using Bowtie RNA sequencing aligner within the MSRG
182 pipeline [34]. Observation of fastQC output for Control Day 7 Replicate 1 small RNA sequencing
183 revealed poor sequencing results, so this replicate was omitted from all analyses in the small
184 RNA focused portions of this study [35].

185 Data availability

186 Sequencing data have been deposited in the GEO depository under accession number
187 GSE165488.

188 **RESULTS/DISCUSSION**

189 *Transcriptome*

190 RNA Sequencing

191 We assayed genome-wide gene expression in pools of *An. stephensi* (Liston strain)
192 experimentally infected with MAYV at 2, 7, and 14 dpi, along with blood fed uninfected
193 negative controls. RNAseq libraries were sequenced on the Illumina NextSeq 550 platform,
194 yielding 20.6 - 28.4 million paired end reads per library. (Supplementary Table 1). Principal
195 components analysis (PCA) performed on read counts of each annotated gene in the *An.*
196 *stephensi* (Indian strain) reference transcriptome (AsteI2) at each time point distributed
197 infected and control samples into distinct groups (Figure 1).

198

199 Figure 1: Principal Components Analysis (PCA) on filtered read counts mapping to annotated
200 genes from the AsteI2 build of the *Anopheles stephensi* genome in Vectorbase. A., B., and C. are
201 read counts from samples in the 2, 7, and 14 dpi groupings respectively. In all PCAs, blue is
202 Mayaro infected, and red are control.
203
204 Differential Expression
205 To determine which genes exhibit differential expression by infection status and
206 between time points a general-linearized model (GLM) was performed on filtered and
207 normalized read counts mapping to the AsteI2 genome (Supplementary Table 2; Figure 2).
208 Differential expression was also computed for a newer chromosome level genome assembly
209 (Supplementary Table 3), however, the results from the older assembly are discussed here as
210 the genes in the older assembly contain better annotations than those in the newer assembly
211 [36]. Contrasts considered in the GLM were infected compared to control at 2, 7, and 14 dpi,
212 and differences between 2 -7 dpi and 7 – 14 dpi for infected treatments correcting for results
213 from control treatments between those time points. Genes were considered significantly
214 regulated if they had a log fold-change (log2FC) value of +/- 1 and P value < 0.05.
215
216 Figure 2: Volcano plots visualizing differential expression of *Anopheles stephensi* transcripts in
217 response to Mayaro infection. The Y-axis shows -log10 transformed P-values, and the X-axis
218 shows log2 transformed fold change values. Red points represent transcripts depleted by more
219 than -1 log2FC in response to infection with a FDR < 0.05, while blue points are transcripts
220 enriched by more than 1 log2FC in response to infection with a P value < 0.05. A. - C. are

221 transcripts regulated in the 2 dpi, 7 dpi, and 14 dpi groupings respectively, while D. and E. are
222 transcripts regulated in the infected treatment between 2 - 7 dpi and 7 - 14 dpi respectively.
223
224 There were 161 (64 enriched, 97 depleted), 45 (29 enriched, 16 depleted), and 204 (149
225 enriched, 55 depleted) of 10,313 annotated genes regulated between control and infected at 2,
226 7, and 14 dpi respectively. 3 genes were regulated in the same direction at each time point, 2
227 enriched (ASTEI09037, ASTEI03083) and 1 depleted (ASTEI04716). The gene with the strongest
228 response to infection at any time point was ASTEI04601 at 2 dpi with a log2FC of -9.8 and the
229 most enriched gene was ASTEI04639 at 14 dpi with log2FC of 5.7. When considering changes
230 between time points for the infected treatment when controlling for the response from the
231 uninfected treatments, there were 96 positively and 44 negatively regulated genes between 2 -
232 7 dpi, and 129 enriched and 32 depleted genes between 7 - 14 dpi. Regulated transcripts for 2-7
233 dpi ranged from -6.2 (ASTEI08168) to 7.5 (ASTEI09252) log2FC in terms of magnitude of
234 expression, and -3.4 (ASTEI10804) to 7.4 (ASTEI04639) log2FC for 7-14 dpi. When considering a
235 FDR threshold as a multiple testing correction very few transcripts in any contrast can be
236 considered significantly regulated; 3 transcripts at 2 dpi (ASTEI04601, ASTEI05497, ASTEI05732)
237 and 2 transcripts at 14 dpi (ASTEI00644, ASTEI08604) fall below a FDR < 0.1 threshold for
238 significance.
239 Gene Ontology
240 A gene ontology (GO) over-representation analysis was performed using g:Profiler on
241 gene IDs which were significantly enriched or depleted in any considered contrast in the GLM
242 described above when using a P value cutoff of 0.05 and any overrepresented GO terms with an

243 FDR < 0.5 were considered significant (Supplementary Table 4) [33]. At 2 dpi significantly
244 regulated molecular function terms were overrepresented by peptidase activity, specifically
245 serine type endopeptidase activity, however this overrepresentation is not observed when just
246 considering depleted or enriched genes it is only present when considering all regulated genes
247 at 2 dpi together. Depleted genes at 2 dpi are overrepresented by cell-cell adhesion terms in
248 the biological function category, and components of the membrane for cellular component,
249 enriched genes are not significantly biased for any terms. At 7 dpi molecular function terms
250 related to odorant binding and carboxylic acid binding were overrepresented in depleted genes,
251 and for enriched genes however with a slightly less than significant FDR of 0.06. Molecular
252 function terms relating to dioxygenases, transferases, and oxidoreductases are
253 overrepresented by depleted genes at 7 dpi, and different types of amino acid binding as well
254 as serine type endopeptidase terms are overrepresented by enriched genes. Biological process
255 terms related to nervous system processes, sensory perception of smell, and mannose
256 metabolism are overrepresented by depleted genes at 7 dpi. At 14 dpi the enriched transcripts
257 were biased for molecular function GO terms related to sensory perception, specifically
258 perception of mechanical, light, and sound stimuli, while depleted transcripts were biased for
259 those related to MAPK/JNK signaling cascades, apoptosis, and amino acid biosynthesis for
260 molecular functions and peroxisome and nucleosome for cellular component terms. From 2 to
261 7 dpi molecular function terms related to serine protease activity were represented by both
262 enriched and depleted genes, but primarily by enriched genes, and enriched genes were biased
263 for cell membrane proteins for cellular component. 7 to 14 dpi had enriched genes biased for
264 molecular functions related to ATP dependent microtubule activity and G-protein coupled

265 receptors while depleted genes were biased for different types of amino acid binding. Biological
266 process terms overrepresented by enriched genes between 7 and 14 dpi were related to
267 sensory perception, and depleted genes were overrepresented for various types of metabolic
268 and biosynthetic processes, NF-kB signaling, and the MAPK/JNK cascade.

269 Endopeptidases, specifically serine proteases were enriched at 7 dpi and from 2 – 7 dpi,
270 suggesting an activation of the Toll pathway as part of the innate humoral response to infection
271 once the virus has had time to establish an infection in the mosquito [33]. Activation of serine
272 proteases is not uncommon in pathogenic infection of insects and has been identified
273 specifically as enriched in *Ae. aegypti* in response to DENV and Zika virus (ZIKV) infection, and in
274 *An. gambiae* and *An. coluzzii* in response to O'nyong'nyong virus (ONNV) infection [16-18, 37].
275 At late stages of infection there was depletion of the autophagic and apoptotic inducing JNK
276 and MAPK cascades in addition to repression of JAK/STAT signaling pathways through
277 repression of MAPK signaling. Autophagy and apoptosis both have demonstrated positive
278 impacts on alphaviral replication [33, 38], suggesting another possible molecular response from
279 the mosquito to prevent viral replication at late stages of infection.

280 *Small RNA*

281 <u>miRNA Identification</u>

282 We next identified novel miRNAs in the small RNA transcriptomes of the MAYV infected
283 samples and controls using miRDeep [31]. We searched for matches in our sequencing reads to
284 all miRNAs in the miRBase database for the species *An. gambiae*, *Aedes aegypti, Culex*
285 *quinquefasciatus, Drosophila melanogaster, Bombyx mori, Apis mellifera,* and *Acyrthosiphon*
286 *pisum*. We found matches to 73 known miRNAs, all from *An. gambiae*, and 78 novel miRNAs

287 identified across all samples, with between 2.2 – 4.0 million reads mapping to identified
288 miRNAs per-sample (Supplementary Table 5 and 6). We found no explicit relationship between
289 diversity of miRNA population and either dpi or infection status. Of the 153 total miRNAs
290 identified across all samples, 83 were present in at least one replicate per treatment (Figure 3).
291 PCA of normalized read counts shows a correlation between infection status and grouping
292 along the PC1/PC2 axis at each time point (Figure 4).
293
294 Figure 3: The top histogram represents the number of miRNAs shared between treatments
295 (intersection size), and each row below the histogram represents a treatment. The lines
296 connecting treatments below the top histogram represent treatments which share that number
297 of miRNAs, and the histogram to the side of the treatments represents the number of miRNAs
298 contained within each treatment.
299
300 Figure 4: Principal Components Analysis (PCA) on read counts mapping to miRNAs identified in
301 the AsteI2 build of the *Anopheles stephensi* genome in Vectorbase. A. - C. are the 2 dpi, 7 dpi,
302 and 14 dpi groupings respectively. In all PCAs, blue is Mayaro infected, and red is control.
303
304 <u>miRNA Differential Expression</u>
305 We next identified known and novel miRNAs that were differentially expressed by
306 infection status (Fig. 5; Table 1, Supplementary Table 7). Contrasts considered in the GLM were
307 infected against control at 2, 7, and 14 dpi, and differences between 2 -7 dpi and 7 – 14 dpi for
308 infected treatments relative to control. miRNAs were considered differentially expressed by

309 having a log fold-change (log2FC) value of +/- 1 and P value < 0.05. There were a total of 8

310 miRNAs differentially regulated in any considered contrast, novel miRNAs as-mirNOV10, as-

311 mirNOV16, and as-mirNOV17 as well as known miRNAs aga-miR-286b, aga-miR-2944a, aga-miR-

312 2944b, aga-miR-307, and aga-miR-309. as-mirNOV10 was enriched at 2 dpi, as-mirNOV16 was

313 enriched at 7 dpi, and as-mirNOV 17 was depleted at 14 dpi and between 7 – 14 dpi. The known

314 miRNAs were depleted as a group at 7 dpi and in the 2 - 7 dpi contrast but enriched in the 7 –

315 14 dpi contrast.

316

317 Table 1: Differentially expressed *Anopheles stephensi* miRNAs in response to Mayaro virus

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

318 infection.

319

320

321 Figure 5: Volcano plots visualizing differential expression of identified *Anopheles stephensi*
322 miRNAs in response to Mayaro infection. The Y-axis shows -log10 transformed P-values, and the
323 X-axis shows log2 transformed fold change values. Red points represent transcripts depleted by
324 more than -1 log2FC in response to infection with a FDR < 0.05, while blue points are transcripts
325 enriched by more than 1 log2FC in response to infection with a FDR < 0.05. A. - C. are the 2 dpi,
326 7 dpi, and 14 dpi groupings respectively, while D. and E. are miRNAs regulated in the infected
327 treatment between 2 - 7 dpi and 7 - 14 dpi respectively.

328

329 The miR-309/286/2944 has been found to be enriched in *An. gambiae* in response to
330 blood feeding [39 – 40], and to be associated with Argonaute proteins post-bloodmeal [39].
331 When experimentally repressed aga- miR-309 was found to retard oocyte development [39],
332 depletion in response to MAYV infection may suggest that viral replication or the host immune
333 response sequesters resources normally requires for host oocyte development, and as a result
334 associated miRNAs are also depleted. Experimental infections have demonstrated that for *Ae.*
335 *aegypti*, infection with CHIKV does not impact the number of eggs laid but it does have a
336 detrimental impact on the viability of the eggs produced, however this is not observed with
337 ZIKV [41, 42]. Infection of *Culex tarsalis* with West Nile Virus demonstrates a decrease in
338 fecundity as measured by egg raft size and number of eggs laid [43].

339 <u>miRNA Target Prediction</u>

340 We next identified putative targets in the *An. stephensi* genome for all known and novel
341 miRNAs identified in all samples [32]. For each of the identified miRNAs we found an average

342 537 potential annotated targets within the Astel2 genome (Supplementary Table 8). Targets for
343 significantly regulated miRNAs were loaded into g:Profiler and any overrepresented GO terms
344 with an FDR < 0.5 were considered significant (Supplementary Table 9) [33].

345 No overrepresented GO terms were identified by the genomic targets of as-mirNOV-10
346 or aga-miR-307. as-mirNOV-16 was significantly enriched in response to infection at 7 dpi and
347 the only GO terms overrepresented by the predicted genomic targets of this miRNA are
348 associated with protein binding. as-mirNOV17 was depleted at 14 dpi and between 7 – 14 dpi
349 and has GO terms related to transmembrane ion channels overrepresented by its genomic
350 targets. aga-mir-2944a and aga-mir-2944b were both depleted at 7 dpi and between 2 – 7 dpi
351 but enriched between 7 – 14 dpi and both have GO terms primarily associated with intracellular
352 signaling and various binding functions, and aga-mir-2944b also appears to be involved with
353 lipid localization and transport. aga-mir-286b and aga-miR-309 were both also depleted at 7 dpi
354 and between 2 and 7 dpi, and were enriched between 7 and 14 dpi; aga-mir-286b only had
355 acetylglucosaminyltransferase activity overrepresented by its genomic targets, and aga-miR-
356 309 had terms related to calcium ion transport, actin filament binding, and catalytic activity
357 overrepresented by its genomic targets.

358 The novel miRNA as-mirNOV-17 has 498 predicted genomic targets, and those targets
359 overlap with 8 enriched and 3 depleted genes at 14 dpi and 6 enriched and 2 depleted genes
360 between 7 and 14 dpi, when as-mirNOV-17 was significantly repressed in response to MAYV
361 infection (Table 1). Neither as-mirNOV-10 or as-mirNOV-16 showed a bias for enriched or
362 depleted genomic targets in the contrasts they were significantly regulated in. The known
363 miRNAs also showed a bias for enriched targets between 2-7 dpi and 7-14 dpi where they are

364 repressed and activated in each contrast respectively. These patterns are consistent with the
365 miRNAs acting as effector molecules for RNAi, except for the known miRNAs between 7 - 14 dpi
366 where their expression is enhanced but they still have a bias for enriched genomic targets [38].
367 Recent studies have demonstrated that through targeting of promotor elements miRNAs can
368 have a positive impact on gene transcription, so this could explain the phenomenon happening
369 between 7 – 14 dpi where miRNA targets are enriched when the miRNAs themselves are also
370 enriched [45].

371 piRNA Identification

372 We identified putative viral piRNAs in the small RNA datasets by tracking the 24-35 nt
373 reads, removing those that were identified as miRNAs, and mapping these reads to the Mayaro
374 Virus NC_003417.1 genome using the Bowtie sequence aligner, all of which was performed
375 within the MSRG pipeline [26 – 27, 34, 52]. There was viral piRNA expression in infected
376 samples with a bias for the positive strand over the negative strand, and the proportion of
377 potential piRNAs mapping to the viral genome remained consistent across time points with no
378 particular peaks or hotspots identified across the viral genome (Figure 6) [53].

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

386

387 Virus-derived piRNA-like small RNAs have been identified in insects and insect cells
388 infected with Flaviviruses, Bunyaviruses and Alphaviruses. Knockdown of the piRNA pathway
389 proteins leads to enhanced replication of arboviruses in mosquito cells, suggesting their
390 potential antiviral properties in mosquitoes [46 – 51]. For example, knockdown of Piwi-4 in *Ae.*
391 *aegypti* Aag2 cell line increased replication of SFV, and silencing of Ago3 and Piwi-5 led to
392 significantly reduced production of piRNAs against Sindbis virus (SINV) [47, 50].

393 <u>siRNA Identification</u>

394 We also identified putative viral siRNAs in the small RNA datasetsby tracking the 18-24
395 nt reads that were not miRNAs and could map to the Mayaro Virus NC_003417.1 genome [26,
396 27, 34, 46]. There was minimal alignment of potential siRNAs to the MAYV genome in the
397 control samples, but in the infected samples there was significant siRNA alignment across the
398 genome (Figure 6). There was a slight negative-strand bias of MAYV siRNAs (Day2 [103+, 124],
399 Day7 [287+, 319-], Day14 [667+, 762-], and the steady increase of siRNAs over time suggests
400 ongoing amplification of the innate immune response with time during the infection process.
401 There were two peaks for siRNA alignment, one at nonstructural protein 1 and another at
402 envelope protein 1 and the number of siRNA reads mapping to the viral genome increased from
403 2 to 14 dpi.

404 The siRNA pathway is thought to be the main antiviral component of immunity in insects
405 at the cellular level [54]. Functional studies in *Aedes* mosquitoes have implicated the siRNA
406 pathway as integral to the antiviral response in the midgut stage of infection as well as at later
407 systemic stages of infection for DENV and SINV [54]. Studies of the siRNA response in *An.*

408 *gambiae* have shown that ONNV does not stimulate a siRNA response at early midgut stages of
409 infection but is present at later stages of infection in the systemic compartments [54]. Our
410 results differ from the findings from *An. gambiae* and ONNV in that the siRNA response is
411 detectable at early stages of infection and is persistent until later stages.

412 *Conclusion*

413 The transcriptomic profiles suggest that MAYV activates the Toll pathway at early and
414 mid-stages of infection as an innate humoral response from the host to fight infection. At later
415 stages of infection there appears to be a repression of JNK and MAPK signaling cascades,
416 potentially impacting autophagic and apoptotic processes as a way to limit MAYV replication.
417 The small RNA profiles suggests a reliance on siRNA silencing as an antiviral immune response.
418 There was a minimal piRNA response to infection with a positive strand bias, but it did not
419 appear to play a major role in antiviral immunity in this study. miRNAs were also elicited in
420 response to infection and some overlap was observed with transcripts identified as regulated in
421 response to infection, but not to the extent that they appear to be strongly regulating
422 transcriptional profiles in response to infection.

423

424 **Acknowledgments.**

425 We thank the UTMB sequencing core for assistance with next generation sequencing.
426 This work was supported by NIH grants R01AI150251, R01AI128201, R01AI116636, USDA Hatch
427 funds (Project #PEN04769), and a grant with the Pennsylvania Department of Health using
428 Tobacco Settlement Funds to JLR, NIH grant R01GM135215 to NCL, BBSRC awards
429 BB/T001240/1 and BB/V011278/1, Royal Society Wolfson Fellowship RSWF\R1\180013, NIH

430 grants R21AI138074, URKI grants 20197 and 85336, EPSRC grant V043911/1, and NIHR grant
431 NIHR2000907 to GLH. CAH was supported by an NSF graduate research fellowship program
432 award (ID 2018258101). SH was supported the Liverpool School of Tropical Medicine Director's
433 Catalyst Fund award. GLH is affiliated to the National Institute for Health Research Health
434 Protection Research Unit (NIHR HPRU) in Emerging and Zoonotic Infections at University of
435 Liverpool in partnership with Public Health England (PHE), in collaboration with Liverpool
436 School of Tropical Medicine and the University of Oxford. GLH is based at LSTM. The views
437 expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, the
438 Department of Health or Public Health England. We would also like to thank Dr Martin Donnelly
439 at Liverpool School of Tropical Medicine for comments on an early version of this manuscript
440 which helped us to identify an error in the analysis.

441

442 **References**

443 1. Anderson CR, Wattley GH, Ahin NW, Downs WG, Reese AA. Mayaro Virus: A New Human
444 Disease Agent. *Am J Trop Med Hyg*. 1957;6(6):1012-1016.
445 doi:10.4269/ajtmh.1957.6.1012
446 2. Esposito DLA, Fonseca BAL da. Will Mayaro virus be responsible for the next outbreak of
447 an arthropod-borne virus in Brazil? *Brazilian J Infect Dis*. 2017;21(5):540-544.
448 doi:10.1016/j.bjid.2017.06.002
449 3. Hassing R-J, Leparc-Goffart I, Blank SN, et al. Imported Mayaro virus infection in the
450 Netherlands. *J Infect*. 2010;61(4):343-345. doi:10.1016/j.jinf.2010.06.009
451 4. Llagonne-Barets M, Icard V, Leparc-Goffart I, et al. A case of Mayaro virus infection

452 imported from French Guiana. *J Clin Virol*. 2016;77:66-68. doi:10.1016/j.jcv.2016.02.013

453 5. Receveur MC, Grandadam M, Pistone T, Malvy D. Infection with Mayaro virus in a French
454 traveller returning from the Amazon region, Brazil, January, 2010. *Euro Surveill*.
455 2010;15(18). http://www.ncbi.nlm.nih.gov/pubmed/20460093. Accessed February 24,
456 2019.

457 6. Neumayr A, Gabriel M, Fritz J, et al. Mayaro virus infection in traveler returning from
458 Amazon Basin, northern Peru. *Emerg Infect Dis*. 2012;18(4):695-696.
459 doi:10.3201/eid1804.111717

460 7. Mota MT de O, Ribeiro MR, Vedovello D, Nogueira ML. Mayaro virus: a neglected
461 arbovirus of the Americas. *Future Virol*. 2015;10(9):1109-1122. doi:10.2217/fvl.15.76

462 8. Abad-Franch F, Grimmer GH, de Paula VS, Figueiredo LTM, Braga WSM, Luz SLB. Mayaro
463 Virus Infection in Amazonia: A Multimodel Inference Approach to Risk Factor
464 Assessment. Weaver SC, ed. *PLoS Negl Trop Dis*. 2012;6(10):e1846.
465 doi:10.1371/journal.pntd.0001846

466 9. Brustolin M, Pujhari S, Henderson CA, Rasgon JL. Anopheles mosquitoes may drive
467 invasion and transmission of Mayaro virus across geographically diverse regions.
468 Christofferson RC, ed. *PLoS Negl Trop Dis*. 2018;12(11):e0006895.
469 doi:10.1371/journal.pntd.0006895

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

473 11. Smith GC, Francy DB. Laboratory studies of a Brazilian strain of Aedes albopictus as a

474 potential vector of Mayaro and Oropouche viruses. *J Am Mosq Control Assoc*.
475 1991;7(1):89-93. http://www.ncbi.nlm.nih.gov/pubmed/1646286. Accessed February 24,
476 2019.
477 12. Tesh RB, Higgs S, Hausser NL, et al. Experimental Transmission of Mayaro Virus by Aedes
478 aegypti. *Am J Trop Med Hyg*. 2011;85(4):750-757. doi:10.4269/ajtmh.2011.11-0359
479 13. Rezza G, Chen R, Weaver SC. O'nyong-nyong fever: a neglected mosquito-borne viral
480 disease. *Pathog Glob Health*. 2017;111(6):271-275. doi:10.1080/20477724.2017.1355431
481 14. Yadav P, Barde P, Singh D, Mishra A, Mourya D. EXPERIMENTAL TRANSMISSION OF
482 CHIKUNGUNYA VIRUS BY ANOPHELES STEPHENSI MOSQUITOES. *Acta Virol*. 2003;47:45-
483 47.
484 http://www.elis.sk/download_file.php?product_id=8&session_id=jc2533t2t4ld65j13rldd
485 b4r55. Accessed February 24, 2019.
486 15. Franz A, Kantor A, Passarelli A, Clem R. Tissue Barriers to Arbovirus Infection in
487 Mosquitoes. *Viruses*. 2015;7(7):3741-3767. doi:10.3390/v7072795
488 16. Bonizzoni M, Dunn WA, Campbell CL, Olson KE, Marinotti O, James AA. Complex
489 Modulation of the Aedes aegypti Transcriptome in Response to Dengue Virus Infection.
490 Moreira LA, ed. *PLoS One*. 2012;7(11):e50512. doi:10.1371/journal.pone.0050512
491 17. Etebari K, Hegde S, Saldana MA, et al. Global transcriptome analysis of Aedes aegypti
492 mosquitoes in response to Zika virus infection. *bioRxiv*. August 2017:179416.
493 doi:10.1101/179416
494 18. Waldock J, Olson KE, Christophides GK. Anopheles gambiae Antiviral Immune Response
495 to Systemic O'nyong-nyong Infection. Traub-Csekö YM, ed. *PLoS Negl Trop Dis*.

496 2012;6(3):e1565. doi:10.1371/journal.pntd.0001565

497 19. Sim C, Hong YS, Vanlandingham DL, et al. Modulation of Anopheles gambiae gene
498 expression in response to o'nyong-nyong virus infection. *Insect Mol Biol*. 2005;14(5):475-
499 481. doi:10.1111/j.1365-2583.2005.00578.x

500 20. Palatini U, Miesen P, Carballar-Lejarazu R, et al. Comparative genomics shows that viral
501 integrations are abundant and express piRNAs in the arboviral vectors Aedes aegypti and
502 Aedes albopictus. *BMC Genomics*. 2017. doi:10.1186/s12864-017-3903-3

503 21. Saldaña MA, Etebari K, Hart CE, et al. Zika virus alters the microRNA expression profile
504 and elicits an RNAi response in Aedes aegypti mosquitoes. Armstrong PM, ed. *PLoS Negl*
505 *Trop Dis*. 2017;11(7):e0005760. doi:10.1371/journal.pntd.0005760

506 22. Varjak M, Donald CL, Mottram TJ, et al. Characterization of the Zika virus induced small
507 RNA response in Aedes aegypti cells. Olson KE, ed. *PLoS Negl Trop Dis*.
508 2017;11(10):e0006010. doi:10.1371/journal.pntd.0006010

509 23. Carissimo G, Pondeville E, McFarlane M, et al. Antiviral immunity of Anopheles gambiae
510 is highly compartmentalized, with distinct roles for RNA interference and gut microbiota.
511 *Proc Natl Acad Sci U S A*. 2015;112(2):E176-85. doi:10.1073/pnas.1412984112

512 24. Sinka ME, Rubio-Palis Y, Manguin S, et al. The dominant Anopheles vectors of human
513 malaria in the Americas: occurrence data, distribution maps and bionomic précis. *Parasit*
514 *Vectors*. 2010;3(1):72. doi:10.1186/1756-3305-3-72

515 25. Hay SI, Sinka ME, Okara RM, et al. Developing Global Maps of the Dominant Anopheles
516 Vectors of Human Malaria. *PLoS Med*. 2010;7(2):e1000209.
517 doi:10.1371/journal.pmed.1000209

518 26. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence
519 data. *Bioinformatics*. 2014;30(15):2114-2120. doi:10.1093/bioinformatics/btu170
520 27. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner.
521 *Bioinformatics*. 2013;29(1):15-21. doi:10.1093/bioinformatics/bts635
522 28. Liao Y, Smyth GK, Shi W. The R package Rsubread is easier, faster, cheaper and better for
523 alignment and quantification of RNA sequencing reads. *Nucleic Acids Res*.
524 2019;47(8):e47-e47. doi:10.1093/nar/gkz114
525 29. Chen Y, Mccarthy D, Ritchie M, Robinson M, Smyth GK. *EdgeR User's Guide*.; 2008.
526 https://www.bioconductor.org/packages/release/bioc/vignettes/edgeR/inst/doc/edgeR
527 UsersGuide.pdf. Accessed September 24, 2018.
528 30. Uri Reimand J¨, Arak T, Adler P, et al. g:Profiler-a web server for functional interpretation
529 of gene lists (2016 update). *Nucleic Acids Res*. 2016;(1). doi:10.1093/nar/gkw199
530 31. GitHub - rajewsky-lab/mirdeep2: Discovering known and novel miRNAs from small RNA
531 sequencing data. https://github.com/rajewsky-lab/mirdeep2. Accessed October 8, 2019.
532 32. Betel D, Wilson M, Gabow A, Marks DS, Sander C. The microRNA.org resource: targets
533 and expression. *Nucleic Acids Res*. 2008;36(Database issue):D149-53.
534 doi:10.1093/nar/gkm995
535 33. Brackney DE. Implications of autophagy on arbovirus infection of mosquitoes. *Curr Opin*
536 *Insect Sci*. 2017;22:1-6. doi:10.1016/j.cois.2017.05.001
537 34. Ma Q, Srivastav SP, Gamez S, et al. A mosquito small RNA genomics resource reveals
538 dynamic evolution and host responses to viruses and transposons. *Genome Res*.
539 2021;31(3):gr.265157.120. doi:10.1101/gr.265157.120

540 35. Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data
541 [Online]. Available online
542 at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
543 36. Chakraborty M, Ramaiah A, Adolfi A, et al. Hidden genomic features of an invasive
544 malaria vector, Anopheles stephensi, revealed by a chromosome-level genome assembly.
545 *BMC Biol*. 2021;19(1):1-16. doi:10.1186/s12915-021-00963-z
546 37. Carissimo G, Pain A, Belda E, Vernick KD. Highly focused transcriptional response of
547 Anopheles coluzzii to O'nyong nyong arbovirus during the primary midgut infection. *BMC*
548 *Genomics*. 2018;19(1):526. doi:10.1186/s12864-018-4918-0
549 38. Samuel GH, Adelman ZN, Myles KM. Antiviral Immunity and Virus-Mediated Antagonism
550 in Disease Vector Mosquitoes. 2017. doi:10.1016/j.tim.2017.12.005
551 39. Fu X, Dimopoulos G, Zhu J. Association of microRNAs with Argonaute proteins in the
552 malaria mosquito Anopheles gambiae after blood ingestion. *Sci Rep*. 2017;7(1).
553 doi:10.1038/s41598-017-07013-1
554 40. Biryukova I, Ye T, Levashina E. Transcriptome-wide analysis of microRNA expression in
555 the malaria mosquito Anopheles gambiae. *BMC Genomics*. 2014;15(1):557.
556 doi:10.1186/1471-2164-15-557
557 41. Resck MEB, Padilha KP, Cupolillo AP, et al. Unlike Zika, Chikungunya virus interferes in the
558 viability of Aedes aegypti eggs, regardless of females' age. *Sci Rep*. 2020;10(1):1-9.
559 doi:10.1038/s41598-020-70367-6
560 42. da Silveira ID, Petersen MT, Sylvestre G, et al. Zika Virus Infection Produces a Reduction
561 on Aedes aegypti Lifespan but No Effects on Mosquito Fecundity and Oviposition

562 Success. *Front Microbiol*. 2018;9:3011. doi:10.3389/fmicb.2018.03011

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

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

567 45. Xiao M, Li J, Li W, et al. MicroRNAs activate gene transcription epigenetically as an
568 enhancer trigger. *RNA Biol*. 2017;14(10):1326-1334.
569 777doi:10.1080/15476286.2015.1112487

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

573 47. Schnettler E, Donald CL, Human S, et al. Knockdown of piRNA pathway proteins results in
574 enhanced semliki forest virus production in mosquito cells. *J Gen Virol*.
575 2013;94(PART7):1680-1689. doi:10.1099/vir.0.053850-0

576 48. Morazzani EM, Wiley MR, Murreddu MG, Adelman ZN, Myles KM. Production of Virus-
577 Derived Ping-Pong-Dependent piRNA-like Small RNAs in the Mosquito Soma. Ding S-W,
578 ed. *PLoS Pathog*. 2012;8(1):e1002470. doi:10.1371/journal.ppat.1002470

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

582 50. Miesen P, Girardi E, Van Rij RP. Distinct sets of PIWI proteins produce arbovirus and
583 transposon-derived piRNAs in Aedes aegypti mosquito cells. *Nucleic Acids Res*.

584 2015;43(13):6545-6556. doi:10.1093/nar/gkv590

585 51. Schnettler E, Ratinier M, Watson M, et al. RNA interference targets arbovirus replication

586 in Culicoides cells. *J Virol*. 2013;87(5):2441-2454. doi:10.1128/JVI.02848-12\

587 52. Espósito DLA, da Fonseca BAL. Complete genome sequence of Mayaro virus (Togaviridae,

588 Alphavirus) strain BeAr 20290 from Brazil. *Genome Announc*. 2015;3(6).

589 doi:10.1128/genomeA.01372-15

590 53. Lavergne A, Thoisy B de, Lacoste V, et al. Mayaro virus: Complete nucleotide sequence

591 and phylogenetic relationships with other alphaviruses. *Virus Res*. 2006;117(2):283-290.

592 doi:10.1016/j.virusres.2005.11.006

593 54. Carissimo G, Pondeville E, McFarlane M, et al. Antiviral immunity of Anopheles gambiae

594 is highly compartmentalized, with distinct roles for RNA interference and gut microbiota.

595 *Proc Natl Acad Sci U S A*. 2015;112(2):E176-E185. doi:10.1073/pnas.1412984112

596

597

598

599

600

601

602

603

604

605

606

607 Supplementary Table 1: Information related to infection of *Anopheles stephensi* with Mayaro

608 virus. Includes number of mosquitoes in each treatment and time point and associated

609 mortality, nanodrop readings for all RNA extractions collected, pooling scheme for sequencing

610 of mRNA and small RNA, and qPCR data from each sample using primers specific for Mayaro

611 virus strain BeAn to confirm infection status.

612

613 Supplementary Table 2: Differentially expressed transcripts from the *Anopheles stephensi*

614 AsteI2 genome.

615

616 Supplementary Table 3: Differential expression results from the *Anopheles stephensi*

617 PRJNA629843 genome.

618

619 Supplementary Table 4: GO term overrepresentation for differentially regulated transcripts.

620

621 Supplementary Table 5: Read counts mapping to the identified *Anopheles stephensi* miRNAs in

622 each small RNA sample sequenced, both raw counts and counts normalized to library size

623 (reads per million, RPM) are presented.

624

625 Supplementary Table 6: All known and novel mature and precursor miRNA sequences

626 identified.

627

628 Supplementary Table 7: Differential expression of *Anopheles stephensi* miRNAs.

629

630 Supplementary Table 8: Genomic targets from the *Anopheles stephensi* AsteI2 genome for all

631 identified miRNAs.

632

633 Supplementary Table 9: Overrepresented GO terms represented by targets of significantly

634 regulated miRNAs.

635

636

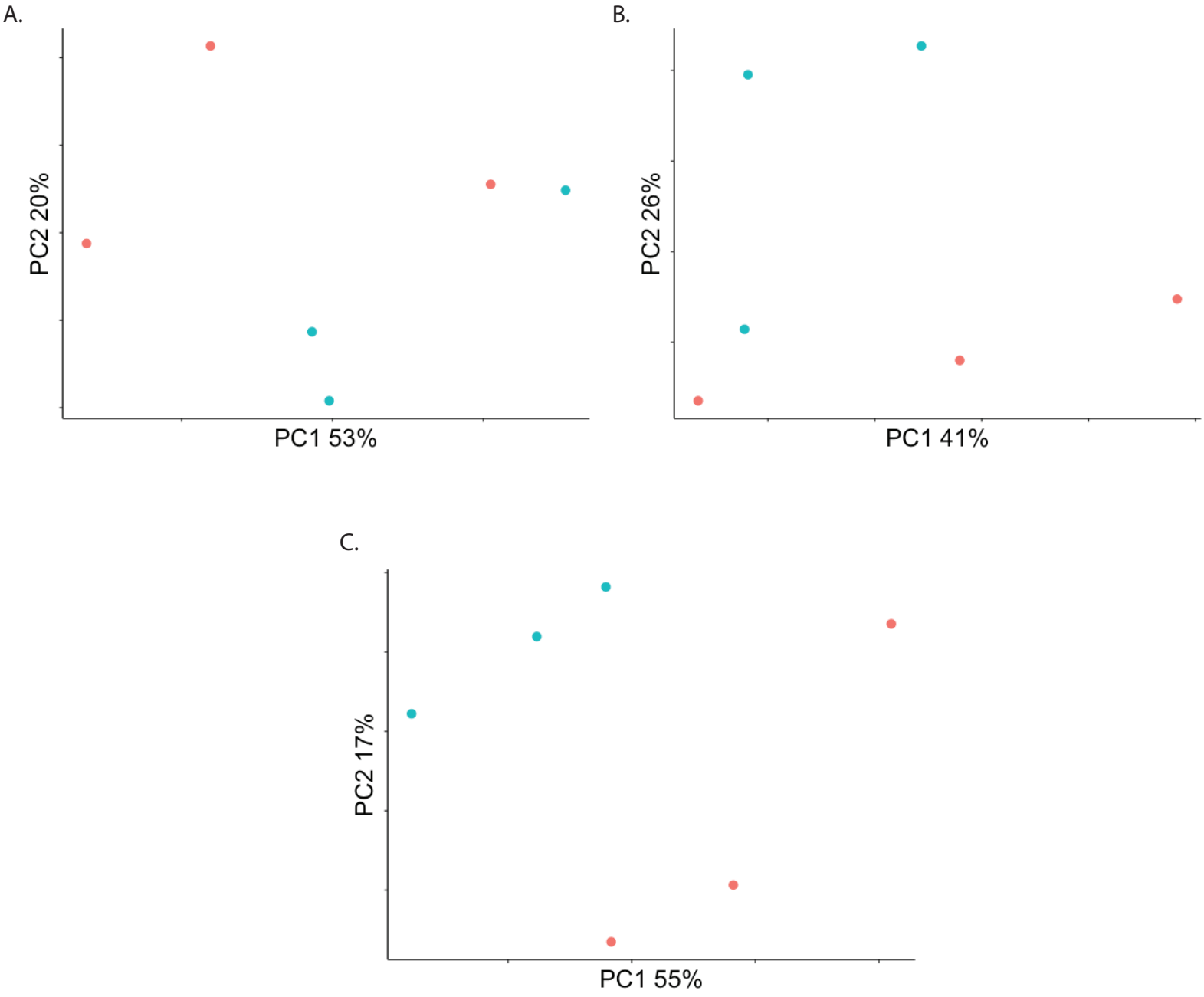

Figure 1: Principal Components Analysis (PCA) on filtered read counts mapping to annotated genes from the AsteI2 build of the *Anopheles stephensi* genome in Vectorbase. A., B., and C. are read counts from samples in the 2, 7, and 14 dpi groupings respectivey. In all PCAs, blue are Mayaro infected, and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

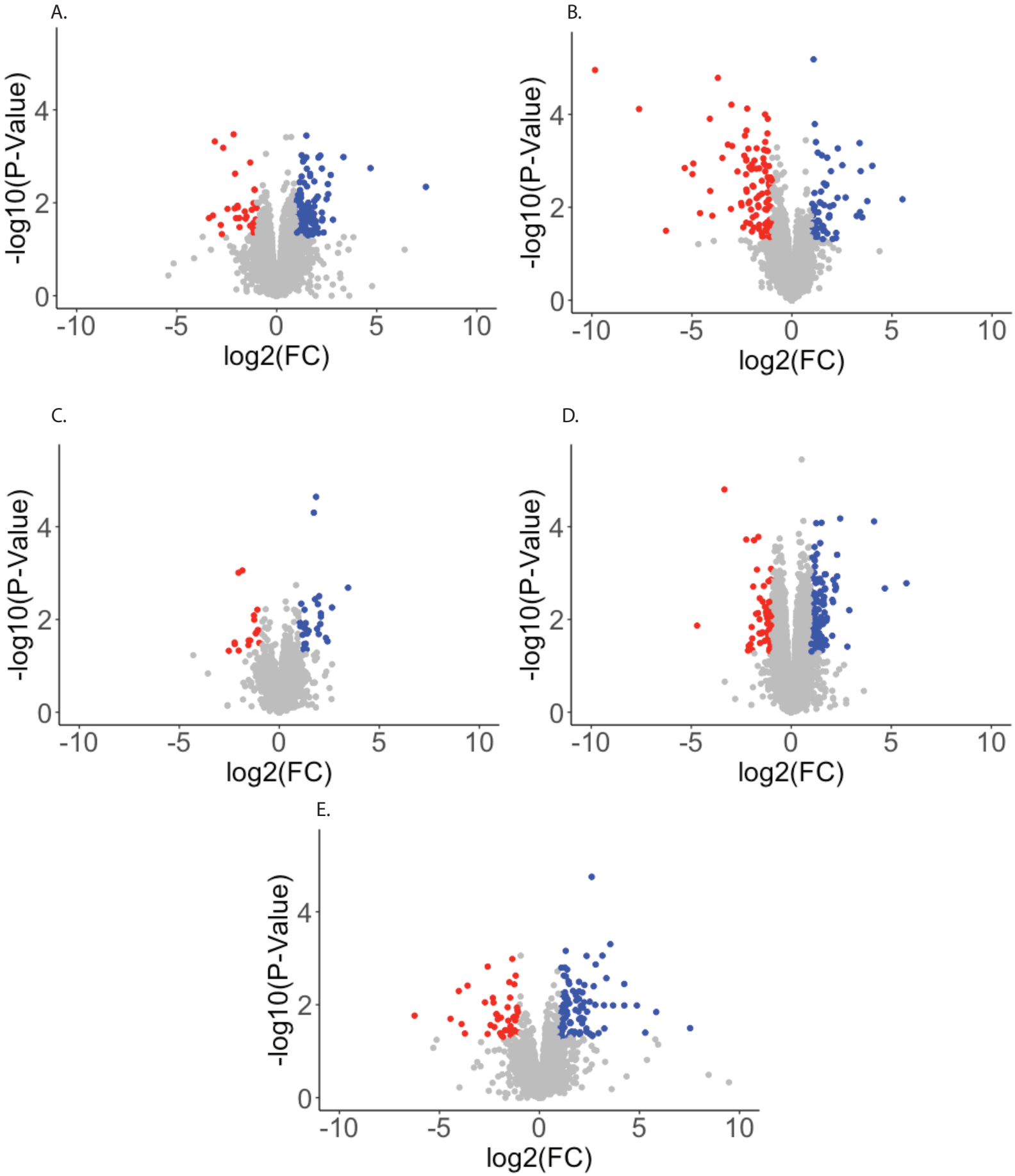

Figure 2: Volcano plots visualizing differential expression of Anopheles stephensi transcripts in response to Mayaro infection. The Y-axis shows -log10 transformed P-values, and the X-axis shows log2 transformed fold change values. Red points represent transcripts downregulated by more than -1 logFC 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.

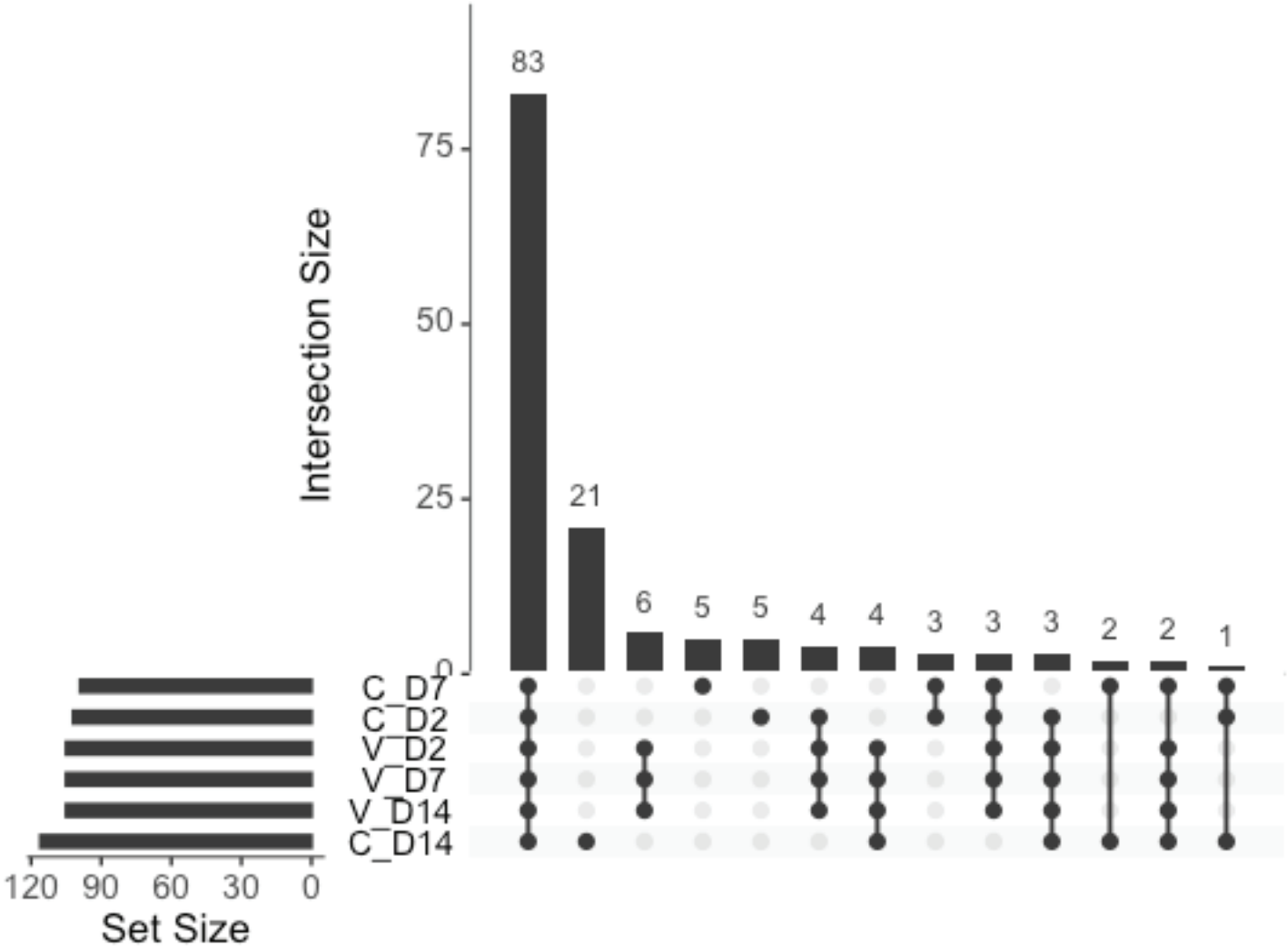

Figure 3: The top histogram represents the number of miRNAs shared between treatments (intersection size), and each row below the histogram represents a treatment. The lines connecting treatments below the top histogram represent treatments which share that number of miRNAs, and the histogram to the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

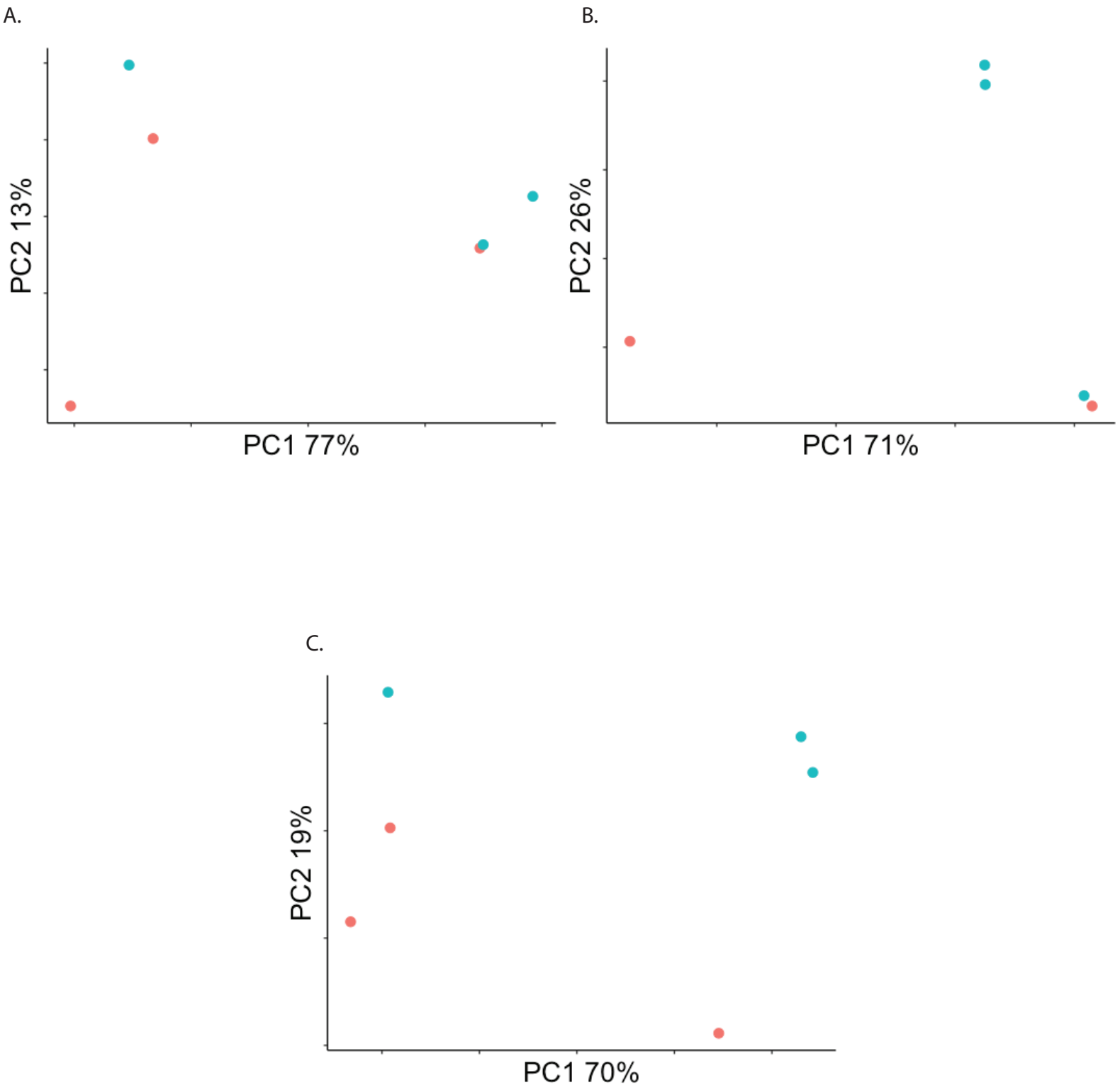

Figure 4: Principal Components Analysis (PCA) on read counts mapping to miRNAs identified in the AsteI2 build of the *Anopheles stephensi* genome in Vectorbase. A. - C. are the 2 dpi, 7 dpi, and 14 dpi groupings respectively. In all PCAs, blue is Mayaro infected, and red is control.

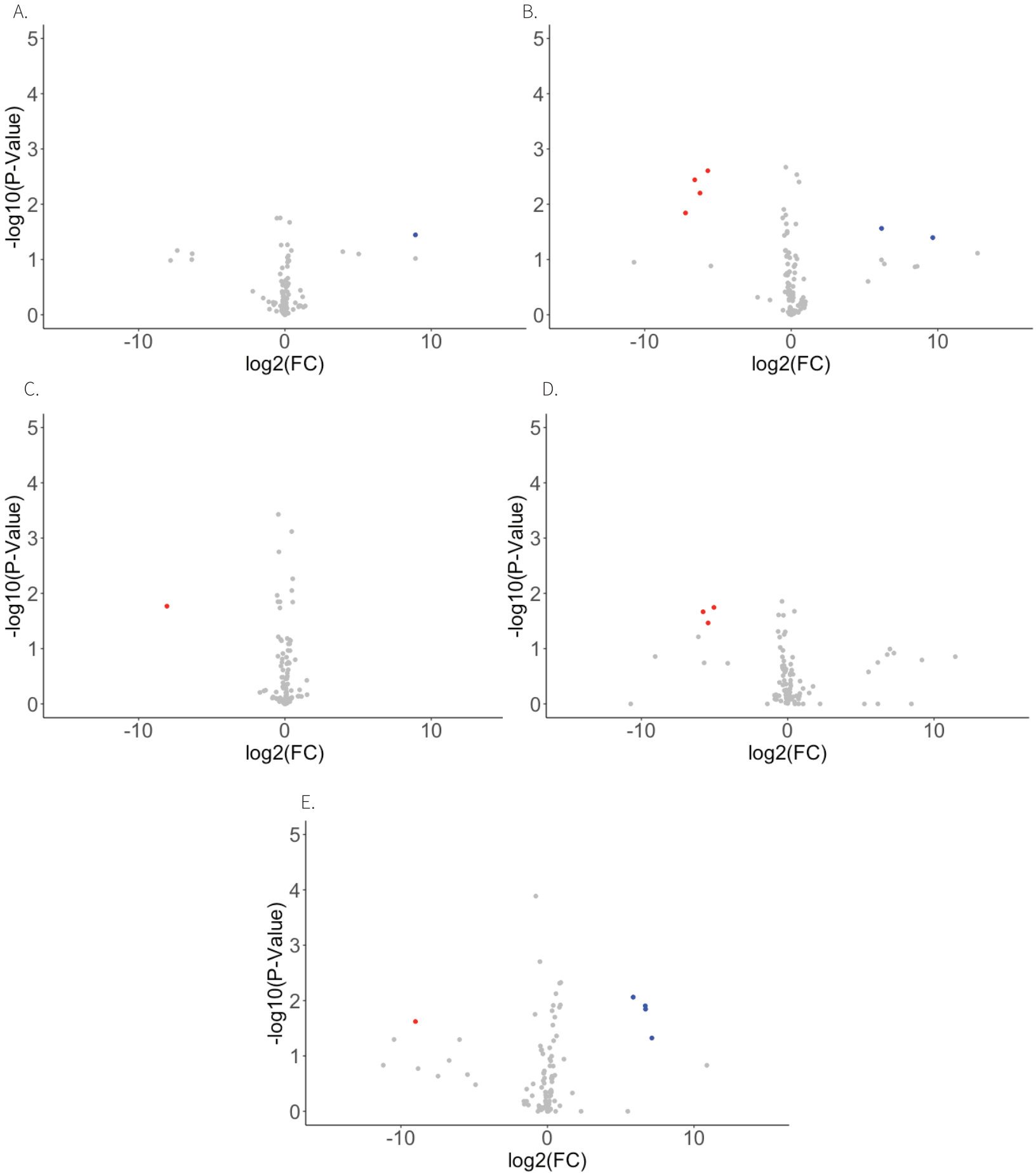

Figure 5: Volcano plots visualizing differential expression of identified *Anopheles stephensi* miRNAs in response to Mayaro infection. The Y-axis shows -log10 transformed P-values, and the X-axis shows log2 transformed fold change values. Red points represent transcripts downregulated by more than -1 logFC in response to infection with a P Value < 0.05, while blue points are transcripts upregulated by more than1logFC 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.

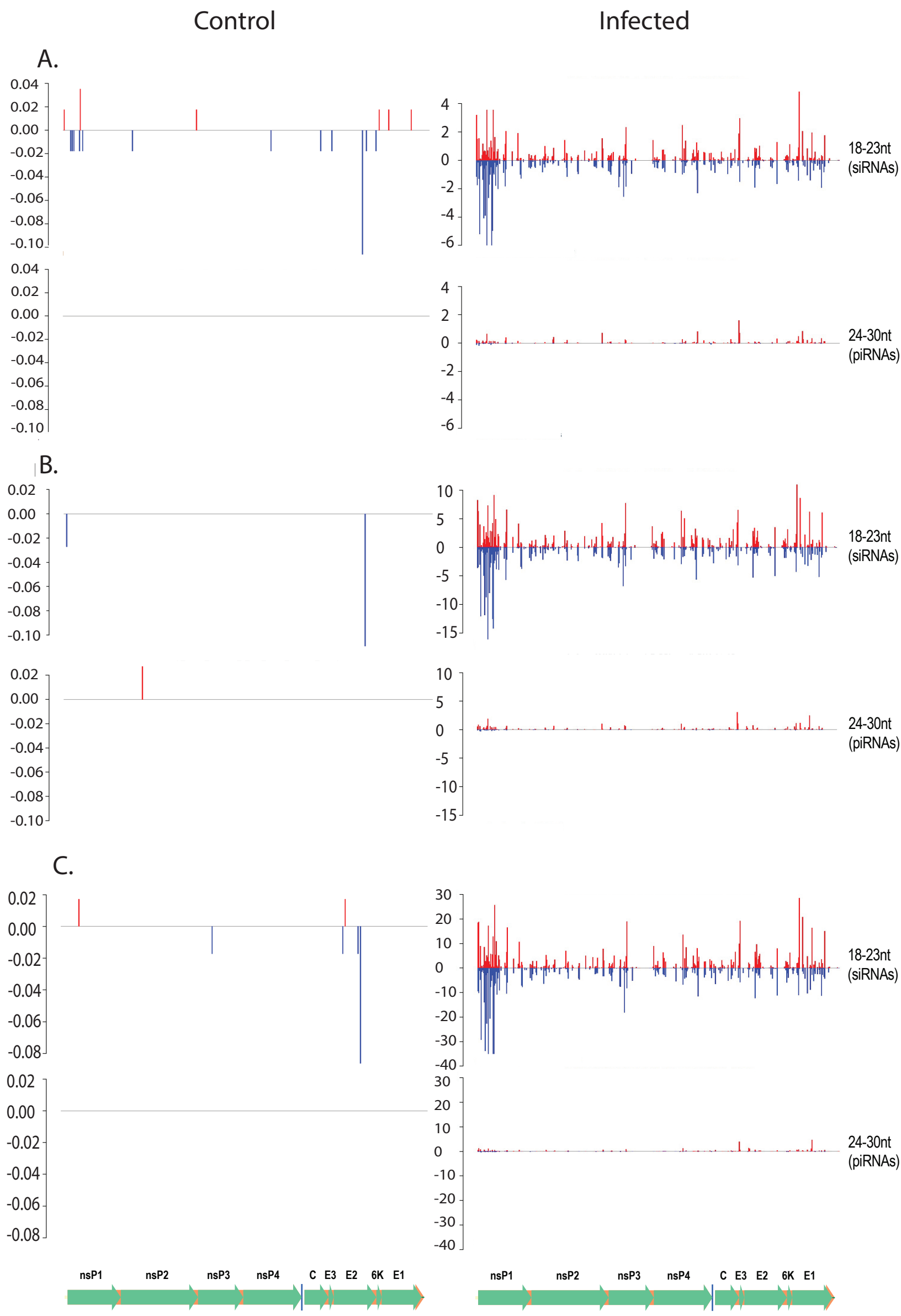

Figure 6: Histograms demonstrating read depth across the Mayaro virus genome for reads with a piRNA size profile (24 - 35 nt) and a siRNA size profile (18 – 23). Y-axis is read depth in reads per million, and X-axis is position in viral genome. Blue demonstrates reads for that sample mapping to the negative strand, while red demonstrates those mapping 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.