



HAL
open science

Global genomics of *Aedes aegypti* unveils widespread and novel infectious viruses capable of triggering a small RNA response

Shruti Gupta, Rohit Sharma, Adeline E Williams, Irma Sanchez-Vargas, Noah H Rose, Chao Zhang, Alexander Crosbie-Villaseca, Zheng Zhu, Gargi Dayama, Andrea Gloria-Soria, et al.

► To cite this version:

Shruti Gupta, Rohit Sharma, Adeline E Williams, Irma Sanchez-Vargas, Noah H Rose, et al.. Global genomics of *Aedes aegypti* unveils widespread and novel infectious viruses capable of triggering a small RNA response. 2024. pasteur-04709182

HAL Id: pasteur-04709182

<https://pasteur.hal.science/pasteur-04709182v1>

Preprint submitted on 25 Sep 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

1 **Global genomics of *Aedes aegypti* unveils widespread and novel infectious** 2 **viruses capable of triggering a small RNA response**

3 **Authors:**

4 Shruti Gupta^{1*}, Rohit Sharma^{1*}, Adeline E. Williams^{2,3}, Irma Sanchez-Vargas³, Noah H. Rose^{4,5},
5 Chao Zhang⁶, Alexander Crosbie-Villaseca¹, Zheng Zhu¹, Gargi Dayama¹, Andrea Gloria-Soria⁷,
6 Doug E. Brackney⁷, Jessica Manning², Sarah S. Wheeler⁸, Angela Caranci⁹, Trinidad Reyes¹⁰,
7 Massamba Sylla¹¹, Athanase Badolo¹², Jewelna Akorli¹³, Ogechukwu B. Aribodor¹⁴, Diego
8 Ayala¹⁵, Wei-Liang Liu¹⁶, Chun-Hong Chen¹⁶, Chalmers Vasquez¹⁷, Cassandra Gonzalez
9 Acosta¹⁸, Alongkot Pontlawat¹⁹, Tereza Magalhaes²⁰, Brendan Carter²¹, Dawn Wesson²¹,
10 Darred Surin²², Meg A. Younger²², Andre Luis Costa-da-Silva²³, Matthew DeGennaro²³,
11 Alexander Bergman²⁴, Louis Lambrechts²⁴, Carolyn S. McBride⁴, Ken E. Olson³, Eric Calvo²,
12 and Nelson C. Lau^{1,25†}

13 **Affiliations:**

- 14 1. Department of Biochemistry and Cell Biology, Boston University Chobanian and Avedisian School of Medicine,
15 Boston MA 02118
- 16 2. Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National
17 Institutes of Health, Rockville, MD 20852, USA
- 18 3. Center for Vector-Borne Infectious Diseases, Department of Microbiology, Immunology, and Pathology, Colorado
19 State University, Fort Collins, CO 80523, USA
- 20 4. Department of Ecology & Evolutionary Biology, Princeton University, Princeton, NJ 08544, USA; Princeton
21 Neuroscience Institute, Princeton University, Princeton, NJ 08544, USA
- 22 5. Department of Ecology, Behavior & Evolution, University of California, San Diego
- 23 6. Boston University Chobanian and Avedisian School of Medicine, Department of Medicine Section of
24 Computational Biomedicine
- 25 7. Department of Entomology, The Connecticut Agricultural Experiment Station
- 26 8. Sacramento-Yolo Mosquito & Vector Control District, 8631 Bond Road, Elk Grove CA 95624
- 27 9. Northwest Mosquito & Vector Control District, 1966 Compton Ave. Corona, CA 92881
- 28 10. Madera County Mosquito & Vector Control District, 3105 Airport Drive, Madera, Ca 93637
- 29 11. Laboratory Vector & Parasites, Department of Livestock Sciences and Techniques, Sine Saloum University El
30 Hadji Ibrahima NIASS, Kaffrine Campus, Senegal
- 31 12. Laboratory of Fundamental and Applied Entomology, Université Joseph Ki-Zerbo, Ouagadougou, Burkina Faso.
- 32 13. Department of Parasitology, Noguchi Memorial Institute for Medical Research, University of Ghana, Accra,
33 Ghana
- 34 14. Department of Zoology, Nnamdi Azikiwe University, Awka, Anambra State, Nigeria
- 35 15. UMR MIVEGEC, IRD, CNRS, Univ. Montpellier, 911 avenue Agropolis, BP 64501, 34394 Montpellier, France; Le
36 Centre International de Recherches Médicales de Franceville, BP 769, Franceville, Gabon
- 37 16. National Mosquito-Borne Diseases Control Research Center, National Health Research Institutes, Miaoli County,
38 Taiwan
- 39 17. Miami Dade Mosquito Control 8901 NW 58 Street, Miami, FL 33178, USA
- 40 18. Coordinacion de Enfermedades Transmitidas por Vector y Zoonosis, Servicios de Salud de Morelos, Morelos,
41 Mexico; Centro Nacional de Programas Preventivos y Control de Enfermedades (CENAPRECE). Mexico City,
42 Mexico
- 43 19. Armed Forces Research Institute of Medical Sciences (AFRIMS), Department of Entomology, Bangkok, Thailand
- 44 20. Department of Entomology, Texas A&M University, College Station TX, 77843
- 45 21. Department of Tropical Medicine, Tulane School of Public Health and Tropical Medicine
- 46 22. Department of Biology, Boston University, Boston MA 02215
- 47 23. Department of Biological Sciences & Biomolecular Sciences Institute, Florida International University, Miami, FL
48 33199, USA
- 49 24. Institut Pasteur, Université Paris Cité, CNRS UMR2000, Insect-Virus Interactions Unit, 75015 Paris, France
- 50 25. Genome Science Institute, Boston University Chobanian and Avedisian School of Medicine, Boston MA 02118

51 *equal significant contribution

52 †Corresponding author

53 **Keywords:** mosquito virus, *Aedes aegypti*, RNAi, small RNAs, siRNAs, piRNAs

54 **ABSTRACT**

55 The mosquito *Aedes aegypti* is a prominent vector for arboviruses, but the breadth of mosquito
56 viruses that infects this species is not fully understood. In the broadest global survey to date of
57 over 200 *Ae. aegypti* small RNA samples, we detected viral small interfering RNAs (siRNAs)
58 and Piwi interacting RNAs (piRNAs) arising from mosquito viruses. We confirmed that most
59 academic laboratory colonies of *Ae. aegypti* lack persisting viruses, yet two commercial strains
60 were infected by a novel tombus-like virus. *Ae. aegypti* from North to South American locations
61 were also teeming with multiple insect viruses, with Anphevirus and a bunyavirus displaying
62 geographical boundaries from the viral small RNA patterns. Asian *Ae. aegypti* small RNA
63 patterns indicate infections by similar mosquito viruses from the Americas and reveal the first
64 wild example of dengue virus infection generating viral small RNAs. African *Ae. aegypti* also
65 contained various viral small RNAs including novel viruses only found in these African
66 substrains. Intriguingly, viral long RNA patterns can differ from small RNA patterns, indicative of
67 viral transcripts evading the mosquitoes' RNA interference (RNAi) machinery. To determine
68 whether the viruses we discovered via small RNA sequencing were replicating and
69 transmissible, we infected C6/36 and Aag2 cells with *Ae. aegypti* homogenates. Through blind
70 passaging, we generated cell lines stably infected by these mosquito viruses which then
71 generated abundant viral siRNAs and piRNAs that resemble the native mosquito viral small
72 RNA patterns. This mosquito small RNA genomics approach augments surveillance approaches
73 for emerging infectious diseases.

74 INTRODUCTION

75

76 The yellow fever mosquito, *Aedes aegypti*, has a global reach across most major human
77 locales, and to this day remains a major health scourge as a prominent vector for arthropod-
78 borne viral (arboviral) diseases like dengue fever. Municipal vector control organizations
79 conduct routine molecular surveillance of arboviruses from trapped *Ae. aegypti* mosquitoes but
80 are only able to assay known viruses¹. Recent research efforts applying high-throughput RNA
81 sequencing have now led to an explosion in the mosquito virome lists²⁻⁵. However, the
82 persistence of these novel insect viruses within mosquito populations and their vertical
83 transmission has not been fully examined in detail.

84 Vertical transmission and infection rates of arboviruses in *Ae. aegypti* have been examined
85 for dengue and Zika viruses, among others⁶. Even when infection rates are high, the vertical
86 transmission rates can be low if virus infection does not extend to the gonads⁷⁻¹⁰. Vertical
87 transmission of arboviruses like Zika virus in laboratory *Ae. aegypti* infections is possible, but at
88 low rates^{11,12}. In addition, tracking prevalence of virus infections in mosquitoes in the wild is a
89 complex epidemiological challenge when there are fluctuating levels of virus infection rates even
90 during an outbreak¹³⁻¹⁵.

91 The ability of a virus to persist via vertical transmission in mosquito populations raises the
92 concern of a species jump to humans^{16,17}. Yet, tracking persistent viruses in wild and small
93 pools of mosquitoes can be difficult with rapid surveillance screening because RT-PCR can be
94 too imprecise to distinguish between strong viral RNA expression and low viral loads. Therefore,
95 another molecular measure is needed to provide a better representation of persistent virus
96 infections in mosquito populations.

97 Mosquitoes and mosquito cell cultures naturally mount a robust innate immune response to
98 viruses with the RNA-interference (RNAi) pathway. During virus replication, viral double-
99 stranded RNAs are processed into viral small interfering RNAs (siRNAs, ~18-23 nucleotides
100 long) that are loaded into Argonaute (AGO) proteins¹⁸. A second RNAi pathway that mainly
101 regulates Transposable Elements (TEs) involves the Piwi proteins that bind Piwi-interacting
102 RNAs (piRNAs, ~24-35nt long), which have sequences complementary to TEs¹⁹. Since many
103 TEs are evolutionarily related to viruses, mosquito Piwi proteins may generate viral piRNAs
104 similar to how TE piRNAs are made. The viral siRNAs that are antisense to the viral mRNA will
105 trigger AGO cleavage of viral mRNAs. While viral piRNAs may have antiviral properties in
106 mosquitoes^{20,21}, this response may be complicated by possible virus-to-virus interactions²².

107 When we first developed a Mosquito Small RNA Genomics (MSRG) pipeline²³, we detected
108 widespread insect virus persistence in mosquito cell cultures and in a few laboratory
109 mosquitoes. A recent mosquito small RNA sequencing survey by the Marques lab profiled *Ae.*
110 *aegypti* and *Ae. albopictus* samples from South America, Asia and Africa locales²². This study
111 presented an emerging collection of insect viruses generating a viral small RNA response,
112 potentially suggesting virus persistence in *Aedes* mosquitoes. However, this study then focused
113 on two insect viruses, Humaita-Tubiaca virus (HTV) and the bunyavirus Phasi-Charoen-Like
114 virus (PCLV), and their role in potentially influencing dengue virus (DENV) competence²². The
115 question then remains: How prevalent globally are *Ae. aegypti* mosquito viruses that trigger a
116 small RNA response?

117 Here, we present the most comprehensive global survey of *Ae. aegypti* mosquito small
118 RNAs to date, comprised of new laboratory strains and wild-caught and lab colonies from the
119 Americas, Asia, and Africa continents. Our results significantly expand the catalog of insect
120 viruses capable of propagating in mosquitoes and generating a small RNA response. We

121 demonstrate experimental evidence for the transmissibility of these novel insect viruses to
122 mosquito cell cultures to generate an RNAi response. Because these mosquito viruses can
123 possibly develop tropism for mammalian cells, our small RNA genomics approach is a valuable
124 addition to current molecular surveillance efforts to detect newly emerging pathogens vectored
125 by mosquitoes and other arthropods.

126

127 RESULTS

128 The collection of *Ae. aegypti* mosquitoes was organized through collaborators shipping
129 mosquito samples that were either preserved (frozen or submerged in an RNA-protectant
130 solution) or shipped as live animals for total RNA extraction and small RNA library preparation.
131 Most mosquito RNAs were analyzed from whole mosquitoes to optimize for throughput.
132 Published small RNA libraries from Olmo et al²² and our previous study²³ were integrated into
133 our analysis of viral small RNA profiles. We compiled >280 small and long RNA sequencing
134 datasets of *Ae. aegypti* from a select set of laboratory strains and from locations in North,
135 Central and South America, Asia, and Africa (**Figure 1**). All datasets were analyzed by our
136 MSRG pipeline, but this study primarily focused on Insect Specific Viruses (ISVs) and
137 arboviruses while TEs and piRNA cluster loci analyses will be presented in a future study.

138 To address library quality control (QC), we first only selected libraries for analysis with at
139 least 1 million reads. There was an average depth of ~40M reads across all libraries (see
140 **Supplemental Table S1 and S2** for library sequencing depth statistics). We also tracked an
141 Endogenous Viral Element (EVE) in *Ae. aegypti* named AEFE1²⁴, whose small RNAs are
142 expressed in all mosquito samples we analyzed and are represented by primarily antisense
143 piRNAs seen as blue peaks in coverage plots (**Supplemental Figure S1A**). Lastly, we
144 inspected the read length distribution profiles for each library to track the expected peaks for
145 miRNAs and piRNAs (**Supplemental Figure S2**). A virus needed to show at least 10 small RNA
146 reads per million (RPM) to register on the bubble plots; if no viruses are detected, these QC
147 measures confirmed a small RNA library was trustworthy.

148

149 *Augmenting the MSRG pipeline with improved virus discovery*

150 Mosquitoes' notable somatic piRNAs²⁵ lend to the broad diversity of small RNA sequences
151 in whole animal total RNA. This breadth means that small RNA libraries cannot be queried
152 efficiently or specifically against databases such as the GenBank Virus Reference List (GBVRL)
153 that have become flooded with recent massive virome surveys^{4,5} and coronavirus variant
154 genomes²⁷. In GBVRL version-249, there were ~7.4 million records and 53.7% of these were
155 beta-coronaviruses. To overcome this inefficiency and provide scalability, MSRG utilizes
156 curated lists of viruses and TEs for small RNAs to be mapped against using BowTie (v1)²⁶.

157 To address the limitations of the manually curated virus database in the MSRG, we devised
158 an auxiliary pipeline to first analyze all *Ae. aegypti* small RNA libraries with the VirusDetect
159 program²⁸ (**Fig. 1B**). The VirusDetect program first performs *de novo* assembly of small RNAs
160 into long contigs that can then query GBVRL efficiently and specifically to yield the count and
161 coverage of the contigs against the most updated virus list in GBVRL (**Supp. Fig. S1B,C**). In
162 this study, we updated the GBVRL in VirusDetect and locked it at version-249 from June 2022.
163 The VirusDetect results confirmed the presence of viruses already present in the 2019 MSRG
164 virus database while also revealing new viruses that we then added to our MSRG virus
165 database list (**Supplemental Table S3**).

166 We then set up additional VirusDetect queries of our mosquito small RNA datasets against
167 the Third-Party Assembly (TPA) and Metagenomics Assemble Genomes (MAG) databases that
168 are distinct from GBVRL. This step was needed because when VirusDetect was run on GBVRL
169 alone, a false-positive call was made for a tombus-like virus isolated from a *Hyposignathus*
170 *monstrosus* bat²⁹ due to the low contig coverage (**Supp. Fig. S1D,E**). The additional TPA and
171 MAG runs discovered the Tiger Mosquito Bi-segmented Tombus-Like virus³⁰ (TMBTLV) as the
172 true source of these small RNAs, as well as revealing a new *Aedes partiti*-like virus³⁰. In total,
173 the VirusDetect analysis added 94 new viruses to our MSRG virus list (**Supp. Table S3**).

174

175 *Academic and commercial laboratory strains of Ae. aegypti*

176 Limited insect viruses were found in academic lab strains from these new *Ae. aegypti* small
177 RNA libraries as well as from previously examined lab strain datasets (**Figure 2A**). Our first
178 MSRG pipeline study²³ also revealed very few insect viruses in these strains, but the expanded
179 virus list and low viral reads in this study boosted our confidence in the sparseness of viruses in
180 these strains. Main figure bubble plots show all small RNAs (18-35nt), while breakdowns of the
181 siRNAs (18-23nt) and piRNAs (24-35nt) are shown in **Supplemental Figure S3**.

182 In contrast to these academic lab strains, two independent commercial lab strains from
183 Benzon Research Labs (BZL) and Bayer AG (GP)³¹ contained massive amounts of TMBTLV
184 small RNAs, with viral siRNAs more abundant than viral piRNAs (**Fig. 2B**). Coverage patterns of
185 the TMBTLV small RNAs differed greatly between the two commercial strains. The BZL strain
186 viral small RNAs were biased for the plus strand and mainly derived from the RNA dependent
187 RNA Polymerase (RdRP) gene, while the GP strain displayed a much greater accumulation of
188 minus strand siRNAs across the entire TMBTLV genome and had piRNAs still biased for the
189 plus strand. All GP strain samples were co-infected with Zika virus (ZIKV), but control non-Zika-
190 infected GP strains small RNA data were not available, so ZIKV's contribution to the massive
191 TMBTLV siRNAs cannot be resolved until future experiments can be conducted.

192 Commercial labs often raise *Ae. aegypti* alongside many other arthropods for the testing of
193 pesticides. Benzon Research Labs states that their BZL strain was derived from an USDA
194 "Gainesville" strain from 1994. The continuous rearing in their facility for over 25 years may
195 have lent to this strain contracting additional viruses not seen in academic lab strains (**Fig.**
196 **2C,D**). These viruses must persist by vertical transmission, because the same viral small RNAs
197 in adult BZL females were also detected in BZL eggs (**Fig. 2E**), although the eggs had lower
198 levels and more minus strand viral small RNAs. For some mosquito viruses, the small RNAs
199 originated from all genes in the viral genomes, but for a partiti-like virus and TMBTLV, the RdRP
200 genes served as the primary source of small RNAs.

201

202 *Diverse profiles of viral small RNAs in Ae. aegypti across the two American continents*

203 With the help of municipal vector control departments in California and Miami, Florida, along
204 with collections and lab-maintained colonies from the DeGennaro, Olson, and Lambrechts labs,
205 we assembled a diverse collection of *Ae. aegypti* samples from the American continents to
206 generate and sequence small RNAs from whole mosquitoes. We also integrated previously
207 published *Ae. aegypti* small RNA libraries from Suriname and Brazil²².

208 There was widespread prevalence of mosquito viruses generating robust viral small RNAs in
209 North, Central and South America (**Figure 3A,B**). We detected other mosquito viruses related to
210 plant viruses, similar to TMBTLV, from the two commercial lab strains (**Fig. 2B**). A Liverpool
211 Tombus-like virus (LTLV) was also reported in the same metagenomics study discovering

212 TMBTLV³⁰, whereas Verdadero virus is a partitivirus first described infecting *Ae. aegypti* as well
213 as *Drosophila*³². Partitiviruses were originally discovered to infect plants, protozoans, and fungi,
214 while tombusviruses are a group representing the tomato bushy stunt viruses^{33,34}. Viral siRNAs
215 and piRNAs for LTLV were much higher in the midguts versus the total abdomens of an *Ae.*
216 *aegypti* strain from Recife, Brazil (**Supplemental Figure S4A**), suggesting that mosquitoes
217 feeding on plant sap nutrients could be a route for the transfer of plant-related viruses to an
218 insect host.

219 Even though many of the samples in these regions contained the same viruses, the
220 coverage pattern of these viruses displayed remarkable differences. There were variations in
221 the patterns of Verdadero viral small RNAs between a Mexico isolate and several different
222 samples of Miami, Florida isolates (**Supp. Fig. S4B**). The Binegev-like virus and Renna virus
223 were found from the west to the east coasts of the United States, and they also displayed
224 interesting variations in viral small RNA patterns between samples (**Fig. 3A, Supp. Fig. S4D,E**).

225 The viruses most widespread across the *Ae. aegypti* small RNA samples of the Americas
226 were PCLV and HTV (as noted previously in Olmo, et al²²), along with the *Ae. aegypti*
227 Anphevirus first described in a Florida strain²⁵. Remarkably, PCLV and HTV were most
228 frequently found in *Ae. aegypti* of the southern fraction of the Americas while Anphevirus
229 displayed a bias for the northern fraction of the Americas (**Fig. 3C,D**). Florida, Mexico, and
230 Caribbean locales reflected a mixing zone for these three viruses. Viral small RNA patterns and
231 abundances fluctuated widely between individual mosquito isolates (**Fig. 3E,F**), but Anphevirus
232 and HTV small RNAs had a clear plus strand bias, especially for the viral structural genes. In
233 contrast, PCLV small RNAs displayed a stronger antisense bias and mostly originated from the
234 'Small' and 'Medium' segments. A principal component analysis of the viral small RNAs
235 supports the delineation of North American group mosquitoes from the Central and South
236 American mosquitoes (**Supp. Fig. S4F**).

237 Miami, Florida and various places in Brazil are sites of recent dengue fever outbreaks³⁵, yet
238 no DENV small RNAs were detected in any of these samples, including those from *Ae. aegypti*
239 sampled by the Florida Department of Health (FDOH) in the vicinity of known dengue fever
240 patients. The lack of DENV small RNAs mirrors the sporadic and challenging detection efforts of
241 DENV in surveillance of mosquitoes to precede a human disease outbreak^{1,13-15}. However, all of
242 these samples revealed that multiple insect viruses can simultaneously infect and generate a
243 robust RNAi response of viral small RNAs in a small mosquito cohort (**Fig. 3A, B**).

244

245 *A snapshot into Asian Ae. aegypti viral small RNA patterns.*

246 We generated a new collection of small RNA libraries from *Ae. aegypti* mosquitoes across
247 eastern Asia (**Figure 4A,B**) and observed a significant diversity of insect viruses generating
248 significant small RNA responses. Several mosquito viruses in Asia were common with those in
249 American samples (**Fig. 3**), such as HTV and PCLV (**Fig. 4C,D**). This also included the Partiti-
250 like virus-1 Jane strain, which notably persisted in the ThaiKP colony that was propagated for
251 >40 generations since its isolation in 2010³⁶. Interestingly, the ThaiKP Partiti-like virus-1 viral
252 small RNAs were strongly biased for the plus strand of the RdRP gene in contrast to both plus
253 and minus strand siRNAs against this virus in a Singaporean wild isolate (**Fig. 4E**). The
254 molecular implications of these and similar diverse fluctuations in viral small RNAs from HTV
255 (**Fig. 4C**) remain unknown.

256 Surprisingly, our MSRG pipeline detected the first case of dengue viral small RNAs from a
257 wild mosquito isolate – a Singaporean sample whose RNA was sequenced in Olmo, et al²² (**Fig.**

258 **4F)**. DENV1 siRNAs in this Singaporean *Ae. aegypti* were more abundant than viral piRNAs,
259 reflecting the same trend as DENV infections in mosquito Aag2 cells²³. These natural DENV1
260 small RNAs also accumulated to the same level as a DENV2 infection of a lab *Ae. aegypti*
261 experiment³⁷, reflecting the likelihood of significant DENV1 infection in this Singaporean sample.

262

263 *African Ae. aegypti colonies harbor a persistent African-specific mosquito virus.*

264 Most *Ae. aegypti* samples in this study thus far belong to the *Ae. aegypti aegypti* subspecies
265 that have higher human biting proclivity and better ZIKV vector competence compared to the
266 contemporary African subspecies *Ae. aegypti formosus*^{38,39}. Most African colonies in this study
267 are subspecies *formosus*, but some are subspecies *aegypti* (THI, NGO, CVerd), while others
268 are substantially admixed (KUM, OGD). Distinct genetic backgrounds between *aegypti* and
269 *formosus* subspecies are one likely explanation behind the differences in these subspecies³⁸,
270 but the persistent viruses generating viral small RNAs in *formosus* subspecies were previously
271 unknown.

272 Our small RNA profiles revealed that PCLV's global reach extends into several African
273 colony strains primarily on the western African coast, which may be biased to the limited access
274 to colonies only from these African locales (**Figure 5A,B**). The PCLV small RNA coverage
275 patterns were also more biased towards the 'S' and 'M' segments with plentiful antisense viral
276 piRNAs, while the 'L' segment mainly generated sense viral piRNAs (**Fig. 5C**). Unlike Asian and
277 Central and Southern American *aegypti* strains where HTV frequently accompanied PCLV in
278 causing a small RNA response in the same mosquito sample, *formosus* strains only had a few
279 HTV small RNA signatures, all of which were found exclusive of PCLV (**Fig. 5A**).

280 Formosus virus displayed the most striking persistent small RNA response in 12 samples
281 from native-range colonies (**Fig. 5A**). This rhabdovirus has a ~12.2kb genome that was initially
282 absent from GBVRL database. It was deposited in the TPA (accession BK059424) by Parry, et
283 al³⁰, from a metagenomics assembly of transcriptomes from a different African lab colony not in
284 this study originating from Bundibugyo, Uganda⁴⁰. Importantly, four additional African *Ae.*
285 *aegypti* samples from Olmo et al²², that were established and maintained separately from the
286 main set of native-range colonies, also displayed Formosus virus small RNAs with similar
287 coverage patterns (**Fig. 5D**, i.e. KED_02_JM and KED_Female_LM). Since most of the native-
288 range colonies here were sampled after at least seven generations, we conclude that the
289 Formosus virus is persisting and being transmitted vertically within the colonies.

290 Formosus virus primarily generates viral piRNAs from the sense strands of the NP, HP, and
291 GP genes in the 5' half of the viral genome, while the RdRP gene seemingly evades piRNA
292 production (**Fig. 5D**). Conversely, long RNA sequencing indicated the plus-strand RdRP gene is
293 transcribed as highly as NP, HP, and GP genes (**Fig. 5E**). We also detected robust minus
294 strand long RNAs indicative of replication of the negative-strand rhabdovirus genome, yet much
295 fewer viral piRNAs were generated from this minus strand (**Fig. 5D**). Future studies will
296 investigate how the RdRP gene transcript and negative-strand genomic transcript evade the
297 RNAi machinery. Lastly, we observed some sex-specific Formosus viral piRNA accumulation
298 patterns in males at the 5' end of the NP gene, even though long viral RNA patterns look similar
299 between the sexes.

300

301 *Insights into mosquito virus replication and RNAi dynamics from matched long RNA sequencing*

302 Next, we asked if the Formosus virus example of the interplay between viral small RNAs and
303 long RNAs was indicative of a broader picture of dynamics between viral replication and viral
304 small RNA biogenesis for other viruses in other *Ae. aegypti* mosquitoes. We compared in a
305 scatterplot the log-transformed long and small RNA levels between ~75 matched samples for
306 each mosquito virus (**Figure 6A**). We noted four interesting virus groups in this analysis. The
307 first group is AEFE1 EVE measurements from adults all clustering together, which displayed
308 more small RNAs than long RNAs, and reflects the care we took to generate and sequence
309 these RNA libraries as reproducibly as possible.

310 The next two groups were viruses in samples that displayed significant viral long RNAs but
311 few small RNAs, and vice versa significant viral small RNAs without much long RNAs. In several
312 Florida isolates that were captured by municipal vector control surveillance, the *Aedes Toti*-like
313 virus appeared to replicate and express viral genes effectively, perhaps before the mosquitoes
314 could mount an RNAi response with small RNAs (**Fig. 6B**). Notably, these Florida mosquitoes
315 had robust small RNA responses to endogenous TEs like *R1-El4* as well as other viruses like
316 Verdadero virus, validating that the lack of *Aedes Toti*-like virus small RNAs is not merely a
317 technical error (**Supp. Fig. S4B**). More interestingly, there were some virus cases like
318 Formosus virus in the Entebbe, Uganda colony (ENT), in which both males and females only
319 generated viral small RNAs. This may be linked to the complete loss of the viral long RNAs (**Fig.**
320 **6C**) despite the long RNA libraries still tracking the TE long RNAs.

321 The fourth standout group is viruses in the BZL strain eggs which displayed massive levels
322 of long RNA reads from both plus and minus strands of the various viruses in this strain (**Fig.**
323 **6D**). For both TMBTLV and an *Aedes Binegev*-like virus (ABLV), the long RNA reads were more
324 than an order of magnitude greater than the small RNAs in both the eggs and whole adult
325 females. Despite the production of abundant viral piRNAs and siRNAs in the adult whole female
326 and maternal contribution of these viral small RNAs to the eggs, virus silencing does not appear
327 effective, which could allow for efficient vertical virus transmission from female to egg.

328

329 *Infectious capacity of metagenomically assembled virus entries from RNA sequencing*

330 Although TMBTLV and Formosus virus were in the TPA and MAG databases within
331 GenBank³⁰, they were not included in the GBVRL at the time of our initial analysis, raising the
332 question of whether these entries are truly infectious viruses. To bolster our RNA sequencing
333 findings, we sought to isolate viruses from filtered mosquito homogenates to infect mosquito cell
334 cultures (**Figure 7A**). If the virus infections in cell cultures were deemed stable, we could then
335 sequence the virus genomes for GBVRL submission and sequence viral small RNAs from
336 infected cells to compare against the mosquito viral small RNAs.

337 We followed established serial propagation and blind passaging procedures using C6/36 *Ae.*
338 *albopictus* cells and Aag2 *Ae. aegypti* cells to test for virus infection from the homogenates of
339 the BZL lab strain (**Fig. 7B**), a colony strain from Poza Rica, Mexico (PZR, **Supplemental**
340 **Figure S5A**), and two African colony strains AWK and KIN (**Supp. Fig. S5B**). After 5-to-6
341 rounds of blind passaging, we were able to use RT-PCR and primers to consistently detect the
342 stable infections of TMBTLV RNAs S1 and S2, the RdRP and Capsid genes of Verdadero virus,
343 and a Totivirus and Formosus virus amplicon. These results suggested we could establish
344 stably infected mosquito cells as new stocks of these mosquito viruses, albeit at moderate virus
345 copy levels, since we have not yet established a titer regime for these viruses.

346 For TMBTLV, Totivirus, and Formosus virus, we examined infection kinetics in mosquito
347 cells and virus tropism. TMBTLV infection rates proceeded faster and to a greater extent in

348 C6/36 cells than Aag2 cells (**Fig. 7C**), likely because of the *Dcr2* mutation in C6/36 cells that
349 reduces antiviral RNAi and makes C6/36 cells the most common cell culture system to isolate
350 viruses from insects^{41,42}. In contrast, Formosus virus exhibited efficient replication only in Aag2
351 cells, suggesting that this virus has more restricted species tropism than TMBTLV. Totivirus
352 replicated in both cell types, although replication kinetics were a little faster in Aag2 cells. Lastly,
353 we tested if these three viruses can also replicate in mammalian Vero E6 and Huh7.5 cells that
354 are susceptible to mosquito-borne human viruses like ZIKV and DENV. Surprisingly, blind
355 passage of BZL mosquito homogenates into these mammalian cells was effective at showing
356 TMBTLV replication, and some cytopathic effects were observed (**Supp. Fig. S5C, D**). A fainter
357 signal for the Formosus virus was also observed in some of the blind passaging in the
358 mammalian cells.

359 The transfer of these mosquito viruses into cell cultures enabled us to clone TMBTLV and
360 Formosus virus genomic fragments for sequencing and GBVRL submission. When comparing
361 the genomes of our TMBTLV and Formosus virus isolates to the initial genomes assembled by
362 Parry et al,³⁰ we detected 25 total protein-coding mutations in the ~4.5kb TMBTLV genome, but
363 only 11 protein-coding mutations in the ~12.2kb Formosus virus genome (**Supp. Figure S6**).
364 Together, these results provide strong evidence supporting the infectious capacity of these
365 persistent mosquito viruses even though the source mosquitoes displayed an RNAi response
366 with viral small RNAs.
367

368 *Small RNA responses in virus-infected mosquito cells are comparable to the whole mosquito*

369 We subjected our virus-infected C6/36 cells and Aag2 cells to small RNA sequencing to
370 determine whether new viral small RNA patterns in these cell cultures could recapitulate the
371 patterns in the whole mosquito. These C6/36 and Aag2 cell lines were obtained from other
372 mosquito laboratories and already had known viruses persisting and generating small RNAs<sup>23,43-
373 45</sup> (**Figure 8A,B**). We were able to reconfirm the presence of Cell Fusing Agent Virus (CFAV)
374 and PCLV in both of our 'mock' C6/36 and Aag2 cells²³, as well as discover Sobemo-like viruses
375 and densoviruses in our C6/36 and Aag2 cells, respectively.

376 In addition to the preexisting viruses in the background-strain cell lines, these infected
377 mosquito cells in Figures 7 and S5 exhibited viral small RNAs from TMBTLV (from BZL
378 mosquitoes), Anphevirus (from PZR mosquitoes), and Formosus virus (from AWK and KIN
379 mosquitoes). The C6/36 cells generated a strong viral piRNA response to all three viruses with
380 remarkable resemblance to the primarily plus strand viral piRNA patterns observed in the
381 cognate whole mosquito sample (**Fig. 8C**). The *Dcr2* mutations in C6/36 cells prevent these
382 cells from generating conventional siRNAs from a double-stranded RNA (dsRNA) intermediate
383 made during virus replication⁴¹, which could partially explain the lack of virus minus strand
384 siRNAs. It is possible that apparent virus plus strand siRNAs in C6/36 cells could also be
385 piRNAs that have been extensively trimmed to this shorter siRNA-like length.

386 The TMBTLV infection of Aag2 cells, which have an intact *Dcr2* gene, showed a clear
387 pattern of viral siRNAs processed from a dsRNA intermediate since minus and plus strand
388 siRNAs evenly covered the entire virus genome (**Fig. 8D**). This intact antiviral RNAi response in
389 Aag2 cells against TMBTLV explains the slower infection kinetics in Aag2 cells versus C6/36
390 cells (**Fig. 7C**). However, Anphevirus and Formosus virus both did not generate antisense
391 siRNAs in Aag2 cells, but instead produced mainly plus strand piRNAs from all the viral genes
392 except the RdRP gene, in striking similarity to the small RNA patterns in the whole mosquito
393 (**Fig. 8D**). Our results set the stage for utilizing these cell cultures and new mosquito virus

394 stocks to study the biochemical mechanisms underlying the effective transmission of these
395 viruses despite the triggering of small RNA responses in mosquito cells and animals.

396

397 **DISCUSSION**

398 We present here the most comprehensive global mosquito virus small RNA survey to date
399 and demonstrate how strikingly frequent are the robust small RNA responses to the persisting
400 viruses within wild mosquitoes and in recently established colonies. In addition to analyzing new
401 small and long RNA sequencing datasets, our MSRG pipeline also integrated previously
402 published small RNA datasets²² to yield a more complete picture of the geographic distributions
403 of mosquito viruses. Our analysis confirms the global reach of PCLV (and to a lesser extent
404 HTV) infecting *Ae. aegypti*, but geographical boundaries of mosquito viruses based on small
405 RNA profiles can still be observed such as virus delineations in the Americas continents and
406 African-specific viruses like the Formosus virus. By comparing viral long RNAs to small RNAs,
407 we uncovered new insights into the dynamics of virus replication intermediates that can evade
408 RNAi interactions, mechanisms which we can explore in the future by infecting these viruses
409 into mosquito cell cultures while knocking down mosquito host genes.

410

411 *Limitations and future opportunities of the study*

412 Small RNA sequencing is a low throughput assay with more complicated library construction
413 methods compared to long RNA sequencing. RNA sequencing is also less practical for routine
414 virus surveillance conducted by municipal vector control organizations, which rely on primer
415 panels in RT-PCR assays due to speed and low cost. Our current MSRG bioinformatics
416 analysis is now more streamlined from the initial iteration^{23,26}, but it remains a challenge to keep
417 pace with the deluge of new virome entries and metagenomics studies^{4,5,27}. Because of
418 throughput limitations with small RNA sequencing scalability, our study is a snapshot in time
419 and limited to specific locales despite our efforts to profile mosquitoes globally. Some of the
420 narrow sampling issues are logistical and political restrictions on access to mosquitoes from
421 certain countries.

422 It is possible that some of the viruses detected in colonized mosquitoes could result from a
423 lab-acquired infection when multiple colonies share the same insectary space. Colonies that
424 have been perpetuated for many generations may begin to reflect the environment that they are
425 raised in instead of the locale that they originate from and are meant to represent. For instance,
426 we detected the American-prevalent Anphevirus and CFAV in the Lambrechts lab's Lope strain
427 and the McBride lab's ZIK strain, respectively. These African-originating strains had the most
428 generations of lab propagation (>15 generations) compared to the other African colonies
429 (average ~8 generations, Fig. 8A).

430 Climate change is expanding the ecological niches for mosquitoes like *Ae. aegypti*, which
431 have now invaded further into North America and Europe³⁵. Broadening our small RNA
432 genomics surveys to include animal captures in new locations would illuminate the potential
433 disease vectoring threat that these invading mosquitoes would pose to human populations.
434 Small RNA sequencing provides extra rigor beyond the RT-PCR assay because the virus is
435 expected to have replicated at high enough levels to provide the long RNA precursors needed
436 for the mosquito RNAi machinery to convert into small RNAs. Thus, it was particularly surprising
437 to observe the first case of DENV small RNAs in a wild mosquito isolate (**Fig. 4F**). It remains an
438 open question whether wild mosquitoes displaying arboviral small RNAs are still actively

439 infectious or if the RNAi response is now suppressing virus transmission. Only sequencing
440 approaches, not RT-PCR, can discover new viruses and RNAi responses in mosquitoes, and
441 our study demonstrates the utility of this approach to augment arbovirus surveillance programs.
442

443 *Could persistent insect viruses in wild mosquitoes be a source for emerging pathogens?*

444 It is practical for academic lab mosquito strains to be clean of persistent viruses that would
445 complicate genetic analyses. However, to better model arbovirus vector competence in the wild
446 where human populations are affected, we need to study further the impact of persistent viruses
447 that generate abundant viral small RNAs in mosquitoes. The infectivity of the mosquito viruses
448 into cell cultures was surprising considering the apparent viral small RNA responses exhibited
449 by the mosquito animals and the cell cultures. We did have occasional issues with mosquito
450 viruses transferring between culture flasks by possible aerosolization within the biosafety
451 cabinet during the handling of multiple mosquito homogenates (**Fig. 8**). This infectivity could
452 explain the persistence and widespread detection of these viruses in so many of the mosquito
453 samples profiled in this study.

454 An open question is whether persistent mosquito viruses can be transmitted from the
455 mosquito to the bitten animal during blood feeding. For instance, it is unclear if massive vertical
456 virus transmission of TMBTLV via maternal contribution to the eggs (**Fig. 6D**) would reflect a
457 meaningful reservoir of viruses in the mosquito salivary glands and proboscis. On the other
458 hand, the LTLV had very high levels of small RNAs in the midgut (**Fig. 3B, Supp. Fig. S4A**),
459 suggesting this virus could likely mix with an initial bloodmeal and be transmitted during
460 subsequent blood feedings. Notably, several of these mosquito viruses generating an RNAi
461 response are related to plant viruses, suggesting that mosquito feeding on plant sap could be
462 another modality of virus transmission, and perhaps these viruses are suppressing RNAi with
463 factors related to P19 found in plant tombusviruses³⁴. Lastly, our preliminary evidence shows
464 that TMBTLV could be transmitted to mammalian Vero E6 and Huh7.5 cells. This surprising
465 result was reproduced across multiple independent replications of this infection. This finding
466 points to the possibility that arboviruses like DENV and ZIKV also started as mosquito-specific
467 flaviviruses like CFAV until they gained the capability to infect humans to facilitate new avenues
468 of viral transmission.

469

470 *Can the Formosus virus and other mosquito viruses affect arbovirus vector competence?*

471 Previous studies on genetic variations causing lower ZIKV vector competence in *Ae. aegypti*
472 *formosus* compared to *Ae. aegypti aegypti* were done prior to our new results on the Formosus
473 virus persisting and being transmitted vertically in these African colonies³⁸. It is possible that the
474 *formosus-aegypti* crosses conducted in the previous study may have also transmitted Formosus
475 virus, and insect virus interactions with arboviruses are documented. For example, co-infection
476 with CFAV reduces DENV and ZIKV replication in cell culture and transmission in mosquitoes⁴⁶.
477 Our next priority will be to test how these mosquito viruses interact with DENV and ZIKV in co-
478 infection experiments.

479 In addition, these mosquito viruses may also be useful agents for mosquito transduction and
480 control in a similar vein as insect densovirus^{47,48}. Densoviruses have been engineered to
481 carry reporter genes and toxic gene knockdown cassettes as prototype genetic tools to
482 manipulate mosquitoes, but technical challenges must be overcome to realize the potential of
483 densoviruses as mosquito genetic tools. We observed that, like the insect viruses investigated

484 in this study, densoviruses also engage with the mosquito RNAi pathway to generate viral small
485 RNAs (**Fig. 8B** and Ma, et al²³). Rivaling densovirus in genome compactness is TMBTLV at just
486 ~4.5kb across two RNA segments. We are still characterizing the viral genomic RNAs' 5' and 3'
487 ends to build and test infectious clones, but there is potential for new mosquito genetic tools to
488 be garnered from such a survey.

489 Previous studies have tried to harness RNAi to control mosquito populations and curtail
490 arbovirus transmission⁴⁹⁻⁵⁴, but dsRNA delivery by injection has a low throughput, and RNAi by
491 dsRNA ingestion is also limited in mosquitoes⁵⁵. Mosquito viruses like TMBTLV that can readily
492 trigger dsRNA and siRNAs in mosquitoes and cells (**Fig. 2B, 8D**) may solve the siRNA delivery
493 problem. Future small RNA sequencing surveys like this can potentially discover new virus-
494 vector biology, facilitating the development of new tools to combat the health threat that
495 arboviruses pose to humans.
496

497 **MATERIALS AND METHODS**

498

499 **Construction of small and total RNA libraries from whole mosquitoes and cells**

500 Total RNA was extracted from whole frozen mosquitoes or mosquitoes preserved in RNA
501 protection solutions. Zirconium beads (3.0mm) in a bead beater were used to homogenize
502 mosquitoes before proceeding with NEB Monarch Total RNA Miniprep Kit (NEB #T2010). The
503 on-column DNase I treatment was performed during RNA extraction. Small RNA libraries were
504 made using NEBNext Small RNA Library Prep Set (NEB #E7330) with up to 5µg of RNA input.
505 During library amplification, samples were put through up to 25 total PCR cycles. Total RNA
506 libraries were made using Zymo-Seq RiboFree Total RNA Library Kit using up to 250ng RNA
507 input, and following manufacturer's protocol. All libraries were checked on an Agilent
508 Bioanalyzer 2100 using either DNA 1000 or High Sensitivity DNA kit and sequenced at the
509 BUMC Microarray and Sequencing Resource on an Illumina Next Seq 2000 using P3 flow cells
510 for 50SE and 50PE reads for small RNA and total RNA libraries, respectively.

511

512 **Running VirusDetect with TPA and MAG database modifications**

513 The source code for VirusDetect Version 1.7²⁸ was downloaded from the GitHub repository
514 (<https://github.com/kentnf/VirusDetect>). Additionally, the preprocessed VRL virus database
515 Version 248 was obtained from the VirusDetect webpage
516 (http://bioinfo.bti.cornell.edu/ftp/program/VirusDetect/virus_database/v248/). The source code
517 and the corresponding database were deployed on the Shared Computing Cluster of Boston
518 University. To meet VirusDetect requirements, the following packages were installed on the
519 cluster: perl version 5.28.1, bioperl version 1.7.2, and python3 version 3.8.10.

520 The Third Party Annotation (TPA) GenBank files were sourced from the NCBI database
521 (<https://ftp.ncbi.nlm.nih.gov/tpa/tsa/>). An in-house R pipeline was utilized to extract GenBank
522 records, preserving 45,601 entries with the taxonomy 'VIRUS' for downstream processing. For
523 each entry, both the genome sequence and corresponding protein sequences were retrieved
524 from the NCBI database using the ACCESSION ID and saved in database files formatted for
525 VirusDetect. Two id-mapping files were also generated following VirusDetect's instructions. In
526 addition to the GBVRL and TPA datasets, partial virus sequences in GenBank were annotated
527 as Metagenome Assembled Genome (MAG). All MAG records downloaded from GenBank
528 underwent the same processing steps as the TPA database. A total of 65,167 entries were
529 retained to construct the modified VirusDetect database.

530

531 **Sequencing and bioinformatics analysis of small and long RNAs from mosquitoes**

532 RNA Libraries were selected if they were above 1 million reads to run through the Mosquito
533 Small RNA Genomics (MSRG) pipeline^{23,26}. Outputs from MSRG include the alignment to a
534 curated virus database. This includes breakdowns (in reads per million) of reads coming from
535 siRNAs (18-23 nt) and piRNAs (24-35 nt), which strand reads map to, and a ratio of peaks to
536 the average distance between peaks (used as a metric of coverage across viral sequence).

537 Some of the libraries in the Olmo et al study²² were made using a modified version of the
538 NEBNext Multiplex Small RNA Library Prep Set where a random 6-mer was attached to the 5'
539 adapter. A Cutadapt trimming step in MSRG was adapted to remove the first six bases of each
540 read. This modified pipeline was then used to run MSRG on these samples.

541 Total RNA libraries were sequenced as 50PE reads but were processed to look like small RNA
542 reads so that they could also be run using the MSRG pipeline. All reads were trimmed to 35nt,
543 R1 was reverse complemented to be in the same orientation as R2, and then the reverse
544 complemented R1 was merged with R2. The resulting FASTQ file was then run through MSRG
545 and processed just like the small RNA libraries.

546 Viruses were only considered to be present in a sample if it had at least 10 reads per million
547 (RPM) and a coverage ratio greater than 0.75. Additionally, a strand bias score was calculated
548 for each virus/sample pair by taking the ratio of reads mapping to the top versus bottom strand.
549 For plots made using R/RStudio, packages used include: ggplot2, tidyverse, and ggpubr.

550 In the bubble plot, log₁₀ of the strand bias score was mapped to the color, and log₁₀ of the
551 RPM was mapped to the size. For the strand bias, any value greater than 1 (a greater than 10-
552 fold strand bias) was squished to be the “max” color on the scale. For the scatterplots
553 comparing the small and total RNAs, the 10 RPM ratio was ignored in order to completely
554 compare the two libraries. Sample/virus pairs were kept if at least one of the two libraries was at
555 least 10 RPM.

556

557 **Cell Cultures**

558 C6/36 and Aag2-TC mosquito cell cultures were propagated in DMEM (Gibco) supplemented
559 with 10% FBS and 1% Tryptose Phosphate Broth (TPB) and maintained at 28° C and 5% CO₂.
560 HUH7.5 and Vero E6 TC cell cultures were propagated in DMEM (Gibco) supplemented with
561 10% FBS and 1% P/S (Penicillin/Streptomycin) and maintained at 37° C and 5% CO₂. The
562 mammalian cell cultures were split by washing with 1X PBS and 0.25% Trypsin EDTA.

563

564 **Cell infection with mosquito viruses**

565 The mosquitoes received from Awka (Nigeria), Kintampo (Ghana), Pozo Rica (Mexico), and
566 Banzon Research (USA) were used to infect the cell lines with viruses. Mosquitoes were
567 homogenized in 1.5 mL DMEM with zirconium beads (3.0mm) in a bead beater. The
568 homogenate was passed through a 0.45µm pore size filter and added into a T-25 flask
569 containing ~90% cell confluency in 5ml of media in T-25 flasks. On the first day of infection
570 only, a cocktail of Primocin, Normocin, and Fungin antibiotics (InvivoGen) was added to the
571 infection media.

572 For mosquito cell blind passage, the medium was collected from the infected cells, passed
573 through the 0.45µm pore size filter, and added into a new T-25 flask containing ~90% cell
574 confluency. For serial passage, the cells infected by homogenizing the mosquitoes were split
575 into a new flask with a fresh medium. This process was repeated every 7th day for both blind
576 and serial passage. The cells collected on the 7th day were used for total RNA extraction to
577 confirm the virus infection.

578 For mammalian cell blind passage experiments, VeroE6 and Huh7.5 cells were infected with
579 non-titered virus stock medium collected from C6/36 cells infected with the mosquito-viruses
580 from Awka, Nigeria (AWK) and Banzon Research (BZL). The collected medium was passed
581 through a 0.45µm pore size filter and was added to a new T25 flask containing ~90%
582 confluency. On the 4th day post-infection, the medium was replaced with a fresh batch of virus
583 stock medium. On the 7th day, medium from the cells was transferred out and cells were

584 washed twice with 1X PBS before trypsinizing cells for total RNA extraction and archiving live
585 cells by cryopreservation. This process was repeated on every 7th day.

586

587 **RNA extraction and RT-qPCR analysis of virus amplicons**

588 Total RNA was extracted with Monarch Total RNA Miniprep Kit (NEB) following the
589 manufacturer's instructions, including the DNase I treatment step. For RT-PCR, 1µg RNA was
590 used for first strand cDNA preparation by using ProtoScript II Reverse Transcriptase kit followed
591 by Phusion High-Fidelity DNA Polymerase (NEB). RT-qPCR was performed with LunaScript RT
592 Supermix (NEB) in a Bio-Rad CFX Opus 96 Real-Time PCR System. The housekeeping genes,
593 *Ae. aegypti* actin (GenBank accession XM_001659913) and *Ae. albopictus* actin (GenBank
594 accession XM_019702203), were used as reference genes to normalize the target gene
595 expression by $2^{-\Delta\Delta C_t}$ methods.

596

597 **Virus Kinetics**

598 The virus kinetics of TMBTLV, Formosus virus, and Totivirus were measured in C6/36 and
599 Aag2-TC cell lines. First, the virus particles were quantified in the virus stock medium by digital-
600 droplet PCR on the Bio-Rad QX-200 instrument. The cells were infected with 5 mL virus stock
601 medium in the T-25 flask. The cells were collected from the flask on the 1st, 3rd, 5th, 7th, and
602 10th day post-infection. At each time point, the cells were resuspended in the medium using a
603 pipette, and one mL of the medium containing the cells was replaced with a fresh medium. The
604 cells were collected and processed from one mL for the total RNA extraction. Subsequently, the
605 virus quantification was done by RT-qPCR, as described above. This experiment was repeated
606 in three independent biological replicates for each time point.

607

608 **Resource availability**

609 **Lead contact**

610 Further information and requests for resources and reagents should be directed to and will be
611 fulfilled by the lead contact, Nelson Lau (nclau@bu.edu).

612 **Materials availability**

613 All unique/stable reagents generated in this study are available from the lead contact. Material
614 transfer agreements with Boston University may apply.

615 **Data and code availability**

616 All sequencing data produced and generated by this study is available on Sequencing Read
617 Archive (SRA) under BioProject PRJNA1104658. See Tables S1 and S2 for specific BioSample
618 and SRA accessions. SRA accessions for publicly available datasets used in this study can be
619 found in Table S1C. The MSRG pipeline code can be found on the Github repository:
620 <https://github.com/laulabbumc/MosquitoSmallRNA>.

621

622 **ACKNOWLEDGEMENTS**

623 We thank Mohsan Saeed and Fabiana Feitosa-Suntheimer for comments on this manuscript.
624 We acknowledge Anubis Vega-Rúa and Silvânia da Veiga Leal as the source of mosquito
625 colonies from Guadeloupe and Cape Verde, respectively. We thank João Marques for
626 assistance in accessing his lab's public datasets. We acknowledge Mark Stenglein and Marylee
627 Kapuscinski for technical assistance in small RNA sequencing.

628 N.C.L.'s lab was funded by NIH/NIGMS (GM135215). L.L. was supported by the French
629 Government's Investissement d'Avenir program, Laboratoire d'Excellence Integrative Biology of
630 Emerging Infectious Diseases (grant ANR-10-LABX-62-IBEID). A.B. was supported by a stipend
631 from the Pasteur - Paris University (PPU) International PhD Program. A.E.W. and K.E.O. were
632 supported by R01 AI130085. A.E.W. and E.C. research was supported by the Intramural
633 Research Program of NIH/NIAID (AI001246). N.H.R. and C.S.M.'s work here was supported by
634 NIH GRANT R00DC012069 and a New York Stem Cell Foundation Robertson Neuroscience
635 Investigator Award. A.L.C.S. and M.D. were supported by U.S. Centers for Disease Control and
636 Prevention (CDC) Cooperative Agreement Number 1U01CK000510, Southeastern Regional
637 Center of Excellence in Vector-Borne Diseases Gateway Program. The CDC did not have a role
638 in the design of the study or the collection, analysis, or interpretation of data. D.E.B.'s work was
639 supported in part by grants from the National Institutes of Health, National Institute of Allergy
640 and Infectious Diseases (AI148477). M.A.Y. is supported by the Searle Scholars Program, the
641 Richard and Susan Smith Family Foundation, the Esther A. & Joseph Klingenstein Fund, and
642 the Simons Foundation.

643 *Author contributions*

644 Conceptualization, N.C.L., S.G., R.S., A.E.W., and Z.Z.; Methodology/Investigation, N.C.L., S.G.,
645 R.S., A.C.V., A.E.W., Z.Z., and G.D.; Formal Analysis, S.G., N.C.L., and R.S.; Data
646 Curation/Software, C.Z., S.G., and Z.Z.; Writing – Original Draft, N.C.L., S.G., and R.S.; Writing
647 – Review & Editing, N.C.L., S.G., A.E.W., L.L., N.H.R., C.S.M., and Z.Z.; Visualization, N.C.L.,
648 S.G., R.S., and Z.Z.; Funding Acquisition, N.C.L., L.L., K.E.O., E.C., C.S.M., M.A.Y., D.E.B., and
649 M.D.; Sample Contribution & Resources, A.E.W., I.S.V., N.H.R., A.G.S., D.E.B., J.M., S.S.W.,
650 A.C., T.R., M.S., A.B., J.A., O.B.A., D.A., W.L.L., C.H.C., C.V., C.G.A., A.P., T.M., B.C., D.W.,
651 D.S., M.A.Y., A.L.C.S., M.D., A.B., L.L., C.S.M., K.E.O., and E.C.

652

653 *Declaration of interests*

654 The authors declare no competing interests.

655 FIGURES and TABLES LEGENDS

656 **Figure 1. Overview of the global mosquito small RNA survey to discover RNA-** 657 **interference (RNAi) responses to mosquito viruses.**

658 (A) Overview of the MSRG pipeline applied to a global survey of whole mosquitoes from
659 Americas, Asia, Africa and laboratory strains. (B) Implementation of the VirusDetect program
660 with updated GBVRL and custom databases for comprehensive mosquito virus detection. (C)
661 Summary tabulation of the samples and RNA libraries analyzed in this study.

662

663 **Supplemental Figure S1. Using the AEFE1 EVE as a constant quality tracker control for** 664 **viral small RNA profiling and implementing VirusDetect to discover new ISVs from small** 665 **RNAs.**

666 (A) Coverage plots on left and bubble plots on right representing the small RNAs from the
667 Endogenous Viral Element (EVE) from *Ae. aegypti* called AEFE1 (Suzuki et al, 2017). (B) The
668 results output table from the VirusDetect program only loaded with the GBVRL database while
669 analyzing the small RNA library from one of the BZL lab strain samples. (C) Coverage of the
670 de-novo small RNA assembly contigs generated by VirusDetect of the corresponding virus
671 determinations from the table in (B). The red asterisks mark a false-positive call against a
672 tombus-like virus entry in GBVRL with limited contig coverage. (D) VirusDetect results
673 analyzing the same BZL lab strain small RNA library in (B) but now with a custom loading of the
674 TPA and MAG databases that are distinct from GBVRL. This result now shows two new viruses
675 with much more complete contig coverage in (E), such as the Tiger Mosquito Bi-segmented
676 Tombus-Like virus (TMBTLV) that is the true hit replacing the false-positive call in (B-C).

677

678 **Supplemental Figure S2. Read length distribution profiles of all the *Ae. aegypti* small** 679 **RNA libraries analyzed in this global survey.**

680 Read length distribution plots as percentages of the libraries with functional classes of reads
681 represented by different colored lines. The groups correspond to (A) Lab strains, (B) North
682 American strains, (C) Central and South American strains, (D) Asian strains, (E) African strains,
683 (F) Cell lines, (G) Analyzed-separately-as-additional-replicates samples.

684

685 **Figure 2. In contrast to commercial *Aedes aegypti* lab strains, most academic lab strains** 686 **are low in viral small RNAs indicative of persistent viruses.**

687 (A) Bubble plot of viral small RNAs from lab strains. Number reads per million is reflected by
688 bubble diameter, and color represents strand bias of reads, red is plus strand biased, blue is
689 minus strand biased. Magenta circles and dashed boxes mark the BZL and GP samples that
690 are commercial lab strains of *Ae. aegypti* and both are infected by Tiger mosquito bi-partite
691 Tombus-like virus (TMBTLV). All the other *Ae. aegypti* samples were reared in academic labs.
692 See metadata in Table S1 for sample details. Lab initials: BZL = Benzon Research, MY = Meg
693 Younger, DB = Doug Brackney, JM = João Marques, ZT = Zhejian Tu, GH = Grant Hughes, TC
694 = Tonya Colpitts, BH = Bruce Hay, GP = Gorben Pijlman (B) Coverage plots of TMBTLV and
695 Zika virus (ZIKV) small RNAs from the two commercial lab strains. Additional coverage plots of
696 (C) two viruses that only generated viral small RNAs in a 2nd batch of female BZL *Ae. aegypti*
697 and (D) other viruses always present in the BZL strain in both females and males. (E) Viral
698 small RNAs and likely persistent viruses are detected in the BZL eggs at lower small RNA levels
699 with a more distinct pattern of antisense viral small RNAs in eggs compared to the parental
700 whole females in (B).

701

702 **Supplemental Figure S3. Complete bubble plots of viral small RNAs sampled in this**
703 **global survey study.**

704 Top plots are the siRNA-length small RNAs (18-23nt), the middle plots are the piRNA-length
705 small RNAs (24-35nt), and the bottom plots are all the small RNAs as shown in the main figures
706 2 – 6. The groups correspond to (A) Lab strains, (B) North American strains, (C) Central and
707 South American strains, (D) Asian strains, and (E) African strains. (F) are additional samples
708 that we considered as replicates of other libraries but did not include in the main study because
709 of concerns of potential cross-sample contamination.

710

711 **Figure 3. *Ae. aegypti* in the Americas display diverse insect virus small RNA responses**
712 **with geographic boundaries delineated by three prevalent mosquito viruses.**

713 (A) Bubble plot of viral small RNAs from North America *Ae. aegypti* samples. Number reads per
714 million is reflected by bubble diameter, and color represents strand bias of reads, red is plus
715 strand biased, blue is minus strand biased. (B) Bubble plot of viral small RNAs from Central and
716 South America *Ae. aegypti* samples. Dashed brown line boxes mark two viruses noted in Figure
717 S4AB as being related to plant viruses. (C) Map of the Americas where the *Ae. aegypti* samples
718 originated from, with geographic boundaries delineated by the distribution of Anphevirus and
719 PCLV displayed in the bubble plot in (D). Example coverage plots of Anphevirus small RNAs (E)
720 and PCLV and HTV small RNAs (F).

721

722 **Figure S4. Plant-related viruses and other prevalent viruses in American *Aedes aegypti***
723 **shared on both east and west coasts.**

724 (A-B) Coverage plots of two insect viruses related to plant-originating viruses marked in Figure
725 3. (C) Maps marking approximate locations of *A. aegypti* samples collected from California,
726 Louisiana and Florida. Coverage plots of viral small RNAs from whole *A. aegypti* from: (D)
727 *Aedes* Binegev-like virus- Cayenne strain from California and Florida; and (E) Renna virus from
728 three states. (F) Principal Component Analysis of virus small RNA levels can discriminate
729 different groups of mosquito samples in between the North American, Central & South
730 American, and Lab strains.

731

732 **Figure 4. Asian *Ae. aegypti* viral small RNA patterns reveal a first case of dengue viral**
733 **small RNAs from a wild isolate and share patterns with mosquitoes from the Americas.**

734 (A) Bubble plot of viral small RNAs from Asia *Ae. aegypti* samples. Number reads per million is
735 reflected by bubble diameter, and color represents strand bias of reads, red is plus strand
736 biased, blue is minus strand biased. Dashed pink line box mark notable viruses whose coverage
737 plots are displayed below. (B) Map of Asia locations where the *Ae. aegypti* samples originated.
738 Coverage plots of HTV (C), PCLV (D) and Partiti-like virus 1 Jane strain (E) viral small RNAs
739 from a selection of Asian *Ae. aegypti*. (F) The DENV viral small RNA coverage from a
740 Singapore isolate compared to other DENV viral small patterns from infections in Aag2 cells and
741 a lab injected infection of *Ae. aegypti* by the Myles lab.

742

743 **Figure 5. African *Ae. aegypti* colony strains carry viral small RNAs unique to their**
744 **continent.**

745 **(A) Bubble plot of viral small RNAs from Africa *Ae. aegypti* colony strains. Number reads**
746 per million is reflected by bubble diameter, and color represents strand bias of reads, red is plus
747 strand biased, blue is minus strand biased. Dashed pink line boxes mark the PCLV and
748 Formosus viruses noted in panels (C) and (D), respectively. (B) Map of Africa locations where
749 the *Ae. aegypti* colonies or samples originated. (C) Coverage plots of PCLV small RNAs from a
750 selection of African *Ae. aegypti* showing high M-fragment piRNAs rivaling the S-fragment
751 piRNAs. (D) The Formosus virus viral small RNA coverage from African *Ae. aegypti* colonies
752 from the McBride lab and an independent Kedougou, Senegal sample from Olmo et al 2023.
753 The black arrow points to female-specific viral piRNA specie. (E) Three examples of Formosus
754 virus long RNAs sequenced from matched samples in (D).

755

756 **Figure 6. Virus transmission and replication dynamics suggested by long RNA**
757 **sequencing compared to matched small RNAs of mosquito viruses and transposable**
758 **elements (TEs).**

759 (A) Scatterplot comparing the particular *Ae. aegypti* samples that enabled matched library
760 construction of small RNAs and long RNAs. Sequencing reads per million with a pseudo-count
761 of 1 are plotted on the logarithmic scales. Dots representing the matched samples are colored
762 by sex, and groups of samples clustering together are noted in the labeled ovals. (B) Coverage
763 plots of two Florida isolates of *Ae. aegypti* exhibiting abundant long RNA signal for the totivirus
764 but negligible viral small RNA in the upper plots that contrast both long and small RNAs against
765 an *R1-Ele4* TE. (C) Coverage plots of long RNAs versus small RNAs for the Formosus virus
766 and *R1-Ele4* TE from both males and females of the ENT African colony. (D) Coverage plots of
767 long RNAs compared to small RNAs for viruses and a TE from the BZL strain of *Ae. aegypti*
768 females and dormant eggs.

769

770 **Figure 7. Novel mosquito viruses infecting and replicating in mosquito cells.**

771 (A) Our methodology to molecularly validate the small RNA detection of ISVs are true viruses
772 that can be isolated and verified for triggering the RNAi response in mosquito cells. (B) RT-PCR
773 detection of TMBTLV RNAs S1 and S2 during multiple rounds of blind passaging, starting with
774 BZL mosquito homogenate as the initial virus infection source placed onto C6/36 and Aag2
775 mosquito cells. (C) Virus replication kinetics measured in C6/36 and Aag2 cells over 1 week.
776 Flasks of 1 million cells were infected on Day 0 with 200 viral qPCR amplicons per infection.
777 Virus stocks are from filtered media from subsequent passage from the experiment in (A). Error
778 bars correspond to the standard error from measuring three biological replicates of infection of
779 the mosquito cells with the same virus stock.

780

781 **Figure S5. Tracking virus presence with RT-PCR during blind passaging of C6/36**
782 **mosquito cells from infections with PZR, AWK and KIN mosquito homogenates – with the**
783 **potential to infect mammalian cells.**

784 (A) RT-PCR of viral amplicons against the Verdadero virus RdRP and capsid genes. Two
785 independent infections with two batches of PZR mosquitoes. (B) RT-PCR of viral amplicons
786 against Tesano virus, Totivirus and Formosus virus from two independent infections with two
787 batches of AWK and KIN mosquitoes. The actin control are primers against *Ae. albopictus* actin.
788 (C) Continued blind passaging of virus stocks from C6/36 cells are then tested for infection of
789 mammalian VeroE6 and Huh7.5 cells followed by RT-PCR detection of viral amplicons. Red
790 arrowheads point to insect virus amplicons detected in the mammalian cells. Asterisk marks a

791 totivirus already present in this lab stock of C636 cells. (D) Brightfield images cells from after 1
792 week of infection from 4th passage in (C) before harvesting total RNA for RT-PCR analysis. The
793 cytopathic effect of infection from the BZL virus stock is most evident on C6/36 and Huh7.5
794 cells.

795

796 **Figure S6. Cloning and sequencing the TMBTLV and Formosus virus from *Ae. aegypti*.**

797 (A) Contig assembly of Sanger sequencing of cloned TMBTLV amplicons from S1 and S2
798 RNAs. (B) Diagrams of SNP mutations from our cloned fragment sequencing versus the initial
799 TMBTLV references BK059489 and BK059490. (C) Contig assembly of Nanopore sequencing
800 of cloned Formosus virus amplicons. (D) Diagrams of SNP mutations from our cloned fragment
801 sequencing versus the initial Formosus references BK059424. Both of our TMBTLV and
802 Formosus virus variants sequences have now been contributed to GenBank with verified
803 sequence accessions that should facilitate its inclusion in the future updates of the GBVRL
804 database.

805

806 **Figure 8. Mosquito viruses isolated in cell culture trigger an RNAi response of viral small
807 RNA patterns that resemble the patterns from whole mosquitoes.**

808 Bubble plots of viral small RNAs sequenced from (A) *Ae. albopictus* C6/36 cells and (B) *Ae.*
809 *aegypti* Aag2 cells, either mock or stably infected after multiple blind passages of virus stocks
810 from mosquito homogenates. Number reads per million is reflected by bubble diameter, and
811 color represents strand bias of reads, red is plus strand biased, blue is minus strand biased.
812 Magenta boxes and dashed circles in (A) and (B) highlight specific samples inspected in
813 coverage plots in (C) and (D). Coverage plots for three viruses, TMBTLV, Anphevirus and
814 Formosus virus in (C) C6/36 cells and (D) Aag2 cells with the intact mosquito from females
815 shown in the middle for comparison.

816

817 **Table S1. Overview of small RNA libraries made and used for this study.**

818 Tab S1a. Mosquitoes. Tab S1b. Cell culture infections. Tab S1c. Previously published datasets.

819

820 **Table S2. Overview of total RNA libraries made and used for this study.**

821 Tab S2a. Mosquitoes. Tab S2b. Cell culture infections.

822

823 **Table S3. Current List of Insect viruses in the MSRG pipeline analyzed in this study.**

824 At the bottom are also some additional viruses that were circumspect, were either already
825 removed or are being considered for removal in future extensions of this study.

826

827 **Table S4. List of oligonucleotides used in this study.**

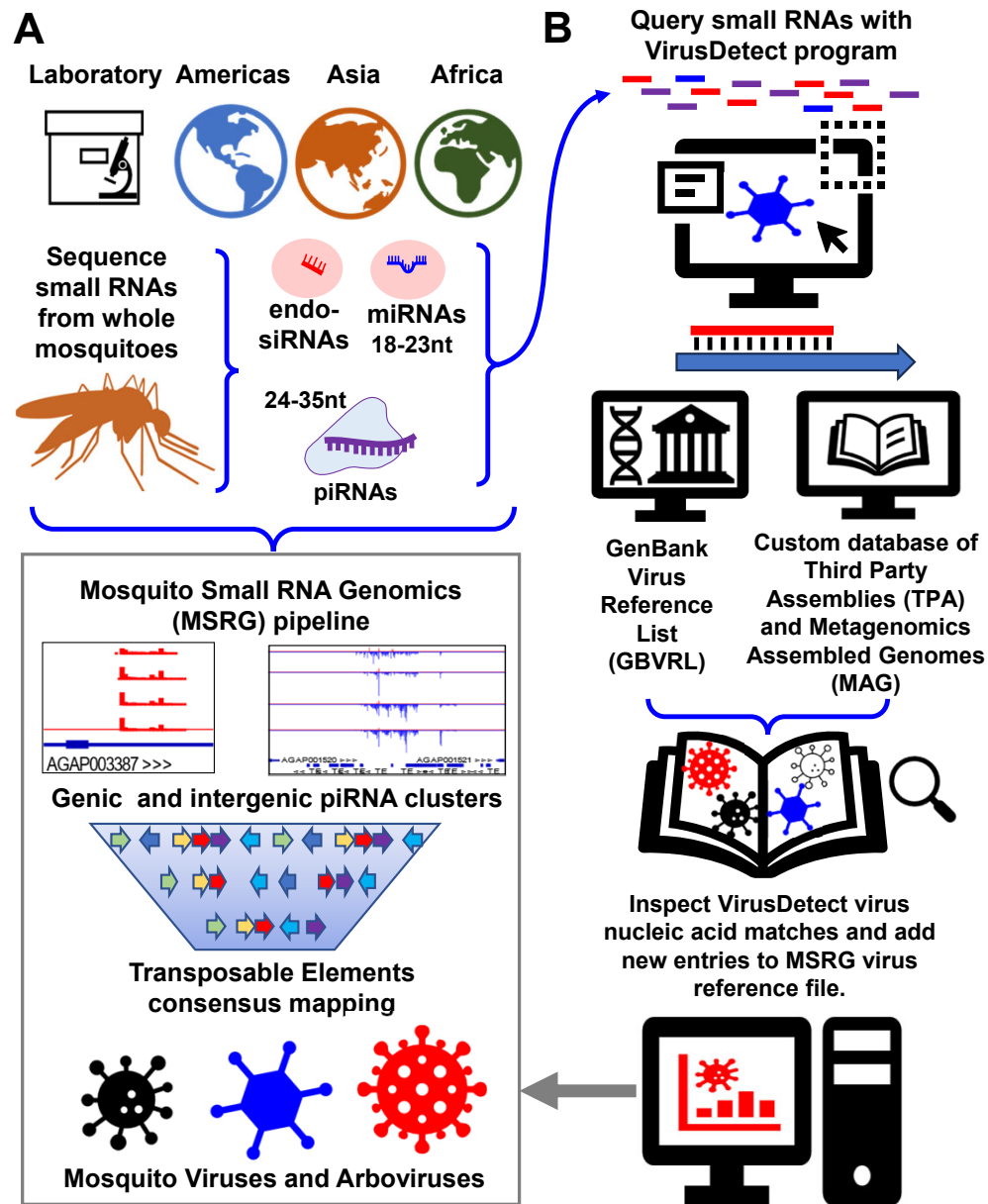
828 REFERENCES

- 829 1. Fernandes, J.N., Moise, I.K., Maranto, G.L., and Beier, J.C. (2018). Revamping Mosquito-borne
830 Disease Control to Tackle Future Threats. *Trends Parasitol* *34*, 359-368.
831 10.1016/j.pt.2018.01.005.
- 832 2. Moonen, J.P., Schinkel, M., van der Most, T., Miesen, P., and van Rij, R.P. (2023). Composition
833 and global distribution of the mosquito virome - A comprehensive database of insect-specific
834 viruses. *One Health* *16*, 100490. 10.1016/j.onehlt.2023.100490.
- 835 3. Pan, Y.F., Zhao, H., Gou, Q.Y., Shi, P.B., Tian, J.H., Feng, Y., Li, K., Yang, W.H., Wu, D., Tang, G., et
836 al. (2024). Metagenomic analysis of individual mosquito viromes reveals the geographical
837 patterns and drivers of viral diversity. *Nat Ecol Evol*. 10.1038/s41559-024-02365-0.
- 838 4. Shi, M., Lin, X.D., Tian, J.H., Chen, L.J., Chen, X., Li, C.X., Qin, X.C., Li, J., Cao, J.P., Eden, J.S., et al.
839 (2016). Redefining the invertebrate RNA virosphere. *Nature* *540*, 539-543.
840 10.1038/nature20167.
- 841 5. Zhang, Y.Z., Shi, M., and Holmes, E.C. (2018). Using Metagenomics to Characterize an Expanding
842 Virosphere. *Cell* *172*, 1168-1172. 10.1016/j.cell.2018.02.043.
- 843 6. Lequime, S., Paul, R.E., and Lambrechts, L. (2016). Determinants of Arbovirus Vertical
844 Transmission in Mosquitoes. *PLoS Pathog* *12*, e1005548. 10.1371/journal.ppat.1005548.
- 845 7. da Costa, C.F., da Silva, A.V., do Nascimento, V.A., de Souza, V.C., Monteiro, D., Terrazas,
846 W.C.M., Dos Passos, R.A., Nascimento, S., Lima, J.B.P., and Naveca, F.G. (2018). Evidence of
847 vertical transmission of Zika virus in field-collected eggs of *Aedes aegypti* in the Brazilian
848 Amazon. *PLoS Negl Trop Dis* *12*, e0006594. 10.1371/journal.pntd.0006594.
- 849 8. Thangamani, S., Huang, J., Hart, C.E., Guzman, H., and Tesh, R.B. (2016). Vertical Transmission of
850 Zika Virus in *Aedes aegypti* Mosquitoes. *Am J Trop Med Hyg* *95*, 1169-1173. 10.4269/ajtmh.16-
851 0448.
- 852 9. Rosen, L., Shroyer, D.A., Tesh, R.B., Freier, J.E., and Lien, J.C. (1983). Transovarial transmission of
853 dengue viruses by mosquitoes: *Aedes albopictus* and *Aedes aegypti*. *Am J Trop Med Hyg* *32*,
854 1108-1119. 10.4269/ajtmh.1983.32.1108.
- 855 10. Feitosa-Suntheimer, F., Zhu, Z., Mameli, E., Dayama, G., Gold, A.S., Broos-Caldwell, A., Troupin,
856 A., Rippee-Brooks, M., Corley, R.B., Lau, N.C., et al. (2022). Dengue Virus-2 Infection Affects
857 Fecundity and Elicits Specific Transcriptional Changes in the Ovaries of *Aedes aegypti*
858 Mosquitoes. *Front Microbiol* *13*, 886787. 10.3389/fmicb.2022.886787.
- 859 11. Comeau, G., Zinna, R.A., Scott, T., Ernst, K., Walker, K., Carriere, Y., and Riehle, M.A. (2020).
860 Vertical Transmission of Zika Virus in *Aedes aegypti* Produces Potentially Infectious Progeny. *Am*
861 *J Trop Med Hyg* *103*, 876-883. 10.4269/ajtmh.19-0698.
- 862 12. Chaves, B.A., Junior, A.B.V., Silveira, K.R.D., Paz, A.D.C., Vaz, E., Araujo, R.G.P., Rodrigues, N.B.,
863 Campolina, T.B., Orfano, A.D.S., Nacif-Pimenta, R., et al. (2019). Vertical Transmission of Zika
864 Virus (Flaviviridae, Flavivirus) in Amazonian *Aedes aegypti* (Diptera: Culicidae) Delays Egg
865 Hatching and Larval Development of Progeny. *J Med Entomol* *56*, 1739-1744.
866 10.1093/jme/tjz110.
- 867 13. Kada, S., Paz-Bailey, G., Adams, L.E., and Johansson, M.A. (2024). Age-specific case data reveal
868 varying dengue transmission intensity in US states and territories. *PLoS Negl Trop Dis* *18*,
869 e0011143. 10.1371/journal.pntd.0011143.
- 870 14. Golding, M.A.J., Noble, S.A.A., Khouri, N.K., Layne-Yarde, R.N.A., Ali, I., and Sandiford, S.L.
871 (2023). Natural vertical transmission of dengue virus in Latin America and the Caribbean:
872 highlighting its detection limitations and potential significance. *Parasit Vectors* *16*, 442.
873 10.1186/s13071-023-06043-1.
- 874 15. Huang, E.Y.Y., Wong, A.Y.P., Lee, I.H.T., Qu, Z., Yip, H.Y., Leung, C.W., Yin, S.M., and Hui, J.H.L.
875 (2020). Infection patterns of dengue, Zika and endosymbiont *Wolbachia* in the mosquito *Aedes*
876 *albopictus* in Hong Kong. *Parasit Vectors* *13*, 361. 10.1186/s13071-020-04231-x.

- 877 16. Flanagan, M.L., Parrish, C.R., Cobey, S., Glass, G.E., Bush, R.M., and Leighton, T.J. (2012).
878 Anticipating the species jump: surveillance for emerging viral threats. *Zoonoses Public Health*
879 *59*, 155-163. 10.1111/j.1863-2378.2011.01439.x.
- 880 17. Hanley, K.A., Cecilia, H., Azar, S.R., Moehn, B.A., Gass, J.T., Oliveira da Silva, N.I., Yu, W., Yun, R.,
881 Althouse, B.M., Vasilakis, N., and Rossi, S.L. (2024). Trade-offs shaping transmission of sylvatic
882 dengue and Zika viruses in monkey hosts. *Nature communications* *15*, 2682. 10.1038/s41467-
883 024-46810-x.
- 884 18. Olson, K.E., and Blair, C.D. (2015). Arbovirus-mosquito interactions: RNAi pathway. *Curr Opin*
885 *Virol* *15*, 119-126. 10.1016/j.coviro.2015.10.001.
- 886 19. Gamez, S., Srivastav, S., Akbari, O.S., and Lau, N.C. (2020). Diverse Defenses: A Perspective
887 Comparing Dipteran Piwi-piRNA Pathways. *Cells* *9*. 10.3390/cells9102180.
- 888 20. Varjak, M., Leggewie, M., and Schnettler, E. (2018). The antiviral piRNA response in mosquitoes?
889 *J Gen Virol* *99*, 1551-1562. 10.1099/jgv.0.001157.
- 890 21. Petit, M., Mongelli, V., Frangeul, L., Blanc, H., Jiggins, F., and Saleh, M.C. (2016). piRNA pathway
891 is not required for antiviral defense in *Drosophila melanogaster*. *Proceedings of the National*
892 *Academy of Sciences of the United States of America* *113*, E4218-4227.
893 10.1073/pnas.1607952113.
- 894 22. Olmo, R.P., Todjro, Y.M.H., Aguiar, E., de Almeida, J.P.P., Ferreira, F.V., Armache, J.N., de Faria,
895 I.J.S., Ferreira, A.G.A., Amadou, S.C.G., Silva, A.T.S., et al. (2023). Mosquito vector competence
896 for dengue is modulated by insect-specific viruses. *Nat Microbiol* *8*, 135-149. 10.1038/s41564-
897 022-01289-4.
- 898 23. Ma, Q., Srivastav, S.P., Gamez, S., Dayama, G., Feitosa-Suntheimer, F., Patterson, E.I., Johnson,
899 R.M., Matson, E.M., Gold, A.S., Brackney, D.E., et al. (2021). A mosquito small RNA genomics
900 resource reveals dynamic evolution and host responses to viruses and transposons. *Genome*
901 *research* *31*, 512-528. 10.1101/gr.265157.120.
- 902 24. Suzuki, Y., Frangeul, L., Dickson, L.B., Blanc, H., Verdier, Y., Vinh, J., Lambrechts, L., and Saleh,
903 M.C. (2017). Uncovering the Repertoire of Endogenous Flaviviral Elements in *Aedes* Mosquito
904 Genomes. *J Virol* *91*. 10.1128/JVI.00571-17.
- 905 25. Lewis, S.H., Quarles, K.A., Yang, Y., Tanguy, M., Frezal, L., Smith, S.A., Sharma, P.P., Cordaux, R.,
906 Gilbert, C., Giraud, I., et al. (2018). Pan-arthropod analysis reveals somatic piRNAs as an
907 ancestral defence against transposable elements. *Nat Ecol Evol* *2*, 174-181. 10.1038/s41559-
908 017-0403-4.
- 909 26. Dayama, G., Bulekova, K., and Lau, N.C. (2022). Extending and Running the Mosquito Small RNA
910 Genomics Resource Pipeline. *Methods in molecular biology* *2509*, 341-352. 10.1007/978-1-
911 0716-2380-0_20.
- 912 27. Volz, E. (2023). Fitness, growth and transmissibility of SARS-CoV-2 genetic variants. *Nature*
913 *reviews. Genetics* *24*, 724-734. 10.1038/s41576-023-00610-z.
- 914 28. Zheng, Y., Gao, S., Padmanabhan, C., Li, R., Galvez, M., Gutierrez, D., Fuentes, S., Ling, K.S.,
915 Kreuze, J., and Fei, Z. (2017). VirusDetect: An automated pipeline for efficient virus discovery
916 using deep sequencing of small RNAs. *Virology* *500*, 130-138. 10.1016/j.virol.2016.10.017.
- 917 29. Bennett, A.J., Bushmaker, T., Cameron, K., Ondzie, A., Niama, F.R., Parra, H.J., Mombouli, J.V.,
918 Olson, S.H., Munster, V.J., and Goldberg, T.L. (2019). Diverse RNA viruses of arthropod origin in
919 the blood of fruit bats suggest a link between bat and arthropod viromes. *Virology* *528*, 64-72.
920 10.1016/j.virol.2018.12.009.
- 921 30. Parry, R., James, M.E., and Asgari, S. (2021). Uncovering the Worldwide Diversity and Evolution
922 of the Virome of the Mosquitoes *Aedes aegypti* and *Aedes albopictus*. *Microorganisms* *9*.
923 10.3390/microorganisms9081653.
- 924 31. Goertz, G.P., van Bree, J.W.M., Hiralal, A., Fernhout, B.M., Steffens, C., Boeren, S., Visser, T.M.,
925 Vogels, C.B.F., Abbo, S.R., Fros, J.J., et al. (2019). Subgenomic flavivirus RNA binds the mosquito
926 DEAD/H-box helicase ME31B and determines Zika virus transmission by *Aedes aegypti*.

- 927 Proceedings of the National Academy of Sciences of the United States of America *116*, 19136-
928 19144. 10.1073/pnas.1905617116.
- 929 32. Cross, S.T., Maertens, B.L., Dunham, T.J., Rodgers, C.P., Brehm, A.L., Miller, M.R., Williams, A.M.,
930 Foy, B.D., and Stenglein, M.D. (2020). Partitiviruses Infecting *Drosophila melanogaster* and
931 *Aedes aegypti* Exhibit Efficient Biparental Vertical Transmission. *J Virol* *94*. 10.1128/JVI.01070-
932 20.
- 933 33. Nibert, M.L., Ghabrial, S.A., Maiss, E., Lesker, T., Vainio, E.J., Jiang, D., and Suzuki, N. (2014).
934 Taxonomic reorganization of family Partitiviridae and other recent progress in partitivirus
935 research. *Virus Res* *188*, 128-141. 10.1016/j.virusres.2014.04.007.
- 936 34. Scholthof, H.B. (2006). The Tombusvirus-encoded P19: from irrelevance to elegance. *Nat Rev*
937 *Microbiol* *4*, 405-411. 10.1038/nrmicro1395.
- 938 35. Zardini, A., Menegale, F., Gobbi, A., Manica, M., Guzzetta, G., d'Andrea, V., Marziano, V.,
939 Trentini, F., Montarsi, F., Caputo, B., et al. (2024). Estimating the potential risk of transmission of
940 arboviruses in the Americas and Europe: a modelling study. *Lancet Planet Health* *8*, e30-e40.
941 10.1016/S2542-5196(23)00252-8.
- 942 36. Fansiri, T., Fontaine, A., Diancourt, L., Caro, V., Thaisomboonsuk, B., Richardson, J.H., Jarman,
943 R.G., Ponlawat, A., and Lambrechts, L. (2013). Genetic mapping of specific interactions between
944 *Aedes aegypti* mosquitoes and dengue viruses. *PLoS genetics* *9*, e1003621.
945 10.1371/journal.pgen.1003621.
- 946 37. Samuel, G.H., Pohlenz, T., Dong, Y., Coskun, N., Adelman, Z.N., Dimopoulos, G., and Myles, K.M.
947 (2023). RNA interference is essential to modulating the pathogenesis of mosquito-borne viruses
948 in the yellow fever mosquito *Aedes aegypti*. *Proceedings of the National Academy of Sciences of*
949 *the United States of America* *120*, e2213701120. 10.1073/pnas.2213701120.
- 950 38. Aubry, F., Dabo, S., Manet, C., Filipovic, I., Rose, N.H., Miot, E.F., Martynow, D., Baidaliuk, A.,
951 Merklings, S.H., Dickson, L.B., et al. (2020). Enhanced Zika virus susceptibility of globally invasive
952 *Aedes aegypti* populations. *Science* *370*, 991-996. 10.1126/science.abd3663.
- 953 39. Rose, N.H., Sylla, M., Badolo, A., Lutomiah, J., Ayala, D., Aribodor, O.B., Ibe, N., Akorli, J., Otoo,
954 S., Mutebi, J.P., et al. (2020). Climate and Urbanization Drive Mosquito Preference for Humans.
955 *Curr Biol* *30*, 3570-3579 e3576. 10.1016/j.cub.2020.06.092.
- 956 40. McBride, C.S., Baier, F., Omondi, A.B., Spitzer, S.A., Lutomiah, J., Sang, R., Ignell, R., and Vosshall,
957 L.B. (2014). Evolution of mosquito preference for humans linked to an odorant receptor. *Nature*
958 *515*, 222-227. 10.1038/nature13964.
- 959 41. Scott, J.C., Brackney, D.E., Campbell, C.L., Bondu-Hawkins, V., Hjelle, B., Ebel, G.D., Olson, K.E.,
960 and Blair, C.D. (2010). Comparison of dengue virus type 2-specific small RNAs from RNA
961 interference-competent and -incompetent mosquito cells. *PLoS Negl Trop Dis* *4*, e848.
962 10.1371/journal.pntd.0000848.
- 963 42. Brackney, D.E., Scott, J.C., Sagawa, F., Woodward, J.E., Miller, N.A., Schilkey, F.D., Mudge, J.,
964 Wilusz, J., Olson, K.E., Blair, C.D., and Ebel, G.D. (2010). C6/36 *Aedes albopictus* cells have a
965 dysfunctional antiviral RNA interference response. *PLoS Negl Trop Dis* *4*, e856.
966 10.1371/journal.pntd.0000856.
- 967 43. Goertz, G.P., Miesen, P., Overheul, G.J., van Rij, R.P., van Oers, M.M., and Pijlman, G.P. (2019).
968 Mosquito Small RNA Responses to West Nile and Insect-Specific Virus Infections in *Aedes* and
969 *Culex* Mosquito Cells. *Viruses* *11*. 10.3390/v11030271.
- 970 44. Dietrich, I., Shi, X., McFarlane, M., Watson, M., Blomstrom, A.L., Skelton, J.K., Kohl, A., Elliott,
971 R.M., and Schnettler, E. (2017). The Antiviral RNAi Response in Vector and Non-vector Cells
972 against Orthobunyaviruses. *PLoS Negl Trop Dis* *11*, e0005272. 10.1371/journal.pntd.0005272.
- 973 45. Weger-Lucarelli, J., Ruckert, C., Grubaugh, N.D., Misencik, M.J., Armstrong, P.M., Stenglein,
974 M.D., Ebel, G.D., and Brackney, D.E. (2018). Adventitious viruses persistently infect three
975 commonly used mosquito cell lines. *Virology* *521*, 175-180. 10.1016/j.virol.2018.06.007.

- 976 46. Baidaliuk, A., Miot, E.F., Lequime, S., Moltini-Conclois, I., Delaigue, F., Dabo, S., Dickson, L.B.,
977 Aubry, F., Merklings, S.H., Cao-Lormeau, V.M., and Lambrechts, L. (2019). Cell-Fusing Agent Virus
978 Reduces Arbovirus Dissemination in *Aedes aegypti* Mosquitoes In Vivo. *J Virol* 93.
979 10.1128/JVI.00705-19.
- 980 47. Johnson, R.M., and Rasgon, J.L. (2018). Densonucleosis viruses ('densovirus') for mosquito and
981 pathogen control. *Current opinion in insect science* 28, 90-97. 10.1016/j.cois.2018.05.009.
- 982 48. Suzuki, Y., Barik, T.K., Johnson, R.M., and Rasgon, J.L. (2015). In vitro and in vivo host range of
983 *Anopheles gambiae* densovirus (AgDENV). *Sci Rep* 5, 12701. 10.1038/srep12701.
- 984 49. Travanty, E.A., Adelman, Z.N., Franz, A.W., Keene, K.M., Beaty, B.J., Blair, C.D., James, A.A., and
985 Olson, K.E. (2004). Using RNA interference to develop dengue virus resistance in genetically
986 modified *Aedes aegypti*. *Insect Biochem Mol Biol* 34, 607-613. 10.1016/j.ibmb.2004.03.013.
- 987 50. Sanchez-Vargas, I., Scott, J.C., Poole-Smith, B.K., Franz, A.W., Barbosa-Solomieu, V., Wilusz, J.,
988 Olson, K.E., and Blair, C.D. (2009). Dengue virus type 2 infections of *Aedes aegypti* are
989 modulated by the mosquito's RNA interference pathway. *PLoS Pathog* 5, e1000299.
990 10.1371/journal.ppat.1000299.
- 991 51. Lopez, S.B.G., Guimaraes-Ribeiro, V., Rodriguez, J.V.G., Dorand, F., Salles, T.S., Sa-Guimaraes,
992 T.E., Alvarenga, E.S.L., Melo, A.C.A., Almeida, R.V., and Moreira, M.F. (2019). RNAi-based
993 bioinsecticide for *Aedes* mosquito control. *Sci Rep* 9, 4038. 10.1038/s41598-019-39666-5.
- 994 52. Magalhaes, T., Bergren, N.A., Bennett, S.L., Borland, E.M., Hartman, D.A., Lymperopoulos, K.,
995 Sayre, R., Borlee, B.R., Campbell, C.L., Foy, B.D., et al. (2019). Induction of RNA interference to
996 block Zika virus replication and transmission in the mosquito *Aedes aegypti*. *Insect Biochem Mol*
997 *Biol* 111, 103169. 10.1016/j.ibmb.2019.05.004.
- 998 53. Adelman, Z.N., Anderson, M.A., Morazzani, E.M., and Myles, K.M. (2008). A transgenic sensor
999 strain for monitoring the RNAi pathway in the yellow fever mosquito, *Aedes aegypti*. *Insect*
1000 *Biochem Mol Biol* 38, 705-713. 10.1016/j.ibmb.2008.04.002.
- 1001 54. Franz, A.W., Sanchez-Vargas, I., Adelman, Z.N., Blair, C.D., Beaty, B.J., James, A.A., and Olson,
1002 K.E. (2006). Engineering RNA interference-based resistance to dengue virus type 2 in genetically
1003 modified *Aedes aegypti*. *Proceedings of the National Academy of Sciences of the United States*
1004 *of America* 103, 4198-4203. 10.1073/pnas.0600479103.
- 1005 55. Romoli, O., Henrion-Lacritick, A., Blanc, H., Frangeul, L., and Saleh, M.C. (2024). Limitations in
1006 harnessing oral RNA interference as an antiviral strategy in *Aedes aegypti*. *iScience* 27, 109261.
1007 10.1016/j.isci.2024.109261.
- 1008



C

Region	Unique locations/labs	Lau lab generated samples	Public datasets	Small RNA datasets	Long RNA datasets
Laboratories	9	18	38	56	14
North America	18	48	4	52	13
Central & South America	11	20	23	43	4
Asia	8	14	4	18	14
Africa	17	30	10	40	30
Total	63	130	79	209	75

Figure 1. Overview of the global mosquito small RNA survey to discover RNA-interference (RNAi) responses to mosquito viruses.

(A) Overview of the MSRG pipeline applied to a global survey of whole mosquitoes from Americas, Asia, Africa and laboratory strains. (B) Implementation of the VirusDetect program with updated GBVRL and custom databases for comprehensive mosquito virus detection. (C) Summary tabulation of the samples and RNA libraries analyzed in this study.

Figure 2.

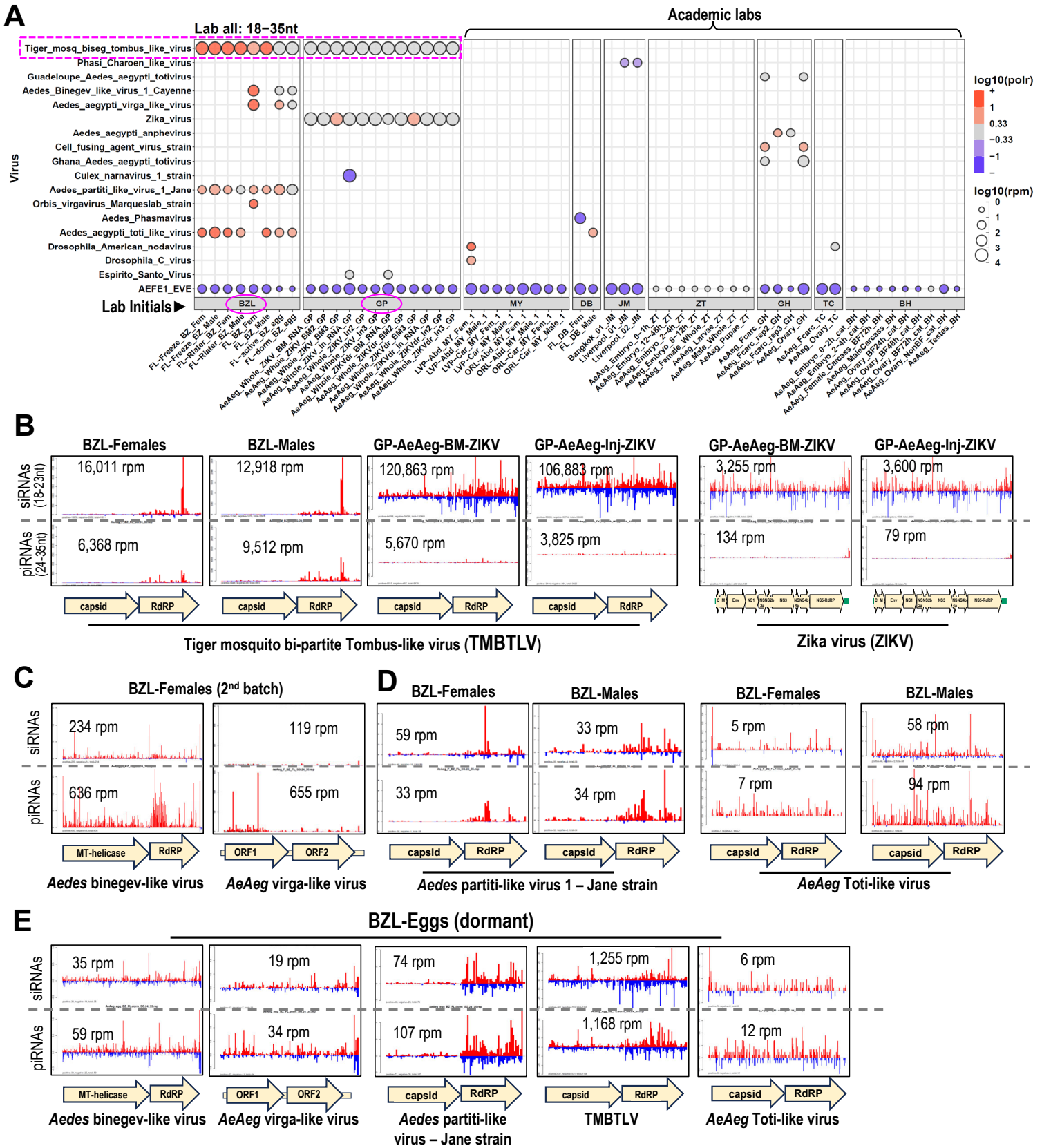


Figure 2. In contrast to commercial *Aedes aegypti* lab strains, most academic lab strains are low in viral small RNAs indicative of persistent viruses. (A) Bubble plot of viral small RNAs from lab strains. Number reads per million is reflected by bubble diameter, and color represents strand bias of reads, red is plus strand biased, blue is minus strand biased. Magenta circles and dashed boxes mark the BZL and GP samples that are commercial lab strains of *Ae. aegypti* and both are infected by Tiger mosquito bi-partite Tombus-like virus (TMBTLV). All the other *Ae. aegypti* samples were reared in academic labs. See metadata in Table S1 for sample details. Lab initials: BZL = Benzon Research, MY = Meg Younger, DB = Doug Brackney, JM = João Marques, ZT = Zhejian Tu, GH = Grant Hughes, TC = Tonya Colpitts, BH = Bruce Hay, GP = Gorben Pijlman (B) Coverage plots of TMBTLV and Zika virus (ZIKV) small RNAs from the two commercial lab strains. Additional coverage plots of (C) two viruses that only generated viral small RNAs in a 2nd batch of female BZL *Ae. aegypti* and (D) other viruses always present in the BZL strain in both females and males. (E) Viral small RNAs and likely persistent viruses are detected in the BZL eggs at lower small RNA levels with a more distinct pattern of antisense viral small RNAs in eggs compared to the parental whole females in (B).

Gupta, Sharma, et al.
Figure 4.

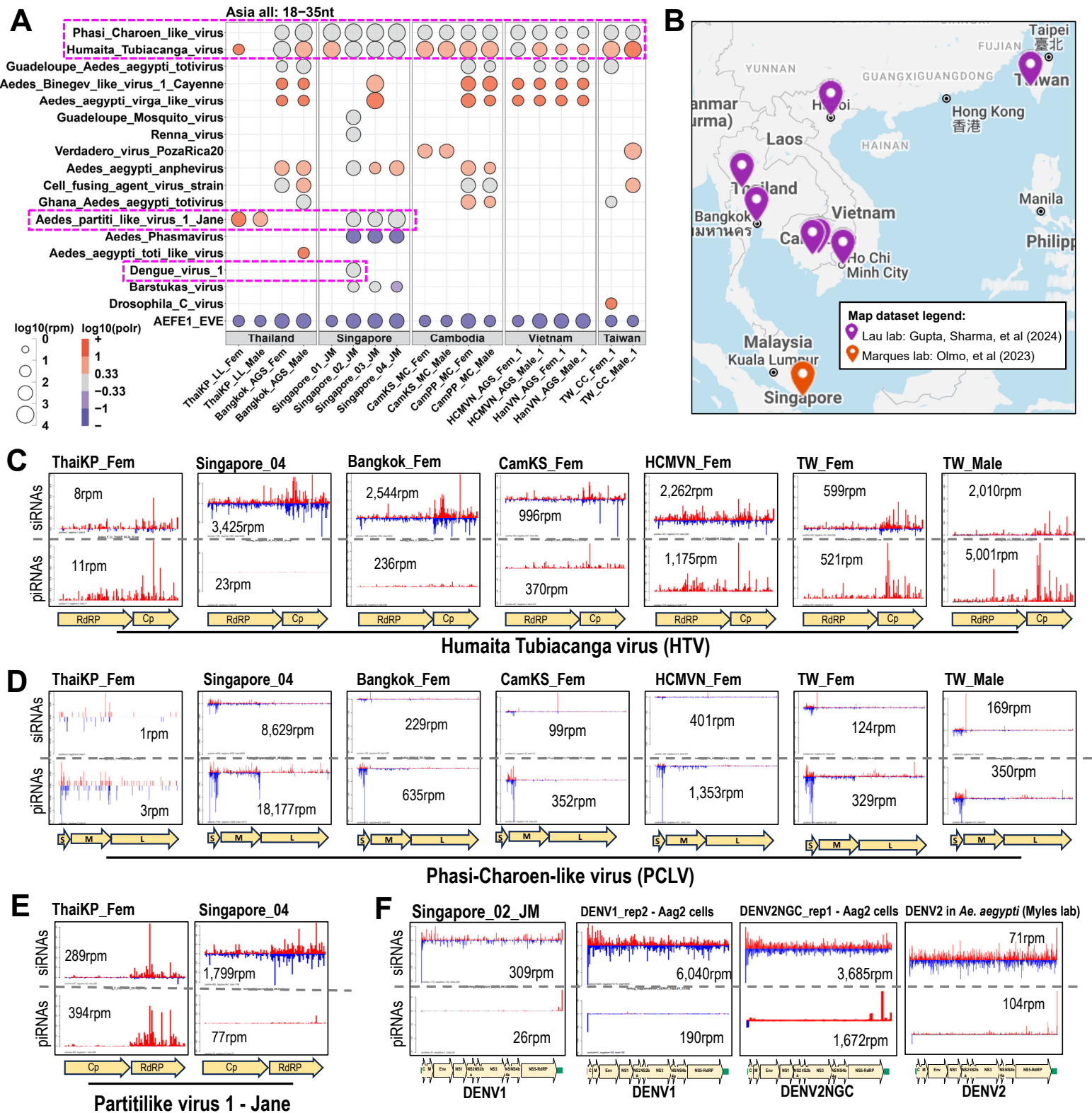


Figure 4. Asian *Ae. aegypti* viral small RNA patterns reveal a first case of dengue viral small RNAs from a wild isolate and share patterns with mosquitoes from the Americas.

(A) Bubble plot of viral small RNAs from Asia *Ae. aegypti* samples. Number reads per million is reflected by bubble diameter, and color represents strand bias of reads, red is plus strand biased, blue is minus strand biased. Dashed pink line box mark notable viruses whose coverage plots are displayed below. (B) Map of Asia locations where the *Ae. aegypti* samples originated. Coverage plots of HTV (C), PCLV (D) and Partiti-like virus 1 Jane strain (E) viral small RNAs from a selection of Asian *Ae. aegypti*. (F) The DENV viral small RNA coverage from a Singapore isolate compared to other DENV viral small patterns from infections in Aag2 cells and a lab injected infection of *Ae. aegypti* by the Myles lab.

Gupta, Sharma, et al.
Figure 5.

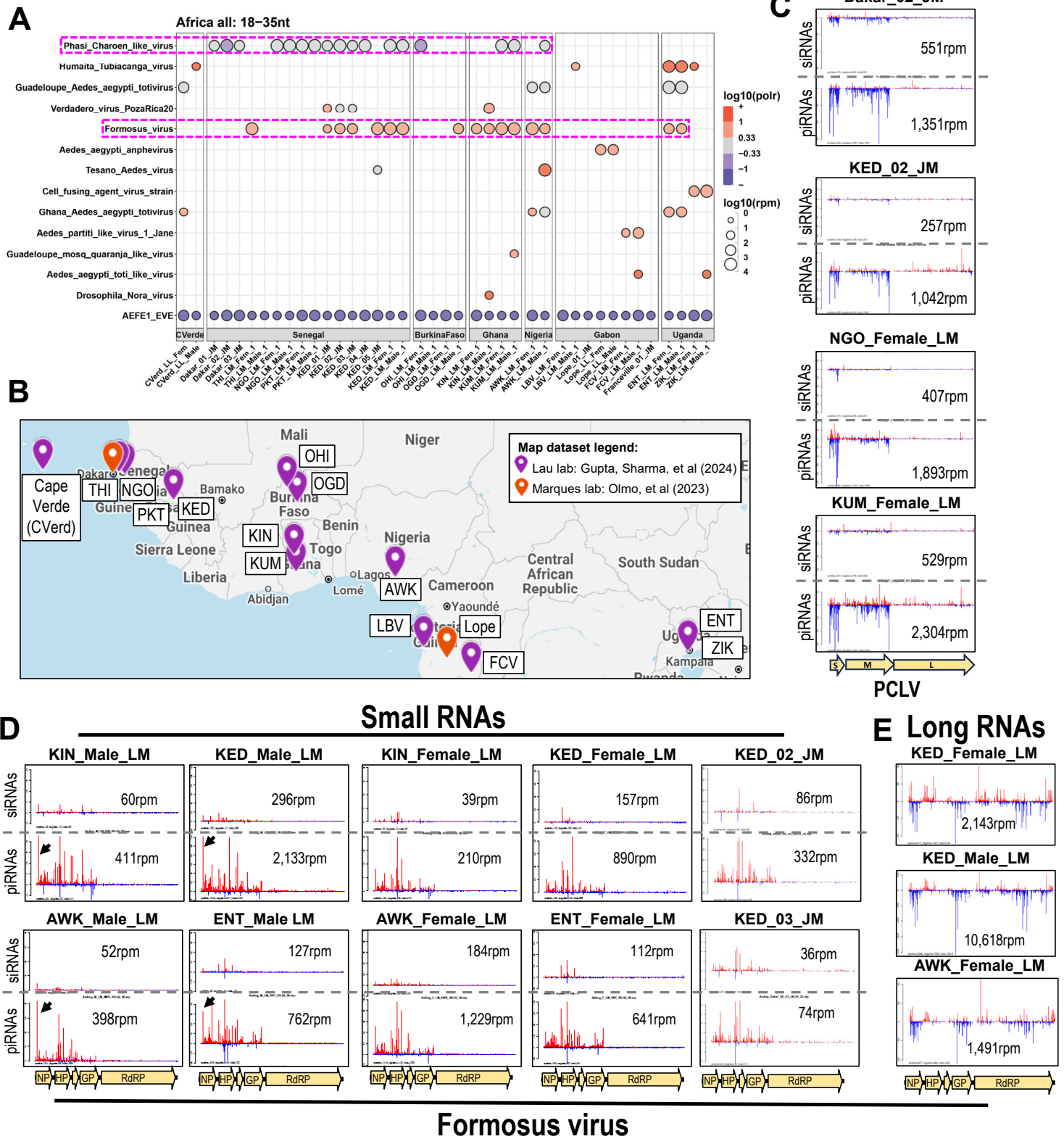


Figure 5. African *Ae. aegypti* colony strains carry viral small RNAs unique to their continent.

(A) Bubble plot of viral small RNAs from African *Ae. aegypti* colony strains. Number reads per million is reflected by bubble diameter, and color represents strand bias of reads, red is plus strand biased, blue is minus strand biased. Dashed pink line boxes mark the PCLV and Formosus viruses noted in panels (C) and (D), respectively. (B) Map of Africa locations where the *Ae. aegypti* colonies or samples originated. (C) Coverage plots of PCLV small RNAs from a selection of African *Ae. aegypti* showing high M-fragment piRNAs rivaling the S-fragment piRNAs. (D) The Formosus virus viral small RNA coverage from African *Ae. aegypti* colonies from the McBride lab and an independent Kedougou, Senegal sample from Olmo et al 2023. The black arrow points to male-specific viral piRNA species. (E) Three examples of Formosus virus long RNAs sequenced from matched samples in (D).

Gupta, Sharma, et al.
Figure 6.

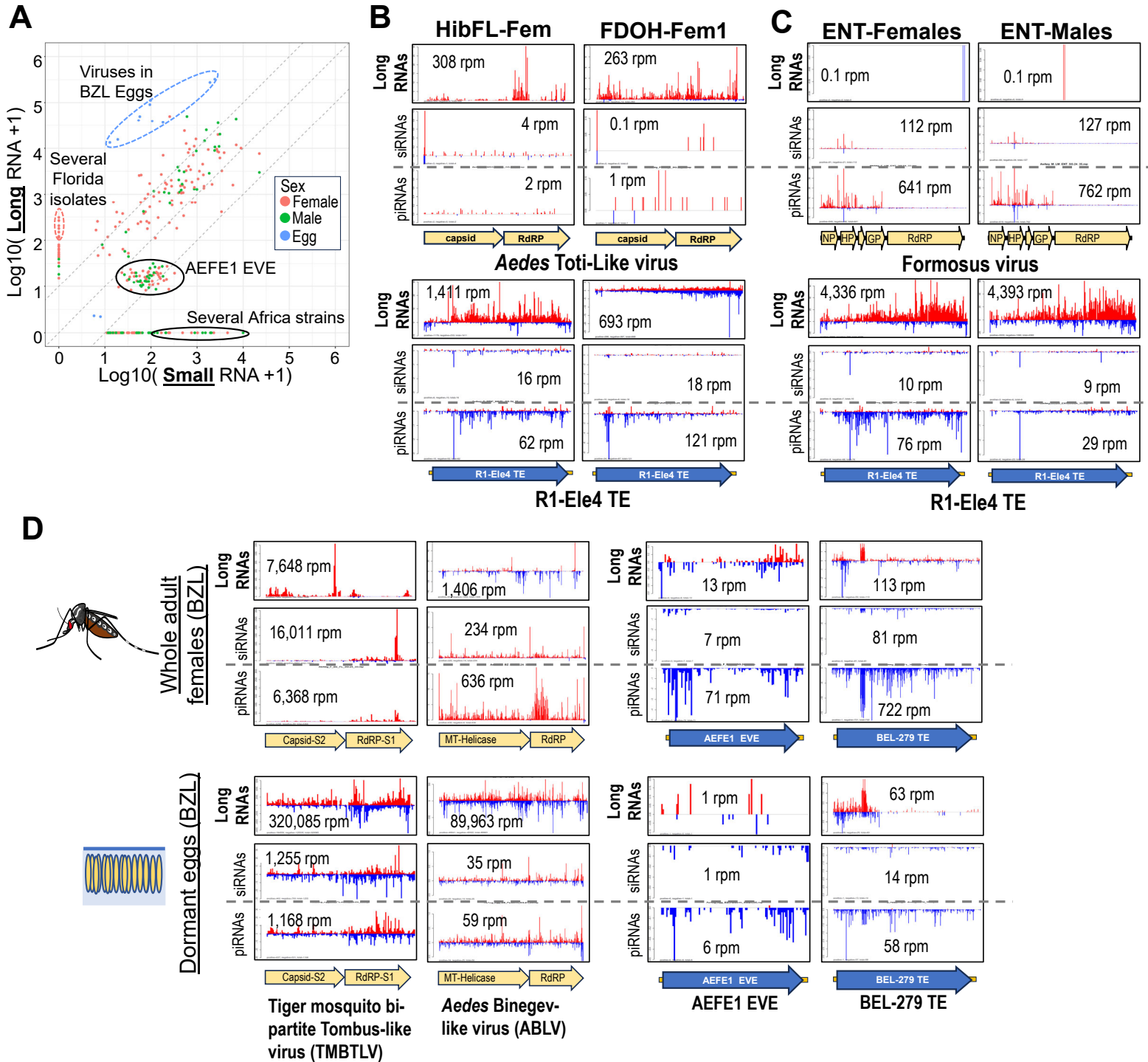


Figure 6. Virus transmission and replication dynamics suggested by long RNA sequencing compared to matched small RNAs of mosquito viruses and transposable elements (TEs).

(A) Scatterplot comparing the particular *Ae. aegypti* samples that enabled matched library construction of small RNAs and long RNAs. Sequencing reads per million with a pseudo-count of 1 are plotted on the logarithmic scales. Dots representing the matched samples are colored by sex, and groups of samples clustering together are noted in the labeled ovals. (B) Coverage plots of two Florida isolates of *Ae. aegypti* exhibiting abundant long RNA signal for the totivirus but negligible viral small RNA in the upper plots that contrast both long and small RNAs against an R1-Ele4 TE. (C) Coverage plots of long RNAs versus small RNAs for the Formosus virus and R1-Ele4 TE from both males and females of the ENT African colony. (D) Coverage plots of long RNAs compared to small RNAs for viruses and a TE from the BZL strain of *Ae. aegypti* females and dormant eggs.

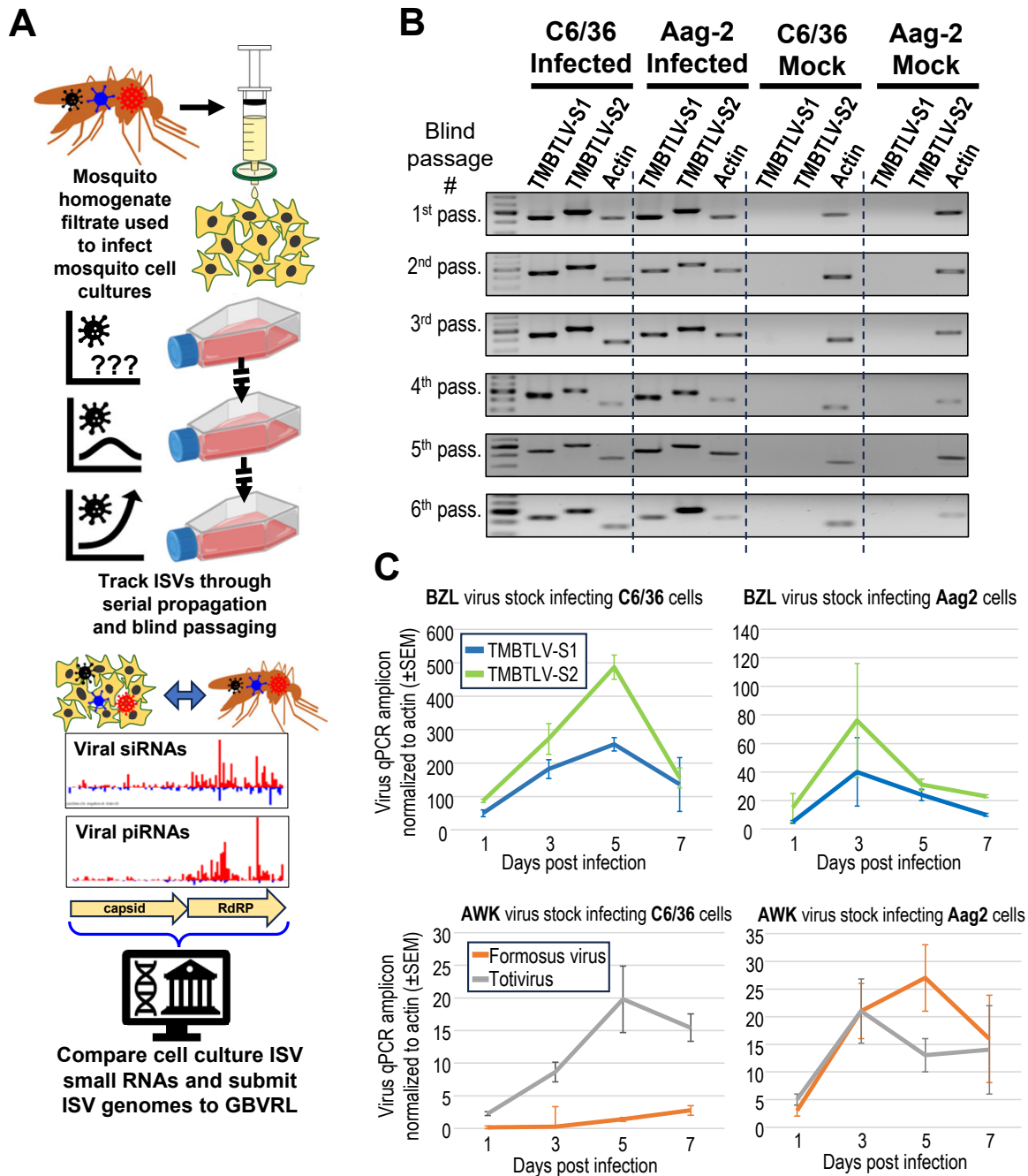


Figure 7. Novel mosquito viruses infecting and replicating in mosquito cells.

(A) Our methodology to molecularly validate the small RNA detection of ISVs are true viruses that can be isolated and verified for triggering the RNAi response in mosquito cells. (B) RT-PCR detection of TMBTLV RNAs S1 and S2 during multiple rounds of blind passaging, starting with BZL mosquito homogenate as the initial virus infection source placed onto C6/36 and Aag2 mosquito cells. (C) Virus replication kinetics measured in C6/36 and Aag2 cells over 1 week. Flasks of 1 million cells were infected on Day 0 with 200 viral qPCR amplicons per infection. Virus stocks are from filtered media from subsequent passage from the experiment in (A). Error bars correspond to the standard error from measuring three biological replicates of infection of the mosquito cells with the same virus stock.

Figure 8.

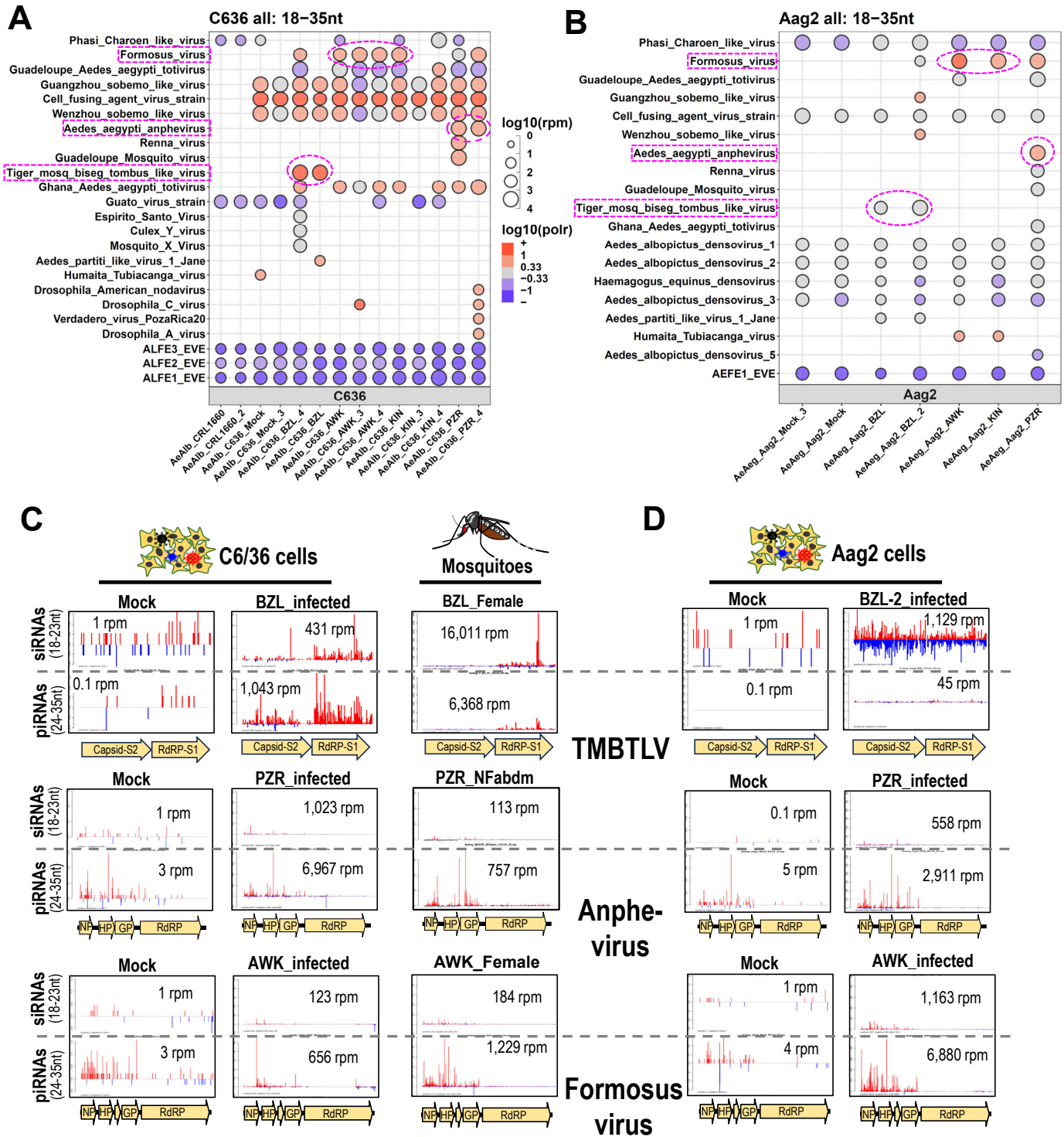


Figure 8. Mosquito viruses isolated in cell culture trigger an RNAi response of viral small RNA patterns that resemble the patterns from whole mosquitoes.

Bubble plots of viral small RNAs sequenced from (A) *Ae. albopictus* C6/36 cells and (B) *Ae. aegypti* Aag2 cells, either mock or stably infected after multiple blind passages of virus stocks from mosquito homogenates. Number reads per million is reflected by bubble diameter, and color represents strand bias of reads, red is plus strand biased, blue is minus strand biased. Magenta boxes and dashed circles in (A) and (B) highlight specific samples inspected in coverage plots in (C) and (D). Coverage plots for three viruses, TMBTLV, Anphevirus and Formosus virus in (C) C6/36 cells and (D) Aag2 cells with the intact mosquito from females shown in the middle for comparison.