SUPPLEMENTARY INFORMATION FILE FOR GUPTA, SHARMA ET AL.
NELSON LAU LAB 2026 NATURE COMMUNICATIONS.

Small RNA genomics of *Aedes aegypti* mosquitoes discover infectious viruses that trigger an RNA interference response

TABLE OF CONTENTS:

SUPPLEMENTARY TEXT DISCUSSION FOR GUPTA, SHARMA ET AL.

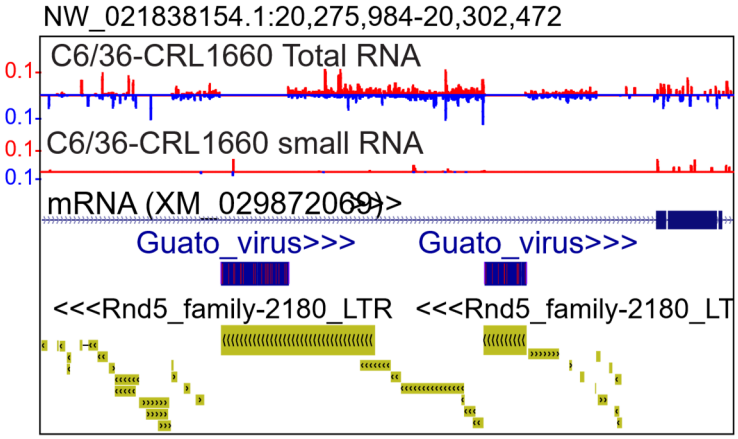
Small RNA genomics of *Aedes aegypti* mosquitoes discover infectious viruses that trigger an RNA interference response

The issue of discerning bona fide viruses from integrated sequences is a complex genomics problem that many other previous studies have focused on detecting these integrated virus elements from genomic DNA sequencing[1-8]. These studies conducted genomic DNA extraction and Whole Genome Sequencing (WGS) or targeted genomic DNA (gDNA) sequencing followed by structural variants analysis of sequencing reads. This approach is non-trivial in effort and sequencing cost, so it is beyond the scope of our study which instead focuses on small RNA transcriptome analysis of viruses.

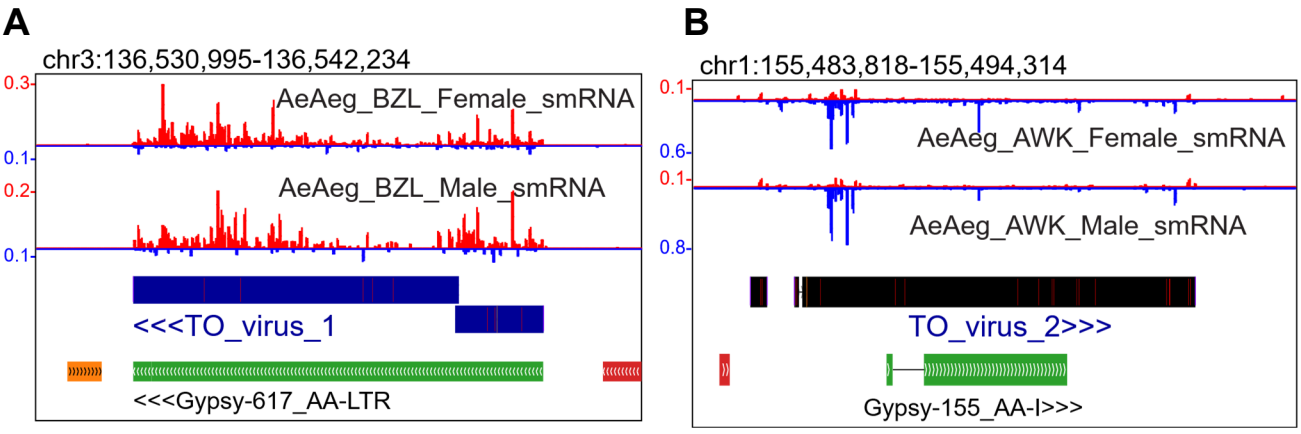
Our study here focuses on transcriptome analyses to discover new infectious mosquito viruses with a demonstrated ability to trigger a small RNA response. The reason why detecting viral sequences integrated into these mosquito sample genomes is beyond the scope of this study is because: (1) there is the biochemical limitation from limited mosquito material that only RNA extraction or gDNA extraction can be done, but not both simultaneously. It remains an ongoing problem that no commercial product can truly optimize extracting both RNA and gDNA from the same sample. (2) The large (2.8GB) *Ae. aegypti* genome also does not afford the economic limitation of our extensive RNA sequencing to include significant WGS to be added to this study.

Some of our VirusDetect runs with small RNA library inputs occasionally detected a putative virus entry that we then manually removed from the MSRG Virus List (see Supplementary Data Table S3) as misannotated entries. The example removal process is illustrated by when our VirusDetect runs initially proposed the detection of a Guato virus Glycoprotein (accession: KT9664840) or an *Aedes* To virus (accessions: MT913595, MT9135967). When we performed BLASTN and BLAT analyses, we observed a perfect match of the Guato virus glycoprotein entry to multiple loci in the *Ae. albopictus* genome (See **Supplementary Discussion Fig. SDF-1**), and perfect matches of the To virus to multiple loci in the *Ae. aegypti* genome where RepeatMasker had annotated a GYPSY-155-AA-LTR transposon (See **Supplementary Discussion Fig. SDF-2**). Because these BLAST and BLAT matches almost perfectly match to repeat elements in these *Aedes* genomes, we can conclude these are likely retroviral transposons generating abundant piRNAs while also being misannotated as a putative transmitted virus.

Supplementary Discussion Fig. SDF-1. A nearly perfect alignment of putative Guato virus GP entry to one *Ae. albopictus* genomic locus that generates piRNA generating locus. There are 165 other genomic loci in the *Ae. albopictus* genome with nearly as extensive of an alignment to this putative Guato virus sequence.

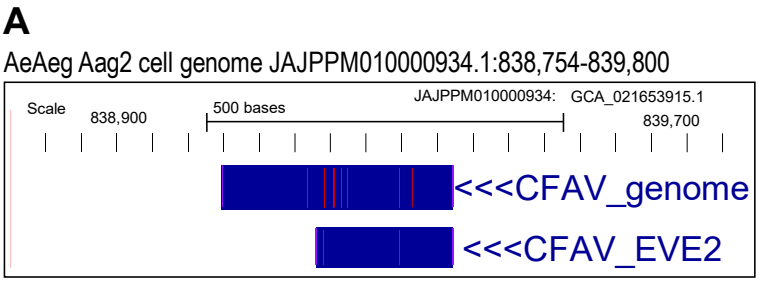


Supplementary Discussion Fig. SDF-2. Nearly perfect alignments of putative TO virus entry to GYPSY transposon loci and transposon piRNAs from *Ae. aegypti*. (A) and (B) are two distinct loci in the *Ae. aegypti* genome where these two entries are likely mis-annotating a retroviral transposon as a virus. There are 19 and 61 other genomic loci with as strong an alignment to the putative TO-1 and TO-2 virus entries, respectively.

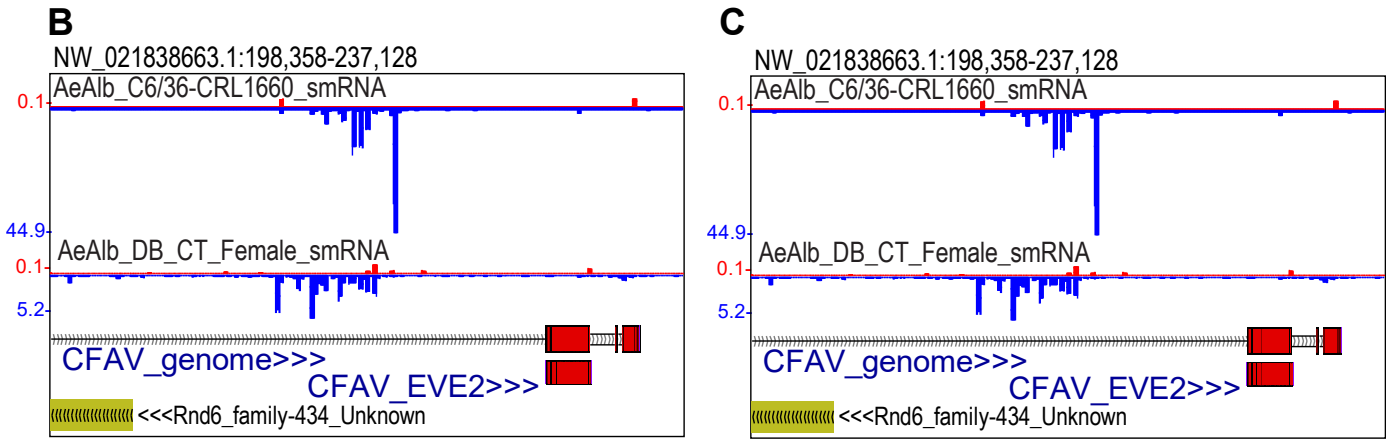


Regarding our analysis of *Ae. aegypti* EVEs, we performed a standard BLAT and BLAST approach similar to other studies[1,4] to examine what parts of these viruses and their small RNAs could be attributed or potentially confounded by integrated viral elements (also called as EVEs). This approach confirmed the identification of the CFAV EVEs from the Aag2 genome and the *Ae. albopictus* genome showing the clearest demonstration of piRNAs emanating from the CFAV EVEs. See **Supplementary Discussion Fig. SDF-3** for these confirmation results. While this approach also readily detected *Ae. aegypti* EVEs against *Aedes* Anphevirus, Aedes Totivirus, Formosus virus (Extended Data Fig. S7e, S7f), repeated BLAT and BLAST tests across multiple genome assemblies failed to detect any EVEs against the Tiger Mosquito Bi-partite Tombus-like virus (TMBTLV).

Supplementary Discussion Fig. SDF-3: CFAV EVE BLAT alignments to *Ae. aegpyti* and *Ae. albopictus* genomic loci. Red lines are sequence mismatches in the BLAT alignment. (A) The CFAV genome and the CFAV EVE2 was readily aligned to the Ae. aegypti Aag2 cell line genome that is only a draft assembly lacking chromosome anchoring coordinates. (B) and (C) are two distinct loci in the *Ae. albopictus* genome that represent EVEs against the CFAV



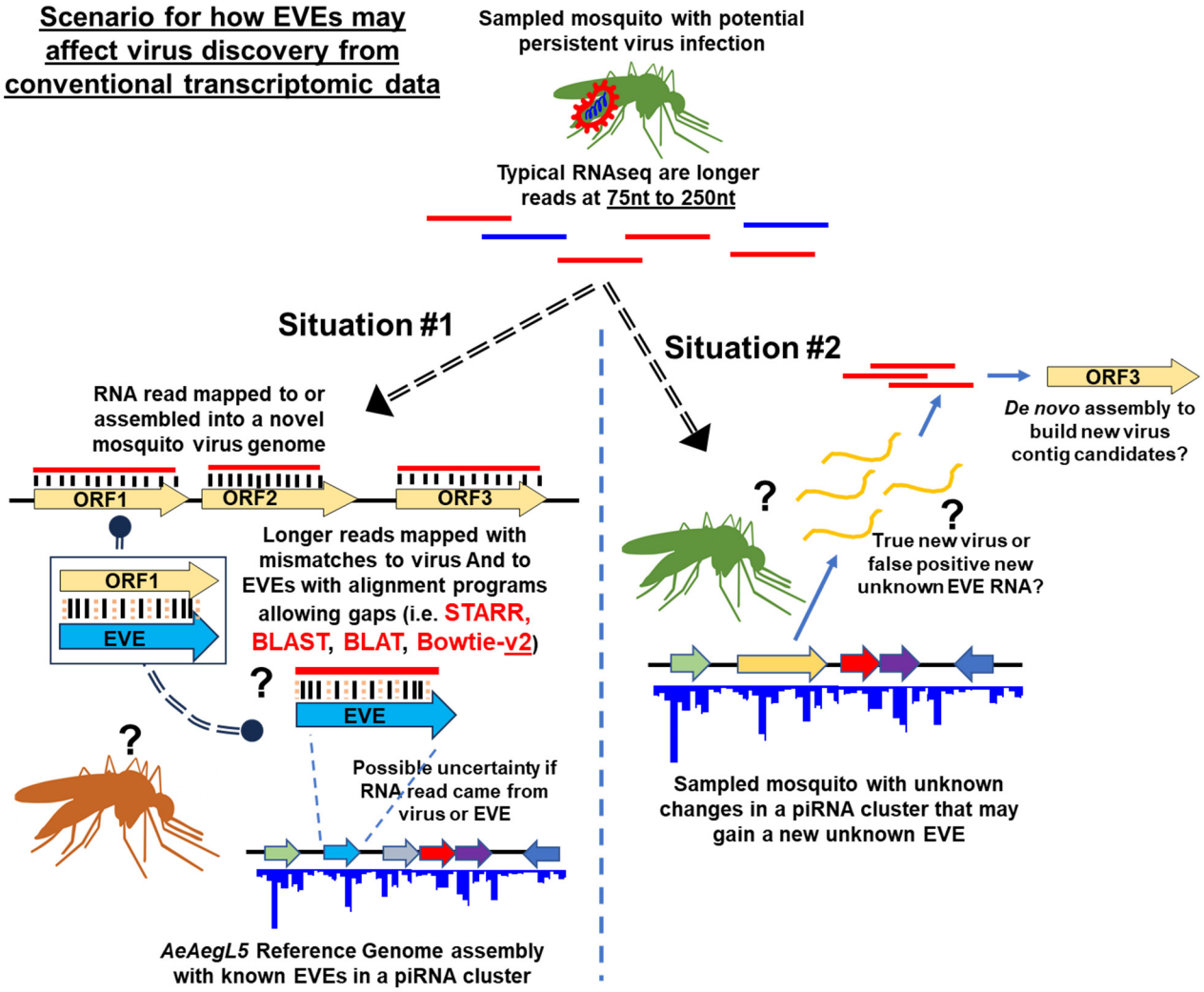
Supplementary Discussion Fig. SDF-3.



Although previous studies have suggested that virus discovery from RNAseq reads could be confounded by variability in EVE presence and expression in *Aedes* mosquitoes[6,9-12], we were mindful of this concern when raised against our virus discovery approach using mosquito small RNA sequencing. The potential confusion between mapping a longer, standard RNAseq read to both a virus genome and to an EVE transcript is because the standard RNAseq alignment programs allow sufficient gaps and mismatches to still yield strong statistical significance in the mapping result to both targets (i.e. Bowtie v2, STARR, BWA and even BLAST and BLAT). This feature is what allows us and other investigators to take a true virus sequence and perform BLAST and BLAT searches to identify EVEs in the *Ae. aegypti* genome[1-8] and shown in Supplementary Discussion Fig. SDF-3 and in Extended Data Fig. S7e-S7g. We acknowledge the potential for confusion between discriminating EVE transcripts and novel viruses as we illustrate this issue in **Supplementary Discussion Fig. SDF-4**.

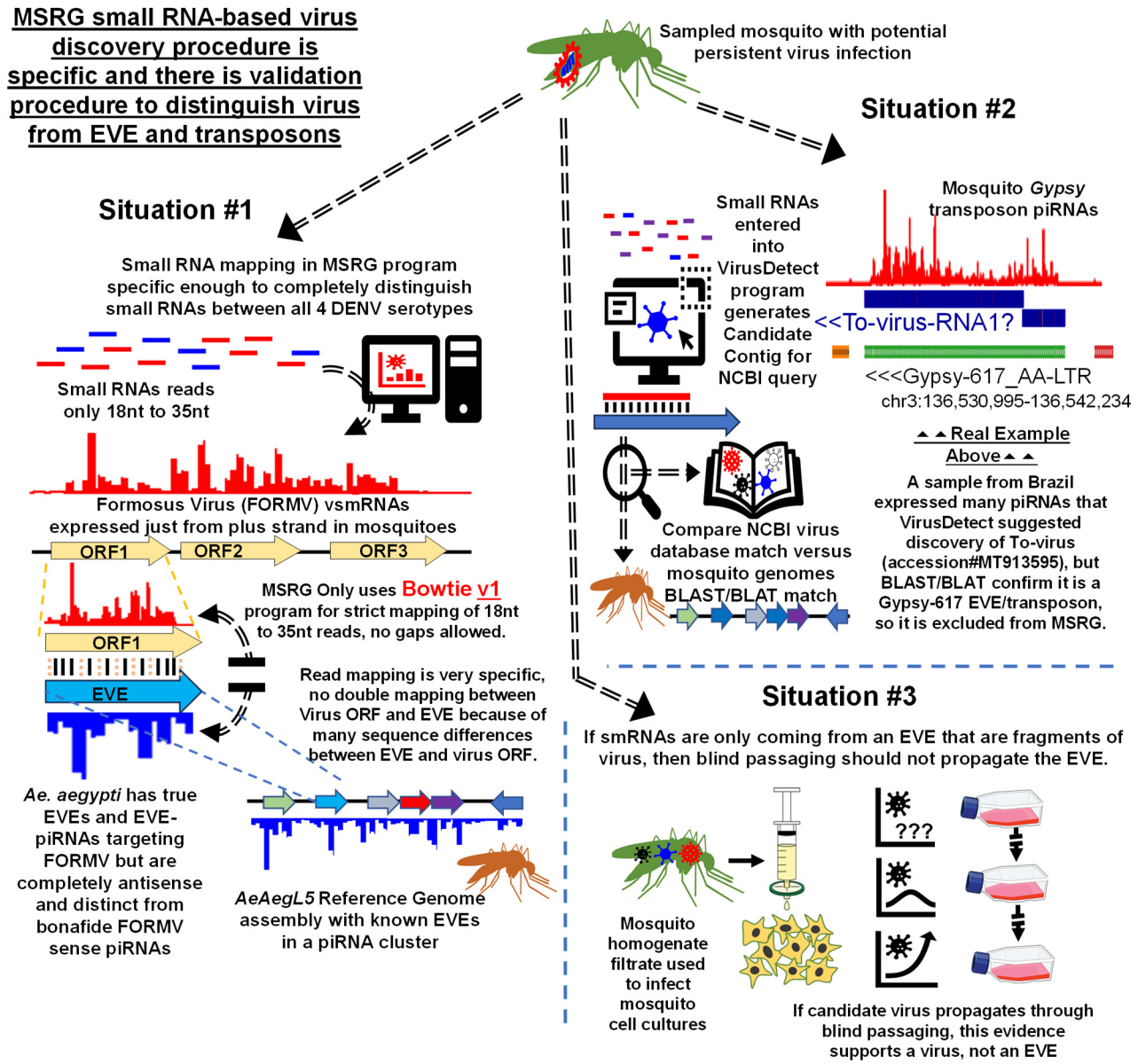
Supplementary Discussion Fig. SDF-4. Diagram explaining how conventional RNAseq analysis for virus discovery could be confounded by EVE/transposons transcripts. Scenario for how EVEs may affect virus discovery from conventional transcriptomic data. Situation #1 indicates when there is an EVE present in both the sampled mosquito represented in the Green silhouette and in the Reference genome sequence of Ae. aegypti, represented by the Brown silhouette. Situation #2 indicates when the confounding EVE in the sampled mosquito in the Green silhouette is not present and characterized in the Reference genome sequence. Note that in conventional transcriptomics data, longer reads mapped with BLAST, BLAT or Bowtie v2 can allow gaps and more mismatches that would contribute to confusion between EVEs and virus discovery.

Supplementary Discussion Fig. SDF-4. Diagram explaining how conventional RNAseq analysis for virus discovery could be confounded by EVE/transposons transcripts.



Our MSRG small RNA-based virus discovery processes are highly specific and there is validation procedure to distinguish viruses from EVE and transposons. Regarding Situation#2 above for when the presence or absence of a particular EVE is not knowable in the tested mosquito, in Fig. 5 and Fig. 6 of our experimental approach, we show that blind passaging to establish stable virus infection followed by small RNA sequencing of stably-infected cells demonstrates that discovered virus are bonafide infectious agents because an EVE that by definition are fragments of viruses would not be able form an infectious virus. This procedure is highlighted in Situation #3 of **Supplementary Discussion Fig. SDF-5.**

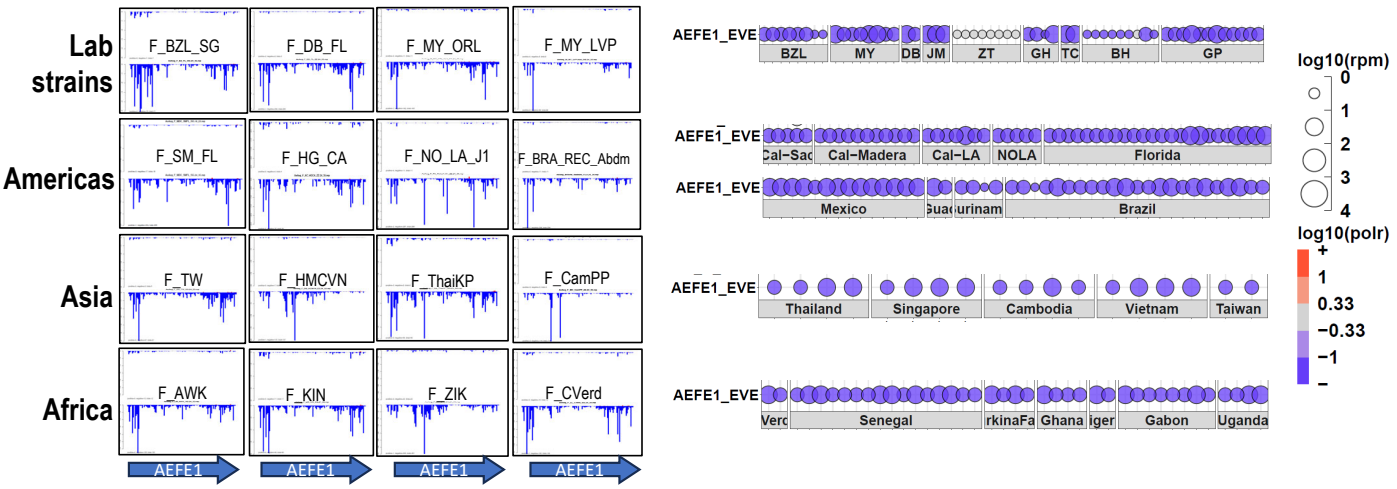
Supplementary Discussion Fig. SDF-5. Diagram explaining MSRG discrimination of bonafide virus discovery from confounding EVE/transposons.
Situation #1 indicates when there is an EVE present in both the sampled mosquito represented in the Green silhouette and in the Reference genome sequence of *Ae. aegypti*, represented by the Brown silhouette. Note that MSRG maps small RNA reads only with Bowtie v1 without gaps, so there is no confusion nor double mapping between EVEs and virus discovery. Situation #2 indicates when there is a confounding EVE in the sampled mosquito in the Green silhouette, we perform BLAST/BLAT queries with the Reference genome sequence and then remove the confounding entry from the MSRG pipeline. Situation #3 adapts the panel from Figure 5A of the blind passaging experiment that can confirm whether a candidate virus discovered by small RNA sequencing is a real virus which should propagate versus a hidden EVE which are just genomic fragments and unlikely to be transmitted from the mosquito animal to the mosquito cell culture. T-flask illustration from NIAID NIH BioArt Source (bioart.niaid.nih.gov/bioart/303).



References for Supplementary Text Discussion

1. Crava, C.M. *et al.* Population genomics in the arboviral vector Aedes aegypti reveals the genomic architecture and evolution of endogenous viral elements. *Mol Ecol* **30**, 1594-1611 (2021).
2. Havill, J., Strasburg, O., Udoh, T., Crawford, J.E. & Gloria-Soria, A. EVE-X: Software to Identify Novel Viral Insertions in Wild-Caught Arthropod Hosts From Next-Generation Short Read Data. *Mol Ecol Resour* **25**, e14026 (2025).
3. Tassetto, M. *et al.* Control of RNA viruses in mosquito cells through the acquisition of vDNA and endogenous viral elements. *Elife* **8**(2019).
4. Whitfield, Z.J. *et al.* The Diversity, Structure, and Function of Heritable Adaptive Immunity Sequences in the Aedes aegypti Genome. *Curr Biol* **27**, 3511-3519 e7 (2017).
5. Suzuki, Y. *et al.* Non-retroviral Endogenous Viral Element Limits Cognate Virus Replication in Aedes aegypti Ovaries. *Curr Biol* **30**, 3495-3506 e6 (2020).
6. Suzuki, Y. *et al.* Uncovering the Repertoire of Endogenous Flaviviral Elements in Aedes Mosquito Genomes. *J Virol* **91**(2017).
7. Palatini, U. *et al.* Comparative genomics shows that viral integrations are abundant and express piRNAs in the arboviral vectors Aedes aegypti and Aedes albopictus. *BMC Genomics* **18**, 512 (2017).
8. Crochu, S. *et al.* Sequences of flavivirus-related RNA viruses persist in DNA form integrated in the genome of Aedes spp. mosquitoes. *J Gen Virol* **85**, 1971-1980 (2004).
9. Zhao, M. *et al.* Genetic diversity of Flaviviridae and Rhabdoviridae EVEs in Aedes aegypti and Aedes albopictus on Hainan Island and the Leizhou Peninsula, China. *Infect Genet Evol* **123**, 105627 (2024).
10. Brait, N. *et al.* A tale of caution: How endogenous viral elements affect virus discovery in transcriptomic data. *Virus Evol* **10**, vead088 (2024).
11. Rozo-Lopez, P., Brewer, W., Kafer, S., Martin, M.M. & Parker, B.J. Untangling an insect's virome from its endogenous viral elements. *BMC Genomics* **24**, 636 (2023).
12. Ter Horst, A.M., Nigg, J.C., Dekker, F.M. & Falk, B.W. Endogenous Viral Elements Are Widespread in Arthropod Genomes and Commonly Give Rise to PIWI-Interacting RNAs. *J Virol* **93**(2019).

a

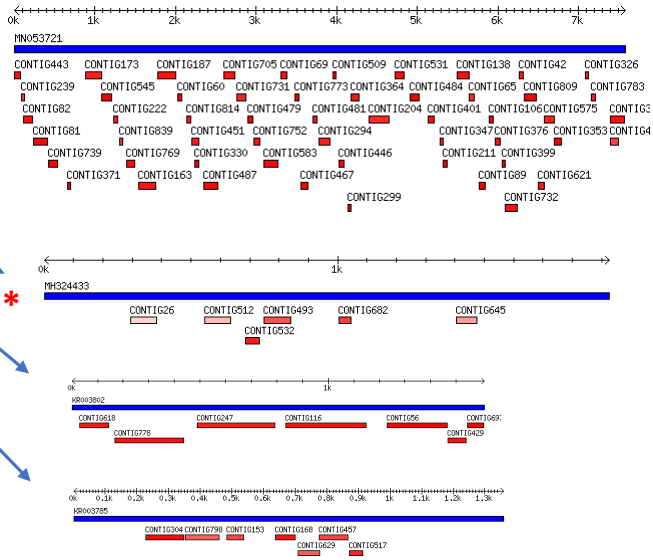


b

VirusDetect result from Genbank Virus Reference List (GBVRL)

Reference	Length	Coverage (%)	#contig	Depth	Depth (Norm)	%Identity	%Iden Max	%Iden Min	Genus	Description
MN053721	7582	5033 (66.4)	54	27.3	0.4	98.65	100	95.00	NA	Aedes aegypti toti-like virus strain 2017-Ab-AAF-1-5 capsid (orf1) and RNA-dependent RNA polymerase (orf2) genes, complete cds.
MH324433 *	1925	429 (22.3)	6	1066.8	17.4	89.30	97.92	83.15	NA	Hypsignathus monstrosus tombus-like virus 1 isolate HMTLV1-RML51 hypothetical protein 1 and hypothetical protein 2 genes, complete cds.
KR003802	1609	1372 (85.3)	7	8.8	0.1	98.69	100	97.70	NA	Humaita-Tubiacanga virus isolate Rio segment capsid Capsid gene, complete cds.
KR003785	1358	547 (40.3)	7	291.5	4.7	95.64	99.16	92.52	phasivirus	Phasi Charoen-like virus isolate Rio segment S Nucleocapsid gene, complete cds.
MZ560693	148	143 (96.6)	1	898.6	14.6	100	100	100	picobirnavirus	Picobirnavirus sp. isolate PBV/Cryptosporidium/NZL/NZ03_CR_S1/2019 RNA-dependent RNA polymerase (RdRp) gene, partial cds.
MT815904	391	194 (49.6)	3	733.6	11.9	92.31	93.75	91.23	pospiviroid	Citrus exocortis viroid isolate CEYVd OH19-2, partial sequence.
MN225577	927	223 (24.1)	1	202.5	3.3	97.31	97.31	97.31	vesiculovirus	UNVERIFIED: Guampa vesiculovirus RNA-dependent RNA polymerase protein-like gene, partial sequence.

c

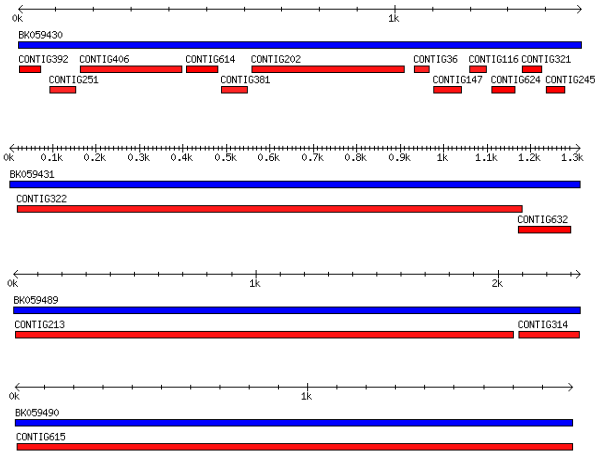


d

VirusDetect result from TPA and MAG databases

Reference	Length	Coverage (%)	#contig	Depth	Depth (Norm)	%Identity	%Iden Max	%Iden Min	Genus	Description
BK059430	1494	1270 (85)	12	23.8	0.4	98.58	100	97.10	TPA_asm: Aedes partiti-like virus 1 isolate Jane RNA-dependent RNA polymerase (RdRP) gene, complete cds	1
BK059431	1313	1275 (97.1)	2	129.5	2.1	98.21	100	98.02	TPA_asm: Aedes partiti-like virus 1 isolate Jane capsid protein gene, complete cds	1
BK059489	2338	2306 (98.6)	2	2774.2	45.2	98.40	98.49	97.61	TPA_asm: Tiger mosquito bi-segmented tombus-like virus isolate Phu Hoa HSP90-containing protein and RdRP genes, complete cds	1
BK059490	1903	1897 (99.7)	1	5922.4	96.5	98.68	98.68	98.68	TPA_asm: Tiger mosquito bi-segmented tombus-like virus isolate Phu Hoa hypothetical protein and putative structural protein genes, complete cds	1

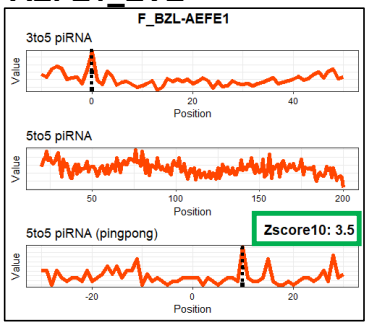
e

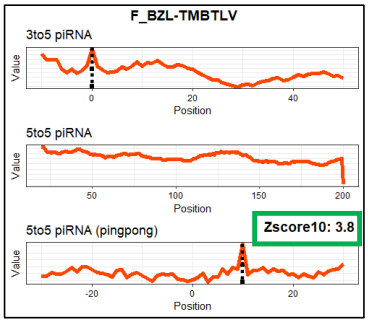


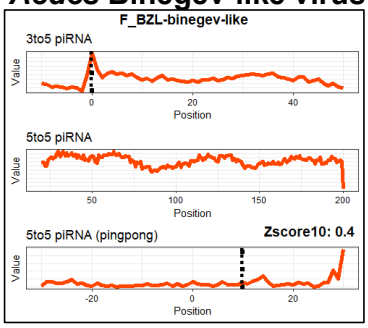
Supplementary Fig. S1. Using the AEFE1 EVE as a constant quality tracker control for vsmRNA profiling, implementing VirusDetect to discover new ISVs from small RNAs, and plotting piRNA biogenesis. (**a**) Coverage plots on left and bubble plots on right representing the small RNAs from the Endogenous Viral Element (EVE) from *Ae. aegypti* called AEFE1 (Suzuki et al, 2017)[25]. (**b**) The results output table from the VirusDetect program only loaded with the GBVRL database while analyzing the small RNA library from one of the BZL lab strain samples. (**c**) Coverage of the de-novo small RNA assembly contigs generated by VirusDetect of the corresponding virus determinations from the table in (b). The red asterisks mark a false-positive call against a tombus-like virus entry in GBVRL with limited contig coverage. (**d**) VirusDetect results analyzing the same BZL lab strain small RNA library in (b) but now with a custom loading of the TPA and MAG databases that are distinct from GBVRL. This result now shows two new viruses with much more complete contig coverage in (**e**), such as the Tiger Mosquito Bi-segmented Tombus-Like virus (TMBTLV) that is the true hit replacing the false-positive call in (b-c).(**f**) Autocorrelation analyses of phasing and ping-pong patterns from viral piRNAs in representative mosquito (**f.i-iii**) and cell (**f.iv**) samples. For each virus, top two plots are for 3'-to-5' phasing and bottom plot is of piRNA ping-pong patterns along with a Z10 score. Z10>3.0 denotes a strong ping-pong piRNA signature.

f

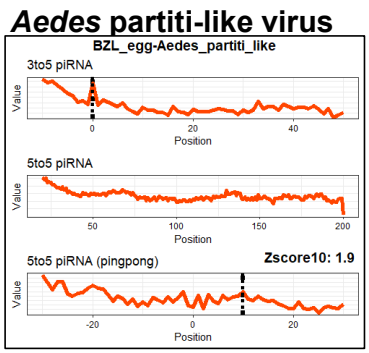
i **FL_BZ_Fem**

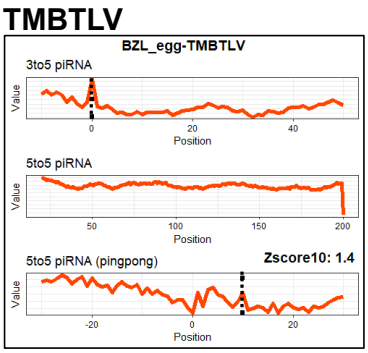


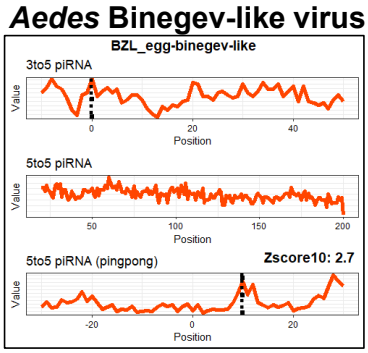




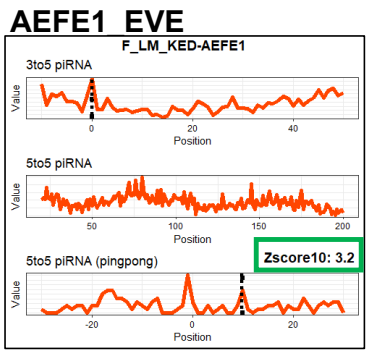
ii **FL-active_BZ_egg**

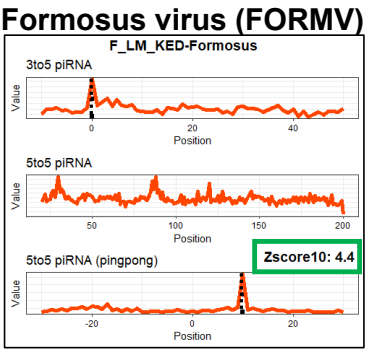


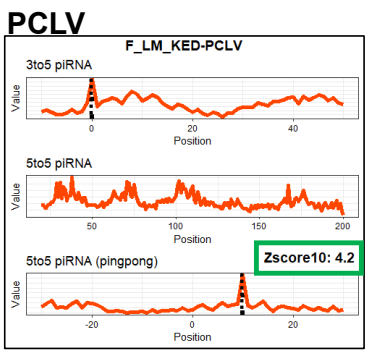




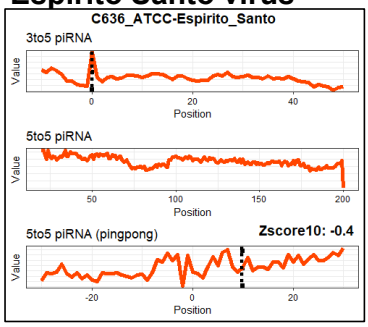
iii **KED_LM_Fem_1**

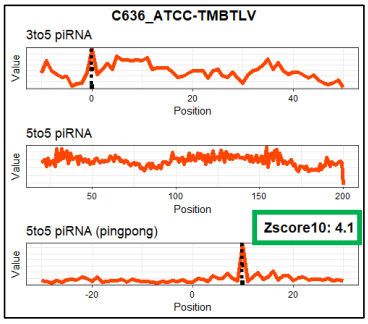






iv **C636-ATCC (BZL)**

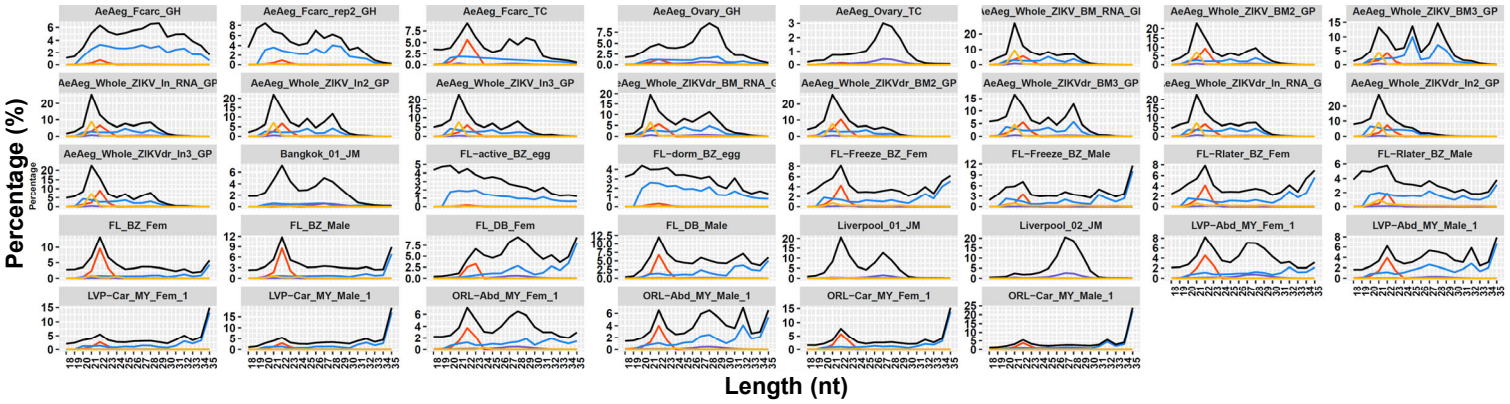




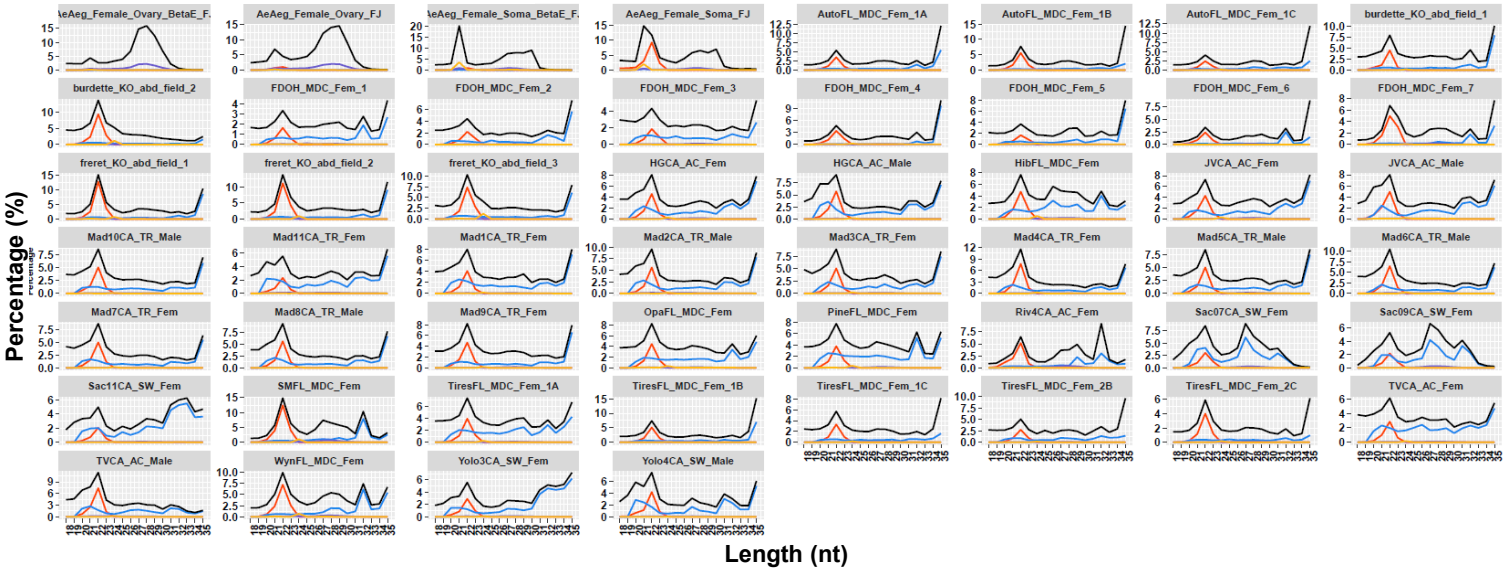
Supplementary Fig. S1. Using the AEFE1 EVE as a constant quality tracker control for vsmRNA profiling, implementing VirusDetect to discover new ISVs from small RNAs, and plotting piRNA biogenesis. (**a**) Coverage plots on left and bubble plots on right representing the small RNAs from the Endogenous Viral Element (EVE) from *Ae. aegypti* called AEFE1 (Suzuki et al, 2017)[25]. (**b**) The results output table from the VirusDetect program only loaded with the GBVRL database while analyzing the small RNA library from one of the BZL lab strain samples. (**c**) Coverage of the de-novo small RNA assembly contigs generated by VirusDetect of the corresponding virus determinations from the table in (b). The red asterisks mark a false-positive call against a tombus-like virus entry in GBVRL with limited contig coverage. (**d**) VirusDetect results analyzing the same BZL lab strain small RNA library in (b) but now with a custom loading of the TPA and MAG databases that are distinct from GBVRL. This result now shows two new viruses with much more complete contig coverage in (**e**), such as the Tiger Mosquito Bi-segmented Tombus-Like virus (TMBTLV) that is the true hit replacing the false-positive call in (b-c).(**f**) Autocorrelation analyses of phasing and ping-pong patterns from viral piRNAs in representative mosquito (**f.i-iii**) and cell (**f.iv**) samples. For each virus, top two plots are for 3'-to-5' phasing and bottom plot is of piRNA ping-pong patterns along with a Z10 score. Z10>3.0 denotes a strong ping-pong piRNA signature.

Gupta, Sharma, et al.
Supplementary Figure S2.

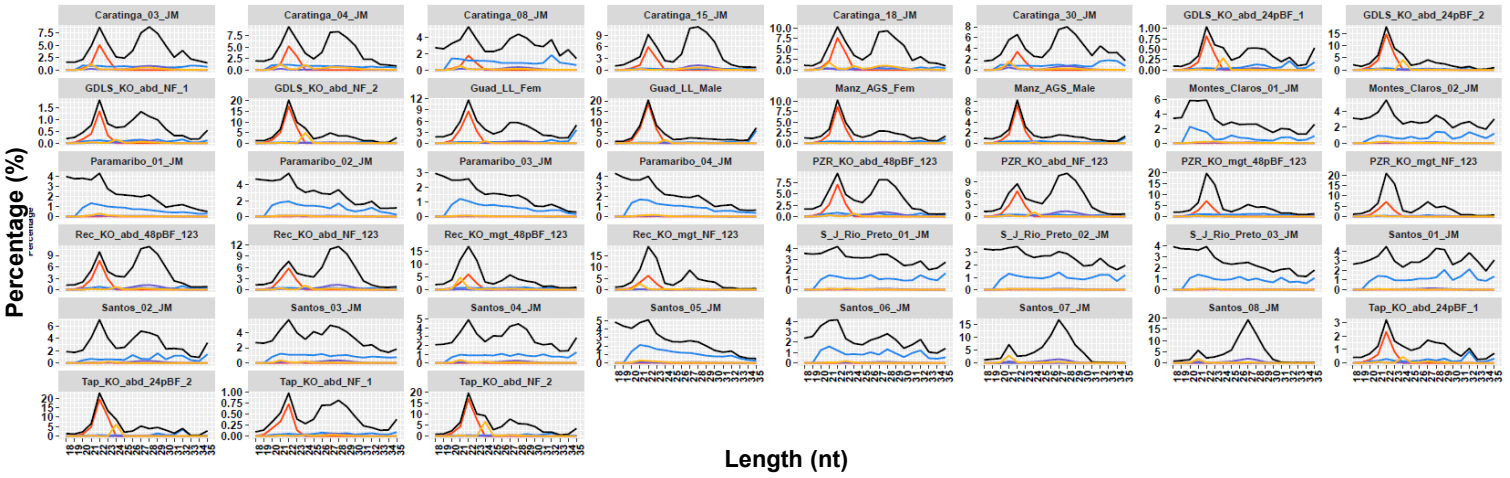
a. Lab strains



b. North American strains



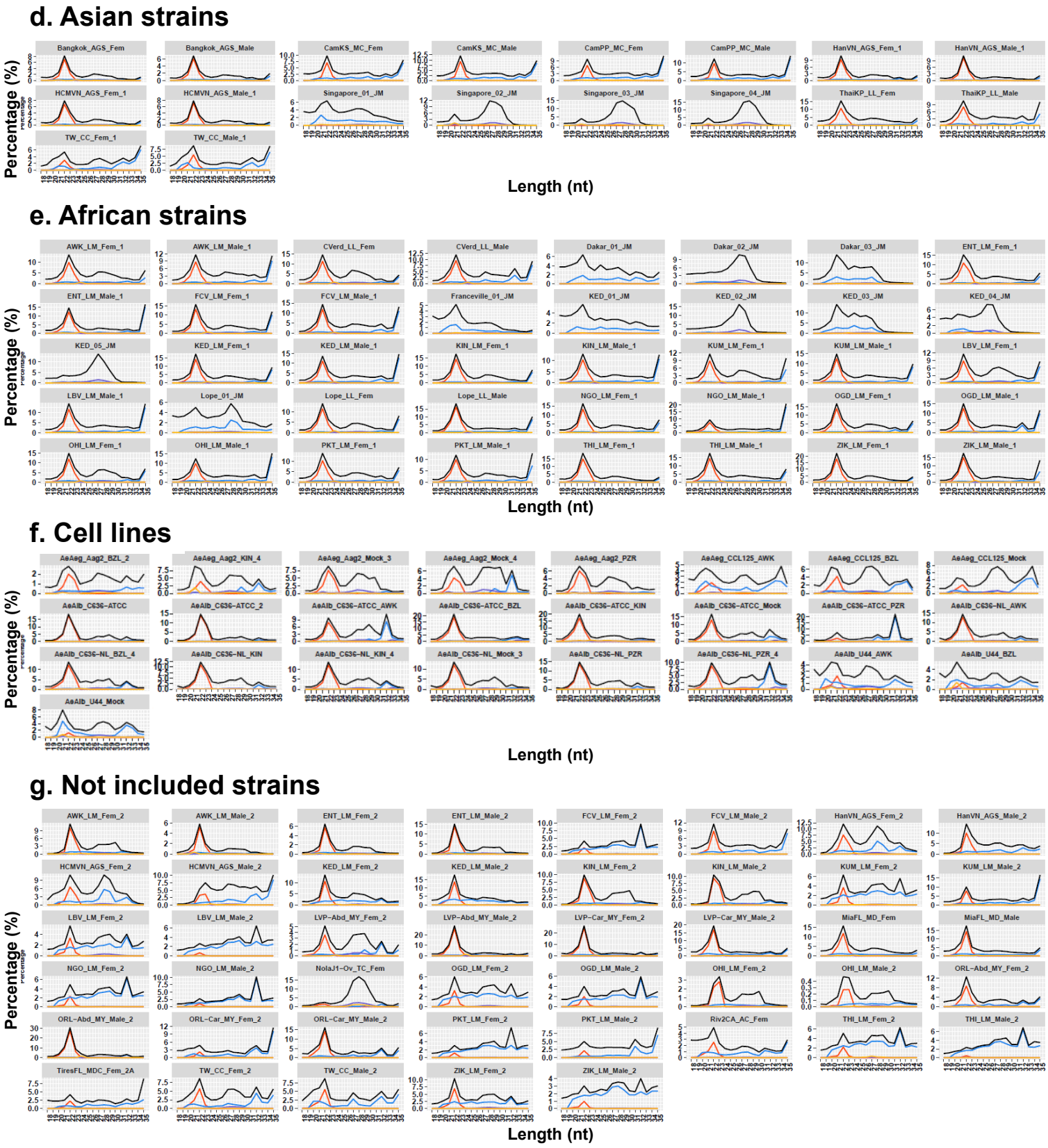
c. Central and South America strains



Type — TotalCount — miRNACount — StructureRNACount — VirusCount — TECount

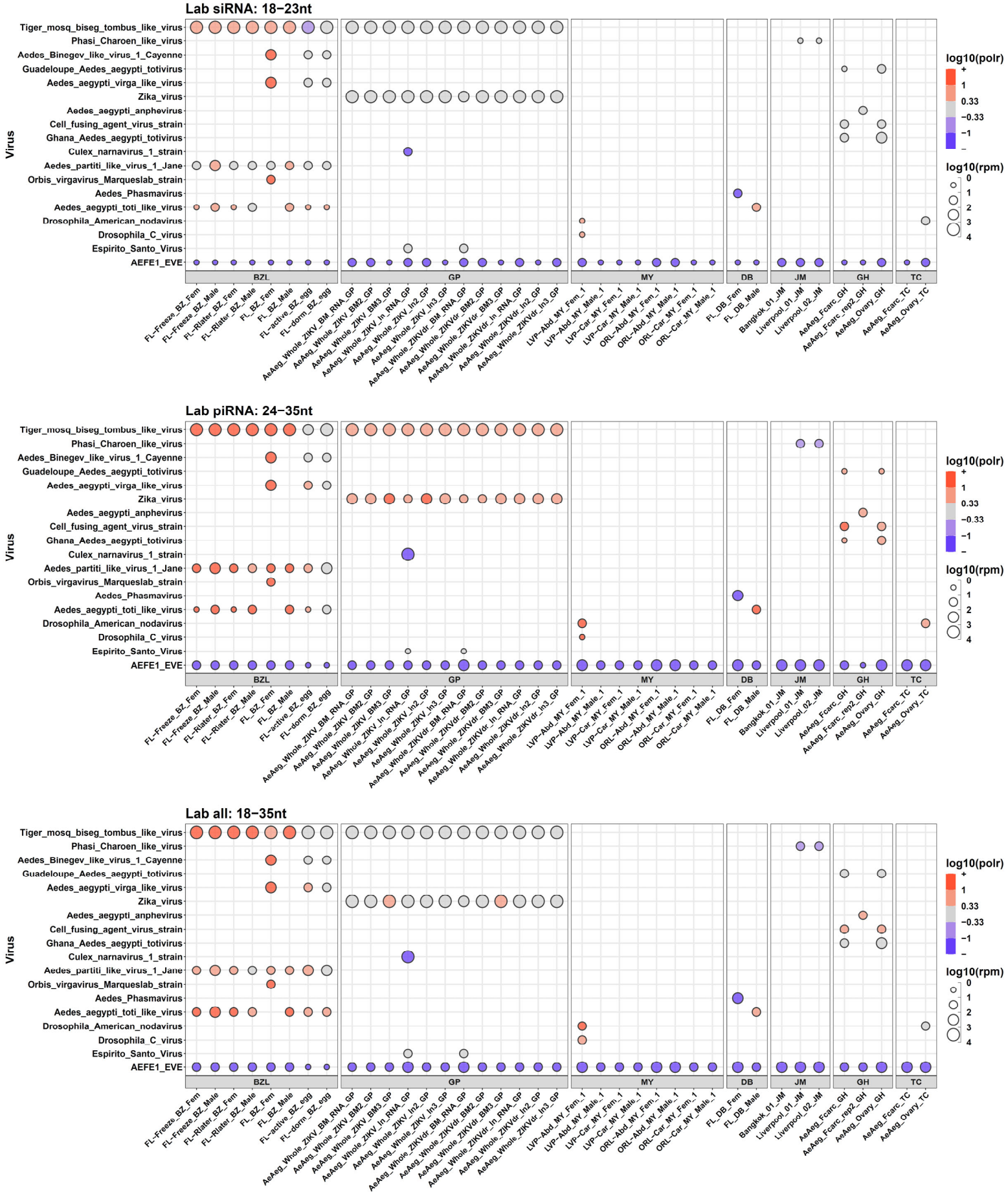
Supplementary Fig. S2. Read length distribution profiles of all the *Ae. aegypti* small RNA libraries analyzed in this global survey. Read length distribution plots as percentages of the libraries with functional classes of reads represented by different colored lines. The groups correspond to (**a**) Lab strains, (**b**) North American strains, (**c**) Central and South American strains, (**d**) Asian strains, (**e**) African strains, (**f**) Cell lines, (**g**) Removed samples.

Gupta, Sharma, et al.
Supplementary Figure S2.



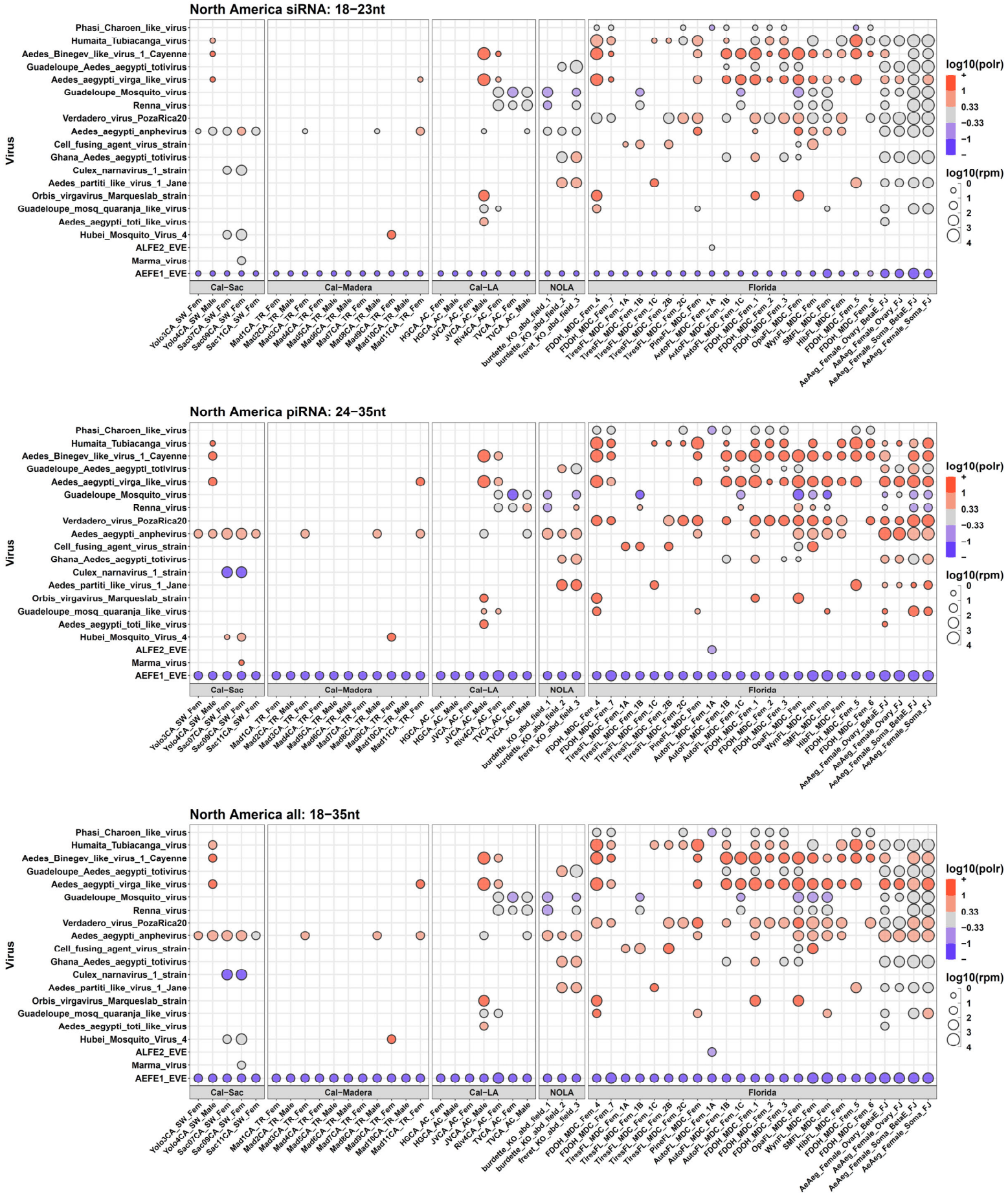
Supplementary Figure S2. Read length distribution profiles of all the *Ae. aegypti* small RNA libraries analyzed in this global survey. Read length distribution plots as percentages of the libraries with functional classes of reads represented by different colored lines. The groups correspond to (a) Lab strains, (b) North American strains, (c) Central and South American strains, (d) Asian strains, (e) African strains, (f) Cell lines, (g) Removed samples.

a



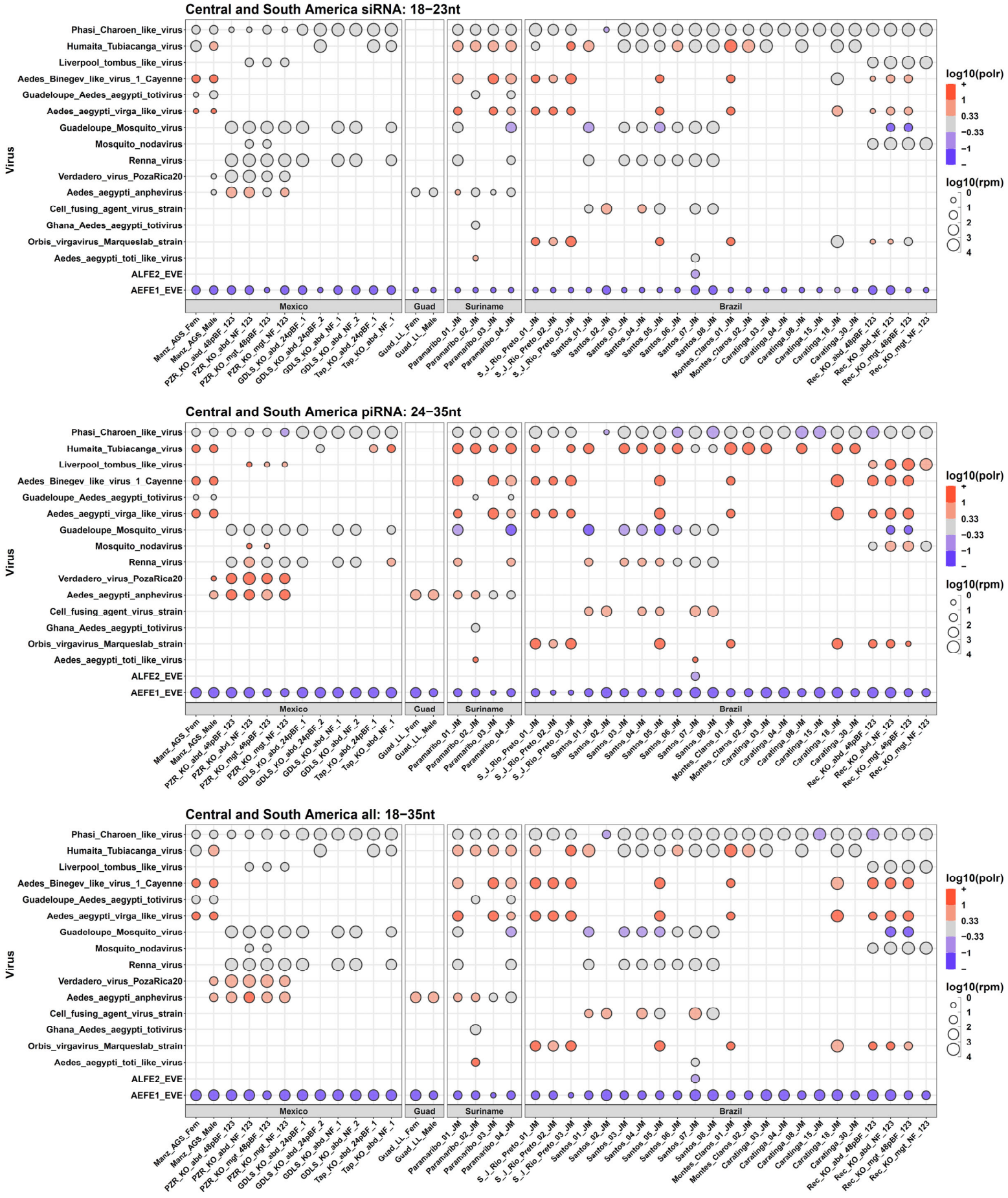
Supplementary Fig. S3. Complete bubble plots of vsmRNAs sampled in this global survey study.
Top plots are the siRNA-length small RNAs (18-23nt), the middle plots are the piRNA-length small RNAs (24-35nt), and the bottom plots are all the small RNAs as shown in the main figures 2 – 6. The groups correspond to (**a**) Lab strains, (**b**) North American strains, (**c**) Central and South American strains, (**d**) Asian strains, and (**e**) African strains. (**f**) Additional samples that we considered as replicates of other libraries but did not include in the main study because of concerns of potential cross-sample contamination.

b



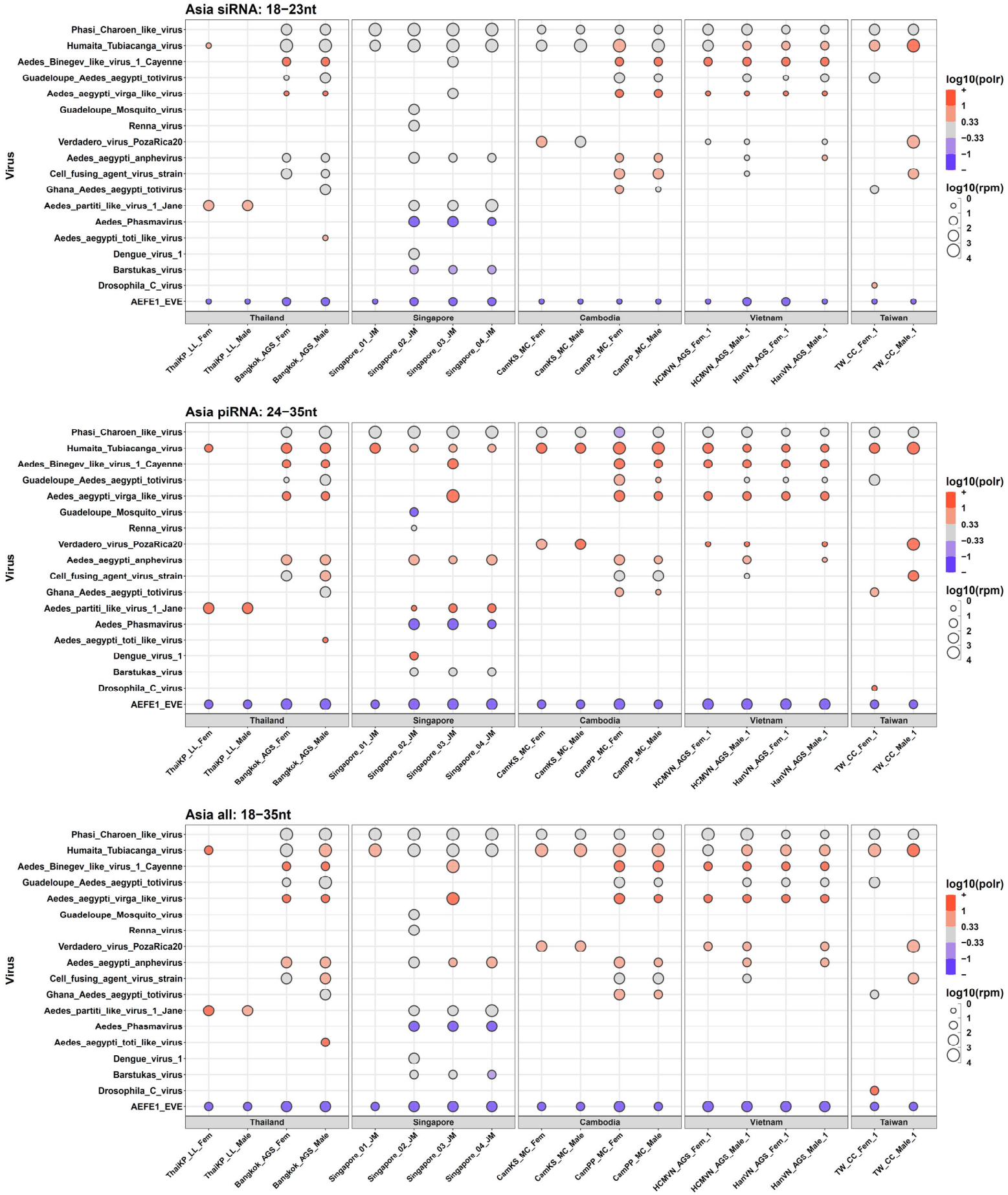
Supplementary Fig. S3. Complete bubble plots of vsmRNAs sampled in this global survey study.
Top plots are the siRNA-length small RNAs (18-23nt), the middle plots are the piRNA-length small RNAs (24-35nt), and the bottom plots are all the small RNAs as shown in the main figures 2 – 6. The groups correspond to (**a**) Lab strains, (**b**) North American strains, (**c**) Central and South American strains, (**d**) Asian strains, and (**e**) African strains. (**f**) Additional samples that we considered as replicates of other libraries but did not include in the main study because of concerns of potential cross-sample virus contamination.

C



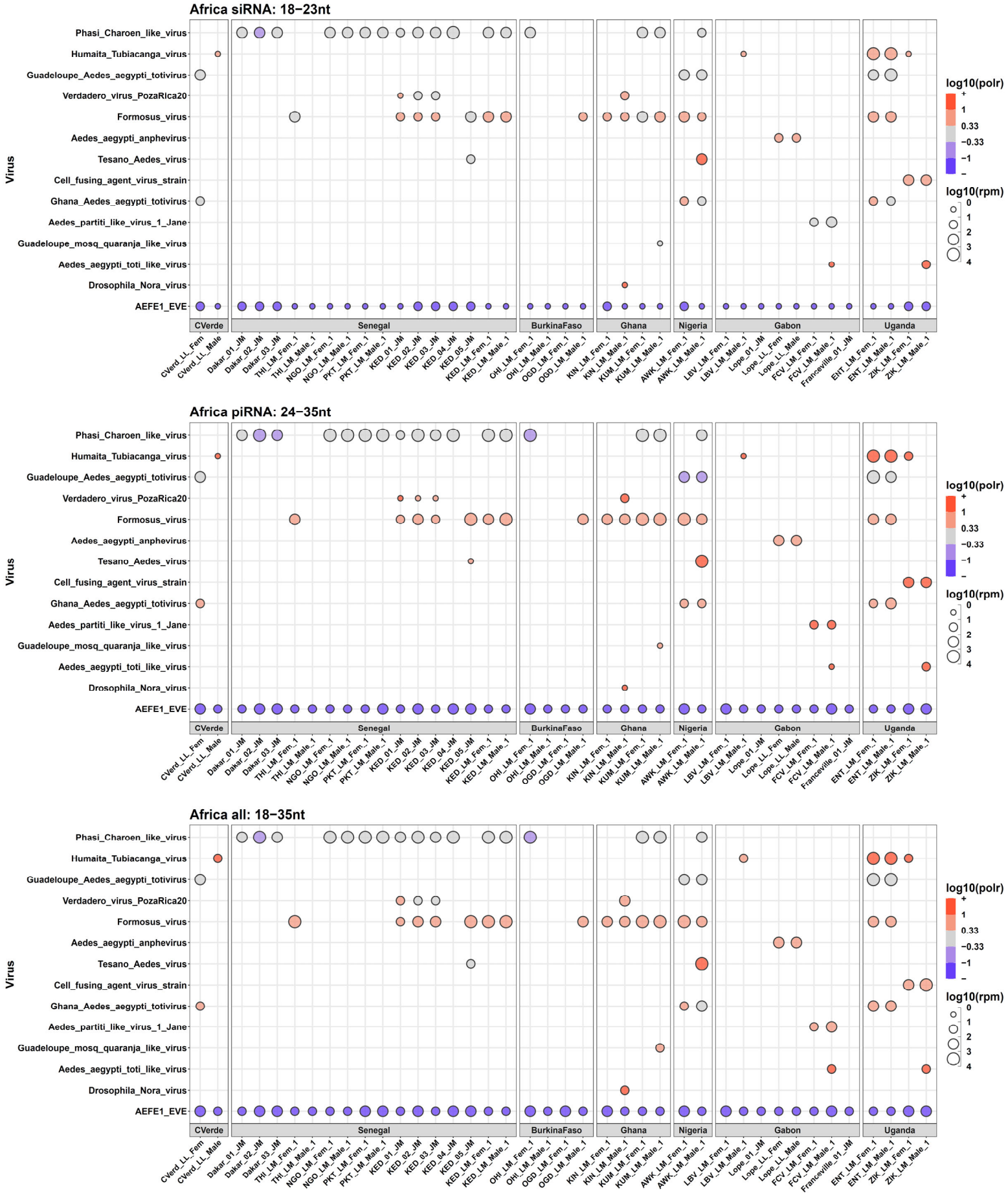
Supplementary Fig. S3. Complete bubble plots of vsmRNAs sampled in this global survey study.
Top plots are the siRNA-length small RNAs (18-23nt), the middle plots are the piRNA-length small RNAs (24-35nt), and the bottom plots are all the small RNAs as shown in the main figures 2 – 6. The groups correspond to (**a**) Lab strains, (**b**) North American strains, (**c**) Central and South American strains, (**d**) Asian strains, and (**e**) African strains. (**f**) Additional samples that we considered as replicates of other libraries but did not include in the main study because of concerns of potential cross-sample virus contamination.

d



Supplementary Fig. S3. Complete bubble plots of vsmRNAs sampled in this global survey study.
Top plots are the siRNA-length small RNAs (18-23nt), the middle plots are the piRNA-length small RNAs (24-35nt), and the bottom plots are all the small RNAs as shown in the main figures 2 – 6. The groups correspond to (**a**) Lab strains, (**b**) North American strains, (**c**) Central and South American strains, (**d**) Asian strains, and (**e**) African strains. (**f**) Additional samples that we considered as replicates of other libraries but did not include in the main study because of concerns of potential cross-sample virus contamination.

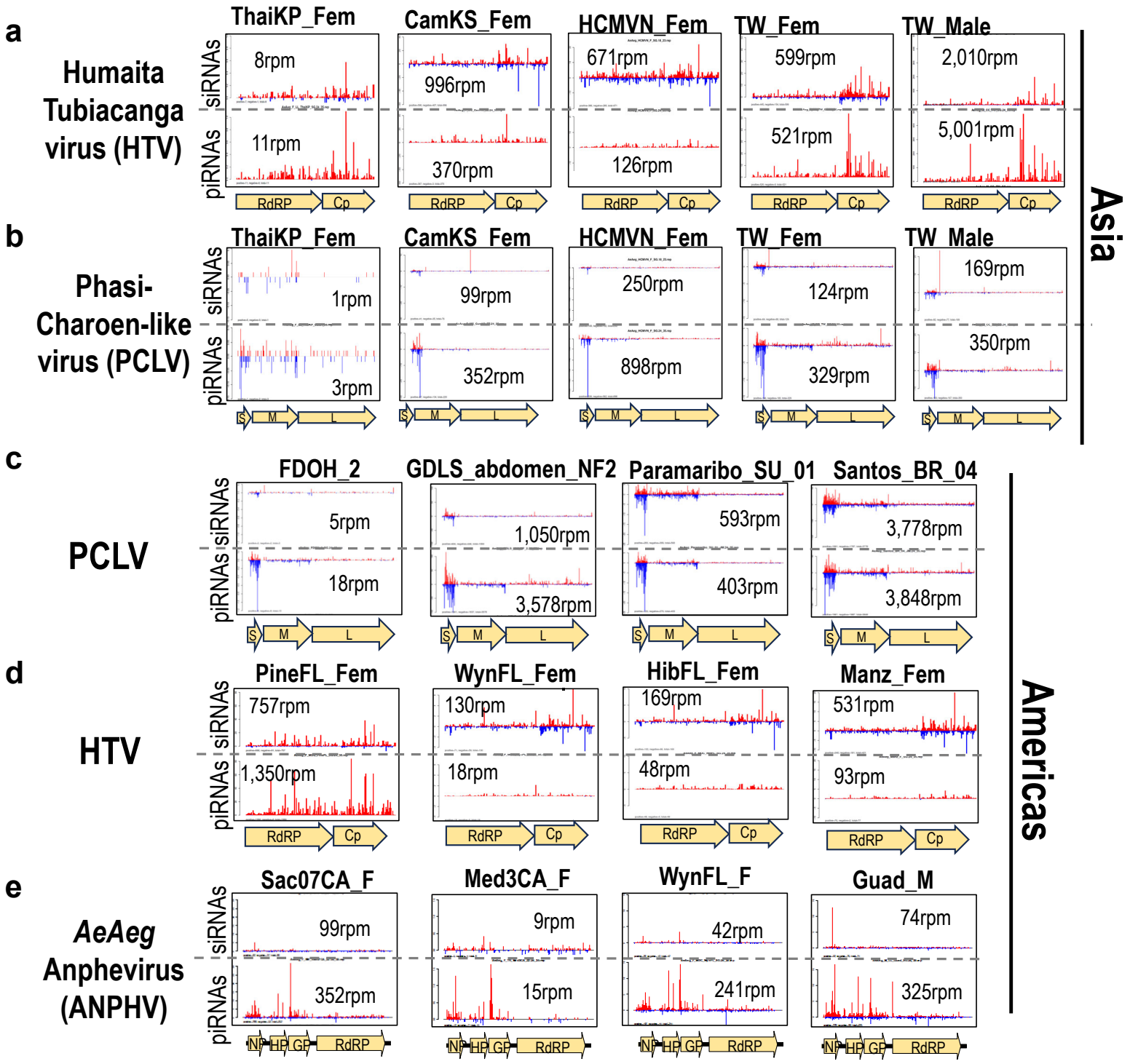
e



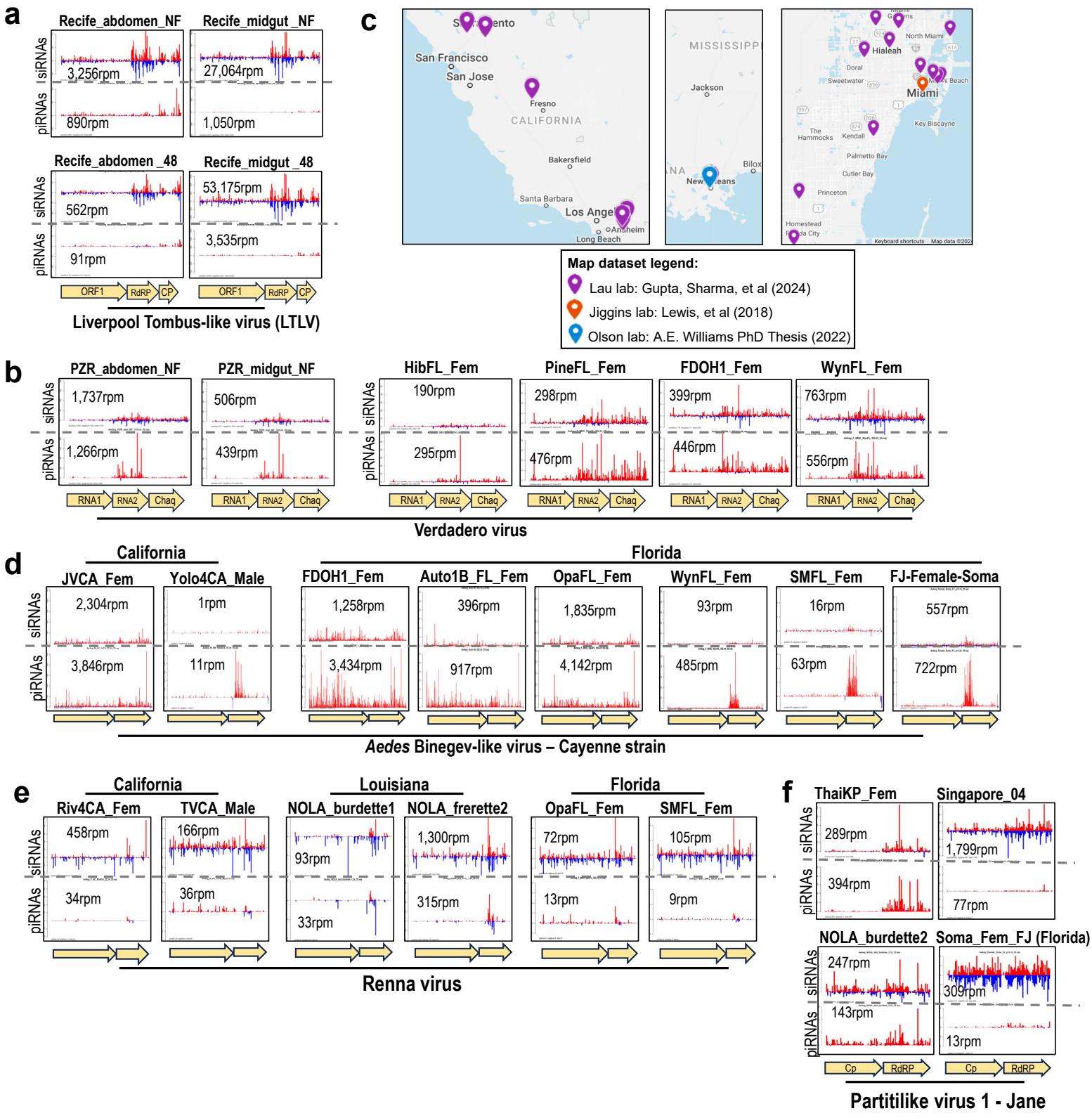
Supplementary Fig. S3. Complete bubble plots of vsmRNAs sampled in this global survey study.
Top plots are the siRNA-length small RNAs (18-23nt), the middle plots are the piRNA-length small RNAs (24-35nt), and the bottom plots are all the small RNAs as shown in the main figures 2 – 6. The groups correspond to (**a**) Lab strains, (**b**) North American strains, (**c**) Central and South American strains, (**d**) Asian strains, and (**e**) African strains. (**f**) Additional samples that we considered as replicates of other libraries but did not include in the main study because of concerns of potential cross-sample virus contamination.

f

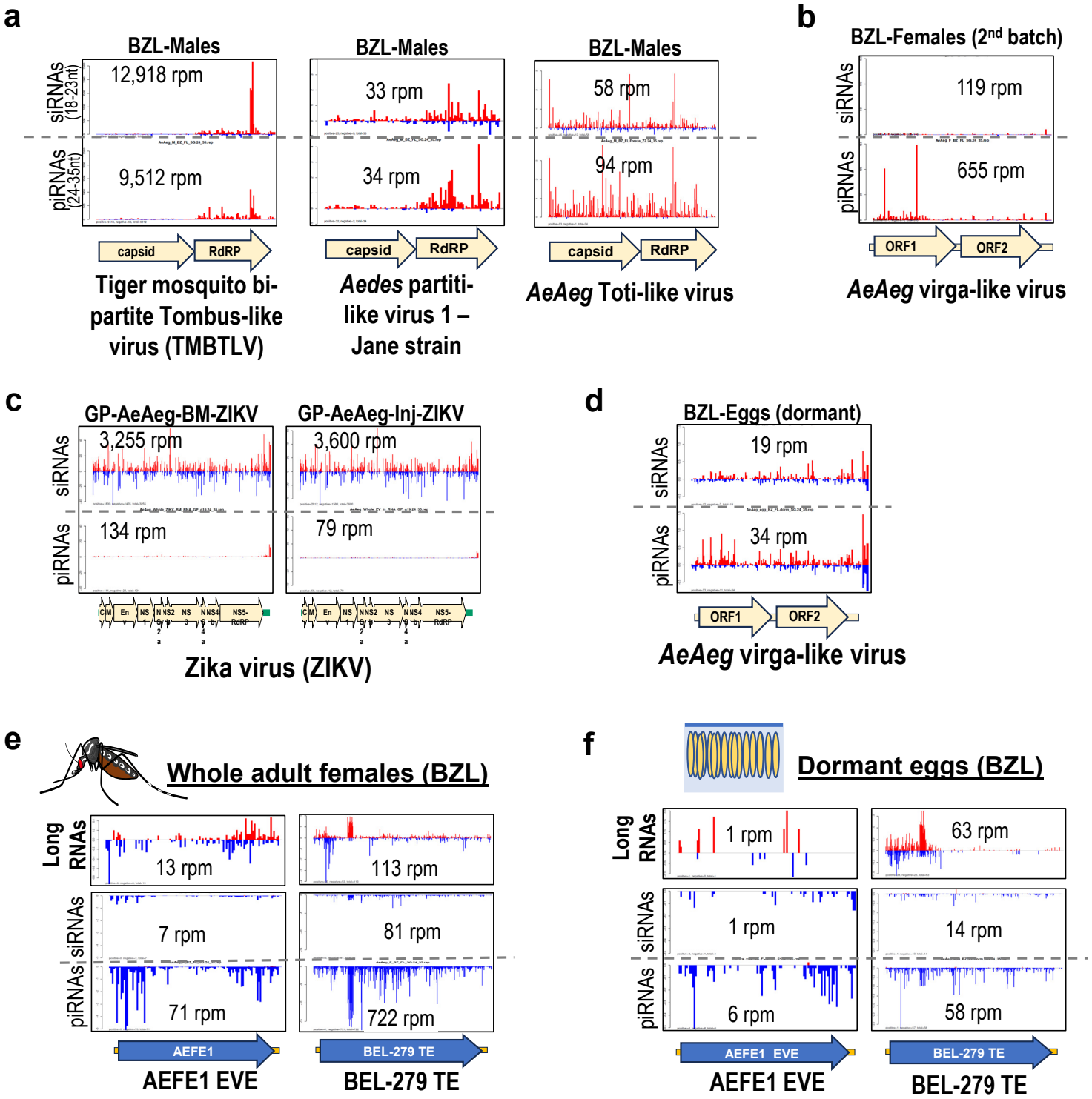
Supplementary Fig. S3. Complete bubble plots of vsmRNAs sampled in this global survey study.
Top plots are the siRNA-length small RNAs (18-23nt), the middle plots are the piRNA-length small RNAs (24-35nt), and the bottom plots are all the small RNAs as shown in the main figures 2 – 6. The groups correspond to (**a**) Lab strains, (**b**) North American strains, (**c**) Central and South American strains, (**d**) Asian strains, and (**e**) African strains. (**f**) Additional samples that we considered as replicates of other libraries but did not include in the main study because of concerns of potential cross-sample virus contamination.



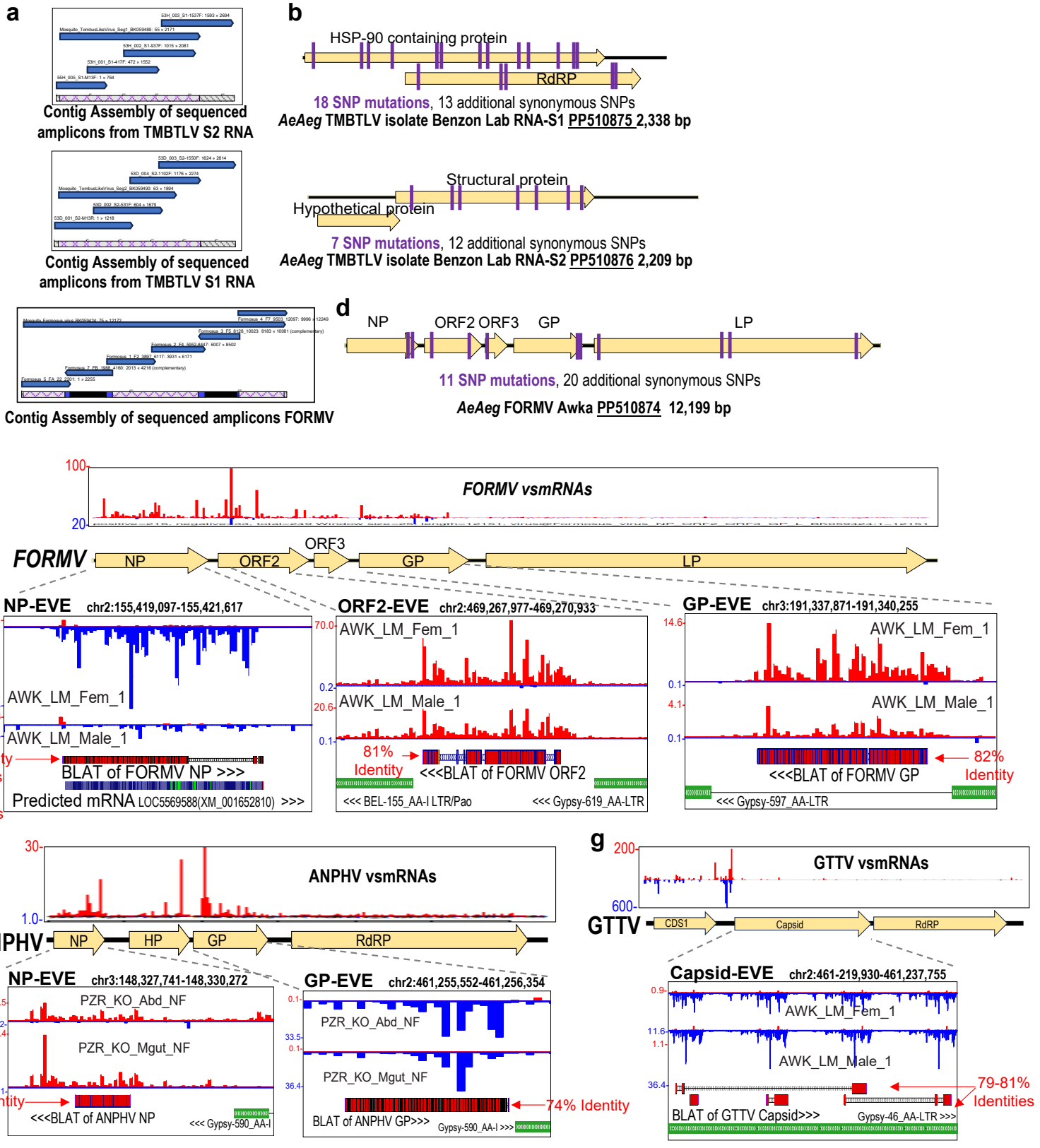
Supplementary Fig. S4. Additional coverage plots for HTV, PCLV, and ANPHV from the Americas and Asia. Additional coverage plots from Asian samples for (**a**) HTV and (**b**) PCLV. Additional coverage plots from American samples for (**c**) PCLV, (**d**) HTV, and (**e**) ANPHV. (**f**) Principal Component Analysis of vsmRNA levels can discriminate different groups of mosquito samples in between the North American, Central & South American, and Lab strains.



Supplementary Fig. S5. Plant-related viruses and other prevalent viruses in American *Aedes aegypti* shared on both east and west coasts.
(**a-b**) Coverage plots of two insect viruses related to plant-originating viruses marked in Figure 3. (**c**) Maps marking approximate locations of *A. aegypti* samples collected from California, Louisiana and Florida. Coverage plots of vsmRNAs from whole *A. aegypti* from: (**d**) *Aedes* Binegev-like virus- Cayenne strain from California and Florida; and (**e**) Renna virus from three states. (**f**) Coverage plots of Partiti-like virus 1 Jane strain vsmRNAs from a selection of Asian and American *Ae. aegypti*.

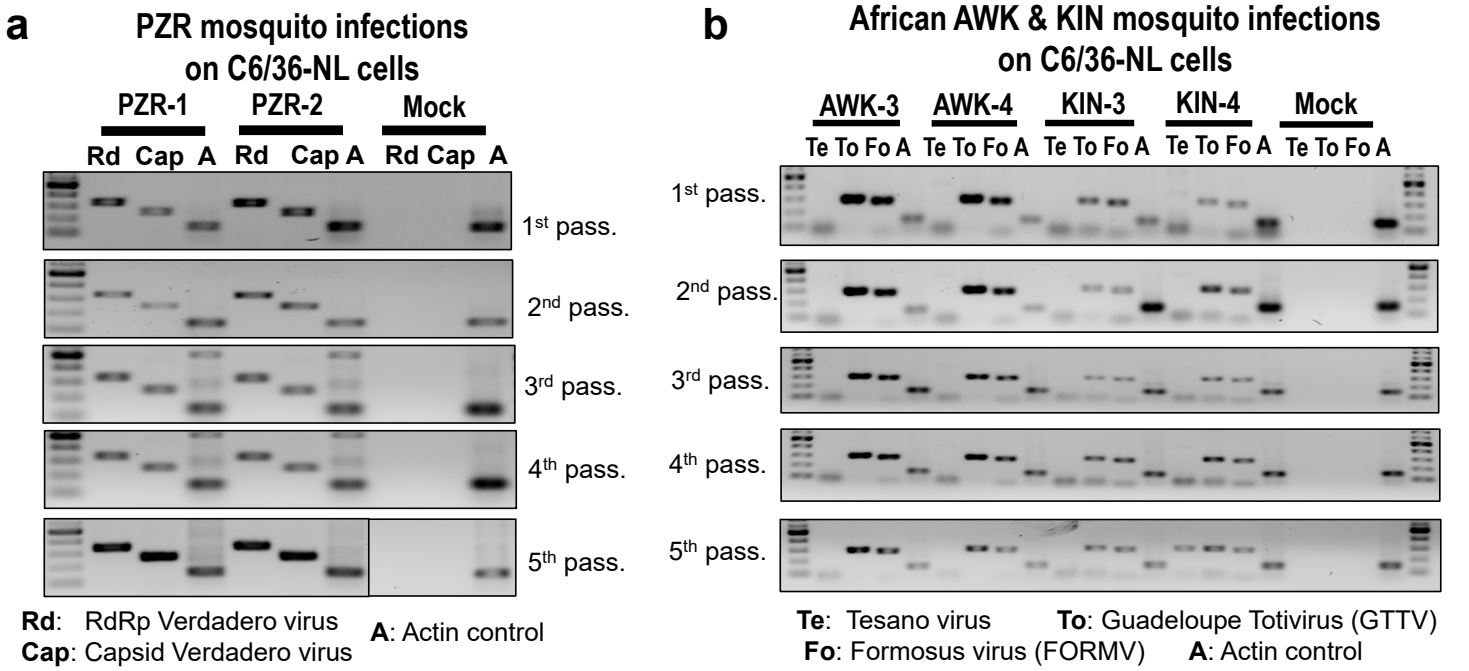


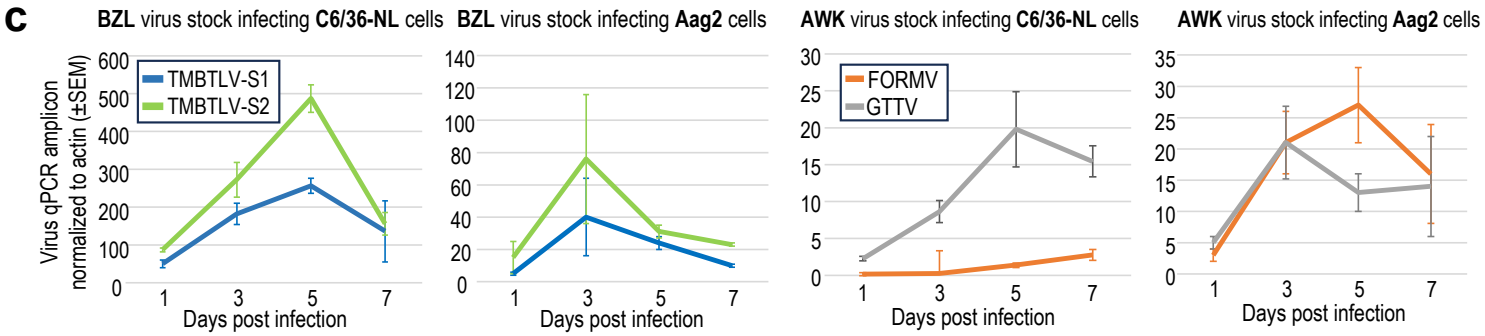
Supplementary Fig. S6. Additional coverage plots for lab and BZL mosquito samples. (**a**-**b**) Additional coverage plots of small RNAs from (a) BZL- Males and (b) 2nd batch of BZL-Females. (**c**) Coverage plots of ZIKV small RNAs from GP strains. (**d**) Additional coverage plot of small RNAs from *AeAeg* virga-like virus in Dormant BZL eggs. (**e**-**f**) Coverage plots of long RNAs versus small RNAs for an EVE and a TE from the BZL strain on (e) *Ae. aegypti* females and (f) dormant eggs.

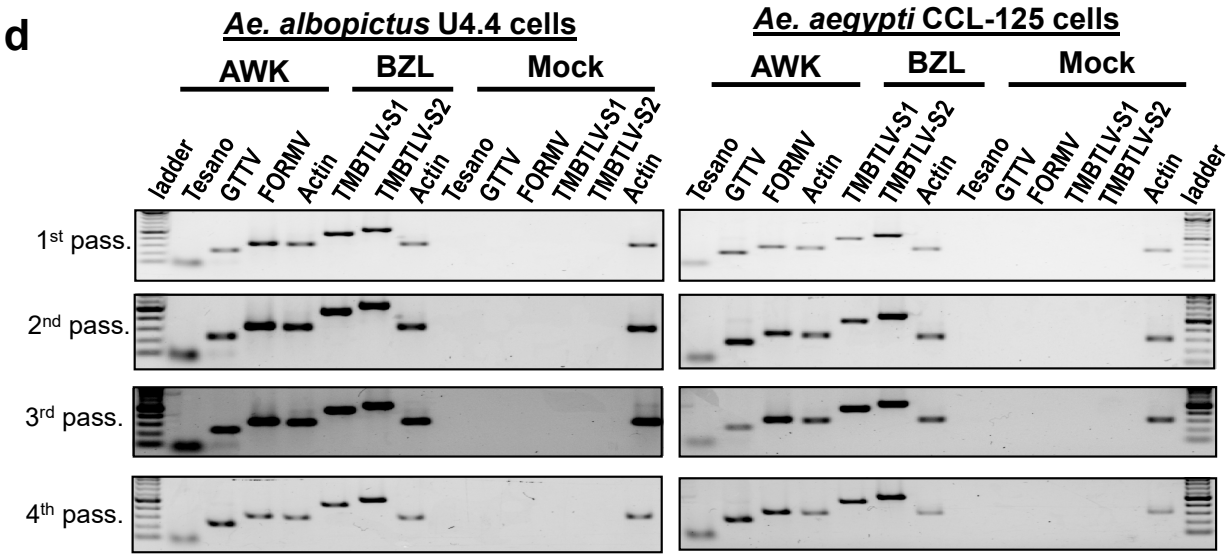


Supplementary Fig. S7. Cloning and sequencing TMBTLV and FORMV from *Ae. aegypti* and identification of EVE smRNAs that can base-pair against FORMV, ANPHV and GTTV

(**a**) Contig assembly of Sanger sequencing of cloned TMBTLV amplicons from S1 and S2 RNAs. (**b**) Diagrams of SNP mutations from our cloned fragment sequencing versus the initial TMBTLV references BK059489 and BK059490. (**c**) Contig assembly of Nanopore sequencing of cloned FORMV amplicons. (**d**) Diagrams of SNP mutations from our cloned fragment sequencing versus the initial FORMV reference BK059424. Both of our TMBTLV and FORMV variants sequences have now been contributed to GenBank with verified sequence accessions that should facilitate its inclusion in the future updates of the GBVRL database. (**e**-**g**) Genome browser snapshots of *Ae. aegypti* EVEs that generate significant smRNAs with base-pairing capacity against FORMV, ANPHV and GTTV. See the Supplementary Text Discussion on our approach to detect these *Ae. aegypti* EVEs and diagrams explaining MSRG discrimination of bonafide virus discovery from confounding EVE/transposons.



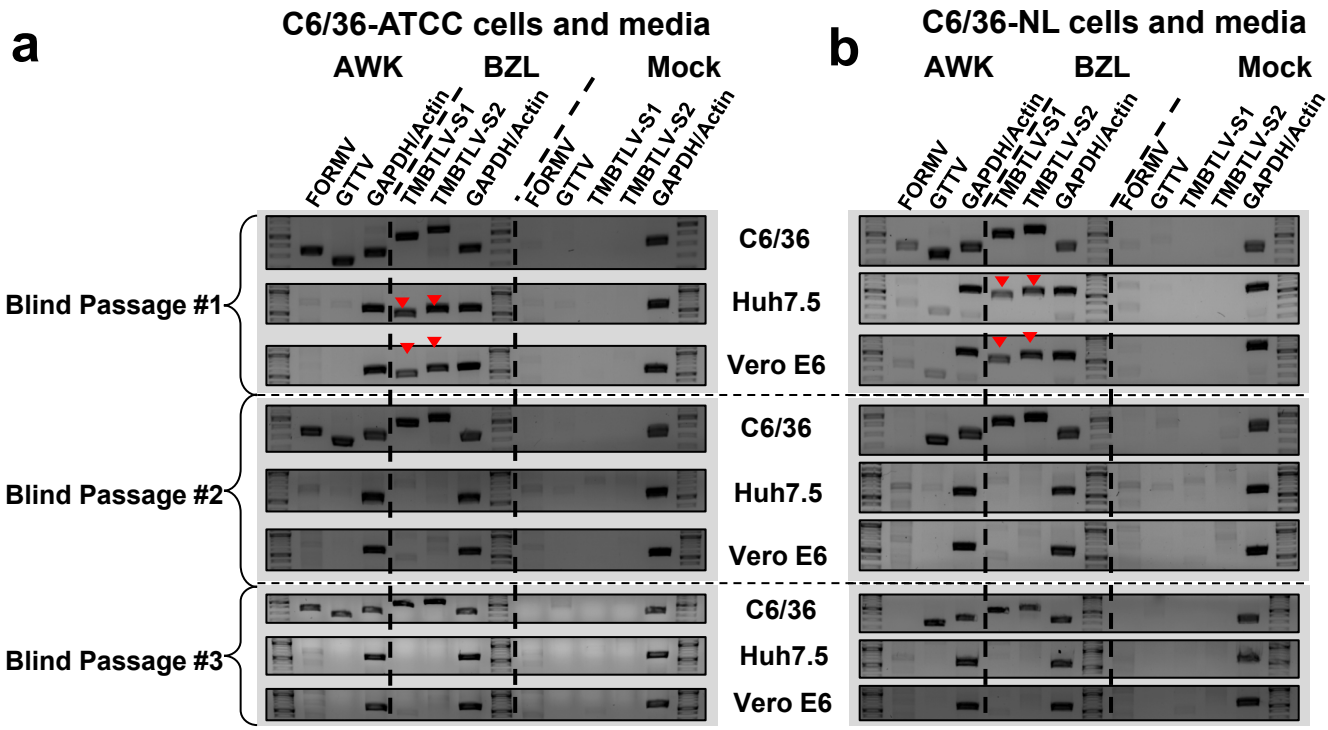


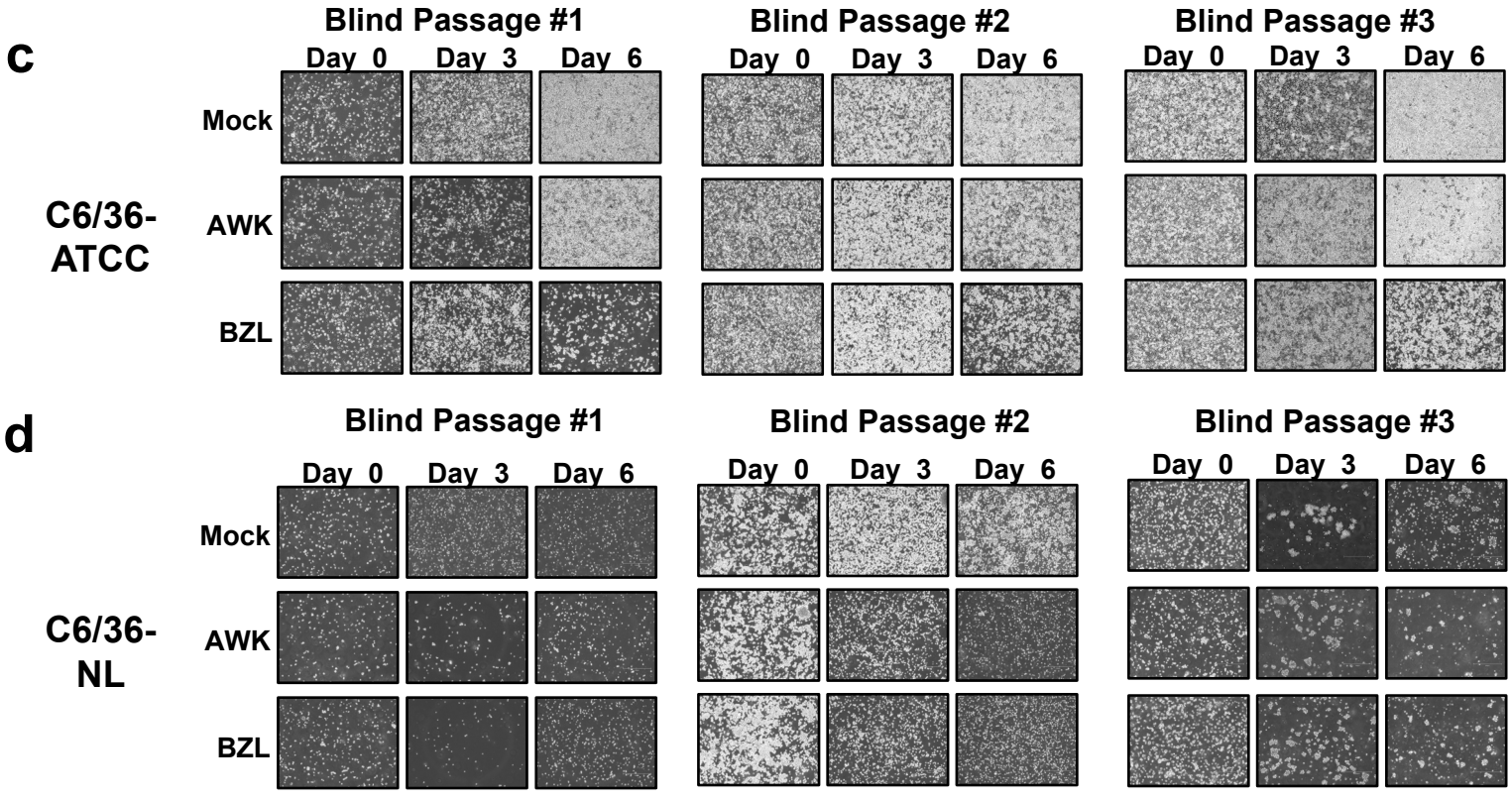


Supplementary Fig. S8. Additional blind passaging infections with PZR, AWK and KIN mosquito homogenates on C6/36-NL, Aag2, U4.4. and CCL-125 mosquito cells.

(**a**) RT-PCR of viral amplicons against the Verdadero virus RdRP and capsid genes. Two independent infections of C6/36-NL cells with two batches of PZR mosquitoes. (**b**) RT-PCR of viral amplicons against Tesano virus, GTTV and FORMV from two independent infections of C6/36-NL cells with two batches of AWK and KIN mosquitoes. The actin control are primers against *Ae. albopictus* actin. (**c**) Virus replication kinetics measured in C6/36-NL 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. (**d**) RT-PCR of viral amplicons against TMBTLV, Tesano virus, GTTV and FORMV from independent infections with two batches of AWK and KIN mosquitoes. The actin control are primers against *Ae. albopictus* actin.

Gupta, Sharma, et al.
Supplementary Figure S9.

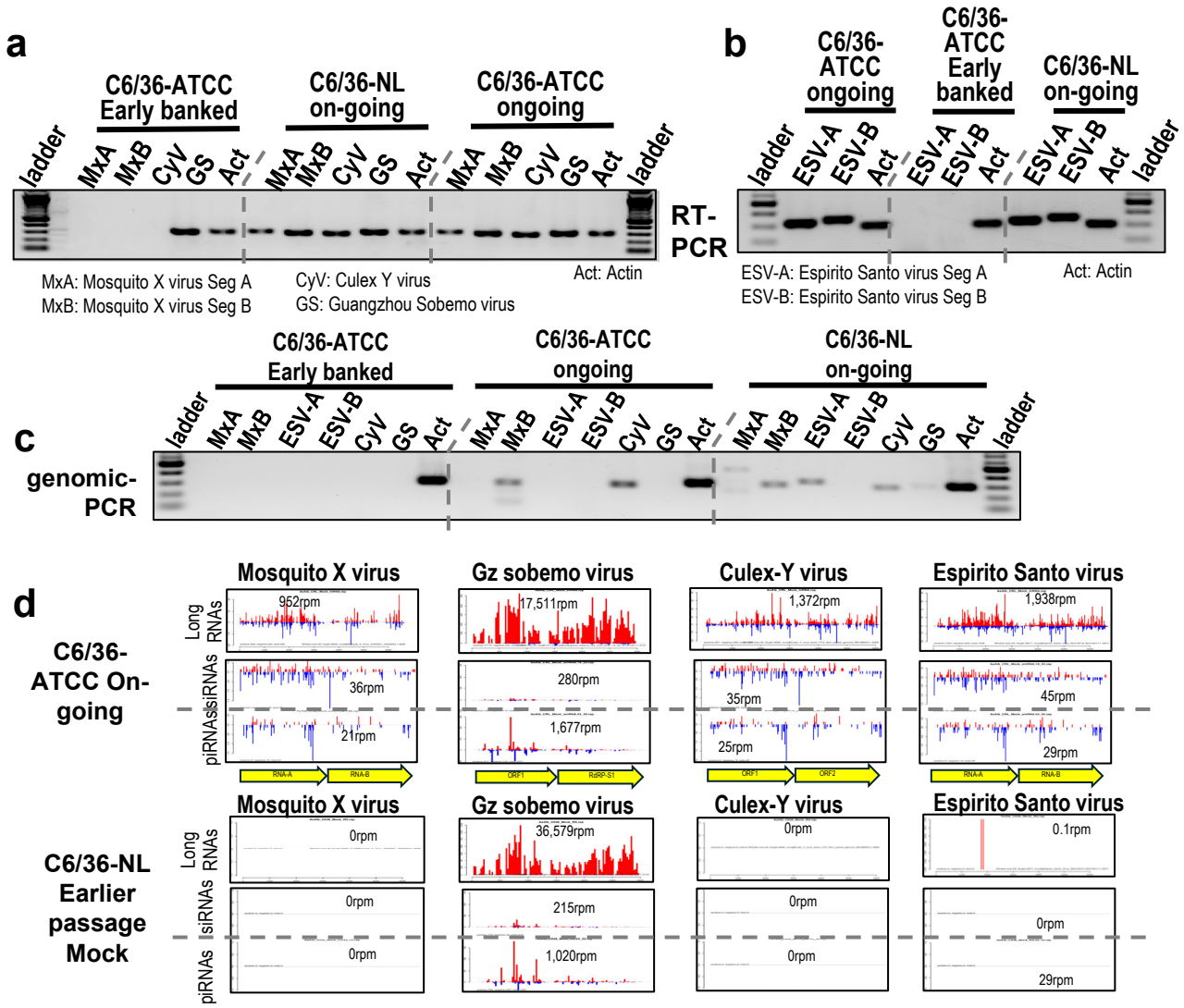




Supplementary Fig. S9. Testing the capacity of the AWK and BZL mosquito virus stocks to infect mammalian cells.

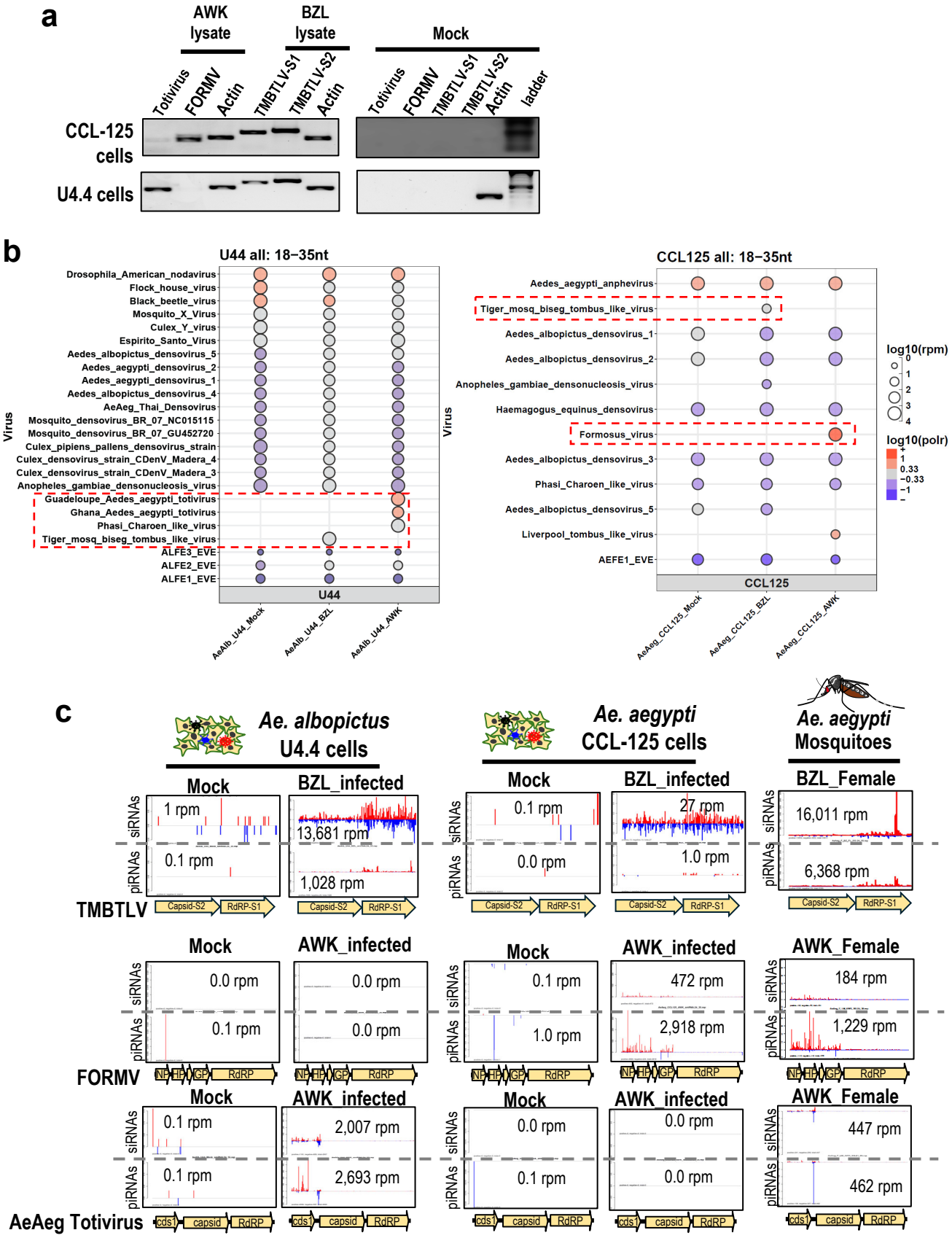
(**a-b**) Continued blind passaging of virus stocks from C6/36-ATCC cells (a) and C6/36-NL cells (b) are then tested for infection of mammalian VeroE6 and Huh7.5 cells followed by RT-PCR detection of viral amplicons. Red arrowheads point to insect virus amplicons detected in the mammalian cells but only in the first passage. Asterisk marks a GTTV already present in this lab stock of C636 cells. (**c-d**) Brightfield microscopy images of the culture growth overview of the C6/36-ATCC cells (c) and C6/36-NL cells (d) during the virus infections from AWK and BZL mosquito virus stocks.The decreased cell culture confluency/density by Day 6 is a proxy for the CytoPathic Effect (CPE) from likely virus propagation.

Gupta, Sharma, et al.
Supplementary Figure S10.



Supplementary Fig. S10. Infectivity of mosquito viruses that elicit a viral RNAi responses during routine culture with C6/36.
(**a-b**) RT-PCR of viral amplicons from the total RNA of three isolates of C6/36 cells, the early bank of C6/36-ATCC cells versus on-going cultures maintained for several months of C6/36-ATCC and C6/36-NL cells for (**a**) Mosquito X-virus, Culex-Y virus, and the Guangzhou Sobemo virus, and (**b**) against Espirito Santo Virus (ESV). The actin control are primers against *Ae. albopictus* actin. (**c**) Genomic PCR of the same viral amplicons from (a-b) but on the genomic DNA template from the same C6/36 cell lines. (**d**) Coverage plots of total RNA-seq and small RNA sequencing for the viruses measured above in (A-C) from two of the C6/36 cell lines. The C6/36-ATCC Early banked cells (not shown here) did not have total RNA sequencing, but small RNA sequence did not detect any mosquito viruses, as reflected in the bubble plot of Fig. 6a.

Gupta, Sharma, et al.
Supplementary Figure S11.



Supplementary Fig. S11. Mosquito viruses infecting U4.4 and CCL-125 cells also elicit a viral small RNA response.
(**a**) RT-PCR of viral amplicons from the total RNA from U4.4 and CCL-125 cells in a subsequent stable re-infection experiment distinct from the blind passaging experiment in Ext. Data Fig. S8. These cells' total RNA was then subjected to small RNA and total RNA sequencing for the analysis in the next panels. (**b**) Bubble plots of the viral smRNAs expression levels from the total RNA of stably infecting U4.4 and CCL-125 cells. Red dashed boxes highlight the viruses from the BZL and AWK mosquito lysates that had stably infected these mosquito cell lines. (**c**) Coverage plots of the viral smRNAs in the stably infected cells compared to the original viral smRNA patterns from the BZL and AWK female mosquitoes used for the lysates to reinfect the mosquito cells.