Mosquito vector competence for dengue is modulated by insect-specific viruses

Roenick Proveti Olmo, Yaovi M. H. Todjro, Eric R. G. R. Aguiar, João Paulo P. de Almeida, Flávia V. Ferreira, Juliana N. Armache, Isaque J. S. de Faria, Alvaro G. A. Ferreira, Siad C. G. Amadou, Ana Teresa S. Silva, et al.

Figure or Table # Please group Extended Data items by type, in sequential order. Total number of items (Figs. + Tables) must not exceed 10.	Figure/Table title One sentence only	Filename Whole original file name including extension. i.e.: Smith_ED_Fig1.jpg	Figure/Table Legend If you are citing a reference for the first time in these legends, please include all new references in the main text Methods References section, and carry on the numbering from the main References section of the paper. If your paper does not have a Methods section, include all new references at the end of the main Reference list.
Extended Data Fig. 1	**Co-occurrence of 139 viral contig clusters identified in *A. aegypti* and *A. albopictus* mosquitoes.**	Olmo_ExtFigure1.jpg	Heatmap represents the small RNA abundance for each of the 139 viral contigs in our 91 small RNA libraries from *A. aegypti* and *A. albopictus* mosquitoes. White indicates absence of small RNAs mapping to that contig. Contig clusters were defined using the dendrogram shown on the heatmap. Clusters that had a RdRp sequence were classified as a putative virus. Virus presence was considered if >50% of contigs belonging to a cluster were represented.
Extended Data Fig. 2	**Phylogeny of viruses identified in *A. aegypti* mosquitoes.**	Olmo_ExtFigure2.jpg	Phylogenetic trees were generated using the RdRp amino acid (aa) or nucleotide (nt) sequences and the substitution models as indicated: **a**, Aslam narnavirus (aa - LG + G); **b**, Nyamuk partiti-like virus (aa - BLOSUM 62); **c**, Orbis virgavirus (aa - BLOSUM62 + F); **d**, Bahianus rhabdovirus (aa - BLOSUM62); **e**, Lactea totivirus (nt - Tamura-Nei 93). Bootstrap confidence is shown close to each clade and values under 60% were omitted.
Extended Data Fig. 3	**Virus-derived small RNA profiles in mosquitoes.**	Olmo_ExtFigure3.jpg	Small RNA size distribution and 5’ base preference is shown on the left while the density of small RNAs (coverage) is shown on the right for representative contig(s) of each of the 12 viruses identified in this study.

Extended Data Fig. 4	**Burden of viruses in mosquitoes from different collection sites.**	Olmo_ExtFigure4.jpg	**a**, Abundance of small RNA sequences in pooled libraries from each location. Each dot represents the small RNA abundance in a contig, and violin plots represent contig clusters (see Extended Data Fig. 1) at different locations with colors matching the mosquito species. Error bars represent the standard deviation of the mean for each location. The number of contigs analyzed per location is indicated above each graph. **b**, Detection of representative contigs of newly detected viruses by RT-qPCR (black bars) in comparison to the detection of small RNAs (20-30 nt length) to the same given contig (gray bars). RT-qPCR detection is normalized against the endogenous constitutive gene *RpL32*. *, indicates detection by conventional RT-PCR. **c**, viral contig detection by conventional RT-PCR using independent sets of primers pairs. Conventional PCR and qPCR were repeated twice on the same samples. The expected size of viral contigs is shown. *ns* indicates a non-specific band. **d**, Sequence variation between viral contigs of Orbis virgavirus in RNA samples originated from Suriname along the region that is complementary to RT-qPCR primers. **e**, Ratio between relative RT-qPCR and small RNA abundance for each virus. The number of independent mosquito samples analyzed per virus is indicated above each graph. **f**, Combined incidence of DENV and ZIKV for each mosquito capture location in the previous, current, and subsequent years of collection

			(represented by -1, 0, and +1, respectively). Data were obtained from public sources for each location.
Extended Data Fig. 5	**Characterization of HTV and PCLV infection in wild and laboratory mosquitoes.**	Olmo_ExtFigure5.jpg	**a**, geographic distribution of mosquitoes carrying HTV and PCLV in the city of Caratinga, Brazil. Maps show the density of adult *A. aegypti* mosquitoes captured from July 2010 until August 2011 estimated from the number of mosquitoes captured in individual traps. All mosquitoes, HTV positive, PCLV positive and double positive individuals are shown. Virus detection was performed by RT-qPCR. Map source: OpenStreetMap. **b**,**c** tissue tropism of HTV and PCLV upon natural and artificial infections in *A. aegypti* mosquitoes. **b**, Scheme of mosquito dissection and tissues tested for virus infection by RT-qPCR. Pie charts show the prevalence of HTV and PCLV infection, assessed in tissues of naturally infected wild mosquitoes or laboratory mosquitoes injected with HTV and PCLV. Individual tissues were tested for virus presence upon dissection at 2-, 4-, 6- and 8-days post injection (*d.p.i.*) by RT-qPCR. **c**, Detection of HTV and PCLV in eggs by RT-PCR. Eggs were either rinsed with distilled water (no treatment group) or washed with bleach (2,5% active chlorine) prior to RNA extraction. The endogenous constitutive gene *RpL32* was used as amplification control. Results are representative of two independent experiments. **d-e**, HTV and PCLV do not grow in mammalian cell culture. VERO cells were exposed to mosquito extracts containing HTV (**d**) and PCLV (**e**), and supernatants were collected at 1-, 3- and 5-days

			post exposure. A spike containing 10^5 pfu of vesicular stomatitis virus (VSV) was added prior RNA extraction and used to normalize the quantification of HTV and PCLV in the supernatant. No statistically significant difference was observed in HTV and PCLV levels at 1-, 3- and 5-days post infection as determined by two-sided one-way ANOVA with Dunns' correction for multiple comparisons. Dots and error bars indicate the mean and the standard error of the mean, respectively. *n* indicates the number of independent tissue culture wells tested for each virus at each time point.
Extended Data Fig. 6	**HTV and PCLV facilitate systemic ZIKV infection in mosquitoes.**	Olmo_ExtFigure6.jpg	**a-c**, Strategy to evaluate the interference of HTV and PCLV for ZIKV infection and replication in natural populations of mosquitoes. (**a**) HTV/PCLV infected and virus-free wild mosquito populations were infected with ZIKV by intrathoracic injection. Viral loads and prevalence of infection were measured in the (**b**) midgut and (**c**) carcass of mosquitoes at 2-, 4- and 8-days post feeding. The prevalence of infection in each group is shown below plots. **d-f**, Laboratory mosquitoes (**d**) were infected artificially with HTV and PCLV and 7 days later were fed on ZIKV-infected mice. Viral loads and prevalence of infection were measured in the (**e**) midgut and (**f**) carcass of mosquitoes at 4- and 8-days post injection. The prevalence of infection in each group is shown below the plots. **g-i**, Laboratory mosquitoes were infected artificially with HTV and PCLV or control (mock) and 7 days later infected with

			ZIKV by intrathoracic injection. Viral loads and prevalence of infection were measured in the (**h**) midgut and (**i**) carcass of mosquitoes at 2-, 4- and 8-days post injection. The prevalence of infection in each group is shown below plots. **j-n**, Wild mosquito populations naturally infected with HTV and PCLV were allowed to feed in mice infected with ZIKV or mock-infected controls. Viral loads of HTV and PCLV were measured in the midgut (**k,l**) and in the carcass (**m,n**) of mosquitoes at 4-, 8- and 14-days post feeding. **o-p**, HTV/PCLV infected or control mosquitoes were exposed to ZIKV-infected mice (**o**). Viral loads and prevalence of infection were measured in salivary glands (**p**) of mosquitoes at the indicated time points. Pie charts below each group indicate the prevalence of ZIKV infection. *d.p.f.* – days post feeding, *d.p.i.* – days post injection, *NS* – non-significant. In box plots of ***b***, ***c***, ***e***, ***f***, ***h***, ***i***, ***k***, ***l***, ***m***, ***n***, and ***p,*** boxes show the second and third interquartile ranges divided by the median while whiskers represent maximum and minimum values. Statistical significance was determined by two-tailed Mann–Whitney U-test. Numbers of infected samples over the total number tested are indicated above each column. Each dot represents an individual sample. Statistical significance of prevalence was determined by two-tailed Fisher's exact test.
Extended Data Fig. 7	**Differential gene expression in wild mosquitoes**	Olmo_ExtFigure7.jpg	**a**, Differential gene expression in the carcass of wild mosquitoes carrying HTV and PCLV or non-infected siblings during DENV infection at 4, 8 and 14 days post

	carrying HTV and PCLV.		feeding. **b**, number of up or down regulated genes regarding the infection with HTV and PCLV at each time point as shown in ***a***. Common genes across time points are shown. **c**, Immune genes regulated during infection with HTV and PCLV in comparison to virus-free siblings at different times after DENV infection. **d-e**, Intrathoracic injection of ZIKV in wild mosquitoes carrying HTV and PCLV or virus free siblings (**d**). Histone H4 levels were quantified in the midgut of mosquitoes at 2, 4, and 8 days post injection with ZIKV (**e**). Error bars represent mean and standard deviations of the mean, and statistical significance was determined by two-sided one-way ANOVA with Tukey's correction for multiple comparisons. **f-g**, Artificial infection of laboratory mosquitoes with HTV and PCLV does not modulate levels of histone H4. **g**, laboratory mosquitoes were artificially infected with HTV and PCLV and histone H4 levels were analyzed at different time points. In box plots of ***e*** and ***g,*** boxes show the second and third interquartile ranges divided by the median while whiskers represent maximum and minimum values. Statistics were performed using two-sided one-way ANOVA with Tukey's correction for multiple comparisons. Each dot represents an individual sample. *CPM* – counts per million, *d.p.i.* – days post infection, *NS* – non-significant.
Extended Data Fig. 8	**Complexity of histone genes in**	Olmo_ExtFigure8.jpg	**a**, *Histone H4* gene copies in the *A. aegypti* genome (Vectorbase version 52) were reannotated using

	the genome of *A. aegypti*.		BLAST similarity search with further confirmation of RNA-seq reads mapping to each gene copy. Along with other *histone* genes currently annotated in Vectorbase, the number of copies in chromosomes or supercontigs are shown and the largest cluster of genes highlighted. **b**, Organization of the largest cluster of histone genes on chromosome 3 as indicated by the gray box. **c**, Weblogo showing the conservation of the amino acid sequence of histone H4 open reading frames, which only varied at the positions 36 and 98, indicated by a circle and an asterisk, respectively. The number of amino acid changes in each position is indicated. **d**, Histone H4 genes organized by nucleotide sequence similarity according to the dendrogram with the expression indicated by the heatmap in different *A. aegypti* tissues. Bootstrap values over 60 are shown. *Histone H4* genes positioned in the cluster at chromosome 3 are indicated by gray boxes and the presence of a polyadenylation signal is indicated. **e**, *Histone H4* gene expression in wild mosquito populations carrying HTV and PCLV or virus free siblings infected with DENV, quantified by RT-qPCR from cDNAs synthesized with random primers (hexamers) or anchored oligo dT_{22}. *d.p.i.*, days post injection. In ***d***, SRA accession numbers in same order as shown in the heatmap: Female whole body (non-BF SRR1585314, SRR1585315, SRR1585316; 48 h post-BF SRR1532683, SRR1532684, SRR1532685, SRR1532693, SRR1532694, SRR1532695); Female brain (48 h post-BF SRR1166497; 96 h post-BF SRR1167481); Male (brain

			SRR1167543); Female salivary glands (SRR2659965, SRR2659966); Female midgut (SRR5288077, SRR5288080, SRR5288082, SRR5288087, SRR5288093, SRR5288100); Female malp. Tubules (non-BF SRR3680433, SRR3680434); Female carcass (12 h post-BF SRR923823; 24 h post-BF SRR923830; 36 h post-BF SRR923835; 48 h post-BF SRR923841; 60 h post-BF SRR923847; 72 h post-BF SRR923736); Fem. Carcass (no ovaries) 24 h post-BF (SRR388683); Fem. Low reprod. Tract (0 h post-mating SRR3213863, SRR3213864; 6 h post-mating SRR3213865, SRR3213866; 24 h post-mating SRR3213867, SRR3213868); Male sperm (early SRR3554588; late SRR3554589); Male testis (SRR6311395, SRR6311396); Embryo (4-8 h SRR1578254, SRR1578255, SRR1578256); 1 day old female ovaries (SRR388680); Female ovaries (non-BF SRR1167515, SRR1167516, SRR1167517, SRR1167518, SRR1167519, SRR1167520; 24 h post-BF SRR388682; 96 h post-BF SRR1167538, SRR1167539). **f-h**, silencing of histone H4 by RNA interference in adult mosquitoes. **f**, strategy for dsRNA mediated gene silencing in adult mosquitoes. **g-h**, Histone H4 levels in the midgut of ISV free laboratory mosquitoes (**g**) or wild mosquitoes carrying HTV and PCLV (**h**) injected with dsRNA targeting GFP (dsGFP) as control or histone H4 (dsH4) at 4 days post feeding. **I**, AGO2 levels in the midgut of mosquitoes carrying HTV and PCLV injected with dsRNA targeting GFP (dsGFP) as control or Ago2 (dsAGO2). Each dot represents an individual sample. In box plots of ***e***, ***g***, ***h***, and ***i,*** boxes show interquartile ranges divided by the median while whiskers represent maximum and minimum values. Statistical significance was determined using two-sided

			one-way ANOVA with Tukey's correction for multiple comparisons.

2

Item	Present?	Filename Whole original file name including extension. i.e.: Smith_SI.pdf. The extension must be .pdf	A brief, numerical description of file contents. i.e.: *Supplementary Figures 1-4, Supplementary Discussion, and Supplementary Tables 1-4.*
Supplementary Information	No		
Reporting Summary	Yes	nr-reporting-summary_Olmo_et_al.pdf	
Peer Review Information	Choose an item.	*OFFICE USE ONLY*	

3

4

5

Type	Number Each type of file (Table, Video, etc.) should be numbered from 1 onwards. Multiple files of the same type should be listed in sequence, i.e.: Supplementary Video 1, Supplementary Video 2, etc.	Filename Whole original file name including extension. i.e.: *Smith_ Supplementary_Video_1.mov*	Legend or Descriptive Caption Describe the contents of the file
Supplementary Table	Supplementary Table 1, Supplementary Table 2, Supplementary Table 3, Supplementary Table 4,	Olmo_et_al_SupplementaryTables.xlsx	**Supplementary Table 1 - Overview of small RNA libraries.** Metadata of small RNA libraries used in our

	Supplementary Table 5, Supplementary Table 6		study (SRA deposit ID, species, mosquito capture location, number of mosquitoes per pool, RNA treatment, and sequencing method). **Supplementary Table 2 - Small RNA assembly metrics.** Detailed information of assembled contigs with length >50nt and length >199 nt per small RNA library. **Supplementary Table 3 - Overview of contigs with similarity to viral sequences deposited in GenBank.** Detailed information of BLASTn or BLASTp hits from contigs matching viral sequences (length >199 nt) per small RNA library. **Supplementary Table 4 - Overview of CDHit clusters.**

			Number of contigs that compose each CDHit cluster. **Supplementary Table 5 - List of oligonucleotides used in this study.** **Supplementary Table 6 - Parameters used to model DENV transmission.**

6

Parent Figure or Table	Filename Whole original file name including extension. i.e.: *Smith_SourceData_Fig1.xls*, or *Smith_ Unmodified_Gels_Fig1.pdf*	Data description i.e.: Unprocessed western Blots and/or gels, Statistical Source Data, etc.
Source Data Fig. 1	Olmo_Source_Data_Figure1.xlsx	Statistical Source Data
Source Data Fig. 2	Olmo_Source_Data_Figure2.xlsx	Statistical Source Data
Source Data Fig. 3	Olmo_Source_Data_Figure3.xlsx	Statistical Source Data
Source Data Fig. 4	Olmo_Source_Data_Figure4.xlsx	Statistical Source Data
Source Data Fig. 5	Olmo_Source_Data_Figure5.xlsx	Statistical Source Data
Source Data Extended Data Fig./Table 1	Olmo_Source_Data_ExtDataFigure1.xlsx	Statistical Source Data
Source Data Extended Data Fig./Table 2	Olmo_Source_Data_ExtDataFigure2.zip	Source Data in fasta and Newick formats
Source Data Extended Data Fig./Table 3	Olmo_Source_Data_ExtDataFigure3.zip	Source Data of coverage plots in BED formats and bar plots in Excel table.
Source Data Extended Data Fig./Table 4	Olmo_Source_Data_ExtDataFigure4.zip	Statistical Source Data, unprocessed gel image

Source Data Extended Data Fig./Table 5	Olmo_Source_Data_ExtDataFigure5.zip	Statistical Source Data, unprocessed gel image
Source Data Extended Data Fig./Table 6	Olmo_Source_Data_ExtDataFigure6.xlsx	Statistical Source Data
Source Data Extended Data Fig./Table 7	Olmo_Source_Data_ExtDataFigure7.xlsx	Statistical Source Data
Source Data Extended Data Fig./Table 8	Olmo_Source_Data_ExtDataFigure8.zip	Statistical Source Data, annotation of histone genes in gff3 format

7

Mosquito vector competence for dengue is modulated by insect specific viruses

Roenick P. Olmo[1,2]*, Yaovi M. H. Todjro[1]*, Eric R. G. R. Aguiar[1,3]*, João Paulo P. de Almeida[1], Flávia V. Ferreira[1], Juliana N. Armache[1], Isaque J. S. de Faria[1], Alvaro G. A. Ferreira[4], Siad C. G. Amadou[1], Ana Teresa S. Silva[1], Kátia P. R. de Souza[1], Ana Paula P. Vilela[1], Antinea Babarit[2], Cheong H. Tan[5], Mawlouth Diallo[6], Alioune Gaye[6], Christophe Paupy[7], Judicaël Obame-Nkoghe[8,9], Tessa M. Visser[10], Constantianus J. M. Koenraadt[10], Merril A. Wongsokarijo[11], Ana Luiza C. Cruz[12], Mariliza T. Prieto[13], Maisa C. P. Parra[14], Maurício L. Nogueira[14,15], Vivian Avelino-Silva[16], Renato N. Mota[17], Magno A. Z. Borges[18], Betânia P. Drumond[12], Erna G. Kroon[12], Mario Recker[19,20], Luigi Sedda[21], Eric Marois[2], Jean-Luc Imler[2] & João T. Marques[1,2#]

[1]Department of Biochemistry and Immunology, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, 31270-901 Belo Horizonte, Brazil.

[2]Université de Strasbourg, CNRS UPR9022, INSERM U1257, 67084 Strasbourg, France.

[3]Department of Biological Science (DCB), Center of Biotechnology and Genetics (CBG), State University of Santa Cruz (UESC), 45662-900, Ilhéus, Brazil.

[4]Mosquitos Vetores: Endossimbiontes e Interação Patógeno-Vetor, Instituto René Rachou-Fiocruz, Belo Horizonte 30190-002, MG, Brazil

[5]Environmental Health Institute, Vector Biology & Control Division, National Environment Agency, 138667, Singapore.

[6]Pôle de Zoologie Médicale, Institut Pasteur de Dakar, Dakar, Senegal

[7]Maladies Infectieuses et Vecteurs : Écologie, Génétique, Évolution et Contrôle (MIVEGEC); Université de Montpellier, Institut de Recherche pour le Développement, CNRS, 34394 Montpellier, France.

[8]Laboratoire de Biologie Moléculaire et Cellulaire, Département de Biologie, Université des Sciences et Techniques de Masuku, 901 Franceville, Gabon.

[9]Écologie des Systèmes Vectoriels, Centre Interdisciplinaire de Recherches Médicales de Franceville, 769 Franceville, Gabon.

[10]Laboratory of Entomology, Wageningen University & Research, 6708, Wageningen, The Netherlands.
[11]Central Laboratory of the Bureau of Public Health, Paramaribo, Suriname.
[12]Department of Microbiology, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais (UFMG), 31270-901, Belo Horizonte, Brazil.
[13]Secretaria Municipal de Saúde, Seção de Controle de Vetores, Santos City Hall, 11013-151, Santos, Brazil.
[14]Laboratory of Research in Virology, Faculdade de Medicina de São José do Rio Preto (FAMERP), 15090-000, São José do Rio Preto, Brazil.
[15]Departament of Pathology, University of Texas Medical Branch, Galveston, TX, USA.
[16]Department of Infectious and Parasitic Diseases, Faculdade de Medicina da Universidade de São Paulo (FMUSP), 01246-903, Cerqueira Cesar, Brazil.
[17]Health Surveillance (Zoonosis Control), Brumadinho City Hall, 35460-000, Brumadinho, Brazil.
[18]Center for Biological and Health Sciences, Universidade Estadual de Montes Claros, 39401-089, Montes Claros, Brazil.
[19]Centre for Ecology and Conservation, University of Exeter, Penryn Campus, Penryn TR10 9FE, United Kingdom
[20]Institute of Tropical Medicine, Universitätsklinikum Tübingen, Tübingen, Germany
[21]Lancaster Medical School, Lancaster University, Lancaster, LA1 4YG, United Kingdom.
* These authors contributed equally to this work.
[#] Corresponding author, email: jtm@ufmg.br; joao.marques@unistra.fr

ABSTRACT

***Aedes aegypti* and *Aedes albopictus* mosquitoes are the main vectors for dengue virus (DENV) and other arboviruses, including Zika virus (ZIKV). Understanding the factors that affect transmission of arboviruses from mosquitoes to humans is a priority, because it could inform public health and targeted interventions. Reasoning that interactions among viruses in the vector insect might affect transmission, we analysed the viromes of 815 urban *Aedes* mosquitoes collected from 12 countries worldwide. Two mosquito-specific viruses, Phasi Charoen-like virus (PCLV) and Humaita Tubiacanga virus (HTV) were the most abundant in *A. aegypti* worldwide. Spatiotemporal analyses of virus circulation in an endemic urban area revealed a 200% increase in chances of having DENV in wild *A. aegypti* mosquitoes when both HTV and PCLV were present. Using a mouse model in the laboratory, we showed that the presence of HTV and PCLV increased the ability of mosquitoes to transmit DENV and ZIKV to a vertebrate host. By transcriptomic analysis, we found that, in DENV infected mosquitoes, HTV and PCLV block the downregulation of histone H4, which we identify as an important pro-viral host factor in vivo.**

MAIN

Dengue fever is the fastest growing vector-borne disease worldwide and causes an estimated 400 million new infections every year[1–4]. In addition, over the past decades, several other arboviruses, including ZIKV and chikungunya (CHKV), have emerged and caused a substantial burden of disease. Increased transmission of arboviruses has been underpinned by increased geographic reach of the main vector mosquitoes, *A. aegypti* and *A. albopictus* [4,5], mainly due to climate change, because warming produces ideal conditions for mosquitoes. Vector abundance, assessed using cross-sectional surveys, has long been used as a proxy for infection risk, but the incidence of arbovirus infection does not directly correlate with mosquito abundance[6]. We still lack a complete understanding of the factors that affect rates of transmission to humans.

Virologic surveillance of adult *Aedes* mosquitoes by metagenomic analysis can lead to early identification of circulating arboviruses, and help raise preparedness to inform public health measures that can curtail or even prevent outbreaks[7]. In addition to arboviruses, these surveillance efforts have also identified an enormously diverse set of insect-specific viruses (ISVs) in *Aedes* mosquitoes [8–12]. Although ISVs do not infect vertebrates, they have been shown to affect the capacity of the mosquito to be infected,

maintain and transmit arboviruses, which together comprise vector competence and will therefore affect the incidence of infection in humans [7,13,14].

In order to carry out a comprehensive characterization of the viromes of mosquitoes that can harbour arboviruses, and inform risk assessment and public health strategies to mitigate arbovirus disease, we collected in the wild and performed metagenomic analysis of over 800 adult *Aedes* mosquitoes, and report our findings here.

RESULTS

Virome analysis of *Aedes* mosquitoes

Adult *A. aegypti* and *A. albopictus* mosquitoes were collected from the field in 12 different sites from six countries on four continents (**Fig. 1a**). In total, 815 adult mosquitoes were pooled according to species, location and date of collection, resulting in 91 samples derived from 69 *A. aegypti* and 22 *A. albopictus.* Details of the pools are described in **Supplementary Table 1**. Whole mosquito samples were used to extract RNA and construct small RNA libraries that were sequenced and analyzed using a shotgun metagenomic strategy previously optimized to detect viruses (See ref. [8]). Briefly, this strategy is based on the detection of virus-derived small RNAs that are used to assemble longer contiguous sequences for further characterization. In total, we identified putative viral 1448 contigs present in our mosquito samples (**Fig. 1b,c**). Data curation (described in **Fig. 2a** and in the methods section) suggested that these contigs represent at least 12 different viruses based on the phylogeny of polymerase genes (**Fig. 2b**), including seven known viruses previously identified as ISVs. Out of these, three remain unclassified while the other four belong to the *Phenuiviridae*, *Xinmoviridae*, *Bunyaviridae* and *Flaviviridae* families **(Fig. 2b)**. No known arboviruses were detected in our metagenomic analysis. The five remaining viral polymerase sequences showed low similarity to the closest known reference in Genbank. Phylogenetic analyses confirmed that they are likely new viral species (**Extended Data Fig. 2**), belonging to the *Partitiviridae*, *Totiviridae*, *Rhabdoviridae*, *Narnaviridae* and *Virgaridae* families (**Fig. 2b**). These viruses were named according to their classification (**Fig. 2b**). All new viruses were most closely related to known ISVs (**Extended Data Fig. 2**), but their final classification requires biological characterization.

All 12 identified viruses, 7 known and 5 new, had RNA genomes, either single-stranded (of positive and negative polarity) or double-stranded (**Fig. 2b**). The small RNA profile observed for these viruses shows clear production of siRNAs (**Extended Data Fig. 3**),

which results from the activity of the RNAi pathway during active viral replication in the mosquito host [8,15–19]. Viruses detected in *Aedes* mosquitoes were strictly species-specific and often associated with specific locations (**Fig. 2c**). Out of the 12 identified viruses, 10 were found in *A. aegypti* and two in *A. albopictus* suggesting a less diverse virome in the latter even when accounting for a lower number of samples. Indeed, looking at the diversity of the mosquito virome per country, a single virus species was detected in each *A. albopictus* population while 4-6 different viruses were present in *A. aegypti* (**Fig. 2d**). In addition, comparing different mosquito species that were collected from the same sites in Caratinga, Montes Claros, Lope, Franceville and Singapore, we observe that *A. aegypti* had higher virome diversity than *A. albopictus* in 4 out of 5 cases (**Fig. 2c**).

Using the small RNAs mapping to each virus contig as a proxy for abundance, we found that viral loads varied for the same virus in different locations and also between different viruses (**Extended Data Fig. 4a**). Detection of these new viruses in the original RNA samples from wild mosquitoes used for the metagenomic analysis was confirmed by RT-qPCR and conventional PCR (**Extended data Fig. 4b,c**). These results validated our small RNA sequencing strategy. Detection of Orbis virgavirus (OVV) by RT-qPCR failed in samples from Suriname, but this was attributed to polymorphisms in the primer annealing region (**Extended data Fig. 4d**). For each virus, the viral load determined by small RNA abundance corresponded with qRT-PCR detection. However, the relative quantification of viruses by qRT-PCR and small RNA abundance was unique for each virus, with some being underestimated (PCLV) or overestimated (Aslam narnavirus) (**Extended data Fig. 4e**).

Biogeography of mosquito viruses

Three of twelve locations had five or more viruses circulating in the local *A. aegypti* population: Santos, Paramaribo and Singapore (**Fig. 2c**). Notably, these are all port cities, which are likely to have a continuous influx of mosquitoes. Three different lines of laboratory *A. aegypti* mosquitoes lacked any viruses according to our analysis (**Fig. 2c**). Most of the viruses we detected were present in mosquitoes at single sites but five were present on at least two continents (**Fig. 2d**). In *A. aegypti*, two known ISVs, PCLV and HTV, were present in more than half of the samples (**Fig. 2c**), with the remaining eight viruses found in less than 20% of the samples. No viruses were found with a prevalence higher than 20% in *A. albopictus* and only one was present in multiple sites (Fig. **2c,d**). HTV and PCLV in *A. aegypti* had the highest viral loads of all viruses in either mosquito species (**Extended Data Fig. 4**) but were not present in any

samples of *A. albopictus*. Notably, HTV and PCLV were either absent, or present at very low viral loads in *A. aegypti* mosquitoes collected in Africa (**Fig. 2c**), where transmission of arboviruses, such as DENV and ZIKV is low (**Extended data Fig. 4f**)[1,2]. High loads of HTV and PCLV were observed in mosquitos sampled in areas with high DENV/ZIKV incidence, namely Asia and South America (**Extended data Fig. 4f**)[1,2]. We hypothesized that there was a positive association between ISVs and arboviruses, which was unexpected (competition between RNA viruses in the same host would be more likely).

Circulation of ISVs and arboviruses in the wild

In order to examine the spatiotemporal dynamics of the two major resident viruses in wild mosquitoes, HTV and PCLV, we chose to focus on one of the 12 sites used for the metagenomic analysis using a collection of archived mosquito RNA samples, 515 *A. aegypti* and 24 *A. albopictus*, previously collected over a year, August 2010 to July 2011, in Caratinga city, southeast Brazil (**Fig. 3a**). This dataset was previously used to assess DENV circulation in an endemic urban area[20]. Based on our metagenomics approach, we detected three viruses in Caratinga mosquitoes, OVV, HTV and PCLV (**Fig. 2c**), which we confirmed using RT-qPCR. OVV, HTV and PCLV were detected in wild *A. aegypti* but were absent from *A. albopictus,* even though both species were often captured in the same traps (**Fig. 3b**). OVV was only detected in three individual mosquitoes, and we focused our analyses on HTV and PCLV, which were present at prevalence of 61% and 83%, respectively (**Fig. 3b,c**). Based on this survey of individual mosquitoes by RT-qPCR, we confirmed that HTV and PCLV were highly prevalent in natural mosquito populations during the whole period of collections and independent of the location within the city (**Fig. 3d** and **Extended Data Fig. 5a**). Moreover, we observed a strong positive association between the presence of HTV and PCLV in mosquitoes (p<1E-10, chi-squared test), suggesting that co-infection might be advantageous for these viruses.

HTV and PCLV are presumed to be ISVs although they are poorly characterized to date [8,21,22]. We detected HTV and PCLV in different tissues, including the salivary glands, which suggested that they could be transmitted by a mosquito bite (**Extended Data Fig. 5b,c**). To assess the possibility that HTV and PCLV could be transmitted to humans, we analyzed human blood samples collected concomitantly with mosquitoes in Caratinga city, southeast Brazil from February to July 2011. Human blood samples and mosquitoes from Caratinga were previously analyzed by RT-qPCR for the presence of DENV [20] and these data were used for comparison. Plotting the numbers of this previous analyses, we observe that DENV was detected in less than

5% of mosquitoes, but in more than 30% of human blood samples (**Fig. 3c**). We did not detect HTV or PCLV in RNA extracted from human blood samples from Caratinga, despite their high prevalence in mosquitoes, suggesting that these viruses are unable to productively infect humans (**Fig. 3c**). Furthermore, HTV and PCLV do not grow in mammalian cell lines, such as Vero cells, reinforcing the idea that they are ISVs (**Extended Data Fig. 5d,e**). HTV and PCLV were detected in mosquito eggs, which suggests that these viruses are maintained in mosquitos by vertical transmission (**Extended Data Fig. 5c**).

Our data indicate that HTV and PCLV are not infectious to humans, but we wanted to understand whether they may affect how arboviruses are transmitted. In the dataset of mosquitoes from Caratinga city analyzed here, we observed a statistically significant enrichment of HTV and PCLV in individuals that also harbored DENV (**Fig. 3E**). Statistical analyses revealed that both HTV alone (OR 2.59, [1.09,716] 95%CI) and HTV/PCLV co-infections (OR 3.06, [1.29,8.46] 95% CI) are both associated with the presence of DENV in mosquitoes, whereas PCLV alone had no statistically significant association with DENV. However, due to positive association between HTV and PCLV that we identified in mosquitoes from Caratinga (p<1E-10, chi-squared test, described above), it is hard to dissect the contribution of each virus. We note that our analysis is based on a small sample of mosquitoes, it and could be affected by physical, ecological or environmental factors. However, in Caratinga, a town that is only 3 km wide and 3 km long, ecological conditions are likely homogenous on the same collection date. Population density was not found to be important in a previous study from our group[23]. In addition, our analyses found neither geographic nor temporal patterns of virus distribution in Caratinga (**Fig. 3d** and **Extended data Fig. 5a**). Thus, our observations using field samples strongly indicate a positive interaction between two ISVs, HTV and PCLV, and the arbovirus DENV in mosquitoes.

HTV and PCLV increase arbovirus replication in mosquitoes

Since we had access to ovitraps from Rio de Janeiro, where HTV and PCLV were also found at high prevalence, we obtained a few hundred eggs from wild *A. aegypti* mosquitoes and reared them in the laboratory for 2 generations. From the F2 of lab-reared population, we pooled eggs from 5 individual females that were either free of any virus, or carried both HTV and PCLV to produce two separate mosquito lines (see Methods for details). We were unable to isolate lines carrying only HTV or PCLV, reinforcing the strong association observed in our wild sample cohort.

We exposed the two separate mosquito lines that were ISV free or co-infected with HTV/PCLV to blood feeding on mice previously infected with DENV or ZIKV using an infectious blood meal. We found that mosquitoes carrying HTV and PCLV had similar prevalence and viral loads of DENV in the midgut at 4, 8 and 14 days post feeding (d.p.f.) to that present in ISV-free mosquitoes (**Fig. 4a,b**). At 14 d.p.f., we observed a trend towards higher viral load of DENV in the midgut of mosquitoes carrying HTV/PCLV but this trend was not statistically significant (**Fig. 4b**). In the carcass of mosquitoes, we observed a five-fold significant increase in DENV levels at 8 and 14 d.p.f. in individuals carrying HTV/PCLV compared to ISV free controls (**Fig. 4c**). Mosquitoes with HTV/PCLV also displayed higher susceptibility to ZIKV than ISV-free mosquitoes (**Fig. 4d-f**). ZIKV RNA levels were significantly increased at 4, 8 and 14 d.p.f. in the midgut of mosquitoes carrying HTV and PCLV compared to ISV free controls (**Fig. 4e**). In the carcass, we observed a fivefold increase in dissemination at 4 d.p.f. and significantly 10-fold higher viral loads at 8 d.p.f. in the presence of HTV and PCLV. Overall, our results demonstrate increased systemic DENV and ZIKV infection in mosquitoes carrying HTV/PCLV compared to virus-free controls.

To further investigate the specific effect of HTV and PCLV during the systemic phase of infection, we directly injected ZIKV into the mosquito haemocele, which bypasses the stage of midgut infection (**Extended Data Fig. 6a**). Here, we also observed increased viral replication in the carcass of HTV/PCLV carrying wild mosquitoes compared to ISV-free controls (**Extended Data Fig. 6a-c**).

Our mosquito colonies established from individuals with and without ISVs were derived from the same but highly heterogenous wild population. It is therefore possible that our selection generated colonies composed of individuals that differed with regards to their genetic backgrounds in addition to the presence of ISVs. To rule out a role for the genetic background, we next performed experiments with laboratory mosquitoes that are genetically more homogenous. Laboratory mosquitoes were artificially infected with HTV and PCLV to test whether ISVs have a direct impact on the susceptibility to arboviruses (**Extended data Fig. 6d**). Notably, HTV and PCLV loads and tissue tropism during artificial injection were similar to naturally infected mosquitoes after 8 days post injection (**Extended Data Fig. 5b**). Artificially infected laboratory mosquitoes had increased systemic ZIKV RNA levels at 8 d.p.f. compared to controls, similar to what we observed for lines carrying HTV and PCLV derived from wild populations (**Extended data Fig. 6d-f**). Increased systemic viral replication was also observed when laboratory mosquitoes were artificially infected with HTV and PCLV before being injected with ZIKV (**Extended Data Fig. 6g-i**). Although artificially infected lab

mosquitoes did not show increased ZIKV replication in the midgut, this can be explained by the fact that naturally infected mosquitoes have more marked effects due to presence of HTV and PCLV throughout development (**Extended Data Fig. 5b**).

Notably, in wild mosquitoes, ZIKV infection had a positive impact on PCLV levels in the midgut of infected mosquitoes (**Extended data Fig. 6j,k**), which suggests a mutual beneficial interaction between these viruses. As mentioned before, HTV was not detected in the midgut even in the presence of ZIKV (**Extended data Fig. 6l**). Neither HTV nor PCLV were consistently affected by ZIKV infection in the carcass, although we observed an increase in PCLV levels and a reduction in HTV levels in single time points (**Extended data Fig. 6m,n**).

Transmission of DENV and ZIKV is increased by ISVs

We tested whether increased ZIKV and DENV levels in mosquitoes carrying HTV and PCLV led to increased amounts of arboviruses in mosquito salivary glands, and increased transmission to a vertebrate host. Wild HTV/PCLV positive mosquitoes showed faster kinetics and higher prevalence of ZIKV infection in salivary glands compared to ISV-free controls (**Fig. 4g** and **Extended Data Fig. 6o-p**). We simulated vectorial transmission in a susceptible animal model using mice deficient for type I and type II interferon receptors[24,25] (**Fig. 4h**). We opted to test ZIKV transmission because the mouse model for this virus is more robust than for DENV. Mice were incubated with ZIKV-infected mosquitoes at 6, 8 and 12 d.p.f. and viremia was analyzed in these animals. No viremia was observed in mice bitten by mosquitoes at 6 d.p.f. (**Fig. 4i**). Mosquitoes were able to efficiently transmit ZIKV to 5 out of 5 mice at 8 d.p.f. but only in the presence of HTV and PCLV (**Fig. 4i**). At 12 d.p.f., mosquitoes with or without HTV/PCLV were able to equally transmit ZIKV to 3 out of 3 mice (**Fig. 4i**). Thus, the presence of ISVs is associated with shortening of the extrinsic incubation period (EIP) of ZIKV, which is the time required for infected mosquitoes to become infectious to a vertebrate host. While mosquitoes carrying HTV and PCLV were able to transmit ZIKV between 7 and 8 days, ISV free individuals required between 9 and 12 days. Thus, the presence of ISVs in mosquitoes could lead to shortening of the EIP between 1 and 5 days although our experiments did not allow us to pinpoint the exact difference. Furthermore, mosquitoes carrying HTV and PCLV and analyzed at both 8 and 12 d.p.f. had significantly higher ZIKV levels compared to ISV free controls (**Fig. 4j**).

To further elucidate the impact of HTV and PCLV infection on the EIP we applied a previously developed mathematical model [26,27]. Our modelling demonstrates that even

small changes in EIP could have a large impact on the number of human infections (**Fig. 4k**). For example, shortening the EIP from 10 days to 8 days, within the range of the difference we observed between mosquitoes with and without ISVs in the laboratory, could lead to a 5-fold increase in the number of infections (**Fig. 4k**), which can be explained by the short average life expectancy of *Aedes* mosquitoes in the wild. As a consequence, arboviral prevalence in mosquitoes is also increased due to both the increased availability of infected humans and faster viral kinetics inside mosquitoes (**Fig. 4l**), providing a link between our field observations and laboratory experiments. Although this model was parameterized for DENV transmission, the results should broadly hold for other arboviral diseases transmitted by the same mosquito vector, including ZIKV.

HTV and PCLV modulate histone H4 expression in mosquitoes

In order to probe the biological mechanisms by which ISVs affect systemic dissemination of arboviruses in *A. aegypti* mosquitoes, we analyzed the transcriptome of RNA harvested from entire mosquito carcasses at different times after DENV infection (4, 8 and 14 d.p.f.). Overall, we found that the presence of HTV and PCLV had little effect on the transcriptome of DENV infected mosquitoes (**Extended data Fig. 7a**). Only 100/ 10000 genes analysed were significantly up- or down regulated and less than 10 were common between time points (**Extended data Fig. 7b**). Of interest, genes associated with known antiviral pathways, such as Toll, IMD, Jak-STAT, autophagy and RNA interference did not show any consistent differences in expression (**Extended data Fig. 7c**). Next, we compared the transcriptome of DENV infected and non-infected individuals from groups of mosquitoes carrying HTV and PCLV, or virus-free controls, using Gene Set Enrichment analysis (GSEA)[28] (**Fig. 5a**). We focused our analysis on the carcass of mosquitoes infected by DENV at 8 and 14 d.p.f. where the presence of HTV and PCLV had the strongest effect (**Fig. 5a**). This analysis identified 7 biological pathways that were significantly affected both by the presence of HTV / PCLV and DENV infection in at least one time point (**Fig. 5b**). Notably, all pathways were downregulated during DENV infection and upregulated by the presence of HTV and PCLV, as indicated by the enrichment score (NES) (**Fig. 5b**). Out of these, four pathways were significantly affected in at least 6 out of 8 comparisons: *nucleosome*, *nucleosome assembly*, *DNA templated transcription initiation* and *protein heterodimerization activity* (**Fig. 5b**). Analysis of genes responsible for the significant enrichment, showed that they were almost the same for these 4 biological pathways (**Fig. 5c**). Indeed, histones represented the majority of genes differentially regulated by DENV infection and the presence of HTV and PCLV,

with histone H4 topping the list (**Fig. 5c**). Thus, we used RT-qPCR to analyze histone H4 expression and validate our observations in independent experiments using ZIKV. Histone H4 expression was significantly downregulated by ZIKV infection in the carcass of infected mosquitoes in a time-dependent manner (**Fig. 5d,e**), which was prevented by the presence of HTV and PCLV (**Fig. 5e**). Furthermore, levels of histone H4 were significantly higher in the presence of HTV and PCLV at every time point tested when compared to ISV free controls (**Fig. 5e**). We also observed upregulation of histone H4 in the midgut, but only at 4 d.p.i. (**Extended Data Fig. 7d,e**). Importantly, differential expression of histone H4 between mosquitoes with or without HTV/PCLV was only observed in the presence of DENV and ZIKV infections (**Fig. 5f,g**). We showed that artificial infection of laboratory mosquitoes with HTV and PCLV alone did not significantly affect histone H4 expression (**Extended Data Fig. 7f,g**).

Histone genes are highly conserved, often found in multiple copies that lack polyadenylation signals [29]. Yet, there are non-canonical histone genes that possess polyadenylation signals. The genome of *A. aegypti* encodes at least 299 histone genes in the assembled chromosomes (chr.) and another 135 copies present in extra Supercontigs (**Extended Data Fig. 8a**). In comparison, humans encode only about 80 histone genes in total[29], despite having a larger genome. Most histone genes (267 out of 299) were found in a single cluster on chr. 3 (**Extended Data Fig. 8a,b**). With regards to histone H4 we identified 66 genes in total and 59 in the cluster on chr. 3 (**Extended Data Fig. 8a**). Histone H4 genes showed high similarity, with almost 100% aminoacid conservation but some sequence variation at the nucleotide level, especially in the copies outside of the chr.3 cluster (**Extended Data Fig. 8c,d**). Histone H4 genes that have clear polyadenylation signals were the most detected in our dataset and in the available transcriptome of mosquito tissues since they all were prepared using polyA selection (**Extended Data Fig. 8d**). Interestingly, the effect of HTV and PCLV on histone H4 RNA levels was not significant when we analyzed polyadenylated copies, which represented less than 10% of all histone H4 expression (**Extended Data Fig. 8e**). This suggests HTV and PCLV may primarily affect non-polyadenylated histones that are coordinately synthesized with DNA replication during the S-phase of cell division and are stable after incorporation into chromatin [30].

Silencing of histone H4 and DENV replication in mosquitoes

We applied dsRNA-mediated gene silencing to knock down histone H4 expression in adult mosquitos prior to infection with DENV or ZIKV (**Fig. 5h** and **Extended Data Fig. 8f,g**). Mosquitos silenced for histone H4 before infection had significantly reduced

DENV levels in the midgut at 4 and 8 d.p.i., although the difference was not significant at the later time point (**Fig. 5i**). As a result, we observed slower kinetics of infection in the carcass with lower prevalence at 4 d.p.i. and reduced viral loads at later time points (**Fig. 5j**). These results indicate that histone H4 is an important pro-viral host factor during DENV infection. As observed for the effect of HTV and PCLV, histone H4 affected the kinetics of DENV infection but was not essential for viral replication. Together, our data suggest that two highly prevalent ISVs, HTV and PCLV, affect mosquito vector competence for DENV and ZIKV by preventing the downregulation of histone H4, a novel proviral host factor. To test this hypothesis, we directly targeted histone H4 in mosquitoes carrying HTV and PCLV by using dsRNA mediated gene silencing. We did not observe any changes in histone H4 expression in mosquitoes injected with cognate dsRNA (**Extended Data Fig. 8h**). This is in stark contrast to the efficient silencing triggered by the same dsRNA sequence in mosquitoes that did not carry HTV and PCLV (**Extended Data Fig. 8g**). Notably, another dsRNA targeting the AGO2 nuclease that is central to the RNA interference pathways was able to trigger efficient silencing in the same mosquitoes carrying HTV and PCLV (**Extended Data Fig. 8i**). These results again point to a specific effect of HTV and PCLV on histone H4 expression in mosquitoes.

DISCUSSION

We report positive interactions between ISVs and arboviruses in mosquitoes in the wild and in the laboratory. Previously ISVs have mainly been reported to interfere with arbovirus replication in mosquitoes (superinfection exclusion) [7,31,32]. As HTV and PCLV are the most abundant ISVs that we detected in wild *A. aegypti* mosquitoes, it is feasible that they can have a substantial impact on the global transmission of DENV and ZIKV.

We also showed that HTV and PCLV increase histone H4 expression during DENV infection and that histone H4 is a pro-viral host factor for the replication of DENV in mosquitoes. We propose that ISVs increase DENV infection through upregulation of histone H4. However, we were unable to establish a direct connection between the regulation of histone H4 expression and the increase in vector competence. Regarding the role of histone H4 as a putative proviral factor, it is worth mentioning that C protein from flaviviruses interacts with histones and is capable of interfering with nucleosome assembly [33]. Recent work further suggests that the C protein of Yellow fever virus and possibly other flaviviruses mimics the tail of histone H4 and regulates gene expression to favour infection [34]. Thus, downregulation of histone H4 may be part of a coordinated host response to limit the ability of flaviviruses to control gene expression, which could

be counteracted by HTV and PCLV. We observed no major changes in gene expression in the presence of HTV and PCLV, which suggests that the proviral role of histone H4 is not achieved through major regulation of gene expression. Rather, replication-dependent histones seem to be preferentially regulated by HTV and PCLV, which could point to a mechanism involving cell division. Notably, replication-dependent histones are targeted by RNA interference [35], which could provide a connection with the role of this pathway in the antiviral defense of mosquitoes [16,36].

Although previous studies have reported interactions between ISVs and arboviruses, most were performed in cell lines[37–40]. One exception is Nhumirim virus (NHUV) and Cell fusing agent virus (CFAV), which were shown to interfere with replication of arboviruses in the same family as West Nile virus, DENV and ZIKV [31,41,42]. Interestingly, in cell lines, PCLV either inhibited or did not affect the replication of ZIKV[21,22]. Also since we could not test the presence of PCLV alone in adult mosquitoes, we cannot rule out that HTV has the predominant proviral effect.

Further work will be needed to understand how HTV and PCLV regulate histone H4 during DENV and ZIKV infection, and whether this mechanism affects other arboviruses such as CHIKV. Understanding how ISVs affect vector competence could reveal alternative strategies for controlling arbovirus transmission.

METHODS

Ethics statement

All procedures involving vertebrate animals were approved by the ethical review committee of the Universidade Federal de Minas Gerais (CEUA 337/2016 and 118/2022 to J.T.M.). Mosquito were collected in Gabon under the research authorization AR0013/17/MESRS/CENAREST/CG/CST/CSAR delivered by CENASREST. Unlinked anonymous testing of human blood samples was approved by the ethics committee in research (COEP) of Universidade Federal de Minas Gerais (number 415/04 to EGK).

Human blood samples

Human samples were previously reported[20]. Forty-four blood samples were collected from patients that sought medical attention in the city of Caratinga between February and July of 2011 by professional nurses as part of a city surveillance plan. Blood samples were mixed with EDTA as an anticoagulant and stored at 4 °C. Serum was obtained from blood samples and inactivated at 56 °C for 30 min. Total RNA extraction from human blood samples was performed using Trizol LS Reagent (Invitrogen).

Data reporting

No statistical methods were used to predetermine sample size. Experiments were not randomized and investigators were not blinded to allocation during experiments and outcome assessment.

Mosquito collection in the field

Locations of mosquito collections are described in the **Supplementary Table 1**. Field traps were used to collect adult mosquitoes that were further identified using morphological characteristics. Whole mosquitoes were ground in TRIzol (Invitrogen) and kept refrigerated prior to RNA extraction. Collection and processing of individual mosquitoes from Caratinga for spatiotemporal analysis of virus circulation were previously reported[20]. RNA samples of these mosquitoes were re-analyzed in this work.

RNA extraction

RNA was isolated using TRIzol reagent (Invitrogen) following the manufacturer protocol with minor modifications. Briefly, individual mosquito samples or tissues were collected in 1,5 mL tubes, 3-5 glass beads (1 mm diameter) and ice-cold TRIzol were

added before being homogenized in a Mini-BeadBeater-16 (Biospec©) for 90 seconds. Glycogen (Ambion) was added (10µg per sample) to facilitate pellet visualization upon RNA precipitation. RNAs were resuspended in RNAse-free water (Ambion) and stored at -80˚C.

Small RNA library construction

Different strategies for library construction were implemented and are indicated in **Supplementary Table 1**. The strategy was determined according to RNA quality evaluated by the 2100 Bioanalyzer system (Agilent). Libraries were built using total RNA or size selected small RNAs (18-30 nt), depending on quality and yield of the sample. In the case of low RNA yield, especially when the source was a single mosquito, total RNA was directly used as input for library preparation. For samples with more than 20 ug of RNA available, small RNAs were selected by size (18–30 nt) on a denaturing PAGE. For samples with more than 20 ug of total RNA that displayed a degradation profile (*i.e.*, lack of sharp ribosomal RNA peaks), total RNA was subjected to oxidation using sodium periodate [43,44], prior to size selection. Oxidized and non-oxidized size selected RNAs (18-30 nt) were used as input for library construction. In all cases, libraries were prepared utilizing the TruSeq® Small RNA Library Prep Kit (Illumina®) or NEBNext® Multiplex Small RNA Library Prep Set for Illumina® (New England BioLabs inc.) following protocols recommended by the manufacturers. Libraries were pooled and sequenced at the GenomEast sequencing platform at the Institut de Génétique et de Biologie Moléculaire et Cellulaire in Strasbourg, France.

Small RNA-based metagenomics for virus identification

After sequencing, raw sequenced reads from small RNA libraries were submitted to adaptor trimming using cutadapt [45] v1.12, discarding sequences with low Phred quality (< 20), ambiguous nucleotides and/or with length shorter than 15 nt. Remaining sequences were mapped to reference sequences of *A. aegypti* (AaeL5) [46] or *A. albopictus* [47] using Bowtie [48] v1.1.2 allowing no mismatches. Size profiles of small RNAs matching reference sequences and 5' nt frequency were calculated using in-house Perl v5.16.3, BioPerl library v1.6.924 and R v3.3.1 scripts. Plots were made in R using ggplot2 v2.2.0 package. Sequences that did not present similarities with bacteria or the host genomes were used for contig assembly and subsequent analyses. Assembly was performed essentially as previously described [8] with the following changes: (1) We replaced Velvet [49] assembler by SPAdes [50] on the second round of contig assembly. (2) Assembled contigs ranging from 50 to 199 nt were

characterized solely based on sequence similarity search against Viral RefSeq Database [51]. (3) Contigs greater than 200nt were characterized based on sequence similarity against the NCBI NT and NR databases and submitted to pattern-based strategies. (4) For manual curation of putative viral contigs, top 5 BLAST [52] hits were analyzed to rule out similarity to other organisms, ORF organization and small RNA size profile (distribution and coverage) were analyzed to differentiate between viruses. Contigs containing truncated ORFs and small RNA profiles without the presence of symmetric small RNA peaks at 21 nt were considered to be EVEs as described [15]. (5) Manually curated viral contigs were grouped using CD-HIT [53] requiring 90% of coverage with 90% of identity to remove redundancy. Representative contigs were used for co-occurrence analysis based on small RNA abundance on each of the small RNA libraries available. Contigs grouped into a single cluster (Hierarchical clustering based on Pearson correlation) were then used as trusted on SPAdes for a re-assembly step using all the libraries in which that viral sequence was found. In total, we assembled 7260 contigs larger than 200 nt (**Fig. 1b,** assembly metrics in **Supplementary Table 2**). 1448/ 7240 contigs were identified as putative viral sequences using sequence similarity searches against non-redundant nucleotide and protein databases (NT and NR, respectively) at GenBank (**Fig. 1b, Supplementary Table 3**). Although the number of contigs assembled per library varied greatly, we observed high abundance and diversity of viral contigs in most samples (**Fig. 1c**). Comparing results from the two mosquito species, the percentage of viral contigs was substantially smaller in *A. albopictus* libraries compared with *A. aegypti* (**Fig. 1c**). In addition, we noted more variation in the number of assembled contigs, and larger proportions of unknown contigs, in libraries from *A. albopictus*, probably because this species is less studied compared to *A. aegypti* (**Fig. 1c**). Most animal genomes contain integrated viral sequences known as endogenous viral elements (EVEs) that are transcribed and generate small RNAs [54–56]. In order to discriminate sequences of viruses from EVEs, we took advantage of the small RNA profile associated with ORF analysis and contig size (**Fig. 2a**) [15]. This filter identified 446 putative EVE sequences that were removed from the initial viral contigs (**Fig. 2a**). The remaining viral contigs, representing putative viruses, were grouped into 158 unique clusters (based on sequence similarity) (**Supplementary Table 4**). In 19 clusters, parts of contigs had significant similarity to two different viruses. Our results suggested that these were misassemblies and thus they were removed from further analyses (**Fig. 2a**). Contigs representing the remaining clusters were used to evaluate co-occurrence in the 91 small RNA libraries from *A. aegypti* and *A. albopictus* (**Extended Data Fig. 1**). The occurrence of contigs in each library was indicated by the normalized number of small

RNA reads mapped to each reference. Across the small RNA libraries, contigs that consistently co-occurred and shared similar expression profiles were considered probable fragments from the same viral genome (**Fig. 2a**). This analysis yielded a total of 12 clusters of co-occurring contigs and 3 single contigs that had no additional partners. Contigs from each cluster were further analyzed based on the closest reference sequence to determine the putative organization of fragments along the viral genome. This analysis showed that clusters #2 and #3 contained contigs belonging to the same virus, PCLV, and were considered together for further analyses (**Extended Data Fig. 1**). To further classify the clusters and single contigs, we focused on the ones that represented sequences encoding viral polymerases. We were able to identify clear polymerase sequences in each of the 11 clusters and in one out the 3 single contigs.

Phylogenetic analyses

Assembled viral contigs were submitted to analysis of conserved domains to identify RdRp-related regions using NCBI Conserved Domain Search (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). For each putative virus, the largest RdRp segment was used to identify virus relatives at NCBI sequence databases (NT and NR) using sequence-similarity searches through BLAST tool. Multiple alignments were performed using the MAFFT online tool [57] available at https://www.ebi.ac.uk/Tools/msa/mafft/. For putative new viruses identified at protein level (*Narnaviridae, Partitiviridae, Rhabdoviridae, Totivirdae* and *Virgaviridae*), amino acid sequences were selected, and phylogenetic analyses were carried out on CIPRES Portal version 3.3 (https://www.phylo.org/portal2) [58]. The best-fit model of protein evolution was selected in ModelTest-NG [59] for each viral species, using Maximum Likelihood (ML) method. For the virus from *Totiviridae* family, an additional strategy was also applied using nucleotide sequences where the best-fit model was defined using MEGA-X tool [60], and the tree was constructed using ML method. For all phylogenetic trees, clade robustness was assessed using bootstrap method (1000 pseudoreplicates) and trees were edited using iTOL version 5.7 [61] (https://itol.embl.de/).

Mosquitoes

Wild *A. aegypti* mosquitoes used in cage experiments were F2 to F5 generations derived from eggs collected in the Rio de Janeiro city (Urca neighborhood) in Brazil and were kindly provided by Dr. Rafael M. de Freitas from Fiocruz-RJ and Dr. Luciano A. Moreira from Fiocruz-MG. The laboratory Red-eye strain [62] was kindly provided by

Prof. Pedro C. Oliveira from the Universidade Federal do Rio de Janeiro - UFRJ. *A. aegypti* mosquitoes were maintained in a climatic chamber at 28°C and 70-80% relative humidity, in a 14:10 hour light:dark photoperiod, and 10% sucrose solution *ad libitum*. Mosquito cages contained individuals that emerged in a 48h-interval.

Generation of HTV+/PCLV+ and HTV–/PCLV– mosquito lines

Mosquito lines persistently infected with PCLV and HTV or non-infected counterparts were generated from F2 generations of wild mosquitoes. Three days after a blood meal, F2 mated females were individually isolated in tubes containing a filter paper and 0.5 cm of water and were allowed to lay eggs for 24 hours. Individual females were collected, and the total RNA was isolated using TRIzol (Invitrogen) following the standard protocol. Detection of HTV and PCLV was performed by RT-qPCR using the primers described in **Supplementary Table 5**. Eggs corresponding to 5 female mosquitoes infected with HTV, PCLV or both viruses were pooled prior to hatching. Pools of eggs from 5 females negative for both viruses were pooled similarly to create virus free lines. Subsequent detection of HTV and PCLV was performed to confirm correct identification of lines. Lines generated from females carrying a single virus (HTV or PCLV) were tested and found to carry both viruses. Therefore, only virus free and double infected lines were expanded for two more generations for experiments described in this work.

Artificial infection of naïve laboratory *A. aegypti* mosquitoes with HTV and PCLV

Extracts of naturally infected *A. aegypti* mosquitoes were used as source for HTV and PCLV since we were not able to produce these viruses in cell culture. Viral stocks were produced from pools of 15 *A. aegypti* naturally infected with HTV and PCLV or non-infected mosquitoes (virus-free controls), that were grinded using pestles in 1200 µL of L-15 Leibovitz medium (Gibco) supplemented with 10% fetal bovine serum (FBS). Samples were centrifuged at 3000 × g for 15 minutes at 4°C for clarification. Supernatants were collected and passed through a 0.22 µm filter, aliquoted, and stored at -80°C prior use. Infection with HTV and PCLV or mock control was performed by microinjecting 69 nL of the extract into naïve laboratory mosquitoes (*A. aegypti* RedEye strain) using a Nanoject II microinjector (Drummond).

Infection of Vero cells with HTV and PCLV

Filtered *A. aegypti* extracts (500 µL) containing HTV and PCLV (obtained as described above) were transferred into T-25 flasks containing 90% confluent Vero cells in non-supplemented DMEM medium. After one hour of viral adsorption, 4.5 mL of DMEM

medium supplemented with penicillin/streptomycin and 10% of FBS were added to cells, that were incubated at 37°C and 5% CO_2. 100 µL aliquots of the supernatant were collected during each passage on days 1, 3 and 5 after exposure to HTV and PCLV. Virus in the supernatant was assessed by RT-qPCR. Vesicular stomatitis virus (VSV) was added as a spike immediately before RNA extraction to be used as an internal control.

DENV and ZIKV propagation

Viral isolates of DENV1 (MV09) and ZIKV (PE243/2015)[63] were propagated in C6/36 *A. albopictus* cells or Vero cells respectively. For DENV1 propagation, C6/36 cells were maintained on L15 medium supplemented with 5% FBS (fetal bovine serum) and 1x Antibiotic-Antimycotic (Gibco) as described[20]. Cells were seeded to 70% confluence and infected at a multiplicity of infection (MOI) of 0.01 and maintained for 6 to 9 days at 28˚C. For ZIKV propagation, a similar procedure was followed using Vero cells that were maintained in DMEM medium supplemented with 10% FBS (fetal bovine serum) and 1x Antibiotic-Antimycotic (Gibco). Vero cells were seeded to a confluence of 70-80% infected with ZIKV at a MOI of 0.01 and maintained for 6 days in culture. For both viruses the supernatant was collected and clarified by centrifugation to generate virus stocks that were kept at -80˚C prior to use. Mock-infected supernatants used as controls were prepared under same procedure without virus infection. Titration of DENV1 was performed in BHK-21 cells while ZIKV was titrated in Vero cells, both using the plaque assay method to determine viral titer.

DENV and ZIKV infection in mosquitoes

Natural infection in mosquitoes was performed using mice deficient for interferon-I and interferon-II receptors as described [19]. Briefly, infection in AG129 mice was established by intraperitoneal injection of approximately 10^6 PFU of DENV1 or 10^6 PFU of ZIKV. Infected mice were anaesthetized 3 days post infection (peak of viraemia) using ketamine/xylazine (80/8 mg per kg) and placed on top of the netting-covered containers with adult mosquito females. Mosquitoes were allowed to feed on mice for 30 minutes to 1 hour, alternating on the same animal between cages every 10 min if two groups were to be compared. Viremia of each mouse was tested by RT-qPCR immediately after feeding and the experiment would be discarded if a discrepancy of more than 10x in viral load was observed. For infections by membrane feeding, 5-6 day old adult females were starved for 24h and fed with a mixture of blood and virus supernatant containing 10^7 PFU/mL of DENV serotype 1 utilizing a glass artificial feeding system covered with pig intestine membrane as previously described [19]. After

blood feeding, fully engorged females were selected and kept in standard rearing conditions until collection at different time points. Mosquitoes infected by injection were anesthetized with cold at 4˚C and kept on ice during the whole procedure. Virus stocks were diluted with L15 medium (Gibco) and injections were carried out using the microinjector Nanoject II (Drummond) with a volume of 69 nL. Each individual mosquito was injected with 16 PFU of DENV or ZIKV. Mosquitoes were harvested at different days post injection for dissection and RNA extraction. Tissues (midguts, salivary glands, and ovaries) were dissected in ice-cold 1x phosphate buffered saline (PBS) containing 0.01% Triton X-100 (Sigma). Remnants of mosquito tissues were considered as carcass, as illustrated in figure schemes. Tissues or mosquitoes were ground in TRIzol (Invitrogen) using glass beads as previously described [19] and kept frozen at –80˚C until RNA extraction as described above.

Mouse model for vectorial transmission of arboviruses

Groups of mosquitoes naturally infected with HTV and PCLV or PCLV/HTV-free siblings were fed on the same ZIKV-infected AG129 mouse as described above. One mouse was used per time point of transmission that was evaluated. Engorged females were kept after feeding and, three days later, allowed to lay eggs for 48 hours in dark cups containing 1 cm of water and a paper sheet attached to their walls. At 6-, 8-, and 12-days post feeding, 10 to 12 mosquito females of each group were allowed to feed on naïve AG129 mice for up to 3 hours. Mice were bled 3 days after mosquito biting and viral RNA levels were quantified by RT-qPCR. Viral loads were also quantified by RT-qPCR in engorged mosquitoes from each group after feeding on a naïve mouse.

dsRNA mediated gene silencing

RNA transcription was performed using T7 Megascript kit (Ambion) following the manufacturer's instructions. Briefly, template DNA containing T7 promoter sequences in both 5' and 3' extremities was obtained by RT-PCR for dsAGO2 and dsH4, or by PCR amplification from plasmid pDSAG (Addgene #62289) for dsGFP. Primer sequences are provided in **Supplementary Table 5**. Adult 4-day-old females were intrathoracically injected with 69 nl of a dsRNA solution (7.2 µg µl−1) diluted in annealing buffer (20 mM Tris-HCl pH 7.5, 100 mM NaCl) using a nano-injector Nanoject II (Drummond Scientific Company). Mosquitoes were allowed to recover for 48 h before further experiments. Once recovered, dsRNA-injected mosquitoes were allowed to feed or were microinjected with virus following the procedures described above.

RT-qPCR and RT-PCR

Total RNA extracted from individual mosquitoes or individual tissues were reverse transcribed using M-MLV reverse transcriptase (RT) using random primers (hexamers) for initiation. Negative controls were prepared following the same protocol without adding reverse transcriptase. cDNA was subjected to polymerase chain reaction (RT-PCR) using the kit GoTaq Hot Start Green PCR Master Mix (Promega) or quantitative polymerase chain reaction (RT-qPCR) utilizing the kit Power SYBR® Green Master Mix (Applied Biosystems) following manufacturer instructions. Results were expressed using the 2-ΔCt method relative to the endogenous control *rpL32*. Primer sequences are listed in **Supplementary Table 5**. In experiments designed to differentiate the expression of Histone H4 polyA transcripts versus non-polyA, cDNAs were synthesized starting from the same amount of total RNA in two independent reactions either using random hexamers or anchored oligo dT_{22} as reverse transcription initiators. Subsequent quantifications by RT-qPCR were performed as described above.

Mathematical model

To investigate the effect of variations in the extrinsic incubation period (EIP) on the cumulative incidence of DENV infections, we used a previously developed spatially explicit, individual-based meta-population model [28]. Briefly, humans and mosquitoes are defined with a unique state representing their current epidemiological status and, in case of humans, their infection history. Human individuals are considered to be either susceptible, exposed, infectious or recovered with respect to each serotype, allowing up to four sequential infections. For mosquitoes, only the susceptible, exposed, and infectious states of the epidemiologically relevant adult life-stage are considered. The sizes of the respective populations are kept constant with deaths being replaced by births. For simplicity, rather than accounting for seasonality through changes in mosquito densities or temperature-dependent variations in EIP, this is done here by varying daily mosquito biting rates and given as

$$a_v(t) = a_0(\beta + (1-\beta)\sin(t\pi/364)^4),$$

where *t* denotes time days and assuming a 364-day year. Both human and vector mortality rates are age-dependent, *i.e.*, the per capita risk of death is assumed to increase with age, which prevents individuals living beyond biologically reasonable ages. For computational efficiency, individuals' life expectancies are assigned at birth. Spatial structure is included by subdividing humans and mosquitoes into a spatially

organised sets of non-overlapping communities, where individuals mix homogeneously. Mosquitoes preferentially bite humans in their own community and adjacent communities but can also bite individuals of non-adjacent communities with low probability to account for (daily) human mobility and associated long-distant transmission events. Parameter values are listed in **Supplementary Table 6** and were chosen to capture the qualitative dynamics of DENV in a high-transmission setting with four co-circulating serotypes (DENV1-4).

The model exhibits pronounced demographic and epidemiological stochasticities that arise from the fully probabilistic nature of state transitions and result in significant inter-annual oscillations in both disease incidence and relative prevalence of the four co-circulating serotypes. To investigate the effect of shortening, or lengthening the extrinsic incubation period, we ran the model for a period of 100 years and recorded the total annual number of infections for the last 50 years, discarding the transient dynamics. Due to the stochastic nature of the model, we averaged the 50-year cumulative incidence over five model runs for each value of the EIP. Mosquito and human prevalence was equally evaluated over a 50-year time period but taken as the proportion of infected individuals on a single day during the seasonal peak, resulting in 50 individual data points.

Poly-A selection, RNA library construction, and transcriptomic analysis

RNA samples from individual mosquitoes were pooled and RNA quality was verified using the 2100 Bioanalyzer system (Agilent). mRNA libraries were constructed using the kits NEBNext® Poly(A) mRNA Magnetic Isolation Module and NEBNext® Ultra™ II Directional RNA Library Prep Kit for Illumina® (New England BioLabs inc.) following the manufacturer protocol. Libraries were pooled and sequenced at the GenomEast sequencing platform at the Institut de Génétique et de Biologie Moléculaire et Cellulaire in Strasbourg, France. Sequenced reads with an average quality score above phred 25 had adaptors removed using Trimmomatic v0.39 and were further mapped to the decoyed transcriptome of *A. aegypti* (Vectorbase release 48) using Salmon v1.3.0 [64,65]. Quasi-mapping quantifications were imported into R v3.6.3 and data normalization was performed using the packages EdgeR v3.28.1 and TMM [66,67]. Differentially expressed genes were inferred using the exactTest function assuming a natural dispersion of 40% in gene expression, whose input was used in the function decideTestsDGE. Fold change plots were created using the package ggplot2 v3.3.6 and Euler diagrams generated with the package venneuler v1.1-3. Heatmaps were generated using the R

packages tydiverse v1.3.1 and gplots v3.1.3. Ranked lists of gene expression for each comparison was used as input for Gene Set Enrichment Analysis (GSEA) [28] using the R package fgsea v1.12.0 [68] and in-house developed gene-sets comprising Gene Ontology annotation, pathways, and genes of interest. Sets with adjusted *p*-value < 0.1 were considered in our analysis.

Analysis of Histone H4 genes in *A. aegypti*

Alignment of putative *histone H4* genes was performed using T-Coffee webserver [69] with the variant M-Coffee that allows to combine multiple outputs from different aligners (MAFFT, Clustal and Muscle). The model (TN93 + I) for the phylogenetic reconstruction was defined using SMS within PhyMl server [70,71]. Finally, the Maximum-likelihood tree was constructed with PhyMl requiring 1000 bootstrap replicates and edited using iTOL server [61]. The *Histone H4* polyadenilation signature was defined according to the presence of a canonical polyadenylation signal (AAUAAA) as previously described [35]. The *histone H4* expression heatmap shown in ***Extended Data Fig. 8d*** was produced in R v4.0.2 using the package gplots v3.1.3 and show transcripts TPM counts and normalized by Z-score (row) and sorted according to the phylogenetic tree. Libraries were obtained at the Sequence Read Archive (SRA/NCBI) and accession numbers are shown.

Statistical analyses

Evaluation of statistical significance was performed using GraphPad Prism 9.0 or R-cran v3.3.1 software unless stated otherwise. Viral loads of RT-qPCR positive-only individual mosquito/tissues were log transformed and subjected to Mann-Whitney *U* test. Pairwise comparisons of virus infection prevalence was evaluated by Fisher's exact test or sequentially by a generalized linear model considering the interaction between time points and their prevalence, followed by ANOVA type II using the car package v3.1-0 on R. Presence and absence of DENV in mosquitoes was modelled with univariate and multivariate zero-inflated binomial model (ZIB) [72,73], since 95% of the collections are zeroes. The covariate or covariates (in case of multivariate model) was/were the same for the logit and logistic parts of the model. In particular, we considered PCLV, HTV and their interaction. Model selection was carried out by AIC and BIC comparison [74]. The Vuong test was conducted a priori to test if the zero inflated binomial model was statistically significant and better (in terms of AIC and BIC) than a non-zero inflated model. Data were analysed using R and the 'pscl' package v1.5.5 [75]. Finally, we tested the presence of spatial autocorrelation in the two viruses – via

variogram analyses [20] - but no significant autocorrelations have been found (results not reported).

Data availability

Small RNA and transcriptome libraries from this study have been deposited in the Sequence Read Archive (SRA) at NCBI (project accession PRJNA722589). Sequences spanning the RdRP region from newly discovered viruses were deposited in GenBank under accession MZ556103-MZ556111. Accession numbers for small RNA libraries are provided in **Supplementary Table 1**.

Code availability

All scripts used in this work were deposited in GitHub and can be accessed on https://github.com/ericgdp/sRNA-virome code version 1.0.

Acknowledgments

We thank all members of the ZIKAlliance consortium, especially Anna-Bella Failloux and Alain Kohl, who helped establish a network of collaborators and contributed with fruitful discussions. We thank current and former members of the Marques laboratory and the M3i unit - Insect Models of Innate Immunity, especially Stephanie Blandin and Nelson Martins, for suggestions and discussions.

Funding

This work of the Interdisciplinary Thematic Institute IMCBio, as part of the ITI 2021-2028 program of the University of Strasbourg, CNRS and Inserm, was supported by IdEx Unistra (ANR-10-IDEX-0002), by SFRI-STRAT'US project (ANR 20-SFRI-0012), and EUR IMCBio (IMCBio ANR-17-EURE-0023) under the framework of the French Investments for the Future Program as well as from the previous Labex NetRNA (ANR-10-LABX-0036) to J.T.M and J.-L.I.. This work has been also supported by grants from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) to J.T.M. and E.R.G.R.A.; Fundação de Amparo a Pesquisa do Estado de Minas Gerais (FAPEMIG), Rede Mineira de Imunobiológicos (grant # REDE-00140-16), Rede Mineira de Biomoléculas (grant # REDE-00125-16), Instituto Nacional de Ciência e Tecnologia de Vacinas (INCTV), fonds régional de coopération pour la recherche FRCT2020 Région Grand-Est – ViroMod and Institute for Advanced Studies of the University of Strasbourg (USIAS fellowship 2019) to J.T.M.; Google Latin American Research Award (LARA 2019) to J.T.M. and J.P.P.A.; FAPESP (Grant #13/21719-3) to M.L.N. This work was also supported by the European Union's Horizon 2020 research and innovation programme under ZIKAlliance grant agreement No 734548 and Investissement d'Avenir Programs (ANR-10-LABX-0036 and ANR- 11-EQPX-0022) to J.-L.I. and J.T.M., and ANR PRC TIGERBRIDGE, grant number: 16-CE35-0010-01 to

C.P. J.T.M., B.P.D., E.G.K and M.L.N. are CNPq Research Fellows. M.L.N. is partly funded by the Centers for Research in Emerging Infectious Diseases (CREID), “The Coordinating Research on Emerging Arboviral Threats Encompassing the Neotropics (CREATE-NEO)” grant U01 AI151807 by the National Institutes of Health (NIH/USA). This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior — Brasil (CAPES) — Finance Code 001 to J.T.M. and V.A.S. R.P.O. received a post-doctoral fellowship from CAPES.

Contributions

R.P.O., E.R.G.R.A., J.-L.I. and J.T.M conceived the project. Methodology: R.P.O., Y.M.H.T., E.R.G.R.A., J.P.P.A., J.N.A., L.S. and J.T.M. designed the experiments and performed computational analysis. R.P.O., Y.M.H.T., I.J.S.F., F.V.F., A.G.A.F., S.C.G.A., A.T.S.S., K.P.R.S., A.P.P.V., A.B. performed experiments. C.H.T., M.D., A.G., C.P., J.O.N., T.M.V., C.J.M.K., M.A.W., A.L.C.C., M.T.P., M.C.P.P., M.L.N., V.A.S., R.N.M., M.A.Z.B., B.P.D., E.G.K., E.M. provided samples and reagents. M.R. performed the mathematical modelling. R.P.O., Y.M.H.T., E.R.G.R.A., J.-L.I. and J.T.M. wrote the manuscript. All authors read and contributed to manuscript editing.

Competing interests

Authors declare that they have no competing interests.

Figure Legends

Fig. 1 The virome of *A. aegypti* and *A. albopictus* mosquitoes.
a, World map indicating sites of mosquito collection. Pie charts show the proportion of *A. aegypti* and *A. albopictus* in red and blue, respectively, at each collection site. Adult mosquitoes were captured either using traps or human baits. Laboratory strains of mosquitoes analyzed in this work are indicated at the bottom. **b**, Overview of our analysis pipeline. Captured mosquitoes were morphologically identified by species and stored in a biobank of RNA samples, that were pooled to prepare small RNA libraries for high-throughput sequencing. Using our metagenomic pipeline, assembled contigs were classified into viral and non-viral sequences based on similarity against reference databases. Sequences that lack any similarity to known references are indicated as unknown. **c**, Individual results from our metagenomic analysis for each of the 91 small RNA libraries in this study. Bars indicate the total number of contigs and the proportion of viral, non-viral and unknown contigs per library.

Fig. 2 A highly diverse and distinct collection of viruses in *A. aegypti* and *A. albopictus.*
a, Overview of the strategy for manual curation of viral contigs to confirm the origin and remove sequences potentially derived from EVEs. Curation consisted of BLAST search for similar viral sequences, inspection of predicted ORF structure including continuity and extension throughout each contig, and evaluation of the small RNA profile for the identification of signatures of siRNA production (symmetrical accumulation of RNAs with 20-23 nt length that mapped to each strand with no 5' base preference) or piRNA production (accumulation of 24-29 nt length RNAs with 5' U preference); and overall contig coverage by small RNAs. Remaining viral contigs were clustered by sequence similarity and co-ocurrence to identify groups of contigs that belong to same viruses. Re-assembly was perfomed within these groups and resulting contigs were analyzed for the presence of domains. Potential polymerases identified were used to classify viruses based on sequence similarity and phylogeny. **b**, Host, virus genomic organization, family, and closest reference on GenBank identified by BLAST similarity searches for each of the 12 viruses identified in our datasets. New viral species are indicated in red while previously known viruses are in black. Sequence similarity and accession number according to the closest viral sequence at the nucleotide (nt) or protein (aa) levels are indicated. **c**, Viral load shown as a heatmap for each of the twelve viruses in mosquito populations from each collection site or laboratory strains. In the heatmap, white indicates absence of a virus and *NA* indicates absence of samples from a given location. Prevalence of each virus is shown on the right as number of samples with detectable virus over the total. Number of individuals per pool and number of species per collection site are described on the **Supplementary Table 1**. **d**, Pie charts represent the overall burden of virus and viral diversity for *A. aegypti* and *A. albopictus* populations in each collection site across the word. X indicates that no viral contigs have been identified.

Fig. 3 Concurrent detection of HTV, PCLV and DENV.
a, Location of the study site, the city of Caratinga in the Southeast of Brazil. **b**, Prevalence of individual and co-infections by OVV, HTV and PCLV tested by RT-qPCR. **c**, Prevalence of DENV, HTV and PCLV separately in individual *A. aegypti* mosquitoes and human blood samples accessed by RT-qPCR. **d**, Monthly prevalence of HTV and PCLV separately in individual *A. aegypti* mosquitoes. **e**, Prevalence of HTV

and PCLV co-infection in DENV infected (DENV+) and DENV non-infected (DENV-) mosquitoes. Statistical significance was determined by two-tailed Fisher's exact test. **f**, Likelihood of DENV infection in mosquitoes carrying HTV or PCLV and HTV shown by odds ratio.

Fig. 4 Effects of insect specific viruses on DENV and ZIKV transmission.
a-g, Strategy to evaluate the interference of HTV and PCLV on DENV (**a-c**) or ZIKV (**d-g**) infection in wild mosquito populations. HTV/PCLV infected or virus free mosquitoes were exposed to DENV (**a**) or ZIKV (**d**) on a blood meal. Viral loads and prevalence of infection were measured in the midgut (**b,e**), carcass (**c,f**) and salivary glands (**g**) of mosquitoes at the indicated days post feeding (d.p.f.). Pie charts below each group indicate the prevalence of DENV or ZIKV infection at each time point. **h**, Strategy to compare the ability of mosquitoes with or without HTV and PCLV to transmit ZIKV to the susceptible AG129 mouse model. **i**, Viremia in mice was determined 3 days after exposure to mosquito bites, comparing animals that were bitten by mosquitoes carrying or not HTV and PCLV at 6, 8 and 12 days post oral infection. *ND* – not detected. Pie charts below each group indicate the prevalence of infection. (**j**) ZIKV RNA levels in mosquitoes used in the transmission experiment in **i**. *NS* – non-significant. **k-l** Spatially explicit, individual-based meta-population model showing the effect of EIP duration on the number of human infections (**k**) and virus prevalence in mosquitoes (**l**) taken as a proportion of infected individuals on a single day during the seasonal peak over a 50 year period (n=50). In box plots of ***b***, ***c***, ***e***, ***f***, ***j***, and ***l,*** boxes show the second and third interquartile ranges divided by the median while whiskers represent maximum and minimum values. Statistical significance was determined by two-tailed Mann–Whitney U-test. In ***k***, lines represent the averaged cumulative incidence over five model runs for each value of the EIP, while shadows depict confidence intervals. Numbers of infected midguts (***b***,***e***), carcasses (***c***,***f***), or whole mosquitoes (***j***) over the total number tested are indicated above each column. Each dot represents an individual sample. Statistical significance of the prevalence of infection was determined in ***b***, ***c***, ***e***, ***f***, ***i***, and ***j*** by two-tailed Fisher's exact test or, in ***g***, using a binomial generalized linear model followed by ANOVA type II testing for time points as factor.

Fig. 5 HTV and PCLV regulate the expression of proviral histone H4 during DENV infection.
a, Overall strategy to identify biological pathways associated with HTV and PCLV interaction with DENV. HTV/PCLV infected and virus free wild mosquito populations were allowed to feed on DENV-infected mice. The transcriptome of DENV infected and non-infected mosquitoes from HTV/PCLV and virus free groups were analyzed separately at 8- and 14-days post feeding. Gene Set Enrichment Analysis (GSEA) was performed for each comparison. **b**, Biological processes significantly enriched (adjusted p-value < 0.01) in the comparisons of DENV-infected versus non-infected mosquitoes and PCLV/HTV infected versus virus free controls are shown. **c**, Overlap of leading-edge genes belonging to the 4 biological processes enriched in at least 6 out of 8 comparisons. Size of each circle represents the number of leading-edge genes. Histogram shows that histone genes represent the majority in the leading edge of significantly enriched processes. **d-e**, Histone H4 levels in the carcass of mosquitoes at 2, 4, and 8 d.p.i. with ZIKV in wild mosquitoes carrying HTV/PCLV or virus free controls (**e**). **f-g**. Differential expression of histone H4 levels in the carcass of mosquitoes in the presence of HTV and PCLV is only observed in DENV infected individuals. **h-j**, Silencing of histone H4 mRNA in adult mosquitoes exposed to DENV. DENV infection was analyzed at 4 and 8 d.p.f. in the midgut (**i**) and carcass (**j**) of silenced (dsH4) and control (dsGFP) mosquitoes. Pie charts below each group

indicate the prevalence of DENV or ZIKV infection. *d.p.f.* – days post feeding. *d.p.i.* – days post injection, *NS* – non-significant. In box plots of ***e***, ***g***, ***h***, and ***j,*** boxes show the second and third interquartile ranges divided by the median while whiskers represent maximum and minimum values. Numbers of infected samples over the total number tested are indicated above each column. Statistical significance was determined by two-tailed one-way ANOVA with Tukey's correction for multiple comparisons (**e**,**g**) or by two-tailed Mann-Whitney U-test (**h**,**i**). Each dot represents an individual sample. Statistical significance of prevalence of infection was determined by two-tailed Fisher's exact test.

992 **References**

993

994 1. Weaver, S. C., Charlier, C., Vasilakis, N. & Lecuit, M. Zika, Chikungunya, and
995 Other Emerging Vector-Borne Viral Diseases. *Annu Rev Med* **69**, 395–408 (2018).
996 2. Bhatt, S. *et al.* The global distribution and burden of dengue. *Nature* **496**, 504–507
997 (2013).
998 3. Messina, J. P. *et al.* Global spread of dengue virus types: mapping the 70 year
999 history. *Trends Microbiol* **22**, 138–146 (2014).
1000 4. Franklinos, L. H. V., Jones, K. E., Redding, D. W. & Abubakar, I. The effect of
1001 global change on mosquito-borne disease. *Lancet Infect Dis* **19**, e302–e312
1002 (2019).
1003 5. Kraemer, M. U. G. *et al.* Past and future spread of the arbovirus vectors Aedes
1004 aegypti and Aedes albopictus. *Nat Microbiol* **4**, 854–863 (2019).
1005 6. Cromwell, E. A. *et al.* The relationship between entomological indicators of Aedes
1006 aegypti abundance and dengue virus infection. *PLoS Negl Trop Dis* **11**, e0005429
1007 (2017).
1008 7. de Almeida, J. P., Aguiar, E. R., Armache, J. N., Olmo, R. P. & Marques, J. T. The
1009 virome of vector mosquitoes. *Curr Opin Virol* **49**, 7–12 (2021).
1010 8. Aguiar, E. R. G. R. *et al.* Sequence-independent characterization of viruses based
1011 on the pattern of viral small RNAs produced by the host. *Nucleic Acids Res* **43**,
1012 6191–6206 (2015).
1013 9. Boyles, S. M. *et al.* Under-the-Radar Dengue Virus Infections in Natural
1014 Populations of Aedes aegypti Mosquitoes. *mSphere* **5**, (2020).
1015 10. Ramos-Nino, M. E. *et al.* High prevalence of Phasi Charoen-like virus from wild-
1016 caught Aedes aegypti in Grenada, W.I. as revealed by metagenomic analysis.
1017 *PLoS One* **15**, e0227998 (2020).

11. Shi, C. *et al.* Stable distinct core eukaryotic viromes in different mosquito species from Guadeloupe, using single mosquito viral metagenomics. *Microbiome* **7**, 121 (2019).
12. Zakrzewski, M. *et al.* Mapping the virome in wild-caught Aedes aegypti from Cairns and Bangkok. *Sci Rep* **8**, 4690 (2018).
13. Patterson, E. I., Villinger, J., Muthoni, J. N., Dobel-Ober, L. & Hughes, G. L. Exploiting insect-specific viruses as a novel strategy to control vector-borne disease. *Curr Opin Insect Sci* **39**, 50–56 (2020).
14. Vasilakis, N. & Tesh, R. B. Insect-specific viruses and their potential impact on arbovirus transmission. *Curr Opin Virol* **15**, 69–74 (2015).
15. Aguiar, E. R. G. R., Olmo, R. P. & Marques, J. T. Virus-derived small RNAs: molecular footprints of host-pathogen interactions. *Wiley Interdiscip Rev RNA* **7**, 824–837 (2016).
16. Morazzani, E. M., Wiley, M. R., Murreddu, M. G., Adelman, Z. N. & Myles, K. M. Production of virus-derived ping-pong-dependent piRNA-like small RNAs in the mosquito soma. *PLoS Pathog* **8**, e1002470 (2012).
17. Myles, K. M., Wiley, M. R., Morazzani, E. M. & Adelman, Z. N. Alphavirus-derived small RNAs modulate pathogenesis in disease vector mosquitoes. *Proc Natl Acad Sci U S A* **105**, 19938–19943 (2008).
18. Frangeul, L., Blanc, H., Saleh, M.-C. & Suzuki, Y. Differential Small RNA Responses against Co-Infecting Insect-Specific Viruses in Aedes albopictus Mosquitoes. *Viruses* **12**, (2020).
19. Olmo, R. P. *et al.* Control of dengue virus in the midgut of Aedes aegypti by ectopic expression of the dsRNA-binding protein Loqs2. *Nat Microbiol* **3**, 1385–1393 (2018).
20. Sedda, L. *et al.* The spatial and temporal scales of local dengue virus transmission in natural settings: a retrospective analysis. *Parasit Vectors* **11**, 79 (2018).

21. Schultz, M. J., Frydman, H. M. & Connor, J. H. Dual Insect specific virus infection limits Arbovirus replication in Aedes mosquito cells. *Virology* **518**, 406–413 (2018).
22. Fredericks, A. C. *et al.* Aedes aegypti (Aag2)-derived clonal mosquito cell lines reveal the effects of pre-existing persistent infection with the insect-specific bunyavirus Phasi Charoen-like virus on arbovirus replication. *PLoS Negl Trop Dis* **13**, e0007346 (2019).
23. Sedda, L., Taylor, B. M., Eiras, A. E., Marques, J. T. & Dillon, R. J. Using the intrinsic growth rate of the mosquito population improves spatio-temporal dengue risk estimation. *Acta Trop* **208**, 105519 (2020).
24. Lin, J.-J. *et al.* Aggressive organ penetration and high vector transmissibility of epidemic dengue virus-2 Cosmopolitan genotype in a transmission mouse model. *PLoS Pathog* **17**, e1009480 (2021).
25. Sun, P. *et al.* A mosquito salivary protein promotes flavivirus transmission by activation of autophagy. *Nat Commun* **11**, 260 (2020).
26. Lourenço, J. & Recker, M. Natural, persistent oscillations in a spatial multi-strain disease system with application to dengue. *PLoS Comput Biol* **9**, e1003308 (2013).
27. Lourenço, J. *et al.* Epidemiological and ecological determinants of Zika virus transmission in an urban setting. *Elife* **6**, e29820 (2017).
28. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545–15550 (2005).
29. Flaus, A., Downs, J. A. & Owen-Hughes, T. Histone isoforms and the oncohistone code. *Curr Opin Genet Dev* **67**, 61–66 (2021).
30. Lyons, S. M. *et al.* A subset of replication-dependent histone mRNAs are expressed as polyadenylated RNAs in terminally differentiated tissues. *Nucleic Acids Res* **44**, 9190–9205 (2016).

31. Baidaliuk, A. *et al.* Cell-Fusing Agent Virus Reduces Arbovirus Dissemination in *Aedes aegypti* Mosquitoes *In Vivo*. *J Virol* **93**, e00705-19, /jvi/93/18/JVI.00705-19.atom (2019).
32. Blitvich, B. J. & Firth, A. E. Insect-specific flaviviruses: a systematic review of their discovery, host range, mode of transmission, superinfection exclusion potential and genomic organization. *Viruses* **7**, 1927–1959 (2015).
33. Colpitts, T. M., Barthel, S., Wang, P. & Fikrig, E. Dengue virus capsid protein binds core histones and inhibits nucleosome formation in human liver cells. *PLoS One* **6**, e24365 (2011).
34. Mourão, D. *et al. A histone-like motif in yellow fever virus contributes to viral replication*. http://biorxiv.org/lookup/doi/10.1101/2020.05.05.078782 (2020) doi:10.1101/2020.05.05.078782.
35. Girardi, E. *et al.* Histone-derived piRNA biogenesis depends on the ping-pong partners Piwi5 and Ago3 in Aedes aegypti. *Nucleic Acids Res* **45**, 4881–4892 (2017).
36. Varjak, M. *et al.* Aedes aegypti Piwi4 Is a Noncanonical PIWI Protein Involved in Antiviral Responses. *mSphere* **2**, e00144-17 (2017).
37. Parry, R. & Asgari, S. Aedes Anphevirus: an Insect-Specific Virus Distributed Worldwide in *Aedes aegypti* Mosquitoes That Has Complex Interplays with *Wolbachia* and Dengue Virus Infection in Cells. *J Virol* **92**, e00224-18, /jvi/92/17/e00224-18.atom (2018).
38. Zhang, G., Asad, S., Khromykh, A. A. & Asgari, S. Cell fusing agent virus and dengue virus mutually interact in Aedes aegypti cell lines. *Sci Rep* **7**, 6935 (2017).
39. Nasar, F., Erasmus, J. H., Haddow, A. D., Tesh, R. B. & Weaver, S. C. Eilat virus induces both homologous and heterologous interference. *Virology* **484**, 51–58 (2015).
40. Kenney, J. L., Solberg, O. D., Langevin, S. A. & Brault, A. C. Characterization of a novel insect-specific flavivirus from Brazil: potential for inhibition of infection of

arthropod cells with medically important flaviviruses. *J Gen Virol* **95**, 2796–2808 (2014).

41. Romo, H., Kenney, J. L., Blitvich, B. J. & Brault, A. C. Restriction of Zika virus infection and transmission in *Aedes aegypti* mediated by an insect-specific flavivirus. *Emerging Microbes & Infections* **7**, 1–13 (2018).
42. Goenaga, S. *et al.* Potential for Co-Infection of a Mosquito-Specific Flavivirus, Nhumirim Virus, to Block West Nile Virus Transmission in Mosquitoes. *Viruses* **7**, 5801–5812 (2015).
43. Alefelder, S., Patel, B. K. & Eckstein, F. Incorporation of terminal phosphorothioates into oligonucleotides. *Nucleic Acids Res* **26**, 4983–4988 (1998).
44. Marques, J. T. *et al.* Loqs and R2D2 act sequentially in the siRNA pathway in Drosophila. *Nat Struct Mol Biol* **17**, 24–30 (2010).
45. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet j.* **17**, 10 (2011).
46. Matthews, B. J. *et al.* Improved reference genome of Aedes aegypti informs arbovirus vector control. *Nature* **563**, 501–507 (2018).
47. Chen, X.-G. *et al.* Genome sequence of the Asian Tiger mosquito, Aedes albopictus, reveals insights into its biology, genetics, and evolution. *Proc Natl Acad Sci U S A* **112**, E5907-5915 (2015).
48. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**, R25 (2009).
49. Zerbino, D. R. Using the Velvet de novo assembler for short-read sequencing technologies. *Curr Protoc Bioinformatics* **Chapter 11**, Unit 11.5 (2010).
50. Bankevich, A. *et al.* SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* **19**, 455–477 (2012).

51. O'Leary, N. A. *et al.* Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res* **44**, D733-745 (2016).
52. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J Mol Biol* **215**, 403–410 (1990).
53. Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**, 3150–3152 (2012).
54. Aguiar, E. R. G. R. *et al.* A single unidirectional piRNA cluster similar to the flamenco locus is the major source of EVE-derived transcription and small RNAs in Aedes aegypti mosquitoes. *RNA* **26**, 581–594 (2020).
55. 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).
56. 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).
57. Katoh, K., Misawa, K., Kuma, K. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* **30**, 3059–3066 (2002).
58. Miller, M. A., Pfeiffer, W. & Schwartz, T. The CIPRES science gateway: a community resource for phylogenetic analyses. in *Proceedings of the 2011 TeraGrid Conference on Extreme Digital Discovery - TG '11* 1 (ACM Press, 2011). doi:10.1145/2016741.2016785.
59. Darriba, D. *et al.* ModelTest-NG: A New and Scalable Tool for the Selection of DNA and Protein Evolutionary Models. *Mol Biol Evol* **37**, 291–294 (2020).
60. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol* **35**, 1547–1549 (2018).

61. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* **47**, W256–W259 (2019).
62. Barletta, A. B. F. *et al.* Microbiota activates IMD pathway and limits Sindbis infection in Aedes aegypti. *Parasit Vectors* **10**, 103 (2017).
63. Donald, C. L. *et al.* Full Genome Sequence and sfRNA Interferon Antagonist Activity of Zika Virus from Recife, Brazil. *PLoS Negl Trop Dis* **10**, e0005048 (2016).
64. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
65. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* **14**, 417–419 (2017).
66. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* **11**, R25 (2010).
67. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
68. Korotkevich, G. *et al. Fast gene set enrichment analysis*. http://biorxiv.org/lookup/doi/10.1101/060012 (2016) doi:10.1101/060012.
69. Di Tommaso, P. *et al.* T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic Acids Res* **39**, W13-17 (2011).
70. Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* **59**, 307–321 (2010).
71. Lefort, V., Longueville, J.-E. & Gascuel, O. SMS: Smart Model Selection in PhyML. *Mol Biol Evol* **34**, 2422–2424 (2017).
72. Lambert, D. Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing. *Technometrics* **34**, 1 (1992).

73. Hilbe, J. M. *Negative Binomial Regression*. (2007).

74. Cameron, A. C., Cameron, A. C., & Cambridge University Press. *Regression analysis of count data*. (Cambridge University Press, 1998).

75. Zeileis, A., Kleiber, C. & Jackman, S. Regression Models for Count Data in *R*. *J. Stat. Soft.* **27**, (2008).

Figure 1

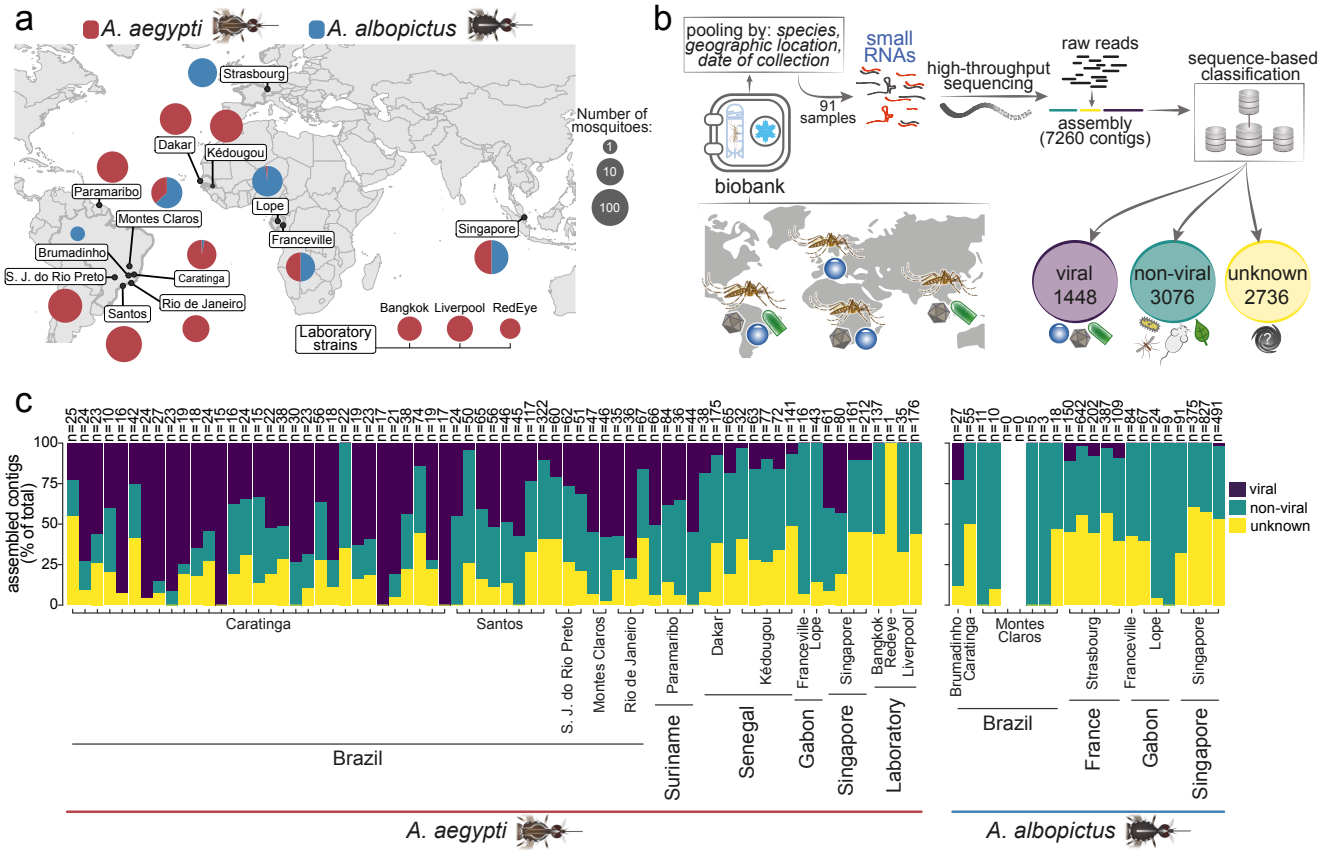

a
A. aegypti
A. albopictus
Strasbourg
Dakar
Kédougou
Paramaribo
Lope
Montes Claros
Franceville
Singapore
Brumadinho
S. J. do Rio Preto
Caratinga
Santos
Rio de Janeiro
Bangkok Liverpool RedEye
Laboratory strains
Number of mosquitoes:
1
10
100
b
pooling by: species, geographic location, date of collection
91 samples
small RNAs
high-throughput sequencing
raw reads
assembly (7260 contigs)
sequence-based classification
biobank
viral 1448
non-viral 3076
unknown 2736
c
assembled contigs (% of total)
viral
non-viral
unknown
Caratinga
Santos
S. J. do Rio Preto
Montes Claros
Rio de Janeiro
Paramaribo
Dakar
Kédougou
Franceville - Lope
Singapore
Bangkok - Redeye Liverpool
Brumadinho - Caratinga
Montes Claros
Strasbourg
Franceville - Lope
Singapore
Brazil
Suriname
Senegal
Gabon
Singapore
Laboratory
Brazil
France
Gabon
Singapore
A. aegypti
A. albopictus

Figure 2

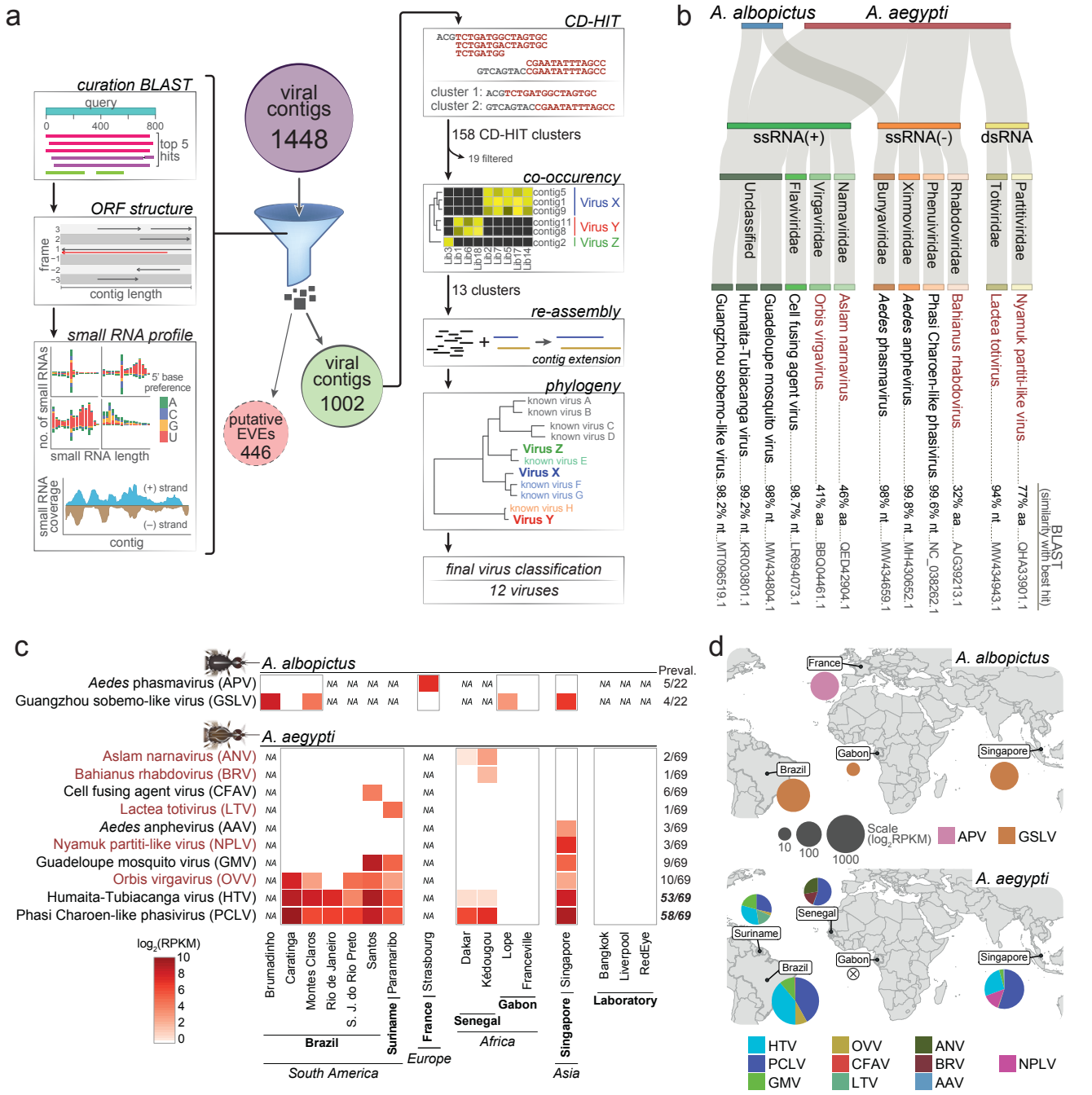

a
curation BLAST
query
top 5 hits
ORF structure
contig length
small RNA profile
small RNA length
5' base preference
(+) strand
(–) strand
contig
viral contigs 1448
putative EVEs 446
viral contigs 1002
CD-HIT
158 CD-HIT clusters
19 filtered
co-occurency
Virus X
Virus Y
Virus Z
13 clusters
re-assembly
contig extension
phylogeny
known virus A
known virus B
known virus C
known virus D
Virus Z
known virus E
Virus X
known virus F
known virus G
known virus H
Virus Y
final virus classification
12 viruses
b
A. albopictus
A. aegypti
ssRNA(+)
ssRNA(-)
dsRNA
Unclassified
Flaviviridae
Virgaviridae
Namaviridae
Bunyaviridae
Xinmoviridae
Phenuiviridae
Rhabdoviridae
Totiviridae
Partitiviridae
Guangzhou sobemo-like virus 98.2% nt MT096519.1
Humaita-Tubiacanga virus 99.2% nt KR003801.1
Guadeloupe mosquito virus 98% nt MW434804.1
Cell fusing agent virus 98.7% nt LR694073.1
Orbis virgavirus 41% aa BBQ04461.1
Aslam narnavirus 46% aa QED42904.1
Aedes phasmavirus 98% nt MW434659.1
Aedes anphevirus 99.8% nt MH430652.1
Phasi Charoen-like phasivirus 99.6% nt NC_038262.1
Bahianus rhabdovirus 32% aa AJG39213.1
Lactea totivirus 94% nt MW434943.1
Nyamuk partiti-like virus 77% aa QHA33901.1
BLAST (similarity with best hit)
c
A. albopictus
Aedes phasmavirus (APV) 5/22
Guangzhou sobemo-like virus (GSLV) 4/22
A. aegypti
Aslam narnavirus (ANV) 2/69
Bahianus rhabdovirus (BRV) 1/69
Cell fusing agent virus (CFAV) 6/69
Lactea totivirus (LTV) 1/69
Aedes anphevirus (AAV) 3/69
Nyamuk partiti-like virus (NPLV) 3/69
Guadeloupe mosquito virus (GMV) 9/69
Orbis virgavirus (OVV) 10/69
Humaita-Tubiacanga virus (HTV) 53/69
Phasi Charoen-like phasivirus (PCLV) 58/69
log2(RPKM)
Brumadinho
Caratinga
Monte Claros
Rio de Janeiro
S. J. do Rio Preto
Santos
Paramaribo
Strasbourg
Dakar
Kédougou
Lope
Franceville
Singapore
Bangkok
Liverpool
RedEye
Brazil
Suriname
France
Gabon
Senegal
Singapore
Laboratory
South America
Europe
Africa
Asia
d
A. albopictus
France
Gabon
Singapore
Brazil
Scale (log2RPKM)
APV
GSLV
A. aegypti
Senegal
Suriname
Brazil
Gabon
Singapore
HTV
PCLV
GMV
OVV
CFAV
LTV
ANV
BRV
AAV
NPLV

Figure 3

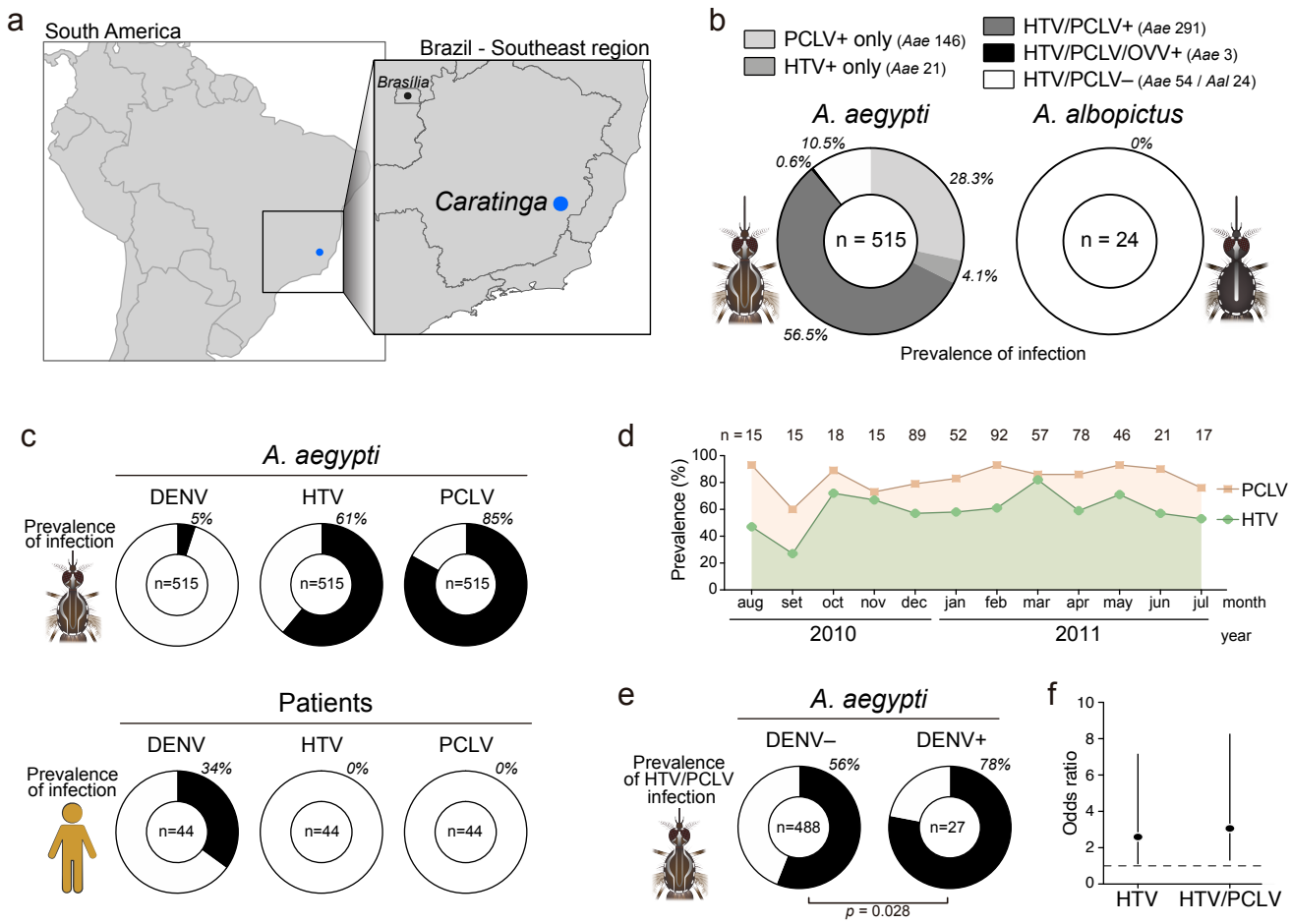

a
South America
Brazil - Southeast region
Brasília
Caratinga
b
PCLV+ only (Aae 146)
HTV+ only (Aae 21)
HTV/PCLV+ (Aae 291)
HTV/PCLV/OVV+ (Aae 3)
HTV/PCLV– (Aae 54 / Aal 24)
A. aegypti
A. albopictus
10.5%
0.6%
28.3%
4.1%
56.5%
n = 515
0%
n = 24
Prevalence of infection
c
A. aegypti
DENV
HTV
PCLV
Prevalence of infection
5%
61%
85%
n=515
n=515
n=515
Patients
DENV
HTV
PCLV
Prevalence of infection
34%
0%
0%
n=44
n=44
n=44
d
n = 15 15 18 15 89 52 92 57 78 46 21 17
Prevalence (%)
100
80
60
40
20
0
aug set oct nov dec jan feb mar apr may jun jul
month
2010
2011
year
PCLV
HTV
e
A. aegypti
DENV–
DENV+
Prevalence of HTV/PCLV infection
56%
78%
n=488
n=27
p = 0.028
f
Odds ratio
10
8
6
4
2
0
HTV
HTV/PCLV

Figure 4

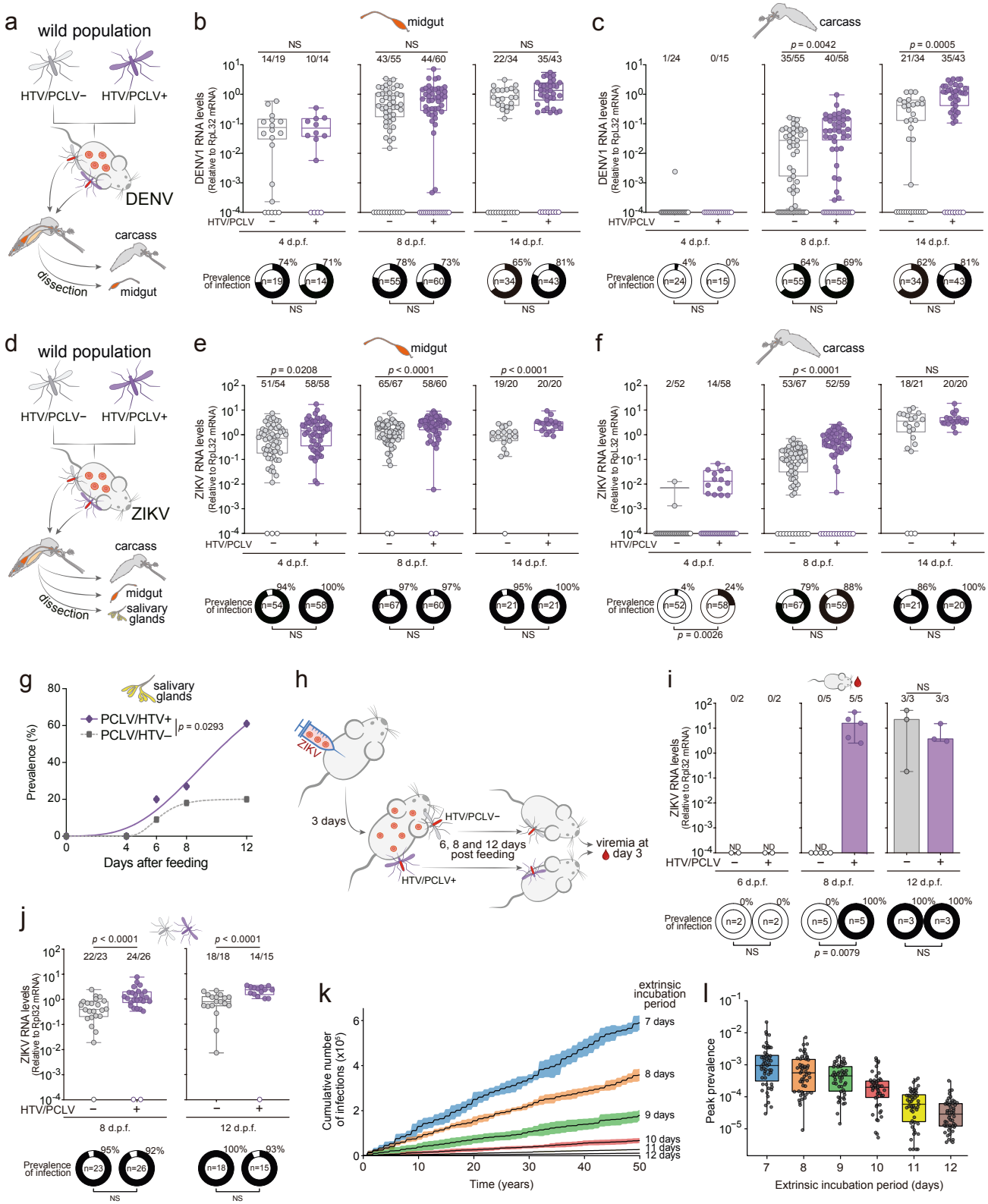
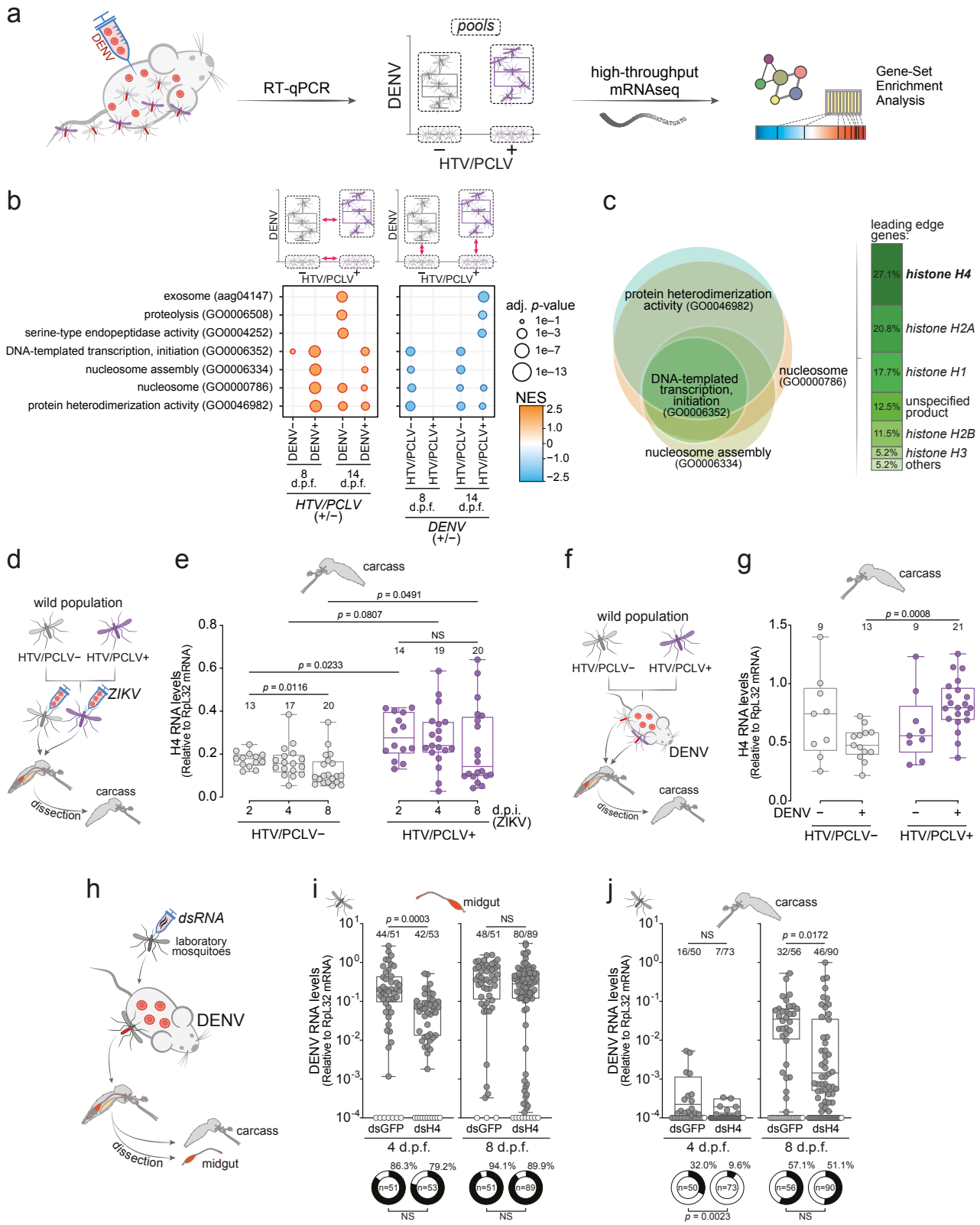
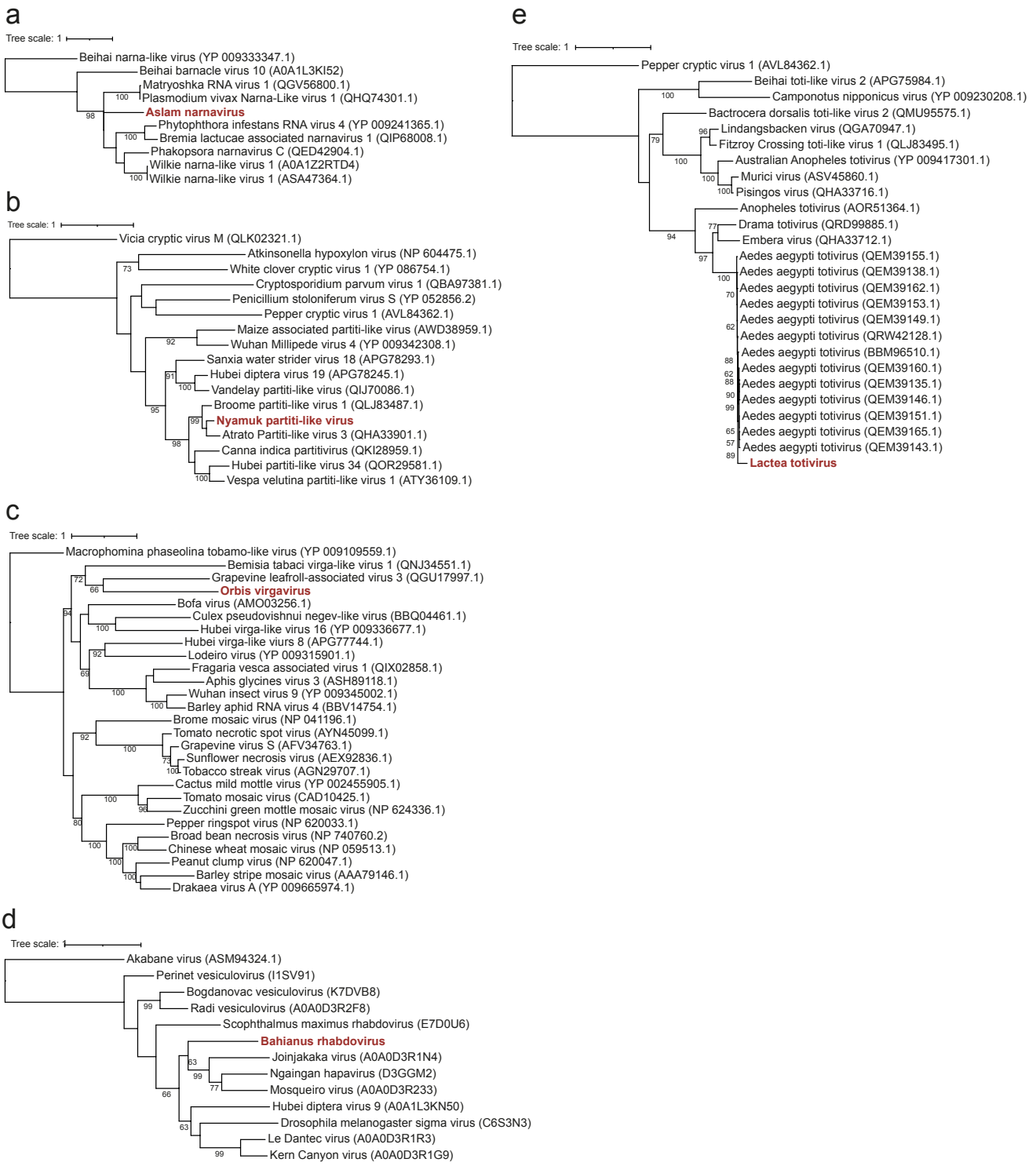

a
wild population
HTV/PCLV−
HTV/PCLV+
DENV
carcass
midgut
dissection
b
midgut
DENV1 RNA levels (Relative to RpL32 mRNA)
NS
14/19
10/14
43/55
44/60
22/34
35/43
HTV/PCLV
4 d.p.f.
8 d.p.f.
14 d.p.f.
Prevalence of infection
74%
71%
78%
73%
65%
81%
n=19
n=14
n=55
n=60
n=34
n=43
c
carcass
1/24
0/15
p = 0.0042
35/55
40/58
p = 0.0005
21/34
35/43
4%
0%
64%
69%
62%
81%
n=24
n=15
n=55
n=58
n=34
n=43
d
wild population
ZIKV
carcass
midgut
salivary glands
dissection
e
midgut
ZIKV RNA levels (Relative to RpL32 mRNA)
p = 0.0208
51/54
58/58
p < 0.0001
65/67
58/60
19/20
20/20
94%
100%
97%
97%
95%
100%
n=54
n=58
n=67
n=60
n=21
n=21
f
carcass
2/52
14/58
53/67
52/59
18/21
20/20
4%
24%
79%
88%
86%
100%
n=52
n=58
n=67
n=59
n=21
n=20
p = 0.0026
g
salivary glands
Prevalence (%)
PCLV/HTV+
PCLV/HTV−
p = 0.0293
Days after feeding
h
ZIKV
3 days
HTV/PCLV−
6, 8 and 12 days post feeding
HTV/PCLV+
viremia at day 3
i
ZIKV RNA levels (Relative to RpL32 mRNA)
0/2
0/2
0/5
5/5
3/3
3/3
ND
6 d.p.f.
8 d.p.f.
12 d.p.f.
0%
100%
n=2
n=5
n=3
p = 0.0079
j
p < 0.0001
22/23
24/26
18/18
14/15
95%
92%
100%
93%
n=23
n=26
n=18
n=15
k
extrinsic incubation period
7 days
8 days
9 days
10 days
11 days
12 days
Cumulative number of infections (x10⁶)
Time (years)
l
Peak prevalence
Extrinsic incubation period (days)

Figure 5

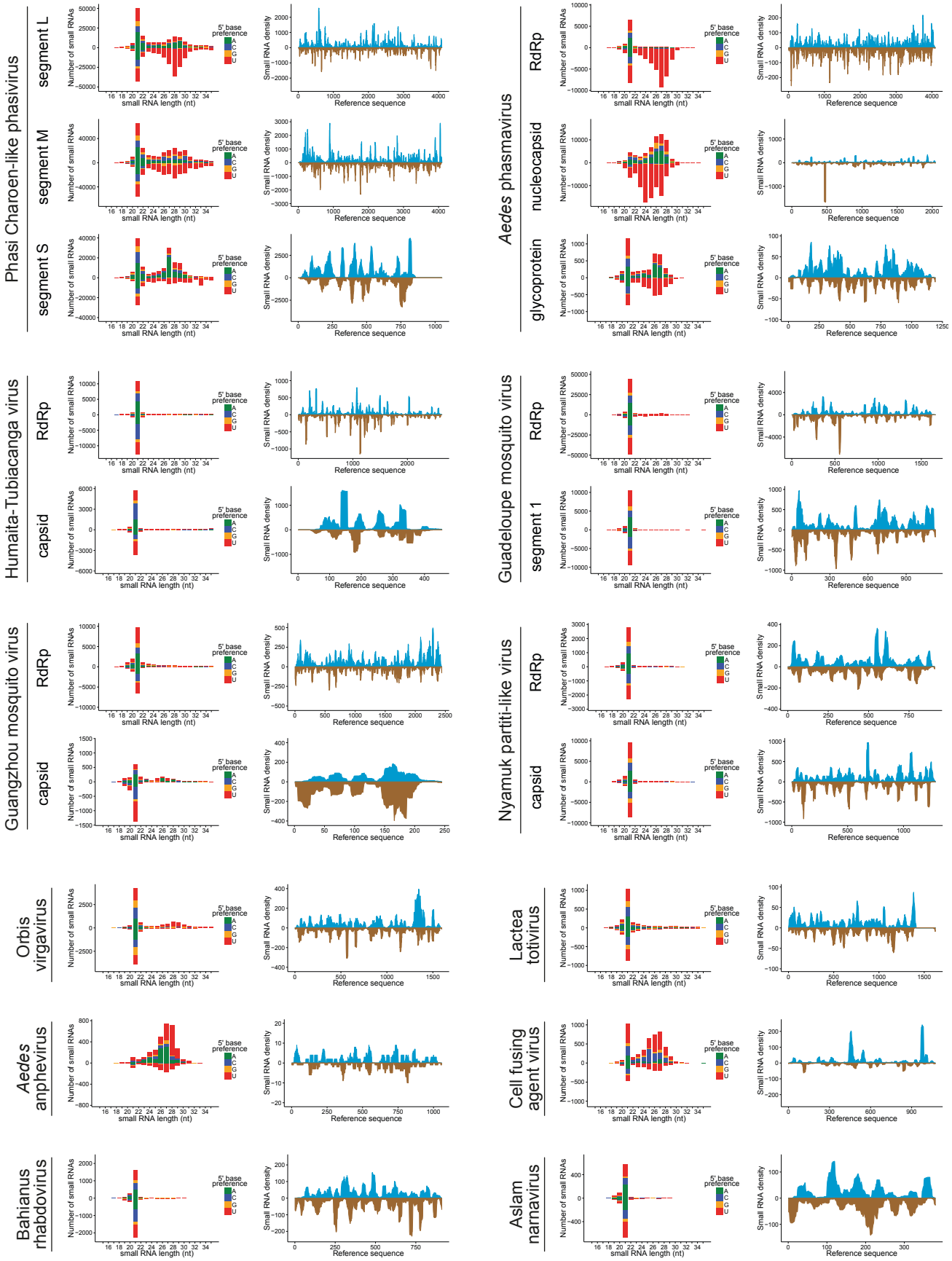

a
DENV
RT-qPCR
pools
DENV
HTV/PCLV
high-throughput mRNAseq
Gene-Set Enrichment Analysis
b
exosome (aag04147)
proteolysis (GO0006508)
serine-type endopeptidase activity (GO0004252)
DNA-templated transcription, initiation (GO0006352)
nucleosome assembly (GO0006334)
nucleosome (GO0000786)
protein heterodimerization activity (GO0046982)
adj. p-value
1e–1
1e–3
1e–7
1e–13
NES
2.5
1.0
0
–1.0
–2.5
DENV–
DENV+
8 d.p.f.
14 d.p.f.
HTV/PCLV (+/–)
HTV/PCLV–
HTV/PCLV+
DENV (+/–)
c
protein heterodimerization activity (GO0046982)
DNA-templated transcription, initiation (GO0006352)
nucleosome (GO0000786)
nucleosome assembly (GO0006334)
leading edge genes:
27.1% histone H4
20.8% histone H2A
17.7% histone H1
12.5% unspecified product
11.5% histone H2B
5.2% histone H3
5.2% others
d
wild population
HTV/PCLV– HTV/PCLV+
ZIKV
carcass
dissection
e
carcass
p = 0.0491
p = 0.0807
NS
p = 0.0233
p = 0.0116
H4 RNA levels (Relative to RpL32 mRNA)
13 17 20
14 19 20
2 4 8
d.p.i. (ZIKV)
HTV/PCLV–
HTV/PCLV+
f
wild population
HTV/PCLV– HTV/PCLV+
DENV
carcass
dissection
g
carcass
p = 0.0008
9 13 9 21
H4 RNA levels (Relative to RpL32 mRNA)
DENV – + – +
HTV/PCLV–
HTV/PCLV+
h
dsRNA
laboratory mosquitoes
DENV
carcass
dissection
midgut
i
midgut
p = 0.0003
NS
44/51 42/53 48/51 80/89
DENV RNA levels (Relative to RpL32 mRNA)
dsGFP dsH4 dsGFP dsH4
4 d.p.f. 8 d.p.f.
86.3% 79.2% 94.1% 89.9%
n=51 n=53 n=51 n=89
NS NS
j
carcass
NS
p = 0.0172
16/50 7/73 32/56 46/90
DENV RNA levels (Relative to RpL32 mRNA)
dsGFP dsH4 dsGFP dsH4
4 d.p.f. 8 d.p.f.
32.0% 9.6% 57.1% 51.1%
n=50 n=73 n=56 n=90
p = 0.0023 NS

Extended Data Figure 1

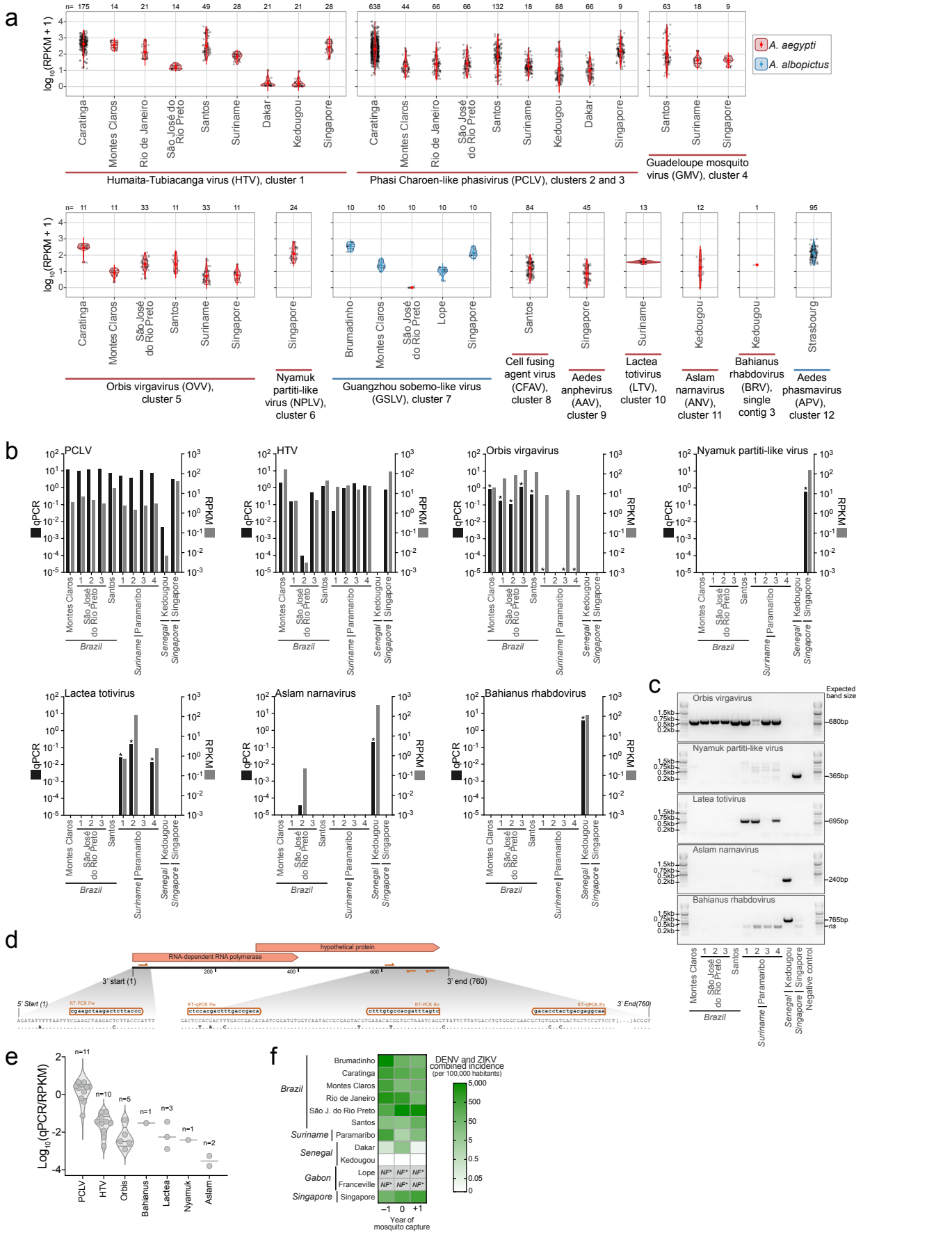
Extended Data Figure 2



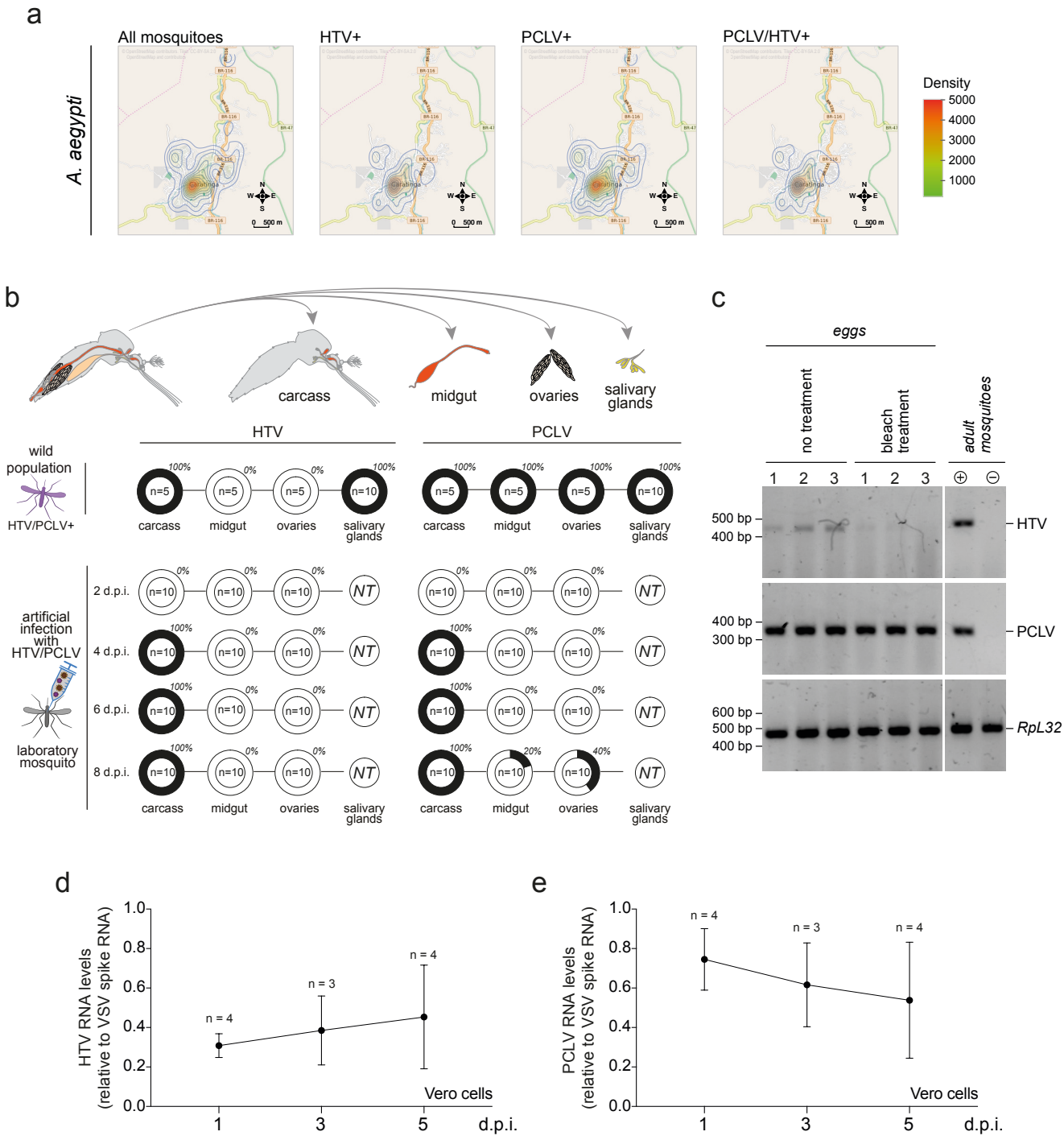
Extended Data Figure 3



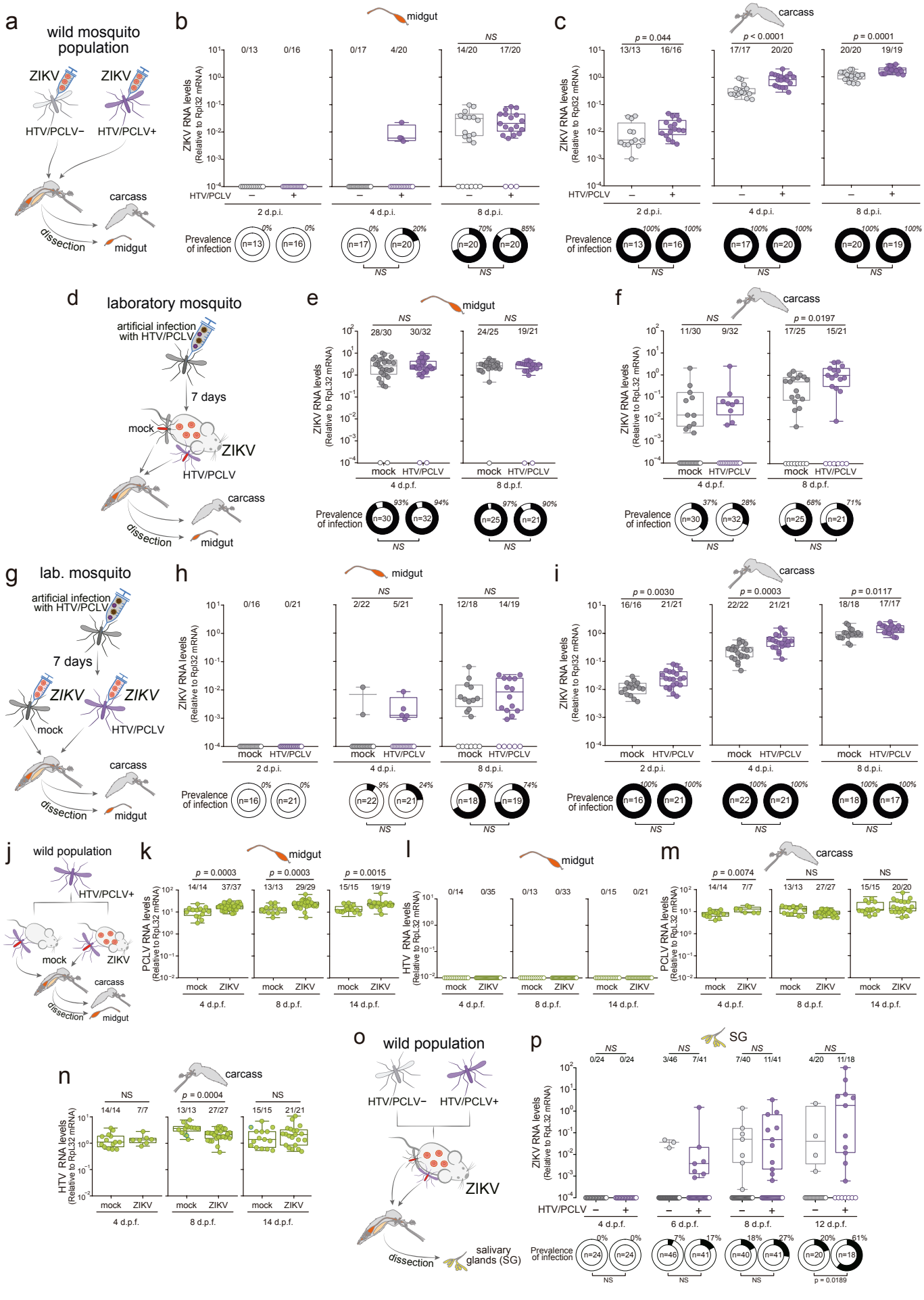
Extended Data Figure 4



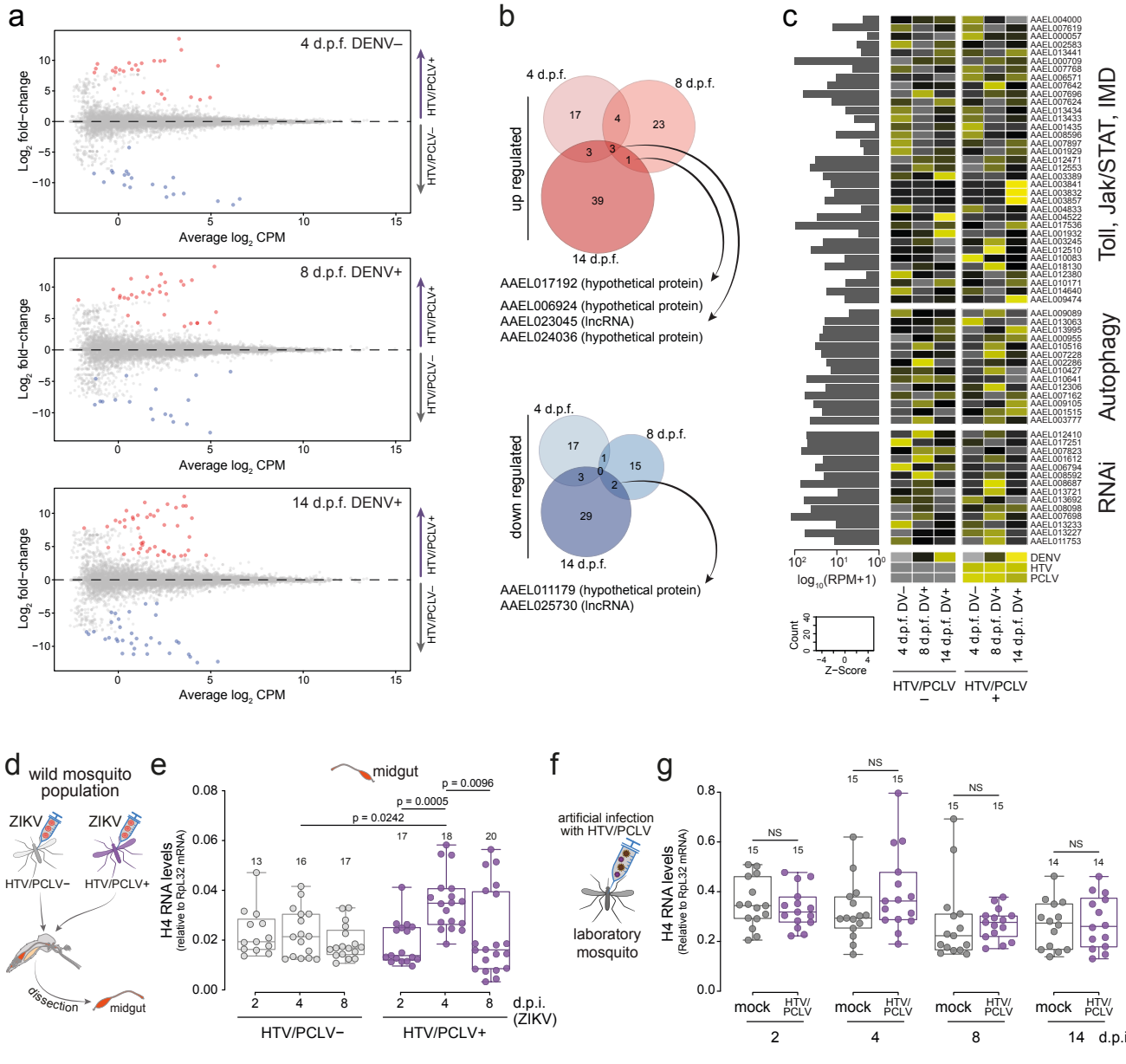
Extended Data Figure 5



Extended Data Figure 6



Extended Data Figure 7



Extended Data Figure 8