Pan-flavivirus analysis reveals that the insect-specific Kamiti River virus produces a new subgenomic RNA and high amounts of 3' UTR-derived siRNAs

Benoit Besson[a*], *Gijs J. Overheul*[a], *Michael T. Wolfinger*[b,c] *Sandra Junglen*[d], *Ronald P. van Rij*[a]#

[a]Department of Medical Microbiology, Radboud Institute for Molecular Life Sciences, Radboud University Medical Center, Nijmegen, The Netherlands

[b]Research Group Bioinformatics and Computational Biology, Faculty of Computer Science, University of Vienna, Vienna, Austria.

[c]Department of Theoretical Chemistry, University of Vienna, Vienna, Austria.

[d]Institute of Virology, Charité-Universitätsmedizin Berlin, Corporate Member of Free University, Humboldt-University and Berlin Institute of Health, Berlin, Germany.

Running title: Mosquito RNAi responses to flaviviruses

#Correspondence to Ronald P. van Rij, ronald.vanrij@radboudumc.nl

*Present address: Benoit Besson, Université de Strasbourg, RNA architecture and reactivity & Insect Models of Innate Immunity, Institut de Biologie Moléculaire et Cellulaire, CNRS, Strasbourg, France.

Benoit Besson and Gijs J Overheul contributed equally to this work. Author order was determined alphabetically.

Word count abstract: 249

Word count text: 5165 (excluding references and figure legends)

ABSTRACT

Flaviviruses subvert the host RNA decay machinery to produce subgenomic flavivirus RNA (sfRNA), products of the 5'-3' exoribonuclease XRN1/Pacman stalling on secondary RNA structures in the 3' untranslated region (UTR) of the viral genome. The classical insect-specific flavivirus (cISF) Kamiti River virus (KRV) has a unique 1.2 kb long 3' UTR, of which only 40% is homologous to its closest family member, cell-fusing agent virus (CFAV). We mapped the 5' end of KRV sfRNAs and found that KRV produces high copy numbers of a long, 1017 nt sfRNA1 and a short, 421 nt sfRNA2, corresponding to two predicted XRN1-resistant elements. Furthermore, we identified a new positive and negative sense 1.5 kb subgenomic RNA species that is colinear with the 3' region of the NS5 gene and the viral 3' UTR, which we tentatively named subgenomic cISF RNA (cifRNA). Expression of both sfRNA1 and sfRNA2 was reduced in *Pacman* deficient *Aedes albopictus* mosquito cells, while expression of the longer cifRNA was *Pacman*-independent. Interestingly, a pan-Flavivirus small RNA analysis in *Aedes albopictus* cells revealed that nearly all KRV-derived siRNAs mapped to the 3' UTR region and that these siRNAs are produced in high quantity. 3' UTR-biased siRNA production appeared to be conserved in other cISFs, albeit to a lesser extent, whereas siRNAs were evenly distributed across the viral genome of other representatives of the *Flavivirus* genus. We suggest that cISFs and particularly KRV developed a unique mechanism to produce high amounts of siRNA as a decoy of the antiviral RNAi response.

IMPORTANCE

The *Flavivirus* genus contains diverse mosquito viruses ranging from insect-specific viruses circulating exclusively in mosquito populations to mosquito-borne viruses that cause disease in humans and animals. Studying the mechanisms of virus replication and antiviral immunity

in mosquitoes is important to understand arbovirus transmission and may inform the development of disease control strategies. In insects, RNA interference (RNAi) provides broad antiviral activity, constituting the main immune response against viruses. Comparing the RNAi response across members of the *Flavivirus* genus, we found that all flaviviruses are targeted by RNAi. However, the insect-specific Kamiti River virus was unique in that small interfering RNAs are highly skewed towards its uniquely long 3' untranslated region. Moreover, we found that Kamiti River virus produces a new subgenomic RNA species in addition to subgenomic flavivirus RNAs previously observed for other flaviviruses. These results suggest that mosquito-specific viruses have evolved unique mechanisms for genome replication and immune evasion.

INTRODUCTION

The genus *Flavivirus* constitutes diverse phylogenetic clades of viruses, found in vertebrates and arthropods including mosquitoes. Mosquito-borne arboviruses are transmitted horizontally between mosquitoes and vertebrates, whereas insect-specific flaviviruses (ISF) are thought to be primarily transmitted vertically and restricted to their arthropod hosts (1, 2). ISFs are further separated into two distinct phylogenetic clades: lineage I or classical ISFs (cISF), a clade that branches at the base of the *Flavivirus* genus, and lineage II or dual-host affiliated ISF (dISF) that forms a separate phylogenetic clade embedded in vector-borne clades (3–5). While the healthcare and economic burden of arboviruses is well established (6), ISFs have been proposed as modulators of arbovirus transmission and are being explored for biotechnological applications such as vaccine development (7–9).

Flaviviruses have an ~11 kb long, positive-sense genomic RNA ((+)gRNA), which circularizes via long range RNA-RNA interactions between their 5' and 3' untranslated regions (UTR) for RNA translation and replication (10, 11). Asymmetric replication is mediated via an antigenomic negative-sense RNA intermediate ((-)gRNA), which serves as a template for replication of the (+)gRNA and is hypothesized to be annealed either to its template and/or to newly synthesized (+)gRNA (12, 13).

Flaviviruses take advantage of the ability of RNA to form regulatory, evolutionarily conserved elements to produce a highly structured subgenomic flavivirus RNA (sfRNA) (14–17). Formation of sfRNA is regulated by exoribonuclease-resistant RNA (xrRNA) structures in the 3' UTR, which typically encompass three-way junctions (3WJ) or stem-loop (SL) elements that adopt a particular fold, mediated by a pseudo knot (18). The tight and complex structure of xrRNAs stalls the 5'-3' exoribonuclease 1 (XRN1, also referred to as Pacman in mosquitoes) and terminates the degradation of viral RNA (19), resulting in the production of

86 sfRNA. Flaviviruses may encode multiple xrRNA-like structures (20), each of which can
87 induce the production of a distinct sfRNA species. While the longest sfRNA generated from
88 the first xrRNA is generally the most abundant, sfRNA production from individual xrRNAs
89 may vary between mammalian and mosquito hosts, suggesting viral adaptation to the host
90 (19, 21–23).
91 It is well established that sfRNA is essential for flavivirus replication and dissemination (15,
92 22, 24, 25), for which several mechanisms have been suggested, in some cases with sfRNA
93 serving as a decoy for the viral genome. For example, sfRNA was shown to inhibit the host
94 RNA decay pathway (26), to control apoptosis (15, 27), to encode a microRNA (28), and to
95 inhibit the mosquito Toll pathway (29). Moreover, sfRNA can be a substrate for small
96 interfering RNA (siRNA) production by Dicer (30) and was proposed to inhibit the RNA
97 interference (RNAi)-based antiviral immune response (26, 31–33), although this was recently
98 disputed (27).
99
100 Mosquitoes have an RNAi-centered immune response, and deficiency in RNAi leads to
101 increased sensitivity to virus infections (34–37). Viral double-stranded RNA (dsRNA) is
102 cleaved by Dicer-2 into 21 nt viral siRNA duplexes (vsiRNAs), which are loaded into the
103 Argonaute-2-containing RISC complex with the help of RNA-binding proteins Loqs and
104 R2D2 (37, 38). Upon loading the duplexes, one of the strands (passenger strand) is degraded
105 and the remaining guide strand is used by Argonaute-2 to recognize and cleave
106 complementary single-stranded viral RNA.
107 In addition to the siRNA pathway, the PIWI-interacting RNA (piRNA) pathway has been
108 implicated in antiviral defense in mosquitoes (37, 39, 40). In this pathway, viral single-
109 stranded RNA is processed into mature 25-30 nt viral piRNAs (vpiRNAs) associating with
110 the PIWI proteins Piwi5 and Ago3, which amplify the piRNA response in a feedforward

mechanism called the ping-pong amplification loop (41–43). While Piwi5 is required for vpiRNA biogenesis in *Aedes aegypti* (41, 42), only *Piwi4* depletion has thus far been shown to affect arbovirus replication (44, 45) and the importance of the piRNA response during acute viral infections remains to be clarified. Yet, endogenous viral elements (EVE) in the genome of *Aedes* mosquitoes give rise to piRNAs that can target cognate viral RNA and reduce viral RNA levels in ovaries (45–48), underlining the antiviral potential of the piRNA pathway.

Having noted that Kamiti River virus (KRV), a cISF originally identified in *Aedes mcintoshi* mosquitoes (49), has a particularly long 3' UTR, we set out to characterize KRV subgenomic RNA species. We mapped two main sfRNAs and identified a new *Pacman*-independent subgenomic RNA, which we refer to as subgenomic cISF RNA (cifRNA). Small RNA sequencing of *Aedes* mosquito cells infected with mosquito-borne and insect-specific flaviviruses revealed that KRV vsiRNAs predominantly derive from the 3' UTR in an sfRNA independent manner. A similar, but less pronounced trend was observed from two other cISFs, Culex flavivirus (CxFV) and cell-fusing agent virus (CFAV), whereas siRNAs mapped across the whole length of the genome for all other flaviviruses tested. We speculate that KRV and likely other cISFs developed a unique mechanism to evade antiviral RNAi.

RESULTS

KRV has a unique 3' UTR

KRV has a 3' UTR of 1208 nt, much longer than in any other member of the *Flavivirus* genus (median of 486 nt), but also longer that the 3' UTRs of members of the cISF clade (median of 663 nt) (Fig. 1A). Structure predictions suggested that KRV 3' UTR is highly structured, comprising evolutionarily conserved elements, alongside RNA secondary structures that appear to be unique to KRV (Fig. 1B). Our model predicted the signature flavivirus regulatory SL at the 3' end of the genome and corroborated the presence of two cISF xrRNA structures (xrRNA1 and xrRNA2) that are highly conserved between KRV, CFAV and Aedes Flavivirus (AEFV) (50), and not conserved in the more distant Culex Flavivirus and Xishuangbanna Aedes Flavivirus (CxFV and XFV, respectively). Moreover, structure predictions of the KRV 3' UTR suggested the presence of simple and branched stem-loop elements, as well as several long hairpins, including the internal 3' stem-loop (i3'SL), previously predicted using a comparative genomics approach (20).

Interestingly, while the 3' terminal 419 nt long sequence of the KRV 3' UTR downstream of xrRNA2 appears to be conserved with other cISFs (20), the 5' terminal 789 nt sequence extending from the stop codon to xrRNA2 appears to be unique to KRV (Fig. B). This 5' sequence of KRV 3' UTR does not seem to share ancestry with AEFV, the cISF with the second longest 3' UTR (Fig. 1A), nor with other flaviviruses, with the exception of xrRNA1 which is highly conserved both in structure and sequence, and was hypothesized to be the result of a self-duplication event (51, 52).

KRV produces multiple subgenomic RNA species

Given the long KRV 3' UTR and the observation that flavivirus 3' UTRs give rise to sfRNAs, we visualized the RNA species produced during KRV infection of *Aedes albopictus*

154 U4.4 cells by northern blot (Fig. 2A left panel). Two sfRNA (sfRNA1 and sfRNA2) were
155 detected, likely the product of XRN1/Pacman stalling on xrRNA structures. Both sfRNAs
156 were visualized only with probes detecting the (+)RNA, displaying strong signals
157 corresponding to the expected sizes (~1000 and ~400 nt), suggesting that KRV sfRNA1 and
158 sfRNA2 outnumber KRV (+)gRNA, as observed for other flaviviruses as well (15, 27, 30).
159 As expected, the (-)gRNA was more difficult to detect than the (+)gRNA, consistent with its
160 lower abundance. Interestingly, we detected a previously unknown viral RNA species of ~1.5
161 kb, which we tentatively refer to as cifRNA for subgenomic cISF RNA (Fig. 2A left panel).
162 Unlike the sfRNAs, cifRNA could be visualized using probes detecting both the (+) and (-)
163 strands, suggesting that they are produced by different mechanisms.
164 We next aimed to quantify the different KRV RNA species by RT-qPCR (Fig. 2B) and found
165 that for each molecule of gRNA, there was a 2.4-fold increase of cifRNA signal. Taking into
166 account that the primer set detecting cifRNA also detects the gRNA, this would correspond to
167 similar levels of gRNA and cifRNA, consistent with the northern blot results. A 10-fold
168 increase of sfRNA1 and 400-fold increase of sfRNA2 relative to gRNA was observed,
169 confirming the presence and high abundance of the two sfRNAs during KRV infection.
170 We characterized the 5' start site of the subgenomic RNA species using a 5'-3' ligation
171 assay. This analysis confirmed that sfRNA1 and sfRNA2 started immediately upstream of
172 xrRNA1 and xrRNA2, resulting in products of 1017 nt and 421-422 nt, respectively (Fig.
173 2C). The 5'-3' ligation assay also identified a hotspot around nt 9831 of the KRV genome,
174 which was found in 60% of the sequenced clones (Fig. 2D). This hotspot would correspond
175 to a 1545 nt subgenomic RNA, presumably the cifRNA detected by northern blot. Consensus
176 secondary structure modelling of the genomic region neighboring the putative cifRNA start
177 within NS5 in KRV, CFAV, and AEFV suggests the presence of a bulged stem-loop element

178 (Fig. S1). We speculate that this structured RNA could be associated with the production of
179 cifRNA.

180 **Pacman-dependent biogenesis of KRV sfRNA**

181 To determine whether biogenesis of the subgenomic RNAs is *Pacman*-dependent, we used
182 CRISPR/Cas9 gene editing to create *Pacman* knockout (KO) U4.4 cell lines. Several putative
183 *Pacman* loci are annotated in the genome of *Aedes albopictus*, of which AALFPA_065179,
184 AALFPA_057530 and AALFPA_079140 contain the conserved 5'-3' exoribonuclease
185 domain (>98% identity) and AALFPA_052256 only contains the SH3-like domain and is
186 unlikely to encode a functional Pacman nuclease (Fig. S2A). Guide RNAs were designed to
187 introduce frameshift mutations leading to premature stop codons in the 5'-3' exoribonuclease
188 domain. Two *Pacman* KO U4.4 cell clones were obtained (g3#3 and g2#13), which were
189 compared to a CRISPR control line (CTRL) that was subjected in parallel to the same
190 treatment without functional guide RNA, and to the wildtype (WT) parental U4.4 cell line.
191 *Pacman* mRNAs containing the 5'-3' exoribonuclease domain were unstable in both *Pacman*
192 KO U4.4 cell clones (Fig. S2B), likely due to nonsense mediated decay induced by the
193 presence of premature stop codons. KRV replicated to similar levels in *Pacman* KO cells as
194 in WT and CTRL cells (Fig. S2C).

195 Using northern blotting, we observed lower signal for KRV sfRNA1 and sfRNA2 in KRV
196 infected *Pacman* KO cells, confirming that their biogenesis is *Pacman*-dependent (Fig. 2B,
197 bottom right panel). Interestingly, two different ~800 nt and ~500 nt subgenomic RNAs were
198 identified in *Pacman* KO cells, which we named sfRNA1' and sfRNA2', likely the products
199 of redundant 5'-3'exoribonucleases stalling on structures downstream of xrRNA1 (53, 54).
200 This is reminiscent of the appearance of other RNA species without loss of sfRNA upon
201 knockdown of XRN1 in human cells (55). The exact 5' start sites of sfRNA1' and sfRNA2'
202 were determined by 5'-3'end ligation to be at nt 10,533 and 10,830 of the KRV genome,

respectively (Fig. 2C). These sites did not correspond to notable predicted structures or RNA motifs (Fig. 1C, data not shown). In contrast to sfRNAs, cifRNA appeared to be *Pacman*-independent, suggesting that it is produced via another mechanism (Fig. 2A, top right panel). In Zika virus (ZIKV) infected wild-type cells, sfRNA was highly abundant but no cifRNA was detectable. Moreover, a significant smear appeared *Pacman* KO cells, suggesting inefficient processing of viral RNA in the absence of Pacman (Fig. S2D).

RNAi response to flavivirus infection in *Aedes* mosquito cells

Given our observation that KRV produces longer subgenomic RNAs than other flaviviruses and the proposed function of sfRNA as a viral escape mechanism, notably in small RNA silencing pathways (31, 56), we analyzed viral small RNAs produced during KRV infection in comparison to other flaviviruses. Representatives of each major clade of mosquito-associated flaviviruses were selected to provide a pan-flavivirus overview of viral siRNA and piRNA profiles in *Ae. albopictus* U4.4 cells. *Culex*-associated arbovirus Saint-Louis encephalitis virus (SLEV, isolate MSI-7) and West Nile virus (WNV), *Aedes*-associated arbovirus dengue virus (DENV serotype 2) and ZIKV, and the dISF Nounané virus (NOUV), *Culex*-associated cISF CxFV and *Aedes*-associated cISF CFAV and KRV were studied (Fig. 3A). Further, the epidemic SLEV-MSI-7 strain was compared to the ancestral strain SLEV-Pal as representatives for cosmopolitan and epidemic versus enzootic mosquito-borne flaviviruses (Fig. S4) (57). All tested flaviviruses replicated to similar levels in U4.4 cells with approximately 10^8 RNA copies/µg of total RNA at 72 hours post infection, except for CxFV and CFAV, which reached $3\text{-}5.10^6$ copies/µg of total RNA (Fig. S3A).

As observed previously (2, 30, 41, 44, 46, 58, 59), size profiles of viral small RNAs are characterized by a prominent peak of 21 nt vsiRNA from both positive- and negative-sense RNA for all tested flaviviruses (Fig. 3B), with a shoulder of predominantly positive-sense

228 RNAs of 25-30 nt. Given similar viral RNA levels (Fig. S3A), differences in scales suggest
229 that CxFV and NOUV elicit an overall weaker siRNA response compared to WNV, SLEV-
230 MSI-7/Pal, DENV or ZIKV, while KRV elicits the strongest siRNA response recorded, and
231 CFAV induces a strong siRNA response despite its relatively low RNA levels in cells (Fig.
232 S3B).
233 The shoulder of 25-30 nt small RNAs in Figure 3B likely represent vpiRNAs associated with
234 1U-/10A-bias as observed for several tested flavivirus (Fig. S5), although we have not
235 formally demonstrated PIWI association. Using the gRNA as reference, the flaviviruses
236 differed from each other in the relative amount of 25-30 nt viral small RNAs (Fig. S3C).
237 Notably, viral piRNA over siRNA ratios were relatively low for *Aedes*-associated arboviruses
238 DENV and ZIKV (0.08 and 0.04, respectively), whereas these ratios were higher for NOUV
239 and CFAV (0.77 and 0.85, respectively) (Fig. 3B).
240
241 **Asymmetric distribution of vsiRNAs across the KRV genome**
242 The distribution of vsiRNAs across the viral genomes (Fig. 4, S4) showed relative uniform
243 mapping of vsiRNAs on both the (+)gRNA and (-)gRNA. A notable exception was KRV, for
244 which most vsiRNAs mapped to the 3' UTR region. This was not due to an artifact, as
245 presentation of the data on a logarithmic scale indicates that siRNAs, albeit extremely lowly
246 abundant, also map to other parts of the genome. This pattern is reminiscent of 3' UTR
247 biased mapping observed for the other cISFs, CxFV, CFAV and AEFV, although the skewed
248 distribution is much more pronounced for KRV (Fig. 4A-B) (2, 46, 60). About 14% of the
249 vsiRNAs of CxFV and CFAV and more than 95% of KRV vsiRNAs derived from their 3'
250 UTRs, in stark contrast to the other flaviviruses for which a median of ~4% of vsiRNA
251 mapped to the 3' UTR (Fig. 4C).

252 The distribution of vsiRNAs on the 3' region of both CFAV and KRV, either expressed as a
253 percentage of the genome-mapping vsiRNAs (Fig. 4D) or as a density of vsiRNA per nt (Fig.
254 4E), was further investigated. The 3' UTR region of CFAV was clearly associated with a
255 higher density of vsiRNAs. In contrast, only a negligible amount of KRV vsiRNA derived
256 from its gRNA-specific regions (<2%), 5% of vsiRNAs mapped to the region specific of
257 cifRNA, whereas 92% of KRV vsiRNA mapped to the sfRNA region in the 3' UTR. Thus,
258 the 3' bias characteristic of cISF and especially KRV-derived vsiRNAs correlated highly
259 with the 3' UTR, yet, the vsiRNA distribution indicates that the substrate for vsiRNA
260 production is longer than its sfRNAs (Fig. 4B).
261 In contrast to siRNAs, vpiRNAs mapped to several discrete hotspots on the viral (+)gRNA
262 (Fig. S6, S4). For each virus analyzed, vpiRNAs mapped to different genome coordinates in a
263 manner that was highly reproducible in replicate experiments, in agreement with previous
264 observations (39, 44, 46, 61). It is worth noting that KRV derived piRNAs mapped at several
265 hotspot across the gRNA and were not enriched at the 3' UTR, which indicates that each
266 pathway processes a different substrate. Altogether, our data illustrate a general antiviral
267 siRNA response to flaviviruses and highlight the unique case of cISFs, especially KRV, for
268 which the skewed distribution of vsiRNA towards the 3' UTR suggests a unique siRNA
269 response to the infection.
270

271 **Viral small RNA production in *Pacman* knockout cells**

272 It is unlikely that sfRNAs are the substrate for vsiRNA production, given that the KRV 3'
273 biased vsiRNA derived from both (+) and (-) RNA in equimolar quantities, whereas sfRNAs
274 exclusively derive from the (+) RNA strand. To strengthen this conclusion, we explored the
275 involvement of sfRNAs in the 3' bias of KRV vsiRNAs by comparing vsiRNA profiles in
276 U4.4 control cells with *Pacman* KO cells, in which two new KRV subgenomic RNAs were

277 produced (sfRNA1' and 2'; Fig. 2A). We also analyzed ZIKV and found that *Pacman*
278 knockout did not affect sfRNA production at the resolution of our northern blot, although a
279 smear of larger RNA fragments was observed (Fig. S2D) and may be specific to mosquito
280 cells (55).

281 Interestingly, total siRNA levels were higher in *Pacman* KO cells than in control cells (Fig.
282 5A), perhaps due to the higher processing of mRNA by the siRNA pathway when the RNA
283 decay pathway was impaired. In contrast, vsiRNA levels decreased slightly for KRV and
284 ZIKV in the absence of Pacman (Fig. 5B, S2C). Moreover, in the segment differentiating
285 KRV sfRNA1 from the *Pacman* KO associated sfRNA1' and 2' (nt 10361-10533), no
286 difference in vsiRNA distribution was observed (Fig. 5C, 5E). Similarly, no major
287 differences were observed for vsiRNA profiles of ZIKV between *Pacman* KO and control
288 cells (Fig. 5D, 5F). These results further strengthen the conclusion that KRV 3' biased
289 vsiRNAs are not produced from its sfRNA species.

DISCUSSION

Within the *Flavivirus* genus, cISFs constitute a unique clade of viruses that evolved independently, only infecting invertebrate hosts in which they are not associated with known symptoms (62). As such, cISFs represent a prime resource to better understand viral infection and antiviral immunity in mosquitoes. In this study, we explored the potential role of the exceptionally long and unique 3' UTR of KRV. We found that KRV produces high quantities of two sfRNAs as well as a putative new subgenomic RNA referred as cifRNA. In a pan-flavivirus small RNA analysis, we found that vsiRNAs generally mapped across the viral genome for most mosquito-specific and mosquito-borne viruses, while there was a strong vsiRNA bias toward the 3' UTR of KRV, the production of which was independent of sfRNAs.

Unique RNAi response toward classical insect specific flaviviruses is sfRNA-independent

RNAi is a cornerstone of mosquito immunity comparable to the importance of the interferon response in mammalian systems, as its deficiency leads to increased sensitivity to viral infections (35, 63–65). Our pan-flavivirus analysis strengthens previous observations in *Anopheles* (66), *Culex* (58, 67) and *Aedes* (41, 63) that mosquito RNAi raises a broad and uniform antiviral response against all assessed mosquito-borne flaviviruses. Yet, cISFs seem to have evolved to produce a unique RNAi response with vsiRNAs biased towards the 3' UTR of the viral genome, which was particularly strong for KRV but also detectable for CxFV, CFAV, and previously for AEFV (2). Interestingly, we did not observe a 3' vsiRNA bias for the dISF NOUV or for the sylvatic SLEV-Pal (57), indicating that the biased vsiRNA production is not required for a mosquito restricted transmission cycle.

The homogeneous distribution across the genome and the absence of a strand bias of vsiRNAs is consistent with processing of flaviviruses dsRNA formed by (+) and (-) gRNA hybrids (35). The 3' vsiRNA bias of KRV and other cISFs suggest a correlation with viral RNA species produced specifically by cISFs, which remain to be elucidated. A 3' vsiRNA bias has previously been suggested to be related to sfRNA and RNA structure of the region (2), but our data do not support such hypothesis. First, KRV vsiRNAs derive equally from sfRNA1 or sfRNA2 regions, while more vsiRNAs would be expected toward the 3' end due to the high abundance of both sfRNA1 and sfRNA2 (22, 27, 30). Second, vsiRNA patterns are not affected in *Pacman*-knockout cells, whereas sfRNA production is affected in those cells. Third, the vsiRNA 3' bias starts upstream of the 3' UTR and the sfRNAs, and, fourth, both (+) and (-) sense vsiRNA are present at equimolar levels, whereas sfRNA is a (+) sense RNA. Thus, the mechanism underlying the 3' bias of vsiRNAs remains to be understood and may be multi-factorial. Perhaps the double-stranded cifRNA is a substrate of vsiRNA biogenesis, but other hypotheses should be considered as well. For example, sfRNAs could bind the 5' end of (-)gRNA as a competitive regulator of RNA replication (68–70), which would generate a double-stranded Dicer substrate.

KRV produces a putative novel subgenomic RNA

As ISFs infect only insect hosts, they evolved to adapt to a single organism compared to vector-borne flaviviruses and present specific features such as the *fairly interesting Flavivirus ORF* derived from a ribosomal frameshift in the NS2A-NS2B coding sequence (71). Here, based on northern blot analyses and end-to-end ligation, we propose that KRV and possibly other cISFs encode a previously unknown subgenomic RNA species, which we named subgenomic cifRNA. The presence of both (+) and (-) sense cifRNA at similar levels with gRNA suggests that cifRNA biogenesis involved negative strand synthesis by the NS5

polymerase (12). Different models have been described for the formation of subgenomic RNA through an (-) RNA intermediate, including long range RNA-RNA interactions as for members of the order *Nidovirales* (72–74), premature termination of replication or internal promoters as for tombusviruses and alphaviruses (75, 76). If similar mechanisms are responsible for cifRNA production remains to be investigated. The evolutionary conserved, predicted RNA structure in the immediate proximity of the 5' end cifRNA would be an interesting candidate to investigate as a potential regulator of cifRNA production.

The generation of cifRNA raises questions about its possible function(s) in cISF replication and vertical transmission in mosquito populations (77, 78). It could potentially be used as a template for new unidentified proteins: the longest predicted ORF could produce a ~10 kDa peptide in frame with the C-terminal region of NS5, shorter than its last functional RdRp domain. Alternatively, cifRNA could serve as a regulator of viral RNA replication, for example by sequestering NS5 or other proteins required for viral genome replication.

Conclusion

As part of the constant arms race between viruses and their hosts (79), cISFs and especially KRV have evolved unique ways to maintain themselves in mosquito populations. KRV's strikingly long 3' UTR representing 10% of its total gRNA, combined with the expression of two highly abundant sfRNAs, a subgenomic cifRNA species, as well as the strong 3' bias of vsiRNAs makes KRV an intriguing model to study the biology of cISFs and the mechanisms of mosquito antiviral immunity.

ACKNOWLEDGEMENTS

We thank Pascal Miesen for preliminary analysis of small RNA sequencing data. We thank Andrea Gamarnik (Fundación Instituto Leloir, Buenos Aires, Argentina) for advice on the 5'-3' ligation method. We thank Beate Kümmerer (University of Bonn) for providing DENV2 strain 16681 and M. Niedrig (Robert-Koch-Institute, Berlin) for providing WNV strain NY-99. CFAV (isolate Rio Piedras), CxFV (isolate Uganda08), KRV (isolate SR-75) and ZIKV isolate H/PF/2013 were provided by Xavier de Lamballerie (Aix Marseille Université) through the European Virus Archive (EVAg), funded by the European Union's Horizon 2020 programme. This study was financially supported by the Deutsche Forschungsgemeinschaft (DFG) within Priority Programme SPP1596 [grant number JU 2857/1-1 and JU2857/1-2 to SJ and grant number VA 827/1-1 and RI 2381/1-2 to RPvR] and by a European Research Council Consolidator Grant under the European Union's Seventh Framework Programme [grant number ERC CoG 615680 to RPvR]. Small RNA sequencing was performed by the GenomEast platform, a member of the "France Génomique" consortium (ANR-10-INBS-0009).

MATERIAL AND METHODS

Cells and viruses

Aedes albopictus C6/36 cells (ECACC General Cell Collection, #89051705) and U4.4 cells (kindly provided by G.P. Pijlman, Wageningen University, the Netherlands) were cultured at 25°C in Leibovitz L15 medium (Gibco) supplemented with 10% heat inactivated fetal calf serum (Sigma), 2% tryptose phosphate broth solution (Sigma), 1x MEM non-essential amino acids (Gibco), and 50 U/ml penicillin and 50 µg/ml streptomycin (Gibco).

CFAV (isolate Rio Piedras), CxFV (isolate Uganda08), KRV (isolate SR-75) and ZIKV (isolate H/PF/2013) were obtained from the European Virus Archive. DENV2 (strain 16681) was provided by Beate Kümmerer, University of Bonn. SLEV (strain MSI-7) was obtained from the National Collection of Pathogenic Viruses, Porton Down, Salisbury, United Kingdom and WNV (strain NY-99) was kindly provided by M. Niedrig, Robert-Koch-Institute, Berlin, Germany. SLEV (Palenque) and NOUV (isolate B3) were previously characterized (57, 80). Reference sequences are listed in Table S1.

Plasmids

Plasmids used as standard to normalize the qPCR data were obtained by blunt end ligation of the flavivirus qPCR products into pGEM-3Z (Promega), using SmaI (NEB) and T4 DNA ligase (Roche). For KRV RNA species quantification, plasmid pUC57-KRV-9000-11375 containing part of NS5 gene and the 3' UTR of KRV (nt 9000-11375) was synthesized by GenScript.

The pAc-Cas9-AalbU6.2 plasmid was generated by replacing the *D. melanogaster U6* promoter in the pAc-sgRNA-Cas9 (gifted by Ji-Long Liu, Addgene plasmid #49330) (81) with the *Ae. albopictus U6.2* promoter for AALFPA_045636 and by introducing XbaI restriction sites using In-Fusion cloning (Takara) on 4 fragments amplified by the primers in

Table S2, according to the manufacturer's instructions. sgRNA sequences targeting *Pacman* were cloned directly after the U6 promoter using SapI restriction, followed by ligation using T4 DNA ligase (Roche) and annealed and phosphorylated complementary oligos (Table S2), replacing the 5'-GGAAGAGCGAGCTCTTCC-3' sequence that was used as negative control for CRISPR/Cas9 knockouts.

CRISPR/Cas9

To obtain the CRISPR control (CTRL) and *Pacman*-knockouts, U4.4 cells were transfected with the pAc-Cas9-AalbU6.2 plasmid. It expresses 3xFLAG-tagged Cas9 with N- and C-terminal SV40 nuclear localization signals, followed by a viral 2A ribosome self-cleavage site and the puromycin N-acetyltransferase coding sequence, driven by the *D. melanogaster Actin5c* promoter. U4.4 cells were seeded in a 24-well plate and transfected the next day with 500 ng of plasmid, using X-tremeGENE HP transfection reagent (Roche) according to the manufacturer's instructions. At 2 days after transfection, puromycin (InvivoGen) was added to the culture medium at a concentration of 20 µg/ml and 4 days later, the cells were transferred to a new plate in a 1:2 dilution, again with 20 µg/ml puromycin. The other half of the cells was used for gDNA isolation (Zymo Research #D3024) and PCR (Promega #M7806) using primers flanking the sequence targeted by the sgRNA (Table S2) to assess editing efficiency. Multiple sgRNA constructs were initially constructed and the constructs with the highest editing efficiency were selected, assessed by size heterogeneity of the PCR products on ethidium bromide-stained agarose gel. Cells transfected with these sgRNA constructs (g2 and g3) were seeded in 96-well plates at a density of a single cell per well in supplemented L15 medium in the absence of puromycin. After 3 weeks, gDNA was isolated from the single-cell clones, followed by PCR and Sanger sequencing of the targeted *Pacman*

425 locus. Based on the sequencing results U4.4 clones g2#13 and g3#3, both containing only
426 out-of-frame deletions in the *Pacman* coding region, were selected for further analyses.
427
428 **Small RNA library preparation and analysis**
429 U4.4 cells were seeded at a density of $2x10^6$ cells per well in 6-well plates and infected the
430 following day with the designated flaviviruses at an MOI of 0.1. Cells were harvested at 72
431 hours post infection (hpi) in RNA-Solv reagent (Omega Biotek R630-02) for total RNA
432 isolation. 1 µg of RNA was used as input for library preparation using the NEBNext
433 Multiplex Small RNA Library Prep Kit for Illumina (NEB E7560S), according to the
434 manufacturer's recommendations. The libraries were size-selected on 6% polyacrylamide/1x
435 TBE gels, quantified using the Agilent 2100 Bioanalyzer System, and pooled for sequencing
436 on an Illumina HiSeq4000 machine by the GenomEast Platform (Strasbourg, France). Viral
437 small RNA sequences were mapped to the designated genome (Table S1) using Bowtie
438 (Galaxy Tool Version 1.1.2) (82) allowing 1 mismatch. The genome distribution of the viral
439 small RNAs was obtained by plotting the 5' ends of mapping reads on the viral genome
440 sequence. Nucleotide biases were plotted using the WebLogo 3 program (Galaxy Tool
441 Version 0.4). All reads were normalized by library size as reads per million. The small RNA
442 sequencing datasets have been deposited at the Sequence Read Archive under accession
443 number PRJNA830662.
444
445 **Reverse transcription and quantitative PCR**
446 For RT-qPCR, 1 µg of total RNA was reverse transcribed in a 20 µl reaction using Taqman
447 RT Reagents (ThermoFisher Scientific) with hexamers at 48°C for quantification of viral
448 RNA copies or SuperScript IV Reverse Transcriptase (ThermoFisher Scientific) with primer
449 5'-AGCGCATTTATGGTATAGAAAAGA-3' at 60°C for quantification of specific KRV

450 RNA species . Quantitative PCR was performed using the GoTaq qPCR SYBR mastermix
451 (Promega) on a LightCycler 480 instrument (Roche). A standard curve of plasmids
452 containing the corresponding viral sequence was used to convert Ct values to relative viral
453 RNA copy numbers. *Pacman* mRNA levels were normalized to house-keeping gene
454 *ribosomal protein L5*. For qPCR primer sequences, see Table S2.
455
456 **Northern Blot**
457 5 µg of total RNA was separated on a 1X MOPS, 5% formaldehyde, 0.8% agarose gel for 5
458 h, transferred on a Hybond NX nylon membrane (Amersham) and cross-linked in the Gene
459 linker UV chamber (Bio-Rad). Viral RNAs were detected with DNA oligonucleotides (Table
460 S2) end-labelled with [^{32}P] γ-adenosine-triphosphate (Perkin Elmer) using T4 polynucleotide
461 kinase (Roche). Hybridization to the oligonucleotide probes was performed overnight at 42°C
462 in Ultra-hyb Oligo hybridization buffer (Ambion). Membranes were then washed three times
463 at 42°C with decreasing concentrations of SDS (0.2 to 0.1%) and exposed to X-ray films
464 (Carestream).
465
466 **5' to 3' end ligation**
467 The 5' ends of KRV RNA species were determined by 5' to 3' end ligation using a method
468 adapted from (10, 83). C6/36 cells were seeded at a density of $4x10^6$ cells per T75 flask and
469 infected the following day with KRV at an MOI of 10. Cells were harvested at 48 hpi and
470 RNA was isolated using RNA-Solv reagent and 10 µg of total RNA was treated with T4
471 RNA ligase (Epicenter). Ligated RNAs were revere transcribed with Taqman Multiscribe
472 (Applied Biosystem) using random hexamers. The 5'-3' junction region was amplified by
473 PCR using a forward primer at the end of the 3' UTR and a reverse primer in the 5' part of

474 KRV RNA species (Table S2), cloned into pGEM-3Z (Promega) using In-Fusion technology
475 (Takara) and Sanger sequenced by the in-house sequencing facility.
476
477 **RNA structure prediction and bioinformatics**
478 RNA secondary structure predictions of the KRV, CFAV, CxFV, and XFV 3'UTR regions
479 was performed with the ViennaRNA Package v.2.5.1 (84). Evolutionarily conserved
480 elements were identified with the help of the viRNA GitHub repository
481 (https://github.com/mtw/viRNA), and used as constraints for RNA structure prediction.
482 Locally stable RNA structures were predicted with RNALfold from the ViennaRNA
483 Package, allowing for a maximal base pair span of 100 nt.
484 A consensus secondary structure of the genomic region upstream of the 3'UTR was
485 computed from a structural LocARNA (85) alignment of the respective regions with
486 RNAalifold from the ViennaRNA Package (86). Consensus xrRNA structure predicted from
487 CFAV SHAPE data (50) was visualized using VARNA-v3.93 (87).
488 Multiple sequence alignments of whole genome and 3' UTR sequences were generated with
489 MAFFT (88), curated with BMGE (89) and a maximum likelihood phylogenetic tree was
490 built with PhyML (90) using NGPhylogeny.fr (91) using default settings. Phylogenetic trees
491 were visualized on iTOL (92). Percentage identities were determined with Mview (93).
492 Viral reference sequences are listed in Table S3.
493
494 **Statistical analysis**
495 Graphical representation and statistical analyses were performed using GraphPad Prism v7
496 software. Differences were tested for statistical significance using one- or two-way ANOVA
497 and Fisher's LSD test.
498

LEGENDS

Figure 1. KRV has a long and unique 3' UTR

(A) Length of the 3' UTR of all members of the *Flavivirus* genus with a RefSeq, a complete coding genome and a 3' UTR of at least 200 nt. Viruses belong to the clades indicated: cISF, classical insect specific flaviviruses; dISF, dual-host affiliated insect specific flaviviruses; MBF, mosquito-borne flaviviruses; NKV, no known vector; TBF, tick-borne flaviviruses. For virus name and accession numbers, see Table S1. **(B)** Secondary structure prediction of the 3'UTR of four insect-specific flaviviruses. Maximum likelihood phylogenic tree and alignment of 3' UTR of listed viruses with conserved regions as described in (94). Branch lengths are proportional to the number of substitutions per site. Evolutionarily conserved RNA elements are highlighted in colour, indicating that elements depicted in the same colour are structurally homologous. Elements without colour represent locally stable RNA structures from single-sequence RNA structure predictions. Exoribonuclease-resistant structures (xrRNA) in KRV, CFAV and AEFV are shown in blue, including reported pseudoknot interactions (16) with sequence regions downstream of the three-way junction structures. Repeat elements a (Ra), and b (Rb) (20) are depicted in olive and orange, respectively. 3' stem-loop elements (3'SL) are shown in dark green. The internal 3'SL element of KRV is predicted to adopt a longer closing stem, which lacks evolutionary support in other viruses. The same applies for the extended closing stems of Ra elements in XFV. Percent nucleotide identities of each virus to KRV are indicated for the region between xrRNA2 and the 3' SL. Lengths of the 3' UTRs are indicated on the right.

Figure 2. KRV produces high quantities of subgenomic RNA species

(A) Northern blot of positive-sense (POS) or negative-sense (NEG) viral RNA in wildtype (WT) or *Pacman* knockout (KO) U4.4 cells mock infected (-) or infected with KRV (+) at an

524 MOI of 0.1 for 72 h. Viral RNA was detected using a pool of oligonucleotide probes for the
525 3' UTR of KRV, between positions 10,361 and 11,375. All images were captured from the
526 same northern blot. For uncropped images, see Fig. S7. **(B)** Relative RT-qPCR quantification
527 of KRV RNA in U4.4 cells infected for 72 h at an MOI of 0.1. Data are expressed relative to
528 gRNA copy numbers and bars indicate means and standard deviation of four replicates. * $p <$
529 0.05; ** $p < 0.01$ by one-Way ANOVA and Fisher's LSD test. **(C)** Position of 5' ends of
530 KRV sfRNA1 and sfRNA2 defined by 5' to 3' end ligation and sequencing, displayed on the
531 consensus xrRNA structure predicted from SHAPE data for CFAV xrRNA1, which is 90%
532 identical to KRV xrRNA1 and xrRNA2 (50). **(D)** Position of 5' ends of KRV cifRNA
533 identified by 5' to 3' end ligation. Positions are indicated relative to the genome sequence.
534

535 **Figure 3. Comparison of flavivirus-derived small RNAs in U4.4 cells**

536 **(A)** Maximum likelihood phylogenic tree based on whole genome sequences of the indicated
537 viruses. Branch lengths are proportional to the number of substitutions per site. MBF,
538 mosquito-borne flavivirus; ISF, insect-specific flavivirus. **(B)** Size profiles of flavivirus-
539 derived small RNAs in read per million (RPM) from U4.4 cells infected for 72 h at an MOI
540 of 0.1. The results are the average of two experiments for all flaviviruses, except for CxFV (n
541 = 1), CFAV (n = 3, of which 2 in WT U4.4 cells and 1 in CRISPR CTRL U4.4 cells), and
542 KRV (n = 3). Error bars are the standard deviation between replicates. Positive-sense RNAs
543 are shown in red, negative-sense RNA in blue. Ratios of viral piRNAs over siRNAs are
544 indicated for each virus.
545

546 **Figure 4. KRV vsiRNAs are strongly biased towards the 3' UTR**

547 **(A)** Distribution of flavivirus-derived vsiRNAs across the genome of each virus in reads per
548 million (RPM) from U4.4 cells infected at a MOI of 0.1 for 72h. Start and end of the 3' UTRs

are indicated by dashed vertical lines. **(B)** Distribution of KRV vsiRNAs on a logarithmic scale with positions of the cifRNA, 3' UTR, and sfRNAs indicated. **(C)** Percentage of vsiRNAs mapping to the 3' UTR compared to the whole genome sequence for the indicated flaviviruses. The dashed horizontal line indicates the median of 3' UTR derived vsiRNA from non-cISFs. **(D-E)** Percentage of vsiRNAs compared to the whole genome sequence **(D)** and average density of vsiRNAs per nucleotide (E) in the indicated regions of the genome of CFAV (left) or KRV (right). The size of each region is indicated in italic as a percentage of the genome size. For CFAV, a start position of 9800 was assumed for the putative cifRNA. (A-E) The results are the average of two experiments for all flaviviruses, except for CxFV (n = 1) and KRV (n = 3). Error bars are the standard deviation between replicates for each individual nucleotide. Positive-strand RNAs are shown in red, negative-strand RNAs in blue. * $p < 0.05$; **** $p < 0.0001$ by two-way ANOVA and Fisher's LSD test.

Figure 5. Pacman knockout does not affect vsiRNA profiles

(A-B) Quantification of total siRNAs (A) and sense (+) and antisense (-) vsiRNAs (B) in wild-type (WT), *Pacman* control (CTRL) or *Pacman* knockout (KO) U4.4 cell lines infected with ZIKV at an MOI of 0.1 or KRV at an MOI of 10 for 72 h. Viral siRNA levels were normalized to viral gRNA levels. Errors bars represent the standard deviation from two independent cell lines. ns, non-significant; *, $p < 0.05$ by two-way ANOVA and Fisher's LSD test. **(C-D)** Distribution of (+) and (-) vsiRNAs across the 3' end of the genomes of KRV (C) or ZIKV (D) (from nt 9000 onwards) in control and *Pacman* KO cells. Top panels show (+) vsiRNAs and lower panels show (-) vsiRNAs. Boundaries of subgenomic RNAs and 3' UTRs are indicated by dashed vertical lines. The results are the average from two independent cell lines. **(E-F)** Percentage of vsiRNAs derived from the indicated regions compared to the entire gRNA. The size of each region is indicated in italic as a percentage of

the viral genome size. Errors bars represent the standard deviation from two independent cell lines. (C-F) Blue, (+) vsiRNA; red, (-) vsiRNA; darker, WT and CTRL cells; lighter, *Pacman*-KO cells.

Supplementary figure 1. Consensus structure prediction of KRV, CFAV and AEFV region neighboring the putative cifRNA start within NS5.

A. Structural alignment of homologous regions in KRV, CFAV, and AeFV represented with the RNA dot-bracket annotation, with unpaired nucleotides as dots and base pairs as brackets. The RNAalifold color scheme overlay indicates observed structure conservation supported by covariation levels of each base pair. Allowed pairs: A-U, G-C and U-G. Red asterisk, putative cifRNA start. Gray bar scale, sequence homology. **B.** Predicted consensus secondary structure.

Supplementary figure 2. Characterization of flavivirus infection of *Pacman*-KO U4.4 cells

(A) *Pacman* KO U4.4 cell lines were generated by CRISPR/Cas9 technology, amplified from single clones, and the edited sites in exon 4 were Sanger sequenced. Sequencing identified three small deletions that all induced out-of-frame mutations in the exoribonuclease domain in both *Pacman* KO clones. Gene structure, conserved domains and primer sets used in (B) are indicated. **(B)** Relative quantification by RT-qPCR of *Pacman* mRNA in *Pacman* CTRL and KO cells infected with KRV at a MOI of 1 for 96h. *Pacman* mRNA levels were normalized to house-keeping gene *ribosomal protein L5* and *Pacman* mRNA in WT cells. **(C)** Quantification of ZIKV or KRV RNA in cells and culture supernatants of the indicated U4.4 cells at 72 h post infection at an MOI of 0.1 or 10. **(D)** Northern blot of ZIKV RNA in

598 wildtype (WT) and *Pacman* KO U4.4 cells infected at a MOI of 0.1 (+) or mock infected (-)
599 for 72 h. Probes detecting the 3' UTR sequence were used.
600
601 **Supplementary figure 3. Relative quantification of gRNA, vsiRNA and vpiRNA.**
602 **(A)** Quantification of viral gRNA copies, **(B)** total vsiRNA over gRNA copy ratios, and **(C)**
603 (+) vpiRNA over gRNA copy ratios in the samples analyzed in Figures 3-4 and S6. U4.4
604 cells were infected with the indicated virus at a MOI of 0.1 for 72h. Error bars are the
605 standard deviation of at least 2 independent experiments. ns, non-significant by one-way
606 ANOVA and Fisher's LSD test.
607
608 **Supplementary figure 4. Comparison of viral small RNA profiles of ancestral SLEV-Pal**
609 **with the pandemic SLEV-MSI-7.**
610 Small RNA profile **(A)**, distribution of vsiRNAs **(B)** and vpiRNAs **(C)** over the genome of
611 SLEV-Pal comparison to SLEV-MSI-7 (data from Figures 3-4 and S6) showed no major
612 differences between the ancestral and pandemic strains. U4.4 cells were infected with the
613 indicated virus at a MOI of 0.1 for 72h. The results are the average of two experiments. Error
614 bars are the standard deviation between replicates for each individual nucleotide. Positive-
615 sense RNAs are shown in red, negative-sense RNAs in blue.
616
617 **Supplementary figure 5. Sequence logos of flavivirus vpiRNAs**
618 Sequence logos of flavivirus derived 25-30 nt vpiRNAs. The 1U-bias characteristic of PIWI
619 protein-associated small RNAs is detectable for vpiRNAs of CFAV, KRV, DENV and
620 SLEV. 10A-bias on the positive strand characteristic of ping-pong piRNA amplification is
621 detectable for vpiRNAs of CFAV, KRV, DENV.
622

623 **Supplementary figure 6. Viral piRNA profiles of flavivirus infected U4.4 cells**

624 **(A)** Distribution of flavivirus derived 25-30 nt vpiRNAs on the genome of each virus in reads
625 per million (RPM) from U4.4 cells infected for 72h at a MOI of 0.1. **(B)** Percentage of
626 vpiRNAs mapping to the 3' UTR compared to the entire gRNA for the indicated flaviviruses.
627 **(C)** Percentage of vpiRNAs mapping to the indicated regions compared to the whole gRNA
628 and **(D)** average density of vpiRNAs per nucleotide in different regions of the genome of
629 CFAV (left) or KRV (right). The size of each region is indicated in italic as a percentage of
630 the gRNA size. In (C) and (D), only (+) vpiRNAs are shown. (**A-D**) The results are the
631 average of two experiments for all flaviviruses, except for CxFV (n=1) and KRV (n=3). Error
632 bars are the standard deviation between replicates. Positive-strand RNAs are shown in red,
633 negative-strand RNA in blue.

634

635 **Supplementary figure 7. Uncropped northern blot images**

636

637 **Supplementary table 1. List of viruses used in pan-flavivirus small RNA analysis**

638

639 **Supplementary table 2. List of oligonucleotides for northern blots, qPCR and cloning**

640

641 **Supplementary table 3. List of genome references used for 3'UTR analysis**

642

643 REFERENCES

644 1. Saiyasombat R, Bolling BG, Brault AC, Bartholomay LC, Blitvich BJ. 2011. Evidence of
645 Efficient Transovarial Transmission of Culex Flavivirus by Culex pipiens (Diptera:
646 Culicidae). Journal of Medical Entomology 48:1031–1038.

647 2. Frangeul L, Blanc H, Saleh M-C, Suzuki Y. 2020. Differential Small RNA Responses
648 against Co-Infecting Insect-Specific Viruses in Aedes albopictus Mosquitoes. Viruses
649 12:468.

650 3. Blitvich B, Firth A. 2015. Insect-Specific Flaviviruses: A Systematic Review of Their
651 Discovery, Host Range, Mode of Transmission, Superinfection Exclusion Potential
652 and Genomic Organization. Viruses 7:1927–1959.

653 4. Moureau G, Cook S, Lemey P, Nougairède A, Forrester NL, Khasnatinov M, Charrel
654 RN, Firth AE, Gould EA, de Lamballerie X. 2015. New Insights into Flavivirus
655 Evolution, Taxonomy and Biogeographic History, Extended by Analysis of Canonical
656 and Alternative Coding Sequences. PLoS One 10:e0117849--30.

657 5. Hall RA, Bielefeldt-Ohmann H, McLean BJ, O'Brien CA, Colmant AMG, Piyasena
658 TBH, Harrison JJ, Newton ND, Barnard RT, Prow NA, Deerain JM, Mah MGKY,
659 Hobson-Peters J. 2016. Commensal Viruses of Mosquitoes: Host Restriction,
660 Transmission, and Interaction with Arboviral Pathogens. Evolutionary Bioinformatics
661 12s2:EBO.S40740.

662 6. Pierson TC, Diamond MS. 2020. The continued threat of emerging flaviviruses.
663 Nature Microbiology 5:796–812.

664 7. Porier DL, Wilson SN, Auguste DI, Leber A, Coutermarsh-Ott S, Allen IC, Caswell CC,
665 Budnick JA, Bassaganya-Riera J, Hontecillas R, Weger-Lucarelli J, Weaver SC,

666 Auguste AJ. 2021. Enemy of My Enemy: A Novel Insect-Specific Flavivirus Offers a
667 Promising Platform for a Zika Virus Vaccine. Vaccines (Basel) 9:1142.
668 8. Hobson-Peters J, Harrison JJ, Watterson D, Hazlewood JE, Vet LJ, Newton ND,
669 Warrilow D, Colmant AMG, Taylor C, Huang B, Piyasena TBH, Chow WK, Setoh YX,
670 Tang B, Nakayama E, Yan K, Amarilla AA, Wheatley S, Moore PR, Finger M, Kurucz
671 N, Modhiran N, Young PR, Khromykh AA, Bielefeldt-Ohmann H, Suhrbier A, Hall RA.
672 2019. A recombinant platform for flavivirus vaccines and diagnostics using chimeras
673 of a new insect-specific virus. Science Translational Medicine 11.
674 9. Vet LJ, Setoh YX, Amarilla AA, Habarugira G, Suen WW, Newton ND, Harrison JJ,
675 Hobson-Peters J, Hall RA, Bielefeldt-Ohmann H. 2020. Protective Efficacy of a
676 Chimeric Insect-Specific Flavivirus Vaccine against West Nile Virus. Vaccines (Basel)
677 8:258.
678 10. Varjak M, Donald CL, Mottram TJ, Sreenu VB, Merits A, Maringer K, Schnettler E,
679 Kohl A. 2017. Characterization of the Zika virus induced small RNA response in Aedes
680 aegypti cells. PLOS Neglected Tropical Diseases 11:e0006010.
681 11. De Falco L, Silva NM, Santos NC, Huber RG, Martins IC. 2021. The Pseudo-Circular
682 Genomes of Flaviviruses: Structures, Mechanisms, and Functions of Circularization.
683 Cells 10:642.
684 12. Mazeaud C, Freppel W, Chatel-Chaix L. 2018. The Multiples Fates of the Flavivirus
685 RNA Genome During Pathogenesis. Frontiers in Genetics 9.
686 13. Fajardo T, Sanford TJ, Mears H V., Jasper A, Storrie S, Mansur DS, Sweeney TR.
687 2020. The flavivirus polymerase NS5 regulates translation of viral genomic RNA.
688 Nucleic Acids Research 48:5081–5093.

689 14. Trettin KD, Sinha NK, Eckert DM, Apple SE, Bass BL. 2017. Loquacious-PD facilitates
690 Drosophila Dicer-2 cleavage through interactions with the helicase domain and
691 dsRNA. Proceedings of the National Academy of Sciences 114:E7939–E7948.
692 15. Pijlman GP, Funk A, Kondratieva N, Leung J, Torres S, van der Aa L, Liu WJ,
693 Palmenberg AC, Shi PY, Hall RA, Khromykh AA. 2008. A Highly Structured,
694 Nuclease-Resistant, Noncoding RNA Produced by Flaviviruses Is Required for
695 Pathogenicity. Cell Host and Microbe 4:579–591.
696 16. MacFadden A, O'Donoghue Z, Silva PAGCGC, Chapman EG, Olsthoorn RC, Sterken
697 MG, Pijlman GP, Bredenbeek PJ, Kieft JS. 2018. Mechanism and structural diversity
698 of exoribonuclease-resistant RNA structures in flaviviral RNAs. Nature
699 Communications 9:119.
700 17. Slonchak A, Khromykh AA. 2018. Subgenomic flaviviral RNAs: What do we know
701 after the first decade of research. Antiviral Research
702 https://doi.org/10.1016/j.antiviral.2018.09.006.
703 18. Vicens Q, Kieft JS. 2021. Shared properties and singularities of exoribonuclease-
704 resistant RNAs in viruses. Computational and Structural Biotechnology Journal
705 19:4373–4380.
706 19. Kieft JS, Rabe JL, Chapman EG. 2015. New hypotheses derived from the structure of
707 a flaviviral Xrn1-resistant RNA: Conservation, folding, and host adaptation. RNA
708 Biology 12:1169–1177.
709 20. Ochsenreiter R, Hofacker I, Wolfinger M. 2019. Functional RNA Structures in the
710 3′UTR of Tick-Borne, Insect-Specific and No-Known-Vector Flaviviruses. Viruses
711 11:298.

712 21. Villordo SM, Filomatori C V, Sánchez-Vargas I, Blair CD, Gamarnik A V. 2015. Dengue
713 Virus RNA Structure Specialization Facilitates Host Adaptation. PLoS Pathogens
714 11:e1004604--22.
715 22. Filomatori C V., Carballeda JM, Villordo SM, Aguirre S, Pallarés HM, Maestre AM,
716 Sánchez-Vargas I, Blair CD, Fabri C, Morales MA, Fernandez-Sesma A, Gamarnik A V.
717 2017. Dengue virus genomic variation associated with mosquito adaptation defines
718 the pattern of viral non-coding RNAs and fitness in human cells. PLOS Pathogens
719 13:e1006265.
720 23. Funk A, Truong K, Nagasaki T, Torres S, Floden N, Balmori Melian E, Edmonds J,
721 Dong H, Shi PY, Khromykh AA. 2010. RNA structures required for production of
722 subgenomic flavivirus RNA. J Virol 84:11407–11417.
723 24. Pallarés HM, Costa Navarro GS, Villordo SM, Merwaiss F, de Borba L, Gonzalez
724 Lopez Ledesma MM, Ojeda DS, Henrion-Lacritick A, Morales MA, Fabri C, Saleh MC,
725 Gamarnik A V. 2020. Zika Virus Subgenomic Flavivirus RNA Generation Requires
726 Cooperativity between Duplicated RNA Structures That Are Essential for Productive
727 Infection in Human Cells. Journal of Virology 94.
728 25. Yeh S-C, Tan W-L, Chowdhury A, Chuo V, Kini RM, Pompon J, Garcia-Blanco MA.
729 2021. The anti-immune dengue subgenomic flaviviral RNA is found in vesicles in
730 mosquito saliva 2 and associated with increased infectivity. bioRxiv
731 https://doi.org/10.1101/2021.03.02.433543.
732 26. Moon SL, Anderson JR, Kumagai Y, Wilusz CJ, Akira S, Khromykh AA, Wilusz J. 2012.
733 A noncoding RNA produced by arthropod-borne flaviviruses inhibits the cellular
734 exoribonuclease XRN1 and alters host mRNA stability. RNA 18:2029–2040.

735 27. Slonchak A, Hugo LE, Freney ME, Hall-Mendelin S, Amarilla AA, Torres FJ, Setoh YX,
736 Peng NYG, Sng JDJ, Hall RA, van den Hurk AF, Devine GJ, Khromykh AA. 2020. Zika
737 virus noncoding RNA suppresses apoptosis and is required for virus transmission by
738 mosquitoes. Nature Communications 11:2205.
739 28. Hussain M, Torres S, Schnettler E, Funk A, Grundhoff A, Pijlman GP, Khromykh AA,
740 Asgari S. 2012. West Nile virus encodes a microRNA-like small RNA in the 3′
741 untranslated region which up-regulates GATA4 mRNA and facilitates virus
742 replication in mosquito cells. Nucleic Acids Research 40:2210–2223.
743 29. Pompon J, Manuel M, Ng GK, Wong B, Shan C, Manokaran G, Soto-Acosta R,
744 Bradrick SS, Ooi EE, Missé D, Shi PY, Garcia-Blanco MA. 2017. Dengue subgenomic
745 flaviviral RNA disrupts immunity in mosquito salivary glands to increase virus
746 transmission. PLoS Pathogens 13:e1006535--27.
747 30. Göertz GP, Fros JJ, Miesen P, Vogels CBF, van der Bent ML, Geertsema C, Koenraadt
748 CJM, van Rij RP, van Oers MM, Pijlman GP. 2016. Noncoding Subgenomic Flavivirus
749 RNA Is Processed by the Mosquito RNA Interference Machinery and Determines
750 West Nile Virus Transmission by Culex pipiens Mosquitoes. J Virol 90:10145–10159.
751 31. Moon SL, Dodd BJT, Brackney DE, Wilusz CJ, Ebel GD, Wilusz J. 2015. Flavivirus
752 sfRNA suppresses antiviral RNA interference in cultured cells and mosquitoes and
753 directly interacts with the RNAi machinery. Virology 485:322–329.
754 32. Phoo WW, Li Y, Zhang Z, Lee MY, Loh YR, Tan YB, Ng EY, Lescar J, Kang C, Luo D.
755 2016. Structure of the NS2B-NS3 protease from Zika virus after self-cleavage. Nature
756 Communications 1–8.
757 33. Schnettler E, Sterken MG, Leung JY, Metz SW, Geertsema C, Goldbach RW, Vlak JM,
758 Kohl A, Khromykh AA, Pijlman GP. 2012. Noncoding Flavivirus RNA Displays RNA

759 Interference Suppressor Activity in Insect and Mammalian Cells. Journal of Virology
760 86:13486–13500.
761 34. Cirimotich CM, Scott JC, Phillips AT, Geiss BJ, Olson KE. 2009. Suppression of RNA
762 interference increases alphavirus replication and virus-associated mortality in Aedes
763 aegypti mosquitoes. BMC Microbiology 9:49.
764 35. Bronkhorst AW, van Rij RP. 2014. The long and short of antiviral defense: small RNA-
765 based immunity in insects. Current Opinion in Virology 7:19–28.
766 36. Marques JT, Imler J-L. 2016. The diversity of insect antiviral immunity: insights from
767 viruses. Current Opinion in Microbiology 32:71–76.
768 37. Rosendo Machado S, van der Most T, Miesen P. 2021. Genetic determinants of
769 antiviral immunity in dipteran insects – Compiling the experimental evidence.
770 Developmental & Comparative Immunology 119:104010.
771 38. Elrefaey AME, Hollinghurst P, Reitmayer CM, Alphey L, Maringer K. 2021. Innate
772 Immune Antagonism of Mosquito-Borne Flaviviruses in Humans and Mosquitoes.
773 Viruses 13:2116.
774 39. Miesen P, Joosten J, van Rij RP. 2016. PIWIs Go Viral: Arbovirus-Derived piRNAs in
775 Vector Mosquitoes. PLOS Pathogens 12:e1006017.
776 40. Blair CD. 2019. Deducing the Role of Virus Genome-Derived PIWI-Associated RNAs
777 in the Mosquito–Arbovirus Arms Race. Frontiers in Genetics 10.
778 41. Miesen P, Ivens A, Buck AH, van Rij RP. 2016. Small RNA Profiling in Dengue Virus 2-
779 Infected Aedes Mosquito Cells Reveals Viral piRNAs and Novel Host miRNAs. PLOS
780 Neglected Tropical Diseases 10:e0004452.

781 42. Miesen P, Girardi E, van Rij RP. 2015. Distinct sets of PIWI proteins produce arbovirus
782 and transposon-derived piRNAs in Aedes aegypti mosquito cells. Nucleic Acids
783 Research 43:6545–6556.
784 43. Joosten J, Miesen P, Ta\c sköprü E, Pennings B, Jansen PWTC, Huynen MA,
785 Vermeulen M, van Rij RP. 2019. The Tudor protein Veneno assembles the ping-pong
786 amplification complex that produces viral piRNAs in Aedes mosquitoes. Nucleic
787 Acids Res 47:2546–2559.
788 44. Varjak M, Donald CL, Mottram TJ, Sreenu VB, Merits A, Maringer K, Schnettler E,
789 Kohl A. 2017. Characterization of the Zika virus induced small RNA response in Aedes
790 aegypti cells. PLoS Neglected Tropical Diseases 11:e0006010.
791 45. Tassetto M, Kunitomi M, Whitfield ZJ, Dolan PT, Sánchez-Vargas I, Garcia-Knight M,
792 Ribiero I, Chen T, Olson KE, Andino R. 2019. Control of RNA viruses in mosquito cells
793 through the acquisition of vDNA and endogenous viral elements. Elife 8.
794 46. Suzuki Y, Baidaliuk A, Miesen P, Frangeul L, Crist AB, Merkling SH, Fontaine A,
795 Lequime S, Moltini-Conclois I, Blanc H, van Rij RP, Lambrechts L, Saleh M-C. 2020.
796 Non-retroviral Endogenous Viral Element Limits Cognate Virus Replication in Aedes
797 aegypti Ovaries. Current Biology 30:3495-3506.e6.
798 47. Palatini U, Miesen P, Carballar-Lejarazu R, Ometto L, Rizzo E, Tu Z, van Rij RP,
799 Bonizzoni M. 2017. Comparative genomics shows that viral integrations are
800 abundant and express piRNAs in the arboviral vectors Aedes aegypti and Aedes
801 albopictus. BMC Genomics 18:512.
802 48. Crava CM, Varghese FS, Pischedda E, Halbach R, Palatini U, Marconcini M, Gasmi L,
803 Redmond S, Afrane Y, Ayala D, Paupy C, Carballar-Lejarazu R, Miesen P, Rij RP,
804 Bonizzoni M. 2021. Population genomics in the arboviral vector Aedes aegypti

805 reveals the genomic architecture and evolution of endogenous viral elements.
806 Molecular Ecology 30:1594–1611.
807 49. Crabtree MB, Sang RC, Stollar V, Dunster LM, Miller BR. 2003. Genetic and
808 phenotypic characterization of the newly described insect flavivirus, Kamiti River
809 virus. Archives of Virology 148:1095–1118.
810 50. MacFadden A, O'Donoghue Z, Silva PAGC, Chapman EG, Olsthoorn RC, Sterken MG,
811 Pijlman GP, Bredenbeek PJ, Kieft JS. 2018. Mechanism and structural diversity of
812 exoribonuclease-resistant RNA structures in flaviviral RNAs. Nature Communications
813 9:119.
814 51. Göertz GP, van Bree JWM, Hiralal A, Fernhout BM, Steffens C, Boeren S, Visser TM,
815 Vogels CBF, Abbo SR, Fros JJ, Koenraadt CJM, van Oers MM, Pijlman GP. 2019.
816 Subgenomic flavivirus RNA binds the mosquito DEAD/H-box helicase ME31B and
817 determines Zika virus transmission by Aedes aegypti. Proceedings of the National
818 Academy of Sciences 116:19136–19144.
819 52. Gritsun DJ, Jones IM, Gould EA, Gritsun TS. 2014. Molecular archaeology of
820 Flaviviridae untranslated regions: duplicated RNA structures in the replication
821 enhancer of flaviviruses and pestiviruses emerged via convergent evolution. PLoS
822 One 9:e92056.
823 53. Łabno A, Tomecki R, Dziembowski A. 2016. Cytoplasmic RNA decay pathways -
824 Enzymes and mechanisms. Biochimica et Biophysica Acta (BBA) - Molecular Cell
825 Research 1863:3125–3147.
826 54. Chen Y-S, Fan Y-H, Tien C-F, Yueh A, Chang R-Y. 2018. The conserved stem-loop II
827 structure at the 3' untranslated region of Japanese encephalitis virus genome is
828 required for the formation of subgenomic flaviviral RNA. PLoS One 13:e0201250--17.

829 55. Akiyama BM, Laurence HM, Massey AR, Costantino DA, Xie X, Yang Y, Shi PY, Nix
830 JC, Beckham JD, Kieft JS. 2016. Zika virus produces noncoding RNAs using a multi-
831 pseudoknot structure that confounds a cellular exonuclease. Science (1979)
832 https://doi.org/10.1126/science.aah3963.
833 56. Göertz GP, van Bree JWMM, Hiralal A, Fernhout BM, Steffens C, Boeren S, Visser
834 TM, Vogels CBFF, Abbo SR, Fros JJ, Koenraadt CJMM, van Oers MM, Pijlman GP.
835 2019. Subgenomic flavivirus RNA binds the mosquito DEAD/H-box helicase ME31B
836 and determines Zika virus transmission by Aedes aegypti. Proceedings of the
837 National Academy of Sciences 116:19136–19144.
838 57. Kopp A, Gillespie TR, Hobelsberger D, Estrada A, Harper JM, Miller RA, Eckerle I,
839 Müller MA, Podsiadlowski L, Leendertz FH, Drosten C, Junglen S. 2013. Provenance
840 and Geographic Spread of St. Louis Encephalitis Virus. mBio 4.
841 58. Göertz G, Miesen P, Overheul G, van Rij R, van Oers M, Pijlman G. 2019. Mosquito
842 Small RNA Responses to West Nile and Insect-Specific Virus Infections in Aedes and
843 Culex Mosquito Cells. Viruses 11:271.
844 59. Saiyasombat R, Bolling BG, Brault AC, Bartholomay LC, Blitvich BJ. 2011. Evidence of
845 Efficient Transovarial Transmission of Culex Flavivirus by Culex pipiens (Diptera:
846 Culicidae). Journal of Medical Entomology 48:1031–1038.
847 60. Rozen-Gagnon K, Gu M, Luna JM, Luo J-D, Yi S, Novack S, Jacobson E, Wang W, Paul
848 MR, Scheel TKH, Carroll T, Rice CM. 2021. Argonaute-CLIP delineates versatile,
849 functional RNAi networks in Aedes aegypti, a major vector of human viruses. Cell
850 Host & Microbe 29:834-848.e13.
851 61. Wang Y, Jin B, Liu P, Li J, Chen X, Gu J. 2018. piRNA Profiling of Dengue Virus Type
852 2-Infected Asian Tiger Mosquito and Midgut Tissues. Viruses 10:213.

853 62. Bolling B, Weaver S, Tesh R, Vasilakis N. 2015. Insect-Specific Virus Discovery:
854 Significance for the Arbovirus Community. Viruses 7:4911–4928.
855 63. Qiu Y, Xu Y-P, Wang M, Miao M, Zhou H, Xu J, Kong J, Zheng D, Li R-T, Zhang R-R,
856 Guo Y, Li X-F, Cui J, Qin C-F, Zhou X. 2020. Flavivirus induces and antagonizes
857 antiviral RNA interference in both mammals and mosquitoes. Science Advances 6.
858 64. Blair C, Olson K. 2015. The Role of RNA Interference (RNAi) in Arbovirus-Vector
859 Interactions. Viruses 7:820–843.
860 65. Lee W-S, Webster JA, Madzokere ET, Stephenson EB, Herrero LJ. 2019. Mosquito
861 antiviral defense mechanisms: a delicate balance between innate immunity and
862 persistent viral infection. Parasites & Vectors 12:165.
863 66. Colmant AMG, Hobson-Peters J, Bielefeldt-Ohmann H, van den Hurk AF, Hall-
864 Mendelin S, Chow WK, Johansen CA, Fros J, Simmonds P, Watterson D, Cazier C,
865 Etebari K, Asgari S, Schulz BL, Beebe N, Vet LJ, Piyasena TBH, Nguyen H-D, Barnard
866 RT, Hall RA. 2017. A New Clade of Insect-Specific Flaviviruses from Australian
867 Anopheles Mosquitoes Displays Species-Specific Host Restriction. mSphere 2.
868 67. Rückert C, Prasad AN, Garcia-Luna SM, Robison A, Grubaugh ND, Weger-Lucarelli J,
869 Ebel GD. 2019. Small RNA responses of Culex mosquitoes and cell lines during acute
870 and persistent virus infection. Insect Biochemistry and Molecular Biology 109:13–23.
871 68. Fan Y-H, Nadar M, Chen C-C, Weng C-C, Lin Y-T, Chang R-Y. 2011. Small Noncoding
872 RNA Modulates Japanese Encephalitis Virus Replication and Translation in trans.
873 Virol J 8:492.
874 69. Alvarez DE, Lodeiro MF, Ludueña SJ, Pietrasanta LI, Gamarnik A V. 2005. Long-
875 Range RNA-RNA Interactions Circularize the Dengue Virus Genome. Journal of
876 Virology 79:6631–6643.

877 70. Roby J, Pijlman G, Wilusz J, Khromykh A. 2014. Noncoding Subgenomic Flavivirus
878 RNA: Multiple Functions in West Nile Virus Pathogenesis and Modulation of Host
879 Responses. Viruses 6:404–427.
880 71. Firth AE, Blitvich BJ, Wills NM, Miller CL, Atkins JF. 2010. Evidence for ribosomal
881 frameshifting and a novel overlapping gene in the genomes of insect-specific
882 flaviviruses. Virology 399:153–166.
883 72. Pasternak AO, Spaan WJM, Snijder EJ. 2006. Nidovirus transcription: how to make
884 sense…? Journal of General Virology 87:1403–1421.
885 73. Sztuba-Solińska J, Stollar V, Bujarski JJ. 2011. Subgenomic messenger RNAs:
886 Mastering regulation of (+)-strand RNA virus life cycle. Virology 412:245–255.
887 74. Slonchak A, Khromykh AA. 2018. Subgenomic flaviviral RNAs: What do we know
888 after the first decade of research. Antiviral Research 159:13–25.
889 75. Jiwan SD, White KA. 2011. Subgenomic mRNA transcription in Tombusviridae. RNA
890 Biology 8:287–294.
891 76. Hertz JM, Huang H V. 1992. Utilization of heterologous alphavirus junction
892 sequences as promoters by Sindbis virus. Journal of Virology 66:857–864.
893 77. Halbach R, Junglen S, van Rij RP. 2017. Mosquito-specific and mosquito-borne
894 viruses: evolution, infection, and host defense. Current Opinion in Insect Science
895 22:16–27.
896 78. Lutomiah JJL, Mwandawiro C, Magambo J, Sang RC. 2007. Infection and Vertical
897 Transmission of Kamiti River Virus in Laboratory Bred Aedes aegypti Mosquitoes.
898 Journal of Insect Science 7:1–7.

899 79. Clarke DK, Duarte EA, Elena SF, Moya A, Domingo E, Holland J. 1994. The red queen
900 reigns in the kingdom of RNA viruses. Proceedings of the National Academy of
901 Sciences 91:4821–4824.
902 80. Junglen S, Kopp A, Kurth A, Pauli G, Ellerbrok H, Leendertz FH. 2009. A New
903 Flavivirus and a New Vector: Characterization of a Novel Flavivirus Isolated from
904 Uranotaenia Mosquitoes from a Tropical Rain Forest. Journal of Virology 83:4462–
905 4468.
906 81. Bassett AR, Tibbit C, Ponting CP, Liu J-L. 2014. Mutagenesis and homologous
907 recombination in Drosophila cell lines using CRISPR/Cas9. Biology Open 3:42–49.
908 82. Afgan E, Baker D, Batut B, van den Beek M, Bouvier D, Čech M, Chilton J, Clements
909 D, Coraor N, Grüning BA, Guerler A, Hillman-Jackson J, Hiltemann S, Jalili V, Rasche
910 H, Soranzo N, Goecks J, Taylor J, Nekrutenko A, Blankenberg D. 2018. The Galaxy
911 platform for accessible, reproducible and collaborative biomedical analyses: 2018
912 update. Nucleic Acids Research 46:W537–W544.
913 83. de Borba L, Villordo SM, Marsico FL, Carballeda JM, Filomatori C V., Gebhard LG,
914 Pallarés HM, Lequime S, Lambrechts L, Sánchez Vargas I, Blair CD, Gamarnik A V.
915 2019. RNA Structure Duplication in the Dengue Virus 3′ UTR: Redundancy or Host
916 Specificity? mBio 10.
917 84. Lorenz R, Bernhart SH, Höner zu Siederdissen C, Tafer H, Flamm C, Stadler PF,
918 Hofacker IL. 2011. ViennaRNA Package 2.0. Algorithms for Molecular Biology 6:26.
919 85. Will S, Reiche K, Hofacker IL, Stadler PF, Backofen R. 2007. Inferring Noncoding RNA
920 Families and Classes by Means of Genome-Scale Structure-Based Clustering. PLoS
921 Computational Biology 3:e65.

922 86. Bernhart SH, Hofacker IL, Will S, Gruber AR, Stadler PF. 2008. RNAalifold: improved
923 consensus structure prediction for RNA alignments. BMC Bioinformatics 9:474.
924 87. Darty K, Denise A, Ponty Y. 2009. VARNA: Interactive drawing and editing of the
925 RNA secondary structure. Bioinformatics
926 https://doi.org/10.1093/bioinformatics/btp250.
927 88. Katoh K, Standley DM. 2013. MAFFT Multiple Sequence Alignment Software Version
928 7: Improvements in Performance and Usability. Molecular Biology and Evolution
929 30:772–780.
930 89. Criscuolo A, Gribaldo S. 2010. BMGE (Block Mapping and Gathering with Entropy): a
931 new software for selection of phylogenetic informative regions from multiple
932 sequence alignments. BMC Evolutionary Biology 10:210.
933 90. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New
934 Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing
935 the Performance of PhyML 3.0. Systematic Biology 59:307–321.
936 91. Lemoine F, Correia D, Lefort V, Doppelt-Azeroual O, Mareuil F, Cohen-Boulakia S,
937 Gascuel O. 2019. NGPhylogeny.fr: new generation phylogenetic services for non-
938 specialists. Nucleic Acids Research 47:W260–W265.
939 92. Letunic I, Bork P. 2021. Interactive Tree Of Life (iTOL) v5: an online tool for
940 phylogenetic tree display and annotation. Nucleic Acids Research 49:W293–W296.
941 93. Brown NP, Leroy C, Sander C. 1998. MView: a web-compatible database search or
942 multiple alignment viewer. Bioinformatics 14:380–381.
943 94. Nicholson J, Ritchie SA, Van Den Hurk AF. 2014. Aedes albopictus (Diptera:
944 Culicidae) as a Potential Vector of Endemic and Exotic Arboviruses in Australia.
945 Journal of Medical Entomology 51:661–669.

946

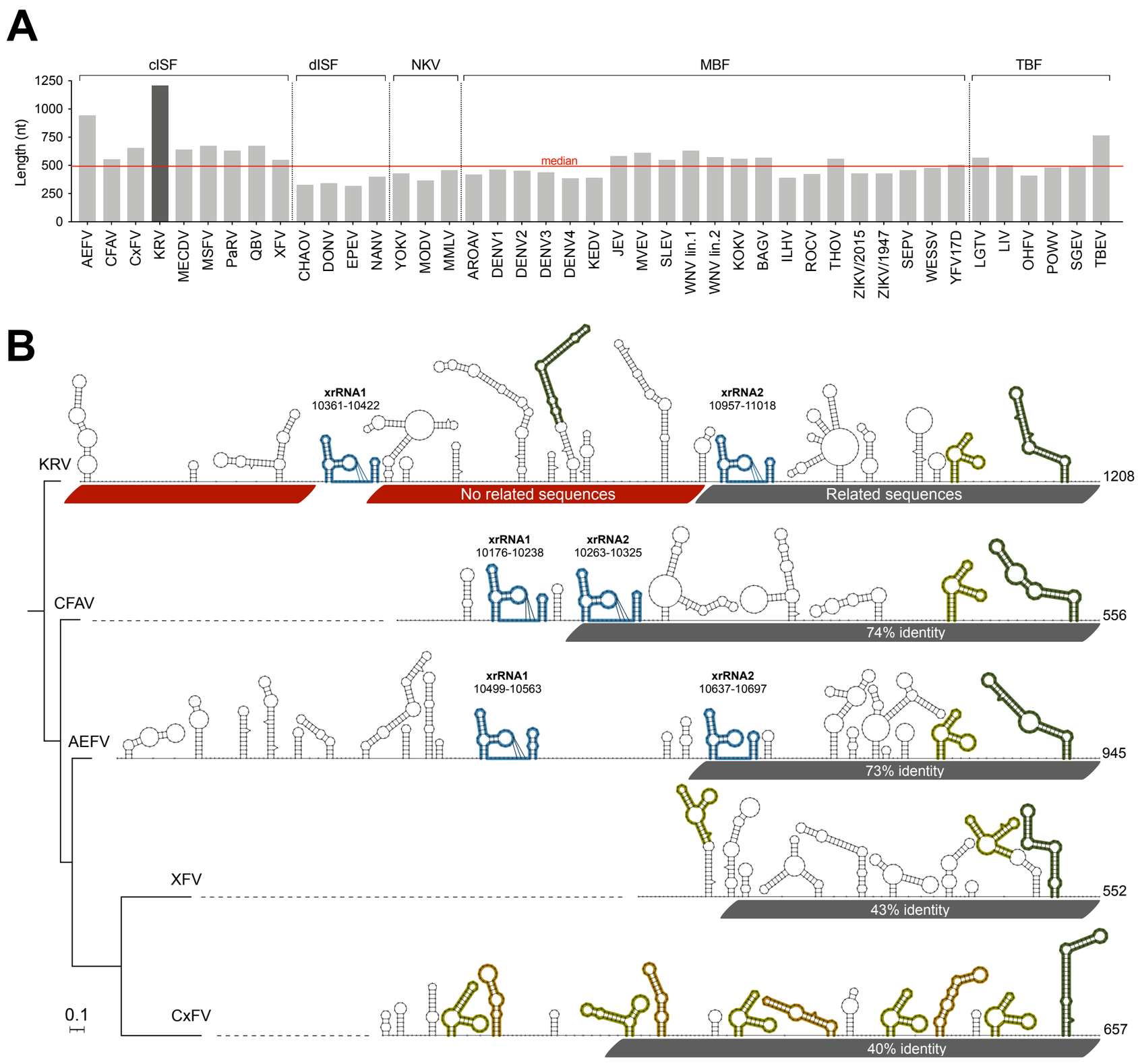
A
cISF
dISF
NKV
MBF
TBF
Length (nt)
1250
1000
750
500
250
0
median
AEFV
CFAV
CxFV
KRV
MECDV
MSFV
PaRV
QBV
XFV
CHAOV
DONV
EPEV
NANV
YOKV
MODV
MMLV
AROAV
DENV1
DENV2
DENV3
DENV4
KEDV
JEV
MVEV
SLEV
WNV lin.1
WNV lin.2
KOKV
BAGV
ILHV
ROCV
THOV
ZIKV/2015
ZIKV/1947
SEPV
WESSV
YFV17D
LGTV
LIV
OHFV
POWV
SGEV
TBEV
B
KRV
xrRNA1
10361-10422
xrRNA2
10957-11018
No related sequences
Related sequences
1208
CFAV
xrRNA1
10176-10238
xrRNA2
10263-10325
74% identity
556
AEFV
xrRNA1
10499-10563
xrRNA2
10637-10697
73% identity
945
XFV
43% identity
552
CxFV
40% identity
657
0.1

Fig. 1

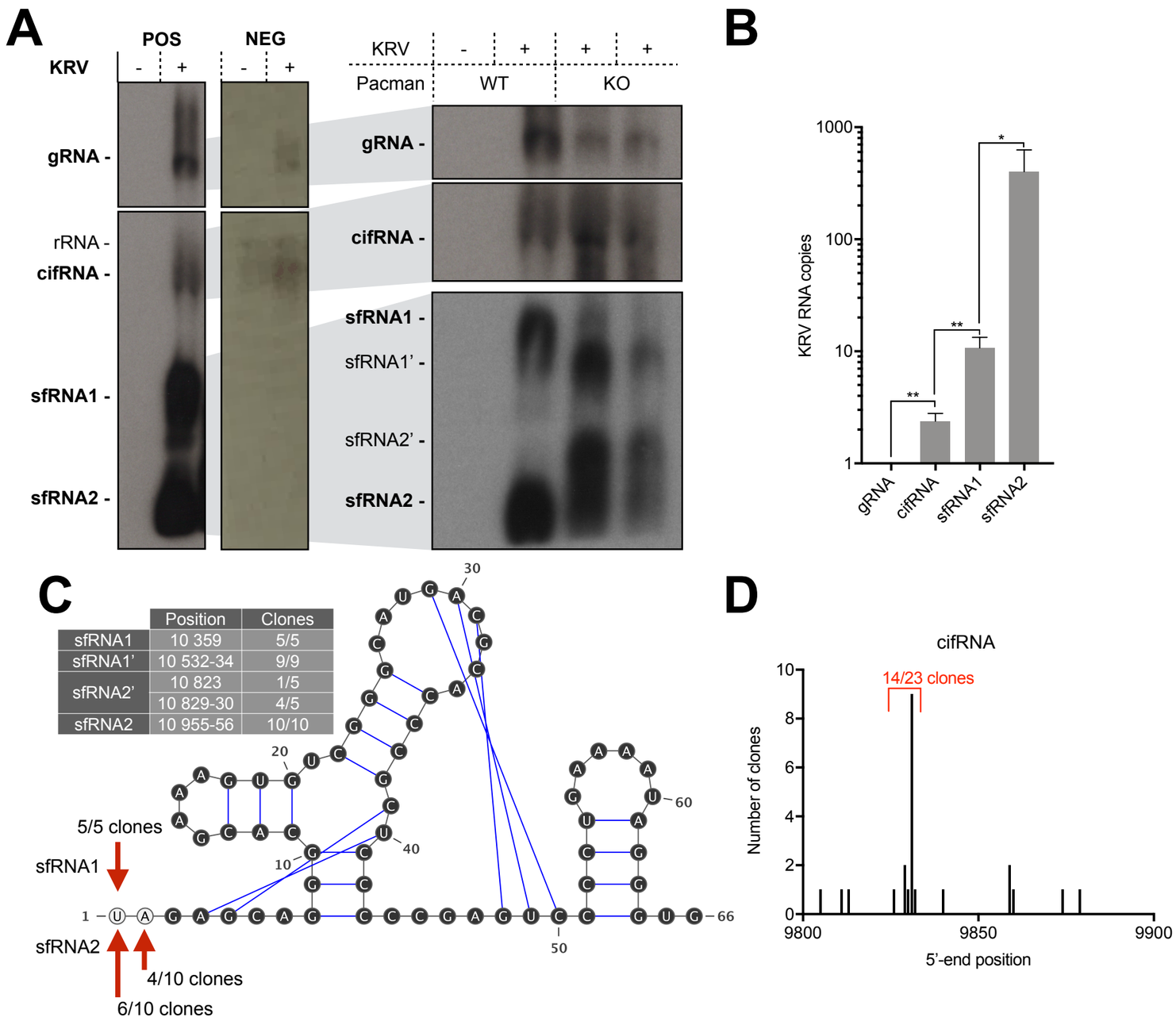
A
POS
NEG
KRV
-
+
gRNA -
rRNA -
cifRNA -
sfRNA1 -
sfRNA2 -
Pacman
WT
KO
sfRNA1'
sfRNA2'
B
KRV RNA copies
1000
100
10
1
gRNA
cifRNA
sfRNA1
sfRNA2
C
Position
Clones
sfRNA1
10 359
5/5
sfRNA1'
10 532-34
9/9
sfRNA2'
10 823
1/5
10 829-30
4/5
sfRNA2
10 955-56
10/10
5/5 clones
sfRNA1
sfRNA2
4/10 clones
6/10 clones
D
cifRNA
14/23 clones
Number of clones
5'-end position
9800
9850
9900

Fig. 2

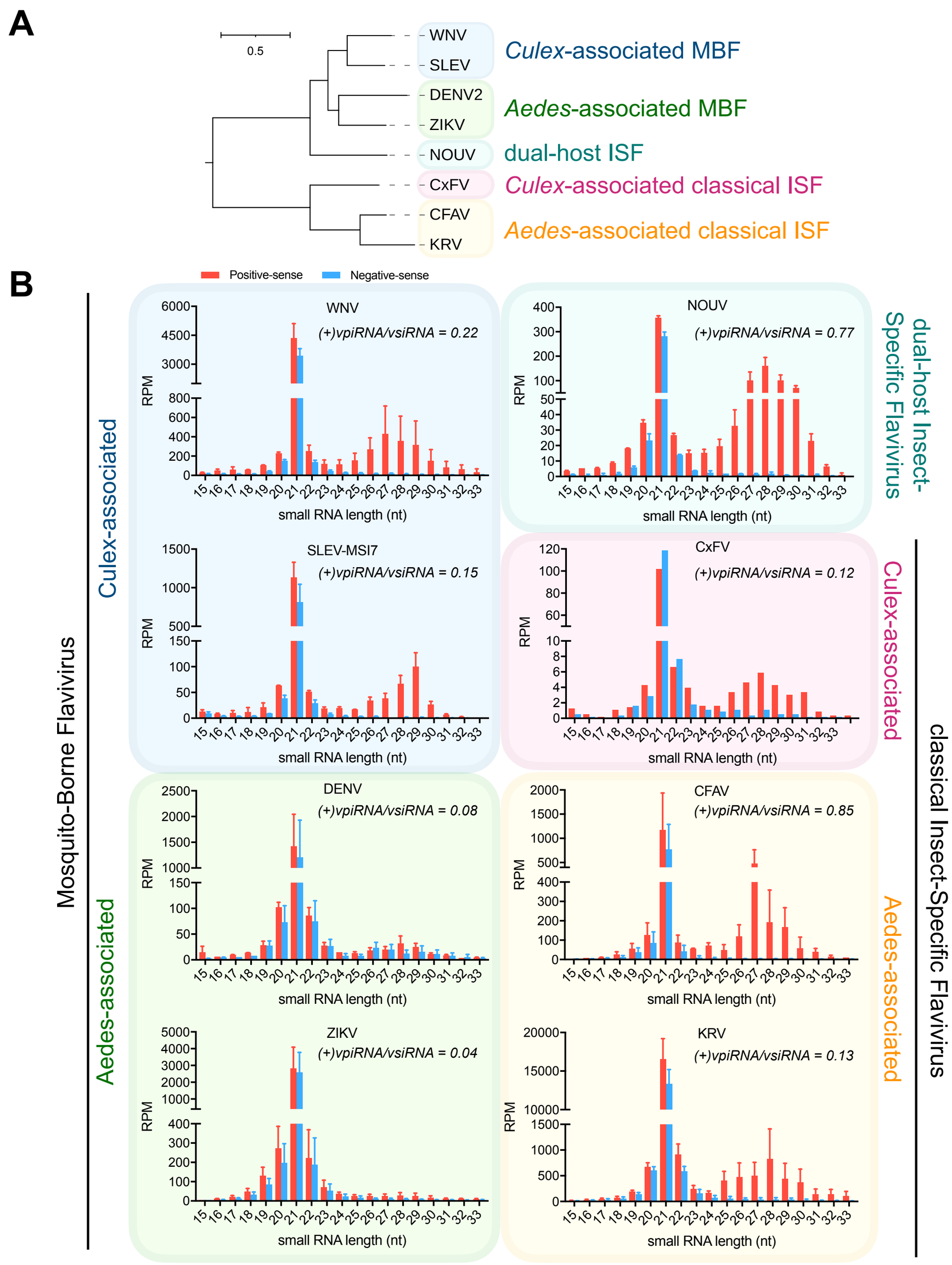
A
0.5
WNV
SLEV
DENV2
ZIKV
NOUV
CxFV
CFAV
KRV
Culex-associated MBF
Aedes-associated MBF
dual-host ISF
Culex-associated classical ISF
Aedes-associated classical ISF
B
Positive-sense
Negative-sense
WNV
(+)vpiRNA/vsiRNA = 0.22
NOUV
(+)vpiRNA/vsiRNA = 0.77
SLEV-MSI7
(+)vpiRNA/vsiRNA = 0.15
CxFV
(+)vpiRNA/vsiRNA = 0.12
DENV
(+)vpiRNA/vsiRNA = 0.08
CFAV
(+)vpiRNA/vsiRNA = 0.85
ZIKV
(+)vpiRNA/vsiRNA = 0.04
KRV
(+)vpiRNA/vsiRNA = 0.13
RPM
small RNA length (nt)
Mosquito-Borne Flavivirus
Culex-associated
Aedes-associated
dual-host Insect-Specific Flavivirus
Culex-associated
Aedes-associated
classical Insect-Specific Flavivirus

Fig. 3

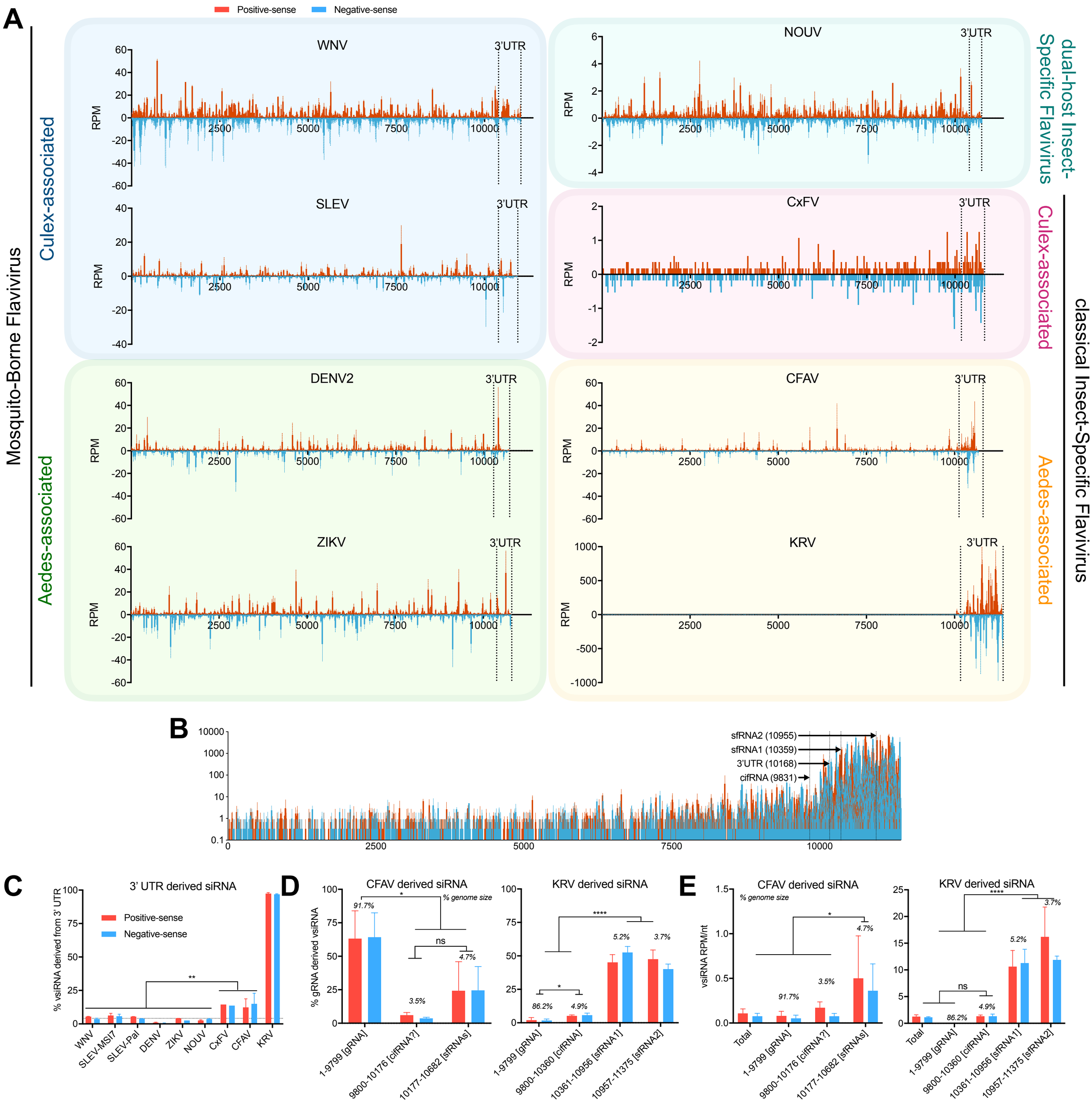
A
Positive-sense
Negative-sense
Mosquito-Borne Flavivirus
Culex-associated
Aedes-associated
WNV
SLEV
DENV2
ZIKV
NOUV
CxFV
CFAV
KRV
3'UTR
RPM
Dual-host Insect-Specific Flavivirus
Culex-associated
Aedes-associated
classical Insect-Specific Flavivirus
B
sfRNA2 (10955)
sfRNA1 (10359)
3'UTR (10168)
cifRNA (9831)
C
3' UTR derived siRNA
% vsiRNA derived from 3' UTR
Positive-sense
Negative-sense
WNV
SLEV-MSI7
SLEV-Pal
DENV
ZIKV
NOUV
CxFV
CFAV
KRV
D
CFAV derived siRNA
KRV derived siRNA
% gRNA derived vsiRNA
% genome size
ns
1-9799 [gRNA]
9800-10176 [cifRNA?]
10177-10682 [sfRNAs]
9800-10360 [cifRNA]
10361-10956 [sfRNA1]
10957-11375 [sfRNA2]
E
CFAV derived siRNA
KRV derived siRNA
vsiRNA RPM/nt
Total

Fig. 4

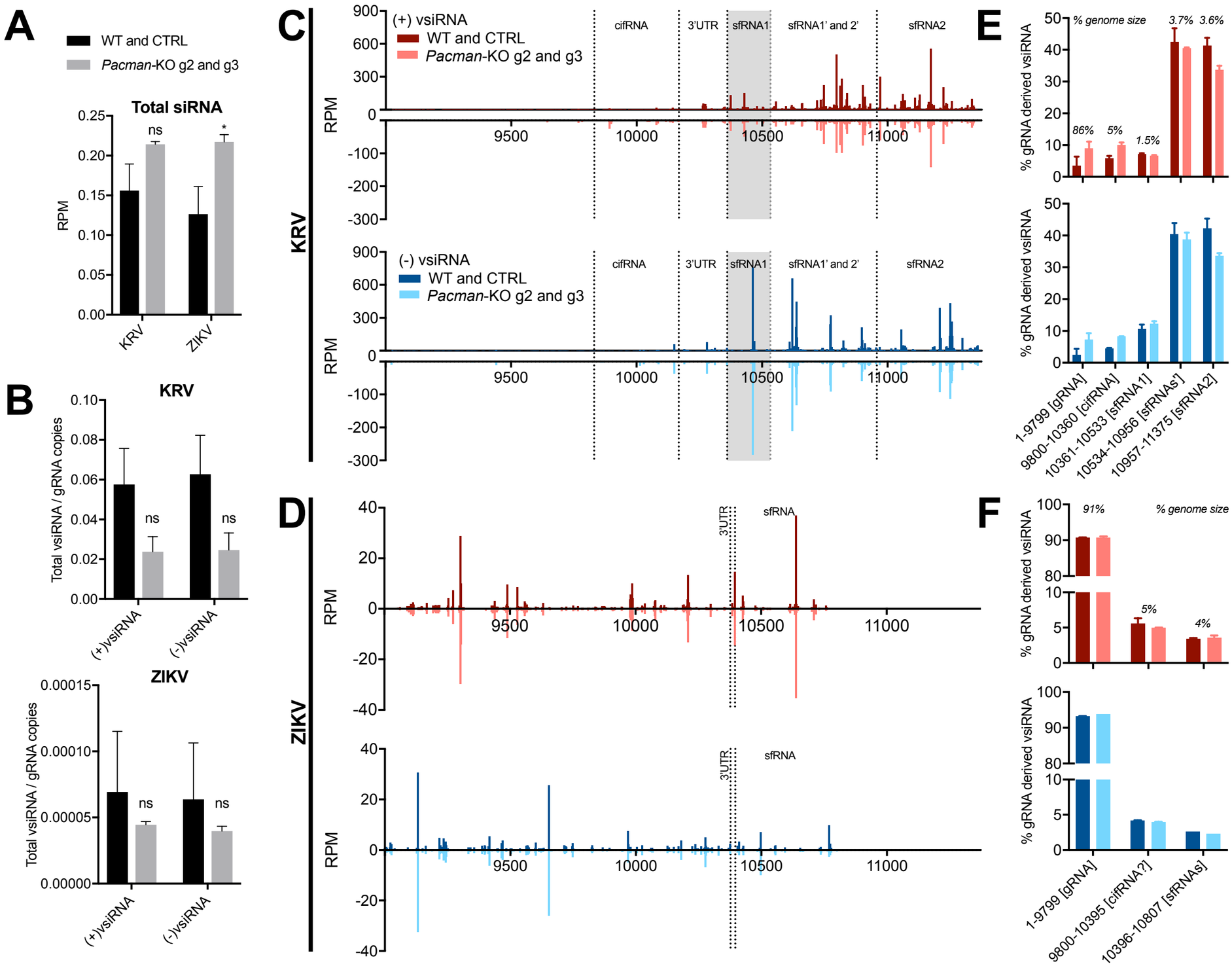
A
WT and CTRL
Pacman-KO g2 and g3
Total siRNA
RPM
ns
*
KRV
ZIKV
B
KRV
Total vsiRNA / gRNA copies
ns
ns
(+)vsiRNA
(-)vsiRNA
ZIKV
Total vsiRNA / gRNA copies
ns
ns
(+)vsiRNA
(-)vsiRNA
C
KRV
(+) vsiRNA
WT and CTRL
Pacman-KO g2 and g3
cifRNA
3'UTR
sfRNA1
sfRNA1' and 2'
sfRNA2
RPM
(-) vsiRNA
WT and CTRL
Pacman-KO g2 and g3
D
ZIKV
3'UTR
sfRNA
RPM
E
% genome size
3.7%
3.6%
86%
5%
1.5%
% gRNA derived vsiRNA
1-9799 [gRNA]
9800-10360 [cifRNA]
10361-10533 [sfRNA1]
10534-10956 [sfRNAs']
10957-11375 [sfRNA2]
F
91%
% genome size
5%
4%
% gRNA derived vsiRNA
1-9799 [gRNA]
9800-10395 [cifRNA?]
10396-10807 [sfRNAs]

Fig. 5

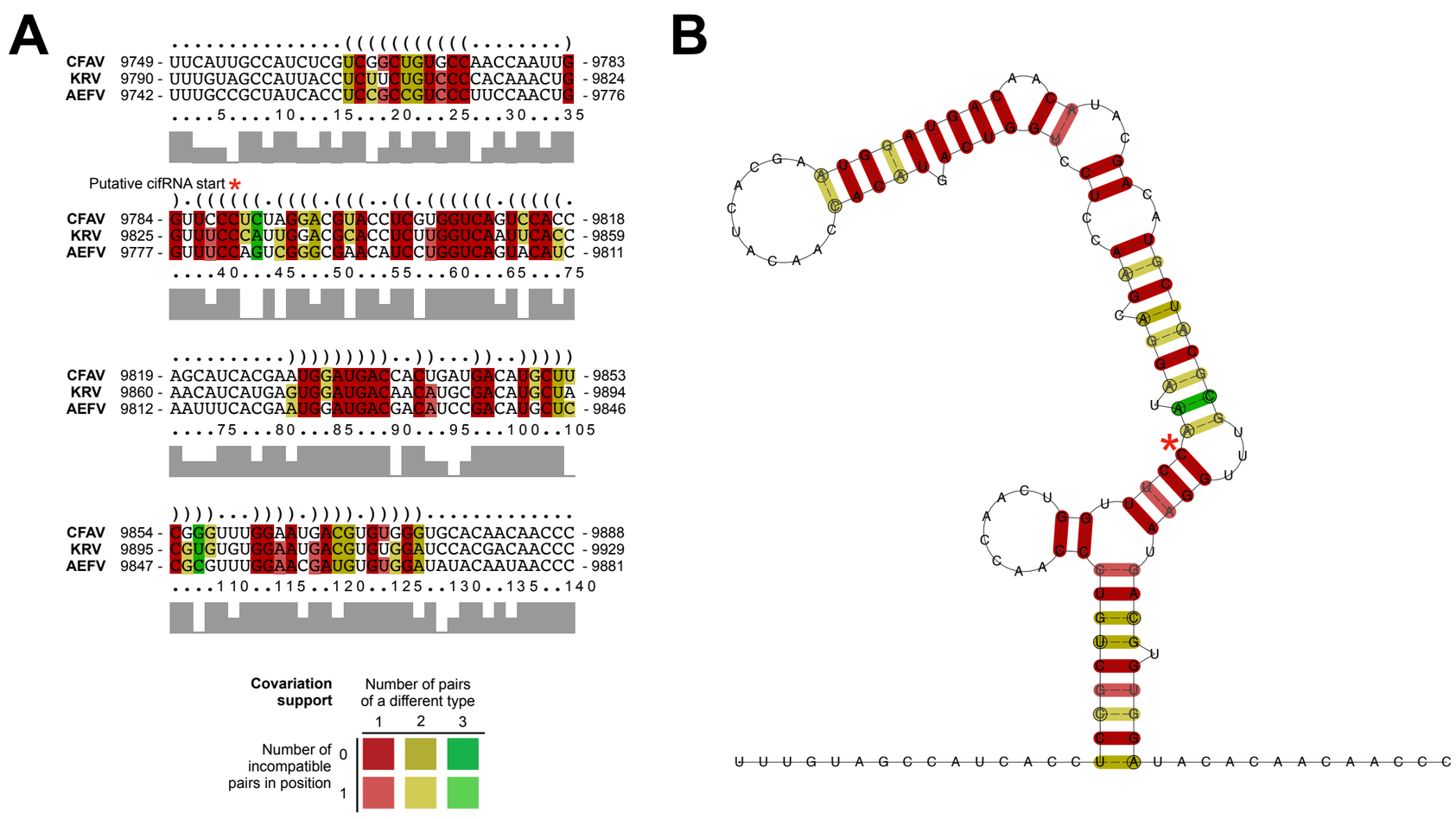
A
B
CFAV
KRV
AEFV
Putative cifRNA start
Covariation support
Number of pairs of a different type
Number of incompatible pairs in position

Fig. S1

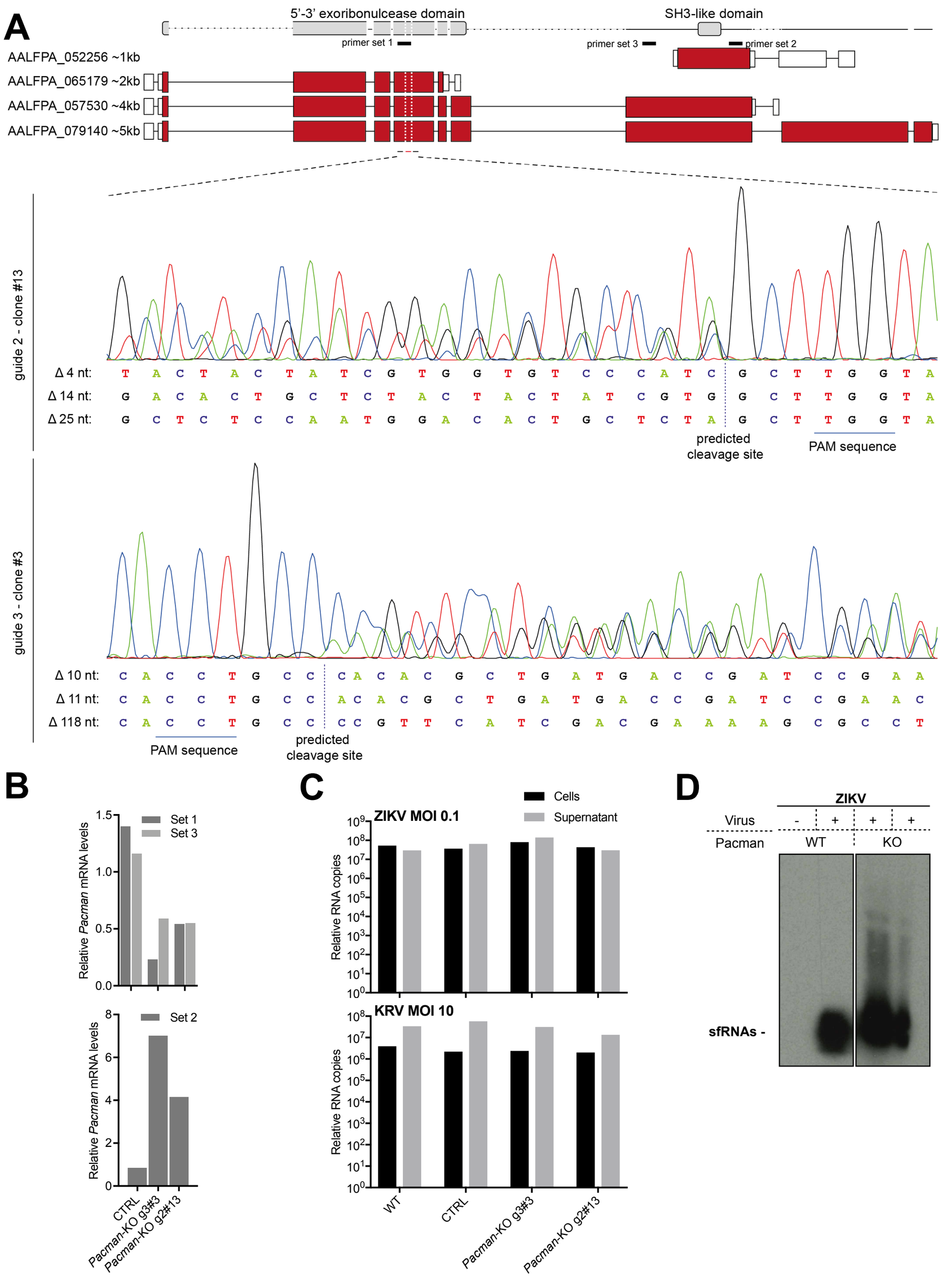
A
5'-3' exoribonulcease domain
SH3-like domain
primer set 1
primer set 3
primer set 2
AALFPA_052256 ~1kb
AALFPA_065179 ~2kb
AALFPA_057530 ~4kb
AALFPA_079140 ~5kb
guide 2 - clone #13
Δ 4 nt: T A C T A C T A T C G T G G T G T C C C A T C G C T T G G T A
Δ 14 nt: G A C A C T G C T C T A C T A C T A T C G T G G C T T G G T A
Δ 25 nt: G C T C T C C A A T G G A C A C T G C T C T A G C T T G G T A
predicted cleavage site
PAM sequence
guide 3 - clone #3
Δ 10 nt: C A C C T G C C C A C A C G C T G A T G A C C G A T C C G A A
Δ 11 nt: C A C C T G C C A C A C G C T G A T G A C C G A T C C G A A C
Δ 118 nt: C A C C T G C C C C G T T C A T C G A C G A A A A G C G C C T
PAM sequence
predicted cleavage site
B
Relative Pacman mRNA levels
Set 1
Set 3
Set 2
CTRL
Pacman-KO g3#3
Pacman-KO g2#13
C
Cells
Supernatant
ZIKV MOI 0.1
KRV MOI 10
Relative RNA copies
WT
CTRL
Pacman-KO g3#3
Pacman-KO g2#13
D
ZIKV
Virus
Pacman
WT
KO
sfRNAs -

Fig. S2

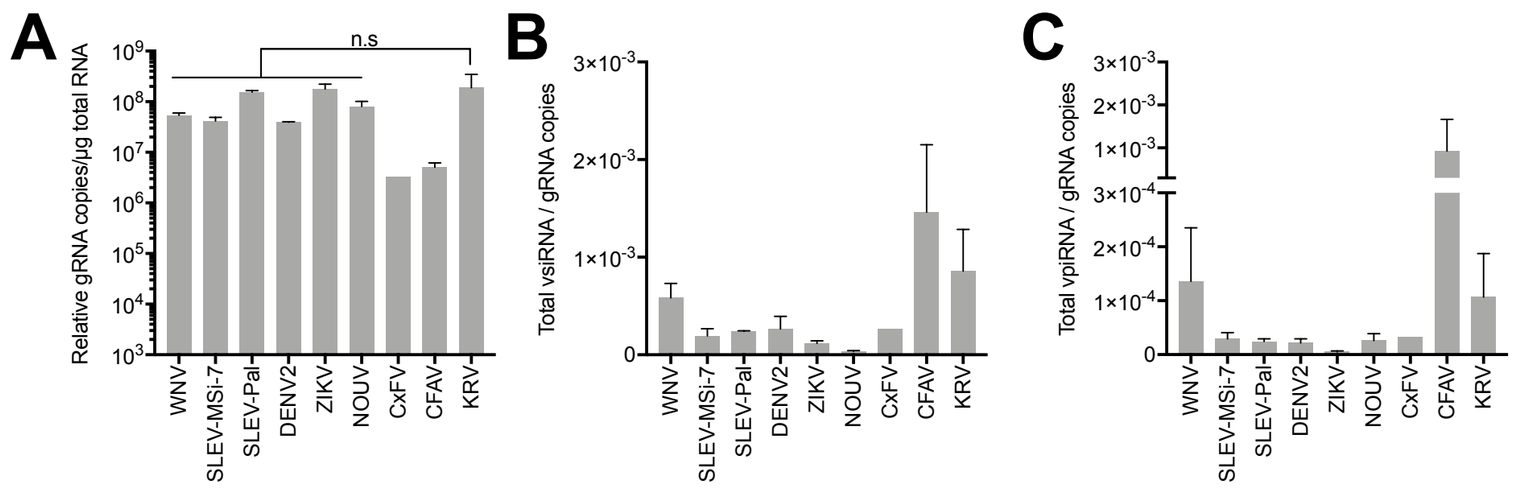
A
n.s
Relative gRNA copies/µg total RNA
10^9
10^8
10^7
10^6
10^5
10^4
10^3
WNV
SLEV-MSi-7
SLEV-Pal
DENV2
ZIKV
NOUV
CxFV
CFAV
KRV
B
Total vsiRNA / gRNA copies
3×10^{-3}
2×10^{-3}
1×10^{-3}
0
WNV
SLEV-MSi-7
SLEV-Pal
DENV2
ZIKV
NOUV
CxFV
CFAV
KRV
C
Total vpiRNA / gRNA copies
3×10^{-3}
2×10^{-3}
1×10^{-3}
3×10^{-4}
2×10^{-4}
1×10^{-4}
0
WNV
SLEV-MSi-7
SLEV-Pal
DENV2
ZIKV
NOUV
CxFV
CFAV
KRV

Fig. S3

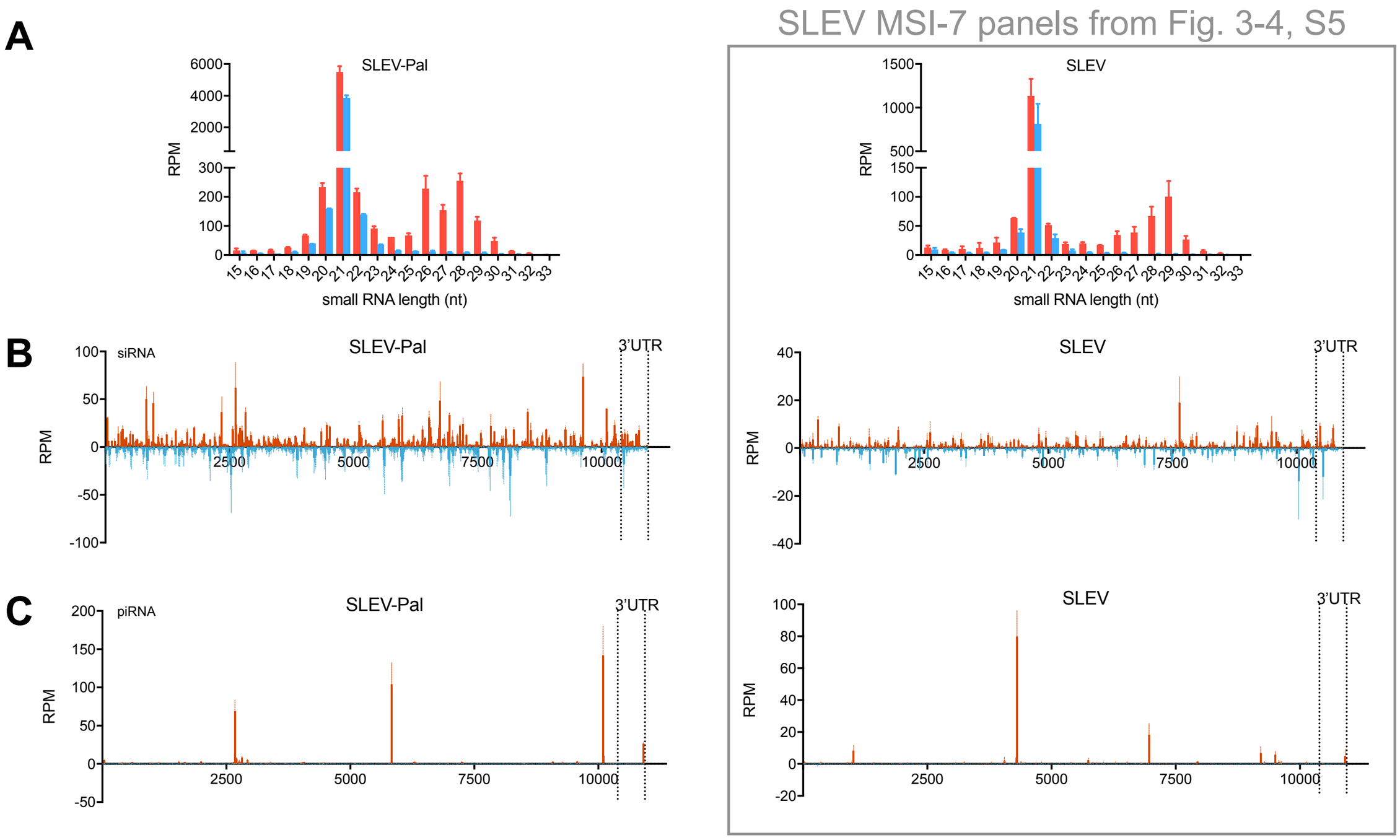
SLEV MSI-7 panels from Fig. 3-4, S5
A
SLEV-Pal
SLEV
RPM
small RNA length (nt)
B
siRNA
SLEV-Pal
SLEV
3'UTR
C
piRNA
SLEV-Pal
SLEV
3'UTR

Fig. S4

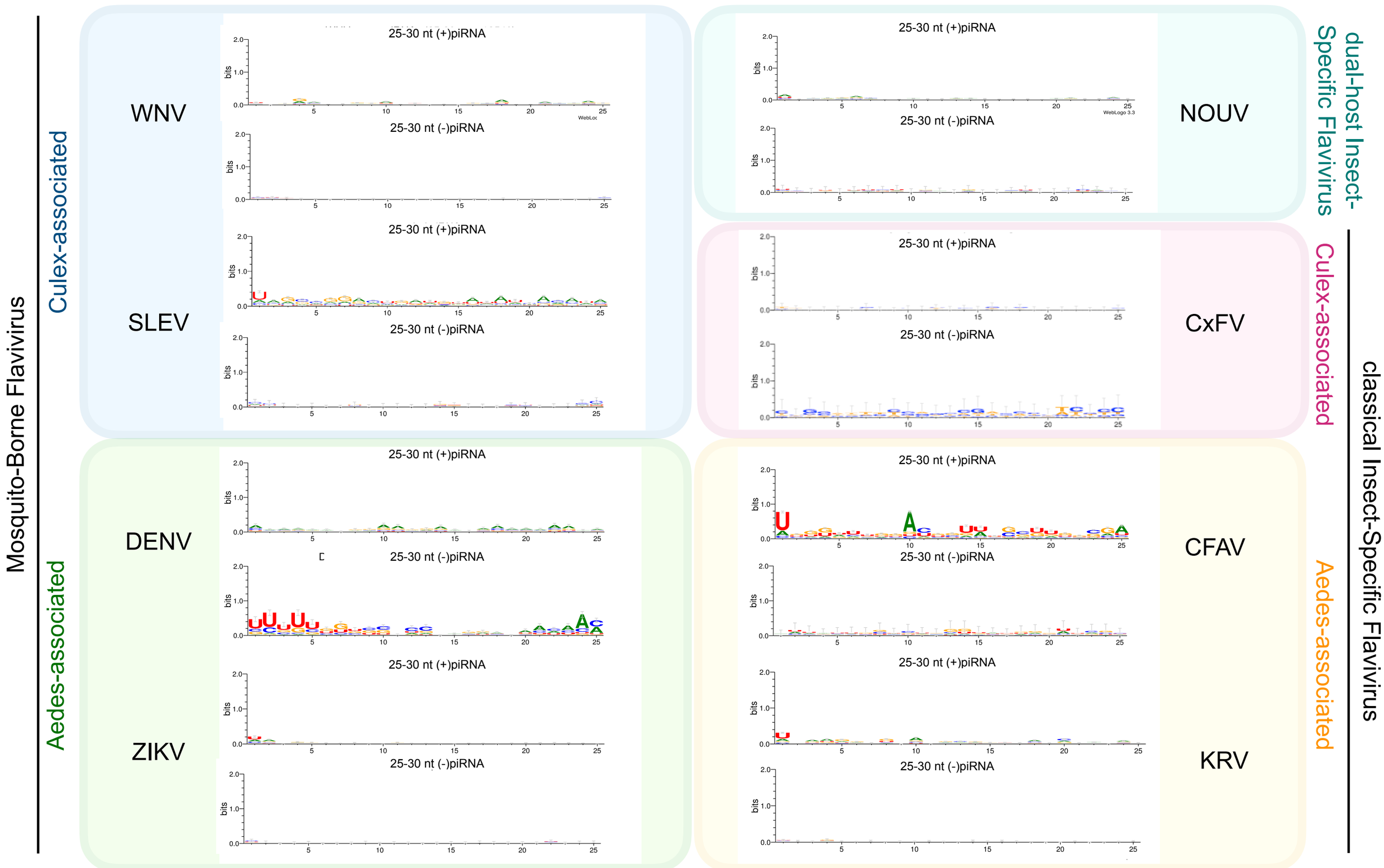
Mosquito-Borne Flavivirus
Culex-associated
Aedes-associated
WNV
SLEV
DENV
ZIKV
25-30 nt (+)piRNA
25-30 nt (-)piRNA
bits
NOUV
CxFV
CFAV
KRV
dual-host Insect-Specific Flavivirus
Culex-associated
Aedes-associated
classical Insect-Specific Flavivirus

Fig. S5

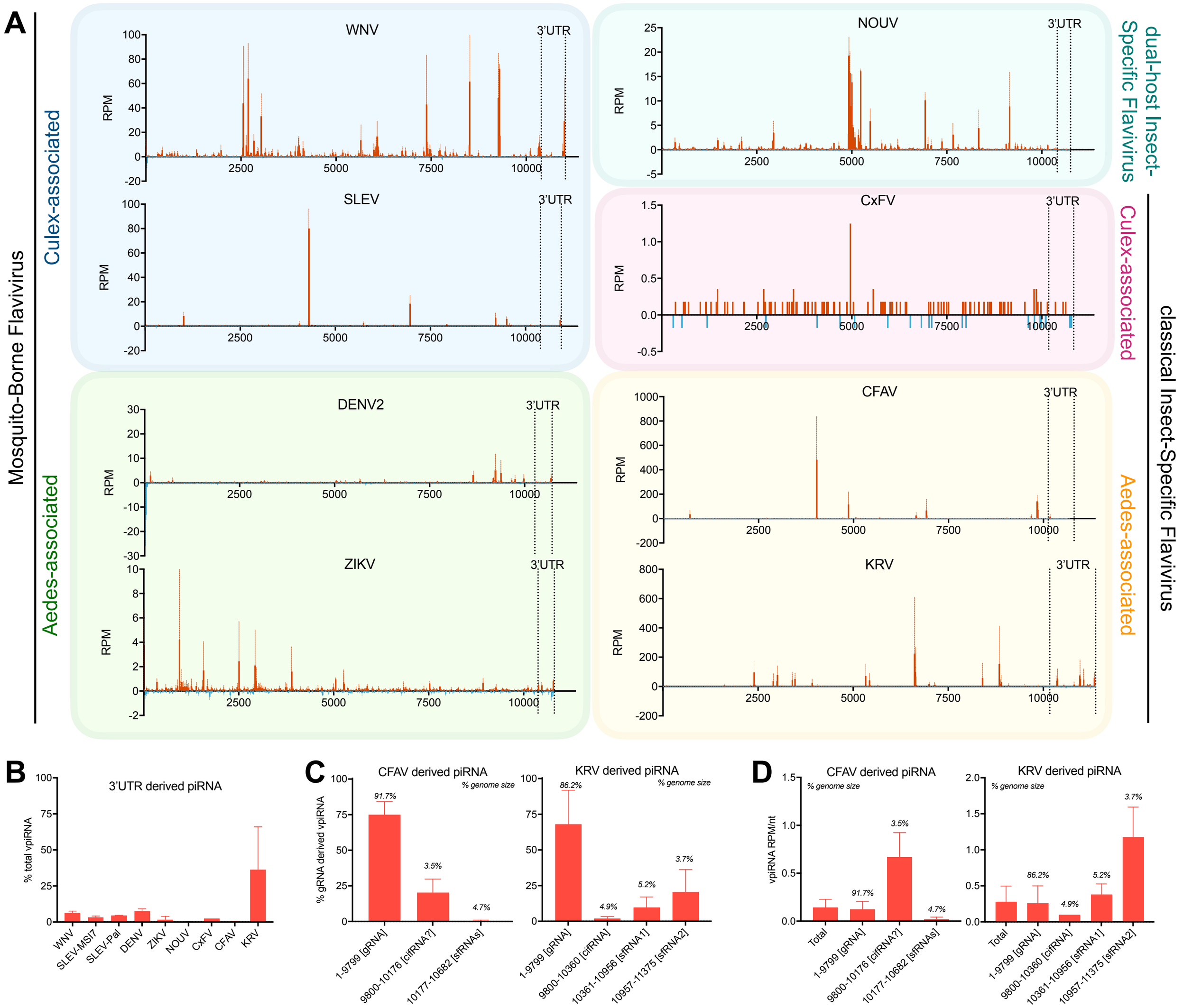
A
Mosquito-Borne Flavivirus
Culex-associated
Aedes-associated
WNV
SLEV
DENV2
ZIKV
NOUV
CxFV
CFAV
KRV
3'UTR
RPM
dual-host Insect-Specific Flavivirus
Culex-associated
Aedes-associated
classical Insect-Specific Flavivirus
B
3'UTR derived piRNA
% total vpiRNA
WNV
SLEV-MSI7
SLEV-Pal
DENV
ZIKV
NOUV
CxFV
CFAV
KRV
C
CFAV derived piRNA
KRV derived piRNA
% genome size
% gRNA derived vpiRNA
91.7%
3.5%
4.7%
86.2%
4.9%
5.2%
3.7%
1-9799 [gRNA]
9800-10176 [cifRNA?]
10177-10682 [sfRNAs]
9800-10360 [cifRNA]
10361-10956 [sfRNA1]
10957-11375 [sfRNA2]
D
vpiRNA RPM/nt
Total

Fig. S6

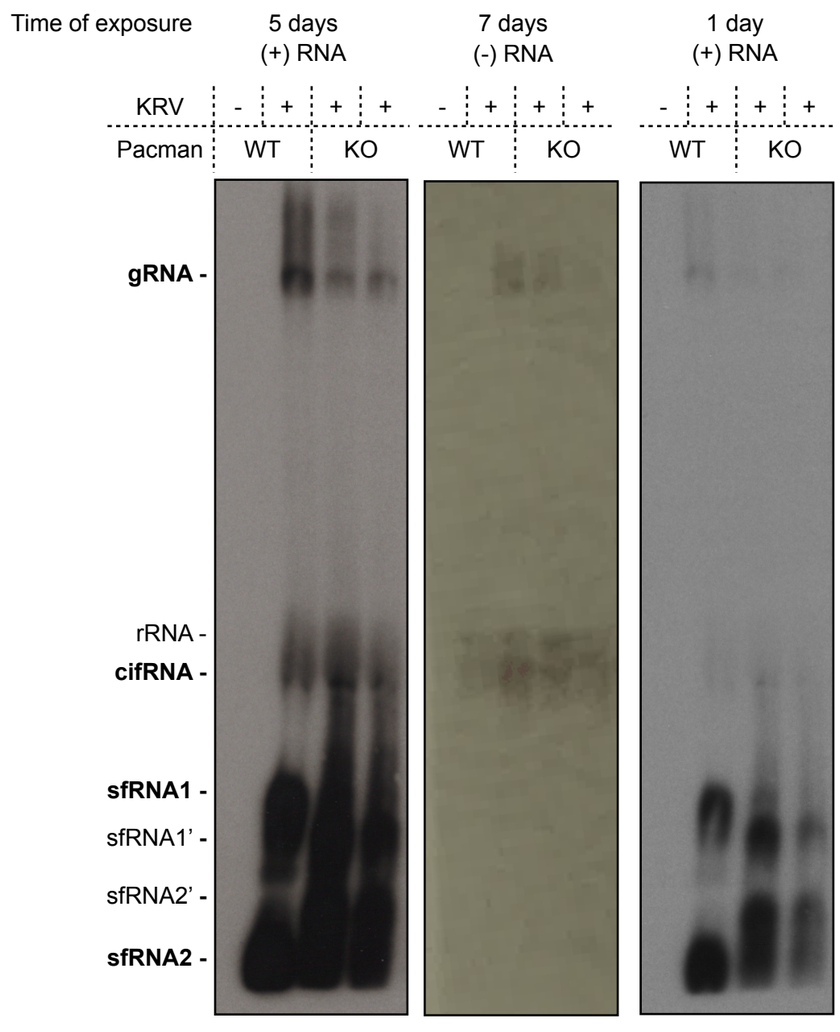
Time of exposure
5 days
(+) RNA
7 days
(-) RNA
1 day
(+) RNA
KRV
- + + + - + + + - + + +
Pacman
WT KO WT KO WT KO
gRNA -
rRNA -
cifRNA -
sfRNA1 -
sfRNA1' -
sfRNA2' -
sfRNA2 -

Fig. S7

Supplementary table 1. List of viruses used in pan-flavivirus small RNA analysis

Virus	Strain/Isolate	Start CDS	End CDS	3'UTR start	End	Reference
CFAV	RioPiedras	104	10129	10127	10682	NC_001564.2
CxFV	Uganda08	92	10183	10181	10837	NC_008604.2
DENV2	16881	97	10272	10270	10723	NC_001474.2
KRV	SR-75	97	10170	10168	11375	NC_005064.1
NOUV	B3	80	10408	10406	10755	NC_033715.1
SLEV	Palenque	99	10391	10389	10938	JQ957869.1
SLEV	MSI-7	99	10391	10389	10939	MSI-7(incomplete): DQ359217.1; Hubbard (closest complete): EU566860.1, Refseq: NC_007580.2
WNV	NY99	97	10398	10396	11029	NC_009942.1
ZIKV	H/PF/2013	108	10379	10377	10807	KJ776791.2

Supplementary table 2. List of oligonucleotides for northern blots, qPCR and cloning

Use	Virus	Probe sequence	Position	Notes
Northern Blot Probes	KRV (+)RNA	CAGGGGACTCGGGGGAGCGGGT	10392, 10988	-
		GAGCGGGTGCGTCASGCCCGACAC	10972	-
		TATCTTTTCTATACCATARATGC	11350	-
	KRV (-)RNA	ACCCGCTCCCCCGAGTCCCCTG	10392, 10988	-
		CTGGTTCTCGCAACTCCAGTCGAA	10599	-
		GTGTCGGGCSTGACGCACCCGCTC	10972	-
		TATCTTTTCTATACCATARATGC	11350	-
	CFAV (+)RNA	TGCCCGACATTTTCATGCCCTGCT	10188	-
		GCCCGACACCAATGTGCCCTGCTC	10273	-
		CGCATCTATGGTATAGAAAAGATA	10671	-
	ZIKV (+)RNA	TGTGGCTGACTAGCAGGCCTGACA	10396	-
		TGCCATGGCGTTCTCGGCCTGACT	10480	-
		AGACCCATGGATTTCCCCACACCG	10784	-
qPCR primers	KRV gRNA For	GGTCAATGAGACCGAACGA	7602	-
	KRV gRNA Rev	GTGTATCCATACACAGACGAC	7758	-
	KRV cifRNA For	AATGTGGCAGTCTTATTGGTTC	10001	-
	KRV cifRNA Rev	AATGACGTCATTCCCCTTCC	10148	-
	KRV sfRNA1 For	GGTGACCTGTCTCATACATG	10503	-
	KRV sfRNA1 Rev	ACATTGCTGATCCTTGTTCC	10637	-
	KRV sfRNA2 For	GTCATAGGCACCTGACCTG	11073	-
	KRV sfRNA2 Rev	TCCGTCCGGTTTTGAAAGC	11232	-
	Aalb pacman KO For	GCAACTCGGGCCGAACAAAC	1321	Sanger sequencing of Cas9 edited region
	Aalb pacman KO Rev	CTGATCGCGATAAATTGCCG	1975	
	Aalb pacman Set1 For	GGGCTACCAGGACTATAATG	1299	AALFPA-065179 / 057530 / 079140
	Aalb pacman Set1 Rev	TCCCGTTGAGATCGGTTTCG	1639	
	Aalb pacman Set2 For	CAGAGGAAGCACGAGTTGGT	3568	AALFPA-052256 /

	Aalb pacman Set2 Rev	ATGTGGACGGCACTTGTGAT	3802	057530 / 079140
	Aalb pacman Set3 For	GCGCTGAAAACTCTGAAGCC	2431	AALFPA-057530 / 079140
	Aalb pacman Set3 Rev	CCAGATTGGGTTCCTCGTGT	2602	
	RPL5 For	TCGCTTACGCCCGCATTGAGGGTGAT	-	Housekeeping gene
	RPL5 Rev	TCGCCGGTCACATCGGTACAGCCA	-	
Cloning of viral sequences as standards for qPCR analysis and qPCR primers	WNV NS5 For	TGAAGAGCCCCAACTAGTGC	8010	-
	WNV NS5 Rev	TTCAAGGACCCGAATCGTCC	8159	
	SLEV-MSI7 NS5 For	GAGAGAAGGGCGTCTCACAG	7802	-
	SLEV-MSI7 NS5 Rev	GAACATGCTTCAGGGTTGCG	7943	
	SLEV-Pal NS5 For	CACGTCCAAGAGGTGAAGGG	7956	-
	SLEV-Pal NS5 Rev	TCACAGCTCGGGTTTGACTC	8115	
	DENV2 NS5 For	GTAGTGGACCTCGGTTGTGG	7798	-
	DENV2 NS5 Rev	GTGTCACACTTTTCTGGCGG	7975	
	ZIKV NS3 For	GAGAGAGTCATTCTGGCTGGA	5925	-
	ZIKV NS3 Rev	TCCCTCAATGGCTGCTACTT	6154	
	NOUV NS5 For	AAGCCTACACGAAAGGAGGC	7992	-
	NOUV NS5 Rev	ACGACTCCCCAATGTCACAC	8122	
	CxFV NS1 For	GATCCGGAGGGTTTGTGTGG	2265	-
	CxFV NS1 Rev	GCATTGTAGGACATCCTCAC	2400	
	CFAV NS1 For	GCAGCGGCGCTTTTGTGTGG	2265	-
	CFAV NS1 Rev	GCACTGCAAGGCATCCTCAC	2400	
	KRV NS5 For	ATCCACAGCTGTAGGCCTTG	7515	-
	KRV NS5 Rev	CAACCCGTCCGTTTGGTTTC	7680	
Cloning of KRV end-to-end ligations for sequencing	cifRNA For	CGACTCTAGAGGATCC-AACCAATAAGACTGCCACATTC	1126	-
	cifRNA Rev	ATTCGAGCTCGGTACC-GCAATGCACTCCTGAGTAG	10000	
	sfRNA1 For	CGACTCTAGAGGATCC-GTATGAGACAGGTCACCACT	1126	-
	sfRNA1 Rev	ATTCGAGCTCGGTACC-GCAATGCACTCCTGAGTAG	10500	
	sfRNA2 For	CGACTCTAGAGGATCC-TAAGGCGCCACTCTTATCC	1126	-
	sfRNA2 Rev	ATTCGAGCTCGGTACC-GCAATGCACTCCTGAGTAG	11092	
	sfRNA1' and 2' For	CGACTCTAGAGGATCC-ATGGCGTTTTCAATGAGATAGG	1126	-
	sfRNA1' and 2' Rev	ATTCGAGCTCGGTACC-GCAATGCACTCCTGAGTAG	10933	

Cloning of guide RNA for CRISPR/Cas9 gene editing	Amp till tracrRNA (XbaI) For	TCCTTCGGTCCTCCGATCGTTG	-	In-Fusion #1
	Amp till tracrRNA (XbaI) Rev	TGCTTTTTTTCTAGAAGATCTGGAAAAATGATGTG		
	tracrRNA till Aalb pU6 (XbaI) For	TCTAGAAAAAAAGCACCGACTCGGTG	-	In-Fusion #2
	tracrRNA till Aalb pU6 (XbaI) Rev	GACGGCAATGAAATGGAAGAGCGAGCTCTTCC		
	Aalb pU6-2 till pAc (XbaI) For	CATTTCATTGCCGTCGCTTC	-	In-Fusion #3
	Aalb pU6-2 till pAc (XbaI) Rev	CGAGATCTGTCTAGACTCAGCTCGAGTGTGGTCTTAG		
	pAc Fw (XbaI) For	TCTAGACAGATCTCGCTGCCTGTTATG	-	In-Fusion #4
	pAc Fw (XbaI) Rev	GCCAAGAATGGAGCGATCGC		
	Aalb pacman guide #2.1 For	AATGTGGTGTCCCATCCTGGGCT	-	-
	Aalb pacman guide #2.1 Rev	AACAGCCCAGGATGGGACACCAC		
	Aalb pacman guide #3.1 For	AATGCGTGTGGTAAGCACTGGGC	-	-
	Aalb pacman guide #3.1 Rev	AACGCCCAGTGCTTACCACACGC		

Supplementary table 3. List of genome references used for 3'UTR analysis

RefSeq	Clade	Full name	Abbreviation	5'UTR size (nt)	CDS size (nt)	3'UTR size (nt)
NC_012932.1	ISFV	Aedes flavivirus	AEFV	96	10026	942
NC_001564.2	ISFV	Cell fusing agent virus	CFAV	103	10026	553
NC_008604.2	ISFV	Culex flavivirus	CxFV	91	10092	654
NC_005064.1	ISFV	Kamiti River virus	KRV	96	10074	1208
NC_027819.1	ISFV	Mercadeo virus	MECDV	88	10212	638
NC_021069.1	ISFV	Mosquito flavivirus	MSFV	111	10080	674
NC_027817.1	ISFV	Parramatta River virus	PaRV	109	10155	629
NC_012671.1	ISFV	Quang Binh virus	QBV	112	10080	673
NC_017086.1	ISFV	Chaoyang virus	CHAOV	99	10308	326
NC_016997.1	ISFV	Donggang virus	DONV	113	10335	343
NC_027999.1	ISFV	Paraiso Escondido virus	EPEV	119	10326	316
NC_040610.1	ISFV	Nanay virus	NANV	106	10299	399
NC_034017.1	ISFV	Xishuangbanna aedes flavivirus	XFV	90	10245	549
NC_009026.2	MBFV	Aroa virus	AROAV	104	10290	421
NC_001477.1	MBFV	Dengue virus 1	DENV1	94	10179	462
NC_001474.2	MBFV	Dengue virus 2	DENV2	96	10176	451
NC_001475.2	MBFV	Dengue virus 3	DENV3	94	10173	440
NC_002640.1	MBFV	Dengue virus 4	DENV4	101	10164	384
NC_012533.1	MBFV	Kedougou virus	KEDV	106	10227	390
NC_001437.1	MBFV	Japanese encephalitis virus	JEV	95	10299	582
NC_000943.1	MBFV	Murray Valley encephalitis virus	MVEV	95	10305	614
NC_007580.2	MBFV	Saint Louis encephalitis virus	SLEV	98	10293	549
NC_009942.1	MBFV	West Nile virus	WNV lin.1	96	10302	631
NC_001563.2	MBFV	West Nile virus	WNV lin.2	96	10293	573
NC_009029.2	MBFV	Kokobera virus	KOKV	83	10233	558
NC_012534.1	MBFV	Bagaza virus	BAGV	94	10281	566
NC_009028.2	MBFV	Ilheus virus	ILHV	92	10275	388
NC_040776.1	MBFV	Rocio virus	ROCV	92	10278	424
NC_034151.1	MBFV	T'Ho virus	THOV	97	10284	556
NC_035889.1	MBFV	Zika virus	ZIKV/2015	107	10272	429
NC_012532.1	MBFV	Zika virus	ZIKV/1947	106	10260	428
NC_008719.1	MBFV	Sepik virus	SEPV	116	10218	459
NC_012735.1	MBFV	Wesselsbron virus	WESSV	118	10218	478
NC_002031.1	MBFV	Yellow fever virus 17D	YFV17D	118	10236	508
NC_005039.1	NKV	Yokose virus	YOKV	150	10278	429
NC_003635.1	NKV	Modoc virus	MODV	109	10125	366
NC_004119.1	NKV	Montana myotis leukoencephalitis virus	MMLV	108	10125	460
NC_003690.1	TBFV	Langat virus	LGTV	130	10245	568

NC_001809.1	TBFV	Louping ill virus	LIV	129	10245	500
NC_005062.1	TBFV	Omsk hemorrhagic fever virus	OHFV	132	10245	410
NC_003687.1	TBFV	Powassan virus	POWV	111	10248	480
NC_027709.1	TBFV	Spanish goat encephalitis virus	SGEV	132	10245	493
NC_001672.1	TBFV	Tick-borne encephalitis virus	TBEV	132	10245	764
					Median	493