The PAZ domain of *Aedes aegypti* Dicer 2 is critical for accurate and high-fidelity size determination of virus-derived small interfering RNAs.

Melinda Reuter[1,2,3,4#], Rhys H. Parry[5#], Melanie McFarlane[1], Rommel J. Gestuveo[6], Rozeena Arif[1], Alexander A. Khromykh[5,7], Benjamin Brennan[1], Margus Varjak[8], Alfredo Castello[1], Lars Redecke[9,10], Esther Schnettler[2,3,4*], Alain Kohl[1,11*]

[1]MRC-University of Glasgow Centre for Virus Research, Glasgow, UK; [2]Bernhard-Nocht-Institute for Tropical Medicine, Hamburg, Germany; [3]German Centre for Infection Research (DZIF), Partner Site Hamburg-Luebeck-Borstel-Riems, Hamburg, Germany; [4]Faculty of Mathematics, Informatics and Natural Sciences, University Hamburg, Hamburg, Germany; [5]School of Chemistry and Molecular Biosciences, The University of Queensland, St. Lucia, Australia; [6]Division of Biological Sciences, University of the Philippines Visayas, Miagao, Iloilo, Philippines; [7]Australian Infectious Diseases Research Centre, Global Virus Network Centre of Excellence, Brisbane, Queensland, Australia; [8]Institute of Technology, University of Tartu, Estonia; [9]University of Lübeck, Institute of Biochemistry, Lübeck, Germany; [10]Deutsches Elektronen Synchrotron (DESY), Photon Science, Hamburg, Germany; [11]Centre for Neglected Tropical Diseases, Departments of Tropical Disease Biology and Vector Biology, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, UK

[#]These authors contributed equally to this manuscript.

*Correspondence: schnettler@bnitm.de and alain.kohl@lstmed.ac.uk

Short title: Analysis of the *Aedes aegypti* Dicer 2 PAZ domain

Keywords: RNA interference; small RNA; Dcr2; arbovirus; mosquito cell; host response.

ABSTRACT

The exogenous siRNA (exo-siRNA) pathway is a critical RNA interference response involved in controlling arbovirus replication in mosquito cells. It is initiated by the detection of viral long double-stranded RNA (dsRNA) by the RNase III enzyme Dicer 2 (Dcr2), which is processed into predominantly 21 nucleotide (nt) virus-derived small interfering RNAs, or vsiRNAs that are taken up by the Argonaute 2 (Ago2) protein to target viral single-stranded RNAs. The detailed understanding of Dicer structure, function and domains owes much to studies outside the context of viral infection and studies in model organisms, and as such how Dcr2 domains contribute to detecting viral dsRNA to mount antiviral responses in infected mosquito cells remains less well understood. Here, we used a Dcr2 reconstitution system in *Aedes aegypti* derived Dcr2 KO cells to assess the contribution of the PAZ domain to induction of the exo-siRNA pathway following infection with Semliki Forest virus (SFV; *Togaviridae*, *Alphavirus*). Amino acids critical for PAZ activity were identified, and loss of PAZ function affected the production of 21 nt vsiRNAs -with enrichment of 22 nt SFV-derived small RNAs observed- and silencing activity. This study establishes PAZ domain's functional contribution to Dcr2 processing of viral dsRNA to 21 nt vsiRNAs.

INTRODUCTION

The emergence and re-emergence of arboviruses pose a continuing threat to human and animal health, with mosquito-transmitted pathogens such as dengue virus (DENV; *Flaviviridae, Flavivirus*), Zika virus (ZIKV; *Flaviviridae, Flavivirus*) and chikungunya virus (CHIKV; *Togaviridae, Alphavirus*) as prominent examples of the threats posed (Zaid et al. 2021; Halstead 2019; Weaver et al. 2018; Weaver and Reisen 2010; Huang et al. 2019; Pierson and Diamond 2020). The mosquito *Aedes aegypti* is an important vector for the transmission of many human-infecting arboviruses, including DENV, ZIKV and CHIKV. It is found across many areas of the tropics and subtropics where it is well adapted to the human environment, preferentially taking bloodmeals from human hosts which can result in arbovirus transmission. Prevention of transmission frequently relies on chemical control

measures that target vectors, though novel approaches such as *Wolbachia* endosymbionts and gene drive-based approaches are promising (Achee et al. 2019; Facchinelli et al. 2023; Carvalho and Moreira 2017; Brady and Hay 2020; Caragata et al. 2021; Wang et al. 2021, 2022).

Arboviruses actively replicate in their vectors, which results in the induction of immune responses. Understanding these is relevant for a better comprehension of vector processes that impact arbovirus transmission and thus targets for intervention (Alonso-Palomares et al. 2018; Lambrechts and Saleh 2019). Among host responses that control arbovirus replication in vectors, the RNA interference response -specifically the exogenous siRNA (exo-siRNA) pathway- plays a critical role. Much of our current understanding of this pathway comes from studies in *Drosophila melanogaster*, including *in vitro* studies, with critical components and mechanisms also conserved in mosquitoes thus, proving the importance of the exo-siRNA pathway across insects. It is initiated following targeting of viral replication induced long double-stranded RNA (dsRNA) by the RNase III enzyme Dicer 2 (Dcr2), which cleaves the dsRNA into virus-derived small interfering RNAs (vsiRNAs). In mosquitoes and *D. melanogaster*, each vsiRNA strand is predominantly 21 nucleotides (nt) in length, with a 19 nt overlap and 2 nt overhangs at the 3´end. In the next step, 21 nt vsiRNAs are taken up by Argonaute 2 (Ago2). Ago2 is part of the RNA-induced silencing complex (RISC) and retains one strand of the siRNA to target and degrade complementary viral RNAs (Blair and Olson 2015; Tikhe and Dimopoulos 2021; Olson and Blair 2015; Bronkhorst and van Rij 2014; Swevers et al. 2018; Prince et al. 2023; Leggewie and Schnettler 2018; Samuel et al. 2018). Indeed, absence of Dcr2 in mosquito cell lines resulted in the loss of 21 nt vsiRNAs and impaired antiviral responses (Brackney et al. 2010; Gestuveo et al. 2022; Scott et al. 2010; Varjak et al. 2017), while *Ae. aegypti* mosquitoes without functional Dcr2 develop disease phenotypes, increased replication and dissemination following arbovirus infection (Merkling et al. 2023; Samuel et al. 2023). These data show the importance of Dcr2 as an initiator of the exo-siRNA pathway, but other functional aspects of this effector protein are less understood especially in the context of viral infection and targeting of viral derived dsRNA.

The domain structure of *Ae. aegypti* Dcr2 (Aaeg Dcr2) is similar to that of *D. melanogaster* Dcr (Dm Dcr2), with helicase-DUF-PAZ-RNase 3A-RNase 3B-dsRBD from N to C terminus of the protein (Fig. 1), though other functional domains are present in the protein, too (Paturi and Deshmukh 2021). Structural studies have revealed that Dicer proteins generally have an L shape in which these different domains are arranged (Paturi and Deshmukh 2021; Zapletal et al. 2023; Ciechanowska et al. 2021). Analyses of Dm Dcr2 have confirmed this shape (Yamaguchi et al. 2022; Sinha et al. 2018; Su et al. 2022) , which was also predicted for Aaeg Dcr2 using AlphaFold (Gestuveo et al. 2022). The bottom (or base) module of the L-shaped Dcr2 structure contains the helicase domain; the core comprises RNase 3A/B and dsRBD domains, and the cap module the PAZ (Piwi/Ago/Zwille) domain. Biochemical and structural studies have investigated the dsRNA binding properties of Dm Dcr2 domains, with the PAZ domain preferably binding dsRNA termini with 3' overhangs (Sinha et al. 2018). The structural arrangements of the 3' nt overhang binding pocket in the PAZ domain and 5' monophosphate in the adjacent Platform domain have been investigated in detail. Importantly, interactions of dsRNA with PAZ and Platform via 3' 2 nt overhang and 5' monophosphate of the dsRNA allows positioning and spacing of the dsRNA for precise dicing activity and siRNA length (Su et al. 2022; Yamaguchi et al. 2022). These structural data consolidate a previous study on the importance of binding of the 5' monophosphate in dsRNA for length fidelity in 21 nt siRNA production from long dsRNA- though not efficiency of small RNA production *per se*; and mutagenesis analysis of the Dm Dcr2 PAZ domain also demonstrated the critical role of 21 nt siRNAs in silencing an inverted repeat transgene in the fly model (Kandasamy and Fukunaga 2016). Interestingly, the presence of 5' monophosphate and two specific arginine residues in Dm Dcr2 PAZ were found to be critical for cleavage of short dsRNA (Cenik et al. 2011; Fukunaga et al. 2014), though not longer ones where (as stated above) accurate length determination of siRNAs is the critical contribution of 5' monophosphate binding to the dicing process.

A limitation of studies on Dm Dcr2 PAZ domain is their reliance on *in vitro* studies with dsRNA substrates or endogenous small RNAs, while the role of this domain in antiviral

responses during infection when targeting viral derived dsRNA in the cell is, to our knowledge, still unclear. The importance of the Dcr2 helicase domain in antiviral responses has been previously shown (Gestuveo et al. 2022; Marques et al. 2013; Donelick et al. 2020), there remains a critical gap because the nature of the viral dsRNA substrate is unknown- does Dcr2 detect long, short, or different types of dsRNA in infected cells, and how does recognition occur? Here we aimed to expand our previous studies on the role of Dcr2 domains in mosquito cell antiviral responses by investigating the role of the Aaeg Dcr2 PAZ domain in mosquito antiviral vsiRNA generation and its requirement for efficient antiviral responses. For this, we used the previously described Aaeg Dcr2 activity reconstitution system in *Ae. aegypti* derived Dcr2 KO cells, with which we demonstrated the importance of the helicase and RNase domains and that 21 nt vsiRNAs are generally produced by a processive production mechanism from viral dsRNA (Gestuveo et al. 2022). To understand the contribution of the PAZ domain to the antiviral activity of Aaeg Dcr2, we identified conserved amino acids in Aaeg Dcr2 PAZ domain which were mutated to assess their impact on dsRNA processing activity. Mutations abrogated antiviral activity against a positive sense RNA virus, Semliki Forest virus (SFV), a member of the *Togaviridae* family, genus *Alphavirus*. PAZ function inactivating mutations resulted in a loss of 21 nt vsiRNA generation, and enrichment of 22 nt SFV-derived small RNAs. These findings indicate that PAZ domain functionality is critical for determination of the usually predominant vsiRNA length of 21 nt, but mutations did not affect processive cleavage of viral dsRNA as such.

RESULTS

Identification of amino acids in Aaeg Dcr2 PAZ domain required for antiviral activity.

Analysis of Aaeg Dcr2 sequence (InterProScan, Geneious) indicated that the PAZ domain was located from positions N824 to G941 (Fig. 1A). This was in line with NCBI domain analysis, which also located the PAZ domain between amino acids 824-941. Next, Dcr2 PAZ domain sequences of *Ae. aegypti* (including recently available sequences, see (Gestuveo et al. 2022)), *Ae. albopictus*, *D. melanogaster* as well as *Anopheles* and *Culex* spec. were

aligned to identify highly conserved amino acids in the PAZ domain and thus with potential functional activities. Several positions, as shown in Fig. 1B, were either fully conserved between at least *Ae. aegypti, Ae. albopictus*, and *D. melanogaster* Dcr2 PAZ domains (amino acids D829, N872, V880, V883, Y909, V923) OR across mosquitoes (amino acid L930) and chosen for mutagenesis in different combinations, which we assumed would give the best chance of success to disrupt PAZ domain functionality. For completeness, corresponding amino acid positions in Dm Dcr2 are indicated in Fig. 2A. Using NCBI domain analysis, the PAZ domain of Dm Dcr2 (1722 amino acids in length) was shown as located from amino acids 843-1004.

Alphafold was used to determine the location/structure of these conserved amino acids within the structure of the Aaeg Dcr2 PAZ domain, which appeared mainly located (with the exception of L930) around its core (Supplemental Fig. S1) when compared to Dm Dcr2 (Yamaguchi et al. 2022), suggesting structural and/ or functional importance. Combinatorial mutagenesis was carried out as described for mutants M1-4, with conserved amino acids replaced in the combinations shown as indicated with the aim of achieving disruption PAZ domain function with a high degree of certainty (Fig. 2A). PAZ domain mutant M1-M4 and WT Dcr2 expression in AF319 Dcr2 KO cells transfected with the corresponding expression constructs was assessed by western blot, showing that all proteins were expressed (Fig. 2B, Supplemental Fig. S2).

Next, we assessed the functionality of the Aaeg Dcr2 mutants for their silencing ability and antiviral activity. WT Dcr2 displayed antiviral activity against SFV-FFLuc in this transfection-based assay (Fig. 3), as previously demonstrated (Gestuveo et al. 2022). M1-M4 Dcr2 mutants lost SFV-FFLuc antiviral activity, compared to WT Dcr2. Similarly, the functional silencing abilities of the mutant Dcr2 were investigated in a reporter plasmid-based assay, where exogenous dsRNA is provided to induce the exo-siRNA pathway against a reporter gene (Gestuveo et al. 2022). Here, M1-M4 Dcr2 mutants only retained residual (although

significant compared to control) ability to silence the reporter gene. This indicated that the mutation combinations in M1-M4 all affected Dcr2 PAZ domain functionality, and thus Dcr2's ability as an effector in the exo-siRNA pathway.

Aaeg Dcr2 PAZ domain loss of function mutations: effects on 21 nt vsiRNA production.

The impact of Aaeg Dcr2 PAZ domain mutations on vsiRNA production was investigated utilising high throughput small RNA sequencing from AF319 cells transiently expressing Dcr2 PAZ domain mutants or WT Dcr2 and infected with SFV. Raw populations of small RNA read lengths prototypical of miRNAs (22 nt), siRNAs (21 nt) and piRNAs (24-30 nt) (Supplemental Fig. S3) were observed in all cells. Small RNA profiles from Dcr2 PAZ domain mutants M1-4 and the eGFP control demonstrated a bias towards a read length of 22 nt rather than 21 nt observed in the WT Dcr2 expressing cells. Read mapping profiles of SFV-derived small RNAs indicated that WT Dcr2 efficiently mediated the production of SFV-derived 21 vsiRNAs (Fig. 4A), which are (as expected, see (Gestuveo et al. 2022)) the predominant length in the 18-22 nt size range and aligned along the length of the SFV genome and antigenome. Almost no such vsiRNAs were detected in the eGFP control. Importantly, M1-M4 Dcr2 mutants showed reduced production of 21 nt vsiRNAs (though those remaining align along the length of SFV4 genome and antigenome) and a more even distribution of read numbers across the 20-22 nt small RNA size range, compared to WT Dcr2, was observed. This indicates that Dcr2 PAZ domain mutants M1-M4 have lost the ability to precisely measure siRNA length and consequently do not predominantly produce 21 nt vsiRNAs; but have retained the ability to process viral dsRNA into small RNAs. Crucially we examined absolute magnitudes of SFV-derived small RNAs in the sizes of 20 nt and 22 nt and observed a clear enrichment for 22 nt, ranging from 2.18-fold (Dcr 2 PAZ domain mutant M3) to 3.02-fold (Dcr2 PAZ domain mutant M4) (Fig. 4B). For the 20 nt size, we found modest or no significant change (0.91-1.45fold) consistent with previous results (Kandasamy and Fukunaga 2016). Critically, AlphaFold 3 based modelling of Aaeg Dcr2

196 PAZ domain interactions with dsRNA suggested that mutations did not abrogate the ability to
197 bind dsRNA (Supplemental Fig. S4). This consolidates the idea that conserved amino acid
198 are required for keeping dsRNA in a precise position for accurate cleavage to occur.
199
200 **Virus-derived small RNAs retain a 2 nt overhang cleavage pattern independently of**
201 **Aaeg Dcr2 PAZ domain functionality.**
202 The predominant 21 nt vsiRNAs produced by Dcr2 cleavage that are commonly observed
203 following arbovirus infection of mosquito cells are generated as duplexes. Assuming that
204 mechanisms observed for *D. melanogaster* also apply in mosquitoes, they consist of two 21
205 nt strands that overlap by 19 nt with a 2 nt overhang at each 3´end. To investigate if and
206 how dsRNA cleavage patterns were affected by the PAZ domain for generating virus-derived
207 small RNAs, we analysed strand overlapping small RNA patterns (Fig. 5A, heat maps that
208 show mean overlap probability z-scores of SFV-derived small RNAs 18-30 nt; Fig. 5B,
209 showing number of overlapping pairs per million mapped reads of 19-23 nt SFV-derived
210 small RNAs). For WT Dcr2, but also Dcr2 PAZ domain mutants M1-M4, all virus derived
211 small RNAs in the range of 18-23 nt all showed a clear pattern, with strand complementarity
212 significantly enhanced for overlaps that result in 2 nt overhangs. For example, 21 nt vsiRNAs
213 overlapped with complementary sequences also by 19 nt (thus confirming observations in *D.*
214 *melanogaster*); virus-derived small RNAs that are 20 nt in length overlapped with sequences
215 that are 18 nt in length etc. This showed that PAZ domain mutations affected Aaeg Dcr2's
216 ability to produce the usually predominant 21 vsiRNA over other lengths but did not affect
217 the cleavage mechanism itself. We did however note a higher frequency of precise 2 nt
218 overhangs for virus-derived small RNAs in AF319 expressing M1-M4 Dcr2 mutants
219 compared to WT Dcr2. Indeed, further analysis noted a marked Z score drop for example for
220 19 nt overlap (in 21 nt vsiRNA) and +2 nt overhangs for WT Dcr2 compared to Dcr2 PAZ
221 domain mutants M1-M4 (Supplemental Fig. S5, left panel). Interestingly, when we
222 reanalysed 21 nt vsiRNA data previously published from SFV infection of the Ago2 KO cell
223 line AF525, compared to parental AF5 cells (Scherer et al. 2021), we also observed higher

frequencies of virus-derived small RNA duplexes with +2 nt overhangs (Supplemental Fig. S5, right panel). This leaves the possibility that failure to efficiently transfer or to Ago2, or fully process small RNA (and failure to degrade passenger strands) contributes to the increased presence of intact virus-derived small RNAs duplexes that are detected in this overlap analysis.

Production of virus-derived PIWI-interacting small RNAs (vpiRNAs), which are generally longer (24-30 nt) with a 10 nt overlap of the sense and antisense piRNAs as a result of the so-called ping-pong amplification mechanism in a Dcr2 independent manner (Varjak et al. 2018), was detected across all conditions.

Loss of Aaeg Dcr2 PAZ domain functionality changes the position and magnitude of vsiRNA cleavage.

We have previously shown that Aaeg Dcr2 produces a diverse but largely consistent pool of 21 nt vsiRNAs following SFV infection; a characteristic lost in helicase domain mutants (Gestuveo et al. 2022). To assess the impact of PAZ domain mutations on virus-derived small RNA population, diversity and cleavage bias, we examined these through two different analyses. First we defined a position-based “harmony” metric which incorporates the strand bias and magnitude of the vsiRNAs and second, using a vsiRNA population approach where each unique siRNA was treated as an individual gene and examined differential vsiRNA expression between treatment groups. While visualization of virus-derived small RNA coverage, as demonstrated in Fig. 4, provides a reasonable indicator of the origin and magnitude of vsiRNAs, a more nuanced analysis of the consistency of Dcr2 cleavage over the SFV-derived dsRNA requires consideration of the position, magnitude and strandedness of 21 nt vsiRNAs. To develop a unified metric that encapsulates these attributes at each genomic position, we calculated the terminal 5’ vsiRNA read depth at each position of the SFV genome and antigenome, and normalized the total depth to one. Subsequently, we computed the normalized differential coverage, d_i, as the difference between the normalised genome (positive-sense) and anti-genome (negative-sense) coverage proportions at each

position. The genome coverage proportion at position i, G_i, is defined as the positive-sense coverage at i divided by the total positive-sense coverage for all N positions. The anti-genome coverage proportion, A_i, is similarly calculated as the negative-sense coverage at i normalized by the total negative-sense coverage. The formula for the normalized differential coverage at each position i is given by:

$$d_i = \frac{G_i}{\sum_{j=1}^{N} G_j} - \frac{A_i}{\sum_{j=1}^{N} A_j}$$

Through this measure, we obtained a ratio reflecting the balance of genome to antigenome virus-derived small RNA coverage, which provides insight into the strand-specific cleavage patterns of WT Dcr2 across the positions of the SFV genome. Using this metric, we generated a series of scatter plots (Fig. 6A) to compare the d_i values for each position of the SFV genome between WT Dcr2, negative control eGFP and the Dcr2 PAZ domain mutants M1-M4. Importantly, if the d_i value was 0 for a position for the comparison between treatments it was excluded from the analysis, keeping only non-zero d_i values for downstream analysis. We utilized linear regression analysis to assess the relationship between d_i values for all positions on the SFV genome from the WT Dcr2 and each mutant, with the coefficient of determination (R^2) reflecting the extent to which both treatments shared related vsiRNA profile “harmony”. The regression analysis revealed varying degrees of correlation, with the lowest correlation between the eGFP control treatments compared to the WT Dcr2 (R^2=0.115). This suggests very little positional cleavage harmony between the WT Dcr2 and eGFP control, which is to be expected. Dcr2 PAZ domain mutants M1-M4 showed a more consistent pattern of Dcr2 cleavage to WT with R^2 values between 0.39-0.41. Finally, we compared d_i values for all 21 nt vsiRNA reads for all treatments and between mutants and summarised the R^2 values as a heat map (Figure 6B). The results indicated that the PAZ domain mutants M1-M4 were more closely related to each other than WT Dcr2 or eGFP control with R^2 between 0.74-0.84, indicating that the introduced

mutations resulted in more harmonious cleavage patterns across the SFV genome but different from wild type.

In addition to analyzing positional cleavage patterns of 21 nt vsiRNAs (Fig. 6A-B), we examined population-level patterns and diversity of SFV-derived small RNA populations for individual lengths of 20 nt, 21 nt, 22 nt, and 23 nt, as well as combined lengths of 20-23 nt by conducting a differential transcript abundance analysis using edgeR (Fig. 6C). Through the generation of a non-redundant list of unique small RNAs, we generated count tables and normalized for library size using the weighted trimmed mean of M-values (TMM) method. Through multi-dimensional scaling (MDS), plots we were able to visualize the relationships between the treatments. MDS plots are a form of dimensionality reduction; visually, the distances between points represent the dissimilarities or differences between the treatments. Specifically, these distances correspond to the leading log-fold changes between the treatments, which reflect the relative expression levels of vsiRNAs across conditions. The MDS plots revealed varying degrees of similarity between the WT Dcr2 and Dcr2 PAZ domain mutant treatments across different virus-derived small RNA size classes. Notably, the 21 nt vsiRNAs displayed distinct clustering patterns, indicating that specific mutations in the Dcr2 PAZ domain contribute to the altered distribution and abundance of these vsiRNA species.

DISCUSSION

While the importance of the helicase domain has been demonstrated for both Dm and Aaeg Dcr2 antiviral responses (Gestuveo et al. 2022; Marques et al. 2013; Donelick et al. 2020), the relevance of the PAZ domain had not yet been investigated for Aaeg Dcr2 and processing of viral dsRNA in the context of arboviral replication in mosquito cells. Here, we identified conserved amino acids in Aaeg Dcr2 that inactivate PAZ domain function and antiviral activity. Critically, we observed that while inactivation did affect the production of 21 nt vsiRNA, it did not affect Aaeg Dcr2's virus-derived small RNA production in general but broadened the length of the produced SFV-derived small RNAs and led to an enrichment of

306 specifically 22 nt SFV-derived small RNAs, similar to previous mutational analysis of Dm
307 Dcr2 PAZ domain (Kandasamy and Fukunaga 2016). Thus, key observations hold true for
308 viral dsRNAs and Aaeg Dcr2. The specific functions that are affected by the mutations we
309 set in Aaeg Dcr2 PAZ will require further investigation. However, structural modelling
310 (Supplemental Fig. S1) suggests positioning mostly in the core of the PAZ domain and thus
311 a role in the binding mechanism of the dsRNA termini. This is further confirmed by modelling
312 Aaeg PAZ domain interactions with dsRNA (Supplemental Fig. S4). Suggesting that the
313 location of dsRNA is not affected but most likely mutations destabilise binding and thus allow
314 flexibility in the interaction, which in turn leads to variable small RNA lengths following
315 cleavage and in particular enrichment of 22 nt SFV-derived small RNAs. We cannot of
316 course fully exclude effects on dsRNA binding or impaired processing that may not be
317 captured by our analysis. Importanly, the critical PAZ domain amino acids we described here
318 in Aaeg Dcr2 are different from those previously identified in Dm Dcr2 (Kandasamy and
319 Fukunaga 2016). Based on our and previous data that demonstrate the importance of the
320 Dcr2 helicase domain in the exo-siRNA pathway, we propose that the helicase binds usually
321 blunt-ended, long viral dsRNA and cleavage of these termini -which is likely to generate
322 small RNAs of variable length- generates the first dsRNA terminus with 2 nt overhang and 5'
323 monophosphate that will be used by the PAZ domain to anchor the dsRNA strand and allow
324 Dcr2 to produce mostly, though not exclusively, 21 nt vsiRNAs, and phase variation from
325 occasional unprecise binding that allows the production of a diverse pool of vsiRNAs from
326 dsRNA, may therefore start from early cleavage events onwards. Importantly, the size of 21
327 nt is critical and the most efficient siRNA length for silencing to target RNAs in the fly system
328 (Kandasamy and Fukunaga 2016; Elbashir et al. 2001). It is intriguing to note that U is
329 enriched in first position in libraries (Supplemental Fig. S3) compared to vsiRNAs where a
330 bias towards C or G can be seen. Indeed this may be due to presence of host derived
331 sequences eg. miRNAs that frequently have a U in the 5' first base position (Ghildiyal et al.
332 2010; Mi et al. 2008; Warf et al. 2011). Indeed in *D. melanogaster*, it has been shown that
333 siRNAs derived from siRNA pathway, for example following processing long dsRNA, give

rise to siRNAs with frequently C in position 1 which has been suggested to facilitate sorting into Ago2 (Ghildiyal et al. 2010) . A similar mechanism may be at play here in the production of vsiRNAs. Our data also suggest that maturation of vsiRNAs -by cleaving the passenger strand- is less efficient with Dcr2 PAZ domain mutants M1-M4. While the exact processes would need further investigations, the marked increases in duplexes (with +2 nt overhangs) similar to what is observed in Ago2 KO cells infected SFV where an accumulation of duplexes is seen (Scherer et al. 2021), might be due due for example to a failure of vsiRNAs to efficiently transfer into Ago2 via dysfunctional Dcr2 PAZ domain, or a processing issue. This requires further investigation, and there is a degree of “noisiness” in such assays that needs to be taken into account. Indeed, in our previous work (Gestuveo et al. 2022) for WT Dcr2, 19 nt overlaps with 2 nt overhangs for example in 21 nt vsiRNAs were easily observed than in this study. This may be due to technical issues, for example small RNA sequencing, transfection levels of Dcr2 expression plasmids etc. and to some extent production of stable Dcr2 expressing cell lines may remediate small differences for future work.

In summary, this study demonstrates the critical role of the Aaeg Dcr2 PAZ domain in the precise production of 21 nt vsiRNAs. Its critical contribution is in accurate vsiRNA length determination, which in turn is critical for the *Ae. aegypti* exo-siRNA to mount an antiviral response against SFV. Moreover, future work will determine how arboviruses of different families are targeted by Dcr2 and give us further insights into the nature of the viral dsRNA substrate.

MATERIALS AND METHODS

Cells

The Dcr2 knockout cell line Aag2-AF319 (subsequently abbreviated to AF319) used here is derived from *Ae. aegypti* Aag2-AF5 cells, a clone derived from the Aag2 cell line (provided by Kevin Maringer, The Pirbright Institute, UK) (Fredericks et al. 2019; Varjak et al. 2017). As previously described AF319 cells were cultured in Leibovitz's L-15 medium with

GlutaMax (Gibco) with 10% tryptose phosphate broth (TPB; Gibco), 10% fetal bovine serum (FBS; Gibco), and penicillin-streptomycin (pen-strep; 100 U/mL-100 µg/mL; Gibco) at 28 °C (Fredericks et al. 2019; Varjak et al. 2017). Cell lines are available as Aag2-AF319 (ECACC 19022602) and Aag2-AF5 (ECACC 19022601) through Public Health England. Baby hamster kidney (BHK-21) cells are a commonly used cell line available at the MRC-University of Glasgow Centre for Virus Research; these cells were grown in Glasgow Minimum Essential Medium (GMEM; Gibco) supplemented with 10% TPB, 10% newborn calf serum (Gibco), and pen-strep at 37 °C with 5% CO_2.

Viruses

Virus production from icDNA and plaque assay titrations were performed as previously described (Ulper et al. 2008; Varjak et al. 2017) to produce SFV4 (abbreviated to SFV; GenBank ID: KP699763) and SFV4(3H)-*FFLuc* (abbreviated to SFV-FFLuc; reporter virus expressing FFLuc).

Plasmid generation

Mutations in the PAZ domain of Aaeg Dcr2 (GenBank ID: AAW48725) were created by In-Fusion cloning of synthetically generated PAZ domain mutant fragments into the previously generated pPUb-myc-Dcr2 (Varjak et al. 2017) following the manufacturer's guidelines. Briefly, pPUb-myc-Dcr2 was linearised by PCR, removing the WT PAZ domain. Mutant PAZ domain fragments were PCR amplified using primers with extensions that were homologous to the linearised vector. In-Fusion cloning was then performed to create pPUb-myc-Dcr2 constructs containing the various mutants. Mutations were based on multiple sequence alignment and identification of conserved amino acids in *Ae. aegypti*, *D. melanogaster* and *Ae. albopictus*.

Virus infections

For investigations of antiviral activity against SFV, AF319 cells were seeded at 2 x 10^5 cells/well in 24-well plates and transfected with 1 µg pPUb plasmid expressing WT or mutant Dcr2 or eGFP (as control) after 24 h using Dharmafect2 (Horizon Discovery) following the recommended protocol. At 24 hpt, the cells were infected with SFV-FFLuc (MOI=0.1) and lysed at 48 hpi using 1X Passive Lysis Buffer (PLB; Promega). FFLuc levels were measured using a Luciferase Assay System (Promega) according to the manufacturer's protocol in GloMax Multi-Detection System with Dual Injectors (Promega). For the production of samples for small RNA sequencing, 2.5 x 10^5 cells/well AF319 cells were seeded in 8 wells of a 24 well plate, transfected as above after 24 h and infected with SFV (MOI=1) at 24 hpt for 48 h. For RNA extraction and follow up small RNA sequencing, cells from 7 wells were combined and lysed using TRIzol (Thermo Fisher Scientific) according to the manufacturer's guidelines, with glycogen as a carrier. The remaining well was lysed in LDS protein lysis buffer and used for confirmatory western blot analysis.

dsRNA production

FFLuc (dsFFLuc) and eGFP (dseGFP) were amplified using T7 RNA polymerase promoter-flanked primers (Varjak et al. 2017; McFarlane et al. 2020) and dsRNAs produced with MEGAscript RNAi kits (Thermo Fisher Scientific) by *in vitro* transcription as per manufacturer's guidelines. Following *in vitro* transcription, products were treated with DNaseI, and RNase A. dsRNA was purified by column purification.

RNAi reporter assay

RNAi reporter assays were performed as previously described (Varjak et al. 2017) to assess the silencing activities of mutant Dcr2 compared to WT Dcr2. For this, 2 x 10^5 AF319 cells/well were seeded in 24-well plates. After 24 h cells were co-transfected with 1 µg pPUb-based expression plasmid, 50 ng pGL3-PUb (Anderson et al. 2010) and 20 ng pPUb-RLuc (Alexander et al. 2023) together with either 20 ng dsRNA targeting FFLuc (dsFFLuc)

or 20 ng dsRNA control (dseGFP) and lysed at 24 hpt in 1X PLB (Promega). To determine FFLuc/RLuc levels, dual luciferase assays were performed using a Dual-Luciferase Reporter Assay System (Promega) according to the manufacturer's protocol in a GloMax Multi-Dectection System with Dual Injectors (Promega).

Western blot analysis

Cell lysis was performed in 1X Bolt sample reducing agent and 1X Bolt LDS sample buffer (Invitrogen). The lysate was separated using a 4-12% Bis-Tris Plus gel (Invitrogen). For transfer onto a nitrocellulose membrane, a Trans-Blot SD Semi-Dry Transfer Cell (BioRad) was used. Membranes were blocked for a minimum of 1 h using 5% (w/v) non-fat dry milk powder in PBS-Tween (PBS with 0.05% Tween 20, PBS-T) and then washed three times for 10 min using PBS-T. As primary antibodies, mouse anti-myc tag antibody (1:2000, Abcam) and mouse anti-α tubulin antibody (1:2000; Sigma-Aldrich) in 5% (w/v) non-fat dry milk powder in PBS-T were used. After overnight incubation at 4 °C, membranes were again washed three times in PBS-T. This was followed by incubation of 1 h with goat anti-mouse DyeLight 800 (1:5000; Invitrogen) secondary antibody conjugated with a near fluorescent dye in 5% (w/v) non-fat dry milk powder in PBS-T. The membranes were again washed three times with PBS-T and a final wash with distilled water and analysed using an Odyssey CLx and Image Studio Lite v.5.2.5 (LI-COR Biosciences).

Small RNA sequencing

To investigate small RNA production, at least 1 µg isolated total RNA per condition were sent for analysis. Small RNA sequencing was performed at BGI-Tech Solutions (DNBSeq, UMI small library, SE50), UMI being unique molecular identifiers. With a minimum of 10 million clean reads, with final libraries resulting between 11-18 million reads (Supplemental Table S2).

Small RNA analysis

Using fastp (v0.23.2) (Chen et al. 2018), adapters and low-quality reads from basecalled fastq files were trimmed, retaining reads between 18-32 nt. The distribution of read sizes and first position bias was calculated using the script '1_fastq_histogram.sh' from the GitHub repository (rhparry/viral_sRNA_tools). Clean, trimmed reads were then mapped to the SFV genome (GenBank ID: KP699763) with Bowtie2 (v2.4.5) (Langmead and Salzberg 2012), using the sensitive mapping flag (—sensitive) as specified in the script '2_mapping_vrnas.sh'. Histograms of mapped read lengths and first base pair bias were generated using samtools (v1.16.1) and '3_bam_sRNA_histogram.sh'. Output BAM files were filtered to include only the 21 nt reads. Coverage statistics across the SFV genome were calculated using '4_viral_srna_coverage.sh', which employs the bedtools genome coverage tool (v.2.27.1) (Quinlan and Hall 2010). For calculating the harmony metric, coverage at the 5' end (-5 flag) of each read was normalized over the whole genome. Normalised coverage files were processed with awk '{f[NR]=$2;r[NR]=$3;sumf+=$2;sumr+=$3} END {for (i=1;i<=NR;i++) {printf "%d %.6e\n", i, (f[i]/sumf)-(r[i]/sumr)}}' to calculate d_i values for each position *i* of the SFV genome. Linear regression analyses for downstream analysis scatter plots of per position d_i values of the SFV genome were analysed in GraphPad Prism (v10.0.2), and R^2 values for all pairwise treatments were computed and presented as a heat map. Overlapping virus-derived small RNA pairs and overlap probabilities (z-score) were calculated iteratively from BAM files using the 'signature.py' small RNA signatures Python script, both single read lengths from 18-30 nt were calculated as per the following conditions (--minquery 18 --maxquery 18 --mintarget 18 --maxtarget 18 --minscope 1 --maxscope 18), collective 18-30 nt read lengths were computed as follows (--minquery 18 --maxquery 30 --mintarget 18 --maxtarget 30 --minscope 1 --maxscope 30) (Antoniewski 2014). To examine the diversity and populations of SFV-derived small RNA species, unique 20-23 nt SFV mapped reads were treated as discrete virus-derived small RNAs. Counts tables for all extracted SFV-derived small RNAs from extracted fasta files from individual libraries were converted, processed, and enumerated in bash (v.3.2). The dimensions of count tables for 20-23 nt virus-derived small RNA species (n

= 75710), 20 nt (n = 13430), 21 nt (n = 27592), 22 nt (n = 20536) and 23 nt (n = 14155) species. Differential gene expression analysis was performed using package edgeR (v.3.30.3) (Robinson et al. 2009) in R Studio (v.1.3.1073), with library size normalization using the weighted TMM method in calcNormFactors() and filtered for a minimum count of 10 with filterByExpr(). Relationships between samples were explored using an MDS plot of normalized count tables with plotMDS(top = 500) of all SFV-derived small RNA species.

Protein structure predictions

The WT sequence of Aaeg Dcr2 (GenBank ID: AAW48725) was used as a query for structure prediction. AlphaFold 2 algorithm (Jumper et al. 2021) was run in monomer mode using a local install of version 2.3.2 without any restrictions. The resulting models were examined and, in accordance with overall model quality predictions, summarised in mean pLDDT value with the highest quality model selected for further analysis. Structural analyses were performed in PyMOL Molecular Graphics System (version 4.5.0, Schrödinger, LLC). Domain positions were assigned and labelled by color based on GenBank sequence annotations, and residues mutated were highlighted.

Ae. aegypti Dcr2-dsRNA interactions were assessed using AlphaFold 3. The WT sequence of *Ae. aegypti* Dcr2 (GenBank ID: AAW48725) and derived mutant sequences (M1-M4) were modelled as interacting with dsRNA. The 53 bp dsRNA sequence was obtained from PDB ID: 7W0E for *D. melanogaster* Dcr2 in the active dicing state (Su et al. 2022). Sequences for WT or M1-M4 *Ae. aegypti* Dcr2 and each mutant were fed into AlphaFold 3 webserver (Abramson et al. 2024) along with the dsRNA sequence to predict protein structures and their interaction with dsRNA. Molecular graphics and analysis of structures were performed using UCSF ChimeraX (Pettersen et al. 2021).

Data analyses

Virus-derived small RNA metrics, coverage and overlapping pairs analysis were visualized using GraphPad Prism (v10.0.2).

DATA DEPOSITION

Data underlying figures are available under http://dx.doi.org/10.5525/gla.researchdata.1479. Small RNA sequencing data generated for this study have been deposited in the NCBI Sequence Read Archive (SRA), available under accession number PRJNA1014654. Scripts utilized for small RNA analysis are available from the viral_sRNA_tools GitHub repository at: https://github.com/rhparry/viral_sRNA_tools

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

COMPETING INTEREST STATEMENT

The authors declare no competing interests. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

ACKNOWLEDGMENTS

For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission.

This study was supported by the UK Medical Research Council (MC_UU_12014/8, MC_UU_00034/4) (A.K.) and (MR/R021562/1, MC_UU_00034/2) (A.C.); European Research Council (ERC) Consolidator Grant ‘vRNP-capture’ N# 101001634 (A.C.); German Centre for Infection Research (DZIF) (TTU 01.708) (E.S.); DFG (project 497659464) (M.R.); Wellcome Trust/Royal Society Sir Henry Dale Fellowship (210462/Z/18/Z) (B.B.); Overseas Scholarship Scheme for PhD by Higher Education Commission of Pakistan (R.A.).

Author contributions are recognised as follows. Conceptualization: M.R., R.P., M.M., R.J.G., A. C., M.V., L.R., E.S., A.K.; Methodology: M.M., M.R., R.P., R.A., R.J.G., L.R.; Formal analysis: M.R., R.P., M.M., R.A., L.R.; Investigation: M.R., R.P., M.M., R. A., L.R.; Validation: M.R., R.P., M.M., R.A., L.R..; Resources: A.A.K., B.B., A.C., L.R., E.S., A.K.; Writing -

Original Draft Preparation: M.R., R.P., M.M., R.J.G., R.A., A.C., L.R., E.S., A.K.; Writing - Review and Editing: M.R., R.P., M.M., R.J.G., R.A., A.A.K., B.B., M.V, A.C., L.R., E.S., A.K..; Visualization: M.R., R.P, M.M., R.A., L.R..; Data Curation: M.R., R.P., M.M., E.S., A.K.; Supervision: M.M., A.A.K., B.B., A.C., E.S., A.K.; Project Administration: M.M., E.S, A.K.; Funding Acquisition: R.A., B.B., A.C., E.S., A.K.

REFERENCES

Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, Ronneberger O, Willmore L, Ballard AJ, Bambrick J, et al. 2024. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*.

Achee NL, Grieco JP, Vatandoost H, Seixas G, Pinto J, Ching-Ng L, Martins AJ, Juntarajumnong W, Corbel V, Gouagna C, et al. 2019. Alternative strategies for mosquito-borne arbovirus control. *PLoS Negl Trop Dis* **13**: e0006822. https://www.ncbi.nlm.nih.gov/pubmed/30605475.

Alexander AJT, Salvemini M, Sreenu VB, Hughes J, Telleria EL, Ratinier M, Arnaud F, Volf P, Brennan B, Varjak M, et al. 2023. Characterisation of the antiviral RNA interference response to Toscana virus in sand fly cells. *PLoS Pathog* **19**: e1011283. https://www.ncbi.nlm.nih.gov/pubmed/36996243.

Alonso-Palomares LA, Moreno-Garcia M, Lanz-Mendoza H, Salazar MI. 2018. Molecular Basis for Arbovirus Transmission by Aedes aegypti Mosquitoes. *Intervirology* **61**: 255–264. https://www.ncbi.nlm.nih.gov/pubmed/31082816.

Anderson MA, Gross TL, Myles KM, Adelman ZN. 2010. Validation of novel promoter sequences derived from two endogenous ubiquitin genes in transgenic Aedes aegypti. *Insect Mol Biol* **19**: 441–449. https://www.ncbi.nlm.nih.gov/pubmed/20456509.

Antoniewski C. 2014. Computing siRNA and piRNA overlap signatures. *Methods Mol Biol* **1173**: 135–146. https://www.ncbi.nlm.nih.gov/pubmed/24920366.

Blair CD, Olson KE. 2015. The role of RNA interference (RNAi) in arbovirus-vector interactions. *Viruses* **7**: 820–843. https://www.ncbi.nlm.nih.gov/pubmed/25690800.

Brackney DE, Scott JC, Sagawa F, Woodward JE, Miller NA, Schilkey FD, Mudge J, Wilusz J, Olson KE, Blair CD, et al. 2010. C6/36 Aedes albopictus cells have a dysfunctional antiviral RNA interference response. *PLoS Negl Trop Dis* **4**: e856. https://www.ncbi.nlm.nih.gov/pubmed/21049065.

Brady OJ, Hay SI. 2020. The Global Expansion of Dengue: How Aedes aegypti Mosquitoes Enabled the First Pandemic Arbovirus. *Annu Rev Entomol* **65**: 191–208. https://www.ncbi.nlm.nih.gov/pubmed/31594415.

Bronkhorst AW, van Rij RP. 2014. The long and short of antiviral defense: small RNA-based immunity in insects. *Curr Opin Virol* **7**: 19–28. https://www.ncbi.nlm.nih.gov/pubmed/24732439.

Caragata EP, Dutra HLC, Sucupira PHF, Ferreira AGA, Moreira LA. 2021. Wolbachia as translational science: controlling mosquito-borne pathogens. *Trends Parasitol* **37**: 1050–1067. https://www.ncbi.nlm.nih.gov/pubmed/34303627.

Carvalho FD, Moreira LA. 2017. Why is Aedes aegypti Linnaeus so Successful as a Species? *Neotrop Entomol* **46**: 243–255. https://www.ncbi.nlm.nih.gov/pubmed/28401481.

Cenik ES, Fukunaga R, Lu G, Dutcher R, Wang Y, Tanaka Hall TM, Zamore PD. 2011. Phosphate and R2D2 restrict the substrate specificity of Dicer-2, an ATP-driven ribonuclease. *Mol Cell* **42**: 172–184. https://www.ncbi.nlm.nih.gov/pubmed/21419681.

Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**: i884–i890. https://www.ncbi.nlm.nih.gov/pubmed/30423086.

Ciechanowska K, Pokornowska M, Kurzynska-Kokorniak A. 2021. Genetic Insight into the Domain Structure and Functions of Dicer-Type Ribonucleases. *Int J Mol Sci* **22**. https://www.ncbi.nlm.nih.gov/pubmed/33435485.

Donelick HM, Talide L, Bellet M, Aruscavage J, Lauret E, Aguiar E, Marques JT, Meignin C, Bass BL. 2020. In vitro studies provide insight into effects of Dicer-2 helicase mutations in Drosophila melanogaster. *RNA*. https://www.ncbi.nlm.nih.gov/pubmed/32843367.

Elbashir SM, Martinez J, Patkaniowska A, Lendeckel W, Tuschl T. 2001. Functional anatomy of siRNAs for mediating efficient RNAi in Drosophila melanogaster embryo lysate. *EMBO J* **20**: 6877–6888. https://www.ncbi.nlm.nih.gov/pubmed/11726523.

589 Facchinelli L, Badolo A, McCall PJ. 2023. Biology and Behaviour of Aedes aegypti in the
590 Human Environment: Opportunities for Vector Control of Arbovirus Transmission.
591 *Viruses* **15**. https://www.ncbi.nlm.nih.gov/pubmed/36992346.
592 Fredericks AC, Russell TA, Wallace LE, Davidson AD, Fernandez-Sesma A, Maringer K.
593 2019. Aedes aegypti (Aag2)-derived clonal mosquito cell lines reveal the effects of pre-
594 existing persistent infection with the insect-specific bunyavirus Phasi Charoen-like virus
595 on arbovirus replication. *PLoS Negl Trop Dis* **13**: e0007346.
596 https://www.ncbi.nlm.nih.gov/pubmed/31693659.
597 Fukunaga R, Colpan C, Han BW, Zamore PD. 2014. Inorganic phosphate blocks binding of
598 pre-miRNA to Dicer-2 via its PAZ domain. *EMBO J* **33**: 371–384.
599 https://www.ncbi.nlm.nih.gov/pubmed/24488111.
600 Gestuveo RJ, Parry R, Dickson LB, Lequime S, Sreenu VB, Arnold MJ, Khromykh AA,
601 Schnettler E, Lambrechts L, Varjak M, et al. 2022. Mutational analysis of Aedes aegypti
602 Dicer 2 provides insights into the biogenesis of antiviral exogenous small interfering
603 RNAs. *PLoS Pathog* **18**: e1010202. https://www.ncbi.nlm.nih.gov/pubmed/34990484.
604 Ghildiyal M, Xu J, Seitz H, Weng Z, Zamore PD. 2010. Sorting of *Drosophila* small silencing
605 RNAs partitions microRNA* strands into the RNA interference pathway. *RNA* **16**: 43–
606 56.
607 Halstead SB. 2019. Travelling arboviruses: A historical perspective. *Travel Med Infect Dis*
608 **31**: 101471. https://www.ncbi.nlm.nih.gov/pubmed/31472285.
609 Huang YS, Higgs S, Vanlandingham DL. 2019. Emergence and re-emergence of mosquito-
610 borne arboviruses. *Curr Opin Virol* **34**: 104–109.
611 https://www.ncbi.nlm.nih.gov/pubmed/30743191.
612 Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K,
613 Bates R, Zidek A, Potapenko A, et al. 2021. Highly accurate protein structure prediction
614 with AlphaFold. *Nature* **596**: 583–589. https://www.ncbi.nlm.nih.gov/pubmed/34265844.
615 Kandasamy SK, Fukunaga R. 2016. Phosphate-binding pocket in Dicer-2 PAZ domain for
616 high-fidelity siRNA production. *Proc Natl Acad Sci U S A* **113**: 14031–14036.
617 https://www.ncbi.nlm.nih.gov/pubmed/27872309.
618 Lambrechts L, Saleh MC. 2019. Manipulating Mosquito Tolerance for Arbovirus Control. *Cell*
619 *Host Microbe* **26**: 309–313. https://www.ncbi.nlm.nih.gov/pubmed/31513769.
620 Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods*
621 **9**: 357–359. https://www.ncbi.nlm.nih.gov/pubmed/22388286.
622 Leggewie M, Schnettler E. 2018. RNAi-mediated antiviral immunity in insects and their
623 possible application. *Curr Opin Virol* **32**: 108–114.
624 https://www.ncbi.nlm.nih.gov/pubmed/30428425.
625 Marques JT, Wang JP, Wang X, de Oliveira KP, Gao C, Aguiar ER, Jafari N, Carthew RW.
626 2013. Functional specialization of the small interfering RNA pathway in response to
627 virus infection. *PLoS Pathog* **9**: e1003579.
628 https://www.ncbi.nlm.nih.gov/pubmed/24009507.
629 McFarlane M, Almire F, Kean J, Donald CL, McDonald A, Wee B, Laureti M, Varjak M, Terry
630 S, Vazeille M, et al. 2020. The Aedes aegypti Domino Ortholog p400 Regulates
631 Antiviral Exogenous Small Interfering RNA Pathway Activity and ago-2 Expression.
632 *mSphere* **5**. https://www.ncbi.nlm.nih.gov/pubmed/32269152.
633 Merkling SH, Crist AB, Henrion-Lacritick A, Frangeul L, Couderc E, Gausson V, Blanc H,
634 Bergman A, Baidaliuk A, Romoli O, et al. 2023. Multifaceted contributions of Dicer2 to
635 arbovirus transmission by Aedes aegypti. *Cell Rep* **42**: 112977.
636 https://www.ncbi.nlm.nih.gov/pubmed/37573505.
637 Mi S, Cai T, Hu Y, Chen Y, Hodges E, Ni F, Wu L, Li S, Zhou H, Long C, et al. 2008. Sorting
638 of Small RNAs into Arabidopsis Argonaute Complexes Is Directed by the 5′ Terminal
639 Nucleotide. *Cell* **133**: 116–127.
640 Olson KE, Blair CD. 2015. Arbovirus-mosquito interactions: RNAi pathway. *Curr Opin Virol*
641 **15**: 119–126. https://www.ncbi.nlm.nih.gov/pubmed/26629932.

642 Paturi S, Deshmukh M V. 2021. A Glimpse of “Dicer Biology” Through the Structural and
643 Functional Perspective. *Front Mol Biosci* **8**: 643657.
644 https://www.ncbi.nlm.nih.gov/pubmed/34026825.
645 Pettersen EF, Goddard TD, Huang CC, Meng EC, Couch GS, Croll TI, Morris JH, Ferrin TE.
646 2021. <scp>UCSF ChimeraX</scp> : Structure visualization for researchers, educators,
647 and developers. *Protein Science* **30**: 70–82.
648 Pierson TC, Diamond MS. 2020. The continued threat of emerging flaviviruses. *Nat*
649 *Microbiol* **5**: 796–812. https://www.ncbi.nlm.nih.gov/pubmed/32367055.
650 Prince BC, Walsh E, Torres TZB, Ruckert C. 2023. Recognition of Arboviruses by the
651 Mosquito Immune System. *Biomolecules* **13**.
652 https://www.ncbi.nlm.nih.gov/pubmed/37509194.
653 Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic
654 features. *Bioinformatics* **26**: 841–842. https://www.ncbi.nlm.nih.gov/pubmed/20110278.
655 Robinson MD, McCarthy DJ, Smyth GK. 2009. edgeR: a Bioconductor package for
656 differential expression analysis of digital gene expression data. *Bioinformatics* **26**: 139–
657 140. https://doi.org/10.1093/bioinformatics/btp616.
658 Samuel GH, Adelman ZN, Myles KM. 2018. Antiviral Immunity and Virus-Mediated
659 Antagonism in Disease Vector Mosquitoes. *Trends Microbiol* **26**: 447–461.
660 https://www.ncbi.nlm.nih.gov/pubmed/29395729.
661 Samuel GH, Pohlenz T, Dong Y, Coskun N, Adelman ZN, Dimopoulos G, Myles KM. 2023.
662 RNA interference is essential to modulating the pathogenesis of mosquito-borne
663 viruses in the yellow fever mosquito Aedes aegypti. *Proc Natl Acad Sci U S A* **120**:
664 e2213701120. https://www.ncbi.nlm.nih.gov/pubmed/36893279.
665 Scherer C, Knowles J, Sreenu VB, Fredericks AC, Fuss J, Maringer K, Fernandez-Sesma A,
666 Merits A, Varjak M, Kohl A, et al. 2021. An Aedes aegypti-aerived Ago2 knockout cell
667 line to investigate arbovirus infections. *Viruses* **13(6)**: 1066.
668 Scott JC, Brackney DE, Campbell CL, Bondu-Hawkins V, Hjelle B, Ebel GD, Olson KE, Blair
669 CD. 2010. Comparison of dengue virus type 2-specific small RNAs from RNA
670 interference-competent and -incompetent mosquito cells. *PLoS Negl Trop Dis* **4**: e848.
671 https://www.ncbi.nlm.nih.gov/pubmed/21049014.
672 Sinha NK, Iwasa J, Shen PS, Bass BL. 2018. Dicer uses distinct modules for recognizing
673 dsRNA termini. *Science (1979)* **359**: 329–334.
674 https://www.ncbi.nlm.nih.gov/pubmed/29269422.
675 Su S, Wang J, Deng T, Yuan X, He J, Liu N, Li X, Huang Y, Wang HW, Ma J. 2022.
676 Structural insights into dsRNA processing by Drosophila Dicer-2-Loqs-PD. *Nature* **607**:
677 399–406. https://www.ncbi.nlm.nih.gov/pubmed/35768513.
678 Swevers L, Liu J, Smagghe G. 2018. Defense Mechanisms against Viral Infection in
679 Drosophila: RNAi and Non-RNAi. *Viruses* **10**.
680 https://www.ncbi.nlm.nih.gov/pubmed/29723993.
681 Tikhe C V, Dimopoulos G. 2021. Mosquito antiviral immune pathways. *Dev Comp Immunol*
682 **116**: 103964. https://www.ncbi.nlm.nih.gov/pubmed/33301792.
683 Ulper L, Sarand I, Rausalu K, Merits A. 2008. Construction, properties, and potential
684 application of infectious plasmids containing Semliki Forest virus full-length cDNA with
685 an inserted intron. *J Virol Methods* **148**: 265–270.
686 https://www.ncbi.nlm.nih.gov/pubmed/18054090.
687 Varjak M, Leggewie M, Schnettler E. 2018. The antiviral piRNA response in mosquitoes? *J*
688 *Gen Virol* **99**: 1551–1562. https://www.ncbi.nlm.nih.gov/pubmed/30372405.
689 Varjak M, Maringer K, Watson M, Sreenu VB, Fredericks AC, Pondeville E, Donald CL,
690 Sterk J, Kean J, Vazeille M, et al. 2017. Aedes aegypti Piwi4 Is a Noncanonical PIWI
691 Protein Involved in Antiviral Responses. *mSphere* **2**.
692 https://www.ncbi.nlm.nih.gov/pubmed/28497119.
693 Wang GH, Du J, Chu CY, Madhav M, Hughes GL, Champer J. 2022. Symbionts and gene
694 drive: two strategies to combat vector-borne disease. *Trends Genet* **38**: 708–723.
695 https://www.ncbi.nlm.nih.gov/pubmed/35314082.

Wang GH, Gamez S, Raban RR, Marshall JM, Alphey L, Li M, Rasgon JL, Akbari OS. 2021. Combating mosquito-borne diseases using genetic control technologies. *Nat Commun* **12**: 4388. https://www.ncbi.nlm.nih.gov/pubmed/34282149.

Warf MB, Johnson WE, Bass BL. 2011. Improved annotation of *C. elegans* microRNAs by deep sequencing reveals structures associated with processing by Drosha and Dicer. *RNA* **17**: 563–577.

Weaver SC, Charlier C, Vasilakis N, Lecuit M. 2018. Zika, Chikungunya, and Other Emerging Vector-Borne Viral Diseases. *Annu Rev Med* **69**: 395–408. https://www.ncbi.nlm.nih.gov/pubmed/28846489.

Weaver SC, Reisen WK. 2010. Present and future arboviral threats. *Antiviral Res* **85**: 328–345. https://www.ncbi.nlm.nih.gov/pubmed/19857523.

Yamaguchi S, Naganuma M, Nishizawa T, Kusakizako T, Tomari Y, Nishimasu H, Nureki O. 2022. Structure of the Dicer-2-R2D2 heterodimer bound to a small RNA duplex. *Nature* **607**: 393–398. https://www.ncbi.nlm.nih.gov/pubmed/35768503.

Zaid A, Burt FJ, Liu X, Poo YS, Zandi K, Suhrbier A, Weaver SC, Texeira MM, Mahalingam S. 2021. Arthritogenic alphaviruses: epidemiological and clinical perspective on emerging arboviruses. *Lancet Infect Dis* **21**: e123–e133. https://www.ncbi.nlm.nih.gov/pubmed/33160445.

Zapletal D, Kubicek K, Svoboda P, Stefl R. 2023. Dicer structure and function: conserved and evolving features. *EMBO Rep* **24**: e57215. https://www.ncbi.nlm.nih.gov/pubmed/37310138.

Figure Legends

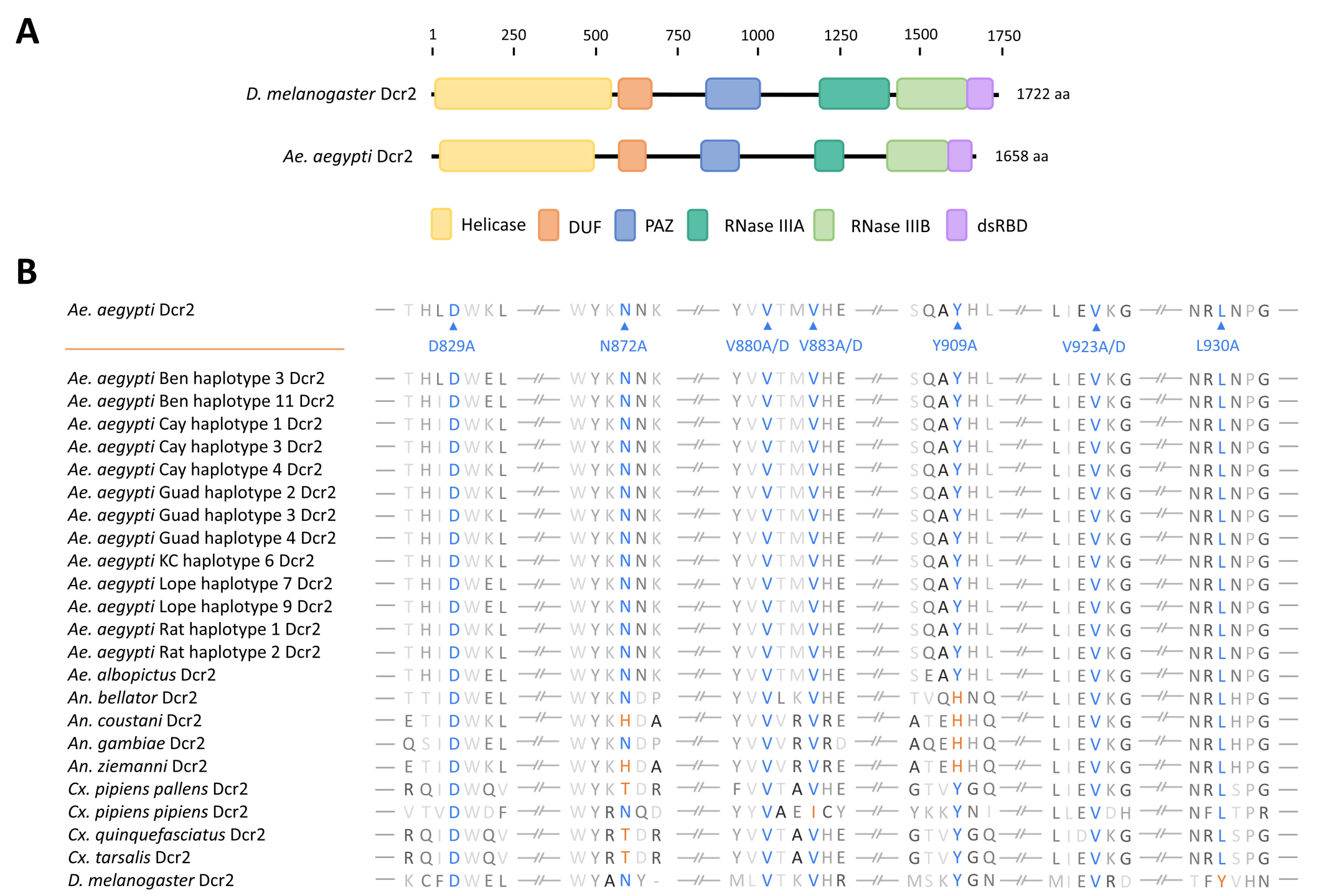
Figure 1. Domain features of Dcr2. (A) Schematic representation of *D. melanogaster* and *Ae. aegypti* Dcr2, including functional domains: Helicase domain, domain of unknown function (DUF), PIWI-Argonaute-Zwille (PAZ) domain, RNase IIIA and IIIB domains, and dsRNA binding domain (dsRBD); aa, amino acid **(B)** Multiple sequence alignment of insect Dcr2 referring to Aaeg Dcr2 AAW48725, including the selected mutations of highly conserved amino acids for potential PAZ domain loss of function. Accession numbers for sequences used are indicated in Supplemental Table S1. Alignment was produced using Benchling (https://www.benchling.com/).

Figure 2. Mutation and expression of Aaeg Dcr2. (A) Table showing produced Aaeg Dcr2 PAZ domain mutants M1-M4, the respective introduced loss-of-function mutations, as well as corresponding amino acid residues in Dm Dcr2; aa, amino acid. **(B)** Assessment of (myc-tagged) Aaeg Dcr2 expression by western blot analysis. AF319 cells were transfected with either pPUb plasmids expressing WT Dcr2 or Dcr2 PAZ domain mutants M1-M4, using pPUb-

734 myc-eGFP as control. Anti-myc and anti-α tubulin (control) antibodies were used.
735 Representative of three independent repeats (other repeats in Supplemental Fig. S2).
736
737 **Figure 3. Functional analysis of Aaeg Dcr2 PAZ domain mutants. (A)** Antiviral activity of
738 Aaeg Dcr2. AF319 cells were transfected with expression plasmids encoding myc-tagged
739 WT Dcr2 or Dcr2 PAZ domain mutants M1-M4 or eGFP (negative control). These cells were
740 infected with SFV-FFLuc (MOI = 0.1) 24 hours post-transfection (hpt). FFLuc-levels were
741 measured at 48 hours post-infection (hpi) and are shown as mean ± SEM relative light units
742 compared to eGFP control set to 1 from 3 independent repeats, performed in technical
743 triplicate with * indicating $p<0.05$ according to Student's t-test. **(B)** Silencing activity of
744 mutant Aaeg Dcr2. AF319 cells were co-transfected with (i) myc-tagged pPUb plasmids
745 expressing WT Dcr2 or Dcr2 PAZ domain mutants M1-M4 (with eGFP as control), (ii) FFLuc
746 and RLuc (internal control) reporter plasmids and (iii) dsRNA targeting FFLuc (FFLuc
747 dsRNA, blue) or eGFP (eGFP dsRNA; non-silencing control, grey). At 24 hpt, FFLuc and
748 RLuc levels were measured and are shown as mean ± SEM relative light units (FFLuc/RLuc)
749 normalized to dseGFP from 3 independent repeats, performed in technical triplicate, with * =
750 $p<0.05$, ** = $p<0.01$, *** = $p<0.001$ versus controls according to two-way ANOVA.
751
752 **Figure 4. Aaeg Dcr2 PAZ domain mutations affected 21 nt vsiRNA magnitude and**
753 **length. (A)** Small RNA sequencing of AF319 cells, transiently expressing Dcr2 PAZ domain
754 mutants M1-M4 or WT Dcr2, or eGFP (as negative control), and infected with SFV (MOI=1)
755 at 48 hpi. Histogram of small RNA reads 18-30 nt in length, mapped to the SFV (SFV4
756 GenBank ID: KP699763) genome (positive numbers) and antigenome (negative numbers)
757 with colours indicating first base nucleotide prevalence per size, shown as mean % mapped
758 reads (Y axis, percentage reads) from two independent experiments +/- SD. Mapping of
759 SFV-derived 21 nt vsiRNAs, mapped along the SFV genome (magenta) or anti-genome
760 (cyan) (Y axis, vsiRNA reads per million) +/- SD. **(B)** Fold change of magnitude of SFV-

761 derived small RNAs of lengths 20-22 nt of Dcr2 PAZ domain mutants M1-M4 compared to
762 WT Dcr2.
763
764 **Figure 5. Aaeg Dcr2 PAZ domain mutations result in dysregulated vsiRNA duplex**
765 **overlap lengths. (A)** Heat maps showing mean overlap probability z-scores of 18-30 nt
766 SFV-derived small RNAs from AF319 transiently expressing Dcr2 PAZ domain mutant M1-
767 M4 or WT Dcr2, or eGFP (control) from 2 independent repeats, as initially characterized in
768 Fig 4. Differing lengths of vsiRNAs, analysed individually, are shown horizontally, and nt
769 overlaps are listed vertically. The red arrow labelled Dcr2 indicates the expected 2 nt overlap
770 from dsRNA cleavage with cells boxed in black. The black arrow labeled pp (ping-pong)
771 shows the expected 10 nt overlap from potential ping-pong amplification in the piRNA
772 pathway. **(B)** Number of overlapping pairs per million mapped reads of 19-23 nt virus-
773 derived small RNAs from AF319 transiently expressing Dcr2 PAZ domain mutants M1-M4 or
774 WT Dcr2, or eGFP (control) infected with SFV. Data are shown as the mean of 2
775 independent repeats with the range of values.
776
777 **Figure 6. Mutations in the PAZ domain of Aaeg Dcr2 change the position and**
778 **magnitude of cleavage over the SFV genome and virus-derived small RNA**
779 **populations compared to WT Dcr2.** Samples from Fig 4 were analysed as follows. **(A)**
780 Scatter plots showing normalized differential coverage (d_i) for each position in the SFV
781 genome. Each panel compares WT Dcr2 with no Dcr2, an eGFP expressing Dcr2 negative
782 treatment, or Dcr2 PAZ domain mutants M1-M4. The linear regression line is depicted within
783 each plot, with the R^2 value indicating the fit of the model. The number of datapoints (n) with
784 $d_i>0$ over the SFV genome used for the calculation is shown on each graph. **(B)** Heatmap
785 displaying pairwise R^2 values from the linear regressions, comparing the similarity of
786 differential coverage profile between treatments, as indicated. **(C)** Multidimensional scaling
787 (MDS) plots representing the relationships between different size classes of vsiRNA reads
788 from WT Dcr2, eGFP control and Dcr2 PAZ domain mutants M1-M4. Each plot corresponds

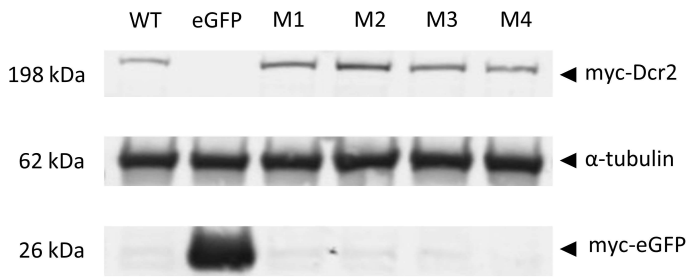
789 to a specific size class of SFV-derived small RNA: 20 nt, 21 nt, 22 nt, and 23 nt, as well as a
790 combined plot of 20-23 nt. The axes represent the leading two log-fold change (logFC)
791 dimensions and the percentage of variation indicated.
792

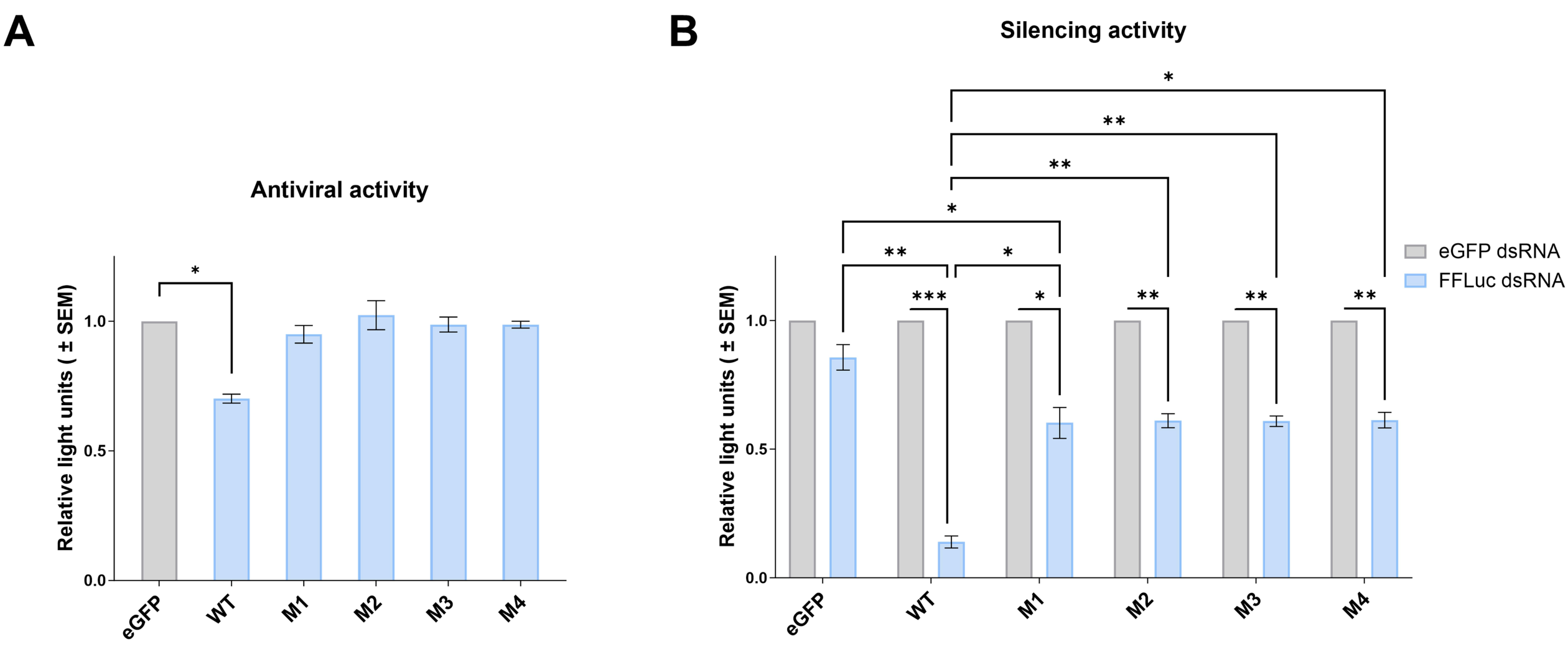

A
1
250
500
750
1000
1250
1500
1750
D. melanogaster Dcr2
1722 aa
Ae. aegypti Dcr2
1658 aa
Helicase
DUF
PAZ
RNase IIIA
RNase IIIB
dsRBD
B
Ae. aegypti Dcr2
D829A
N872A
V880A/D
V883A/D
Y909A
V923A/D
L930A
Ae. aegypti Ben haplotype 3 Dcr2
Ae. aegypti Ben haplotype 11 Dcr2
Ae. aegypti Cay haplotype 1 Dcr2
Ae. aegypti Cay haplotype 3 Dcr2
Ae. aegypti Cay haplotype 4 Dcr2
Ae. aegypti Guad haplotype 2 Dcr2
Ae. aegypti Guad haplotype 3 Dcr2
Ae. aegypti Guad haplotype 4 Dcr2
Ae. aegypti KC haplotype 6 Dcr2
Ae. aegypti Lope haplotype 7 Dcr2
Ae. aegypti Lope haplotype 9 Dcr2
Ae. aegypti Rat haplotype 1 Dcr2
Ae. aegypti Rat haplotype 2 Dcr2
Ae. albopictus Dcr2
An. bellator Dcr2
An. coustani Dcr2
An. gambiae Dcr2
An. ziemanni Dcr2
Cx. pipiens pallens Dcr2
Cx. pipiens pipiens Dcr2
Cx. quinquefasciatus Dcr2
Cx. tarsalis Dcr2
D. melanogaster Dcr2

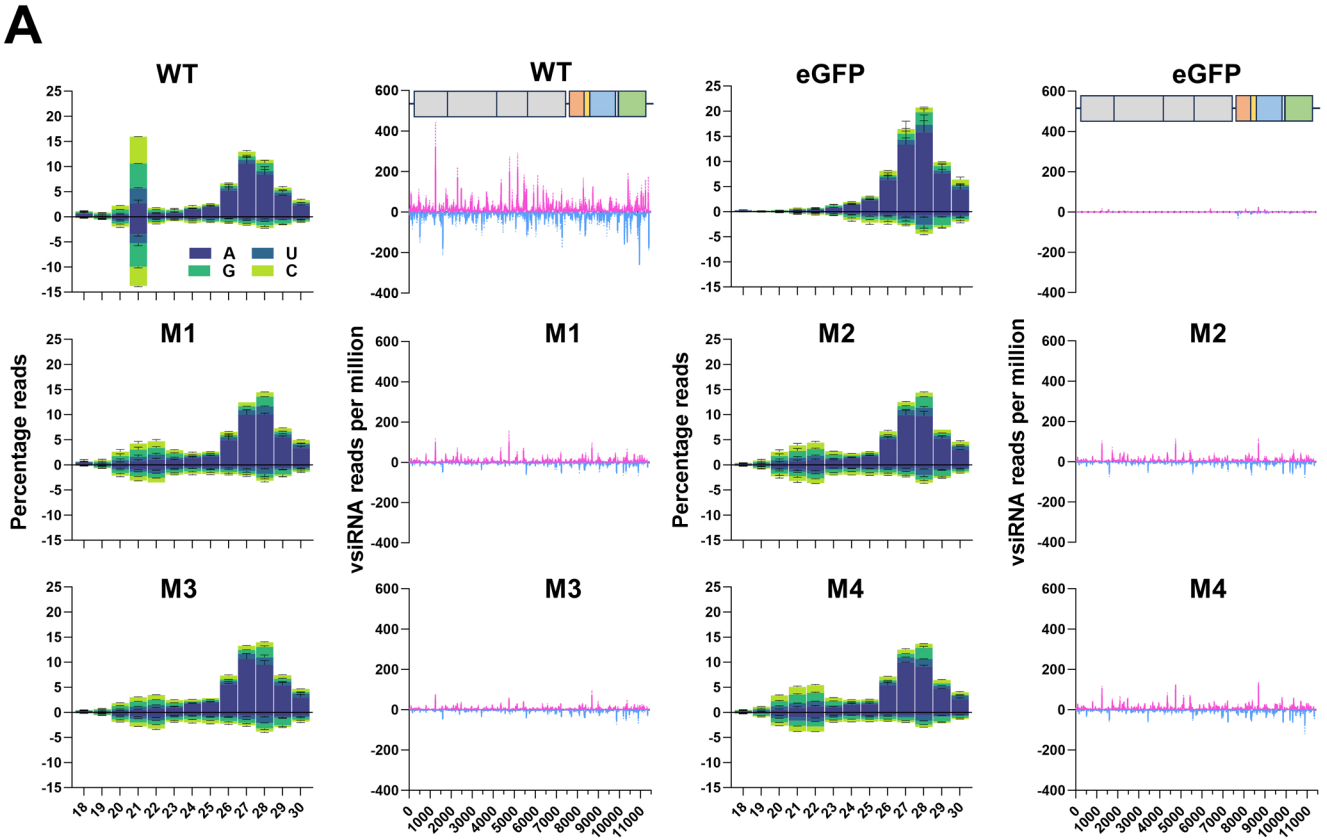
A

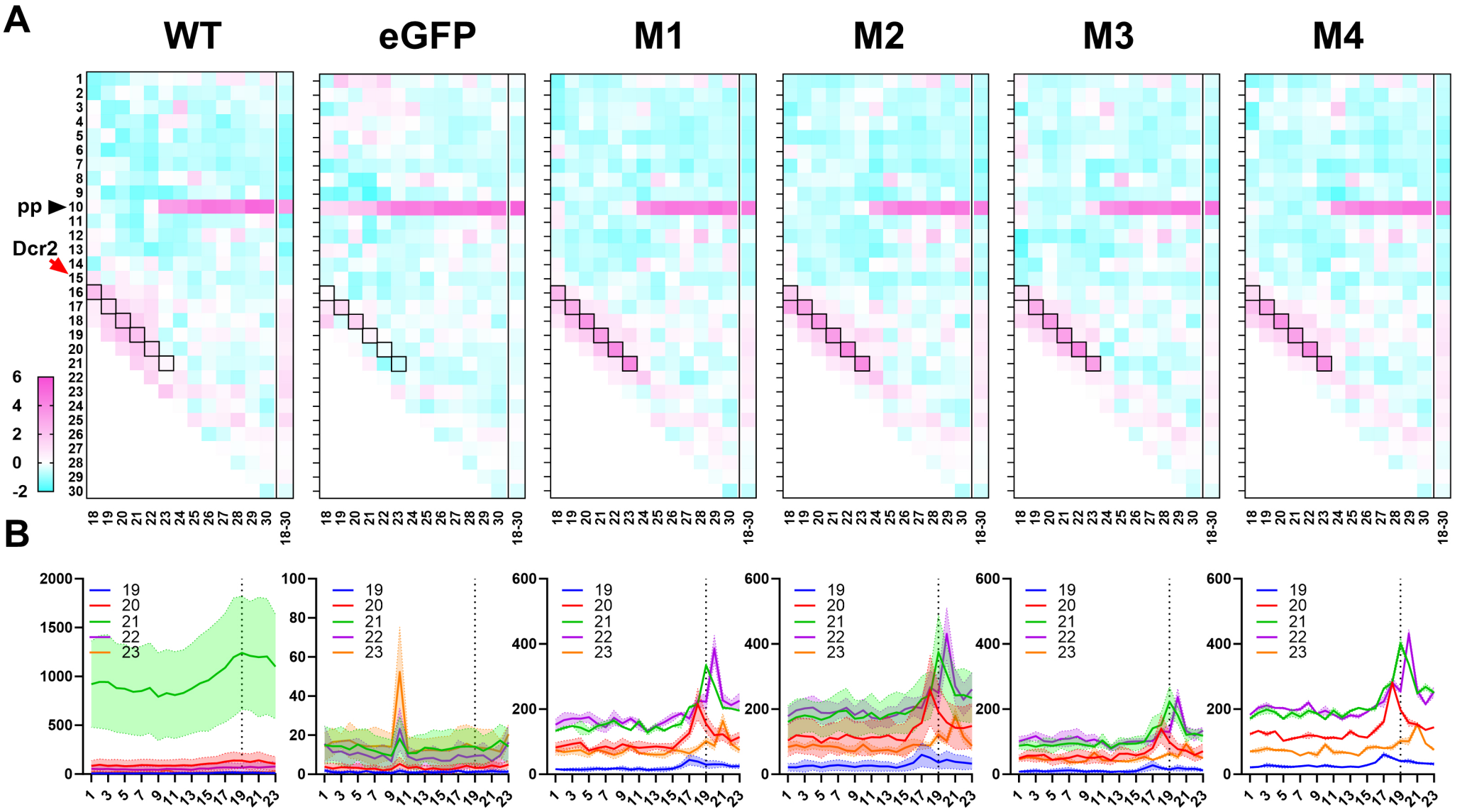
Dcr2 mutant	Dcr2 mutations *Ae. aegypti*	Corresponding aa residues *D. melanogaster*
M1	V880A	V895
	V883A	V898
	Y909A	Y925
	V923A	V942
M2	V880D	V895
	V883D	V898
	Y909A	Y925
	V923D	V942
M3	D829A	D845
	N872A	N888
	V880A	V895
	V883A	V898
	Y909A	Y925
	V923A	V942
	L930A	Y949
M4	D829A	D845
	N872A	N888
	V880D	V895
	V883D	V898
	Y909A	Y925
	V923D	V942
	L930A	Y949

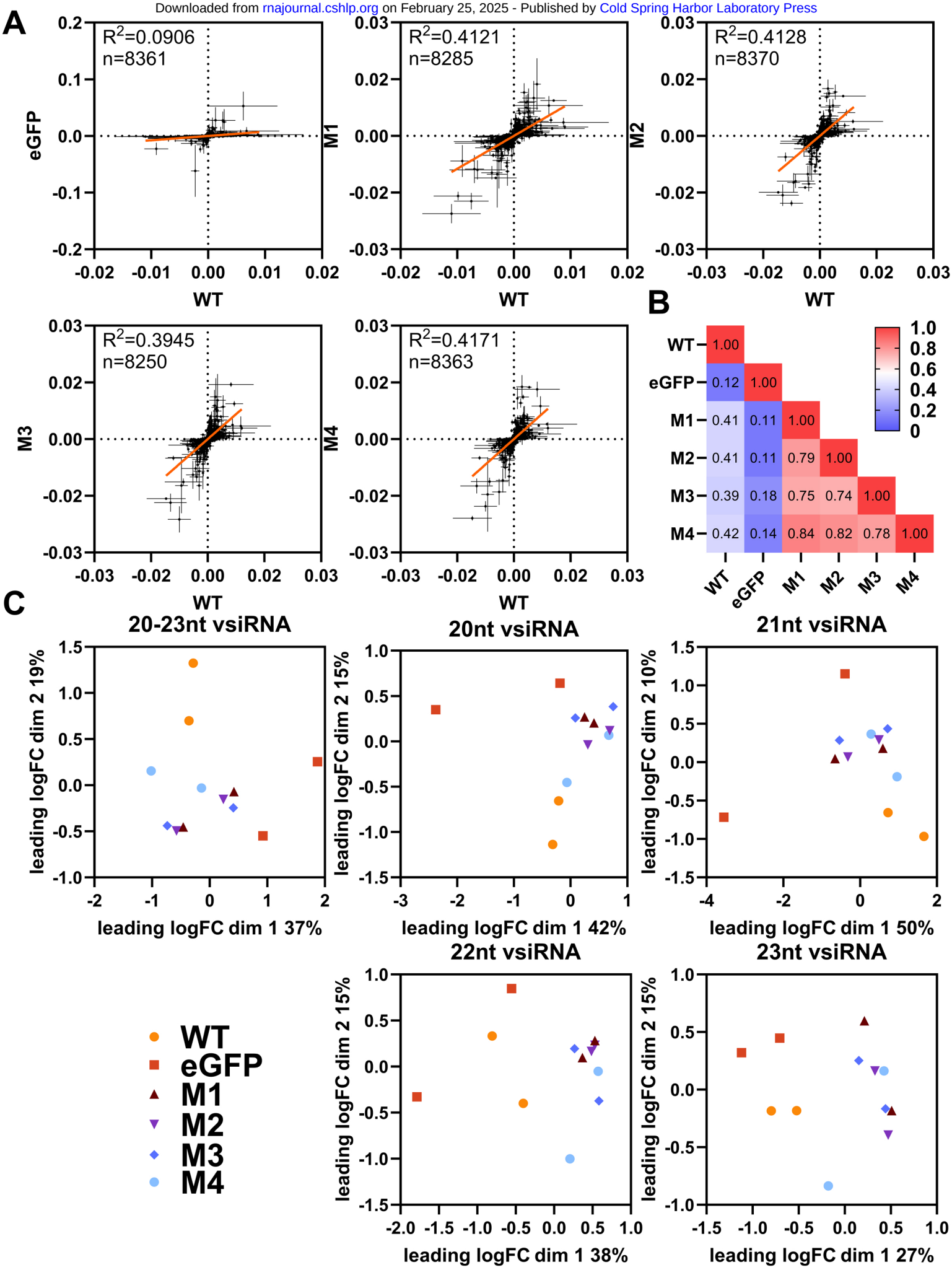
B




A
Antiviral activity
Relative light units (± SEM)
1.0
0.5
0.0
*
eGFP
WT
M1
M2
M3
M4
B
Silencing activity
Relative light units (± SEM)
1.0
0.5
0.0
*
**
**
*
**
*
***
*
**
**
**
eGFP
WT
M1
M2
M3
M4
eGFP dsRNA
FFLuc dsRNA


A
WT
WT
eGFP
eGFP
M1
M1
M2
M2
M3
M3
M4
M4
A
U
G
C
Percentage reads
vsiRNA reads per million
B
20
21
22
1.00 1.16 1.18 0.91 1.45
1.00 0.25 0.24 0.20 0.30
1.00 2.66 2.62 2.18 3.02
3
2
1
0
WT
M1
M2
M3
M4


A
WT
eGFP
M1
M2
M3
M4
pp
Dcr2
6
4
2
0
-2
18 19 20 21 22 23 24 25 26 27 28 29 30 18-30
B
19
20
21
22
23
2000
1500
1000
500
0
100
80
60
40
20
600
400
200
1 3 5 7 9 11 13 15 17 19 21 23



The PAZ domain of Aedes aegypti Dicer 2 is critical for accurate and high-fidelity size determination of virus-derived small interfering RNAs.

Melinda Reuter, Rhys H. Parry, Melanie McDonald, et al.

RNA published online February 13, 2025