

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25

## **An ancient satellite repeat controls gene expression and embryonic development in *Aedes aegypti* through a highly conserved piRNA**

**Rebecca Halbach<sup>1</sup>, Pascal Miesen<sup>1</sup>, Joep Joosten<sup>1</sup>, Ezgi Taşköprü<sup>1</sup>, Bas Pennings<sup>1</sup>, Chantal B.F. Vogels<sup>2,3</sup>, Sarah H. Merklings<sup>4,5</sup>, Constantianus J. Koenraadt<sup>2</sup>, Louis Lambrechts<sup>4,5</sup>, and Ronald P. van Rij<sup>1</sup>**

<sup>1</sup> Department of Medical Microbiology, Radboud University Medical Center, Radboud Institute for Molecular Life Sciences, P.O. Box 9101, 6500 HB Nijmegen, the Netherlands

<sup>2</sup> Laboratory of Entomology, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, the Netherlands

<sup>3</sup> Current affiliation: Department of Epidemiology of Microbial Diseases, Yale School of Public Health, 60 College Street, New Haven, CT 06510, USA

<sup>4</sup> Insect-Virus Interactions Group, Department of Genomes and Genetics, Institut Pasteur, 75015 Paris, France

<sup>5</sup> Evolutionary Genomics, Modeling and Health, Unité Mixte de Recherche 2000, Centre National de la Recherche Scientifique, 75015 Paris, France

**Correspondence:** [ronald.vanrij@radboudumc.nl](mailto:ronald.vanrij@radboudumc.nl)

## 26 **Abstract**

27 Tandem repeat elements such as the highly diverse class of satellite repeats occupy large parts of  
28 eukaryotic chromosomes. Most occur at (peri)centromeric and (sub)telomeric regions and have  
29 been implicated in chromosome organization, stabilization, and segregation<sup>1</sup>. Others are located  
30 more dispersed throughout the genome, but their functions remained largely enigmatic. Satellite  
31 repeats in euchromatic regions were hypothesized to regulate gene expression *in cis* by modulation  
32 of the local heterochromatin, or *in trans* via repeat-derived transcripts<sup>2,3</sup>. Yet, due to a lack of  
33 experimental models, gene regulatory potential of satellite repeats remains largely unexplored. Here  
34 we show that, in the vector mosquito *Aedes aegypti*, a satellite repeat promotes sequence-specific  
35 gene silencing via the expression of two abundant PIWI-interacting RNAs (piRNAs). Strikingly,  
36 whereas satellite repeats and piRNA sequences generally evolve extremely fast<sup>4-6</sup>, this locus was  
37 conserved for approximately 200 million years, suggesting a central function in mosquito biology.  
38 Tandem repeat-derived piRNA production commenced shortly after egg-laying and inactivation of  
39 the most abundant of the two piRNAs in early embryos resulted in an arrest of embryonic  
40 development. Transcriptional profiling in these embryos revealed the failure to degrade maternally  
41 provided transcripts that are normally cleared during maternal-to-zygotic transition. Our results  
42 reveal a novel mechanism in which satellite repeats regulate global gene expression *in trans* via  
43 piRNA-mediated gene silencing, which is fundamental to embryonic development. These findings  
44 highlight the regulatory potential of this enigmatic class of repeats.

45

## 46 **Main**

47 Even though satellite repeats have been discovered nearly 60 years ago<sup>7,8</sup>, and comprise a  
48 substantial portion of eukaryotic genomes, little is known about the functions of this class of  
49 repetitive DNA. Many satellite repeats are actively transcribed, and some of them produce small  
50 interfering (si)RNAs required for the establishment and maintenance of heterochromatic regions<sup>9-16</sup>.  
51 Around two-thirds of the genome of *Aedes aegypti*, the most important vector for arthropod-borne  
52 viruses like dengue, Zika, and yellow fever virus, consists of repetitive elements<sup>17</sup> (Extended Data  
53 Fig 1A), making this mosquito an interesting model to study these sequences. We analyzed small  
54 RNAs derived from unique and repetitive sequences in the genome of *Ae. aegypti* somatic and  
55 germline tissues as well as Aag2 cells. Even though satellite repeats constitute less than 10% of the  
56 genome, they were not only highly covered by siRNAs (Extended Data Fig 1A), but especially by  
57 PIWI-interacting (pi)RNAs (Extended Data Fig 1A). piRNAs are a class of small RNAs that protect  
58 animal genomes from harmful parasitic elements like transposons<sup>18</sup>. In the fruit fly *Drosophila*  
59 *melanogaster*, piRNAs are mostly derived from transposon-rich genomic regions termed piRNA  
60 clusters<sup>19</sup>. Yet, in *Ae. aegypti*, piRNAs from transposable elements (TEs) are underrepresented

61 compared to their abundance in the genome<sup>20</sup>, especially in the soma (Extended Data Fig 1A), but  
62 instead, we found satellite-repeat derived piRNAs to be highly overrepresented in somatic tissues.  
63 Intriguingly, approximately three-quarters of these reads in the soma, and half of the reads in the  
64 germline or Aag2 cells represent only two individual sequences that map to a repeat locus on  
65 chromosome 3. This locus was about 3.5 kb in size and consisted of 20 full and one disrupted repeat  
66 unit organized in a head-to-tail array (Fig 1A, B). These two highly abundant satellite-derived small  
67 RNAs were 30 and 29 nucleotides in size, respectively (Extended Data Fig 1B), and resistant to  $\beta$ -  
68 elimination, suggesting that they are 2'-O-methylated at their 3' end, a common feature of mature  
69 PIWI-bound piRNAs<sup>21-23</sup> (Fig 1C). We named these two sequences tapiR1 and 2 (tandem repeat-  
70 associated piRNA1/2). Expression of tapiR1 and 2 was ubiquitous in both somatic and germline  
71 tissues of adult mosquitoes (Extended Data Fig 2A). In *Ae. aegypti*, the PIWI-interacting RNA  
72 pathway has expanded to include seven PIWI genes (Piwi2-7 and Ago3) compared to three in  
73 flies<sup>24</sup>. Immunoprecipitation (IP) in Aag2 cells of the aedine PIWI proteins that are expressed both  
74 in the soma and gonads (Piwi4-6 and Ago3) followed by northern blotting or deep sequencing  
75 indicates that both tapiR1 and tapiR2 exclusively associate with Piwi4 (Fig 1D, Extended Data Fig  
76 2B, C, Supplementary Fig S1A,B). Indeed, only knockdown of Piwi4, but not of other PIWI or  
77 AGO-clade genes reduced tapiR1 and 2 levels (Extended Figure 2D, E, Supplementary Fig S1C,D).  
78 Thus far, the piRNA repertoire and function of Piwi4 remained unclear. Piwi4 neither associates  
79 with TE nor virus-derived piRNAs<sup>25</sup>, yet was linked to piRNA biogenesis from transposons<sup>25</sup> and to  
80 antiviral defense<sup>26,27</sup>. As nearly 90 % of Piwi4-associated small RNAs only comprise tapiR1, and,  
81 to a much lower extent, tapiR2 (Fig 1D), we hypothesize that tapiR1 not only dominates the piRNA  
82 repertoire, but also shapes downstream functions of Piwi4.

83 Satellite repeats are one of the fastest evolving parts of eukaryotic genomes. Except for a few  
84 examples<sup>28-30</sup>, most satellite repeats display high sequence divergence between species, and can  
85 even be species-specific<sup>6,31,32</sup>, akin to piRNAs<sup>4,5</sup>. Hence, we were surprised to find that the  
86 identified tandem repeat locus in *Ae. aegypti* is conserved in the closely related Asian tiger  
87 mosquito *Ae. albopictus*, and even in the more distantly related southern house mosquito *Culex*  
88 *quinquefasciatus* (Extended Data Fig 3A). This locus is, however, not present in the genome  
89 assembly of the malaria vector *Anopheles gambiae*. The tandem repeat locus differed in the number  
90 of monomers across species, and the monomers exhibited substantial length and sequence  
91 divergence, both between species and between monomers within one species. However, the parts of  
92 the monomer that give rise to tapiR1 and 2 were by far more conserved than the overall monomer,  
93 suggesting that these sequences are under extensive selective constraints (Fig 1E). We further  
94 analyzed whether expression of the two repeat locus-derived piRNAs is conserved among different  
95 mosquitoes, including species for which no genome assembly is available. We analyzed 17 different

96 mosquito species from 5 different genera (*Aedes*, *Culex*, *Culiseta*, *Coquillettidia* and *Anopheles*), as  
97 well as *Culicoides nubeculosus*, a biting midge that also transmits arboviruses, but is only distantly  
98 related to mosquitoes, and the fruitfly *Drosophila melanogaster*. Strikingly, even though piRNAs  
99 are usually not conserved even between closely related species<sup>4,5</sup>, we detected both tapiR1 and 2 in  
100 four genera of the Culicinae subfamily of mosquitoes (Fig 1F, Extended Data Fig 3B). In line with  
101 the absence of the repeat locus in the *Anopheles gambiae* genome, we did not observe tapiR1 or 2  
102 expression in this subfamily of mosquitoes, nor in the two non-mosquito species. This observation  
103 suggests that the locus evolved in the late Triassic after divergence of the Anophelinae from the  
104 Culicinae subfamily of mosquitoes (229-192 mya<sup>33</sup>), but before further divergence of the culicine  
105 genera (226-172 mya<sup>33</sup>). This establishes this repeat locus as one of the very few ancient and deeply  
106 conserved satellite repeats that have hitherto been described<sup>28-30,34</sup>. Conservation of the locus over  
107 million years of mosquito evolution strongly suggests important and conserved functions for the  
108 locus and its associated piRNAs.

109 The satellite repeat locus overlaps with the 3'UTR of the gene AAEL017385 (LOC23687805) of  
110 unknown function in the current genome annotation (Fig 1A). This organization suggests that one  
111 or more splice variants of AAEL017385 are the source of the piRNAs, and that expression of this  
112 gene and piRNA biogenesis might be closely linked. However, knockdown of the different splice  
113 isoforms did not reduce expression of tapiR1 (Extended Data Fig 3C, Supplementary Fig S1E),  
114 arguing against this possibility. Rapid amplification of 3' ends (RACE) of AAEL017385 transcripts  
115 revealed transcription termination sites directly upstream of the first tapiR1 and tapiR2 repeat,  
116 respectively (Extended Data Fig 3D). Even though we cannot exclude that some AAEL017385  
117 transcripts overlap with the satellite repeat, our data strongly suggest that the two piRNAs and  
118 AAEL017385 are not expressed from the same transcriptional unit. Instead, the satellite repeat  
119 locus might be transcribed from an unknown upstream or internal promoter. In support of this  
120 notion, the repeat locus is not associated with the AAEL017385 orthologues in *Ae. albopictus*  
121 (AALF011179) and *Cx. quinquefasciatus* (CPIJ011773).

122 We next characterized the sequence requirements for target silencing by tapiR1, as its expression is  
123 approximately one log higher compared to tapiR2 in Aag2 cells. Using a luciferase reporter  
124 harbouring a fully complementary target site in the 3' UTR, we validated that this piRNA is able to  
125 target RNAs *in trans*. The reporter was silenced more than 10-, or 35-fold compared to a reporter  
126 without target site, or a control reporter with a partially inverted target site, respectively (Fig 2A).  
127 Addition of an antisense oligonucleotide (AO) complementary to tapiR1 relieved silencing in a  
128 concentration-dependent manner (Extended Data Fig 4), confirming that the observed effect is  
129 mediated by the piRNA in a sequence-specific fashion. Unlike most miRNAs<sup>35</sup>, silencing was not  
130 dependent on the position of the target site in the mRNA and was efficient in the open reading

131 frame and both the 5' and the 3' UTR (Fig 2B). During the course of our study we noticed that the  
132 firefly and *Renilla* luciferase genes, which we used as reporter and normalization control,  
133 respectively, contain potential target sites for tapiR1. We confirmed that *Renilla* luciferase is indeed  
134 potently suppressed by tapiR1 and firefly luciferase slightly (Extended Data Fig 5A-C), yet,  
135 mutating these target sites did not affect any of the conclusions reported below, but increased the  
136 observed effects (Extended Data Fig 5D,E).

137 To assess targeting requirements for tapiR1, we introduced mismatches in the piRNA target sites.  
138 Three consecutive mismatches were tolerated unless they were located in the t1 to t9 region of the  
139 piRNA (the nucleotides base-paired to piRNA positions 1 to 9) (Fig 2C, Extended Data Fig 6A),  
140 and single mismatches only impaired silencing at positions t3 to t7 (Extended data Fig 6A,B),  
141 reminiscent of a microRNA seed<sup>35</sup> and comparable to what has been termed the piRNA seed in *C.*  
142 *elegans*<sup>36,37</sup>. Even though tapiR1 targeting requirements resemble those of miRNAs, the results are  
143 unlikely to be due to the piRNA being funnelled into the miRNA pathway. First, tapiR1 biogenesis  
144 is not dependent on Ago1 (Extended Data Fig 2E), and secondly, this piRNA is 2'-O-methylated  
145 (Fig 1C), a feature of siRNAs and mature PIWI-bound piRNAs, but not Ago1-associated miRNAs.  
146 A mismatch at position t1 did not alter silencing, suggesting that the first nucleotide is anchored in a  
147 binding pocket of Piwi4, similar to other Argonaute proteins<sup>38-40</sup>. Unexpectedly, a mismatch at  
148 position t2 was tolerated as well. This was, however, only the case when the rest of the target site  
149 was perfectly complementary, but not when the target contained mismatches outside of the seed  
150 (Extended Data Fig 6C). We further noticed that, in contrast to *C. elegans*<sup>36</sup>, G:U wobble pairs were  
151 not tolerated inside the seed and had the same effect as a mismatch at the same position (Extended  
152 Data Fig 6D). Whereas the 5' seed region is normally absolutely required for targeting<sup>36-38,41,42</sup>, the  
153 3' part of the piRNA might increase specificity and efficiency of the targeting. For this reason we  
154 further assessed the extent of complementarity needed to allow for tapiR1-mediated silencing.  
155 Introduction of increasing numbers of mismatches at the 3' end did not interfere with silencing  
156 when at least half of the piRNA could base pair with the target site (Fig 2D), indicating that the  
157 3' part of the piRNA is not necessarily required, yet the seed region alone not sufficient for  
158 silencing.

159 Taken together, these results suggest that tapiR1 needs relatively low sequence complementarity to  
160 efficiently silence targets<sup>36,41-43</sup>, and that there are no constraints regarding the position of the target  
161 site on the mRNA. As a consequence, tapiR1 may target a plethora of different cellular RNAs that  
162 are perfectly base pairing to the seed and additional matches outside the seed.

163 Considering the fact that the repeat locus is extremely conserved, we hypothesized that this piRNA  
164 regulates cellular gene expression. Some satellite repeats can influence genes by modulation of the  
165 local chromatin environment *in cis*<sup>14,44</sup>, or have been hypothesized to induce silencing of genes

166 with homologous repeat insertions<sup>29</sup>. In contrast, the *tapiR1/2* locus has the potential to silence  
167 expression of a broad range of remote genes *in trans*, independent of repeat insertions in target  
168 genes, and thus, to regulate diverse and highly complex cellular processes.

169 To test this idea, we blocked *tapiR1*-mediated silencing with the *tapiR1* AO in *Aag2* cells and  
170 assessed global gene expression by RNAseq two days after treatment. Intriguingly, expression of  
171 134 genes, amongst which many long non-coding RNAs, was significantly increased up to around  
172 450 fold compared to the treatment with a control oligonucleotide (Fig 2E, Supplementary Table 1).  
173 Transposons were not globally affected, although some elements were up-regulated as well, up to  
174 around 850-fold (Extended Data Fig 7A, Supplementary Table 2). Expression of deregulated genes,  
175 and a transposable element was increased in a concentration-dependent manner upon *tapiR1* AO  
176 treatment as measured by RT-qPCR (Fig 2F), validating our RNAseq results. We then used  
177 RNAHybrid to predict *tapiR1* target sites and verified these sequences in luciferase reporter assays.  
178 Twelve out of 23 sites in protein-coding genes, lncRNAs, or transposable elements were indeed  
179 sufficient to support suppression of the reporter (Extended Data Fig 8), strongly suggesting that  
180 *tapiR1* directly represses these cellular RNAs, and confirming that *tapiR1* only needs limited base  
181 pairing to mediate silencing. Computational prediction of target sites has inherent limitations, and  
182 not all genes with predicted target sites were differentially regulated upon AO treatment, or, *vice*  
183 *versa*, target sites were functional in a reporter context, yet the gene itself was not differentially  
184 expressed (Extended Data Fig 7B,C). Additionally, the minimum free energy of the piRNA/target  
185 duplex was not predictive for the effect size of *tapiR1*-mediated silencing. Thus, similar to  
186 miRNAs<sup>45</sup>, other factors beyond Watson-Crick base pairing seem to play a role in definition of  
187 *bona fide* target genes for *tapiR1*. Nevertheless, our results indicate that *tapiR1* is able to directly  
188 and strongly regulate gene expression and transposon RNA levels in a sequence-dependent manner.

189 The expression of satellite repeats is often regulated in a developmental or stage-specific  
190 manner<sup>29,46</sup>, thus we analysed the expression pattern of *tapiR1* and 2 throughout the mosquito life  
191 cycle. *tapiR1* and 2 were not expressed during the first three hours of embryonic development of  
192 *Ae. aegypti* (Fig 3A), but could be detected in all subsequent life stages (Extended Data Fig 9A,B).  
193 At the very beginning of embryonic development the zygotic genome is transcriptionally quiescent  
194 and the first mitotic divisions are exclusively driven by maternally deposited transcripts and  
195 proteins. Maternal-to-zygotic transition (MZT) is marked by the degradation of these maternal  
196 transcripts and concomitant zygotic genome activation<sup>47</sup>. Destabilization of maternal transcripts  
197 initially occurs through maternal decay activities, and later, after onset of zygotic transcription also  
198 by zygotic components<sup>47,48</sup>. One of the best described mechanisms involves zygotically expressed  
199 miRNAs, for example miR-430, miR-427, and the miR-309 cluster in zebrafish<sup>49</sup>, *Xenopus*<sup>50</sup>, and  
200 flies<sup>51</sup>, respectively. Based on its expression pattern and its strong suppressive ability we

201 hypothesized that tapiR1 could be part of the zygotic degradation pathway in mosquitoes, and  
202 necessary for embryonic development. To test this idea we injected either a tapiR1-specific AO or  
203 control AO into early pre-blastoderm *Ae. aegypti* embryos (before zygotic genome activation and  
204 expression of tapiR1) (Fig 3B), and assessed their development using discrete scoring schemes  
205 (Supplementary Fig 1F). Strikingly, more than 90 % of all tapiR1 AO-injected embryos were  
206 arrested early in development, whereas about half of all control embryos showed obvious signs of  
207 developmental progression (Fig 3C). In accordance, only a small fraction of tapiR1 AO-injected  
208 embryos hatched (Fig 3D) and continued to develop as larvae, suggesting that tapiR1-deficiency  
209 impedes development. RNA sequencing from tapiR1 AO or control AO-injected embryos 20.5 h  
210 after injection revealed massive deregulation of cellular transcripts (Fig 3E, Extended Data Fig 9C,  
211 Supplementary Table 3, 4). Expression of 205 genes, among which 44 lncRNAs, as well as few  
212 transposable elements was increased up to around 1000 and 500 fold in tapiR1 AO-treated embryos,  
213 respectively. Target genes with predicted tapiR1 target sites were more strongly up-regulated than  
214 genes without predicted site, as indicated by a shift of the cumulative distribution of RNA fold  
215 changes (Fig 3F). These findings show that tapiR1 controls regulatory circuits by direct gene  
216 targeting also *in vivo*, and that this function is essential for embryonic development, likely by  
217 promoting mRNA turnover of a subset of maternal transcripts. In line with this conclusion,  
218 confirmed target genes are down-regulated after the onset of tapiR1 expression (Fig 3G), and  
219 tapiR1 targets are overrepresented in transcripts that are maternally provided and degraded during  
220 MZT (Fig 3H).

221 piRNAs have been shown to promote degradation of *nanos*<sup>52</sup> and other transcripts involved in germ  
222 cell development<sup>53</sup>. Yet, this was dependent on transposon-derived piRNAs, and rather depicts a re-  
223 purposing of the existing piRNA pool, which is, due to its large targeting potential, ideal to be used  
224 to degrade a large number of transcripts. In contrast, we propose that, analogous to abundant  
225 miRNAs in other animals<sup>49-51</sup>, Culicinae mosquitoes have evolved a specific piRNA to destabilize a  
226 defined set of maternally deposited transcripts in early embryonic development. To our knowledge,  
227 this is the first example of sequence-specific gene silencing by transcriptional products from a  
228 satellite repeat *in trans*, and underlines the regulatory potential of tandemly repeated DNA.

229

## 230 **Methods**

### 231 **Cell culture**

232 *Aedes aegypti* Aag2 cells were cultured in Leibovitz's L-15 medium (Invitrogen) supplemented with  
233 10 % heat-inactivated Fetal Bovine Serum (PAA Laboratories), 2 % Tryptose Phosphate Broth  
234 Solution (Sigma Aldrich), 1x MEM Non-Essential Amino Acids (Invitrogen) and 50 U/ml  
235 penicillin/ streptomycin (Invitrogen) at 25 °C.

236

### 237 **Mosquito rearing**

238 Injection and northern blot of embryos was performed using a Cell fusion agent virus-free,  
239 isofemale *Aedes aegypti* strain called Jane. This strain was initiated from a field population  
240 originally sampled in the Muang District of Kamphaeng Phet Province, Thailand<sup>54</sup>, and reared for  
241 26 generations at 28±1 °C, 75±5 % relative humidity, 12:12 hour light-dark cycle. Embryos were  
242 hatched under low pressure for 30-60 min. Larvae were grown in dechlorinated tap water and fed  
243 fish food powder (Tetramin) every two days. Adults were maintained in cages with constant access  
244 to a 10% sucrose solution. Female mosquitoes were fed on commercial rabbit blood (BCL) through  
245 a membrane feeding system (Hemotek Ltd.) using pig intestine as membrane. For AO injections,  
246 female mosquitoes were transferred to 25 °C and 70 % humidity for at least two days before forced  
247 to lay eggs, and embryos were then placed back to 28 °C immediately after the injection. For the  
248 time-course experiment in Fig. 3A,G, embryos were kept at 25 °C during the course of the  
249 experiment.

250 All other *in vivo* experiments were performed with the *Ae. aegypti* Rockefeller strain, obtained from  
251 Bayer AG, Monheim, Germany. The mosquitoes were maintained at 27±1 °C with 12:12 hour  
252 light:dark cycle and 70% relative humidity, as described before<sup>55</sup>.

253 Mosquitoes used in Fig 1F were either different laboratory-reared, or wild-caught species: *Aedes*  
254 *aegypti* Liverpool strain, *Culex pipiens*, *Anopheles coluzzii*, *An. quadriannulatus*, *An. stephensi*  
255 mosquitoes, *Culicoides nubeculosus* biting midges, and *D. melanogaster* w<sup>1118</sup> flies were laboratory  
256 strains. The mosquitoes were deep-frozen and stored at -80 °C until use. *Ae. albopictus*, *Ae.*  
257 *cantans*, *Ae. intrudens*, *Ae. pullatus*, *Ae. cinereus*, *Ae. vexans*, *Cx. pusillus*, *Culiseta morsitans*,  
258 *Coquillettidia richiardii*, *An. maculipennis*, *An. claviger*, and *An. coluzzii* were wild-caught  
259 individuals collected in different regions in Italy, Sweden, or the Netherlands between July 2014  
260 and June 2015<sup>56</sup>. Species were identified at the species level, and stored at -20 °C for a maximum of  
261 two years.

262

263

264



## 265 **Gene knockdown**

266 Double-stranded RNA was generated by *in vitro* transcription of T7 promoter-flanked PCR  
267 products with T7 RNA polymerase. Primer sequences are given in Supplementary Table S5. The  
268 reaction was carried out at 37°C for 3 to 4 h, then heated to 80 °C for 10 min and gradually cooled  
269 down to room temperature to facilitate dsRNA formation. The dsRNA was purified with the  
270 GeneElute Total RNA Miniprep Kit (Sigma Aldrich).  
271 Aag2 cells were seeded in 24-well plates the day before the experiment, and transfected with X-  
272 tremeGENE HP Transfection reagent (Roche) according to the manufacturer's instructions, using a  
273 ratio of 4 µL reagent per µg of dsRNA. The transfection medium was replaced after 3 h with fully  
274 supplemented Leibovitz-15 medium and RNA was harvested 48 h later. Knockdown was confirmed  
275 by RT-qPCR.

276

## 277 **RNA isolation**

278 RNA from cells and mosquitoes was isolated with Isol-RNA lysis buffer (5PRIME) according to  
279 the manufacturer's instructions. Briefly, 200 µL chloroform was added to 1 mL lysis buffer, and  
280 centrifuged at 16,060 x g for 20 min at 4 °C. Isopropanol was added to the aqueous phase, followed  
281 by incubation on ice for at least one hour, and centrifugation at 16,060 x g for 10 min at 4 °C. The  
282 pellet was washed three to five times with 85 % ethanol and dissolved in RNase free water. RNA  
283 was quantified on a Nanodrop photospectrometer.

284

## 285 **Periodate treatment and β-elimination**

286 Total RNA was treated with 25 mM NaIO<sub>4</sub> in a final concentration of 60 mM borax and 60 mM  
287 boric acid (pH 8.6) for 30 min at room temperature. In the control, NaIO<sub>4</sub> was replaced by an equal  
288 volume of water. The reaction was quenched with glycerol and β-elimination was induced with a  
289 final concentration of 40 mM NaCl for 90 min at 45 °C. RNA was ethanol precipitated and  
290 analysed with northern blot.

291

## 292 **Generation of antibodies**

293 Custom-made antibodies (Eurogentec) against endogenous PIWI proteins were generated by  
294 immunization of two rabbits per antibody with a mix of two unique peptides (Ago3:  
295 TSGADSSSEDDKQSS, IYKRKQRMSENIQF; Piwi4: HEGRGSPSSRPAYSS,  
296 HHRESSAGGRERSGN; Piwi5: DIVRSRPLDSKVVKQ, CANQGGNWRDNYKRAI; Piwi6:  
297 MADNPQEGSSGGRIR, RGDHRQKPYDRPEQS). After 87 days and a total of four  
298 immunizations (t=0, 14, 28, 56 days), sera of both rabbits were collected, pooled, and purified  
299 against each peptide separately. Specificity of the antibody was confirmed by Western blotting of

300 Aag2 cells stably expressing PTH (ProteinA, TEV cleavage site, 6x His-tag)-tagged PIWI<sup>57</sup> upon  
301 knockdown of the respective PIWI protein, or a control knockdown (dsRLuc) (see Supplementary  
302 Fig S1A).

303

### 304 **Immunoprecipitation and western blotting**

305 Aag2 cells were lysed with RIPA buffer (10 mM Tris-HCl, 150 mM NaCl, 0.5 mM EDTA, 0.1 %  
306 SDS, 1 % Triton-X-100, 10 % DOC, 1x protease inhibitor cocktail), supplemented with 10 %  
307 glycerol and stored at -80 °C until use. The IP was performed with custom-made antibodies against  
308 Piwi4-6 and Ago3 (1:10 dilution) at 4 °C for 4 h on rotation. Protein A/G Plus beads (Santa Cruz)  
309 were added at a dilution of 1:10 and then incubated overnight at 4 °C on rotation. Beads were  
310 washed 3 times with RIPA buffer, and half was used for RNA isolation and protein analysis each.  
311 For RNA extraction, beads were treated with proteinase K for 2 h at 55 °C and isolated with phenol-  
312 chloroform extraction. Equal amounts of RNA for input and IPs were then analysed by northern  
313 blotting. For western blotting, the IP samples were boiled in 2x Laemmli buffer for 10 min at 95 °C,  
314 separated on 7.5 % SDS-polyacrylamide gels, and blotted on 0.2 µm nitrocellulose membranes  
315 (Bio-Rad) in a wet blot chamber on ice. Membranes were blocked for 1 h with 5 % milk in PBS-T  
316 (137 mM NaCl, 12 mM phosphate, 2.7 mM KCl, pH 7.4, 0.1 % (v/v) Tween 20) and incubated with  
317 PIWI-specific (dilution 1:1000) and Tubulin primary antibodies (rat anti-Tubulin alpha, MCA78G,  
318 1:1000, Sanbio) overnight at 4 °C. The next day, membranes were washed three times with PBS-T  
319 and incubated with secondary antibodies conjugated to a fluorescence dye (IRDye 800CW  
320 conjugated goat anti rabbit, 1:10,000, Li-Cor, and IRDye 680LT conjugated goat anti rat, 1:10,000,  
321 Li-Cor) for 1 h at room temperature in the dark. After washing three times in PBS-T, signal was  
322 detected with the Odyssey-CLx Imaging system (Li-Cor).

323

### 324 **Northern blot**

325 piRNAs were detected by northern blot analyses, as published in ref.<sup>58</sup>. Briefly, RNA was denatured  
326 at 80 °C for 2 min in Gel Loading Buffer II (Ambion) and size-separated on 0.5 x TBE (45 mM  
327 Tris-borate, 1 mM EDTA), 7 M Urea, 15 % denaturing polyacrylamide gels. RNA was then blotted  
328 on Hybond-NX nylon membranes (GE Healthcare) in a semi-dry blotting chamber for 45 min at  
329 20 V and 4 °C and crosslinked to the membrane with EDC crosslinking solution (127 mM 1-  
330 methylimidazole (Sigma-Aldrich), 163 mM N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide  
331 hydrochloride (Sigma-Aldrich), pH 8.0) at 60 °C for two hours. Crosslinked membranes were pre-  
332 hybridized in ULTRAHyb-Oligo hybridization buffer (Thermo Scientific) for one hour at 42 °C and  
333 probed with indicated <sup>32</sup>P 5' end-labelled DNA oligonucleotide probes over night at 42 °C.  
334 Membranes were then washed with decreasing concentrations of SCC (300 mM NaCl, 30 mM

335 sodium citrate pH 7.0; 150 mM NaCl, 15 mM sodium citrate; 15 mM NaCl, 1.5 mM sodium citrate)  
336 and 0.1 % SDS, and exposed to Carestream BioMax XAR X-Ray films (Kodak). Probe sequences  
337 can be found in the Supplementary Table S5.

338

### 339 **Reporter cloning and luciferase assay**

340 Reporters were constructed by cloning annealed and phosphorylated oligonucleotides with the  
341 indicated tapiR1 or control target sites in the pMT-GL3 vector<sup>59</sup>. This vector encodes the *Photinus*  
342 *pyralis* firefly luciferase (GL3) under a copper-inducible metallothionein promoter. Sense and  
343 antisense oligonucleotides (Sigma Aldrich) were annealed by heating to 80 °C, and gradually  
344 cooling down to room temperature, phosphorylated with T4 polynucleotide kinase (Roche) at 37 °C  
345 for 30 min, purified and then ligated into the pMT-GL3 vector. For cloning of 3'UTR and 5' UTR  
346 reporters, the target site or the target site and an upstream *BamHI* site were cloned into the *PmeI*  
347 and *SacII*, or *NotI* and *XhoI* restriction sites, respectively. ORF reporters were constructed by  
348 cloning a Kozak sequence followed by the first 45 nucleotides of luciferase and the target site into  
349 *XhoI* and *NcoI* sites. Sequences of the oligonucleotides are provided in Supplementary Table S5.  
350 Where indicated, mutated firefly or *Renilla* luciferase versions were used that harbour synonymous  
351 mutations destroying the predicted target sites for tapiR1 (firefly luciferase: 782  
352 gagtcgcttaatgtatagattgaagaa 810 mutated to 782 **gtgtcgtgcttatgtaccggttcgaggag** 810, and *Renilla*  
353 luciferase 462 tgaatggcctgatattgaagaa 483 mutated to 462 **tgagtggccagatatcgaggag** 483; modified  
354 nucleotides in bold).

355 Aag2 cells were seeded in 96-well plates the day before the experiment and transfected with 100 ng  
356 of the indicated plasmids and 100 ng pMT-*Renilla*<sup>59</sup> per well, using 2 µL X-tremeGENE HP DNA  
357 transfection reagent per 1 µg plasmid DNA according to the manufacturer's instructions.

358 Alternatively, 100 ng reporter plasmid and 100 ng pMT-*Renilla* were co-transfected with the  
359 indicated amounts of unlabelled, fully 2' *O*-methylated antisense RNA oligonucleotide using an  
360 additional amount of 4 µL X-tremeGENE HP DNA transfection reagent (Roche) per 1 µg  
361 oligonucleotide. Medium was replaced 3 h after reporter plasmid transfection with 0.5 mM CuSO<sub>4</sub>  
362 in fully supplemented Leibovitz's L-15 medium to induce the metallothionein promoter. 24 h later,  
363 cells were lysed in 30 µL Passive lysis buffer (Promega) and activity of both luciferases was  
364 measured in 10 µL of the sample with the Dual Luciferase Reporter Assay system (Promega) on a  
365 Modulus Single Tube Reader (Turner Biosystems). Firefly luciferase was normalized to *Renilla*  
366 luciferase activity. For each construct, at least two to three independent clones were measured in  
367 triplicate.

368

369

## 370 **RT-qPCR**

371 1 µg of total RNA was treated with DNaseI (Ambion) for 45 min at 37 °C and reverse transcribed  
372 using the Taqman reverse transcription kit (Applied Biosystems) according to the manufacturer's  
373 protocol. Real-time PCR was performed with the GoTag qPCR Master Mix (Promega) and  
374 measured on a LightCycler480 instrument (Roche) with 5 min initial denaturation and 45 cycles of  
375 5 s denaturation at 95 °C, 10 s annealing at 60 °C and 20 s amplification at 72 °C. Starting  
376 fluorescence values of specific mRNAs were calculated with linear regression method of log  
377 fluorescence per cycle number and LinRegPCR program, version 2015.3, as described in ref<sup>60</sup>.

378

## 379 **3' RACE**

380 3' Rapid Amplification of cDNA Ends (3' RACE) was performed using the FirstChoice RLM-  
381 RACE Kit (Thermo Fischer Scientific) according to the manufacturer's instructions. Amplification  
382 products were separated on agarose gel, purified and Sanger sequenced. Primer sequences can be  
383 found in Supplementary Table S1.

384

## 385 **Blood feeding experiment**

386 Naïve female *Aedes aegypti* (Liverpool strain) mosquitoes were offered human blood (Sanquin  
387 Blood Supply Foundation, Nijmegen, The Netherlands) through a Parafilm membrane using the  
388 Hemotek PS5 feeder (Discovery Workshops). Five engorged females were selected and sacrificed  
389 at each of the indicated time points. RNA was isolated as described above.

390

## 391 **tapiR1 antisense oligonucleotide treatment and injection**

392 Aag2 cells were seeded in 24-well plates the day before the experiment. Cells were treated with  
393 500 nM 5'Cy5-labelled, fully 2'-O-methylated antisense RNA oligonucleotide in 530 µL medium  
394 with 4 µL X-tremeGENE HP DNA transfection reagent (Roche) per 1 µg oligonucleotide. Medium  
395 was replaced after 3 h and cells from three independent experiments were harvested 48 h after  
396 transfection and prepared for RNA sequencing (see below).

397 For injection of embryos, engorged female mosquitoes that were kept at 25 °C and 70 % humidity  
398 were allowed to lay eggs for 45 min. Embryos were desiccated for 1.5 min, covered with  
399 Halocarbon oil (Sigma Aldrich) and injected with 50 µM 5'Cy5-labelled, fully 2'-O-methylated  
400 antisense RNA oligonucleotide with a FemtoJet 4x (Eppendorf) with 1200 hPa pressure. Injected  
401 embryos were then transferred to a wet Whatman paper and kept at 27 °C and 80 % humidity for  
402 the indicated times. Per experiment, 50 to 150 embryos were injected per condition.

403

404

## 405 **Scoring of embryo development and hatching**

406 Injected embryos were allowed to develop for 2.5 days after injection on a moist Whatman paper  
407 and then fixed in 4% paraform aldehyde for 8 h to overnight. Afterwards, the pigment of the  
408 endochorion was bleached with Trpis solution<sup>61</sup> (0.037 M sodium chlorite, 1.45 M acetic acid) for  
409 24 to 48 h. Embryos were washed five times in PBS and images were taken with a EVOS FL  
410 imaging system (Thermo Fisher Scientific). Embryos with evident larval segmentation (head, fused  
411 thoracical elements and abdomen) were scored as fully developed and embryos without any evident  
412 structure of the ooplasm as undeveloped. Individuals that showed first signs of structural  
413 rearrangements of the ooplasm, but did not complete larval segmentation were scored as  
414 intermediate (see Supplementary Fig S1F). To avoid biases, the scoring was performed blindly.  
415 Hatching rate was counted from injected embryos 4 days post injection. Embryos were kept moist  
416 for two days and then allowed to slowly dry for the rest of the period. The embryos were transferred  
417 to water and then forced to hatch by applying negative pressure for a period of 30 min. The number  
418 of hatched L1 larvae was counted immediately afterwards.

419

## 420 **Sequence logo**

421 Repeat monomers from the satellite repeat loci in *Ae. aegypti*, *Ae. albopictus*, and *Cx.*  
422 *quinquefasciatus* were extracted manually from the current genome annotations obtained from  
423 Vectorbase (*Aedes aegypti* Liverpool AaegL5, *Aedes albopictus* Foshan AaloF1, *Culex*  
424 *quinquefasciatus* Johannesburg CpipJ2). A repeat unit was defined as the sequence starting from the  
425 first tapiR1 nucleotide until one nucleotide upstream of the next tapiR1 sequence. Sequences were  
426 aligned using MAFFT (v7.397)<sup>62</sup> (with options `-genafpair -leavegappyregion --kimura 1 --`  
427 `maxiterate 1000 --retree 1`) and the sequence logo was constructed with the R package `ggseqlogo`<sup>63</sup>.

428

## 429 **Small RNA sequencing**

430 Small RNAs from Aag2 cells (input) or PIWI immunoprecipitations were cloned with the TruSeq  
431 small RNA sample preparation kit (Illumina) according to the manufacturer's instructions. For the  
432 input sample, size selected 19-33 nt small RNAs purified from polyacrylamide gel were used to  
433 construct the library as described previously<sup>64</sup>, whereas IP samples were not extracted from gel.  
434 Libraries were sequenced on an Illumina HiSeq 4000 instrument by Plateforme GenomEast  
435 (Strasbourg, France).

436

## 437 **mRNA sequencing**

438 RNA was isolated from Aag2 cells 48 h after AO transfection (three independent experiments), or  
439 embryos 20.5 h after AO injection (50 embryos pooled per experiment from five independent

440 experiments) with RNAsolv reagent following standard phenol-chloroform extraction (see above).  
441 Polyadenylated RNAs were extracted and sequencing libraries were prepared using the TruSeq  
442 stranded mRNA Library Prep kit (Illumina) following the manufacturer's instructions, and  
443 sequenced on an Illumina Hi Seq 4000 instrument (2x50 bases).

444

#### 445 **Analysis of mRNA sequencing**

446 Reads were mapped to the *Ae. aegypti* genome AaegL5 as provided by VectorBase  
447 (<https://www.vectorbase.org>) with STAR (version 2.5.2b)<sup>65</sup> in 2-pass mode: first mapping was  
448 done for all samples (options: --readFilesCommand zcat --outSAMtype None --outSAMattrIHstart  
449 0 --outSAMstrandField intronMotif), identified splice junctions were combined (junctions located  
450 on the mitochondrial genome were filtered out, as these are likely false positives), and this list of  
451 junctions was used in a second round of mapping (with --sjdbFileChrStartEnd) and default  
452 parameters as above. Reads were quantified with the additional option --quantMode GeneCounts to  
453 quantified reads per gene. Alternatively, reads were quantified on TEfam transposon consensus  
454 sequences ([https://tefam.biochem.vt.edu/tefam/get\\_fasta.php](https://tefam.biochem.vt.edu/tefam/get_fasta.php)) with Salmon (v.0.8.2)<sup>66</sup>, default  
455 settings and libType set to "ISR". Statistical and further downstream analyses were performed with  
456 DESeq2<sup>67</sup> from Bioconductor. Significance was tested at an FDR of 0.01 and a log<sub>2</sub> fold change of  
457 0.5. tapiR1 target sites were predicted with the online tool from RNAHybrid<sup>68</sup> with helix constraints  
458 from nucleotide two to seven, and no G:U wobble allowed in the seed. Predictions were made on  
459 the AaegL5.1 geneset as provided by VectorBase, and on TEfam transposon consensus sequences.  
460 For Fig4H, publicly available sequencing datasets<sup>69</sup> (accession numbers: SRR923702, SRR923826,  
461 SRR923837, SRR923853, SRR923704) were mapped and quantified as described above. Genes  
462 were categorized on the basis of their expression in embryos at 0-2 h vs. 12-16 h post egg-laying.  
463 Genes not detected in the 0-2 h sample were defined as purely zygotic, and genes that did not  
464 increase or decrease by more than log<sub>2</sub>(0.5) as maternal stable. Genes that changed in expression by  
465 more than log<sub>2</sub>(0.5), log<sub>2</sub>(2), and log<sub>2</sub>(5) from 0-2 h to 12-16 h were categorized as maternal  
466 unstable fraction (decreased expression), or as genes that are maternally provided but are  
467 transcribed by the zygote as addition to the preloaded maternal pool (increased expression). tapiR1  
468 targets were defined as genes that were significantly upregulated at least two fold in tapiR1 AO  
469 injected embryos and harbour a predicted tapiR1 target site (mfe <= -24).  
470 The code will be made available on GitHub upon publication.

471

#### 472 **Analysis of small RNA sequencing**

473 3' sequencing adapters (TGGAATTCTCGGGTGCCAAGG) were trimmed from the sequence  
474 reads with Cutadapt (version 1.14)<sup>70</sup> and trimmed reads were mapped with Bowtie (version

475 0.12.7)<sup>71</sup> to the *Aedes aegypti* LVP\_AGWG genome sequence AaegL5.1 obtained from VectorBase  
476 with at most 1 mismatch. Reads that mapped to rRNAs or tRNAs were excluded from the analyses.  
477 Alternatively, 3' sequencing adapters ((NNN)TGGAATTCTCGGGTGCCAAGGC) and three  
478 random bases were trimmed from publicly available datasets from *Ae. aegypti* somatic and germline  
479 tissues<sup>72</sup> (SRR5961503, SRR5961504, SRR5961505, SRR5961506) and then processed as  
480 described above. Oxidized libraries, IPs and input sample were normalized to the total number of  
481 mapped reads, all other libraries to the total number of miRNAs (in millions). piRNAs that were at  
482 least two fold enriched in a PIWI-IP compared to the corresponding input sample and were present  
483 with at least 10 rpm in the IP sample were considered PIWI-bound. Mapping positions were  
484 overlapped with basefeatures and repeatfeatures retrieved from VectorBase and counted with  
485 bedtools<sup>73</sup>. Reads that mapped to two or more features were assigned to only one feature with the  
486 following hierarchy: open reading frames > non-coding RNAs (incl. lncRNAs, pseudogenes,  
487 snoRNAs, snRNAs, miRNAs) > LTR retrotransposons > Non-LTR retrotransposons (SINEs,  
488 LINEs, Penelope) > "Cut and paste" DNA transposons > other DNA transposons (Helitrons,  
489 MITEs) > satellite and tandem repeat features > DUST > other /unknown repeats. Accordingly,  
490 reads that mapped to a repeat feature and an intron or UTR were classified as repeat-derived,  
491 whereas all other reads mapping to introns or UTRs were considered as gene-derived. Positions not  
492 overlapping with any annotation were summarized as "other". Results were then visualized with  
493 ggplot2<sup>74</sup> or Gviz<sup>75</sup> in R.

494 The code will be made available on GitHub upon publication.

495

#### 496 **Data availability**

497 Raw sequence data is deposited in the NCBI Sequence Read Archive under the BioProject number  
498 PRJNA482553.

499

#### 500 **References**

- 501 1 Garrido-Ramos, M. A. Satellite DNA: An Evolving Topic. *Genes (Basel)* **8**,  
502 doi:10.3390/genes8090230 (2017).
- 503 2 Brajkovic, J., Feliciello, I., Bruvo-Madaric, B. & Ugarkovic, D. Satellite DNA-like elements  
504 associated with genes within euchromatin of the beetle *Tribolium castaneum*. *G3 (Bethesda)*  
505 **2**, 931-941, doi:10.1534/g3.112.003467 (2012).
- 506 3 Kuhn, G. C., Kuttler, H., Moreira-Filho, O. & Heslop-Harrison, J. S. The 1.688 repetitive  
507 DNA of *Drosophila*: concerted evolution at different genomic scales and association with  
508 genes. *Mol Biol Evol* **29**, 7-11, doi:10.1093/molbev/msr173 (2012).
- 509 4 Girard, A., Sachidanandam, R., Hannon, G. J. & Carmell, M. A. A germline-specific class  
510 of small RNAs binds mammalian Piwi proteins. *Nature* **442**, 199-202,  
511 doi:10.1038/nature04917 (2006).
- 512 5 Lau, N. C. *et al.* Characterization of the piRNA complex from rat testes. *Science (80- )* **313**,  
513 363-367, doi:10.1126/science.1130164 (2006).

- 514 6 Melters, D. P. *et al.* Comparative analysis of tandem repeats from hundreds of species  
515 reveals unique insights into centromere evolution. *Genome Biol* **14**, R10, doi:10.1186/gb-  
516 2013-14-1-r10 (2013).
- 517 7 Kit, S. Equilibrium sedimentation in density gradients of DNA preparations from animal  
518 tissues. *J Mol Biol* **3**, 711-716 (1961).
- 519 8 Sueoka, N. Variation and heterogeneity of base composition of deoxyribonucleic acids: a  
520 compilation of old and new data. *J. Mol. Biol.* **3**, 31-40 (1961).
- 521 9 Hall, I. M. *et al.* Establishment and maintenance of a heterochromatin domain. *Science (80-  
522 )* **297**, 2232-2237, doi:10.1126/science.1076466 (2002).
- 523 10 Menon, D. U., Coarfa, C., Xiao, W., Gunaratne, P. H. & Meller, V. H. siRNAs from an X-  
524 linked satellite repeat promote X-chromosome recognition in *Drosophila melanogaster*. *Proc  
525 Natl Acad Sci U S A* **111**, 16460-16465, doi:10.1073/pnas.1410534111 (2014).
- 526 11 Zakrzewski, F. *et al.* Epigenetic profiling of heterochromatic satellite DNA. *Chromosoma*  
527 **120**, 409-422, doi:10.1007/s00412-011-0325-x (2011).
- 528 12 Fukagawa, T. *et al.* Dicer is essential for formation of the heterochromatin structure in  
529 vertebrate cells. *Nat Cell Biol* **6**, 784-791, doi:10.1038/ncb1155 (2004).
- 530 13 Kanellopoulou, C. *et al.* Dicer-deficient mouse embryonic stem cells are defective in  
531 differentiation and centromeric silencing. *Genes Dev* **19**, 489-501, doi:10.1101/gad.1248505  
532 (2005).
- 533 14 Pezer, Z. & Ugarkovic, D. Satellite DNA-associated siRNAs as mediators of heat shock  
534 response in insects. *RNA Biol* **9**, 587-595, doi:10.4161/rna.20019 (2012).
- 535 15 May, B. P., Lippman, Z. B., Fang, Y., Spector, D. L. & Martienssen, R. A. Differential  
536 regulation of strand-specific transcripts from *Arabidopsis* centromeric satellite repeats. *PLoS  
537 Genet* **1**, e79, doi:10.1371/journal.pgen.0010079 (2005).
- 538 16 Volpe, T. A. *et al.* Regulation of heterochromatic silencing and histone H3 lysine-9  
539 methylation by RNAi. *Science (80- )* **297**, 1833-1837, doi:10.1126/science.1074973 (2002).
- 540 17 Matthews, B. J. *et al.* Improved *Aedes aegypti* mosquito reference genome assembly enables  
541 biological discovery and vector control. *bioRxiv* (2017).
- 542 18 Siomi, M. C., Sato, K., Pezic, D. & Aravin, A. A. PIWI-interacting small RNAs: the  
543 vanguard of genome defence. *Nat Rev Mol Cell Biol* **12**, 246-258, doi:10.1038/nrm3089  
544 (2011).
- 545 19 Czech, B. & Hannon, G. J. One Loop to Rule Them All: The Ping-Pong Cycle and piRNA-  
546 Guided Silencing. *Trends Biochem Sci* **41**, 324-337, doi:10.1016/j.tibs.2015.12.008 (2016).
- 547 20 Arensburger, P., Hice, R. H., Wright, J. A., Craig, N. L. & Atkinson, P. W. The mosquito  
548 *Aedes aegypti* has a large genome size and high transposable element load but contains a  
549 low proportion of transposon-specific piRNAs. *BMC Genomics* **12**, 606, doi:10.1186/1471-  
550 2164-12-606 (2011).
- 551 21 Kawaoka, S., Izumi, N., Katsuma, S. & Tomari, Y. 3' end formation of PIWI-interacting  
552 RNAs in vitro. *Mol Cell* **43**, 1015-1022, doi:10.1016/j.molcel.2011.07.029 (2011).
- 553 22 Saito, K. *et al.* Pimet, the *Drosophila* homolog of HEN1, mediates 2'-O-methylation of Piwi-  
554 interacting RNAs at their 3' ends. *Genes Dev* **21**, 1603-1608, doi:10.1101/gad.1563607  
555 (2007).
- 556 23 Horwich, M. D. *et al.* The *Drosophila* RNA methyltransferase, DmHen1, modifies germline  
557 piRNAs and single-stranded siRNAs in RISC. *Curr Biol* **17**, 1265-1272,  
558 doi:10.1016/j.cub.2007.06.030 (2007).
- 559 24 Miesen, P., Joosten, J. & van Rij, R. P. PIWIs Go Viral: Arbovirus-Derived piRNAs in  
560 Vector Mosquitoes. *PLoS Pathog* **12**, e1006017, doi:10.1371/journal.ppat.1006017 (2016).
- 561 25 Miesen, P., Girardi, E. & van Rij, R. P. Distinct sets of PIWI proteins produce arbovirus and  
562 transposon-derived piRNAs in *Aedes aegypti* mosquito cells. *Nucleic Acids Res* **43**, 6545-  
563 6556, doi:10.1093/nar/gkv590 (2015).



- 564 26 Schnettler, E. *et al.* Knockdown of piRNA pathway proteins results in enhanced Semliki  
565 Forest virus production in mosquito cells. *J Gen Virol* **94**, 1680-1689,  
566 doi:10.1099/vir.0.053850-0 (2013).
- 567 27 Varjak, M. *et al.* Aedes aegypti Piwi4 Is a Noncanonical PIWI Protein Involved in Antiviral  
568 Responses. *mSphere* **2**, doi:10.1128/mSphere.00144-17 (2017).
- 569 28 Plohl, M. *et al.* Long-term conservation vs high sequence divergence: the case of an  
570 extraordinarily old satellite DNA in bivalve mollusks. *Heredity (Edinb)* **104**, 543-551,  
571 doi:10.1038/hdy.2009.141 (2010).
- 572 29 Li, Y. X. & Kirby, M. L. Coordinated and conserved expression of alphoid repeat and  
573 alphoid repeat-tagged coding sequences. *Dev Dyn* **228**, 72-81, doi:10.1002/dvdy.10355  
574 (2003).
- 575 30 Martinez-Lage, A., Rodriguez-Farina, F., Gonzalez-Tizon, A. & Mendez, J. Origin and  
576 evolution of Mytilus mussel satellite DNAs. *Genome* **48**, 247-256, doi:10.1139/g04-115  
577 (2005).
- 578 31 Henikoff, S., Ahmad, K. & Malik, H. S. The centromere paradox: stable inheritance with  
579 rapidly evolving DNA. *Science (80- )* **293**, 1098-1102, doi:10.1126/science.1062939 (2001).
- 580 32 Plohl, M., Mestrovic, N. & Mravinac, B. Satellite DNA evolution. *Genome Dyn* **7**, 126-152,  
581 doi:10.1159/000337122 (2012).
- 582 33 Reidenbach, K. R. *et al.* Phylogenetic analysis and temporal diversification of mosquitoes  
583 (Diptera: Culicidae) based on nuclear genes and morphology. *BMC Evol Biol* **9**, 298,  
584 doi:10.1186/1471-2148-9-298 (2009).
- 585 34 Chaves, R., Ferreira, D., Mendes-da-Silva, A., Meles, S. & Adegas, F. FA-SAT Is an Old  
586 Satellite DNA Frozen in Several Bilateria Genomes. *Genome Biol Evol* **9**, 3073-3087,  
587 doi:10.1093/gbe/evx212 (2017).
- 588 35 Bartel, D. P. MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215-233,  
589 doi:10.1016/j.cell.2009.01.002 (2009).
- 590 36 Zhang, D. *et al.* The piRNA targeting rules and the resistance to piRNA silencing in  
591 endogenous genes. *Science (80- )* **359**, 587-592, doi:10.1126/science.aao2840 (2018).
- 592 37 Shen, E. Z. *et al.* Identification of piRNA Binding Sites Reveals the Argonaute Regulatory  
593 Landscape of the *C. elegans* Germline. *Cell* **172**, 937-951 e918,  
594 doi:10.1016/j.cell.2018.02.002 (2018).
- 595 38 Matsumoto, N. *et al.* Crystal Structure of Silkworm PIWI-Clade Argonaute Siwi Bound to  
596 piRNA. *Cell* **167**, 484-497 e489, doi:10.1016/j.cell.2016.09.002 (2016).
- 597 39 Wang, Y. *et al.* Structure of an argonaute silencing complex with a seed-containing guide  
598 DNA and target RNA duplex. *Nature* **456**, 921-926, doi:10.1038/nature07666 (2008).
- 599 40 Wang, Y., Sheng, G., Juranek, S., Tuschl, T. & Patel, D. J. Structure of the guide-strand-  
600 containing argonaute silencing complex. *Nature* **456**, 209-213, doi:10.1038/nature07315  
601 (2008).
- 602 41 Mohn, F., Handler, D. & Brennecke, J. Noncoding RNA. piRNA-guided slicing specifies  
603 transcripts for Zucchini-dependent, phased piRNA biogenesis. *Science (80- )* **348**, 812-817,  
604 doi:10.1126/science.aaa1039 (2015).
- 605 42 Reuter, M. *et al.* Miwi catalysis is required for piRNA amplification-independent LINE1  
606 transposon silencing. *Nature* **480**, 264-267, doi:10.1038/nature10672 (2011).
- 607 43 Goh, W. S. *et al.* piRNA-directed cleavage of meiotic transcripts regulates spermatogenesis.  
608 *Genes Dev* **29**, 1032-1044, doi:10.1101/gad.260455.115 (2015).
- 609 44 Feliciello, I., Akrap, I. & Ugarkovic, D. Satellite DNA Modulates Gene Expression in the  
610 Beetle *Tribolium castaneum* after Heat Stress. *PLoS Genet* **11**, e1005466,  
611 doi:10.1371/journal.pgen.1005466 (2015).
- 612 45 Grimson, A. *et al.* MicroRNA targeting specificity in mammals: determinants beyond seed  
613 pairing. *Mol Cell* **27**, 91-105, doi:10.1016/j.molcel.2007.06.017 (2007).
- 614 46 Ugarkovic, D. Functional elements residing within satellite DNAs. *EMBO Rep* **6**, 1035-  
615 1039, doi:10.1038/sj.embor.7400558 (2005).

- 616 47 Tadros, W. & Lipshitz, H. D. The maternal-to-zygotic transition: a play in two acts.  
617 *Development* **136**, 3033-3042, doi:10.1242/dev.033183 (2009).
- 618 48 Thomsen, S., Anders, S., Janga, S. C., Huber, W. & Alonso, C. R. Genome-wide analysis of  
619 mRNA decay patterns during early Drosophila development. *Genome Biol* **11**, R93,  
620 doi:10.1186/gb-2010-11-9-r93 (2010).
- 621 49 Giraldez, A. J. *et al.* Zebrafish MiR-430 promotes deadenylation and clearance of maternal  
622 mRNAs. *Science (80- )* **312**, 75-79, doi:10.1126/science.1122689 (2006).
- 623 50 Lund, E., Liu, M., Hartley, R. S., Sheets, M. D. & Dahlberg, J. E. Deadenylation of maternal  
624 mRNAs mediated by miR-427 in *Xenopus laevis* embryos. *RNA* **15**, 2351-2363,  
625 doi:10.1261/rna.1882009 (2009).
- 626 51 Bushati, N., Stark, A., Brennecke, J. & Cohen, S. M. Temporal reciprocity of miRNAs and  
627 their targets during the maternal-to-zygotic transition in *Drosophila*. *Curr Biol* **18**, 501-506,  
628 doi:10.1016/j.cub.2008.02.081 (2008).
- 629 52 Rouget, C. *et al.* Maternal mRNA deadenylation and decay by the piRNA pathway in the  
630 early *Drosophila* embryo. *Nature* **467**, 1128-1132, doi:10.1038/nature09465 (2010).
- 631 53 Barckmann, B. *et al.* Aubergine iCLIP Reveals piRNA-Dependent Decay of mRNAs  
632 Involved in Germ Cell Development in the Early Embryo. *Cell Rep* **12**, 1205-1216,  
633 doi:10.1016/j.celrep.2015.07.030 (2015).
- 634 54 Fansiri, T. *et al.* Genetic mapping of specific interactions between *Aedes aegypti*  
635 mosquitoes and dengue viruses. *PLoS Genet* **9**, e1003621,  
636 doi:10.1371/journal.pgen.1003621 (2013).
- 637 55 Goertz, G. P., Vogels, C. B. F., Geertsema, C., Koenraadt, C. J. M. & Pijlman, G. P.  
638 Mosquito co-infection with Zika and chikungunya virus allows simultaneous transmission  
639 without affecting vector competence of *Aedes aegypti*. *PLoS Negl Trop Dis* **11**, e0005654,  
640 doi:10.1371/journal.pntd.0005654 (2017).
- 641 56 Mohlmann, T. W. R. *et al.* Community analysis of the abundance and diversity of mosquito  
642 species (Diptera: Culicidae) in three European countries at different latitudes. *Parasit*  
643 *Vectors* **10**, 510, doi:10.1186/s13071-017-2481-1 (2017).
- 644 57 Joosten, J. *et al.* The Tudor protein Veneno assembles the ping-pong amplification complex  
645 that produces viral piRNAs in *Aedes* mosquitoes. doi:<https://doi.org/10.1101/242305>  
646 (2018).
- 647 58 Pall, G. S. & Hamilton, A. J. Improved northern blot method for enhanced detection of  
648 small RNA. *Nat Protoc* **3**, 1077-1084, doi:10.1038/nprot.2008.67 (2008).
- 649 59 van Rij, R. P. *et al.* The RNA silencing endonuclease Argonaute 2 mediates specific  
650 antiviral immunity in *Drosophila melanogaster*. *Genes Dev* **20**, 2985-2995,  
651 doi:10.1101/gad.1482006 (2006).
- 652 60 Ramakers, C., Ruijter, J. M., Deprez, R. H. & Moorman, A. F. Assumption-free analysis of  
653 quantitative real-time polymerase chain reaction (PCR) data. *Neurosci Lett* **339**, 62-66  
654 (2003).
- 655 61 Trpis, M. A new bleaching and decalcifying method for general use in zoology. *Can J Zool*  
656 **48**, 892-893 (1977).
- 657 62 Katoh, K., Misawa, K., Kuma, K. & Miyata, T. MAFFT: a novel method for rapid multiple  
658 sequence alignment based on fast Fourier transform. *Nucleic Acids Res* **30**, 3059-3066  
659 (2002).
- 660 63 Wagih, O. ggseqlogo: a versatile R package for drawing sequence logos. *Bioinformatics* **33**,  
661 3645-3647, doi:10.1093/bioinformatics/btx469 (2017).
- 662 64 van Cleef, K. W. *et al.* Mosquito and *Drosophila* entomobirnaviruses suppress dsRNA- and  
663 siRNA-induced RNAi. *Nucleic Acids Res* **42**, 8732-8744, doi:10.1093/nar/gku528 (2014).
- 664 65 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21,  
665 doi:10.1093/bioinformatics/bts635 (2013).

- 666 66 Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and  
667 bias-aware quantification of transcript expression. *Nat Methods* **14**, 417-419,  
668 doi:10.1038/nmeth.4197 (2017).
- 669 67 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion  
670 for RNA-seq data with DESeq2. *Genome Biol* **15**, 550, doi:10.1186/s13059-014-0550-8  
671 (2014).
- 672 68 Rehmsmeier, M., Steffen, P., Hochsmann, M. & Giegerich, R. Fast and effective prediction  
673 of microRNA/target duplexes. *RNA* **10**, 1507-1517, doi:10.1261/rna.5248604 (2004).
- 674 69 Akbari, O. S. *et al.* The developmental transcriptome of the mosquito *Aedes aegypti*, an  
675 invasive species and major arbovirus vector. *G3 (Bethesda)* **3**, 1493-1509,  
676 doi:10.1534/g3.113.006742 (2013).
- 677 70 Martin, M. Cutadapt removes adapters from high-throughput sequencing reads.  
678 *EMBnet.journal* **17**, 10-12.
- 679 71 Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient  
680 alignment of short DNA sequences to the human genome. *Genome Biol* **10**, R25,  
681 doi:10.1186/gb-2009-10-3-r25 (2009).
- 682 72 Lewis, S. H. *et al.* Pan-arthropod analysis reveals somatic piRNAs as an ancestral defence  
683 against transposable elements. *Nature ecology & evolution* **2**, 174-181, doi:10.1038/s41559-  
684 017-0403-4 (2018).
- 685 73 Quinlan, A. R. BEDTools: The Swiss-Army Tool for Genome Feature Analysis. *Curr*  
686 *Protoc Bioinformatics* **47**, 11 12 11-34, doi:10.1002/0471250953.bi1112s47 (2014).
- 687 74 Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer-Verlag New York,  
688 2016).
- 689 75 Hahne, F. & Ivanek, R. in *Statistical Genomics. Methods in Molecular Biology* Vol. 1418  
690 (eds E. Mathé & S. Davis) (Humana Press, New York, NY, 2016).
- 691

## 692 **Supplementary Information**

- 693 Supplementary Table S1: Differentially expressed genes upon tapiR1 AO treatment in Aag2 cells.
- 694 Supplementary Table S2: Differentially expressed transposable elements upon tapiR1 AO treatment  
695 in Aag2 cells.
- 696 Supplementary Table S3: Differentially expressed genes upon tapiR1 AO treatment in embryos.
- 697 Supplementary Table S4: Differentially expressed transposable elements upon tapiR1 AO treatment  
698 in embryos.
- 699 Supplementary Table S5: Oligonucleotide sequences used in this study.

700

## 701 **Acknowledgments**

702 We thank past and current members of the Van Rij laboratory for discussions. We are grateful to  
703 Anna Beth Crist and Artem Baidaliuk for their help with mosquito rearing and embryo injections,  
704 and to Catherine Bourguoin and Nicolas Puchot for assistance with the microinjection apparatus.

705 We thank Bas Dutilh for his support with analyzing target site enrichments, and Geert-Jan van  
706 Gemert for kindly providing *An. stephensi* mosquitoes. Tim Möhlmann is acknowledged for  
707 providing wild-caught mosquito samples.

708 This work is financially supported by a Consolidator Grant from the European Research Council  
709 under the European Union's Seventh Framework Programme (grant number ERC CoG 615680) and  
710 a VICI grant from the Netherlands Organization for Scientific Research (grant number  
711 016.VICI.170.090). A stay of R.H. at Pasteur Institute, Paris, France was supported by  
712 ERASMUS+.

713

#### 714 **Author contributions**

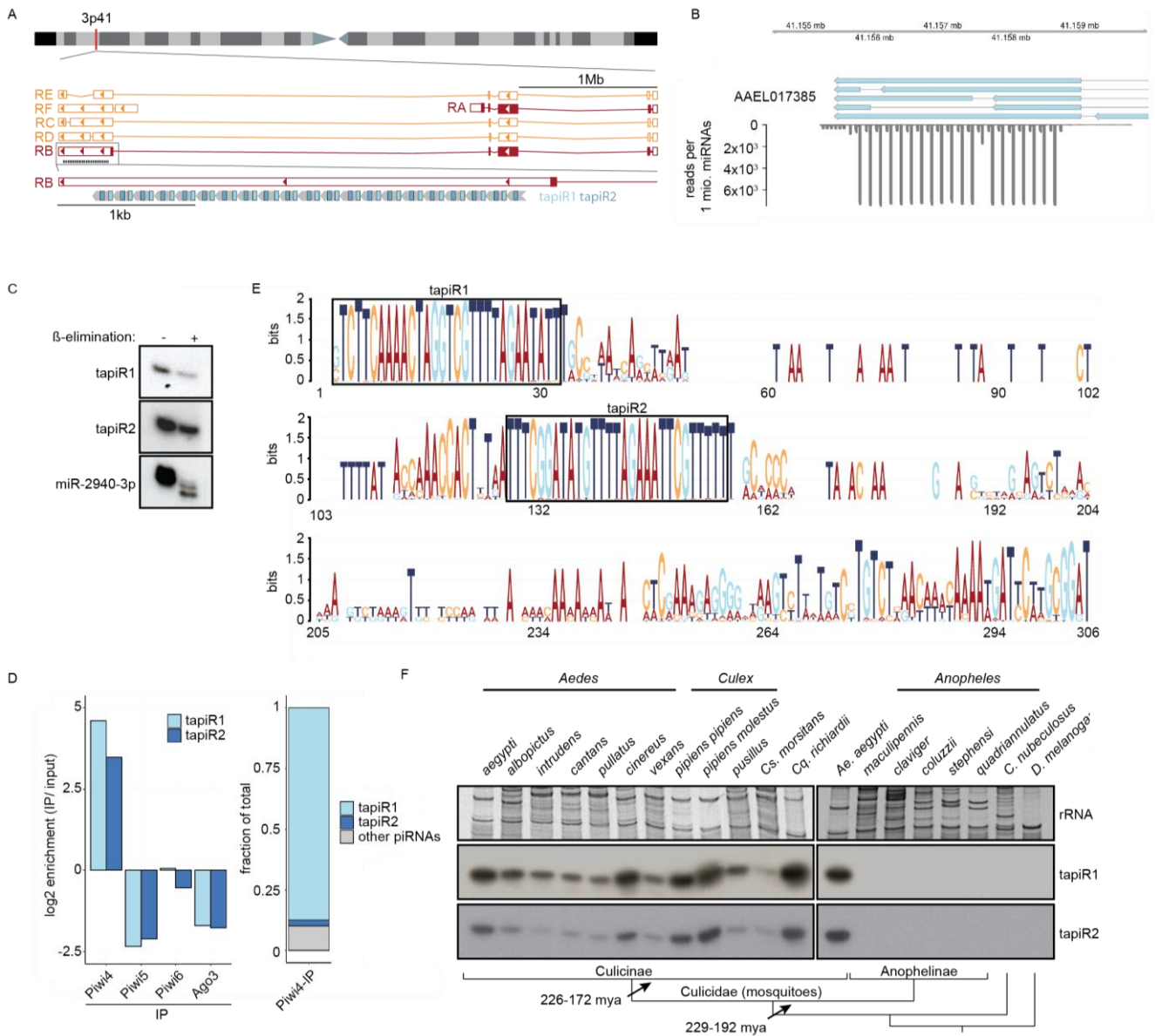
715 R.H., P.M., and R.P.v.R designed the experiments and analyzed the data. R.H. performed the  
716 computational analyses and most of the experiments, except for PIWI-IPs for small RNA  
717 sequencing (J.J. and E.T.), design and validation of PIWI antibodies (B.P.), and tissue isolations  
718 and blood feeding experiment (C.B.F.V. and C.J.K.). S.H.M. and L.L. helped with optimizing  
719 embryo injections. R.H. and R.P.v.R. wrote the paper. All authors read and contributed to the  
720 manuscript.

721

#### 722 **Author information**

723 The authors declare no competing financial interests. Correspondence and requests for materials  
724 should be addressed to R.P.v.R. ([ronald.vanrij@radboudumc.nl](mailto:ronald.vanrij@radboudumc.nl)).

725



726

727 **Figure 1:** Conserved piRNAs are expressed from a satellite repeat and associate with Piwi4.

728 (A) Current annotation of the gene AEEL017385 and its splice variants (RA-RF) on  
729 chromosome 3, with the position of the satellite repeat locus and tapiR1/ 2 piRNAs indicated.

730 (B) Read coverage of the satellite repeats locus. Depicted are exons of AEEL017385 (blue), and  
731 small RNAs per million mapped miRNAs in Aag2 cells.

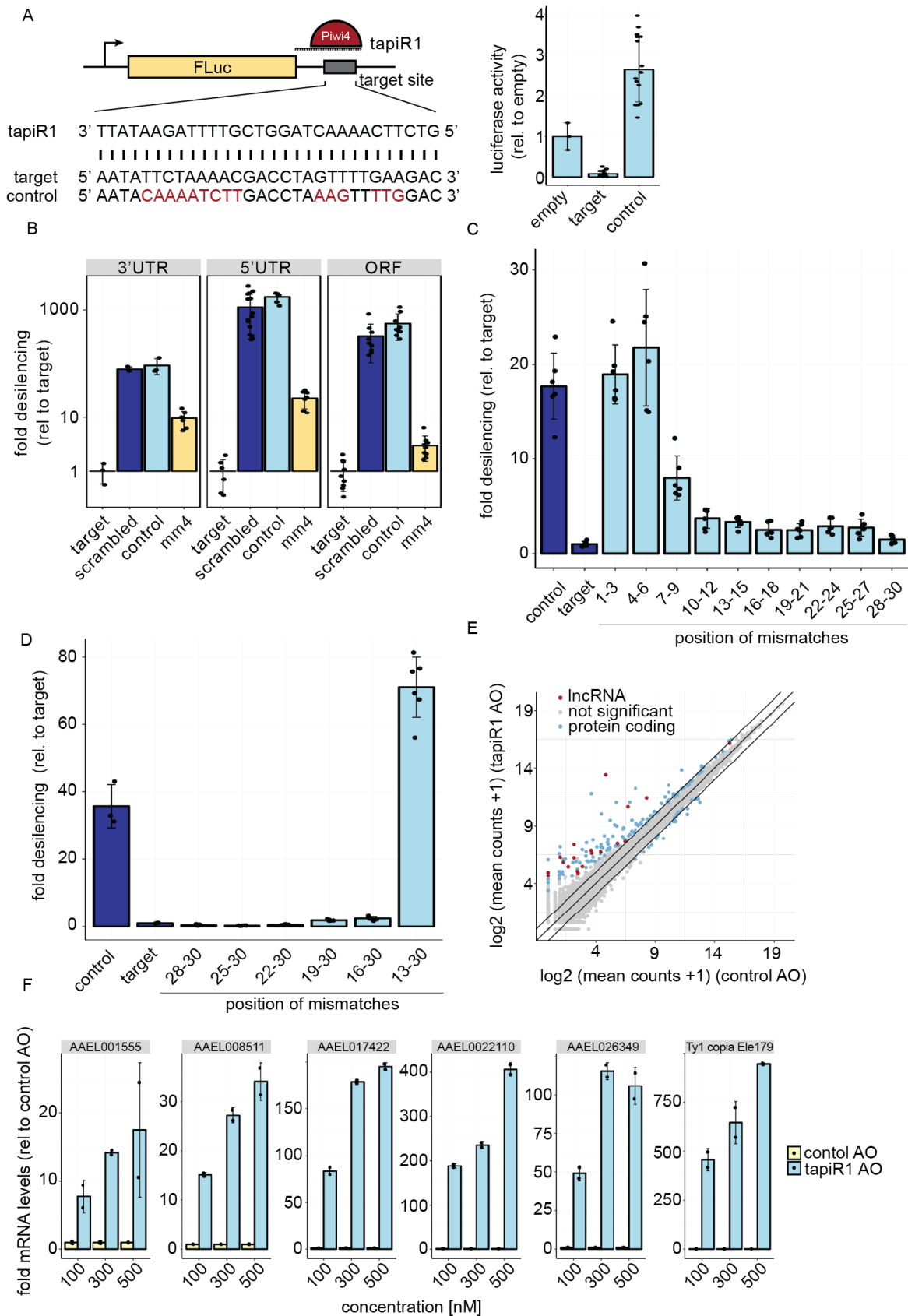
732 (C) Small RNA Northern blot of tapiR1 and 2 upon  $\beta$ -elimination in Aag2 cells. miR-2940-3p serves  
733 as positive control for the treatment.

734 (D) Enrichment or depletion of tapiR1/2 compared to input sample in the indicated PIWI-IP small  
735 RNA sequencing libraries (left panel), and fraction of total reads enriched in Piwi4  
736 (right panel).

737 (E) Sequence conservation of the satellite repeat monomers. All individual repeat monomers from  
738 *Ae. aegypti*, *Ae. albopictus* and *Cx. quinquefasciatus* were used to generate the sequence logo.

739 Boxes highlight the position of tapiR1 and 2 in the monomer.

740 (F) Northern blot analysis of *tapiR1/2* in the indicated mosquito species (genera *Aedes*, *Culex*,  
741 *Culiseta*, *Coquillettidia*, and *Anopheles*) and other insects (*Culicoides* and *Drosophila*). Ethidium  
742 bromide-stained rRNA serves as loading control. For comparison, *Ae. aegypti* was included twice.  
743 Schematic representation of the phylogenetic relationships are indicated in the bottom panel. Bar  
744 lengths are arbitrary and do not reflect evolutionary distances.  
745



746

747 **Figure 2:** tapiR1 silences target RNAs *in trans* through seed-mediated base pairing.

748 (A) Schematic representation of the firefly luciferase (FLuc) reporter constructs (left panel) and

749 luciferase assay in Aag2 cells transfected with reporters containing no target site (empty), a fully

750 complementary target site to tapiR1, or a control target site.

751 (B) Luciferase assay of reporters with tapiR1 target sites, mismatched sites (mm4), or control  
752 sequences located at different positions in the reporter mRNA.

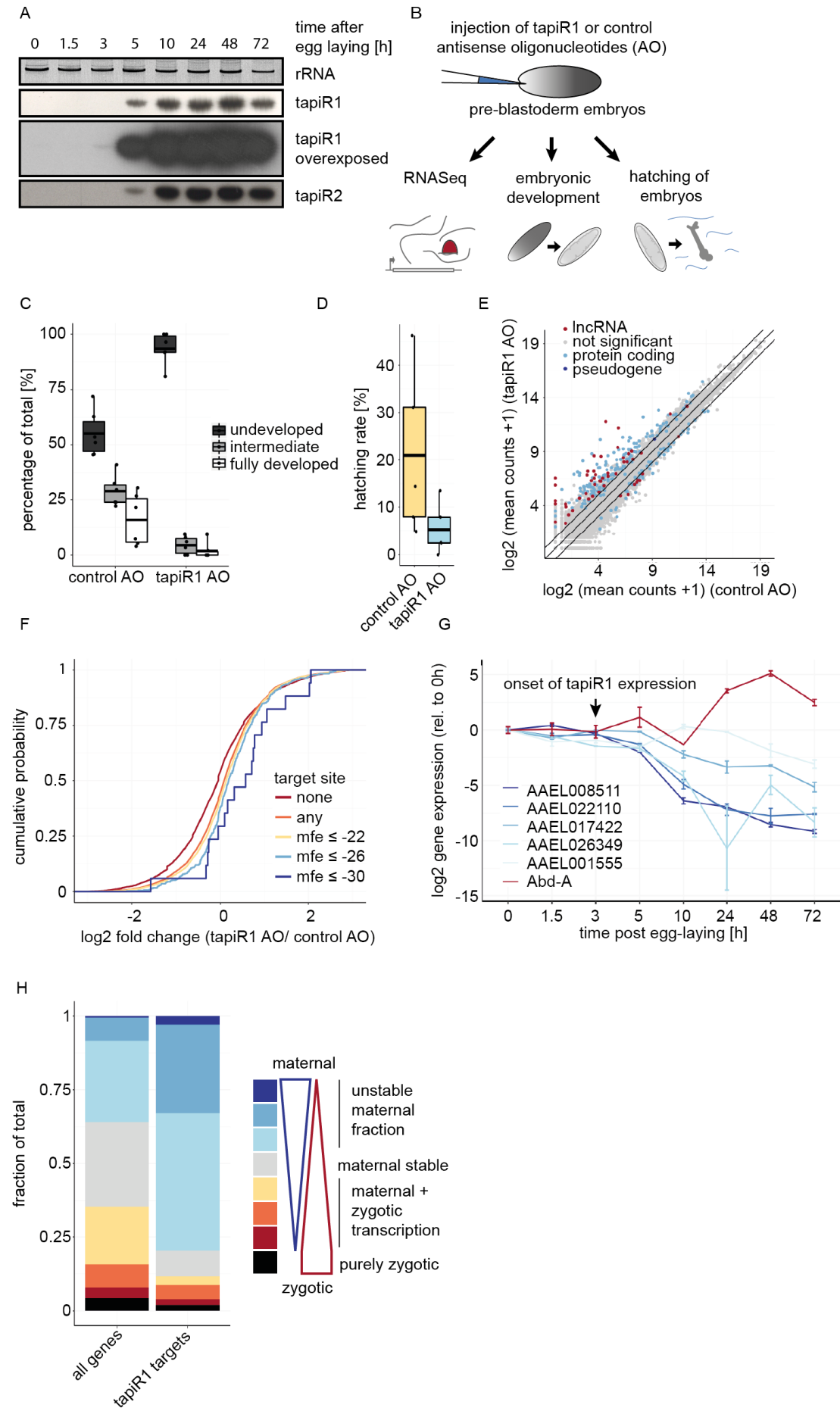
753 (C, D) Luciferase assay of tapiR1 reporters harbouring three consecutive mismatches (C), or  
754 increasing number of mismatches (D) at the indicated positions of the piRNA target site in the 3'  
755 UTR of firefly luciferase. Firefly luciferase activity was normalized to the activity of a co-  
756 transfected *Renilla* luciferase reporter. Indicated are mean, standard deviation, and individual  
757 measurements of a representative experiment performed with two to three independent clones per  
758 construct and measured in triplicates.

759 (E) log<sub>2</sub> expression of mRNAs and lncRNAs in Aag2 cells upon treatment with a tapiR1 specific or  
760 control antisense oligonucleotide (AO). Depicted are average read counts in three biological  
761 replicates. A pseudo-count of one was added to all values in order to plot values of zero. Diagonal  
762 lines indicate a fold change of two. Significance was tested at a false discovery rate (FDR) of 0.01  
763 and a log<sub>2</sub> fold change of 0.5 as indicated by coloured dots.

764 (F) RT-qPCR of tapiR1 target genes upon transfection of Aag2 cells with tapiR1 specific or control  
765 AO. Depicted are mean, standard deviation, and individual measurements of one experiment  
766 measured in technical duplicates.

767





769 **Figure 3:** tapiR1 is essential for embryonic development *in vivo* by promoting turnover of  
770 maternally deposited transcripts.

771 (A) Expression of tapiR1 and 2 as analysed by northern blot in *Ae. aegypti* embryos. Time indicates  
772 the age of the embryos after a 30 min egg laying period. For each time point, around 50 to 150 eggs  
773 were pooled.

774 (B) Outline of the experimental procedure.

775 (C) Percent of embryos injected with either tapiR1 or control AO that reached the indicated  
776 developmental stages 2.5 days post injection. Individual embryos were scored as either  
777 undeveloped, intermediate or fully developed, as shown in Supplementary Figure S1F.

778 (D) Percent of embryos injected with tapiR1 or control AO that hatched four days post injection.  
779 Box-whiskers plot represents mean, first and third quartile and maximum and minimum of the data.  
780 Points show the individual experiments with 20 to 60 (C), or 50 to 150 (D) embryos per group.

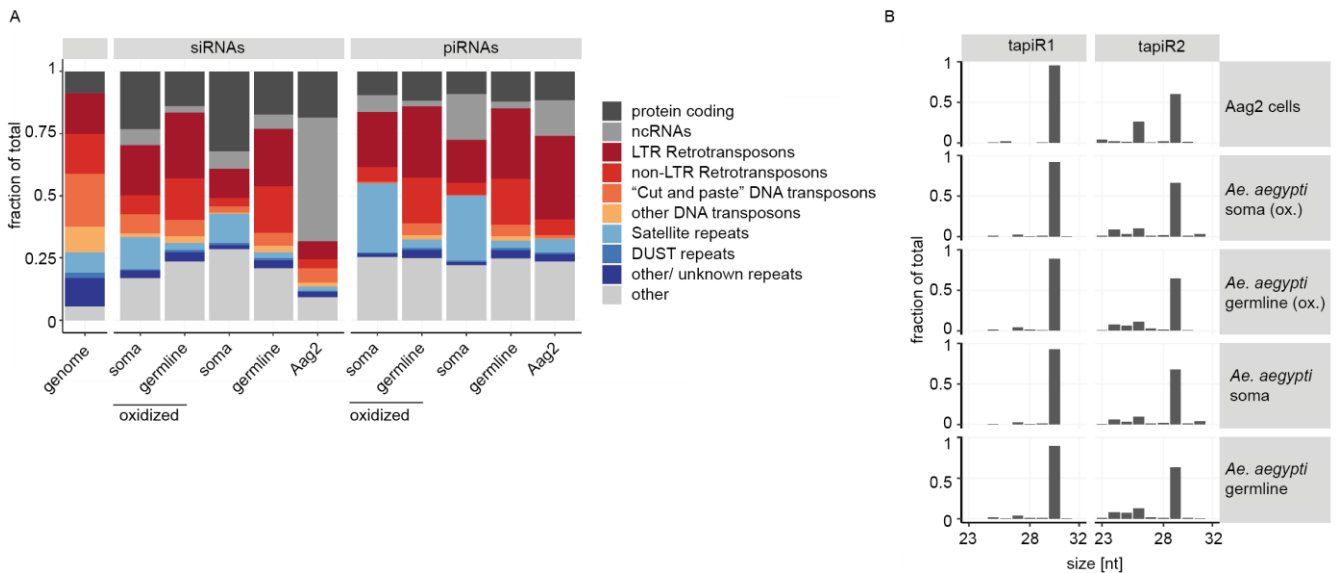
781 (E) log<sub>2</sub> expression of genes in embryos injected with tapiR1 specific or a control AO at 20.5 h post  
782 injection. Mean counts from five biological replicates plus a pseudo-count of one are plotted. Per  
783 replicate, 50 embryos per group were pooled. Significance was tested at an FDR of 0.01 and a log<sub>2</sub>  
784 fold change of 0.5. Diagonal lines highlight a fold change of two.

785 (F) Experimental cumulative distribution of log<sub>2</sub> fold changes of genes without or with predicted  
786 target sites for tapiR1. Target sites were grouped based on the predicted minimum free energy of  
787 the piRNA/target duplex.

788 (G) Expression of tapiR1 target genes in *Ae. aegypti* embryos. RT-qPCR was performed on samples  
789 shown in (A). Abd-A is a gene not targeted by tapiR1 and serves as negative control.

790 (H) Fraction of genes in different classes of genes expressed in embryos between 0 and 16 h post  
791 egg-laying.

792



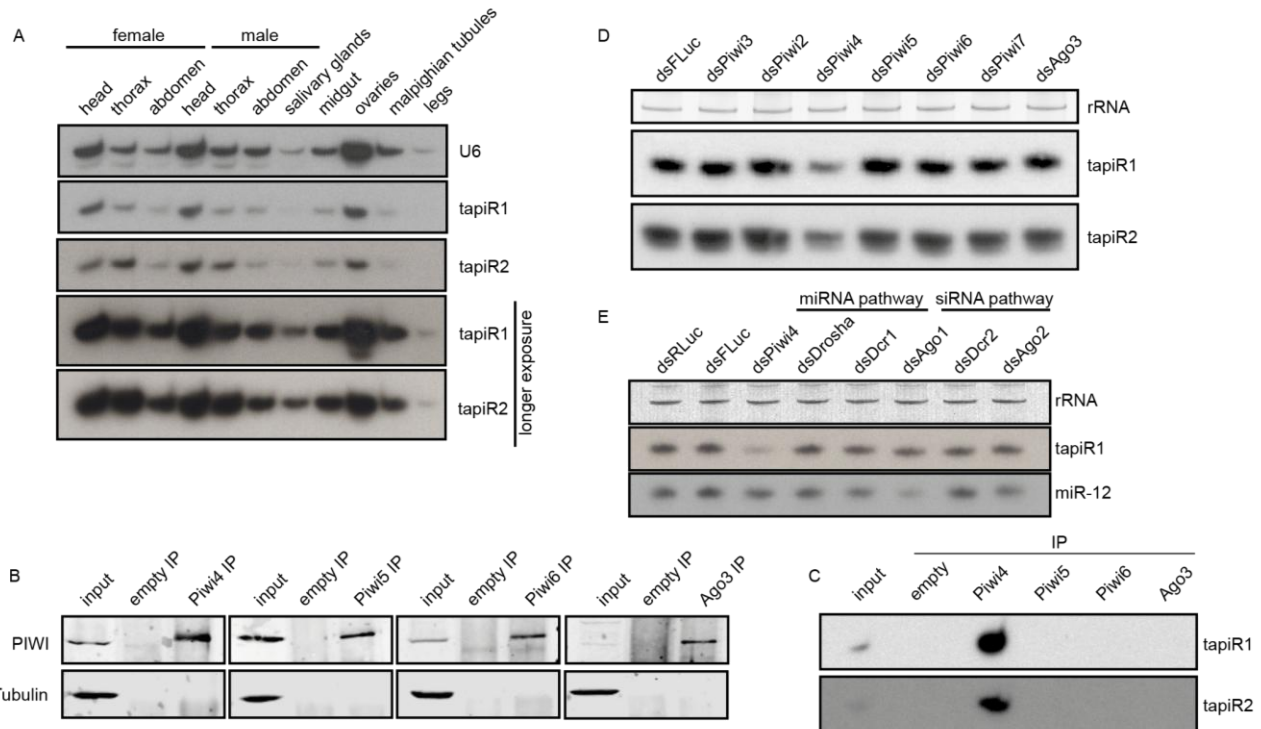
793

794 **Extended Data Figure 1:** Expression of piRNAs from a satellite repeat locus.

795 (A) Fraction of siRNAs and piRNAs mapping on genomic features in libraries derived from  
 796 germline or somatic *Ae. aegypti* adult tissues. Small RNAs that overlapped multiple features were  
 797 assigned to only one category (see Methods). Leftmost bar depicts the abundance of each feature  
 798 category in the genome.

799 (B) Read length distribution of tapiR1 and 2 in libraries from Aag2 cells, and adult germline and  
 800 somatic tissues (oxidized or untreated).

801



802

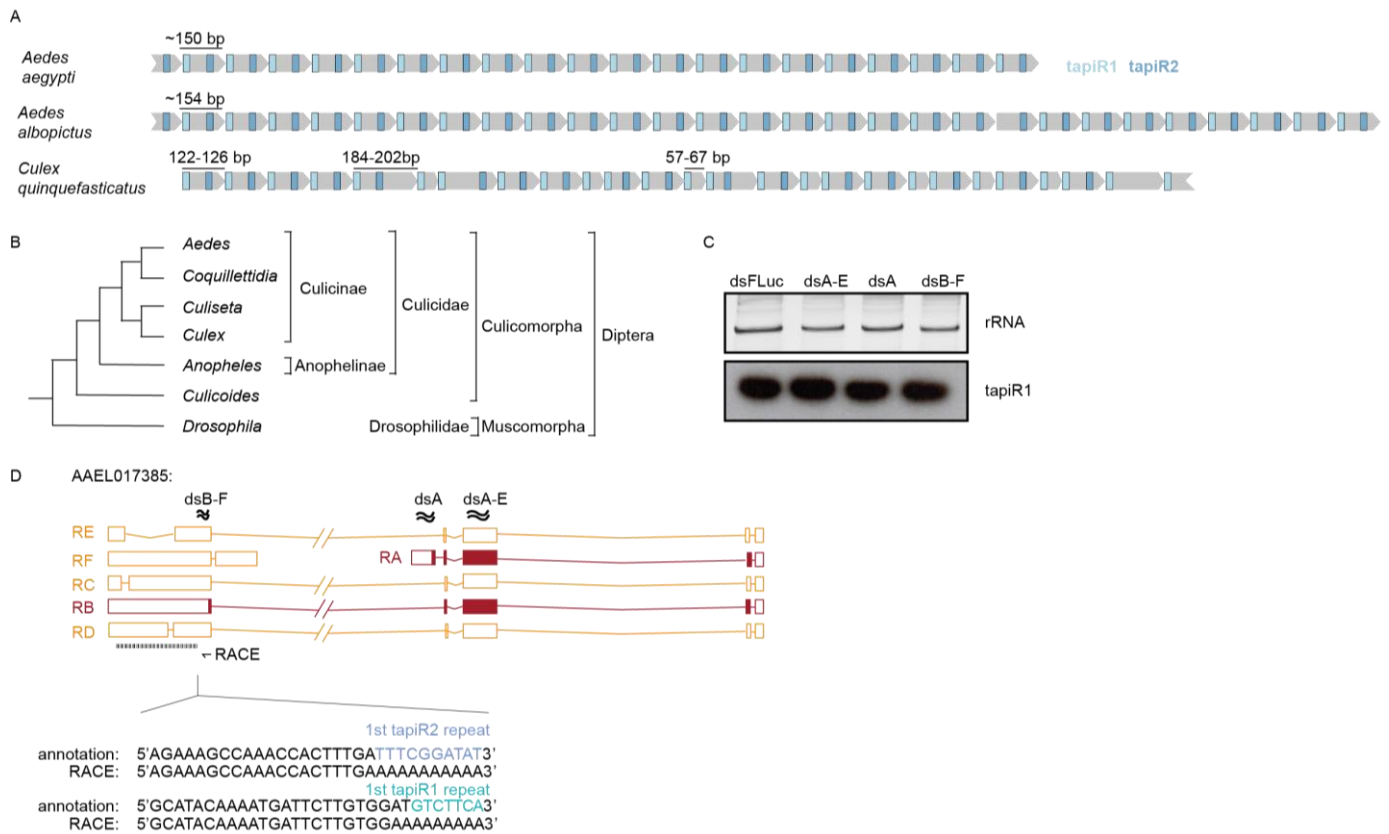
803 **Extended Data Figure 2:** *tapiR1* and 2 are expressed in *Ae. aegypti* mosquitoes and associate with  
 804 PIWI proteins.

805 (A, D, E) Northern blot of *tapiR1* and 2 in different tissues of adult mosquitoes (A), upon dsRNA-  
 806 mediated knockdown of individual PIWI genes (D), and upon knockdown of miRNA and siRNA  
 807 pathway genes (E), or a control (dsFLuc and dsRLuc) in Aag2 cells. Ethidium bromide-stained  
 808 rRNA, or U6 snRNA served as loading control.

809 (B) Western blot analysis of the indicated PIWI proteins before (input) and after  
 810 immunoprecipitation (IP) used for the small RNA northern blot of panel C. An IP with empty beads  
 811 serves as negative control. Tubulin was used to control for non-specific binding.

812 (C) Immunoprecipitation of the indicated PIWI proteins from Aag2 cells followed by northern blot  
 813 analyses for *tapiR1* and 2.

814



815  
816 **Extended Data Figure 3:** Expression of tapiR1 is independent of AAEL017385.

817 (A) Schematic representation of the tapiR1/2 satellite repeat locus in *Ae. aegypti*, *Ae. albopictus* and  
818 *Cx. quinquefasciatus*. Numbers indicate lengths of the repeats, and, for *Cx. quinquefasciatus*, also  
819 length of deviating repeat monomers.

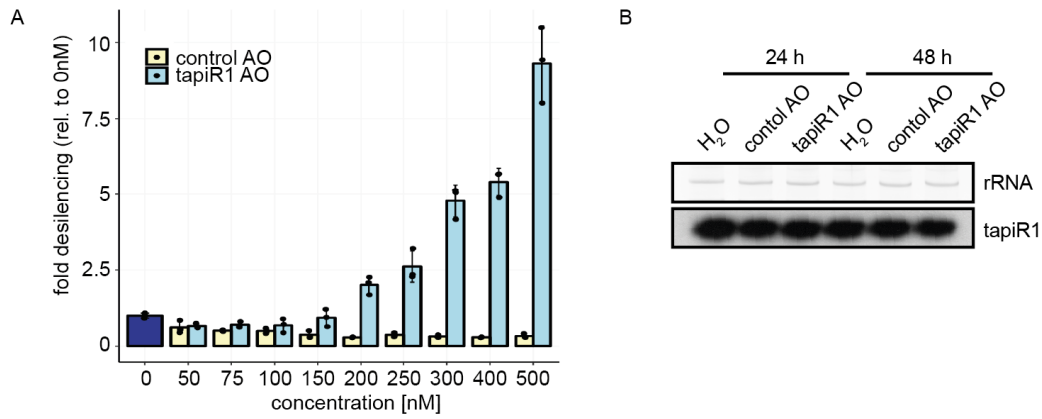
820 (B) Evolutionary relationships of dipterous genera based on ref<sup>33</sup>. Bar lengths are arbitrary and do  
821 not reflect evolutionary distances.

822 (C) Northern blot of tapiR1 in Aag2 cells treated with control dsRNA targeting different transcripts  
823 of AAEL017385, or, as control, firefly luciferase (FLuc). Ethidium bromide stained rRNA serves as  
824 loading control.

825 (D) Top panel: Schematic representation of the AAEL017385 locus and satellite repeat. The primer  
826 used for 3' RACE and positions targeted by dsRNA in panel C are indicated with an arrow and  
827 wavy lines, respectively.

828 Bottom panel: 3' RACE analysis of AAEL017385. Indicated are sequences from the current  
829 AaegL5 genome annotation and RACE PCR products. The sequences of the 5' terminal tapiR1 and  
830 2 repeats are highlighted with colours.

831



832

833 **Extended Data Figure 4:** An antisense oligonucleotide relieves tapiR1-mediated silencing.

834 (A) Luciferase assay of a reporter with a fully complementary target site for tapiR1 in the 3' UTR.

835 Cells were co-transfected with the reporter and increasing amounts of a fully 2' *O*-methylated

836 antisense RNA oligonucleotide (AO), or a control AO. Firefly luciferase activity was normalized to

837 the activity of a co-transfected *Renilla* luciferase reporter. Indicated are mean, standard deviation

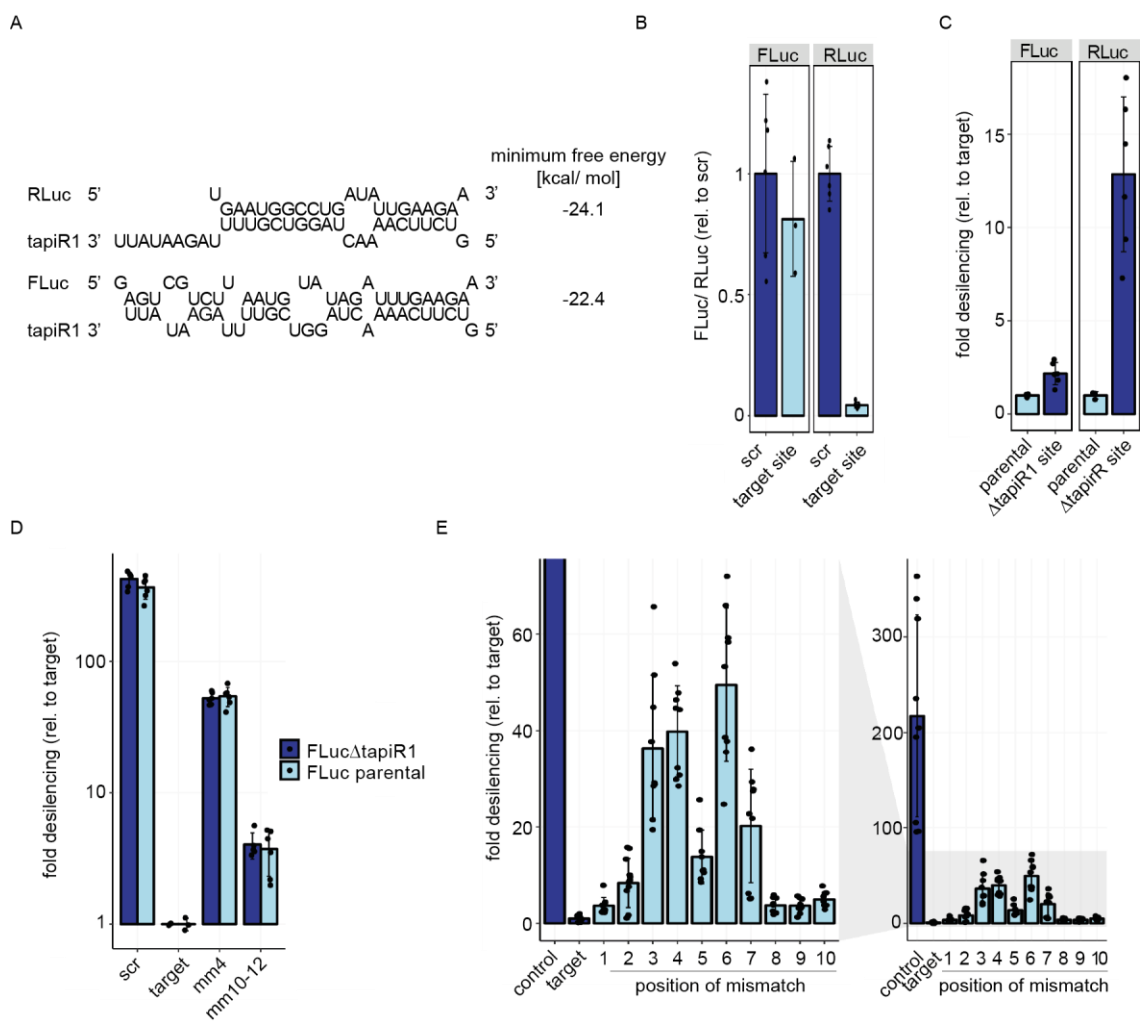
838 and individual measurements from a representative experiment measured in triplicate.

839 (B) Northern blot detection of tapiR1 in Aag2 cells upon treatment with tapiR1 or control AO in

840 Aag2 cells. Cells were harvested after the indicated time points. Ethidium bromide-stained rRNA

841 serves as loading control.

842



843

844 **Extended Data Figure 5: *Renilla* luciferase contains a functional tapiR1 target site.**

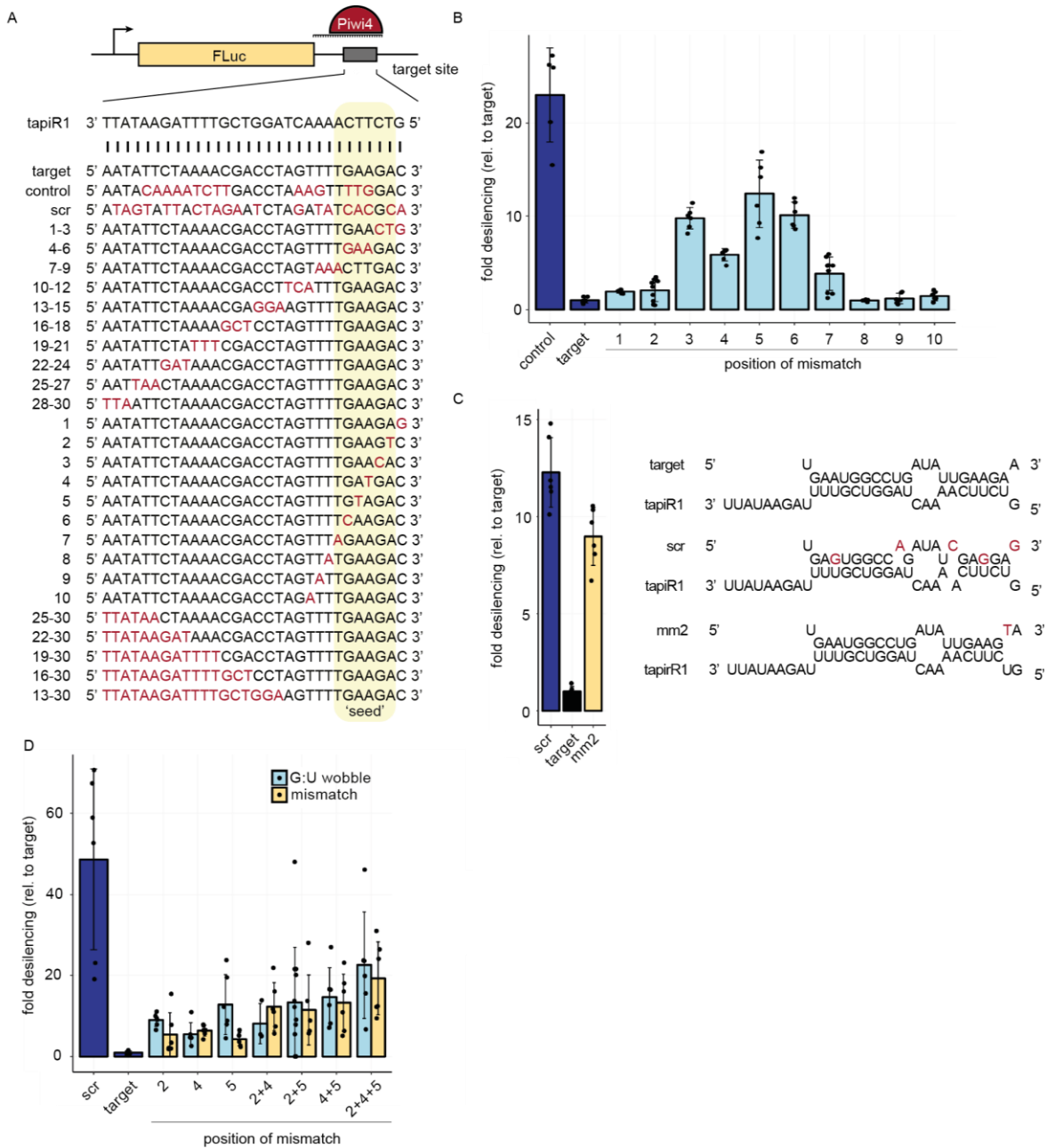
845 (A) Schematic representation of predicted tapiR1 target sites and minimum free energy of the  
 846 indicated structures in the coding sequences of *Renilla* luciferase (RLuc) or firefly luciferase  
 847 (FLuc).

848 (B) Luciferase assay of Aag2 cells transfected with reporters carrying either a scrambled (scr) site  
 849 or the predicted target site from firefly luciferase (left panel), or from *Renilla* luciferase (right  
 850 panel) from panel (A) in the 3'UTR of FLuc.

851 (C) Luciferase activity of firefly luciferase or *Renilla* luciferase construct with synonymous  
 852 mutations introduced into the predicted tapiR1 target site ( $\Delta$ tapiR1 site) or the parental clones.

853 (D) Luciferase reporter assay of reporters carrying target sites for tapiR1 as indicated in panel A in  
 854 the 3'UTR of either the parental firefly luciferase, or the  $\Delta$ tapiR1 firefly luciferase version.

855 (E) Reporter assay with luciferase carrying tapiR1 target sites with single mismatches in the 3' UTR  
 856 as used in Extended Data Fig 6B, using RLuc with a mutated tapiR1 target site ( $\Delta$ tapiR1 site) for  
 857 normalization. Left panel is a zoom to the x-axis of the right panel. Shown are mean, standard  
 858 deviation and individual measurements from representative experiments performed with at least two  
 859 different clones per construct, and measured in triplicate.



860

861 **Extended Data Figure 6:** tapiR1 uses a G:U wobble sensitive seed sequence for target recognition.

862 (A) Schematic representation of the reporter constructs used in panel B and Figure 2. Numbers

863 indicate the position of the mismatch relative to the 5' end of the piRNA.

864 (B) Luciferase assay of reporters carrying a tapiR1 target site with single mismatches in the 3'

865 UTR.

866 (C) Luciferase activity of reporters with the tapiR1 target site from RLuc and indicated mismatches

867 in the 3' UTR of FLuc (left panel). The tapiR1 target duplexes and mutants are presented in the

868 right panel.

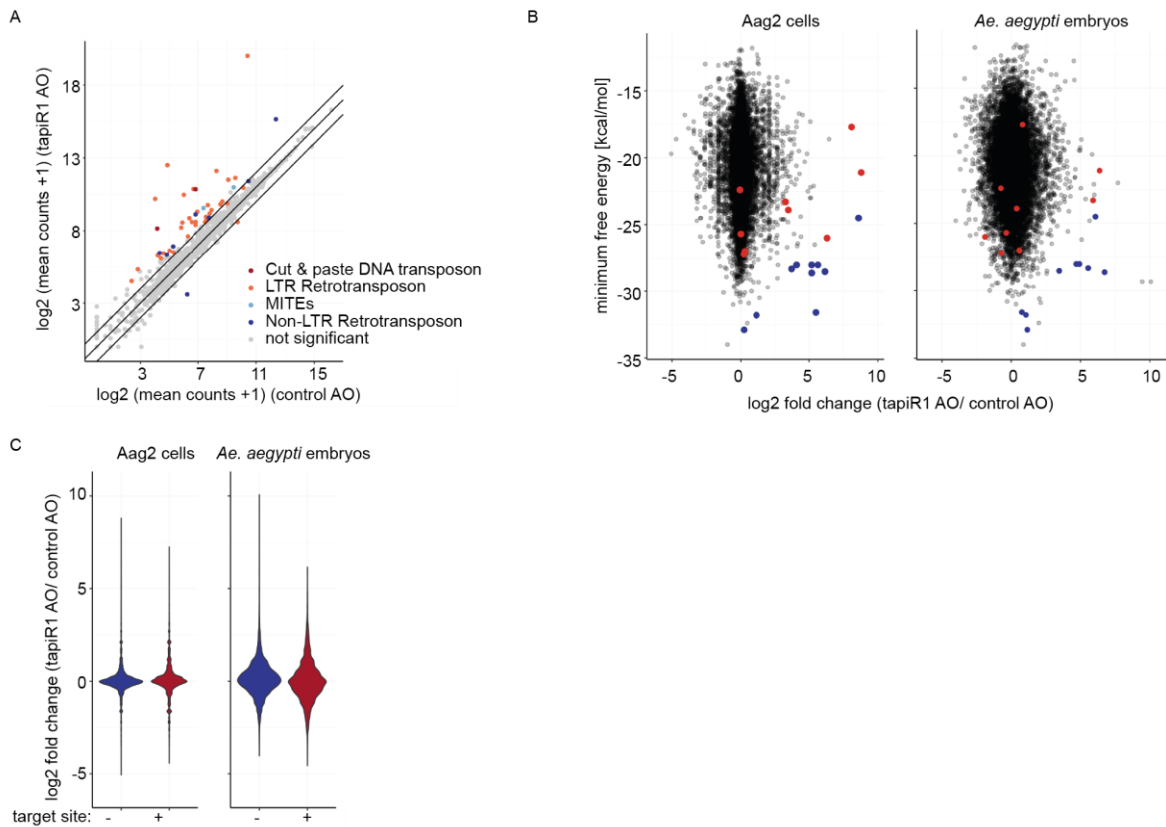
869 (D) Luciferase activity of tapiR1 reporters carrying mismatches or G:U wobble base pairs at the

870 indicated positions. Firefly luciferase activity was normalized to the activity of a co-transfected

871 *Renilla* luciferase reporter to control for differences in transfection efficiencies. Data represent



872 mean, standard deviation and individual measurements of representative experiments with two  
873 independent clones per construct and measured in triplicates.  
874



875

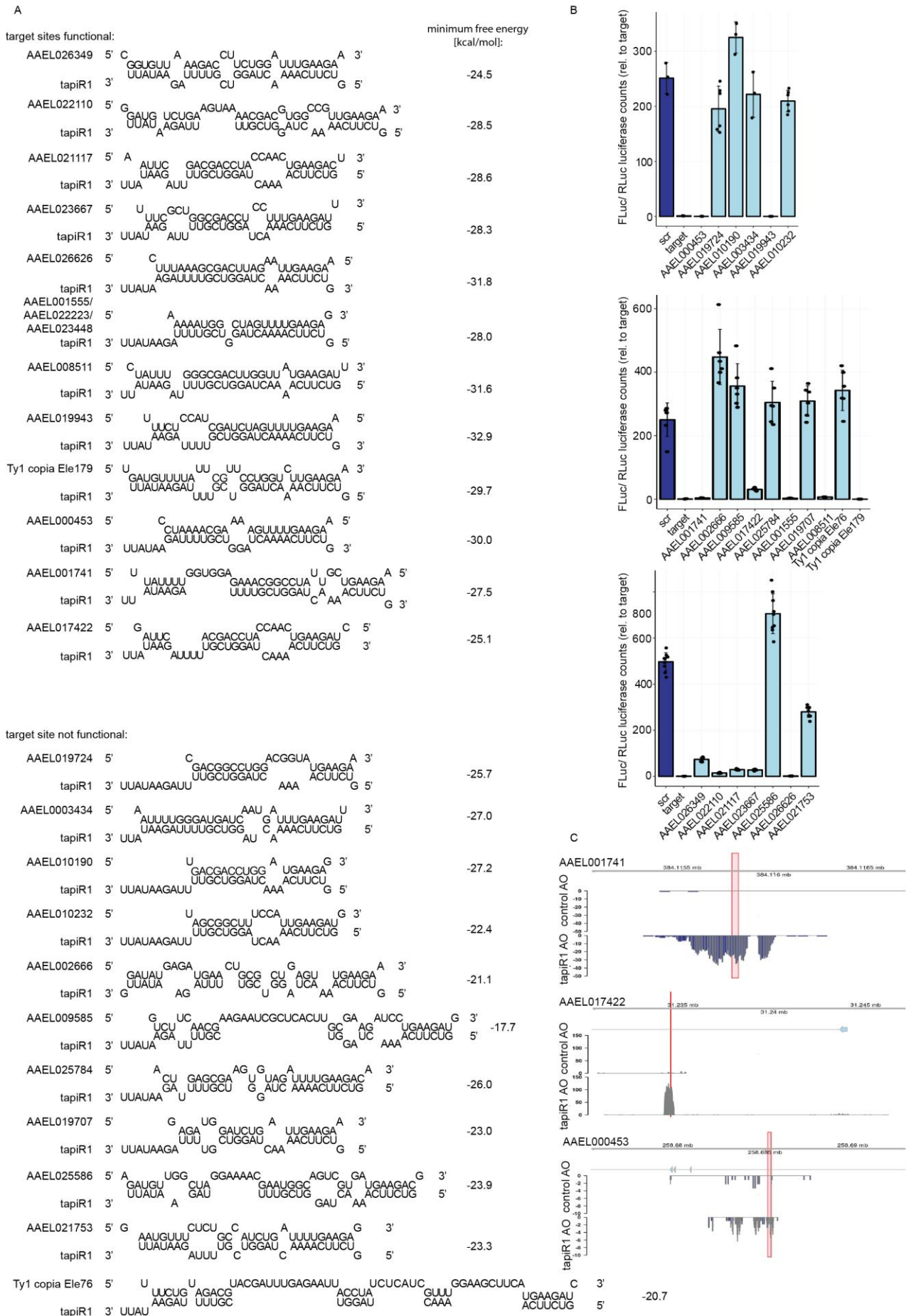
876 **Extended Data Figure 7:** tapiR silences gene expression in Aag2 cells.

877 (A) log<sub>2</sub> mRNA expression of transposable elements in Aag2 cells treated with a tapiR1 specific  
878 antisense oligonucleotide (AO) or control AO. Depicted are the means of three biological replicates.  
879 A pseudo-count of one was added to all values in order to plot values of zero. Diagonal lines  
880 represent a fold change of two. Significance was tested at an FDR of 0.01 and a log<sub>2</sub> fold change of  
881 0.5.

882 (B) log<sub>2</sub> fold changes of genes upon treatment with tapiR1 or control AO in Aag2 cells (left) and  
883 mosquito embryos (right) plotted against the minimum free energy of predicted tapiR1-target  
884 duplexes. Blue dots indicate target sites that were confirmed to be functional, and red dots indicate  
885 target sites that were not functional in luciferase reporter assays (see Extended Data Fig 8).

886 (C) Violin plot of log<sub>2</sub> fold changes of all genes in Aag2 cells (left) and mosquito embryos (right),  
887 either with or without predicted tapiR1 target site.

888



891 **Extended Data Figure 8:** Validation of tapiR1 target genes.

892 (A) Predicted structures and minimum free energy of tapiR1/target duplexes analysed in panel B.

893 (B) Luciferase assay of reporters carrying the predicted target site from panel A in the 3' UTR of

894 firefly luciferase. Firefly luciferase activity was normalized to the activity of a co-transfected

895 *Renilla* luciferase reporter to control for differences in transfection efficiencies. Indicated are mean,

896 standard deviation and individual measurements from representative experiments performed with

897 two to three independent clones per construct and measured in triplicates.

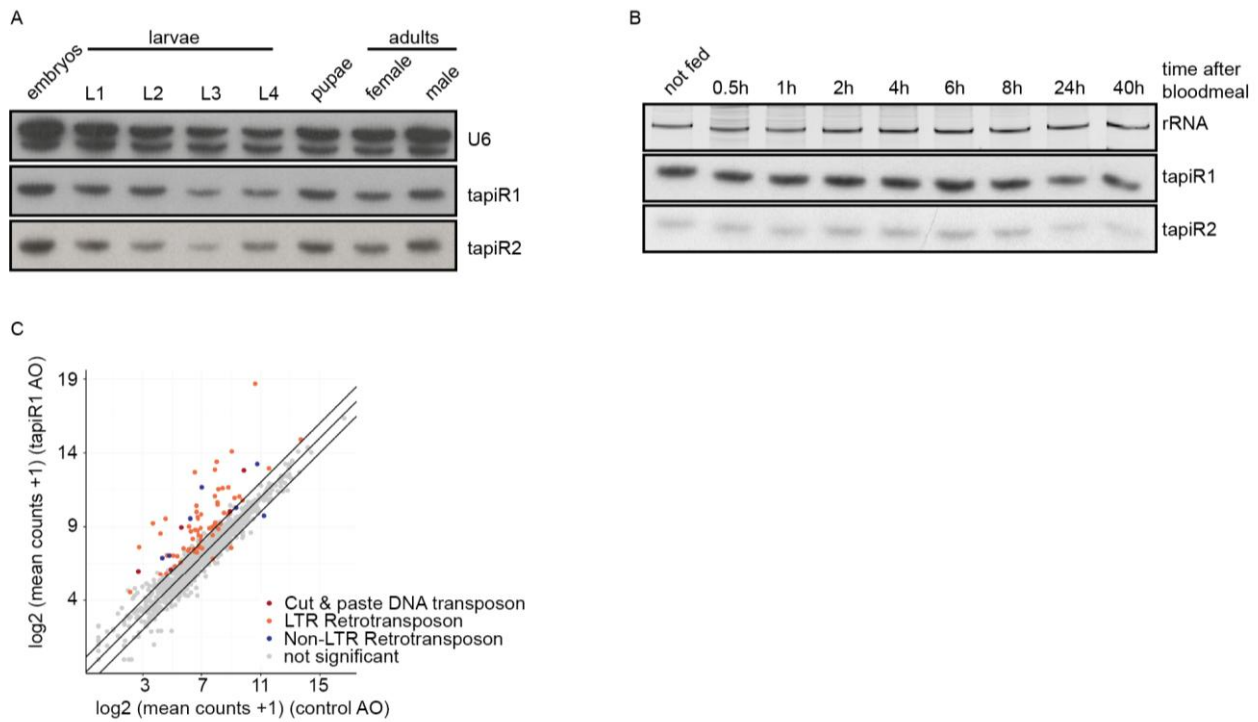
898 (C) AAEL017422, AAEL001741, and AAEL000453 were annotated in the previous AaegL3 gene

899 set, but not in the current AaegL5 gene set. Read coverage in tapiR1 AO and control AO treated

900 Aag2 cells at these genomic regions suggests that these regions are actively transcribed, but

901 repressed by tapiR1. Red boxes indicate the positions of tapiR1 target sites.

902



903

904 **Extended Data Figure 9:** *tapiR1* regulates gene expression in mosquito embryos.

905 (A, B) Northern blot analysis of *tapiR1* and 2 in developmental stages of *Ae. aegypti*

906 mosquitoes (A), or at different time points after blood feeding (B). U6 snRNA (A) or ethidium

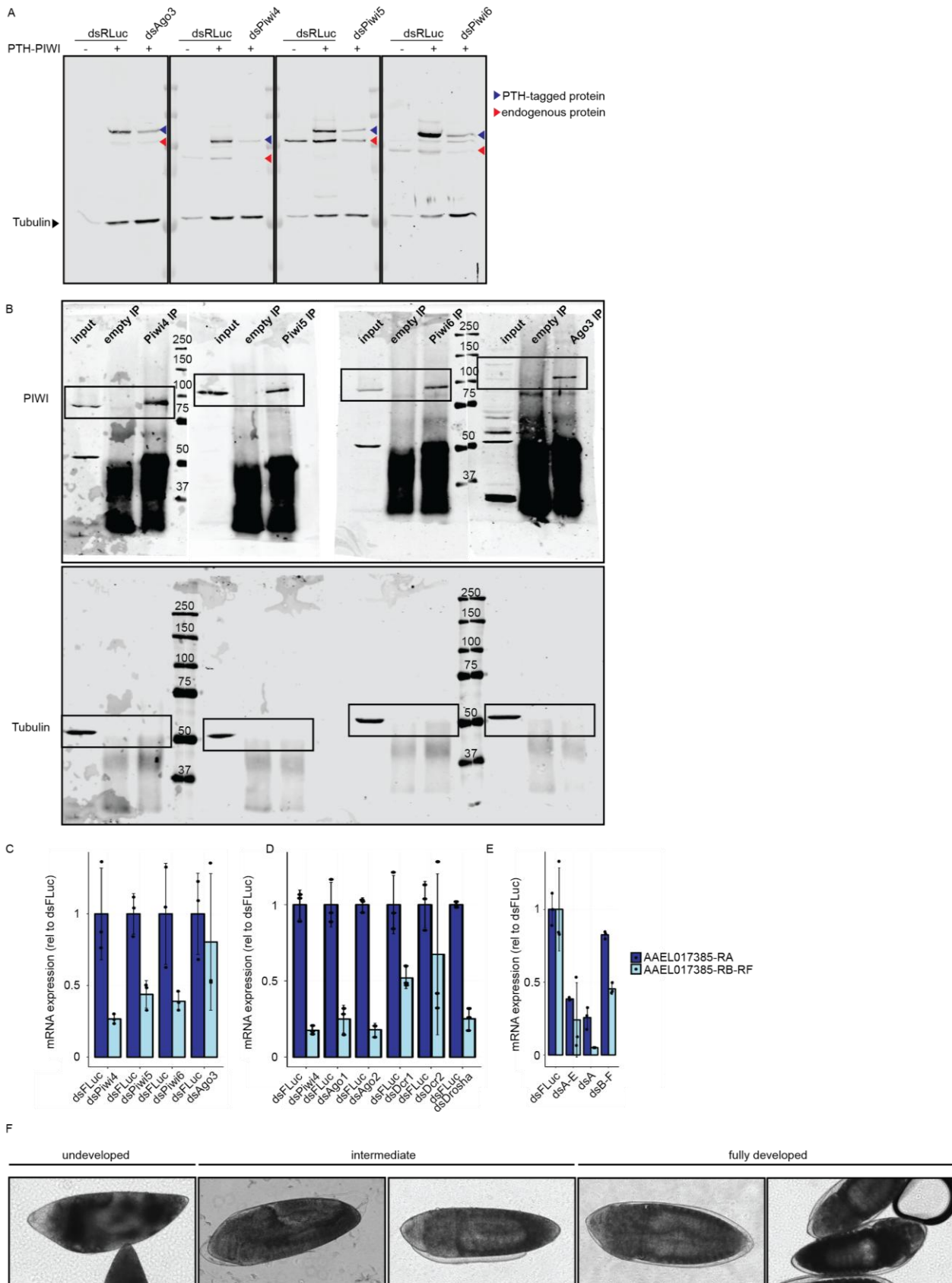
907 bromide-stained rRNA (B) were analyzed to verify equal loading.

908 (C) log<sub>2</sub> mRNA expression of transposable elements in embryos injected with *tapiR1*-specific or

909 control AO. Mean counts of five biological replicates are shown. Significance was tested at an FDR

910 of 0.01 and a log<sub>2</sub> fold change of 0.5. Diagonal lines indicate a fold change of two.

911



912

913 **Supplementary Data Figure S1:** Antibody validation, uncropped Western blot images,  
914 knockdown efficiencies, and scoring scheme for the development of *Ae. aegypti* embryos.

915 (A) Validation of *Ae. aegypti* PIWI antibodies. Specificity was confirmed by detection of an  
916 additional band in PTH-tagged PIWI-expressing Aag2 cells, and loss of signal upon dsRNA-

917 mediated knockdown. Knockdown with dsRNA targeting RLuc (dsRLuc) serves as negative  
918 control.

919 (B) Uncropped Western blot images corresponding to Extended Data Fig 2B.

920 (C-E) Knockdown efficiencies of PIWI genes shown in Extended Data Fig 2D (D), siRNA and  
921 miRNA pathway genes shown in Extended Data Fig 2E (E), and AAEL017385 isoforms in the  
922 experiment shown in Extended Data Fig 3C (F)

923 (F) Representative images of embryos scored as either undeveloped, intermediate or fully  
924 developed at 2.5 days post injection with antisense RNA oligonucleotides.