"Embryonic piRNAs target horizontally transferred vertebrate transposons in assassin bugs"

Tarcísio Fontenele de Brito[1], Maira Arruda Cardoso[1], Nazerke Atinbayeva[3,4], Ingrid Alexandre de Abreu Brito[1], Lucas Amaro da Costa[1], Nicola Iovino[3] and Attilio Pane[1,2*]

Affiliations:

[1] Instituto de Ciências Biomédicas, Universidade Federal do Rio de Janeiro, 21941-902 Rio de Janeiro, Brasil.

[2] Instituto Nacional de Ciência e Tecnologia em Entomologia Molecular, Brasil.

[3]Department of Chromatin Regulation, Max Planck Institute of Immunobiology and Epigenetics, D-79108 Freiburg, Germany.

[4]Albert-Ludwigs-Universität Freiburg, Fahnenbergplatz, 79085 Freiburg im Breisgau, Germany

* Corresponding author

E-mail: apane@icb.ufrj.br (AP)

Keywords

Rhodnius, HTT, piRNA pathway, transposon, siRNA

Abstract

Piwi proteins and the associated Piwi-interacting RNAs (piRNAs) coordinate a surveillance system that protects the animal genome from DNA damage induced by transposable element (TE) mobilization. Here, we investigated the adaptation of the

piRNA pathway to horizontally transferred transposons (HTTs) in the assassin bug *Rhodnius prolixus*, a primary vector of Chagas disease. *Rhodnius* acquired specific classes of HTTs by feeding on bats, opossums and squirrel monkeys. By analyzing the temporal dynamics of piRNA cluster expression and piRNA production during critical stages of *Rhodnius* development, we show that peak levels of ~28 nt long piRNAs correlate with reduced HTT and resident TE expression primarily during embryogenesis. Strikingly, while resident TEs piRNAs seem to engage in a typical ping-pong amplification mechanism, sense and antisense HTT piRNAs instead overlap by ~20 nt or do not display ping-pong signatures. These features are explained at least in part by the low number of HTT copies inserted into the piRNA clusters and might point to a non-canonical mechanism of biogenesis. Our data reveal that the piRNA, but not the siRNA pathway, responded to HTTs that were recently transferred from vertebrate tetrapods to a hematophagous insect of medical relevance.

Introduction

Transposable elements (TEs) are selfish genetic elements that can mobilize and insert in new positions of eukaryotic genomes. These sequences often comprise a substantial proportion of the host genome reaching more than half of the human genome and 90% of the maize genome [1–3]. Although TEs are usually passed on to the next generations via vertical transmission, a growing body of evidence is showing that horizontal transposons transfer (HTT) is more widespread than previously appreciated and occurs not only between closely related species, but also across distant phyla [4–6]. One of the most remarkable examples is represented by the horizontal transfer of DNA transposons between the triatomine insect *Rhodnius prolixus* and vertebrates, which occurred within the past 50 millions years [7]. *Rhodnius* is a hematophagous insect

51 species and a primary vector of *Trypanosoma cruzi*, the causative agent of Chagas
52 disease. Transposons of the *SPACE INVADERS* (*SPIN*), *OposCharlie1* (*OC1*), *hAT1*
53 and *ExtraTerrestrial* (*ET*) in the genome of this insect were found to share up to ~98.1%
54 sequence identity with those described in opossums, squirrel monkeys and bats of
55 South America [7]. *Rhodnius* likely acquired these HTTs because of its blood-feeding
56 habit and the broad range of vertebrate hosts. The same habit is responsible for the
57 transmission of *T. cruzi* to humans and underlies the etiology of Chagas disease.
58 *Rhodnius* also harbors other promiscuous transposons like *RTE-X*, *Gypsy, Helitron, I,*
59 *Maverick* and *Mariner*, with the latter being the predominant family in the *Rhodnius*
60 genome [8]. These TEs are often transmitted horizontally across the animal kingdom and
61 even between animals and plants [9]. How cells respond to newly invading transposons
62 has been investigated in the fruit fly *Drosophila melanogaster*. The best characterized
63 example of HTT involves the horizontal transfer of P-elements from *Drosophila willistoni*
64 to *Drosophila melanogaster* and their subsequent spreading across natural *D.*
65 *melanogaster* populations over the second half of the 20th century [10]. P-element
66 invasion of a naive genome was soon associated with a class of developmental defects
67 collectively known as hybrid dysgenesis [11,12]. Hybrid dysgenesis leads to sterility but
68 only manifests itself in the progeny that paternally inherits the P-element. The discovery
69 of the piRNA pathway and its crucial role as a defense system against transposon-
70 induced DNA damage allowed to shed light on the molecular underpinnings of hybrid
71 dysgenesis [13]. Piwi-interacting RNAs or piRNAs are a class of small non-coding RNAs
72 ranging in size from 18 to 30 nucleotides that are typically found in complex with PIWI
73 proteins [14–16]. The pathway has been best described in *Drosophila melanogaster* and a
74 few other holometabolous insect species [17]. In *Drosophila* somatic cells, piRNA
75 precursor transcripts originate from regions of the genome harboring a variety of
76 transposon remnants known as piRNA clusters [18,19]. These piRNA precursors are

77 processed in cytoplasmic organelles known as Yb-bodies to generate mature mostly
78 antisense piRNAs [20]. In turn, antisense piRNAs are loaded into the Piwi protein that
79 translocates to the nucleus and employs them as guides to identify and silence actively
80 transcribed transposons in the genome [21,22]. Among the primary targets of the somatic
81 pathway are the *Zam*, *Gypsy* and *Idefix* transposons [19,23,24]. The transcription of the
82 germline piRNA clusters instead occurs from both genomic strands (i.e. dual-strand
83 piRNA clusters) [19] and is regulated by a dedicated machinery [25–28]. Cluster transcripts
84 produce primary antisense piRNAs as in the somatic cells [29]. However, a fraction of
85 these RNAs promotes a feedforward amplification loop known as the "ping-pong" cycle,
86 that is regulated by Aubergine (Aub) and Argonaute3 (Ago3) [18,19] . The loop starts when
87 Aub bound to primary antisense piRNAs elicits the formation of sense piRNAs by
88 cleaving the transcripts of active transposons. In turn, sense piRNAs are bound by
89 Ago3, which can now recognize cluster transcripts to produce more antisense piRNAs.
90 The slicing activities of Aub and Ago3 leave a typical hallmark in the secondary piRNA
91 complement known as "ping-pong signature", whereby sense and antisense piRNAs
92 overlap by 10 nt at their 5' end. Also, because antisense piRNAs display a 5' Uracil bias
93 (1U), the sense piRNA mates are enriched with an Adenine residue at the 10th position
94 (10A). Antisense piRNAs associate with Piwi also in the germline and guide the protein
95 to the nucleus. In addition to the piRNA pathway, endogenous ~22 nt long siRNAs
96 (siRNAs) have also been implicated in transposon downregulation in *Drosophila* [30,31].
97 siRNAs are generated through the cleavage double-strand RNA molecules by the
98 ribonuclease Dicer2 (Dcr2) and loaded into Argonaute2 (Ago2) to form the RISC
99 complex [32,33]. Transposon-related siRNAs can originate from piRNA clusters or from
100 transposable elements [30]. In *Drosophila*, the siRNA pathway downregulates certain TEs
101 in the somatic cells, while the piRNA pathway is crucial in germline tissues.

102 Since the discovery of horizontally transmitted P-elements in *Drosophila* species, a
103 wealth of studies have reported the exchange of transposable elements between non-
104 hybridizing insect species [5,12,34]. The piRNA pathway has been proposed to coordinate
105 an adaptive defense mechanism against newly invading transposons [35]. A naive
106 genome would not express piRNAs against a new TE, which can therefore insert in
107 several positions in the genome. However, when a transposon copy eventually
108 integrates in a piRNA cluster, it becomes a source of piRNAs that guide the PIWIs to
109 silence in *trans* the active elements dispersed in the genome. The observation that
110 piRNAs are maternally loaded in the developing eggs also pointed out a mechanism to
111 explain the hybrid dysgenesis syndrome [13]. The progeny of a cross between a female fly
112 bearing a P-element and a naive male will develop normally because it inherits both the
113 TE and the cognate piRNA set. In contrast, the reciprocal cross will prove sterile
114 because the P-element is transmitted by the father to a progeny that lacks the
115 appropriate piRNA set. Aside from the Drosophilids however, much less is known about
116 the mechanisms that allow insect species to repress newly invading transposons. The
117 piRNA pathway is evolutionarily conserved in animals, but remarkable differences can
118 be observed even when the investigation is restricted to species of the same genus
119 [17,36,37]. For instance, it was recently shown that *D. eugracilis*, which diverged only 10
120 million years ago from *D. melanogaster*, neither uses a ping-pong amplification
121 mechanism nor it expresses the critical YB protein [36]. The expression patterns and size
122 distribution of the piRNAs also varies when *D. melanogaster* is compared to other
123 holometabolous insects like *Tribolium casteneum*, *Bombyx mori* and *Apis mellifera* [38 39–
124 41]. In hemimetabolous insects the characterization of the pathway is still in its infancy,
125 but it was shown that *Blattella germanica* expresses ~28 nt piRNAs during
126 embryogenesis as well as stage-specific piRNA pools during post embryonic
127 development [39]. The investigation of the piRNA pathway in more basal insects therefore

128 might reveal novel mechanisms and functions and shed light on the arms race between
129 transposons and related defense systems.

130 *Rhodnius prolixus* is a hemimetabolous insect belonging to the Triatomine subfamily
131 and undergoes 5 nymph stages during development before turning into a fertile winged
132 adult [42]. Blood feeding is not only necessary for ecdysis, but also to trigger oogenesis in
133 adult females. It was estimated that between 19% and 23% of the *Rhodnius* genome is
134 composed of transposon sequences with the *Mariner* family accounting for a large
135 fraction of the mobilome [8,43,44]. The genome of this hemipteran insect harbors four *PIWI*
136 genes and we have previously shown that *Rp-piwi2*, *Rp-piwi3* and *Rp-ago3*, but not *Rp-*
137 *piwi1*, are expressed in ovaries and are necessary for female adult fertility [45].
138 Furthermore, our transcriptomic analyses revealed that orthologs of the *Drosophila*
139 piRNA and siRNA pathways components are also expressed in *Rhodnius* ovaries [46]. In
140 this study, we characterize the temporal dynamics of the piRNAs and their source *loci* in
141 the *Rhodnius* genome. Importantly, we provide the first evidence of the ability of the
142 piRNA pathway to respond and adapt to HTTs transmitted from vertebrates to a
143 hematophagous insect responsible for the transmission of Chagas disease.

144

145

146 **Results**

147 **Small RNA profiling in *Rhodnius prolixus***

148 In order to investigate the dynamics of piRNA expression in oogenesis and early
149 stages of *Rhodnius* development, we generated and sequenced small RNA libraries
150 from embryos (Emb) and 1st instar nymphs (Nym) and analyzed them together with
151 small RNA datasets from previtellogenic stages of oogenesis (PVS) and mature eggs
152 (Egg), that we previously produced [47] (Fig. 1A). After trimming and quality checking, we
153 mapped the reads to the RproC3 version of *Rhodnius* genome available at VectorBase
154 [44,48]. The percentage of reads that mapped to the genome ranged from ~86% for the
155 PVS1 replicate to ~95% for the Egg2 replicate (Supplementary Table 1). The length
156 distribution of the reads reveals a bimodal distribution with two major peaks at ~22 nt
157 and ~28 nt (Fig. 1B). The former is more prominent in PVS, Egg1 and Nym samples,
158 while the latter is more apparent in Egg2 and Emb samples. To gain insight into the
159 complexity of *Rhodnius* small RNAs, we contrasted our small RNA datasets against
160 *Rhodnius* annotated features available at VectorBase [48]. As expected, a substantial
161 proportion of the mapped reads (i.e. 36.16% for PVS and 65.52% for Nym), match
162 annotated miRNAs (Fig.1C). The abundance of reads corresponding to repetitive
163 sequences shows a reciprocal distribution when compared to the miRNAs. In fact, this
164 category is highly expressed in embryos, while it's progressively lower in PVS (12.07%)
165 and Nym (10.6%). Once again, the Egg replicates display two non-overlapping
166 distributions with Egg2 being more comparable to the Emb replicates. Other genomic
167 features, including tRNAs, are below 8% in all the samples. A fraction of the reads,
168 ranging from ~12% in Nym to ~34% in Emb, could not be classified based on annotated
169 features (i.e. "unassigned"). Next, we profiled the length distribution of the reads
170 associated with transposons and repetitive sequences (Fig. 1D). All the stages analyzed

171 displayed an apparent peak at 21-22 nt, that likely comprises siRNAs, and a ~28 nt
172 peak corresponding to piRNAs. The abundance of these classes of small RNAs varies
173 in the different samples. The ~22 nt population is predominant in PVS, while the ~28 nt
174 is dramatically expressed in embryonic stages and largely exceeds the ~22 nt set (Fig.
175 1D). In nymph stages, the levels of ~22 nt and ~28 nt small RNAs seem comparable.
176 Once again, the Egg1 and Egg2 replicates provided two different size distributions. Only
177 Egg1 is compatible with maternally loaded piRNAs generated in PVS, while the second
178 replicate seems more similar to the profile obtained in embryos. The easiest explanation
179 for this observation is that the eggs dissected from the adult females used for the Egg2
180 replicate included a proportion of developing embryos that were retained by the mother.
181 We decided to keep these replicates as they might help shed light on the very initial
182 stages of embryogenesis.
183 It was shown that four DNA transposon families in *Rhodnius*, that is *SPIN, OC1, hAT*
184 and *ET* originated by horizontal transfer from vertebrate tetrapods [7]. We therefore
185 wondered whether piRNAs related to these HTTs are expressed in *Rhodnius*. To
186 answer this question, we attempted to isolate all the RNAs with 18-30 nt matching HTTs
187 sequences. Strikingly, we find both sense and antisense piRNAs for 6 HTTs reported
188 by Guilbert and collaborators (Fig. 1E) [7]. In addition to the HTTs, 46 transposon
189 families produce both sense and antisense piRNAs with 14 families accounting for more
190 than 90% of the piRNAs. Among them, we could detect abundant piRNAs *LTR/Gypsy*,
191 the LINE elements *Jockey* and *RTE-X*, the rolling-circle *Helitron* and, as expected, for
192 DNA transposons of the *Mariner* family (Fig. 1E). The latter explains 8.9% of the total
193 piRNAs. Among the top most piRNA-producing sequences, we find a class of
194 "Unknown" elements, which might comprise novel transposons or repetitive sequences.
195 In agreement with the general dynamics of the piRNA population in the different stages
196 analyzed, the expression of both sense and antisense piRNAs for the HTTs and other

197 resident TEs appears to dramatically increase during embryogenesis, with the antisense
198 population being more prominent (Fig. 1D and E).
199 We then analyzed the expression patterns of TEs in oogenesis and early *Rhdonius*
200 development (Fig. 1F). To this aim we combined ovarian and embryonic RNA-Seq
201 datasets that we and others previously generated together with newly sequenced
202 transcriptomes of 1st instar nymphs [46,49]. The temporal expression dynamics of the TEs
203 appear to inversely correlate with piRNA abundance in each stage. The highest TEs
204 expression levels are observed in nymphs, while the lowest are detected in embryonic
205 stages concomitant with peak levels of piRNAs. These results suggest that the 28 nt
206 piRNA population might drive TE silencing during *Rhodnius* embryogenesis, while TE
207 mobilization might be more tolerated in postembryonic development.
208
209 **piRNAs in *Rhodnius prolixus* are produced from uni-strand clusters**
210 In order to identify the piRNA source *loci* in *Rhodnius*, we narrowed down our search to
211 piRNAs that map uniquely to the genome [19]. For this study, we adopted a new version
212 of the *Rhodnius* genome generated with the Hi-C technique and available at
213 (https://www.dnazoo.org/assemblies/Rhodnius_prolixus). The Hi-C genome comprises
214 11 major scaffolds that likely represent the X chromosome and the 10 autosomes of
215 *Rhodnius*. Scaffold 10 was proposed to actually correspond to the X chromosome. We
216 found that unique piRNAs clusterize at 1276 different positions in the genome mainly
217 along the major scaffolds 1-11 (1176) (Fig. 2A). The top 10 clusters account for ~74% of
218 the total unique piRNAs expressed in *Rhodnius* embryogenesis (Fig. 2B). Only 16
219 RPCLs (i.e. *Rhodnius prolixus* clusters) display RPM > 100 and 6 of them display RPM
220 > 1000. These RPCL are hosted on scaffold 6 (RPCL1, RPCL3 and RPCL4), scaffold
221 11 (RPCL2) and the putative X chromosome scaffold 10 (RPCL5 and RPCL6). We find
222 that transposons of the *LINE RTE-X*, *Jockey* and *Gypsy* families together with *Mariner-*

223 like DNA transposons are the most abundant in the top 10 piRNA clusters
224 (Supplementary Fig. 1). It is worth noting however that the most abundant sequence
225 belongs to the "unknown" category and deserves further investigation in the future. The
226 clusters vary in size from a few kilobases to over 700 kb for the RPCL1 locus (Fig. 2C).
227 For the major RPCL1 cluster, we can observe a clear anticorrelation between the levels
228 of piRNAs and cluster transcripts (Fig. 2C). High piRNA levels in embryos coincide with
229 low levels of cluster transcripts. Instead, cluster transcripts appear to accumulate in
230 PVS and, especially, in nymphs, where piRNA expression is at the lowest. Similar
231 expression patterns are observed for the majority of the piRNA clusters (Fig. 2D).
232 Next, we investigated the expression of the cluster-derived piRNAs in PVS, Egg, Emb
233 and Nym stages (Fig. 3). Once again, we took into account only uniquely mapping
234 piRNAs. Consistent with the general piRNA temporal profile, the top 10 piRNA clusters
235 display a stark increase in the expression of piRNAs during embryonic stages (Fig. 3A
236 and 3B). Importantly, these *loci* give rise to sense and antisense piRNA populations with
237 the latter generally being more abundant (Fig. 3A and 3B). The alignment of the unique
238 mappers demonstrates that the majority of the clusters are of the uni-strand type (Fig.
239 3C). Among the major clusters, we were unable to unambiguously identify dual-strand
240 or bidirectional *loci* in *Rhodnius*. RPCL2, which spans over 200 kb, initially appeared to
241 be a bidirectional *locus* in that part of the cluster seems to be expressed from (+) strand
242 and the other from the (-) strand of the genome (Supplementary Fig. 2). Yet, upon
243 closer inspection we realized that the region between the two parts of the cluster with
244 opposite profiles was only partially sequenced. Thus, RPCL2, like other instances of
245 seemingly bidirectional clusters, might represent a genome assembly error, but a more
246 precise characterization will require an improved assembly of the *Rhodnius* genome.
247 We then looked at the orientation of the transposons within the top 20 piRNA clusters
248 (Fig. 3D). Different from the *D. melanogaster flam* locus, where transposon fragments

249 are polarized, *Rhodnius* clusters harbor transposons on both genomic strands. Thus,
250 these *loci* can produce both sense and antisense piRNA, even though transcription
251 occurs from one strand of the genome.
252 Because Rhodnius embryos produce piRNAs targeting the HTTs, we wondered whether
253 these transposons or their fragments are inserted into the piRNA clusters. To answer
254 this question, we aligned the *SPIN*, *hAT*, *OC1* and *ET* sequences with the *Rhodnius*
255 HiC genome and counted the number of occurrences within the piRNA clusters. In
256 agreement with a recent horizontal transfer from vertebrates, the number of HTT copies
257 in the clusters is low and ranges from less than 5 for *ET* and *SPIN* to ~40 for *OC1* (Fig.
258 3E). The observation that HTTs sequences are inserted in the major clusters albeit with
259 very low copy numbers prompted us to ask whether these copies are expressed and
260 produce piRNAs. We observe that the *SPIN*, *OC1* and *hAT* sequences generate both
261 sense and antisense piRNAs, while the *ET* element seems to produce mostly sense
262 piRNAs (Fig. 3F-I). These piRNAs however represent only a fraction of the HTT
263 piRNAs. When we compare the cluster-derived piRNAs for resident TEs and HTTs, we
264 observe a striking difference. While for the resident TEs between ~60% and ~80% of
265 the piRNAs originate from the clusters (Fig. 3J), less than 15% of the HTT piRNAs are
266 produced from these loci (Fig. 3K). Thus HTT piRNAs are mostly generated by the
267 transposon copies scattered in the genome.
268

269 **HTT piRNAs display unique features**

270 In order to investigate the characteristics of the HTT piRNAs, we first profiled the
271 piRNAs along the transposon sequences. Their distribution is not homogenous and they
272 seem to accumulate at certain positions (Fig. 4A-D). We then investigated the overlap
273 between the 5' ends of sense and antisense sequences for the paired piRNAs (Fig. 4A'-
274 D'). The 10 nt overlap typical of the ping-pong mechanism is either modest or it is

275 confused among other peaks, as in the cases of *OC1* and *ET* (Fig. 4A' and 4B').
276 Instead, the *SPIN* and *hAT* piRNA mates display an apparent ~20 nt overlap (Fig. 4C'
277 and 4D'), that is characteristic of siRNA heteroduplexes produced by the activity of the
278 Dcr2 enzyme. Yet, these piRNA pairs are over 28 nt long. We next investigated the
279 nucleotide distribution frequencies along the HTT piRNAs. We find that both sense and
280 antisense piRNAs display a 1U bias, while the 10A bias in sense piRNAs is less
281 apparent (Fig. 4A"-D"). The most abundant piRNA pairs for *hAT* and *Spin* further
282 underscore the unique properties of the HTTs piRNAs (Fig. 4C-D inset). These pairs
283 comprise piRNAs of over 28 nt in length, which overlap by 19 nt and 21 nt respectively.
284 Remarkably both sense and antisense piRNAs display 1U and 10A residues for the *hAT*
285 pair, while only the sense *Spin* piRNA has a 10A residue.
286

287 **piRNAs in *Rhodnius prolixus* display ping-pong signatures**

288 The HTTs have been suggested to have invaded the *Rhodnius* genome over the past
289 50 million years. Thus, the piRNAs associated with these elements likely display
290 peculiar characteristics that are not shared by other resident TEs. To shed light on this
291 aspect, we filtered all the small RNA sequences ranging from 24 to 30 nt in length,
292 which yielded ~36 million putative piRNA sequences (Supplementary Table 2). We then
293 investigated whether a ping-pong amplification system is active during *Rhodnius*
294 development. (Fig. 5). First, we estimated the fraction of paired piRNAs in all the stages
295 and found that it ranged from ~19% in PVS to ~29% in Embryos (Fig. 5A). As expected
296 the proportion of unpaired piRNAs is more abundant and likely comprises primary
297 piRNAs. Different from the HTTs, we found that a substantial fraction of the pairs
298 (~37%) overlap by exactly 10 nt at their 5' ends (Fig. 5B). The analysis of nucleotide
299 distribution shows that antisense piRNAs display a clear 1U bias in all the samples and

300 that sense piRNAs are enriched with a 10A residue (Fig. 5C). Sense piRNAs also bear
301 a 1U bias, which was also described in *Tribolium castaneum*, although its meaning
302 remains unclear. Thus, piRNAs engage in a ping-pong amplification mechanism in
303 *Rhodnius,* which is not restricted to the germline as in *D. melanogaster*, but is also
304 active in embryos and 1st instar nymphs.

305 The discovery of a ~28 nt piRNA population in embryos and their involvement in
306 a ping-pong amplification loop prompted us to investigate what *Rp-PIWI* genes are
307 expressed in this stage as well as in first instar nymphs. We have previously shown that
308 *Rp-piwi2*, *Rp-piwi3* and *Rp-ago3*, but not *Rp-piwi1* are expressed in and important for
309 oogenesis in *Rhodnius* [45]. In order to expand this analysis, we combined *in silico* and
310 wet biology analyses (Fig. 4D and E). First, we interrogated our RNA-Seq datasets
311 together with those produced by Pascual and Rivera-Pomar [49]. Consistent with our
312 previous studies, *Rp-piwi2*, *Rp-piwi3* and *Rp-ago3* are expressed in *Rhodnius*
313 oogenesis (Fig. 4D) [45]. Interestingly, these genes are also upregulated in the first 48h of
314 embryonic development, while their levels drop substantially once the 1st instar nymphs
315 hatch (Fig. 4D). During postembryonic development, the levels of *Rp-piwi3* are below
316 detectable levels. *Rp-piwi1* is apparently not expressed in any of the stages analyzed.
317 In order to validate these observations, we performed RT-qPCR assays on total RNA
318 extracted from PVS, Egg, Emb (0-48h) and Nym stages. We used oligonucleotide pairs
319 specific for each *Rp-PIWI* gene and for *Rp-rp49*, which served as internal control (Fig.
320 5E) [45]. The Fold Changes of each gene are in line with RPKM values. Interestingly, the
321 *Rp-piwi2* levels dramatically increase in embryogenesis compared to the other stages
322 analyzed and largely exceed those of the remaining *Rp-PIWIs* pointing to an important
323 role in the biogenesis and/or function of the ~28nt piRNA population. Of note, the
324 expression of *Rp-piwi1* as per RT-qPCR is low, but clearly detectable in 1st instar
325 nymphs (Fig. 5E). Whether *Rp-piwi1* transcripts exert any functions in postembryonic

326 development remains to be elucidated. Our results reveal that *Rp-PIWI* genes are
327 therefore temporally regulated during oogenesis and early *Rhodnius* development.
328
329 **TE siRNAs are predominantly expressed in oogenesis**
330 Although the piRNA pathway represents a major surveillance system keeping mobile
331 elements in check, it was shown that Dcr2-dependent siRNAs control transposon
332 silencing in *Drosophila* somatic cells [50]. To investigate whether this mechanism is
333 conserved in *Rhodnius*, we interrogated the transposon-related sequences in the 18 to
334 23 nt size range and analyzed the ping-pong signal, nucleotide bias and strandness
335 (Fig. 6). We find a clear population of sense and antisense small RNAs centered on ~22
336 nt in length that is particularly abundant in PVS and Egg compared to the remaining
337 samples (Fig. 6A and B). Surprisingly, the two Egg replicates display remarkably similar
338 siRNA profiles. Together with the piRNA analysis, this observation suggests that the
339 Egg2 replicate might be capturing the early embryonic stages where the siRNA levels
340 are still high, but the piRNA production from the zygotic genome is building up. The
341 degree of overlap between sense and antisense sequences reveals that the ~22 nt
342 small RNAs expressed in PVS, Egg and to a lesser extent in Emb and Nym are siRNAs,
343 while Emb samples comprise ~22 nt piRNAs (Fig. 6C). Consistent with this, the
344 nucleotide distribution shows that the 1U together with a modest 10A bias are detected
345 only in embryos, but not in the other stages analyzed (Fig. 6D). Finally, we wondered
346 whether TE siRNAs can originate from the piRNA clusters. The analysis of uniquely
347 mapping reads reveals that a substantial proportion of siRNAs can indeed be produced
348 by these *loci* in all the stages (Fig. 6E). Accurate quantification of siRNAs produced
349 from clusters versus TEs or other regions of the genome is confounded by the presence
350 of piRNAs in the same size range. Surprisingly, we do not find sequences
351 corresponding to the HTTs in the siRNA population. These observations reveal that

352 siRNAs and piRNAs expression patterns are temporally regulated and partially overlap
353 during oogenesis and early stages of *Rhodnius* development.
354

355 **Discussion**

356 Studies in *Drosophila melanogaster* and its sibling species highlighted the adaptability
357 of the piRNA pathway to newly invading transposons [51]. Aside from the Drosophilids
358 however, how animal cells respond to TE invasion is still largely unknown. In this study,
359 we show a striking example of piRNA pathway adaptation to TEs that were transferred
360 from vertebrate tetrapods to the assassin bug *Rhodnius prolixus*. *Rhodnius* belongs to
361 the Triatomine subfamily and has been firmly connected to the transmission of Chagas
362 disease. It was shown that the *SPIN*, *OC1*, *hAT* and *ET* elements were horizontally
363 transmitted to *Rhodnius* from bats, opossums and squirrel monkeys [7]. Because of the
364 hematophagous habit of *Rhodnius* however, HTT might have occurred while the insect
365 was feeding on the blood of the vertebrate hosts. This event provided the unique
366 opportunity to investigate the response of the piRNA pathway to transposons
367 exchanged between two different phyla in the past 50 million years. When we compare
368 the complement of piRNAs targeting resident transposons and HTTs in *Rhodnius*, we
369 find some similarities, but also puzzling differences. A major population of ~28 nt long
370 piRNAs is expressed at high levels during *Rhodnius* embryogenesis and comprises
371 sequences of 46 transposon families. Some examples are the *Gypsy*, *RTE-X*, *Helitron*,
372 *Maverick*, *I*, *Penelope* and the abundant class of *Mariner* transposons [8]. We have
373 referred to all these transposons as resident TEs, but it should be noted that they are
374 frequently involved in horizontal transfer events. For these elements, we find primary
375 piRNAs with the typical 1U bias and secondary piRNAs with Adenine at 10th position
376 (10A) for the sense piRNAs. Also, consistent with an active ping-pong amplification

377 mechanism, sense and antisense piRNA mates overlap by 10 nt. Similar to more basal
378 holometabolous and hemimetabolous insects, but unlike *D. melanogaster*, the ping-
379 pong signal is not restricted to the gonads in *Rhodnius*, but it is especially active in
380 embryonic and in post-embryonic developmental stages. The piRNA complement
381 targeting the HTTs instead displays unique properties. Even though HTT piRNAs are
382 ~28 nt long and display a 1U bias, the ping-pong mechanism is ambiguous. The 10 nt
383 overlap is modest or confused among other signals. Instead, the most abundant *SPIN*
384 and *hAT* piRNA mates display a ~20 nt overlap at their 5' ends. Yet, the 10th position of
385 the sense piRNAs is an Adenine as observed for canonical ping-pong piRNA pairs.
386 These observations suggest that the biogenesis of HTT piRNAs in *Rhodnius* differs
387 from the more canonical mechanism that drives resident TE piRNA maturation and that
388 was described in detail in the fruit fly *D. melanogaster*. The lack of ping-pong signature
389 for the HTTs might reside at least in part in the genomic sources of their piRNAs. In this
390 study, we show that piRNA clusters in *Rhodnius* are mostly of the uni-strand type, but
391 unlike the *Drosophila flam* locus [19], their TEs are not polarized. For this reason,
392 *Rhodnius* clusters can produce both sense and antisense piRNAs, even though the
393 transcription occurs from one genomic strand. We find that more than 50% of the
394 unique piRNAs targeting resident TEs originate from the piRNA clusters, while these
395 *loci* account for less than 15% of the HTT piRNAs. This observation is consistent with
396 the low number of HTT copies inserted into the clusters. These sequences therefore,
397 might not produce sufficient levels of antisense transcripts to robustly activate the ping-
398 pong mechanism. The *ET* element provides an extreme case, whereby all the cluster
399 copies generate sense piRNAs, thus preventing the production of secondary piRNAs.
400 How the *hAT* and *SPIN* piRNA mates overlapping by ~20 nt are generated and what
401 factors are involved in the mechanism remains to be elucidated. Interestingly, our
402 results reveal that *Rhodnius* expresses siRNAs for resident TEs, but not for the HTTs.

403 Hence, primary piRNAs produced from HTT copies or fragments scattered in the
404 genome, rather than siRNAs, might provide the first line of defense against newly
405 invading transposons in this insect species.

406 TE siRNA are expressed from piRNA clusters and TE copies dispersed in the genome
407 during oogenesis and to a lesser extent during embryonic and postembryonic stages.
408 RNAi might contribute to reducing transposon transcripts during oogenesis, although it
409 is unable to completely silence these elements [46]. It was recently shown that maternally
410 provided siRNAs loaded in the developing eggs promote piRNA cluster activation in the
411 *Drosophila* embryo [52]. Unlike *Drosophila*, *Rhodnius* does not produce abundant piRNAs
412 during oogenesis. Thus in *Rhodnius* and other more basal insects, the siRNA
413 complement might be especially critical for jumpstarting piRNA production in early
414 embryonic stages. Transposon downregulation is subsequently taken over by the ~28 nt
415 piRNA pool during embryogenesis. Transposon transcripts maternally deposited in the
416 developing oocytes might feed the ping-pong amplification loop at this stage together
417 with cluster-derived antisense piRNAs. piRNA biogenesis seems to result in the
418 consumption of the cluster-derived piRNA precursors and the downregulation of TEs. In
419 post-embryonic stages instead both siRNAs and piRNAs are reduced, while TE and
420 piRNA cluster transcripts appear to accumulate, thus pointing to a minor sensitivity of
421 the 1st instar nymphs to transposon expression. In support of these conclusions, the
422 levels of piRNAs and siRNAs seem to inversely correlate with those of TEs and piRNA
423 cluster transcripts in all the stages analyzed (Fig. 6F).

424 *Rhodnius* harbors four *PIWI* genes and the knockdown of *Rp-piwi2*, *Rp-piwi3* and *Rp-*
425 *ago3* by parental RNAi results in partial or complete female adult sterility [45,53]. Our
426 assays show that these *Rp-PIWI* genes are expressed in embryos and nymphs with *Rp-*
427 *piwi2* being dramatically upregulated in embryos coinciding with peak levels of ~28 nt

428 piRNAs. Thus, *Rp-piwi2* likely exerts a crucial role in piRNA biogenesis or function.
429 Since Ago3 was shown to coordinate the ping-pong cycle together with Aub in
430 *Drosophila* [18,19], it is tempting to speculate that the protein products of the only *ago3*
431 ortholog in *Rhodnius* (i.e. *Rp-ago3*) and *Rp-piwi2* genes might regulate a similar
432 mechanism in *Rhodnius* embryos. However, *Rp-ago3* might actually act in concert with
433 *Rp-piwi3* in this mechanism. The observation that *Rp-piwi3* expression levels are low in
434 oogenesis and almost undetectable in nymphs, where the ~28 nt piRNAs are barely
435 produced, seem to lend support to this hypothesis. The cell lines that were recently
436 developed from *Rhodnius* embryos will likely provide a valuable tool to dissect the
437 function of the different *Rp-PIWI* genes in piRNA biogenesis and transposon silencing
438 as well as shed light on the transcriptional regulation of the piRNA clusters [54].

439 Horizontal transposon transfer has been well characterized in Drosophilids,
440 where numerous instances of P-element transfer from *D. willinstoni* and *D. simulans* to
441 *D. melanogaster* were described [55,56]. Fruit flies therefore provided the opportunity to
442 dissect the response of the piRNA pathway to P-element invasion [57]. To our knowledge
443 however, this is the first study to link piRNAs to HTTs exchanged between vertebrates
444 and an hematophagous insect of medical interest. Studies in *Rhodnius* will therefore
445 help shed light on the evolution of the piRNA pathway and on the initial steps of the
446 arms race between invading TEs and small RNA-based defense systems.

448 **Materials and Methods**

449 **Ethic statement and insect handling**

450 All the experimental protocols involving animals were carried on in accordance with the
451 guidelines of the Committee for Evaluation of Animal Use for Research (Universidade
452 Federal do Rio de Janeiro, CAUAP-UFRJ) and the NIH Guide for the Care and Use of

453 Laboratory Animals (ISBN 0-309-05377-3). Protocols were approved by CAUAP-UFRJ
454 under registry #IBQM155/13. Specialised technicians at the animal facility (Instituto de
455 Bioquímica Médica Leopoldo de Meis, UFRJ) handled the rabbit husbandry under strict
456 guidelines and the supervision of veterinarians.
457 Insects were maintained at controlled temperature (28°C) and relative humidity (70-
458 80%) and fed with rabbit blood at regular intervals to promote ecdysis and oogenesis.
459 After the first blood feeding, all adult insects were fed every 21 days. Dissection of the
460 adult females, where necessary, was carried on 7 days after the first or second blood
461 meal. All animal care and experimental protocols were performed according to the
462 ethics guidelines.
463
464 **Total RNA extraction and small RNA purification**
465 Total RNA for small RNA library preparation was extracted from 30 embryos and 20 1st-
466 instar nymphs for each biological replicate. Embryos were collected over a two-day
467 period spanning the initial stages of embryogenesis up to the formation of head and
468 thoracic segments [58]. Two biological replicates were produced for each developmental
469 stage. Ovaries were dissected in ice-cold 1X PBS and immediately stored on ice.
470 Previtellogenic tissues and mature eggs were teased apart under light microscopy and
471 subjected to total RNA extraction using Trizol Reagent (Invitrogen) as per
472 manufacturer's instructions. For the production of small RNA libraries, 20 µg of total
473 RNA was separated on polyacrylamide gel containing 15M Urea (Pane et al., 2011). A
474 gel slice corresponding to 18 to 30 nucleotide RNAs was manually retrieved from the
475 gel and the RNAs were eluted in 3M NaCl overnight. Finally, RNA was precipitated
476 using 0.7 volumes of Isopropanol and 1 ul Glycogen and resuspended in RNAse-free
477 water. Small RNA library preparation and sequencing were carried out at the Max
478 Planck Institute facility in Freiburg (Germany) with Truseq Small RNA library preparation

479 kit as per manufacturer's protocol (Illumina). Total RNA extracted from 1st instar nymphs
480 was also used to produce and sequence RN-Seq libraries (Macrogen, Korea). RNA-Seq
481 datasets corresponding to four embryonic stages of *Rhodnius* development were
482 produced by others and are available at NCBI-SRA [49] (SRR14606474, SRR14606475
483 and SRR14606473).
484
485 **Bioinformatic analysis**
486 Small RNA-Seq libraries were subjected to adaptor removal and trimming using
487 TrimGalore! [59]. Reads with length lower than 18 nt were filtered out. Trimmed reads
488 were mapped with Bowtie v1.2 [60] to *Rhodnius prolixus* reference genome available at
489 VectorBase [48] allowing up to 3 mismatches and no more than 50 mappings per read (-v
490 3 -m 50 -a --best --strata). Qualities were assessed with FastQC software [61], genomic
491 coverages were compared with deepTools [62] and final quality reports were generated
492 using multiQC [63]. Library characterization was performed using featureCounts from the
493 subread package [64], where counts to each annotated feature were normalized by
494 number of mappings and overlapping features (-O -M --fraction). Repeats and TEs
495 sequences were de novo identified using RepeatModeler v2.0.1 [65]and genomic regions
496 obtained with RepeatMasker v4.1.2 [66]. Sequences of the HTTs have been previously
497 published [7]. Reads overlapping with TEs, with length distribution of 24 to 32 nucleotides
498 and not overlapping with miRNAs, rRNAs, snRNAs, simple repeats or low complexity
499 regions were characterized as piRNAs. Sense and antisense piRNAs were separated
500 by overlapping piRNA mappings with TEs annotation using bedtools intersect and
501 forcing strandness (-s and -S). Fasta sequences were extracted using samtools [67] and
502 sequence logos constructed using ggseqlogo [68]. For ping-pong signal analysis, each
503 piRNA mapped on the positive strand of the genome was used to search for 5' to 5'
504 overlapping piRNAs and the size of the overlap counted for each match. For piRNA

505 cluster identification, the chromosome-length genome assembly of *Rhodnius*
506 (https://www.dnazoo.org/assemblies/Rhodnius_prolixus) was used and all the steps
507 described previously were applied to this version as well. First, annotations of a sliding
508 window with 5 kb in size and steps of 1 kb were generated using the makewindow tool
509 from bedtools v2.31 [69]. Then, bedtools coverage was used to access the uniquely
510 mapped piRNA coverage in each window and only windows with a density greater than
511 5 piRNAs/kb covering at least 12% of the window were retained. Finally, overlapping
512 windows were merged and merged windows with a distance less than 20 kb from each
513 other were combined to produce the final piRNA clusters. RNA-seq data were pre-
514 processed and aligned using snakePipes 62. Genes and transposable element read
515 counts were produced using featureCounts allowing fraction counts for overlapping
516 features. Final expression was calculated as Reads Per Kilobase per Million (RPKM).
517

518 **Quantitative Rt-PCR assays**

519 For quantitative Rt-PCR assays, 1 ug of total RNA extracted with Trizol (Invitrogen) was
520 submitted to DNAse treatment with Turbo DNAse free kit (Ambion) and cDNA synthesis
521 using the MultiScribe Reverse Transcriptase (Thermo Fisher Scientific) as per
522 manufacturer's instructions. The resulting cDNAs were used for qPCR assays with the
523 set of oligonucleotides previously described [45]. Roughly 50 ng of cDNA for each sample
524 was mixed with oligonucleotides specific to target genes and SYBR green reagent (Life
525 Technologies). RT-qPCR was carried on QuantStudio 7 Flex (ThermoFisher). Fold
526 change differences were plotted as $2^{-\Delta Ct}$.

527

528 **Acknowledgments**

529 We would like to thank Bernardo Carvalho, Pedro Lagerblad de Oliveira, Ana Cristina

530 Bahia Nascimento, Marcos Farina de Souza e Marcia Cury El-Cheikh for their constant
531 support. We are grateful to Eric Aguiar and members of the Functional Genomics
532 Laboratory for helpful suggestions and critical reading of the manuscript. We thank Kelli
533 Cristina Melquiades Mendes, Daniela Sodré Leal and Graciela Venturi for the invaluable
534 technical support.

535

536 **Funding Statement**

537 This work was supported by the National Counsel of Technological and Scientific
538 Development (CNPq) (428100/2018-0) (AP), the Research Support Foundation of the
539 State of Rio de Janeiro (FAPERJ) E-26/210.912/2019 (AP); E-26/010.002720/2019
540 (AP); E-26/010.001877/2015 (AP); E-26/201.703/2021 (TB), Wellcome Trust grant
541 207486Z17Z (AP/ABC), CAPES and INCT/Enem fundings (AP, TB, MAC, IB). N.A. is
542 supported by the Max Planck Society and IMPRS program. N. I. is supported by the
543 Max Planck Society; DFG:CRC992, Project B06; Behrens-Weise Stiftung; CIBSS - EXC
544 2189. Also, this project has received funding from the European Research Council
545 (ERC) under the European Union's Horizon 2020 research and innovation programme
546 (grant agreement No.819941) ERC CoG, EpiRIME. The funders had no role in study
547 design, data collection and analysis, decision to publish, or preparation of the
548 manuscript.

549

550 **Data availability**

551 Small RNA sequencing datasets from embryonic (Emb1 and Emb2 replicates) and 1st
552 instar nymphs (Nym1 and Nym2 replicates) were deposited at NCBI SRA database
553 (accession numbers).

554

555 **Author contributions**

556 Tarcísio Fontenele de Brito: Data curation, Formal analysis, Investigation, Methodology,
557 Validation, Visualization, Writing – original draft, Writing – review & editing. Maira
558 Arruda Cardoso: Data curation, Investigation, Methodology, Validation, Writing – original
559 draft, Writing – review & editing. Nazerke Atinbayeva: Investigation, Writing – review &
560 editing. Ingrid Alexandre de Abreu Brito: Investigation, Methodology, Writing – original
561 draft. Lucas Amaro da Costa: Methodology, Investigation, Writing – original draft. Nicola
562 Iovino: Data curation, Funding acquisition, Resources, Writing – original draft, Writing –
563 review & editing. Attilio Pane, Conceptualization, Data curation, Formal analysis,
564 Funding acquisition, Investigation, Methodology, Project administration, Resources,
565 Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

566

567 **Conflict of interests**

568 All the authors declare that they have no conflict of interest.

569

570 **Figure Legends and Tables**

571 **Figure 1. *Rhodnius prolixus* embryos express ~28 nt long piRNAs targeting**
572 **Horizontally Transferred Transposons (HTT) and resident TEs.** A) Developmental
573 stages of *Rhodnius prolixus* analyzed in the current study. For oogenesis, we
574 investigate the previtellogenic stage (PVS) comprising the tropharium and early follicles
575 independently of the mature chorionated eggs dissected from the female abdomen
576 (Egg). Embryos (Emb) were collected over a 48h period encompassing the cleavage
577 stage until the formation of the head and thoracic segments after gastrulation. 1st instar

578 nymphs (Nym) were also analyzed to investigate postembryonic development. B)
579 Length distribution of all mapped reads for each replicate displayed in RPM. The ~22 nt
580 and ~28 nt peaks were anchored to the X-axis (Length) for a better visualization. Y-axis
581 displays RPM values. For each sample, we produced two replicates, which were
582 analyzed separately. B) Small RNA profiling based on percentage of reads mapped to
583 annotated and unannotated (unassigned) regions of *Rhodnius* genome. miRNAs are
584 particularly abundant in oogenesis and postembryonic development (red), while
585 embryonic stages are dominated by repeat-associated small RNAs (green). C) Length
586 distribution of reads mapped to transposable and repetitive sequences of *Rhodnius*
587 displayed in RPM. A major peak comprising ~28 nt piRNAs is clearly detectable in the
588 two embryo replicates (Emb1 and Emb2) and in one of the mature egg replicates
589 (Egg2). This class of small non-coding RNAs is less abundant in previtellogenic stages
590 of oogenesis (PVS1 and PVS2 replicates), in one of the mature egg replicates (Egg1)
591 and in 1st instar nymphs (Nym1 and Nym2 replicates). D) Heatmaps displaying the
592 expression levels of sense and antisense piRNAs for each TE subfamily in PVS, Egg,
593 Emb and Nym stages. Sense and antisense piRNAs for the horizontally transmitted
594 transposons (HTT) ET, SPIN, OC1 and hAT and for the resident TEs are shown
595 separately. Colorbar values display $\log_{10}$(RPM + 1). F) Heatmaps showing TE
596 expression levels as per RNA-Seq in PVS, Egg, Emb and Nym stages. Expression
597 levels for the HTTs ET, SPIN, OC1 and hAT and of the resident TEs are displayed in
598 separate heatmaps. Colorbar values display RPKM. *Rhodnius* acquired the HTTs from
599 opossums, squirrel monkeys and bats (right). For the SPIN and hAT families, Gilbert
600 and coworkers reported two transposon sequences for SPIN and hAT, namely
601 SPIN_NA_13_RP, SPIN_NA_12_RP, hAT_NA_1_RP and hAT_NA_2_RP. One
602 transposon sequence was reported for ET and OC1, which were labeled ET_NA_1_RP
603 and OPSCHARLIE1_NA_1_RP, respectively. For simplicity, we refer to these

604 transposable elements as SPIN, hAT, OC1 and ET throughout the article. Images of
605 bat, squirrel monkey and opossum with Public Domain license were obtained from
606 Wikimedia Commons.

607

608 **Figure 2. piRNA clusters are a major source of piRNAs during *Rhodnius prolixus***
609 **embryonic development.** A) Ideograms of the larger scaffolds obtained from the HiC
610 genome assembly that likely represent the 10 autosomes and the X chromosome
611 (scaffold 10) of *Rhodnius*. Each ideogram is scaled to the size of each scaffold, with a
612 scale in megabases. Black vertical lines show the distribution of the piRNA clusters. The
613 major cluster RPCL1 is highlighted (black box).

614 B) Cumulative percentage of piRNAs in clusters for each replicate. The top 10 RPCLs
615 account for 74% of the total uniquely mapping piRNAs.

616 C) IGV [70] visualization of the RPCL1 cluster region on the HiC_scaffold_6. Uniquely
617 mapped piRNAs for each stage are profiled for each sample (small RNA-Seq profiles).
618 The Y-axis displays RPM (0 to 100 interval). Expression levels of the RPCL1 (RNA-
619 Seq) in all the stages are also displayed. Of note, we included in this analysis the RNA-
620 Seq datasets for embryonic stages 0h, 12h, 24h and 48h post egg laying produced by
621 Pacual and coworkers. The Y-axis displays the RPKM range 0-20. Orange bar
622 highlights the position of the RPCL1 cluster. The distribution of DNA and LTR
623 transposons together with LINE elements and unknown sequences along the RPCL1
624 region are reported with black boxes.

625 (D) Heatmaps showing the temporal regulation of the cluster-derived mature piRNAs
626 (Small RNA-seq) and piRNA precursor transcripts (RNA-Seq) in PVS, Egg, Emb and

627 Nym stages. Mature piRNA levels inversely correlate with the expression levels of
628 cluster transcripts suggesting that piRNA biogenesis consumes cluster-derived
629 transcripts during embryonic stages. Small RNA-seq values (left side of the panel) are
630 displayed as $\log_{10}$(RPM + 1), while RNA-Seq values (right side of the panel) are
631 expressed as RPM. Each line of the heatmap represents a piRNA cluster with the major
632 clusters at the top of the panels.

633

634 **Figure 3. *Rhodnius* piRNA clusters are uni-strand and produce piRNAs against**
635 **HTTs.** A-B) Heatmaps showing sense (A) and antisense (B) piRNA levels for the top 10
636 piRNA clusters in PVS, Egg, Emb and Nym stages. Colorbar displays $\log_{10}$(RPM + 1).
637 C) Scatterplot reporting the strandness of the piRNA clusters. piRNA clusters primarily
638 expressed from the +strand (blue dots) and -strand (red dots) of the genome are
639 displayed. piRNA clusters with apparently unbiased expression are indicated (dark grey
640 dots). A cluster was considered unbiased if less than 75% of its piRNAs could be
641 attributed to a specific strand of the genome. Y-axis values are expressed as RPM. The
642 top 10 piRNA clusters are highlighted. D) Scatterplot showing the orientation of the TE
643 fragment within each piRNA cluster. piRNA clusters harboring TEs predominantly on the
644 +strand (blue dots) or -strand (red dots) of the genome are reported. These categories
645 are mostly formed by small clusters with a low number of TEs. The vast majority of the
646 piRNA clusters are characterized by unbiased distribution of TEs (dark grey dots). The
647 top 10 piRNA clusters are highlighted. E) Number of HTT copies inserted into the piRNA
648 clusters. F-I) piRNAs uniquely mapping to the cluster HTTs. The ET copies (F) seem to
649 exclusively produce sense piRNAs, while OC1 (G), hAT (H) and SPIN (I) generate both
650 sense and antisense piRNAs. Y-axis displays RPM values. X-axis shows the replicates
651 for the PVS, Egg, Emb and Nym samples. J) Percentage of uniquely mapping piRNAs

652 for the resident TEs originating from within or outside the piRNA clusters. K) Percentage
653 of uniquely mapping piRNAs for the HTTs originating from within or outside the piRNA
654 clusters.

655

656 **Figure 4. HTT piRNAs do not engage in a canonical ping-pong amplification**
657 **mechanism.** A-D) piRNA profiling along the sequences of each HTT with sense
658 piRNAs displayed with positive RPM values (green) and antisense piRNAs with
659 negative RPM values (red). For the hAT and SPIN elements, the most abundant
660 sense/antisense piRNA pairs are shown. Of note, the sense piRNAs pairs do not
661 display the typical ping-pong signature, but overlap by ~20 nt. Yet, the sense mate for
662 each pair bears an Adenine residue at the 10th position (red box). A'-D') 5' to 5' overlap
663 length frequency reported as z-score values for each HTT. A"-D") Sequence logos
664 displaying nucleotide bias for sense and antisense piRNAs mapped to the HTTs.

665

666 **Figure 5. The ping-pong amplification loop is active in *Rhodnius prolixus*.** A)
667 Percentage of paired and unpaired piRNAs for PVS, Egg, Emb and Nym replicates. B)
668 5' to 5' overlap length frequency displayed in z-score for piRNAs mapped to resident
669 transposons for each stage. C) Sequence logo showing nucleotide bias for each
670 position of antisense and sense piRNAs mapped to transposable elements in all the
671 replicates. D) RNA-Seq analysis of the *Rp-PIWI* gene expression in oogenesis and
672 early *Rhodnius* development. Colorcode displays RPKM values. E) qRT-PCR analysis
673 of the temporal expression levels of the *Rp-PIWI* genes during oogenesis and early
674 stages of *Rhodnius prolixus* development. PVS=previtellogenic stage, Egg=mature
675 chorionated eggs, Emb=embryos and Nym=1st instar nymphs. Y-axis shows the mean

676 Fold Change. Error bars display standard deviation from the mean.

677

678 **Figure 6. piRNA clusters generate ~22nt siRNAs in *Rhodnius prolixus*.** A-B)
679 Heatmaps showing sense (A) and antisense (B) ~22 nt long siRNAs mapped to the
680 *Rhodnius* TE subfamilies. Colorbar displays $\log_{10}$(RPM + 1). siRNAs appear to be
681 expressed at higher levels in oogenesis compared to embryonic and postembryonic
682 stages. C) Z-scores of the 5' to 5' overlap length frequency for PVS, Egg, Emb and Nym
683 samples. (D) Sequence logos showing nucleotide bias for each position of antisense
684 and sense siRNAs for each stage. The 1U bias typical of the piRNAs is only detectable
685 in embryos. E) Percentage of uniquely mapping siRNAs originating from the piRNA
686 clusters (green bars) versus other regions of the genome (white) for PVS, Egg, Emb
687 and Nym stages. F) Schematic diagram depicting the dynamics of piRNAs (red bar),
688 piRNA cluster transcripts (purple bar), TEs (dark green), HTTs (blue bar) and siRNAs
689 (light green bar) during oogenesis, embryogenesis and postembryonic development of
690 *Rhodnius prolixus*.

691 **Supplementary Fig 1. Abundance of transposon families in the top 10 piRNA**
692 **clusters.** X-axis displays the number of elements found for each transposon family (Y-
693 axis) in each of the top 10 piRNA clusters. Since the "Unknown" elements are the most
694 abundant ones, they are shown separately at the bottom.

695

696 **Supplementary Fig 2. Partially sequenced region in RPCL2.** Positive (blue) and
697 negative (red) read coverage along the piRNA cluster RPCL2 (top). At first inspection,
698 RPCL2 seems like a bidirectional cluster. However, when looking at the exact region

699 where the expression shifts (bottom), a partially sequenced region is notable and might
700 represent an assembly error.

701 **Supplementary Table 1. General mapping statistics**

702

703 **Supplementary Table 2. piRNA identification and pairing information**

704

705 **References**

706 1. SanMiguel, P. *et al.* Nested retrotransposons in the intergenic regions of the maize
707 genome. *Science* **274**, 765–768 (1996).
708 2. Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature*
709 **409**, 860–921 (2001).
710 3. Hubley, R. *et al.* The Dfam database of repetitive DNA families. *Nucleic Acids Res.*
711 **44**, D81–9 (2016).
712 4. Kofler, R., Hill, T., Nolte, V., Betancourt, A. J. & Schlötterer, C. The recent invasion
713 of natural Drosophila simulans populations by the P-element. *Proc. Natl. Acad. Sci.*
714 *U. S. A.* **112**, 6659–6663 (2015).
715 5. Peccoud, J., Loiseau, V., Cordaux, R. & Gilbert, C. Massive horizontal transfer of
716 transposable elements in insects. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 4721–4726
717 (2017).
718 6. Zhang, H.-H., Peccoud, J., Xu, M.-R.-X., Zhang, X.-G. & Gilbert, C. Horizontal
719 transfer and evolution of transposable elements in vertebrates. *Nat. Commun.* **11**,
720 1362 (2020).
721 7. Gilbert, C., Schaack, S., Pace, J. K., II, Brindley, P. J. & Feschotte, C. A role for

722 host–parasite interactions in the horizontal transfer of transposons across phyla.
723 *Nature* **464**, 1347–1350 (2010).
724 8. Filée, J., Rouault, J.-D., Harry, M. & Hua-Van, A. Mariner transposons are sailing in
725 the genome of the blood-sucking bug Rhodnius prolixus. *BMC Genomics* **16**, 1061
726 (2015).
727 9. Gao, D. *et al.* Horizontal Transfer of Non-LTR Retrotransposons from Arthropods to
728 Flowering Plants. *Mol. Biol. Evol.* **35**, 354–364 (2017).
729 10. Kidwell, M. G., Kimura, K. & Black, D. M. Evolution of hybrid dysgenesis potential
730 following P element contamination in Drosophila melanogaster. *Genetics* **119**, 815–
731 828 (1988).
732 11. Kidwell, M. G., Kidwell, J. F. & Sved, J. A. Hybrid Dysgenesis in DROSOPHILA
733 MELANOGASTER: A Syndrome of Aberrant Traits Including Mutation, Sterility and
734 Male Recombination. *Genetics* **86**, 813–833 (1977).
735 12. Daniels, S. B., Peterson, K. R., Strausbaugh, L. D., Kidwell, M. G. & Chovnick, A.
736 Evidence for horizontal transmission of the P transposable element between
737 Drosophila species. *Genetics* **124**, 339–355 (1990).
738 13. Brennecke, J. *et al.* An epigenetic role for maternally inherited piRNAs in
739 transposon silencing. *Science* **322**, 1387–1392 (2008).
740 14. Yin, H. & Lin, H. An epigenetic activation role of Piwi and a Piwi-associated piRNA
741 in Drosophila melanogaster. *Nature* **450**, 304–308 (2007).
742 15. Cox, D. N., Chao, A. & Lin, H. piwi encodes a nucleoplasmic factor whose activity
743 modulates the number and division rate of germline stem cells. *Development* **127**,
744 503–514 (2000).
745 16. Czech, B. *et al.* piRNA-Guided Genome Defense: From Biogenesis to Silencing.
746 *Annu. Rev. Genet.* **52**, 131–157 (2018).
747 17. Santos, D. *et al.* What Are the Functional Roles of Piwi Proteins and piRNAs in

748 Insects? *Insects* **14**, (2023).
749 18. Gunawardane, L. S. *et al.* A slicer-mediated mechanism for repeat-associated
750 siRNA 5' end formation in Drosophila. *Science* **315**, 1587–1590 (2007).
751 19. Brennecke, J. *et al.* Discrete small RNA-generating loci as master regulators of
752 transposon activity in Drosophila. *Cell* **128**, 1089–1103 (2007).
753 20. Saito, K. *et al.* Roles for the Yb body components Armitage and Yb in primary
754 piRNA biogenesis in Drosophila. *Genes Dev.* **24**, 2493–2498 (2010).
755 21. Sienski, G., Dönertas, D. & Brennecke, J. Transcriptional silencing of transposons
756 by Piwi and maelstrom and its impact on chromatin state and gene expression. *Cell*
757 **151**, 964–980 (2012).
758 22. Le Thomas, A. *et al.* Piwi induces piRNA-guided transcriptional silencing and
759 establishment of a repressive chromatin state. *Genes Dev.* **27**, 390–399 (2013).
760 23. Zanni, V. *et al.* Distribution, evolution, and diversity of retrotransposons at the
761 flamenco locus reflect the regulatory properties of piRNA clusters. *Proc. Natl. Acad.*
762 *Sci. U. S. A.* **110**, 19842–19847 (2013).
763 24. Pélisson, A. *et al.* Gypsy transposition correlates with the production of a retroviral
764 envelope-like protein under the tissue-specific control of the Drosophila flamenco
765 gene. *EMBO J.* **13**, 4401–4411 (1994).
766 25. Pane, A., Jiang, P., Zhao, D. Y., Singh, M. & Schüpbach, T. The Cutoff protein
767 regulates piRNA cluster expression and piRNA production in the Drosophila
768 germline. *EMBO J.* **30**, 4601–4615 (2011).
769 26. Mohn, F., Sienski, G., Handler, D. & Brennecke, J. The rhino-deadlock-cutoff
770 complex licenses noncanonical transcription of dual-strand piRNA clusters in
771 Drosophila. *Cell* **157**, 1364–1379 (2014).
772 27. Klattenhoff, C. *et al.* The Drosophila HP1 homolog Rhino is required for transposon
773 silencing and piRNA production by dual-strand clusters. *Cell* **138**, 1137–1149

774 (2009).
775 28. Andersen, P. R., Tirian, L., Vunjak, M. & Brennecke, J. A heterochromatin-
776 dependent transcription machinery drives piRNA expression. *Nature* **549**, 54–59
777 (2017).
778 29. Malone, C. D. *et al.* Specialized piRNA pathways act in germline and somatic
779 tissues of the Drosophila ovary. *Cell* **137**, 522–535 (2009).
780 30. Czech, B. *et al.* An endogenous small interfering RNA pathway in Drosophila.
781 *Nature* **453**, 798–802 (2008).
782 31. Rosenkranz, D., Rudloff, S., Bastuck, K., Ketting, R. F. & Zischler, H. Tupaia small
783 RNAs provide insights into function and evolution of RNAi-based transposon
784 defense in mammals. *RNA* **21**, 911–922 (2015).
785 32. Bernstein, E., Caudy, A. A., Hammond, S. M. & Hannon, G. J. Role for a bidentate
786 ribonuclease in the initiation step of RNA interference. *Nature* **409**, 363–366 (2001).
787 33. Lee, Y. S. *et al.* Distinct roles for Drosophila Dicer-1 and Dicer-2 in the
788 siRNA/miRNA silencing pathways. *Cell* **117**, 69–81 (2004).
789 34. Carvalho, T. L. *et al.* Horizontal Transposon Transfer and Their Ecological Drivers:
790 The Case of Flower-breeding Drosophila. *Genome Biol. Evol.* **15**, (2023).
791 35. Aravin, A. A., Hannon, G. J. & Brennecke, J. The Piwi-piRNA pathway provides an
792 adaptive defense in the transposon arms race. *Science* **318**, 761–764 (2007).
793 36. Chary, S. & Hayashi, R. The absence of core piRNA biogenesis factors does not
794 impact efficient transposon silencing in Drosophila. *PLoS Biol.* **21**, e3002099
795 (2023).
796 37. Dowling, D. *et al.* Phylogenetic Origin and Diversification of RNAi Pathway Genes
797 in Insects. *Genome Biol. Evol.* **8**, 3784–3793 (2016).
798 38. Aravin, A. A. *et al.* Double-stranded RNA-mediated silencing of genomic tandem
799 repeats and transposable elements in the D. melanogaster germline. *Curr. Biol.* **11**,

800 1017–1027 (2001).
801 39. Llonga, N., Ylla, G., Bau, J., Belles, X. & Piulachs, M.-D. Diversity of piRNA
802 expression patterns during the ontogeny of the German cockroach. *J. Exp. Zool. B*
803 *Mol. Dev. Evol.* **330**, 288–295 (2018).
804 40. Ninova, M., Griffiths-Jones, S. & Ronshaugen, M. Abundant expression of somatic
805 transposon-derived piRNAs throughout Tribolium castaneum embryogenesis.
806 *Genome Biol.* **18**, 184 (2017).
807 41. Wang, W., Ashby, R., Ying, H., Maleszka, R. & Forêt, S. Contrasting Sex-and
808 Caste-Dependent piRNA Profiles in the Transposon Depleted Haplodiploid
809 Honeybee Apis mellifera. *Genome Biol. Evol.* **9**, 1341–1356 (2017).
810 42. Buxton, P. A. THE BIOLOGY OF A BLOOD-SUCKING BUG, RHODNIUS
811 PROLIXUS. *Transactions of the Royal Entomological Society of London* **78**, 227–
812 256 (1930).
813 43. Montiel, E. E., Panzera, F., Palomeque, T., Lorite, P. & Pita, S. Satellitome Analysis
814 of Rhodnius prolixus, One of the Main Chagas Disease Vector Species. *Int. J. Mol.*
815 *Sci.* **22**, (2021).
816 44. Mesquita, R. D. *et al.* Genome of Rhodnius prolixus, an insect vector of Chagas
817 disease, reveals unique adaptations to hematophagy and parasite infection. *Proc.*
818 *Natl. Acad. Sci. U. S. A.* **112**, 14936–14941 (2015).
819 45. Brito, T. *et al.* Transcriptomic and functional analyses of the piRNA pathway in the
820 Chagas disease vector Rhodnius prolixus. *PLoS Negl. Trop. Dis.* **12**, e0006760
821 (2018).
822 46. Coelho, V. L. *et al.* Analysis of ovarian transcriptomes reveals thousands of novel
823 genes in the insect vector Rhodnius prolixus. *Sci. Rep.* **11**, 1918 (2021).
824 47. Brito, T. F. de *et al.* Transovarial transmission of a core virome in the Chagas
825 disease vector Rhodnius prolixus. *PLoS Pathog.* **17**, e1009780 (2021).

826 48. Giraldo-Calderón, G. I. *et al.* VectorBase: an updated bioinformatics resource for
827 invertebrate vectors and other organisms related with human diseases. *Nucleic*
828 *Acids Res.* **43**, D707–13 (2015).
829 49. Pascual, A. & Rivera-Pomar, R. Dynamics of maternal gene expression in
830 Rhodnius prolixus. *Sci. Rep.* **12**, 6538 (2022).
831 50. Kawamura, Y. *et al.* Drosophila endogenous small RNAs bind to Argonaute 2 in
832 somatic cells. *Nature* **453**, 793–797 (2008).
833 51. Adaptation to P Element Transposon Invasion in Drosophila melanogaster. *Cell*
834 **147**, 1551–1563 (2011).
835 52. Luo, Y., He, P., Kanrar, N., Fejes Toth, K. & Aravin, A. A. Maternally inherited
836 siRNAs initiate piRNA cluster formation. *Mol. Cell* **83**, 3835–3851.e7 (2023).
837 53. Lu, H.-L. *et al.* Expansion of genes encoding piRNA-associated argonaute proteins
838 in the pea aphid: diversification of expression profiles in different plastic morphs.
839 *PLoS One* **6**, e28051 (2011).
840 54. Penrice-Randal, R. *et al.* New Cell Lines Derived from Laboratory Colony Triatoma
841 infestans and Rhodnius prolixus, Vectors of Trypanosoma cruzi, Do Not Harbour
842 Triatoma Virus. *Insects* **13**, (2022).
843 55. Kidwell, M. G. Hybrid Dysgenesis in DROSOPHILA MELANOGASTER: Factors
844 Affecting Chromosomal Contamination in the P-M System. *Genetics* **104**, 317–341
845 (1983).
846 56. Anxolabéhère, D., Kidwell, M. G. & Periquet, G. Molecular characteristics of diverse
847 populations are consistent with the hypothesis of a recent invasion of Drosophila
848 melanogaster by mobile P elements. *Mol. Biol. Evol.* **5**, 252–269 (1988).
849 57. Asif-Laidin, A. *et al.* Modeling early germline immunization after horizontal transfer
850 of transposable elements reveals internal piRNA cluster heterogeneity. *BMC Biol.*
851 **21**, 117 (2023).

852 58. Berni, M. *et al.* Toll signals regulate dorsal-ventral patterning and anterior-posterior
853 placement of the embryo in the hemipteran Rhodnius prolixus. *Evodevo* **5**, 38
854 (2014).
855 59. Babraham Bioinformatics - Trim Galore!
856 https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/.
857 60. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-
858 efficient alignment of short DNA sequences to the human genome. *Genome Biol.*
859 **10**, R25 (2009).
860 61. Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput
861 Sequence Data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
862 62. Ramírez, F. *et al.* deepTools2: a next generation web server for deep-sequencing
863 data analysis. *Nucleic Acids Res.* **44**, W160–5 (2016).
864 63. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis
865 results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–
866 3048 (2016).
867 64. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose
868 program for assigning sequence reads to genomic features. *Bioinformatics* **30**,
869 923–930 (2014).
870 65. Flynn, J. M. *et al.* RepeatModeler2 for automated genomic discovery of
871 transposable element families. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 9451–9457
872 (2020).
873 66. RepeatMasker Home Page. http://www.repeatmasker.org.
874 67. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**,
875 2078–2079 (2009).
876 68. Wagih, O. ggseqlogo: a versatile R package for drawing sequence logos.
877 *Bioinformatics* **33**, 3645–3647 (2017).

878 69. Quinlan, A. R. BEDTools: The Swiss-Army Tool for Genome Feature Analysis.
879 *Curr. Protoc. Bioinformatics* **47**, 11.12.1–34 (2014).
880 70. Robinson, J. T. *et al.* Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26
881 (2011).

882

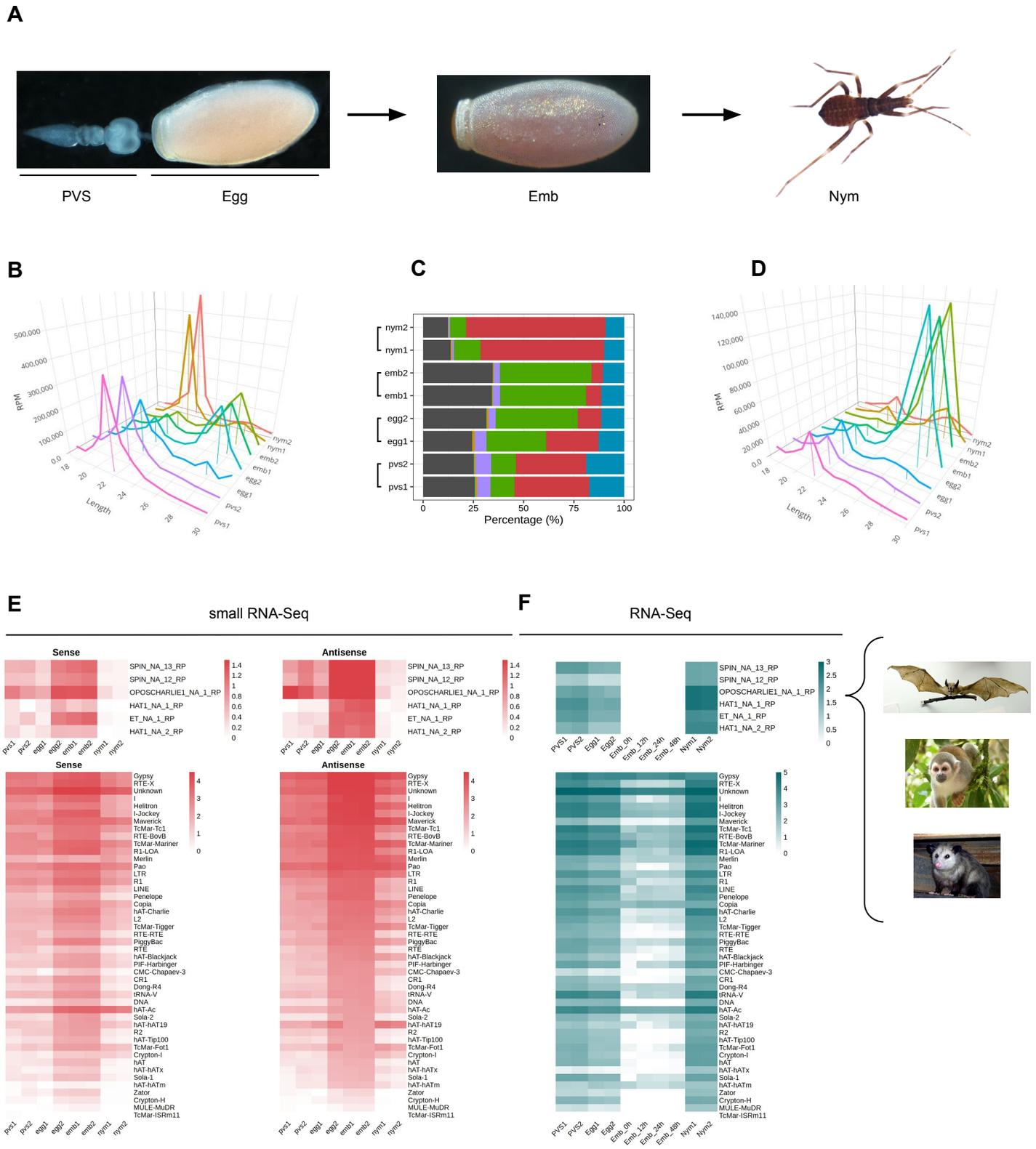

Figure 1. Rhodnius prolixus embryos express ~28 nt long piRNAs targeting Horizontally Transferred Transposons (HTT) and resident TEs. A) Developmental stages of Rhodnius prolixus analyzed in the current study. For oogenesis, we investigate the previtellogenic stage (PVS) comprising the tropharium and early follicles independently of the mature chorionated eggs dissected from the female abdomen (Egg). Embryos (Emb) were collected over a 48h period encompassing the cleavage stage until the formation of the head and thoracic segments after gastrulation. 1st instar nymphs (Nym) were also analyzed to investigate postembryonic development. B) Length distribution of all mapped reads for each replicate displayed in RPM. The ~22 nt and ~28 nt peaks were anchored to the X-axis (Length) for a better visualization. Y-axis displays RPM values. For each sample, we produced two replicates, which were analyzed separately. B) Small RNA profiling based on percentage of reads mapped to annotated and unannotated (unassigned) regions of Rhodnius genome. miRNAs are particularly abundant in oogenesis and postembryonic development (red), while embryonic stages are dominated by repeat-associated small RNAs (green). C) Length distribution of reads mapped to transposable and repetitive sequences of Rhodnius displayed in RPM. A major peak comprising ~28 nt piRNAs is clearly detectable in the two embryo replicates (Emb1 and Emb2) and in one of the mature egg replicates (Egg2). This class of small non-coding RNAs is less abundant in previtellogenic stages of oogenesis (PVS1 and PVS2 replicates), in one of the mature egg replicates (Egg1) and in 1st instar nymphs (Nym1 and Nym2 replicates). D) Heatmaps displaying the expression levels of sense and antisense piRNAs for each TE subfamily in PVS, Egg, Emb and Nym stages. Sense and antisense piRNAs for the horizontally transmitted transposons (HTT) ET, SPIN, OC1 and hAT and for the resident TEs are shown separately. Colorbar values display log10(RPM + 1). F) Heatmaps showing TE expression levels as per RNAseq in PVS, Egg, Emb and Nym stages. Expression levels for the HTTs ET, SPIN, OC1 and hAT and of the resident TEs are displayed in separate heatmaps. Colorbar values display RPKM. Rhodnius acquired the HTTs from opossums, squirrel monkeys and bats (right). For the SPIN and hAT families, Gilbert and coworkers reported two transposon sequences for SPIN and hAT, namely SPIN_NA_13_RP, SPIN_NA_12_RP, hAT_NA_1_RP and hAT_NA_2_RP. One transposon sequence was reported for ET and OC1, which were labeled ET_NA_1_RP and OPSCHARLIE1_NA_1_RP, respectively. For simplicity, we refer to these transposable elements as SPIN, hAT, OC1 and ET throughout the article. Images of bat, squirrel monkey and opossum were obtained from Wikimedia Commons.

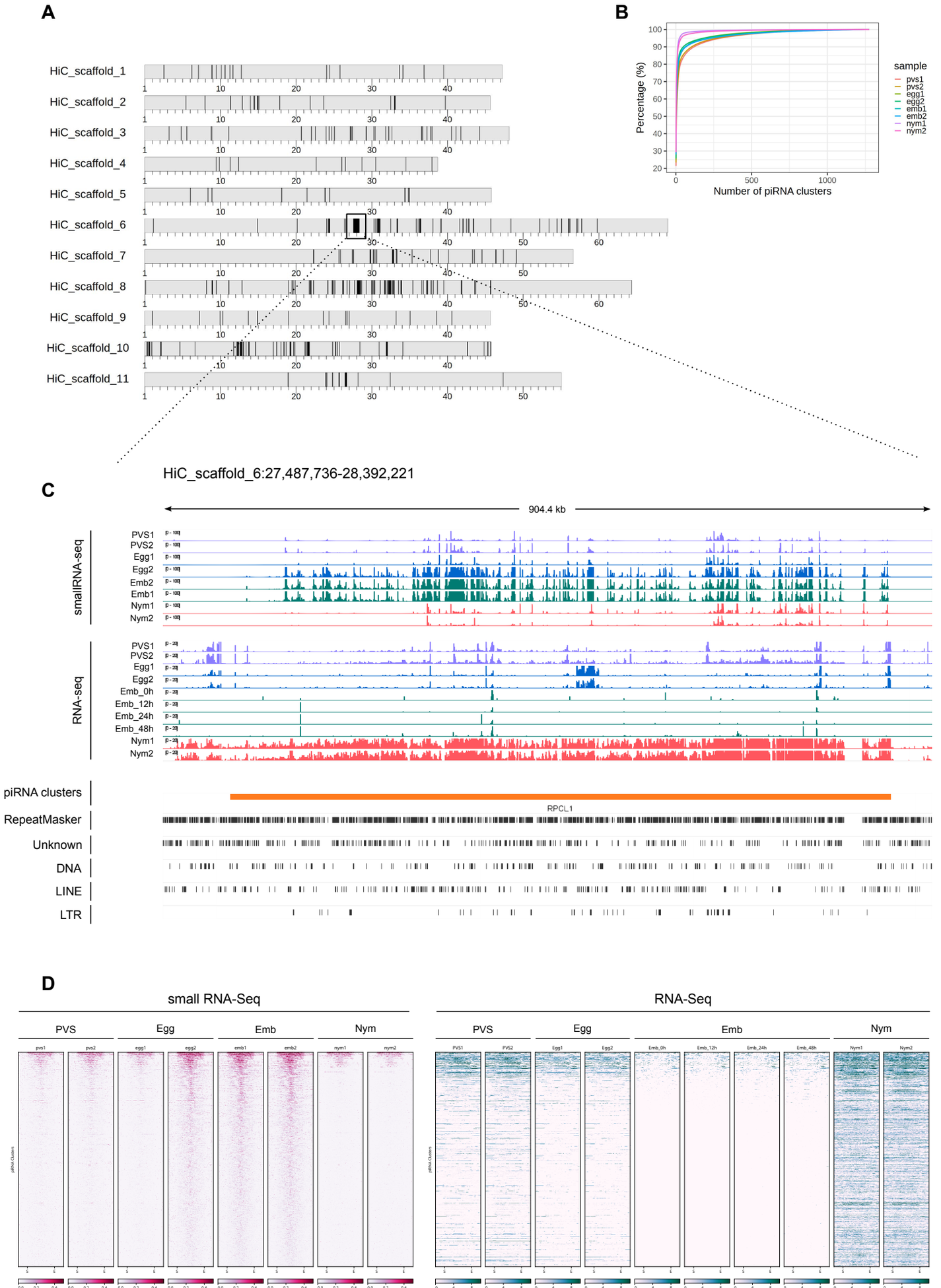

Figure 2. piRNA clusters are a major source of piRNAs during Rhodnius prolixus embryonic development. A) Ideograms of the larger scaffolds obtained from the HiC genome assembly that likely represent the 10 autosomes and the X chromosome (scaffold 10) of Rhodnius. Each ideogram is scaled to the size of each scaffold, with a scale in megabases. Black vertical lines show the distribution of the piRNA clusters. The major cluster RPCL1 is highlighted (black box). B) Cumulative percentage of piRNAs in clusters for each replicate. The top 10 RPCLs account for 74% of the total uniquely mapping piRNAs. C) IGV visualization of the RPCL1 cluster region on the HiC_scaffold_6. Uniquely mapped piRNAs for each stage are profiled for each sample (small RNA-seq profiles). The Y-axis displays RPM (0 to 100 interval). Expression levels of the RPCL1 (RNA-seq) in all the stages are also displayed. Of note, we included in this analysis the RNA-seq datasets for embryonic stages 0-12h, 12-24h, 24-36h and 36-48h produced by Pacual and coworkers. The Y-axis displays the RPKM range 0-20. Orange bar highlights the position of the RPCL1 cluster. The distribution of DNA and LTR transposons together with LINE elements and unknown sequences along the RPCL1 region are reported with black boxes. D) Heatmaps showing the temporal regulation of the cluster-derived mature piRNAs (Small RNA-seq) and piRNA precursor transcripts (RNA-seq) in PVS, Egg, Emb and Nym stages. Mature piRNA levels inversely correlate with the expression levels of cluster transcripts suggesting that piRNA biogenesis consumes cluster-derived transcripts during embryonic stages. Small RNA-seq values (left side of the panel) are displayed as log10(RPM + 1), while RNA-seq values (right side of the panel) are expressed as RPM. Each line of the heatmap represents a piRNA cluster with the major clusters at the top of the panels.

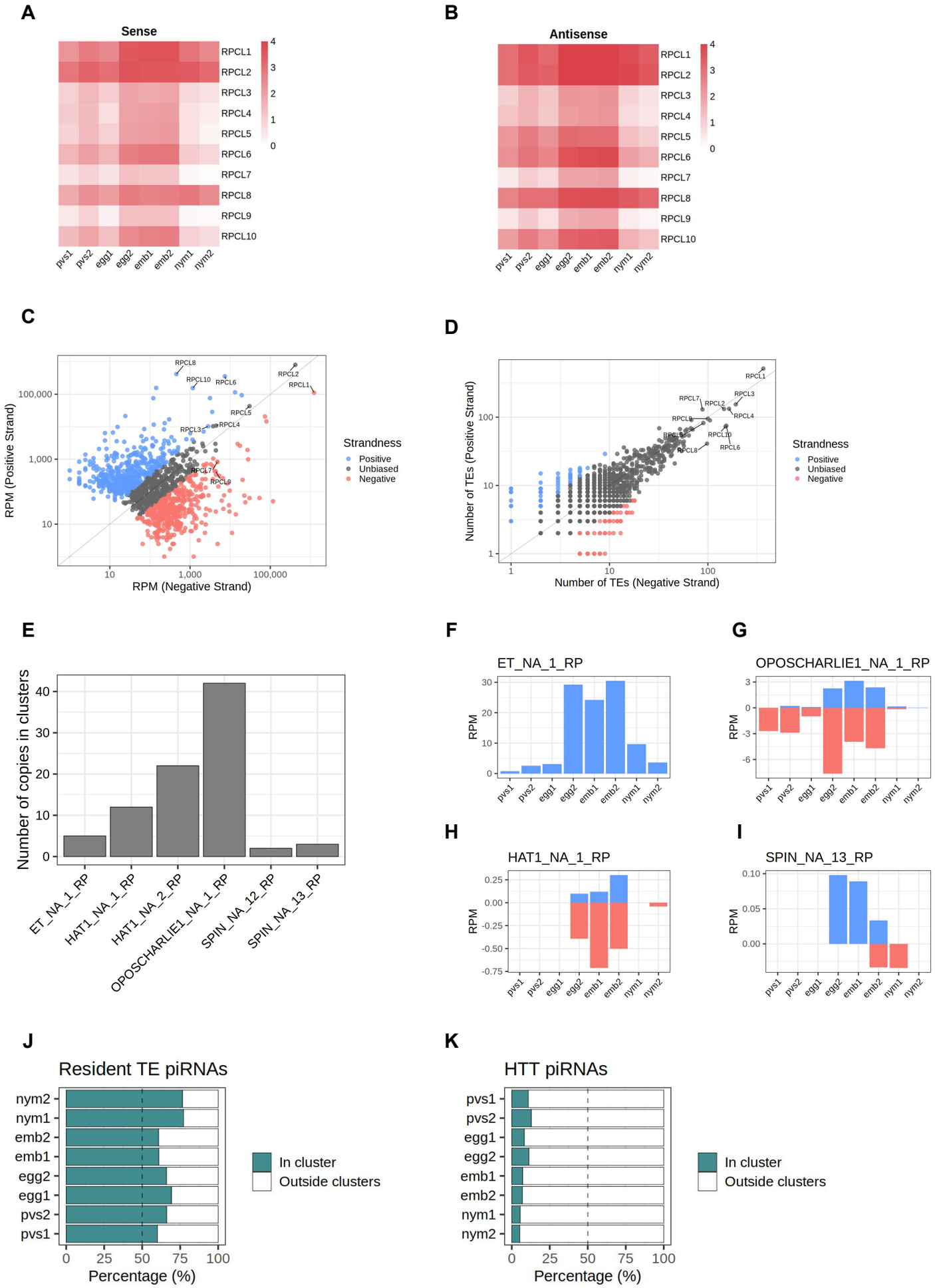

Figure 3. Rhodnius piRNA clusters are uni-strand and produce piRNAs against HTTs. A-B) Heatmaps showing sense (A) and antisense (B) piRNA levels for the top 10 piRNA clusters in PVS, Egg, Emb and Nym stages. Colorbar displays log10(RPM + 1). C) Scatterplot reporting the strandness of the piRNA clusters. piRNA clusters primarily expressed from the +strand (blue dots) and -strand (red dots) of the genome are displayed. piRNA clusters with apparently unbiased expression are indicated (dark grey dots). A cluster was considered unbiased if less than 75% of its piRNAs could be attributed to a specific strand of the genome. Y-axis values are expressed as RPM. The top 10 piRNA clusters are highlighted. D) Scatterplot showing the orientation of the TE fragment within each piRNA cluster. piRNA clusters harboring TEs predominantly on the +strand (blue dots) or -strand (red dots) of the genome are reported. These categories are mostly formed by small clusters with a low number of TEs. The vast majority of the piRNA clusters are characterized by unbiased distribution of TEs (dark grey dots). The top 10 piRNA clusters are highlighted. E) Number of HTT copies inserted into the piRNA clusters. F-I) piRNAs uniquely mapping to the cluster HTTs. The ET copies (F) seem to exclusively produce sense piRNAs, while OC1 (G), hAT (H) and SPIN (I) generate both sense and antisense piRNAs. Y-axis displays RPM values. X-axis shows the replicates for the PVS, Egg, Emb and Nym samples. J) Percentage of uniquely mapping piRNAs for the resident TEs originating from within or outside the piRNA clusters. K) Percentage of uniquely mapping piRNAs for the HTTs originating from within or outside the piRNA clusters.

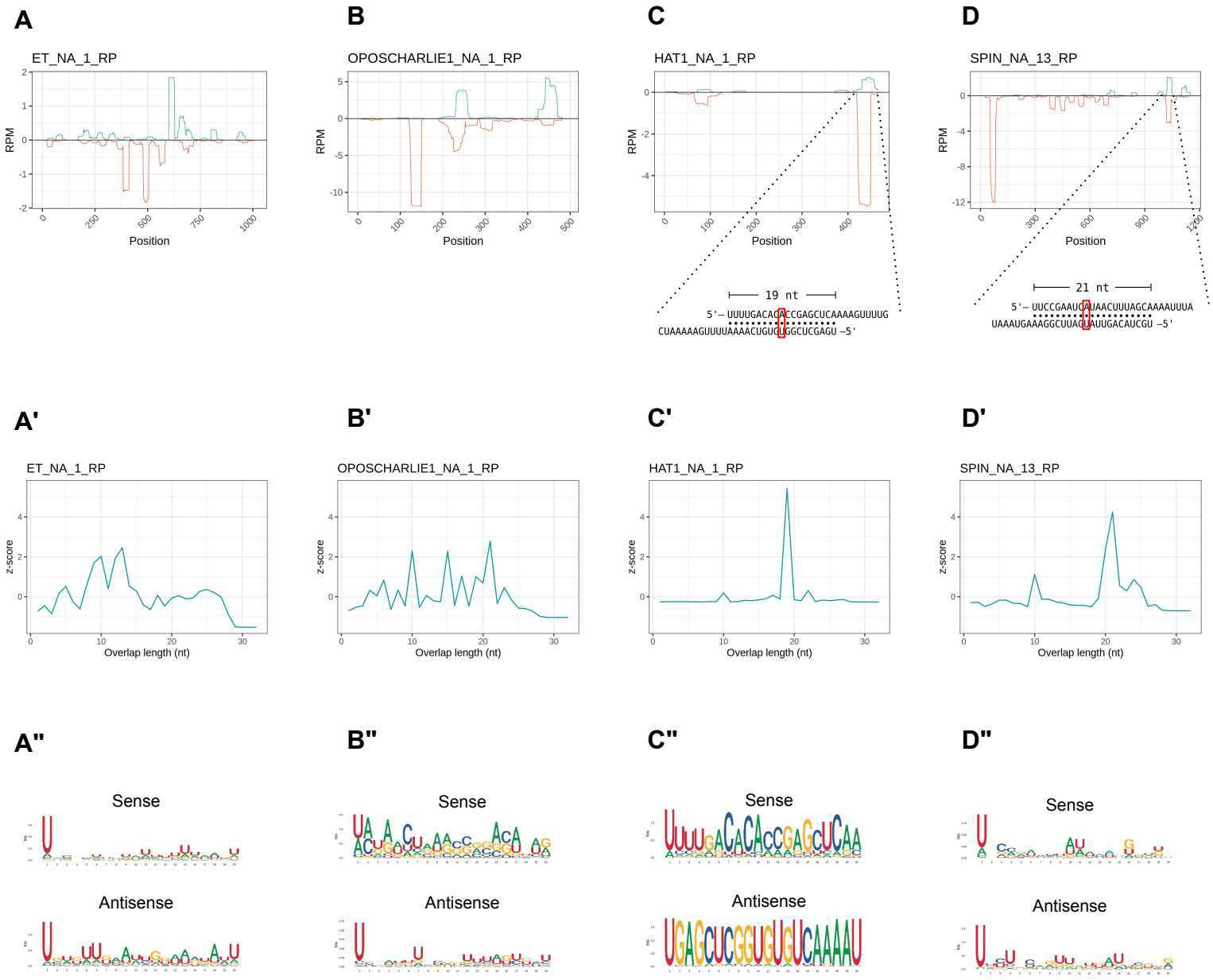

Figure 4. HTT piRNAs do not engage in a canonical ping-pong amplification mechanism. A-D) piRNA profiling along the sequences of each HTT with sense piRNAs displayed with positive RPM values (green) and antisense piRNAs with negative RPM values (red). For the hAT and SPIN elements, the most abundant sense/antisense piRNA pairs are shown. Of note, the sense piRNAs pairs do not display the typical ping-pong signature, but overlap by ~20 nt. Yet, the sense mate for each pair bears an Adenine residue at the 10th position (red box). A'-D') 5' to 5' overlap length frequency reported as z-score values for each HTT. A"-D") Sequence logos displaying nucleotide bias for sense and antisense piRNAs mapped to the HTTs.

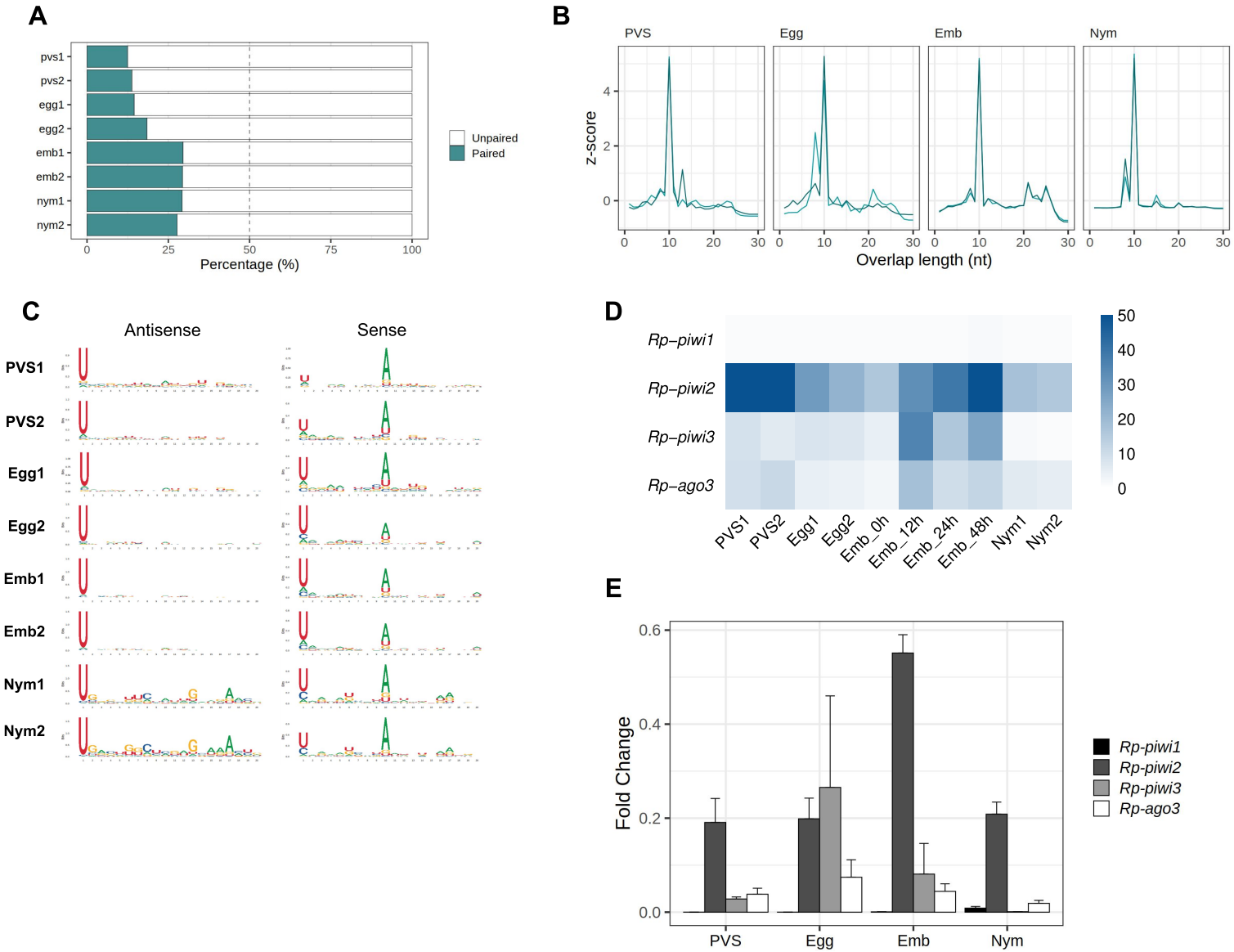

Figure 5. The ping-pong amplification loop is active in Rhodnius prolixus. A) Percentage of paired and unpaired piRNAs for PVS, Egg, Emb and Nym replicates. B) 5' to 5' overlap length frequency displayed in z-score for piRNAs mapped to resident transposons for each stage. C) Sequence logo showing nucleotide bias for each position of antisense and sense piRNAs mapped to transposable elements in all the replicates. D) qRT-PCR analysis of the temporal expression levels of the Rp-PIWI genes during oogenesis and early stages of Rhodnius prolixus development. PVS=previtellogenic stage, Egg=mature chorionated eggs, Emb=embryos and Nym=1st instar nymphs. Y-axis shows the mean Fold Change. Error bars display standard deviation from the mean.

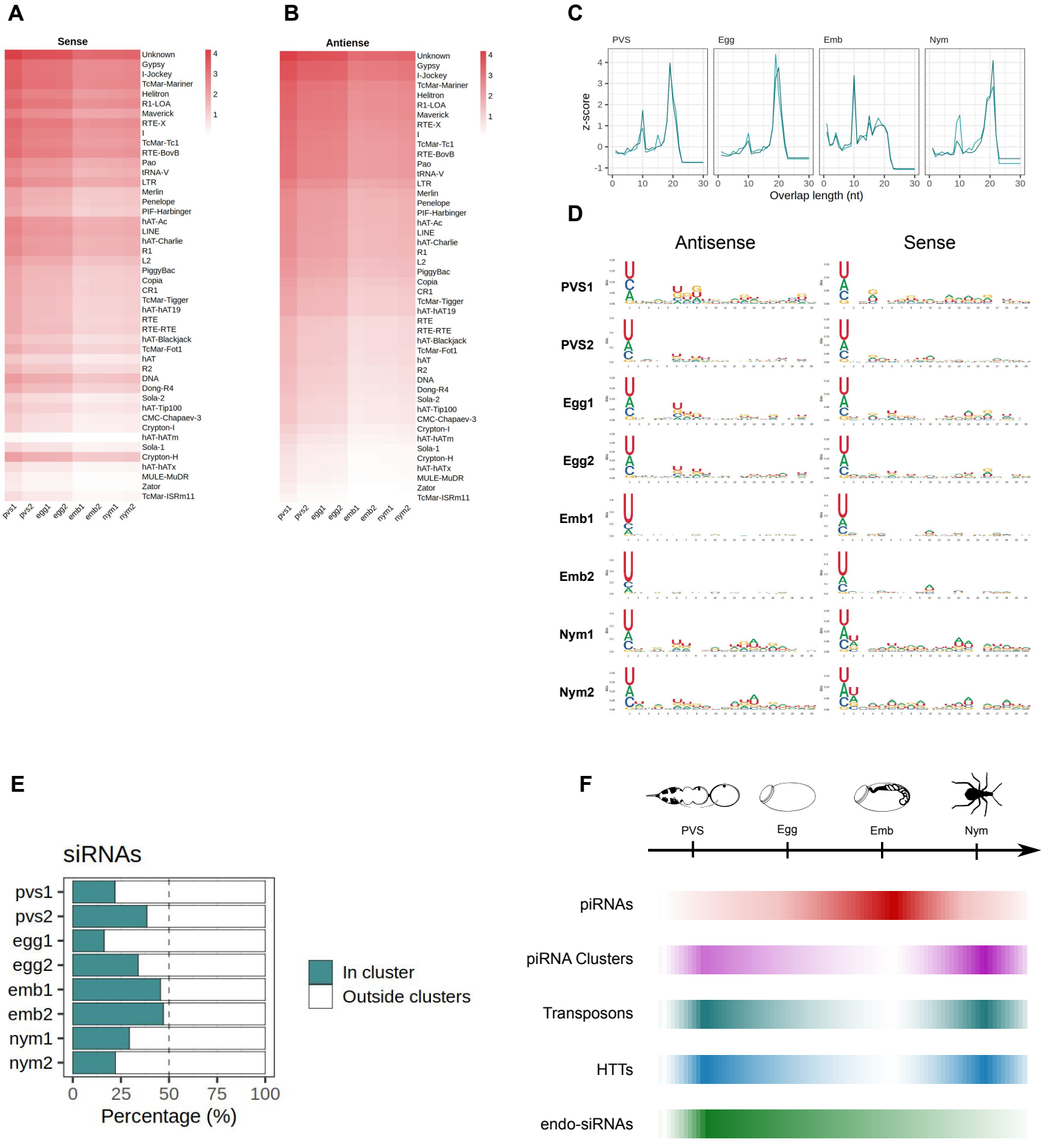

Figura 6. piRNA clusters generate ~22nt siRNAs in Rhodnius prolixus. A-B) Heatmaps showing sense (A) and antisense (B) ~22 nt long siRNAs mapped to the Rhodnius TE subfamilies. Colorbar displays log10(RPM + 1). siRNAs appear to be expressed at higher levels in oogenesis compared to embryonic and postembryonic stages. C) Z-scores of the 5' to 5' overlap length frequency for PVS, Egg, Emb and Nym samples. (D) Sequence logos showing nucleotide bias for each position of antisense and sense siRNAs for each stage. The 1U bias typical of the piRNAs is only detectable in embryos. E) Percentage of uniquely mapping siRNAs originating from the piRNA clusters (green bars) versus other regions of the genome (white) for PVS, Egg, Emb and Nym stages. F) Schematic diagram depicting the dynamics of piRNAs (red bar), piRNA cluster transcripts (purple bar), TEs (dark green), HTTs (blue bar) and siRNAs (light green bar) during oogenesis, embryogenesis and postembryonic development of Rhodnius prolixus.