

1

2 Deep experimental profiling of microRNA diversity, deployment, and evolution

3 across the *Drosophila* genus

5 Jaaved Mohammed^{1-3, 5, 6}, Alex S. Flynt^{3,4,6,7}, Alexandra M. Panzarino³, Md Mosharrof Hussein
6 Mondal⁴, Adam Siepel⁵ and Eric C. Lai^{2,3,7}

⁸ ¹Cornell University, Department of Biological Statistics and Computational Biology, Ithaca NY
⁹ 14853, USA

¹⁰ ¹¹ ¹² ²Tri-Institutional Training Program in Computational Biology and Medicine, New York NY 10021,
USA

³Sloan-Kettering Institute, Department of Developmental Biology, 1275 York Ave, Box 252, New York NY 10065, USA

¹⁶ University of Southern Mississippi, Department of Biological Sciences, Hattiesburg MS 39406,
¹⁷ USA

¹⁶ University of Southern Mississippi, Department of Biological Sciences, Hattiesburg MS 39406,
¹⁷ USA

¹⁹
²⁰ ⁵Cold Spring Harbor Laboratory, Simons Center for Quantitative Biology, 1 Bungtown Rd, Cold
²¹ Spring Harbor NY 11724 USA

⁶These authors contributed equally.

24 Author for correspondence:
25 Eric C. Lai
26 Phone: (212) 639-5578
27 Fax: (212) 717-3604
28 Email: laie@mskcc.org

30 Keywords: microRNA evolution mirtron microRNA birth and death *Drosophila*

31 Running head: miRNA flux across the Drosophilid phylogeny

32 **Abstract**

33 Comparative genomic analyses of microRNAs (miRNAs) have yielded myriad insights
34 into their biogenesis and regulatory activity. While miRNAs have been deeply annotated in a
35 small cohort of model organisms, evolutionary assessments of miRNA flux are clouded by the
36 functional uncertainty of orthologs in related species, and insufficient data regarding the extent
37 of species-specific miRNAs. We address this by generating a comparative small RNA (sRNA)
38 catalog of unprecedented breadth and depth across the *Drosophila* genus, extending our extant
39 deep analyses of *D. melanogaster* with sRNA data from multiple tissues of 11 other fly species.
40 Aggregate analysis of several billion sRNA reads permits curation of accurate and holistic
41 compendia of miRNAs across this genus, providing abundant opportunities to identify species-
42 and clade-specific variation in miRNA identity, abundance, and processing. Amongst well-
43 conserved miRNAs, we observe unexpected cases of clade-specific variation in 5' end
44 precision, occasional antisense loci, and some putatively non-canonical loci. We also employ
45 strict criteria to identify a massive set (649) of novel, evolutionarily-restricted miRNAs. Amongst
46 the bulk collection of species-restricted miRNAs, two notable subpopulations of rapidly-evolving
47 miRNAs are splicing-derived mirtrons and testis-restricted, clustered (TRC) canonical miRNAs.
48 We quantify rates of miRNA birth and death using our annotation and a phylogenetic model for
49 estimating rates of miRNA turnover in the presence of annotation uncertainty. We show striking
50 differences in birth and death rates across miRNA classes defined by biogenesis pathway,
51 genomic clustering, and tissue restriction, and even identify variation heterogeneity amongst
52 *Drosophila* clades. In particular, distinct molecular rationales underlie the distinct evolutionary
53 behavior of different miRNA classes. We broaden observations made from *D. melanogaster* as
54 Drosophilid-wide principles for opposing evolutionary viewpoints for miRNA maintenance.
55 Mirtrons are associated with a high rate of 3' untemplated addition, a mechanism that impedes
56 their biogenesis, whereas TRC miRNAs appear to evolve under positive selection. Altogether,
57 these data reveal miRNA diversity amongst *Drosophila* species and permit future discoveries in
58 understanding their emergence and evolution.

59 **Introduction**

60 MicroRNAs (miRNAs) are ~22 nucleotide (nt) RNAs that play important regulatory roles
61 in diverse eukaryotic species by promoting transcript degradation or by translational repression
62 [1,2]. In the canonical metazoan pathway, primary miRNA (pri-miRNA) transcripts bearing
63 hairpins are first cleaved in the nucleus by the RNase III cleavage enzyme Drosha. Upon export
64 to the cytoplasm, these precursor miRNA (pre-miRNA) hairpins are further processed into
65 miRNA duplexes by Dicer, another RNase III enzyme. One duplex strand, termed the "mature"
66 miRNA strand, is preferentially retained in an Argonaute (AGO) complex and guides it to
67 complementary mRNA targets. Its partner miRNA* strand, or "star" species, is preferentially
68 degraded, although functional capacity of star strands has been documented. Beyond the
69 canonical pathway, diverse non-canonical biogenesis pathways involving RNases that function
70 in other cellular processes have been uncovered [3]. Chief amongst these is the "mirtron"
71 pathway, in which the Drosha cleavage is substituted by the spliceosome to define either or
72 both pre-miRNA hairpin termini [4-6].

73 In the evolutionary context, comparative and population genomics have been crucial to
74 our understanding of the functional roles of miRNAs. Such efforts helped define features such
75 as ultra-conservation of the miRNA "seed" sequence (positions ~2-8 of the miRNA strand that
76 mediate target recognition), and the overall higher constraint upon the miRNA and star strands
77 compared to other partitions of the pre-miRNA hairpin [7-9]. With the availability of *Drosophila*
78 population data, primarily from *D. melanogaster*, deeper insights into miRNA evolution on a
79 more recent timescale emerged, such as the accelerated, adaptive evolution of miRNAs within
80 clusters of testes-restricted expression [10-12]. Additionally, advances in high-throughput
81 sequencing and the availability of miRNA prediction software that leverages this data [13,14],
82 there has been a surge in the annotation of miRNAs across taxa, many of which comprise
83 recently-evolved loci. Overall, miRNAs have now been associated with phenotypic diversity, and
84 their expansion has been correlated with organismal complexity, body-plan innovation, and life
85 cycle [1,2,15].

86 miRNA catalogs are continually augmented across broad phylogenetic branches, but
87 these have mostly been assessed at the level of presence/absence of miRNA loci [15-17]. Much
88 remains to be explored about miRNA evolutionary features across sets of related species, such
89 as patterns and rates of gene emergence, decay, and expansion, and consistency in processing
90 across orthologs. To date, only a few deep comparisons have been conducted. These include
91 analysis of four *Caenorhabditis* species [18,19] (sRNAs from additional nematode species have
92 been cloned [20], but not analyzed for miRNA emergence), up to six Mammalian species

93 [21,22], and three *Drosophila* species [23,24]. A net gain estimate of 12 miRNA genes/million
94 years (Myr) was first estimated in *Drosophila* [23]. However, this estimate was later revised to
95 0.82-1.6 genes/Myr using a refined, conservative collection of miRNA genes [24], which proved
96 relatively concordant with a subsequent estimate of 0.83 genes/Myr in mammals [21]. Since
97 these studies, the annotation of *Drosophila* and mammalian miRNAs has expanded by many
98 fold [25]. Nevertheless, despite now thousands of miRNAs collected within the miRBase
99 repository for constituent species of these clades, numbers of pan-*Mammalian* (94) [21], or pan-
100 *Drosophilid* (123) miRNAs [8] have not changed much over the past decade. Thus, there are
101 perhaps an order of magnitude more recently-evolved miRNAs than well-conserved loci in some
102 metazoans.

103 A relevant consideration is that most previous studies of miRNA evolutionary flux
104 considered them as a unitary class. However, our recent studies provide evidence that
105 subclasses of miRNAs exhibit distinct evolutionary parameters [8,11,17]. For example, we noted
106 that *Drosophila* mirtrons evolve more quickly than canonical miRNAs [24]. This observation was
107 substantiated by our recent strict annotation of ~500 mirtrons in both mouse and human, nearly
108 all of which are specific to rodents or primates, respectively [26]. One can recognize, then, that
109 expression levels of the minority of conserved miRNA loci dwarf those of the collective majority
110 of miRNA loci, which include both canonical and non-canonical loci. We speculate from this that
111 there should be diverse mechanisms that can drive characteristic evolutionary behaviors of
112 various miRNA classes. However, a foundation to study these may require a deep empirical
113 analysis of species-specific miRNAs across a phylogeny.

114 In this study, we set out to characterize class-specific properties of miRNAs within 12
115 species of the *Drosophila* genus, which diverged from the common Dipteron ancestor ~60 Myr.
116 Advantages of the *Drosophila* system include its wealth of phenotypic diversity, straightforward
117 culture of a wide species collection, access to high-quality whole-genome assemblies, and most
118 importantly, the increased power of fine comparative assessment of evolutionary features within
119 sub-groups of closely-related species [27]. Building on a collection of approximately 1.9 billion
120 sRNA sequences from *D. melanogaster* (summarized in [28,29]), we sequenced an additional
121 ~1.5 billion sRNAs from embryos, heads, male bodies and female bodies of the other 11
122 species. This comparative dataset permits evolutionary miRNA analysis at an unprecedented
123 scale. In particular, we elaborate myriad features of miRNA annotation and evolution, and show
124 how these differ with respect to miRNA biogenesis types, tissues with an animal, and between
125 different branches of the fruitfly phylogeny.

126 **Results**

127

128 **Compendia of sRNA data across 12 *Drosophila* species**

129 We previously annotated *D. melanogaster* miRNAs from ~1.9 billion small RNA
130 reads, spanning >100 different developmental stages, tissue types, cell lines, and
131 genetic and environmental manipulations [28,29]. While this scale is not currently
132 feasible to achieve across other Drosophilids, we sought parameters of data collection
133 that would permit deep annotations in other species, and thus serve as an empirical
134 foundation for comparative analyses of miRNA evolution.

135 Our experience with *D. melanogaster* suggested that mixed embryos, adult
136 heads, male bodies and female bodies are efficacious for broad capture of miRNA
137 diversity. To test this, we performed recovery analyses of *D. melanogaster* miRNAs by
138 subsampling data from these four tissue types (see **Methods**). At an aggregate depth of
139 100M reads, we recovered 94-98% of conserved (128) miRNAs with at least 30 mature
140 miRNA reads and three miRNA* reads from 100 simulation experiments (**Figure 1A**). Of
141 the 135 miRNAs that emerged recently in the melanogaster-group, we recovered 21-
142 27% of miRNAs using these miRNA/star thresholds. Increasing this depth to 200M reads
143 resulted in a maximal recovery gain to 30-34% of newly-evolved *D. melanogaster*
144 miRNAs (**Figure 1B and/or Supplementary Figure S1**). In other words, doubling the
145 sequencing depth allowed recovery of at most only 7% more newly-evolved miRNAs.
146 Thus, we considered 100M reads from the four tissue types as a strong empirical
147 foundation across these species that would permit broad insights into miRNA evolution.

148 We deeply sequenced 52 small RNA libraries (~1.5 billion total reads) from these
149 four tissue types across 11 species, exceeding 100M reads for nearly all species
150 (**Figure 1C**). As expected, read lengths of most libraries peaked at 21-22 nts,
151 representing miRNAs, and most body libraries showed an additional 24-28nt peak
152 representing the piRNA population (**Supplementary Table S1, Supplementary Figure**
153 **S2**). These datasets broadly extend the limited collection of publicly available sRNA data
154 from other fly species, primarily *D. simulans* (~200M reads) and *D. virilis* (~700M reads),
155 which we aggregated with our libraries (**Figure 1D, Supplementary Table S2**).

156

157 **A genus-wide catalog of *Drosophila* miRNA annotations**

158 Several strategies to annotate miRNAs from small RNA data have been
159 developed over the years. These collectively have distinct merits, but no single strategy

160 suffices to discover the full range of confident miRNAs, especially ones with atypical
161 structures and/or non-canonical biogenesis. Therefore, we deployed a multi-pronged
162 framework (see **Methods**) including (1) miRDeep2 (to cast a wide net of candidate
163 hairpins with evidence of cloned small RNA duplexes), (2) an independent set of
164 predicted hairpin structures (especially useful for identifying miRNAs from extended
165 hairpin precursors that are disallowed by miRDeep2), (3) intron annotations (to identify
166 mirtrons and tailed mirtrons, which are systematically overlooked by canonical miRNA
167 finders such as miRDeep2), and (4) whole-genome alignments to identify putative
168 miRNA orthologs across multiple *Drosophila* species (to "rescue" miRNA loci from the
169 candidate pool that have confidently cloned orthologs). Since all initial computational
170 scans include substantial false positives, we subsequently utilized stringent criteria (e.g.,
171 abundance, distribution, and patterns of sRNA read alignments indicative of Drosha and
172 Dicer cleavage) and systematic visual inspection of all loci before assigning final
173 annotations to various categories (**Supplementary Figure S3**).

174 We first queried miRBase (v21) loci for Drosophilid orthologs whose cloned small
175 RNAs had not previously been explicitly identified. This exercise served as an initial
176 check on the overall quality of the datasets, since a basic inference from genomic
177 conservation of a miRNA locus is that it is likely processed into mature small RNAs. Our
178 data support the first cloning evidence for 592 unannotated orthologs of conserved
179 Drosophilid miRNA loci (**Figure 2A**). 512 of these loci were cloned at thresholds of at
180 least 30 miR reads and 3 miR* reads, which would have supported high-confidence de
181 novo annotation. The remainder were cloned at lower thresholds, and would initially
182 have been segregated as "candidates", but could be recovered based on their orthology
183 to loci well-cloned in other species (80 candidate-rescued, and 42 candidate miRNAs).
184 This supported a rationale to "rescue" certain candidate miRNA loci that fall below high
185 stringency thresholds, but that could be reasonably considered as genuine based on
186 high sequence orthology to a confident miRNA annotation in one or more other species.
187 On the other hand, our deep datasets supported our decision to demote 47 annotations
188 from Drosophilid miRBase loci. These were primarily *D. pseudoobscura* loci for which
189 our present data indicate previous annotations were based on non-miRNA reads
190 (**Supplementary Figure S4, Supplementary Table S3**). Our reassessment of these
191 miRBase loci emphasizes the rigor of current miRNA scoring criteria.

192 For the remainder of our analysis, we grouped the newly-cloned orthologs of
193 miRBase loci along with extant miRBase miRNAs, so as to emphasize the truly novel

194 collection of miRNAs that is unique to our *de novo* annotation effort. In particular, our
195 data support the annotation of 649 novel, confident miRNAs across 12 Drosophilid
196 species (**Figure 2A**). Perhaps not unsurprisingly, many of the highest expressed novel
197 miRNAs are from the *virilis* subclade, which is most distant from *D. melanogaster*. The
198 example of *dvi_264* was cloned at >100,000 reads (**Figure 3A**). At such depth recovery,
199 we observe not only precision of 5p and 3p reads, but we also observe cloned loop
200 reads, which provide evidence of a dominant diced product and indicate 3'-trimming of
201 the predominantly cloned 20 nt species (**Figure 3A**). However, perhaps unexpected was
202 that even in species relatively close to *D. melanogaster*, we still recovered scores of
203 novel confident miRNAs, even though we sampled their small RNAs at <1/10 the read
204 depth and tissue/cell diversity assayed in *D. melanogaster*. For example, *der_50* is but
205 one example of a *melanogaster* subclade miRNA expressed only in *D. erecta* and
206 recovered at a depth of >1000 reads, but absent from *D. melanogaster* (**Figure 3A**).
207 Such observations provide indications of evolutionary flux that we explore later in this
208 study.

209 **Figure 3B** illustrates *dsi_14614* as a novel non-canonical miRNA of the splicing-
210 derived "mirtron" class. *D. melanogaster* mirtron-3p reads are associated with a high
211 rate 3' untemplated uridylation [30], and nearly all *dsi_14614-3p* reads are uridylated
212 (**Figure 3B**). Our comparative sRNA data afforded us the unique capability of recovering
213 candidate miRNAs orthologous to confident ones, which we termed as "candidate-
214 rescued" miRNAs; i.e., that we consider as genuine miRNA products. As an example,
215 *dse_989* was classified as a candidate mirtron originally due to the lack of miR* read
216 recovery (**Figure 3B**), however, its synteny to *dsi_14614* allowed us the ability to elevate
217 the confidence of this mirtron. From an initial collection of 765 "candidate" loci bearing
218 small RNA evidence reminiscent of miRNAs, we re-classified 82 as candidate-rescued
219 miRNAs or mirtrons that were clearly orthologous to confident loci (**Figure 2A**,
220 **Supplementary Figure S5**).

221 The remaining 683 "candidate" loci have small RNA evidence that are
222 reminiscent of miRNAs, but do not meet minimum criteria and are therefore set aside
223 (**Supplementary Figure S5**). We do not presently consider these as genuine miRNA
224 species, even though the veracity of some of these may emerge from deeper or
225 specialized sequencing, and/or they may provide fodder for eventual evolutionary
226 emergence. Nevertheless, our massive set of small RNA data allows us to expand the
227 collection of 1965 known and unannotated miRBase loci of confident and rescued status

228 with 732 novel miRNAs and mirtrons to arrive at a final collection of 2697 total miRNAs
229 and mirtrons present within the *Drosophila* genus. A master table of these loci and their
230 properties is provided as **Supplementary Table S4**. These annotations can be explored
231 in the supplemental website, which provides extensive information regarding read
232 pileups, secondary structures, aligned sequences in other Drosophilid species and so
233 forth (http://compgen.cshl.edu/mirna/12flies/12flies_alignments.html). This miRNA
234 compendium represents one of the largest in any single genus, and we sought to exploit
235 these empirical data to derive insights into miRNA biogenesis and evolution.
236

237 **miRNA classes, alignments, and expression**

238 We segregated our aggregate annotations into loci of distinct biogenesis and
239 genomic clustering classes. We first separated canonical miRNAs from mirtrons, a group
240 of small hairpin introns that mimic pre-miRNAs [4,5,31], and further partitioned canonical
241 miRNAs to segregate Testes-restricted, Recently-evolved, Clustered miRNAs (TRC
242 miRNAs) [11]. We previously utilized *D. melanogaster*-centric miRNA annotations and *D.*
243 *melanogaster* population data to provide evidence that miRNAs of these three classes
244 evolve by distinct selective pressures, as discerned from their patterns of precursor and
245 “seed” sequence conservation, copy number, and signatures of positive selection [11].
246 Our highly expanded pan-Drosophilid miRNA annotations provided opportunities to test
247 some of these notions later in this study.

248 More than half (374 loci; 51.1%) of the novel miRNAs identified in our study and
249 72.6% (1427) of known and unannotated miRBase loci were solo canonical miRNAs
250 (**Figure 2B**). On the other hand, TRC miRNAs comprised 31.7% (126) of our novel
251 collection, and mirtrons accounted for 17.2% (232). These findings affirm that specific
252 pools of hairpins, namely non-canonical and testis-restricted loci, contribute
253 disproportionately to the aggregate catalog of miRNA substrates. When all miRNAs were
254 evaluated together, we observed 4.5 times more non-TRC canonical miRNAs (66.8%)
255 than TRC miRNAs (15.9%), an enrichment that reproduced across species
256 (**Supplementary Figure S6**).

257 Curation of accurate miRNA orthologs was paramount to our comparative
258 analyses. We grouped miRNA orthologs into alignments by building upon previous,
259 manual alignments for *D. melanogaster* miRNAs [8], and assigning new miRNAs into
260 groups via genome-wide homology identification and multi-species whole genome
261 alignments (see **Methods**). Altogether, we grouped 2697 miRNAs into 1004 miRNA

262 alignments. From these alignments, we observed that species-specific miRNAs
263 (comprising the majority of loci newly annotated in this study) were the dominant class
264 (65.4%), and miRNAs with one cloned ortholog the next largest class (15.2%). On the
265 other end, 115 alignments (11.4% of loci) contained 10 or more members, and thus were
266 present at the base of the Drosophilid phylogeny (**Figure 2C**).

267 We next assessed miRNA expression across orthologous miRNAs to understand
268 the range and variation of miRNA expression among alignments. To compare miRNA
269 expression and to assign an expression value per alignment, we computed the
270 $\log_2(\text{reads per million mapped miRNA reads, RPMM})$ score for all loci and recorded the
271 maximum expression level per miRNA in any individual library. We evaluated this metric,
272 instead of the "average" expression of miRNAs, to account for the tissue-specific
273 deployment of many miRNAs. In comparisons of miRNA age, we observed that 92.3% of
274 conserved miRNA alignments had a maximum expression >32 RPMM [i.e. $\log_2(5)$],
275 while 24.1% of newly-evolved miRNA achieved such an expression (**Figure 2D**). At a
276 relaxed cutoff threshold, however, we observed that 80.2% of novel miRNA alignments
277 achieved expression of >1 RPMM in at least one library. On the other hand, in
278 comparisons of miRNA class, we observed that TRC miRNA alignments outperformed
279 other classes at conservative maximum expression cutoff (>32 RPMM) (e.g., 45% for
280 TRC versus 19.7% for mirtron and 34.1% for other canonical miRNA) (**Figure 2E**).
281

282 Novel, deeply-conserved miRNAs

283 Catalogs of well-conserved miRNAs are considered largely complete, as it is
284 generally believed that the set of "clonable" hairpins with miRNA-like evolutionary
285 signatures were exhausted years ago. However, some conserved miRNAs continue to
286 be found, many of which derive from unusual genomic locations or non-canonical
287 pathways, perhaps explaining why they were overlooked earlier. For example, deeply-
288 conserved, non-canonical, *dme-mir-10404* was recently reported to be processed from
289 the internal spacer regions of highly repetitive rRNA loci [32]. Indeed, we find this miRNA
290 is well-cloned from across the Drosophilid phylogeny (**Supplementary Figure S7**).

291 Amongst our novel miRNA annotations, a handful of loci appeared to be cloned
292 from a broad range of Drosophilid species (**Supplementary Figure S8**). The behavior of
293 *pasha* 5' UTR hairpins was instructive. A feedback loop in which Drosha cleaves 5' UTR
294 foldbacks in *pasha/DGCR8* is conserved from mammals to fruitflies, and we previously
295 validated this *in vivo* regulatory interaction using engineered transgenes [33,34]. It was

296 reported that mammalian *DGCR8* 5' UTR hairpin products of Drosha cleavage are
297 retained in the nucleus, thereby preventing further maturation by Dicer [33].
298 Nevertheless, sufficient mature reads of *DGCR8* hairpins are found in mammalian deep
299 sequencing data to justify annotation of *mir-3618* and *mir-1306*. We identified small RNA
300 duplexes from the corresponding species ranges for both the deeply conserved
301 (*dme_422*) and melanogaster group-restricted (*dme_474*) *pasha* 5' UTR hairpins
302 (**Supplementary Figure S8**). Although their resultant small RNAs are not particularly
303 abundant, the coupling with our prior evidence of *in vivo* cleavage of these hairpins by
304 Drosha provides de facto indication of the depth of our deep small RNA profiling.

305 Amongst novel conserved miRNAs, a notable pair is *dme_373* and *dme_164*,
306 which are clustered in the first intron of *clockwork orange* (*cwo*). Both loci are highly
307 conserved, exhibit greater loop divergence relative to the hairpin arms that is diagnostic
308 of evolutionarily constrained miRNAs, and are processed into small RNA duplexes
309 across the Drosophilids (**Figure 3C**). However, both loci exhibit atypical features:
310 *dme_164* harbors a conserved, A-rich lower stem that likely precludes Drosha access,
311 while *dme_373* contains an unusually large (~50 nt) loop (mean loop size of 209 *D.*
312 *melanogaster* miRNAs is 22 nt). While corresponding reads were detected throughout
313 the Drosophilid phylogeny, we note that their accumulation was modest and efforts to
314 detect them by Northern blotting were negative (data not shown). Thus, while the deep
315 conservation of these hairpins implies functional utility, it remains to be seen if *cwo*
316 miRNAs are matured via non-canonical mechanisms, or if they serve another regulatory
317 role but happen to be sampled in small RNA sequencing. Additional examples of
318 conserved miRNAs are shown in **Supplementary Figure S8** with read details in
319 **Supplementary Figure S9**.

320

321 **Evolutionary shifted processing of some conserved miRNA loci**

322 It is generally assumed that genomic conservation of miRNA loci goes hand-in-
323 hand with conserved processing of mature small RNAs, which in turn are locked into
324 conserved regulatory networks. Since even a 1-nt shift in miRNA 5' identity can redirect
325 its target network, conserved miRNAs are inferred to maintain precise processing.
326 However, in the absence of systematic small RNA sequencing analysis across a genus,
327 the tenets of this assumption have not been challenged by empirical data.

328 We investigated the consistency in 5' end processing for 129 well-conserved
329 Drosophilid miRNAs using our comparative data. As expected, the strong majority of

330 conserved miRNAs exhibit 5' identities (**Supplementary Figure S10**). Of note, some
331 miRNA loci are documented to generate substantial iso-miR species bearing distinct 5'
332 ends. In general, we observed concordance in iso-miR abundance across these
333 *Drosophila* genomes. For example, two iso-miRs from the 3' arm of *pre-mir-79*
334 accumulated at a consistent abundance of approximately 3:1 ratio in all species (**Figure**
335 **4A**). Such conservation in iso-miR 5' end processing supports the notion that multiple
336 species from a single locus are incorporated into evolutionarily constrained regulatory
337 networks. By contrast, the heterogeneous processing of *D. melanogaster mir-193* [29] is
338 not consistent across the Drosophilid phylogeny (**Figure 4B**). In particular, its 5' end
339 shifts by two nucleotides in the virilis subclade relative to the other species. Thus, this
340 conserved miRNA locus appears to be altering its targeting capacity in different species.

341 We identified other cases of clade-specific shifts in processing register for
342 conserved mature strand miRNAs. For example, all *drosophila*-group species
343 consistently processed *mir-959* into a particular species, but a different iso-miR appears
344 substantially in the *Dana/Dpse/Dper/Dwil* ancestor, while Sophophora-group species
345 dominantly accumulate a completely distinct, third iso-miR-959 (**Figure 4B**). We also
346 observe evolutionary shifts for non-canonical loci. This is illustrated by clade-specific
347 shifts in 5' processing for mirtron-derived miR-1014-3p (**Figure 4B**). Such alterations in
348 targeting capacity of conserved miRNAs were hidden until the availability of deep,
349 evolutionary profiling of small RNAs across the genus, and remind us that miRNA
350 processing cannot be well predicted from genomic sequence alone.

351

352 **Antisense miRNAs**

353 Certain miRNA loci are transcribed and processed on both strands [35]. A
354 marquee example in *Drosophila* is *mir-iab-4/mir-iab-8*, for which sense and antisense
355 miRNAs have distinct and genetically overt neural functions [36,37]. Since
356 sense/antisense transcription of this miRNA locus has been observed in beetles [38],
357 one might assume this is a conserved feature throughout the Drosophilids. Indeed, we
358 confidently annotated *mir-iab-4* and *mir-iab-8* in all *Drosophila* species, except for *D.*
359 *persimilis* and *D. grimshawi* where the lower-expressed *mir-iab-8* locus had reads but
360 had to be recovered through the "candidate-rescue" pipeline (**Figure 4C**).

361 We observed 18 other confident sense/antisense miRNA pairs, as well as
362 several dozen candidate antisense miRNAs (**Supplementary Table S5**). A few of these
363 involve conserved miRNAs. The most broadly conserved antisense locus was *mir-307-*

364 AS, which was confidently or candidately detected in seven related Drosophilid species
365 (**Figure 4C**). The modest, but clear, cross-species accumulation of *mir-307-AS* might
366 simply reflect low expression, but alternatively it may have spatially or temporally
367 restricted deployment. However, most antisense miRNAs were poorly conserved, and in
368 fact a substantial fraction of them were present in species other than *D. melanogaster*
369 (**Supplementary Figure S11**). For example, *D. mojavensis* generated a novel
370 sense/antisense locus *dmo_105/dmo_309* adjacent to the deeply conserved intronic
371 miRNAs *mir-994/mir-318* (**Figure 4D**). We present several additional example of
372 sense/antisense miRNA pairs in **Supplementary Figure S11**, including loci that
373 originated apparently de novo in individual species. Overall, it appears that only
374 sense/antisense miRNA pairs have originated across all branches of the Drosophilid
375 phylogeny, but few instances have been retained over substantial periods of evolution.
376

377 **Tailed mirtrons**

378 Amongst miRNAs of non-canonical biogenesis, mirtrons form the dominant class.
379 In the originally described pathway, splicing directly generates a pre-miRNA mimic that
380 is appropriate for export from the nucleus and then cleavage by Dicer [4,5]. Later,
381 alternative "tailed mirtrons" were described, for which further trimming from either the 5'
382 or 3' end must occur to generate the pre-miRNA substrate [31]. In *D. melanogaster*, we
383 identified one 3'-tailed mirtron (*mir-1017*) that is broadly conserved in other flies, along
384 with a handful of other 3'-tailed mirtron candidates [6]. Recently, we appreciated that
385 mammals express a few dozen conventional mirtrons and 3'-tailed mirtrons, but
386 hundreds of hundreds of 5'-tailed mirtrons [26]. Thus, there is differential utilization of
387 tailed mirtron pathways between invertebrates and mammals.

388 Of the 236 novel mirtrons annotated across the *Drosophila* genus in this study
389 (**Figure 2B**), nearly all are of conventional subtype. Thus, flies do not seem to have
390 propensity for tailed mirtrons. Moreover, while our compendia of candidates includes
391 some potential 5'-tailed mirtrons, we do not feel confident in the read patterns to classify
392 any of this class in any *Drosophila* species. Thus, this seems to be a substantial
393 distinction from mammals. Amongst 3'-tailed mirtrons, we recovered *mir-1017* in all
394 species excepting *D. grimshawi*, which has a slightly lower total library size (*mir-1017* is
395 known to be restricted to the nervous system). We also recovered the other five previous
396 proposed 3'-tailed mirtrons in our forward annotation pipeline [6], all of which were
397 known supported by much higher read depth.

398 We also recovered 6 novel confident 3' tailed mirtrons (**Supplementary Table 4**).
399 An example is *D. sechellia dse_288*, which requires removal of a ~12 nt tail following
400 splicing and debranching (**Figure 4E**). Although it is clearly supported by >110 mature
401 and >20 star reads, this locus is not processed as a tailed mirtron in its closely related
402 sister species in the simulans clade. We provide details of another novel 3' tailed mirtron
403 from *D. erecta der_70* in **Supplementary Figure 12**. *Der_70* is notable not only as an
404 atypical mirtron, but also for bearing a non-canonical GC splice donor within the
405 corresponding splice junctions of the CYLD gene of all five melanogaster-group
406 orthologs; the other Drosophilid species harbor a typical GT splice donor. This is the first
407 recognized example of non-canonical mirtron splicing. Curiously, none of the eleven
408 confident 3'-tailed mirtrons we detected have conserved processing, even in closely
409 related sister species.
410

411 **Massive evolutionary flux of testes-restricted miRNA clusters**

412 The second largest class of novel miRNAs we annotated classified as testis-
413 restricted, clustered (TRC) miRNAs, based on their residence in genomic clusters and
414 preferred or exclusive accumulation in male body/testis libraries relative to other tissue
415 libraries. We identified 126 novel TRC miRNAs, all of which were recently-evolved,
416 which represented 17% of all new miRNA annotations. Collectively, this abundance of
417 novel miRNAs clustered into nine novel genomic regions within the genomes of different
418 Drosophila species (**Supplementary Figures S13-16**).

419 Strikingly, we find one or more novel TRC clusters specific to each major
420 Drosophilid branch. Beyond the previously described miRNA clusters composed of
421 conserved and recently-emerged testis-expressed miRNAs [11], we discovered many
422 new TRC specific to *D. ananassae* (2 TRC, containing 34 miRNAs) (**Figure 5A**) or *D.*
423 *willistoni* (1 TRC, containing 19 miRNAs), and orthologous clusters that were only
424 traceable between closely related sister species, such as between *D. virilis* and *D.*
425 *mojavensis* (2 TRC shared by these species, containing 31-34 miRNAs, with 2 additional
426 *virilis*-specific clusters containing 18 miRNAs), or between *D. pseudoobscura* and *D.*
427 *persimilis* (2 TRC shared by these species, containing 47-50 miRNAs) (**Supplementary**
428 **Figures S13-16**). Some of these clusters dwarf the largest previously known fly miRNA
429 clusters. For example, we expanded the membership of the *D. pseudoobscura*
430 *dps_3416→dps-mir-2536* TRC to 36 miRNAs, and identified an orthologous 26-member
431 *D. persimilis* cluster (**Figure 5B**). Small RNA expression showed significantly higher

432 expression in male body and testis libraries than other tissues (**Figure 5B, C**).
433 Interestingly, while miRNAs in the 3' region of this cluster preserve their order between
434 the two species, miRNAs near the 5' end the cluster evolved rapidly via both local gene
435 duplication and *de novo* miRNA emergence, as evident from precursor and miRNA
436 sequence alignments (**Figure 5B** family assignments and **Supplementary Figure S15**).

437 The wealth of male-body sRNA libraries for all 12 genomes and testis data for *D.*
438 *pseudoobscura* and *D. virilis* permitted the evaluation of the relative expression of TRC
439 miRNAs to that of age-matched solo canonical miRNA cohorts. In *D. virilis*, the species
440 with the greatest number of TRC miRNAs (66 in total including known and novel TRC
441 miRNAs), we observed significantly higher expression for TRC miRNAs than for solo
442 canonical miRNAs (Mann-Whitney U-test $p < 10^{-8}$) (**Figure 5D**). Although we observed a
443 similar shift for substantially increased average expression of TRC miRNAs in *D.*
444 *pseudoobscura*, this did not quite achieve significance over age-matched non-TRC
445 canonical miRNAs due to a small number of highly-expressed loci in the latter category
446 (**Supplementary Figure S17**).

447 Altogether, the massive flux of clustered testis miRNA loci across the Drosophilid
448 phylogeny generalizes their distinct evolutionary features that we had established from
449 studies based on a *D. melanogaster*-centric viewpoint. These data provide strong
450 evidence that TRC miRNAs are unlikely to be evolving along a purifying selection route,
451 but instead may be utilized for adaptive regulatory purposes.
452

453 **Distinct rates of gain and losses between miRNAs classes**

454 Estimating rates of evolutionary turnover for genomic elements, be they non-
455 coding RNAs, *cis*-regulatory elements, or protein-coding genes, remains an active area
456 of investigation. miRNA turnover rates have been estimated within *Drosophila*, but these
457 efforts have been limited by uncertainty in miRNA annotations and the number of
458 species considered [23,24,39]. Consequently, many newly-evolved loci have been
459 unaccounted for. In light of new evidence that miRNA sequence and structure evolution
460 are influenced by multiple factors, including biogenesis pathways, clustering state, and
461 testes-biased expression [2,8,11,29], we sought to understand whether such factors
462 influence miRNA birth and death rates.

463 Using our updated *Drosophila* miRNA collection, we set out to characterize and
464 compare rates across our three miRNA classes. To accomplish this, we developed a
465 phylogenetic probabilistic graphical model with the intention of estimating miRNA gene

466 birth and death rates by maximum likelihood (see **Methods**, **Figure 6A**). This method
467 allows us (1) to infer universal, clade-, or branch-specific birth (λ) and death (μ) rate
468 parameters, (2) to predict node-wise ancestral miRNA presence or absence; and (3) to
469 estimate expected counts of edge-wise gain and loss events. We acknowledge that
470 some small RNAs were sampled more deeply in certain species, especially in *D.*
471 *melanogaster* (**Figure 1**). However, besides *D. melanogaster*, there is in general not a
472 strict correlation between the depth of sampling and the number of confidently annotated
473 miRNAs per species. This is due in part to the "rescue" approach (**Figure 3**). Therefore,
474 we chose to apply our estimates of miRNA flux using our full collection of annotations,
475 rather than by attempting to make a new set of annotations by subsampling a lower,
476 fixed number of reads from across the species, since this would inevitably decrease
477 annotation confidence.

478 We applied our method to the pooled collection of miRNA families from our three
479 classes and estimated model parameters (i.e. λ , μ). Using these rate estimates, we then
480 reconstructed node-wise presence and absence of ancestral miRNAs, and subsequently
481 branch-wise miRNA birth and death events, for each miRNA alignment family and
482 across all miRNA classes (summarized in **Figure 6B-D**, see **Supplementary Figure**
483 **S18** for examples of all possible tree configurations, and **Supplementary Figures S19-**
484 **S21** for trees per miRNA alignments across three miRNA classes). Consistent with
485 previous studies, the canonical miRNA class contained the largest number of ancient
486 miRNAs, that is, those present at the root of the *Drosophila* phylogeny. Of 236 *D.*
487 *melanogaster* canonical miRNAs, 106 were clearly present in the Drosophilid ancestor
488 (**Figure 6B**). As mentioned, there are more canonical miRNAs annotated in *D.*
489 *melanogaster* than any other fly species owing to its depth of sequencing; otherwise, the
490 majority of canonical miRNAs that are not testis-restricted are conserved (**Figure 6B**).

491 The tables are turned when examining the fraction of conserved loci in the other
492 categories of miRNAs. The strong majority of testes-restricted canonical miRNAs across
493 the Drosophilid phylogeny, most of which are arranged in genomic clusters, are not
494 conserved. Indeed, only 13 such miRNAs are deeply conserved, and include specific
495 members of the *mir-972*–*979* cluster and the *mir-959*–*964* cluster. Otherwise, most fly
496 species harbor dozens of lineage-restricted TRC miRNAs (**Figure 6C**). Similarly, the
497 mirtron class also contains very few conserved loci. This notion was suggested earlier
498 [24], but we now broadly extend this principle using empirical annotation of mirtrons
499 across 12 *Drosophila* species. While the genomes of several well-profiled and deeply-

500 mined, extant species contained >45 mirtrons (e.g., *D. melanogaster* with 52, *D.*
501 *pseudoobscura* with 50, *D. virilis* with 47 mirtrons), only 7 mirtrons were present at the
502 root of the Drosophilid phylogeny (**Figure 6D**).

503 Next, we computed rates of miRNA birth and death by first aggregating birth and
504 death events across important clades (melanogaster-group, obscura-subgroup and
505 virilis-subgroup) or the entire phylogeny for each miRNA class, and normalizing them by
506 both the total branch length [in Millions of years (Myr)] and by conserved members of
507 each class. These clade-specific and tree-wide rate estimates permitted additional intra-
508 miRNA-class comparisons of total miRNA flux (i.e. birth plus death) in each of these
509 representative clades or across the entire phylogeny. Interestingly, we saw striking rate
510 variation across the three classes of miRNAs (**Figure 6E**). In general, canonical non-
511 testes-restricted miRNAs exhibited the lowest rates of birth, death, and total miRNA flux
512 in each clade and across the *Drosophila* phylogeny when compared to the two other
513 classes. At the other end of the spectrum, mirtrons exhibited the highest rate estimates.
514 This is attributable not only to the large collection of single-species mirtron annotation in
515 our collection (**Figure 6D**), but also to certain atypical patterns of mirtron presence within
516 extant species that do not group along clade boundaries. For example, *dps_22* is
517 present within both obscura-subgroup species and *D. virilis*, but absent in other
518 Drosophilids. This and other cases are shown in **Supplementary Figure S22**.

519 Testes-restricted canonical miRNAs exhibited birth, death and total flux rates in
520 between those of canonical non-testes-restricted miRNA and mirtrons, but showed a
521 significantly elevated death rate within the obscura-subgroup when compared to the
522 other miRNA classes. This observation adds to our previous findings of TRC miRNA
523 death within this clade. That is, orthologs of both *D. pseudoobscura* and *D. persimilis* are
524 missing in part or entirety for two pan-Drosophilid TRC miRNA clusters (*dme-mir-959* →
525 *964* and *dme-mir-972* → *979*) [11].

526 Altogether, these findings support the previous hypothesis for the non-
527 homogeneity in mirtron and canonical miRNA rates of evolution [2]. Moreover, we
528 highlight differential behavior along certain lineages; i.e., with accelerated flux of TRC
529 miRNAs within the obscura-group species.

530
531 **Multiple mechanisms underlie distinct flux behaviors of different miRNA classes**

532 We explored several molecular strategies that could underlie the distinct
533 evolutionary behaviors of different miRNA classes using these comprehensive novel
534 miRNA annotations.

535 **1. cis-mutations.** The impact of nucleotide changes themselves, especially
536 those that are sparse among orthologous pre-miRNAs sequences, are little known on
537 miRNA expression, genesis or decay. We identified compelling sets of miRNA orthologs
538 for experimental tests, including ones exhibiting large variations in apparent biogenesis
539 despite sometimes full genomic identity in the mature miRNA species.

540 For example, the melanogaster-subgroup-specific locus *mir-4984* is well-aligned
541 across five genomes and exhibits only a few substitutions restricted to the miRNA*
542 region of its pre-miRNA, yet large expression differences between orthologs (**Figure**
543 **7A**). Only *D. melanogaster*, *D. simulans*, and *D. sechellia* orthologs are processed, and
544 no reads were mapped to the *D. yakuba* and *D. erecta* orthologs are not expressed.
545 Even between expressed orthologs, there is a >10-fold reduced normalized RPM
546 expression in *D. simulans* (0.06) as compared to *D. melanogaster* (0.66) and *D.*
547 *sechellia* (0.82) orthologs that may be driven by a single miRNA* A-to-G substitution. We
548 validated these expression levels deduced from the sRNA-seq data with independent
549 Northern blot assays (**Figure 7B**). That is, we detected both the pre-miRNA and mature
550 products for *dme-mir-4984*, yet failed to detect mature species for both *dya-mir-4984*
551 and *dsi-mir-4984*. Another example is the alignment of *dps_41* and *dpe_2484*, which are
552 specific to the obscura-subgroup. *dps_41* is >100 times higher expressed than *dpe_41*
553 in the sRNA data (**Figure 7C**) and *dpe_41* is undetectable in the corresponding Northern
554 blot (**Figure 7D**).

555 We broadened this analysis using *D. pseudoobscura/persimilis*, which had
556 advantages for being a closely related (0.93 Myr) sister pair for which we had identified
557 numerous novel miRNA annotations that might potentially be subject to expression
558 fluctuation. We labeled miRNA alignments with ≥ 6 -fold $\log_{10}(\text{RPMM})$ expression
559 difference as differentially-expressed, and identified six conserved and nine newly-
560 evolved miRNAs as such (**Supplementary Figure S23**). The conserved loci included
561 several members of the *mir-309* cluster and seemed to be a sampling artifact given they
562 are expressed as a highly stage-specific operon [40]. Otherwise, there was a high
563 correlation (0.938) between the remaining 205 *D. pseudoobscura* and *D. persimilis*
564 ortholog pairs (i.e. < 6-fold change) (**Supplementary Figure S23**).

565 We sought to investigate the sequence features between the 15 differentially-
566 expressed, obscura-group ortholog pairs by investigating the patterns of the miR:miR*
567 duplex conservation. In all cases, orthologs shared high sequence and structure
568 similarity apart from a few substitutions. For example, we identified *dps-mir-2567* and
569 *dpe_2484* as differentially-expressed, recently-evolved orthologs (*dps-mir-2567* = 12.97
570 RPM; *dpe_2484* = 0.11 RPM) (**Supplementary Figure S23**). Between these two
571 species, we observed five duplex substitutions, one of which resides in the seed region
572 of the dominant 5' arm. In light of these apparent substitutions, we asked if duplex
573 substitutions were more prevalent between differentially-expressed miRNAs than non-
574 differentially expressed ones for both conserved and newly-evolved miRNAs. Indeed,
575 within the obscura-group we saw significantly more differentially expressed miRNAs with
576 duplex substitutions within the newly-evolved group (Fisher's Exact Test [FET] P <
577 0.003), and within the conserved group (FET P = 0.03) (**Figure 7E**).

578 **2. Adaptive seed mutations of TRC miRNAs.** Functional miRNAs are not
579 expected not to diverge between closely related species, especially within seed regions.
580 However, we previously used *D. melanogaster* population data and melanogaster group
581 species orthologs to provide evidence for adaptive evolution of TRC miRNAs in this
582 clade, including within seeds [11]. There is limited population data in other Drosophilids,
583 but we investigated polymorphisms from whole genome sequences of 11 North
584 American *D. pseudoobscura* strains and 2 *D. pseudoobscura bogotana* sub-species
585 (data available from <http://pseudobase.biology.duke.edu/>) [41]. To identify unambiguous
586 divergences between species, we focused on miRNAs with clear 1-to-1 orthologs, such
587 as the miRNAs within the 3' sub-cluster region of the *dps_3416*–*dps-mir-2536* cluster.

588 We identified two TRC miRNAs with seed divergences in this sub-cluster. For
589 example, the mature (3') arm of *dps-mir-2523* contained a G-to-T substitution at the 8th
590 seed position relative to its *D. persimilis* ortholog *dpe_106* (**Figure 7F**). Analysis of the
591 *D. pseudoobscura* population data indicated that all individuals were monomorphic for
592 the 'T' allele (i.e. a fixed difference). We also observed several other non-seed positions
593 of divergence within the star strand even within the *D. pseudoobscura* population, which
594 is unusual and suggests fast evolution. As another example, we observe that both
595 mature and star arms of *dps-mir-2542-1* exhibit multiple positions of seed divergence
596 with its ortholog *dpe_101* (**Figure 7F**). Despite the impracticality of applying formal tests
597 for evidence of natural selection given such a small sample, it is evident from these
598 examples that several obscura TRC loci defy conventional behavior for purifying

599 selection of seed regions and instead are quickly altering their seed regions between
600 closely related species while still maintaining miRNA biogenesis.

601 **3. Preferential 3' untemplated uridylation of mirtrons.** Approximately 54%
602 (232/428) of mirtron and tailed-mirtron annotations in *Drosophila* are new to our study,
603 and are recently-emerged. This massive set of novel spliced miRNAs prompted us to
604 ask if they exhibit characteristic properties of 3'-untemplated uridylation, as we reported
605 in *D. melanogaster* [42,43]. In comparisons of 1777 canonical miRNAs and 289 mirtrons
606 (**Supplementary Table S6**), we observed that mirtrons exhibited a massively greater
607 rate of 3' untemplated uridylation than canonical miRNAs (**Figure 7G, H**). This was
608 evident on a per-locus basis on CDF plots (**Supplementary Figure S24A**), and was also
609 the case even after conditioning on canonical miRNAs whose 3' arm read ended in AG
610 dinucleotide as with mirtrons. However, conditioning on 3'-AG enhanced the frequency
611 of 3' uridylation observed on canonical miRNA-3p species (**Supplementary Figure**
612 **S24A**), consistent with the notion that the mirtron uridylyltransferase Tailor has some
613 intrinsic preference for hairpins terminating in 3'-G/AG [42,43].

614 Next, we examined whether the elevated frequency of uridylation at canonical
615 miRNAs and mirtrons whose 3' read ended with G was consistent across conserved and
616 newly-evolved loci. For the conserved loci, we recapitulated previous signatures of
617 uridylation (**Figure 7I**). Namely, mirtrons exhibited a significantly higher frequency of
618 uridylation than canonical miRNAs (Mann-Whitney Test [M.W.T] $p < 10^{-21}$), and
619 comparisons among canonical miRNAs revealed that loci whose 3' arm read ended with
620 G were more uridylated than loci ending in a base other than G (i.e. IUPAC 'H' ambiguity
621 character) (M.W.T $p < 10^{-12}$). Of note, newly-evolved mirtrons and canonical miRNAs
622 also exhibited the same signature as conserved loci (**Figure 7J**), and comparisons of
623 loci within individual species revealed similar significant results in many species
624 (**Supplementary Figure S24B**). Altogether, these findings from small RNA sequencing
625 across the Drosophilid phylogeny broadly support the notion that adventitious access of
626 splicing-derived hairpins to Dicer (i.e., mirtrons) continually creates a cohort of largely
627 undesirable miRNA substrates, that are suppressed via Tailor-mediated uridylation that
628 is in part sensitive to terminal hairpin "G".

629

630 **Discussion**

631

632 **A deep and broad empirical analysis of miRNA flux across the *Drosophila* genus**

633 We are now 15 years into molecular evolutionary analysis of *Drosophila* miRNAs,
634 but until now, there have been only limited attempts to address using the empirical data
635 sampled throughout the *Drosophila* genus. In this study, we extended our curation of
636 miRNAs from ~1.9 billion *D. melanogaster* sRNA reads by sequencing ~1.5 billion
637 sRNAs from 11 other *Drosophila* species, assessing a diversity of samples optimized for
638 miRNA discovery. Beyond the first experimental cloning of 592 orthologs of conserved
639 miRNAs, we used a multitude of annotation pipelines and rigorous scoring criteria to
640 annotate 649 completely novel miRNAs across 12 genomes. We carefully assessed
641 these for ortholog relationships as a foundation for assessing evolutionary flux, including
642 meticulous manual assignment of orthologs, paralogs, and newly-emerging members of
643 genomic clusters (i.e. rapidly evolving TRC loci).

644 Overall, these data yield myriad insights into miRNA loci that are hidden from
645 genomic alignments. These include the surprising existence of novel conserved
646 miRNAs, unexpected clade-specific shifts in processing register, and post-transcriptional
647 modifications of miRNAs. Beyond conserved loci, we uncover hundreds of "young"
648 miRNAs that could not be identified by genomic sequence alone. These data allow us to
649 quantify distinct rates of miRNA flux according to biogenesis type, genomic locale, tissue
650 restriction, and evolutionary clade. We identify patterns of structural change and
651 associated with flux in expression of evolutionarily nascent canonical miRNAs, providing
652 a mechanistic basis for their instability. We also develop a new phylogenetic model to
653 characterize rates of small RNA evolution in the presence of annotation uncertainty.
654 Overall, we solidify the perspective that miRNAs do not comprise a unitary class, but
655 encompass a diversity of functional loci with distinct evolutionary imperatives.

656 While our study provides the deepest perspective of miRNA evolutionary novelty
657 across a genus to date, clear challenges remain for the future. We previously described
658 evolutionary nascent miRNA-like loci in *D. melanogaster* that defy clear annotation by
659 current standards [28,29]. In the future, analysis of a greater diversity of Drosophilid
660 tissue and cell samples, combined with AGO1-IP sequencing, will be critical to interpret
661 the earliest stages of miRNA emergence. At the same time, it worth appreciating that the
662 current depth and breadth of sequencing is sufficient to identify hundreds of species-
663 restricted, and even species-specific miRNAs with high confidence.

664

665 **Divergent rationales for rapid evolution of mirtrons and TRC miRNAs**

666 Amongst our extensive collection of recently-emerged miRNAs, we discern two
667 major subclasses of rapidly evolving loci, splicing-derived miRNAs (i.e., mirtrons) and
668 testis-restricted clustered (i.e., TRC) miRNAs. We propose divergent functional
669 explanations for their distinct evolutionary behavior, relative to the bulk collection of
670 recently-emerged miRNAs that either evolve under mild purifying selection or lack
671 substantial utility and evolve neutrally [8].

672 Mirtrons mature via the dominant non-canonical mechanism that bypasses the
673 Drosha/DGCR8 "Microprocessor", which otherwise serves as a molecular gatekeeper for
674 generation of specific and accurate Dicer substrate hairpins. Mirtrons occasionally yield
675 regulatory species that incorporate into beneficial regulatory networks, but the vast
676 majority are not retained during evolution. We hypothesize that most mirtrons are
677 adventitious Dicer substrates whose net regulatory capacity may be undesirable. By
678 analogy, expressing synthetic, random, siRNAs are often used as a control situation, but
679 probably one can imagine that they would impart fortuitous gene regulation that would
680 not be tolerated over evolution. Indeed, molecular mechanisms involving uridylation
681 have recently been shown to selectively suppress splicing-mediated miRNA biogenesis,
682 which can accelerate the evolutionary flux of these miRNA substrates [42,43].

683 Our current studies across the *Drosophila* genus broadly confirms that the
684 accelerated evolutionary dynamics of mirtrons is well-correlated with their tremendously
685 high rates of 3' uridylation and evolutionary turnover. Indeed, only six of the >400
686 mirtrons we annotated across the Drosophilid phylogeny were present in the fruitfly
687 ancestor. We and others recently showed that the mechanism is mediated by the
688 uridyltransferase Tailor, which has capacity to recognize hairpins bearing 3'-(A)G, which
689 is characteristic for splicing-derived hairpins [42,43]. We hypothesized that this may
690 have carryover effect on suppressing the evolutionary emergence of canonical miRNAs
691 that happen to end in 3'-(A)G. Indeed, our broad survey using de novo miRNA
692 annotation in eleven new *Drosophila* species provides evidence that newly-emerging
693 canonical miRNA hairpins that end in 3'-G have more uridylation than ones that end in
694 the other three nucleotides. This can explain the rapid evolutionary turnover of mirtrons,
695 as well as the preferential uridylation of canonical miRNA-3p species ending in G of
696 various evolutionary ages, which we observed to be depleted in the collection of deeply
697 conserved, canonical miRNAs [42,43]. This supports the overall view of how molecular

698 mechanisms that restrict the biogenesis of non-canonical miRNA substrates can affect
699 the evolution of both non-canonical and canonical miRNA evolution.

700 We note also that almost all our annotations of the related 3'-tailed mirtrons,
701 excepting *mir-1017*, have been lost from the Drosophilid phylogeny. In this case, since
702 their hairpin tails are trimmed, they lack a characteristic 3' nucleotide for potential
703 identification. We have also annotated a class of confident novel antisense miRNAs,
704 although excepting *mir-iab-4/mir-iab-8* and possibly *mir-307/mir-307AS*, it appears that
705 all other instances of these have been purged over evolution. Thus, each of these
706 specialized biogenesis pathways seems largely to exist to manufacture one or perhaps
707 two miRNAs, even though biochemical capacity to make more exists in most species.
708 Therefore, we speculate that there may be additional strategies to suppress miRNA
709 biogenesis than we currently appreciate, especially that are suited to recognize aberrant
710 miRNA substrates. For example, at least some 3' tailed mirtrons are associated with
711 untemplated 3' additions (**Supplementary Figure S12**). It remains to be seen if this is
712 also associated with an inhibitory role as with conventional mirtrons, and if so, how 3'-
713 tailed substrates are recognized.

714 On the other hand, the extraordinarily rapid dynamics of TRC miRNAs in all
715 subclades of the *Drosophila* genus provides strong evidence for their positive selection
716 and adaptive evolution. Not only do TRC miRNA sequences evolve more quickly than
717 canonical miRNA substrates of matched age, the total flux in TRC miRNA numbers
718 between *Drosophila* subclades outpaces that of canonical miRNA loci. For example, a
719 clear sequence ortholog of 106/497 canonical miRNAs not in the TRC class were
720 present in the pan-Drosophilid ancestor, whereas this is only true of 13/265 TRC miRNA
721 loci (**Figure 6C**). We entertain an alternative interpretation that some TRC miRNAs,
722 owing to positive selection, have evolved in primary sequence so quickly that their
723 ancestral relationships are not possible to assess. In any case, it is clear that the
724 wholesale appearance and disappearance of massive TRC loci in different clades
725 reflects a fundamentally different usage of these miRNAs than for maintenance of
726 conserved seed-driven target networks as with typical canonical miRNAs.

727 Moreover, the atypical dynamics of TRC miRNAs are dramatically accelerated in
728 both species examined in the obscura subclade. In fact, *D. pseudoobscura* and *D.*
729 *persimilis* themselves exhibit substantial differences in their TRC repertoire, underlying
730 nearly an order of magnitude greater birth estimate in the obscura branch than other
731 branches of the phylogeny. The functional underpinnings of this remain to be tested, but

732 they go hand-in-hand with the observation of dramatic proliferations of testis-restricted
733 AGO2 paralogs specifically in the obscura subclade, and not in other Drosophila
734 subclades [44,45].

735 Overall, our study provides a wealth of small RNA data that can guide functional
736 studies of miRNA biogenesis, regulation of miRNA processing, and will underlie
737 discovery of novel small RNA types (such as siRNAs and piRNAs). In addition, our deep
738 and broad sampling across an entire genus provides myriad insights into the distinct
739 evolutionary trajectories of multiple miRNA subtypes, affirming that miRNAs cannot be
740 considered a unitary class with respect to their functional impact and utilization.

741 **Materials and Methods**

742

743 ***Drosophila* samples and small RNA library sequencing**

744 To analyze miRNA evolution in *Drosophila* species, we obtained cultures of
745 whole-genome sequenced *D. simulans*, *D. sechellia*, *D. yakuba*, *D. erecta*, *D.*
746 *ananassae*, *D. pseudoobscura*, *D. persimilis*, *D. willistoni*, *D. virilis*, and *D. mojavensis*
747 strains from the UCSD *Drosophila* Species Stock Center. Adult *D. grimshawi* samples
748 were a gift of Dr. Kevin White (University of Chicago). Small RNAs (~18-28nt) were
749 isolated from male bodies, female bodies, heads, and mixed embryos using
750 polyacrylamide gel electrophoresis, and we prepared libraries as described. Libraries
751 were sequenced on Illumina GAIIx or Hi-Seq 2000 instruments.

752

753 **Recovery of *D. melanogaster* miRNAs from simulated libraries**

754 We simulated four libraries each representing 25 million randomly sampled reads
755 from the pooled collection of *D. melanogaster* sRNAs from male bodies, female bodies,
756 mixed embryos, and heads respectively. We created 100 samples of 100M reads each
757 (25M X 4 libraries), and determined the recoverability of conserved and newly-evolved
758 *D. melanogaster* miRNAs at varying minimum mature and star read expression
759 thresholds. These artificial libraries, allowed us to investigate the amount of miRNAs at
760 two different age groups (i.e. conserved and newly-evolved) that could be recovered in
761 the other 11 *Drosophila* genomes. We defined conserved *D. melanogaster* miRNAs as
762 those with unambiguous orthologs in the obscura-group species, *D. willistoni*, or the
763 *Drosophila*-group species. Recently-evolved *D. melanogaster* miRNAs were defined as
764 those with orthologs within the melanogaster-group species only.

765

766 **Annotation of miRNA genes**

767 To identify novel miRNAs and assess the expression levels of all miRNA loci, we
768 first mapped reads from each of the 11 *Drosophila* species unto their reference
769 genomes. All reference genomes, except for *D. simulans*, were obtained from Flybase.
770 We utilized a revised *D. simulans* genome assembly created from an isogenic *w501*
771 female within our analysis [46]. Reads were mapped using the Bowtie program by
772 allowing for up to 3 mismatches (parameters: -v 3 -k 20 --best --strata). Perfectly
773 mapped reads, and reads with 3' end mismatches characteristic of untemplated
774 additions, were used for the identification of miRNAs.

775 We supplemented existing *Drosophila* miRNA annotations from miRBase v21
776 with novel miRNAs and mirtrons identified in this study using a multi-stage pipeline [25].
777 First, canonical miRNA and mirtrons were predicted using miRDeep2 using default
778 software settings [13]. To identify short mirtron and long pre-miRNA hairpins, two
779 classes systematically missed by miRDeep2, we mapped sRNA datasets to introns
780 obtained from FlyBase gene annotations, and predicted hairpin structures from the
781 *einverted* program of the EMBOSS package [47] in a genome-wide manner per species.
782 We used the *invert_it.pl* utility script from the ShortStack program [48] to filter the
783 *einverted* results. The parameters specified to this script were: -f 0.6 -p 30. Introns
784 or hairpin structures with at least one mapped read were retained and ranked by p-
785 values calculated from a Random Forest classifier. We trained this classifier with a
786 balanced set of positive training examples comprised of known *D. melanogaster* and *D.*
787 *pseudoobscura* miRNAs download from miRBase (v21), and a negative training set
788 composed on non-miRNA predictions identified manually in this study. We used a total
789 of 37 features per training case representing sequence, structure, and sRNA read
790 alignment features (**Supplementary Table S7**). Minimum free energy, and sub-optimal
791 secondary structures were predicted using RNAfold and RNAsubopt in the Vienna RNA
792 Software [49].

793 All miRNAs predicted from this pipeline were vetted manually and
794 bioinformatically for miRNA and mirtron candidacy. In the manual phase, all miRDeep2
795 predictions, and intron and hairpin structures with $p > 0.5$ were examined for evidence of
796 cleavage by Drosha and Dicer based on the sRNA read alignment, a hairpin secondary
797 structure, and synteny with other miRNA predictions. Putative canonical miRNA were
798 further classified bioinformatically using criteria based on (1) expression, (2) clonability of
799 Drosha/Dicer products, such as the miR, miR*, loop, or MOR sequences, (3) structure
800 pairing of the miR:miR* duplex, (4) 5' end consistency of miR and miR* reads, and (6)
801 ratio of background to miRNA reads (**Supplementary Figure S3**). Mirtrons were
802 classified using the same features as for canonical miRNAs, but an additional criterion
803 for untemplated modifications of the 3' arm reads was specified. Canonical miRNA or
804 mirtron predictions that met all criteria were labeled as "confident" while those that failed
805 some or all criteria were labeled as "candidate" or "FALSE," respectively. "Candidate"
806 loci that were orthologous to "confident" annotations were re-classified as "candidate-
807 rescued." Confidence classifications for all miRNA and mirtrons are provided in
808 **Supplementary Table S4**. Finally, novel "confident", "candidate-rescued", and

809 “candidate” miRNAs and mirtrons were segregated from miRBase annotation and their
810 unannotated orthologs.

811

812 Identification of miRNA clusters and testes-restricted miRNAs

813 miRNA clusters were identified by grouping genes within a 10 Kb window of each
814 other. Mirtrons were excluded from this classification. The majority of miRNA clusters
815 identified in this study comprised genes with testes-restricted expression. Testes-
816 restricted miRNAs were characterized as genes with > 4-fold \log_{10} (RPMM) testis or
817 male-body expression enrichment when compared against all other tissue and
818 developmental-timepoint libraries. If >75% of miRNA genes within a cluster were
819 classified as testes-restricted, then all genes within said cluster were labeled Canonical,
820 Testes-restricted, Clustered miRNAs.

821

822 Identification of miRNA orthologs and alignments

823 miRNA orthologs were identified using the LASTZ program with the following
824 parameters: H=2000 Y=3400 L=4000 K=2200 Q=HoxD55.q [50]. Hits were ranked by a
825 score based on the consistency, continuity, and percent identity metrics from LASTZ. A
826 12-species sequence alignment were created for each miRNA prediction using best
827 scoring orthologs and the Fast Statistical Aligner program [51]. Paralogs were a
828 byproduct of this procedure because they attained lower rank during orthology
829 assignments. All orthologs and paralogs were automatically included in our annotation
830 pipeline, and were vetted by the same criteria.

831 Our process of collating accurate miRNA synteny information played a crucial
832 role within our annotation procedure because it allowed us to “rescue” low-evidence,
833 candidate miRNA and mirtron predictions contingent on the availability of one or more
834 confident orthologs. Automated approaches were inept for identifying testes-restricted,
835 clustered miRNA orthologs due to their close genomic proximity, low sequence
836 conservation, and high rates of tandem duplication. For these miRNAs, orthologs and
837 paralogs were identified from multiple sequence alignments of all members of a
838 particular cluster in all species with homologs (see **Supplementary Figures S9** for
839 examples). MiRNA alignments were further grouped into three classes: (1) TRC
840 miRNAs, (2) solo canonical miRNAs, and (3) mirtrons.

841

842 Birth and Death Model

843 To assess birth and death rate variation across classes of miRNAs and across
844 *Drosophila* clades of interest, we designed and implemented a phylogenetic probabilistic
845 graphical model. This model permits estimation of parameters of gene birth (λ) and
846 death (μ) (**Figure 6A**) based upon our assignments of miRNA presence and absence in
847 each species per miRNA family alignment. Parameter estimation required two sets of
848 precomputed data. The first datum needed was a binary encoding of miRNA presence
849 (1) and absence (0) as leaf node labels of the phylogenetic model. In this regard, we
850 labeled non-miRNAs and “candidate” miRNAs as cases of absences, and “confident” or
851 “candidate-rescued” miRNAs as cases of presences. We assigned each miRNA for
852 which no orthologs were identified to its own singleton miRNA alignment, to be counted
853 as an independent birth event. The second datum needed was phylogenetic branch-
854 length estimates for the 12 *Drosophila* species phylogeny [27]. We estimated branch
855 lengths (τ) in units of substitutions per site, using fourfold degenerate sites (i.e. sites
856 within a codon in which all four possible nucleotide substitutions are synonymous) and
857 the maximum-likelihood program RaXML [52]. Fourfold degenerate sites were extracted
858 from a *de novo* 12 *Drosophila* species whole genome alignment constructed using the
859 LASTZ and MULTIZ programs and the chaining and netting protocol used for the UCSC
860 Genome Browser. The resulting maximum-likelihood newick-formatted tree was:
861 ((((((dm3:0.055153,(droSim1:0.027716,droSec1:0.023941):0.024430):0.050893,(droYak
862 2:0.090814,droEre2:0.079010):0.032754):0.328435,droAna3:0.466508):0.162763,(dp4:0
863 .018457,droPer1:0.018684):0.407262):0.120336,droWil1:0.593093):0.118858,((droVir3:
864 0.244781,droMoj3:0.335567):0.082788,droGri2:0.319783):0.118858);

865 Given these two datasets, we used our model to infer maximum-likelihood
866 parameter estimates (i.e. λ , μ) using the standard belief-propagation algorithm to
867 compute likelihoods, and the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm to
868 obtain maximum-likelihood parameter estimates [53]. Parameter estimates were
869 computed for the merged miRNA collection (final estimate: $\lambda = 0.292$, $\mu = 0.694$), which
870 we later used to compute (1) ancestral gene presence or absence states, and (2)
871 probabilities of observable edge-wise birth and death events (**Supplementary Figure**
872 **S18-21**). To assess cumulative counts of observable birth and death events per miRNA
873 class or *Drosophila* clade, we computed edge-wise joint posterior probabilities (i.e.
874 $P(\text{child}, \text{parent})$) by belief propagation. For simplicity, we called birth [$P(1, 0)$], death
875 [$P(0, 1)$], and “no change” events [$P(0, 0)$ or $P(1, 1)$] if these probability estimates were

876 ≥ 0.5 (**Figure 6B-D**). This method is implemented as a Java software package and
877 available upon request.

878

879 **Data access**

880 *Drosophila* sRNA sequencing data are under submission for access via the NCBI
881 Gene Expression Omnibus. Read pileup, structure prediction, and 12-fly sequence
882 alignments of all *Drosophila* miRNAs are provided as Supplementary material via an
883 online website (http://compgen.cshl.edu/mirna/12flies/12flies_alignments.html).

884

885 **Acknowledgements**

886 We thank Matthew DeCruz for technical assistance and the UCSD *Drosophila*
887 Species Stock Center for fly stocks. Work in E.C.L.'s group was supported by the
888 National Institutes of Health (R01-GM083300 and R01-NS083833) and MSK Core Grant
889 P30-CA008748. J.M. was supported in part by the Tri-Institutional Training Program in
890 Computational Biology and Medicine (via NIH training grant 1T32-GM083937).

891 **Bibliography**

- 892 1. Flynt AS, Lai EC (2008) Biological principles of microRNA-mediated regulation:
893 shared themes amid diversity. *Nature reviews Genetics* 9: 831-842.
- 894 2. Axtell MJ, Westholm JO, Lai EC (2011) Vive la différence: biogenesis and evolution
895 of microRNAs in plants and animals. *Genome biology* 12: 221-221.213.
- 896 3. Yang JS, Lai EC (2011) Alternative miRNA biogenesis pathways and the
897 interpretation of core miRNA pathway mutants. *Molecular cell* 43: 892-903.
- 898 4. Okamura K, Hagen JW, Duan H, Tyler DM, Lai EC (2007) The mirtron pathway
899 generates microRNA-class regulatory RNAs in *Drosophila*. *Cell* 130: 89-100.
- 900 5. Ruby JG, Jan CH, Bartel DP (2007) Intronic microRNA precursors that bypass
901 Drosha processing. *Nature* 448: 83-86.
- 902 6. Flynt AS, Chung WJ, Greimann JC, Lima CD, Lai EC (2010) microRNA biogenesis
903 via splicing and exosome-mediated trimming in *Drosophila*. *Molecular cell* 38:
904 900-907.
- 905 7. Lai EC, Tomancak P, Williams RW, Rubin GM (2003) Computational identification of
906 *Drosophila* microRNA genes. *Genome biology* 4: R42.41-R42.20.
- 907 8. Mohammed J, Flynt AS, Siepel A, Lai EC (2013) The impact of age, biogenesis, and
908 genomic clustering on *Drosophila* microRNA evolution. *RNA* 19: 1295-1308.
- 909 9. Lim L, Lau N, Weinstein E, Abdelhakim A, Yekta S, et al. (2003) The microRNAs of
910 *Caenorhabditis elegans*. *Genes & development* 17: 991 - 1008.
- 911 10. Lu J, Fu Y, Kumar S, Shen Y, Zeng K, et al. (2008) Adaptive evolution of newly-
912 emerged microRNA genes in *Drosophila*. *Molecular biology and evolution*.
- 913 11. Mohammed J, Bortolamiol-Becet D, Flynt AS, Gronau I, Siepel A, et al. (2014)
914 Adaptive evolution of testis-specific, recently-evolved, clustered miRNAs in
915 *Drosophila*. *RNA* 20: 1195-1209.
- 916 12. Lyu Y, Shen Y, Li H, Chen Y, Guo L, et al. (2014) New microRNAs in *Drosophila*--
917 birth, death and cycles of adaptive evolution. *PLoS genetics* 10: e1004096.
- 918 13. Friedlander MR, Mackowiak SD, Li N, Chen W, Rajewsky N (2012) miRDeep2
919 accurately identifies known and hundreds of novel microRNA genes in seven
920 animal clades. *Nucleic acids research* 40: 37-52.
- 921 14. Friedlander MR, Chen W, Adamidi C, Maaskola J, Einspanier R, et al. (2008)
922 Discovering microRNAs from deep sequencing data using miRDeep. *Nature
923 biotechnology* 26: 407-415.
- 924 15. Christodoulou F, Raible F, Tomer R, Simakov O, Trachana K, et al. (2010) Ancient
925 animal microRNAs and the evolution of tissue identity. *Nature* 463: 1084-1088.
- 926 16. Grimson A, Srivastava M, Fahey B, Woodcroft BJ, Chiang HR, et al. (2008) Early
927 origins and evolution of microRNAs and Piwi-interacting RNAs in animals. *Nature*
928 455: 1193-1197.
- 929 17. Mohammed J, Siepel A, Lai EC (2014) Diverse modes of evolutionary emergence
930 and flux of conserved microRNA clusters. *RNA* 20: 1850-1863.
- 931 18. de Wit E, Linsen SE, Cuppen E, Berezikov E (2009) Repertoire and evolution of
932 miRNA genes in four divergent nematode species. *Genome research* 19: 2064-
933 2074.

- 934 19. Shi Z, Montgomery TA, Qi Y, Ruvkun G (2013) High-throughput sequencing reveals
935 extraordinary fluidity of miRNA, piRNA, and siRNA pathways in nematodes.
936 Genome research 23: 497-508.
- 937 20. Sarkies P, Selkirk ME, Jones JT, Blok V, Boothby T, et al. (2015) Ancient and novel
938 small RNA pathways compensate for the loss of piRNAs in multiple independent
939 nematode lineages. PLoS biology 13: e1002061.
- 940 21. Meunier J, Lemoine F, Soumillon M, Liechti A, Weier M, et al. (2012) Birth and
941 expression evolution of mammalian microRNA genes. Genome research
942 10.1101/gr.140269.112.
- 943 22. Berezikov E, Thuemmler F, van Laake LW, Kondova I, Bontrop R, et al. (2006)
944 Diversity of microRNAs in human and chimpanzee brain. Nature genetics 38:
945 1375-1377.
- 946 23. Lu J, Shen Y, Wu Q, Kumar S, He B, et al. (2008) The birth and death of microRNA
947 genes in Drosophila. Nature genetics 40: 351-355.
- 948 24. Berezikov E, Liu N, Flynt AS, Hedges E, Rooks M, et al. (2010) Evolutionary flux of
949 canonical microRNAs and mirtrons in Drosophila. Nature genetics 42: 6-9.
- 950 25. Kozomara A, Griffiths-Jones S (2014) miRBase: annotating high confidence
951 microRNAs using deep sequencing data. Nucleic acids research 42: D68-73.
- 952 26. Wen J, Ladewig E, Shenker S, Mohammed J, Lai EC (2015) Analysis of Nearly One
953 Thousand Mammalian Mirtrons Reveals Novel Features of Dicer Substrates.
954 PLoS computational biology 11: e1004441.
- 955 27. Clark AG, Eisen MB, Smith DR, Bergman CM, Oliver B, et al. (2007) Evolution of
956 genes and genomes on the Drosophila phylogeny. Nature 450: 203-218.
- 957 28. Wen J, Mohammed J, Bortolamiol-Becet D, Tsai H, Robine N, et al. (2014) Diversity
958 of miRNAs, siRNAs and piRNAs across 25 Drosophila cell lines. Genome
959 research 24: 1236-1250.
- 960 29. Berezikov E, Robine N, Samsonova A, Westholm JO, Naqvi A, et al. (2011) Deep
961 annotation of Drosophila melanogaster microRNAs yields insights into their
962 processing, modification, and emergence. Genome research 21: 203-215.
- 963 30. Westholm JO, Ladewig E, Okamura K, Robine N, Lai EC (2012) Common and
964 distinct patterns of terminal modifications to mirtrons and canonical microRNAs.
965 RNA 18: 177-192.
- 966 31. Westholm JO, Lai EC (2011) mirtrons: microRNA biogenesis via splicing. Biochimie
967 93: 1897-1904.
- 968 32. Chak LL, Mohammed J, Lai EC, Tucker-Kellogg G, Okamura K (2015) A deeply
969 conserved, noncanonical miRNA hosted by ribosomal DNA. RNA 21: 375-384.
- 970 33. Han J, Pedersen JS, Kwon SC, Belair CD, Kim YK, et al. (2009) Posttranscriptional
971 crossregulation between Drosha and DGCR8. Cell 136: 75-84.
- 972 34. Smibert P, Bejarano F, Wang D, Garaulet DL, Yang JS, et al. (2011) A Drosophila
973 genetic screen yields allelic series of core microRNA biogenesis factors and
974 reveals post-developmental roles for microRNAs. RNA 17: 1997-2010.
- 975 35. Tyler DM, Okamura K, Chung WJ, Hagen JW, Berezikov E, et al. (2008)
976 Functionally distinct regulatory RNAs generated by bidirectional transcription and
977 processing of microRNA loci. Genes & development 22: 26-36.
- 978 36. Garaulet DL, Castellanos MC, Bejarano F, Sanfilippo P, Tyler DM, et al. (2014)
979 Homeotic Function of Drosophila Bithorax-Complex miRNAs Mediates Fertility by

- 980 Restricting Multiple Hox Genes and TALE Cofactors in the CNS. *Developmental*
981 cell 29: 635-648.
982 37. Picao-Osorio J, Johnston J, Landgraf M, Berni J, Alonso CR (2015) MicroRNA-
983 encoded behavior in *Drosophila*. *Science* 350: 815-820.
984 38. Hui JH, Marco A, Hunt S, Melling J, Griffiths-Jones S, et al. (2013) Structure,
985 evolution and function of the bi-directionally transcribed iab-4/iab-8 microRNA
986 locus in arthropods. *Nucleic acids research* 41: 3352-3361.
987 39. Nozawa M, Miura S, Nei M (2010) Origins and evolution of microRNA genes in
988 *Drosophila* species. *Genome biology and evolution* 2: 180-189.
989 40. Bushati N, Stark A, Brennecke J, Cohen SM (2008) Temporal Reciprocity of
990 miRNAs and Their Targets during the Maternal-to-Zygotic Transition in
991 *Drosophila*. *Curr Biol* 18: 501-506.
992 41. McGaugh SE, Heil CS, Manzano-Winkler B, Loewe L, Goldstein S, et al. (2012)
993 Recombination modulates how selection affects linked sites in *Drosophila*. *PLoS*
994 biology 10: e1001422.
995 42. Bortolamiol-Becet D, Hu F, Jee D, Wen J, Okamura K, et al. (2015) Selective
996 Suppression of the Splicing-Mediated MicroRNA Pathway by the Terminal
997 Uridyltransferase Tailor. *Molecular cell* 59: 217-228.
998 43. Reimao-Pinto MM, Ignatova V, Burkard TR, Hung JH, Manzenreither RA, et al.
999 (2015) Uridylation of RNA Hairpins by Tailor Confines the Emergence of
1000 MicroRNAs in *Drosophila*. *Molecular cell* 59: 203-216.
1001 44. Lewis SH, Salmela H, Obbard DJ (2016) Duplication and Diversification of Dipteran
1002 Argonaute Genes, and the Evolutionary Divergence of Piwi and Aubergine.
1003 *Genome biology and evolution* 8: 507-518.
1004 45. Lewis SH, Webster CL, Salmela H, Obbard DJ (2016) Repeated Duplication of
1005 Argonaute2 Is Associated with Strong Selection and Testis Specialization in
1006 *Drosophila*. *Genetics* 204: 757-769.
1007 46. Hu TT, Eisen MB, Thornton KR, Andolfatto P (2012) A second generation assembly
1008 of the *Drosophila simulans* genome provides new insights into patterns of
1009 lineage-specific divergence. *Genome research* 10.1101/gr.141689.112.
1010 47. Rice P, Longden I, Bleasby A (2000) EMBOSS: the European Molecular Biology
1011 Open Software Suite. *Trends Genet* 16: 276-277.
1012 48. Shahid S, Axtell MJ (2014) Identification and annotation of small RNA genes using
1013 ShortStack. *Methods* 67: 20-27.
1014 49. Lorenz R, Bernhart SH, Honer Zu Siederdissen C, Tafer H, Flamm C, et al. (2011)
1015 ViennaRNA Package 2.0. *Algorithms for molecular biology* : AMB 6: 26.
1016 50. Harris RS (2007) Improved pairwise alignment of genomic DNA. University Park,
1017 PA: The Pennsylvania State University.
1018 51. Bradley RK, Roberts A, Smoot M, Juvekar S, Do J, et al. (2009) Fast statistical
1019 alignment. *PLoS computational biology* 5: e1000392.
1020 52. Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-
1021 analysis of large phylogenies. *Bioinformatics* 30: 1312-1313.
1022 53. Zhu C, Byrd RH, Lu P, Nocedal J (1997) L-BFGS-B: Fortran subroutines for large-
1023 scale bound-constrained optimization. *ACM Transactions on Mathematical*
1024 *Software* 23: 550-560.

1026 **Figure Legends**

1027

1028 **Figure 1:** Summary of small RNA sequencing data and analysis of sequencing depth
1029 sufficiency. **(A)** The recovery rate of known *D. melanogaster* miRNAs using sets of 100 million
1030 (M) total reads sampled randomly from *D. melanogaster* head, mixed embryo, male-body,
1031 female-body public libraries. These libraries mimic those we sought to create for the 11 other
1032 *Drosophila* genomes. Bars represent the fraction of conserved or newly-evolved *D.*
1033 *melanogaster* miRNAs recovered at various miRNA and miRNA* minimum read thresholds, and
1034 error bars represent the standard error of the recovery rate across 100 different samples of
1035 100M reads. **(B)** Saturation curve of miRNA (mature and star) strand recovery at varying
1036 minimum read depth cutoffs. Based upon these results, we sequenced 100M reads per species.
1037 **(C)** Actual read sequencing depth and library conditions profiled within this study, or **(D)**
1038 acquired from public repositories for 11 *Drosophila* species. Datasets for *D. melanogaster* are
1039 not shown.

1040

1041 **Figure 2:** Summary of all known and novel miRNAs recovered within 12 *Drosophila* genomes.
1042 **(A)** Counts of known and novel miRNAs recovered or identified, respectively, at our two highest
1043 confidence classes – “confident” and “candidate-rescued.” MiRNAs from a third confidence
1044 class - “candidate” miRNAs - are shown in *Supp. Fig S5*. **(B)** The proportion of miRNAs
1045 recovered within three classes defined by biogenesis pathway, and testes-restricted, clustered
1046 status. Pie charts are provided for all novel or known annotations, and for the merged collection.
1047 **(C)** The distribution of alignment sizes upon assignment of all miRNAs into 1031 alignments.
1048 Paralogous miRNAs were assigned to single species alignment. The majority of miRNAs
1049 identified are singletons (species-specific) or doubletons (clade-specific). **(D)** Cumulative
1050 distribution function of alignment expression. Alignments are segregated based upon age and
1051 miRNA class. Empirical CDFs are plotted using the maximum expression values computed
1052 across all constitutive members of each alignment. RPMM = Reads per Million mapped miRNA
1053 reads.

1054

1055 **Figure 3:** Examples of novel miRNAs identified in this study. **(A)** *dvi_264* and *der_50* are
1056 examples of confident, canonical miRNAs. Small RNAs were cloned from both arms of their pre-
1057 miRNA, and an alignment of these sequences reveal patterns of precise 5' end cleavage and 1-
1058 2nt 3' blunt end overhang in the hairpin structure. These are all signatures of Drosha- and Dicer-
1059 mediated cleavage. **(B)** An example of orthologous confident (*dsi_14614*) and a candidate-

1060 rescued (*dse_989*) mirtrons within the two sister species *D. simulans* and *D. sechellia*. Total
1061 cloned reads for *dse_939* fell below our threshold for “confident” classification and was
1062 effectively placed into the “candidate” confidence group. We “rescued” this locus due to its
1063 synteny with *dsi_14614*, a confident mirtron. **(C)** Two novel, conserved, canonical miRNAs
1064 identified within the host gene *cwo*. *dme_164* sequence alignment indicates higher conservation
1065 of the miR and miR* arms in comparison to the loop and flanking basal stem regions, a classic
1066 signature of pre-miRNA sequence conservation. However, it possesses conserved A-rich
1067 regions that flank the duplex, which are seemingly incompatible with a lower-stem needed for
1068 Drosha processing. *dme_373* exhibits a large terminal loop atypical of most canonical miRNAs.
1069

1070 **Figure 4:** Shifted processing and alternate biogenesis pathways of Drosophila miRNAs. **(A-B)**
1071 Consistency in 5' end processing for conserved miRNAs. **(A)** sRNA read alignment for *mir-79*,
1072 represented compactly in these bubble plots, show two miR sequences with unique 5' ends.
1073 These represent two seed-distinct iso-miRs, that are both produced in several *Drosophila*
1074 species. Position 0 represents the proportion of reads that begin with the base of the most
1075 abundant 5' arm sequence at either the 5' strand (miR*) and 3' strand (miR) for all 12
1076 *Drosophila* genomes. Proportions shown at positions less than or greater than 0 represent
1077 proportion of reads with shifted processing. For *mir-79*, two iso-miRs are produced in similar
1078 proportions. **(B)** Panels of bubble plots depicting the heterogeneity of 5' end processing for the
1079 miR sequence of other conserved miRNAs and mirtrons. Greater than 4 alternate iso-miR-193
1080 sequences in *D. melanogaster* were noted previously. This heterogeneity is preserved in the
1081 genomes of the other Drosophilids, and a conserved, dominant iso-miRs is not apparent. We
1082 identified clade-specific iso-miRs for one canonical miRNA (*mir-969*) and one mirtron (*mir-*
1083 *1014*). Specifically, two unique iso-miR-969 sequences are each preferentially abundant in the
1084 *Sophophora*-group and *Drosophila*-group species, respectively, and for *mir-1014*, the
1085 melanogaster-group species produces one iso-miR-1014 sequence that is distinct from the
1086 dominant iso-miR of other Sophophorans. **(C)** *mir-iab-4/8* and *mir-307/mir-307-as* are the only
1087 two reasonably conserved miRNAs with sense and antisense transcription and processing
1088 based upon our genus-wide data. **(D)** Dozens of other recently-evolved antisense miRNAs were
1089 identified in our study however, such as *dmo_105/dmo_309* that arose adjacent to the
1090 conserved *mir-994/318* cluster. **(E)** Example of an atypical splicing derived, 3' tailed mirtron,
1091 identified from our data, *dse_288*. It requires trimming of ~12 nts from the debranched 3' splice
1092 site to generate the pre-miRNA hairpin.
1093

1094 **Figure 5:** Testes-restricted, Recently-evolved, Clustered (TRC) canonical miRNAs in
1095 *Drosophila*. **(A)** An example of a novel TRC miRNA cluster (*dan_86* → *dan_373*) in *D.
1096 ananassae*. The majority of miRNAs show high expression in the male-body libraries. **(B)** An
1097 example of a TRC miRNA cluster (*dps_3416* → *dps-mir-2536*) in the obscura-group species.
1098 The *D. pseudoobscura* cluster contains 36 miRNAs while its sister species, *D. persimilis*,
1099 contains 26 miRNAs. MiRNAs within the 3' end region of these orthologous clusters (orange
1100 highlight) have preserved their order while miRNAs within the 5' region show high gene
1101 duplication. Colored circles and numbers represent miRNAs of the same family. **(C)** Expression
1102 heatmap for all *D. pseudoobscura* copies reveals a predominant testes-restricted profile. **(D)**
1103 Comparison of expression difference between TRC and solo canonical miRNAs present in *D.
1104 virilis* alone or within the virilis/mojavensis clade alone. TRC miRNAs of the virilis-subgroup
1105 show significantly higher expression than their age-matched solo canonical cohorts (*Mann-
1106 Whitney Test* $p < 10^{-8}$). All Drosophilid subclades have their own distinct TRC loci, and details of
1107 all the novel TRC loci cloned in this study are provided in Supplementary Figures S13-S17.

1108

1109 **Figure 6.** Estimation of miRNA birth and death rates in *Drosophila*. **(A)** A probabilistic,
1110 phylogenetic graphical model for estimating rates of gene birth and death. The model takes
1111 binary data representing miRNA presence or absence at the leaves of the tree and uses
1112 maximum likelihood and numerical optimization methods to estimate model parameter (μ, λ)
1113 values. Branch lengths (τ) are fixed. Maximum likelihood parameters are then used to
1114 reconstruct node-wise miRNA counts and edge-wise birth and death events. **(B-D)** Summary of
1115 estimated ancestral miRNA content and edge-wise birth and death events for three classes of
1116 miRNAs. MiRNA classes are canonical miRNAs **(B)**, Testes-restricted, canonical miRNAs
1117 (TRC) **(C)**, and mirtrons **(D)**. Estimates of edge wise birth and death events are shown in green
1118 and red, respectively. Net emergence rate (i.e. total birth - death events / Myr) are shown in
1119 each class for the melanogaster-group, obscura-group, and virilis-subgroup species. **(E)** Net
1120 miRNA gain rate for three clades- melanogaster-group, obscura-group, and virilis subgroup –
1121 are shown. Note that mirtrons and TRC miRNAs exhibit much higher rates of flux than do
1122 canonical non-testis-restricted miRNAs.

1123

1124 **Figure 7.** Multiple distinct cis-molecular signatures associated with miRNA flux. **(A-D)** Duplex
1125 alterations that affect miRNA processing. **(A)** *mir-4984* hairpin is similar across 5 related
1126 melanogaster subgroup species and its mature (green) arm is identical in all these species.
1127 However, small RNA sequencing indicates substantial accumulation only in *Dmel*, very modest

1128 in *Dsim/Dyak*, and not in *Dyak/Dere*. **(B)** Experimental tests of *UAS-DsRed-mir-4984* expression
1129 constructs transfected into S2 cells shows that only the *Dmel* construct was effectively
1130 processed into miRNAs. DsRed expression confirms that all constructs were expressed. **(C)**
1131 *dps_41/dpe_2484* ortholog pair, with only a few duplex divergences, exhibits divergent
1132 expression between very closely related species. **(D)** Experimental tests in S2 cells confirm
1133 differential biogenesis of these miRNAs. **(E)** Transcriptome comparison of miRNAs differentially
1134 expressed between sister species *Dpse* and *Dper*. In general, significantly more duplex
1135 divergent miRNAs are differentially-expressed miRNAs for both newly-evolved and conserved
1136 miRNAs. **(F)** Adaptive evolution of seed regions of testis-restricted, clustered (TRC) miRNAs.
1137 Shown are examples of 1-to-1 orthologs of TRC miRNAs between *Dpse* and *Dper*, including
1138 available *Dpse* population data. Highlighted are examples of seed divergence between
1139 expressed TRC miRNA orthologs between these closely related species, indicating adaptive
1140 evolutionary behavior. **(G-J)** Impact of terminal uridylation system on evolutionary suppression
1141 of mirtrons and behavior of canonical miRNAs. **(G-H)** Compared to canonical miRNAs **(G)**,
1142 mirtrons **(H)** in every Drosophilid species acquire extraordinarily high rates of terminal
1143 untemplated uridylation (purple) on the 3' ends of their 3p species, compared to any other
1144 nucleotide modifications. **(I-J)** 3' uridylation of canonical miRNAs is sensitive to terminal hairpin
1145 nucleotide. In these graphs, miRNA loci are divided by biogenesis type (canonical vs. splicing-
1146 derived), by terminal nucleotide (3'-G vs. 3'-A/U/C, i.e. "3'-H"), and by evolutionary age. Analysis
1147 of deeply conserved miRNA loci **(I)** and recently-evolved loci **(J)** shows that canonical miRNA
1148 hairpins that end in G acquire higher levels of 3' uridylation than do other canonical miRNA
1149 hairpins.

Mohammed and Flynt et al

Deep experimental profiling of microRNA diversity, deployment, and evolution across the *Drosophila* genus

Supplementary Figures and Tables

Supplementary Table S1: Small RNA libraries for 11 *Drosophila* species created for and analyzed in this study.

Supplementary Table S2: Small RNA libraries for *Drosophila* species acquired from public repositories and analyzed in this study. (A) Libraries for *D. melanogaster*. (B) Libraries for *D. simulans*, *D. yakuba*, *D. erecta*, *D. pseudoobscura*, and *D. virilis*.

Supplementary Table S3: miRBase Drosophilid loci that were demoted from miRNA status.

Supplementary Table S4: Master list of all known and novel miRNAs in *Drosophila*.

Supplementary Table S5: List of all sense and antisense miRNA pairs identified and analyzed in this study.

Supplementary Table S6: 3' Untemplated nucleotide addition counts to canonical miRNA-3p and mirtron-3p species.

Supplementary Table S7: miRNA sRNA read, sequence, and structure features for Random Forest and logistic regression classifiers.

Supplementary Figures Legend

Supplementary Figure S1: Conserved and newly-evolved *D. melanogaster* miRNAs recovered at varied read depth thresholds using *in silico* simulated libraries. These libraries are composed of randomly sampled reads across all *D. melanogaster* sRNA-seq male-body, female-body, head, and mixed embryo libraries used within this study. miRNA recovery rates are computed per read-depth sample at various miR or miR* read thresholds. Error bars depict the standard error of the recovery rate across 100 simulations.

Supplementary Figure S2: Read length distribution for all small RNA libraries sequenced in this study. We extended our previous broad and deep analysis of *D. melanogaster* by sampling 11 additional *Drosophila* species as listed to the right, by analyzing mixed embryos, adult heads, male bodies and female bodies; a testis library was also generated for *D. simulans*. A subset of libraries were sequenced in replicates, especially ones where the expected dominant miRNA-sized peak (21-22 nt) peak was not initially observed. A piRNA peak is seen in most of the body libraries. Due to the technical difficulty in culturing *D. grimshawi*, it was only feasible to generate two libraries for male and female bodies.

Supplementary Figure S3: Detailed flow-chart of miRNA and mirtron identification pipeline and scoring criteria.

Supplementary Figure S4: Drosophilid miRbase loci demoted for lack of compelling small RNA evidence. Examples of 47 demoted miRBase (v21) miRNAs and mirtrons. 37 of these represent piRNAs within *D. pseudoobscura* and *D. virilis* previously classified as miRNAs.

Supplementary Figure S5: Total miRNA and mirtron annotation count within 12 *Drosophila* species. Annotations are further subdivided within (1) three confidence categories- “confident”, “candidate-rescued, and candidate”, and (2) between known and novel annotations. Note that “candidate” annotations were not utilized for analyses of miRNA flux in this study.

Supplementary Figure S6: Distribution of miRNAs for three classes of miRNAs within each *Drosophila* species. These classes are defined by biogenesis pathway and canonical miRNAs are further divided by their testes-restricted expression. Only “confident” and “candidate-rescued” loci are included; loci that are considered “candidate” only and lack further rationale to be rescued based on a confidently processed miRNA ortholog are not included in these pie charts.

Supplementary Figure S7: Read alignments for mir-10404, a conserved non-canonical miRNA generated from the ITS1 spacer in ribosomal RNA, across the Drosophilid phylogeny.

Supplementary Figure S8: Other well-conserved miRNAs identified within this study. Alignment and representative hairpin structure for each conserved miRNA. Included is dme_474 which is not well-conserved but is one of the Drosha-cleaved hairpins within the pasha 5' UTR. Note that unlike most other conserved miRNAs, these loci generally accumulate modest amounts of small RNAs and/or have atypical structural features. This might reflect that their processing is atypical and/or regulated, or that their conservation reflects a role other than, or in addition to miRNA-type function. For example, besides the pasha 5' UTR hairpins, two of these loci are located in CDS or 3'UTR, and thus cleavage could mediate host mRNA downregulation.

Supplementary Figure S9: Alignment and small RNA read details for novel conserved miRNAs annotated in this study.

Supplementary Figure S10: Evolutionary patterns of 5' end cleavage precision for all conserved *D. melanogaster* miRNAs.

Supplementary Figure S11: Supplementary Figure S11: Additional examples of novel sense/antisense miRNA pairs. (A) Example of novel sense/antisense miRNA pair from *D. willistoni* (dwi_62/dwi-98). (B) Example of novel sense/antisense miRNA pair from *D. ananassae* (dan_100/dan_244).

Supplementary Figure S12: Example of a novel, atypical 3' tailed mirtron. A 3' tailed mirtron from *D. erecta*, Der_70. This locus produces a dominant 3p miRNA, which is trimmed by ~13 from the splice acceptor site. Note that there is 3' untemplated uridylation (purple) associated with a subset of Der_70-3p reads. While Der_70-5p reads are modest, there are also apparent partially diced products that include the terminal loop that phase precisely with the Der_70-5p species, supporting this as a Dicer position. Bottom alignment shows that Der_70 tailed mirtron resides in the conserved gene CYLD, and is associated with a non-canonical splice “GC” donor in the five melanogaster group species, which is instead a conventional “GT” splice donor in most other insects. Note *D. willistoni* seems to have lost both splice sites.

Supplementary Figure S13: Annotation of novel Testes-restricted, recently-evolved, clustered (TRC) miRNA clusters identified within *D. ananassae*.

Supplementary Figure S14: Annotation of a novel Testes-restricted, recently-evolved, clustered (TRC) miRNA cluster identified in *D. willistoni*. Tissue code indicates the miRNAs are all highest expressed in male body libraries.

Supplementary Figure S15: Annotation of novel Testes-restricted, recently-evolved, clustered (TRC) miRNA clusters identified in obscura group species, *D. pseudoobscura* and *D. persimilis*. (A) Genomic organization and small RNA read density of orthologous *dps_3416* → *dps-mir-2536* TRC clusters in *D. pseudoobscura* and *D. persimilis*. (B) Genomic organization and small RNA read density of orthologous *dps-mir-2510* → *dps_23* TRC clusters in *D. pseudoobscura* and *D. persimilis*. Tissue code indicates the miRNAs are all highest expressed in male body/testis libraries. Note that there are also additional copies of some of these TRC loci located outside of these clusters.

Supplementary Figure S16: Annotation of novel Testes-restricted, recently-evolved, clustered (TRC) miRNA clusters identified in virilis clade species. (A) The *dvi_66* → *dvi_40* cluster. This cluster has a clear homolog in *mojavensis*. The small RNA mappings to their respective genomic loci are shown. (B) The *dvi_24637* → *dvi_197* cluster has clear homologs in *D. mojavensis*. The small RNA mappings to their respective genomic loci are shown. (C) Annotation of three additional novel Testes-restricted, recently-evolved, clustered (TRC) miRNA clusters identified in virilis clade species. The *dvi_43* → *dvi_207* cluster has two copies in *D. virilis* (i.e. roughly similar members can be found on scaffold_12723 and scaffold_12963). The other cluster is *D. mojavensis dmo_62* and *dmo_330*.

Supplementary Figure S17: Expression difference between *D. pseudoobscura*-specific or obscura-group-specific TRC and solo canonical miRNAs. Points reflect the maximum expression per locus assessed over all *D. pseudoobscura* libraries. P-value computed from two-tailed Wilcoxon Rank Sum Test.

Supplementary Figure S18: All possible phylogenetic reconstruction of ancestral miRNA presence and absence for 3 miRNA classes using a phylogenetic probabilistic graphical model with universal parameters of $\lambda = 0.292$ and $\mu = 0.694$. These parameters were computed by running the phylogenetic reconstruction algorithms on all mirtrons and miRNAs pooled together. These trees illustrate how the method's maximum likelihood reconstruction performs for all possible configurations of extant miRNAs presence and absence per alignment. Blue text indicates count of alignments with this particular configuration in each class. Summary of miRNA birth and death (Figure 6) are based upon these estimates of ancestral miRNA presence and absence.

Supplementary Figure S19: Individual phylogeny of extant and inferred ancestral miRNA presence and absence for all canonical miRNAs that are not in the testis-restricted clustered subclass.

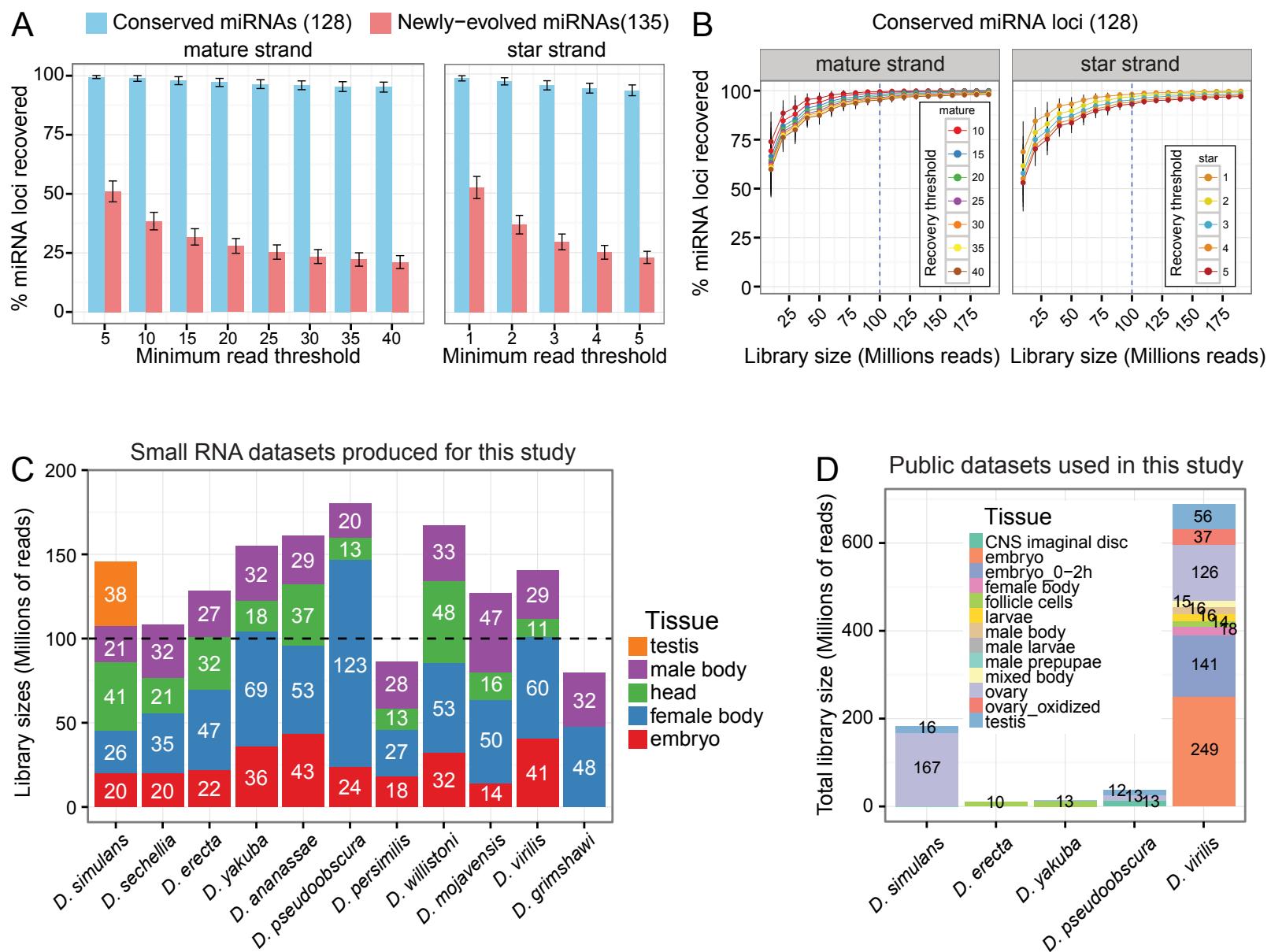
Supplementary Figure S20: Individual phylogeny of extant and inferred ancestral miRNA presence and absence for all mirtrons.

Supplementary Figure S21: Individual phylogeny of extant and inferred ancestral miRNA presence and absence for all testis-restricted clustered (TRC) canonical miRNAs.

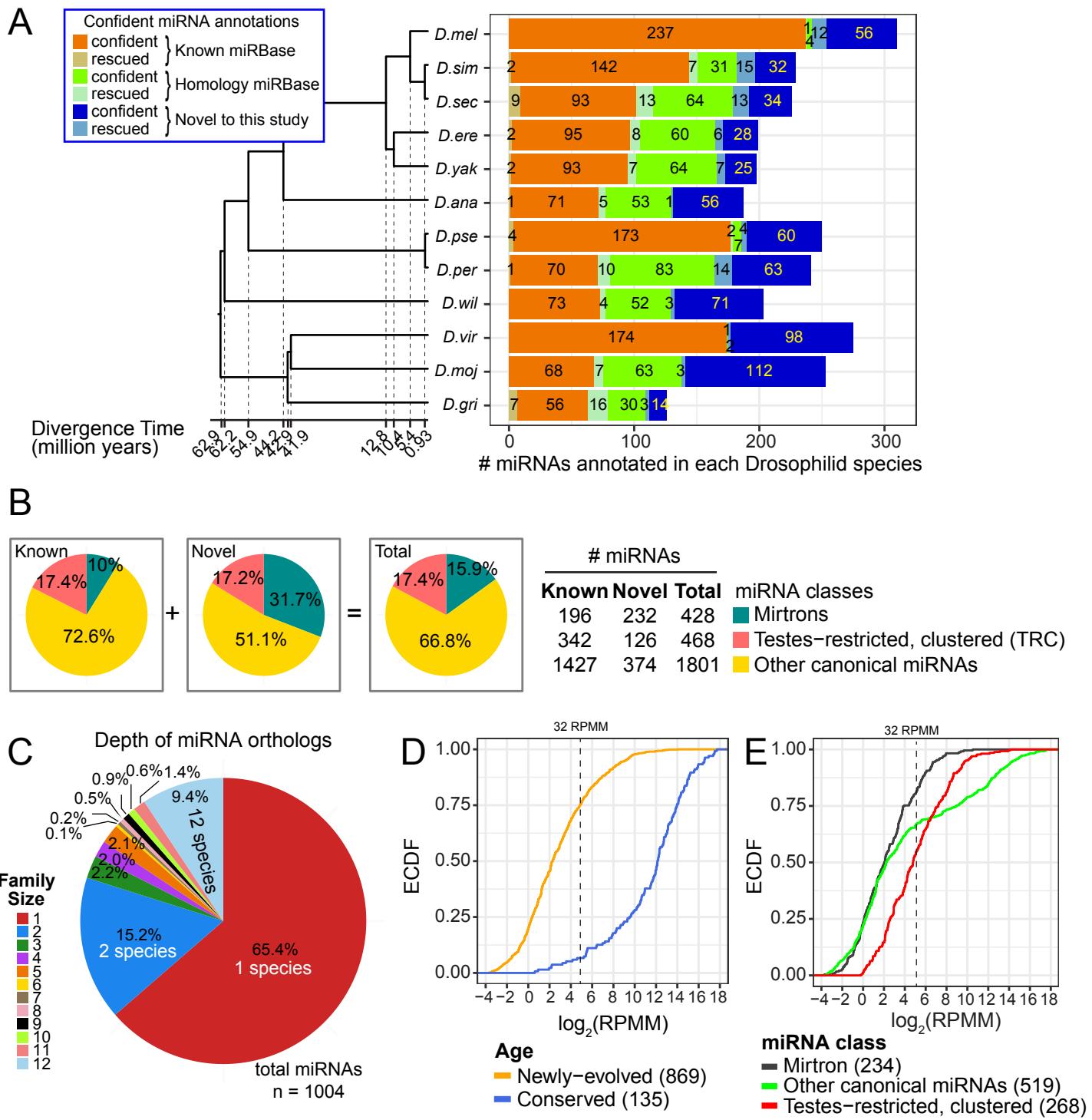
Supplementary Figure S22: Examples of mirtrons with atypical emergence and decay patterns. Expression profiles and mirtron alignments are shown per example to highlight the non-clade specificity of mirtron presence.

Supplementary Figure S23: Differential expression of miRNAs between obscura group species. Scatterplot depicting the correlation of miRNA expression of all *D. pseudoobscura* and *D. persimilis* ortholog pairs (RPMM = Reads Per Million Mapped MiRNA Reads). All miRNA alignments with orthologs in both species are shown. Points that lie on or near the diagonal represent similarly expressed ortholog pairs. Orthologs with > 6-fold log₁₀(RPMM) difference (denoted by the blue-dashed line and labeled points) are examples of significantly differentially expressed orthologs. Points are colored by miRNA age, and shapes represent miRNAs with or without miR:miR* duplex region substitutions (fraction of duplex sites with substitutions are labeled). Note that the mir-309 cluster (yellow) loci are expected to be expressed in the very early embryo, and given that the embryo development and timing were not controlled in library preparation, their differential accumulation may not be genuine. Amongst loci changed by >6-fold, dme-mir-2b-1 and dme-mir-310 are deeply conserved, but all others are specific to the obscura group species.

Supplementary Figure S24: 3' end untemplated nucleotide additions for canonical miRNAs and mirtrons in 12 Drosophila species. (A) Proportion of AG ending 3' arm miRNAs and mirtrons that contain reads within mono-A, C, G or U untemplated additions. Error bars represent the standard error of the mean. More mirtrons contain untemplated uridylation than comparable 3' end AG-ending canonical miRNAs. (B) Species-specific empirical cumulative distribution function of mono-uridylation for mirtrons and canonical miRNAs with 3' end 'G' nucleotide or non-'G' nucleotides (i.e. IUPAC ambiguity code 'H'). P-value computed from two-tailed Wilcoxon Rank Sum Test between canonical 3'-end 'H' miRNAs and mirtrons. Significant differences in mono-uridylation distributions between these two classes are noted in blue text. P-values from comparisons between canonical 3' end 'H' miRNAs and 3' end 'G' miRNAs are all non-significant and not shown.



Mohammed et al
Figure 1



Mohammed et al
Figure 2

A

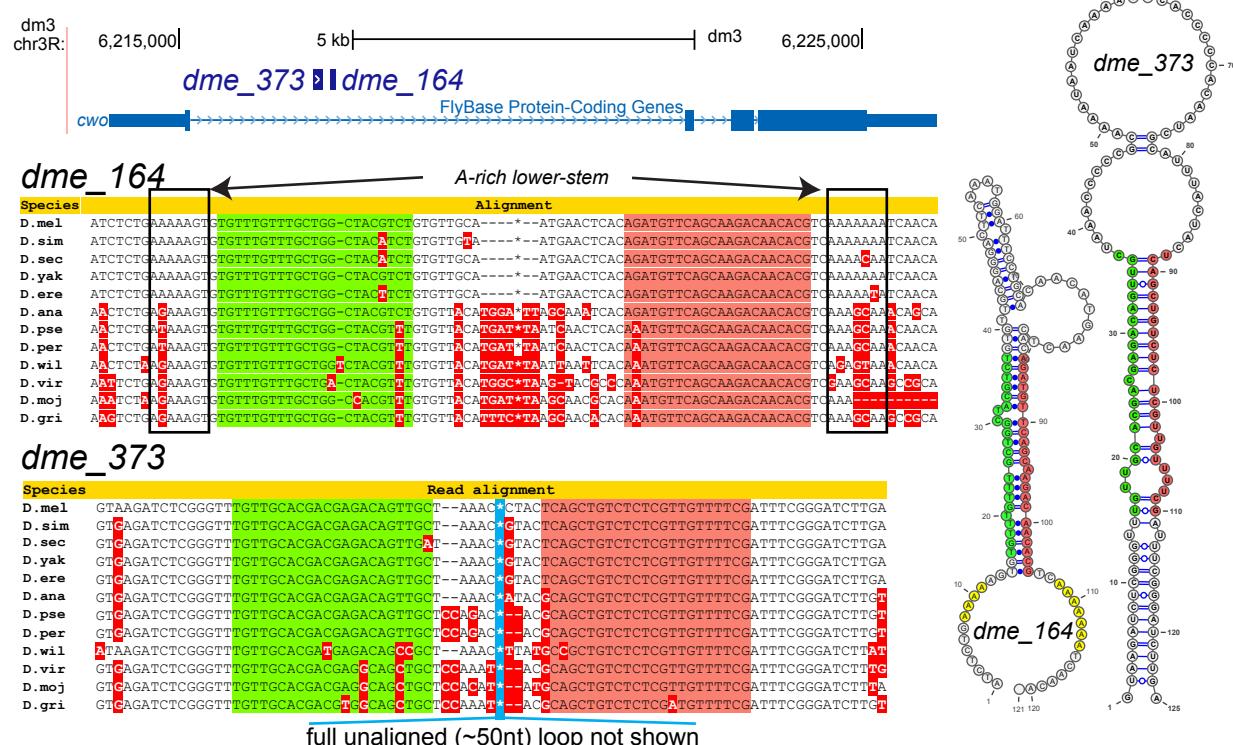
Novel Confident, species-specific, solo canonical miRNAs

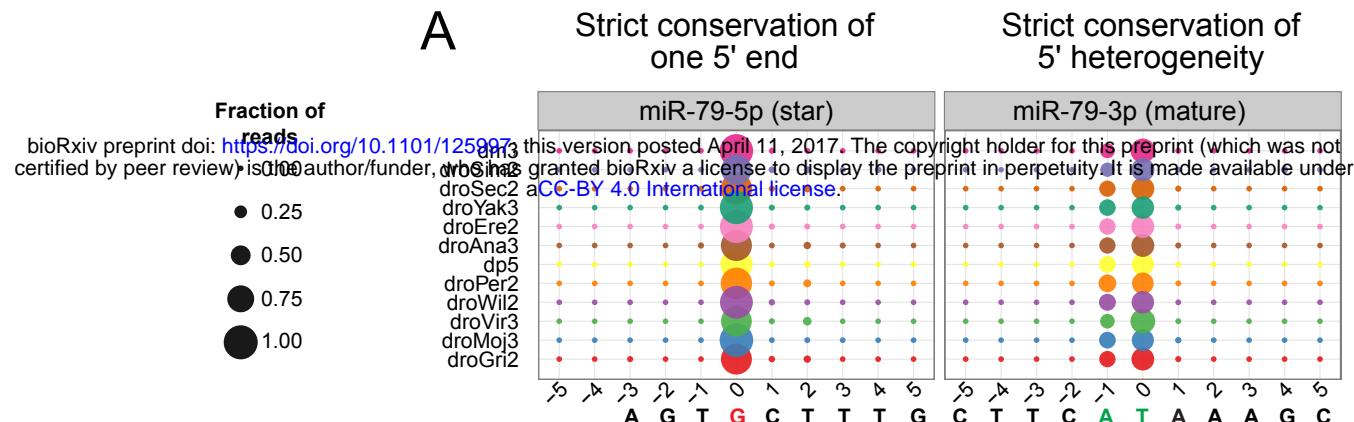
	Size		Total		Total	
			Size	Hits	Norm	RPM
der_50	20	1	41804.00	177.17		
der_50	22	1	38728.00	188.54		
der_50	22	1	22875.00	74.33		
der_50	21	1	11824.00	67.67		
der_50	23	1	6451.00	38.06		
der_50	23	1	8.00	0.39		
der_50	22	1	7.00	0.42		
			Total		145487.50	345.43
dsi_14614	21	1	567	6.45		
dsi_14614	22	1	344	3.91		
dsi_14614	20	1	260	3.06		
dsi_14614	20	1	122	1.39		
dsi_14614	21	1	85	1.00		
dsi_14614	22	1	14	0.17		
dsi_14614	23	1	4	0.09		
dsi_14614	21	1	2	0.09		
			Total		1747	18.7

B

Novel mirtron and “rescued” ortholog

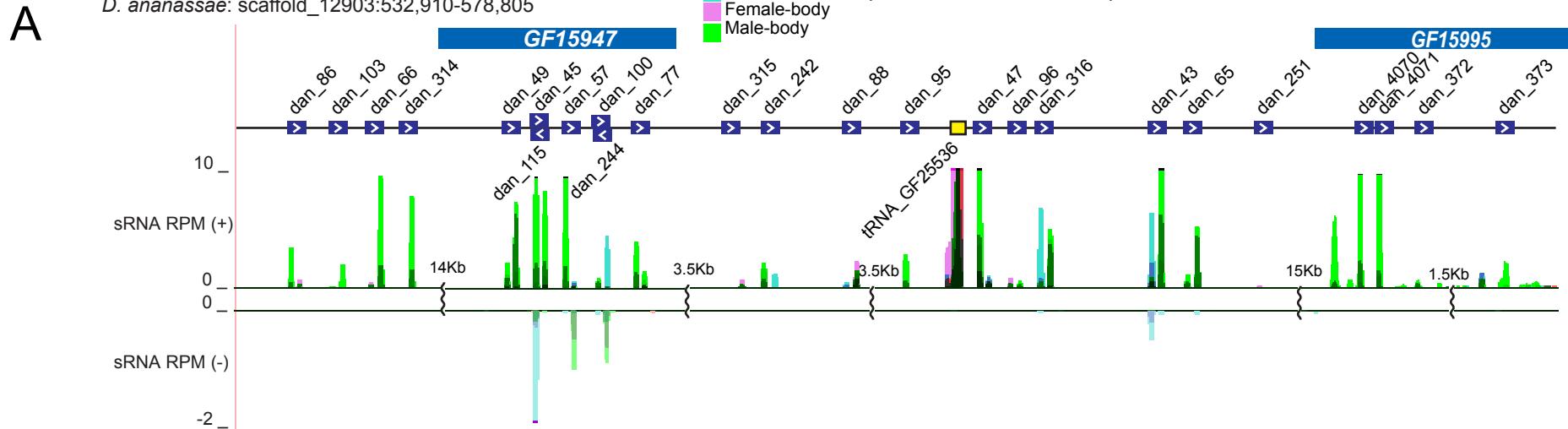
	Read alignment			CG2199		
	Size	Hits	Reads	Total	Norm	RPM
Novel confident	23	1	9.00	9.00	0.08	
Novel confident	22	1	4.00	4.00	0.11	
Novel confident	21	1	3.00	3.00	0.06	
Novel confident	24	1	3.00	3.00	0.19	
Novel confident	23	1	3.00	3.00	0.04	
Novel confident	22	1	2.00	2.00	0.08	
Novel confident	20	1	1.00	1.00	0.12	
Novel confident	23	1	1.00	1.00	0.07	
Novel confident	24	1	1.00	1.00	0.03	
Novel confident	20	1	1.00	1.00	0.07	
Novel confident	23	1	1.00	1.00	0.03	
Novel confident	24	1	1.00	1.00	0.08	
Novel confident	21	1	1.00	1.00	0.05	
Novel confident	21	1	1.00	1.00	0.28	
Novel confident	26	1	1.00	1.00	0.03	
Novel confident	22	1	1.00	1.00	0.08	
Novel confident	24	1	1.00	1.00	0.05	
Novel confident	21	1	1.00	1.00	0.07	
Novel confident	26	1	1.00	1.00	0.03	
Novel confident	21	1	1.00	1.00	0.08	
Novel confident	23	1	1.00	1.00	0.03	
Novel confident	20	1	1.00	1.00	0.03	
Novel confident				Total	38.00	0.19
Novel candidate-rescued	22	1	5.00	5.00	0.16	
Novel candidate-rescued	21	1	3.00	3.00	0.20	
Novel candidate-rescued	23	1	2.00	2.00	0.13	
Novel candidate-rescued	25	1	2.00	2.00	0.13	
Novel candidate-rescued	23	1	2.00	2.00	0.10	
Novel candidate-rescued	23	1	1.00	1.00	0.05	
Novel candidate-rescued	23	1	1.00	1.00	0.06	
Novel candidate-rescued	25	1	1.00	1.00	0.05	
Novel candidate-rescued	22	1	1.00	1.00	0.05	
Novel candidate-rescued	24	1	1.00	1.00	0.07	
Novel candidate-rescued	22	1	1.00	1.00	0.07	
Novel candidate-rescued	23	1	1.00	1.00	0.07	
Novel candidate-rescued	24	1	1.00	1.00	0.07	
Novel candidate-rescued	23	1	1.00	1.00	0.07	
Novel candidate-rescued	24	1	1.00	1.00	0.07	
Novel candidate-rescued				Total	22.00	0.32

CNovel Conserved *Drosophila* miRNAsMohammed et al
Figure 3



dan_86 → dan_373 cluster

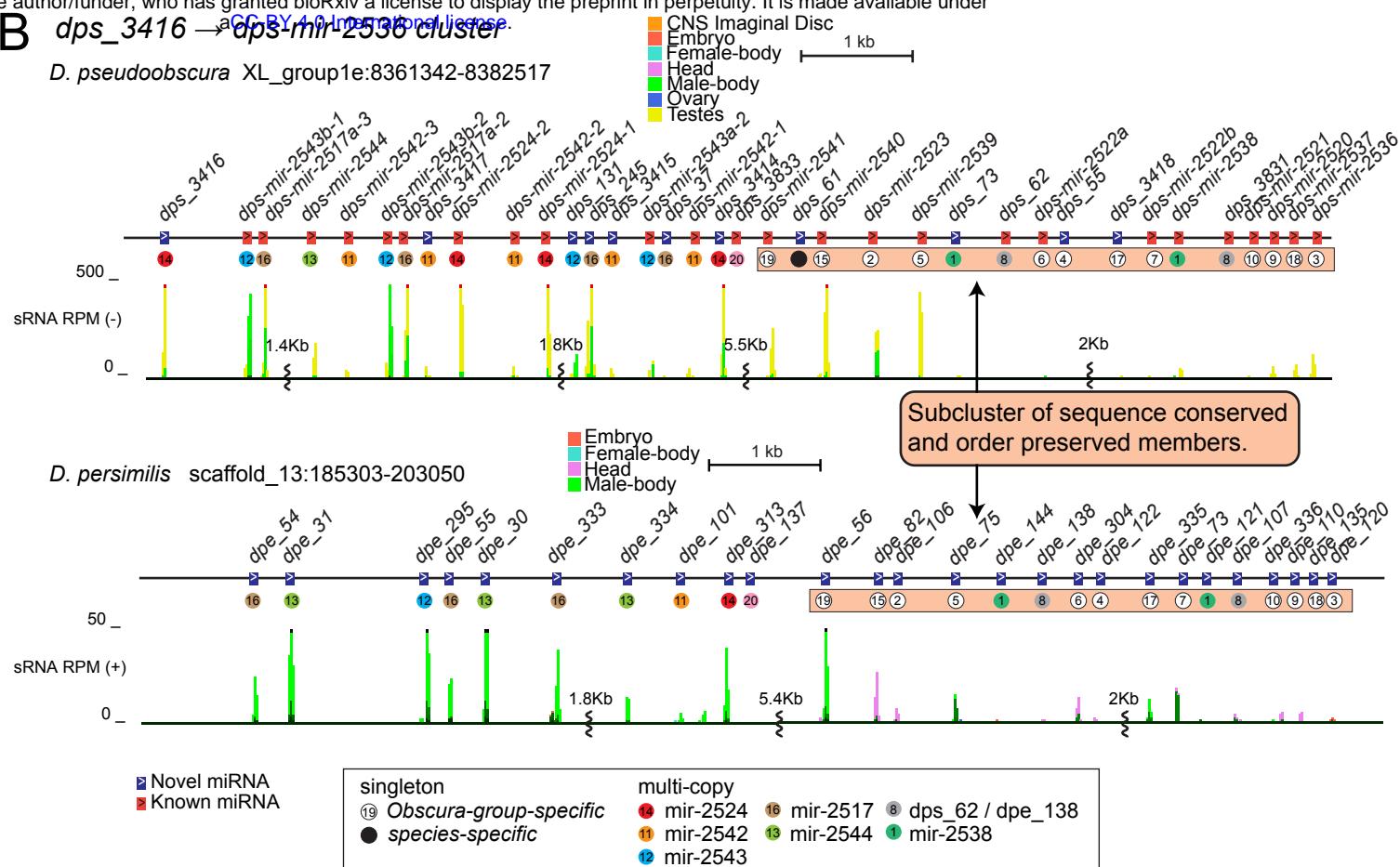
D. ananassae: scaffold_12903:532,910-578,805



bioRxiv preprint doi: <https://doi.org/10.1101/125997>; this version posted April 11, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

B *dps_3416 → dps-mir-2530 cluster*

D. pseudoobscura XL_group1e:8361342-8382517



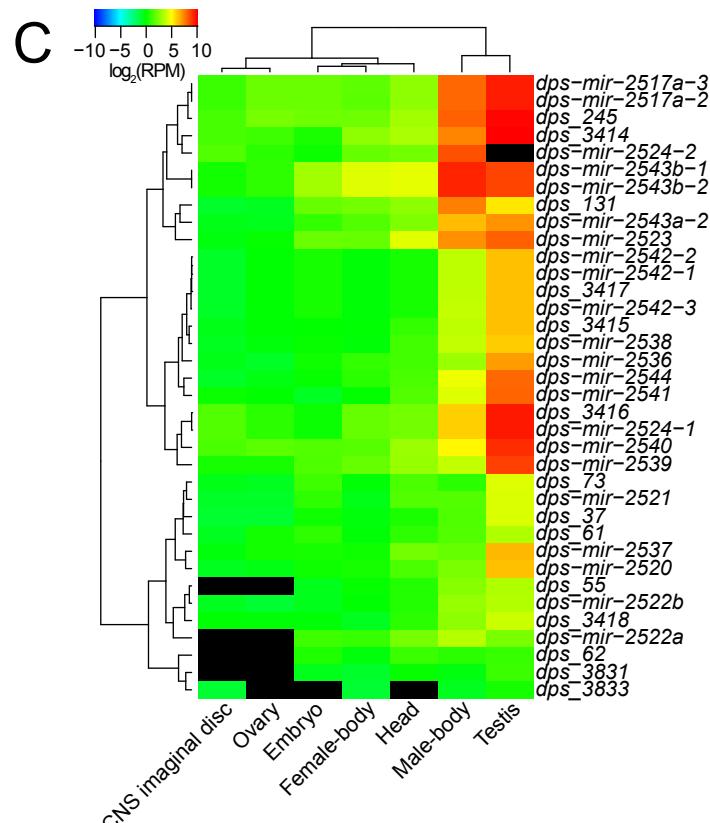
D. persimilis scaffold_13:185303-203050

Novel miRNA
Known miRNA

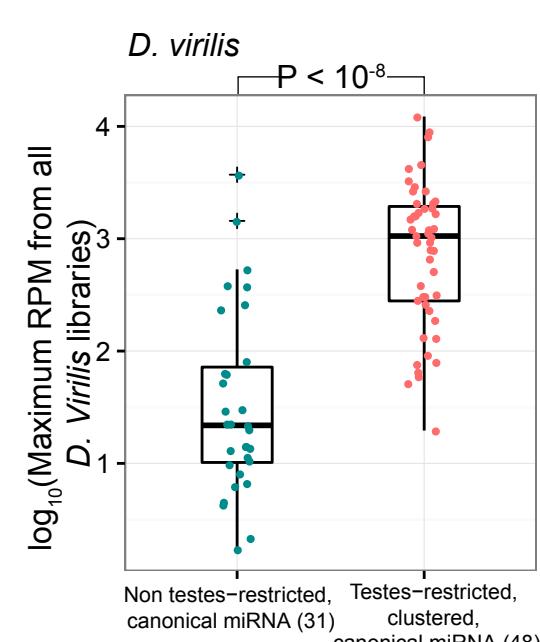
singleton
Obscura-group-specific
● species-specific

multi-copy
⑯ mir-2524 ⑯ mir-2517 ⑧ dps_62 / dpe_138
⑪ mir-2542 ⑬ mir-2544 ⑩ mir-2538
⑫ mir-2543

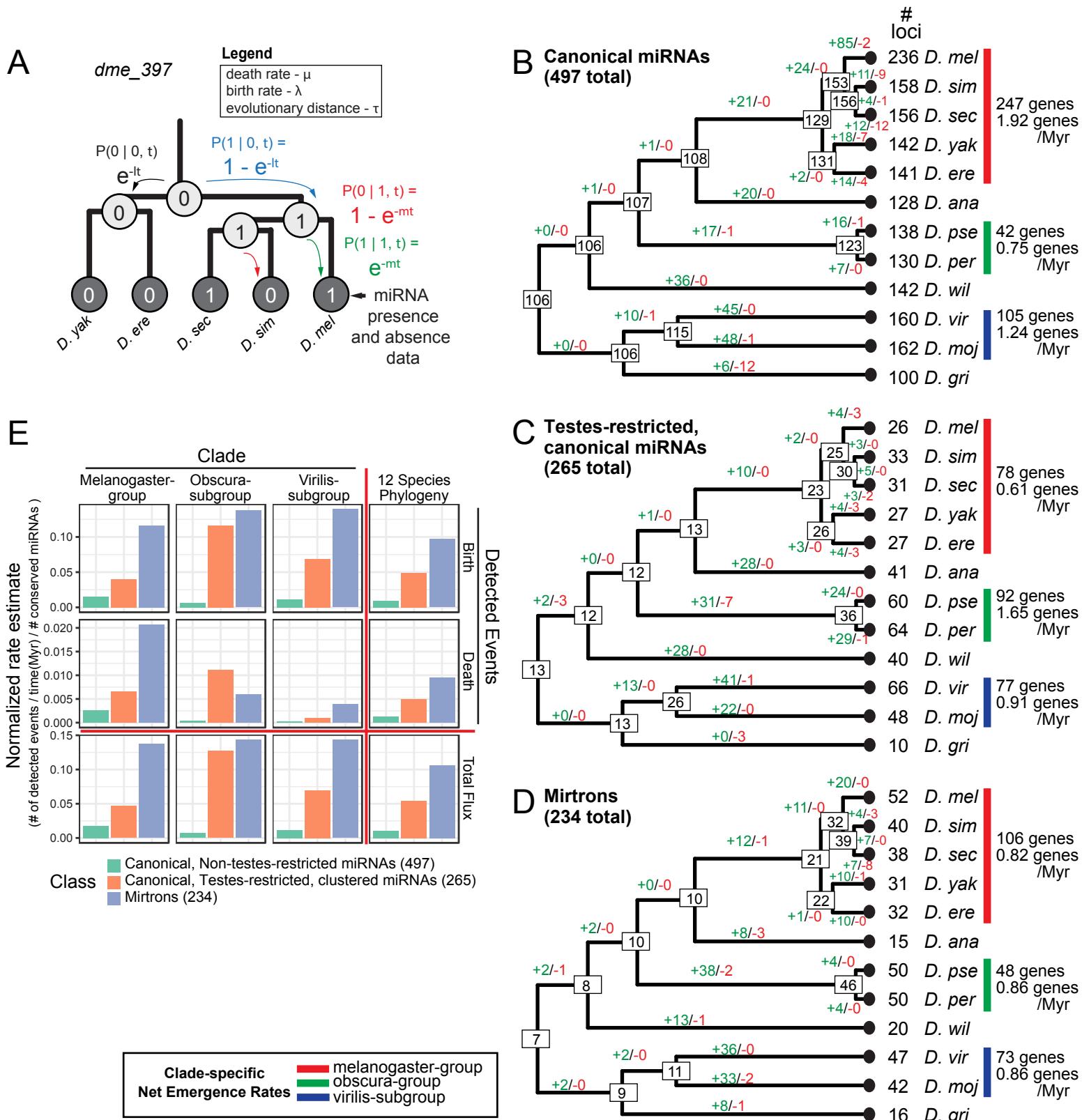
Subcluster of sequence conserved and order preserved members.



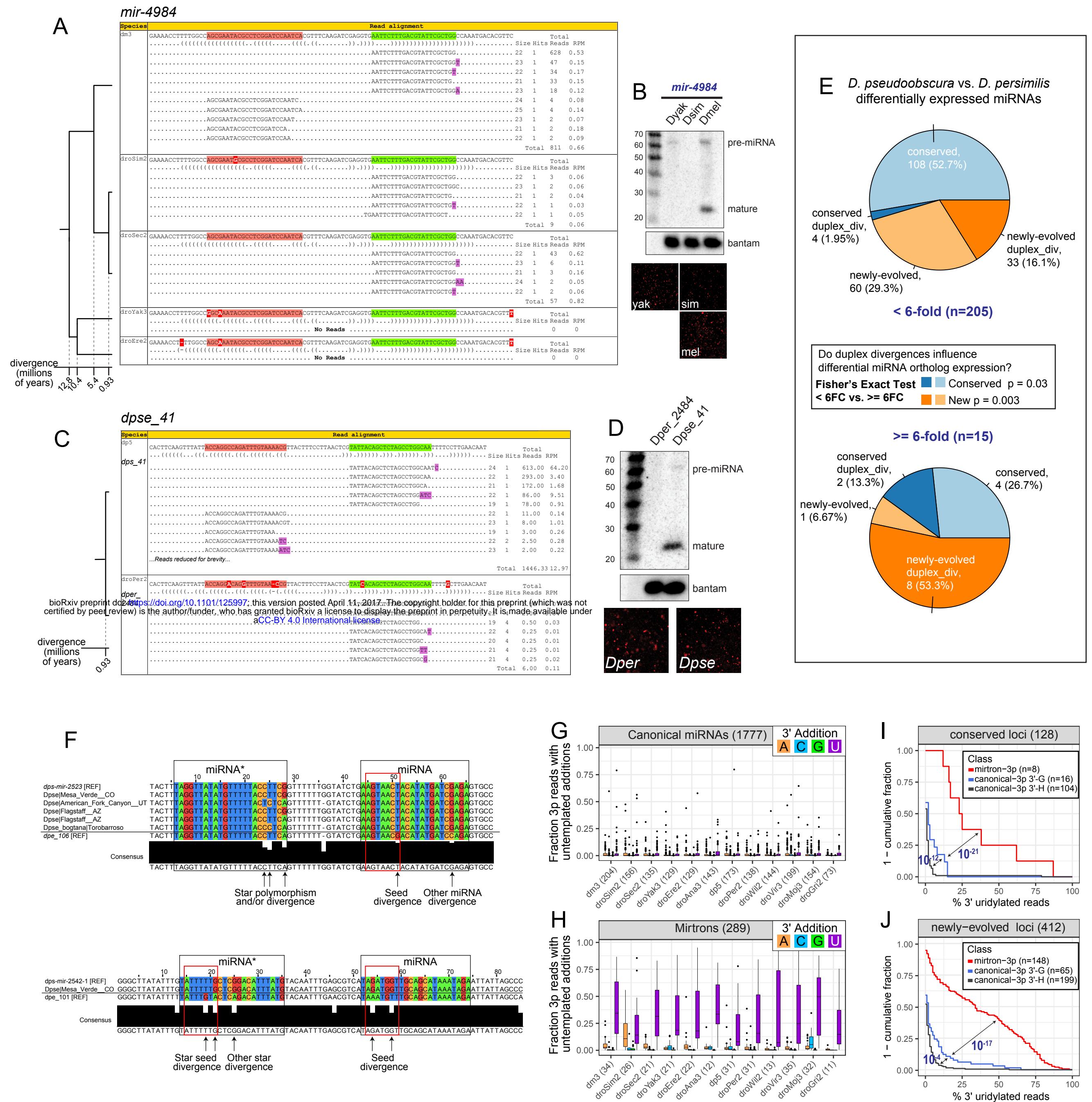
D



Mohammed et al
Figure 5

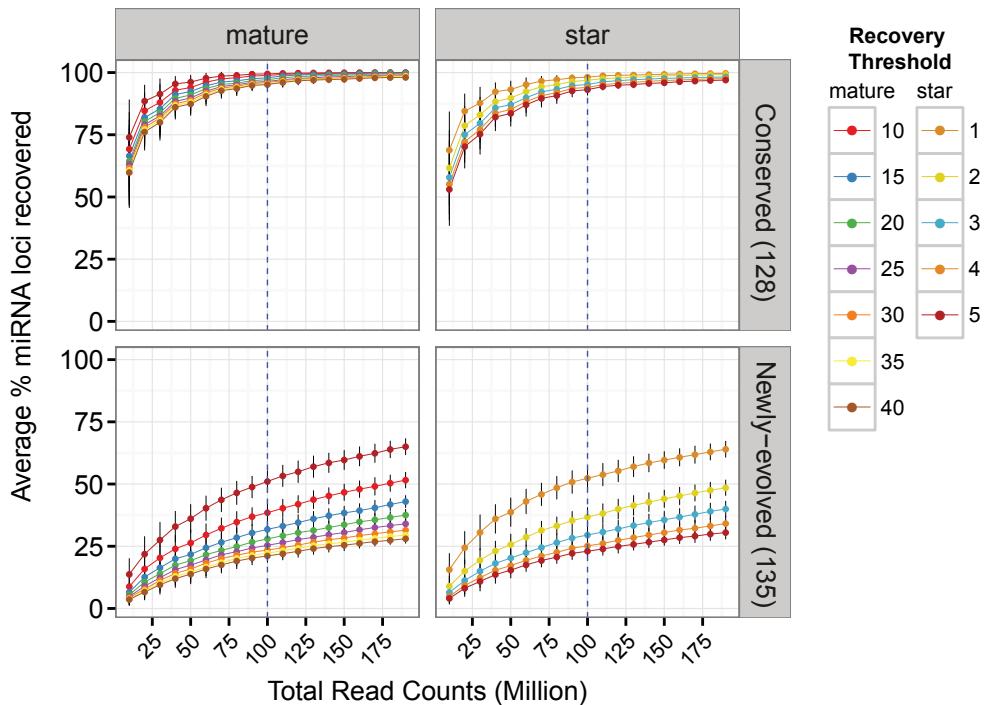


Mohammed et al
Figure 6

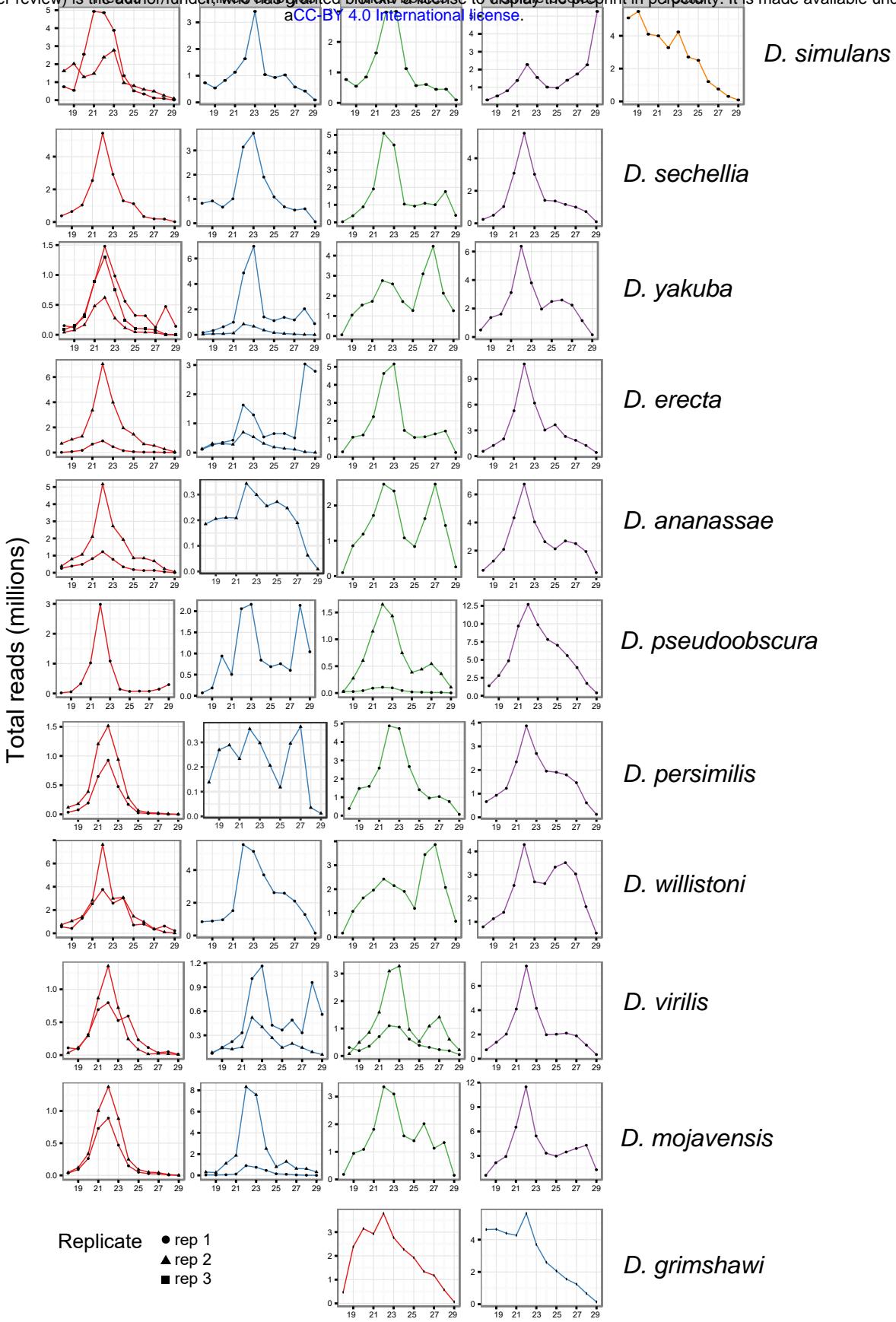


Mohammed et al

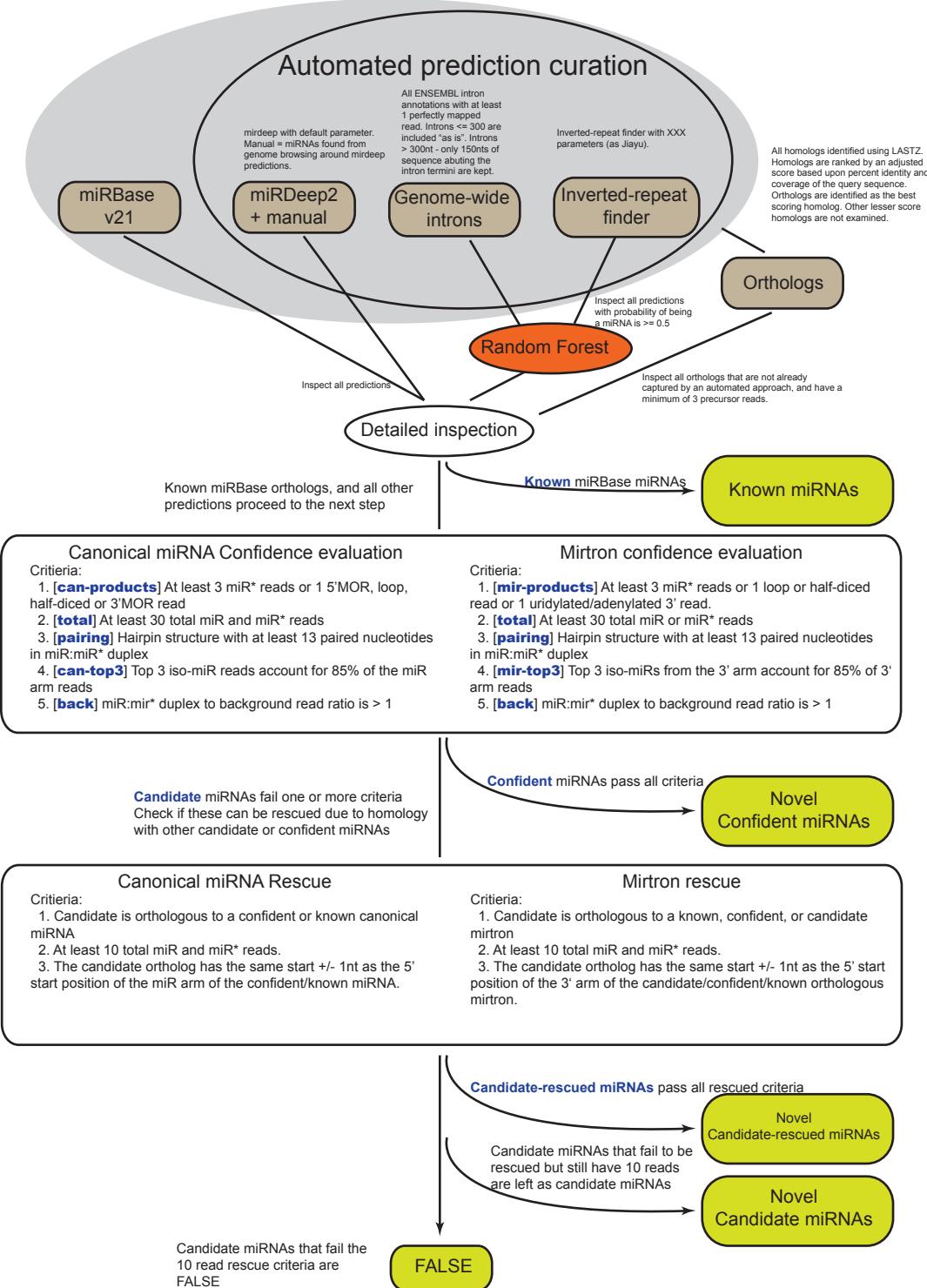
Figure 7



Supplementary Figure S1: Conserved and newly-evolved *D. melanogaster* miRNAs recovered at varied read depth thresholds using in silico simulated libraries. These libraries are composed of randomly sampled reads across all *D. melanogaster* sRNA-seq male-body, female-body, head, and mixed embryo libraries used within this study. miRNA recovery rates are computed per read-depth sample at various miR or miR* read thresholds. Error bars depict the standard error of the recovery rate across 100 simulations.

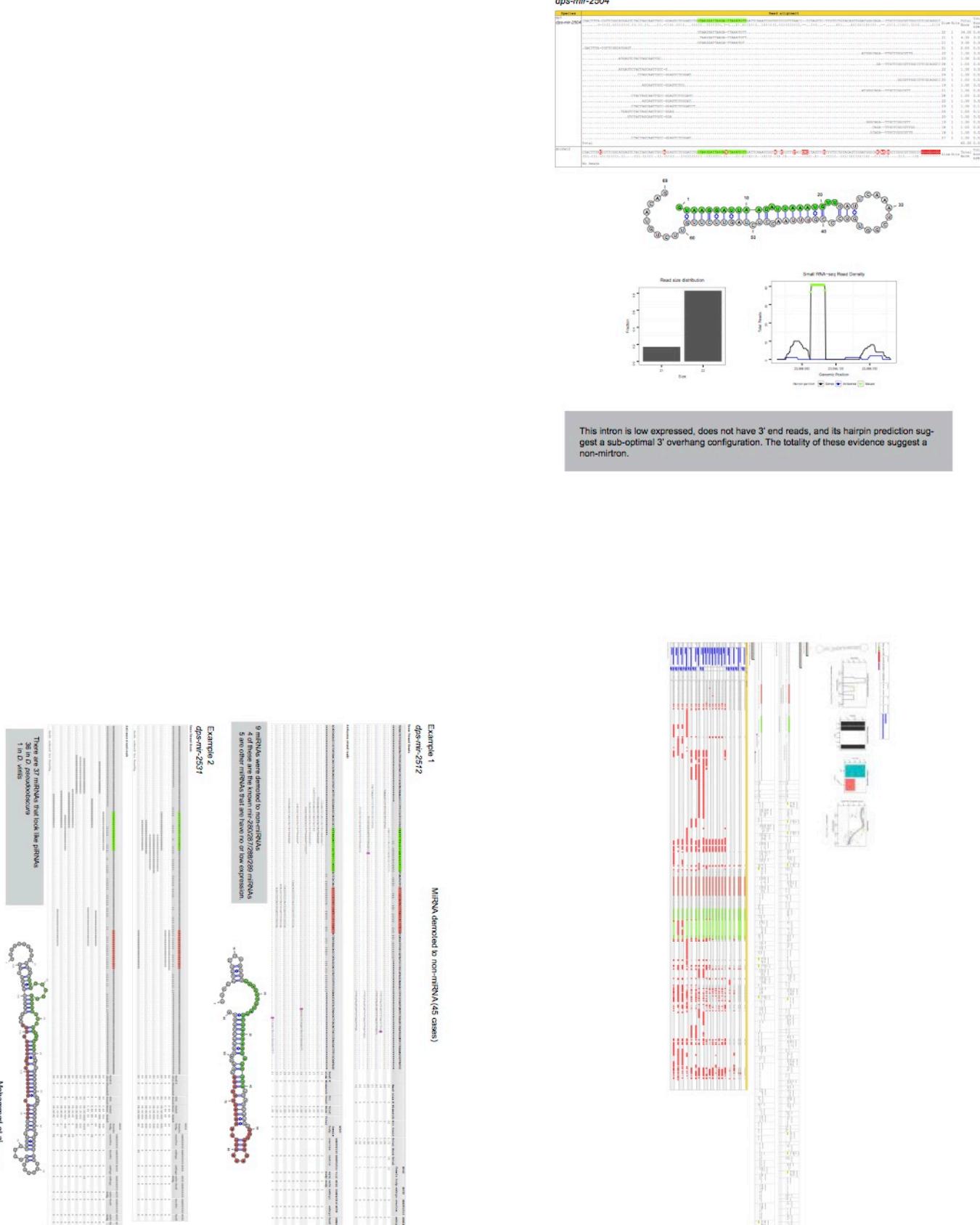


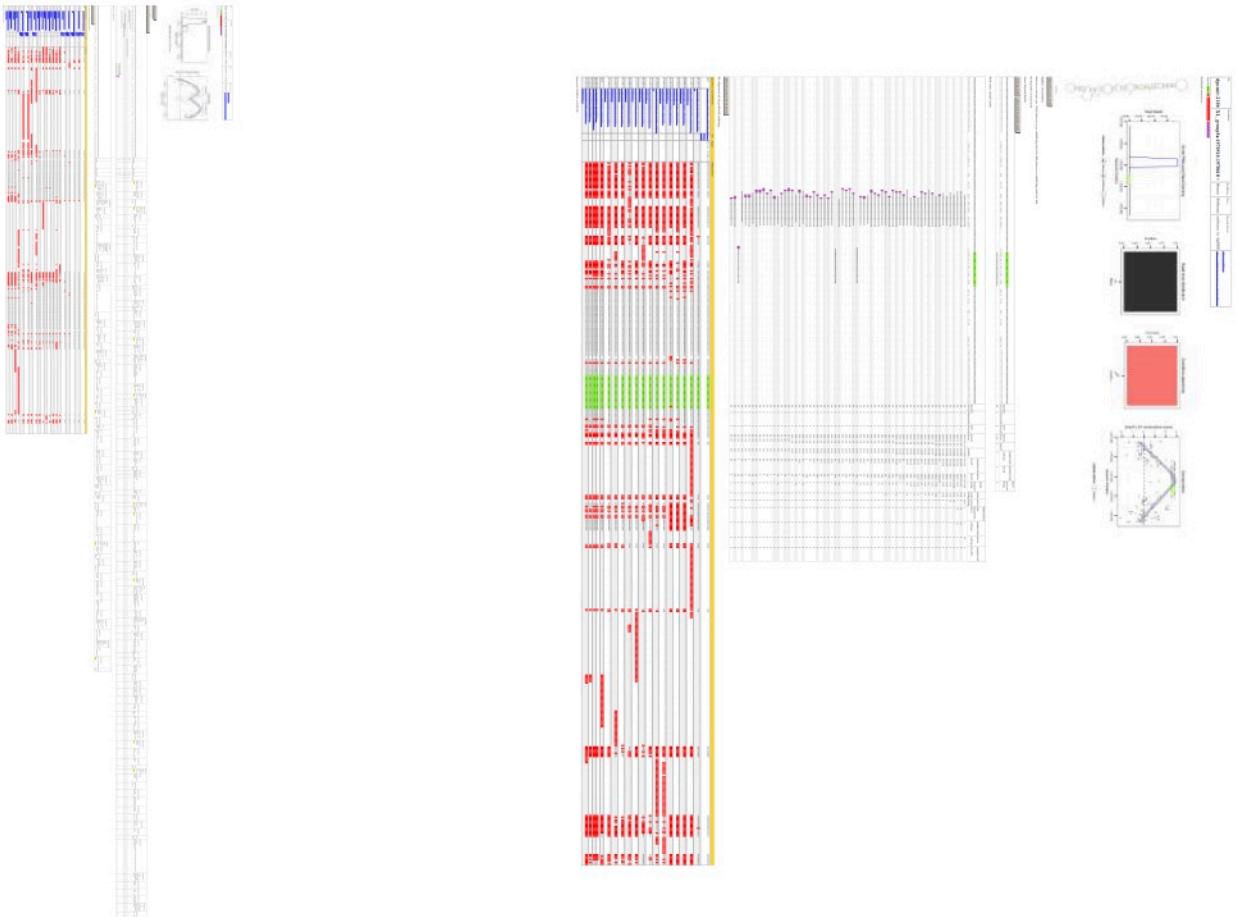
Supplementary Figure S2: Read length distribution for all small RNA libraries sequenced in this study. We extended our previous broad and deep analysis of *D. melanogaster* by sampling 11 additional *Drosophila* species as listed to the right, by analyzing mixed embryos, adult heads, male bodies and female bodies; a testis library was also generated for *D. simulans*. A subset of libraries were sequenced in replicates, especially ones where the expected dominant miRNA-sized peak (21-22 nt) peak was not initially observed. A piRNA peak is seen in most of the body libraries. Due to the technical difficulty in culturing *D. grimshawi*, it was only feasible to generate two libraries for male and female bodies.

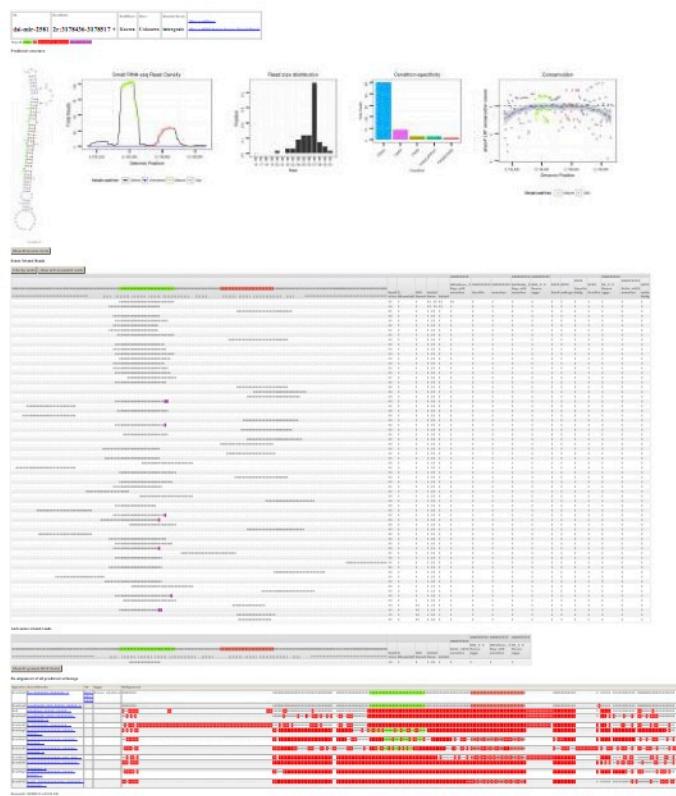
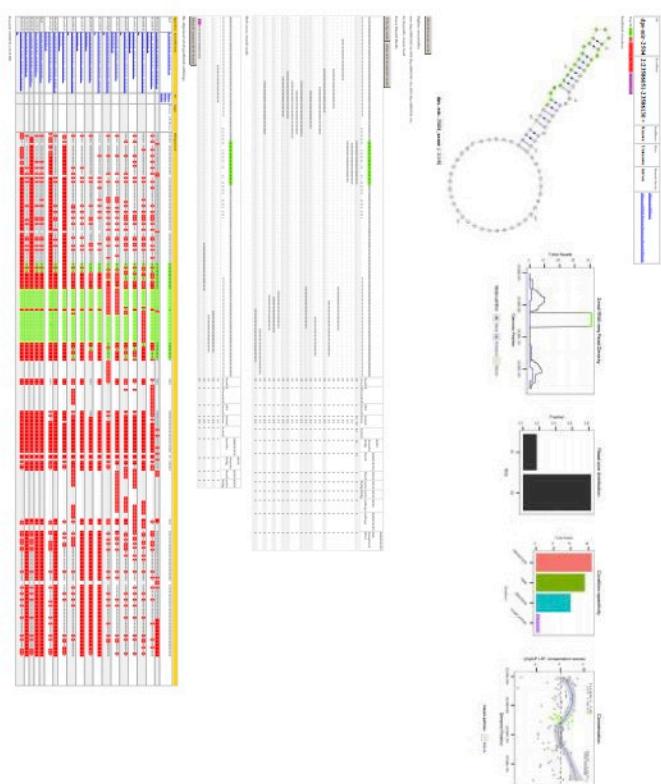


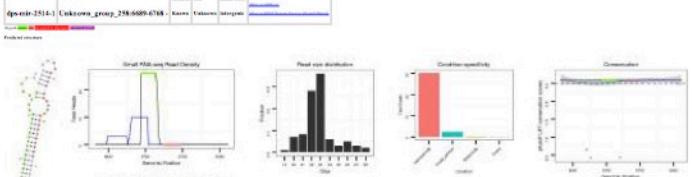
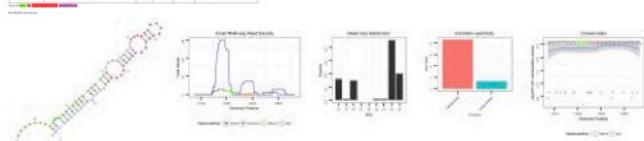
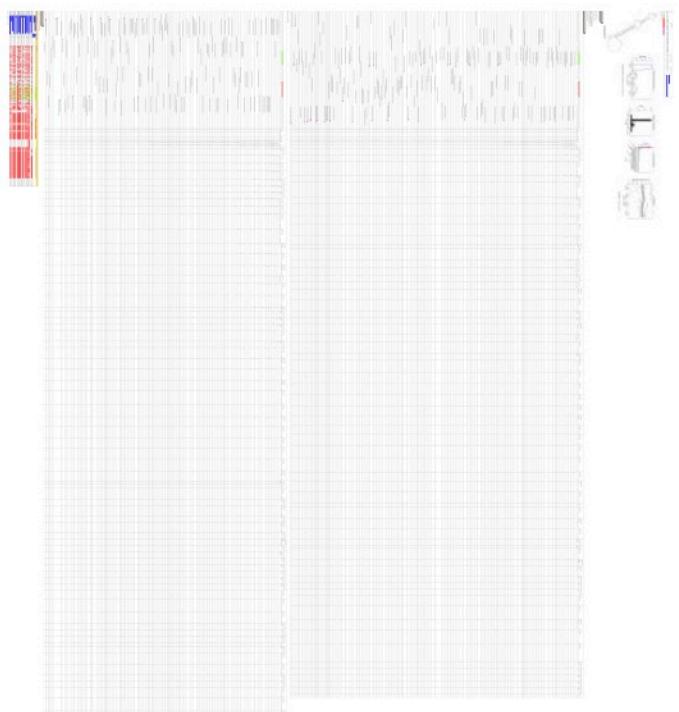
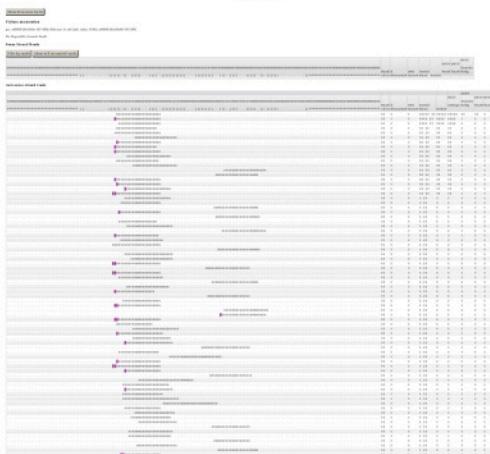
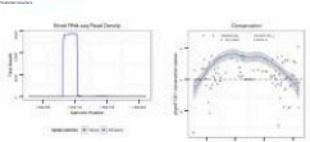
Supplementary Figure S3: Detailed flow-chart of miRNA and mirtron identification pipeline and scoring criteria.

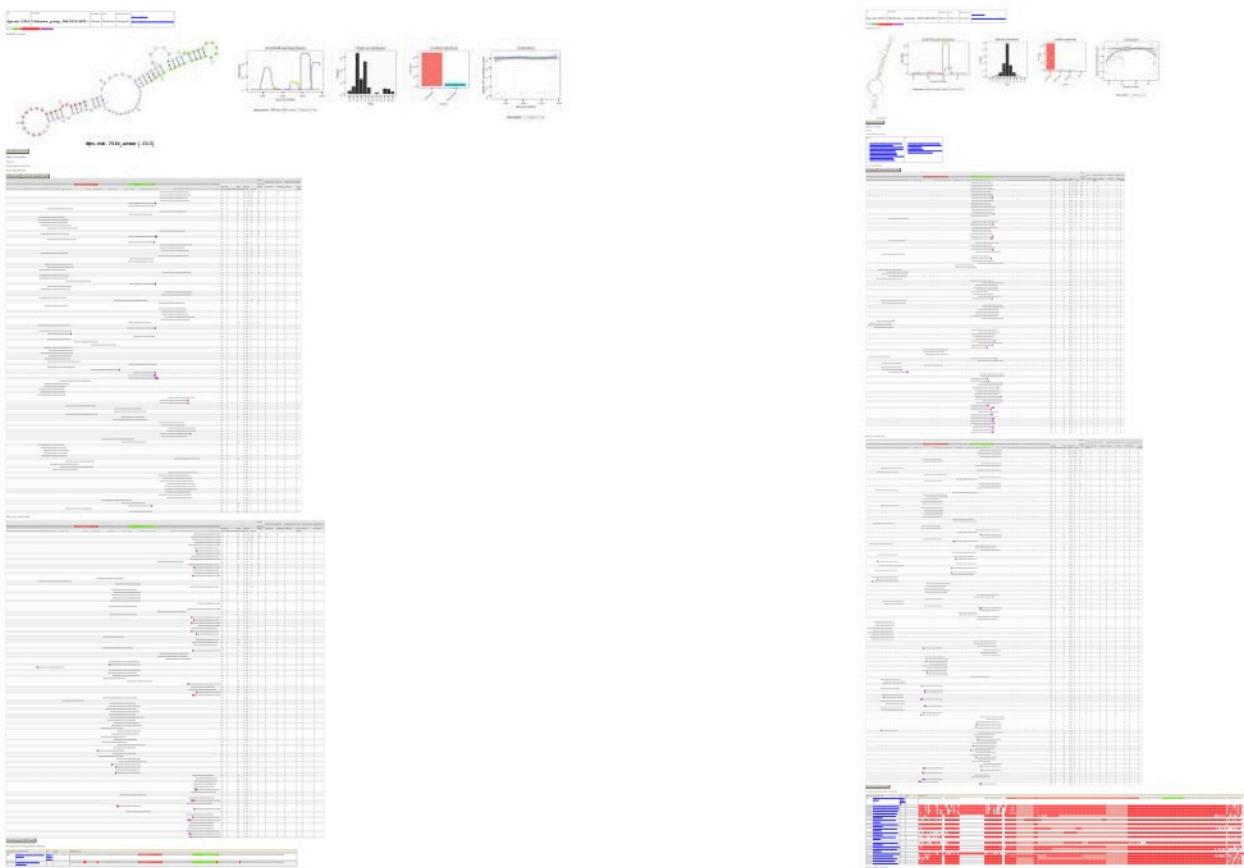
Supplementary Figure S4.
Drosophilid miRbase loci demoted for lack of compelling small RNA evidence.

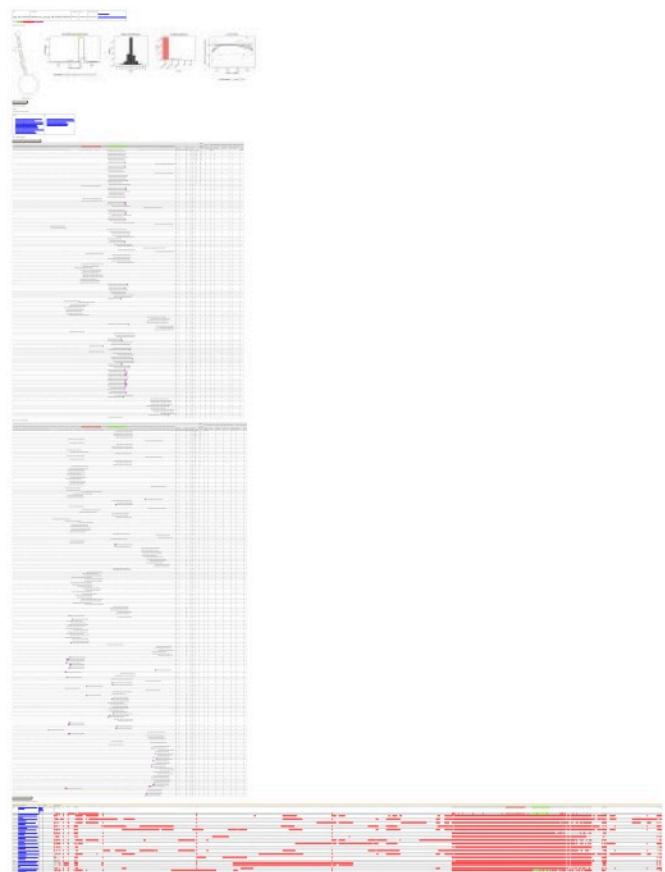






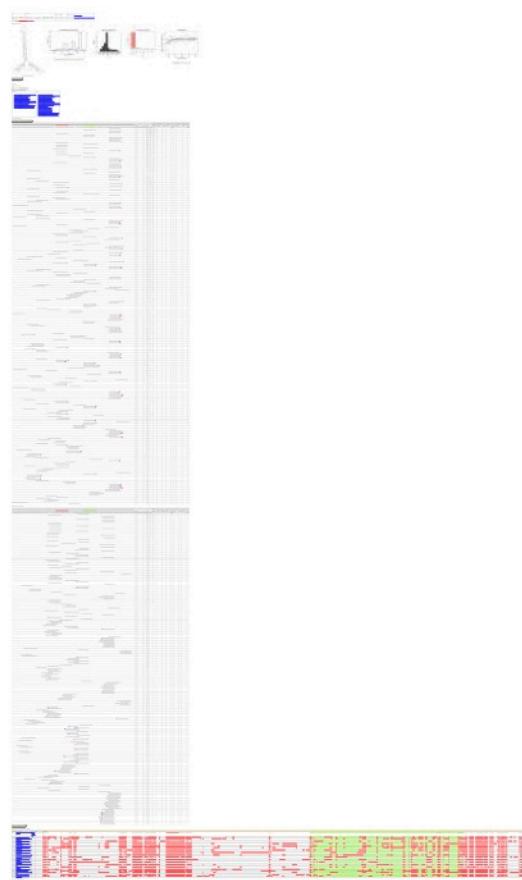
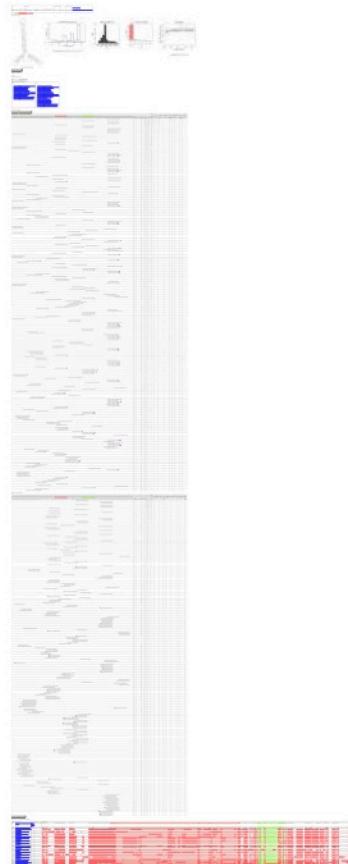


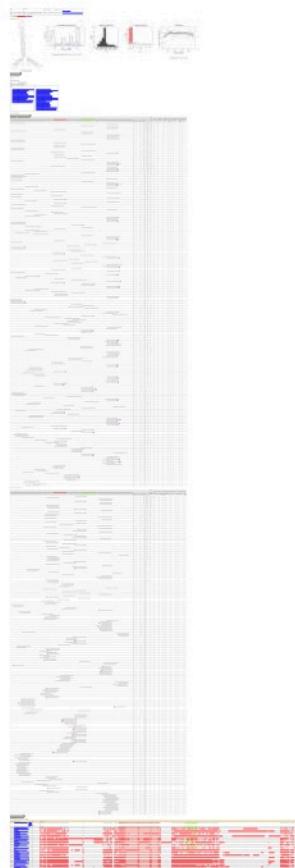
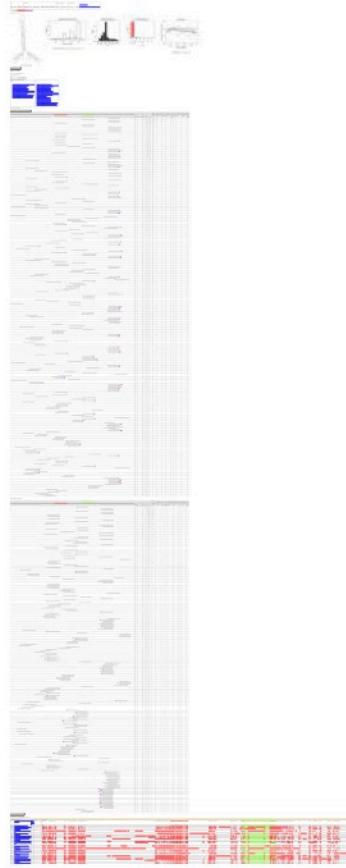
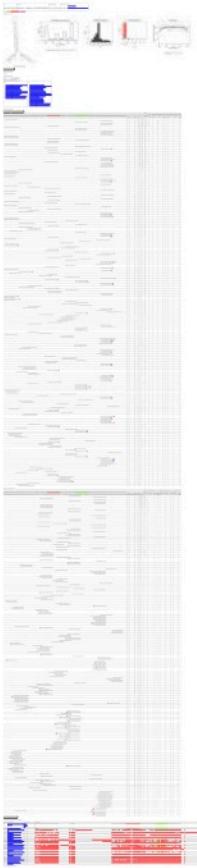


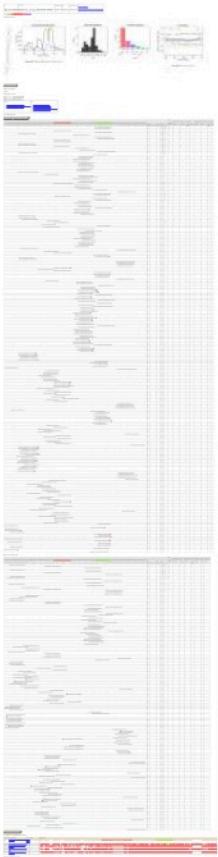




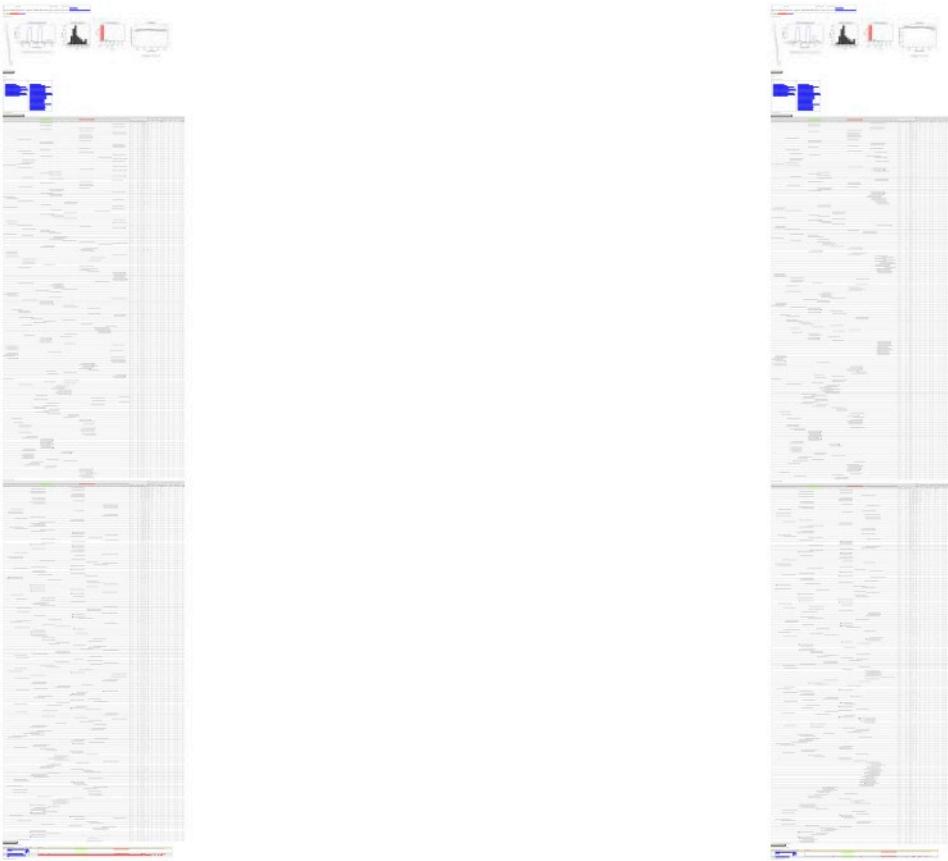


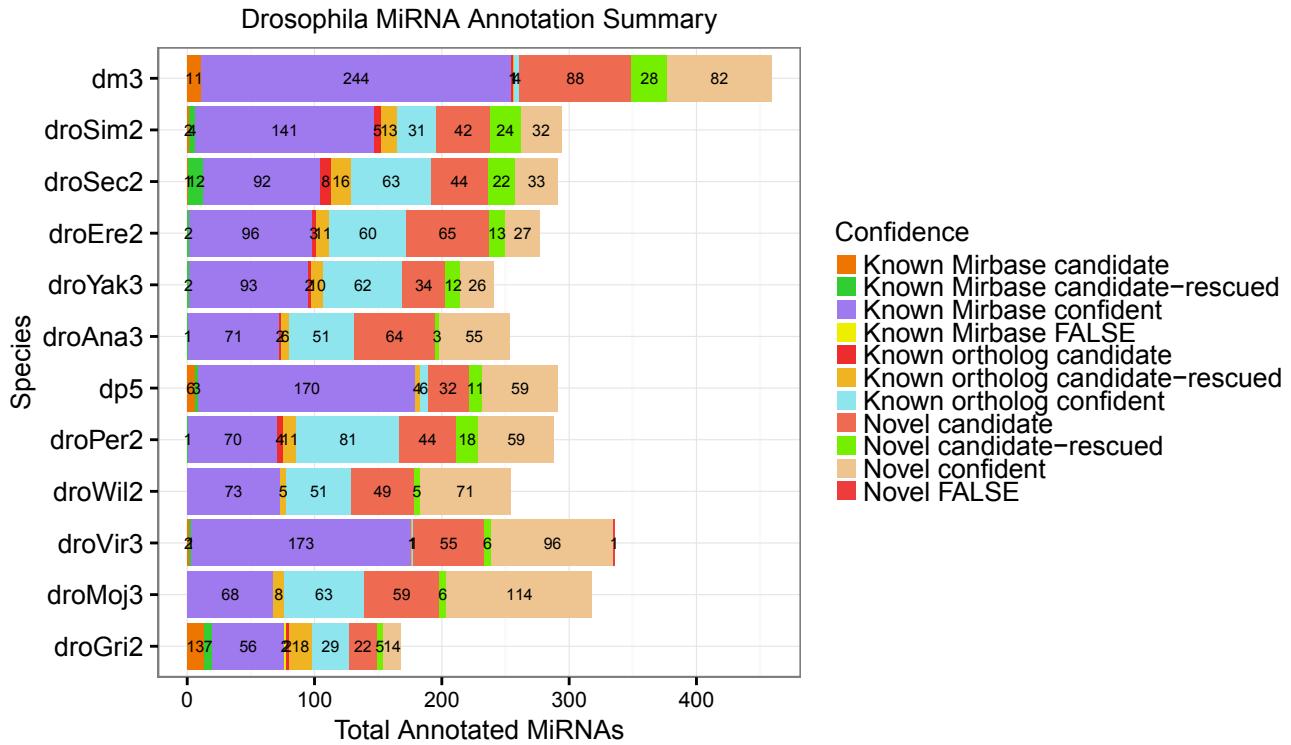




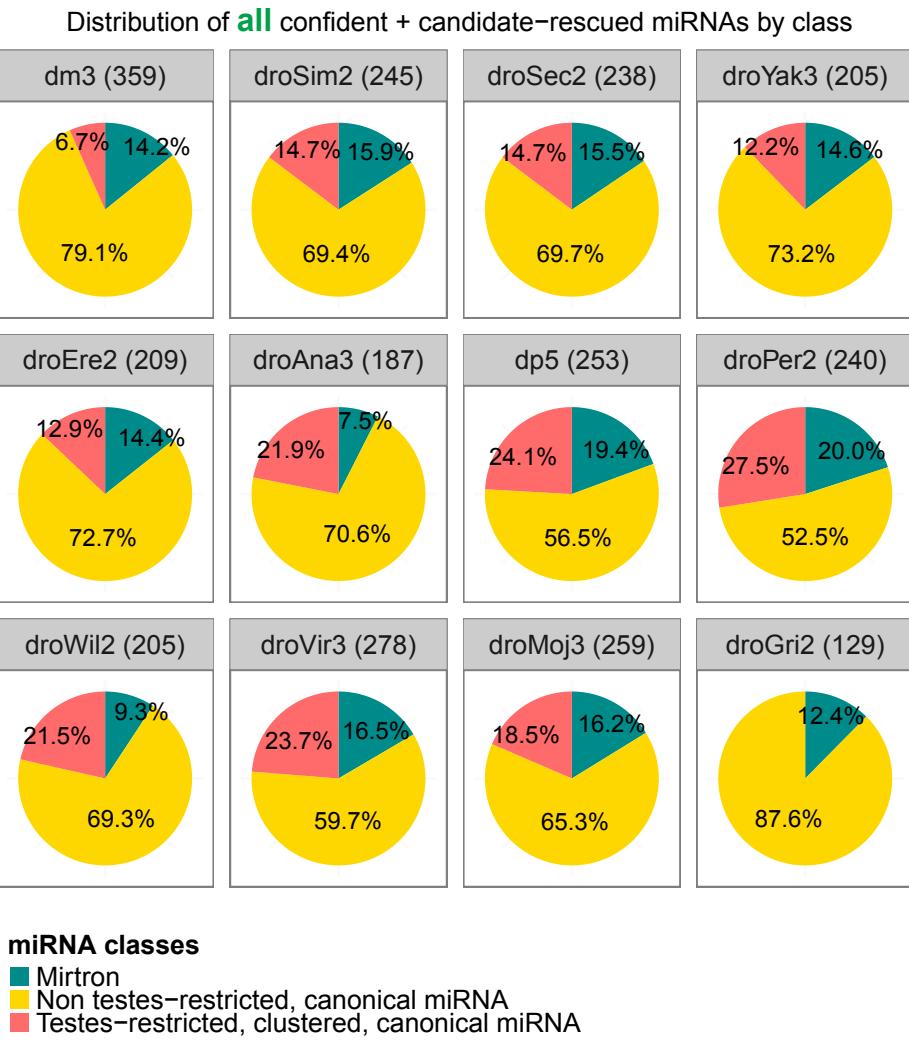








Supplementary Figure S5: Total miRNA and mirtron annotation count within 12 Drosophila species. Annotations are further subdivided within (1) three confidence categories- “confident”, “candidate-rescued, and candidate”, and (2) between known and novel annotations. Note that “candidate” annotations were not utilized for analyses of miRNA flux in this study.



Supplementary Figure S6: Distribution of miRNAs for three classes of miRNAs within each Drosophila species. These classes are defined by biogenesis pathway and canonical miRNAs are further divided by their testes-restricted expression. Only “confident” and “candidate-rescued” loci are included; loci that are considered “candidate” only and lack further rationale to be rescued based on a confidently processed miRNA ortholog are not included in these pie charts.

Supplementary Figure S7.

Read alignments for mir-10404, a conserved non-canonical miRNA generated from the ITS1 spacer in ribosomal RNA, across the Drosophilid phylogeny.

dme-mir-10404 Complete Alignment

CATTTATTGAAGGAATTGATATATGC.....	26	20	4.55	0.03
.....TGTATTTTAATTCTTTCAA.....	21	20	4.50	0.11
.....TGAAGGAATTGATATAATGCCAGTAAAT.....	28	20	4.35	0.02
.....TTATTGAAGGAATTGATATATGCCA.....	25	20	4.05	0.02
.....GTATTTTAATTCTTTCAATA.....	22	20	4.00	0.03
.....TAAAATGGTGTATTTTAATTCTT.....	25	20	3.90	0.21
.....T-ACATTTATTGAAGGAATTGATATATG.....	27	20	3.90	0.05
.....TTTATTGAAGGAATTGATATAATG.....	23	20	3.85	0.02
.....TATTGAAGGAATTGATATATA <ins>A</ins>	21	20	3.80	0.03
.....TTGAAGGAATTGATATATGCCAGTAA.....	26	20	3.60	0.03
.....ATTGAAGGAATTGATATATGCCA <ins>TC</ins>	25	20	3.55	5.80
.....TTATTGAAGGAATTGATAT.....	19	20	3.55	0.02
.....ACATTTATTGAAGGAATTGATATA.....	24	20	3.55	0.02
.....TATTGAAGGAATTGATAT.....	18	20	3.55	0.03
.....TTGAAGGAATTGATATA <ins>TG</ins>	23	20	3.50	0.03
.....TTGAAGGAATTGATATAATGCCAGT.....	24	20	3.35	0.03
.....ATTTATTGAAGGAATTGATATATGCCAG.....	28	20	3.35	0.04
.....TTGAAGGAATTGATATA <ins>TG</ins>	23	20	3.25	0.03
.....TGAAGGAATTGATATAATGCCAGT.....	23	20	3.25	0.02
.....TTTATTGAAGGAATTGATAT.....	20	20	3.20	0.02
.....ATTTATTGAAGGAATTGATAT.....	21	20	3.15	0.02
.....TTGAAGGAATTGATATATGCCAGTA.....	25	20	3.10	0.02
.....ATTTATTGAAGGAATTGAT.....	19	20	3.00	0.02
.....TATTGAAGGAATTGATATA <ins>TG</ins>	24	20	3.00	0.04
.....T-ACATTTATTGAAGGAATTGATATAT.....	26	20	2.90	0.02
.....T-ACATTTATTGAAGGAATTGATATATGC.....	28	20	2.85	0.03
.....GTATTTTAATTCTTTCAATA <ins>C</ins>	24	20	2.80	0.04
.....TTGAAGGAATTGATATATGCCAG.....	23	20	2.80	0.02
.....TTTATTGAAGGAATTGATA.....	19	20	2.75	0.03
.....TATTGAAGGAATTGATATA <ins>TG</ins>	25	20	2.65	0.02
.....TATTTTTAATTCTTTCAATAAA.....	23	20	2.65	0.03
.....TGTATTTTAATTCTTTCA.....	20	20	2.65	0.03
.....TGTATTTTAATTCTTTCAATAAAA.....	26	20	2.55	0.03
.....T-ACATTTATTGAAGGAATTGAT.....	22	20	2.55	0.02
.....TTTATTGAAGGAATTGATATA <ins>TG</ins>	26	20	2.50	0.02
.....ATTTATTGAAGGAATTGATATATG.....	24	20	2.50	0.01
.....ATTTATTGAAGGAATTGATATAT.....	23	20	2.50	0.02
.....TGGTGTATTTTAATTCTTTCAATAA.....	27	20	2.45	0.07
.....AAGGAATTGATATAATGCCAGTAAATGG.....	28	20	2.45	0.02
.....TAAAATGGTGTATTTTAATTCTTTC.....	27	20	2.40	0.07
.....TATTTTTAATTCTTTCAATAA.....	22	20	2.35	0.03
.....TGTATTTTAATTCTTTCAATAAAAAA.....	27	20	2.35	0.04
.....TATTGAAGGAATTGATATATGCCAGTAA.....	28	20	2.30	0.01
.....T-ACATTTATTGAAGGAATTGATAT.....	24	20	2.25	0.02
.....TATTGAAGGAATTGATATA <ins>CG</ins>	20	20	2.25	0.02
.....AAT-ACATTTATTGAAGGAATTGATA.....	25	20	2.20	0.02
.....TTGAAGGAATTGATATATGCT <ins>T</ins>	21	20	2.05	0.02
.....ATTTATTGAAGGAATTGATA.....	20	20	2.05	0.02
.....TATTGAAGGAATTGATATA <ins>TG</ins>	21	20	2.00	0.02
.....ATTTATTGAAGGAATTGATATA.....	22	20	2.00	0.01
.....TTATTGAAGGAATTGATATATGCCAGTA.....	28	20	2.00	0.02
.....TTGAAGGAATTGATATATGCC <ins>G</ins>	22	20	2.00	0.03
.....TGAAGGAATTGATATATGCCAGTA.....	24	20	1.90	0.02
.....TTATTGAAGGAATTGATATATGC <ins>A</ins>	24	20	1.85	0.03
.....TATTGAAGGAATTGATATA <ins>G</ins>	19	20	1.85	0.01
.....TTGAAGGAATTGATATA <ins>TG</ins>	21	20	1.80	0.01
.....ATTTATTGAAGGAATTGATATATGCCA.....	27	20	1.80	0.01
.....TTTATTGAAGGAATTGATATATGCC.....	25	20	1.80	0.05
.....TTATTGAAGGAATTGATA.....	18	20	1.75	0.02
.....TATTTTTAATTCTTTCAATA.....	21	20	1.75	0.02
.....TAAAATGGTGTATTTTAATTCTT.....	24	20	1.65	0.02
.....TGGTGTATTTTAATTCTTTCAA.....	24	20	1.60	0.12
.....TTGATATATGCCAGTAAATGGTG.....	24	20	1.60	0.03
.....TTTATTGAAGGAATTGAT.....	18	20	1.55	0.03
.....ACATTTATTGAAGGAATTGATATATGC.....	27	20	1.50	0.02
.....GAAGGAATTGATATATGC.....	18	20	1.45	0.02
.....GTATTTTTAATTCTTTCAA.....	20	20	1.45	0.02

.....TTGAAGGAATTGATATATGCC T	22	3	53.33	0.31
.....TGAAGGAATTGATATATGCCAGTA.....	24	3	52.33	0.25
.....TGATTTTAATTCTTCATAAAAAA A	28	3	48.00	0.36
.....GTATTTTAATTCTTC.....	18	4	46.25	0.68
.....ATTGAAGGAATTGATATATGCCA.....	23	3	44.00	0.28
.....AAGGAATTGATATATGCCAGTAA.....	23	4	42.25	1.56
.....TGTATTTTAATTCTTCAA A	22	4	39.25	0.33
.....TATTGAAGGAATTGATATATGCCAG.....	25	3	35.67	0.19
.....TGAAGGAATTGATATATGCCAGT.....	23	3	34.67	0.22
.....ATTATTGAAGGAATTGATAT.....	21	3	34.67	0.23
.....TTATTGAAGGAATTGATAT.....	19	3	32.67	0.18
.....TATTGAAGGAATTGATATATGCCA A	25	3	32.67	0.21
.....TTGAAGGAATTGATATATGCCAGTAA.....	26	3	31.67	0.20
.....TATTGAAGGAATTGTA.....	15	4	29.75	0.25
.....GTATTTTAATTCTTCAT.....	21	4	29.50	0.15
.....TATTGAAGGAATTGATATATGCCAGT.....	26	3	26.00	0.12
.....TTGAAGGAATTGATATATGCCAGTAAA.....	24	4	24.50	0.24
.....TTTATTGAAGGAATTGATATAT.....	28	3	24.33	0.17
.....TTGAAGGAATTGATA.....	22	3	23.00	0.17
.....ATTGAAGGAATTGATATAT.....	15	3	22.33	0.18
.....AAGGAATTGATATATGCCAGTAAA.....	19	3	21.67	0.11
.....TATTGAAGGAATTGATATAT T	25	4	21.50	0.27
.....TATTGAAGGAATTGATATAT GAA	22	3	21.33	0.11
.....GAAGGAATTGATATATGC.....	23	4	21.25	0.15
.....AAGGAATTGATATATGCCAGTAAA.....	18	4	21.25	0.50
.....TTTATTGAAGGAATTGATATATGC.....	24	4	21.00	0.62
.....TTTATTGAAGGAATTGATATATGCC.....	24	3	17.33	0.11
.....ATTTATTGAAGGAATTGATATAT.....	25	3	17.33	0.14
.....ATTTATTGAAGGAATTGATATATGCC.....	23	3	17.00	0.13
.....GAAGGAATTGATATATGCC.....	26	3	17.00	0.11
.....TGTATTTTAATTCTTCATAAA C	20	4	16.50	0.33
.....ATTTTTAATTCTTCATAAA.....	25	4	16.25	0.10
.....TTGAAGGAATTGATATATGCCA A	22	4	15.50	0.13
.....TTATTGAAGGAATTGATATATGCCAGTAA.....	23	3	15.33	0.10
.....ATTTTTAATTCTTCATAAA.....	28	3	15.00	0.17
.....TATTGAAGGAATTGATATATGCCA AA	21	4	14.50	0.11
.....ATTTATTGAAGGAATTGATATATGCCAG.....	26	3	14.33	0.09
.....GAAGGAATTGATATA.....	28	3	14.00	0.95
.....T-AATTATTGAAGGAATTGATATA.....	15	4	13.50	0.68
.....TATTGAAGGAATTGATATAT A	25	2	12.50	0.15
.....ATTTATTGAAGGAATTGATATATGC.....	21	3	12.33	0.09
.....TTATTGAAGGAATTGATATATGC A	25	3	12.00	0.08
.....TGTATTTTAATTCTTCATAAA T	24	3	12.00	0.11
.....GTGTATTTTAATTCTTCAA.....	25	4	11.75	0.08
.....TATTTTAATTCTTCAA.....	22	4	11.75	0.24
.....GGTGTATTTTAATTCTTCAA.....	19	4	11.75	0.19
.....TGTATTTTAATTCTTCATAAT T	22	4	11.50	0.34
.....GGTGTATTTTAATTCTTCATAAA.....	23	4	11.50	0.09
.....GAAGGAATTGATATATGCCAG.....	26	4	11.50	0.34
.....ATTTTTAATTCTTCATAAA.....	21	4	11.50	0.33
.....ATTGAAGGAATTGATATATGCC T	23	4	11.25	0.09
.....GAAGGAATTGATATA.....	23	3	11.00	0.21
.....AATTGATATATGCCAGT.....	17	4	11.00	0.56
.....TTTATTGAAGGAATTGTA.....	18	4	10.75	0.32
.....TTGAAGGAATTGATATATGCCAGTAA.....	17	4	10.75	0.09
.....TTGAAGGAATTGATATATGC A	27	3	10.67	0.13
.....TTATTGAAGGAATTGATATATGCAG.....	21	3	10.67	0.07
.....TTTATTGAAGGAATTGATA.....	19	3	10.00	0.06
.....TTTATTGAAGGAATTGATATA.....	21	3	9.67	0.08
.....TTATTGAAGGAATTGATATATGCCAG.....	26	3	9.67	0.06
.....ATTTTTAATTCTTCATAAA.....	20	4	9.00	0.08
.....TTATTGAAGGAATTGATATATGCCAGT.....	27	3	8.67	0.06
.....TTTATTGAAGGAATTGATAT.....	20	3	8.67	0.07
.....ATTGAAGGAATTGATATATGC A	22	3	8.67	0.06
.....ATTTATTGAAGGAATTGATATA.....	22	3	8.67	0.13
.....TTATTGAAGGAATTGATA.....	16	4	8.50	0.07
.....GAAGGAATTGATATATGCC.....	19	4	8.50	0.15

.....TTGAAGGAATTGATATATGCC.....	21	20	55.70	1.58
.....TATTGAAGGAATTGATATATGCCA.....	24	20	31.25	0.91
.....TTGAAGGAATTGATATATGC.....	20	20	23.60	0.44
.....ATTGAAGGAATTGATATATGC.....	21	20	16.05	0.30
.....TTATTGAAGGAATTGATATATG.....	22	20	14.45	0.95
.....ATTGAAGGAATTGATATATGCC.....	22	20	9.65	0.25
.....TTATTGAAGGAATTGATATATGC.....	23	20	8.65	0.45
.....TGTATTTTAATTCTTCATAAA.....	25	20	6.90	0.36
.....TGAAGGAATTGATATATGCCA.....	21	20	4.75	0.30
.....TATTGAAGGAATTGATATATG <color>T</color>	22	20	4.70	0.07
.....TTTATTGAAGGAATTGATATATGC.....	24	20	2.50	0.13
.....ATTTATTGAAGGAATTGATATATGC.....	25	20	2.35	0.16
.....GTATTTTAATTCTTCATAAA.....	22	20	2.35	0.12
.....TTGAAGGAATTGATATATG.....	19	20	2.30	0.07
.....TGAAGGAATTGATATATGCC.....	20	20	2.15	0.11
.....TTATTGAAGGAATTGATATATGCCA.....	25	20	1.60	0.05
.....GTATTTTAATTCTTCATAAA.....	24	20	1.25	0.07
.....TTATTGAAGGAATTGATATAT.....	21	20	1.15	0.07
.....GTATTTTAATTCTTCATAAA.....	23	20	1.10	0.06
.....TTATTGAAGGAATTGATA.....	18	20	0.85	0.06
.....TGAAGGAATTGATATATG <color>C</color>	19	20	0.80	0.02
.....TTATTGAAGGAATTGATATATGCC.....	24	20	0.75	0.04
.....TATTGAAGGAATTGATATATGCCAGT.....	26	20	0.75	0.02
.....TGAAGGAATTGATATATGCCAG.....	22	20	0.70	0.04
.....TGAAGGAATTGATATATGCCAGAA.....	26	20	0.65	0.02
.....AATTATTGAAGGAATTGATATATGC.....	26	20	0.65	0.02
.....TTATTGAAGGAATTGATATATGATA.....	20	20	0.65	0.02
.....TATTGAAGGAATTGATATATG <color>A</color>	22	20	0.60	0.03
.....GAAGGAATTGATATATGCCA.....	20	20	0.60	0.01
.....TATTGAAGGAATTGATATA.....	19	20	0.60	0.03
.....TGAAGGAATTGATATATGCCAGT.....	23	20	0.55	0.02
.....TATTGAAGGAATTGATATATG <color>T</color>	23	20	0.50	0.01
.....ATTGAAGGAATTGATATAT.....	19	20	0.50	0.03
.....TGAAGGAATTGATATATGCCAGT.....	24	20	0.45	0.01
.....TATTGAAGGAATTGATATATG <color>C</color>	20	20	0.40	0.01
.....TATTGAAGGAATTGATATATG <color>S</color>	19	20	0.40	0.01
.....TTTATTGAAGGAATTGATA.....	19	20	0.40	0.01
.....TATTGAAGGAATTGATATATGCCAG.....	25	20	0.35	0.02
.....TGTATTTTAATTCTTCATAAT <color>T</color>	23	20	0.35	0.01
.....GAAGGAATTGATATATGCC.....	19	20	0.35	0.02
.....TTTATTGAAGGAATTGATATATGCC.....	25	20	0.35	0.01
.....TTTATTGAAGGAATTGATATATG.....	23	20	0.35	0.01
.....ATTTATTGAAGGAATTGATATATGCC.....	26	20	0.30	0.01
.....TTGAAGGAATTGATATATGCCAGTAAA.....	27	20	0.30	0.01
.....ATTGATATATGCCAGTAAATG.....	22	20	0.30	0.01
.....AAGGAATTGATATATGCCA.....	19	20	0.30	0.02
.....TGAAGGAATTGATATATGCCAGAAAA.....	27	20	0.25	0.02
.....TTTATTGAAGGAATTGATATATGCCA.....	26	20	0.25	0.01
.....GAAGGAATTGATATATG.....	18	20	0.25	0.01
.....TATTGAAGGAATTGATATATGCCA <color>A</color>	25	20	0.25	0.02
.....TGTATTTTAATTCTTCATA.....	20	20	0.20	0.01
.....ATTTATTGAAGGAATTGATA.....	20	20	0.20	0.01
.....TTTATTGAAGGAATTGAT.....	18	20	0.20	0.01
.....AAATTATTGAAGGAATT.....	18	20	0.20	0.01
.....TTTATTGAAGGAATTGATATATGCCAG.....	27	20	0.20	0.01
.....CAGTAAATGGTGTATTTTAAT.....	23	20	0.20	0.01
.....TGAAGGAATTGATATATGCCAGTA.....	25	20	0.20	0.01
.....TTGAAGGAATTGATATATGCCA <color>T</color>	23	20	0.15	0.01
.....TGTATTTTAATTCTTCATAAT <color>T</color>	24	20	0.10	0.01
.....TGAAGGAATTGATATATGCCA <color>T</color>	22	20	0.10	0.01
.....ATTTATTGAAGGAATTGATATATG.....	24	20	0.10	0.01
.....GTATTTTAATTCTTCATA.....	19	20	0.10	0.01
.....AAGGAATTGATATATGCC.....	18	20	0.10	0.01
.....TGTATTTTAATTCTTCATAATAAAA <color>A</color>	28	20	0.10	0.01
.....TGTATTTTAATTCTTCATAA <color>A</color>	22	20	0.10	0.01
.....ATTTTAATTCTTCATAATAAC <color>C</color>	22	20	0.10	0.01
.....TTGAAGGAATTGATATATGCC <color>G</color>	22	20	0.10	0.00

TTGAAGGAATTGATATATGCCA	23	20	0.10	0.00
TTATTGAAGGAATTGATATATGCCAGT	28	20	0.10	0.00
TTGAAGGAATTGATATATGCCA	21	20	0.10	0.01
TGAAGGAATTGATATATGC	18	20	0.10	0.01
TATTGAAGGAATTGATATA	20	20	0.10	0.01
TGAAGGAATTGATATATGCCAGTAAAT	28	20	0.10	0.01
TTTATTGAAGGAATTGATAT	20	20	0.10	0.01
ATTTATTGAAGGAATTGATATA	22	20	0.10	0.01
T-AATTTATTGAAGGAATTGATATATGC	28	20	0.10	0.01
GTATTTTAATTCTTCATAAA	24	20	0.10	0.01
TTATTGAAGGAATTGATATATGCCAGT	27	20	0.10	0.01
ATTGAAGGAATTGATATATGC	22	20	0.10	0.01
TTGAAGGAATTGATATA	19	20	0.10	0.01
AGGTATAAAA-T-AAATTATTGAAGGA	27	20	0.05	0.00
TTATTGAAGGAATTGATATATGC	23	20	0.05	0.00
AAATTATTGAAGGAATTGATATATGC	26	20	0.05	0.00
AATTATTGAAGGAATTGATATA	23	20	0.05	0.00
GAAGGAATTGATATATGCCAG	21	20	0.05	0.00
TATTTTTAATTCTTCATA	21	20	0.05	0.00
AAGGAATTGATATATGCCAGTAA	23	20	0.05	0.00
ATTTATTGAAGGAATTGATATATGCCA	27	20	0.05	0.00
ATTGAAGGAATTGATATATGCCAC	25	20	0.05	0.00
TATTGAAGGAATTGATATATGCCA	25	20	0.05	0.00
GTGTATTTTAATTCTTCAA	23	20	0.05	0.00
TATTGAAGGAATTGATATATGC	23	20	0.05	0.00
TTAGGTATAAAA-T-AAATTATTGAAGG	28	20	0.05	0.00
TTGAAGGAATTGATATATGT	20	20	0.05	0.00
AAAA-T-AAATTATTGAAGGA	22	20	0.05	0.00
TATTGAAGGAATTGATATATGCCAGT	27	20	0.05	0.00
A-T-AAATTATTGAAGGAATTGA	22	20	0.05	0.00
TGAAGGAATTGATATATGCC	21	20	0.05	0.00
ATTGAAGGAATTGATATATGC	21	20	0.05	0.00
TATTGAAGGAATTGATATGCC	21	20	0.05	0.00
GTGTATTTTAATTCTTCATA	23	20	0.05	0.00
TTTAATTCTTCATAAA	19	20	0.05	0.00
GAAGGAATTGATATATGCCAGTAAAT	27	20	0.05	0.00
A-T-AAATTATTGAAGGAATTGAT	23	20	0.05	0.00
ATTGAAGGAATTGATATATGCCAG	24	20	0.05	0.00
TATTGAAGGAATTGATATA	19	20	0.05	0.00
TTGAAGGAATTGATATATGC	21	20	0.05	0.00
TGTATTTTAATTCTTCATAAA	26	20	0.05	0.00
AA-T-AAATTATTGAAGGA	19	20	0.05	0.00
ATTATTGAAGGAATTGATATATGCC	27	20	0.05	0.00
TATTGAAGGAATTGATATA	21	20	0.05	0.00
Total			541.15	7.78

Total

			Size	Hits	Total	Total
					Norm	Norm
					RPM	
TTATTTAGGTATA	AAAA	-T-ACAT	TATTGAAGGAATTGATATATGCAGTAAATGGTGTATTTAATTCTTCATAAAA	ATTGAC-----A	ATATATA	CAAA
.....(((((.....((-(-.(((((((((.((((((.)))))))....))))))))....))))....))))....)).						
.....TATTGAAGGAATTGATATATG	21	20	251.35	3.67
.....TATTGAAGGAATTGATATATGC	22	20	242.90	3.85
.....TATTGAAGGAATTGATATATGCC	23	20	92.60	0.93
.....TTGAAGGAATTGATATATGCCA	22	20	34.15	0.40
.....TTGAAGGAATTGATATATGCC	21	20	26.45	0.41
.....TTATTGAAGGAATTGATATATGC	23	20	15.50	0.19
.....ATTGAAGGAATTGATATATGCC	22	20	14.65	0.23
.....TATTGAAGGAATTGATATAT	20	20	13.30	0.15
.....TTGAAGGAATTGATATATGC	20	20	13.25	0.15
.....TTATTGAAGGAATTGATATATG	22	20	12.80	0.14
.....ATTGAAGGAATTGATATATGC	21	20	12.20	0.21
.....TGTATTTTAATTCTTCATAA	23	20	8.25	0.17
.....TGTATTTTAATTCTTCATAAA	25	20	7.60	0.15
.....TGAAAGGAATTGATATATGCCA	21	20	5.55	0.07
.....ATTGAAGGAATTGATATATG	20	20	5.35	0.10
.....TATTGAAGGAATTGATATA	19	20	3.60	0.06
.....TTTATTGAAGGAATTGATATATGC	24	20	3.40	0.04
.....TGTATTTTAATTCTTCATA	22	20	3.00	0.10
.....TGTATTTTAATTCTTCATAAA	24	20	2.70	0.05

.....ATTTATTGAAGGAATTGATATATGC.....	25	20	2.45	0.03
.....TTGAAGGAATTGATATATG.....	19	20	2.40	0.03
.....TATTGAAGGAATTGATATATGCC T	24	20	2.35	0.04
.....GTATTTTAATTCTTCAAT.....	21	20	2.30	0.05
.....TGAAGGAATTGATATATGCC.....	20	20	2.30	0.03
.....TTATTGAAGGAATTGATATATGCC.....	24	20	2.20	0.03
.....TTGAAGGAATTGATATATGCC T	22	20	2.10	0.03
.....TTATTGAAGGAATTGATATA.....	20	20	1.85	0.02
.....GTATTTTAATTCTTCAATAA.....	24	20	1.80	0.07
.....TATTGAAGGAATTGATATATGCCA.....	24	20	1.80	0.03
.....TATTGAAGGAATTGATATATGC A	23	20	1.70	0.03
.....TGTATTTTAATTCTTCAATAAA.....	26	20	1.60	0.06
.....T-ACATTATTGAAGGAATTGATATA.....	25	20	1.55	0.03
.....TGAAGGAATTGATATATGC.....	19	20	1.45	0.02
.....TATTGAAGGAATTGATATATGCCAG.....	25	20	1.35	0.02
.....A-T-ACATTATTGAAGGAATTGATAT.....	25	20	1.35	0.03
.....TATTGAAGGAATTGATATATGCCAGT.....	26	20	1.20	0.02
.....CATTATTGAAGGAATTGATATATGC.....	26	20	1.20	0.02
.....TATTGAAGGAATTGATATATGA A	22	20	1.15	0.01
.....TTATTGAAGGAATTGATATAT.....	21	20	1.10	0.03
.....GAAGGAATTGATATATGCC.....	19	20	1.05	0.01
.....TATTGAAGGAATTGATATATGT T	22	20	0.95	0.02
.....ATTGAAGGAATTGATATA.....	19	20	0.95	0.01
.....GTATTTTAATTCTTCAATAA.....	23	20	0.80	0.03
.....ATTTATTGAAGGAATTGATA.....	20	20	0.80	0.01
.....TATTGAAGGAATTGATATAT T	21	20	0.80	0.01
.....TTTATTGAAGGAATTGATATATGCC.....	25	20	0.75	0.01
.....A-T-ACATTATTGAAGGAATTGATA.....	24	20	0.70	0.03
.....TGTATTTTAATTCTTCAA.....	21	20	0.65	0.01
.....TTATTGAAGGAATTGATATATGCCA.....	25	20	0.60	0.01
.....TTTATTGAAGGAATTGATATATGCCAG.....	27	20	0.55	0.02
.....ATTGATATATGCCAGTAAAATGGTGTA.....	27	20	0.50	0.01
.....GATATATGCCAGTAAAATG.....	19	20	0.50	0.02
.....TATTGAAGGAATTGATATATGC T	23	20	0.50	0.02
.....GAAGGAATTGATATATGC.....	18	20	0.50	0.01
.....TTTATTGAAGGAATTGATATATGCCA.....	26	20	0.50	0.01
.....TGTATTTTAATTCTTCAATAA T	25	20	0.50	0.01
.....TGAAGGAATTGATATATGCCAGT.....	23	20	0.50	0.01
.....TTATTGAAGGAATTGATATATGCCAGT.....	27	20	0.45	0.01
.....TTATTGAAGGAATTGATAT.....	19	20	0.45	0.01
.....TTTATTGAAGGAATTGATATA.....	21	20	0.45	0.01
.....ATTATTATTGAAGGAATTGATATAT.....	23	20	0.45	0.01
.....A-T-ACATTATTGAAGGAATTGATAT.....	27	20	0.45	0.02
.....TGAAGGAATTGATATATGCCAG.....	22	20	0.45	0.01
.....TTGAAGGAATTGATATATGCCA T	23	20	0.40	0.01
.....ATTGAAGGAATTGATATATGCCA.....	23	20	0.40	0.01
.....TATTTTTAATTCTTCAATAAA.....	23	20	0.40	0.01
.....TGAAGGAATTGATATATGCCAGT.....	24	20	0.40	0.01
.....ATTGATATATGCCAGTAAAATG.....	22	20	0.40	0.01
.....GTATTTTAATTCTTCAATAAAA.....	25	20	0.35	0.01
.....A-T-ACATTATTGAAGGAATT.....	20	20	0.35	0.01
.....TGTATTTTAATTCTTCAATAAAA.....	27	20	0.35	0.01
.....TGTATTTTAATTCTTCAATAA C	25	20	0.35	0.01
.....ATTTTTAATTCTTCAATAAAA.....	22	20	0.35	0.01
.....TGTATTTTAATTCTTCAAT T	23	20	0.30	0.01
.....T-ACATTATTGAAGGAATTGATA.....	23	20	0.30	0.01
.....TTTATTGAAGGAATTGATA.....	19	20	0.30	0.01
.....A-T-ACATTATTGAAGGAATTGAT.....	23	20	0.30	0.00
.....ATTGAAGGAATTGATATATGCC T	23	20	0.30	0.01
.....TTTATTGAAGGAATTGATATATG.....	23	20	0.30	0.01
.....AA-T-ACATTATTGAAGGAATTGATAT.....	26	20	0.30	0.01
.....ATAAAAA-T-ACATTATTGAAGGAATT.....	26	20	0.30	0.01
.....GTATTTTAATTCTTCAATAA T	24	20	0.30	0.01
.....ATTGAAGGAATTGATATATGCCAGT.....	25	20	0.30	0.01
.....GAAGGAATTGATATATGCCA.....	20	20	0.25	0.01
.....TATTTTTAATTCTTCAAT.....	20	20	0.25	0.01
.....TGAAGGAATTGATATATGCCAGTAAA.....	26	20	0.25	0.01

.....TATTGAAGGAATTGATATATGCCAGTA.....	27	20	0.25	0.01
.....CATTTATTGAAGGAATTGATA.....	21	20	0.25	0.00
.....TTATTGAAGGAATTGATA.....	18	20	0.25	0.01
.....AA-T-ACATTATTGAAGGAAT.....	20	20	0.25	0.01
.....AGGTATAAAAA-T-ACATTATTGAAGGAAT.....	29	20	0.25	0.01
.....TGATTTTAATTCTTCATAA.....	24	20	0.25	0.01
.....TTTTAATTCTTCATAAAA.....	21	20	0.25	0.01
.....ATTATTGAAGGAATTGATAT.....	21	20	0.25	0.00
.....TGAAGGAATTGATATATGCCAGTA.....	25	20	0.25	0.00
.....ATTGAAGGAATTGATATATGCA.....	22	20	0.25	0.01
.....TTGATATATGCCAGTAAAATGGTGT.....	26	20	0.25	0.01
.....ACATTATTGAAGGAATTGATATATGCC.....	28	20	0.20	0.00
.....T-ACATTATTGAAGGAATT.....	19	20	0.20	0.00
.....AAGGAATTGATATATGCC.....	18	20	0.20	0.01
.....TGATATATGCCAGTAAAATG.....	20	20	0.20	0.00
.....GATATATGCCAGTAAAATGGTGT.....	23	20	0.20	0.00
.....ATTATTGAAGGAATTGATATATGCC.....	26	20	0.20	0.01
.....TTGAAGGAATTGATATATGCA.....	21	20	0.20	0.01
.....A-T-ACATTATTGAAGGAATTG.....	21	20	0.20	0.00
.....ATTATTGAAGGAATTGATATATG.....	24	20	0.20	0.01
.....ATTTTAATTCTTCATAA.....	20	20	0.20	0.01
.....ATAAAAA-T-ACATTATTGAAGGA.....	23	20	0.20	0.00
.....ATTATTGAAGGAATTGATATATGCCAGT.....	29	20	0.20	0.00
.....CATTATTGAAGGAATTGATATATGCC.....	27	20	0.20	0.00
.....ATTTTAATTCTTCATAAA.....	23	20	0.20	0.00
.....TTGATATATGCCAGTAAAATGGTGT.....	25	20	0.15	0.00
.....ATATATGCCAGTAAAATGGTGT.....	22	20	0.15	0.01
.....TTGAAGGAATTGATATATGCCAGT.....	24	20	0.15	0.00
.....T-ACATTATTGAAGGAATTGATATG.....	28	20	0.15	0.00
.....T-ACATTATTGAAGGAATTGATAT.....	24	20	0.15	0.00
.....TGAAGGAATTGATATATG.....	18	20	0.15	0.01
.....GAAGGAATTGATATATGCCAGTAAAAT.....	27	20	0.15	0.01
.....TTGAAGGAATTGATATATGCCA.....	23	20	0.15	0.01
.....TTATTGAAGGAATTGATATATGCA.....	24	20	0.15	0.01
.....GATATATGCCAGTAAAATGGTGT.....	22	20	0.15	0.00
.....TTGAAGGAATTGATATATGCCAGTAAA.....	28	20	0.15	0.00
.....TATTGAAGGAATTGATAT.....	18	20	0.15	0.02
.....GTATTTTAATTCTTCATAA.....	22	20	0.15	0.01
.....GTATTTTAATTCTTCAA.....	20	20	0.15	0.01
.....AA-T-ACATTATTGAAGGAATT.....	21	20	0.15	0.01
.....A-T-ACATTATTGAAGGAATTGATA.....	26	20	0.15	0.01
.....CATTATTGAAGGAATTGATATATG.....	25	20	0.15	0.01
.....ATAAAAA-T-ACATTATTGAAGGAATTGAT.....	29	20	0.15	0.01
.....AA-T-ACATTATTGAAGGAATTGATATAT.....	28	20	0.15	0.01
.....TAAAAA-T-ACATTATTGAAGGAATTG.....	27	20	0.15	0.00
.....TGATTTTAATTCTTCATAA.....	22	20	0.15	0.00
.....GGAATTGATATATGCCAGTAAAATGGT.....	27	20	0.10	0.00
.....ACATTATTGAAGGAATTGATATATG.....	27	20	0.10	0.00
.....AAA-T-ACATTATTGAAGGAATTGATA.....	26	20	0.10	0.00
.....ATTATTGAAGGAATTGATATA.....	22	20	0.10	0.00
.....ATTGAAGGAATTGATATATGCA.....	21	20	0.10	0.00
.....AAA-T-ACATTATTGAAGGAATT.....	22	20	0.10	0.00
.....TATTGAAGGAATTGATATATGCCAGC.....	26	20	0.10	0.01
.....ATTTTAATTCTTCATAA.....	19	20	0.10	0.00
.....TTGAAGGAATTGATATATGCCAGTA.....	26	20	0.10	0.00
.....ATATATGCCAGTAAAATGGTGT.....	23	20	0.10	0.01
.....AGGTATAAAAA-T-ACATTATTGAAG.....	25	20	0.10	0.00
.....GATATATGCCAGTAAAATGGTGT.....	24	20	0.10	0.00
.....TATTGAAGGAATTGATATATGCA.....	21	20	0.10	0.02
.....TTGAAGGAATTGATATATGCCAGTAAA.....	29	20	0.10	0.00
.....GGTATAAAAA-T-ACATTATTGAAGGA.....	26	20	0.10	0.00
.....CAGTAAAATGGTGTATTTTAAT.....	23	20	0.10	0.00
.....TTGAAGGAATTGATATATGCCAGTA.....	25	20	0.10	0.00
.....TTGATATATGCCAGTAAAATGGTGTAT.....	27	20	0.10	0.00
.....ACATTATTGAAGGAATTGATA.....	22	20	0.10	0.00
.....AATTGATATATGCCAGTAAAATGGTGT.....	27	20	0.10	0.00
.....ATATGCCAGTAAAATGGTGT.....	20	20	0.10	0.00

.....A-T-ACATTATTGAAGGAATTGATAT.....GTATTTTAATTCTTCATAAA.....	25	20	0.35	0.01
.....GTATTTTAATTCTTCATAAA.....GTATTTTAATTCTTCATAAA.....	24	20	0.35	0.01
.....GTATTTTAATTCTTCATAAA.....GTATTTTAATTCTTCATAAA.....	23	20	0.35	0.01
.....A-T-ACATTATTGAAGGAATT.....TATTTTAATTCTTCATAAA.....	20	20	0.35	0.01
.....TGAAGGAATTGATATATG.....TATTTTAATTCTTCATAAA.....	23	20	0.35	0.01
.....TGAAGGAATTGATATATGCCAGTA.....TATTTTAATTCTTCATAAA.....	18	20	0.30	0.01
.....TATTTTAATTCTTCATAAA.....GTATTTTAATTCTTCATAAA.....	24	20	0.30	0.01
.....ATTGAAAGGAATTGATATATGCCAGT.....GTATTTTAATTCTTCATAAA.....	21	20	0.30	0.01
.....ATTGAAAGGAATTGATATATGCCAGT.....ATTGAAAGGAATTGATATATGCCAGT.....	20	20	0.30	0.01
.....TTTATTGAAGGAATTGATATAT.....ATTGAAAGGAATTGATATATGCCAGT.....	23	20	0.30	0.01
.....ATTGAAAGGAATTGATATATGCCAGT.....TTTATTGAAGGAATTGATATAT.....	20	20	0.30	0.01
.....ATTGAAAGGAATTGATATATGCCAGT.....ATTGAAAGGAATTGATATATGCCAGT.....	23	20	0.30	0.01
.....TATTGAAGGAATTGATATATGC.....TATTGAAGGAATTGATATATGC.....	20	20	0.30	0.01
.....ATAAAAAA-T-ACATTATTGAAGGA.....ATTTTTAATTCTTCATAAA.....	23	20	0.30	0.01
.....TTATTGAAGGAATTGATATATGTC.....ATTTTTAATTCTTCATAAA.....	20	20	0.30	0.01
.....TATTGAAGGAATTGATATATGTC.....TGTATTTTAATTCTTCATAAA.....	23	20	0.25	0.01
.....ATTTATTGAAGGAATTGATATATGCCA.....ATTTTTAATTCTTCATAAA.....	21	20	0.25	0.01
.....ATTGAAAGGAATTGATATATGCCA.....TGTATTTTAATTCTTCATAAA.....	25	20	0.25	0.01
.....ATTGAAAGGAATTGATATATGCCA.....ATTGAAAGGAATTGATATATGCCA.....	27	20	0.25	0.01
.....ATTGAAAGGAATTGATATATGCCA.....ATTGAAAGGAATTGATATATGCCA.....	19	20	0.25	0.01
.....ATTGAAAGGAATTGATATATGCCA.....ATTGAAAGGAATTGATATATGCCA.....	23	20	0.25	0.01
.....TTGAAAGGAATTGATATATGTC.....ATTGAAAGGAATTGATATATGTC.....	20	20	0.20	0.01
.....ATTGAAAGGAATTGATATATGTC.....ATTGAAAGGAATTGATATATGTC.....	18	20	0.20	0.00
.....ATTGAAAGGAATTGATATATGTC.....ATAAAAACAAAATTGAC-.....	20	20	0.20	0.02
.....TGTATTTTAATTCTTCATAAA.....ATT.....	20	20	0.20	0.01
.....TATTGAAGGAATTGATATATGCCA.....TGTATTTTAATTCTTCATAAA.....	25	20	0.20	0.00
.....TTATTGAAGGAATTGATATATGCCA.....TGTATTTTAATTCTTCATAAA.....	27	20	0.20	0.00
.....TGTATTTTAATTCTTCATAAA.....TGTATTTTAATTCTTCATAAA.....	20	20	0.20	0.01
.....TGTATTTTAATTCTTCATAAA.....TGTATTTTAATTCTTCATAAA.....	25	20	0.20	0.01
.....ATAAAAAA-T-ACATTATTGAAG.....ATAAAAAA-T-ACATTATTGAAG.....	21	20	0.20	0.00
.....A-T-ACATTATTGAAGGAATTG.....ACATTATTGAAGGAATTGATAT.....	21	20	0.20	0.00
.....ACATTATTGAAGGAATTGATAT.....TGATATATGCCAGTAAAATG.....	23	20	0.20	0.00
.....TGATATATGCCAGTAAAATG.....ATTGATATATGCCAGTAAAATG.....	20	20	0.20	0.01
.....ATTGATATATGCCAGTAAAATG.....T-ACATTATTGAAGGAATTGATA.....	22	20	0.20	0.00
.....T-ACATTATTGAAGGAATTGATA.....T-ACATTATTGAAGGAATTGATA.....	23	20	0.15	0.00
.....ATAAAAAA-T-ACATTATTGAAGGA.....TTGAAAGGAATTGATATATGCCA.....	25	20	0.15	0.01
.....TTGAAAGGAATTGATATATGCCA.....TTTATTGAAGGAATTGATATA.....	23	20	0.15	0.00
.....TTTATTGAAGGAATTGATATA.....TTTATTGAAGGAATTGATATA.....	21	20	0.15	0.01
.....TTTATTGAAGGAATTGATATA.....TTTATTGAAGGAATTGATATA.....	22	20	0.15	0.00
.....TTTATTGAAGGAATTGATATA.....CAGTAAAATGGGTATTTTAAT.....	23	20	0.15	0.00
.....CAGTAAAATGGGTATTTTAAT.....TGAAGGAATTGATATATGCCAGTAA.....	26	20	0.15	0.01
.....TGAAGGAATTGATATATGCCAGTAA.....TGTATTTTAATTCTTCATAAT.....	23	20	0.15	0.00
.....TGTATTTTAATTCTTCATAAT.....TATTGAAGGAATTGATATATGCCAGTAA.....	28	20	0.15	0.01
.....TATTGAAGGAATTGATATATGCCAGTAA.....TGTATTTTAATTCTTCATAATAAAA.....	28	20	0.15	0.00
.....TGTATTTTAATTCTTCATAATAAAA.....CAGTAAAATGGGTATTTTT.....	20	20	0.15	0.01
.....CAGTAAAATGGGTATTTTT.....TTGAAGGAATTGATATATGCCAGTAA.....	26	20	0.15	0.00
.....TTGAAGGAATTGATATATGCCAGTAA.....GAAGGAATTGATATATGCCA.....	20	20	0.15	0.00
.....GAAGGAATTGATATATGCCA.....GAAGGAATTGATATATGCCAGTAAAAT.....	27	20	0.15	0.01
.....GAAGGAATTGATATATGCCAGTAAAAT.....TATTGAAGGAATTGATATATGCCA.....	25	20	0.15	0.00
.....TATTGAAGGAATTGATATATGCCA.....TATTGAAGGAATTGATATATGCCA.....	23	20	0.15	0.01
.....A-T-ACATTATTGAAGGAATTGATA.....A-T-ACATTATTGAAGGAATTGATA.....	24	20	0.15	0.01
.....ATAAAAAA-T-ACATTATTGAAGGA.....TTTATTGAAGGAATTGATATATGCC.....	26	20	0.15	0.00
.....TTTATTGAAGGAATTGATATATGCC.....CATTTATTGAAGGAATTGATATATGCC.....	25	20	0.15	0.00
.....CATTTATTGAAGGAATTGATATATGCC.....TTATTGAAGGAATTGATATATGTC.....	27	20	0.15	0.01
.....TTATTGAAGGAATTGATATATGTC.....TTATTGAAGGAATTGATATATGTC.....	24	20	0.15	0.03
.....TTATTGAAGGAATTGATATATGTC.....GATATATGCCAGTAAAATGGGTAA.....	24	20	0.15	0.00
.....GATATATGCCAGTAAAATGGGTAA.....ATTGAAGGAATTGATATATGCA.....	22	20	0.15	0.00
.....ATTGAAGGAATTGATATATGCA.....TATTGAAGGAATTGATATATGCC.....	24	20	0.15	0.01
.....TATTGAAGGAATTGATATATGCC.....ATTATTATGCCAGTAAAATG.....	18	20	0.15	0.01
.....ATTATTATGCCAGTAAAATG.....ATTATTATGCCAGTAAAATG.....	28	20	0.15	0.01
.....ATTATTATGCCAGTAAAATG.....GATATATGCCAGTAAAATG.....	18	20	0.15	0.05
.....GATATATGCCAGTAAAATG.....GTGTATTTTAATTCTTCATAAT.....	19	20	0.15	0.00
.....GTGTATTTTAATTCTTCATAAT.....TTTATTGAAGGAATTGATATATGCA.....	23	20	0.15	0.00
.....TTTATTGAAGGAATTGATATATGCA.....TTTATTGAAGGAATTGATATATGCA.....	21	20	0.10	0.01
.....TTTATTGAAGGAATTGATATATGCA.....TTTATTGAAGGAATTGATATATGCA.....	18	20	0.10	0.00

dme-mir-10404 Complete						
.....AA-T-ACATTATTGAAGGAATT.....TATTTTAATTCTTCAATAA.....	21	20	0.10	0.00	
.....ATTATTGAAGGAATTGATATG.....TATTTTAATTCTTCAATAA.....	22	20	0.10	0.00	
.....ACATTATTGAAGGAATTGATATA.....TATTTTAATTCTTCAATAA.....	22	20	0.10	0.01	
.....A-T-ACATTATTGAAGGAATTGATATA.....TATTTTAATTCTTCAATAA.....	24	20	0.10	0.00	
.....CATTTATTGAAGGAATTGATAT.....TGTTGTATTTTAATTCTTCAATAAT.....	27	20	0.10	0.03	
.....GGTGTATTTTAATTCTTCAATAAAA.....TGTTGTATTTTAATTCTTCAATAAT.....	22	20	0.10	0.00	
.....TTGAAGGAATTGATATATGCCA.....TTGAAGGAATTGATATATGCCA.....	28	20	0.10	0.02	
.....ATTTTTAATTCTTCAATAAAA.....ATTTTTAATTCTTCAATAAAA.....	27	20	0.10	0.00	
.....CAGTAAAAATGGTGTATTTTAATTTC.....CAGTAAAAATGGTGTATTTTAATTTC.....	23	20	0.10	0.00	
.....TTGAAGGAATTGATATATGCC.....TTGAAGGAATTGATATATGCC.....	26	20	0.10	0.01	
.....AAAA-T-ACATTATTGAAGGAATTGATA.....AAAA-T-ACATTATTGAAGGAATTGATA.....	22	20	0.10	0.01	
.....TATTGAAGGAATTGATATATGC.....TATTGAAGGAATTGATATATGC.....	28	20	0.10	0.01	
.....ATTTTTAATTCTTCAATAAA.....ATTTTTAATTCTTCAATAAA.....	24	20	0.10	0.00	
.....TTATTGAAGGAATTGATATATG.....TTATTGAAGGAATTGATATATG.....	21	20	0.10	0.00	
.....ATTTATTGAAGGAATTGATATATGCCAGT.....ATTTATTGAAGGAATTGATATATGCCAGT.....	23	20	0.10	0.00	
.....AA-T-ACATTATTGAAGGAATTGATA.....AA-T-ACATTATTGAAGGAATTGATA.....	29	20	0.10	0.02	
.....TATTGAAGGAATTGATATATGC.....TATTGAAGGAATTGATATATGC.....	25	20	0.10	0.00	
.....AA-T-ACATTATTGAAGGAATTGATA.....AA-T-ACATTATTGAAGGAATTGATA.....	24	20	0.10	0.02	
.....TGTATTTTAATTCTTCAATAAAA.....TGTATTTTAATTCTTCAATAAAA.....	26	20	0.10	0.01	
.....ACATTATTGAAGGAATTG.....ACATTATTGAAGGAATTG.....	19	20	0.10	0.00	
.....GTGTATTTTAATTCTTCAATAAAA.....GTGTATTTTAATTCTTCAATAAAA.....	27	20	0.10	0.00	
.....ATTGAAGGAATTGATATATG.....ATTGAAGGAATTGATATATG.....	21	20	0.10	0.01	
.....CATTTATTGAAGGAATTGATA.....CATTTATTGAAGGAATTGATA.....	21	20	0.10	0.00	
.....TTTATTGAAGGAATTGATATATGCCAG.....TTTATTGAAGGAATTGATATATGCCAG.....	27	20	0.10	0.01	
.....TATTGAAGGAATTGATATATGCCAGTA.....TATTGAAGGAATTGATATATGCCAGTA.....	27	20	0.10	0.00	
.....ATTGAAGGAATTGATATAT.....ATTGAAGGAATTGATATAT.....	19	20	0.10	0.00	
.....TTGAAGGAATTGATATATGCCAG.....TTGAAGGAATTGATATATGCCAG.....	24	20	0.10	0.01	
.....TAGGTATAAAAA-T-ACATTTA.....TAGGTATAAAAA-T-ACATTTA.....	20	20	0.10	0.00	
.....CATTTATTGAAGGAATTG.....CATTTATTGAAGGAATTG.....	18	20	0.10	0.00	
.....TAAAAAA-T-ACATTTATTGAAGGAATT.....TAAAAAA-T-ACATTTATTGAAGGAATT.....	25	20	0.10	0.00	
.....CATTTATTGAAGGAATTGATATA.....CATTTATTGAAGGAATTGATATA.....	23	20	0.10	0.00	
.....GTATTTTAATTCTTCAATAAA.....GTATTTTAATTCTTCAATAAA.....	24	20	0.10	0.00	
Total						800.80 7.59

droAna3 dan 4035 confiden

.....TGAAGGAATTGATATATGCCA.....	21	20	2.10	0.03
.....ATTTATTGAAGGAATTGATATATGCC.....	26	20	2.10	0.05
.....TATTGAAGGAATTGATATATGCC T	24	20	1.95	0.04
.....TGTATTTTAATTCTTCAAA C	22	20	1.95	0.08
.....ATTGAAGGAATTGATATATGCCA.....	23	20	1.80	0.03
.....TTATTGAAGGAATTGATATA.....	20	20	1.75	0.04
.....AAATTATTGAAGGAATTGATATATGCC.....	28	20	1.65	0.06
.....ATTATTGAAGGAATTGATATA TG	25	20	1.65	0.02
.....TTTATTGAAGGAATTGATATATGCCAGT.....	28	20	1.50	0.02
.....TTTATTGAAGGAATTGATATA TG	24	20	1.40	0.03
.....TGAAGGAATTGATATATGC.....	19	20	1.35	0.02
.....GAAGGAATTGATATATGC.....	18	20	1.30	0.02
.....AATTATTGAAGGAATTGATATATGCC.....	27	20	1.15	0.04
.....TTTATTGAAGGAATTGATATA.....	21	20	1.10	0.05
.....TTATTGAAGGAATTGATATA A	20	20	0.90	0.02
.....AAATATAATGACGAATTGT.....	24	20	0.90	0.03
.....TGTATTTTAATTCTTCAAA C	25	20	0.85	0.03
.....GTATTTTAATTCTTCAATAA.....	23	20	0.85	0.01
.....TTTATTGAAGGAATTGATATATGCCA.....	26	20	0.85	0.01
.....TTGAAGGAATTGATATATGCC T	22	20	0.85	0.04
.....TTGAAGGAATTGATATATGCCAGTT.....	25	20	0.80	0.01
.....GTATTTTAATTCTTCAATAAA.....	24	20	0.75	0.01
.....ATTGAAGGAATTGATATA.....	19	20	0.75	0.01
.....TGAAGGAATTGATATATGCCAGTT.....	24	20	0.75	0.01
.....TATTGAAGGAATTGATATA TG	23	20	0.70	0.01
.....TATTGAAGGAATTGATATA C	21	20	0.70	0.02
.....TTATTGAAGGAATTGATAT.....	19	20	0.70	0.01
.....TTGAAGGAATTGATATA TG	21	20	0.70	0.01
.....T-AATTATTGAAGGAATTGATATATGCC.....	29	20	0.65	0.03
.....ATTGAAGGAATTGATATA A	22	20	0.60	0.02
.....TATTGAAGGAATTGATAT T	19	20	0.60	0.02
.....TTATTGAAGGAATTGATATATGCCAGC.....	27	20	0.60	0.02
.....AAAAATATAATGACGAATT.....	20	20	0.60	0.02
.....TATTGAAGGAATTGATATATGCCA.....	24	20	0.50	0.02
.....ATTATTGAAGGAATTGATA.....	20	20	0.50	0.03
.....TATTTTTAATTCTTCAATA.....	21	20	0.50	0.01
.....GAAGGAATTGATATATGCCA.....	20	20	0.50	0.01
.....ATAAAAATATAATGACGAATTGT.....	24	20	0.45	0.01
.....TTATTGAAGGAATTGATATATGCCAGTT.....	28	20	0.45	0.01
.....TTTATTGAAGGAATTGATATA TG	23	20	0.45	0.02
.....TTATTGAAGGAATTGATATATGCCA.....	25	20	0.45	0.01
.....TTGATATATGCCAGTTAACGGTGT.....	26	20	0.45	0.02
.....ATAAAAATATAATGACGAATTGTAT.....	26	20	0.40	0.02
.....AATATAATGACGAATTGT.....	19	20	0.40	0.01
.....TTATTGAAGGAATTGATATA TG	26	20	0.40	0.01
.....TTATTGAAGGAATTGATATA C	21	20	0.35	0.01
.....AAAAATATAATGACGAATT.....	28	20	0.35	0.01
.....TATTGAAGGAATTGATATATGCCAGC.....	19	20	0.35	0.01
.....TATTGAAGGAATTGATATATGCCAGC C	26	20	0.35	0.01
.....ATAAAAATATAATGACGAAT.....	26	20	0.35	0.01
.....AA-T-AAATTATTGAAGGAATT.....	21	20	0.35	0.01
.....TTTATTGAAGGAATTGATATATGCCAGC.....	28	20	0.35	0.01
.....AAAAATATAATGACGAAT.....	19	20	0.35	0.01
.....TATTGAAGGAATTGATATATGCCAGC.....	26	20	0.35	0.01
.....ATAAAA-T-AAATTATTGAAGGA.....	23	20	0.35	0.01
.....TATTGAAGGAATTGATAT GC	20	20	0.35	0.01
.....T-AAATTATTGAAGGAATTGATA.....	25	20	0.30	0.01
.....TATTGAAGGAATTGATATATGCCAGTTA.....	22	20	0.30	0.02
.....ATTATTGAAGGAATTGATA.....	28	20	0.30	0.01
.....GTATTTTAATTCTTCAAT.....	22	20	0.30	0.01
.....AAAATATAATGACGAATTGT.....	21	20	0.30	0.01
.....TATTTTTAATTCTTCAATAAA.....	23	20	0.30	0.01
.....ATTGATATA TGCCAGTTAAATG	22	20	0.30	0.01
.....AATTGATATATGCCAGTTAACGGTGT.....	28	20	0.30	0.01
.....AAAAA-T-AAATTATTGAAGGAATT.....	24	20	0.30	0.01
.....TGAAGGAATTGATATATGCCAGTTA.....	26	20	0.30	0.01
.....ATAAAATGACGAATTGTAT.....	18	20	0.25	0.01
.....ATAAAATGACGAATTGTATAT-AT-----T.....	23	20	0.25	0.01
.....GATATATGCCAGTTAACGGTGTAT.....	25	20	0.25	0.01

	TGTATTTTAATTCTTCA.	20	20	0.25	0.01
..TTGAAGGAATTGATATATGT	..	20	20	0.25	0.01
..TGAAGGAATTGATATATGCCAGTTAA	..	27	20	0.25	0.00
..	TATTTTTAATTCTTCAAT	20	20	0.25	0.00
..ATTTATTGAAGGAATTGATAT	..	21	20	0.25	0.01
..TATTGAAGGAATTGATATGCCAG	..	25	20	0.25	0.00
..TTATTGAAGGAATTGATATGT	..	22	20	0.25	0.01
..A-T-AAATTATTGAAGGAATTGATA	..	23	20	0.25	0.01
..	ATAAAAATATAATGACGAATT	24	20	0.25	0.01
..TTTATTGAAGGAATTGATA	..	22	20	0.25	0.01
..TATTGAAGGAATTGATATGCCAGTT	..	19	20	0.25	0.01
..AAATTATTGAAGGAATTGATATGC	..	27	20	0.25	0.01
..	AGTTAAATGGTGTATTTTAATT	27	20	0.25	0.01
..TTGAAGGAATTGATATGCCAAA	..	23	20	0.25	0.01
..TTTATTGAAGGAATTGATAT	..	18	20	0.25	0.01
..AAA-T-AAATTATTGAAGGAATT	..	25	20	0.25	0.01
..TATTGAAGGAATTGATATGCCAG	..	20	20	0.25	0.01
..ATTGAAGGAATTGATATGCCAGT	..	24	20	0.20	0.01
..TTGAAGGAATTGATATGCCAGT	..	23	20	0.20	0.00
..TAAAAA-T-AAATTATTGAAGGAATTGA	..	27	20	0.20	0.01
..	ATATAATGACGAATTGTAT	20	20	0.20	0.01
..TATTGAAGGAATTGATATGCCG	..	24	20	0.20	0.01
..	TGTATTTTAATTCTTCAATAA	23	20	0.20	0.01
..T-AAATTATTGAAGGAATTGATATATGC	..	19	20	0.20	0.01
..	TTTTAATTCTTCAATA	28	20	0.20	0.01
..ATTGAAGGAATTGATATGCCAGT	..	18	20	0.20	0.01
..TTGAAGGAATTGATATGCCAGT	..	25	20	0.20	0.01
..TGAAGGAATTGATATGCCAG	..	24	20	0.20	0.01
..GATATATGCCAGTTAAATG	..	21	20	0.20	0.00
..ATTGATATATGCCAGTTAAAT	..	25	20	0.20	0.00
..GATATATGCCAGTTAAAT	..	23	20	0.20	0.01
..	CAGTTAAATGGTGTATTTTAATT	22	20	0.20	0.01
..TTGAAGGAATTGATATATGCCAGTTAA	..	19	20	0.20	0.00
..ATTTTAGGTATAAAAA-T-AAATT	..	28	20	0.20	0.01
..GTATAAAAA-T-AAATTATTGAAGGA	..	22	20	0.15	0.01
..TTGAAGGAATTGATATATGA	..	25	20	0.15	0.01
..	AAAAATATAATGACGAATTG	20	20	0.15	0.01
..AAAA-T-AAATTATTGAAGGAATTGATA	..	22	20	0.15	0.01
..TTATTGAAGGAATTGATATATGCCAGTTA	..	27	20	0.15	0.00
..TTGAAGGAATTGATATGCCAGT	..	29	20	0.15	0.01
..	GTATTTTAATTCTTCA	24	20	0.15	0.01
..TTGAAGGAATTGATATGCCAG	..	19	20	0.15	0.01
..	AATATAATGACGAATTGTAT	23	20	0.15	0.01
..ATTTATTGAAGGAATTGATATAT	..	21	20	0.15	0.01
..	AAATGGTGTATTTTAATTTC	23	20	0.15	0.01
..ATTGAAGGAATTGATATATGA	..	22	20	0.15	0.01
..GAAGGAATTGATATATGCCAGT	..	21	20	0.15	0.00
..TATTGAAGGAATTGATATG	..	22	20	0.15	0.00
..	ATTTTTAATTCTTCAATAAA	19	20	0.15	0.00
..TATGCCAGTTAAATGGTGA	..	22	20	0.15	0.01
..GATATATGCCAGTTAAATGGTGA	..	24	20	0.15	0.01
..GATATATGCCAGTTAAATGGTGA	..	25	20	0.15	0.01
..GATATATGCCAGTTAAATGGTGA	..	22	20	0.15	0.01
..AAATTTATTGAAGGAATTGATA	..	22	20	0.15	0.00
Total		759.25	7.76		

Total

.....TTGAAGGAATTGATATATGCC.....	21	2	843.00	6.21
.....TTGAAGGAATTGATATATGCCA.....	22	2	359.00	2.65
.....TTGAAGGAATTGATATATGC.....	20	2	286.00	2.11
.....ATTGAAGGAATTGATATATGC.....	21	2	229.00	1.69
.....TATTGAAGGAATTGATATATGCCA.....	24	2	163.50	1.30
.....TTATTGAAGGAATTGATATATGC.....	23	2	147.50	1.29
.....TGTATTTTAATTCTTCAATAAA.....	25	2	117.00	1.00
.....TGTATTTTAATTCTTCAATA.....	23	2	113.00	0.91
.....ATTGAAGGAATTGATATATGCC.....	22	2	106.50	0.86
.....TATTGAAGGAATTGATATAT.....	20	2	99.00	0.81
.....TTGAAGGAATTGATATATG.....	19	2	97.00	0.75
.....ATTGAAGGAATTGATATATG.....	20	2	87.00	0.76
.....TGTATTTTAATTCTTCAAT.....	22	2	82.00	0.66
.....TATTGAAGGAATTGATATATGCCAG.....	25	2	73.00	0.65
.....TATTGAAGGAATTGATATATGCC <ins>T</ins>	24	2	60.50	0.49
.....TATTGAAGGAATTGATATATGCA <ins>A</ins>	23	2	59.50	0.51
.....TATTGAAGGAATTGATATATG <ins>T</ins>	22	2	56.00	0.49
.....TGAAGGAATTGATATATGCCA.....	21	2	53.50	0.51
.....TGTATTTTAATTCTTCAATAAA.....	24	2	49.00	0.43
.....TGTATTTTAATTCTTCAATAAA <ins>A</ins>	26	2	43.50	0.48
.....TTATTGAAGGAATTGATATATG.....	22	2	42.00	0.31
.....TTATTGAAGGAATTGATATATGCC.....	24	2	41.50	0.41
.....TTATTGAAGGAATTGATATATGCCAG.....	26	2	41.00	0.41
.....TGAAGGAATTGATATATG.....	19	2	36.50	0.40
.....TGAAGGAATTGATATATGCC.....	20	2	33.00	0.34
.....TTTATTGAAGGAATTGATATATG.....	24	2	29.00	0.32
.....GTATTTTAATTCTTCAATA.....	22	2	26.50	0.28
.....TATTGAAGGAATTGATATA.....	19	2	26.50	0.25
.....TATTGAAGGAATTGATATATG <ins>A</ins>	22	2	23.50	0.22
.....GTATTTTAATTCTTCAATAAA.....	24	2	23.50	0.22
.....ATTTATTGAAGGAATTGATATATG.....	25	2	23.50	0.25
.....GTATTTTAATTCTTCAAT.....	21	2	21.00	0.18
.....TATTTTTAATTCTTCAATAAA.....	23	2	17.50	0.17
.....TTTATTGAAGGAATTGATATATGCCAG.....	20	2	15.00	0.16
.....AATTTATTGAAGGAATTGATATATG.....	27	2	14.50	0.14
.....ATTTTTAATTCTTCAATA.....	26	2	13.50	0.56
.....TAAAATTATTGAAGGAATTGATA.....	20	2	13.00	0.13
.....TATTGAAGGAATTGATATATGCCAGT.....	24	2	13.00	0.15
.....TTATTGAAGGAATTGATATATGCCA.....	26	2	12.00	0.12
.....TTTATTGAAGGAATTGATATATGCC.....	25	2	12.00	0.13
.....TTTATTGAAGGAATTGATATATGCC.....	25	2	12.00	0.14
.....GAAGGAATTGATATATGCC.....	24	2	11.50	0.13
.....TATTTTTAATTCTTCAATA.....	19	2	11.00	0.13
.....TGTATTTTAATTCTTCAATAAA <ins>C</ins>	21	2	11.00	0.12
.....AAAGACATATCTGAC--AAATAAATGAC.....	25	2	11.00	0.11
.....ATTTTTAATTCTTCAATAAA.....	19	2	10.50	0.13
.....TATTTTTAATTCTTCAATAAA.....	22	2	10.50	0.13
.....GAAGGAATTGATATATG.....	22	2	10.00	0.09
.....AAAGACATATCTGAC--AAATAAATGA.....	18	2	10.00	0.10
.....TATTGAAGGAATTGATATATG <ins>T</ins>	25	2	9.50	0.11
.....ATTGAAGGAATTGATATATGCCA.....	23	2	9.50	0.08
.....GTATTTTAATTCTTCAATAAA.....	23	2	9.50	0.11
.....TTGAAGGAATTGATATATGCC <ins>T</ins>	23	2	9.00	0.10
.....TGTATTTTAATTCTTCAAA.....	22	2	9.00	0.09
.....TGAAGGAATTGATATATGCCAG.....	22	2	9.00	0.09
.....AAAGACATATCTGAC--AAATAAATGAA.....	27	2	8.00	0.09
.....ATAAAAGACATATCTGAC--AAATA.....	22	2	8.00	0.10
.....TTATTGAAGGAATTGATATA.....	20	2	7.50	0.08
.....ATAAAAGACATATCTGAC--AAATAAATGA.....	27	2	7.00	0.07
.....AAAGACATATCTGAC--AAATAAAT.....	23	2	6.50	0.08
.....ATTGATATATGCCAGTAAAATG.....	22	2	6.50	0.07
.....TTGAAGGAATTGATATATGCA <ins>A</ins>	21	2	6.00	0.07
.....ATTGAAGGAATTGATATATGCCAG.....	24	2	6.00	0.07
.....ATTGAAGGAATTGATATATG <ins>T</ins>	21	2	6.00	0.06
.....TATTTTTAATTCTTCAAT.....	20	2	6.00	0.07
.....TTGAAGGAATTGATATATG <ins>T</ins>	20	2	5.50	0.06

.....TATTGAAGGAATTGATATATGCCG.	TTTTTAATTCTTCAATAAA.	21	2	5.50	0.07
.....TATTGAAGGAATTGATATATGCCAGTATGTTTTAATTCTTCAATTC.	24	2	5.50	0.06
.....TTATTGAAGGAATTGATATATGCCAGT.AAGACATATCTGAC--AAATAAATGAAT	27	2	5.50	0.07
.....GATATATGCCAGTAAAATG.AAAGACATATCTGAC--AAATAAATGAAT	23	2	5.50	0.07
.....TAAAATTATTGAAGGAATTGATATA.AAAGACATATCTGAC--AAATAAATGAAT	27	2	5.00	0.06
.....TATTGAAGGAATTGATATATGCCA[A]AAAGACATATCTGAC--AAATAAATGAAT	19	2	5.00	0.07
.....TGTATTTTAATTCTTCAATAA[T]AAAGACATATCTGAC--AAATAAATGAAT	26	2	5.00	0.05
.....AAATTATTGAAGGAATTGATATATGCAAAGACATATCTGAC--AAATAAATGAAT	25	2	5.00	0.06
.....AAATTATTGAAGGAATTGATATATGCAAAGACATATCTGAC--AAATAAATGAAT	24	2	4.50	0.05
.....TTATTGAAGGAATTGATATATGCAAAGACATATCTGAC--AAATAAATGAAT	22	2	4.50	0.38
.....ATTATTGAAGGAATTGATATATGCC.AAAGACATATCTGAC--AAATAAATGAAT	21	2	4.00	0.06
.....TTGAAGGAATTGATATATGCCAGTAAAAGACATATCTGAC--AAATAAATGAAT	27	2	4.00	0.04
.....ATTGAAGGAATTGATATATGCC[T]AAAGACATATCTGAC--AAATAAATGAAT	27	2	4.00	0.25
.....CAGTAAAATGGGTATTTTAATAAAGACATATCTGAC--AAATAAATGAAT	22	2	4.00	0.05
.....TGAAGGAATTGATATATGAAAGACATATCTGAC--AAATAAATGAAT	25	2	3.50	0.08
.....ATTGAAGGAATTGATATATAAAGACATATCTGAC--AAATAAATGAAT	23	2	3.50	0.05
.....TTGAAGGAATTGATATATGCCAGT.AAAGACATATCTGAC--AAATAAATGAAT	19	2	3.50	0.05
.....ATTGAAGGAATTGATATATGCCA[A]AAAGACATATCTGAC--AAATAAATGAAT	23	2	3.50	0.04
.....TGTATTTTAATTCTTCAATAA[A]AAAGACATATCTGAC--AAATAAATGAAT	23	2	3.50	0.04
.....TTTAATTCTTCAATAAA.AAAGACATATCTGAC--AAATAAATGAAT	20	2	3.50	0.04
.....TATTTTAATTCTTCAATAA[C]AAAGACATATCTGAC--AAATAAATGAAT	19	2	3.50	0.05
.....TATTGAAGGAATTGATATATGCCG.AAAGACATATCTGAC--AAATAAATGAAT	24	2	3.50	0.04
.....TATTGAAGGAATTGATATATGCAAAGACATATCTGAC--AAATAAATGAAT	20	2	3.50	0.04
.....ATTGATATATGCCAGTAAAATGGTGTAAAGACATATCTGAC--AAATAAATGAAT	27	2	3.50	0.04
.....GTATTTTAATTCTTCAATAAA[A]AAAGACATATCTGAC--AAATAAATGAAT	25	2	3.50	0.05
.....TTTTAATTCTTCAATAAA.AAAGACATATCTGAC--AAATAAATGAAT	20	2	3.50	0.04
.....TATTTTAATTCTTCAATAA[C]AAAGACATATCTGAC--AAATAAATGAAT	23	2	3.50	0.04
.....TATTGAAGGAATTGATATATGCCA[A]AAAGACATATCTGAC--AAATAAATGAAT	24	2	3.00	0.04
.....TTTTAATTCTTCAATAAA.AAAGACATATCTGAC--AAATAAATGAAT	20	2	3.00	0.03
.....TATTGAAGGAATTGATATATGCAAAGACATATCTGAC--AAATAAATGAAT	22	2	3.00	0.04
.....ATTGAAGGAATTGATATATGCAAAGACATATCTGAC--AAATAAATGAAT	25	2	3.00	0.03
.....ATTTTAATTCTTCAATAAA[A]AAAGACATATCTGAC--AAATAAATGAAT	18	2	3.00	0.04
.....TTCTTCATAAAAGACATATCTGAC--AAAGACATATCTGAC--AAATAAATGAAT	22	2	3.00	0.04
.....TTGAAGGAATTGATATATGCCA[A]AAAGACATATCTGAC--AAATAAATGAAT	23	2	3.00	0.04
.....TGAAGGAATTGATATATGCCAGTAAAAGACATATCTGAC--AAATAAATGAAT	24	2	3.00	0.04
.....TATTTTAATTCTTCAATAAA[A]AAAGACATATCTGAC--AAATAAATGAAT	24	2	3.00	0.04
.....TGATATATGCCAGTAAAATGAAAGACATATCTGAC--AAATAAATGAAT	20	2	3.00	0.04
.....TTATTGAAGGAATTGATATATGCCAGTAAAAGACATATCTGAC--AAATAAATGAAT	28	2	3.00	0.04
.....ATAAAAT-TAAAATTATTGAAGGA.AAAGACATATCTGAC--AAATAAATGAAT	29	2	2.50	0.03
.....TGAAGGAATTGATATATGCCAGTA.AAAGACATATCTGAC--AAATAAATGAAT	24	2	2.50	0.21
.....TTGAAGGAATTGATATATGCCAG.AAAGACATATCTGAC--AAATAAATGAAT	26	2	2.50	0.15
.....TTTATTGAAGGAATTGATATATGAAAGACATATCTGAC--AAATAAATGAAT	23	2	2.50	0.04
.....TTGAAGGAATTGATATATGCCAGTAAA.AAAGACATATCTGAC--AAATAAATGAAT	20	2	2.50	0.03
.....TTTATTGAAGGAATTGATATATGCCA.AAAGACATATCTGAC--AAATAAATGAAT	28	2	2.50	0.03
.....TTGAAGGAATTGATATATGCCAGTA.AAAGACATATCTGAC--AAATAAATGAAT	26	2	2.50	0.03
.....TATTGAAGGAATTGATATATGCCAGTA.AAAGACATATCTGAC--AAATAAATGAAT	25	2	2.50	0.03
.....AAAGACATATCTGAC--AAATAAATGAATAAAGACATATCTGAC--AAATAAATGAAT	28	2	2.50	0.04
.....TATTGAAGGAATTGATATATGAAAGACATATCTGAC--AAATAAATGAAT	24	2	2.50	0.04
.....TATTGAAGGAATTGATATATGAAAGACATATCTGAC--AAATAAATGAAT	21	2	2.50	0.16
.....ATTGATATATGCCAGTAAAATGGTGTAAAGACATATCTGAC--AAATAAATGAAT	28	2	2.50	0.03
.....GTATTTTAATTCTTCAATAAA[A]AAAGACATATCTGAC--AAATAAATGAAT	24	2	2.50	0.03
.....GTATTTTAATTCTTCAAA.AAAGACATATCTGAC--AAATAAATGAAT	20	2	2.50	0.04
.....AAAATTATTGAAGGAATTGATATATGCAAAGACATATCTGAC--AAATAAATGAAT	28	2	2.00	0.03
.....AAAGACATATCTGAC--AAATAAATGAATAAAGACATATCTGAC--AAATAAATGAAT	18	2	2.00	0.02
.....AAAGACATATCTGAC--AAATAAATGAATAAAGACATATCTGAC--AAATAAATGAAT	20	2	2.00	0.02

.....TTTATTGAAGGAATTGATATATG.....	23	2	2.00	0.02
.....AAGACATATCTGAC--AAATAAATGAAT---A.....	27	2	2.00	0.03
.....TATTGAAGGAATTGATATATGCCCT.....	25	2	2.00	0.03
.....TTGAAGGAATTGATATATGCC.....	22	2	2.00	0.09
.....TTGTATATGCCAGTAAAATG.....	21	2	2.00	0.03
.....AAGACATATCTGAC--AAATAAATG.....	23	2	2.00	0.03
.....TTATTGAAGGAATTGATAT.....	19	2	2.00	0.03
.....TTTATTGAAGGAATTGATATGCCAGT.....	28	2	2.00	0.03
.....AAAAT-TAAAATTATTGAAGGAATT.....	25	2	2.00	0.17
.....AGACATATCTGAC--AAATAAATGA.....	23	2	2.00	0.03
.....AATTATTGAAGGAATTGATATATGCC.....	27	2	2.00	0.17
.....TTGAAGGAATTGATATATGCCAGTAAA.....	27	2	2.00	0.02
.....TGTATTTTAATTCTTCAATAA.....	24	2	2.00	0.02
.....AAAT-TAAAATTATTGAAGGA.....	21	2	2.00	0.08
.....T-TAAAATTATTGAAGGAATTGATA.....	25	2	2.00	0.02
.....ATAAAAT-TAAAATTATTGAAGGAATT.....	27	2	2.00	0.17
.....TCAATAAGACATATCTGA.....	19	2	2.00	0.02
.....GAAGGAATTGATATATGCCA.....	20	2	2.00	0.03
.....GATATATGCCAGTAAAATGGTGT.....	23	2	2.00	0.03
.....TGTATTTTAATTCTTCAAA.....	22	2	2.00	0.03
.....AAGACATATCTGAC--AAATAAA.....	20	2	2.00	0.03
.....TATTGAAGGAATTGATATATGCCAA.....	26	2	2.00	0.02
.....TGAAGGAATTGATATATGCCAGT.....	23	2	2.00	0.03
.....TAAAAT-TAAAATTATTGAAGGAAT.....	25	2	2.00	0.03
.....TTTTAATTCTTCAATAA.....	21	2	1.50	0.13
.....GACATATCTGAC--AAATAAATGAAT---CG.....	26	2	1.50	0.13
.....TAAAGACATATCTGAC--AAATAAATGAAT.....	28	2	1.50	0.02
.....ATATCTGAC--AAATAAATGA.....	19	2	1.50	0.02
.....TTGAAGGAATTGATATATGCCAA.....	24	2	1.50	0.02
.....TTATTGAAGGAATTGATATATGCA.....	24	2	1.50	0.13
.....ATTGAAGGAATTGATATATGCA.....	21	2	1.50	0.02
.....TGCCAGTAAAATGGTGTATTTTAATT.....	27	2	1.50	0.13
.....TTTATTGAAGGAATTGATATA.....	21	2	1.50	0.02
.....GGGTATTTTAATTCTTCAAT.....	24	2	1.50	0.02
.....ATTATTGAAGGAATTGATATATG.....	24	2	1.50	0.02
.....TATTGAAGGAATTGATATATGCCAGA.....	26	2	1.50	0.02
.....GAAGGAATTGATATATGCCAGT.....	22	2	1.50	0.08
.....GAAGGAATTGATATATGCCAGTAAAAT.....	27	2	1.50	0.09
.....AAGACATATCTGAC--AAATAAA.....	21	2	1.50	0.02
.....TGAAGGAATTGATATATGCCAGTAA.....	25	2	1.50	0.02
.....TGTATTTTAATTCTTCAATAA.....	25	2	1.50	0.02
.....GTGTATTTTAATTCTTCAAT.....	23	2	1.50	0.02
.....TGAAGGAATTGATATATGCCAGTAAAATG.....	29	2	1.50	0.13
.....TAAAGACATATCTGAC--AAATA.....	21	2	1.50	0.02
Total			11193.50	82.44

droPer2	dpe_2477	confident	T-C-TTTAGG-A-TAA-AAT-TATTGAAGGAATTGATATATGCCAGTAAAATGGTGTATTTTAATTCTTCAATAA-AACATA-TCTGAC--AAATAAATGAAT---AA---A	Size	Hits	Total Norm	Total Norm RPM
-	-	-	((((((((((((((.((((((.)))))))))))))))))))))))))))))))	22	12	256.75	3.97
		TATTGAAGGAATTGATATATGC.....	21	13	167.23	2.58
		TATTGAAGGAATTGATATATGCC.....	23	12	84.00	1.30
		TTGAAGGAATTGATATATGCC.....	21	12	45.25	0.70
		TTGAAGGAATTGATATATGCCA.....	22	12	39.25	0.61
		TTGAAGGAATTGATATATGCA.....	20	12	26.67	0.41
		TATTGAAGGAATTGATATATGCCA.....	24	12	16.08	0.25
		ATTGAAGGAATTGATATATGCA.....	21	12	15.92	0.25
		TATTGAAGGAATTGATATATGCCAG.....	25	12	15.83	0.32
		TTATTGAAGGAATTGATATATGCA.....	23	12	15.75	0.24
		TATTGAAGGAATTGATATAT.....	20	13	13.15	0.20
		TTATTGAAGGAATTGATATATGCCAG.....	26	12	11.92	0.54
		TGTATTTTAATTCTTCAATAAA.....	26	2	10.50	0.54
		TATTGAAGGAATTGATATATGCA.....	22	13	8.38	0.16
		TTGAAGGAATTGATATATG.....	19	13	8.00	0.15
		TTATTGAAGGAATTGATATATG.....	22	13	7.54	0.12
		TGAAGGAATTGATATATGCA.....	19	12	6.42	0.10
		TATTGAAGGAATTGATATATGCA.....	23	12	6.33	0.10
		TGTATTTTAATTCTTCAAT.....	22	15	6.00	0.16

.....TTTATTGAAGGAATTGATATATGCCAG.....	27	12	6.00	0.14
.....TGAAGGAATTGATATATGCCA.....	21	12	5.67	0.11
.....ATTGAAGGAATTGATATATG.....	20	13	5.15	0.08
.....TGTATTTTAATTCTTCAATAAA.....	25	15	4.80	0.15
.....TGTATTTTAATTCTTCAATA.....	23	15	4.73	0.14
.....TATTGAAGGAATTGATATATGCC T	24	12	4.42	0.07
.....ATTGAAGGAATTGATATATGCC.....	22	12	4.33	0.07
.....TATTGAAGGAATTGATATA.....	19	13	4.15	0.06
.....TTATTGAAGGAATTGATATATGCC.....	24	12	3.50	0.07
.....TTATTGAAGGAATTGATATATGC.....	24	12	3.50	0.06
.....AATTATTGAAGGAATTGATATATGC.....	26	12	3.25	0.21
.....ATTTATTGAAGGAATTGATATATGC.....	25	12	3.00	0.08
.....TGTATTTTAATTCTTCAATAAA.....	24	15	2.67	0.08
.....TGTATTTTAATTCTTCAATAAAAA.....	27	2	2.50	0.13
.....TTATTGAAGGAATTGATATA.....	20	13	2.00	0.03
.....TGAAGGAATTGATATATGCC.....	20	12	1.67	0.03
.....TATTGAAGGAATTGATATATGCCAGTA.....	27	12	1.42	0.06
.....TATTGAAGGAATTGATATATGCC T	23	12	1.33	0.02
.....GAAGGAATTGATATATGC.....	18	12	1.33	0.06
.....TTGAAGGAATTGATATATGCCAGTAAAA.....	28	12	1.25	0.06
.....ATTGAAGGAATTGATATATGCCAG.....	24	12	1.08	0.06
.....GTATTTTAATTCTTCAAT.....	21	15	1.07	0.05
.....GTATTTTAATTCTTCAATAAAA.....	25	2	1.00	0.05
.....TTTATTGAAGGAATTGATATATGCC.....	25	12	1.00	0.03
.....TGTATTTTAATTCTTCAATAAAA A	28	2	1.00	0.05
.....TATTGAAGGAATTGATATATGATAT.....	18	13	1.00	0.03
.....TTATTGAAGGAATTGATATAT.....	21	13	1.00	0.02
.....TATTGAAGGAATTGATATATGCCAGT.....	26	12	1.00	0.02
.....ATTGATATATGCCAGTAAAATG.....	22	13	0.85	0.02
.....TTGAAGGAATTGATATATGCC T	22	12	0.83	0.01
.....TGAAGGAATTGATATATGCCAG.....	22	12	0.83	0.03
.....ATTGAAGGAATTGATATATGCCA.....	23	12	0.83	0.02
.....TTTATTGAAGGAATTGATATATGCCA.....	26	12	0.83	0.02
.....GTATTTTAATTCTTCAATA.....	22	15	0.80	0.03
.....CAGTAAAATGGGTGATTTTAAT.....	23	14	0.79	0.02
.....TTGAAGGAATTGATATATGCCAGTAAA.....	27	12	0.75	0.03
.....TGTATTTTAATTCTTCAA.....	21	15	0.73	0.03
.....ATTGATATATGCCAGTAAAATGGTGT.....	21	14	0.71	0.02
.....TGTATTTTAATTCTTCAATAAA C	26	13	0.69	0.03
.....GAAGGAATTGATATATGCC.....	19	12	0.67	0.03
.....ATATCTGAC--AAATAAATG.....	18	14	0.64	0.05
.....ATTATTGAAGGAATTGAT.....	19	13	0.62	0.04
.....TTATTGAAGGAATTGATA.....	18	13	0.62	0.01
.....TATTGAAGGAATTGATATATGCC C	24	12	0.58	0.02
.....TTATTGAAGGAATTGATATATGCCA.....	25	12	0.58	0.02
.....TGAAGGAATTGATATATGCCAGT.....	20	14	0.57	0.02
.....AAAAACATATCTGAC--AAA.....	23	12	0.50	0.02
.....TTTTAATTCTTCAATAAAA.....	18	2	0.50	0.04
.....ATTTTAATTCTTCAATAAAAA.....	22	2	0.50	0.03
.....ATCTGAC--AAATAAATGAAT.....	24	2	0.50	0.03
.....TATTGAAGGAATTGATATATGCCAGTA.....	19	14	0.50	0.02
.....TTTAATTCTTCAATAAAA.....	28	12	0.50	0.03
.....TTTTTAATTCTTCAATAAAA.....	21	2	0.50	0.03
.....ATCTGAC--AAATAAATGAAT.....	25	2	0.50	0.03
.....TTATTGAAGGAATTGATATATGCCAGT.....	27	12	0.50	0.02
.....TATTCTGAC--AAATAAATGA.....	19	14	0.50	0.01
.....TATTTTAATTCTTCAATAAAA.....	24	2	0.50	0.19
.....TATTGAAGGAATTGATATATGCC C	24	12	0.50	0.03
.....GTATTTTAATTCTTCAATAAAA.....	26	2	0.50	0.03
.....TTATTGAAGGAATTGATATATGATAT.....	19	13	0.46	0.01
.....TTGAAGGAATTGATATATGAT T	20	13	0.46	0.01
.....ATTGAAGGAATTGATATATGCC A	22	12	0.42	0.02
.....TTGAAGGAATTGATATATGCC C	22	12	0.42	0.01
.....AAATTATTGAAGGAATTGATATATGCC.....	27	12	0.42	0.03
.....TATTGAAGGAATTGATATATGCC A	25	12	0.42	0.02

.....ATTGAAGGAATTGATATAT.....TATTTTAATTCTTCAATAAA.....	23	15	0.40	0.02
.....ATTGAAGGAATTGATA.....ATTTTAATTCTTCAATA.....	20	15	0.40	0.02
.....TGATATATGCCAGTAAAATG.....TATTTTAATTCTTCAATA.....	21	15	0.40	0.02
.....ATTTATTGAAGGAATTGATA.....ACATATCTGAC--AAATAAATGAAT-----A.....	19	13	0.38	0.02
.....TGAAAGGAATTGATATATGCCAG.....TCTGAC--AAATAAATGAAT-----A.....	19	13	0.38	0.01
.....TATTGAAGGAATTGATATATGCT <ins>T</ins>	20	13	0.38	0.01
.....TGAAGGAATTGATATATGCCAGTAAAA.....	20	13	0.38	0.01
.....ATTTATTGAAGGAATTGATATATGCC.....ACATATCTGAC--AAATAAATGAAT-----A.....	24	14	0.36	0.01
.....ATTTATTGAAGGAATTGATATATGCC.....TCTGAC--AAATAAATGAAT-----A.....	18	14	0.36	0.03
.....TGAAAGGAATTGATATATGCCAG.....	23	12	0.33	0.02
.....TATTGAAGGAATTGATATATGCT <ins>T</ins>	24	12	0.33	0.02
.....TGAAGGAATTGATATATGCCAGTAAAA.....	27	12	0.33	0.02
.....ATTTATTGAAGGAATTGATATATGCC.....	26	12	0.33	0.01
.....ATTTATTGAAGGAATTGATATATGCC.....GTATTTTAATTCTTCAATAA.....	27	12	0.33	0.01
.....TGAAAGGAATTGATATATGCCAGTAA.....	26	12	0.33	0.01
.....ATATATGCCAGTAAAATG.....	18	13	0.31	0.01
.....AATTATTGAAGGAATTGATATA.....	23	13	0.31	0.01
.....TTTATTGAAGGAATTGATATATG.....	23	13	0.31	0.01
.....TATTGAAGGAATTGATATATA <ins>A</ins>	21	13	0.31	0.01
.....AATTATTGAAGGAATTGATA.....	21	13	0.31	0.03
.....TTTATTGAAGGAATTGATATA.....	21	13	0.31	0.01
.....TGATTTTAATTCTTCAATAA.....	23	15	0.33	0.01
.....TTGAAAGGAATTGATATATGCCAGTAA.....	23	15	0.33	0.02
.....ATATATGCCAGTAAAATG.....	18	13	0.31	0.01
.....AATTATTGAAGGAATTGATATA.....	23	13	0.31	0.01
.....TTTATTGAAGGAATTGATATATG.....	23	13	0.31	0.01
.....TATTGAAGGAATTGATATATA <ins>A</ins>	21	13	0.31	0.01
.....AATTATTGAAGGAATTGATA.....	21	13	0.31	0.03
.....TTTATTGAAGGAATTGATATA.....	21	13	0.31	0.01
.....TGATTTTAATTCTTCAATAA.....	23	15	0.27	0.01
.....TTGAAAGGAATTGATATATGCCAGTAA.....	22	15	0.27	0.01
.....GTATTTTAATTCTTCAATAAA.....	24	15	0.27	0.01
.....GAATTGATATATGCCAGTAAAATG.....	25	12	0.25	0.01
.....AGGAATTGATATATGCCAGTAAAATG.....	26	12	0.25	0.01
.....TTATTGAAGGAATTGATATATGCCAGTAA.....	28	12	0.25	0.01
.....TATTGAAGGAATTGATATATGCCAG <ins>A</ins>	26	12	0.25	0.01
.....TGAAGGAATTGATATATGCCAGTAA.....	25	12	0.25	0.01
.....GAAGGAATTGATATATGCCAG.....	21	12	0.25	0.01
.....TTGAAGGAATTGATATATGCCAGT.....	24	12	0.25	0.01
.....ATTATTGAAGGAATTGATATATGCCAG.....	28	12	0.25	0.01
.....TAAAATTATTGAAGGAATTGATATA.....	26	13	0.23	0.01
.....ATTATTGAAGGAATTGTA.....	18	13	0.23	0.01
.....TATTGAAGGAATTGATAT <ins>GC</ins>	20	13	0.23	0.01
.....TAAAATTATTGAAGGAATTGATA.....	24	13	0.23	0.01
.....TTTATTGAAGGAATTGATAT.....	20	13	0.23	0.01
.....TATTGAAGGAATTGATATATGCT <ins>T</ins>	21	13	0.23	0.09
.....TATTGAAGGAATTGATATATC <ins>C</ins>	21	13	0.23	0.01
.....ATTTATTGAAGGAATTGATA.....	22	13	0.23	0.02
.....GATATATGCCAGTAAAATGGTGT.....	23	13	0.23	0.01
.....TGAAGGAATTGATATATG.....	18	13	0.23	0.01
.....TATCTGAC--AAATAAATGA.....	18	14	0.21	0.01
.....ATCTGAC--AAATAAATGAAT-----A.....	20	14	0.21	0.01
.....CATATCTGAC--AAATAAATGAAT-----A.....	23	14	0.21	0.02
.....ATATCTGAC--AAATAAATGAAT-----A.....	22	14	0.21	0.02
.....ATTGAAGGAATTGATATATGCCAGTAA.....	26	12	0.17	0.01
.....TATTGAAGGAATTGATATATGCCA <ins>AC</ins>	26	12	0.17	0.01
.....AGGAATTGATATATGCCAGTAAAAT.....	25	12	0.17	0.01
.....TATTGAAGGAATTGATATATGCCA <ins>AA</ins>	26	12	0.17	0.01
.....TTGAAGGAATTGATATATGCCA <ins>AA</ins>	24	12	0.17	0.01
.....TTGAAGGAATTGATATATGCT <ins>T</ins>	21	12	0.17	0.02
.....TTGAAGGAATTGATATATGCCAGTA.....	25	12	0.17	0.01
.....TTGAAGGAATTGATATATGCCA <ins>T</ins>	23	12	0.17	0.04
.....GAAGGAATTGATATATGCCAGTAA.....	24	12	0.17	0.01
.....AAGGAATTGATATATGCCA.....	19	12	0.17	0.06
.....ATTGAAGGAATTGATATATGCCAGT.....	25	12	0.17	0.00
.....ATTTATTGAAGGAATTGATAT.....	21	13	0.15	0.01
.....TATATGCCAGTAAAATGGTGT.....	21	13	0.15	0.01
.....TATTGAAGGAATTGATATATA <ins>A</ins>	20	13	0.15	0.00
.....AAATTATTGAAGGAATT.....	18	13	0.15	0.01
.....TATTGAAGGAATTGATATATG <ins>C</ins>	22	13	0.15	0.01
.....AATTATTGAAGGAATTGATATATG.....	25	13	0.15	0.01
.....ATTGATATATGCCAGTAAAAT.....	21	13	0.15	0.01
.....TATTGAAGGAATTGATATATG <ins>C</ins>	19	13	0.15	0.01
.....AAATTATTGAAGGAATTGATA.....	22	13	0.15	0.01

		TGTATTTTAATTCTTCAAA	22	20	0.35	0.01
	TATTCTTCAAT.....	20	20	0.35	0.01
	TGTATTTTAATTCTTCAATAA	24	20	0.35	0.02
	TGTATTTTAATTCTTCAATAA	25	20	0.35	0.01
	AATTGATATATGCCAGTAAAATG	23	20	0.35	0.01
	GTATTTTAATTCTTCAAT.....	21	20	0.35	0.01
	TGTATTTTAATTCTTCAATAA	25	20	0.35	0.01
	TATTGAAGAAATTGATATA	19	3	0.33	0.02
	ATTGAAGAAATTGATATA	21	3	0.33	0.01
	T-TGGGT-TAAGA-----TTTTATTGAA.....	22	20	0.30	0.01
	AAAAAATTATTGACGTAATG.....	21	20	0.30	0.01
	TGATATATGCCAGTAAAATGGTGTAT.....	26	20	0.25	0.01
	ATAAAAATTATTGACGT.....	19	20	0.25	0.01
	TATTTTAATTCTTCAATAAA.....	23	20	0.25	0.01
	ATTTTAATTCTTCAAT.....	19	20	0.25	0.01
	TATTTTAATTCTTCAATAAA.....	22	20	0.25	0.01
	GATATATGCCAGTAAAATGGT.....	21	20	0.25	0.01
	ATTGATATATGCCAGTAAAATGGTGT.....	26	20	0.20	0.01
	TGTATTTTAATTCTTCAATAA	24	20	0.20	0.01
	GGT-TAAGA-----TTTTATTGAA.....	18	20	0.20	0.01
	GTGTATTTTAATTCTTCAATA.....	24	20	0.15	0.00
	GTATTTTAATTCTTCAATA.....	22	20	0.15	0.01
	ATAAAAATTATTGACGTAATGAAAT.....	27	20	0.15	0.01
	GTAAAATGGGTATTTTAAT.....	21	20	0.15	0.00
TCGT	-TTGGGT-TAAGA-----TTTTATTGAA.....GTAAAATGGGTATTTTA.....	26	20	0.15	0.01
	GTATTTTAATTCTTCAAA.....	19	20	0.15	0.01
	ATATTATTGACGTAATGAA.....	20	20	0.15	0.01
	ATTTTAATTCTTCAATAAA.....	19	20	0.15	0.01
	TTAATTCTTCAATAAA.....	22	20	0.15	0.01
	TTGGGT-TAAGA-----TTTTATTGA.....	19	20	0.10	0.00
	GTATTTTAATTCTTCAAT.....	20	20	0.10	0.00
	GGTGTATTTTAATTCTTCAAT.....	22	20	0.10	0.00
	ATAAAAATTATTGACGTAATG.....	24	20	0.10	0.00
	TATTTTAATTCTTCAATAA	23	20	0.10	0.00
	AAATATTATTGACGTAATG.....	19	20	0.10	0.00
	GTGTATTTTAATTCTTCAAT.....	23	20	0.10	0.00
	GTATTTTAATTCTTCA.....	19	20	0.10	0.00
	TTGATATATGCCAGTAAAATGGTGTAT.....	27	20	0.10	0.00
	TTTTAATTCTTCAATA.....	22	20	0.10	0.00
	AAATATTATTGACGTAATGAA.....	18	20	0.10	0.00
	TGGGTATTTTAATTCTTCAATA.....	21	20	0.10	0.00
	TAAAATATTATTGACGTAAT.....	26	20	0.10	0.00
	TATATGCCAGTAAAATGGTGT.....	21	20	0.10	0.00
	AAATGGGTATTTTAAT.....	22	20	0.10	0.00
	AATTGATATATGCCAGTAAAATGGTGT.....	27	20	0.10	0.00
	TATGCCAGTAAAATGGTGT.....	20	20	0.10	0.00
	AAATATTATTGACGTAATG.....	18	20	0.10	0.00
	TTCTTCAATAAAAATTATTGACGTA.....	28	20	0.10	0.00
	AAAATATTATTGACGTAATGAA.....	23	20	0.10	0.00
	AAATATTATTGACGTAATG.....	20	20	0.05	0.00
	GT -TTGGGT-TAAGA-----TTTTATTG.....	21	20	0.05	0.00
	TTTTATTG.....	27	20	0.05	0.00
	ATTGATATATGCCAGTAAAATGGTGT.....	27	20	0.05	0.00
	TGTATTTTAATTCTTCAATAAA	27	20	0.05	0.00
	GATATATGCCAGTAAAATGGTGT.....	23	20	0.05	0.00
	AATAAAAATTATTGACGTA.....	21	20	0.05	0.00
	GTGTATTTTAATTCTTCA.....	21	20	0.05	0.00
	ATATATGCCAGTAAAATGG.....	19	20	0.05	0.00
	TGGGT-TAAGA-----TTTTATTGAA.....	19	20	0.05	0.00
	AAAATATTATTGACGTAAT.....	20	20	0.05	0.00
	GTGTATTTTAATTCTTCTT.....	19	20	0.05	0.00
	AATTCTTCAATAAAA.....	18	20	0.05	0.00
	GTATTTTAATTCTTCAATAAAA.....	26	20	0.05	0.00
	CAGAAAATGGGTATTTTAATTCTTCTT.....	29	20	0.05	0.00
	TTTAATTCTTCAATAAAA.....	21	20	0.05	0.00

	ATGCCAGTAAATGGTGA	19	20	0.05	0.00
	GGTGTATTTTAATTCTTC.	22	20	0.05	0.00
TTGGGT-TAAGA-----	TTTTATTGAAGC	23	20	0.05	0.00
	TGATATATGCCAGTAAAT.	19	20	0.05	0.00
	GGTGTATTTTAATTCTTC.	21	20	0.05	0.00
	GTATTTTAATTCTTCAA	21	20	0.05	0.00
	ATTGATATATGCCAGTAAATGGT.	25	20	0.05	0.00
	ATGGTGTATTTTAATTCTTC.	24	20	0.05	0.00
	TGGTGTATTTTAATTCTTC.	19	20	0.05	0.00
	CAGTAAATGGTGTATTTTAATT.	25	20	0.05	0.00
	AAATATTATTGACGTAATGA.	20	20	0.05	0.00
	TGTATTTTAATTCTTCAA	23	20	0.05	0.00
CGT-TTGGGT-TAAGA-----	TTTTATTG	22	20	0.05	0.00
	TTGATATATGCCAGTAAATGGT.	25	20	0.05	0.00
.T-TTGGGT-TAAGA-----	TTTTATTG.	20	20	0.05	0.00
	GTGTATTTTAATTCTTCAA	22	20	0.05	0.00
	AAAATATTATTGACGTAATGA.	22	20	0.05	0.00
TCGT-TTGGGT-TAAGA-----	TTTTATTGA.	24	20	0.05	0.00
	ATATATGCCAGTAAATGGT.	20	20	0.05	0.00
	TTATTGACGTAATGAAAT.	18	20	0.05	0.00
	TGGTGTATTTTAATTCTTCAA	28	20	0.05	0.00
	TGATATATGCCAGTAAA.	18	20	0.05	0.00
	ATTGATATATGCCAGTAAA.	20	20	0.05	0.00
.GT-TTGGGT-TAAGA-----	TTTTATTGAA	24	20	0.05	0.00
	C.				
	AAAAATATTATTGACGTAATGAA.	24	20	0.05	0.00
TTGGGT-TAAGA-----	TTTTATTG.	19	20	0.05	0.00
.T-TTGGGT-TAAGA-----	TTTTATTGA.	21	20	0.05	0.00
	TGGTGTATTTTAATTCTTC.	25	20	0.05	0.00
	TGTATTTTAATTCTTCAA	26	20	0.05	0.00
	TATTTTAATTCTTCAA.	19	20	0.05	0.00
	AATGGTGTATTTTAATTCTC.	21	20	0.05	0.00
	CAGTAAATGGTGTATTTAA.	22	20	0.05	0.00
	TTTAATTCTTCAA	20	20	0.05	0.00
	AAAAATATTATTGACGTA.	19	20	0.05	0.00
.GGGT-TAAGA-----	TTTTATTGAAG.	20	20	0.05	0.00
	TGGTGTATTTTAATTCTTC.	27	20	0.05	0.00
	AAATATTATTGACGTAATGAA.	23	20	0.05	0.00
	CAGTAAATGGTGTATTTAA.	26	20	0.05	0.00
	TTTAATTCTTCAA	20	20	0.05	0.00
	TTTTAATTCTTCAA	21	20	0.05	0.00
	TAAAATGGTGTATTTAA.	20	20	0.05	0.00
	TTTTAATTCTTCAA	19	20	0.05	0.00
	TGATATATGCCAGTAAATGGT.	24	20	0.05	0.00
	AGTAAATGGTGTATTTAA.	25	20	0.05	0.00
	ATATGCCAGTAAATGGT.	21	20	0.05	0.00
	TGTATTTTAATTCTTCAA	24	20	0.05	0.00
	AATTGATATATGCCAGTAAATGGT.	25	20	0.05	0.00
	GTATTTTAATTCTTCAA	25	20	0.05	0.00
Total				143.57	1.22

droVir3 dvi 24632 candidate-rescued

TATTGAAGGAATTGATATATGC	A	23	20	18.90	0.39
TATTGAAGGAATTGATATA	.	19	20	15.55	0.12
.	TGTATTTTAATTCTTCAATAAA	25	20	14.55	0.29
TATTGAAGGAATTGATATATGCCAGTAAAATGGGTATTTT	.	42	20	13.30	0.63
TATTGAAGGAATTGATATATGCCA	A	25	20	12.95	0.23
TTATTGAAGGAATTGATATATG	.	22	20	11.45	0.22
.	TGTATTTTAATTCTTCAATAAA	26	20	11.25	0.13
TTGAAGGAATTGATATATGC	.	20	20	11.05	0.25
.	TGTATTTTAATTCTTCAAT	22	20	9.50	0.30
TGAAGGAATTGATATATGCCA	.	21	20	7.65	0.15
ATTGAAGGAATTGATATATG	.	20	20	6.10	0.18
TTATTGAAGGAATTGATATATGCCAGT	.	27	20	5.20	0.06
TATTGAAGGAATTGATATATGCCAGTAAAATGGGTATTTTAA	.	45	20	4.60	0.13
ATTGAAGGAATTGATATATGCC	.	22	20	4.40	0.08
TTGAAGGAATTGATATATG	.	19	20	4.15	0.06
.	GTATTTTAATTCTTCAATA	22	20	3.80	0.08
TTTTATTGAAGGAATTGATATATGC	.	25	20	3.70	0.08
TATTGAAGGAATTGATATATGCC	T	24	20	3.60	0.10
TATTGAAGGAATTGATATATGCCAGTAAAATGGGTATTTT	.	40	20	3.60	0.07
TATTGAAGGAATTGATATATGCCA	AA	26	20	3.50	0.08
TATTGAAGGAATTGATATATGCC	C	24	20	3.25	0.05
TATTGAAGGAATTGATATATGCCAGTAAAATGGGTATTTT	.	41	20	2.75	0.07
TTGAAGGAATTGATATATGCCAGT	.	24	20	2.70	0.10
.	GTATTTTAATTCTTCAATAAA	25	20	2.60	0.04
TATTGAAGGAATTGATATATGCCAGTAAAATGGGTATTTT	T	43	20	2.50	0.04
.	TGTATTTTAATTCTTCA	20	20	2.50	0.04
TATTGAAGGAATTGATATATGC	T	23	20	2.45	0.04
.	GTATTTTAATTCTTCAATAAA	24	20	2.35	0.03
GTATTTTAATTCTTCAAT	.	21	20	2.25	0.11
TTATTGAAGGAATTGATATATGCCA	.	25	20	2.20	0.06
TATTGAAGGAATTGATATATG	AA	23	20	2.20	0.05
TTTATTGAAGGAATTGATATATGCCAG	.	27	20	2.15	0.06
TATTGAAGGAATTGATATATGCCAGT	.	26	20	1.90	0.02
TTTTATTGAAGGAATTGATATATGCCAG	.	28	20	1.85	0.04
.	TGTATTTTAATTCTTCAATAAA	27	20	1.85	0.04
TATTGAAGGAATTGATATATGC	A	24	20	1.85	0.04
.	TATTTTTAATTCTTCAATAAA	23	20	1.85	0.06
.	TGTATTTTAATTCTTCAATAAAACA	29	20	1.80	0.13
TTATTGAAGGAATTGATATATGCCAG	.	26	20	1.80	0.04
TGAAGGAATTGATATATGCCAG	.	22	20	1.75	0.02
ATTGAAGGAATTGATATATGCCA	.	23	20	1.75	0.03
TGAAGGAATTGATATATGCC	T	22	20	1.70	0.03
TTTATTGAAGGAATTGATATATGCCA	.	26	20	1.70	0.08
TTGAAGGAATTGATATATGCC	C	22	20	1.50	0.03
TTTATTGAAGGAATTGATATATGCCAGT	.	28	20	1.50	0.14
.	TGTATTTTAATTCTTCAATAAA	25	20	1.40	0.05
TATTGAAGGAATTGATATAT	.	21	20	1.40	0.04
TTTATTGAAGGAATTGATATATG	C	24	20	1.30	0.02
TTGAAGGAATTGATATATGCCAGTAAAAT	.	29	20	1.20	0.04
TTGAAGGAATTGATATATGCCAGT	.	25	20	1.10	0.01
.	TGTATTTTAATTCTTCAATAAA	25	20	1.10	0.05
TTTTAGGTATAAAC	--AATTATTGAAGGAAT	31	1	1.00	0.08
.	TGAAGGAATTGATATATGCCAGTA	25	20	0.95	0.07
.	TGTATTTTAATTCTTCAA	21	20	0.95	0.01
.	TGACGTAATGAAATAAAC	23	20	0.90	0.02
.	TGTATTTTAATTCTTCAATAA	24	20	0.80	0.06
TATTGAAGGAATTGATATATGCCAGTA	.	28	20	0.80	0.02
TATTGAAGGAATTGATATATGCC	G	23	20	0.80	0.02
TTTATTGAAGGAATTGATATA	.	21	20	0.80	0.01
TTTATTGAAGGAATTGATA	.	19	20	0.75	0.03
TTTATTGAAGGAATTGATATA	TGCCAGTAAAATGGGTATTTT	45	20	0.70	0.02
TTATTGAAGGAATTGATA	.	20	20	0.70	0.02
TTGAAGGAATTGATATATGCCAG	.	23	20	0.70	0.02
.	ATGGTGTATTTTAATTCTTCAAT	26	20	0.70	0.04
ATAAAC	--AATTATTGAAGGAAT	24	20	0.70	0.10
ATTTATTGAAGGAATTGATATATGC	.	26	20	0.70	0.01
.	TTGAAGGAATTGATATATGC	21	20	0.70	0.02

.....TTTTATTGAAGGAATTGATATATGCCA.....	27	20	0.65	0.03
.....TATTGAAGGAATTGATATATGCC.....	20	20	0.65	0.02
.....TTGAAGGAATTGATATAT.....	18	20	0.65	0.05
.....TTATTGAAGGAATTGATATATGCC.....	24	20	0.60	0.01
.....TAGGTATAAAC---AATTATTGAAGGAA.....	27	20	0.60	0.06
.....TGACGTAATGAAATAAAC---CAA.....	21	20	0.55	0.06
.....AACATTGTGACGTAATGAAATAAAC.....	26	20	0.55	0.01
.....TTGAAGGAATTGATATAATGCCA <ins>T</ins>	23	20	0.55	0.01
.....TTTATTGAAGGAATTGATATATGCCAGTAAAATGGTGTATT.....	43	20	0.55	0.01
.....ATTGAAGGAATTGATATAT.....	19	20	0.55	0.01
.....TTTATTGAAGGAATTGATATATGCCAGT.....	29	20	0.50	0.01
.....ACATTGTGACGTAATGAAATAAAC---CA.....	27	20	0.50	0.04
.....TATTGAAGGAATTGATATAATGCCAG.....	25	20	0.50	0.01
.....TATTGAAGGAATTGATATATAA <ins>A</ins>	23	20	0.50	0.03
.....GTATTTTAATTCTTTCAATAAAAA.....	26	20	0.50	0.01
.....TAGGTATAAAC---AATTATTGAAGGAA.....	28	20	0.45	0.01
.....TATTGAAGGAATTGATATATA <ins>A</ins>	21	20	0.45	0.01
.....TGAAGGAATTGATATATGCCAGTAAA.....	26	20	0.45	0.02
.....TTATTGAAGGAATTGATATATGCCAGTAAAATGGTGTATTTTT.....	43	20	0.45	0.01
.....TATTGAAGGAATTGATATATGC <ins>G</ins>	22	20	0.45	0.01
.....GTATTTTAATTCTTTCAATC <ins>C</ins>	22	20	0.45	0.01
.....TATTGAAGGAATTGATATATGCCAGTAAAATGGTGTATT.....	39	20	0.45	0.02
.....TGACGTAATGAAATAAAC---C.....	19	20	0.45	0.03
.....TTTATTGAAGGAATTGATATAATGCC.....	25	20	0.45	0.04
.....TGAAGGAATTGATATATGCC.....	20	20	0.45	0.02
.....TTTGTGACGTAATGAAATAAAC---C.....	23	20	0.45	0.04
.....TATTGAAGGAATTGATATATGCCAGTA.....	27	20	0.45	0.01
.....ATTGAAGGAATTGATATATG <ins>A</ins>	21	20	0.40	0.01
.....TGTATTTTAATTCTTTCAAT <ins>T</ins>	23	20	0.40	0.01
.....GTATTTTAATTCTTTCAATAA <ins>T</ins>	24	20	0.40	0.01
.....GAAGGAATTGATATATGC.....	18	20	0.40	0.06
.....AACATTGTGACGTAATGAAAT.....	22	20	0.40	0.02
.....GTATTTTAATTCTTTCAATAA.....	23	20	0.40	0.02
.....TATTGAAGGAATTGATATATGCCA <ins>T</ins>	25	20	0.40	0.05
.....TTGTGACGTAATGAAATAAAC---CAATA.....	26	20	0.40	0.01
.....ATTGAAGGAATTGATATA.....	18	20	0.35	0.01
.....ATTTTTAATTCTTTC.....	16	20	0.35	0.04
.....TGAAGGAATTGATATATGCCAGTAAA.....	27	20	0.35	0.03
.....AAAAACATTGTGACGTAAT.....	20	20	0.35	0.10
.....TTTATTGAAGGAATTGATAT.....	20	20	0.35	0.01
.....TTTTTAATTCTTTCAATAAAAACATT.....	27	20	0.35	0.03
.....TTTATTGAAGGAATTGATAT.....	18	20	0.35	0.02
.....ATTGATATATGCCAGTAAAATGGTGT.....	27	20	0.35	0.01
.....TAAAC---AATTATTGAAGGAATTGAT.....	26	20	0.35	0.03
.....TATTGAAGGAATTGATATC <ins>G</ins>	19	20	0.35	0.02
.....TGAAGGAATTGATATATG.....	18	20	0.30	0.02
.....TATTGAAGGAATTGATATATC <ins>C</ins>	21	20	0.30	0.06
.....TGACGTAATGAAATAAAC---CAAT.....	22	20	0.30	0.05
.....TATTGAAGGAATTGATATATGC <ins>AG</ins>	24	20	0.30	0.01
.....AACATTGTGACGTAATGAAATA.....	23	20	0.30	0.01
.....TTTATTGAAGGAATTGATATA.....	22	20	0.30	0.01
.....ATTGTGACGTAATGAAATA.....	20	20	0.30	0.01
.....TGAAGGAATTGATATATGC.....	19	20	0.25	0.01
.....AACATTGTGACGTAATGAA.....	20	20	0.25	0.02
.....AGGTATAAAC---AATTATTGAAGGAAT.....	27	20	0.25	0.01
.....AATTATTGAAGGAATTGATATATGC.....	27	20	0.25	0.00
.....TTTTAATTCTTTCAAT.....	17	20	0.25	0.03
.....TTATTGAAGGAATTGATA.....	18	20	0.25	0.02
.....TTGAAGGAATTGATATATGCCA <ins>T</ins>	27	20	0.25	0.00
.....AAAAACATTGTGACGTAATGAAAT.....	23	20	0.25	0.01
.....ATTTATTGAAGGAATTGATA.....	25	20	0.25	0.01
.....GAAGGAATTGATATATGCCAGTAAA.....	21	20	0.25	0.01
.....TATTGAAGGAATTGATATATAA <ins>A</ins>	22	20	0.25	0.03
.....AAC---AATTATTGAAGGAAT.....	21	20	0.25	0.03
.....ATTGATATATGCCAGTAAAATG.....	22	20	0.25	0.02
.....TATTGAAGGAATTGATATATGCCAGTAAAATGGTGTATTTTA <ins>A</ins>	45	20	0.25	0.01

	TTTTAATTCTTCAATAAAAACATTGT.....	29	20	0.25	0.02
	TATTTTAATTCTTCAATAAA.....	22	20	0.20	0.01
	TTAGGTATAAAC---AATTATTGAAAGGAAT.....	29	20	0.20	0.01
	TCTTCAATAAAAACATTG.....	20	20	0.20	0.01
	TTTTAATTCTTCAATAAAAACA.....	25	20	0.20	0.02
	AAGGAATTGATATATGCCAGTAAAA.....	25	20	0.20	0.01
	GTATTTTAATTCTTCAATAA.....	24	20	0.20	0.02
	AACATTGTGACGTAATGAAATA.....	24	20	0.20	0.01
	TATTGAAGGAATTGATATATGCCAGTAAAATGGTGTATTT.....	43	20	0.20	0.01
	TGTATTTTAATTCTTCAATAA.....	23	20	0.20	0.01
	TTATTGAAGGAATTGATATA.....	22	20	0.20	0.03
	AACATTGTGACGTAATGAAAT.....	23	20	0.20	0.06
	GATATATGCCAGTAAAATGGTGTAT.....	25	20	0.20	0.03
	ATTGTGACGTAATGAAAT.....	19	20	0.20	0.02
	TTTTATTGAAGGAATTGATA.....	20	20	0.20	0.02
	ATTGATATATGCCAGTAAAA.....	20	20	0.20	0.01
	AACATTGTGACGTAATGAAATAA.....	25	20	0.20	0.01
	TTTTATTGAAGGAATTGATATA.....	24	20	0.20	0.01
	TTTTAATTCTTCAATAAAAACATTG.....	28	20	0.20	0.01
	GAAGGAATTGATATATGCCA.....	20	20	0.20	0.02
	TTGAAGGAATTGATATATGCCAGTAAAATGGTGTATTT.....	21	20	0.20	0.02
	AAACATTGTGACGTAATGAA.....	38	20	0.20	0.01
	ATTGAAGGAATTGATATA.....	23	20	0.20	0.01
	TTGAAGGAATTGATATATGCC.....	28	20	0.20	0.01
	AACATTGTGACGTAATGAAAT.....	22	20	0.20	0.02
	AACATTGTGACGTAATGAAAT.....	16	20	0.20	0.01
	AACATTGTGACGTAATGAAAT.....	22	20	0.20	0.01
	AACATTGTGACGTAATGAA.....	19	20	0.20	0.04
	GAATTGATATATGCCAGTAAAATGGTGT.....	28	20	0.15	0.01
	ATTTTTAATTCTTCAATAAAA.....	23	20	0.15	0.01
	C---AATTATTGAAGGAATTGATATA.....	25	20	0.15	0.01
	Total				1950.45	3.09

droMoJ3	dmo_3124	candidate-rescued		Size	Hits	Total Norm	Total RPK
		TG---TTAGGTATAAAA-TTTTTTATTGAAGGAATTGATATATGCCAGTAAAATGGTATTTTAATTCTTCAATAAA.....					
	CATATA.....TG TGACGTAATGAAATAAC----CAAAAA.....					
	TATTGAAGGAATTGATATATGC.....	22	7	739.14	7.22	
	TATTGAAGGAATTGATATATGCC.....	23	7	470.00	4.59	
	TATTGAAGGAATTGATATATG.....	21	7	350.00	3.42	
	TTGAAGGAATTGATATATGCC.....	21	7	119.29	1.17	
	TTGAAGGAATTGATATATGCCA.....	22	7	89.14	0.87	
	TGTATTTTAATTCTTCAAT.....	22	7	70.43	0.87	
	TGTATTTTAATTCTTCAATA.....	23	7	58.71	0.76	
	TATTGAAGGAATTGATATATGCCA.....	24	7	56.86	0.56	
	ATTGAAGGAATTGATATATGCC.....	22	7	45.43	0.44	
	ATTGAAGGAATTGATATATG.....	21	7	43.29	0.42	
	TTATTGAAGGAATTGATATATGC.....	23	7	42.00	0.41	
	TTGAAGGAATTGATATATGC.....	20	7	37.57	0.37	
	TATTGAAGGAATTGATATA.....	20	7	31.43	0.31	
	TGTATTTTAATTCTTCAATAA.....	25	7	29.14	0.38	
	TTTATTGAAGGAATTGATATA.....	24	7	23.43	0.24	
	TTTTATTGAAGGAATTGATATA.....	25	7	20.14	0.20	
	GTATTTTAATTCTTCAAT.....	21	7	16.86	0.23	
	TTTTATTGAAGGAATTGATATA.....	26	6	15.50	0.16	
	TATTGAAGGAATTGATATATGCC.....	24	7	15.43	0.15	
	TATTGAAGGAATTGATATATGCA.....	23	7	14.86	0.15	
	CATATATGTGACGTAATGAAATAAC----.....	26	5	14.60	0.20	
	TGTATTTTAATTCTTCAATAA.....	24	7	14.29	0.19	
	TGTATTTTAATTCTTCAAA.....	21	7	13.86	0.19	
	ATATATGTGACGTAATGAA.....	19	6	13.83	0.21	
	ATATATGTGACGTAATGAAAT.....	21	5	13.20	0.19	
	TTATTGAAGGAATTGATATA.....	22	7	12.29	0.12	
	TTGAAGGAATTGATATATG.....	19	7	12.14	0.12	
	GTATTTTAATTCTTCAATA.....	22	7	11.43	0.12	
	AACATATATGTGACGTAAT.....	20	6	10.00	0.14	
	AACATATATGTGACGTAATGAAAT.....	25	5	9.60	0.13	
	TGAAGGAATTGATATATGCCA.....	21	7	9.43	0.09	
	ATTGAAGGAATTGATATATG.....	20	7	8.86	0.09	

.....TTTTTATTGAAGGAATTGATATATGC	27	6	8.83	0.09
.....TATTGAAGGAATTGATATACT	21	7	8.57	0.09
.....TTGAAGGAATTGATATATGCC	22	7	8.43	0.08
.....TATTGAAGGAATTGATATA	19	7	7.71	0.08
.....TGTATTTTAATTCTTCAATAAC	25	7	7.00	0.09
.....ATATATGTGACGTAATGAAATA	22	5	6.00	0.08
.....TGAAGGAATTGATATATGCC	20	7	6.00	0.06
.....AACATATATGTGACGTAATGAA	23	6	5.67	0.08
.....TTATTGAAGGAATTGATATATGCC	24	7	5.57	0.06
.....TTATTGAAGGAATTGATATAATGCCAG	27	7	5.57	0.06
.....TTATTGAAGGAATTGATATAATGCCAG	26	7	5.57	0.06
.....TTTATTGAAGGAATTGATATAATGCCAG	28	7	5.57	0.06
.....TTTTTATTGAAGGAATTGATATAATGC	28	6	5.17	0.05
.....GTATTTTAATTCTTCAATAAA	24	7	5.14	0.07
.....TGAAGGAATTGATATATGC	19	7	5.14	0.05
.....AACATATATGTGACGTAATGAAATA	26	5	4.80	0.07
.....TATTGAAGGAATTGATATATGCCAG	25	7	4.71	0.05
.....TGTATTTTAATTCTTCAATT	23	7	4.57	0.06
.....AACATATATGTGACGTAATGA	22	6	4.50	0.09
.....TGTATTTTAATTCTTCAATAA	25	7	4.43	0.06
.....TTTATTGAAGGAATTGATA	20	7	4.43	0.05
.....TGTATTTTAATTCTTCA	20	7	4.29	0.08
.....AACATATATGTGACGTAAT	19	6	4.17	0.09
.....TTTATTGAAGGAATTGATATAATGCC	26	7	4.14	0.04
.....TTTTTATTGAAGGAATTGATATA	25	6	3.83	0.04
.....TGAAGGAATTGATATAATGCC	22	7	3.71	0.04
.....TATTGAAGGAATTGATATATGT	22	7	3.71	0.04
.....TTTATTGAAGGAATTGATATAATGCC	25	7	3.43	0.04
.....TATTGAAGGAATTGATATAATGCCAGT	26	7	3.43	0.04
.....TTTATTGAAGGAATTGATA	19	7	3.43	0.04
.....TTTATTGAAGGAATTGATA	21	7	3.29	0.04
.....TTATTGAAGGAATTGATATA	21	7	3.29	0.03
.....AACATATATGTGACGTAATGAAAT	24	5	3.20	0.05
.....AACATATATGTGACGTAAT	21	6	3.17	0.05
.....TTTATTGAAGGAATTGATATATG	23	7	3.14	0.05
.....TTTTTATTGAAGGAATTGATATAATGCC	27	6	3.00	0.04
.....TTTATTGAAGGAATTGATAT	21	7	3.00	0.04
.....TATTGAAGGAATTGATATAATGCC	24	7	2.86	0.03
.....ATATATGTGACGTAATGAAATAAAC	25	5	2.80	0.03
.....TTTATTGAAGGAATTGATATAAT	23	7	2.71	0.04
.....TTATTGAAGGAATTGATA	20	7	2.71	0.03
.....TATTGAAGGAATTGATATATG	22	7	2.71	0.03
.....CATATATGTGACGTAATGAAATAAAC	27	5	2.60	0.04
.....GAAGGAATTGATATATGC	18	7	2.57	0.04
.....TGTATTTTAATTCTTCAA	22	7	2.57	0.04
.....ATTGAAGGAATTGATATAATGCCA	23	7	2.57	0.03
.....TTTATTGAAGGAATTGATATAATGCCAGT	28	7	2.43	0.03
.....TTTATTGAAGGAATTGATATATG	24	7	2.43	0.03
.....ATTGATATATGCCAGTAAAATG	22	7	2.43	0.04
.....GTATTTTAATTCTTCAATAA	23	7	2.43	0.03
.....GAAGGAATTGATATAATGCC	19	7	2.43	0.03
.....CATATATGTGACGTAATGAAATAAAC	28	5	2.40	0.03
.....ATATATGTGACGTAATGAAATAAAC	26	5	2.40	0.03
.....ATAAACATATATGTGACGTAAT	22	6	2.33	0.04
.....TTTATTGAAGGAATTGATA	22	7	2.29	0.03
.....TTTTTATTGAAGGAATTGATA	21	6	2.17	0.03
.....GTATTTTAATTCTTCAA	20	7	2.14	0.04
.....TTATTGAAGGAATTGATATAATGCCA	25	7	2.14	0.03
.....GATATATGCCAGTAAAATG	19	7	2.00	0.04
.....ATTGAAGGAATTGATATAT	19	7	2.00	0.04
.....TTATTGAAGGAATTGATAT	19	7	2.00	0.03
.....AACATATATGTGACGTAATGAA	22	6	2.00	0.03
.....TTTTTATTGAAGGAATTGATA	22	6	2.00	0.03
.....TTTTTATTGAAGGAATTGATAT	24	6	2.00	0.02
.....TATTGAAGGAATTGATATC	19	5	2.00	0.03
.....ATATGTGACGTAATGAAAT	19	7	1.86	0.03
.....TTATTGAAGGAATTGATATAATGCCA	26	7	1.86	0.03

.....TTATTGAAGGAATTGATATATGCCAGT.....	27	7	1.86	0.02
.....TTTTTATTGAAGGAATTGATAT.....	23	6	1.83	0.03
.....AACATATATGTGACGTAATGAAATA.....	25	5	1.80	0.03
.....CATATATGTGACGTAATGAAATAAC---CAA.....	29	5	1.80	0.02
.....TATTGAAGGAATTGATATATGCT.....	23	7	1.71	0.02
.....TTTTATTGAAGGAATTGATATATGCCAGT.....	29	7	1.71	0.02
.....ATTGAAGGAATTGATATATGCCCT.....	23	7	1.71	0.02
.....TTTTTATTGAAGGAATTGATA.....	23	6	1.67	0.02
.....TTTTTATTGAAGGAATTGATATATGCC.....	28	6	1.67	0.02
.....ATATATGTGACGTAATGAAATAAC---CAA.....	28	5	1.60	0.02
.....ACATATATGTGACGTAATGAAAT.....	23	5	1.60	0.03
.....ATTGAAGGAATTGATATATGCA.....	22	7	1.57	0.02
.....TTTATTGAAGGAATTGATATAT.....	22	7	1.57	0.02
.....ATAAACATATATGTGACGTC.....	19	7	1.57	0.03
.....TTTATTGAAGGAATTGATAT.....	20	7	1.57	0.02
.....TTTTTTATTGAAGGAATTGATATA.....	26	6	1.50	0.02
.....TTTTTATTGAAGGAATTG.....	19	6	1.50	0.02
.....TTTTTATTGAAGGAATTGATATATGCCA.....	27	7	1.43	0.02
.....TGTATTTTAATTCTTCAATA.....	24	7	1.43	0.02
.....TATTTTTAATTCTTCAAT.....	20	7	1.43	0.03
.....AACATATATGTGACGTAATGAAATAAC.....	29	5	1.40	0.02
.....ATATATGTGACGTAATGAAA.....	20	6	1.33	0.02
.....TTTTTATTGAAGGAATTGATATA.....	23	6	1.33	0.01
.....TTTTTATTGAAGGAATTGATATA.....	24	6	1.33	0.02
.....TTTTTTTATTGAAGGAATTGATATATGC.....	29	6	1.33	0.03
.....TGTATTTTAATTCTTCAAA.....	22	7	1.29	0.03
.....TATTGAAGGAATTGATATATA.....	21	7	1.29	0.01
.....CATAATATGTGACGTAATGAAAT.....	22	5	1.20	0.03
.....ATATATGTGACGTAATGA.....	18	6	1.17	0.02
.....CATAATATGTGACGTAATGAA.....	20	6	1.17	0.02
.....AACATATATGTGACGTAATGA.....	21	6	1.17	0.02
.....AACATATATGTGACGTAATG.....	20	6	1.17	0.02
.....TTTTTATTGAAGGAATTGATATATG.....	26	6	1.17	0.02
.....TTGAAGGAATTGATATATGCC.....	23	7	1.14	0.01
.....TGTATTTTAATTCTTCAATA.....	24	7	1.14	0.02
.....TATTTTTAATTCTTCAATA.....	21	7	1.14	0.02
.....CAGTAAAATGGTGTATTTTAAT.....	23	7	1.14	0.02
.....TGTATTTTAATTCTTCAAT.....	23	7	1.14	0.02
.....AAGGAATTGATATATGCC.....	18	7	1.14	0.02
.....ATTTTTAATTCTTCAATA.....	20	7	1.14	0.02
.....CAGTAAAATGGTGTATTTTAATTCTT.....	28	7	1.14	0.02
.....TTGAAGGAATTGATATATGCCAG.....	23	7	1.14	0.02
.....TTTTTAATTCTTCAATA.....	19	7	1.00	0.01
.....ATTGATATATGCCAGTAAAATGGTGT.....	27	7	1.00	0.02
.....TATGTGACGTAATGAAATA.....	19	5	1.00	0.01
.....ATAAACATATATGTGACGTA.....	20	6	1.00	0.02
.....TATATGTGACGTAATGAAATAAC.....	24	5	1.00	0.02
.....ATATATGTGACGTAATGAAATAAC---CAA.....	29	5	1.00	0.02
.....TTTTTATTGAAGGAATTGATATATGCCAG.....	29	6	1.00	0.02
.....CAGTAAAATGGTGTATTTT.....	20	7	1.00	0.02
.....TGAAGGAATTGATATATGCCAG.....	22	7	1.00	0.01
.....ATTGAAGGAATTGATATATGCCAG.....	24	7	1.00	0.01
.....ATAAACATATATGTGACGTAATGAAAT.....	27	5	1.00	0.02
.....TATATGTGACGTAATGAAAT.....	20	5	1.00	0.02
.....TATTTTTAATTCTTCAATAA.....	22	7	1.00	0.01
.....ATAAACATATATGTGACGTAATGAAATA.....	28	5	1.00	0.02
.....TTTTTTATTGAAGGAATTG.....	20	6	1.00	0.01
.....TTGATATATGCCAGTAAAATG.....	21	7	0.86	0.01
.....GTGTATTTTAATTCTTCAAT.....	23	7	0.86	0.02
.....GTATTTTAATTCTTCAAT.....	22	7	0.86	0.02
.....ATTGATATAATGCCAGTAAAATGGTGT.....	26	7	0.86	0.01
.....TGAAGGAATTGATATATGCCAGTA.....	24	7	0.86	0.01
.....TATTGAAGGAATTGATATATGCCA.....	26	7	0.86	0.01
.....GAATTGATATATGCCAGTAAAATG.....	24	7	0.86	0.01
.....GAAGGAATTGATATATGCCAGTAAAATG.....	28	7	0.86	0.01
.....GTATTTTAATTCTTCAATAA.....	24	7	0.86	0.01
.....GAAGGAATTGATATATGCCA.....	20	7	0.86	0.01

	ATAAACATATATGTGACGTAATGA	24	6	0.83	0.02
TTTTTTTATTGAAGGAATT	.ACATATATGTGACGTAATGA	19	6	0.83	0.02
	.CATATATGTGACGTAATGA	20	6	0.83	0.02
	.TTCTTCAATAAACATATATGTGAC	19	6	0.83	0.02
TTTTTATTGAAGGAATTGATATAT	.AACATATATGTGACGTAATGAAATAAC	25	6	0.83	0.02
	---C	24	6	0.83	0.02
	.CATATATGTGACGTAATGAAATA	29	5	0.80	0.02
	.ATATATGTGACGTAATGAAATAAC	23	5	0.80	0.02
	---CA	27	5	0.80	0.02
	.ATATATGTGACGTAATGAAATAA	23	5	0.80	0.02
	.TATATGTGACGTAATGAAATA	21	5	0.80	0.02
CAGTAAAATGGGTATTTTA		21	7	0.71	0.02
TGAAGGAATTGATATATGCCAGT		23	7	0.71	0.01
CAGTAAAATGGGTATTTTAATTCT		27	7	0.71	0.01
ATTGAAGGAATTGATATATGCCAGT		25	7	0.71	0.01
ATTTTTAATTCTTCAATAAA		22	7	0.71	0.01
ATTCTTCAATAAACATATATGTGAC		27	6	0.67	0.01
AATAAACATATATGTGACGT		20	6	0.67	0.01
TTTTTTTATTGAAGGAATTGA		22	6	0.67	0.01
A-TTTTTTTATTGAAGGAATTGATAT		26	6	0.67	0.01
TTTTTTTATTGAAGGAATTGATATATGCC		29	6	0.67	0.01
TTTTTTTATTGAAGGAATTGATATATG		27	6	0.67	0.01
	AAACATATATGTGACGTAATGAAA	24	6	0.67	0.01
	.CATATATGTGACGTAATGAAA	21	6	0.67	0.01
	.TTCTTCAATAAACATATATGTGACGT	27	6	0.67	0.01
TTTTTATTGAAGGAATTGATATATG		25	6	0.67	0.01
A-TTTTTTTATTGAAGGAAT		20	6	0.67	0.01
	ATGTGACGTAATGAAATAAC	22	5	0.60	0.01
	---C	24	5	0.60	0.01
	.ATATATGTGACGTAATGAAATAA	24	5	0.60	0.01
	.ACATATATGTGACGTAATGAAATA	19	7	0.57	0.01
TTTATTGAAGGAATTGAT		26	7	0.57	0.01
	AGTAAAATGGGTATTTTAATTCT	23	7	0.57	0.01
	.TATTTTAATTCTTCAATAAA	22	7	0.57	0.01
TATTGAAGGAATTGATATATCT		25	7	0.57	0.01
	.GGAATTGATATATGCCAGTAAATG	18	7	0.57	0.01
TTGAAGGAATTGATATAT		23	7	0.57	0.01
	TATTTTAATTCTTCAATAAC	27	7	0.57	0.01
TATTGAAGGAATTGATATATGCCAGTA		23	7	0.57	0.01
TTGAAGGAATTGATATATGCCAT		28	7	0.57	0.01
TGAAGGAATTGATATATGCCAGTAAAT		26	7	0.57	0.01
	CAGTAAAATGGGTATTTTAATTCT	2792.35	27.26		
Total					

.....TGAAGGAATTGATATATGCCA.....	21	16	0.94	0.02
..TTATTGAAGGAATTGATATATGCCA.....	25	16	0.88	0.02
....TTTATTGAAGGAATTGATA.....	19	16	0.88	0.02
.....TGTATTTTAATTCTTCAATAA.....	24	14	0.86	0.02
.....GTATTTTAATTCTTCAATA.....	22	14	0.86	0.02
....TTATTGAAGGAATTGATATATGCCAG.....	26	16	0.81	0.01
....TTGAAGGAATTGATATATGCC.....	22	16	0.81	0.01
....TTTATTGAAGGAATTGATATAATGC.....	24	16	0.69	0.01
.....TGTATTTTAATTCTTCA.....	20	14	0.57	0.02
....TTTATTGAAGGAATTGATATATGC.....	25	16	0.56	0.01
....TTATTGAAGGAATTGATATA.....	21	16	0.50	0.01
....TTTATTGAAGGAATTGATA.....	20	16	0.50	0.01
....TTATTGAAGGAATTGATATATGCC.....	24	16	0.44	0.01
....TGAAGGAATTGATATATGCC.....	20	16	0.44	0.01
....TTATTGAAGGAATTGATA.....	18	16	0.44	0.01
....TTTATTGAAGGAATTGATATA.....	21	16	0.38	0.01
....TATTGAAGGAATTGATATATGCCAGT.....	26	16	0.38	0.01
....TTTATTGAAGGAATTGATATAATG.....	23	16	0.38	0.01
....TGAAGGAATTGATATATGCCAG.....	22	16	0.38	0.01
.....GTATTTTAATTCTTCA.....	19	14	0.36	0.01
.....TGTATTTTAATTCTTCAATAAA.....	25	14	0.36	0.01
....TATTGAAGGAATTGATATATGC.....	23	16	0.31	0.01
....ATTTTATTGAAGGAATTGATATATGC.....	26	16	0.31	0.01
....TTTATTGAAGGAATTGATATA.....	22	16	0.31	0.01
.....TATATTTTGAC--AATGAAATAAAC---A.....	25	16	0.31	0.01
....TGAAGGAATTGATATATGCCAGT.....	23	16	0.31	0.01
.....GTATTTTAATTCTTCAAA.....	20	14	0.29	0.01
....GAAGGAATTGATATATGCCA.....	20	16	0.25	0.01
....TTGAAGGAATTGATATATGCCAG.....	23	16	0.25	0.00
....A----ATTTATTGAAGGAATTGATATATGC.....	27	16	0.25	0.01
....TTATTGAAGGAATTGATAT.....	19	16	0.25	0.00
....ATTTTATTGAAGGAATTGATATATGCC.....	27	16	0.25	0.01
....TTTATTGAAGGAATTGATATATGCC.....	26	16	0.25	0.01
....TTTATTGAAGGAATTGATATATGCCAG.....	27	16	0.25	0.00
....ATTGAAGGAATTGATATATGCCAG.....	24	16	0.25	0.00
....TATTGAAGGAATTGATAT.....	19	16	0.19	0.00
....TTGATATATGCCAGTAAAATG.....	21	16	0.19	0.01
....TATTGAAGGAATTGATAT.....	18	16	0.19	0.01
....TATTGAAGGAATTGATATATGCCAGTA.....	27	16	0.19	0.01
....TTTATTGAAGGAATTGATATATGCCA.....	27	16	0.19	0.00
....ATTGAAGGAATTGATATATGCC.....	23	16	0.19	0.00
....TATTGAAGGAATTGATATATGCC.....	23	16	0.19	0.01
....GAAGGAATTGATATATGCC.....	19	16	0.19	0.01
....TTTATTGAAGGAATTGATATATGCCA.....	26	16	0.19	0.01
....TGAAGGAATTGATATATGC.....	19	16	0.19	0.00
....GAAGGAATTGATATATGC.....	18	16	0.19	0.01
....TTTATTGAAGGAATTGATATATGCC.....	25	16	0.19	0.00
.....TTTTAATTCTTCAATA.....	18	14	0.14	0.00
....ATGGGTATTTTAATTCTTCAAT.....	26	14	0.14	0.00
....TGTATTTTAATTCTTCAATAA.....	25	14	0.14	0.00
....TATTTTAATTCTTCAAT.....	20	14	0.14	0.00
....AATGGGTATTTTAATTCTTCAAT.....	27	14	0.14	0.00
....GTATTTTAATTCTTCAATAA.....	23	14	0.14	0.00
....ATTTTAATTCTTCAATA.....	20	14	0.14	0.00
....ATTTTAATTCTTCAATAA.....	21	14	0.14	0.00
....TGTATTTTAATTCTTCAATA.....	23	14	0.14	0.00
....ATTTTAATTCTTCAAT.....	19	15	0.13	0.00
....TATTGAAGGAATTGATAT.....	20	16	0.13	0.01
....TTATTGAAGGAATTGATATATGCCAGTA.....	28	16	0.13	0.00
....TATTGAAGGAATTGATATATGCC.....	24	16	0.13	0.01
....ATTGAAGGAATTGATATATGCCA.....	23	16	0.13	0.00
....TCA----ATTTATTGAAGGAATT.....	20	16	0.13	0.00
....ATTGAAGGAATTGATATATGCCAGT.....	25	16	0.13	0.00
....A----ATTTATTGAAGGAATTGATA.....	22	16	0.13	0.00
....TGAAGGAATTGATATATGCCA.....	22	16	0.13	0.00
....TTTATTGAAGGAATTGATAT.....	20	16	0.13	0.00
....ATTGATATATGCCAGTAAAATG.....	22	16	0.13	0.00

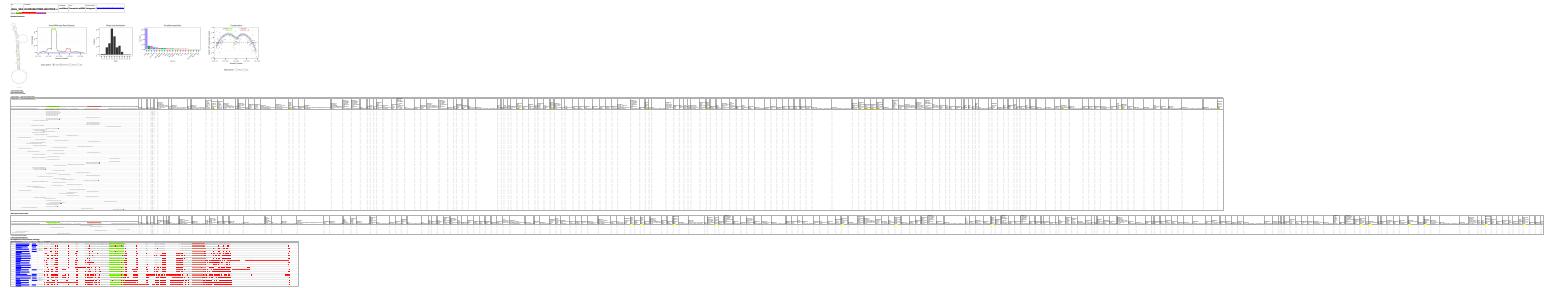
TATTGAAGGAATTGATATATT	21	16	0.13	0.01
TATTGAAGGAATTGATATATGC A	23	16	0.13	0.01
TATTGAAGGAATTGATAT T	19	16	0.13	0.01
TTGAAGGAATTGATATATGCCAGTA	25	16	0.13	0.01
.....TTTTAATTCTTTCAAT	18	16	0.13	0.00
.....GATATATGCCAGTAAATG	19	16	0.13	0.00
.....GTGTATTTAATTCTTT	19	16	0.13	0.00
TTGAAGGAATTGATATAATGCCAGT	24	16	0.13	0.00
ATTTTATTGAAGGAATTGATATATGCCA	28	16	0.13	0.00
TTTATTGAAGGAATTGAT	19	16	0.13	0.01
TTTATTGAAGGAATTGATATAT	22	16	0.13	0.00
TATTGAAGGAATTGATAT A	21	16	0.13	0.00
TATTGAAGGAATTGATATATGT	22	16	0.13	0.01
ATTTTATTGAAGGAATTGATATA	23	16	0.13	0.01
ATTGAAGGAATTGATATA	18	16	0.13	0.00
ATTGAAGGAATTGATATAT	19	16	0.13	0.00
TATTGAAGGAATTGATATATGCCAGTAA	28	16	0.13	0.01
.....GTATTTTAATTCTTTCAATAAA	24	14	0.07	0.00
.....TGTATTTTAATTCTTTCAATA C	25	14	0.07	0.00
.....TTTTAATTCTTTCAATA	19	14	0.07	0.00
.....TATTTTAATTCTTTCAATAAA	23	14	0.07	0.00
.....ATTCTTTCAATAAAATATA	19	14	0.07	0.00
.....TATTTTAATTCTTTCAATA	21	14	0.07	0.00
.....TGTATTTTAATTCTTTCAATA T	24	14	0.07	0.00
.....TGTATTTTAATTCTTTCAAT G	23	14	0.07	0.00
.....TGTATTTTAATTCTTTCAATA C	24	14	0.07	0.00
.....TTCTTTCAATAAAATATTTTGAC--A	26	14	0.07	0.00
.....TTCTTTCAATAAAATATTTTGAC--AAT	28	14	0.07	0.00
.....TGTATTTTAATTCTTTCAATA C	25	14	0.07	0.00
.....TGTATTTTAATTCTTTCAATA G	25	14	0.07	0.00
.....ATGGGTATTTTAATTCTTTCAATA	27	14	0.07	0.00
.....ATGGTATTTTAATTCTTTCA	21	14	0.07	0.00
.....ATATATGCCAGTAAATG	20	16	0.06	0.00
.....TTTTTGAC--AATGAAATAAAC---AAA	23	16	0.06	0.00
.....GAAGGAATTGATATATGCCAGT	22	16	0.06	0.00
.....TTTATTGAAGGAATTGATATATGCCAG	28	16	0.06	0.00
.....TATTGAAGGAATTGATATATGCC G	24	16	0.06	0.00
.....AAATGGGTATTTTAATTCTTT	24	16	0.06	0.00
.....ATTGAAGGAATTGATATATGCC G	22	16	0.06	0.00
.....AATATATTTTGAC--AATGAAATAAAC---A	27	16	0.06	0.00
.....TGAAGGAATTGATATATGCCAGTAA	25	16	0.06	0.00
.....TTGATATATGCCAGTAAATGGTGT	26	16	0.06	0.00
.....TATTGAAGGAATTGATAT C	20	16	0.06	0.00
.....TCA---ATTTTATTGAAGGAATTGATA	24	16	0.06	0.00
.....GAAGGAATTGATATATGCCAG	21	16	0.06	0.00
.....TATTGAAGGAATTGATATA A	20	16	0.06	0.00
.....TTGAAGGAATTGATATATGCC A	24	16	0.06	0.00
.....ATATATGCCAGTAAATGGTGT	23	16	0.06	0.00
.....TTGAAGGAATTGATATAT	18	16	0.06	0.00
.....AAAAATGGGTATTTTAATT	20	16	0.06	0.00
.....A---ATTTTATTGAAGGAATTGATATAT	25	16	0.06	0.00
.....CA---ATTTTATTGAAGGAATTGATATATGC	28	16	0.06	0.00
.....TTTATTGAAGGAATTGAT	18	16	0.06	0.00
.....CAGTAAAATGGGTATTTT	19	16	0.06	0.00
.....AATATATTTTGAC--AATGAA	20	16	0.06	0.00
.....TTATTGAAGGAATTGATA CG	20	16	0.06	0.00
.....GATATATGCCAGTAAATGGTGT	23	16	0.06	0.00
.....TTGAAGGAATTGATATATGCCAGTA	26	16	0.06	0.00
.....TTGAAGGAATTGATATATGCC CC	24	16	0.06	0.00
.....ATTTTATTGAAGGAATTG	18	16	0.06	0.00
.....TATGCCAGTAAATGGTG	18	16	0.06	0.00
.....TTTATTGAAGGAATTGATATAT	23	16	0.06	0.00
.....ATTGAAGGAATTGATATATGCCAGTA	26	16	0.06	0.00
.....GAATTGATATATGCCAGTAAATGGTGT	28	16	0.06	0.00
.....TGAAGGAATTGATATATGCC T	21	16	0.06	0.00
.....TATTGAAGGAATTGATATATG A	22	16	0.06	0.00

		TGAAGGAATTGATATATGCCAG C	23	16	0.06	0.00
		TTATTGAAGGAATTGATATATT T	22	16	0.06	0.00
		ATTTTATTGAAGGAATTGATA	21	16	0.06	0.00
		AAGGAATTGATATATGCCAG	20	16	0.06	0.00
		AAGGAATTGATATATGCCA	19	16	0.06	0.00
		ATTGATATATGCCAGTAAA	19	16	0.06	0.00
		.A---ATTTTATTGAAGGAATT	18	16	0.06	0.00
		TTATTGAAGGAATTGATATA TG T	23	16	0.06	0.00
		TTTATTGAAGGAATTGATATATGCCAGT	28	16	0.06	0.00
		.GATATATGCCAGTAAAATGGT	21	16	0.06	0.00
		TATTGAAGGAATTGATATATGC T	24	16	0.06	0.00
		TTTGAC--AATGAAATAAAC----AAAAA	24	16	0.06	0.00
		TTGAC--AATGAAATAAAC----A	18	16	0.06	0.00
		TTTGAC--AATGAAATAAAC----A	20	16	0.06	0.00
		.AAATCA---ATTTTATTGAAGGAATT	22	16	0.06	0.00
		.A---ATTTTATTGAAGGAATTGATATATGCC	28	16	0.06	0.00
		TATTGAAGGAATTGATATATGCCAGT T	27	16	0.06	0.00
		.TGTATTTTAATTCTTT	18	17	0.06	0.00
		.AAATGGTGTATTTAAT	18	17	0.06	0.00
		TATTGAAGGAATTGATA AG	19	19	0.05	0.00
	Total				199.35	3.42

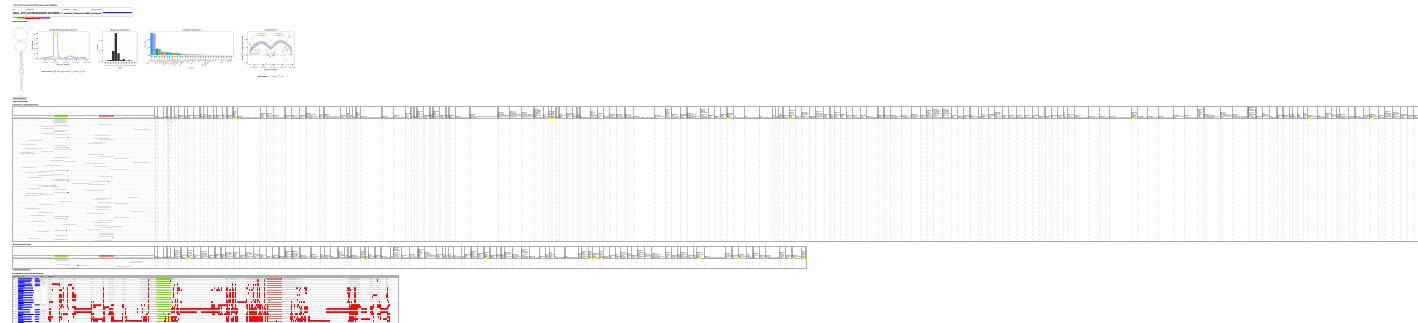
Generated: 08/02/2016 at 01:56 AM

Supplementary Figure S9.

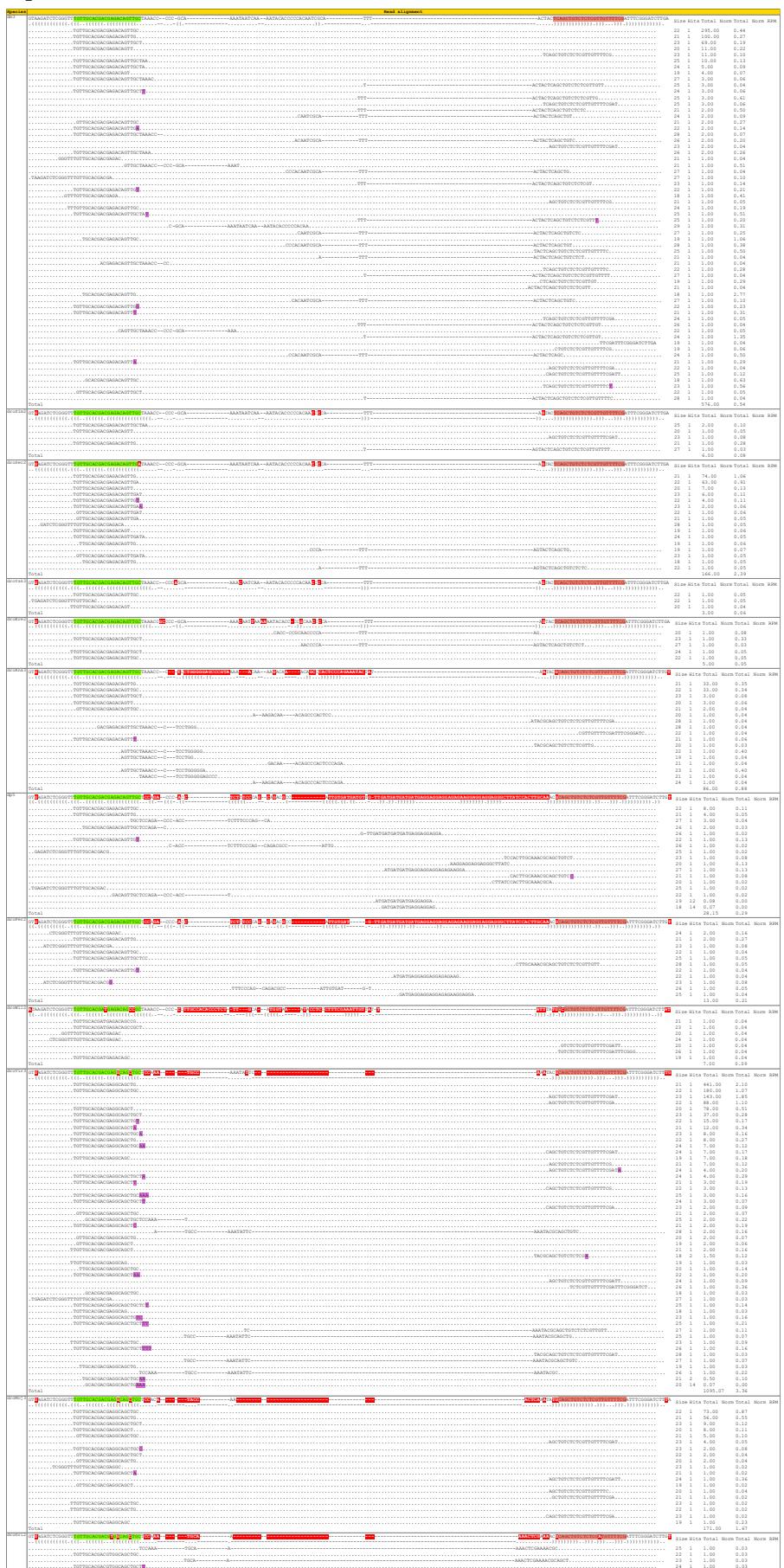
Alignment and small RNA read details for novel conserved miRNAs annotated in this study.

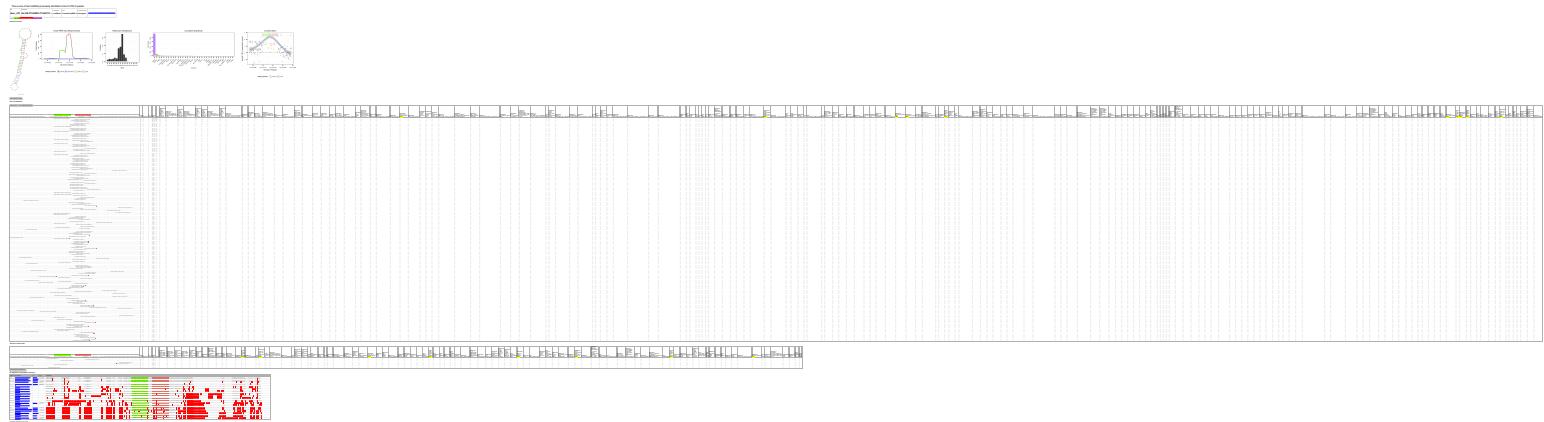




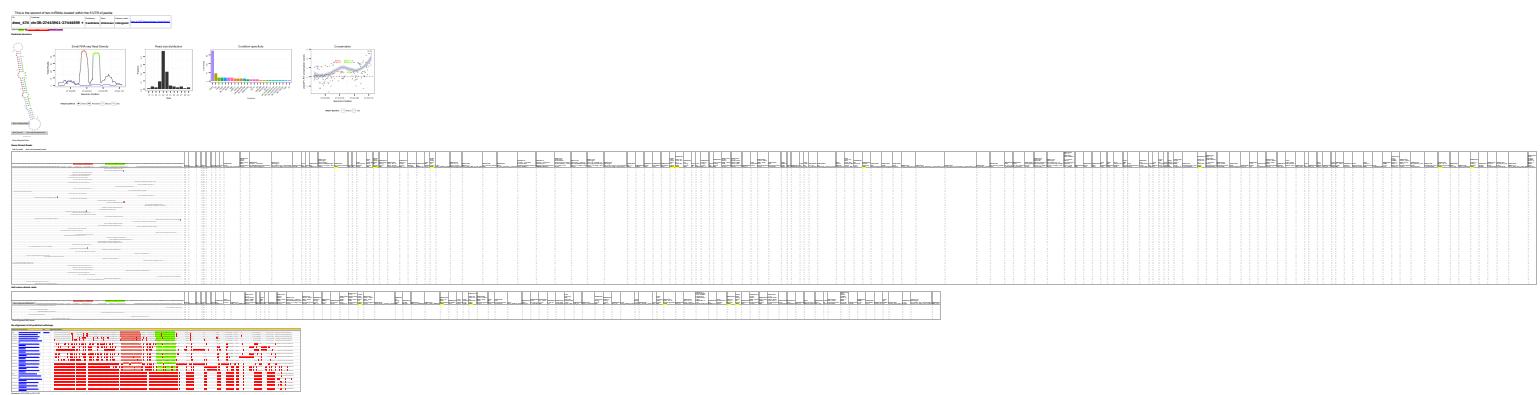


dme_373









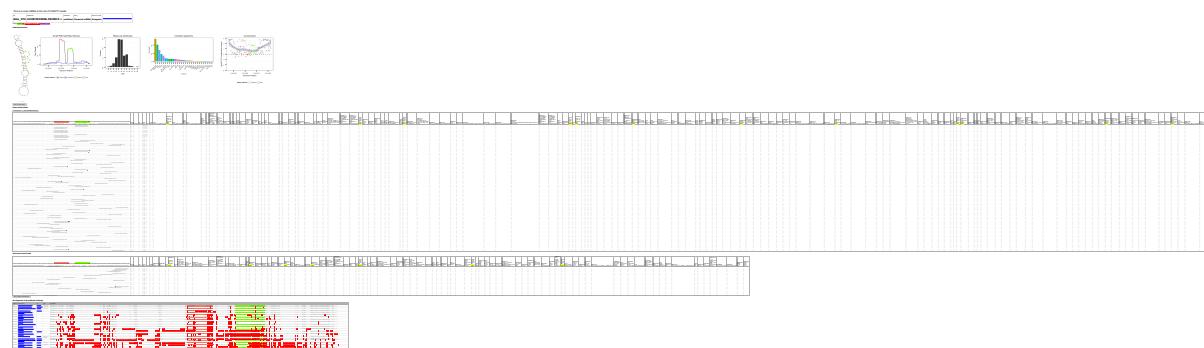
dme 474

Species	Read alignment									
	Size	Hits	Total	Norm	Total	Norm	RPM			
dm3	GTACTTGAATTGCGC AGCTTTGCTTGGGATTTGTCC-GA--CA-ATTAGGCAG AATCCAT CGCAAAGACGCAGC AAGTAGCATCGTTTA(((((((((.....((((((.(((-.-.-.-.))))))))))))....))))))))....	22	1	31.00	0.12					
AATCCATAGCGAAAGACGCAGC.....	22	1	9.00	0.10					
AATCCATAGCGAAAGACGCAGC.....	23	1	8.00	0.08					
CGCAGCTTTGCTTGGGATTT.....	23	1	5.00	0.09					
AGCTTTGCTTGGGATTTGT.....	22	1	5.00	0.12					
TGCGCAGCTTTGCTTGGGAT.....	23	1	4.00	0.27					
GCTTTGCTTGGGATTTGT.....	22	1	4.00	0.05					
TGCGCAGCTTTGCTTGGGAT.....	25	1	3.00	0.06					
AGACGCAGCAAGTAGCATCGTT.....	22	1	2.00	1.03					
GCTTTGCTTGGGATTTGT.....	21	1	2.00	0.09					
TGCGCAGCTTTGCTTGGG.....	21	1	2.00	0.07					
CGCAGCTTTGCTTGGGAT.....	20	1	1.00	0.05					
TTAGGCAGAACATCCATAGCGAA.....	21	1	1.00	0.16					
TCCATAGCGAAAGACGCAGC.....	22	1	1.00	0.05					
TTAGGCAGAACATCCATAGCGAAAGACGCAG.....	29	1	1.00	0.04					
CAGCTTTGCTTGGGATTTGT.....	23	1	1.00	0.06					
CGCAGCTTTGCTTGGGAT.....	21	1	1.00	0.28					
TTTTGCTTGGGATTTGTCC-GA--C.....	22	1	1.00	0.04					
CAGAACATCCATAGCGAAAGACGC.....	24	1	1.00	0.05					
AGGCAGAACATCCATAGCGAA.....	19	1	1.00	0.04					
AATTGCGCAGCTTTGCTTGGGAT.....	25	1	1.00	0.30					
AGCTTTGCTTGGGATTTTG.....	21	1	1.00	0.06					
AGAACATCCATAGCGAAAGAGA.....	18	1	1.00	0.12					
GCAGCTTTGCTTGGGATTTGTCC-GA--C.....	29	1	1.00	0.04					
ATCCATAGCGAAAGACGCAGC.....	21	1	1.00	0.05					
AATCCATAGCGAAAGACGCAG.....	21	1	1.00	0.38					
TTGCCAGCTTTGCTTGGGATTTGT.....	28	1	1.00	0.04					
AGCTTTGCTTGGGATTTGTCC-.....	24	1	1.00	0.04					
GCTTTGCTTGGGATTTGTCC-.....	23	1	1.00	0.04					
TTAGGCAGAACATCCATAGCGAAAGACCGA.....	27	1	1.00	0.04					
TTGCCAGCTTTGCTTGGGATTT.....	26	1	1.00	0.04					
TGCTTGGGATTTGTCC-GA--.....	20	1	1.00	0.50					
TTGCCAGCTTTGCTTGGG.....	22	1	1.00	0.04					
TACTTGAAATTGCGCAGCTTTGCTT.....	26	1	1.00	1.10					
AATCCATAGCGAAAGACGC.....	19	1	1.00	0.28					
TTTGCTTGGGATTTGTCC-.....	19	1	1.00	0.24					
TTTGCTTGGGATTTGTCC-.....	21	1	1.00	0.04					
TGAATTGCGCAGCTTTGCTTGG.....	24	1	1.00	0.51					
TCCATAGCGAAAGACGCAGCAAG.....	23	1	1.00	0.05					
TAGCGAAAGACGCAGCAAGTAGCATCG.....	27	1	1.00	0.08					
TTAGGCAGAACATCCATAGCGAAAGACGC.....	27	1	1.00	0.04					
	Total			106.00	0.14					
droSim2	GTACTTGAATTGCGC AGCTTTGCTTGGGATTTGTCC-GA--CA-ATTAGGCAG AATCCAT CGCAAAGACGCAGC AAGTAGCA CGTTTA(((((((((.....((((((.(((-.-.-.-.))))))))))))....))))-)))....	18	1	1.00	0.07					
TCC-GA--CA-ATTAGGCAGA.....	22	1	1.00	0.03					
AATCCATCGCAGAACACGCAGC.....	21	1	1.00	0.05					
TTTGCTC-GA--CA-ATTAGGC.....	19	1	1.00	0.05					
AATCCATCGCAGAACACGC.....	19	1	1.00	0.05					
AATCCATCGCAGAACACGCAGC-.....	22	1	1.00	0.05					
	Total			6.00	0.06					
droSec2	GTACTTGAATTGCGC AGCTTTGCTTGGGATTTGTCC-GA--CA-ATTAGGCAG AATCCAT CGCAAAGACGCAGC AAGTAGCA CGTTTA(((((((((.....((((((.(((-.-.-.-.))))))))))))....))))-)))....	22	2	1.00	0.07					
AATCCATCGCAGAACACGCAGC-.....	20	2	0.50	0.03					
CGCAGCTTTGCTTGGGAT.....			1.50	0.05					
	Total			1.50	0.05					
droYak3	CCACTTGAATTGCGC AGCTTTGCTTGGGATTTGTCC-GA--CA-ATTAGGCAG AATCCAT CGCAAAGACGCAGC AAGTAGCA CGTTTA(((((((((.....((((((.(((-.-.-.-.))))))))))))....))))-)))....	25	1	1.00	0.04					
C-AA--CAGATTAGGCAGAACATCCATCGC.....	20	1	1.00	0.37					
GCGCAGCTTTGCTTGGG.....	23	1	1.00	0.05					
AGCTTTGCTTGGGATTCGTG.....			3.00	0.06					
	Total			3.00	0.06					
droEre2	GTACTTGAATTGCGC TGCCTTGCTTGGGATTTGTCC-GA--CA-ATTAGGCAG AATCCAT CGCAAAGACGCAGC AAGTAGCA CGTTTA(((((((((.....((((((.(((-.-.-.-.))))))))))))....))))-)))....	22	1	2.00	0.65					
AATCCATCGCAGAACACGCAGC.....	19	1	1.00	0.05					
CTGCTTTGCTTGGGAT.....	19	1	1.00	0.05					
AGGCAGAACATCCATCGCAGA.....			4.00	0.17					
	Total			4.00	0.17					
droAna3	CACTTG CGCG-GC CACTTTGCTTGGGATTTGTCC-GA-CCA-CGCGACA-AATCCAT CGTAAAGAGTAAGCGCGAAGCAACR-AA(((((((-((.((((((.(((-.-.-.-.))))))))))))....))))-)))....	23	1	1.00	0.06					
CCACTTTGCTTGGCGTTTGT.....			1.00	0.06					
	Total			1.00	0.06					

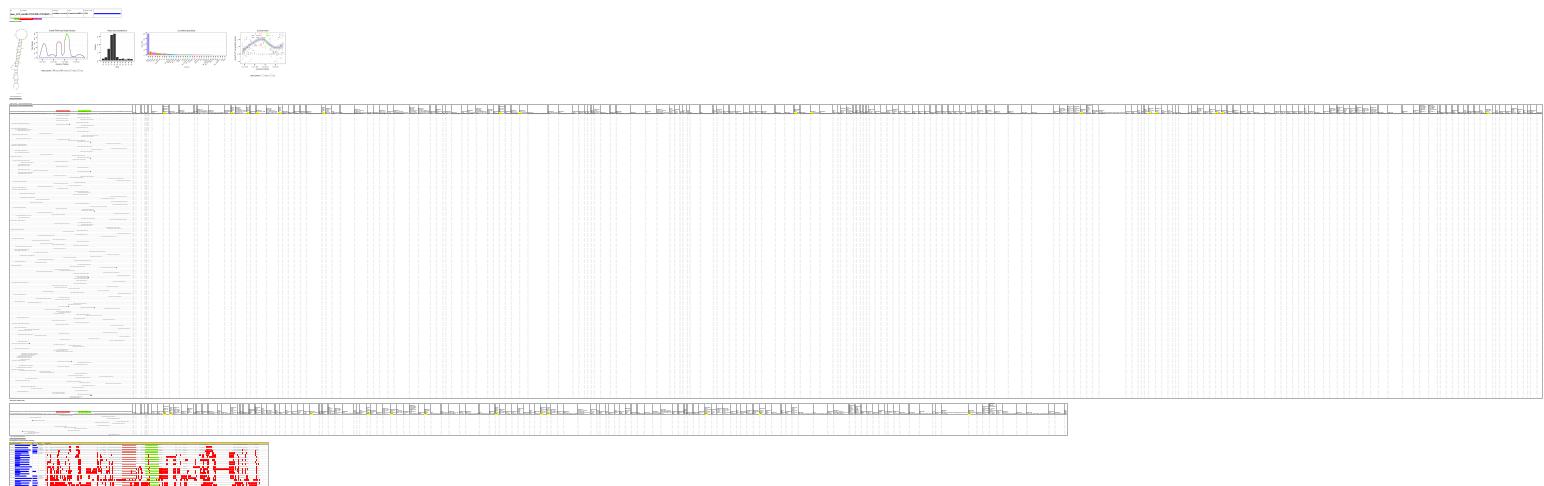


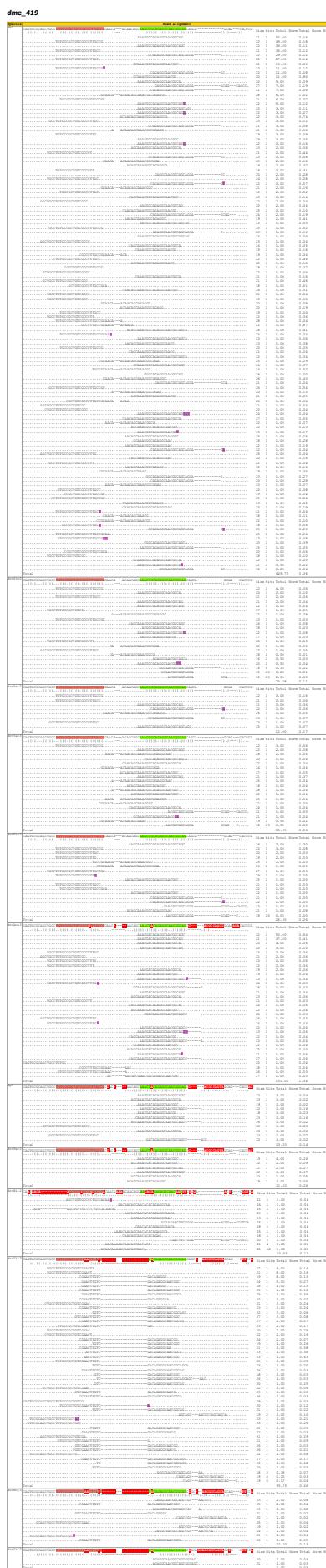
dme_245





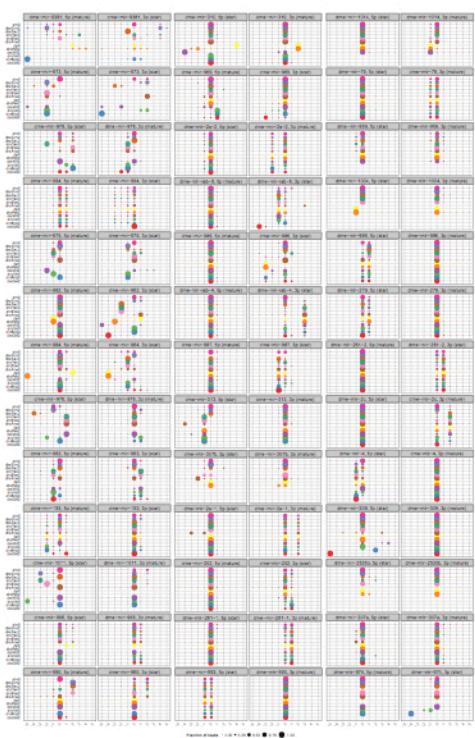






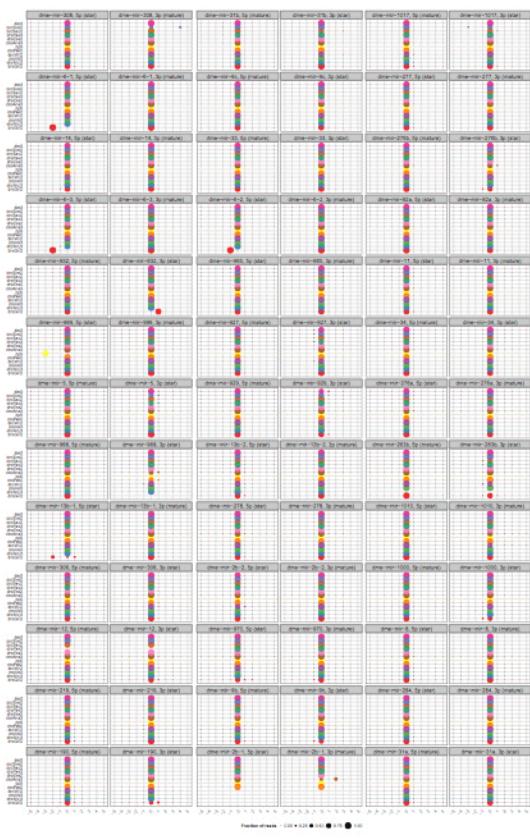


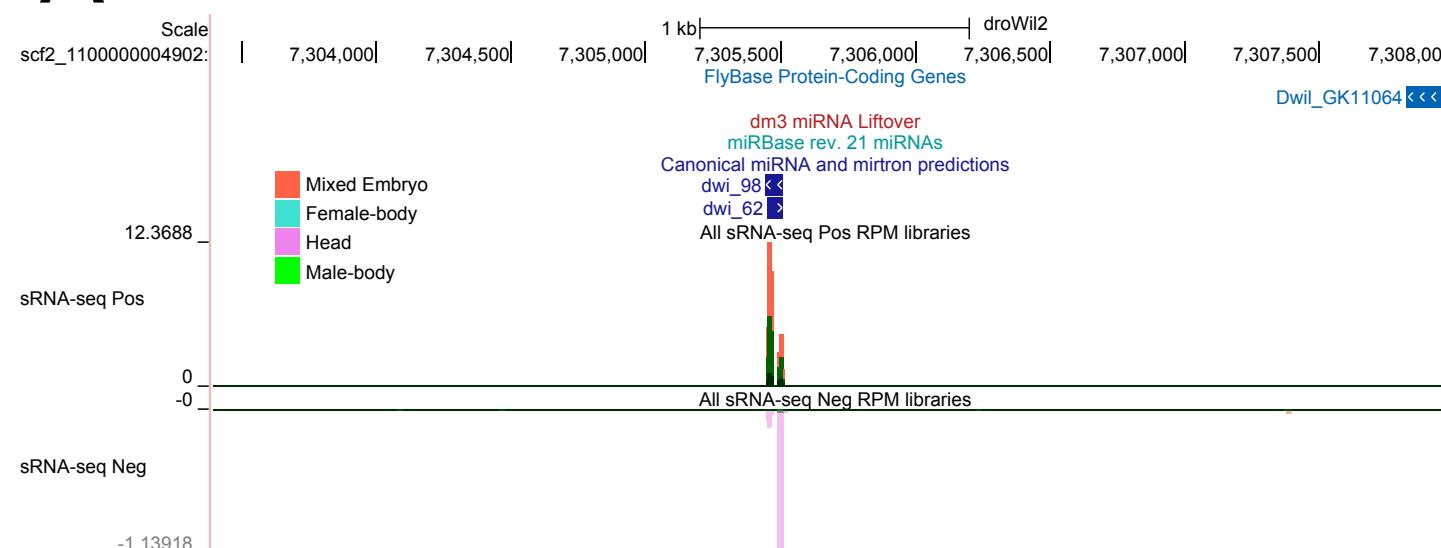
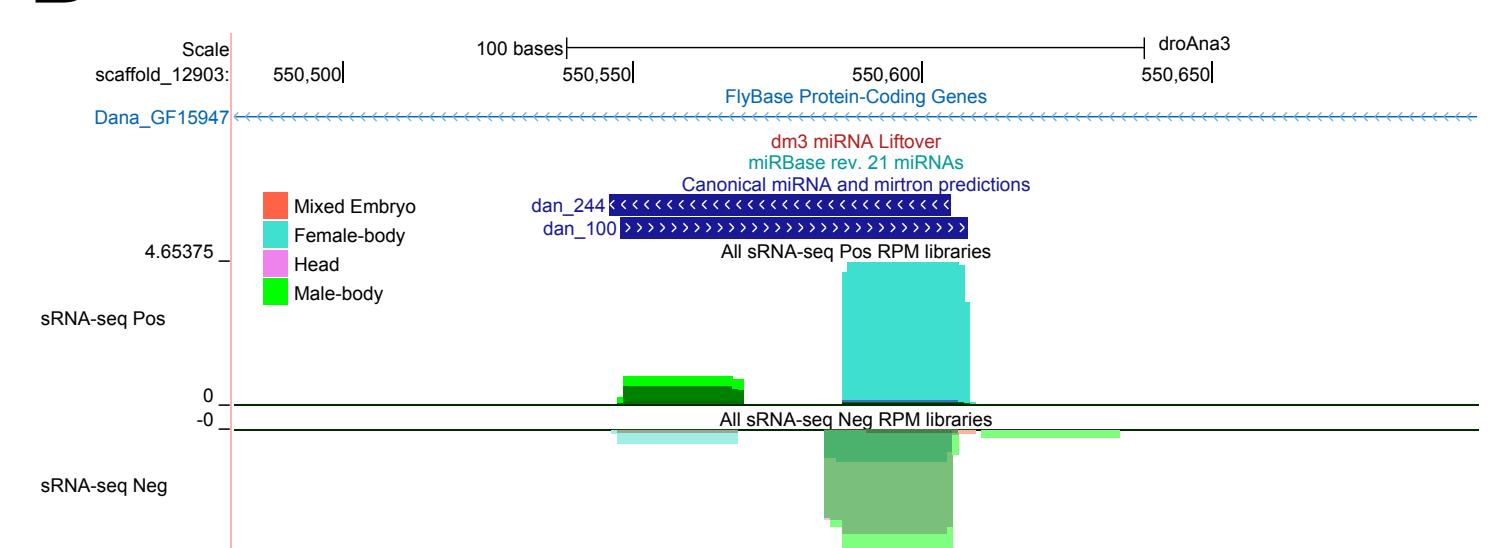




Supplementary Figure S10: Evolutionary patterns of 5' end cleavage precision for all conserved *D. melanogaster* miRNAs.

Mohammed et al
Supplementary Figure S10



A**B**

dwi_62

Sequence	Read	Hit	Total	Total	Size	Count	Norm	RPM
CTCTCGGTAGACTCCATGCCCCATGGACCATGCTACGGTTCACTTGTAAAATATTACGAGTCTAGAAGAGAA	22	1	498.00	4.96				
.....AATGCCATGGACCATGCTACGT.....	22	1	105.00	0.90				
.....GTAGTTGGTCATGGCATTG.....	22	1	92.00	0.92				
.....AATGCCATGGACCATGCTACGG.....	21	1	33.00	0.28				
.....AATGCCATGGACCATGCTACGTT.....	23	1	10.00	0.13				
.....TAGTTGGTCCATGGCATTGGA.....	22	1	8.00	0.10				
.....GTAGTTGGTCCATGGCATTGAT.....	20	1	6.00	0.14				
.....TAGTTGGTCCATGGCATTGG.....	21	1	3.00	0.06				
.....GTAGTTGGTCCATGGCATTGG.....	22	1	3.00	0.11				
.....AATGCCATGGACCATGCTACGG.....	22	1	3.00	0.11				
.....AATGCCATGGACCATGCTAC.....	20	1	3.00	0.06				
.....GTAGTTGGTCCATGGCATTGG.....	23	1	2.00	0.04				
.....TAGTTGGTCCATGGCATTG.....	20	1	2.00	0.05				
.....GTAGTTGGTCCATGGCATTG.....	19	1	2.00	0.12				
.....CGTAGTTGGTCCATGGCATTG.....	22	1	1.00	0.06				
.....GTAGTTGGTCCATGGCATTG.....	21	1	1.00	0.04				
.....TAGTTGGTCCATGGCATTG.....	21	1	1.00	0.04				
.....GTAGTTGGTCCATGGCATTG.....	23	1	1.00	0.04				
.....AATGCCATGGACCATGCTACGT.....	21	1	1.00	0.04				
.....CAATGCCATGGACCATGCTAC.....	21	1	1.00	0.04				
.....AATGCCATGGACCATGCTACG.....	23	1	1.00	0.04				
.....AATGCCATGGACCATGCTACG.....	23	1	1.00	0.04				
.....GTAGTTGGTCCATGGCATTG.....	22	1	1.00	0.06				

dwi_98

Sequence	Read	Hit	Total	Total	Size	Hits	Total	Total	Size	Count	Norm	RPM
TCTCTCTAGACTCCATGCCCCATGGACCAACTACGGTTATATTTCAGTAAAACTAGCATGGTCATGGCATGGAGTCTACGCAGAGA	22	1	36.00	0.38								
.....AATGCCATGGACCAACTACGT.....	22	1	10.00	0.44								
bioRxiv preprint doi: https://doi.org/10.1101/125997 ; this version posted April 11, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.												
.....GTAGTTGGTCATGGCATTGGAGT.....	22	1	1.00	0.04								
.....GTAGCATGGTCATGGCATTGG.....	22	1	1.00	0.04								
.....ATGCCATGGACCAACTACGT.....	22	1	1.00	0.06								
.....TAGCATGGTCATGGCATTGGAG.....	23	1	1.00	0.04								

dan_100

Sequence	Read	Hit	Total	Total	Size	Count	Norm	RPM
TATATCTTAATTATCAGTAAAGCCAAAGGGAGGTGGATTTCATAATGGATAATCAAACCTCCCTTTGGCTCTACTACTAATTGGATAATCAA	22	1	99.00	1.35				
.....CCTCCCTTTGGCTCTACTACT.....	22	1	36.00	0.51				
.....AGTAGAACCCAAAAGGATGTG.....	21	1	26.00	0.27				
.....CTCCCTTTGGCTCTACTACT.....	21	1	6.00	0.19				
.....CTCCCTTTGGCTCTACTACT.....	20	1	5.00	0.16				
.....CAGTAGAGGCCAAAAGGATGTG.....	22	1	5.00	0.15				
.....CCTCCCTTTGGCTCTACTACT.....	23	1	3.00	0.10				
.....AGTAGAACCCAAAAGGATGTG.....	19	1	2.00	0.06				
.....CCTCCCTTTGGCTCTACTACT.....	23	1	2.00	0.06				
.....AGTAGAACCCAAAAGGATGTG.....	23	1	2.00	0.06				
.....AGTAGAACCCAAAAGGATGTG.....	21	1	2.00	0.09				
.....GTAGAGGCCAAAAGGATGTG.....	20	1	1.00	0.03				
.....CCTCCCTTTGGCTCTACTACT.....	22	1	1.00	0.04				
.....AGTAGAACCCAAAAGGATGTG.....	21	1	1.00	0.06				
.....AGTAGAACCCAAAAGGATGTG.....	20	1	1.00	0.06				
.....AGTAGAACCCAAAAGGATGTG.....	20	1	1.00	0.06				

dan_244

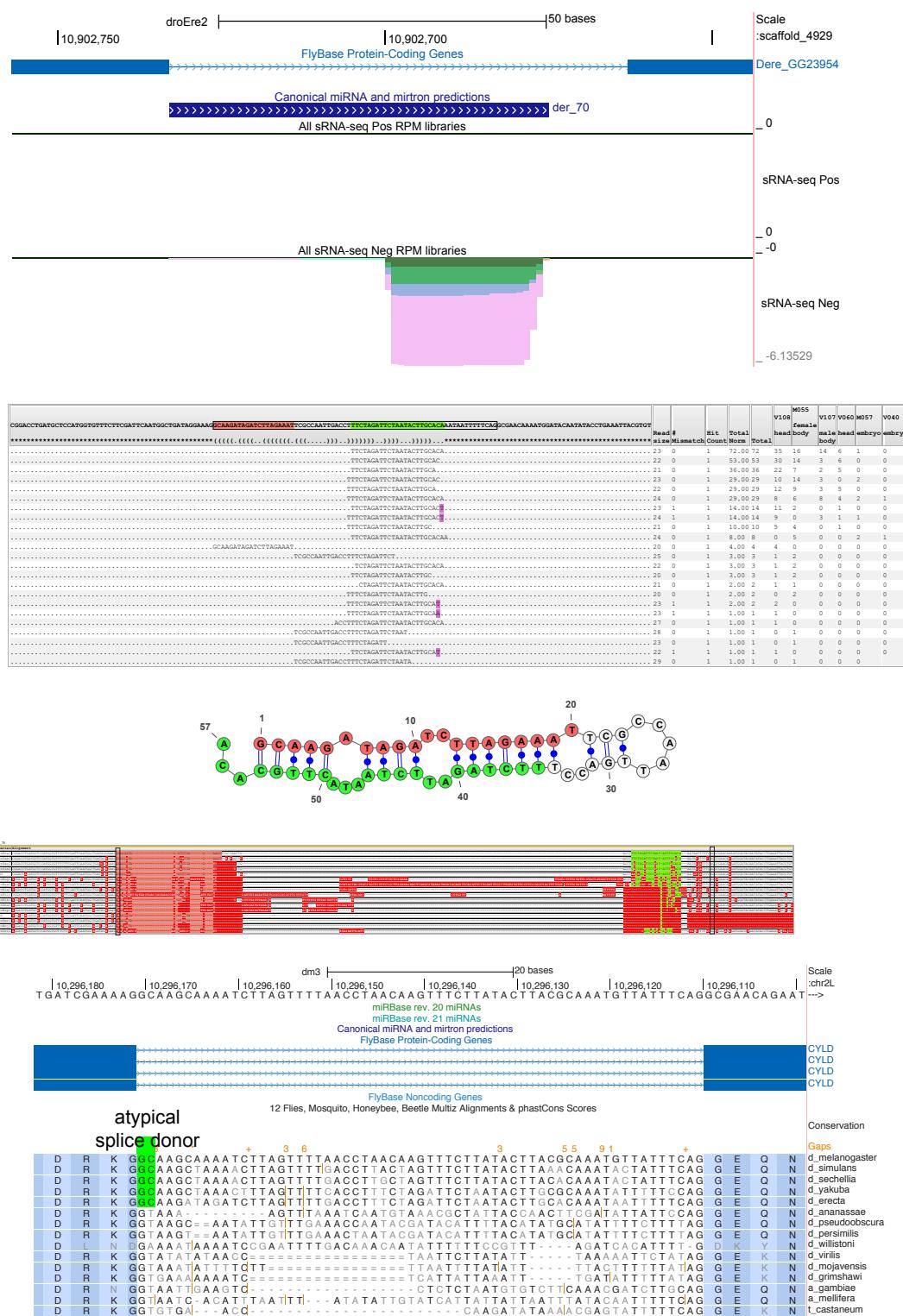
Sequence	Read	Hit	Total	Total	Size	Hits	Total	Total	Size	Count	Norm	RPM
GATTACCAATTAAAGTGTAGAGGCCAARGGGAGGTGGATTTCATAATGGATAATACAACATCCCTTTGGCTCTACTACTAATAATTAGATAATAC	22	1	24.00	0.37								
.....AGTAGAACCCAAAAGGAGGTGG.....	21	1	4.00	0.06								
.....AGTAGAACCCAAAAGGAGGTGG.....	19	1	4.00	0.12								
.....AGTAGAACCCAAAAGGAGGTG.....	19	1	4.00	0.12								
.....ACATCCCTTTGGCTCTACTACT.....	21	1	3.00	0.05								
.....TAGTAGAACCCAAAAGGAGGTG.....	20	1	3.00	0.18								
.....AGTAGAACCCAAAAGGAGGTG.....	22	1	2.00	0.09								
.....AGTAGAACCCAAAAGGAGGTG.....	21	1	1.00	0.06								
.....AAGTAGAACCCAAAAGGAGGTG.....	19	1	1.00	0.04								
.....AGTAGAACCCAAAAGGAGGTG.....	20	1	1.00	0.03								
.....ACATCCCTTTGGCTCTACTGA.....	22	1	1.00	0.03								
.....TAGTAGAACCCAAAAGGAGGTG.....	23	1	1.00	0.06								

Supplementary Figure S11: Additional examples of novel sense/antisense miRNA pairs.

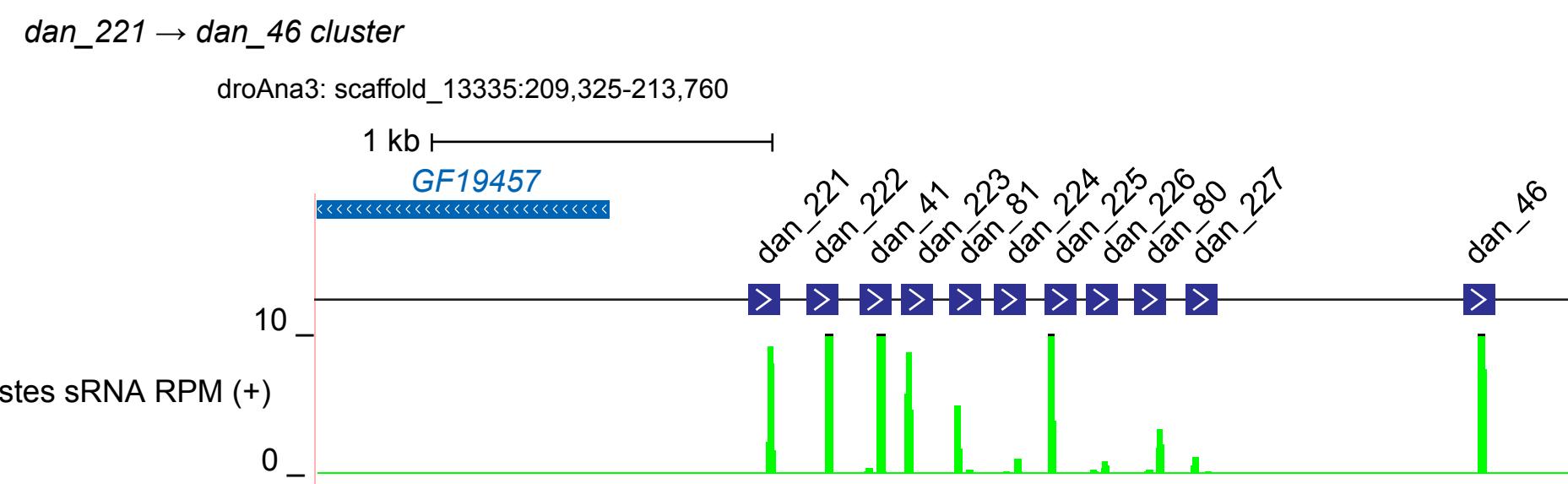
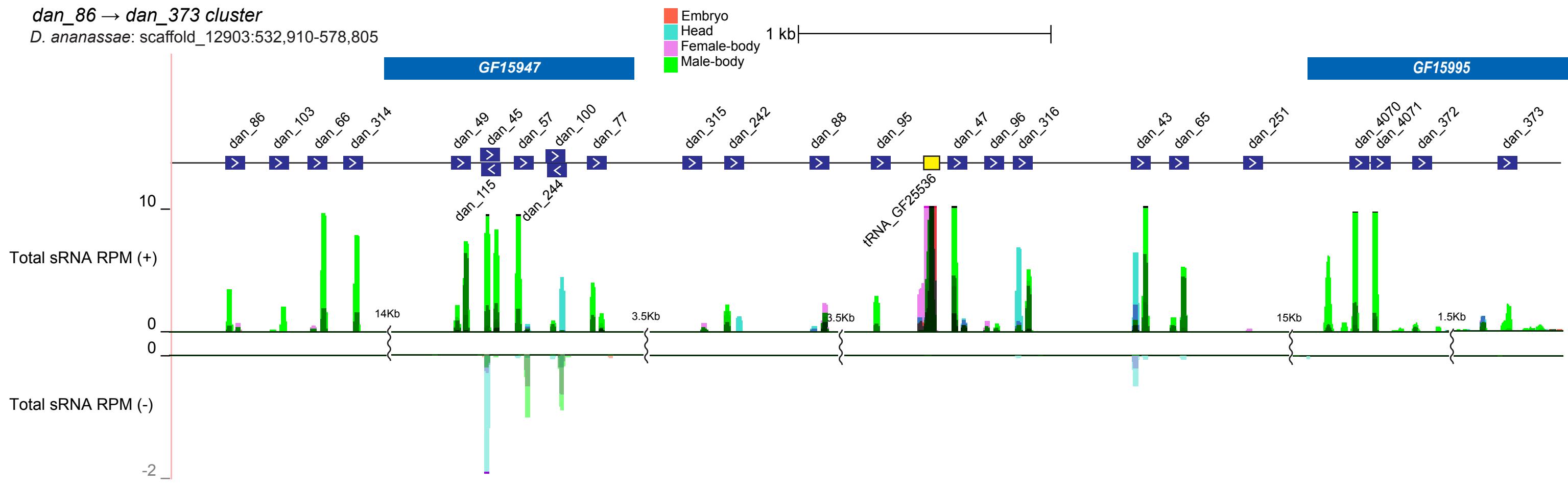
(A) Example of novel sense/antisense miRNA pair from *D. willistoni* (dwi_62/dwi-98). (B) Example of novel sense/antisense miRNA pair from *D. ananassae* (dan_100/dan_244).

Mohammed et al
Supplementary Figure S11

der_70
3'-tailed-mirtron



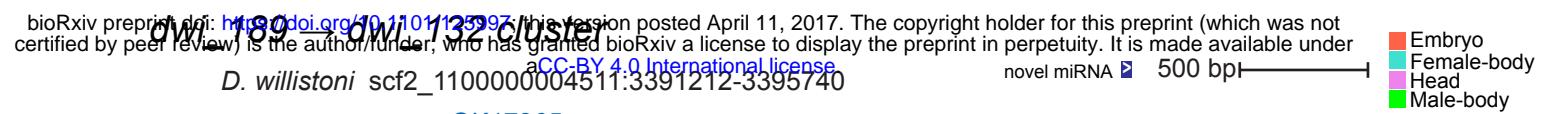
Supplementary Figure 12. Example of a novel, atypical 3' tailed mirtron. A 3' tailed mirtron from *D. erecta*, *Der_70*. This locus produces a dominant 3p miRNA, which is trimmed by ~13 from the splice acceptor site. Note that there is 3' untemplated uridylation (purple) associated with a subset of *Der_70*-3p reads. While *Der_70*-5p reads are modest, there are also apparent partially diced products that include the terminal loop that phase precisely with the *Der_70*-5p species, supporting this as a Dicer position. Bottom alignment shows that *Der_70* tailed mirtron resides in the conserved gene *CYLD*, and is associated with a non-canonical splice "GC" donor in the five melanogaster group species, which is instead a conventional "GT" splice donor in most other insects. Note *D. willistoni* seems to have lost both splice sites.



Supplementary Figure S13: Annotation of novel Testes-restricted, recently-evolved, clustered (TRC) miRNA clusters identified within *D. ananassae*.

bioRxiv preprint doi: <https://doi.org/10.1101/125997>; this version posted April 11, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

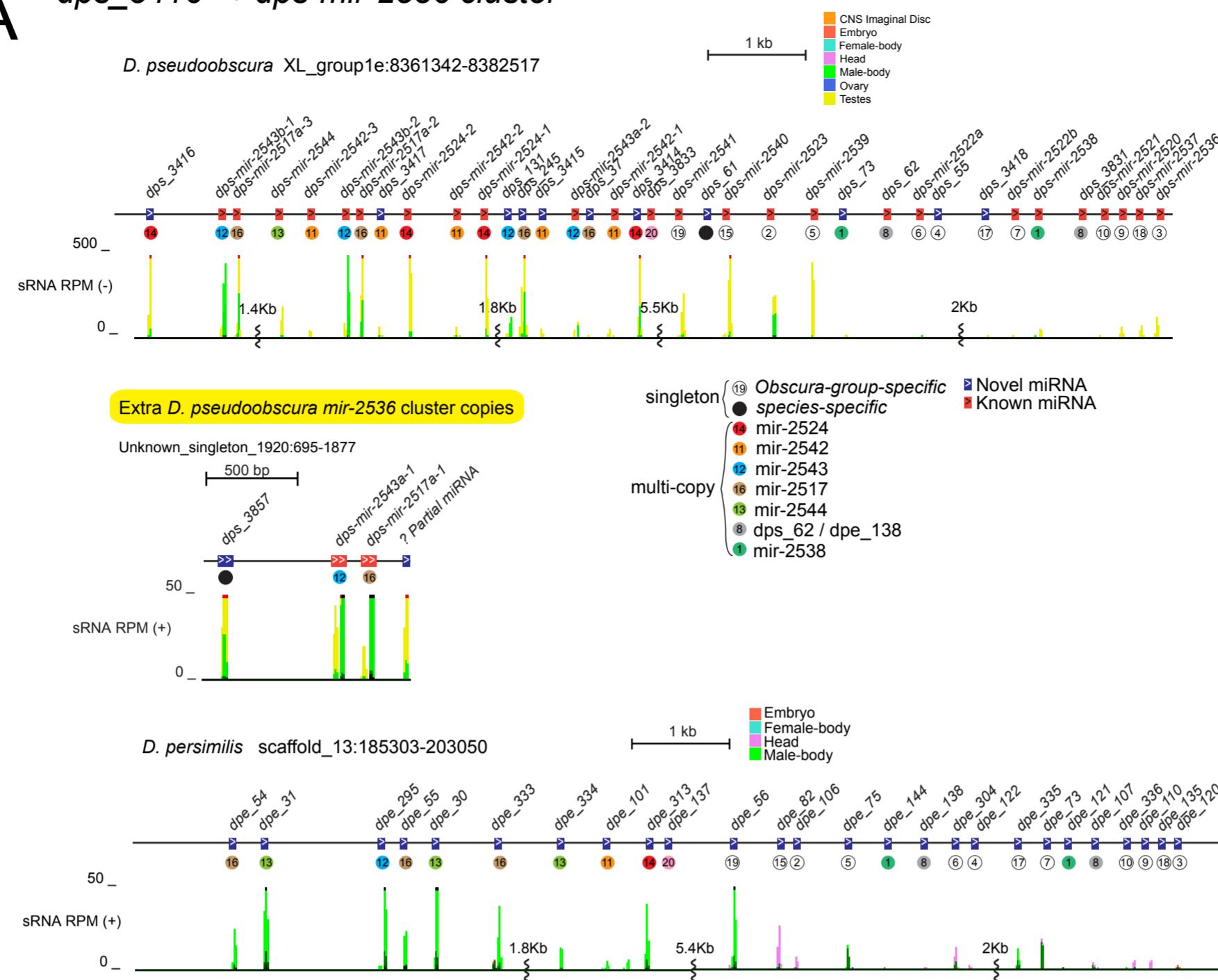
Mohammed et al
 Supplementary Figure S13



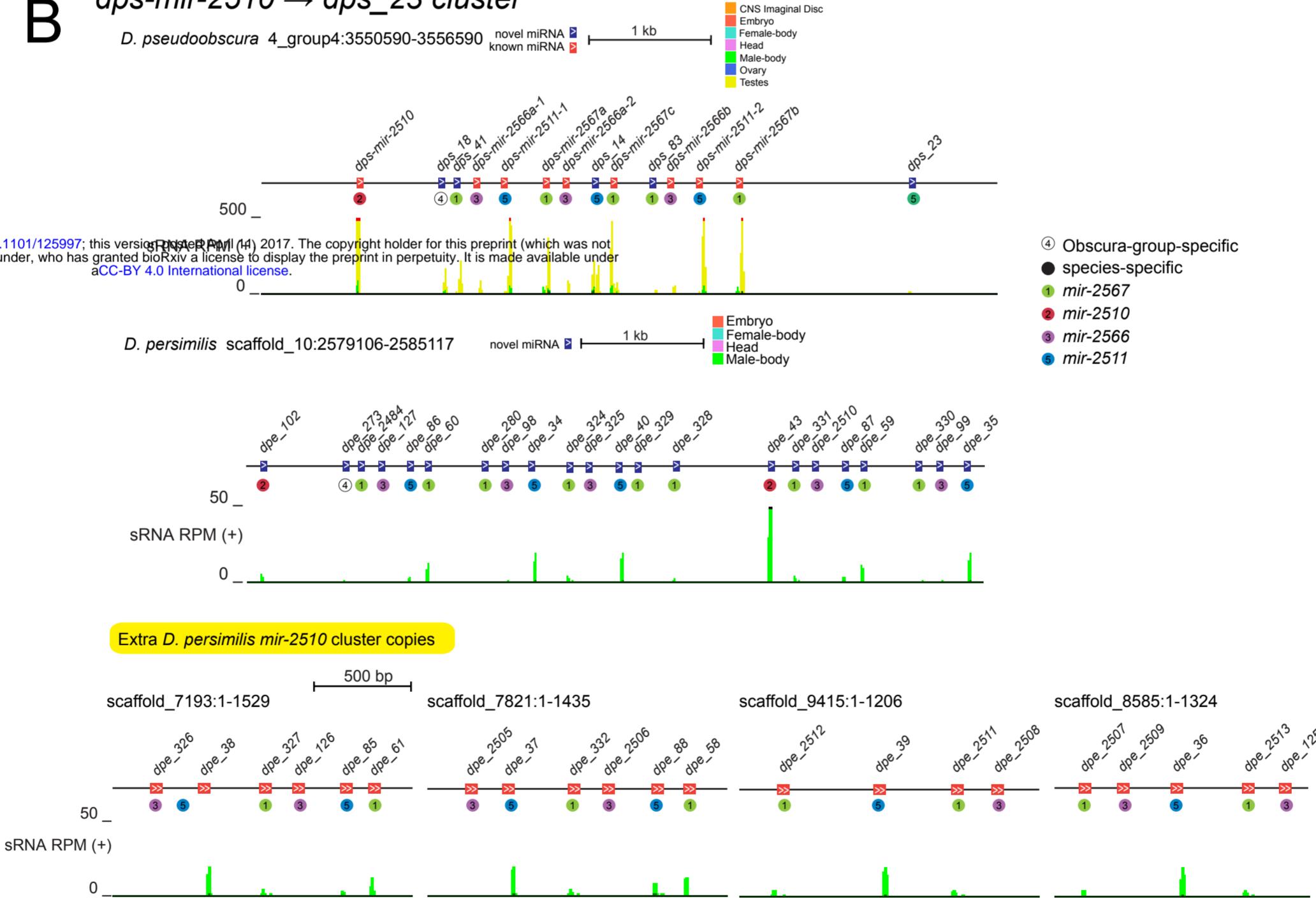
Supplementary Figure S14: Annotation of a novel Testes-restricted, recently-evolved, clustered (TRC) miRNA cluster identified in *D. willistoni*. Tissue code indicates the miRNAs are all highest expressed in male body libraries.

Mohammed et al
Supplementary Figure S14

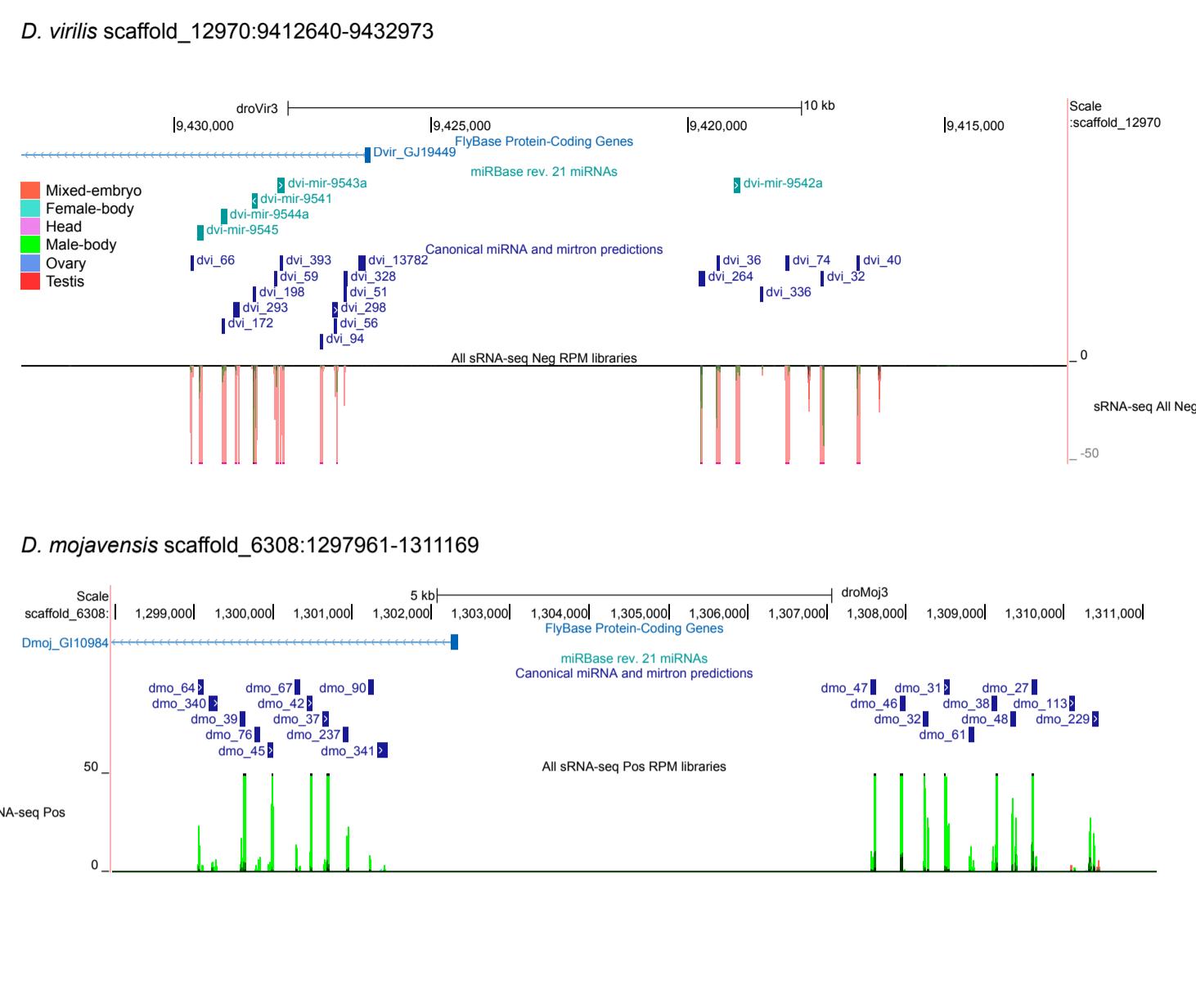
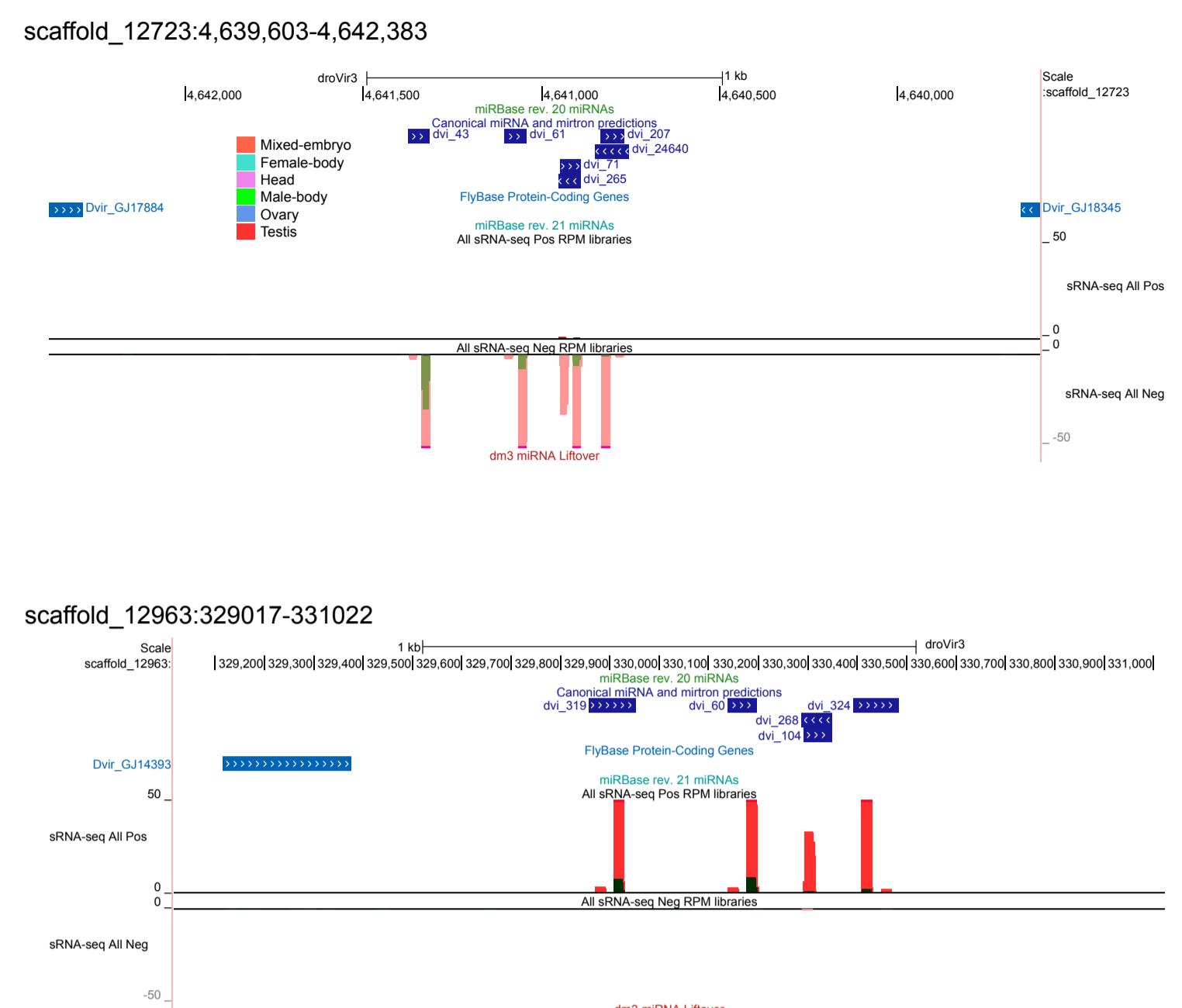
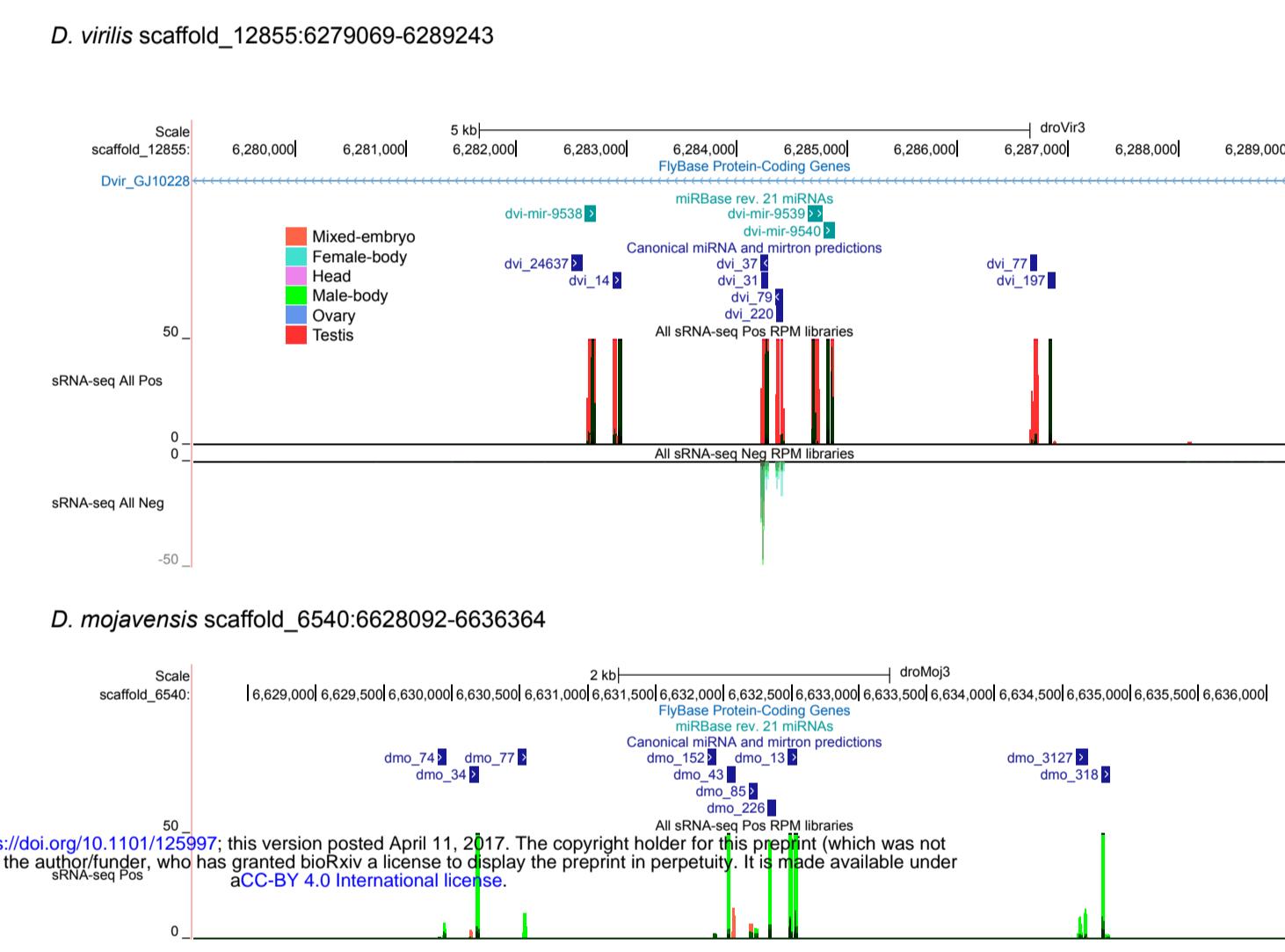
A *dps_3416 → dps-mir-2536 cluster*



B *dps-mir-2510 → dps_23 cluster*



Supplementary Figure S15: Annotation of novel Testes-restricted, recently-evolved, clustered (TRC) miRNA clusters identified in obscura group species, *D. pseudoobscura* and *D. persimilis*. (A) Genomic organization and small RNA read density of orthologous *dps_3416 → dps-mir-2536* TRC clusters in *D. pseudoobscura* and *D. persimilis*. (B) Genomic organization and small RNA read density of orthologous *dps-mir-2510 → dps_23* TRC clusters in *D. pseudoobscura* and *D. persimilis*. Tissue code indicates the miRNAs are all highest expressed in male body/testis libraries. Note that there are also additional copies of some of these TRC loci located outside of these clusters.

A*dvi_66 → dvi_40 cluster
mir-9545 cluster***C***dvi_43 → dvi_207 cluster
mir-9704 cluster***B***dvi_24637 → dvi_197 cluster
mir-9538 cluster*

Supplementary Figure S16: Annotation of novel Testes-restricted, recently-evolved, clustered (TRC) miRNA clusters identified in virilis clade species.

- A. The *dvi_66 → dvi_40* cluster. This cluster has a clear homologs in *mojavensis*. The small RNA mappings to their respective genomic loci are shown.
 B. The *dvi_24637 → dvi_197* cluster has clear homologs in *D. mojavensis*. The small RNA mappings to their respective genomic loci are shown.

C. Annotation of three additional novel Testes-restricted, recently-evolved, clustered (TRC) miRNA clusters identified in virilis clade species.

The *dvi_43 → dvi_207* cluster has two copies in *D. virilis* (i.e. roughly similar members can be found on scaffold_12723 and scaffold_12963). The other cluster is *D. mojavensis* *dmo_62* and *dmo_330*.

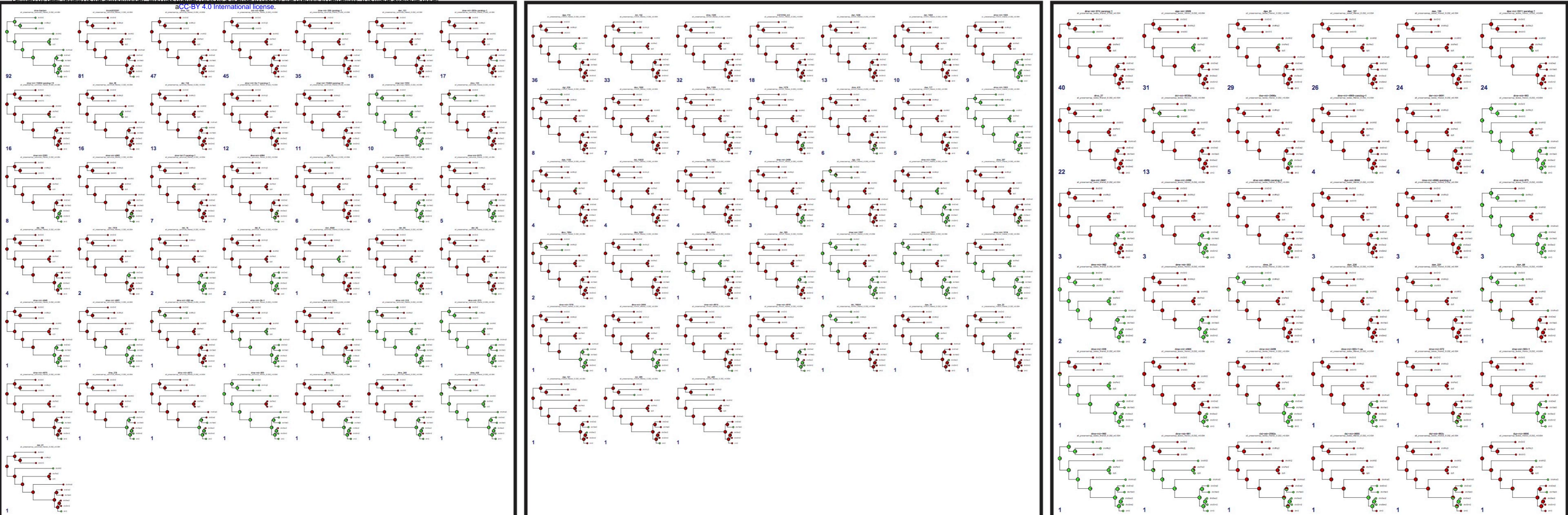
Mohammed et al
Supplementary Figure 16

D. pseudoobscura

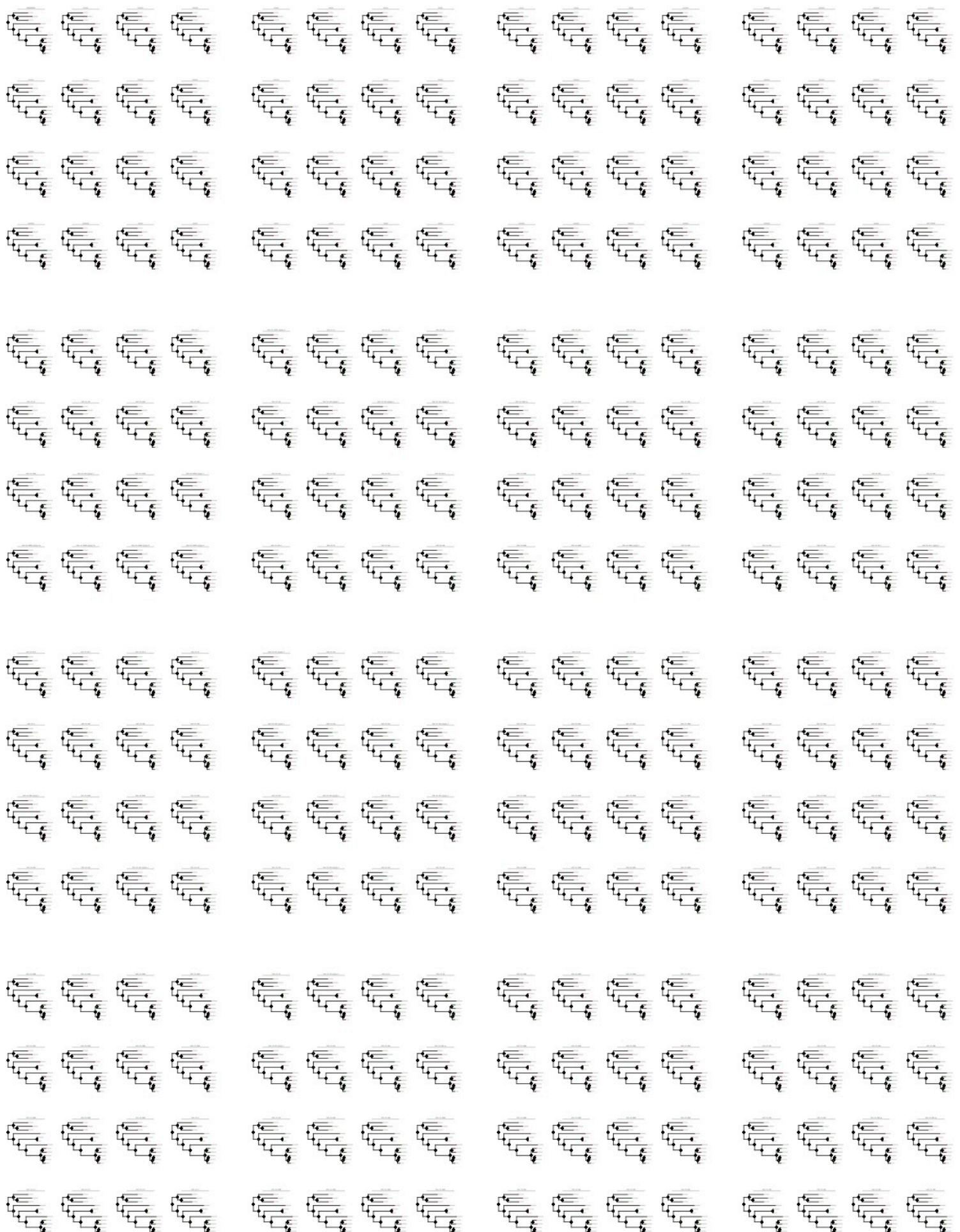


Supplementary Figure S17: Expression difference between *D. pseudoobscura*-specific or obscura-group-specific TRC and solo canonical miRNAs. Points reflect the maximum expression per locus assessed over all *D. pseudoobscura* libraries. P-value computed from two-tailed Wilcoxon Rank Sum Test.

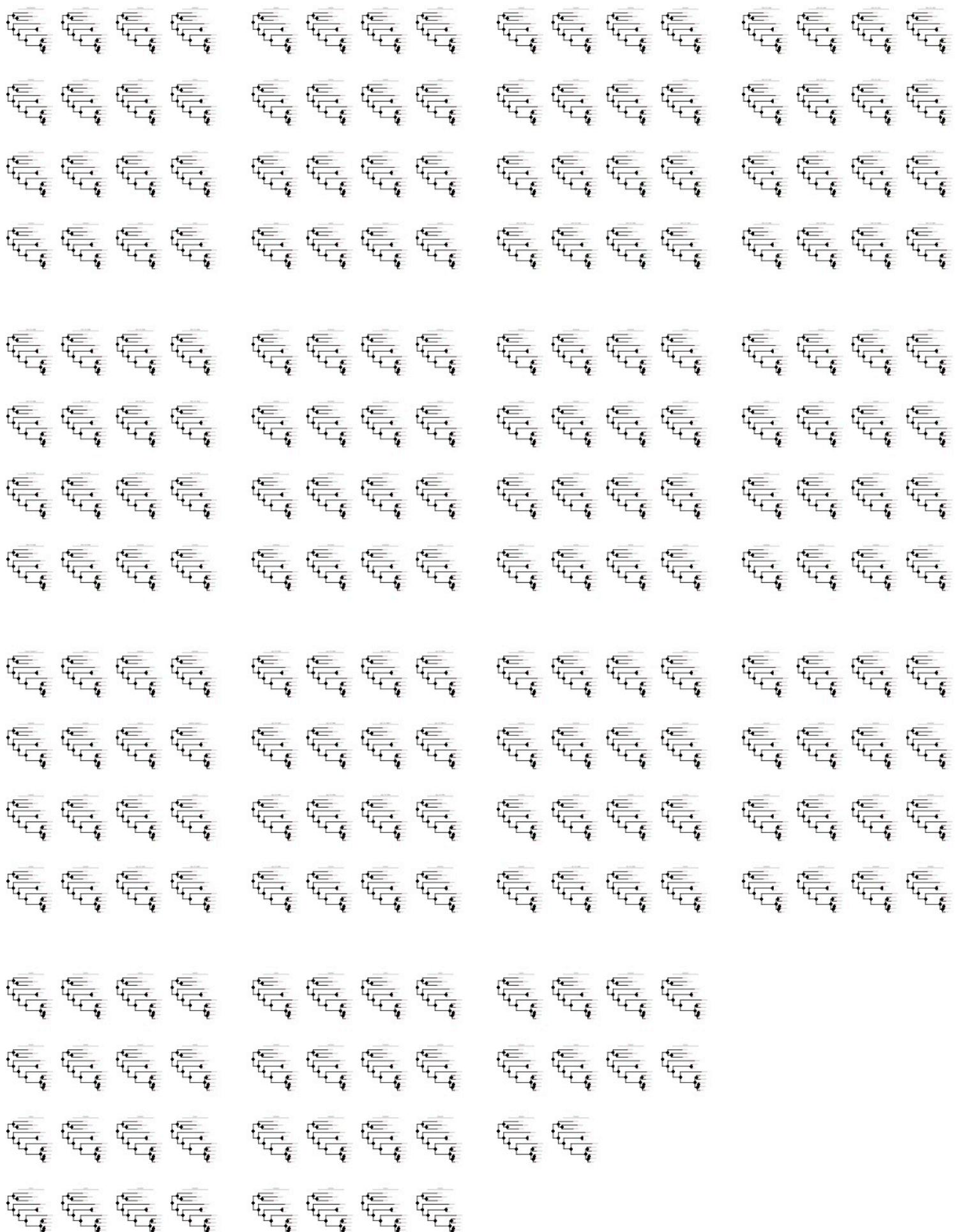
Mohammed et al
Supplementary Figure S17

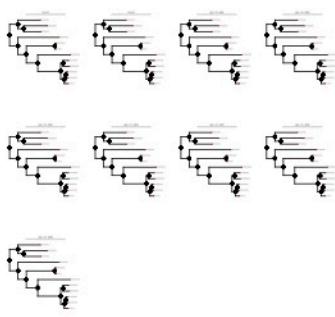


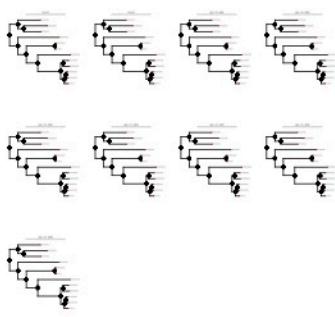
Supplementary Figure S18: All possible phylogenetic reconstruction of ancestral miRNA presence and absence for 3 miRNA classes using a phylogenetic probabilistic graphical model with universal parameters of $\lambda = 0.292$ and $\mu = 0.694$. These parameters were computed by running the phylogenetic reconstruction algorithms on all mirtrons and miRNAs pooled together. These trees illustrate how the method's maximum likelihood reconstruction performs for all possible configurations of extant miRNAs presence and absence per alignment. Blue text indicates count of alignments with this particular configuration in each class. Summary of miRNA birth and death (Figure 6) are based upon these estimates of ancestral miRNA presence and absence.











dps 22 Complete Alignment

Supplementary Figure S22: Examples of mirtrons with atypical emergence and decay patterns. Expression profiles and mirtron alignments are shown per example to highlight the non-clade specificity of mirtron presence.

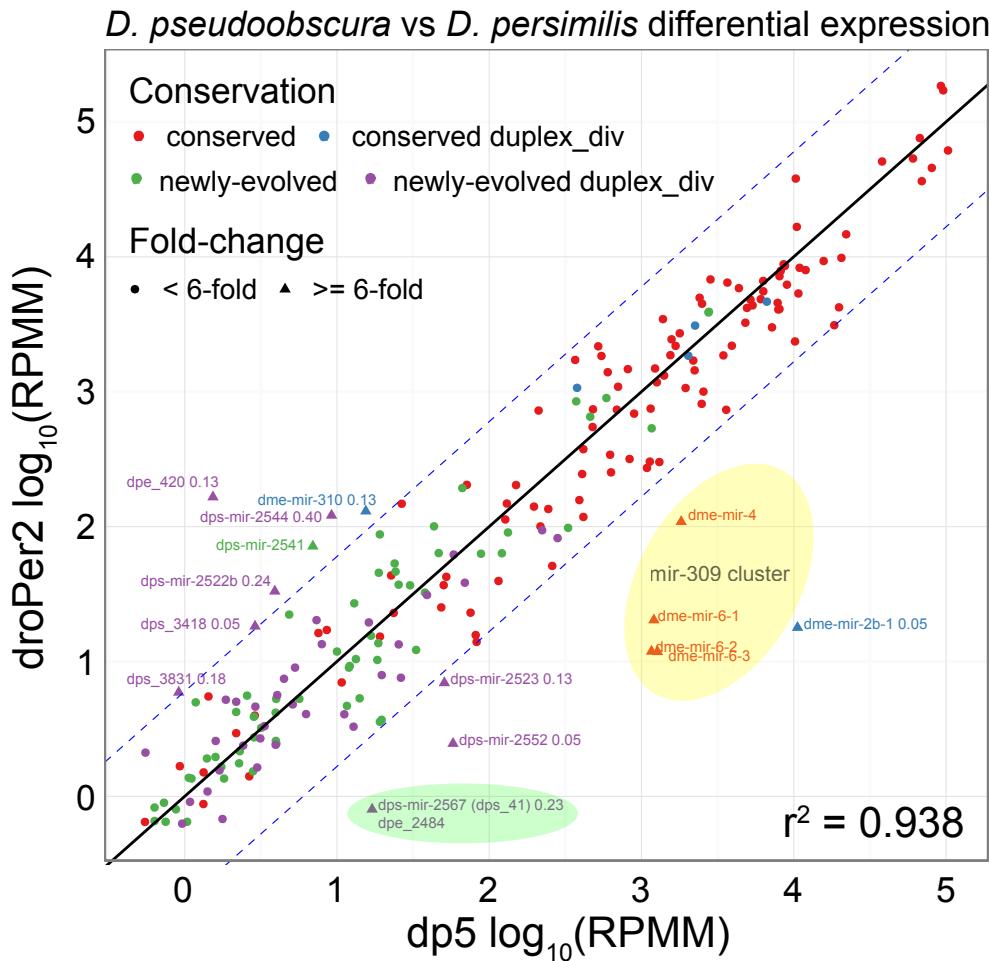
dvi 400 Complete Alignment

Generated: 08/02/2016 at 02:09 AM

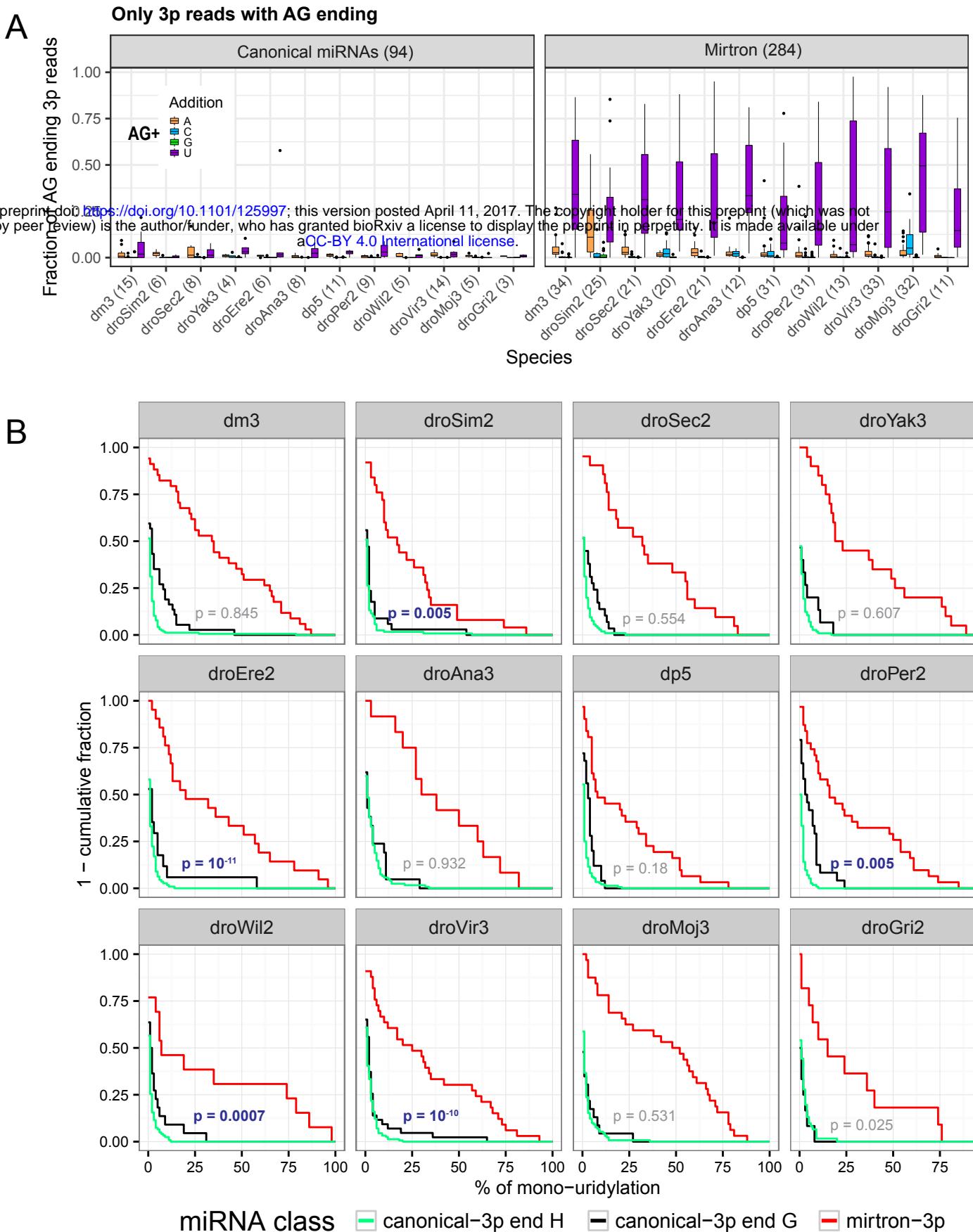
dse 121 Complete Alignment

Generated: 08/02/2016 at 02:06 AM

dps_15 Complete Alignment



Supplementary Figure S23. Differential expression of miRNAs between obscura group species. Scatterplot depicting the correlation of miRNA expression of all *D. pseudoobscura* and *D. persimilis* ortholog pairs (RPMM = Reads Per Million Mapped MiRNA Reads). All miRNA alignments with orthologs in both species are shown. Points that lie on or near the diagonal represent similarly expressed ortholog pairs. Orthologs with > 6 -fold $\log_{10}(\text{RPMM})$ difference (denoted by the blue-dashed line and labeled points) are examples of significantly differentially expressed orthologs. Points are colored by miRNA age, and shapes represent miRNAs with or without miR:miR* duplex region substitutions (fraction of duplex sites with substitutions are labeled). Note that the mir-309 cluster (yellow) loci are expected to be expressed in the very early embryo, and given that the embryo development and timing were not controlled in library preparation, their differential accumulation may not be genuine. Amongst loci changed by >6 -fold, dme-mir-2b-1 and dme-mir-310 are deeply conserved, but all others are specific to the obscura group species.



Supplementary Figure S24: 3' end untemplated nucleotide additions for canonical miRNAs and mirtrons in 12 Drosophila species. (A) Proportion of AG ending 3' arm miRNAs and mirtrons that contain reads within mono-A, C, G or U untemplated additions. Error bars represent the standard error of the mean. More mirtrons contain untemplated uridylation than comparable 3' end AG-ending canonical miRNAs. (B) Species-specific empirical cumulative distribution function of mono-uridylation for mirtrons and canonical miRNAs with 3' end 'G' nucleotide or non-'G' nucleotides (i.e. IUPAC ambiguity code 'H'). P-value computed from two-tailed Wilcoxon Rank Sum Test between canonical 3'-end 'H' miRNAs and mirtrons. Significant differences in mono-uridylation distributions between these two classes are noted in blue text. P-values from comparisons between canonical 3' end 'H' miRNAs and 3' end 'G' miRNAs are all non-significant and not shown.