Drosophila primary microRNA-8 encodes a microRNA encoded peptide (miPEP) acting in parallel of *miR-8*

Audrey Montigny[1], Patrizia Tavormina[1], Carine Duboe[1], Hélène San Clémente[1], Marielle Aguilar[1], Philippe Valenti[2], Dominique Lauressergues[1], Jean-Philippe Combier[1] & Serge Plaza[1*].

Affiliations

[1]Université de Toulouse, UPS, CNRS, UMR5546, Laboratoire de Recherche en Sciences Végétales, 31320 Auzeville-Tolosan, France.

[2]Centre de Biologie du Développement (CBD), Centre de Biologie Intégrative (CBI), Université de Toulouse, CNRS, Bat 4R3, 118 route de Narbonne, F-31062, Toulouse, France

* **Correspondence:** S.Plaza, e-mail: serge.plaza@lrsv.ups-tlse.fr, Tel: +33 (0)5 34 32 38 27

Key words: *Drosophila*, sORF, lncRNA, *miR-8*, miPEP, small peptides

Summary

Background:

In the last decades, genome-wide studies of many species have revealed the existence of a myriad of RNAs differing in size, coding potential and function. Among these are the long non-coding RNAs (lncRNAs), some of them producing functional small peptides via the translation of short ORFs (sORFs). It now appears that any kind of RNA presumably has a potential to encode small peptides. Accordingly, our team recently discovered that plant primary transcripts of microRNAs (*pri-miRNAs*) produce small regulatory peptides (miPEPs) involved in auto-regulatory feedback loops enhancing their cognate microRNA expression which in turn controls plant development. Here we investigate whether this regulatory feedback loop is conserved in *Drosophila melanogaster*.

Results:

We performed a survey of ribosome profiling data and revealed that many *pri-miRNAs* exhibit ribosome translation marks. Focusing on *miR-8*, we showed that *pri-miR-8* can produce a miPEP-8. Functional assays performed in *Drosophila* revealed that miPEP-8 affects development when over-expressed or knocked down. Combining genetic and molecular approaches as well as genome-wide transcriptomic analyses, we showed that *miR-8* expression is independent of miPEP-8 activity and that miPEP-8 acts in parallel of *miR-8* to regulate the expression of hundreds of genes.

Conclusion:

Taken together, these results reveal that several *Drosophila pri-miRNAs* exhibit translation potential. Contrasting with the mechanism described in plants, these data shed light on the function of yet un-described microRNA encoded peptides in *Drosophila* and their regulatory potential on genome expression.

Background

More than twenty years after the first genome annotation, it is now becoming clear that the protein-centric view of gene content strongly underestimates the number of DNA regions that are expressed and fulfil important functions for development and physiology since the majority of the genome is in fact transcribed [1]. A first discovery was the importance of hundreds of small non-coding RNAs, such as microRNAs (*miR*) playing regulatory roles in the silencing of genes and transposable elements. More recently, genome-wide transcript profiling has disclosed the existence of numerous RNAs referred to as long non-coding RNAs (lncRNAs or lincRNAs) since they lack the classical hallmarks of protein-coding genes.

Although the functions of all lncRNAs remain largely unknown, there are several experimental cases illustrating their key role as functional RNAs in various steps of the control of genome expression [2]. In association with other molecules, lncRNAs can coordinate several physiological processes and their dysfunction impacts several pathologies, including cancer and infectious diseases. lncRNAs control genetic information, such as chromosome structure modulation, transcription, splicing, messenger RNA (mRNA) stability, mRNA availability and post-translational modifications. They also act as scaffolds, bearing interaction domains for DNA, mRNAs, miRs and proteins, depending on both their primary sequence and secondary structure [3]. In addition, while lncRNAs annotated as non-coding cannot produce large-sized proteins, they all contain myriads of short open reading frames (sORF) [4-7] and a surprising result was the discovery for a subset of them of their translation into small functional peptides [8-11].

MicroRNAs define a class of small, non-coding RNAs able to down-regulate the expression of target mRNA by binding to the 3'-ends inducing mRNA degradation and/or translation repression. Intergenic microRNAs are produced from the sequential cleavage of long precursors named primary transcripts of microRNA (*pri-miRs*) (frequently annotated as lncRNAs) by Drosha and Dicer into 22nt miRNA duplexes associated with the RISC protein complex. Identified in a broad spectrum of living species, they are transcribed from coding genes or lncRNAs by the RNA polymerase II. MicroRNAs are critical for normal animal development and are involved in many biological processes [12]. Due to their role in

86 silencing, *miRs*, and in particular the *pri-miR*s they come from, have always been considered
87 as non-coding. This dogma was recently overturned by the discovery that plant *pri-miR*s
88 encode small regulatory peptides, called miPEPs for microRNA-encoded peptides [13]. In
89 plants, miPEPs specifically increase transcription of their primary transcript impacting the
90 level of the mature miR produced and consequently affecting the control of the entire *miR*
91 Gene Regulatory Network (GRN). To date, this regulation has been extended to several miRs
92 in various plants [14-17]. In human cells, only few reports present evidences of pri-miR
93 translation [18-21]. Pri-miR-22 host gene endogenously produces a miPEP for which the
94 function is unknown [19]. miR-200 might produce a miPEP able to control the Vimentin, a
95 miR200 target [18]. miPEP155 was described to control major histocompatibility complex
96 class II-mediated antigen presentation by disrupting the HSC70-HSP90 machinery [20].
97 However, whether these miPEPs control the expression of their cognate miR was not
98 investigated. More recently, miPEP133 was discovered in miR34a as a tumor suppressor
99 localized in the mitochondria. It enhanced p53 transcriptional activation which, in turn,
100 induces miR-34a expression [21]. In summary, whereas few human pri-miRNA appear
101 translatable, it remains not clear whether the regulation found in plants exists in animals.
102 Here we addressed this question in *Drosophila melanogaster*. We show that several
103 intergenic *pri-miRs* contain marks of ribosome profiling. To investigate whether *Drosophila*
104 can produce miPEPs, we focused on *miR-8*, a previously well-characterized microRNA that
105 sustains many developmental traits [22]. *miR-8* controls organism physiology, tissue growth
106 and survival [23-26], stem cell renewal [27-30], central nervous system development [30-
107 32], signalling and developmental pathways [33-41]. Consequently, *miR-8* loss or gain of
108 function impinges on fly development and survival. In the *pri-miR-8*, we located a small ORF
109 encoding a potential 71 amino acid peptide we called miPEP-8. We showed that this sORF is
110 translatable *in vitro* and *in vivo*. While our attempts to reveal an auto-regulatory loop
111 remained unsuccessful, we showed by genetic and transcriptomic approaches that miPEP-8
112 and *miR-8* act in parallel in controlling wing development and in regulating the expression of
113 distinct sets of genes.

114 # Results

115

116 **Translation potential of sORFs present in *Drosophila pri-microRNAs***

117 As plants *pri-miRs* contain sORFs (miORFs) producing functional miPEPs involved in a
118 positive feedback loop on *pri-miR* expression (Fig. 1A), we first asked whether *D.*
119 *melanogaster pri-miRs* contain significant levels of miORFs. We scored the number of sORFs
120 present in intergenic *miR* genes and their *pri-miR* and compared this number with other
121 classes of RNAs, coding as well as non-coding (Fig. 1B). *Pri-miRs* present the highest
122 enrichment of sORFs/kb when compared with the 5'UTR of coding genes, sequences known
123 to contain translatable short open reading frames, but contain similar amounts of sORFs
124 when compared with lncRNAs, previously reported to be widely bound by ribosomes [42].
125 Ribosome profiling experiments were developed to study translatability of RNAs by scoring
126 the sequences bound by ribosomes [7, 43, 44] including studies conducted on *Drosophila* [4,
127 45, 46]. We searched if and how many *Drosophila pri-miRs* were widely bound by ribosomes.
128 Briefly, we searched in the rib-seq databases marks of ribosome binding in predicted *pri-*
129 *miR*. To avoid difficulties for the interpretation with *miRs* embedded in host coding genes,
130 we focused only on the intergenic *miRs*. We found many marks of ribosome profiling and
131 identified hundreds of potentially translated sORF peptides within dozens of *pri-miRs*,
132 suggesting that, as observed in plants, *Drosophila pri-miRs* are potentially translated (Fig. S1;
133 Additional file 1).

134

135 ***Drosophila pri-miR8* encodes a miPEP-8 translated *in vivo***

136 To further characterize the potential translation of *pri-miRs*, we focused on the *miR-8*
137 primary transcript (*pri-miR-8*). *MiR-8* is an intergenic miR and is likely produced from the
138 expression of long non-coding CR43650 spanning over the *pre-miR-8* (Fig. 1C). In flybase, two
139 CR43650 ncRNAs were predicted, a long and a short form defining putative *pri-miR-8*
140 transcripts, independently identified by two different teams and likely initiated from
141 different promoters [47, 48]. As shown in figure 1C, many marks of potential translation
142 were found along the CR43650 transcripts. We first performed 5'RACE as well as RT-PCR
143 assays to determine which isoform was preferentially produced. While we successfully
144 detected the short isoform in flies and S2 cells, we did not succeed in amplifying the long

145 isoform, suggesting that the short transcript is the most abundant isoform expressed, a
146 result confirmed by RNA-seq data generated during this study (Fig. S2). Our RNA-seq data
147 further suggest that the long transcript defines a different transcription unit since no
148 overlapping reads were detected in the promoter region of the short transcript (Fig. S3A). In
149 agreement with this, a GAL4 enhancer trap recapitulating *miR-8* expression is inserted just
150 upstream of the short transcrip*t* [24]. Of note, we noticed that these flies are homozygous
151 viable, showing that the insertion is not deleterious. In addition, these flies express *miR-8* at
152 a level comparable to control flies (Fig S3B), showing that this insertion disrupting the co-
153 linearity of the *miR-8* locus is not detrimental for *miR-8* expression. Finally, we verified that
154 this short transcript (referred hereafter as *pri-miR-8*) is functional since it efficiently
155 produces a functional mature *miR-8* able to down regulate a *miR-8* sensor (see below Fig.
156 3B, D).

157 We therefore looked for a potential open reading frame within the *pri-miR-8* gene.
158 Focusing on the 5' leader sequence of *pri-miR-8*, we found one ORF located upstream of the
159 *pre-miR-8.* This ORF is the longest ORF present 5' to the *pre-miR* which potentially encodes a
160 miPEP of 71 amino acids in length if initiated from the first ATG (ATG1) (Fig. 1C and S2A).
161 However, the presence of a second ATG (ATG2), located downstream, gives the possibility to
162 produce a shorter peptide. To determine whether the open reading frame is translated and
163 which initiation codon is used, we generated and characterized specific antibodies (Fig. S4).
164 In parallel, we generated different deletion constructs and performed *in vitro* translation
165 experiments using insect cell extracts. As shown in Figure 1D (left panel), we observed an
166 efficient translation from the longest construct (CTG) consisting in an extended genomic
167 region of the defined 5' leader sequence of *pri-miR-8.* Deletion experiments revealed a
168 stronger and efficient translation from ATG1 but not from ATG2. A higher product, possibly
169 initiated at an upstream codon present in the construct was also detected but was not
170 further investigated since it was not present in *pri-miR-8*.

171 We next generated translatable and untranslatable miPEP-8 forms placed in optimal
172 translational Kozak (K) or mutated kozak (KMT) contexts or in the miPEP-8 natural ATG1-
173 initiated translational context. We then expressed these miPEP-8 constructs in *Drosophila* S2
174 cells (Fig. 1D, middle panel). As revealed by western blot experiments, these different
175 constructs produced the same level of miPEP-8 when the ATG was placed in an optimal or in
176 its natural translational context whereas no miPEP-8 was detected from mutated ATG

177 constructs. This result reveals that the natural nucleotide context of miPEP-8 miORF is in a
178 favorable translational context. We next questioned whether *pri-miR-8* was able to produce
179 miPEP-8. We transfected S2 cells with wild type or ATG-mutated *pri-miR-8* constructs and
180 performed western blot experiments. We observed that only the wild-type *pri-miR-8* was
181 able to produce miPEP-8. Finally, we examined whether a peptide corresponding to miPEP-8
182 was detectable in fly extracts by performing a western blot experiment on young adult flies.
183 As revealed in Figure 1D (right panel), we observed a signal co-migrating with the *in vitro*
184 synthetized miPEP-8, corresponding to endogenous miPEP-8 as demonstrated by the lack of
185 this band in *miR-8* deleted Δ2 mutant flies. Sequence alignment analyses revealed that this
186 peptide is poorly conserved amongst *Drosophila* species. Some homologies are detected in
187 *Drosophila* melanogaster group but no conservation was found in more distant *Drosophilae*
188 species.
189 Altogether, our results reveal that the *pri-miR-8* transcript carries, in addition to the
190 *miR-8* sequence, at least one translated ORF located upstream of the *pre-miR,* able to
191 express a miPEP-8 both *in vitro* and *in vivo*.
192
193 **Expression of miPEP-8 impinges on *Drosophila* development.**
194 To study the function of miPEP-8, we generated flies able to express a translatable
195 and untranslatable version of miPEP-8 (ATG mutated). *miR-8* sustains many biological
196 functions in *Drosophila* and either its loss of function or its over-expression leads to
197 detrimental outcomes in cells, tissues or the whole organism [23-27, 29, 31-37, 40, 41, 49],
198 providing a useful readout to assess miPEP-8 activity. To test our hypothesis that miPEP-8
199 controls *miR-8* expression and modulates its activity, we used an over-expression assay.
200 Using the *miR-8* GAL4 driver, a GAL4 insertion in the endogenous *miR-8* promoter reported
201 to mimic *miR-8* expression [23, 24, 33, 40], we first asked whether over-expression of miPEP-
202 8 impinges on fly viability. As reported, driving UAS-*miR-8* over-expression results in
203 increased fly lethality [23] (Fig. 2A). Over-expression of a UAS-miPEP-8 translatable construct
204 also affected fly viability whereas the untranslatable form did not. This indicates that, like
205 *miR-8*, the translatable form of miPEP-8 is able to interfere with development, although with
206 a weaker effect compared to *miR-8*.
207 By loss or gain of function experiments, *miR-8* was shown to induce a « small wing »
208 phenotype [24, 26, 40]. We therefore questioned whether over-expression of miPEP-8 also

209 induced a wing phenotype (Fig. 2B). Using the wing driver line MS1096, *miR-8* over-
210 expression induced several wing defects, from a reduced size to the loss of wing vein,
211 sensory organs, miss shaped, depending on the transgene/promoter strength [40, 41].
212 Quantifying the wing size appeared the most reliable criteria, and we compared this
213 phenotype with the phenotype observed with miPEP-8 over-expression. Consistently with
214 our above result, miPEP-8 over-expression induced a slight, albeit significant wing size
215 reduction, revealing yet again a weaker activity compared with *miR-8*. Importantly, miPEP-8-
216 induced wing reduction was dependent on the integrity of the translation codon since the
217 same construct with the mutated ATG did not induce any phenotype.
218 Altogether, these experiments show that miPEP-8 appears to be biologically active
219 but induces a milder phenotype compared to *miR-8*.
220
221 **In *Drosophila*, *miR-8* expression is independent of miPEP-8 expression.**
222 In plants, miPEPs positively auto-regulate the expression of their own *miR* by
223 regulating the expression of their cognate *pri-miR*. To test whether miPEP-8 regulates *miR-8*
224 expression in *Drosophila*, we monitored the level of *pri-miR-8* and *miR-8* through
225 quantitative PCR experiments on S2 cells and on flies expressing the translatable form of
226 miPEP-8. We first set up experimental conditions in S2 cells by transfecting the *pri-miR-8*
227 construct and quantified the level of the exogenous *pri-miR* and mature miR produced.
228 When *pri-miR-8* was over-expressed, both over-expression of *pri-miR-8* and *miR-8* was
229 detectable (Fig. 3B). We then over-expressed miPEP-8 and quantified the endogenous level
230 of *pri-miR-8* and mature *miR-8*. Whereas we unambiguously detected the over-expression of
231 the miPEP-8 construct, we did not see any change in the levels of endogenous *pri-miR-8* or
232 mature *miR-8* expression (Fig. 3C). As observed in S2 cells, no change in *pri-miR-8* or miR-8
233 levels was observed in flies upon miPEP-8 expression using the *miR-8* GAL4 driver (Fig. S5).
234 We next questioned the potential regulatory role of miPEP-8 on *miR-8* expression by
235 testing miPEP-8 overexpression on endogenous miR-8 activity level in the presence or
236 absence of. One way of challenging this question is the use of a sensor of miR-8 activity,
237 whether endogenous or resulting from over-expression. Thus, we designed a miR-8
238 luciferase reporter, bearing a 3'UTR from the escargot gene (esg), previously shown to be
239 directly regulated by *miR-8* [27]. Over-expression of *pri-miR-8* in S2 cells was able to repress
240 the miR-8 sensor to the same extent as *miR-8* (Figure 3D, left panel), hence validating our

241 miR-8 sensor. As mentioned above, this also indicates that the *pri-miR-8* construct is able to
242 generate a functional and mature *miR-8*. Clearly, however, over-expression of miPEP-8 did
243 not reveal any modulation of the luciferase reporter (Fig. 3D right panel). We performed
244 similar experiments *in vivo* using a miR-8 GFP sensor in wing imaginal discs where *miR-8* was
245 previously shown to be functional [33]. Expressing *miR-8* under the patched (ptc) GAL4
246 promoter led to the repression of the GFP in the ptc domain (Fig. 3E ptc>miR8 panel).
247 Consistently, miPEP-8 over-expression had no effect on the miR-8 GFP sensor in vivo (Fig. 3E
248 ptc>miPEP-8 panel). Altogether, these results indicate that miPEP-8 is not able to control
249 *miR-8* expression, or activity for that matter (see more below). We obtained similar
250 conclusions on endogenous miR-8 target in S2 cells and in wing discs (Fig. S6).

251 Finally, we asked whether the miPEP-regulation of *pri-miR* observed in plants is
252 system specific by testing whether a plant *pri-miR*, up regulated by its miPEP in plant cells,
253 could be up-regulated in *Drosophila* cells. Reciprocally, we tested whether miPEP-8 is able to
254 up-regulate its *pri-miR-8* in plant cells. To that end, we expressed the plant *Arabidopsis*
255 *thaliana pri-miR165a* and miPEP-165a in S2 cells using the actin promoter and measured the
256 level of *pri-miR* produced in the absence or presence of the miPEP. Reciprocally, the
257 *Drosophila pri-miR-8* and its miPEP-8 were cloned in plant expression vectors and
258 agroinfiltrated in *Nicotiana benthamiana* leaves as performed previously [13]. Whereas the
259 up-regulation of the *A. thaliana pri-miR-165a* by miPEP-165a was observed in *N.*
260 *benthamiana*, we did not detect any up-regulation, but rather a down-regulation of *pri-miR-*
261 *165a* in *Drosophila* S2 cells (Fig. S7). Reciprocally, we could detect a slight but significant
262 increase of *pri-miR-8* expression upon miPEP-8 over-expression in *N. bentamiana* leaves,
263 suggesting that a difference of regulation occurs between plant cells and insect cells (Fig.
264 S7). Although, *miR-8* expression appears to be miPEP-8 independent in *Drosophila*, these
265 results further suggest that, like for plants miPEPs, animal miPEPs might nonetheless have
266 the potential of autoregulating the expression of their cognate pri-miR.
267

268 **Endogenous miPEP-8 alteration reveals *in vivo* activity.**

269 To investigate the functional requirement of miPEP-8 *in Drosophila*, we tried several
270 times to edit the miPEP-8 in flies using CRISPR/Cas9, but unsuccessfully. In contrast, it was
271 possible from the first attempt to delete the entire *miR-8* locus, showing that the failure to
272 obtain a specific miPEP-8 edited line is not due to trivial technical problems. We therefore

273 created a specific P landing platform in place of *pri-miR-8* transcript to perform Knock In
274 strategies (Fig. 4A). This edited line exhibits the previously *miR-8* reported phenotypes [26],
275 including a strong developmental lethality with only few escaping flies exhibiting a reduced
276 size (including wings (Fig. 4C and Fig. S11)) and leg defects (not shown). We further knocked
277 in the wild type *pri-miR-8* and the *pri-miR-8* miPEP-8 untranslatable form (mt) at the P
278 landing site and analyzed the outcomes. For both constructs, we observed a nearly total
279 rescue since the theoretical expected 33,3 % homozygotes (and 66,6 % of CyO flies) in the
280 progeny was almost reached (Fig. 4B left panel). These rescued flies appeared phenotypically
281 normal and re-expressed *miR-8* at levels close to *miR-8* endogenous expression (Fig. 4B right
282 panel). Both pri-miR-8 constructs restored significantly wing sizes when compared to the
283 *ΔmiR-8* CRISPR line (Fig. 4C). Interestingly, a significant difference was observed in wings
284 between the *pri-miR-8* wild type construct and the *pri-miR-8 mt* in which the miPEP-8
285 translatability was disrupted (Fig. 4C).
286 As a second approach, we took advantage of a polymorphism mutation detected in
287 *Drosophila* Gene Reference Panel (DGRP) lines generating a premature stop codon leading to
288 a 24 amino acid C-terminal miPEP-8 truncation called miPEP-8alt [50]. We outcrossed the
289 miPEP-8alt DGRP line into white background and analyzed the consequence of the miPEP-
290 8alt mutation. Whereas no significant difference was observed for miR-8 level between
291 these the two miPEPs variants (Fig. 4D), flies homozygous for miPEP-8alt exhibit a significant
292 wing size reduction when compared with the white flies expressing the miPEP-8 (Fig. 4E). We
293 further analyzed the resulting wing phenotypes in different genetic contexts. The phenotype
294 is also present when miPEP-8alt mutation was tested over a deletion of the *miR-8* gene (the
295 Δ2 and the Δ*miR-8* CRISPR line generated in this study) (Fig. 4E), suggesting that the
296 observed phenotype is a consequence of miPEP-8 loss of function. To test this, we
297 performed rescue experiments by expressing miPEP-8 using the *miR-8* GAL4 driver line in
298 miPEP-8alt/Δ*miR-8* background. Flies expressing miPEP-8 in miPEP-8alt/ Δ*miR-8* CRISPR
299 restored the wing size phenotype contrasting with the sibling control flies carrying no miPEP-
300 8 transgene (absence of expression of wild type miPEP-8).
301 Therefore, altogether, these experiments revealed an *in vivo* miPEP-8 function.
302
303 **The function of miPEP-8 is uncoupled from *miR-8* expression and activity**.

304 The above experiments suggest that in *Drosophila* miPEP-8 is not involved in a
305 positive auto-regulatory feedback loop as observed in plants. However, due to the
306 similarities of the phenotypes observed between *miR-8* and miPEP-8, we questioned
307 whether miPEP-8 could be involved in the *miR-8* pathway through another mechanism, or
308 whether it acts in parallel of miR-8. As both *miR-8* and miPEP-8 affected wing formation, we
309 developed a genetic assay to test whether miPEP-8 acts through miR-8 using a previously
310 validated miR-8 sponge, which titrates *miR-8* hence rescuing *miR-8*-induced phenotypes [29,
311 33]. Using this rescue assay, we asked whether the miR-8 sponge could also compensate the
312 miPEP-8 induced phenotype. Co-expressing *miR-8* together with a miR sponge scramble (as a
313 control and to maintain the number of UAS transgenes identical) using the MS1096 GAL4
314 driver led to wing size reduction. This phenotype was efficiently rescued by co-expressing
315 *miR-8* with the effective miR-8 sponge (Fig. 5). In contrast, when miPEP-8 was co-expressed
316 with the *miR-8* sponge, no compensation of the miPEP-8-induced wing reduction was
317 observed. Therefore, this result strongly suggests that miPEP-8 acts in parallel of *miR-8*.
318 We thus reasoned that the effect of miPEP-8 on wing development could be linked to
319 the modulation of gene expression independent of *miR-8*. To identify these putative miPEP-
320 8-regulated genes and to compare them with the *miR-8*-regulated transcriptome, we
321 overexpressed *miR-8* or miPEP-8 in S2 cells and performed RNA-seq 48h after transfection
322 (Fig. 5). Clearly, the transcriptomes appeared different (Fig. 6A). The assays performed on
323 *miR-8* over-expressing cells successfully retrieved previously identified *miR-8* targets both at
324 the RNA (Fig. S8B and Additional file2) and protein level (Fig. S6), hence validating our
325 experimental conditions. GO term enrichment identified biological pathways fitting with
326 miR-8 activity such as “regulation of organism or cell growth and differentiation”, “wing
327 development”, “apoptosis”, “regulation of actin cytoskeleton” (Fig. S9B). As for miPEP-8
328 controlled genes, strikingly, the majority of them were miPEP-8 specific (76%) (Fig. 6B and C)
329 since only 24% appeared co-regulated (Fig. 6B and E). In both cases, we found activated and
330 repressed genes (Fig. 6A, C, D, E). Remarkably, miPEP-8-modulated genes were frequently
331 more strongly modulated than *miR-8*-modulated genes (Fig. 6C, D). Increasing the Fold
332 change (FC>1,5) led to a decrease of the number of genes but the respective proportions
333 and conclusions remained unchanged (Fig. S8A). Our analyses of GO term enrichment clearly
334 identified shared functions for *miR-8* and miPEP-8 (Fig. S9A and S10A), some of which being
335 related to wing morphogenesis (such as cell junction organization actin filament-based

336 processes, epithelial cell morphogenesis, cell differentiation) or developmental processes
337 (such as neurogenesis, cell migration, embryonic morphogenesis) (Additional file 3).
338 However, *miR-8* and miPEP-8 also exhibit specific biological functions such as snRNA
339 modification and leucine metabolic process for miR-8 or K48 linked ubiquitination and
340 chromatin-mediated maintenance of transcription for miPEP-8 (Fig. S9B and S10B).
341 Altogether, these experiments suggest that *miR-8* and miPEP-8 independently control
342 similar biological processes, while regulating functions specific to one or the other.
343

344 ## Discussion

345 In the present study, we investigated whether a small ORF present in *Drosophila pri-*
346 *miR-8* was capable of producing a miPEP-8 and we propose that animal miPEPs are able to
347 act in parallel of their corresponding miRs. Several studies performed in a broad range of
348 organisms have revealed the prevalence of translated small/short open reading frames
349 (smORFs/sORFs) [7, 43, 44, 51-56]. Although sORF peptides were initially identified as being
350 encoded by unusual long non-coding RNAs, to date, it turns out that virtually all classes of
351 RNAs can produce these peptides. Therefore, sORF-encoded peptides (SEPs) are emerging as
352 an unexplored reservoir of putative regulators. However, while a growing body of evidence
353 further supports the importance of sORFs and associated peptides in development,
354 physiology and diseases [8, 54, 57, 58], the number of SEPs that have been characterized so
355 far still remains limited. Therefore, the current challenge resides in deciphering the full
356 repertoire of their functions and molecular modes of action, an issue largely dependent on
357 experimental approaches.

358 We show with several experimental data that a miPEP-8 is indeed produced from
359 *Drosophila pri-miR-8*. First, we found a signal of ribosome binding in the *pri-miR* of several
360 microRNAs and in particular, *miR-8*. Second, we show that the initiation codon of the miORF
361 present within *pri-miR-8* is in a favorable translational context. Third, after having generated
362 specific antibodies, we detected a peptide co-migrating with *in vitro* translated miPEP-8 in fly
363 extracts. Fourth, forced expression or loss of function of this peptide led to a significant
364 developmental phenotype in *Drosophila* and induced significant variations of cellular gene
365 expression. Therefore, the poor conservation detected amongst *Drosophila* species indicates
366 that this sORF-encoded peptide differs from the few conserved ones characterized so far
367 and shed light on it by its recent invention.

368 Here, we tackled the question of whether the miPEP auto-regulatory function was
369 identical to that of plants. While we did not detect any auto-regulatory loop (miPEP
370 increasing the expression of its own *pri-miR* and miR), we observed that the action of miPEP-
371 8 is uncoupled from *miR-8* regulation. On the one hand, our data suggest that this peptide
372 could control similar developmental outcomes or developmental pathways and share the
373 regulation of identical subsets of genes. In this context, we further analyzed whether we
374 could detect a significant miPEP-8 activity in other *miR-8* developmental processes such as

375 intestinal stem cell differentiation [27] and eye morphogenesis [37] (data not shown).
376 However, no significant activity was detected, suggesting that, in the experimental
377 conditions tested, miPEP-8 does not act in all *miR-8* developmental pathways. Such an
378 example was observed in S2 cells in which *miR-8* is expressed at detectable levels whereas
379 endogenous miPEP-8 is not. On the other hand, we reveal that miPEP-8 likely has a
380 regulatory function all of its own, independently of *miR-8*. Indeed, our RNA-seq data
381 indicates that miPEP-8 regulates specific genes and biological processes (i.e. independent of
382 *miR-8* activity). This also occurs *in vivo* since few candidates of the top list of miPEP-8 specific
383 regulated genes identified in S2 cells are also modulated in miPEP-8 loss of function in adult
384 flies (Fig. S12). Future loss of function and expression pattern analyses throughout
385 development should bring further insight into miPEP-8-specific regulatory functions.

386 Is the uncoupling of miPEP activity from miR regulation a general feature of animal
387 miR genes? The study performed here suggests that the mechanisms involved in animals
388 might be different from the miPEP auto-regulatory mechanism observed in plants. As such, a
389 recent study on human miR155 revealed an activity for a miPEP155 that is not correlated to
390 miR155 control [20] and on human miR34 where a miPEP133 mitochondrial function
391 impinging on p53 activity was reported [21]. In light of these results, of course, we cannot
392 affirm that the mechanisms described here are common to all *Drosophila miR* genes. It
393 remains possible that some of them might be auto-regulated by their miPEPs as described in
394 plants. In addition, since ribosome occupancy were not found in all *Drosophila* microRNA
395 genes, it remains possible that some *pri-miRs* are unable to produce miPEPs. Therefore,
396 additional studies will be required to determine whether miPEP-dependent *pri-miR* auto-
397 regulation is specific or widespread amongst *miR* genes.

398 Is the *pri-miR* coding capacity conserved throughout the animal kingdom? In a search
399 for non-coding RNAs able to express sORF-encoded peptides, Razooki and co-workers found
400 that human *miR-22* host gene (*pri-miR-22*) produces a potential miPEP-22 that is induced
401 during viral infection [19]. sORFs have also recently been identified in *miR-200a* and *miR-*
402 *200b pri-miRs*, the human orthologs of the *Drosophila miR-8*. Like *miR-200a* and *miR-200b*,
403 miPEP-200a and b over-expression in prostate cancer cells inhibits migration of these cells
404 by regulating the vimentin-mediated pathway, suggesting that the miPEP-coding function of
405 *pri-miRs* is conserved in humans [18]. Accordingly, most recently, micropeptides encoded by
406 MIR155HG and MIR34HG were described to be involved in autoimmune inflammation by

407 controlling antigen presentation [20] and mitochondrial function respectively [21] via their
408 interaction with different HSP proteins. It is interesting to note however, that these miPEPs
409 appear to be involved in infections/pathologic-conditions, hence suggesting that revealing
410 miPEP function might be largely dependent on the biological context.

411 Ribosome-associated lncRNAs has been considered to constitute a hallmark of
412 protein translation. Here we found a signal of ribosome binding in the *pri-miR* of several
413 Drosophila *microRNAs* genes. Furthermore, we showed that the initiation codon of the
414 miORF present within *pri-mR-8* is in a favorable translational context. Indeed, after having
415 generated specific antibodies, we detected a peptide co-migrating with *in vitro* translated
416 miPEP-8 in fly extracts. However, an alternative possibility proposed by others is that
417 ribosome marks illustrate a mechanism for cellular control of lncRNA levels through
418 ribosome degradation-promoting activity [56, 59]. It will be of interest to investigate further
419 whether the short ORFs present in *pri-miRs* are able to influence their regulation by
420 controlling their stability and degradation as it has been shown for coding genes. Finally, the
421 molecules that give rise to miR-8/miPEP-8 are probably not the same ones since Drosha
422 processing would separate the ORF from the poly(A) tail and thereby cause rapid decapping
423 and degradation of the ORF-containing fragment. In light with these considerations, it is
424 difficult to conclude on a pervasive coding capacity of *pri-miRs* in *Drosophila*. Future work
425 will determine both in plants and animals whether all of them are sources of miPEPs and to
426 what extent their auto-regulatory capacity and/or modes of action are diverse and specific.
427

428 Conclusion:

429 Many studies performed recently have led to functional characterization of a handful
430 of additional SEPs in the plant and animal kingdom. Illustrating the diversity of functions of
431 these new players, these SEPs were identified from different sources of RNAs and play
432 different roles [9, 60, 61]. Among these, contrasting with their initial definition as non-
433 coding, pioneer works in plants showed that even precursors transcript of miRNAs produces
434 SEPs involved in an autoregulatory feedback loop. By addressing the conservation of this
435 mechanism in animals, our findings combined with others confirm that miR-encoded genes
436 probably represent evolutionary conserved bi-functional RNAs carrying coding and non-
437 coding functions. However, contrasting with the mechanism described in plants, our data

438 shed light on the diverse functions fulfilled by microRNA-encoded-peptides despite their
439 poor conservation among *Drosophila species*.

440 **Methods**

441

442 **Fly strains and Genetics**

443 *Drosophila* flies were maintained on standard cornmeal-yeast medium (Dutscher).
444 Experiments were performed at 25°C when *miR-8* GAL4 (NP5247) was used as driver. For the
445 experiments of wing phenotype of flies expressing transgenes under the control of MS1096
446 Gal4, crosses were placed at 28°C. UAS-*pri-miR*-8 and UAS-miPEP-8 transgenic lines were
447 inserted in attP86F site through PhiC31-mediated integration. Injections were performed by
448 Bestgene Inc (USA). Generating pri-miR-8 fly founder line: pri-miR-8 fly founder line was
449 designed and generated by inDroso Functional Genomics (Rennes, France) using
450 CRISPR/Cas9. The pri-miR-8 fly founder line was generated by excising from position
451 16830745 to 16831521 on Chromosome 2R arm and replacing it by an attP::loxP::3xP3-
452 dsRED::loxP cassette (Fig. S11). The two following guide RNA sequences were used to cut on
453 either side of the *pri-miR-8*: CACATATG|CAACGGAAAGAG and GTTGGTGG|TACTGAAGGTTA.
454 The edited region was verified by sequencing. The two pri-miR-8 constructs in pattB were
455 inserted at the Δ*miR-8* created P site. Three independent transformants were used for
456 analyses. The same strategy was used to generate the *miR-8* GAL4 driver, Δ*miR-8.* The
457 miPEP-8 alternative form creating a premature stop codon in miPEP-8 was derived from the
458 DGRP-859 line, outcrossed into white recipient flies and kept over CyO. Experiments are the
459 sum of at least 3 independent crosses. n indicates the number of individuals analysed. For
460 wing measurements, young flies (2-5 days) of the appropriate genotypes were stored in
461 Ethanol. For analysis of wings, females adult wings were removed in wash buffer (PBS and
462 0.1% Triton X-100), and mounted on a slide in 80% glycerol in PBS as described [62]. Wings
463 or wing discs images were acquired on a Zeiss Axiozoom stereomicroscope. Measurements
464 of wing size were performed using IMAGE J software.

465

466 **Molecular methods:**

467 For cloning procedure, miPEP-8, miR-8 or pri-miR-8 plasmids were constructed from PCR
468 amplification of genomic DNA, gene synthesis or by RNA reverse transcription from S2 cells
469 or adult *Drosophila* RNA and cloned in pUAS-attB vector constructs using the In-fusion HD
470 cloning kit (Takarabio) according to the manufacturer specification. All constructs were
471 verified by sequencing. For quantitative PCR experiments, total RNA was isolated from

472 young adult fruit flies (2-5 days) or S2 cells using TRI Reagent (Sigma) according to the
473 manufacturer specifications, followed by RQ1DNase treatment (promega) according the
474 manufacturer specifications. The cDNA template was synthesized using SuperScript III
475 (Invitrogen) with oligo-dT18 as anchor primers. Quantitative real-time PCR was performed
476 on the LightCycler 480 Instument II (Roche Life Science) using LightCycler480 SYBR GREEN I
477 master (Roche Life Science). The mRNA abundance of the examined genes was estimated by
478 qPCR. For the endogenous *pri-miR* or coding genes, RP49 and tubulin genes were used as
479 reference genes and used for normalization. For quantifying mature miRNA, stem loop PCR
480 conditions were set up and the small RNAs U14 and Sno442 were used as reference. Datas
481 presented are the same whatever the reference gene used. When the S2 cells are transiently
482 transfected, the co-transfected pActin-GAL4 vector (Addgene # 24344) was used to monitor
483 transfection efficiency. Typically, each experiment presented was performed with four
484 replicates processed independently and was repeated in time at least three times. All
485 experiments were taken into account and "n" indicate the total number of biological
486 replicates used for analyses. All primers used in the qPCR are listed in the supplemental
487 material (Table S1).
488
489 **RNA analysis**
490 For Rib-seq analyses, Dmel primiR three frames translations have been performed with
491 transeq (Emboss suite 6.6.0). A homemade script written in Perl was generated to compare
492 the resulting translated peptides to the ribo-seq sORF encoded peptides described in [4].
493 RNA-seq was processed by genewiz (Germany). Each dataset contains five independent
494 biological replicates of control *miR-8* and miPEP-8 over-expressing S2 cells RNA-seq. The
495 reads were subjected to standard quality control (QC) and filtered according to the following
496 parameters: (1) trimming and cleaning reads that aligned to primers and/or adaptors, (2)
497 reads with over 50% of low-quality bases (quality value ≤15) in one read, and (3) reads with
498 over 10% unknown bases (N bases). We used Trimmomatic software (v0.36) to remove
499 primers and bad quality reads. After filtering, we removed short reads (parameters were
500 used with default values. Gene and PSI lists for each dataset were compared to identify
501 common events between them. For RNAseq analysis, htseq-counts files were analyzed using
502 the version 3.24.3 of package EdgeR [63], in order to normalize raw counts by "trimmed
503 mean of M-values" (TMM), and test differential expression using the negative binomial

504 distribution. RNA-seq analysis : raw p-values were adjusted with the Benjamini–Hochberg
505 procedure to control the False Discovery Rate (FDR). A gene was declared differentially
506 expressed if it's adjusted p value ≤ 0.05. Heat map parameters applied: row-by-row
507 normalisation by standardisation (Mean and Standard deviation). GO term analysis was
508 performed with PANTHER (http://pantherdb.org/). Sashimi plots were created with IGV
509 (Integrative Genomics Viewer, https://igv.org/). Statistical analyses were performed using
510 the version 3.5.2 of R software and Bioconductor packages. For QPCR analysis, the
511 version 1.3-1 of package Agricolae was used.
512
513 **Cell culture and western blot and luciferase assays**
514 *Drosophila* S2 cells were maintained in Schneider's medium (Invitrogen) supplemented with
515 10% fetal bovine serum (Sigma), 50 U/ml penicillin and 50 µg/ml streptomycin (Invitrogen)
516 at 25°C. For western blot experiments, miPEPs sequences cloned into pF25A ICE T7 Flexi
517 vector were expressed in vitro using TnT® T7 Insect Cell Extract Protein Expression System
518 (Promega). For cells extracts and *Drosophila* extracts, we directly freeze them in nitrogen
519 just before western blot. Proteins were prepared in Laemli buffer (63 mM Tris HCl pH7.5, 2%
520 SDS, 5% 2-mercaptoethanol) and run on SDS-PAGE according to [13]. Primary antibodies
521 used for western were: rabbit anti-miPEP-8 were raised against the sequence
522 KQSDKQNSKERKKNTQI (generated and affinity purified by Agro-bio, France), mouse anti-
523 GAPDH (ThermoFisher AM4300), rabbit anti-Sra-1 (1/1000, provided by A. Giangrande,
524 IGBMC CNRS, France), mouse anti-peanut (1/100, DSHB, USA) and rabbit anti-ABP-1 (1:250,
525 provided by Michael Kessels Jena University Hospital, Germany). HRP conjugated secondary
526 antibodies are from Santa Cruz Biotechnology (1/10000 sc-516102). For luciferase assays, in
527 each experiment, S2 cells were transfected in quadruplicate, in 24-wells plates (700000
528 cells/well) using FuGene HD transfection Reagent (Promega). Experiments were repeated
529 timely independently at least 3 times. After 48h of transfection, cells were washed with PBS
530 and lysed with 100µL Passive Lysis Buffer (Dual luciferase Reporter Assay System, Promega).
531 Firefly luciferase (FL) and Renilla luciferase (RL) activities were then quantified with DUAL
532 luciferase reporter assay (Promega) using 50µl of reagents/well and measure using a Greiner
533 luminometer instrument.
534
535 **Statistical analyses**

Statistical analyses were performed using GraphPad Prism and illustrated as follow: * p-value<0,05 ; ** p-value<0,01 ; *** p-value<0,001; **** p-value<0,0001. In all experiments, results represent mean ±s.e.m. n represente the number of biological independent replicates. Normality test were first performed using D'Agostino Pearson test. If the distribution is Gaussian and in order to detect a global difference between all groups, one-way ANOVA was performed using one-way analyses of variances followed by Bartlett's test for equal variance and Bonferroni's multiple comparison tests. In other cases, when variance or sample sizes are not equal, non-parametric analyses were performed using Kruskal-Wallis test to detect a global difference between all groups followed by comparisons between two groups performed using adjustments for multiple comparisons. When only two groups were compared, a Mann & Witney test was performed.

Declarations

- Ethics approval and consent to participate

Not applicable

- Consent for publication

Not applicable

- Availability of Data and Materials

The RNAseq datasets generated during the current study are available in the NCBI repository under the accession number BioProject ID PRJNA645280 [http://www.ncbi.nlm.nih.gov/bioproject/645280]
The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

- Competing interests

The authors declare that they have no competing interests

- Funding

This work was funded by the French ANR project BiomiPEP (ANR-16-CE12-0018-01) and the Association de Recherche sur le Cancer (ARC PGA1RF2018206987). It was carried out in the Laboratoire de Recherche en Sciences Végétales, which belongs to the Laboratoire d'Excellence entitled TULIP (ANR-10-LABX-41). PT was supported by a PhD fellowship from the ANR BiomiPEP and AM from the ARC respectively.

- Authors' contributions

SP and JPC conceived the project. SP designed the research. AM, PT, CD, PV, DL SP performed the experiments. HSC and MA performed bioinformatics analyzes. SP, AM, PT, JPC wrote the paper.

- Acknowledgements

We thank C.Dozier and the « Peptides & small RNAs » team members for critical reading and their helpful comments on the manuscript. We thank H.Boukhatmi, A.Giangrande, S.Cohen, F.Juge for sharing flies, constructs, and antibodies. We thank the FRAIB RIO Imaging facility for microscopy. We acknowledge the Bloomington *Drosophila* Stock Center (NIH P40OD018537) and the Drosophila Stock Center, KYOTO Stock Center (DGRC) in Kyoto Institute of Technology for providing flies stocks, the Developmental Studies Hybridoma Bank (DSHB) for antibodies, and DNASU and Addgene for plasmids.

Additional File 1: list of sORF peptides defined from Rib-seq analyses

Additional File 2: normalized RNA-seq values from S2 cells.

Additional File 3: PANTHER Overrepresentation Test.

599
600
601
602 **Figure Legends**
603
604 **Figure 1: Translatability of *pri-miR-8*.**
605 **A:** Model of miPEP regulation in plants. **B:** Box plot representation of the number of ORFs in
606 different classes of RNAs. 3'UTR, 5'UTR and CDS represent coding RNAs, whereas lncRNAs
607 and pri-miRs represent non-coding RNAs. An ORF was defined as starting with an ATG and
608 coding for a minimum of 10 amino acids. *Pri-miRs* reveal comparable numbers of ORFs/kb
609 as lncRNAs. **C:** GWIPS-vis [64] genome viewer of the *Drosophila miR-8* locus. Top: genomic
610 positions and ORFs in the three reading frames. Green bars define ATGs and red bars stop
611 codons. Below, RNA-seq profile is shown in green and ribosome profiling is shown in red.
612 The blue horizontal lines represent the two CR43650 non-coding RNA transcripts and
613 potential *miR-8 pri-miRNAs*. In black is schematized the transcript we identified as
614 detectable *pri-miR-8*. Bottom: miPEP-8 amino acid sequence is shown. **D:** western blot
615 experiments using the anti-miPEP-8 antibody. Left panel: *in vitro* synthetized miPEP-8
616 corresponding to the constructs indicated on top. The asterisk indicates the upstream
617 initiated peptide. The arrow indicates the miPEP-8 initiated at ATG1. Middle panel: detection
618 of miPEP-8 in S2 cells over-expressing miPEP-8 placed in different translational contexts;
619 Kozak (optimal); K mt (ATG mutated into TGA); mt (ATG mutated into AGT). Note that the
620 *pri-miR* is translated and endogenous miPEP-8 expression is undetectable in S2 cells. Right
621 panel: anti miPEP-8 western blot of adult *Drosophila* extracts and in the *miR-8* deleted line
622 Δ2/Δ2 [26] in which no miPEP-8 is detected. We noticed the presence of non-specific bands
623 as well as additional specific bands representing possibly miPEP-8 multimeric forms or PTM
624 modifications. Ctrl corresponds to cell extracts transfected with an empty vector.
625
626 **Figure 2: *Drosophila* miPEP-8 is biologically active during development.**
627 **A:** Lethality assay on flies over-expressing *miR-8*, miPEP-8 or miPEP-8mt (ATG mutated) using
628 the *miR-8* GAL4 driver. Left: details of the genetic cross and expected percentage depending
629 on the effect (neutral, deleterious or advantageous) on *Drosophila* development. Right:
630 graph indicating the percentage of hatched flies over-expressing the different constructs.

631 *white* flies (*w*) crossed with the driver line were used as a control. Expressing *miR-8* resulted
632 in developmental lethality since less than 20% of flies hatched (expected value 50%). A
633 significant decrease occurred following miPEP-8 over-expression but not with the
634 untranslatable miPEP-8mt construct. Number of independent crosses: for *w* and *miR-8* n=
635 23; for miPEP-8 wt and mt n=24. **B:** Same as in **A** except the constructs were expressed in
636 wings using the MS1096 driver and the phenotypes scored on wing size. Number of wings
637 analyzed: for *w* n= 20; for *miR-8* n=27; for miPEP-8 wt and mt n= 27. * or ns: Significant
638 differences are indicated relative to white recipient flies. AU: Arbitrary Units.
639
640 **Figure 3: *Pri-miR-8* expression is independent of miPEP-8 control/activity.**
641 **A:** schematic representation of constructs tested on *miR-8* expression and activity levels.
642 Arrows locate the primers used in the qPCR experiments determining miPEP and *pri-miR*
643 relative expression levels. **B:** the characterized *pri-miR-8* produces a mature *miR-8*. S2 cells
644 were transfected with a vector expressing *pri-miR-8*. Left: detection of the over-expression
645 level of *pri-miR-8* by qPCR. Right: detection of the over-expression level of mature *miR-8*
646 using the same RNA samples, n=11 **C:** miPEP-8 lacks repressive activity towards *miR-8*
647 expression. Left: level of miPEP-8 over-expression. Middle and right panels: quantification of
648 *pri-miR-8* and mature *miR-8* in miPEP-8 over-expressing cells compared to control
649 transfected cells (ctrl). n= 13 for the ctrl and 14 for miPEP-8. **D** and **E**: Insensitivity of *miR-8*
650 sensor to miPEP-8 over-expression in S2 transfected cells (n=16) (D) or in wing imaginal discs
651 when miPEP-8 is expressed under the *ptc*-GAL4 promoter (E). In D, a *miR-8* construct (n=12)
652 [17] was used as a positive control repressing the *miR-8* luciferase sensor [20]. Of note, *pri-*
653 *miR-8* (n=21) also repressed the *miR-8* luciferase sensor. In E, first panel to the left: *ptc* GAL4
654 crossed with a UAS mCherry. Second panel to the left: expression pattern of the GFP *miR-8*
655 sensor alone. Scale bars (white) indicate 100mm. A repressive activity is observed with *miR-8*
656 expressed in the *ptc* domain but not with miPEP-8. A representative disc is shown out of ten
657 analysed.
658
659 **Figure 4: targeting miPEP-8 *in vivo* in *Drosophila* induces a wing phenotype.**
660 **A:** Strategy for endogenous miPEP-8 edition. The *pri-miR-8* gene region was deleted by
661 CRISPR and a P landing site was created. Wild type and miPEP-8 ATG mutated *pri-miR-8* in
662 pattB were inserted at the P landing site. **B**: Similar rescue efficiency was observed in at least

663 three independent transgenic lines (left panel). qPCR on mature miR-8 in wild type and
664 mutant (mt) *pri-miR-8* Knock In (KI) lines showed similar miR-8 levels (n=4). (right panel) **C**:
665 wing phenotype in *miR-8* deletion edited line. The *pri-miR-8* miPEP-8 mutated (mt) shows a
666 reduced wing size compared to the wild type *pri-miR-8*. (n= 15 and 28 respectively). D to F:
667 analyses in miPEP-8 mutant identified in DGRP polymorphism **D**: miR-8 level determined by
668 qPCR in white recipient flies (w) and in white flies carrying the miPEP-8 truncated form
669 (miPEP-8alt). (n=6 and 8 respectively) **E** and **F**: Wing size determination in different genetic
670 contexts. miPEP-8alt homozygotes or over *miR-8* deficiencies revealed significant reduced
671 wing size relative to the white recipient flies (w, n=19; miPEP-8alt, n= 21; miPEP-8alt/ miR-8
672 deltions, n=40). Expressing miPEP-8 rescued the wing phenotype of miPEP-8alt flies relative
673 to sibling flies not expressing miPEP-8 (n= 18 and 28 respectively). Significant (*) or non
674 significant (ns) differences are indicated either relative to white recipient flies or between
675 the two groups.
676
677 **Figure 5: Uncoupled activity of miR-8 and miPEP-8**
678 **A**: Rescue assay of *miR-8*- or miPEP-8- induced wing phenotype in flies co-over-expressing
679 *miR-8* or miPEP-8 along with a miR-8 sponge (miR-8sp) or a miR-8 scramble (miR-8scr). Only
680 miR-8-sp (and not miR-8scr) compensates for *miR-8*-induced wing size reduction, hence
681 efficiently titrating *miR-8*, while it has no effect on miPEP-8-induced wing phenotype. **B**:
682 Quantification of A. “ctrl” (MS1096/+) n=19; "mir-8; mir-8scr" n=20; "mir-8; mir-8sp" n=21;
683 "miPEP-8; mir-8scr" n=23; "miPEP-8; mir-8sp" n=19. * p<0,05.
684
685 **Figure 6: miR-8 and miPEP-8 control distinct set of genes**
686 **A**: Heatmap representing the RNA-seq results obtained from S2 cells over-expressing either
687 *miR-8* or miPEP-8. Significant sets of genes are modulated in response to *mirR-8* or miPEP-8
688 over-expression, when compared to control transfected cells (ctrl). N=5. **B:** Venn diagram
689 representing the *miR-8* versus miPEP-8 modulated genes. **C, D, E**: different subgroups are
690 distinguished; miPEP-8 specific (**C**), *miR-8* specific (**D**), and co-regulated by miPEP-8 and *miR-*
691 *8* (**E**).
692
693

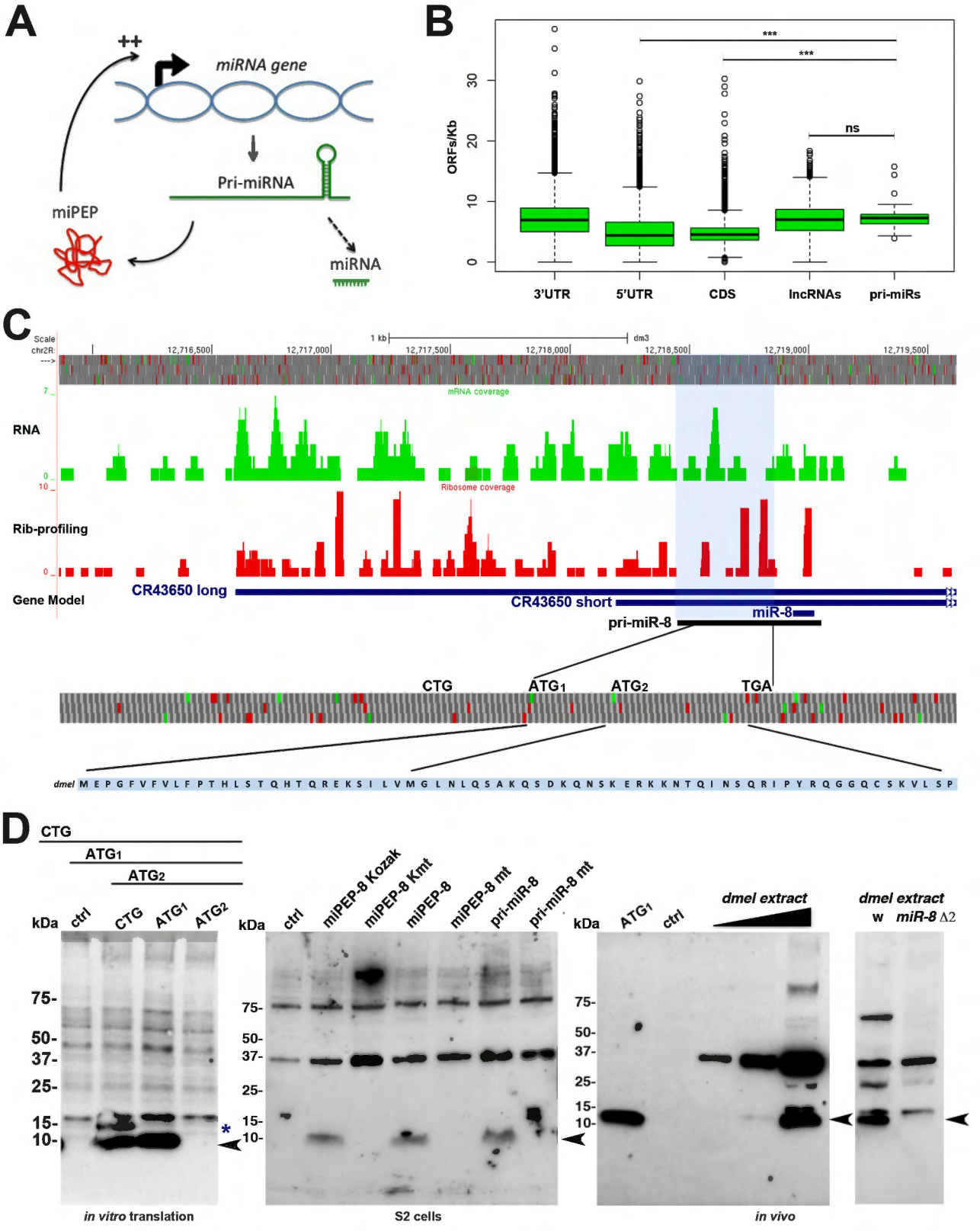
696 ## References

697 1. Cech TR, Steitz JA: **The noncoding RNA revolution-trashing old rules to forge**
698 **new ones.** *Cell* 2014, **157:**77-94.
699 2. Gardini A, Shiekhattar R: **The many faces of long noncoding RNAs.** *FEBS J* 2015,
700 **282:**1647-1657.
701 3. Ransohoff JD, Wei Y, Khavari PA: **The functions and unique features of long**
702 **intergenic non-coding RNA.** *Nat Rev Mol Cell Biol* 2018, **19:**143-157.
703 4. Aspden JL, Eyre-Walker YC, Phillips RJ, Amin U, Mumtaz MA, Brocard M, Couso JP:
704 **Extensive translation of small Open Reading Frames revealed by Poly-Ribo-**
705 **Seq.** *Elife* 2014, **3:**e03528.
706 5. Bazzini AA, Johnstone TG, Christiano R, Mackowiak SD, Obermayer B, Fleming ES,
707 Vejnar CE, Lee MT, Rajewsky N, Walther TC, Giraldez AJ: **Identification of small**
708 **ORFs in vertebrates using ribosome footprinting and evolutionary**
709 **conservation.** *EMBO J* 2014, **33:**981-993.
710 6. de Andres-Pablo A, Morillon A, Wery M: **LncRNAs, lost in translation or licence**
711 **to regulate?** *Current Genetics* 2017, **63:**29-33.
712 7. Ingolia NT, Lareau LF, Weissman JS: **Ribosome profiling of mouse embryonic**
713 **stem cells reveals the complexity and dynamics of mammalian proteomes.**
714 *Cell* 2011, **147:**789-802.
715 8. Plaza: **In search of lost small peptides.** *Annual reviews* 2017.
716 9. Rathore A, Martinez TF, Chu Q, Saghatelian A: **Small, but mighty? Searching for**
717 **human microproteins and their potential for understanding health and**
718 **disease.** *Expert Rev Proteomics* 2018, **15:**963-965.
719 10. Ruiz-Orera J, Alba MM: **Translation of Small Open Reading Frames: Roles in**
720 **Regulation and Evolutionary Innovation.** *Trends Genet* 2019, **35:**186-198.
721 11. Yeasmin F, Yada T, Akimitsu N: **Micropeptides Encoded in Transcripts**
722 **Previously Identified as Long Noncoding RNAs: A New Chapter in**
723 **Transcriptomics and Proteomics.** *Front Genet* 2018, **9:**144.
724 12. O'Brien J, Hayder H, Zayed Y, Peng C: **Overview of MicroRNA Biogenesis,**
725 **Mechanisms of Actions, and Circulation.** *Front Endocrinol (Lausanne)* 2018,
726 **9:**402.
727 13. Lauressergues D, Couzigou JM, Clemente HS, Martinez Y, Dunand C, Becard G,
728 Combier JP: **Primary transcripts of microRNAs encode regulatory peptides.**
729 *Nature* 2015, **520:**90-93.
730 14. Couzigou JM, Andre O, Guillotin B, Alexandre M, Combier JP: **Use of microRNA-**
731 **encoded peptide miPEP172c to stimulate nodulation in soybean.** *New Phytol*
732 2016, **211:**379-381.
733 15. Sharma S, Badola PK, Bhatia C, Sharma D, Trivedi PK: **miRNA-encoded peptide,**
734 **miPEP858, regulates plant growth and development in Arabidopsis.** *BioXiv*
735 2019, **642561**.
736 16. Chen QJ, Deng BH, Gao J, Zhao ZY, Chen ZL, Song SR, Wang L, Zhao LP, Xu WP,
737 Zhang CX, Ma C, Wang SP: **An miRNA-encoded small peptide, vvi-**
738 **miPEP171d1, regulates adventitious root formation.** *Plant Physiol* 2020.
739 17. Zhang QL, Su LY, Zhang ST, Xu XP, Chen XH, Li X, Jiang MQ, Huang SQ, Chen YK,
740 Zhang ZH, Lai ZX, Lin YL: **Analyses of microRNA166 gene structure,**
741 **expression, and function during the early stage of somatic embryogenesis**
742 **in Dimocarpus longan Lour.** *Plant Physiol Biochem* 2020, **147:**205-214.

743 18. Fang J, Morsalin S, Rao V, Reddy ES: **Decoding of Non-Coding DNA and Non-**
744 **Coding RNA: Pri-Micro RNA-Encoded Novel Peptides Regulate Migration of**
745 **Cancer Cells.** *J Pharm Sci Pharmacol* 2017, **3:**23-27.
746 19. Razooky BS, Obermayer B, O'May JB, Tarakhovsky A: **Viral Infection Identifies**
747 **Micropeptides Differentially Regulated in smORF-Containing lncRNAs.**
748 *Genes (Basel)* 2017, **8**.
749 20. Niu LM, Lou FZ, Sun Y, Sun LB, Cai XJ, Liu ZY, Zhou H, Wang H, Wang ZK, Bai J, Yin
750 QQ, Zhang JX, Chen LJ, Peng DH, Xu ZY, Gao YY, Tang SB, Fan L, Wang HL: **A**
751 **micropeptide encoded by lncRNA MIR155HG suppresses autoimmune**
752 **inflammation via modulating antigen presentation.** *Science Advances* 2020, **6**.
753 21. Kang M, Tang B, Li JX, Zhou ZY, Liu K, Wang RS, Jiang ZY, Bi FF, Patrick D, Kim D,
754 Mitra AK, Yang-Hartwich Y: **Identification of miPEP133 as a novel tumor-**
755 **suppressor microprotein encoded by miR-34a pri-miRNA.** *Molecular Cancer*
756 2020, **19**.
757 22. Sander M, Herranz H: **MicroRNAs in Drosophila Cancer Models.** *Adv Exp Med*
758 *Biol* 2019, **1167:**157-173.
759 23. Jin H, Kim VN, Hyun S: **Conserved microRNA miR-8 controls body size in**
760 **response to steroid signaling in Drosophila.** *Genes Dev* 2012, **26:**1427-1432.
761 24. Karres JS, Hilgers V, Carrera I, Treisman J, Cohen SM: **The conserved microRNA**
762 **miR-8 tunes atrophin levels to prevent neurodegeneration in Drosophila.**
763 *Cell* 2007, **131:**136-145.
764 25. Umegawachi T, Yoshida H, Koshida H, Yamada M, Ohkawa Y, Sato T, Suyama M,
765 Krause HM, Yamaguchi M: **Control of tissue size and development by a**
766 **regulatory element in the yorkie 3'UTR.** *Am J Cancer Res* 2017, **7:**673-687.
767 26. Hyun S, Lee JH, Jin H, Nam J, Namkoong B, Lee G, Chung J, Kim VN: **Conserved**
768 **MicroRNA miR-8/miR-200 and its target USH/FOG2 control growth by**
769 **regulating PI3K.** *Cell* 2009, **139:**1096-1108.
770 27. Antonello ZA, Reiff T, Ballesta-Illan E, Dominguez M: **Robust intestinal**
771 **homeostasis relies on cellular plasticity in enteroblasts mediated by miR-8-**
772 **Escargot switch.** *EMBO J* 2015, **34:**2025-2041.
773 28. Antonello ZA, Reiff T, Dominguez M: **Mesenchymal to epithelial transition**
774 **during tissue homeostasis and regeneration: Patching up the Drosophila**
775 **midgut epithelium.** *Fly (Austin)* 2015, **9:**132-137.
776 29. Boukhatmi H, Bray S: **A population of adult satellite-like cells in Drosophila is**
777 **maintained through a switch in RNA-isoforms.** *Elife* 2018, **7**.
778 30. Morante J, Vallejo DM, Desplan C, Dominguez M: **Conserved miR-8/miR-200**
779 **defines a glial niche that controls neuroepithelial expansion and neuroblast**
780 **transition.** *Dev Cell* 2013, **27:**174-187.
781 31. Loya CM, McNeill EM, Bao H, Zhang B, Van Vactor D: **miR-8 controls synapse**
782 **structure by repression of the actin regulator enabled.** *Development* 2014,
783 **141:**1864-1874.
784 32. Lu CS, Zhai B, Mauss A, Landgraf M, Gygi S, Van Vactor D: **MicroRNA-8 promotes**
785 **robust motor axon targeting by coordinate regulation of cell adhesion**
786 **molecules during synapse development.** *Philos Trans R Soc Lond B Biol Sci*
787 2014, **369**.
788 33. Bolin K, Rachmaninoff N, Moncada K, Pula K, Kennell J, Buttitta L: **miR-8**
789 **modulates cytoskeletal regulators to influence cell survival and epithelial**
790 **organization in Drosophila wings.** *Dev Biol* 2016, **412:**83-98.

791 34. Choi IK, Hyun S: **Conserved microRNA miR-8 in fat body regulates innate**
792 **immune homeostasis in Drosophila.** *Dev Comp Immunol* 2012, **37:**50-54.
793 35. Etebari K, Asgari S: **Conserved microRNA miR-8 blocks activation of the Toll**
794 **pathway by upregulating Serpin 27 transcripts.** *RNA Biol* 2013, **10:**1356-
795 1364.
796 36. Kennell JA, Cadigan KM, Shakhmantsir I, Waldron EJ: **The microRNA miR-8 is a**
797 **positive regulator of pigmentation and eclosion in Drosophila.** *Dev Dyn*
798 2012, **241:**161-168.
799 37. Kennell JA, Gerin I, MacDougald OA, Cadigan KM: **The microRNA miR-8 is a**
800 **conserved negative regulator of Wnt signaling.** *Proc Natl Acad Sci U S A* 2008,
801 **105:**15417-15422.
802 38. Lee GJ, Hyun S: **Multiple targets of the microRNA miR-8 contribute to**
803 **immune homeostasis in Drosophila.** *Dev Comp Immunol* 2014, **45:**245-251.
804 39. Lee GJ, Jun JW, Hyun S: **MicroRNA miR-8 regulates multiple growth factor**
805 **hormones produced from Drosophila fat cells.** *Insect Mol Biol* 2015, **24:**311-
806 318.
807 40. Sander M, Eichenlaub T, Herranz H: **Oncogenic cooperation between Yorkie**
808 **and the conserved microRNA miR-8 in the wing disc of Drosophila.**
809 *Development* 2018, **145**.
810 41. Vallejo DM, Caparros E, Dominguez M: **Targeting Notch signalling by the**
811 **conserved miR-8/200 microRNA family in development and cancer cells.**
812 *EMBO J* 2011, **30:**756-769.
813 42. Zeng C, Fukunaga T, Hamada M: **Identification and analysis of ribosome-**
814 **associated lncRNAs using ribosome profiling data.** *BMC Genomics* 2018,
815 **19:**414.
816 43. Hsu PY, Calviello L, Wu HL, Li FW, Rothfels CJ, Ohler U, Benfey PN: **Super-**
817 **resolution ribosome profiling reveals unannotated translation events in**
818 **Arabidopsis.** *Proc Natl Acad Sci U S A* 2016, **113:**E7126-7135.
819 44. Ingolia NT, Ghaemmaghami S, Newman JR, Weissman JS: **Genome-wide analysis**
820 **in vivo of translation with nucleotide resolution using ribosome profiling.**
821 *Science* 2009, **324:**218-223.
822 45. Dunn JG, Foo CK, Belletier NG, Gavis ER, Weissman JS: **Ribosome profiling**
823 **reveals pervasive and regulated stop codon readthrough in Drosophila**
824 **melanogaster.** *Elife* 2013, **2:**e01179.
825 46. Kronja I, Whitfield ZJ, Yuan BB, Dzeyk K, Kirkpatrick J, Krijgsveld J, Orr-Weaver
826 TL: **Quantitative proteomics reveals the dynamics of protein changes during**
827 **Drosophila oocyte maturation and the oocyte-to-embryo transition.**
828 *Proceedings of the National Academy of Sciences of the United States of America*
829 2014, **111:**16023-16028.
830 47. Enderle D, Beisel C, Stadler MB, Gerstung M, Athri P, Paro R: **Polycomb**
831 **preferentially targets stalled promoters of coding and noncoding**
832 **transcripts.** *Genome Res* 2011, **21:**216-226.
833 48. Qian J, Zhang Z, Liang J, Ge Q, Duan X, Ma F, Li F: **The full-length transcripts and**
834 **promoter analysis of intergenic microRNAs in Drosophila melanogaster.**
835 *Genomics* 2011, **97:**294-303.
836 49. Lucas KJ, Roy S, Ha J, Gervaise AL, Kokoza VA, Raikhel AS: **MicroRNA-8 targets**
837 **the Wingless signaling pathway in the female mosquito fat body to regulate**
838 **reproductive processes.** *Proc Natl Acad Sci U S A* 2015, **112:**1440-1445.

839 50. Mackay TFC, Richards S, Stone EA, Barbadilla A, Ayroles JF, Zhu DH, Casillas S,
840 Han Y, Magwire MM, Cridland JM, Richardson MF, Anholt RRH, Barron M, Bess C,
841 Blankenburg KP, Carbone MA, Castellano D, Chaboub L, Duncan L, Harris Z, Javaid
842 M, Jayaseelan JC, Jhangiani SN, Jordan KW, Lara F, Lawrence F, Lee SL, Librado P,
843 Linheiro RS, Lyman RF*, et al*: **The Drosophila melanogaster Genetic Reference**
844 **Panel.** *Nature* 2012, **482:**173-178.
845 51. Dinger ME, Pang KC, Mercer TR, Mattick JS: **Differentiating protein-coding and**
846 **noncoding RNA: challenges and ambiguities.** *PLoS Comput Biol* 2008,
847 **4:**e1000176.
848 52. Ji Z, Song R, Regev A, Struhl K: **Many lncRNAs, 5'UTRs, and pseudogenes are**
849 **translated and some are likely to express functional proteins.** *Elife* 2015,
850 **4:**e08890.
851 53. Ladoukakis E, Pereira V, Magny EG, Eyre-Walker A, Couso JP: **Hundreds of**
852 **putatively functional small open reading frames in Drosophila.** *Genome Biol*
853 2011, **12:**R118.
854 54. Pauli A, Valen E, Schier AF: **Identifying (non-)coding RNAs and small peptides:**
855 **challenges and opportunities.** *Bioessays* 2015, **37:**103-112.
856 55. Raj A, Wang SH, Shim H, Harpak A, Li YI, Engelmann B, Stephens M, Gilad Y,
857 Pritchard JK: **Thousands of novel translated open reading frames in humans**
858 **inferred by ribosome footprint profiling.** *Elife* 2016, **5:**e13328.
859 56. Smith JE, Alvarez-Dominguez JR, Kline N, Huynh NJ, Geisler S, Hu W, Coller J,
860 Baker KE: **Translation of small open reading frames within unannotated**
861 **RNA transcripts in Saccharomyces cerevisiae.** *Cell Rep* 2014, **7:**1858-1866.
862 57. Andrews SJ, Rothnagel JA: **Emerging evidence for functional peptides**
863 **encoded by short open reading frames.** *Nat Rev Genet* 2014, **15:**193-204.
864 58. Saghatelian A, Couso JP: **Discovery and characterization of smORF-encoded**
865 **bioactive polypeptides.** *Nat Chem Biol* 2015, **11:**909-916.
866 59. Carlevaro-Fita J, Rahim A, Guigo R, Vardy LA, Johnson R: **Cytoplasmic long**
867 **noncoding RNAs are frequently bound to and degraded at ribosomes in**
868 **human cells.** *RNA* 2016, **22:**867-882.
869 60. Chugunova A, Loseva E, Mazin P, Mitina A, Navalayeu T, Bilan D, Vishnyakova P,
870 Marey M, Golovina A, Serebryakova M, Pletnev P, Rubtsova M, Mair W,
871 Vanyushkina A, Khaitovich P, Belousov V, Vysokikh M, Sergiev P, Dontsova O:
872 **LINC00116 codes for a mitochondrial peptide linking respiration and lipid**
873 **metabolism.** *Proc Natl Acad Sci U S A* 2019.
874 61. Khitun A, Ness TJ, Slavoff SA: **Small open reading frames and cellular stress**
875 **responses.** *Mol Omics* 2019.
876 62. Gault WJ, Olguin P, Weber U, Mlodzik M: **Drosophila CK1-gamma, gilgamesh,**
877 **controls PCP-mediated morphogenesis through regulation of vesicle**
878 **trafficking.** *J Cell Biol* 2012, **196:**605-621.
879 63. McCarthy DJ, Chen Y, Smyth GK: **Differential expression analysis of**
880 **multifactor RNA-Seq experiments with respect to biological variation.**
881 *Nucleic Acids Res* 2012, **40:**4288-4297.
882 64. Michel AM, Fox G, A MK, De Bo C, O'Connor PB, Heaphy SM, Mullan JP, Donohue
883 CA, Higgins DG, Baranov PV: **GWIPS-viz: development of a ribo-seq genome**
884 **browser.** *Nucleic Acids Res* 2014, **42:**D859-864.
885


A
++
miRNA gene
Pri-miRNA
miPEP
miRNA
B
ORFs/Kb
0
10
20
30
***
***
ns
3'UTR
5'UTR
CDS
lncRNAs
pri-miRs
C
Scale
chr2R
1 kb
dm3
12,716,500
12,717,000
12,717,500
12,718,000
12,718,500
12,719,000
12,719,500
mRNA coverage
RNA
Ribosome coverage
Rib-profiling
Gene Model
CR43650 long
CR43650 short
pri-miR-8
miR-8
CTG
ATG1
ATG2
TGA
dmel M E P G F V F V L F P T H L S T Q H T Q R E K S I L V M G L N L Q S A K Q S D K Q N S K E R K K N T Q I N S Q R I P Y R Q G G Q C S K V L S P
D
CTG
ATG1
ATG2
kDa
ctrl
CTG
ATG1
ATG2
75-
50-
37-
25-
15-
10-
*
in vitro translation
ctrl
miPEP-8 Kozak
miPEP-8 Kmt
miPEP-8
miPEP-8 mt
pri-miR-8
pri-miR-8 mt
S2 cells
ATG1
ctrl
dmel extract
in vivo
dmel extract
w
miR-8 Δ2

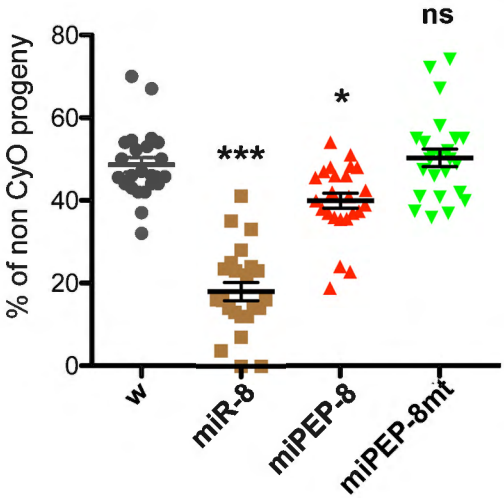
A

♂ UAS-construct (hz) X ♀ *miR-8* GAL4 / CyO

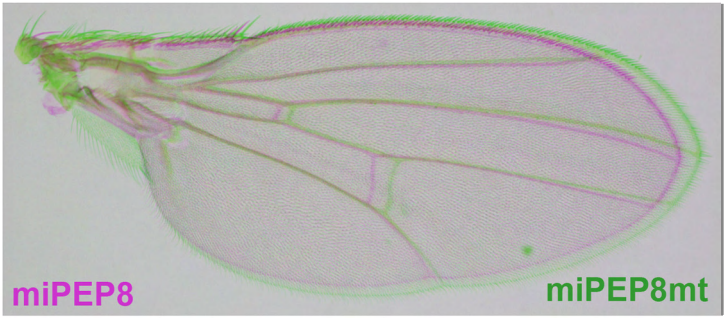
↓

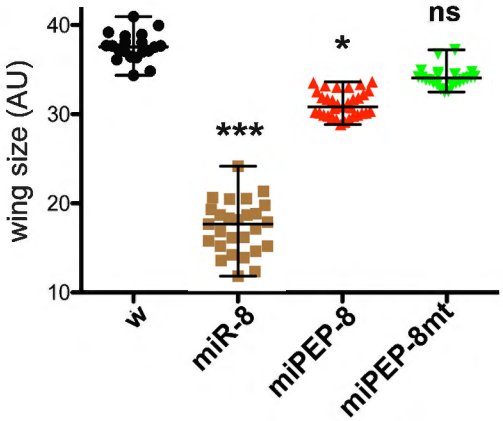
Progeny: number of non CyO/total, if expression is

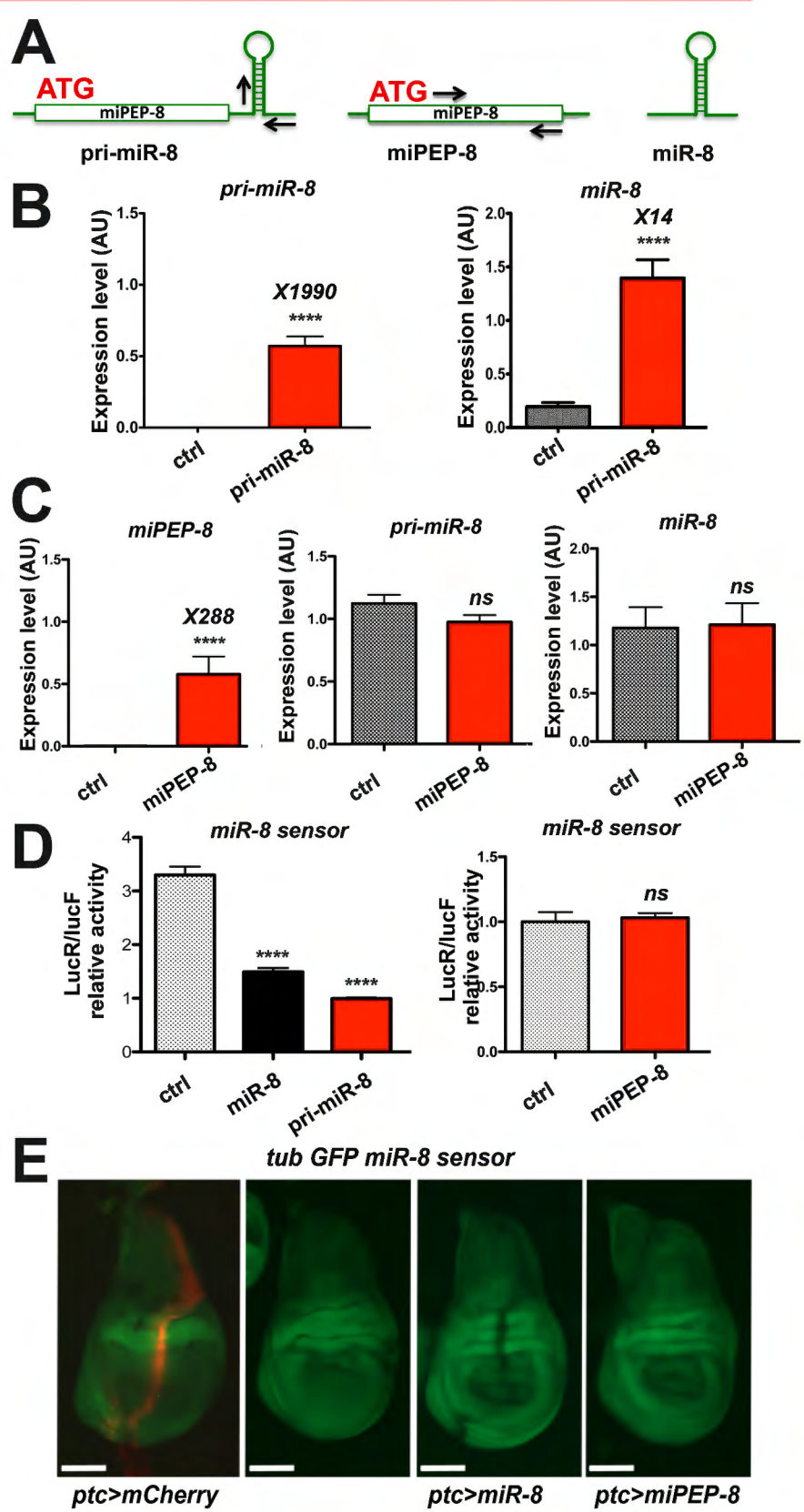
- neutral, expected value 50%
- deleterious < 50%
- Advantageous > 50%

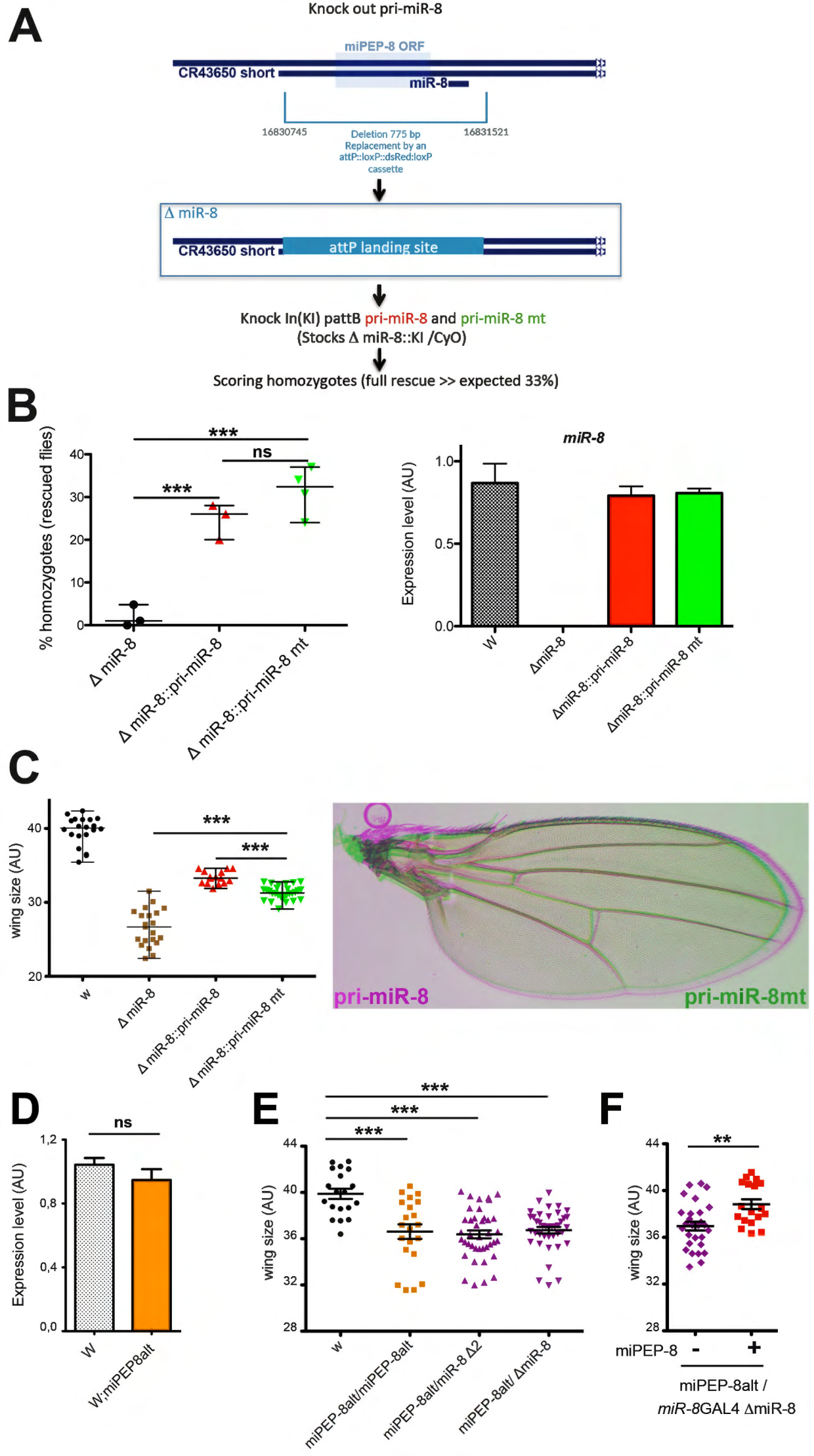


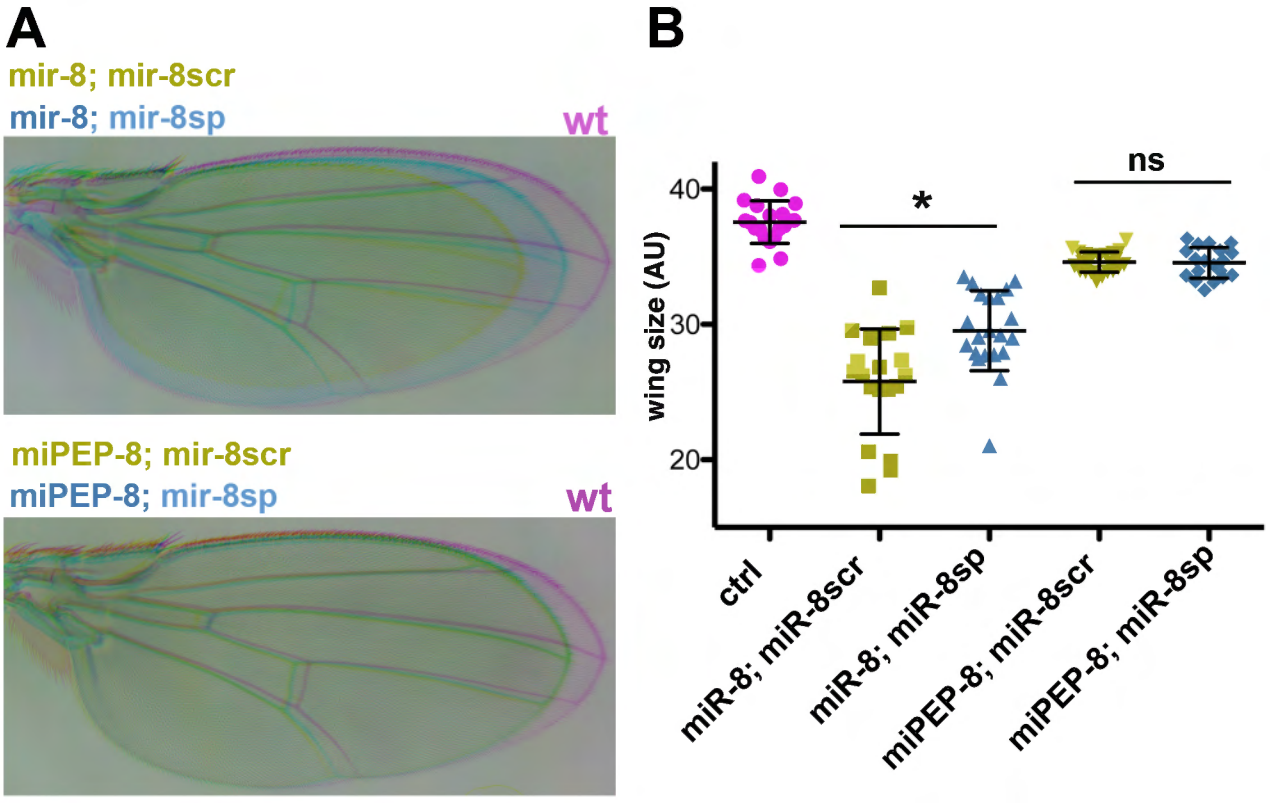
B

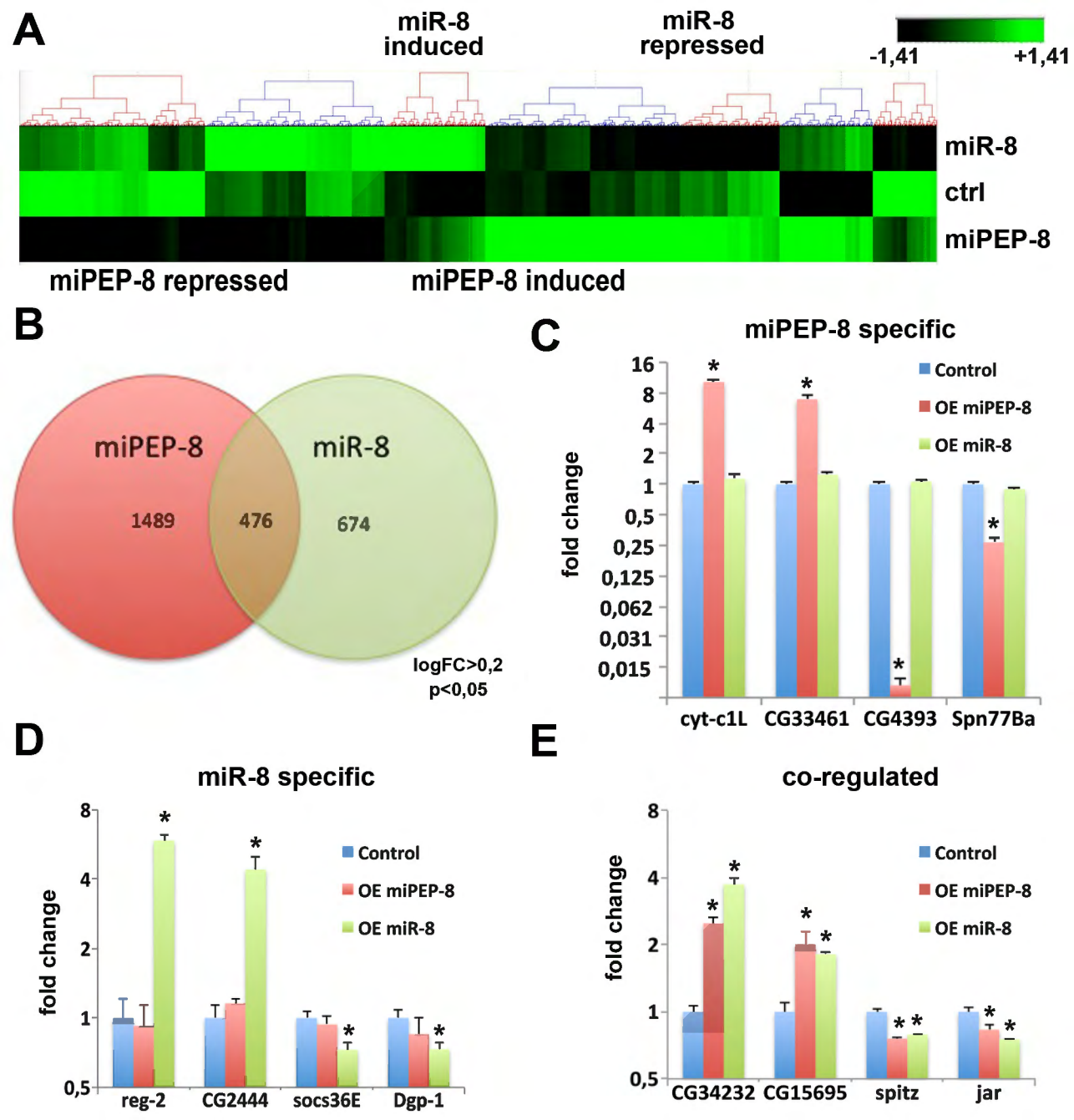





A
ATG
miPEP-8
pri-miR-8
ATG
miPEP-8
miPEP-8
miR-8
B
pri-miR-8
Expression level (AU)
1.5
1.0
0.5
0.0
X1990
****
ctrl
pri-miR-8
miR-8
Expression level (AU)
2.0
1.5
1.0
0.5
0.0
X14
****
ctrl
pri-miR-8
C
miPEP-8
Expression level (AU)
1.5
1.0
0.5
0.0
X288
****
ctrl
miPEP-8
pri-miR-8
Expression level (AU)
1.5
1.0
0.5
0.0
ns
ctrl
miPEP-8
miR-8
Expression level (AU)
2.0
1.5
1.0
0.5
0.0
ns
ctrl
miPEP-8
D
miR-8 sensor
LucR/lucF relative activity
4
3
2
1
0
****
****
ctrl
miR-8
pri-miR-8
miR-8 sensor
LucR/lucF relative activity
1.5
1.0
0.5
0.0
ns
ctrl
miPEP-8
E
tub GFP miR-8 sensor
ptc>mCherry
ptc>miR-8
ptc>miPEP-8


A
Knock out pri-miR-8
miPEP-8 ORF
CR43650 short
miR-8
16830745
16831521
Deletion 775 bp
Replacement by an
attP::loxP::dsRed:loxP
cassette
Δ miR-8
CR43650 short
attP landing site
Knock In(KI) pattB pri-miR-8 and pri-miR-8 mt
(Stocks Δ miR-8::KI /CyO)
Scoring homozygotes (full rescue >> expected 33%)
B
% homozygotes (rescued flies)
***
ns
***
Δ miR-8
Δ miR-8::pri-miR-8
Δ miR-8::pri-miR-8 mt
miR-8
Expression level (AU)
w
ΔmiR-8
ΔmiR-8::pri-miR-8
ΔmiR-8::pri-miR-8 mt
C
wing size (AU)
***
***
w
Δ miR-8
Δ miR-8::pri-miR-8
Δ miR-8::pri-miR-8 mt
pri-miR-8
pri-miR-8mt
D
ns
Expression level (AU)
W
W;miPEP8alt
E
***
***
***
wing size (AU)
w
miPEP-8alt/miPEP-8alt
miPEP-8alt/miR-8 Δ2
miPEP-8alt/ ΔmiR-8
F
**
wing size (AU)
miPEP-8
-
+
miPEP-8alt /
miR-8GAL4 ΔmiR-8


A
mir-8; mir-8scr
mir-8; mir-8sp
wt
miPEP-8; mir-8scr
miPEP-8; mir-8sp
wt
B
wing size (AU)
40
30
20
*
ns
ctrl
miR-8; miR-8scr
miR-8; miR-8sp
miPEP-8; miR-8scr
miPEP-8; miR-8sp


A
miR-8 induced
miR-8 repressed
-1,41
+1,41
miR-8
ctrl
miPEP-8
miPEP-8 repressed
miPEP-8 induced
B
miPEP-8
miR-8
1489
476
674
logFC>0,2
p<0,05
C
miPEP-8 specific
fold change
16
8
4
2
1
0,5
0,25
0,125
0,062
0,031
0,015
Control
OE miPEP-8
OE miR-8
cyt-c1L
CG33461
CG4393
Spn77Ba
D
miR-8 specific
fold change
8
4
2
1
0,5
Control
OE miPEP-8
OE miR-8
reg-2
CG2444
socs36E
Dgp-1
E
co-regulated
fold change
8
4
2
1
0,5
Control
OE miPEP-8
OE miR-8
CG34232
CG15695
spitz
jar