Uridylation of RNA Hairpins by Tailor Confines the Emergence of MicroRNAs in *Drosophila*

**Highlights**
- Tailor is a small RNA uridylyltransferase in *Drosophila*
- Tailor uridylates pre-miRNAs and regulates miRNA maturation
- Tailor prevents the maturation of non-canonical miRNAs, i.e., mirtrons
- Tailor may act as a barrier for the de novo creation of miRNAs

**Accession Numbers**
- GSE66213
- GSE67646

**In Brief**
Reimão-Pinto et al. report a small RNA-specific terminal uridylyltransferase in *Drosophila*, which acts on precursor miRNAs to regulate miRNA maturation and may serve as a barrier for the de novo creation of miRNAs in flies.

**Correspondence**
stefan.ameres@imba.oeaw.ac.at
Uridylation of RNA Hairpins by Tailor Confines the Emergence of MicroRNAs in Drosophila

Madalena M. Reimão-Pinto, 1 Valentina Ignatova, 1, 2 Thomas R. Burkard, 1 Jui-Hung Hung, 2 Raphael A. Manzenreither, 1 Ivica Sowemimo, 1 Veronika A. Herzog, 1 Brian Reichholf, 1 Sara Farina-Lopez, 1 and Stefan L. Ameres 1, *

1 Institute of Molecular Biotechnology of the Austrian Academy of Sciences (IMBA), 1030 Vienna, Austria
2 Institute of Bioinformatics and Systems Biology, National Chiao Tung University, Hsin-Chu 300, Taiwan
* Correspondence: stefan.ameres@imba.oeaw.ac.at
http://dx.doi.org/10.1016/j.molcel.2015.05.033

SUMMARY

Uridylation of RNA species represents an emerging theme in post-transcriptional gene regulation. In the microRNA pathway, such modifications regulate small RNA biogenesis and stability in plants, worms, and mammals. Here, we report Tailor, an uridylyltransferase that is required for the majority of 3’ end modifications of microRNAs in Drosophila and predominantly targets precursor hairpins. Uridylation modulates the characteristic two-nucleotide 3’ overhang of microRNA hairpins, which regulates processing by Dicer-1 and destabilizes RNA hairpins. Tailor preferentially uridylates mirtron hairpins, thereby impeding the production of non-canonical microRNAs. Mirtron selectivity is explained by primary sequence specificity of Tailor, selecting substrates ending with a 3’ guanosine. In contrast to mirtrons, conserved Drosophila precursor microRNAs are significantly depleted in 3’ guanosine, thereby escaping regulatory uridylation. Our data support the hypothesis that evolutionary adaptation to Tailor-directed uridylation shapes the nucleotide composition of precursor microRNA 3’ ends. Hence, hairpin uridylation may serve as a barrier for the de novo creation of microRNAs in Drosophila.

INTRODUCTION

MicroRNAs (miRNAs) mediate post-transcriptional gene silencing in plants and animals and control organismal development, physiology, and disease (Bartel, 2009). Most miRNAs derive from hairpin-containing transcripts (primary miRNAs; pri-miRNAs) that are sequentially processed by RNase III enzymes into mature small RNAs (Kim et al., 2003). First, Drosha liberates an ~60 nt hairpin (precursor miRNA; pre-miRNA) from the pri-miRNA transcript in the nucleus (Lee et al., 2004; Gregory et al., 2004; Denli et al., 2004). Upon Exportin-5-directed export of pre-miRNAs from the nucleus to the cytoplasm, Dicer (Dcr-1 in flies) converts pre-miRNAs into mature ~22-nt small RNA duplexes that are subsequently loaded into an Argonate (Ago1 in flies) (Lee et al., 2004; Okamura et al., 2004), where they guide post-transcriptional silencing of complementary mRNA (Huntzinger and Izaurralde, 2011).

Alternative pathways bypass Drosha or Dicer processing for the production of non-canonical miRNAs (Yang and Lai, 2011). The most prevalent example is the splicing of short introns followed by lariat debranching, which produces Drosha-independent pre-miRNA-like hairpins, referred to as mirtrons (Okamura et al., 2007; Ruby et al., 2007). Mirtron hairpins are exported by Exportin-5 and serve as substrates for Dicer, generating mature miRNAs that are functionally indistinguishable from products of canonical miRNA biogenesis (Okamura et al., 2007; Ruby et al., 2007; Babiarz et al., 2008; Berezikov et al., 2010). While numerous mirtrons have been annotated, they rarely accumulate to levels that compare to canonical miRNAs (Chung et al., 2011), and mirtrons more frequently appear and disappear in evolution compared to canonical miRNAs (Berezikov et al., 2010). The spurious contribution of mirtrons to small RNA profiles is apparently at odds with the conservation of the mirtron pathway in flies, worms, and mammals (Chung et al., 2011; Ladewig et al., 2012; Westholm and Lai, 2011). This raises the question of whether mechanisms exist that dampen the contribution of non-canonical pathways to the pool of small RNAs.

Perhaps because of its enormous regulatory potential, miRNA-mediated gene regulation is tightly controlled (Ha and Kim, 2014). At the post-transcriptional level, the addition of uridine(s) to the 3’ end of miRNAs and their precursors recently emerged as a hallmark for the regulation of miRNA biogenesis and turnover (Kim et al., 2010; Ameres and Zamore, 2013). Small RNA-modifying terminal uridylyltransferases (TUTases) have been reported in plants and animals (Scott and Norbury, 2013). In Arabidopsis, the methyltransferase HEN1 protects miRNAs and siRNAs from HESO 1-directed uridylation, a modification that destabilizes miRNAs (Li et al., 2005; Ren et al., 2012; Zhao et al., 2012). In animals, more versatile functions of uridylation have been described: in worms and mammals, the RNA-binding protein Lin28 recruits the TUTase ZCCHC11/PUP2 to pre-let-7, resulting in pre-miRNA oligouridylation (Heo et al., 2009; Hagan et al., 2009; Lehrbach et al., 2009). Pre-let-7 uridylation prevents dicing and triggers degradation by the exonuclease Dis3L2 (Chang et al., 2013). In the absence of Lin28, the non-processive addition of a single uridine to a selected class of pre-miRNA by
the TUTases ZCCHC11 and ZCCHC6 enhances dicing in mammalian cells because it restores the two-nucleotide 3′ overhangs of pre-miRNAs (Heo et al., 2012). A similar mechanism may also contribute to the identification of defective pre-miRNAs that lack intact 3′ overhangs and trigger their destruction via the exosome (Liu et al., 2014). Finally, destabilization of mature miRNAs upon binding to highly complementary targets is associated with small RNA uridylation (and adenylation) in flies and mammals (Ameres et al., 2010; Xie et al., 2012). While diverse mechanisms employ TUTases in the regulation of small RNAs, their overall impact on shaping the miRNA repertoire remains unclear. In flies, no miRNA-specific TUTase has been identified. Nevertheless, meta-analyses of small RNA libraries from flies indicate that—like in mammals—miRNAs are frequently subjected to modifications. The predominant class of such miRNAs consists of mirtrons, which are frequently modified (Figure 1D) and predominantly carry uridine modifications (Figure S1D).

**The Cytoplasmic TNTase CG1091 Is Required for miRNA Uridylation and Normal Fertility in Flies**

The template-independent incorporation of ribonucleoside monophosphate to the 3′ hydroxyl end of RNAs is catalyzed by RNA-specific TNTases. Members of this enzyme family exhibit a characteristic domain architecture consisting of a TNTase domain (NT_PAP_TUTase), often paired with a PAP/25A-associated domain (PAP_assoc) (Norbury, 2013). Based on domain homology search, we identified seven putative TNTases in the *Drosophila* genome, five of which were expressed in S2 cells (Figure 2A). We depleted candidate enzymes by RNAi in S2 cells followed by detection of the abundantly expressed miR-184 by high-resolution northern hybridization. Although it is not among the most frequently modified miRNAs (~3%; Figure S1A), miR-184 produced higher-molecular-weight signals in northern hybridization experiments that reflected post-transcriptional modifications in high-throughput sequencing datasets (Figure 2B). Among the tested candidate TNTases, only depletion of CG1091 affected miR-184 tailing (Figure 2C), suggesting that CG1091 is required for the post-transcriptional modification of miR-184.

To characterize the global effect of CG1091 on miRNA modifications, we performed high-throughput sequencing of 18- to 30-nt small RNAs from S2 cells depleted of CG1091 by RNAi, as well as adult male flies, carrying two independently generated frame-shift deletions in the third exon of CG1091 (CG1091-c4-1 and CG1091-c4-6, Figure S2A), introduced by CRISPR/Cas9 (Figure 2D). Depletion of CG1091 was confirmed by western blot analysis (Figure 2E). Upon depletion of CG1091, the frequency at which miRNAs contained non-genome-matching nucleotide additions was significantly reduced in S2 cells when compared to control RNAi (p < 10⁻⁴; Wilcoxon matched pairs signed rank test) or untreated cells (p < 3 × 10⁻⁴), and in whole male flies homozygous for CG1091-c4-1/6 when compared to heterozygous siblings (p < 10⁻⁴) or w¹¹¹⁸ flies (p < 10⁻⁴) (Figures 2F and 2G and Figure S2B). In S2 cells, we detected a statistically significant reduction in the post-transcriptional modification of 48 out of 79 miRNAs (p < 0.05, FDR < 0.1). For 19 of these, tailing was reduced by more than 2-fold (red dots, Figure 2F). In whole male flies, 39 out of 107 miRNAs were consistently reduced in tailing by more than 2-fold in the two independent CG1091-c4 allele-carrying flies (Figures 2F and S2B).

Analysis of the nucleotide identity in miRNA 3′ tails revealed that CG1091 is required for the majority of miRNA uridylation (Figure 2H); upon depletion of CG1091 in S2 cells, uridine-containing extensions decreased by 2.6-fold when compared to the control, while all other nucleotide additions stayed unchanged. Similarly, uridylation decreased by 1.9-fold in CG1091-c4-1/6 heterozygous, and by 4.2-fold in homozygous, flies when compared to w¹¹¹⁸ control flies. Furthermore, depletion of CG1091 affected adenylation, although to a lesser extent (1.2- and 1.4-fold decrease in CG1091-c4-1/6 and CG1091-c4-6, respectively), while the addition of cytidine or guanosine was not affected. Note that under CG1091 depletion conditions we still detected residual miRNA uridylation above background.
indicating that CG1091 may not be the sole small RNA-uridylation enzyme. CG1091 acts preferentially on mirtrons, which were significantly more depleted in post-transcriptional modifications compared to canonical miRNAs, both in S2 cells (p < 10^{-4}) and in flies (p < 0.03) (Figure 2I).

While CG1091 is ubiquitously expressed, mRNA levels are highest in the ovary and testis (Graveley et al., 2011), prompting us to investigate any loss-of-function effects on fertility. We observed a significant defect in male (p < 10^{-4}; Figure 2J) and female (p < 0.01; Figure 2K) fertility in CG1091^{c4-1/c4-6} mutant flies when compared to control flies (w^{118B}). In females, fertility defects were not due to impaired fecundity, since mutant and control flies laid similar numbers of eggs (Figure 2K, top graph). We confirmed the observed phenotypes in trans-heterozygous flies, carrying one CRISPR allele (CG1091^{c4-6}) over a publicly available piggy-bac transposon insertion in the coding sequence of tailor (CG1091^{f05717}; Figure 2D) (Thibault et al., 2004), excluding any secondary perturbations linked to the use of CRISPR/Cas9.

Finally, expression of Myc-tagged CG1091 in S2 cells revealed predominant—if not exclusive—localization of CG1091
Figure 2. The Cytoplasmic TNTase CG1091/Tailor Is Required for miRNA Uridylation and Normal Fertility in Flies

(A) Domain organization of known and putative TNTases in flies based on InterPro database. The characteristic domain structure consists of a nucleotidyl-transferase domain (red box) and a PAP/25A-associated domain (blue). Expression in S2 cells was based on modENCODE mRNA sequencing (Cherbas et al., 2011).

(B) Post-transcriptional modification of miR-184-3p in S2 cells is detectable by high-resolution northern hybridization. Higher-molecular-weight bands of miR-184 (> 23 nt) correspond to prefix-matching reads (red) as evidenced by small RNA sequencing.

(legend continued on next page)
dispersed throughout the cytoplasm, without detectable signal in the nucleus (Figure 2L and Figures S2D and S2E).

In conclusion, the cytoplasmic, putative TNTase CG1091 is required for normal fertility in flies and is responsible for the majority of uridylation events associated with miRNAs in S2 cells and in flies, with a particular specificity for mirtron uridylation. Because it mediates the selective uridylation of a subset of miRNAs, we refer to CG1091 as Tailor.

Tailor Regulates miRNA Abundance and Function and Prevents the Accumulation of Mirtrons

Post-transcriptional uridylation can trigger RNA destabilization (Scott and Norbury, 2013). We therefore tested if Tailor depletion impacts miRNA abundance. In S2 cells depleted of Tailor, more than 40% of all miRNAs (32 out of 79 miRNAs) exhibited a statistically significant change in abundance (p < 0.05, FDR < 0.1) when compared to control RNAi conditions (Figure 3A). The majority of these (28 out of 32) increased in abundance, suggesting Tailor generally has a negative impact on miRNA accumulation. Mirtrons exhibited a significantly higher accumulation upon Tailor depletion compared to canonical miRNAs (p < 10^{-4}; Figure 3B), suggesting that Tailor prevents mirtron accumulation.

miRNAs that significantly changed in abundance upon depletion of Tailor were more frequently tailed under unperturbed conditions (p < 1 × 10^{-3}; Figure 3C and Figure S3A), and tailing of these miRNAs was significantly more dependent on Tailor than 40% of all miRNAs (32 out of 79 miRNAs) exhibited a statistically significant change in abundance (p < 0.05, FDR < 0.1) when compared to control RNAi conditions (Figure 3A). The majority of these (28 out of 32) increased in abundance, suggesting Tailor generally has a negative impact on miRNA accumulation. Mirtrons exhibited a significantly higher accumulation upon Tailor depletion compared to canonical miRNAs (p < 10^{-4}; Figure 3B), suggesting that Tailor prevents mirtron accumulation.

Figure 3. Tailor-Dependent Uridylation Impacts Mature miRNA Levels and Prevents Mirtron Accumulation

(A) Greater than 40% of all miRNAs are significantly changed in abundance upon Tailor depletion in S2 cells (p < 0.05, Student’s t test; FDR < 0.1; Benjamini and Hochberg, 1995). Data represented as mean ± SD.

(B) Mirtrons significantly increase in abundance upon depletion of Tailor. p value (Mann-Whitney test) is indicated.

(C) Changes in miRNA abundances correlate with post-transcriptional modifications. The fraction at which each miRNA is modified is plotted. S2 cell data is indicated for miRNAs that do, or do not, change significantly in abundance upon depletion of Tailor. p values (Mann-Whitney test) are indicated.

(D) Changes in miRNA abundance correlate with Tailor-directed tailing. Changes in miRNA tailing between Tailor-depleted and control-dsRNA-treated S2 cells is shown for miRNAs that do, or do not, change significantly in abundance upon depletion of Tailor. p value (Mann-Whitney test) is indicated. See also Figure S3.
(p < 0.05; Figure 3D) when compared to miRNAs that remain unchanged in abundance. We concluded that Tailor-mediated uridylation regulates the abundance of miRNAs and prevents the accumulation of mirtrons.

Notably, four miRNAs decreased in abundance upon depletion of Tailor (Figure 3A), including miR-184. We confirmed this effect by northern hybridization, where miR-184 showed a significant, 1.5-fold decrease in abundance upon depletion of Tailor in S2 cells (Figures S3B and S3C). In selected cases Tailor therefore promoted the accumulation of miRNAs.

Finally, high-throughput mRNA sequencing revealed that the observed changes in miRNA levels upon depletion of Tailor significantly impact miRNA-mediating gene regulation (Figures S3D–S3I).

Taken together, Tailor regulates the abundance and function of mature miRNAs and prevents the accumulation of mirtrons.

### Tailor-Directed Modification of Pre-miRNA 3’ Ends Regulates miRNA Maturation

To understand how Tailor regulates miRNA abundance we performed gain-of-function experiments by generating clonal S2 cells stably expressing Tailor (CG1091OE) and monitored miRNA modification and abundance by northern hybridization experiments. Among six tested miRNAs, five significantly changed in abundance upon Tailor expression when compared to control S2 cells (p < 0.05; Figures 4A and 4B): two increased (1.4-fold for both, miR-8-3p and bantam-3p), and three decreased in abundance (2.3-fold for miR-184-3p, 1.2-fold for miR-33-3p, and 3.3-fold for miR-252-5p). But only a subset of the deregulated miRNAs—the ones derived from the 3p arm of the respective pre-miRNA—showed tailing signatures, manifested as higher-molecular-weight bands in northern hybridization experiments (Figure 4A). This indicated that Tailor modifies pre-miRNAs. We tested this hypothesis by analyzing pre-miRNAs in northern hybridization experiments: upon depletion of Dcr-1 by RNAi (enriching pre-miRNAs to detectable amounts) we observed an accumulation of higher-molecular-weight isoforms of both pre-miR-184 and pre-bantam in CG1091OE cells compared to control S2 cells (Figure 4C). While both pre-miRNAs showed tailing signals, this had differential consequences on the accumulation of the respective mature miRNA species—an increase of bantam-3p and a decrease of miR-184-3p (Figure 4B).

Accurate 2-nt 3’ overhangs are hallmarks for pre-miRNA processing by Dicer proteins: changes in pre-miRNA 3’ overhangs impact miRNA processing by Dicer (Tsutsumi et al., 2011). To test if this may explain the observed effects, we cloned and Sanger sequenced pre-miR-184 and pre-bantam to characterize Tailor-directed changes in pre-miRNA 3’ ends (Newman et al., 2011). In control S2 cells, 87% of all pre-miR-184 reads mapped to the predicted 3’ end, producing a 2-nt 3’ overhang, while the accuracy decreased in CG1091OE cells to 66%, with 14% containing elongated 3-nt 3’ overhangs (Figure 4D). To test if changes in 3’ end accuracy impact processing, we performed in vitro diying assays (Tsutsumi et al., 2011): affinity-purified Dcr-1 processed pre-miR-184 with 3-nt 3’ overhang at significantly lower rates when compared to canonical 2-nt 3’ overhangs (p < 10⁻⁴; Figures 4E and S4A–S4C), explaining the observed decrease in miR-184-3p in CG1091OE cells.

For pre-bantam we observed the opposite: while only 25% of cloned pre-bantam mapped to the predicted 2-nt 3’ overhang in S2 cells (> 40% of all reads supported a blunt 3’ end structure), the accuracy increased in CG1091OE cells, where 53% of all reads mapped to the predicted 3’ end (Figure 4F). In vitro, Dcr-1-directed processing of pre-bantam was 5-fold more efficient in the presence of a 2-nt 3’ overhang when compared to a blunt end (p < 10⁻³; Figures 4G and S4D–S4F), explaining the increase in mature bantam-3p in CG1091OE cells (Figure 4B).

We concluded that Tailor-mediated uridylation of pre-miRNAs changes the accuracy of 3’ ends, which impacts the rates at which mature miRNAs are produced by Dcr-1, ultimately regulating the abundance of mature miRNAs.

### Tailor-Dependent Uridylation Destabilizes Pre-miRNAs and Prevents Hairpin Dicing

The conclusion that Tailor uridylates pre-miRNAs, as proposed by gain-of-function experiments, warranted further experimental support under conditions where Tailor is expressed at physiological levels. To this end, we cloned 40- to 100-nt-long RNAs from S2 cells using an unbiased cloning approach. As expected, the majority (~78 Mio reads in these libraries mapped to abundant, non-coding RNA species that overlapped in size with pre-miRNAs, i.e., tRNA, snoRNA, and tRNA (Figure 5A). While only less than 0.1% of all reads mapped to pre-miRNAs, we recovered sufficient depth (> 70,000 reads) for an unbiased analysis of pre-miRNAs: using a cutoff of > 7 reads, we detected 89 pre-miRNAs, representing the precursors of 73 (50 miR and 23 miR*) out of 79 (55 miR and 24 miR*) abundant pre-miRNAs. Among the remaining six, three mapped to duplicated miRNA loci and could not be uniquely assigned to one locus in small RNA sequencing datasets (miR-13-b1, miR-276b, and miR-2c); two derived from an unusual, 97-nt-long precursor (miR-998-5p and miR-998-3p), which may have been excluded by size selection; and one was present in small RNA libraries only at low levels (< 250 ppm in small RNA libraries; miR-263a). This analysis confirms that cloning of 40- to 100-nt RNAs at sufficient depth recovers the cellular pre-miRNA pool at near saturation. This conclusion was further supported by the fact that we recovered 38 pre-miRNAs (in average > 50 reads), which were not detected in mature small RNA libraries (Table S1).

For further analysis we focused on the 58 pre-miRNAs that were sequenced at sufficient depth to reproducibly recover species with non-genome-matching 3’ additions (Figures 5B and SSA–SSD). Among these, mirtrons exhibited a significantly higher fraction of non-genome-matching additions compared to canonical pre-miRNAs (p < 10⁻³; Figure 5C). Like in mature small RNAs, uridylation (42%) and adenylation (35%) dominated post-transcriptional modifications in pre-miRNAs over cytosine (17%) and guanosine additions (5%) (Figure 5D). Uridylation was likewise enriched in heavily modified pre-miRNAs (Figure 5D) and dominated modifications associated with mirtrons (62%; Figure S5E). Finally, we detected a statistically significant correlation of pre-miRNA tailing and mature miRNA-3p (Pearson’s correlation coefficient r = 0.49; p < 3 × 10⁻³), but not -5p, modifications (r = 0.08, n.s.), indicating that a significant fraction of post-transcriptional modifications detected in mature 3p-miRNAs originated from pre-miRNA tailing (Figure S5F).
To test if Tailor mediates post-transcriptional uridylation of pre-miRNAs, we sequenced 40- to 100-nt RNAs from S2 cells depleted of Tailor by RNAi. Tailor depletion caused a significant decrease in tailed pre-miRNAs when compared to untreated S2 cells in northern hybridization experiments. MicroRNAs originating from the 3p or 5p arm of pre-miRNAs are indicated. 2S rRNA represents loading control.

Among the non-genome-matching additions to pre-miRNAs, solely uridylation decreased by more than 3-fold upon Tailor depletion when compared to S2 cells treated with control dsRNA (Figure 5G). In comparison to canonical pre-miRNAs, mirtrons were significantly more depleted in post-transcriptional modifications (p < 0.03; Figure 5H).

Finally, we tested the consequences of loss of uridylation on pre-miRNA abundance. We found a significant increase in pre-miRNA levels upon depletion of Tailor for the top 15 (p < 0.02), top 10 (p < 0.02), and top 5 (p < 0.009) Tailor-modified pre-miRNAs (Figure 5J). To determine the impact of uridylation on pre-miRNA processing, we determined in vitro dicing rates of pre-miR-1003, a mirtron that was highly modified in pre-miRNA and mature miRNA libraries (Figures 1B, 2F, and 5B) and was significantly upregulated in its mature form upon depletion of Tailor (Figure 3A): addition of a single uridine to pre-miR-1003 decreased dicing efficiency by more than 2-fold (p < 10^{-4}), and di- or tri-uridylation reduced processing to almost undetectable levels (p < 10^{-8} and p < 10^{-11}, respectively; Figures 5J and 5K).

Furthermore, pre-miRNA tailing correlated significantly with the observed change in mature miRNA abundance upon depletion.
Figure 5. Pre-miRNAs, Particularly Mirtron Hairpins, Are Physiological Substrates for Tailor-Directed Destabilization through Tailing
(A) Mapping results of high-throughput sequencing of 40- to 100-nt RNAs from S2 cells.
(B) Abundance of genome-matching and prefix-matching pre-miRNAs in S2 cells. Data presented as mean of three independent biological replicates ± SD. Pre-miRNAs with highest fraction of prefix-matching reads (> 5%) are indicated in black.
(C) Mirtrons are significantly more tailed compared to canonical pre-miRNAs. The fraction at which pre-miRNAs are tailed is shown for all miRNAs and the indicated classes of pre-miRNAs. The number of pre-miRNAs in each group is indicated. p value (Mann-Whitney test) is indicated.
(D) Pre-miRNA tails consist mostly of uridine and adenine. The nucleotide composition of non-genome-matching additions was determined for each pre-miRNA and averaged across all pre-miRNAs.
(E) Depletion of Tailor by RNAi affects tailing of pre-miRNAs in S2 cells. The fraction at which each pre-miRNA is modified is plotted. miRNAs that show >2-fold reduction in tailing are indicated in black.
(F) Depletion of Tailor by RNAi in S2 cells impacts pre-miRNA tailing. p values (Wilcoxon matched-pairs signed rank test) are indicated.
(G) Depletion of Tailor by RNAi impedes addition of uridine, but not any other nucleotides, to pre-miRNAs. Abundance and composition of nucleotide additions was determined for each pre-miRNA, normalized to control treatment, and averaged across all pre-miRNAs.
(H) Mirtron-hairpin uridylation requires Tailor. Change in tailing upon Tailor depletion in S2 cells by RNAi was determined for classified pre-miRNAs and normalized to control treatment. p value is indicated (Mann-Whitney test).
(I) Tailor-directed uridylation destabilizes pre-miRNAs. Upon Tailor depletion, the levels of pre-miRNAs that are tailed by Tailor under unperturbed conditions increase significantly, p values are indicated (Mann-Whitney test).
(J) Uridylation inhibits pre-miRNA dicing. In vitro dicing assays employed SBP-tagged affinity-purified Dcr-1 and the indicated 5’ radiolabelled synthetic pre-miRNA.
(K) Quantification of data shown in (J). Data presented as mean ± SD. p values are indicated (Student’s t test). See also Figure S5 and Table S1.
of Tailor (Pearson’s r = 0.51, p < 10^{-4}; Figures S5G and S5H), with a correlation coefficient that was higher than what we observed for the correlation of mature miRNA tailing with change in mature miRNA abundance (Pearson’s r = 0.23, p < 0.05; Figure S3A). Together, these data establish pre-miRNAs—in particular mirtron-hairpins—as major targets for Tailor-dependent uridylation, a modification that destabilizes pre-miRNAs and prevents efficient hairpin dicing.

Tailor Is a Bona Fide, RNA-Specific TUTase with Unique Targeting Properties

For biochemical characterization we immunopurified FLAG-tagged Tailor upon expression in S2 cells and confirmed enzyme purity by western blotting (Figure S6B). When incubated with a 22-nt, 5’ radiolabelled RNA in the presence of ribonucleotide triphosphates (rNTPs), Tailor extended the length of the substrate RNA by up to ~10 nucleotides within 5 min. This activity was abolished when we replaced one of three highly conserved aspartates in the conserved catalytic pol β superfamily motif with alanine (Figures S6A and S6C). In the presence of individual rNTPs, Tailor catalyzed the efficient incorporation of uridine and—to a lesser extent—cytidine and adenosine, but rarely guanosine (Figures 6A, S6D, and S6E). Tailor-directed terminal nucleotidyl transfer is consistent with the previously proposed two-divalent-cation-dependent catalytic mechanism of TNTases (Martin et al., 2008): the divalent cation-chelating agent EDTA inhibited, while addition of 5mM Mg^{2+} rescued, tailing activity (Figure 6B). Together, these experiments classify Tailor as a bona fide RNA-specific TUTase.

To dissect the enzymatic properties of Tailor in more detail, we developed a high-throughput biochemical assay for the characterization of RNA-specific TNTases. To this end we employed as a substrate a 37-nt RNA containing four random 3’ nucleotides, GEDC, and subjected it to in vitro tailing reactions using immunopurified Tailor for 2 or 5 min, followed by 3’ adaptor ligation and high-throughput sequencing.

![Figure 6. Tailor Is a Bona Fide TUTase with Unique Targeting Properties](image)

(A) Immunopurified Tailor exhibits TNTase activity. FLAG-tagged Tailor was expressed in S2 cells, immunopurified, and incubated with a 22-nt, 5’ radiolabelled RNA in the presence of the indicated ribonucleotide triphosphates. (B) Tailor-directed nucleotide transfer requires Mg^{2+}. Tailoring reactions were performed using a 5’ radiolabeled substrate RNA described in (A) in the presence of rNTPs. Addition of EDTA inhibited the tailing reaction, whereas excess Mg^{2+} rescued the activity.

(C–G) High-throughput biochemical characterization of Tailor-directed RNA tailing. A 37-nt RNA substrate containing four random nucleotides at the 3’ end was subjected to in vitro tailing reactions using immunopurified Tailor for 2 or 5 min, in the presence of rNTPs, followed by 3’ adaptor ligation and high-throughput sequencing. (C) Validation of high-throughput tailing assay. Product of in vitro tailing reactions was resolved by denaturing PAGE (left) or analyzed by high-throughput sequencing (right). Sequencing reveals high selectivity for UTP incorporation despite even concentration of rNTPs in the tailing reaction. (D) Tailor-directed tailing efficiency is not influenced by substrate secondary structures. Substrates were binned according to secondary structure stability (effective free energy; EFE) and analyzed for the fraction at which each substrate was tailed, averaged across two time points (2 and 5 min). Median (white line), inner quartile range (IQR; dark gray area), and 1.5 IQR of lower and upper quartile (light gray area) are shown. No statistically significant difference was detected. (E) Tailor-directed tailing efficiency is influenced by substrate secondary structures. Substrates were binned according to secondary structure stability (effective free energy; EFE) and analyzed for the fraction at which each substrate was tailed, averaged across two time points (2 and 5 min). Median (white line), inner quartile range (IQR; dark gray area), and 1.5 IQR of lower and upper quartile (light gray area) are shown. No statistically significant difference was detected. (F) Tailor directed tailing efficiency is influenced by substrate secondary structures. Substrates were binned according to secondary structure stability (effective free energy; EFE) and analyzed for the fraction at which each substrate was tailed, averaged across two time points (2 and 5 min). Median (white line), inner quartile range (IQR; dark gray area), and 1.5 IQR of lower and upper quartile (light gray area) are shown. No statistically significant difference was detected. (G) Tailor directed tailing efficiency is influenced by substrate secondary structures. Substrates were binned according to secondary structure stability (effective free energy; EFE) and analyzed for the fraction at which each substrate was tailed, averaged across two time points (2 and 5 min). Median (white line), inner quartile range (IQR; dark gray area), and 1.5 IQR of lower and upper quartile (light gray area) are shown. No statistically significant difference was detected. (C–G) See also Figure S6.
incubated it with immunopurified Tailor for 2 min or 5 min in the presence of equal concentrations of rNTPs, and subjected the product to high-throughput sequencing (Figure S6F). Analysis of the resulting libraries (~7 Mio reads per library; Figure S6H) revealed that the tailing product exhibited a length distribution similar to the one observed in denaturing PAGE (Figure 6C). This confirmed a quantitative recapitulation of the tailing reaction by sequencing, and that Tailor catalyzes the incorporation of nucleotides to the 3’ end of an RNA substrate. Notably, Tailor exhibited nearly exclusive selectivity for uridine incorporation: less than 0.02% and 0.25% of all incorporated nucleotides consisted of a nucleotide other than uridine at 2 and 5 min, respectively. The ability of Tailor to efficiently catalyze the incorporation of cytosine or adenosine in the sole presence of CTP or ATP (see Figures S6D and S6E) is therefore apparently without relevance in the presence of UTP.

The randomized 3’ end of the substrate RNA enabled us to simultaneously analyze 256 different substrates, all of which were recovered, and their abundance stayed constant over the course of the assay (Figure S6G). We first determined the secondary structure stability of each substrate, binned them into effective free energy groups (EFE groups) from highly structured (group 1) to single-stranded (group 6), and tested each group for a significant over- or underrepresentation in the mean fraction at which it was tailed (Figure 6D) (note that the substrate was designed to avoid secondary structures not involving the random 3’ nucleotides; see Figures S6I and S6J). This showed that Tailor acted on RNA substrates, irrespective of whether the 3’ end is embedded in secondary structures. In contrast, the number of nucleotides that were added was strongly affected by secondary structures: structured RNAs (EFE groups 1–3) contained significantly shorter tails ($p < 10^{-4}$; Mann-Whitney test), whereas single-stranded substrates were tailed significantly longer (EFE group 6; $p < 10^{-2}$) (Figure 6E). We concluded that substrate secondary structures impact the processivity of Tailor-directed uridylation.

Finally, we analyzed the impact of primary sequence on tailing efficiency: we found that highly tailed species were enriched in guanosine (Figure 6F). Analysis of each individual position of the randomized four nucleotides revealed that only the 3’ terminal position significantly discriminates tailing efficiency (Figure 6G): while 3’ adenosine or 3’ cytidine substrates were significantly less frequently tailed ($p < 10^{-4}$), a 3’ terminal uridine significantly enhanced the tailing reaction ($p < 10^{-4}$), and 3’ guanosine-containing substrates were tailed most efficiently ($p < 10^{-2}$) at both time points (Figures 6G and S6L). While additional 3’-terminal dinucleotide analysis indicated a trend toward more frequent tailing of substrates ending in AG-3’ (as well as UG-3’ and AU-3’), this further increase in tailing was not statistically significant when compared to the effect of the 3’-terminal nucleotide alone (Figures S6M and S6N).

In summary, high-throughput biochemical characterization of Tailor uncovered substrate structure-dependent processivity and primary sequence specificity, with a preference for RNAs ending in 3’ guanosine (and 3’ uridine). Our data revealed unique targeting properties of a previously uncharacterized RNA-specific TUTase.

Tailor Uridylates Pre-miRNAs Ending in 3’G

To assess primary sequence specificity in Tailor-directed modification of miRNAs, we analyzed the position preceding the non-genome-matching tail in (small) RNA libraries. In both S2 cells and in flies we detected an enrichment of 3’G in Tailor-modified miRNAs (Figure 7A): notably, this effect was most pronounced in pre-miRNAs, the major Tailor substrates.

Selective tailing of substrates ending in 3’G also explained the high degree of Tailor-dependent uridylation of mirtron hairpins, whose 3’ ends are defined by the splice acceptor consensus sequence, ending in 3’AG. We confirmed this by in vitro tailing assays using a synthetic precursor of the mirtron miR-1003. While 65% ± 2% of pre-miR-1003 with wild-type sequence (ending in 3’G) was tailed within 15 min, changing the terminal nucleotide to 3’U caused a significant decrease in tailing efficiency (35% ± 2% after 15 min; $p < 10^{-4}$), and mutations to 3’A or C nearly abolished mirtron tailing (9.5% ± 1.8% and 9.3% ± 1.7% tailed, respectively, after 15 min; $p < 10^{-3}$) (Figures 7B and 7C). Single-stranded miR-1003 exhibited a similar 3’ nucleotide preference, albeit with an overall reduced tailing efficiency: while Tailor modified 65% of wild-type pre-miR-1003, only 31% of mature single-stranded miR-1003 was modified within 15 min (Figures S7A and S7B). The intrinsic substrate preference of Tailor for pre-miRNA substrates ending in 3’G therefore explains why mirtron hairpins are preferentially uridylated in flies.

Conserved Pre-miRNAs, but Not Newly Emerging Hairpins, Are Depleted in 3’G

While more than 30 mirtrons have been identified in flies, only few produce abundant, mature miRNAs, perhaps because a mechanism exists that prevents the efficient conversion of such hairpins into small RNAs (Figure 3) (Chung et al., 2011; Westholm and Lai, 2011). We speculated that uridylation of hairpins, such as mirtrons, may represent a regulatory threshold for the accumulation of de novo-generated Dicer substrates. If this were true, one would expect an adaptation to regulatory uridylation in conserved miRNAs. To test this, we investigated the 3’ terminal genome-matching nucleotide of Drosophila pre-miRNAs, expecting that the hallmark of Tailor substrates—a 3’ terminal G—would be underrepresented in conserved pre-miRNAs. We investigated the 3’ end of 143 Drosophila melanogaster miRNAs annotated with high confidence in miBase, for which sufficient experimental evidence exists to indubitably annotate the 3’ end of the underlying pre-miRNA (Kozomara and Griffiths-Jones, 2014). When compared to the nucleotide distribution in the Drosophila genome, we detected a statistically significant underrepresentation of guanosine at the 3’ end of pre-miRNAs (21% in the genome compared to 15% in pre-miRNA 3’ ends; $p < 0.02$: binomial test) (Figures 7D and 7E). This depletion was more pronounced in the 108 pre-miRNAs that are conserved among all sequenced Drosophila species, with only 12% of all pre-miRNAs ending in 3’G ($p < 0.007$). In contrast, among the 35 non-conserved pre-miRNAs, 26% contained a 3’ terminal guanosine, a frequency that was not significantly different from the occurrence of guanosine in the Drosophila genome. Therefore, our data support the hypothesis that evolutionary adaptation to...
pre-miRNA uridylation shapes the nucleotide composition of pre-miRNA 3' ends and may serve as a barrier for the efficient de novo creation of hairpins that can be converted into mature miRNAs.
strategy. Notably, 14 out of 18 non-conserved hairpins were only detected as pre-miRNAs, but not in mature small RNA sequencing (Figures 7G and S7D), and the few non-conserved miRNAs that we detected in both pre-miRNA and mature miRNA cloning were significantly more abundant upon depletion of Tailor compared to conserved miRNAs (p < 0.03; Figure 7H).

Together our data support the hypothesis that Tailor targets non-conserved miRNA hairpins and prevents their efficient maturation.

**DISCUSSION**

Post-transcriptional modification by TUTases lies at the core of mechanisms that regulate small RNA biogenesis and stability in plants, worms, and mammals (Ameres and Zamore, 2013). Our study reveals the first miRNA-modifying TUTase in flies—Tailor. Tailor is responsible for the majority of uridylation signature observed in miRNAs of S2 cells and flies, predominantly acts on precursor hairpins, and is required for the normal accumulation of mature miRNAs. Uridylation generally has a dampening role for miRNA maturation, since Tailor depletion increased the abundance of > 35% of all miRNAs. Much of this effect may be explained by the fact that uridylation inhibits hairpin processing by Dcr-1 and targets pre-miRNAs for decay. Our observations are therefore consistent with a destabilizing role of uridylation in general RNA metabolism (Scott and Norbury, 2013) and resemble the Lin28-TUT4-Dis3L2 pathway in mammalian ES cells, where the 3′-to-5′ exoribonuclease Dis3L2 degrades pre-let-7 following TUT4-directed uridylation (Chang et al., 2013; Faehnle et al., 2014). Both Lin28 (CG17334) and Dis3L2 (CG16940) are conserved in flies and interact with Tailor based on high-throughput mapping of Drosophila protein interactions (Guruharsha et al., 2011), but Lin28 is not expressed in S2 cells, suggesting alternative targeting strategies for pre-miRNA uridylation in these cells.

Tailor-directed uridylation of pre-miRNAs is not always detrimental to miRNA biogenesis; in fact, our data support a model in which Tailor modulates the two-nucleotide 3′ overhangs in pre-miRNAs, which in turn regulates the efficiency at which Dicer converts these substrates into mature miRNAs. While the limited depth of our pre-miRNA sequencing datasets prevented a global validation of this model, it may explain why a few miRNAs require Tailor for their efficient accumulation. Such a proof-reading concept has been proposed to shape pre-miRNA biogenesis in mammals, where the TUTases ZCCHC11 and ZCCHC6 restore inaccurate pre-miRNA 3′ overhangs to facilitate substrate recognition by Dicer (Heo et al., 2012; Liu et al., 2014). How TUTases would be able to sense “defective” pre-miRNA end structures is unclear, but biochemical characterization of Tailor suggests that structured 3′ ends—found in blunt-ended pre-miRNAs—are tailed less processively. Addition of shorter tails (< 2 nt) to structured ends may contribute to pre-miRNA 3′ end restoration. Notably, the effect of Tailor on pre-miRNA maturation is dosage dependent, as exemplified by miR-184: normal accumulation of miR-184 requires physiological TUTase levels, since both depletion and overexpression of Tailor decreases miR-184 levels. In that respect, differential expression of Tailor may be involved in the fine-tuning of miRNA expression levels.

The most intriguing function of Tailor is its role in preventing the maturation of an entire class of miRNAs, namely mirtrons. Selective mirtron uridylation is a consequence of biochemical targeting features that to our knowledge are unique among TUTases. Both in vitro and when inspecting uridylated miRNA hairpins in S2 cells and in flies Tailor exhibited a strong preference for RNA substrates ending in 3′G. This is a hallmark of mirtron-hairpin sequences, whose 3′ ends are defined by the conserved splice acceptor sequence AG-3′. Note that a function of CG1091/Tailor in the regulation of mirtrons was also reported in a companion paper (Bortolamiol-Becet et al., 2015). Why does Tailor prevent mirtron maturation? Among the growing number of alternative pathways that give rise to bona fide small RNAs (Yang and Lai, 2011), the mirtron pathway generates the most numerous miRNAs (Westholm and Lai, 2011). Mirtrons are most prevalent in flies and worms, where intron length distributions match those of canonical pre-miRNAs (Lim and Burge, 2001; Yandell et al., 2006). Because of their length, introns may provide a rich source for the rapid de novo generation of miRNAs, a notion that is supported by the fact that mirtrons evolve more rapidly compared to canonical miRNAs in flies (Berrezikov et al., 2010). But the fact that mirtrons are less likely to be maintained in the course of evolution may also suggest that the sudden appearance of large quantities of a novel miRNA may frequently be detrimental, rather than beneficial. In this regard, a regulatory threshold for the production of miRNAs from mirtron hairpins may be useful. Notably, mirtron regulation may be conserved in worms and mammals, where frequent uridylation of mirtrons was described (Westholm et al., 2012).

Consistent with a pronounced negative effect of Tailor-directed uridylation on general miRNA biogenesis, we find that miRNA hairpins tend to evade uridylation in Drosophila because their 3′ ends are significantly depleted in 3′G. This effect is even stronger in conserved pre-miRNAs, while non-conserved, evolutionary “young” hairpins are not depleted in 3′G. Therefore, our data support the hypothesis that evolutionary adaptation to pre-miRNA uridylation shapes the nucleotide composition of pre-miRNA 3′ ends. Hence, hairpin uridylation may serve as a barrier for the de novo creation of miRNAs in Drosophila.

**EXPERIMENTAL PROCEDURES**

**General Methods**

Total RNA from flies and S2 cells was purified using the mirVana Kit (Ambion) or TRIZol (Invitrogen). High-resolution northern hybridization was as described (Irin et al., 2011). Stable and clonal cultured S2 cell lines were generated as described (Ameres et al., 2010). A monoclonal antibody against Tailor was generated against full-length recombinant protein expressed in E. coli. For further information see Supplemental Experimental Procedures.

**RNAi in S2 Cells**

S2 cells (2 × 10^6) were soaked with 10 μg dsRNA for 1 hr in ExpressFive SFM medium (Invitrogen) supplemented with 200 mM glutamine (Invitrogen) in the absence of FBS, followed by the addition of medium containing 10% FBS. Soaking was repeated after 4 days, and cells were harvested and processed at day 7.

**RNA Library Construction**

Small RNA libraries were constructed as described (Ameres et al., 2013) with modifications detailed in Supplemental Experimental Procedures.
**In Vitro Tailing Assay**

FLAG-tagged Tailor was immunopurified upon expression in S2 cells. In vitro tailing assays employed 10-nM 5’ radiolabelled substrate RNA and FLAG-Tailor under RNAi assays conditions (Ameres et al., 2010), but omitting the ATP-regeneration system. rNTPs were added to a final concentration of 500 μM each. See Supplemental Experimental Procedures for details.

**Data Analysis**

Gel images were quantified using ImageQuant TL v7.0 (GE Healthcare). Curve fitting was performed according to the integrated rate law for a first-order reaction in Prism v6.0b (GraphPad).

**Bioinformatics Analyses and Statistics**

For pre-miRNA conservation analysis, high-confidence miRNAs were extracted from mirBase 21 (Kozomara and Griffiths-Jones, 2014) and curated as described in Supplemental Experimental Procedures. Nucleotide composition of the Drosophila genome dm3 was determined excluding chromosome MT, U, and Uextra.

Mapping of sequencing libraries to Drosophila genome dm3 was performed as described (Ameres et al., 2010). Annotations were derived from FlyBase r5.55_FB2014_01. For details see Supplemental Experimental Procedures. Statistical tests were performed in Prism v6.0b (GraphPad) or Excel v14.4.3 (Microsoft). For multiple hypothesis testing a false-discovery rate (FDR) threshold was set to < 0.1 according to Benjamini and Hochberg (Benjamini and Hochberg, 1995).

**Supplemental Information**

Supplemental Information includes seven figures, two tables, and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.molcel.2015.05.033.

**Supplemental Information**

The accession numbers for the high-throughput sequencing datasets reported in this paper are GEO: GSE66213 and GEO: GSE67646.

**Acknowledgments**

We thank Y. Tomari for Dcr-1 expression vector; R. Fukunaga for pre-bantam-transcription template plasmid; P. Duchek for CRISPR/Cas9 genome editing in flies; E.C. Lai for discussions on the evolutionary conservation of pre-miRNAs; Exelixis (Harvard Medical School) for fly stocks; and all members of the S.L.A. lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raised by the lab for input and discussions. Monoclonal antibodies were raise...
Huntzinger, E., and Izaurralde, E. (2011). Gene silencing by microRNAs: contributions of translational repression and mRNA decay. Nat. Rev. Genet. 12, 99–110.

Kim, V.N., Han, J., and Siomi, M.C. (2009). Biogenesis of small RNAs in animals. Nat. Rev. Mol. Cell Biol. 10, 126–139.

Kim, Y.K., Heo, I., and Kim, V.N. (2010). Modifications of small RNAs and their associated proteins. Cell 143, 703–709.

Kozomara, A., and Griffiths-Jones, S. (2014). miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res. 42, D68–D73.

Ladewig, E., Okamura, K., Flynt, A.S., Westholm, J.O., and Lai, E.C. (2012). Discovery of hundreds of mirtrons in mouse and human small RNA data. Genome Res. 22, 1634–1645.

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S., and Kim, V.N. (2003). The nuclear RNase III Drosha initiates microRNA processing. Nature 425, 415–419.

Lee, Y.S., Nakahara, K., Pham, J.W., Kim, K., He, Z., Sontheimer, E.J., and Carthew, R.W. (2004). Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. Cell 117, 69–81.

Lehrbach, N.J., Armisen, J., Lightfoot, H.L., M urfitt, K.J., Bugaut, A., Balasubramanian, S., and Miska, E.A. (2009). LIN-28 and the poly(U) polymerase PUP-2 regulate let-7 microRNA processing in Caenorhabditis elegans. Nat. Struct. Mol. Biol. 16, 1016–1020.

Li, J., Yang, Z., Yu, B., Liu, J., and Chen, X. (2005). Methylation protects miRNAs and siRNAs from a 3′-end uridylation activity in Arabidopsis. Curr. Biol. 15, 1501–1507.

Lim, L.P., and Burge, C.B. (2001). A computational analysis of sequence features involved in recognition of short introns. Proc. Natl. Acad. Sci. USA 98, 11193–11198.

Liu, X., Zheng, Q., Vrettos, N., Maragkakis, M., Alexiou, P., Gregory, B.D., and Mourelatos, Z. (2014). A MicroRNA precursor surveillance system in quality control of MicroRNA synthesis. Mol. Cell 55, 868–879.

Martin, G., Doublié, S., and Keller, W. (2008). Determinants of substrate specificity in RNA-dependent nucleotidyl transferases. Biochim. Biophys. Acta 1779, 206–216.

Mohammed, J., Siepel, A., and Lai, E.C. (2014). Diverse modes of evolutionary emergence and flux of conserved microRNA clusters. RNA 20, 1850–1863.

Newman, M.A., Mani, V., and Hammond, S.M. (2011). Deep sequencing of microRNA precursors reveals extensive 3′ end modification. RNA 17, 1795–1803.

Norbury, C.J. (2013). Cytoplasmic RNA: a case of the tail wagging the dog. Nat. Rev. Mol. Cell Biol. 14, 643–653.

Okamura, K., Ishizuka, A., Siomi, H., and Siomi, M.C. (2004). Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. Genes Dev. 18, 1655–1666.

Okamura, K., Hagen, J.W., Duan, H., Tyler, D.M., and Lai, E.C. (2007). The mirtron pathway generates microRNA-class regulatory RNAs in Drosophila. Cell 130, 89–100.

Ren, G., Chen, X., and Yu, B. (2012). Uridylation of miRNAs by hen1 suppressor1 in Arabidopsis. Curr. Biol. 22, 695–700.

Ruby, J.G., Jan, C.H., and Bartel, D.P. (2007). Intrinsic miRNA precursors that bypass Drosha processing. Nature 448, 83–86.

Scott, D.D., and Norbury, C.J. (2013). RNA decay via 3′ uridylation. Biochim. Biophys. Acta 1829, 654–665.

Thibault, S.T., Singer, M.A., Miyazaki, W.Y., Milash, B., Dompe, N.A., Singh, C.M., Buchholz, R., Demsky, M., Fawcett, R., Francis-Lang, H.L., et al. (2004). A complementary transposon tool kit for Drosophila melanogaster using P and piggyBac. Nat. Genet. 36, 283–287.

Tsutsumi, A., Kawamata, T., Izumi, N., Seitz, H., and Tomari, Y. (2011). Recognition of the pre-miRNA structure by Dicer-1. Nat. Struct. Mol. Biol. 18, 1153–1158.

Westholm, J.O., and Lai, E.C. (2011). Mirtrons: microRNA biogenesis via splicing. Biochimie 93, 1897–1904.

Westholm, J.O., Ladewig, E., Okamura, K., Robine, N., and Lai, E.C. (2012). Common and distinct patterns of terminal modifications to mirtrons and canonical microRNAs. RNA 18, 177–192.

Xie, J., Ameres, S.L., Friedline, R., Hung, J.H., Zhang, Y., Xie, Q., Zhong, L., Su, Q., He, R., Li, M., et al. (2012). Long-term, efficient inhibition of microRNA function in mice using rAAV vectors. Nat. Methods 9, 403–409.

Yandell, M., Mungall, C.J., Smith, C., Prochnik, S., Kaminker, J., Hartzell, G., Lewis, S., and Rubin, G.M. (2008). Large-scale trends in the evolution of gene structures within 11 animal genomes. PLoS Comput. Biol. 2, e15.

Yang, J.S., and Lai, E.C. (2011). Alternative miRNA biogenesis pathways and the interpretation of core miRNA pathway mutants. Mol. Cell 43, 920–932.

Zhao, Y., Yu, Y., Zhai, J., Ramachandran, V., Dinh, T.T., Meyers, B.C., Mo, B., and Chen, X. (2012). The Arabidopsis nucleotidyl transferase HESO1 uridylates unmethylated small RNAs to trigger their degradation. Curr. Biol. 22, 689–694.
Uridylation of RNA Hairpins by Tailor Confines the Emergence of MicroRNAs in *Drosophila*

Madalena M. Reimão-Pinto, Valentina Ignatova, Thomas R. Burkard, Jui-Hung Hung, Raphael A. Manzenreither, Ivica Sowemimo, Veronika A. Herzog, Brian Reichholf, Sara Fariña-Lopez, and Stefan L. Ameres
Figure S1. Detailed presentation of post-transcriptional modifications of miRNAs in S2 cells. Related to Figure 1. 
(A) The extent to which each miRNA (> 100 ppm) is tailed is shown. Values represent the mean of three independent biological replicates ±SD. 
(B) Total number of prefix-matching reads, i.e. reads with non-genome matching nucleotide addition. Values represent the mean of three independent biological replicates ±SD. 
(C) Mean length of non-genome matching nucleotide additions. Values represent the mean of three independent biological replicates ±SD. 
(D) Composition of non-genome-matching nucleotides for each miRNA. Values represent the mean of three independent biological replicates ±SD. 
(E) Composition of non-genome-matching nucleotides in classified miRNAs (see Fig. 1C). Tailing of mirtrons is predominantly associated with uridylation.
Figure S2. Characterization of Tailor in S2 cells and in vivo in flies. Related to Figure 2.

(A) Schematic representation of the predicted structure of Tailor/CG1091 gene and RNA transcripts. The location of a 7 bp deletion introduced by CRISPR/Cas9 genome engineering in flies (CG1091<sup>c4-1/6</sup>) is indicated and confirmation by Sanger sequencing of a PCR product prepared from the locus of WT (w<sup>1118</sup>) flies and flies homozygous for CG1091<sup>c4-1</sup> is shown. The frameshift causes premature stop of translation, resulting in a predicted 12.5 kDa protein.

(B) Two independently generated CRISPR/Cas9 alleles (CG1091<sup>c4-1</sup> and CG1091<sup>c4-6</sup>) similarly reduce post-transcriptional modifications in mature miRNAs. Change in the fraction at which each miRNA is tailed in homozygous CG1091<sup>c4-1/6</sup> compared to control flies (w<sup>1118</sup>) is compared between CG1091<sup>c4-1</sup> and CG1091<sup>c4-6</sup>.

(C) CG1091-RA

(D) DAPI

(E) S2 cells

AFMW-Tailor

AFMW-Tailor

Myc-Tailor

Myc-Tailor

DAPI

DAPI

75 kDa

Tailor

Actin

100 kDa

25

50

75

37

25

50

37

100 kDa

25

50

75

37

100 kDa

25

50

75

37

CG1091 f05717

Change in fraction tailed (CG1091<sup>c4-6</sup> vs w<sup>1118</sup>)

Change in fraction tailed (CG1091<sup>c4-1</sup> vs w<sup>1118</sup>)

r<sup>2</sup> = 0.86

p < 1 × 10<sup>-4</sup>
A significant correlation was detected (Pearson correlation coefficient; $r^2=0.86; p < 10^{-4}$).

(C) CG10910-RB is predominantly expressed in S2 cells. Coverage plot of the CG1091 locus in polyA+ RNAseq datasets from Drosophila S2 cells. Reads mapping to the first exon of CG1091-RB, but not the very 5' UTR of CG1091-RA were detectable.

(D) In S2 cells, Myc-tagged CG1091 localizes to the cytoplasm without detectable signal in the nucleus. FLAG-Myc-tagged CG1091 expression was driven by an Actin5C promoter (pAFMW-CG1091) upon transient transfection in S2 cells, followed by immunostaining (anti-Myc) and imaging. Single color channel images show total inversions for DAPI and Myc-CG1091. Scale bar = 5 µm.

(E) Transfection of S2 cells with pAFMW-CG1091/Tailor produces full-length, Myc-tagged Tailor protein of the expected size without truncations. Western blot analysis of lysate produced from untreated S2 cells or cells expressing Myc-tagged Tailor is shown. Endogenous Tailor and Actin served as loading controls.
Figure S3. Tailor-directed uridylation regulates miRNA abundance and function. Related to Figure 3.

(A) MicroRNAs that significantly changed in abundance upon depletion of Tailor were more frequently tailed under unperturbed conditions. The mean absolute change in miRNA abundance and their respective mean tailing status in control-treated S2 cells are shown. Pearson’s correlation coefficient is indicated. Error bars represent SD.
(B) Northern hybridization experiments confirm a loss in tailing of miR-184 upon depletion of CG1091/Tailor in S2 cells, which coincides with a decrease in miR-184 abundance when compared to control dsRNA (GFP and LUC) or untreated samples. 28S rRNA and bantam serve as loading controls.

(C) Tailor is required for the normal accumulation of miR-184. Quantification of three independent biological replicates of experiment shown in (B) (left) or in small RNA libraries (right). Abundance was normalized to untreated S2 cells. Tailed isoforms were excluded in quantifications. Mean ±SD is reported.

(D) Schematic representation of CRISPR/Cas9-directed targeting of the CG1091/Tailor locus in Drosophila S2 cells. Four guide RNAs targeting the third exon of CG1091B are indicated. Targeting was confirmed by Surveyor assay (see E) using the primers 1091c-fwd and 1091c-rev (see Table S2).

(E) Surveyor assay confirming genomic editing at the CG1091/Tailor locus in S2 cells. Cells were selected for 7 days for expression of Cas9 and four guide RNAs. Expected size of PCR-product using non-edited genomic DNA is 531 bp. The observed decrease in PCR-product-length using genomic DNA from S2-CG1091CRIPSR-mix cells is attributed to genomic deletions as a consequence of targeting multiple sites simultaneously.

(F) Targeting of CG1091/Tailor in S2 cells by CRISPR depleted Tailor-protein to levels that were not detectable by Western blotting. A monoclonal antibody raised against CG1091 was used. Actin represents loading control.

(G) Abundance (in ppm) of selected, abundant miRNAs that decrease (miR-184-3p, top and miR-980, middle) or increase (miR-1003, bottom) in abundance upon depletion of CG1091/Tailor by RNAi in Drosophila S2 cells when compared to control-treated cells. Values represent the mean of three independent biological replicates ±SD.

(H) Cumulative distribution of mRNA-Seq changes upon depletion of CG1091/Tailor by CRISPR/Cas9-genome editing in Drosophila S2 cells compared to untreated cells. Plotted are distributions of all abundantly expressed genes in untreated cells (>1 FPKM, black; 8,444 genes) and genes targeted by miR-184 (red, top; 23 genes), miR-980 (red, middle; 30 genes) or miR-1003 (red, bottom; 18 genes) as predicted by targetscan (http://www.targetscan.org/fly_12/) (Ruby et al., 2007). The median fold-change of the respective target mRNA-set is reported in red. For miR-184 targets, a median 1.3-fold increase in target mRNA levels was detected upon depletion of CG1091. For miR-980, a miRNA that likewise decreased in abundance upon Tailor-depletion, albeit to a lesser extent compared to miR-184, a median 1.1-fold increase in target mRNAs was observed upon depletion of Tailor. Finally, target mRNAs of miR-1003, a mirtron that increased in abundance upon depletion of Tailor decreased in median abundance by 0.95-fold upon depletion of Tailor.

(I) Fold-change of individual, predicted target-mRNAs for miR-184 (top; 14 out of 23 target mRNAs with statistically significant increase), miR-980 (middle; 13 out of 30 targets with statistically significant increase and only three with significant decrease) and miR-1003 (bottom; five out of 18 targets with statistically significant decrease and only three with increase) upon depletion of CG1091/Tailor by CRISPR/Cas9-genome editing in Drosophila S2 cells compared to untreated cells. Mean fold-change ±SD is reported. Adjusted p-values (according to Benjamini-Hochberg’s procedure) are shown.
Figure S4. The two-nucleotide 3′ overhang in pre-miRNAs is an important feature for efficient processing by Dcr-1. Related to Figure 4.

(A and D) In vitro dicing assay. 5′ radiolabelled pre-miR-184 (A) or pre-bantam (D) with the indicated 3′ end modification was incubated with immunopurified, Strep-tagged Dcr-1. Reaction was incubated for the indicated time and separated on a 15% denaturing polyacrylamide gel.

(B and E) Quantification of three independent replicates of the experiment shown in (A) or (D), respectively. Accumulation of mature miR-184 or bantam was determined relative to the sum of pre- and mature-miRNA. Reported are mean values ±SD.

(C and F) Pre-miR-184 or pre-bantam containing a 2nt 3′ overhang is diced at significantly higher rates compared to alternative 3′ ends (p-value determined by Student’s t-test). Dicing rates were determined by burst kinetics in three experimental replicates and reported as mean ±SD.
Figure S5. Detailed presentation of post-transcriptional modifications of pre-miRNAs in S2 cells. Related to Figure 5.
(A) Extent of tailing is shown for each pre-miRNA. Values represent the mean of three independent biological replicates ±SD.
(B) Mean length of non-genome matching nucleotide additions for each pre-miRNA. Values represent the mean of three independent biological replicates ±SD.
(C) Total number of prefix-matching reads, i.e. reads with non-genome matching nucleotide addition. Values represent the mean of three independent biological replicates ±SD.
(D) Composition of non-genome-matching nucleotides for each pre-miRNA.
(E) Composition of non-genome-matching nucleotides in classified miRNAs (see Fig. 1C). Tailing of mirtron hairpins is predominantly associated with uridylation.
(F) Correlation of tailing status observed for pre-miRNAs and mature miRNAs, derived from the 3p (left) or 5p (right) arm of the respective precursor. Values represent the mean of three independent biological replicates ±SD. Pearson’s correlation coefficient r and p-values are indicated.
(G) Changes in mature miRNA abundance correlate with CG1091-directed tailing of the underlying pre-miRNAs. Correlation of the absolute change in mature miRNA abundance upon depletion of Tailor in S2 cells compared with the tailing status of the respective pre-miRNA in control treated S2 cells. Values represent the mean of three independent biological replicates ±SD. Pearson’s correlation coefficient and p-value are indicated.
(H) Changes in mature miRNA abundance correlate with CG1091-directed tailing of the underlying pre-miRNAs. Changes in miRNA tailing between CG1091-depleted and control-dsRNA-treated S2 cells is shown for miRNAs that do, or do not, change significantly in abundance upon depletion of CG1091. P-Value (Whitney-Mann test) is indicated.
Figure S6. Biochemical characterization of Tailor. Related to Figure 6.

(A) Domain organization of Tailor. Zoom-in shows the sequence of the conserved Polβ-superfamily catalytic motif (green box). The three conserved

---

**Table:**

| Sample   | Time (min) | Reads    | Adapter stripped | Unambig. reads excl. N | Background corrected | Tailed reads | Tail composition (%) | Relative fraction tailed |
|----------|------------|----------|------------------|------------------------|----------------------|--------------|----------------------|--------------------------|
| Input    | -          | 6,997,950| 6,878,306        | 6,685,476              | 6,643,116            | 0            | n.a.                 | 99.75                    |
| NTP      | 2          | 7,054,769| 7,034,203        | 6,734,922              | 6,417,022            | 2,974,524    | 0.020                | 99.87                    |
| NTP      | 5          | 7,034,769| 7,005,566        | 6,690,813              | 6,153,349            | 3,530,222    | 0.054                | 99.75                    |

*See Experimental Procedures for details.*
aspartate residues (grey boxes) predicted to form the catalytic site are shown and one of these is replaced by alanine to generate a catalytic mutant form.

(B) Wild-type FLAG-tagged CG1091 or the catalytic mutant derivative was expressed in S2 cells and immunopurified. Western blot analysis using anti-FLAG antibody confirms purification. Actin confirms enzyme purity.

(C) Incubation of Tailor/CG1091 with a 22 nt, 5’ radiolabelled small RNA in the presence of ribonucleotide triphosphates extended the length of the substrate RNA by up to ~10 nucleotides within 5 min. The activity was abolished in the catalytic mutant.

(D) In vitro tailing reaction using immunopurified FLAG-Tailor/CG1091 in the presence of all, or individual ribonucleotide triphosphates, as indicated, after 2 or 5 min incubation.

(E) Quantification of the experiment shown in (D). Tailor exhibited highest activity in the presence of UTP. While CTP was efficiently employed, incorporations resulted only in short tails, indicating low processivity (see D). The activity was lower in the presence of ATP and almost undetectable in the presence of GTP.

(F) Schematic overview of high-throughput assay for the biochemical characterization of Tailor/CG1091. The assay employs an RNA substrate with four randomized 3’ nucleotides (sequence is shown). Following in vitro tailing assays, the product was gel-purified, subjected to 3’ adapter ligation, followed by reverse transcription, PCR amplification, and high-throughput sequencing.

(G) Abundance of the 256 substrate RNAs in high-throughput sequencing datasets generated via the assay described in (F). Abundance of all substrates in input samples and 2 min or 5 min tailing reactions are shown.

(H) Sequencing statistics of high-throughput biochemical characterization assay.

(I) Secondary structure prediction of substrate RNA was generated using RNAfold (Gruber et al., 2008).

(J) Examples for substrate RNAs that fold into secondary structures with high stability as predicted using RNAfold (Gruber et al., 2008).

(K) The fraction at which the indicated substrates are tailed over time is represented as median (light grey line) and inner quartile range (colored area). Stars indicate significantly different tailing efficiency compared to all substrates (p-values determined by Mann-Whitney test).

(L) Cumulative distributions of the relative fraction tailed for the indicated 3’ terminal dinucleotide-containing substrate RNAs is shown.

(M) Box-plot depiction of data shown in (L). Three stars represent a p-value of <0.001 (KS-test).
Figure S7. Tailor-directed uridylation of hairpins ending in 3’G and its implications for miRNA evolution. Related to Figure 7.

(A) 5’ radiolabeled miR-1003 containing the indicated 3’-terminal nucleotide was incubated with immunopurified Tailor for the indicated time, followed by denaturing gel electrophoresis.

(B) Quantification of experiment shown in (A).

(C) Non-conserved pre-miRNAs expressed in S2 cells more frequently end in 3’G compared to conserved miRNA-hairpins.

(D) Data shown in Figure 7G including information on conservation among different Drosophila species.
Table S1. List of pre-miRNAs recovered by high-throughput sequencing of 40-100 nt RNAs from Drosophila S2 cells, without evidence for abundant expression (>100 ppm) in 18-30 nt small RNA libraries. Conservation of pre-miRNAs was described previously (Mohammed et al., 2014). Related to Figure 5.

| Pre-miRNA   | Abundance (ppm) | Type     | Conservation              |
|------------|-----------------|----------|---------------------------|
| pre-miR-1005 | 297             | mirtron  | Melanogaster subgroup     |
| pre-miR-1009 | 359             | mirtron  | Melanogaster subgroup     |
| pre-miR-1013 | 850             | mirtron  | Melanogaster subgroup     |
| pre-miR-2489 | 15675           | mirtron  | Melanogaster subgroup     |
| pre-miR-2497 | 212             | mirtron  | Melanogaster subgroup     |
| pre-miR-3645 | 210             | canonical| Melanogaster subgroup     |
| pre-miR-4910 | 230             | mirtron  | Melanogaster subgroup     |
| pre-miR-4914 | 686             | mirtron  | Melanogaster subgroup     |
| pre-miR-4918 | 433             | mirtron  | Melanogaster subgroup     |
| pre-miR-4939 | 295             | canonical| Melanogaster subgroup     |
| pre-miR-4950 | 133             | canonical| Melanogaster subgroup     |
| pre-miR-4950 | 131             | canonical| Melanogaster subgroup     |
| pre-miR-4962 | 4416            | canonical| Melanogaster subgroup     |
| pre-miR-927  | 120             | canonical| Pan-Drosophilid           |
| pre-miR-954  | 1135            | canonical| Melanogaster subgroup     |
| pre-miR-1002 | 1255            | canonical| Pan-Drosophilid           |
| pre-miR-4912 | 813             | mirtron  | Melanogaster subgroup     |
| pre-miR-2279 | 617             | canonical| Melanogaster subgroup     |
| pre-miR-4949 | 344             | canonical| Melanogaster subgroup     |
| pre-miR-4968 | 286             | canonical| Melanogaster subgroup     |
| pre-miR-4974 | 210             | canonical| Melanogaster subgroup     |
| pre-miR-2501 | 201             | mirtron  | Melanogaster subgroup     |
| pre-miR-3643 | 174             | mirtron  | Melanogaster subgroup     |
| pre-miR-994  | 174             | canonical| Pan-Drosophilid           |
| pre-miR-2492 | 170             | canonical| Melanogaster subgroup     |
| pre-miR-977  | 159             | canonical| Pan-Drosophilid           |
| pre-miR-975  | 134             | canonical| Pan-Drosophilid           |
| pre-miR-4987 | 133             | canonical| Pan-Drosophilid           |
| pre-miR-31a  | 133             | canonical| Pan-Drosophilid           |
| pre-miR-1000 | 131             | canonical| Pan-Drosophilid           |
| pre-miR-193  | 121             | canonical| Pan-Drosophilid           |
| pre-miR-2493 | 118             | canonical| Melanogaster subgroup     |
| pre-miR-4951 | 116             | canonical| Sophophoran               |
| pre-miR-3641 | 114             | mirtron  | Melanogaster subgroup     |
| pre-miR-983-1| 110             | canonical| Melanogaster subgroup     |
| pre-miR-983-2| 110             | canonical| Melanogaster subgroup     |
| pre-miR-3642 | 109             | mirtron  | Melanogaster subgroup     |
| pre-miR-4966 | 104             | canonical| Melanogaster subgroup     |
Table S2. Oligonucleotides used in this study. Related to Experimental Procedures.

Guide RNA for CRISPR/Cas9 genome editing of the CG1091 locus in flies.

| Name              | DNA sequence (5'-3')       |
|-------------------|---------------------------|
| gRNA              | GACGACGAATCCGCCGCAGC      |

Guide RNA-oligonucleotides used for CRISPR/Cas9 genome editing of the CG1091 locus in Drosophila S2 cells (capital letters indicate targeting sequence; lowercase letters indicate restriction-site-compatible overhangs).

| Name               | DNA sequence (5'-3')       |
|--------------------|----------------------------|
| gRNA-1-fwd         | ttcgGCCAGGTGGGTCTAGCTG     |
| gRNA-1-rev         | aacCAGCTAGACCCGACCTGGCc    |
| gRNA-2-fwd         | ttcgCCACCATAATGGCTCCAGC    |
| gRNA-2-rev         | aacCGTGGAGCGCATATGGTGGc    |
| gRNA-3-fwd         | ttcgGCCGAGGATGGCCGCAC      |
| gRNA-3-rev         | aacGGTCCGGCATCTTGCCGCCc    |
| gRNA-4-fwd         | ttctTTATTGTCTGCAGCTTCCTG   |
| gRNA-4-rev         | aacCAGGAGCTTGAGCACAATAAAc  |

PCR primers for the confirmation of CG1091-Cas9 alleles in flies.

| Name               | DNA sequence (5'-3')       |
|--------------------|----------------------------|
| CG1091 fwd         | CGCATTCGATTAACCTTTCAGC     |
| CG1091 rev         | ATTCCTGTCCAAGTAGCTCTCG     |

PCR primers for Surveyor assay to confirm CG1091-targeting by CRISPR in Drosophila S2 cells.

| Name               | DNA sequence (5'-3')       |
|--------------------|----------------------------|
| CG1091'-fwd        | GGTAAACAAAGGCCACGAAG       |
| CG1091'-rev        | ATGCTGACGAGCTTCACC         |

Oligonucleotides for cloning of CG1091-RB and catalytic mutant version.

| Name               | DNA sequence (5'-3')       |
|--------------------|----------------------------|
| CG1091-CDS-fwd     | CACCATGCGGATCGAATGCCAGCG   |
| CG1091-CDS-rev     | TTAAGATGCATAGGCGTGCAGCG    |
| CG1091_CM_FWD      |GGAGCGGCCAGGCAGTTGAGACGC   |
| CG1091_CM_REV      | GGACACCAATGAGTGAGTTTAGTC  |

Oligonucleotides for the generation of templates for T7 transcription of dsRNA.

| Name               | DNA sequence (5'-3')       |
|--------------------|----------------------------|
| T7-CG1091 fwd      | TAATACGACTCACTATAGGGCTTTGTAGCGATCGCAACTGTT |
| T7-CG1091 rev      | TAATACGACTCACTATAGGGCTTTGTAGCGATCGCAACTGTT |
| T7-GFP fwd         | TAATACGACTCACTATAGGGCAACAGTTCAGCGATTCGCC |
| T7-GFP rev         | ATTACAGACTCACTATAGGGCAACAGTTCAGCGATTCGCC |
| T7-Luc fwb         | TAATACGACTCACTATAGGGGATCCGCCTTCAGCATCAAG |
| T7-Luc rev         | TAATACGACTCACTATAGGGGATCCGCCTTCAGCATCAAG |
| T7-Trf4-fwd        | TAATACGACTCACTATAGGGCAACTGTCATAGGGGATCCGCCTTCAGCATCAAG |
| T7-Trf4-rev        | TAATACGACTCACTATAGGGCAACTGTCATAGGGGATCCGCCTTCAGCATCAAG |
| T7-CG11418 fwd      | TAATACGACTCACTATAGGGCAACTGTCATAGGGGATCCGCCTTCAGCATCAAG |
| T7-CG11418 rev      | TAATACGACTCACTATAGGGCAACTGTCATAGGGGATCCGCCTTCAGCATCAAG |
| T7-Mkg-p fwb        | TAATACGACTCACTATAGGGGAGCCCCAATAGGGGATCCGCCTTCAGCATCAAG |
| T7-Mkg-p rev        | TAATACGACTCACTATAGGGGAGCCCCAATAGGGGATCCGCCTTCAGCATCAAG |
Probes for Northern hybridization.

| Name             | DNA sequence (5'→3')       |
|------------------|---------------------------|
| dme-miR-8-3p     | GACATCTTTACCTGACAGTATTA   |
| dme-miR-184-3p   | GCCCTTATCGTCTCCGTCACA     |
| dme-bantam-3p    | ATGAAAAACAGTCAAACAAATCGAAAAACGG |
| dme-miR-33-5p    | GACAATGCGACTACAATGCAC     |
| dme-miR-252-5p   | ATGAAAAACAGTCAAACAAATCGAAAAACGG |
| dme-miR-34-5p    | AACCAGCTAACCACACTGCA      |
| dme-pre-miR-184  | GCCCTTATCGTCTCCGTCAGTTGTCTTTAAGTGCTCA |
| dme-pre-bantam   | AATCAGCTTTCAAAATGATCTCACTTGTATGAAAAAC |
| dme-2SrRNA       | TACAACCCTCAACCATATGTCATCCAAGCA |

Linkers used for cloning of RNA libraries.

| Name      | Type   | Sequence (5'→3')       |
|-----------|--------|------------------------|
| 3’ linker | DNA    | rApp/NNNAGATCGGAAGAGCACGTCT/ddC/ |
| 5’ linker | RNA    | ACACUCUUUUCCCCUACACGACGGCUCUUCGCAUCUNNNN |

Oligonucleotides for small-scale pre-miRNA cloning.

| Name       | RNA sequence (5'→3')       |
|------------|---------------------------|
| pre-miR-184-twd | CCTATATCTCCCTGCCCCCC     |
| pre-bantam-twd | CCGTTTTTCTGATTTGTTTAGCT     |
| 3’ Linker-rev | GAGACGCTGTGCTTCCGATCT     |

RNA oligonucleotides used for in vitro tailing and dicing assays.

| Name       | RNA sequence (5'→3')       |
|------------|---------------------------|
| miR-184-5p | CCUUAUCAUUCUCUGCCCGC     |
| miR-184-loop | UGUGCAGUAAAGCAAC     |
| miR-184-3p | UGGACCGGAGAACUGAAAGGGC |
| miR-184-3p-U | UGGACCGGAGAACUGAAAGGCU |
| 22 nt substrate | UGAGGUAGUAGGUUGUAGUA |
| 37 nt substrate with 4 random 3'nts | ACACUCUUUCCCCUACACGACGCUCUCGCCAUCUNNNN |
| miR-1003-3p WT | UUCUCAUUCAUCAUCAAG     |
| miR-1003-3p A  | UUCUCAUUCAUCAUACAA     |
| miR-1003-3p U  | UUCUCAUUCAUCAUACAU     |
| miR-1003-3p C  | UUCUCAUUCAUCAUACAC     |
| miR-1003-3p-loop | GUGGGUACUCGUGUGGUUGGUUGCUUGCGCGUCC |
| miR-1003-3p-U | UUCUCAUUCAUCAUACAGU     |
| miR-1003-3p-UU | UUCUCAUUCAUCAUACAGUU     |
| miR-1003-3p-UU | UUCUCAUUCAUCAUACAGUU     |

DNA-splint oligonucleotides used to generate pre-miR-1003 and –miR-184 variants.

| Name       | DNA sequence (5'→3')       |
|------------|---------------------------|
| miR-1003 splint | GTGAATATGTAAATGTGAGAGGACCGCCAGGCAACCACATCCAGA TACC |
| miR-184 splint | GCCCTTATCGTCTCCGTCATTTAAGTGACACGGGGCGAGAATGATAAGG |
**Supplemental Experimental Procedures**

**Fly stocks**

Generation of CG1091 (FBgn0037470) mutant flies by CRISPR/Cas9 genome engineering was performed as described (Gokcezade et al., 2014). Briefly, isogenised $w^{1118}$ embryos were injected with the plasmid pDCC6 (Addgene) containing a gRNA sequence (Table S2) targeting the second exon of the CG1091 locus. Hatched flies were crossed to balancer flies on the third chromosome and F1 resulting males were screened for frameshift mutations by PCR amplification of the targeted CG1091 locus using the primers CG1091-fwd and CG1091-rev (Table S2) followed by sanger sequencing (see Fig. S2A). Depletion of CG1091 protein was confirmed by Western blotting (see Fig. 2E). Progeny resulting from two distinct injected embryos carrying a 7nt deletion were used for further experiments. Flies between 0-5 days old were used for experiments and $w^{1118}$ flies were used as the wild type control.

**Fertility assays**

Male fertility was assayed by crossing single 2-day-old males of the indicated genotype with five 3-to-4-day-old wild-type ($w^{1118}$) virgin females, allowing them to mate for a 48h period in fresh vials. Subsequently, the male was removed, the inseminated females were transferred to a fresh vial and flipped to fresh vials every 2-3 days for a period of 7 days. Eclosing offspring were counted for 19 days after flipping. For each tested male genotype, 8 independent crosses were performed and quantified.
Female oviposition assays were performed by crossing ten 5-day-old females of each genotype with five 2-day-old wild-type (w1118) males in small fly cages with an apple juice containing a dab of fresh yeast and allowed to mate for a 48h period. Subsequently, the apple juice plate was replaced and collected after a 6h period. The number of eggs laid per female in this 6h-period were counted and reported as average fecundity (number of eggs/female). Plates were incubated at 25°C for another 24h and the number of hatched eggs was counted to determine hatching rates [(number hatched eggs/total laid eggs)*100]. This procedure was repeated for four consecutive days for each cross. All oviposition assays were performed in the same 6h period (9am-3pm) for each of the four consecutive days, and eggs were counted at the same time points for both oviposition and hatching rate determinations. For each tested female genotype, three independent crosses were performed.

**Plasmids**

Tailor-CDS was PCR-amplified from S2 cell cDNA, cloned by directional TOPO cloning according to manufacturer’s instructions (Invitrogen). For expression in S2 cells, Tailor-CDS was subcloned into pAFMW by LR recombination (Invitrogen). A catalytic mutant point mutation was introduced by site-directed-mutagenesis using the oligonucleotides CG1091_CM_FWD and CG1091_CM_REV. For oligonucleotides see Table S2.

For cloning of gRNAs targeting the CG1091/Tailor locus by CRISPR/Cas9 in *Drosophila* S2 cells, four pairs of gRNA oligos (gRNA-1 to -4; see Table S1) were annealed and ligated to BspQI-digested pAc-sgRNA-Cas9 (Bassett et al., 2014).
**Generation of clonal S2 cell lines**

For the generation of clonal S2 cell lines, S2 cells were co-transfected with the indicated plasmid and ~1/10 the amount of a plasmid providing resistance to Hygromycin using Cellfectin II (LifeTechnologies) according to the instructions of the manufacturer. 48 h post-transfection, Hygromycin (Sigma) was added to a final concentration of 300 µg/ml and cells were diluted ~1/5 each time they reach a density of 8-10 x 10⁶ cells/ml (approximately once a week). Selection for stable integration was typically accomplished within 2 – 3 weeks. For clonal selection, cells were then diluted serially 1/4 in a 96-well plate starting with ~20,000 cells/well in clonal selection medium (80% fresh Schneider Medium [containing 10% FCS], 20 % conditioned Schneider medium, and 300 µg/ml Hygromycin) and incubated 10 – 14 days. Single, round-shaped clones of cells emerging from wells with permissive cell density were picked, expanded, and transgene expression was confirmed by Western blotting.

**RNAi in S2 Cells**

Regions targeted by double-stranded RNA were from (Dietzl et al., 2007). DNA templates for in vitro transcription were amplified from genomic DNA by PCR using primers incorporating the T7 promoter sequence (see Table S2). After isopropanol precipitation, PCR products were used as templates for transcription by T7 RNA polymerase. DsRNA products were purified using MEGA clear RNA purification kit (Ambion). S2 cells (2 x 10⁶) were soaked with 10 µg dsRNA on day 0 for 1 h in ExpressFive SFM medium (Invitrogen) supplemented with 200 mM Glutamine (Invitrogen) in the absence of FBS, followed by the addition of
medium containing 10% FBS. Soaking was repeated at day 4 and cells were harvested and processed at day 7.

**CRISPR/Cas9 genome editing in Drosophila S2 cells**

S2 cells were transfected with a mixture of four pAc-sgRNA-Cas9 plasmids encoding Cas9 and each one sgRNAs targeting a 160 bp region in the second exon of CG1091B (Fig. S3D). Cells were selected for transgene expression using Puromycine (5μg/ml) for one week. Tailor/C1091-depletion was determined by Western blot analysis (Fig. S3F) and genomic targeting was confirmed by Surveyor assay (Fig. S3E). Total RNA was prepared for three independent replicates using Trizol (LifeTechnologies).

**Surveyor Assay**

Genomic DNA was extracted from ~2 x 10^6 cells using gDNA-prep lysis buffer (100mM Tris pH8, 50mM Nacl, 50mM EDTA, 1% SDS), followed by the amplification of a 531bp region including the CRISPR-targeting sites using the primer 1091-fwd and 1091-rev using Taq polymerase (Fig. S3D and Table S2). The PCR product was denatured and re-annealed in a PCR machine (95°C for 5 min; Ramp to 85°C [-2°C/sec]; Ramp to 25°C [-0.1°C/sec]) followed by treatment of 0.5μl PCR-product with 5 Units of T7 Endonuclease I (New England Biolabs) for 20 min at 37°C. The product was resolved on a 2% agarose gel.
Antibody

A monoclonal antibody against CG1091/Tailor was raised by the MFPL monoclonal antibody facility (S. Schuechner and E. Ogris). Briefly, 6 x HIS-tagged CG1091B was expressed in *E. coli* BL21 and purified under denaturing conditions by Ni-affinity chromatography. Balb/c mice were immunized subcutaneously three times (every 2-3 weeks) with 50 µg of purified antigen mixed at a ratio of 1:1 with adjuvant, before a final intravenous immunization with 30 µg purified antigen (adjuvant-free). Splenic B-cells were fused with X63-Ag8.653 mouse myeloma cells and clones were tested by immunoblot analysis for the detection of CG1091/Tailor. Clone 4C5-B7, producing IgG1-Kappa, yielded the best signal-to-noise performance.

Western blotting

Lysates were separated on 10% SDS PAGE and transferred to PVDF membranes (Immobilon-P, Millipore). Antibodies were used at a dilution of 1/500 for α-Tailor (mouse), 1/10,000 for α-M2-FLAG (mouse; Sigma) and 1/10,000 for α-Actin (rabbit, Sigma) and detected by secondary HRP-antibody-conjugates G21234 or G21040 (Invitrogen; dilution 1/10,000). Primary antibodies were incubated at 4°C over-night and secondary antibodies were incubated at room-temperature for two hours. Images were acquired on a ChemiDoc MP Imaging System (BioRad) using ImageLab v5.1.1 (BioRad).
**Immunocytochemistry**

IF-staining of S2 cells was performed as described previously, with the following modifications (Rogers and Rogers, 2008): Cells, transiently transfected with pAFMW-CG1091B were fixed in 4% PFA (ThermoScientific) in 1 x PBS for 15 min at room temperature. Blocking was performed with 1% BSA (AppliChem) solution in 1 x PBST (containing 0.1% Triton X-100) for 1-hrn at room temperature. Anti-Myc antibody (M4439, Sigma) was used at 1:2,000 dilution and incubated overnight at 4°C. Secondary Alexa Fluor 488 goat anti-mouse IgG antibody (Thermo Fischer Scientific) was used at a 1:2,000 dilution and incubated for 1 h at room temperature. All images were acquired on an LSM 780 instrument (Zeiss) and images were processed using Fiji ImageJ v1.49 (Schindelin et al., 2012).

**Generation of RNA substrates**

Pre-miR-1003 substrates were generated by splint-ligation using the oligos miR-1003-3p (WT/A/U/C) and miR-1003-5p+loop (see Table S2). Briefly, 30 pmol miR-1003-3p variants were 5’ radiolabeled using γ-32P-ATP (6000 Ci/mmol, PerkinElmer) and T4 polynucleotide kinase (NEB) and purified using G25 spin columns (GE Healthcare). 15 pmol were directly gel purified on a 15% denaturing PAA gel (single-stranded substrates). 15 pmol were subjected to splint ligation: Annealing was performed after adding 35 pmol miR-1003-5p+loop RNA and 20 pmol miR-1003-DNA splint oligo in 1 x Lysis buffer (30 mM HEPES-KOH pH 7.4, 100 mM KOAc, 2 mM MgOAc) by heating to 70°C for 5 min and slowly cooling to room temperature. Then the equal volume 2 x ligation buffer...
(132mM Tris-HCl pH7.6, 20 mM MgCl₂, 2 mM DTT, 2 mM ATP, 15% PEG-8000) was added together with 1 µl T4 DNA Ligase (2,000U/µl, NEB). Ligation was performed at 25°C for 2 h, followed by addition of 1 µl RQ1 RNase-free DNase (Promega) and incubation for 15 min at 37°C. Ligation product was gel-purified on an 8% denaturing PAA gel.

Pre-miR-184 and uridylated variants were generated by splint-ligation as described above, using 5’-32P-radiolabeled miR-184-5p, miR-184-loop, miR-184-3p or miR-184-3p-U and miR-184-DNA-splint oligonucleotide.

Pre-bantam was generated by T7 transcription, as described previously (Fukunaga et al., 2012).

**Expression and purification of tagged proteins**

For lysate preparation a confluent 75 cm² flask of S2 cells stably expressing FLAG-Tailor was harvested by centrifugation and the pellet washed twice in 1 x PBS. The pellet was resuspended in 2 cell pellet volumes 1 x Lysis IP buffer (30 mM HEPES-KOH pH 7.4, 100 mM KOAc, 2 mM MgOAc, 0.5 % NP40, 5 mM DTT, complete ETDA-free proteinase inhibitor cocktail [Roche]) and cells were disrupted with ~50 strokes of a Dounce homogenizer using a “B” pestle followed by centrifugation at 18,000 x g for 30 min at 4°C.

Immunoprecipitation of FLAG-Tailor was performed using Protein G Dynabeads (LifeTechnologies) and FLAG M2 monoclonal antibody (Sigma) using the following protocol: The desired amount of Protein G dynabeads is added to a 1.5 ml tube (e.g. for 400 µl lysate [~ 4-8 mg total protein] we used 300 µl of
resuspended Protein G Dynabeads). Beads were washed 3 x in 1 ml 1 x Lysis-IP buffer (see above). Beads were resuspended in 400 μl Lysis-IP buffer and the desired amount of FLAG M2 monoconal antibody (e.g. for 400 μl lysate 20 μl antibody was used) was added and the tubes rotated for 1-2 h at room temperature. Antibody-coupled beads were washed 3-times with 1 ml 1 x Lysis-IP buffer. Lysate was diluted 2-fold by adding 1 x lysis buffer and added to antibody coupled beads and rotated for > 4 h at 4°C. Beads were captured on magnet and washed extensively (> 5 times) with 1 x Lysis-IP buffer containing 500 mM NaCl (1 ml each wash step) followed by an additional 5 wash steps with 1 x Lysis-IP buffer without extra salt. Beads were resuspended in 50 μl FLAG-Elution buffer (1 x Lysis-IP buffer containing 5% Glycerol and 4 mg/ml 3 x FLAG peptide; pH adjusted to 7.4) and rotated for 30 min at 4°C. Elution step was repeated 3 times. Eluted FLAG-Tailor was frozen in liquid nitrogen in 10 μl aliquots, stored at -80 °C and used only once after thawing.

Streptavidine-binding-peptide (SBP)-tagged Dcr-1 was expressed in S2 cells and purified as described with some modifications (Tsutsumi et al., 2011): Briefly, S2 cells stably expressing SBP-Dcr-1 were lysed two times in two cell-pellet volumes of 1 x Lysis-IP buffer, added to 40 μl Streptavidin Sepharose High Performance (GE healthcare) and rotated at 4°C for 3 h. Beads were washed five times with 1 x Lysis-IP buffer containing 0.5 M NaCl, and three times with 1 x Lysis-IP buffer. SBP-Dcr-1 was eluted two times in 50 μl 1 x Lysis-IP buffer containing 5 mM biotin, pH 8.0, 50% Glycerol, and 0.01% BSA.
In vitro tailing assays

In vitro tailing assays were performed using 10 nM RNA and 1 µl immunopurified FLAG-Tailor under the conditions used for in vitro RNAi assays (Haley et al., 2003; Ameres et al., 2010) except for omitting the ATP-regeneration system and ATP. rNTPs were added as indicated at a final concentration of 500 µM. A typical tailing reaction had a total volume of 10 µl consisting of 1 µl 100 nM RNA, 1 µl H2O, 3 µl 40 x reaction mix (20 mM DTT, 1.67 mM UTP, 160 mM KOAc, 11.2 mM MgAc), 4 µl 1 x Lysis buffer (30 mM HEPES-KOH pH 7.4, 100 mM KOAc, 2 mM MgOAc, 5 mM DTT) and 1 µl immunopurified FLAG-Tailor. Reactions were carried out at 25°C, aliquots were removed from the reaction at the indicated timepoints (if not indicated otherwise, reaction was incubated for 5 min) and stopped in 5 µl Formamide loading buffer (95% Formamide, 18 mM EDTA, 0.025% SDS, Xylene Cyanol and Bromphenol Blue; Ambion). Samples were boiled at 95°C for 5 min and separated on a 15 % (for single-stranded substrates) or 12 % sequencing (for pre-miRNA substrates) denaturing PAA gel, dried, exposed to storage phosphor screen (PerkinElmer), imaged on a Typhoon TRIO variable mode imager (Amersham Biosciences), and quantified using ImageQuant TL v7.0 (GE Healthcare).

In vitro dicing assays

Freshly purified, SBP-tagged Dcr-1 was incubated with 10 nM 5’-32P-radiolabeled pre-miRNA under standard RNAi conditions (Haley et al., 2003; Ameres et al., 2010) for the indicated times, as described previously (Han et al., 2011). Products
were resolved on a 15 % denaturing PAA gel, dried, exposed to storage phosphor screens (Perkin-Elmer), imaged on a Typhoon TRIO variable mode imager (Amersham Biosciences), and quantified using ImageQuant TL v7.0 (GE Healthcare). Dicing rates were determined by burst kinetics, i.e. linear fitting to early timepoints of individual replicates and normalized to the indicated control samples.

**Small-scale pre-miRNAs cloning**

20 μg total RNA was size selected (40-100 nt) on a 12% denaturing PAA gel and subjected to 3´ adapter ligation using T4-RNL21-249, K227Q, followed by reverse transcription using Superscript III (LifeTechnologies) and PCR amplified (2 min 95°C; [30 sec 95°C, 30 sec 56°C, 15 sec 72°C]x25; 15 min 72°C) using pre-miRNA-184-fwd or pre-bantam-fwd together with 3´Linker-rev and Taq polymerase (see Table S2 for primer sequences). PCR products were gel purified and subjected to TOPO-TA cloning. ~100 colonies were picked, plasmids were prepared and subjected to Sanger sequencing. Mixed clones were excluded from the analysis.

**Library Construction for High-Throughput Sequencing**

*Small RNA and pre-miRNA libraries:*

Libraries were constructed as described in previously with the following modifications (Ameres et al., 2010): Briefly, 20 μg total RNA were resolved on a 15 %denaturing polyacrylamide gel and 18-30 nt (small RNA sequencing) or 40-100 nt (pre-miRNA sequencing) RNAs were extracted and eluted from the gel. Small RNA samples were depleted of 2SrRNA as described (Seitz et al., 2008).
Small RNAs were ligated to 3′ and 5′ adapters containing 4 random nucleotides at the ligation interface to minimize ligation bias (Jayaprakash et al., 2011) (see Table S2 for adapter sequences). Ligation products were reverse transcribed using SuperScript III Reverse Transcriptase (Invitrogen) and cDNA samples were PCR amplified using KAPA Real-Time Library Amplification Kit (PeqLab). Amplified cDNA was purified on 2 % agarose gels, followed by library quality control and high-throughput sequencing on a HiSeq 2000 instrument (Illumina) (SR50 for small RNA libraries and SR100 for pre-miRNA libraries) performed by the CSF NextGen Sequencing facility.

*Libraries of in vitro tailing assays:*

The substrate RNA used in the in vitro tailing assay was identical to the 5′ adaptor used in small RNA library constructions (for sequence see Table S2). Upon in vitro tailing assays, the product was gel-purified and subjected to 3′ adapter ligation, as described for small RNA cloning (see above). The product was gel purified and reverse transcribed using SuperScript III Reverse Transcriptase (Invitrogen) and cDNA samples were PCR amplified using KAPA Real-Time Library Amplification Kit (PeqLab), followed by SR50 sequencing on an HiSeq 2000 instrument (Illumina) performed by the CSF NextGen Sequencing facility.

*Libraries of poly-adenylated mRNA:*

Three independent biological replicates of total RNA from untreated *Drosophila* S2 cells or S2 cells stably depleted of Tailor/CG1091 by CRISPR/Cas9 genome
engineering were subjected to polyA selection using the NEBNext Poly(A) mRNA Magnetic Isolation Module (New England Biolabs), followed by library preparation using the NEBNext Ultra Directional RNA Library Prep Kit for Illumina (New England Biolabs), according to the instructions of the manufacturer. Library preparation and sequencing was performed at the CSF NGS facility (http://www.csf.ac.at/facilities/next-generation-sequencing).

**Bioinformatics analyses**

Small RNA- and pre-miRNA-library reads were recovered by adapter clipping: The adapter was cut once with cutadapt v1.2.1 (Martin, 2011). The random 4mers on the 5’and 3’ ends of recovered reads were removed with fastx_trimmer from the fastx-toolkit v0.0.13 (http://hannonlab.cshl.edu/ fastx_toolkit/). Processed reads were size-selected (i.e. 18-30 nt for small RNA libraries and >40 nt for pre-miRNA libraries). The algorithm for extracting genome-matching and prefix-matching reads for small RNA- and pre-miRNA-sequencing datasets was essentially as described with the following exceptions (Ameres et al., 2010): The original data structure implemented (i.e., the suffix tree) was replaced with a compressed index based on the Burrow-Wheeler transform, namely the FM index (Ferragina and Manzini, 2005). The FM-index data structure was used to find the occurrence of the exact and near-exact substrings for the prefix of reads. Once a set of the possible occurrence of the prefix was located, all their suffixes were checked for non-templated nucleotides. This was achieved by backtracking to the previous position according to the FM-index to efficiently enumerated all these prefix matches. A filtering step to remove reads with poor Phred score at
the 3’ end was applied to gain more confidence on the tails identified. Note that alignments with mismatches in the seed and the middle portion were discarded. BEDtools (Quinlan and Hall, 2010) was used to annotated miRNAs according to the coordinates of the occurrences. The statistic summary of tails of each miRNA species was generated by shell scripts (detailed method will be published elsewhere). To counteract internal mismatches, aligned reads were filtered out if their ratios of tail length to sequence length were above 0.12 for miRNAs and 0.05 for pre-miRNA. For analysis of pre-miRNA sequencing (Fig. 5A-D), datasets derived from untreated and control dsRNA treated (i.e. dsGFP and dsLUC) samples were pooled. For analysis of CG1091-depletion in vivo, the two different CG1091c4 alleles, which produced similar phenotypes (see Fig. S2B), were merged in Fig. 2F to I. For small RNA sequencing datasets only miRNAs with an abundance of >100 ppm were considered.

For in vitro tailing libraries, 3’ adaptor sequences were clipped (see above) and only reads that contained ≥ 4 nt remaining sequence and no ambiguous nucleotide were considered further (see Fig. S6H for sequencing statistics). In the remaining sequence the first 4 nt were considered as substrates and any following nucleotides represented the tail. Since the input sample was not expected to contain tailed substrates it was used to train a model of maximal fractions of tail contamination (see Fig. S5H). The total counts (c) of each substrate (s) and the counts of a substrate with a specific tail of length (l) greater than zero nt were computed. The maximal tailing fraction (f_{reference}) over all substrates was computed for each tail length:
The fractions were used to correct the counts \( (c') \) in the samples treated with tailing enzyme. It was calculated by subtracting the estimated maximal contamination of a certain tail length from the observed counts:

\[
c'_{l,s} = \begin{cases} 
0, & \text{integer}(c_{l,s} - c_s \times f_{\text{reference},l}) < 0 \\
\text{integer}(c_{l,s} - c_s \times f_{\text{reference},l}) & \text{otherwise}
\end{cases}
\]

Analysis of poly(A)-selected total RNA libraries was performed as follows: The strand specific reads were screened for ribosomal RNA by aligning with BWA (v0.6.1) against known rRNA sequences (RefSeq) (Li and Durbin, 2009). The rRNA subtracted reads were aligned with TopHat (v1.4.1) against the Drosophila melanogaster genome (FlyBase r5.44) and a maximum of 6 mismatches (Trapnell et al., 2009). Introns between 20-150,000 bp were allowed, based on FlyBase statistics. Maximum multi-hits was set to 1 and InDels, as well as Microexon-search was enabled. Additionally, a gene model was provided as GTF (FlyBase r5.44). snRNA, rRNA, tRNA, snoRNA and pseudogenes were masked for downstream analysis. Aligned reads were subjected to FPKM estimation with Cufflinks (v1.3.0) (Trapnell et al., 2010; Roberts et al., 2011). At this step, bias detection and correction was performed. Furthermore, only those fragments compatible with FlyBase annotation (r5.44) were allowed and counted towards the number of mapped hits used in the FPKM denominator. The aligned reads were counted with HTSeq (0.6.1p1) and the polyA containing transcripts were subjected to differential expression analysis with DESeq (v1.10.1) (Anders and...
Huber, 2010). Only mRNAs with an average abundance of >1 FPKM in untreated S2 cell datasets were considered in the analysis (a total of 8,444 mRNAs). Predicted miRNA targets were derived from targetscan database (http://www.targetscan.org/fly_12/) (Ruby et al., 2007).

For evolutionary conservation analysis of pre-miRNA 3´ termini, pre-miRNAs annotated with high confidence in mirBase v21 (Kozomara and Griffiths-Jones, 2014) were extracted and manually curated based on pre-miRNA cloning datasets or the most-abundant annotated miRNA-3p species. The nucleotide composition of the Drosophila melanogaster genome was calculated based on flybase genome-annotation r5 excluding chromosome MT, U, and Uextra.
**Supplemental References**

Ameres, S. L. et al. (2010). Target RNA-directed trimming and tailing of small silencing RNAs. Science 328, 1534-1539.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol 11, R106.

Bassett, A. R. et al. (2014). Mutagenesis and homologous recombination in Drosophila cell lines using CRISPR/Cas9. Biology Open 3, 42-49.

Dietzl, G. et al. (2007). A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. Nature 448, 151-156.

Ferragina, P., and Manzini, G. (2005). Indexing compressed text. Journal of the ACM 52, 552-581.

Fukunaga, R. et al. (2012). Dicer Partner Proteins Tune the Length of Mature miRNAs in Flies and Mammals. Cell 151, 533-546.

Gokcezade, J. et al. (2014). Efficient CRISPR/Cas9 Plasmids for Rapid and Versatile Genome Editing in Drosophila. G3 4, 2279-2282.

Gruber, A. R. et al. (2008). The Vienna RNA websuite. Nucleic Acids Res 36, W70-W74.

Haley, B. et al. (2003). In vitro analysis of RNA interference in *Drosophila melanogaster*. Methods 30, 330-336.

Han, B. W. et al. (2011). The 3′-to-5′ exoribonuclease Nibbler shapes the 3′ ends of microRNAs bound to *Drosophila* Argonaute1. Curr Biol 21, 1878-1887.

Jayaprakash, A. D. et al. (2011). Identification and remediation of biases in the activity of RNA ligases in small-RNA deep sequencing. Nucleic Acids Res 39, e141.

Kozomara, A., and Griffiths-Jones, S. (2014). miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res 42, D68-D73.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754-1760.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. journal 17, pp. 10-pp. 12.

Mohammed, J. et al. (2014). Diverse modes of evolutionary emergence and flux of conserved microRNA clusters. RNA 20, 1850-1863.

Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841-842.

Roberts, A. et al. (2011). Improving RNA-Seq expression estimates by correcting for fragment bias. Genome Biology 12, R22.

Rogers, S. L., and Rogers, G. C. (2008). Culture of Drosophila S2 cells and their use for RNAi-mediated loss-of-function studies and immunofluorescence microscopy. Nature Protocols 3, 606-611.
Ruby, J. G. et al. (2007). Evolution, biogenesis, expression, and target predictions of a substantially expanded set of *Drosophila* microRNAs. Genome Res 17, 1850-1864.

Schindelin, J. et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat Methods 9, 676-682.

Seitz, H. et al. (2008). Argonaute loading improves the 5’ precision of both microRNAs and their miRNA* strands in flies. Curr Biol 18, 147-151.

Trapnell, C. et al. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol 28, 511-515.

Trapnell, C. et al. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25, 1105-1111.

Tsutsumi, A. et al. (2011). Recognition of the pre-miRNA structure by *Drosophila* Dicer-1. Nat Struct Mol Biol 18, 1153-1158.