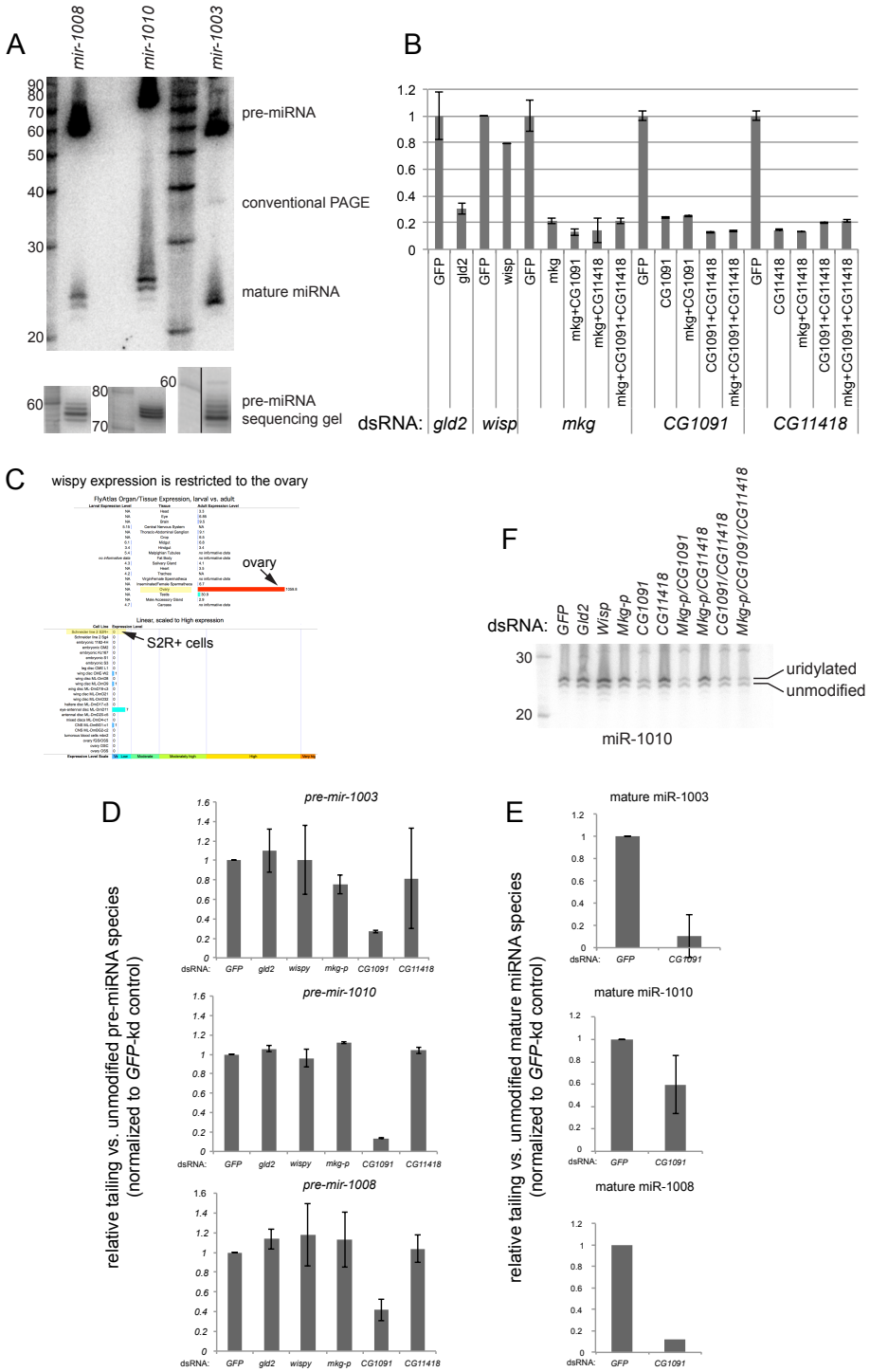

Figure S1. Additional analysis of mirtron uridylation (relevant to Figure 1).
(A) S2 cells were transfected with expression constructs for the mirtrons *mir-1003*, *mir-1008*, *mir-1010*, and total RNAs were analyzed by Northern blotting to resolve mature and pre-miRNA products. Both analyses revealed a series of bands. Especially for pre-miRNA hairpins, for which mirtron sequences are defined by splicing and therefore should not be alternative, the existence of multiple species implies the existence of terminal modifications to mirtron hairpins.
(B) qPCR validation of TUT depletions in single, double, or triple knockdown conditions. All of the cognate TUT transcripts were strongly depleted, even when other dsRNAs were included in the mixture. The exception was w*ispy;* however, its signals were not confidently detected, as they appeared only ~cycle 37.
(C) Inspection of FlyAtlas and modENCODE data show that *wispy* is endogenously expressed basically only in the ovary, and is not present in S2 cells, thus explaining failure to detect it or its depletion.
(D) Quantification of replicate sequencing gel Northern blot experiments shows that depletion of CG1091 specifically and reproducibly causes loss of tailed hairpin species for these three mirtrons tested. Error bars are shown as SD.
(E) Quantification of replicate Northern blot experiments shows that depletion of CG1091 reproducibly causes loss of tailed mature miRNA species for these three mirtrons tested. Error bars are shown as SD.
(F) Single, double and triple knockdowns of various TUTs in S2 cells transfected with *mir-1010* expression construct, and tested for miR-1010 expression. Only those sample depleted for CG1091 showed loss of the upper miR-1010 hybridized band, and no synergistic effects between CG1091 and other TUTs were detected..

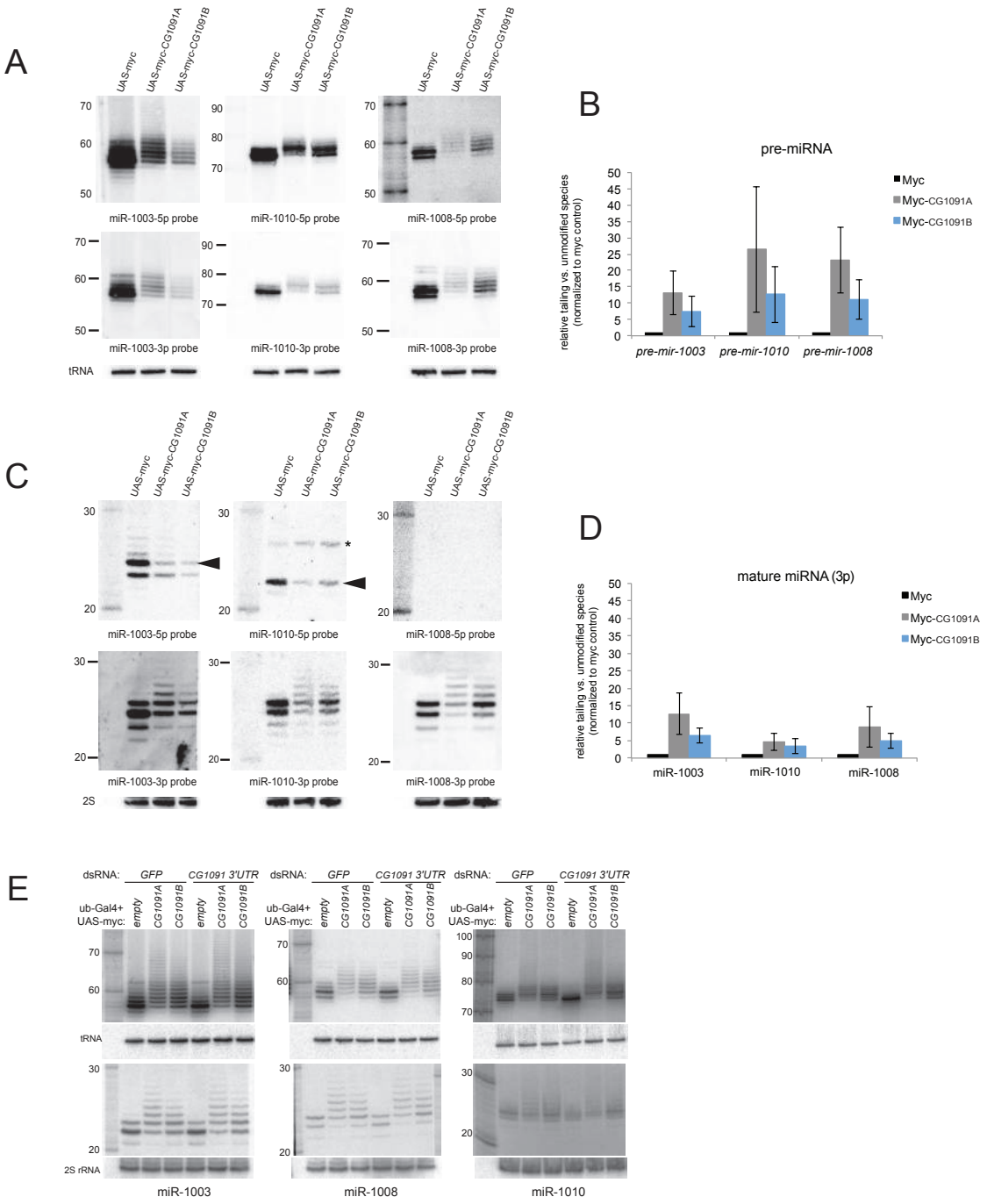

Figure S2. Gain-of-function analyses of CG1091/Tailor (relevant to Figure 2).
(A) S2 cells were cotransfected with control or CG1091 expression constructs and the indicated mirtron constructs, and analyzed by Northern blotting. Blots were sequentially probed for mirtron-5p, mirtron-3p and control tRNA. The mirtron-5p and mirtron-3p probes detect comparable amounts of ectopic tailing induced by both CG1091 isoforms.
(B) Quantification of ectopic hairpin tailing induced by transfected CG1091 constructs. The ratios of tailed to unmodified species in the presence of ectopic CG1091 were normalized to parallel myc control conditions. Error bars are shown as SD.
(C) RNA samples generated and analyzed as in (A), but analyzed for the mature miRNA population. In contrast to the hairpin population, the mature mirtron-3p species are tailed but mature mirtron-5p species are not (arrowheads). Asterisk on miR-1010-5p denotes an alternate 5p isoform (see main Figure 1), not a tailed species, and its level is not enhanced by ectopic CG1091. miR-1008-5p was not detected, presumably due to strong strand asymmetry observed in small RNA sequencing (see Figure 1B). Note that the mirtron-3p probing data are shown in main Figure 2, but are shown here with the stripped and reprobed mirtron-5p blots for direct comparison.
(D) Quantification of ectopic tailing on mature mirtron-3p species miRNAs induced by transfected CG1091 constructs. Overall, we observe a lesser degree of tailing than in the hairpin population. Error bars are shown as SD.
(E) Activity of ectopic CG1091 does not depend on endogenous CG1091. S2 cells were treated with GFP dsRNA or CG1091 3'UTR dsRNA, and then transfected with control or CG1091 expression constructs. The efficacy of CG1091 depletion can be observed by the loss of tailed species in the corresponding control myc transfections of GFP/CG1091 knockdown cells. Both CG1091-A and CG1091-B can induce ectopic mirtron hairpin and mature mirtron-3p tailing in CG1091-depleted cells.

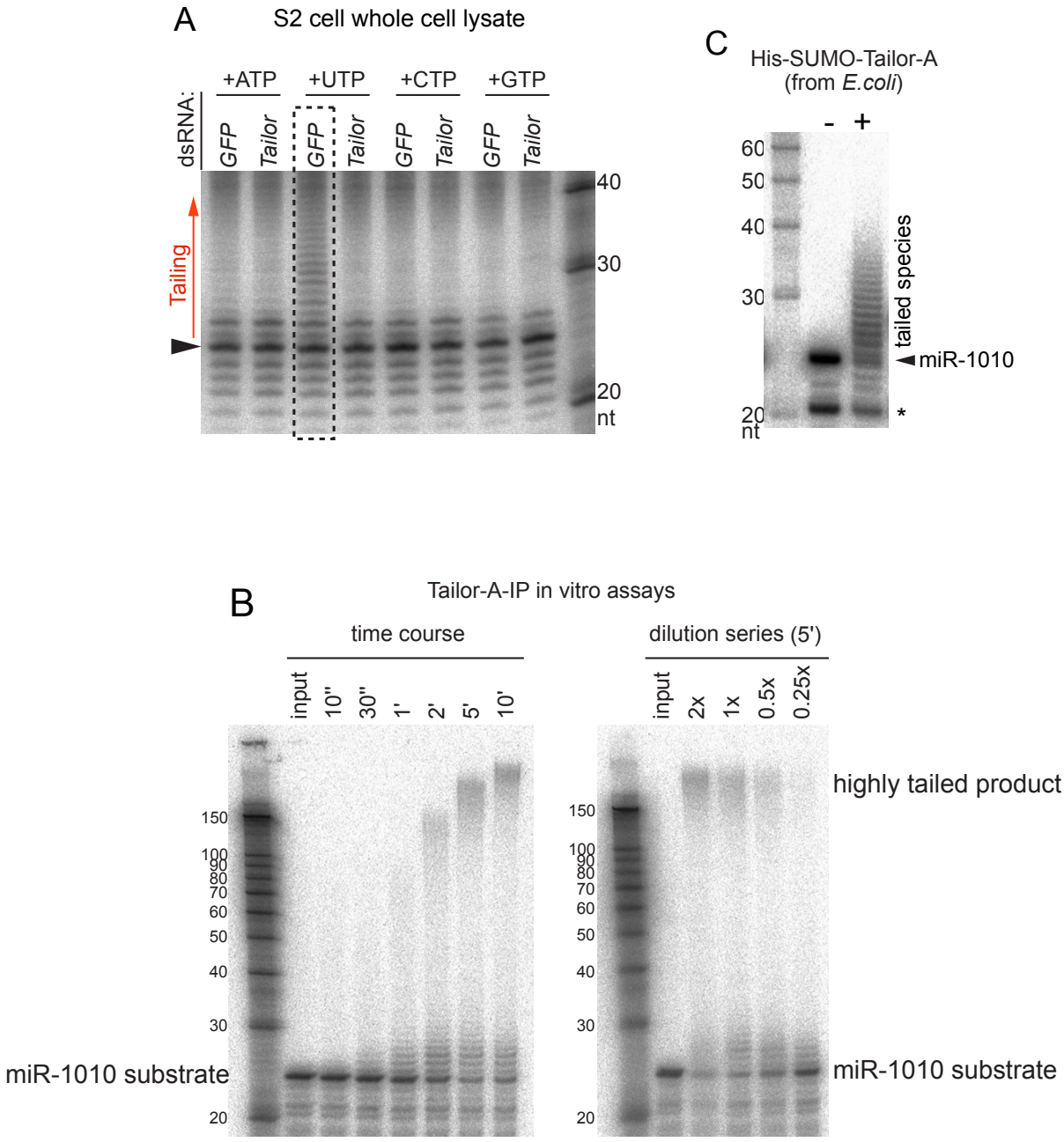

Figure S3. In vitro uridylation assays of Tailor activity (relevant to Figure 3).
(A) Lysates were prepared from S2 cells treated with *GFP* or *Tailor* dsRNAs, and tailing was performed using a radiolabeled 24nt miR-1010 substrate and one of the four ribonucleotides. The shorter species represent degraded RNAs that were evident in all incubations with S2 cell lysate. Tailed species were seen in GFP-dsRNA cells supplemented with UTP but not other nucleotides. The tailed species were not seen in a parallel reaction using lysate treated with *Tailor*-dsRNA.
(B) In vitro uridylation assays using immunoprecipitated myc-Tailor-A and radiolabeled miR-1010 substrate (Left) Time course assay showing the rapid appearance of highly tailed miR-1010 substrates within a few minutes. (Right) Dilution series using various amounts of immunoprecipitated material shows that with less CG1091, less of the substrate was modified. However, the substrate that was modified still acquired relatively similarly long tails.
(C) Recombinant Tailor-A prepared from *E. coli* exhibits tailing activity on miR-1010. Asterisk points to a degradation product of the substrate that is present in the input.

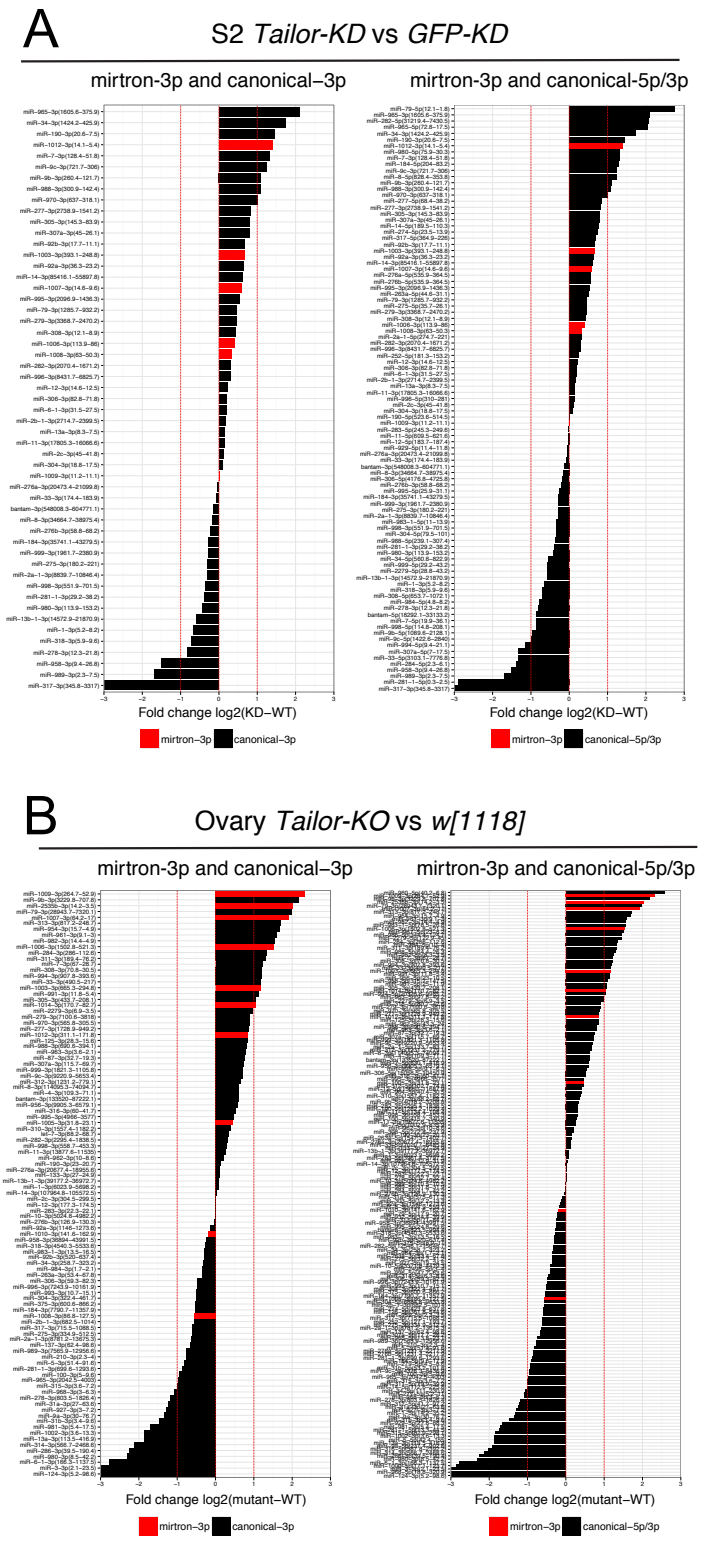

Figure S4. Effects of Tailor on miRNA accumulation (relevant to Figure 5).
(A) We determined the normalized levels of various miRNA species from small RNA libraries from S2 cells treated with Tailor or GFP dsRNA, and plotted their expression differences between these conditions. The left graph depicts mirtron-3p and canonical miRNA-3p species, while the right graph also includes canonical miRNA-5p species. In both cases, there is biased behavior of most mirtron-3p species to be upregulated upon depletion of Tailor. In contrast, there is no directional behavior in the overall accumulation of canonical miRNA species.
(B) Similar analysis as above, using small RNA libraries from Tailor-KO and w[1118] ovaries. We observe overall similar trends for selective increases in mirtron-3p accumulation in the absence of Tailor. Since more mirtrons are expressed in ovaries than in S2 cells, the directional shift of mirtrons is more apparent in these datasets.

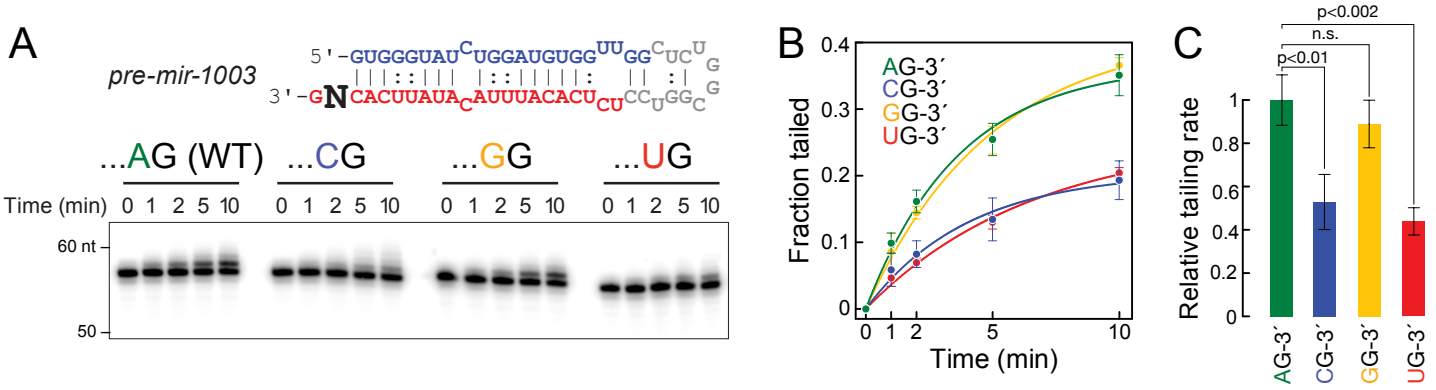

Figure S5. Effect of penultimate 3' nt identity on Tailor-mediated modification (relevant to Figure 6).
(A) In vitro tailing assays on pre-mir-1003 variants bearing the different possible penultimate 3' nts (NG-3'). We used this substrate instead of pre-mir-1010 used in other assays because it is not as robustly modified, a feature that enhanced our ability to visualize minor differences in substrate preference.
(B) Analysis of replicate tailing experiments shows that AG-3' and GG-3' substrates were preferred substrates. Errors include standard deviations.
(C) Relative tailing rates are influenced by the penultimate nucleotide identity in the order AG>GG>>C-G>UG, although the differences between AG and GG were not statistically significant as with the other variants. Tailing efficiency was determined by burst kinetics (i.e. linear regression to early timepoints; 1 and 2 min) and normalized to AG-3´ (WT) substrate. Each of three replicates was independently analyzed and mean values and standard deviations are depicted.

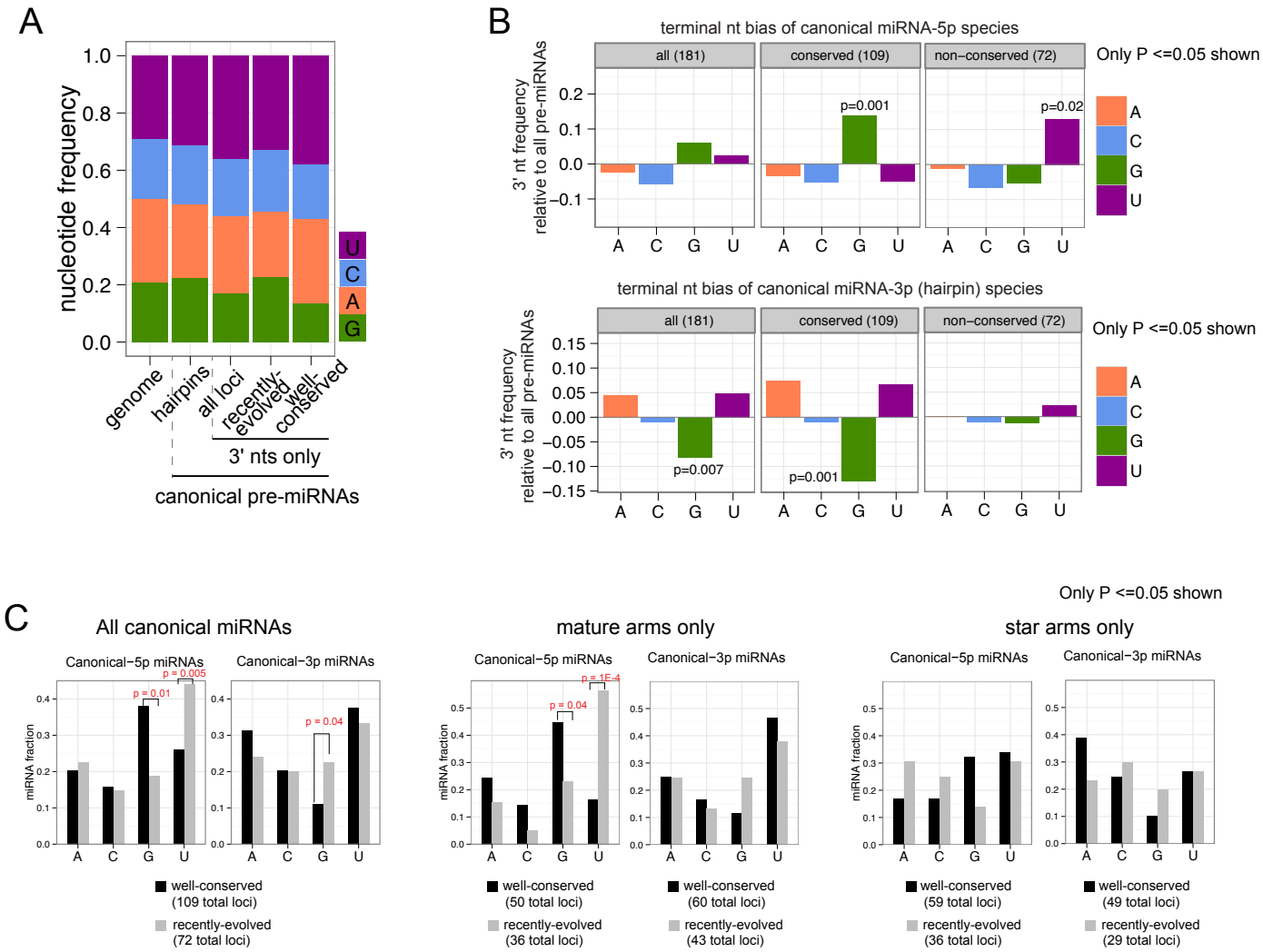

Figure S6. 3' terminal properties of canonical miRNAs over evolution. Relevant to Figure 7.
(A) The nucleotide frequencies across the Drosophila genome, across all canonical pre-miRNA hairpins, and across 3' termini of various subsets of canonical pre-miRNA hairpins (all loci, 72 recently-evolved (melanogaster subgroup-specific) loci, and 109 well-conserved loci (present in all fly species) are shown. There is depletion of 3'-G amongst the well-conserved pre-miRNA hairpins, relative to recently-evolved pre-miRNA hairpins, as well as relative to the total hairpin distribution and genome distribution of nucleotide identities.
(B) Analysis of 3' nt identity bias across various subsets of canonical miRNA species. Plotted are the differences in the test category nucleotide frequency relative to the background nucleotide frequency across all canonical pre-miRNA hairpins. Significance was determined by binomial test. There is significant depletion of 3'-G in conserved canonical miRNA-3p species but not in conserved miRNA-5p species. In fact, there is enrichment of 3'-G amongst conserved miRNA-5p species. Note the right-hand two panels of the 3p species analysis, which also correspond to pre-miRNA hairpin analysis, are also shown in main Figure 7F.
(C) Direct comparison of 3' nt identity bias between well-conserved and recently-evolved Drosophila miRNAs. The direct comparison provides a stronger basis for assessing potential differences in their distribution of terminal nucleotide identities, by application of Fisher's exact test. We analyzed the total age-groups of miRNAs as above, but also divided these into the hairpins for which the mature or the star arm lies on the 5p or 3p arm. The depletion of 3'-G amongst conserved miRNAs is found in both those hairpins for which the 3p arm is the mature or the star product, but the latter category does not reach statistical significance. The trend of increased frequency of 3'-G amongst canonical miRNA-5p species is also mostly due to the behavior of mature arm species. There is also an enrichment of 3'-U amongst recently-evolved canonical miRNA-5p species.

Supplemental Spreadsheets

Table S1, relevant to Figures 1, 5 and 6.

Table S2, relevant to Figure 6.

Table S3, relevant to Figure 6.

Table S4, relevant to Figure 7.

Table S5, relevant to Experimental Procedures

Supplemental Experimental Procedures

Northern blotting

Northern blotting for small RNAs followed previously described probes and procedures (Okamura et al., 2007). In order to obtain singe-nucleotide resolution of pre-miRNA species (50-80nt long), 8 to 15 µg of TRIzol-extracted total RNAs were loaded on 12% polyacrylamide-urea sequencing gels (40cm long). Gel were pre-run in 1x TBE for 30mn at 50mA, and samples were run at 20mA for 20mn then at 50mA for 5 to 6 hours. We transferred them to GeneScreen Plus Hybridization Transfer Membranes (Perkin Elmer) in 0.5x TBE for an hour at 300mA on a semi dry transfer system and then pre-hybridized and hybridized them using standard procedures (Okamura et al., 2007).

Hybrid mirtron expression constructs

Using the scheme and primer sets in **Table S5**, we performed initial amplifications using primer pairs 1+2 and 3+4 and *UAS-dsRed-mir-278* as template (Bejarano et al., 2012) to obtain the 5' and 3' part of the construct, sharing the sequence of the loop region. If no primer "2" is mentioned, mir-278 mir-100X-AG_2 was used to obtain the mutant clone. A second PCR was performed using the gel-purified first-round products as template for primers 1 and 4 only. These products were then cloned into UAS-DsRed using NotI and XbaI sites.

Luciferase sensor assays

For assaying loss-of-function conditions, either *Tailor* or *GFP* control dsRNA was used for two rounds of knockdown for 4 days each at a concentration of 7.5µg dsRNA/10^6 S2 cells. Cells were seeded in 96 well plates after initiating second round of knockdown. After 24 hours of the second round knockdown, cells were transfected with 25ng/well of UAS-mirtron or control UAS-dsRed, 25ng/well of psiCHECK luciferase sensor plasmid, and 12.5ng/well of Ub-Gal4 using Effectene (Qiagen). Three days later, luciferase readings were taken using Dual-Glo Luciferase (Promega).

For assaying gain-of-function conditions, S2 cells were plated in 96 well plates and transfected with 12.5ng/well of either UAS-Myc or UAS-Myc-Tailor, 12.5ng/well of UAS-mirtron or empty UAS-dsRed, 12.5ng/well of psiCHECK plasmid, and 6.25ng/well of Ub-Gal4. Luciferase values were measured 3 days after transfection.

We normalized transfection efficiencies using control firefly luciferase carried within psiCHECK, and fold repression was normalized against empty UAS-DsRed

(pJB38) and psiCHECK empty plasmid. We present representative data from quadruplicate sensor assays, for which each set was performed at least three times and found to yield qualitatively similar results.

Preparation of substrates for tailing and dicing assays

We radiolabeled substrates by incubating 2µl of 1.5µM synthetic miR-1010-3p (UUUCACCUAUCGUUCCAUUUGCAG), or its terminal nt variants (-AA, -AU, -AC), with 3µl 10x PNK buffer (NEB), 3µl of gamma-32P-ATP (6000Ci/mmole, Perkin Elmer) and 1µl of PNK (NEB) in a final volume of 30µl and incubated for an hour at 37ºC. Free gamma-32P-ATP was removed using a G25 column (GE Healthcare) and PNK was inactivated by 5' incubation at 95ºC. To synthesize *pre-mir-1010*, 10µl of radiolabeled miR-1010-3p was mixed with 2µl of 1.5µM non-phosphorylated acceptor (i.e. miR-1010-5P+loop: GUAAGUGGUGUAGAUGAAACAAAUUUACCAACAAUUUUGUUGGAUUG), 2µl of 1.5µM bridge DNA (i.e. sequence complementary to *pre-mir-1010*: CTGCAAATGGAACGATAGGTGAAACAATCCAACAAAATTGTTGGTAAATTTGTTTCATCTACACCACTTAC) and 3µl of PNK buffer in a final volume of 30µl. The mix was heated to 65ºC and then slowly brought to room temperature to anneal the RNA molecules to the DNA bridge. Ligation of the two RNA molecules as performed by incubating the reaction at 25ºC for 3h after addition of 1µl T4 DNA ligase buffer (NEB), 2µl of 10mM ATP and 1µl of T4 DNA ligase in a final volume of 40µl. The DNA bridge was digested with 1µl of DNase I (NEB) for 15' at 37ºC. We purified full-length *pre-mir-1010* or mature miR-1010-3p using a 15% polyacrylamide-urea sequencing gel. RNA was extracted from gel pieces by overnight incubation in 400mM NaCl followed by column filtration.

In vitro tailing assay

To assay uridylation activity in total cell lysates, we incubated labeled miR-1010 in 5µl S2 cell lysates supplemented with 3.2mM $MgCl_2$, 1µl of RNase out (Invitrogen) and 0.25mM of the indicated ribonucleotide in a final volume of 10µl for 3h at room temperature. RNAs were then extracted using phenol chloroform and loaded on 20% polyacrylamide-urea gel.

To assay uridylation using immunoprecipitated proteins, we adapted a published method (Heo et al., 2009). After 3 days of transfection with myc-tagged Tailor, cells were washed with PBS and resuspended in 5 volumes lysis buffer (500mM NaCl, 1mM EDTA,

10mM Tris pH8.0, 1% Triton-X). Cell lysate was obtained by 20' incubation on ice followed by 5 passages through a 25G needle. We cleared the lysate by two 10' centrifugations at 10,000G. The immunopurified proteins were obtained by incubating cleared lysate with beads coated with anti-Myc and rotated for 2h at 4ºC. Beads were washed three times with lysis buffer and 4 times with wash buffer (200mM KCl, 10mM Tris pH8.0, 0.1mM EDTA), then resuspended in wash buffer. 15µl of beads were mixed with 15µl of 2x reaction mix (6.4mM $MgCl_2$, 2mM DTT, 0.5mM rNTP, 1000cpm of radiolabeled RNA substrate) and incubated at 25ºC for the indicated time. Reactions were stopped by addition of 170µl of Stop buffer (20mM Tris pH7.5, 250mM NaCl, 10mM EDTA) and RNAs were extracted with acidic phenol chloroform. RNAs were run on polyacrylamide-urea gels and transferred to GeneScreen Plus® Hybridization Transfer Membranes (Perkin Elmer).

To assay terminal nucleotide preferences, we utilized a stable S2 cell line bearing an actin-Flag-Myc-Tailor-B (FM-Tailor-B) expression construct. Immunoprecipitates of these Tailor complexes were less active than those recovered from transient transfections, and were better suited to observe substrate preferences. From one confluent 75 cm2 flask, cells were homogenized in 2 cell pellet volumes of 1x 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 FM-Tailor-B was immunoprecipitated using Protein G dynabeads (LifeTechnologies) and FLAG M2 monoclonal antibody (Sigma). FM-Tailor-B was recovered by 4x50 µl elutions in 1x Lysis-IP buffer containing 5% Glycerol and 4 mg/ml 3 x FLAG peptide, and flash-frozen in 10 µl aliquots. In vitro tailing assays were performed using 10 nM radiolabeled mature or pre-miRNA substrate (described above) and 1 µl FM-CG1091B. We used conditions described for in vitro RNAi assays (Ameres et al., 2010) except that the ATP-regeneration system and ATP were omitted, and UTP was added to a final concentration of 500 µM, the approximate physiological cellular concentration of UTP (Traut, 1994).

In vitro dicing assay

These assays utilized radiolabeled *pre-mir-1010*, *pre-mir-1010+2U* (2nt uridine tail at 3' end) and *pre-mir-1010+4U* (4nt uridine tail at 3' end) substrates prepared as described above. For preparation of recombinant Dcr-1 and Loqs-PB proteins, Sf9 cells were infected with the recombinant Baculovirus encoding His-tagged Dcr-1 or Loqs-PB (Jiang et al., 2005). Proteins were purified according to the previously published protocol

(Ye and Liu, 2008). In brief, infected cells were lysed by a Dounce homogenizer in hypotonic buffer (10 mM KOAc, 10 mM HEPES, pH 7.4, 2 mM Mg(OAc)2 , 5 mM β -mercaptoethanol) and the protein was purified by Nickel column chromatography followed by SP- and Q-sepharose columns. Fractions were checked by SDS-PAGE and Coomassie blue staining, and the fractions having high protein concentrations were pooled. 10% glycerol was added to the pooled protein samples and small aliquots were stored at -80C.

Small RNA analysis

After removing 3' adaptor sequences with cutadapt (Martin, 2011), we mapped small RNA reads to the *D. melanogaster* genome release 5.3. Unmapped reads were iteratively trimmed one nucleotide each iteration retaining a read length of >=17nt, and then mapped to the genome using Bowtie with no mismatches, up to 30 iterations. The reads were required to match to the defined small RNA sequences with at least 15 nt overlap, and within 2 nt of the 5' end. Trimmed reads were required to match within 4 nt of the defined 3' end; with no restriction on 3' end of tailed reads in order to analyze tailing events. The trimmed reads were not considered for the tailing analysis.

Mirtron hairpin boundaries were defined on the basis of intron splice donor (GU) and acceptor (AG) sequences. The boundaries of 85 high-confidence Drosha/Dicer product duplexes were defined previously (Westholm et al., 2012), on the basis of invariant 5'-nts, 2-nt duplex 3' overhangs, and moR reads. Additional canonical pre-miRNA hairpins were based on miRBase v19 with sequences defined by the most abundant matching read from over >300 aggregated small RNA libraries. For TE-siRNA, cis-NAT-siRNA, piRNA, and hpRNA loci, we used our recently reported annotations (Wen et al., 2014), but with sequences defined by the reads aggregated over >300 small RNA libraries with $\geq$1000 clones for TE-siRNAs and with $\geq$50 clones for other classes of sRNAs. snoRNA and tRNA annotations were obtained from FlyBase version FB2014_04.

To quantify mononucleotide tailings, we calculated the percentage of reads that extended exactly one nucleotide beyond the defined sRNA sequence, relative to the total number of reads (mononucleotide tailed reads and unmodified reads). Mirtrons and canonical miRNAs with $\geq$10 reads in both control and *Tailor* and libraries were included. We normalized miRNAs by the total mapped miRNA reads and the trimmed mean of M-values (TMM) normalization method in the edgeR/Limma Bioconductor library (Oshlack et al., 2010).

Supplementary References

Ameres, S.L., Horwich, M.D., Hung, J.H., Xu, J., Ghildiyal, M., Weng, Z., and Zamore, P.D. (2010). Target RNA-directed trimming and tailing of small silencing RNAs. Science *328*, 1534-1539.

Bejarano, F., Bortolamiol-Becet, D., Dai, Q., Sun, K., Saj, A., Chou, Y.T., Raleigh, D.R., Kim, K., Ni, J.Q., Duan, H.*, et al.* (2012). A genome-wide transgenic resource for conditional expression of Drosophila microRNAs. Development *139*, 2821-2831.

Heo, I., Joo, C., Kim, Y.K., Ha, M., Yoon, M.J., Cho, J., Yeom, K.H., Han, J., and Kim, V.N. (2009). TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. Cell *138*, 696-708.

Jiang, F., Ye, X., Liu, X., Fincher, L., McKearin, D., and Liu, Q. (2005). Dicer-1 and R3D1-L catalyze microRNA maturation in Drosophila. Genes & development *19*, 1674-1679.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. Embnet.journal *7*, 10-12.

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.

Oshlack, A., Robinson, M.D., and Young, M.D. (2010). From RNA-seq reads to differential expression results. Genome biology *11*, 220.

Traut, T.W. (1994). Physiological concentrations of purines and pyrimidines. Molecular and cellular biochemistry *140*, 1-22.

Wen, J., Mohammed, J., Bortolamiol-Becet, D., Tsai, H., Robine, N., Westholm, J.O., Ladewig, E., Dai, Q., Okamura, K., Flynt, A.S.*, et al.* (2014). Diversity of miRNAs, siRNAs and piRNAs across 25 Drosophila cell lines. Genome research *24*, 1236-1250.

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.

Ye, X., and Liu, Q. (2008). Expression, purification, and analysis of recombinant Drosophila Dicer-1 and Dicer-2 enzymes. Methods in molecular biology *442*, 11-27.