Published in final edited form as:
Nature. 2008 June 5; 453(7196): 803–806. doi:10.1038/nature07015.

The *Drosophila* hairpin RNA pathway generates endogenous short interfering RNAs

Katsutomo Okamura[1], **Wei-Jen Chung**[1], **J. Graham Ruby**[2], **Huili Guo**[2], **David P. Bartel**[2], and **Eric C. Lai**[1]

[1]Sloan-Kettering Institute, Department of Developmental Biology, 521 Rockefeller Research Laboratories, 1275 York Avenue, Box 252, New York, New York 10065, USA

[2]Howard Hughes Medical Institute and Department of Biology, Massachusetts Institute of Technology, and Whitehead Institute for Biomedical Research, Cambridge, Massachusetts 02142, USA

Abstract

In contrast to microRNAs and Piwi-associated RNAs, short interfering RNAs (siRNAs) are seemingly dispensable for host-directed gene regulation in *Drosophila*. This notion is based on the fact that mutants lacking the core siRNA-generating enzyme Dicer-2 or the predominant siRNA effector Argonaute 2 are viable, fertile and of relatively normal morphology[1,2]. Moreover, endogenous *Drosophila* siRNAs have not yet been identified. Here we report that siRNAs derived from long hairpin RNA genes (hpRNAs) programme Slicer complexes that can repress endogenous target transcripts. The *Drosophila* hpRNA pathway is a hybrid mechanism that combines canonical RNA interference factors (Dicer-2, Hen1 (known as CG12367) and Argonaute 2) with a canonical microRNA factor (Loquacious) to generate ~21-nucleotide siRNAs. These novel regulatory RNAs reveal unexpected complexity in the sorting of small RNAs, and open a window onto the biological usage of endogenous RNA interference in *Drosophila*.

Artificial, long-inverted repeat transcripts are efficiently processed by a Dicer-2 (Dcr-2)/Argonaute 2 (AGO2)-driven RNA interference (RNAi) pathway in transgenic *Drosophila* [1,3]. We hypothesized that this might reflect the existence of an endogenous pathway that accepts long, inverted repeat transcripts. To test this idea, we searched for inverted repeats using EINVERTED[4] and selected putative hairpins containing mapped small RNA reads (see Methods). Out of 8,132 candidate regions, most consisted of the terminal inverted repeats of individual transposable elements or long terminal repeats of tandem inverted transposable elements. The remaining loci corresponded to inverted tandem duplications of messenger RNA- or transfer RNA-encoding genes, a microRNA (miRNA) gene (*mir-997*), a novel tandem pair of short hairpins (*chou39-1* and *chou39-2*, Supplementary Fig. 1), and a handful of single-gene annotations and unannotated regions.

We analysed the size distribution of cloned RNAs from all of the non-transposable-element EINVERTED hits. Although these mostly exhibited a broad length distribution across the ~18–

Correspondence and requests for materials should be addressed to E.C.L. (Iaie@mskcc.org).

Author Contributions J.G.R. identified hp-CG4068 and hpRNA1. W.-J.C. performed the EINVERTED analysis and identified the additional hpRNA loci and their targets. H.G. performed initial hpRNA northern analysis; all other experiments were designed and carried out by K.O. All authors contributed to the preparation of the manuscript.

Author Information The imaginal disc/brain sample described in paper has been deposited in the NCBI GEO under accession number GSM275691. Reprints and permissions information is available at www.nature.com/reprints

26-nucleotide cloning range, indicative of degradation fragments (Supplementary Fig. 1b), seven genomic regions specifically generated 21–22-nucleotide RNAs (Supplementary Fig. 1c–f). These included genes annotated as CG18854, CG32207, CR32205 and pncr009 (also known as pncr009:3L), a series of 20 repeats that partially overlap the 3′ untranslated region (UTR) of CG4068, and an intergenic region adjacent to CG4770 (Supplementary Figs 2–7). Except for CG4068, the coding potential of all of these loci is limited[5,6]. Still, we chose to introduce an 'hp' prefix to these six loci to distinguish the small-RNA-generating hairpins ('hpRNAs') from the potential protein-encoding segments of these transcripts.

The hpRNA hairpins were collectively much longer than typical animal pre-miRNAs, and several were even longer than plant miRNAs[7]. All hpRNA loci produced dominant small RNAs that presented duplexes with 2-nucleotide 3′ overhangs, implying RNase III processing (Fig. 1 and Supplementary Figs 2–7). Such an origin was more evident with the hp-CG18854 and hp-CG4068 hairpins, from which many consecutive, phased, small RNA duplexes were cloned (Fig. 1). The 20 tandem repeats at the hp-CG4068 locus were suggestive of local duplications, and created potential for a vast array of higher-order hairpin conformations (Fig. 1b and Supplementary Fig. 3). In addition, hp-CR32205, hp-pncr009 and hp-CG32207 were related in sequence and located within a 70-kb interval (Supplementary Figs 5–8). Thus, hpRNAs, like miRNAs, can apparently evolve as local genomic duplications.

We probed the consequences of dsRNA-mediated knockdown of candidate factors on hpRNA biogenesis. We first confirmed the potency of these knockdowns by analysing Bantam, the premiRNA and/or mature miRNA of which were sensitive to Drosha, Pasha, Dcr-1, Loquacious (Loqs), Exportin-5 (Exp5, also known as Ranbp21) and Argonaute 1 (AGO1), as expected (Fig. 2a). The behaviour of hp-CG4068B/D/G and hp-CG18854A contrasted sharply with that of bantam (Fig. 2a and Supplementary Fig. 10a). Consistent with their apparent derivation from phased cleavage of long inverted repeats, their processing was unaffected by Dcr-1 depletion, but was strongly dependent on Dcr-2. In addition, ~21-nucleotide hpRNA products were markedly reduced when AGO2 was depleted. Incidentally, the Dcr-2/AGO2-dependent accumulation of ~21-nucleotide (siRNA) and ~42-nucleotide (terminal loop) hp-CG4068D isoforms (Fig. 2a) provided evidence for the *in vivo* processing of both 'single-repeat' and 'double-repeat' (or higher-order) forms of the hp-CG4068 hairpins (Fig. 1b).

Several other aspects of hpRNA biogenesis deserve mention. First, we were surprised that hpRNA processing was very strongly dependent on the Dcr-1-cofactor Loqs. This was especially unexpected in light of the recent realization that the *loqs* null condition only mildly compromises the maturation of many miRNAs[8], such as Bantam (Fig. 2a). Second, mature hpRNA products declined reproducibly in AGO1-deficient cells, which suggested the possible involvement of both AGO proteins in hpRNA biogenesis and/or function. Third, knockdown of Dcr-2, AGO2 and, to a lesser extent, AGO1 resulted in a ladder of hybridizing bands consistent with impaired hairpin processing (Fig. 2a and Supplementary Fig. 10a). This suggested that, in addition to Dcr-2, AGO proteins might also participate in hpRNA biogenesis. A role for AGO proteins has also been suggested for the maturation of siRNA duplexes and some pre-miRNAs[2,9–12]. Analysis of mutant animals corroborated this picture of hpRNA biogenesis, because mature hpRNA products were strongly reduced in *Dcr-2*, *loqs* and *AGO2* homozygous mutants (Fig. 2b).

We next analysed the termini of hpRNA-derived small RNAs. β-elimination of RNAs with two free hydroxyl groups at their 3′ termini increases their mobility in denaturing Polyacrylamide gel electrophoresis, whereas treatment with calf intestinal phosphatase (CIP) reduces the mobility of 5′ monophosphorylated RNAs[13]. Accordingly, miRNAs run faster after β-elimination and slower after CIP treatment (Fig. 2c and Supplementary Fig. 10b). CIP tests also indicated the presence of 5′ phosphates on hpRNA products (Supplementary Fig. 10b),

but all of them were resistant to β-elimination indicating modification of the 3′-terminal ribose (Fig. 2c). *Drosophila* Hen1 methylates Piwi-associated RNAs (piRNAs) and exogenous siRNAs at their 3′ termini[14,15]. We found that *hen1* mutants exhibited lower levels of mature hpRNA products (Fig. 2b), and these were now fully sensitive to β-elimination (Fig. 2c). These data supported the classification of hpRNA products not as miRNAs, but as siRNAs.

We tested the regulatory activity of hpRNA-derived siRNAs using artificial luciferase transcripts linked to target sites that were complementary to various hp-CG4068- and hp-CG18854-derived siRNAs. Their activity was analysed in cells that overexpressed hp-CG4068 or hp-CG18854, with non-cognate pairs controlling for the generic effect of hpRNA overexpression. These tests revealed the specific repression of hp-CG4068B and hp-CG4068C sensors by single- and double-repeat hp-CG4068 expression constructs (Fig. 3a), and of the hp-CG18854B sensor by hp-CG18854 (Fig. 3b). However, a sensor for hp-CG4068D was not affected by ectopic hp-CG4068, consistent with its lower read count compared to hp-CG4068B and hp-CG4068C.

To address the activity of endogenous hpRNAs expressed by S2 cells, we asked whether 2′-O-methyl antisense oligonucleotides (ASOs) could derepress these sensors. Indeed, ASO-hp-CG18854B (but not ASOs to hp-CG18854A or hp-CG4068B) induced approximately twofold derepression of the hp-CG18854B sensor (Fig. 3c). Reciprocally, we observed that ASO-hp-CG4068B (but not other ASOs) resulted in a approximately twofold activation of the hp-CG4068B sensor (Fig. 3d). Thus, both exogenous and endogenous hpRNAs generate inhibitory RNAs.

Some miRNAs are partially loaded into AGO2 (refs [16] and [17]), but hpRNA products are the first endogenous *Drosophila* small RNAs known to be preferentially sorted to AGO2 as a class. This provided an opportunity to ask whether endogenously programmed AGO2 functions by means of slicing, translational repression, or both. We prepared hp-CG4068B sensors carrying tandem perfect sites, centrally bulged sites, or bulged plus seed mismatched sites. Both mutant sensors were strongly derepressed, and to roughly the same extent, relative to the perfect sensor (Fig. 3e). In fact, the activity of the mutant sensors was comparable to the perfect sensor in the presence of cognate ASO (Fig. 3d). These data support the notion that hp-CG4068B exerts its major regulatory effect by slicing, with comparably little contribution from translational repression by AGO2.

With these functional regulatory data in hand, we searched for endogenous targets. We first considered whether hp-CG4068 might regulate the overlapping CG4068 3′ UTR (Supplementary Fig. 3), but gain- and loss-of-function tests were negative (Supplementary Fig. 11). Searches for *trans*-encoded targets were hindered by the fact that hp-CG4068 was only identifiable in the related species *D. melanogaster*, *Drosophila simulans* and *Drosophila sechellia* (Supplementary Fig. 3). However, hp-CG4068B, the most abundant siRNA product of hp-CG4068, contains 20 nucleotides of antisense complementarity (including three G•U pairs) to the coding region of *mus308* (Fig. 3f). This transcript encodes a DNA polymerase/helicase required for DNA repair after exposure to crosslinking agents[18]. Compensatory co-variation between *D. melanogaster* and *D. sechellia* hp-CG4068B and *mus308* target sites were suggestive of functional conservation (Fig. 3f). Consistent with this, we observed that *mus308* levels were increased ~2–3-fold in cells depleted of Dcr-2 or AGO2 (Supplementary Fig. 12), and that a luciferase-*mus308* sensor was specifically repressed by single-repeat and double-repeat hp-CG4068 expression constructs (Fig. 3a).

Because our data indicated that hpRNAs generate functional siRNAs that are primarily dependent on AGO2, we tested whether endogenous hp-CG4068B complexes exhibited Slicer activity. Endogenously programmed complexes indeed cleaved a perfect hp-CG4068B target

substrate in a manner that was competed away by ASO-hp-CG4068B but not ASO-CG18854B (Fig. 3g). We also found that a *mus308* target was cleaved by endogenous hp-CG4068B with similar specificity (Fig. 3g). We conclude that *mus308* is an endogenous Slicer target of hp-CG4068B.

We also searched for targets of hp-CG18854. The gene annotated as CG18854 is a possible pseudogene, because its open reading frame is short and poorly conserved[5]. CG18854 exhibits significant homology to the chromodomain gene CG8289 (Supplementary Fig. 13), and some of the abundant hp-CG18854-derived siRNAs exhibited extensive antisense complementarity to CG8289. When tested individually, siRNA-complementary sites from CG8289 did not mediate significant repression (data not shown). We therefore examined whether hp-CG18854 could regulate a translational fusion of CG8289 containing an extended complementary sequence. We transfected S2 cells with either tub-GFP or a tub-CG8289:GFP plasmid along with various hpRNA expression constructs, and observed that hp-CG18854 specifically repressed the accumulation of CG8289:GFP (Supplementary Fig. 13). These data demonstrate that hpRNA products can repress endogenous targets.

In plants, long hairpin RNAs from transgenes and long, extensively paired (and presumably very young) miRNA hairpins are substrates of DICER-LIKE4 (refs [19] and [20]), and thus mature through a pathway distinct from that of canonical miRNA hairpins, which are substrates of DICER-LIKE1 (ref. [21]). Likewise, we have shown that *Drosophila* hpRNAs enter a pathway distinct from that of miRNAs. Their derivation from unexpectedly long hairpins serves as an important caution for efforts to identify inverted-repeat small RNA genes. For example, some hpRNA-derived clones were recently reported but attributed incorrectly[22], because only short genomic precursors were considered in that study.

The *Drosophila* pathway combines canonical RNAi (Dcr-2, Hen1 and AGO2) and miRNA (Loqs) biogenesis factors—a revelation that highlights the incomplete nature of our current understanding of small-RNA-sorting mechanisms. Together with concurrent studies that identify endogenous siRNAs from transposons and *cis*-natural antisense pairs in *Drosophila* 22[–24], this work sets the stage for directed studies of the genetic requirements for host-directed RNAi in this organism.

METHODS SUMMARY

We used EINVERTED[4] to identify candidate genomic hairpins contained within 10-kb windows that satisfied a cutoff score ≥80 and had ≥70% pairing within the duplex region. Their expression as small RNAs was analysed using ten 454 libraries[5], a Solexa female head library[23], and a new set of Solexa imaginal disc/brain library (NCBI-GEO GSM275691). We defined candidate hpRNA loci as non-transposon inverted repeats for which the duplex region generated more than three times as many 21–22-nucleotide RNAs than all other-sized RNAs combined (Supplementary Figs 2–9). For functional tests, we followed published protocols for soaking RNAi in S2 cells and northern blotting[25] from knockdown samples or pharate adult flies. For sensor tests, four-copy-site targets (hp-CG4068B, hp-CG4068C and hp-CG4068D sensors) or a two-copy-site target (*mus308* sensor) were prepared by inserting oligonucleotides into a modified version of psiCHECK2 (ref. [25]). For hpRNA expression constructs, one or two hp-CG4068 repeats were cloned into the 3′ UTR of UAS-DsRed; CG18854 fragments were amplified from genomic DNA (CG18854 genomic) or a LD34273 clone (CG18854 cDNA) and were cloned similarly. A CG8289:GFP translational sensor consisted of a CG8289 fragment from genomic DNA and cloned into the KpnI site of tub-GFP[26]. RNA 3′ termini were analysed using periodate treatment in borax/boric-acid buffer followed by NaOH treatment (β-elimination) as described[15]. RNA 5′ termini were analysed using CIP as described[13]. For cleavage assays, we prepared labelled gel-purified templates using α-[32] P-GTP and capping

enzyme (Ambion). Cleavage reactions were performed as described[2] using S2-R+ cell extract. For detailed bioinformatic and molecular methods, see Methods.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

METHODS

Bioinformatics

We used EINVERTED[4] to identify candidate genomic hairpins contained within 10-kb windows that satisfied a cutoff score ≥80 and had ≥70% pairing within the duplex region. These criteria eliminated all but one of the annotated *Drosophila* miRNAs (*mir-997*)[27]. We kept those candidates with small RNA reads in the following data sets: 10 libraries analysed using 454 pyrosequencing[5], a female head library analysed using Solexa[23], and a new set of imaginal disc/brain small RNA sequences analysed using Solexa. We removed those loci in which the predicted duplex overlapped an annotated transposable element[28], and calculated the size distribution of reads from each of the remaining loci. We considered those loci for which the duplex region generated more than three times as many 21–22-nucleotide RNAs than all other-sized RNAs combined as hpRNA candidates (Supplementary Figs 2–9).

RNA interference

Segments of Pasha, Drosha and CG8273 were amplified using the indicated primers and were cloned into XhoI-XbaI sites of Litmus28i vector (NEB); other plasmids were described previously[25,29]. These templates were used to generate dsRNA, and soaking RNAi was performed as described[25]. S2-R+ cells were resuspended in serum-free medium at 3×10^6 cells ml^{-1} density and dsRNA was added to a concentration of 15 μg ml^{-1}. After 30 min incubation, an equal volume of Schneider's medium supplemented with 20% FBS was added. dsRNA treatment was repeated 4 days after the first treatment, and RNA samples were collected 4 days after the second soaking. XhoI-CG8273-279+, AGAGCTCGAGTCAGACAAATCCTCCGGTTC; XbaI-CG8273-682−, AGAGGTCTAGATTCGCCATCTGACTTGGTC; XhoI-Pasha-1240+, AGAGCTCGAGGGAGGTGGAGCAACAAAAGA; XbaI-Pasha-1725−, AGAGGTCTAGATACTCGTGCAGGATGCAGAC; XhoI-Drosha3522+, AGAGCTCGAGGCCGGACATTCCCTACTACA; XbaI-Drosha3943−, AGAGGTCTAGAGGGCATTGTTGGACTCTTG.

Northern blotting

Northern blotting was performed as described[25] using total RNA isolated from S2 cells or pharate adult *Drosophila*. All of the mutant strains used were described previously: *Dcr-2*^{L811fsX} (ref. [1]), *loqs*-KO (ref. [30]), *AGO2*^{414} (ref. [2]) and *hen1*^{f00810} (refs [14] and [15]). The sequences of the probes used in this study are listed below. DNA and locked nucleic acid (LNA) probes were obtained from IDT and Exiqon, respectively, *bantam* probe (DNA), AATCAGCTTTCAAAATGATCTCA; 2S rRNA probe (DNA), TACAACCCTCAACCATATGTAGTCCAAGCA; hp-CG4068-D (LNA), GTGACTTCCGGCGGTTAAGATTT; hp-CG4068-B (LNA), GGAGCGAACTTGTTGGAGTCAA; hp-CG4068-G (LNA), AGTTGGACTCAAACAAGTCCCT; hp-CG18854-A (LNA), TCATTTGATCCATAGTTTCCCGT; hp-CG18854-B (LNA), GGAGGGCGAAATGTTCAAGATCA; miR-8 (LNA), A.

Analysis of RNA chemical structure

RNA 3′ termini were analysed as described[15]. 10 μl of 61.5 mM $NaIO_4$ in borax/boric-acid buffer (60 mM borax and 60 mM boric acid, pH 8.6) was added to 10 μg total RNA in 14.6 μl water, and the samples were incubated for 30 min at room temperature (22 °C). 2.5 μl of 500 mM NaOH was added to each sample and incubation was continued for 90 min at 45 °C. The reactions were stopped by addition of 200 μl of 300 mM NaCl, 10 μg glycogen and 600 μl absolute ethanol. RNA was collected by centrifugation after 30 min incubation on ice.

RNA 5′ termini were analysed as described[13]. RNA samples were incubated with 2 units CIP (New England Biolabs) in 1 × buffer 3 (NEB, 100 mM NaCl, 50 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM dithiothreitol) for 2 h at 37 °C. RNA was purified by phenol/chloroform extraction followed by ethanol precipitation.

Sensor assays

Four-copy-site targets (hp-CG4068B, hp-CG4068C and hp-CG4068D sensors) or a two-copy-site target (*mus308* sensor) were prepared using the oligonucleotides listed below. Target sequences were inserted into SalI-XhoI (four-copy sensors) or NotI-XhoI (two-copy sensor) cloning sites of a modified version of psiCHECK2 (ref. [25]). CG4068 3′ UTR sensor was constructed by inserting a CG4068 3′ UTR fragment amplified with primers CG4068A and B (containing one repeat of the hpRNA repeat) into NotI-XhoI sites of the modified psiCHECK2. CG18854 fragments were amplified from genomic DNA (CG18854 genomic) or LD34273 clone (CG18854 cDNA) and cloned into NotI-XhoI sites of UAS-DsRed. The CG8289 fragment was amplified from genomic DNA and cloned into KpnI site of tub-GFP plasmid[26].

Luciferase sensor assays were performed as described previously[25]. We performed quadruplicate transfections of 25 ng target, 12.5 ng ub-Gal4 and 25 ng UAS-DsRed-hpRNA plasmids into 1×10^5 S2 cells in 96-well format. For 2′-O-methyl antisense-mediated de-silencing assays (inhibitor sequences listed below), we introduced 25 ng target plasmid and l0 pmol of 2′-O-methyl oligo-nucleotides for each well. Three days later, we lysed the cells and subjected them to the dual luciferase assay (Promega) and analysed these on a Veritas plate luminometer (Turner Biosystems). KpnI-CG8289 targetF, ggggtaccgccgccaccatgGTTGCTGAAAAGGATTCG; KpnI-CG8289, ggggtaccTTCGAGGAGCGTTCAATACGAT targetR; NotI-LD34273-1 +, AGAGgcggccgcAGTGCTGAGCATACCTAAGC; XhoI-LD34273-2390−, AGAGctcgagGTTCCCACATCGACTGGAAT; CG4068_A, aggggcggccgcACAAGCCAAAATCGTAtagg; CG4068_B, agggctcgagTTTTGCGTGGACTCATTCCC; hp-CG4068B_targ_A, tcgacaaaaGGAGCGAACTTGTTGGAGTCAAaagaac; hp-CG4068B_targ_B, tcgagttcttTTGACTCCAACAAGTTCGCTCCttttg; hp-CG4068C_targ_A, tcgacaaaaTTCCAGCGCCTGTGAAGCGCCAgagaac; hp-CG4068C_targ_B, tcgagttctcTGGCGCTTCACAGGCGCTGGAAttttg; hp-CG4068D_targ_A, tcgacaaaaGTGACTTCCGGCGGTTAAGATTTagaac; hp-CG4068D_targ_B, tcgagttctAAATCTTAACCGCCGGAAGTCACttttg; hp-CG4068A_si2x_A, GGCCGCGGAGCGAACTTGTTGGAGTCAAaaatcacGGAGCGAACTTGTTGGAGTCA AaC; hp-CG4068A_si2x_B, TCGAGtTTGACTCCAACAAGTTCGCTCCgtgatttTTGACTCCAACAAGTTCGCT GC; hp-CG4068A_mi2x_A, GGCCGCGGAGCGAACTACATCCACTGAAaaatcacGGAGCGAACTACATCCACTGA AaC; hp-CG4068A_mi2x_B, TCGAGtTTCAGTGGATGTAGTTCGCTCCgtgatttTTCAGTGGATGTAGTTCGCT GC; hp-CG4068A_mimut2x_A,

GGCCGCGGAGCGAACTACATGGAGTCAAaaatcacGGAGCGAACTACATGGAGTCA AaC; hp-CG4068A_mimut2x_B, TCGAGtTTGACTCCATGTAGTTCGCTCCgtgatttTTGACTCCATGTAGTTCGCT GC; mus308-target1not, ggccatggGCGAGCTTGTTGGAGTCAgggtgatttggGCGAGCTTGTTGGAGTCAggc; mus308-target2, tcgagcctgactccaacaagctcgcccaaatcacccTGACTCCAACAAGCTCGCccat; ASO-hp-CG4068-B, 5′AACAUggagcgaacttgttggagtcaaAUACA 3′; ASO-hp-CG4068-C, 5′ AACAUttccagcgcctgtgaagcgccagUCACU 3′; ASO-hp-CG4068-D, 5′ AACAUGUGACUUCCGGCGGUUAAGAUUUAUACA 3′; ASO-hp-CG18854-A, 5′ AACAUtgGCCAAGGTACGTGGTCGACCGAAUACU 3′; ASO-hp-CG18854-B, 5′ AACAUGGAGGGCGAAATGTTCAAGATCAUCACU3′.

For the green fluorescent protein (GFP) sensor assay, 250 ng target, 125 ng ub-Gal4 and 250 ng UAS-DsRed-hpRNA plasmids were transfected to 2×10^6 cells in 6-well format. Three days later, transfected cells were harvested and lysed with 2× SDS–PAGE sample buffer. Western blotting was performed using rabbit anti-GFP (Molecular Probes) or mouse anti-α-tubulin (DM1A, Sigma).

In vitro cleavage assay

Templates for *in vitro* cleavage targets were prepared by treating annealed oligonucleotides with Taq polymerase. The oligonucleotides A and B or A and C (below) were used for the template preparation for hp-CG4068-B target or mus308 target, respectively. Target RNAs were *in vitro* transcribed using Megascript T7 kit (Ambion) and purified by acrylamide gel electrophoresis. Purified RNA was labelled by α-^{32} P-GTP using capping enzyme (Ambion) according to the manufacturer's instructions. The cleavage reaction was performed as described[2] using S2-R+ cell extract. S2-R+ cells were resuspended in hypotonic buffer (30 mM HEPES-KOH, pH 7.4, 2 mM magnesium acetate, 5 mM DTT, 1× Complete mini EDTA free (Roche)) and lysed by five passages through a 25-gauge needle. The lysate was cleared by a centrifugation for 25 min at 14,600g at 4 °C, and was flash-frozen in 10-μl aliquots. Approximately 2,000 counts per min cap-labelled RNA was incubated in a reaction mixture (50% S2 lysate, 0.5 mM ATP, 5mM DTT, 100 mM KOAc, 0.1 U μl^{-1} RNaseOut (Invitrogen)) for 1 h at room temperature. 2′-O-methyl-ASO inhibitors (Integrated DNA Technologies) were added at 100 nM concentration and preincubated with the reaction mixture at room temperature for 20 min before the addition of the cap-labelled target RNA. Reactions were stopped by the addition of stop buffer (50 mM sodium chloride, 50 mM EDTA, 1% SDS, and 100 μg ml^{-1} proteinase K). RNA was recovered by phenol/chloroform extraction and ethanol precipitation. A, LucLet7_3′region_AS, attaattcctatGAGGTAGTAGGTTGTATAGTTCGAAGTATTCCGCGTACGTT; B, T7_hp-CG40682B_Luc_sense, TAATACGACTCACTATAGGGAGCGAACTTGTTGGAGTCAAattcAACGTACGCGG AATAC; C, T7_Mus308tar_Luc_ sense, TAATACGACTCACTATAGggtggGCGAGCTTGTTGGAGTCAgggAACGTACGCGGA ATAC.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Lee YS, et al. Distinct roles for *Drosophila* Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. Cell 2004;117:69–81. [PubMed: 15066283]

2. Okamura K, Ishizuka A, Siomi H, Siomi MC. Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. Genes Dev 2004;18:1655–1666. [PubMed: 15231716]
3. Kennerdell JR, Carthew RW. Heritable gene silencing in *Drosophila* using double-stranded RNA. Nature Biotechnol 2000;18:896–898. [PubMed: 10932163]
4. Rice P, Longden I, Bleasby A. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet 2000;16:276–277. [PubMed: 10827456]
5. Wilson RJ, Goodman JL, Strelets VB. FlyBase: integration and improvements to query tools. Nucleic Acids Res 2008;36:D588–D593. [PubMed: 18160408]
6. Tupy JL, et al. Identification of putative noncoding polyadenylated transcripts in *Drosophila melanogaster*. Proc. Natl Acad. Sci. USA 2005;102:5495–5500. [PubMed: 15809421]
7. Jones-Rhoades MW, Bartel DP, Bartel B. MicroRNAs and their regulatory roles in plants. Annu. Rev. Plant Biol 2006;57:19–53. [PubMed: 16669754]
8. Liu X, et al. Dicer-1, but not Loquacious, is critical for assembly of miRNA-induced silencing complexes. RNA 2007;13:2324–2329. [PubMed: 17928574]
9. Leuschner PJ, Ameres SL, Kueng S, Martinez J. Cleavage of the siRNA passenger strand during RISC assembly in human cells. EMBO Rep 2006;7:314–320. [PubMed: 16439995]
10. Matranga C, Tomari Y, Shin C, Bartel DP, Zamore PD. Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. Cell 2005;123:607–620. [PubMed: 16271386]
11. Diederichs S, Haber DA. Dual role for Argonautes in microRNA processing and posttranscriptional regulation of microRNA expression. Cell 2007;131:1097–1108. [PubMed: 18083100]
12. Grishok A, et al. Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. Cell 2001;106:23–34. [PubMed: 11461699]
13. Vagin VV, et al. A distinct small RNA pathway silences selfish genetic elements in the germline. Science 2006;313:320–324. [PubMed: 16809489]
14. Saito K, et al. Pimet, the *Drosophila* homolog of HEN1, mediates 2′-O-methylation of Piwi-interacting RNAs at their 3′ ends. Genes Dev 2007;21:1603–1608. [PubMed: 17606638]
15. Horwich MD, et al. The *Drosophila* RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. Curr. Biol 2007;17:1265–1272. [PubMed: 17604629]
16. Seitz H, Ghildiyal M, Zamore PD. Argonaute loading improves the 5′ precision of both microRNAs and their miRNA strands in flies. Curr. Biol 2008;18:147–151. [PubMed: 18207740]
17. Forstemann K, Horwich MD, Wee L, Tomari Y, Zamore PD. *Drosophila* microRNAs are sorted into functionally distinct Argonaute complexes after production by Dicer-1. Cell 2007;130:287–297. [PubMed: 17662943]
18. Harris PV, et al. Molecular cloning of *Drosophila mus308*, a gene involved in DNA cross-link repair with homology to prokaryotic DNA polymerase I genes. Mol. Cell. Biol 1996;16:5764–5771. [PubMed: 8816490]
19. Dunoyer P, Himber C, Voinnet O. DICER-LIKE 4 is required for RNA interference and produces the 21-nucleotide small interfering RNA component of the plant cell-to-cell silencing signal. Nature Genet 2005;37:1356–1360. [PubMed: 16273107]
20. Rajagopalan R, Vaucheret H, Trejo J, Bartel DP. A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*. Genes Dev 2006;20:3407–3425. [PubMed: 17182867]
21. Reinhart BJ, Weinstein EG, Rhoades MW, Bartel B, Bartel DP. MicroRNAs in plants. Genes Dev 2002;16:1616–1626. [PubMed: 12101121]
22. Ghildiyal M, et al. Endogenous siRNAs derived from transposons and mRNAs in *Drosophila* somatic cells. Science. published online 10 April 2008
23. Czech B, et al. An endogenous small interfering RNA pathway in *Drosophila*. Nature. (this issue)
24. Kawamura Y. *Drosophila* endogenous small RNAs bind to Argonaute 2 in somatic cells. Nature. (this issue)
25. Okamura K, Hagen JW, Duan H, Tyler DM, Lai EC. The mirtron pathway generates microRNA-class regulatory RNAs in *Drosophila*. Cell 2007;130:89–100. [PubMed: 17599402]

26. Stark A, Brennecke J, Russell RB, Cohen SM. Identification of *Drosophila* microRNA targets. PLoS Biol 2003;1:E60. [PubMed: 14691535]

27. Ruby JG, et al. Evolution, biogenesis, expression, and target predictions of a substantially expanded set of *Drosophila* microRNAs. Genome Res 2007;17:1850–1864. [PubMed: 17989254]

28. Karolchik D, et al. The UCSC Genome Browser Database: 2008 update. Nucleic Acids Res 2008;36:D773–D779. [PubMed: 18086701]

29. Forstemann K, et al. Normal microRNA maturation and germ-line stem cell maintenance requires Loquacious, a double-stranded RNA-binding domain protein. PLoS Biol 2005;3:e236. [PubMed: 15918770]

30. Park JK, Liu X, Strauss TJ, McKearin DM, Liu Q. The miRNA pathway intrinsically controls self-renewal of *Drosophila* germline stem cells. Curr. Biol 2007;17:533–538. [PubMed: 17320391]

Acknowledgements

We are grateful to R. Carthew, Q. Liu, H. Siomi and S. Cohen for plasmids and *Drosophila* strains. K.O. was supported by the Charles H. Revson Foundation. H.G. was supported by A*STAR, Singapore. D.P.B. is an HHMI investigator, and work in his laboratory was supported by a grant from the NIH (GM067031). E.C.L. was supported by grants from the Leukemia and Lymphoma Foundation, the Burroughs Wellcome Foundation, the V Foundation for Cancer Research, the Sidney Kimmel Foundation for Cancer Research, and the NIH (GM083300).

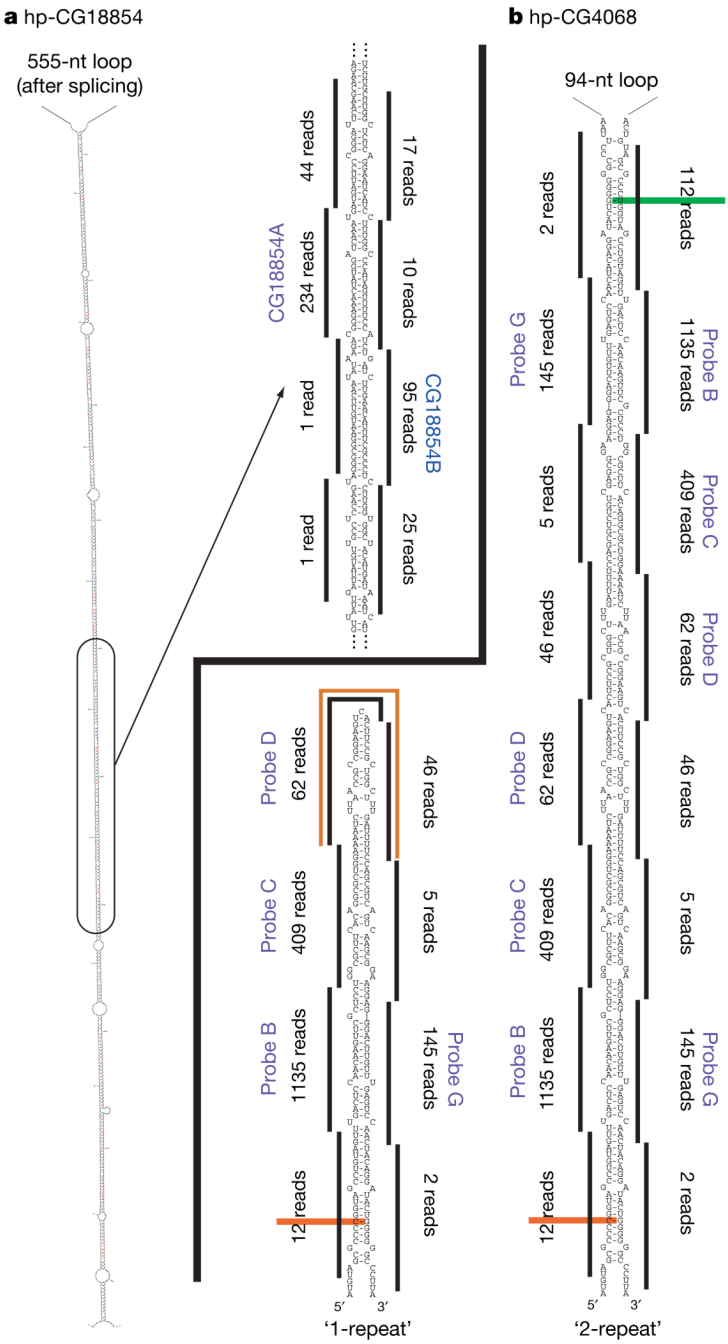

Figure 1. Examples of *Drosophila* hpRNA transcripts

a, hp-CG18854 contains a >400-bp duplex separated by a large loop; the enlarged region highlights the phased nature of small RNA duplexes. Northern probes were designed against the RNAs labelled in blue. **b**, The hp-CG4068 locus consists of 20 tandem repeats that partially overlap the 3′ UTR of CG4068 (Supplementary Fig. 3) and generate phased small RNA duplexes. Each repeat adopts a hairpin structure, but higher-order hairpins are possible because repeats are complementary to each other; '1-repeat' and '2-repeat' isoforms are shown. Distinct small RNAs were cloned from related repeats with minor sequence differences (for example, RNAs highlighted in red and green).

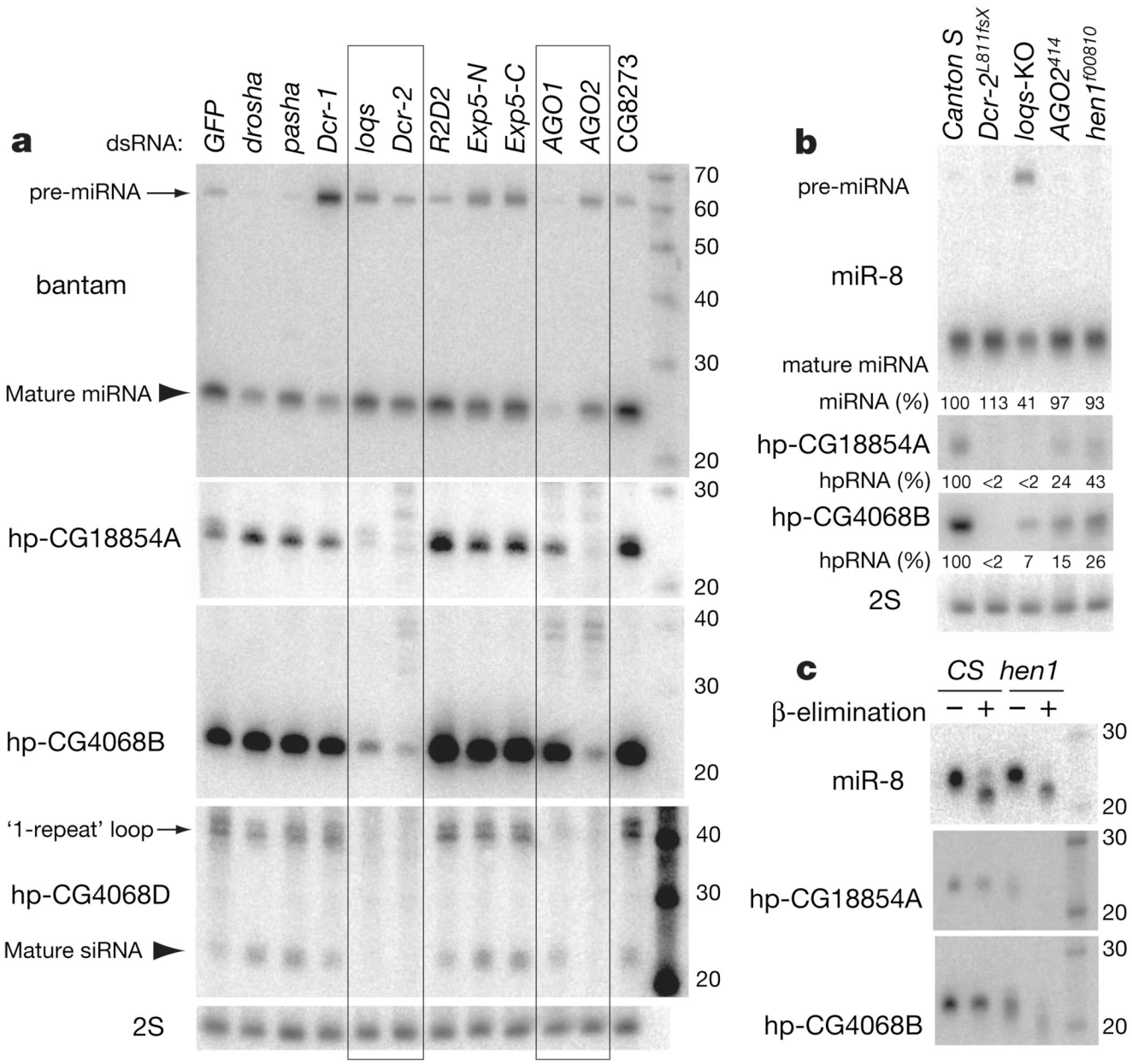

Figure 2. Distinct biogenesis pathways for miRNAs and hpRNAs

a, Unlike miRNAs (for example, Bantam), hpRNA biogenesis in S2 cells is highly dependent on Dcr-2 and AGO2; Loqs and AGO1 suppression affect both miRNA and hpRNA biogenesis. **b,** miRNA and hpRNA biogenesis in pharate adult *Drosophila*. miR-8 was affected only in the *loqs* mutant, whereas hpRNA products were strongly decreased in the *Dcr-2* and *loqs* mutants (<10%), and significantly affected in the *AGO2* and *hen1* mutants. **c,** Modification of hpRNA-derived small RNAs is mediated by Hen1. Note that hpRNA products from *hen1* mutants run as a range of faster-migrating species after β-elimination. CS, Canton S (a strain of fruitfly); KO, knockout deletion strain.

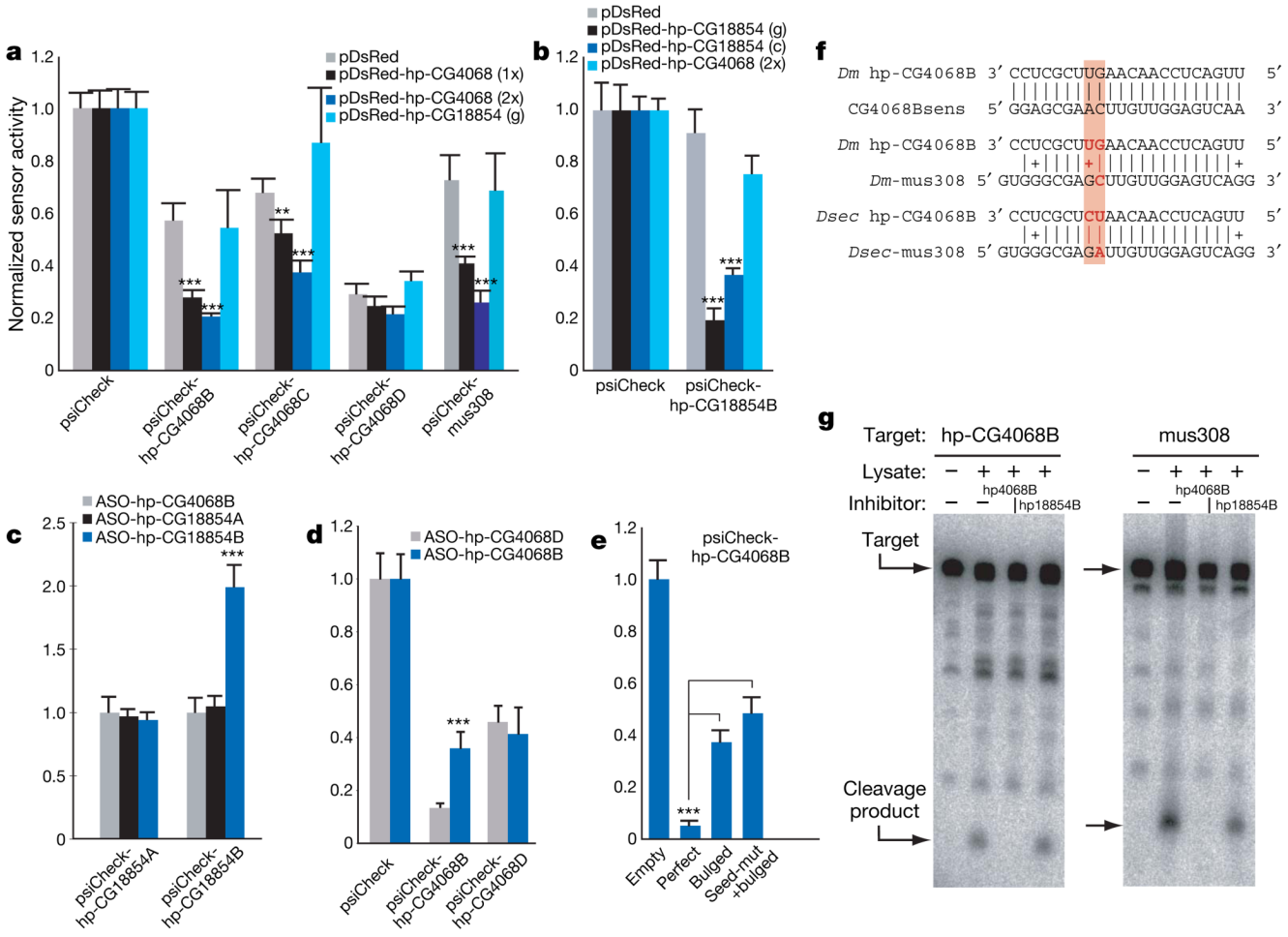

Figure 3. hpRNAs generate regulatory RNAs that can repress endogenous targets

a, In S2 cells, both single-repeat (1×) and double-repeat (2×) hp-CG4068 constructs specifically repressed the hp-CG4068B, hp-CG4068C and mus308 sensors; the hp-CG4068D sensor was unaffected. **b,** Both genomic (g) and cDNA (c) hp-CG18854 expression constructs specifically repressed the hp-CG18854B sensor. **c,** The 2′-O-methyl ASO against hp-CG18854B specifically derepressed the hp-CG18854B sensor, whereas ASO-CG4068B specifically relieved endogenous repression of the hp-CG4068B sensor (**d**). **e,** Compared to a perfect hp-CG4068B sensor, mutant sensors with a central bulge or central bulge plus seed mutations exhibited the same level of derepression. Error bars depict the standard deviation of eight transfections; statistical comparisons were performed with the unequal variance Students *t*-test; $^{**}P < 6 \times 10^{-5}$, $^{***}P < 1 \times 10^{-8}$. **f,** Compensatory covariation (pink shaded box) between hp-CG4068B and mus308 target sites of *D. melanogaster* (*Dm*) and *D. sechellia* (*Dsec*). Red font, nucleotides that have evolved; sens, sensor. **g,** Endogenous RNA-induced silencing complex from S2R cleaved both perfect hp-CG4068B and *mus308* sensors, and this activity was specifically competed away by cognate ASO.