2010-03-11

Endogenous Small RNAs in the *Drosophila* Soma: A Dissertation

Megha Ghildiyal
University of Massachusetts Medical School

Endogenous small RNAs in the *Drosophila* soma

A Dissertation Presented

By

MEGHA GHILDIYAL

Submitted to the Faculty of the
University of Massachusetts Graduate School of Biomedical Sciences, Worcester
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY

March 11th, 2010

INTERDISCIPLINARY GRADUATE PROGRAM

ENDOGENOUS SMALL RNAs IN THE *DROSOPHILA* SOMA

A Dissertation Presented By

Megha Ghildiyal

The signatures of the Dissertation Defense Committee signifies completion and approval as to style and content of the Dissertation

Phillip Zamore, Ph.D., Thesis Advisor

Michael Brodsky, Ph.D., Member of Committee

Kirsten Hagstrom, Ph.D., Member of Committee

Nicholas Rhind, Ph.D., Member of Committee

Yukihide Tomari, Ph.D., Member of Committee

The signature of the Chair of the Committee signifies that the written dissertation meets the requirements of the Dissertation Committee

Craig Mello, Ph.D., Chair of Committee

The signature of the Dean of the Graduate School of Biomedical Sciences signifies that the student has met all graduation requirements of the School.

Anthony Carruthers, Ph.D.
Dean of the Graduate School of Biomedical Sciences

Interdisciplinary Graduate Program

March 11, 2010

DEDICATION

This thesis is dedicated to Kamala bua.

I wish I was there to say bye, but I was too late.

I will always miss you.

ACKNOWLEDGEMENTS

My thesis work would not have been possible without the support, encouragement and cooperation by many. First of all I owe my gratitude to my thesis adviser Phil Zamore. He has been a wonderful mentor, and has guided me at every step, to be a better scientist. I immensely appreciate the attention he pays to details, how we present our data, how we interact with other researchers and colleagues and how we perceive others and ours science. I have infinitely benefited from the direction provided by Phil, and will continue to do so throughout my scientific career. I am grateful to him for providing me with the opportunity to work in his lab and to all the avenues it opened for me.

I owe my gratitude to all past and present members of the Zamore lab; they were not just my colleagues but also my family for the last six years. I want to thank Klaus for his advice, patience and guidance. I am enormously grateful to Tingting, Vasia and Hervé, for sharing with me their secrets to a successful experiment. They were my partners in crime, both inside and outside the lab, and also my greatest support system. I wish them all the best and success in everything. I want to thank Jen for being the elder 'American' sister I always dreamt of, I admire you for your resilience and courage. My most wonderful time in the day was spent in my bay or 8th floor kitchen. For that I want to thank all my past and present bay-mates, Elif, Wee, Ryuya and Zhao; and my talented companions Fabian and Tracey, who gulped a lot of coffee with me. You guys made me smile even through the hardest of times. I am also humbled by the generosity bestowed on me by Tiff, Shengmei, Gwen and Alicia; I always had my needs met and as a result, I

am a spoilt brat. Ira, I am greatly obliged to you for being the one I could run to with my problems and you were always there to listen. Especially, I would like to express my heartfelt gratitude to Stefan. Not only did you teach me all those cool experiments but also made me realize my passion for science. Thanks for being my friend and companion.

I am also grateful to my collaborators, Prof. Zhiping Weng and her lab members, Jia and Soo, for their scientific input and for improving my work. I would also like to thank my TRAC members (Craig Mello, Michael Brodsky, Nicholas Rhind, Kirsten Hagstrom and Yukihide Tomari) for their insightful suggestions, guidance for my research and for always keeping me in track.

I was very fortunate to receive a lot of love and support from my friends, both in India and US. I want to thank Divya, Ambika, Arunima, Mohita, Niti and Jaya for sticking with me through thick and thin. I made some great friends here in US, especially Madhavi, Sagar, Ermelinda (also my housemate), Prachi, Charu, Mukhi, Archana, Geoffrey, Ami, Katerina, Ling, Dimitra; thanks for all the fun times we had and the smiles we shared. Varun, you were with me through stress and sickness, and I can never thank you enough for all the happiness and love you brought to my life.

Above all, I am indebted to my family and all my relatives back home, for loving and supporting me. I love you all and miss you. To my parents, who taught me how to dream and then set me free to pursue those dreams. You made a lot sacrifices for me and I cannot even visit you enough. I apologize for the distance between us but I love you both most in the world, and your love and wisdom are with me in every step I take in life. And thank-you God for all your blessings, I am one lucky girl.

COPYRIGHT INFORMATION

ABSTRACT

Since the discovery in 1993 of the first small silencing RNA, a dizzying number of small RNAs have been identified, including microRNAs (miRNAs), small interfering RNAs (siRNAs) and Piwi-interacting RNAs (piRNAs). These classes differ in their biogenesis, modes of target regulation and in the biological pathways they regulate.

Historically, siRNAs were believed to arise only from exogenous double-stranded RNA triggers in organisms lacking RNA-dependent RNA polymerases. However, the discovery of endogenous siRNAs in flies expanded the biological significance of siRNAs beyond viral defense. By high throughput sequencing we identified *Drosophila* endo-siRNAs as 21 nt small RNAs, bearing a 2´-*O*-methyl group at their 3´ ends, and depleted in *dicer-2* mutants.

Methylation of small RNAs at the 3´ end in the soma, is a consequence of assembly into a mature Argonaute2-RNA induced silencing complex. In addition to endo-siRNAs, we observed certain miRNAs or their miRNA* partners loading into Argonaute2. We discovered, that irrespective of its biogenesis, a miRNA duplex can load into either Argonaute (Ago1 or Ago2), contingent on its structural and sequence features, followed by assignment of one of the strands in the duplex as the functional or guide strand. Usually the miRNA strand is selected as the guide in complex with Ago1 and miRNA* strand with Ago2.

In our efforts towards finding 3´ modified small RNAs in the fly soma, we also discovered 24-28nt small RNAs in certain fly genotypes, particularly *ago2* and *dcr-2*

mutants. 24-28nt small RNAs share many features with piRNAs present in the germline, and a significant fraction of the 24-28nt small RNAs originate from similar transposon clusters as somatic endo-siRNAs. Therefore the same RNA can potentially act as a precursor for both endo-siRNA and piRNA-like small RNA biogenesis. We are analyzing the genomic regions that spawn somatic small RNAs in order to understand the triggers for their production. Ultimately, we want to attain insight into the underlying complexity that interconnects these small RNA pathways.

Dysregulation of small RNAs leads to defects in germline development, organogenesis, cell growth and differentiation. This thesis research provides vital insight into the network of interactions that fine-tune the small RNA pathways. Understanding the flow of information between the small RNA pathways, a great deal of which has been revealed only in the recent years, will help us comprehend how the pathways compete and collaborate with each other, enabling each other's optimum function.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

Introduction

Small silencing RNAs, 20–29 nucleotides (nt) long, are the master-regulators of several biological processes and fine-tune many developmental aspects of eukaryotes. They serve as specificity determinants for Argonaute (Ago) proteins, which they guide to their targets, typically resulting in reduced expression of target genes. Small RNAs exercise their regulation by base pairing with target mRNAs and repress their expression, via transcriptional or post-transcriptional silencing. Beyond these defining features, different small RNA classes guide diverse and complex schemes of gene regulation. Some small silencing RNAs, such as siRNAs, derive from double-stranded RNA (dsRNA), whereas others, such as piRNAs, do not. These different classes of regulatory RNAs also differ in the proteins required for their biogenesis, the constitution of the Argonaute-containing complexes that execute their regulatory functions, their modes of gene regulation, and the biological functions in which they participate (Table 1). New small RNA classes and new examples of existing classes continue to be discovered**.** The discovery of the overwhelming diversity between the small RNA pathways is constantly accompanied with evidence of their interaction and inter-dependence. There is a growing realization that these distinct small RNA pathways are interconnected and that small RNA pathways compete and collaborate as they regulate genes and protect the genome from external and internal threats.

Table I-1. Types of small silencing RNAs.

Name	Organism	Length (nt)	Proteins	Source of Trigger	Function	References
miRNA	Plants, algae, animals, viruses, protists	20-25	Drosha (animals only) + Dicer	Pol II transcription (pri-miRNAs)	Regulation of mRNA stability, translation	1-6
casiRNA	Plants	24	DCL3	Transposons, repeats	Chromatin modification	7-13
tasiRNA	Plants	21	DCL4	miRNA-cleaved *TAS* RNAs	Post transcriptional regulation	14-18
natsiRNA	Plants primary secondary	24 21	DCL2 DCL1	Bidirectional transcripts induced by stress	Regulate stress response genes	19,20
Exo-siRNA	Animals, fungi, protists Plants	~21 21 & 24	Dicer	Transgenic, viral or other exogenous dsRNA	Post transcriptional regulation, anti-viral defense	21-24
Endo-siRNA	Plants, algae, animals, fungi, protists,	~21	Dicer (Except secondary siRNAs in *C. elegans*, which are products of RdRP transcription, and are therefore not technically siRNAs.)	Structured loci, convergent and bi-directional transcription, mRNAs paired to antisense pseudogene transcripts	Post transcriptional regulation of transcripts and transposons Transcriptional gene silencing	1,2,25-34
piRNA germ line	*Drosophila melanogaster*, mammals, zebrafish	24–30	Dicer-independent	Long, primary transcripts?	Transposon regulation, unknown functions	35-43

piRNA-like (soma)	*Drosophila melanogaster*	24–30	Dicer-independent	In *ago2* mutants in *Drosophila*	Unknown	26
21U-RNA piRNAs	*Caenorhabditis elegans*	21	Dicer-independent	Individual transcription of each piRNA?	Transposon regulation, unknown functions	44-47
26G RNA	*Caenorhabditis elegans*	26	RdRP?	Enriched in sperm	Unknown	44

The Discovery of RNAi

RNA silencing was inadvertently triggered when two groups attempted to make petunia leaves more purple by over expressing chalcone synthase (CHS) from a highly expressed transgene; instead, pigmentation was lost or reduced in 25–40% of the plants[48,49]. Because expression of both the endogenous and transgenic CHS genes was reduced, the phenomenon was called "co-suppression." Co-suppression was also held responsible for inducing anti-viral resistance in plants following introduction of a virally derived transgene[50-52]. A follow up study in plants suggested nucleic acid as a possible mediator of co-suppression because of its ability to act as a systemic signal and specifically target complementary RNAs[53].

In parallel, paradoxical results were reported for the nematode, *Caenorhabditis elegans*: introduction of either sense or antisense RNA was able to repress expression of the corresponding gene[54]. In 1998, Fire and Mello, in their Nobel prize-winning work, established double-stranded RNA as the silencing trigger in *C. elegans*[55]. Their experiments overturned the contemporary view that antisense RNA induced silencing by base pairing to its mRNA counterpart, thereby preventing its translation into protein. In worms and other animals, siRNA-mediated silencing is known as RNA interference (RNAi). Remarkably, RNAi is systemic in both plants and nematodes, spreading from cell to cell[56]. In *C. elegans*, RNAi is also heritable: silencing can be transferred to the progeny of the worm originally injected with the trigger dsRNA[57]. Viral infection, inverted repeat transgenes, or aberrant transcription products, all lead to the production of dsRNA. dsRNA is converted to siRNAs that direct RNAi. siRNAs were discovered in

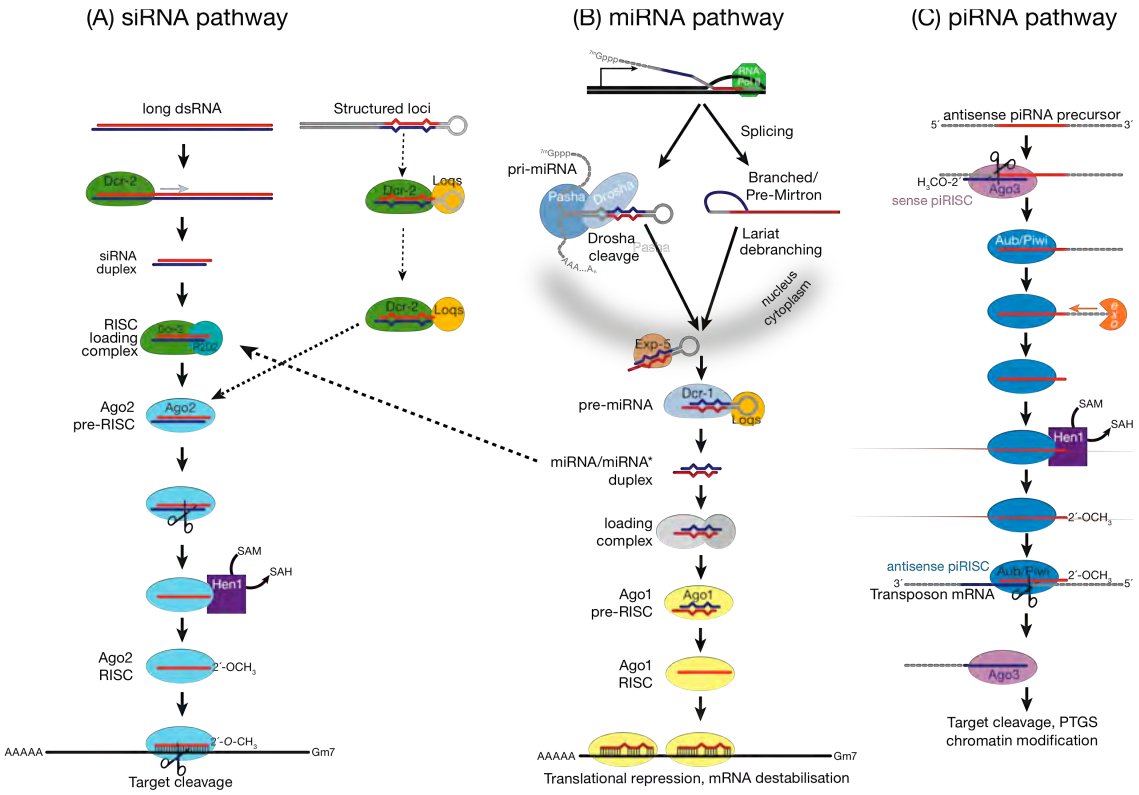
plants[21] and later shown in animal extracts to serve as guides that direct endonucleolytic cleavage of their target RNAs[22,58]. siRNAs can be classified according to the proteins involved in their biogenesis, their mode of regulation or their size. Here, we differentiate the major types of siRNAs according to the molecules that trigger their production, a classification scheme that best captures the biological distinctions among small silencing RNAs.

siRNAs derived from exogenous agents.

Early examples of RNAi were triggered by exogenous dsRNA. In these cases, long, exogenous dsRNA is cleaved into double-stranded siRNAs by Dicer (Dcr), a dsRNA-specific RNase III family ribonuclease[59] (Fig. 1). siRNA duplexes produced by Dicer comprise two ~21 nt strands, each bearing a 5′ phosphate and 3′ hydroxyl group, paired in a way that leaves two-nucleotide overhangs at the 3′ ends[22,24,60]. The strand that directs silencing is called the guide, whereas the other strand, which is ultimately destroyed, is the passenger. Target regulation by siRNAs is mediated by the RNA-induced silencing complex (RISC)**,** the generic name for an Argonaute-small RNA complex[58]. In addition to an Argonaute protein and a small RNA guide, RISC may also contain auxiliary proteins that extend or modify its function, for example, proteins that re-direct the target mRNA to a site of general mRNA degradation[61].

Figure I-1. Small RNA silencing pathways in *Drosophila*. The three small RNA silencing pathways in flies are the siRNA, miRNA and piRNA pathways. These pathways differ in their substrates, biogenesis, effector proteins and modes of target regulation. (i) DsRNA precursors are processed by Dcr-2 to generate siRNA duplexes containing guide and passenger strands. Dcr-2 along with R2D2, loads the duplex into Ago2. A subset of endo-siRNAs exhibit Loqs dependence, rather than R2D2. The passenger strand is later destroyed and the guide strand directs Ago2 to the target RNA. (ii) miRNAs are encoded in the genome and are transcribed to yield a pri-miRNA transcript, which is cleaved by Drosha to yield a short pre-miRNA. Alternatively, miRNAs can be present in introns that are liberated following splicing to yield authentic pre-miRNAs. pre-miRNAs are exported from the nucleus to cytoplasm, where they are further processed by Dcr-1 to generate a miRNA/miRNA* duplex. Once loaded into Ago1, the miRNA strand guides translational repression of target RNAs. (iii) piRNAs are thought to derive from single-stranded RNA precursors and made without a dicing step. piRNAs are mostly antisense, but a small fraction is in the sense orientation. Antisense piRNAs are preferentially loaded into Piwi and Aub, whereas sense piRNAs associate with Ago3. Piwi and Aub collaborate with Ago3 to mediate an inter-dependent amplification cycle that generates additional piRNAs, preserving the bias towards antisense. The antisense piRNAs likely direct transposon mRNA cleavage or chromatin modification at transposon loci.

Figure I-1.


(A) siRNA pathway
(B) miRNA pathway
(C) piRNA pathway
long dsRNA
Structured loci
Dcr-2
Loqs
siRNA duplex
RISC loading complex
Ago2 pre-RISC
SAM
SAH
Hen1
Ago2 RISC
2´-OCH3
AAAAA
Gm7
2´-O-CH3
Target cleavage
Splicing
pri-miRNA
Pasha
Drosha
Drosha cleavge
Branched/ Pre-Mirtron
Lariat debranching
nucleus
cytoplasm
Exp-5
pre-miRNA
Dcr-1
miRNA/miRNA* duplex
loading complex
Ago1 pre-RISC
Ago1
Ago1 RISC
Translational repression, mRNA destabilisation
antisense piRNA precursor
H3CO-2´
Ago3
sense piRISC
Aub/Piwi
antisense piRISC
Transposon mRNA
Target cleavage, PTGS
chromatin modification

Mammals and *C. elegans* each have a single Dicer that makes both miRNAs and siRNAs[62-65], whereas *Drosophila* species has two Dicers: Dcr-1 makes miRNAs, whereas Dcr-2 is specialized for siRNA production[66]. The fly RNAi pathway defends against viral infection, and Dicer specialization may reduce competition between pre-miRNAs and viral dsRNAs for Dicer. Alternatively, Dcr-2 and Ago2 specialization might reflect the evolutionary pressure on the siRNA pathway to counter rapidly evolving viral strategies to escape RNAi. In fact, *dcr-2* and *ago2* are among the most rapidly evolving *Drosophila* genes[67]. *C. elegans* may achieve similar specialization with a single Dicer by using the double-stranded RNA-binding protein, RDE-4, as the gatekeeper for entry into the RNAi pathway[68]. However, no natural virus infection has been documented in *C. elegans*[69]. By contrast, mammals may not use the RNAi pathway to respond to viral infection, having evolved an elaborate, protein-based immune system[70-72].

The relative thermodynamic stabilities of the 5´ ends of the two siRNA strands in the duplex determines the identity of the guide and passenger strands[73-75]. In flies, this thermodynamic difference is sensed by the dsRNA-binding protein R2D2, the partner of Dcr-2 and a component of the RISC Loading Complex (RLC)[76,77]. The RLC recruits Argonaute2 (Ago2), to which it transfers the siRNA duplex. Ago2 can then cleave the passenger strand as if it were a target RNA[78-82]. Ago2 always cleaves its RNA target at the phosphodiester bond that lies between the nucleotides paired to guide nucleotides 10 and 11[24,60]. Release of the passenger strand after its cleavage converts pre-RISC to mature RISC, which contains only single-stranded guide RNA. In flies, the guide strand is 2´-*O*-methylated at its 3´ end by the *S*-adenosyl methionine-dependent

methyltransferase, Hen1, completing RISC assembly[83,84]. In plants, both miRNAs and siRNAs are terminally methylated, which is crucial for their stability[85-87].

Plants exhibit a surprising diversity of small RNA types and the proteins that generate them. The diversification of RNA silencing pathways in plants may reflect the need of a sessile organism to cope with biotic and abiotic stress. The number of RNA silencing proteins can vary enormously among animals too, with *C. elegans* producing 27 distinct Argonaute proteins compared with 5 in flies. Phylogenetic data suggest that nearly all of these 'extra' *C. elegans* Argonautes act in the secondary siRNA pathway, perhaps because endogenous, secondary siRNAs are so plentiful in worms[88]. *Arabidopsis thaliana* has four Dicer-like (DCL) proteins and 10 Argonautes, with both unique and redundant functions. In plants, inverted repeat transgenes or co-expressed sense and antisense transcripts produce two sizes of siRNAs, 21 nt and 24 nt[10,89]. The 21 nt siRNAs are produced by DCL4, but in the absence of DCL4, DCL2 can substitute, making 22 nt siRNAs[13,90-93]. The DCL4-produced 21-mers typically associate with AGO1 and guide mRNA cleavage. The 24-mers associate with AGO4 (major) and AGO6 (surrogate), and promote the formation of repressive chromatin[94].

In plants, exogenous sources of siRNAs are not confined to dsRNAs**.** Single-stranded sense transcripts from tandemly repeated or highly expressed single-copy transgenes are converted to dsRNA by RDR6, a member of the RNA-dependent RNA Polymerase (RdRP) family, transcribe single-stranded RNA from an RNA template[95] **(Box 1)**. RDR6 and RDR1, also convert viral single-stranded RNA into dsRNA, initiating an anti-viral RNAi response[96]. The resulting dsRNA is cleaved by Dicer into siRNAs that

are terminally 2´-*O*-methylated by HEN1[87]. Why plants RNAs expressed from transgenes are converted by RDR6 into dsRNA, but abundant, endogenous mRNAs are not, is poorly understood. Recent evidence that some housekeeping exonucleases compete with plant RNA silencing pathways for aberrant RNAs suggests that substandard RNA transcripts—e.g. those lacking a 5´ cap or 3´ poly(A) tail—act as substrates for RdRPs. Highly expressed transgenes might overwhelm normal RNA quality control pathways, escape destruction, and be converted to dsRNA by RdRPs[96-98].

Endogenous siRNAs (endo-siRNAs)

The first endo-siRNAs were detected in plants and *C. elegans*[9,10,99]. Plants too produce a variety of endo-siRNAs, and, the recent discovery of endo-siRNAs in flies and mammals suggests that endo-siRNAs are ubiquitous among higher eukaryotes.

Plant endo-siRNAs. In plants, *cis*-acting siRNAs (casiRNAs) originate from transposons, repetitive elements, and tandem repeats such as 5S rRNA genes, and comprise the bulk of endo-siRNAs[13] (Fig. 2). CasiRNAs are predominantly 24 nt long and methylated by HEN1. Their accumulation requires DCL3 and the RNA polymerases, RDR2 and POL IV, and either AGO6 (primarily) or AGO4, which act redundantly[7,9,13,100-107]. CasiRNAs promote heterochromatin formation by directing DNA methylation and histone modification of the loci from which they originate[7-13].

Another class of plant endo-siRNAs illustrates how distinct small RNA pathways interact. *Trans*-acting siRNAs (tasiRNAs) are endo-siRNAs generated by the

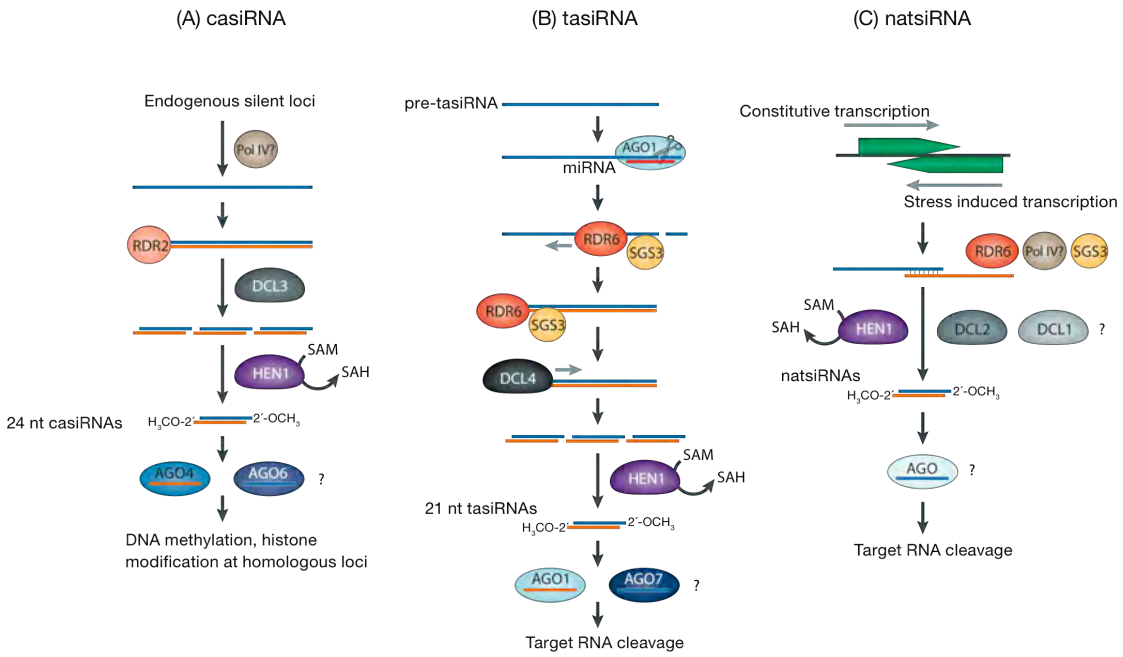
convergence of the miRNA and siRNA pathways in plants[14-18](Fig. 2). miRNA-directed cleavage of certain transcripts recruits the RdRP enzyme, RDR6. RDR6 then copies the cleaved transcript into dsRNA, which DCL4 dices into tasiRNAs that are phased. This phasing suggests that DCL4 begins dicing precisely at the miRNA cleavage site, making a tasiRNA every 21 nt[18]. The site of miRNA cleavage is critical, because in determining the entry point for Dicer, it establishes the target specificity of the tasiRNAs produced. One of the determinants that seems to predispose a transcript to produce tasiRNAs after its cleavage by a miRNA is the presence of a second miRNA or siRNA complementary site on the transcript. Of special mention is the *TAS3* locus, whose RNA transcript has two binding sites for miR-390. Only one of these sites is efficiently cleaved by miR-390, but binding of the miRNA to both appears to be required to initiate conversion of the *TAS3* transcript to dsRNA by RDR6[108,109].

Natural antisense transcript-derived siRNAs (natsiRNAs) are produced in response to stress in plants[19,20](Fig. 2). They are generated from a pair of convergently transcribed RNAs: typically, one transcript is expressed constitutively, whereas the complementary RNA is transcribed only when the plant is subject to environmental stress, such as high salt. Production of 21- and 24-nt siRNAs from region of overlap of the two transcripts requires DCL2 and/or DCL1, RDR6, and SGS3 (SUPPRESSOR OF GENE SILENCING3, probably an RNA-binding protein)[110] and Pol IV[19,20]. The natsiRNAs then direct cleavage of one of the mRNAs of the pair, and in one such case, trigger the DCL1-dependent production of 21 nt secondary siRNAs[20]. In addition to natsiRNAs, “long” siRNAs (lsiRNAs) in *Arabidopsis* also originate from NAT pairs and

are stress-induced. Unlike natsi-RNAs, lsiRNAs are 30–40 nts long and require DCL1, DCL4, AGO7, RDR6 and POL IV for their production[111].

Figure I-2. Plant endo-siRNA biogenesis. Casi, tasi and natsiRNAs are derived from distinct loci. Several of the proteins involved in their biogenesis are genetically redundant, while others have specialized roles. (i) CasiRNAs are the most abundant endogenously produced siRNAs in plants. POL IV and RDR2 are proposed to generate dsRNA precursors, which are then diced by DCL3 to generate 24 nt casiRNAs. These small RNAs load into AGO4 and perhaps AGO6; they promote heterochromatin assembly by targeting DNA methylation and histone modification at the corresponding loci. (ii) TasiRNA biogenesis requires miRNA-mediated cleavage of TAS transcripts, which triggers the production of dsRNA by RDR6. The dsRNA is diced into 21 nt tasiRNAs by DCL4 and acts through either AGO1 or AGO7. (iii) NatsiRNAs are derived from overlapping regions of convergent transcripts and require DCL1 and or DCL2, POL IV, RDR6 and SGS3 for their biogenesis.

Figure I-2.


(A) casiRNA
(B) tasiRNA
(C) natsiRNA
Endogenous silent loci
Pol IV?
RDR2
DCL3
SAM
HEN1
SAH
24 nt casiRNAs
H3CO-2'
2'-OCH3
AGO4
AGO6
?
DNA methylation, histone modification at homologous loci
pre-tasiRNA
AGO1
miRNA
RDR6
SGS3
DCL4
21 nt tasiRNAs
AGO7
Target RNA cleavage
Constitutive transcription
Stress induced transcription
DCL2
DCL1
natsiRNAs
AGO

Animal endo-siRNAs. Plant and worm endo-siRNAs are typically produced through the action of RdRPs (**Box 1**). The genomes of flies and mammals do not seem to encode such RdRP proteins, so the recent discovery of endo-siRNAs in flies and mice was unexpected.

The first mammalian endo-siRNAs to be reported corresponded to the long interspersed nuclear element (L1) retrotransposon and were detected in cultured human cells[25]. Full length LINE-1 (L1) contains both sense and antisense promoters in its 5´ untranslated region (UTR) that could, in principle, drive bi-directional transcription of L1, producing overlapping, complementary transcripts to be processed into siRNAs by Dicer, but the precise mechanism by which transposons trigger siRNA production in mammals remains unknown.

More recently, endogenous siRNAs have been detected in *Drosophila* somatic and germ cells and in mouse oocytes. High throughput sequencing of small RNAs from germ-line and somatic tissues of *Drosophila* and of Ago2 immunoprecipitates revealed a small RNA population that could readily be distinguished from miRNAs and piRNAs[26-29,112,113]. These small RNAs are nearly always exactly 21 nt long, are present in both sense and antisense orientations, have modified 3´ ends, and, unlike miRNAs and piRNAs, are not biased toward beginning with uracil. Production of the 21-mers requires Dcr-2, although in the absence of Dcr-2 a remnant of the endo-siRNA population inexplicably persists. Expression of transposon mRNAs increases in both *dcr-2* and *ago2* mutants, implicating an endogenous RNAi pathway in the silencing of transposons in flies, as reported previously for *C. elegans*[30,31].

Endo-siRNAs have also been identified in mouse oocytes[33,34]. As in flies, mouse endo-siRNAs are 21 nt long, Dicer-dependent, and derived from a variety of genomic sources (see Discussion in Chapter II and Fig. II-6). The mouse endo-siRNAs were bound to Ago2, the sole mammalian Ago protein thought to mediate target cleavage, although it is not known if they also associate with any of the other three mouse Ago proteins. (Mammalian Ago2 is not, however, the ortholog of fly Ago2, whose sequence is considerably diverged from other Ago proteins.)

A subset of mouse oocyte endo-siRNAs maps to regions of protein-coding genes capable of pairing to their cognate pseudogenes and to regions of pseudogenes capable of forming inverted-repeat structures. Pseudogenes can no longer encode proteins, yet they drift from their ancestral sequence more slowly than would be expected if they were simply junk. Perhaps some pseudogene sequences are under evolutionary selection to retain the ability to produce antisense transcripts that can pair with their cognate genes so as to produce endo-siRNAs[114].

miRNAs

The first microRNA, *lin-4*, was identified in a screen for genes required for post-embryonic development in *C. elegans*[115]. The *lin-4* locus produces a 22 nt RNA that is partially complementary to sequences in the 3´ UTR of its regulatory target, the *lin-14* mRNA[116-118]. miRNA binding to partially complementary sites in mRNA 3´ UTRs is now considered a hallmark of animal miRNA regulation. In 2001, tens of miRNAs were identified in humans, flies, and worms by small RNA cloning and sequencing,

establishing miRNAs as a new class of small silencing RNAs[3-5]. miRBase (Release 12.0), the registry that coordinates miRNA naming, now lists 1,638 distinct miRNAs in plants and 6,930 in animals and their viruses[119].

miRNA Biogenesis. miRNAs derive from precursor transcripts called primary miRNAs (pri-miRNAs), which are typically transcribed by Polymerase II[120-123]. Several miRNA genes are present as clusters in the genome and probably derive from a common pri-miRNA transcript. Liberating a 20–24 nt miRNA from its pri-miRNA requires the sequential action of two RNase III endonucleases, assisted by their double-stranded RNA-binding domain (dsRBD) partner proteins (Fig. 1). First, the pri-miRNA is processed in the nucleus into a 60–70 nt long pre-miRNA by Drosha, acting with its dsRBD partner, called DGCR8 in mammals and Pasha in flies[120,124-128]. The resulting pre-miRNA has a hairpin structure: a loop flanked by base-paired arms that form a stem. Pre-miRNAs have a two-nt overhang at their 3′ ends and a 5′ phosphate group, which are indicative of their production by an RNase III. The nuclear export protein Exportin-5 carries the pre-miRNA to the cytoplasm bound to Ran, a GTPase that moves RNA and proteins through the nuclear pore[129-132].

In the cytoplasm, Dicer and its dsRBD partner protein, TRBP in mammals and Loqs in flies, cleaves the pre-miRNA[59,62-64,133-137]. Drosha and Dicer differ in that Dicer—like Argonaute proteins, but unlike Drosha—contains a PAZ domain, presumably allowing it to bind the two-nucleotide, 3′ overhanging end left by Drosha. Dicer cleavage generates a duplex containing two strands, the miRNA and miRNA*, corresponding to

the two sides of the base of the stem. These roughly correspond to the guide and passenger strands of an siRNA, and similar thermodynamic criteria influence the choice of miRNA versus miRNA*[73,74]. miRNAs can arise from either arm of the pre-miRNA stem, and some pre-miRNAs produce mature miRNAs from both arms, whereas others show such pronounced asymmetry that the miRNA* is rarely detected even in high throughput sequencing experiments[44]**(Box 2)**.

In flies, worms and mammals, a few pre-miRNAs are produced by the nuclear pre-mRNA splicing pathway instead of Drosha processing[138-142]. These pre-miRNA-like introns, "mirtrons," are spliced out of mRNA precursors whose sequence suggests they encode proteins. The spliced introns first accumulate as lariat products that require 2´-5´ debranching by the lariat debranching enzyme. Debranching yields an authentic pre-miRNA, which can then enter the standard miRNA biogenesis pathway.

In plants, DCL1 fills the roles of both Drosha and Dicer, converting pri-miRNAs to miRNA/miRNA* duplexes[13,143-145]. DCL1, assisted by its dsRBD partner HYL1, converts pri-miRNAs to miRNA/miRNA* duplexes in the nucleus, after which the miRNA/miRNA* duplex is thought to be exported to the cytoplasm by HASTY, an Exportin-5 homolog (*HASTY* mutants develop precociously, hence their name)[15,145-147]. Unlike animal miRNAs, plant miRNAs are 2´-*O*-methylated at their 3´ ends by HEN1[85,143,148]. HEN1 protects plant miRNAs from 3´ uridylation, thought to be a signal for degradation[87]. HEN1 likely acts before miRNAs are loaded into AGO1, because both miRNA* and miRNA strands are modified in plants[85].

Target regulation by miRNAs. The mechanism by which a miRNA regulates its mRNA target reflects both the specific Argonaute protein into which the small RNA is loaded and the extent of complementarity between the miRNA and the mRNA[149-151]. A few miRNAs in flies and mammals are nearly fully complementary to their mRNA targets; these direct endonucleolytic cleavage of the mRNA[152-156]. Such extensive complementarity is considered the norm in plants, as target cleavage was thought to be the main mode of target regulation in plants[11,89,157]. However, in flies and mammals, most miRNAs pair with their targets through only a limited region of sequence at the 5´ end of the miRNA, the "seed"; these repress translation and direct degradation of their mRNA targets[158-163]. The "seed" region of all small silencing RNAs contributes most of the energy for target binding[164,165]. Thus, the seed is the primary specificity determinant for target selection. The small size of the seed means that a single miRNA can regulate many—even hundreds—of different genes[166,167]. Intriguingly, recent data suggests that the nuclear transcriptional history of an mRNA influences if a miRNA represses its translation at the initiation or the elongation step[168].

As plant miRNAs are highly complementary to their mRNA targets, they can direct mRNA target cleavage. Nonetheless, AGO1-loaded plant miRNAs can also block translation, suggesting a common mechanism between plant and animal miRNAs, despite the absence of specific miRNAs shared between the two kingdoms[169].

Functions of miRNAs. Like transcription factors, miRNAs regulate diverse cellular pathways, and are widely believed to regulate most biological processes, in both plants

and animals, ranging from housekeeping functions to responses to environmental stress. The cited reviews cover this vast body of work and provide valuable insight[170-172].

The study of miRNA pathway mutants provided early evidence for the influence of miRNAs on biological processes in both plants and animals. Loss of Dicer or miRNA-associated Argonaute proteins is nearly always lethal in animals, and such mutants show severe developmental defects in both plants and animals. In *Drosophila*, *dcr-1* mutant germ-line stem cell clones divide slowly; in *Arabidopsis*, embryogenesis is abnormal in *dcl1* mutants; in *C. elegans*, *dcr-1* mutants display defects in germ-line development and embryonic morphogenesis; zebrafish lacking both maternal and zygotic Dicer are similarly defective in embryogenesis; and mice lacking Dicer die as early embryos, apparently devoid of stem cells[65,143,173-176]. Loss of Dicer in mouse embryonic fibroblasts causes increased DNA damage and consequently, the up-regulation of $p19^{Arf}$ and p53 signaling that induces premature senescence[177].

Many miRNAs function in specific biological processes, in specific tissues, and at specific times[178]. The importance of small silencing RNAs goes far beyond the RNA silencing field: long-standing questions about the molecular basis of pluripotency, tumorogenesis, apoptosis, cell identity, etc. are finding answers in small RNAs[170,179].

piRNAs: the longest small RNAs

piRNAs function in the germ-line. Piwi-interacting RNAs (piRNAs) are the most recently discovered class of small RNAs, and, as their name suggests, they bind to the Piwi clade of Argonaute proteins. (Animal Argonaute proteins can be subdivided by

sequence relatedness into Ago and Piwi sub-families.) The Piwi clade comprises Piwi, Aubergine (Aub) and Ago3 in flies, MILI, MIWI and MIWI2 in mice, and HILI, HIWI1, HIWI2 and HIWI3 in humans.

piRNAs were first proposed to ensure germ line stability by repressing transposons when Aravin and colleagues discovered in flies a class of longer small RNAs (~25–30 nt) associated with silencing of repetitive elements[35]. Later, these 'repeat associated small interfering RNAs'—subsequently renamed piRNAs—were found to be distinct from siRNAs: they bind Piwi proteins and do not require Dcr-1 or Dcr-2 for their production, unlike miRNAs and siRNAs[84,180,181]. Moreover, they are 2´-*O*-methylated at their 3´ termini, unlike miRNAs, but like siRNAs in flies[83,180,182-184].

High throughput sequencing of vertebrate piRNAs revealed a class of piRNAs unrelated to repetitive sequences[34,36-42]. Mammalian piRNAs can be divided into pre-pachytene and pachytene piRNAs, according to the stage of meiosis at which they are expressed in developing spermatocytes. Like piRNAs in flies, pre-pachytene piRNAs predominantly correspond to repetitive sequences and are implicated in silencing transposons, such as L1 and intracisternal A-particle (IAP)[39]. In male mice, gametic methylation patterns are established when germ cells arrest their cell cycle 14.5 days postcoitum, resuming cell division 2–3 days after birth[185,186]. Both MILI and MIWI2 are expressed during this period, and *miwi2* and *mili* deficient mice lose DNA methylation marks on transposons[187]. The pre-pachytene piRNAs, which bind MIWI2 and MILI, may serve as guides to direct DNA methylation of transposons. In contrast to pre-pachytene

piRNAs, the pachytene piRNAs mainly arise from unannotated regions of the genome, not transposons, and their function remains unknown[39].

Three recent studies report that the previously discovered germ-line '21U' RNAs in *C. elegans* are piRNAs[44-47]. These small RNAs were initially identified by high throughput sequencing[44]. They are precisely 21 nt long, begin with a uridine 5´-monophosphate, and are 3´ modified. They bind Piwi-Related Gene-1 (PRG-1), a *C. elegans* Piwi protein. Each 21U-RNA may be transcribed separately, as all are flanked by a common upstream motif. Like piRNAs in *Drosophila*, the 21U-RNAs are required for maintenance of the germ line and fertility, and like *Drosophila* Aub and other piRNA pathway components, PRG-1 is found in specialized granules, P granules, associated with germ-line function, in a cytoplasmic, perinuclear ring called "nuage." Worm piRNAs resemble pachytene piRNAs in mammals: their targets and functions are largely unknown.

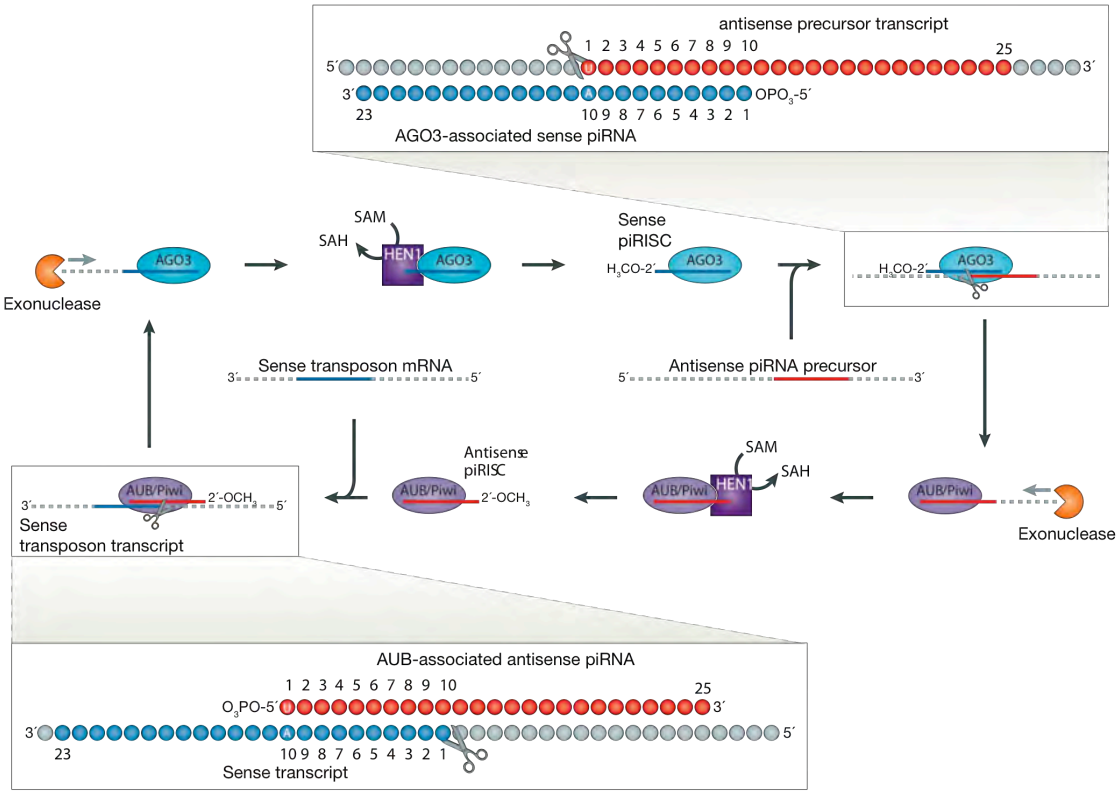
piRNA Biogenesis. piRNA sequences are stunningly diverse, with more than 1.5 million distinct piRNAs identified thus far in flies, but collectively they map to a few hundred genomic clusters27,29,43,113,181,188,189. The best-studied cluster is the *flamenco* locus. *flamenco* was identified genetically as a repressor of the *gypsy*, *ZAM* and *Idefix* transposons[84,190-194]. Unlike siRNAs, *flamenco* piRNAs are mainly antisense, suggesting that piRNAs arise from long, single-stranded precursor RNAs. In fact, disruption of *flamenco* by insertion of a P-element near the 5´ end of the locus blocks the production of

distal piRNAs up to 168 kbp away. Thus, an enormously long, single-stranded RNA transcript appears to be the source of those piRNAs that derive from the *flamenco* locus[43].

The current model for piRNA biogenesis was inferred from the sequences of piRNAs bound to Piwi, Aubergine and Ago3[43,195]. piRNAs bound to Piwi and Aubergine are typically antisense to transposon mRNAs, whereas Ago3 is loaded with piRNAs corresponding to the transposon mRNAs themselves (Fig. 1). Moreover, the first 10 nucleotides of antisense piRNAs are frequently complementary to the sense piRNAs found in Ago3. This unexpected sequence complementarity has been proposed to reflect a feed-forward amplification mechanism—"piRNA ping-pong"—that is activated only transcription of transposon mRNA (Fig. 3)[43,195]. A similar amplification loop has been inferred from high throughput piRNA sequencing in vertebrates, implying its conservation through evolution[36,187]. Many aspects of the ping-pong model remain speculative. Why Ago3 appears to bind only sense piRNAs derived from transposon mRNAs is unknown. An untested idea is that different forms of RNA Pol II transcribe primary piRNA transcripts and transposon mRNAs and that the specialized RNA Pol II that transcribes the primary piRNA precursor recruits Piwi and Aub, but not Ago3. How the 3´ ends of piRNAs are made is also not known.

Figure I-3. Feed-forward or "ping-pong" model for piRNA amplification. According to this model, antisense piRNAs in Piwi or Aub first bind transposon mRNAs and cleave them across from position 10 of the antisense piRNA guide. The 5´ end of the cleaved product is proposed to then load into Ago3 and generate an Ago3-bound sense piRNA. The sense piRNA can, in turn, guide cleavage of an antisense piRNA precursor transcript, fueling the feed-forward amplification loop. A key postulates of the model is that the intracellular concentrations of piRNA-loaded Piwi and Aub are much greater than of piRNA-loaded Ago3. The amplification loop is proposed to facilitate piRNA surveillance of transposon transcription in the germ-line.

Figure I-3.


antisense precursor transcript
1 2 3 4 5 6 7 8 9 10
25
5′
3′
3′
OPO3-5′
23
10 9 8 7 6 5 4 3 2 1
AGO3-associated sense piRNA
SAM
SAH
HEN1
AGO3
Exonuclease
Sense
piRISC
H3CO-2′
Sense transposon mRNA
Antisense piRNA precursor
Antisense
piRISC
AUB/Piwi
2′-OCH3
Sense
transposon transcript
Exonuclease
AUB-associated antisense piRNA
O3PO-5′
Sense transcript

piRNA Function and Regulation. Piwi family proteins are indispensable for germ-line development in many, perhaps all, animals; but they have thus far been most extensively studied in *Drosophila*. Piwi is restricted to the nucleoplasm of *Drosophila* germ cells and adjacent somatic cells. Piwi is required to maintain germ line stem cells and to promote their division; the protein is required in both the somatic niche cells that support germ-line stem cells and in the stem cells themselves[196,197]. In the male germ line, Aub is required for the silencing of the repetitive *Stellate* locus, which would otherwise cause male sterility. Expression of *Stellate* is controlled by the related, repetitive *Suppressor of Stellate* locus, the source of antisense piRNAs that act through Aub to repress *Stellate*[35,180,198].

aub was originally identified because it is required for specification of the embryonic axes[199]. The loss of anterior-posterior and dorsal-ventral patterning in embryos from mothers lacking Aub is an indirect consequence of the double-stranded DNA breaks that occur in the oocyte in its absence[200]. The breaks appear to activate a DNA-damage checkpoint that disrupts patterning of the oocyte and, consequently, of the embryo. The defects in patterning, but not in silencing repetitive elements, are rescued by mutations that bypass the DNA damage signaling pathway, suggesting the breaks are caused by transposition. That activation of a DNA damage checkpoint should inappropriately reorganize embryonic polarity was most unexpected, but further underscores the vital role piRNAs play in germ-line development.

piRNAs outside the germ line? The role of piRNAs in the fly soma is hotly debated. Piwi and Aub are required to silence tandem arrays of *white*, a gene required to produce red eye pigment[201]. It is not understood if piRNAs are produced in the soma as well as in the germ line, or if piRNAs present during germ-line development deposit long-lived chromatin marks that exert their effects days later.

Both piRNAs and endo-siRNAs repress transposons in the germ line, where mutations caused by transposition, of course, would propagate to the next generation. siRNAs—that is, the RNAi pathway—likely provide a rapid response to the introduction of a new transposon into the germ line, a challenge not dissimilar to a viral infection. In contrast, the piRNA system appears to provide a more robust, permanent solution to the acquisition of a transposon. In the soma, however, endo-siRNAs are the predominant transposon-derived small RNA class, and their loss in *dcr-2* and *ago2* mutants increases transposon expression[26,27,29,113]. Somatic piRNA-like small RNAs have been observed in *ago2* mutant flies[26]. Perhaps, in the absence of endo-siRNAs, piRNAs are produced somatically and resume transposon surveillance. Such a model implies significant cross talk between the piRNA and endo-siRNA–generating machineries.

Intertwined pathways

The RNAi, miRNA and piRNA pathways were initially believed to be independent and distinct. However, the lines distinguishing them continue to fade. These pathways interact and rely on each other at several levels, competing for and sharing substrates, effector proteins and cross-regulating each other.

Competition for substrates during loading. Both the siRNA and miRNA pathways load dsRNA duplexes containing a 19 bp double-stranded core flanked by 2 nt 3´ overhangs. An siRNA duplex contains guide and passenger strands and is complementary throughout its core; a miRNA/miRNA* duplex contains mismatches, bulges and G:U wobble pairs. In *Drosophila*, biogenesis of small RNA duplexes is uncoupled from its loading into Ago1 or Ago2[202,203]. Instead, loading is governed by the structure of the duplex: duplexes bearing bulges and mismatches are sorted into the miRNA pathway and hence loaded into Ago1; duplexes with greater double-stranded character partition into Ago2, the Argonaute protein associated with RNAi.

The partitioning of small RNAs between Ago1 and Ago2 also has implications for target regulation. Ago1 primarily represses translation whereas Ago2 represses by target cleavage, reflecting the faster rate of target cleavage by Ago2 compared to Ago1[203]. Sorting creates competition between the two pathways for substrates[202,203]. In *Drosophila* loading of a small RNA duplex into one pathway decreases its association with other pathway.

Different dsRNA precursors require distinct combinations of proteins to produce small silencing RNAs. For example, *Drosophila* endo-siRNAs derived from structured loci require Loqs, rather than R2D2 [27-29]. We presume that under some circumstances the endo-siRNA and miRNA pathways might therefore compete for Loqs. The endo-siRNA and RNAi pathways likely also compete for shared components.

In contrast to *Drosophila*, plants load small RNAs into Argonautes according to the identity of the 5´ nt of the small RNA[108,204]. AGO1 is the main effector Argonaute for

miRNAs, and the majority of miRNAs begin with uridine; and AGO4 is the major effector of the heterochromatic pathway and is predominantly loaded with small RNAs beginning with an adenosine[205]. AGO2 and AGO5, however have no characterized function in plants[205]. Changing the 5´ nt from A to U shifts the loading bias of a plant small RNA from AGO2 to AGO1, and vice versa. Similarly, *Arabidopsis* AGO4 binds small RNAs that begin with adenosine, while AGO5 prefers cytidine.

Aub- and Piwi-bound piRNAs typically begin with U, whereas those bound to Ago3 show no 5´ nucleotide bias. It remains to be determined if this reflects a 5´ nucleotide preference like the situation for the plant AGOs or some feature of an as-yet-discovered piRNA loading machinery that sorts piRNAs between Piwi proteins.

Cross talk. Small RNA pathways are often entangled. TasiRNA biogenesis in *Arabidopsis* is a classic example of such cross talk between pathways. miRNA-directed cleavage of tasiRNA-generating transcripts initiates tasiRNA production and subsequent regulation of tasiRNA targets[14-18]. In *C. elegans*, at least one piRNA has been implicated in initiating endo-siRNA production[45,46], and in flies, the endo-siRNA pathway may repress expression of piRNAs in the soma[26]. Moreover, small RNA levels may be buffered by negative feedback loops in which small RNAs from one pathway alter the expression levels of RNA silencing proteins that act in the same or in other RNA silencing pathways. In flies, endo-siRNAs have been identified that target siRNA pathway genes like *ago2*[26]. Similarly, *let-7* and perhaps other miRNAs repress Dicer expression in mammals, creating a negative feedback loop that buffers miRNA

levels[206,207]. *Arabidopsis* miRNAs can also regulate their own biogenesis: miR-168 controls *AGO1* levels and miR-162 can target *DCL1*[208-210].

Understanding the flow of information between the small RNA pathways, a great deal of which has been revealed only in the recent years, will help us comprehend how the pathways compete and collaborate with each other, enabling each other's optimum function.

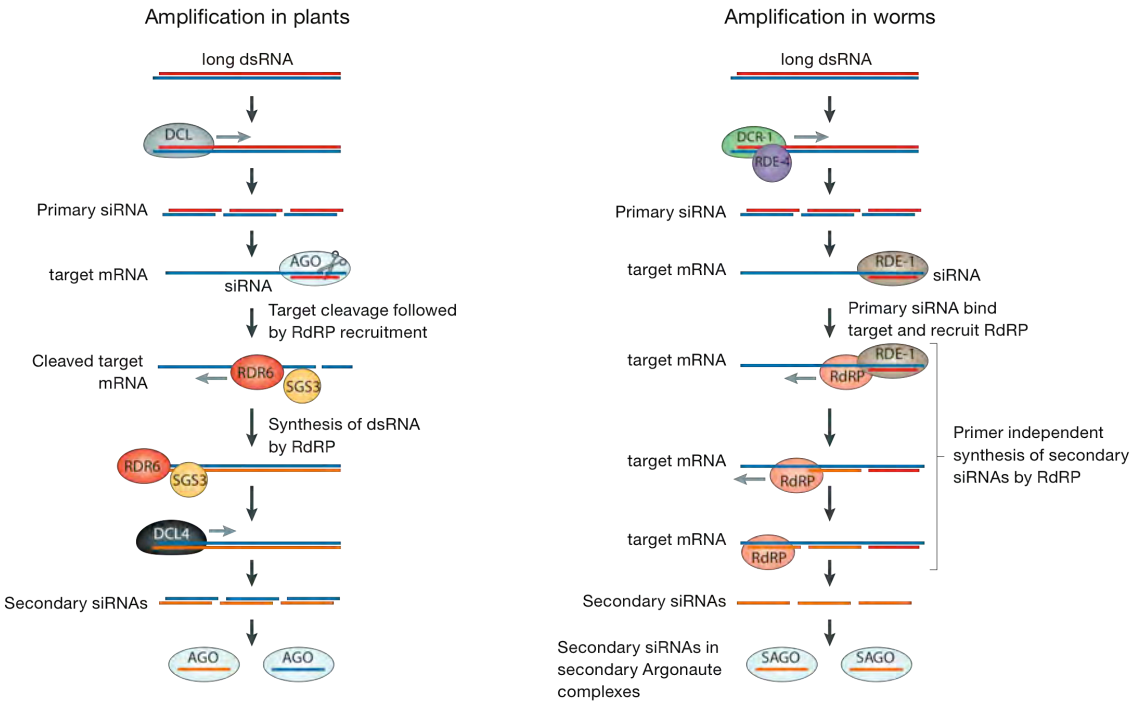
Box 1. Amplifying silencing.

RNA-dependent RNA polymerases (RdRPs) amplify the silencing response. Primary siRNAs derived from exogenous triggers by Dicer processing bind their mRNA targets and direct cleavage by AGO complexes[211]. In plants, RdRPs uses these cleaved transcript fragments as templates to synthesize long dsRNA; the dsRNA is then diced into secondary siRNAs[13,91,92,95,212-214]. Secondary siRNAs are formed both 5′ and 3′ of the primary targeted interval, suggesting that mRNA cleavage per se, rather than priming of RdRP by primary siRNAs is the signal for siRNA amplification. Other data suggest that production of secondary siRNAs in *Arabidopsis* may sometimes be primed[215]. RdRP amplification of siRNAs is especially important in defending plants against viral infection.

In *C. elegans*, primary siRNAs are amplified into secondary siRNAs by a different mechanism[211]. In worms, primary siRNAs are bound to RDE-1, a "primary Argonaute"[216,217]. The primary siRNAs guide RDE-1 to the target mRNA, to which it recruits RdRPs that synthesize secondary siRNAs[218,219]. Worm secondary siRNAs have a 5′ di- or triphosphate, indicating that they are produced by transcription rather than dicing[216,217,220], and, at least in vitro, secondary siRNA production does not require Dicer[219]. How the length of siRNA transcription is controlled is perplexing, but in vitro, the *Neurospora* RdRP, QDE1, can directly transcribe short RNA oligomers ~22 nt long from a much longer template[220]. As a consequence of their production by an RdRP, secondary siRNAs in *C. elegans* are exclusively antisense to their mRNA targets[216,221].

Secondary siRNAs act bound to secondary Argonautes, such as CSR-1, which can cleave its mRNA targets just like fly and human Ago2 proteins[219].

The presence of siRNA amplification in plants, worms, fungi, and according to some early reports, in flies, led to the speculation that RdRPs are a universal feature of RNAi. An amplification step in human RNAi could produce secondary siRNAs bearing homology to other genes, a significant impediment to the use of RNAi as a target discovery tool or as therapy for human diseases. However, the success of allele-specific RNAi in cultured human cells and in mice makes it unlikely that an RdRP-catalyzed amplification step occurs in mammals[222-226]. Similarly, extensive biochemical and genetic studies have demonstrated that the fly RNAi pathway does not use an RdRP enzyme[22,89,227-230].



Box 2. High throughput sequencing and small RNA discovery.

Much of the credit for the identification of small RNAs rests with advances in high throughput sequencing. Presently, there are three commercial "high depth" sequencing systems: Roche's 454 GS FLX Genome Analyzer, Illumina's Solexa Analyzer and, most recently, Applied Biosystem's SOLiD System. Reference 231 describes how each method works. Whereas 454 has the advantage of sequencing >250 bp per read, compared to ~35–50 bp for Solexa and SOLiD, these two platforms provide 70- to 400-fold greater sequencing depth. All three platforms have been used successfully to identify novel small RNA species and to discover new small RNA classes in mutant plants and animals. Using less than 10 μg total RNA, high throughput sequencing, together with advances in small RNA library preparation, has revealed the length distribution, sequence identity, terminal structure, sequence and strand biases, isoform prevalence, genomic origins, and mode of biogenesis for millions of small RNAs. Initial small RNA sequencing experiments sought simply to identify novel small RNA species and classes. Increasingly, high throughput sequencing is being used to profile small RNA expression across the stages of development and in different tissues and disease states. Profiling by deep sequencing provides quantitative information about small RNA expression, like PCR- or microarray-based approaches, but can also precisely detect subtle changes in small RNA sequence or length.

Perhaps the most problematic step in small RNA sequencing is preparing the small RNA library. The most frequently employed cloning protocols require the small RNAs to have 5´ phosphate and a 3´ hydroxyl groups, the hallmarks of Dicer products.

This approach identifies small RNAs with the expected termini, but alternative methods must be used to find small RNAs, such as *C. elegans* secondary siRNAs, with other terminal structures. Additionally, finding every possible small RNA in a cell using exhaustive deep sequencing is a game with diminishing returns. For example, while many miRNAs have been sequence 100,000's or even a million times, the *C. elegans* miRNA *lsy-6*, which is apparently expressed in less than ten cells of the adult, has so far eluded high depth sequencing[232]

CHAPTER II

Endogenous siRNAs derived from transposons and mRNAs in *Drosophila* somatic cells

The following chapter is a collaborative effort. The author conceived the experimental plan and performed experiments for all figures, except Figure 3. Figure 4 was collaboration between the author, Michael Horwich and Tingting Du.Tingting Du did Ago2 knockdown in S2 cells for Figure 4. Hervé Seitz, Soohyun Lee, Jia Xu and Zhiping Weng performed bioninformatic analyses. The author, Hervé Seitz and Phillip Zamore wrote the paper. This chapter appeared in Science. 2008 May 23;320(5879):1077-81.

Summary

Small interfering RNAs (siRNAs) direct RNA interference (RNAi) in eukaryotes. In flies, somatic cells produce siRNAs from exogenous double-stranded RNA as a defense against viral infection. We identified endogenous siRNAs (endo-siRNAs), 21 nucleotides in length that correspond to transposons and heterochromatic sequences in the somatic cells of *Drosophila melanogaster*. We also detected endo-siRNAs complementary to messenger RNAs (mRNAs); these siRNAs disproportionately mapped to the complementary regions of overlapping mRNAs predicted to form double-stranded RNA in vivo. Normal accumulation of somatic endo-siRNAs requires the siRNA-generating ribonuclease Dicer-2 and the RNAi effector protein Argonaute2 (Ago2). We propose that

endo- siRNAs generated by the fly RNAi pathway silence selfish genetic elements in the soma, much as piRNAs do in the germ line.

Introduction

Three RNA-silencing pathways have been identified in flies and mammals: RNA interference (RNAi), guided by small interfering RNAs (siRNAs) derived from exogenous double-stranded RNA (dsRNA); the microRNA (miRNA) pathway, in which endogenous small RNAs repress partially complementary mRNAs; and the Piwi-interacting RNA (piRNA) pathway, whose small RNAs repress transposons in the germ line[43,180,195] and can activate transcription in heterochromatin[233].

Endogenous siRNAs (endo-siRNAs) silence retrotransposons in plants[10,234], and siRNAs corresponding to the L1 retrotransposon have been detected in cultured mammalian cells[25]. Genetic and molecular evidence suggests that in addition to suppressing viral infection, the RNAi pathway silences selfish genetic elements in the fly soma: Mutations in the RNAi gene, *rm62*[235], suppress mutations caused by retroelement insertion[236]; depletion of the Argonaute proteins Ago1 or Ago2 increases transposon expression in cultured *Drosophila* Schneider 2 (S2) cells[237]; small RNAs have been detected in *Drosophila* Kc cells for the *1360* transposon[238] and are produced during transgene silencing in flies[239]; and siRNAs have been proposed to repress germ-line expression of *suffix*, a short interspersed nuclear element (SINE)[240].

The defining properties of *Drosophila* siRNAs are their production from long double-stranded RNA by Dicer-2 (Dcr-2), which generates 5´-monophosphate termini;

their loading into Argonaute2 (Ago2); and their Ago2-dependent, 3´ terminal, 2´-*O*-methylation by the methyltransferase Hen1[83,84,182], unlike most miRNAs[241]. In vivo (Fig. 1A, rightmost panel) and in vitro[228], nearly all siRNAs produced by Dcr-2 from exogenous dsRNA are 21 nucleotides (nt) in length.

Results

High throughput pyrosequencing reveals endo-siRNAs in soma

We characterized the somatic small RNA content of S2 cells[242] and of heads expressing an RNA hairpin silencing the *white* gene by RNAi[243]. To identify endo-siRNA candidates, we analyzed two types of RNA libraries. For total 18–30 nt RNA libraries, 89% (S2 cells) and 96% (heads) mapped to annotated miRNA loci. In contrast, libraries enriched for small RNAs bearing a 3´ terminal, 2´-*O*-methyl modification[244] were depleted of miRNAs: only 19% (S2 cells) and 49% (heads) of reads and 2.4% (S2 cells; 58,681 reads; 12,036 sequences) and 12% (heads; 22,685 reads; 2,929 sequences) of unique sequences mapped to miRNA loci.

Figure 1 shows the length distribution and sequence composition of the four libraries. The total RNA samples were predominantly miRNAs, a bias reflected in their modal length (22 nt) and pronounced tendency to begin with uracil. Excluding miRNAs, revealed a class of small RNAs with a narrow length distribution and no tendency to begin with uracil. Except for an unusual cluster of X chromosome small RNAs (Fig. S1) and a miRNA-like sequence with an unusual putative precursor on chromosome 2 (Fig. S2), few of these small RNAs are likely to correspond to novel miRNAs: none lie in the

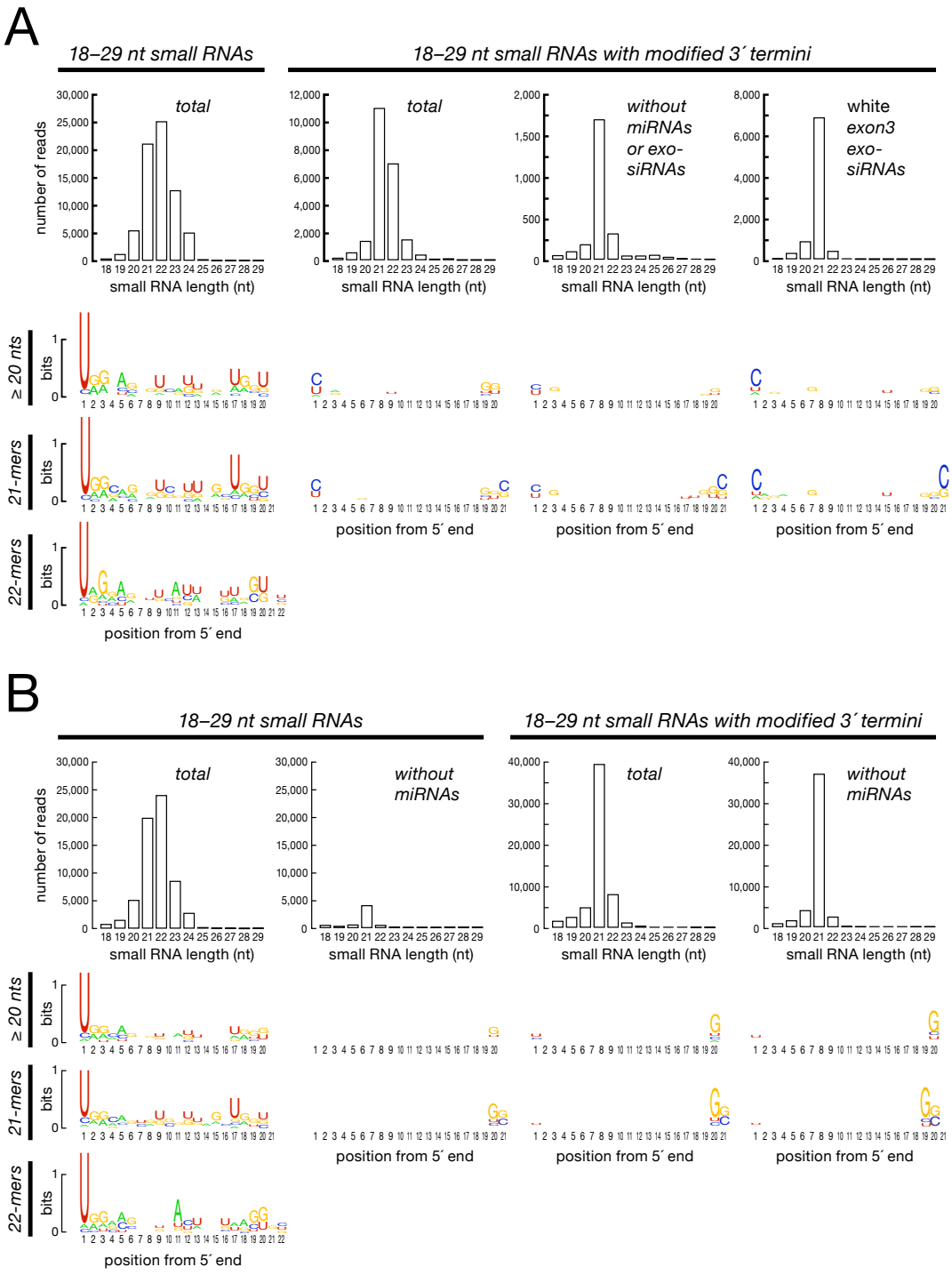
arms of hairpins predicted to be as stable as most pre-miRNAs (i.e., < -15 kcal/mol). However, the small RNAs derived form these clusters were indeed endo-siRNAs derived from structured loci which can fold to form hairpin shaped precursors, albeit much longer in length that miRNA precursors (see Discussion)[27-29]. These were missed in our study, due to folding of only short genomic precursors (Figs. S1 and S2).

After excluding known miRNAs, 64% for heads (Fig. 1A) and 78% for S2 cells of sequences in the libraries enriched for 3´ terminally modified small RNAs—i.e., those likely to be Ago2-associated—were 21 nt long (Fig. 1B). For fly heads, 37% (8,404 reads) derived from the *white* dsRNA hairpin. The abundance of these exo-siRNAs can be estimated by comparing them to the number of reads for individual miRNAs in the total small RNA library, where 1.6% (660 antisense and 491 sense reads) were 21-mers and matched the *white* sequences in the dsRNA- expressing transgene. The collective abundance of all *white* exo-siRNAs was less than the individual abundance of the ten most abundant miRNAs in this sample; the median abundance of any one exo-siRNA species was 2 reads. The *white*-IR transgene phenocopies a nearly null mutation in *white*, yet the sequence of the most abundant exo-siRNA was read just 37 times.

In heads, the sequence composition of the 21 nt, 3´ terminally modified small RNAs closely resembled that of exo-siRNAs, which tended to begin and end with cytosine. In heads and S2 cells, the 21-mers lacked the sequence features of piRNAs, which either begin with uracil (Aub- and Piwi-bound) or contain an adenine at position 10 (Ago3-bound) and are 23–29 nt long. These data suggest that the 21 nt small RNAs are somatic endo-siRNAs.

Figure II-1. High throughput pyrosequencing revealed 3´ terminally modified, 21-nt RNAs in the fly soma. (A) Length and sequence composition of the small RNA sequences from a library of total small RNA from the heads of flies expressing an inverted repeat (IR) silencing the *white* gene and for a parallel library enriched for RNAs modified at their 3´ ends. (B) Similar analysis for small RNA sequences from *Drosophila* S2 cells. Without miRNAs: pre-miRNA matching sequences were removed computationally.

Figure II-1.


A
18–29 nt small RNAs
18–29 nt small RNAs with modified 3´ termini
total
total
without miRNAs or exo-siRNAs
white exon3 exo-siRNAs
number of reads
small RNA length (nt)
≥ 20 nts
21-mers
22-mers
bits
position from 5´ end
B
18–29 nt small RNAs
18–29 nt small RNAs with modified 3´ termini
total
without miRNAs
total
without miRNAs
number of reads
small RNA length (nt)
≥ 20 nts
21-mers
22-mers
bits
position from 5´ end

Endo-siRNAs correspond to transposons and mRNAs.

In S2 cells, endo-siRNAs mapped largely to transposons (86%); in fly heads they mapped about equally to transposons, intergenic and unannotated sequences, and mRNAs. 41% mapped to mRNAs without mapping to transposons, suggesting that endo-siRNAs may regulate mRNA expression. Endo-siRNAs mapping to mRNAs were > 10-fold more likely than expected by chance (5.22 x10-161 < *p*-value < 8x10-151) to derive from genomic regions annotated to produce overlapping, complementary transcripts (Table 1 and Table S1). These data suggest that such overlapping, complementary transcripts anneal in vivo to form dsRNA that is diced into endo-siRNAs. We note that among the mRNAs for which we detected complementary 21-mers was *ago2* itself.

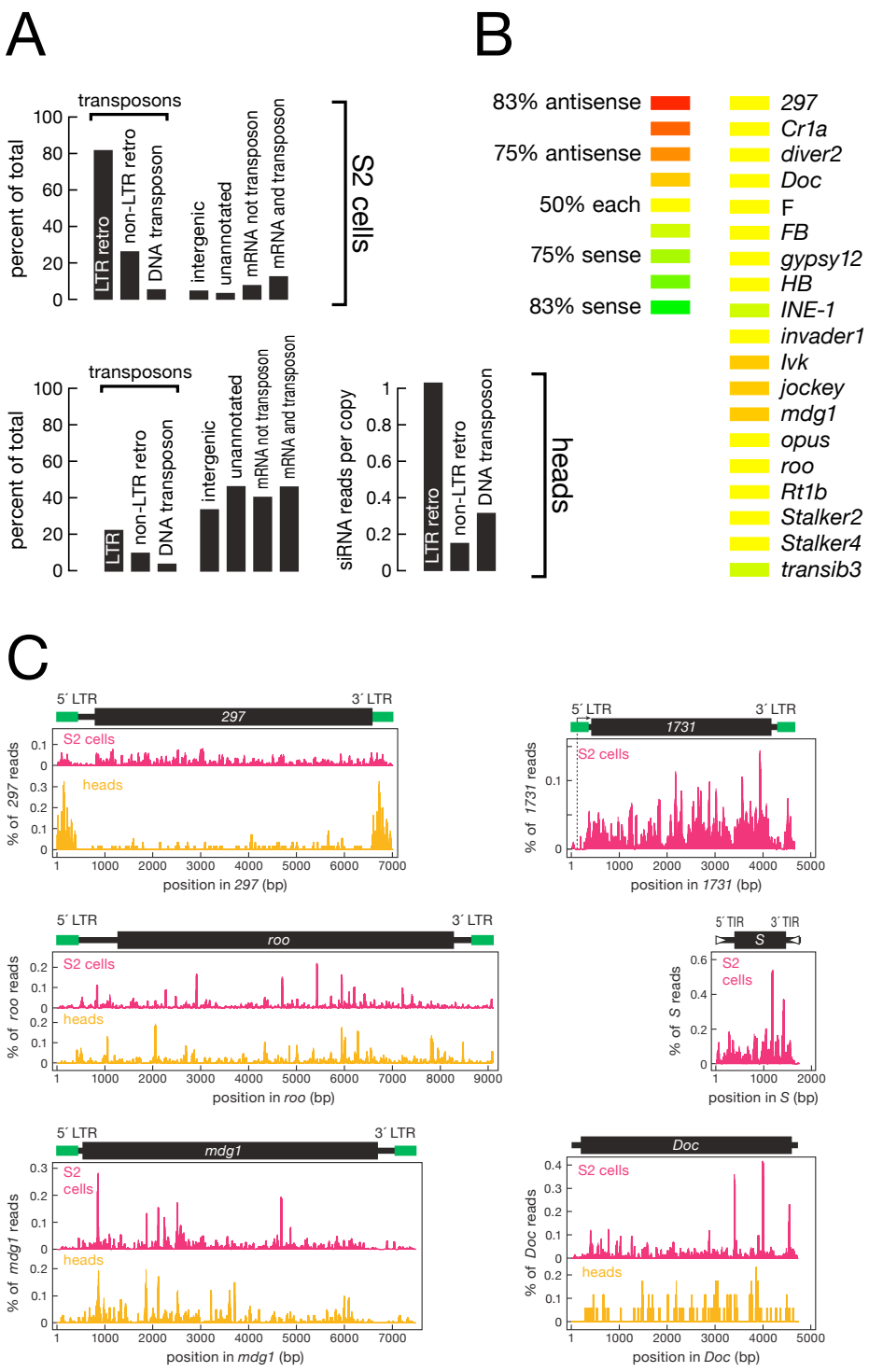
Table II-1. Endo-siRNAs preferentially map to overlapping, complementary mRNAs.

sample	enrichment	enrichment after randomization		*Z*-score	*p*-value
		mean	standard deviation		
Fly heads	10.9	1.0	0.38	26.1	7.9×10^{-151}
S2 cells	12.3	1.1	0.42	27.0	5.2×10^{-161}

Endo-siRNAs mapped to all three large chromosomes (Figs. S3, S4, and S5). siRNAs corresponding to the three transposon types in *Drosophila* were detected, but long terminal repeat (LTR) retrotransposons, the dominant class of selfish genetic elements in flies, were over-represented even after accounting for their abundance in the genome (Fig. 2A; Table S2). Unlike piRNAs, which are disproportionately antisense to transposons, but like siRNAs derived from exogenous dsRNA, about equal numbers of sense and antisense transposon-matching endo-siRNAs were detected (Fig. 2B and Fig. S6[22,43,180,195]. Like piRNAs, endo-siRNAs map to large genomic clusters (Table S3). Of 172 endo-siRNA clusters in S2 cells, four coincided with previously identified piRNA clusters (cluster #1, at 42A of chromosome 2R; clusters 7 and 10 in unassembled genomic sequence; and cluster #15 in the chromosome 3L heterochromatin). In heads, we detected 17 clusters; five corresponded to clusters found in S2 cells, but only one was shared with the germ-line piRNAs: the *flamenco* locus, consistent with recent genetic evidence that a Piwi-independent but *flamenco*-dependent pathway represses the *Idefix* and *ZAM* transposons in the soma[192]. That both endo-siRNAs and piRNAs can arise from the same region suggests either that a single transcript can be a substrate for both piRNA and siRNA production or that distinct classes of transcripts arise from a single locus.

Figure II-2. Endo-siRNAs correspond to transposons. (A) Distribution of annotations for the genomic matches of endo-siRNA sequences. Bars total more than 100 percent because some siRNAs match both LTR- and non-LTR retrotransposons or match both mRNA and transposons. (B) Transposon-derived siRNAs with more than fifty 21-nt reads mapped about equally to sense and antisense orientations. (C) Alignment of endo-siRNA sequences to *Drosophila* transposons. The abundance of each sequence is shown as a percentage of all transposon-matching siRNA sequences. LTR, long terminal repeat; TIR, terminal inverted repeat. Here and in subsequent figures, data from high throughput pyrosequencing and sequencing-by-synthesis were pooled for wild-type heads.

Figure II-2.

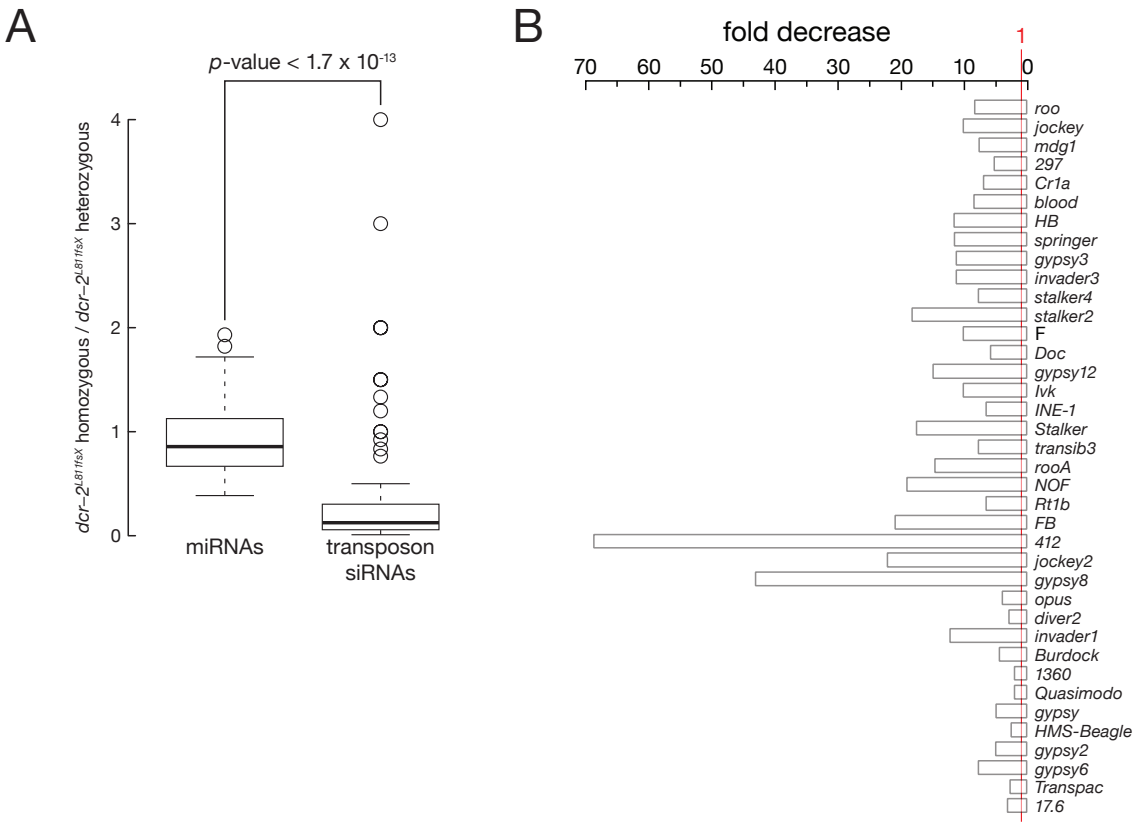

A
B
C
transposons
percent of total
LTR retro
non-LTR retro
DNA transposon
intergenic
unannotated
mRNA not transposon
mRNA and transposon
S2 cells
heads
siRNA reads per copy
83% antisense
75% antisense
50% each
75% sense
83% sense
297
Cr1a
diver2
Doc
F
FB
gypsy12
HB
INE-1
invader1
Ivk
jockey
mdg1
opus
roo
Rt1b
Stalker2
Stalker4
transib3
5´ LTR
3´ LTR
S2 cells
heads
position in 297 (bp)
position in 1731 (bp)
position in roo (bp)
5´ TIR
3´ TIR
position in S (bp)
position in mdg1 (bp)
position in Doc (bp)

Endo-siRNAs are Dcr-2 dependent

Statistically significant reductions in siRNA abundance were observed in *dcr-2*^{L811fsX} null mutant heads relative to heads from heterozygous siblings for 38 transposons (Fig. 3 and Table S4). Normalized for sequencing depth, sequencing results from homozygous *dcr-2* mutant heads yielded 3.1 times fewer 21-mers overall and 6.3 times fewer 21-mers corresponding to transposons than their heterozygous siblings (*p*-value < 2.2 x 10-16; chi-squared test). In contrast, overall miRNA abundance—normalized to sequencing depth—was essentially unchanged between *dcr-2* heterozygotes and homozygotes (Fig. 3 and Table S5). These data suggest that endo-siRNAs are produced by Dcr-2, but we do not yet know why some endo-siRNAs persist in *dcr-2*^{L811fsX} mutants.

Figure II-3. Transposon-matching siRNAs, but not miRNAs, are significantly changed in heads from *dcr-2*^{L811fsX} homozygous flies, compared to their heterozygous siblings (*dcr-2*^{L811fsX}/CyO). (A) Box plots for the ratio of reads for all miRNAs and transposon-matching siRNAs, normalized to sequencing depth, for the two genotypes. Only miRNAs whose sequence was read ≥ 100 times in at least one of the two genotypes were evaluated. Because miRBase does not always report the most abundant isoform of each miRNA, up to 9 nts were tolerated between the termini of each observed miRNA read and the miRBase entry, provided the miRNA matched the pre-miRNA perfectly. *p*-value calculated using Wilcoxon test. (B) The fold decrease for transposon-derived siRNAs for which ≥ 20 reads were detected in *dcr-2*^{L811fsX}/CyO. The changes of all transposons were statistically significant (*p*-value < 0.029, Fisher's exact test); the *p*-values for the change in individual miRNA and siRNA abundance are listed in Tables S4 and S5.

Figure II-3.


A
p-value < 1.7 x 10-13
dcr–2L811fsX homozygous / dcr–2L811fsX heterozygous
0
1
2
3
4
miRNAs
transposon siRNAs
B
fold decrease
1
70
60
50
40
30
20
10
0
roo
jockey
mdg1
297
Cr1a
blood
HB
springer
gypsy3
invader3
stalker4
stalker2
F
Doc
gypsy12
Ivk
INE-1
Stalker
transib3
rooA
NOF
Rt1b
FB
412
jockey2
gypsy8
opus
diver2
invader1
Burdock
1360
Quasimodo
gypsy
HMS-Beagle
gypsy2
gypsy6
Transpac
17.6

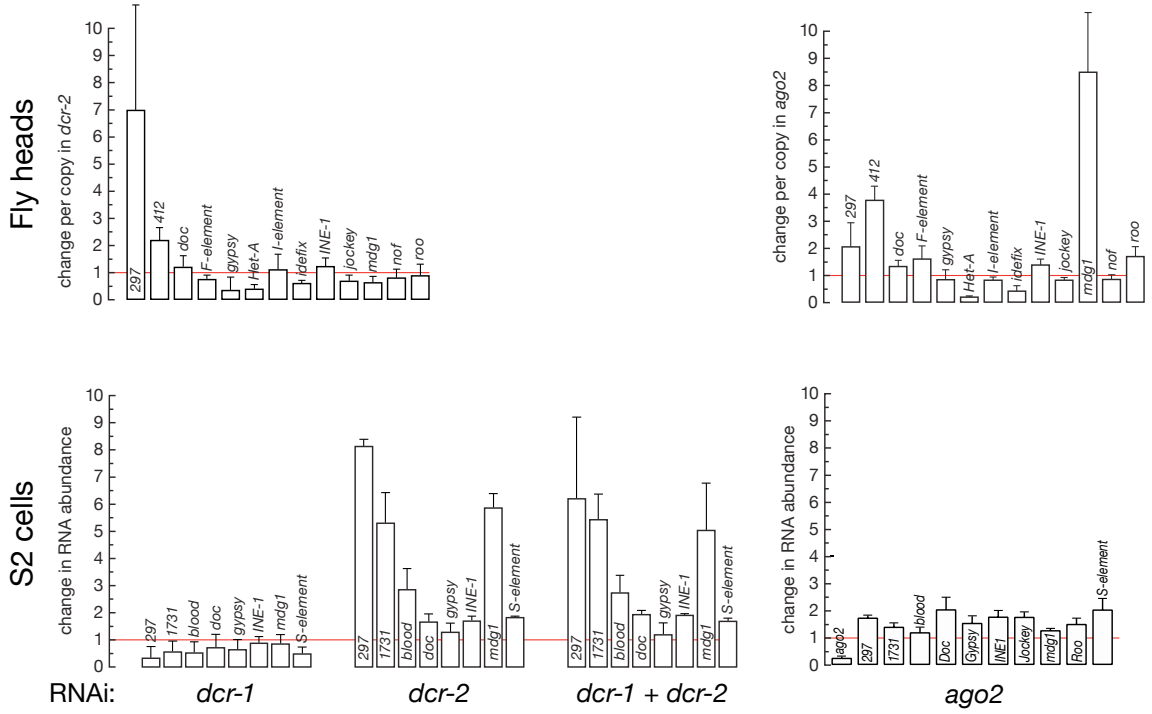
Transposon silencing requires Dcr-2 and Ago2

Transposon expression in the soma reflects both the silencing of transposons—potentially by either or both post-transcriptional and transcriptional mechanisms—and the tissue specificity of transposon promoters. *Drosophila* somatic cells may contain siRNAs targeting transposons that would not be highly expressed even in the absence of those siRNAs, because the promoters of those transposons are not active in some or all somatic tissues or because they are repressed by additional mechanisms. We analyzed the expression of a panel of transposons in heads from *ago2* and *dcr-2* mutants and in S2 cells depleted of Dcr-1, Dcr-2, or Ago2 by RNAi (Fig. 4). We found that the steady-state abundance of RNA from the LTR-retrotransposons *297* and *412* increased in heads from *dcr-2*^{L811fsX} null mutants (Fig. 4A). Similarly, the steady-state abundance of RNA from the LTR-retrotransposons, *297*, *412*, *mdg1*, and *roo*, the non-LTR retrotransposon, *F-element*, and the SINE-like element *INE-1* increased in *ago2*$_{414}$ mutant heads (Fig. 4B).

In S2 cells, RNA expression from the LTR-retrotransposons *297*, *1731*, *mdg1*, *blood*, and *gypsy*, and the DNA transposon, *S-element*, all increased significantly ($0.00001 < p\text{-value} < 0.002$) when Dcr-2 or Dcr-2 and Dcr-1 together, but not Dcr-1 alone, was depleted (Fig. 4C). *ago2(RNAi)* in S2 cells similarly desilenced transposons, including nine LTR- and non-LTR retrotransposons and the DNA transposon, *S-element* (Fig. 4).

Figure II-4. Transposon silencing requires Dcr-2 and Ago2, but not Dcr-1. The change in mRNA expression (mean ± SD, $N = 3$) for each transposon between *dcr-2*^{L811fsX} (A) or *ago2*^{414} (B) heterozygous and homozygous heads was measured by qRT- PCR. The data were corrected for differences in transposon copy number between the paired genotypes. (C) The change in transposon expression (mean ± SD, $N = 3$) in S2 cells was measured for the indicated RNAi depletion, relative to a control dsRNA.

Figure II-4.

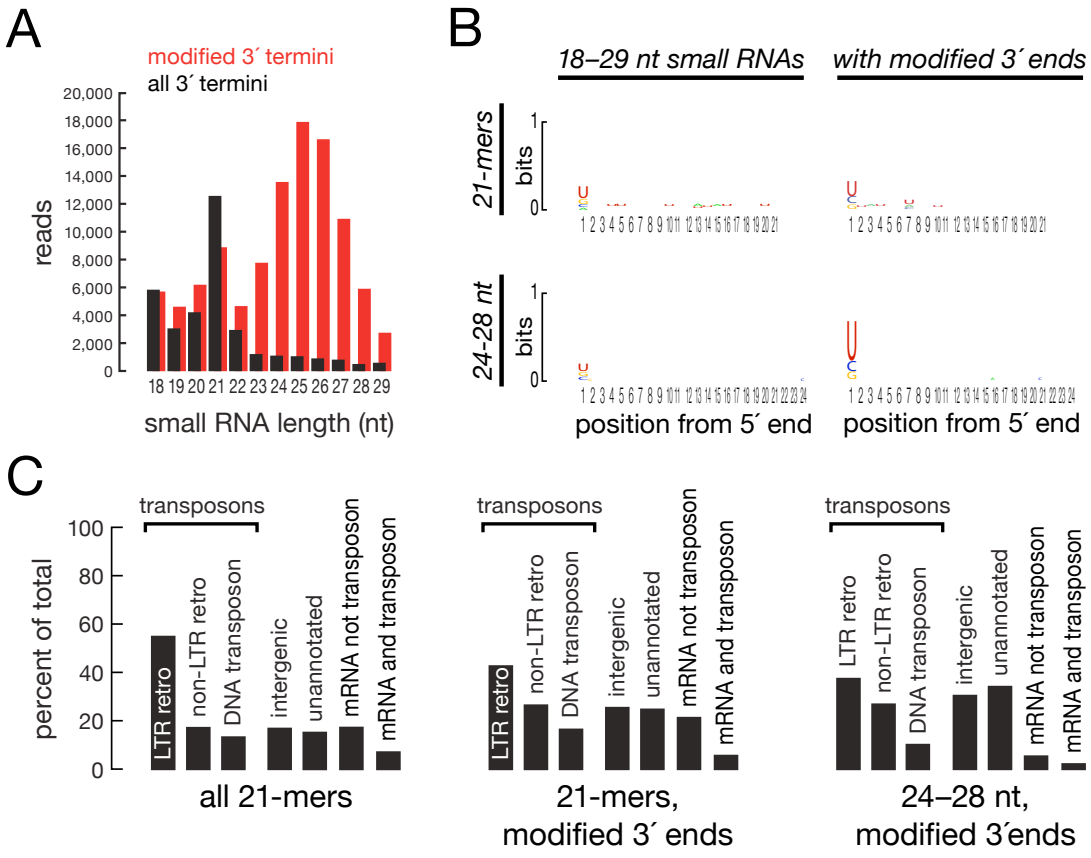

Fly heads
change per copy in dcr-2
297
412
doc
F-element
gypsy
Het-A
I-element
idefix
INE-1
jockey
mdg1
nof
roo
change per copy in ago2
297
412
doc
F-element
gypsy
Het-A
I-element
idefix
INE-1
jockey
mdg1
nof
roo
S2 cells
change in RNA abundance
297
1731
blood
doc
gypsy
INE-1
mdg1
S-element
RNAi:
dcr-1
dcr-2
dcr-1 + dcr-2
ago2
Dago2
Dblood
Doc
Gypsy
INE1
Jockey
Roo

The composition of somatic small RNAs is altered in the absence of Ago2

Is Ago2 required for the production or accumulation of endogenous siRNAs? We sequenced 18–29 nt small RNAs from *ago2*^{414} homozygous fly heads and the same small RNA treated to enrich for 3′ terminally modified RNAs. After computationally removing miRNAs, the sequences from the untreated library contained a prominent 21 nt peak (Fig. 5A) that predominantly began with uracil (Fig. 5B), much like miRNAs and unlike siRNAs in wild-type heads, which often began with cytosine (Fig. 1A). Perhaps in the absence of Ago2, only a subpopulation of endo-siRNAs that can bind Ago1 accumulates. The small RNAs from the *ago2*^{414} library enriched for 3′ terminally modified sequences were predominantly 24–27 nt long and often began with uracil, a length distribution and sequence bias characteristic of piRNAs, which, like siRNAs, are 2′-*O*-methylated at their 3′ ends. Both the 21-nt small RNAs and the piRNA-like RNAs in the *ago2* mutant heads mapped to transposons, unannotated heterochromatic and unassembled sequences, but the piRNA-like sequences mapped to mRNAs far less frequently than either the 21-mers or wild-type endo-siRNAs (Fig. 5C). How these piRNA-like small RNAs are generated and if they contribute to transposon silencing in the fly soma remains to be answered.

Figure II-5. The composition of somatic small RNAs is altered in the absence of Ago2. Size distribution (A) and sequence composition (B) of sequences from a library of total 18–29 nt RNA from the heads of *ago2* null mutant flies or a library enriched for 3´ terminally modified RNAs. Reads matching pre-miRNA sequences were removed. (C) Distribution of annotations for the genomic matches of small RNA sequences from the two *ago2* libraries.

Figure II-5.


A
modified 3′ termini
all 3′ termini
reads
small RNA length (nt)
B
18–29 nt small RNAs
with modified 3′ ends
21-mers
24-28 nt
bits
position from 5′ end
C
transposons
percent of total
LTR retro
non-LTR retro
DNA transposon
intergenic
unannotated
mRNA not transposon
mRNA and transposon
all 21-mers
21-mers, modified 3′ ends
24–28 nt, modified 3′ends

Discussion

The abundance and distribution of endo-siRNAs across the sequences of individual transposon species reflected when the elements entered the fly genome, but not their mechanism of transposition (Fig. 2C). The retrotransposon *297* (80 copies per haploid genome) is the second most abundant retroelement in flies. *297* entered *Drosophila* recently through the ancestor of the *melanogaster* species group 44 million years ago[245]. Compared to flies, *297* has expanded dramatically in S2 cells[246]. *297* matching siRNAs represent 29.2% of all endo-siRNAs in S2 cells, but only 3.3% of endo-siRNAs in heads (Table S2). Remarkably, many of the siRNAs that correspond to *297* in heads map to its LTRs (Fig. 2C). It is difficult to imagine that antisense transcription arising in an adjacent protein-coding gene or an adjacent transposon could produce a precursor dsRNA that would lead to the production of siRNAs so tightly constrained to the LTR sequences. The LTRs of retrotransposons are direct repeats, so intramolecular pairing between LTRs within an RNA transcript—as has been proposed for the terminal inverted repeats (TIRs) of the DNA transposon *TC1* in *C. elegans*[32]—also cannot explain the peculiar pattern of siRNA production from *297*. Perhaps endo-siRNAs arise from an orphaned *297* LTR sequence in flies, but from one or more complete *297* elements in S2 cells. Moreover, somatic siRNAs are not generally confined to specific regions of the other transposons examined (Fig. 2C). (Notably, the endo-siRNAs derived from the DNA transposon, *S-element*, do not appear to arise from intramolecular base-pairing between the complementary 5´ and 3´ TIRs, as occurs for *TC1* in *C. elegans*[32].) The *1731* element has also expanded in S2 cells, from a single active copy in the fly to many highly active

copies in the cultured cell line[247]. Our endo-siRNA data reflects this expansion: *1731* matching siRNAs represent 39% of all endo- siRNAs in S2 cells, but only 0.02% in fly heads, where we found only a single *1731*- matching siRNA (*p*-value < 2.2 x 10-16, chi-square test).

Fly endo-siRNAs are not only limited to transposons, but are also derived from heterochromatic sequences, intergenic regions, long RNA transcripts with extensive structure, and, most interestingly, from mRNAs (Fig. 6). siRNAs derived from mRNAs are >10 times more likely to come from regions predicted to produce overlapping, convergent transcripts than expected by chance, suggesting that endo-siRNAs originate from endogenous dsRNA formed when these complementary transcripts pair.

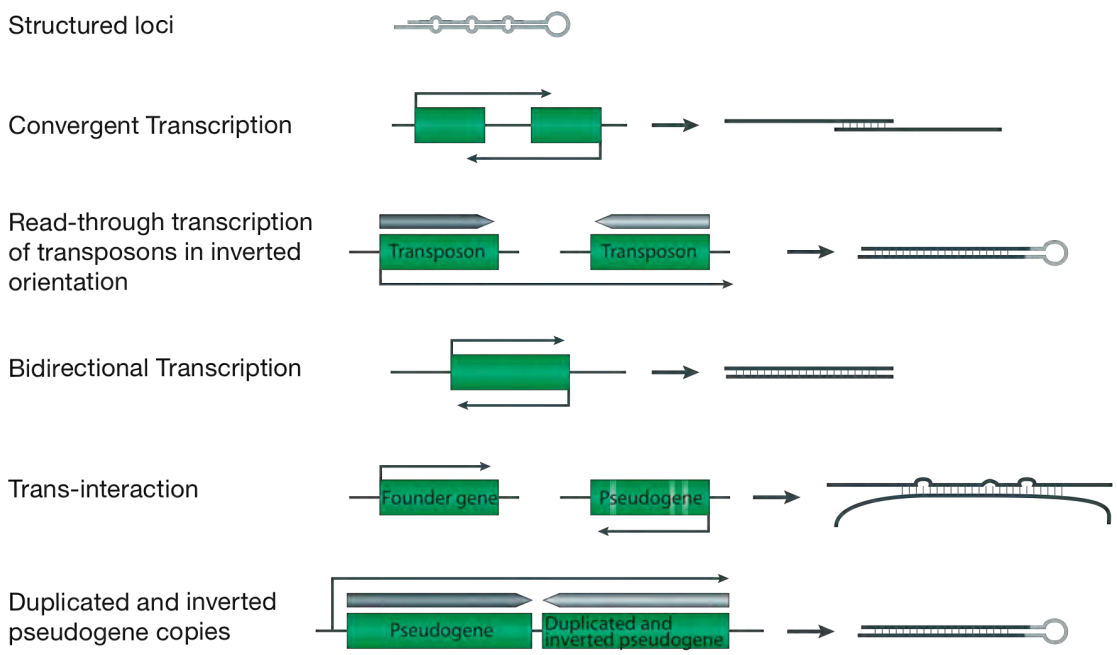
A subset of fly endo-siRNAs derive from "structured loci" whose RNA transcripts can fold into long, intramolecularly paired hairpins[27-29]. Accumulation of these siRNAs requires Dcr-2 and the dsRNA-binding protein Loquacious (Loqs)—typically considered the partner of Dcr-1, the dicer that produces miRNA—rather than R2D2[248], the usual partner of Dcr-2. While surprising, a role for Loqs in the biogenesis of endo-siRNAs from structured loci was anticipated by the earlier finding that Loqs plays a role in the production of siRNAs from transgenes designed to produce long, intramolecularly paired inverted repeat transcripts so as to trigger RNAi in flies[134].

Endo-siRNAs have also been identified in mouse oocytes and a subset of them are derived from pseudogenes[33,34]. Taking all the endo-siRNA studies into consideration, we can speculate about various potential precursors for endo-siRNAs. As flies and mammals

don't have a RdRP, the endo-siRNA precursors are genomic loci that can form dsRNA structures, which can then act as substrates for Dcr-2 (Fig. 6).

Figure II-6. Genomic Sources of dsRNA triggers for endo-siRNAs in flies and mammals. siRNAs are derived from dsRNA precursors. Endo-siRNAs can arise from structured loci that can pair intramolecularly to produce long dsRNA, complementary overlapping transcripts, and bidirectionally transcribed loci. Endo-siRNAs may also originate form protein-coding genes that can pair with their cognate pseudogenes and from regions of pseudogenes that can form inverted-repeat structures.

Figure II-6.


Structured loci
Convergent Transcription
Read-through transcription of transposons in inverted orientation
Transposon
Transposon
Bidirectional Transcription
Trans-interaction
Founder gene
Pseudogene
Duplicated and inverted pseudogene copies
Pseudogene
Duplicated and inverted pseudogene

A key challenge for the future will be to understand the biological function of endo-siRNAs, especially those that can pair with protein-coding mRNAs. Do they regulate mRNA expression? Can endo-siRNAs act like miRNAs, tuning the expression of large numbers of genes? Recent evidence implied a role for endo-siRNAs in robust development of *Drosophila* embryo[249]. This study demonstrated a requirement for Dcr-2 and Ago2 for normal segmentation of embryos exposed to differential temperatures at their anterior and posterior halves. Ago2 has also been implicated in early embryogenesis, assembly of centric heterchromatin, nuclear division and migration, and germ-cell formation[250]. Moreover, in our lab we have observed that *ago2*, *dcr-2* or *r2d2* homozygotes are up to 5 times less observed than expected, relative to heterozygotes. Since, siRNA pathways mutants are viable but probably are less likely to hatch than their heterozygous couterparts, the requirement of the pathway may be manifested only under unfavorable environmental conditions, similar to what has been observed for miRNA-mediated gene regulation[251].

Materials and Methods

General methods

RNA was isolated as described[244] from heads of Oregon R flies or *white*-IR flies[243] or from Schneider 2 (S2) cells, a phagocytic, cultured cell line derived from late-stage *Drosophila* embryos[242]. S2 cells were a clonal cell line containing a stably integrated GFP transgene, pKF63, transiently transfected with dsRNA targeting GFP[203]. dsRNA was prepared[134] and transfected into S2 cells as described[83].

High throughput sequencing

High throughput pyrosequencing was as described[244]. Libraries were constructed using a method that selects for RNAs bearing 5´ monophosphates[4]. For pyrosequencing, the total small RNA libraries yielded 63,315 (S2 cells) and 71,268 (heads) reads corresponding to 4,971 (S2 cells) and 1,884 (heads) unique sequences. High throughput sequencing-by-synthesis (Genome Analyzer, Illumina, San Diego, CA, USA) was as for pyrosequencing except that RNA Ligase 2 [Rnl2(1-249)K227Q] (Addgene, Cambridge, MA, USA) was used for 3´ ligation. Linkers and primers for sequencing-by-synthesis were: 5´ adaptor, 5´-rGrUrU rCrArG rArGrU rUrCrU rArCrA rGrUrC rCrGrA rCrGrA rUrC-3´ (Dharmacon, Lafayette, CO, USA); 3´ preadenylated linkers, 5´-rAppdCdT dGdTdA dGdGdC dAdCdC dAdTdC dAdAdT ddC-3´.

After linker addition, the cDNA was synthesized using a reverse-transcriptase primer corresponding to the 3´ adapter and amplified by PCR using forward (5´-dAdAdT dGdAdT dAdCdG dGdCdG dAdCdC dAdCdC dGdAdC dAdGdG dTdTdC dAdGdA dGdTdT dCdTdA dCdAdG dTdCdC dGdA -3´) and reverse (5´-dCdAdA dGdCdA dGdAdA dGdAdC dGdGdC dAdTdA dCdGdA dAdTdT dGdAdT dGdGdT dGdCdC dTdAdC dAdG-3´) primers. The PCR pool was gel purified (4% Metaphor Agarose, Cambrex, East Rutherford, NJ, USA) with Qiaex II (Qiagen, Valencia, CA, USA) then sequenced (Genome Analyzer, Illumina) according to the manufacturer's protocol.

Quantitative RT-PCR analysis

Two micrograms of total RNA was treated with RQ1 DNase (Promega, Madison,

WI, USA) or Turbo DNase (Ambion, Austin, TX, USA) according to manufacturer's instructions and then reverse transcribed using oligo(dT) primer and Superscript III and Superscript II reverse transcriptases (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's directions. The resulting cDNA was analyzed by quantitative RT-PCR performed in a DNA Engine OPTICON 2 (MJ Research, Bio-Rad, Hercules, CA, USA) or an iQ5 (Bio-Rad, Hercules, CA, USA) instrument using a SYBR Green PCR kit (Qiagen or Bio-Rad) according to manufacturer's instructions. Relative steady-state mRNA levels were determined from the threshold cycle for amplification using the ΔΔCT method[252] or DART-PCR[253]. Table S6 lists the PCR primer sequences.

Computational analyses

For each transposon, reads mapping to at least one genomic copy of that transposon were aligned on the transposon consensus sequence using WU-BLAST (http://blast.wustl.edu/) at low stringency (word size, 1; expectancy threshold, 100). For each aligned read, the top-scoring segment pair was selected; if N segment pairs were equally high-scoring, they were all selected, and were weighted by 1/N (especially true for LTR-matching reads). When the segment pair alignment did not reach the extremities of the read, the alignment was extended in order to cover the complete read. Where reads are reported normalized to sequencing depth, the number of genome-matching reads was used for normalization. Total small RNA data sets correspond to all reads matching the *Drosophila* genome after excluding annotated non-coding RNAs such as ribosomal RNA, snRNAs, snoRNAs, etc. Other computational methods were as described[43]. Programs are

freely available upon request. Sequencing statistics are in Table S7.

Enrichment of endo-siRNAs in regions of overlapping transcripts

The annotated transcriptome (defined as the genomic regions of all annotated mRNAs, including exons and introns) was first divided into the regions that produced overlapping, complementary transcripts and regions that produce transcripts only from one strand. Then all allowable positions that can be the starting position of a non-transposon-overlapping 21-mer were separately determined for the plus and minus strands. The scope of double-stranded regions was defined as the union of the allowable positions for which the anti-sense positions are also allowable. The scope of single-stranded regions is defined as the union of the remaining allowable positions.

We then mapped the endo-siRNAs from wild-type fly heads or S2 cells onto the transcriptome and computed an enrichment score: [(total number of mapped endo-siRNAs whose 5´-end position falls in the scope of double-stranded regions)*(size of the scope of double-stranded regions + size of the scope of single-stranded regions)]/[(total number of endo-siRNAs)*(size of the scope of double- stranded regions)].

To determine the statistical significance of the resulting enrichment scores, we randomly selected the same number of allowable positions as the number of endo-siRNAs in the sample and recomputed the enrichment score, 100 times per sample. The random distribution had a mean ~1, as expected. The *p*-values of the actual enrichment scores of the two libraries were determined with reference to the normal

distribution. Results are summarized in Table 1.

Supplemental Material

Supplemental Figures

Figure II-S1. An unusual small RNA that maps to 17 (13 exact matches and 4 with one mismatch) stable hairpins on the X chromosome (ΔΔG = -22.90 kcal/mol). The small RNA sequence was enriched in the oxidized, β-eliminated library, suggesting it is 2´-*O*-methylated in cultured S2 cells. (A) The sequence of the 13 identical hairpins containing the unusual small RNA. Their extraordinary conservation may indicate a recent series of gene duplication events. (B) The genomic locations of the sequence on the minus strand of the X chromosome.

Note added in proof: The loci described here in Figs. S1 and S2 correspond to endo-siRNA–generating hairpins recently identified in [27-29].

Figure II-S1.

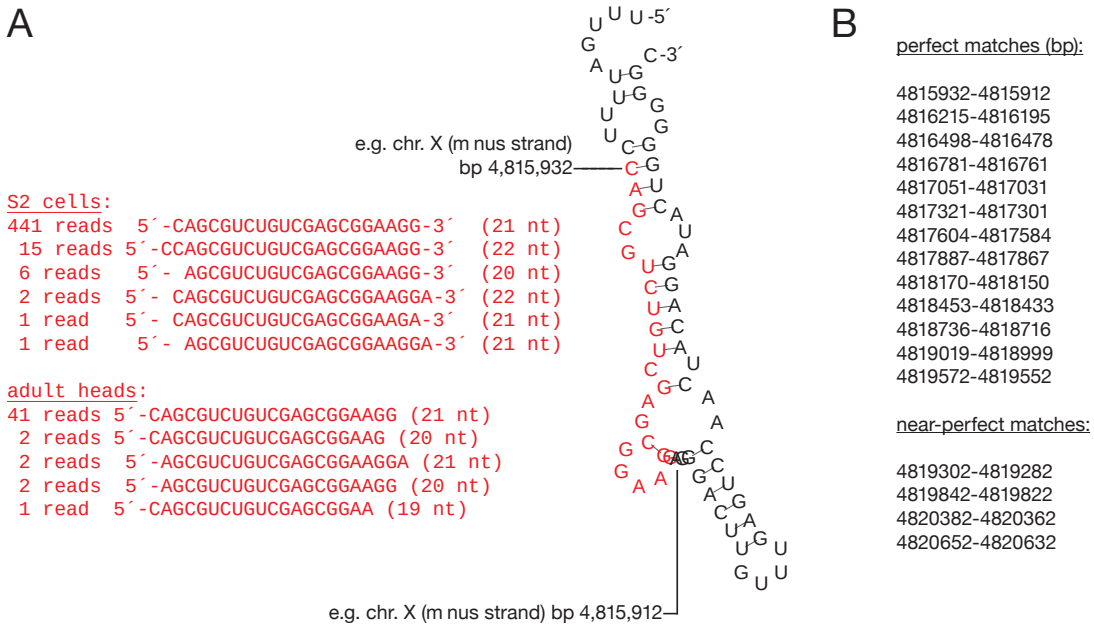
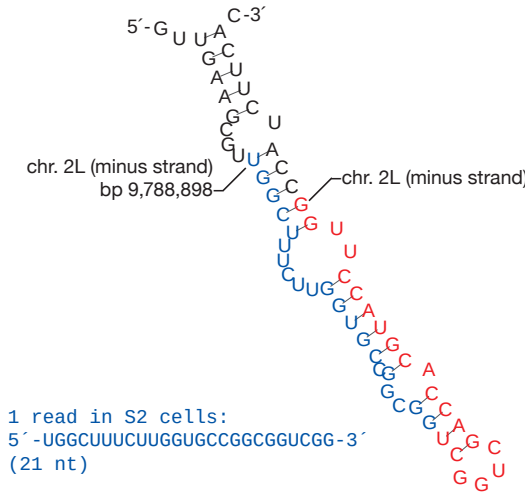

A
S2 cells:
441 reads 5´-CAGCGUCUGUCGAGCGGAAGG-3´ (21 nt)
15 reads 5´-CCAGCGUCUGUCGAGCGGAAGG-3´ (22 nt)
6 reads 5´- AGCGUCUGUCGAGCGGAAGG-3´ (20 nt)
2 reads 5´- CAGCGUCUGUCGAGCGGAAGGA-3´ (22 nt)
1 read 5´- CAGCGUCUGUCGAGCGGAAGA-3´ (21 nt)
1 read 5´- AGCGUCUGUCGAGCGGAAGGA-3´ (21 nt)
adult heads:
41 reads 5´-CAGCGUCUGUCGAGCGGAAGG (21 nt)
2 reads 5´-CAGCGUCUGUCGAGCGGAAG (20 nt)
2 reads 5´-AGCGUCUGUCGAGCGGAAGGA (21 nt)
2 reads 5´-AGCGUCUGUCGAGCGGAAGG (20 nt)
1 read 5´-CAGCGUCUGUCGAGCGGAA (19 nt)
e.g. chr. X (m nus strand) bp 4,815,932
e.g. chr. X (m nus strand) bp 4,815,912
B
perfect matches (bp):
4815932-4815912
4816215-4816195
4816498-4816478
4816781-4816761
4817051-4817031
4817321-4817301
4817604-4817584
4817887-4817867
4818170-4818150
4818453-4818433
4818736-4818716
4819019-4818999
4819572-4819552
near-perfect matches:
4819302-4819282
4819842-4819822
4820382-4820362
4820652-4820632

Figure II-S2. An unusual small RNA derived from a stable hairpin (ΔΔG = -24.20 kcal/mol) on chromosome 2L. The small RNA sequence was enriched in the oxidized, β-eliminated libraries, suggesting it is 2´-*O*-methylated in cultured S2 cells and adult fly heads.

Note added in proof: The loci described here in figs. S1 and S2 correspond to endo-siRNA–generating hairpins recently identified in[27-29].

Figure II-S2.



S2_cells
350 reads: 5´-UCGGUCGACCACGUACCUUGG-3´ (21 nt)
162 reads: 5´-CGGUCGACCACGUACCUUGG-3´ (20 nt)
81 reads: 5´-UCGGUCGACCACGUACCUUGGC-3´(22 nt)
1 read: 5´-GUCGGUCGACCACGUACCUUGG-3´ (22 nt)

Adult_Heads
300 reads: 5´-CGGTCGACCACGTACCTTGGC-3´ (21 nt)
98 reads: 5´-TCGGTCGACCACGTACCTTGGC-3´ (22 nt)
37 reads: 5´-TCGGTCGACCACGTACCTTGG-3´ (21 nt)
14 reads: 5´-CGGTCGACCACGTACCTTGG-3´ (20 nt)
13 reads: 5´-CGGTCGACCACGTACCTTGGCC-3´(22 nt)
10 reads: 5´-GGTCGACCACGTACCTTGGCC-3´(21 nt)
2 reads: 5´-TCGGTCGACCACGTACCTTGGCC-3´(23 nt)
2 reads: 5´-GGTCGACCACGTACCTTGGC-3´ (20 nt)

Figure II-S3. Endogenous siRNAs from adult fly heads. Small RNAs mapping to more than one genomic location were attributed to each site to which they were complementary, but normalized for the number of sites. piRNA data are from Brennecke et al. (*Cell* 2007). The figure was drawn using pooled wild-type head data as indicated in Table S7.

Figure II-S3.

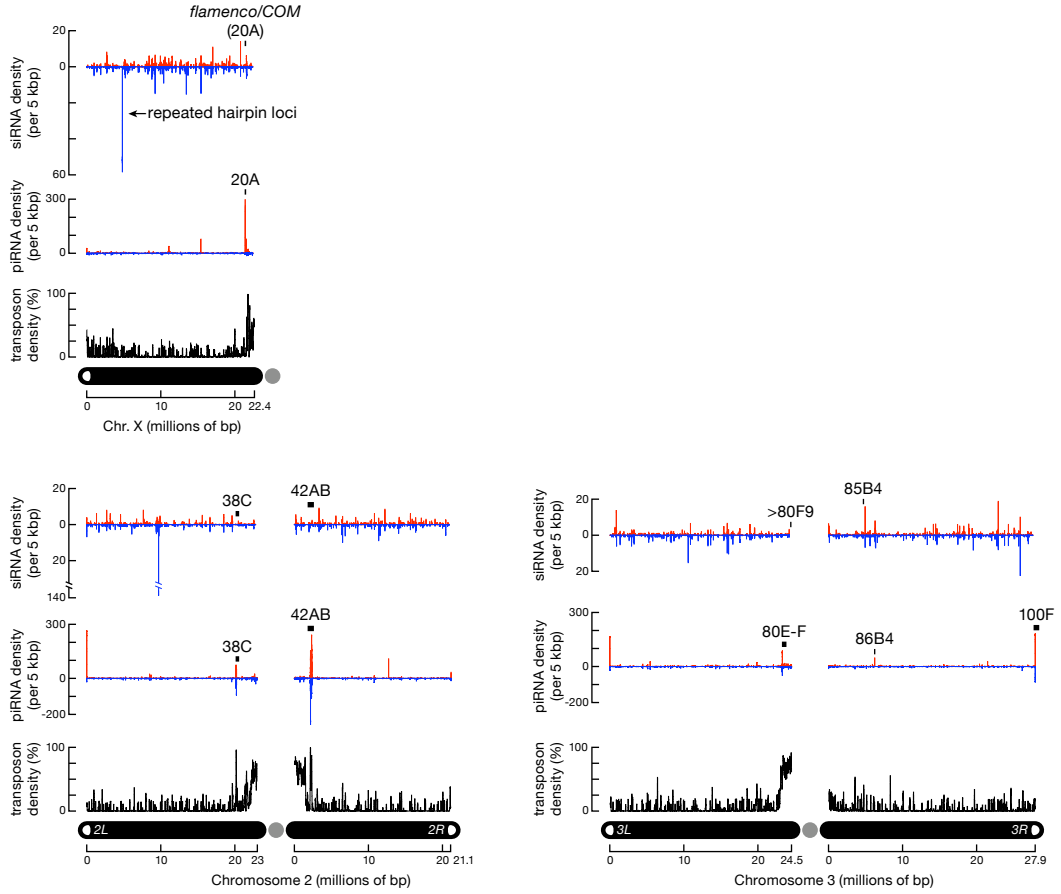

flamenco/COM
(20A)
siRNA density (per 5 kbp)
repeated hairpin loci
20A
piRNA density (per 5 kbp)
transposon density (%)
Chr. X (millions of bp)
38C
42AB
Chromosome 2 (millions of bp)
>80F9
85B4
80E-F
86B4
100F
Chromosome 3 (millions of bp)

Figure II-S4. Endogenous siRNAs from cultured S2 cells. Small RNAs mapping to more than one genomic location were attributed to each site to which they were complementary, but normalized for the number of sites. piRNA data are from Brennecke et al. (*Cell* 2007).

Figure II-S4.

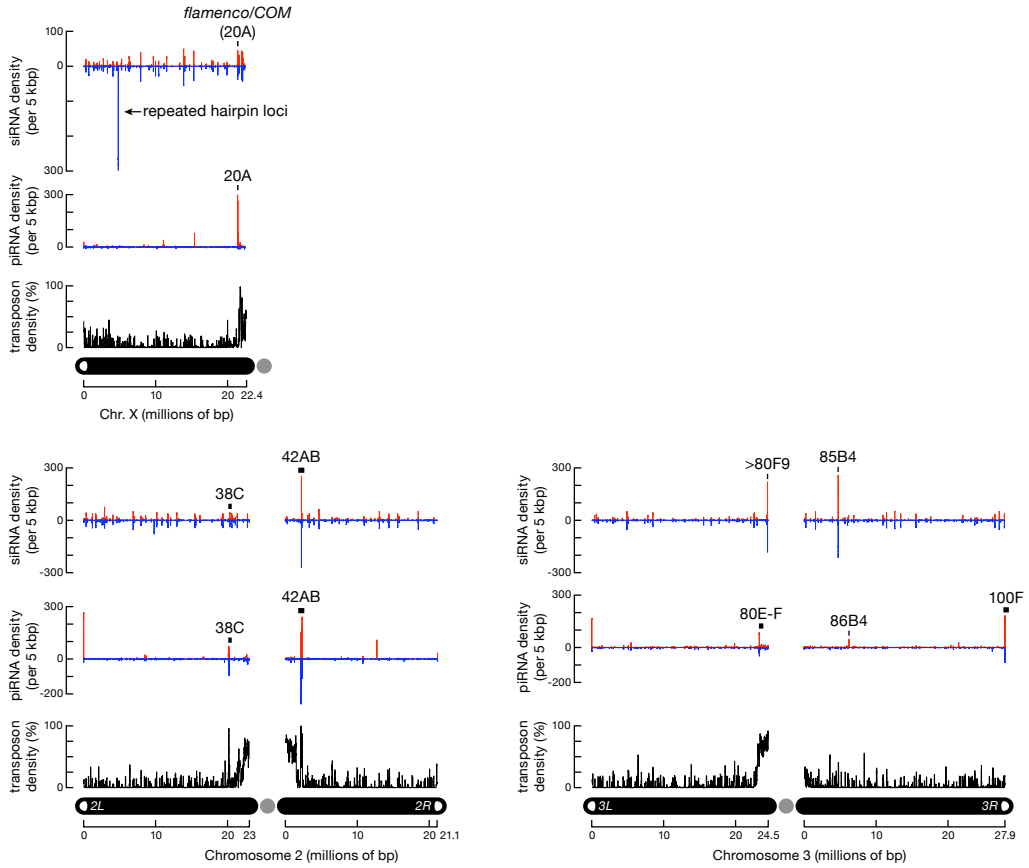

flamenco/COM
(20A)
siRNA density (per 5 kbp)
repeated hairpin loci
20A
piRNA density (per 5 kbp)
transposon density (%)
Chr. X (millions of bp)
42AB
38C
2L
2R
Chromosome 2 (millions of bp)
>80F9
85B4
80E-F
86B4
100F
3L
3R
Chromosome 3 (millions of bp)

Figure II-S5. Uniquely mapping endogenous siRNAs from cultured S2 cells.

Figure II-S5.

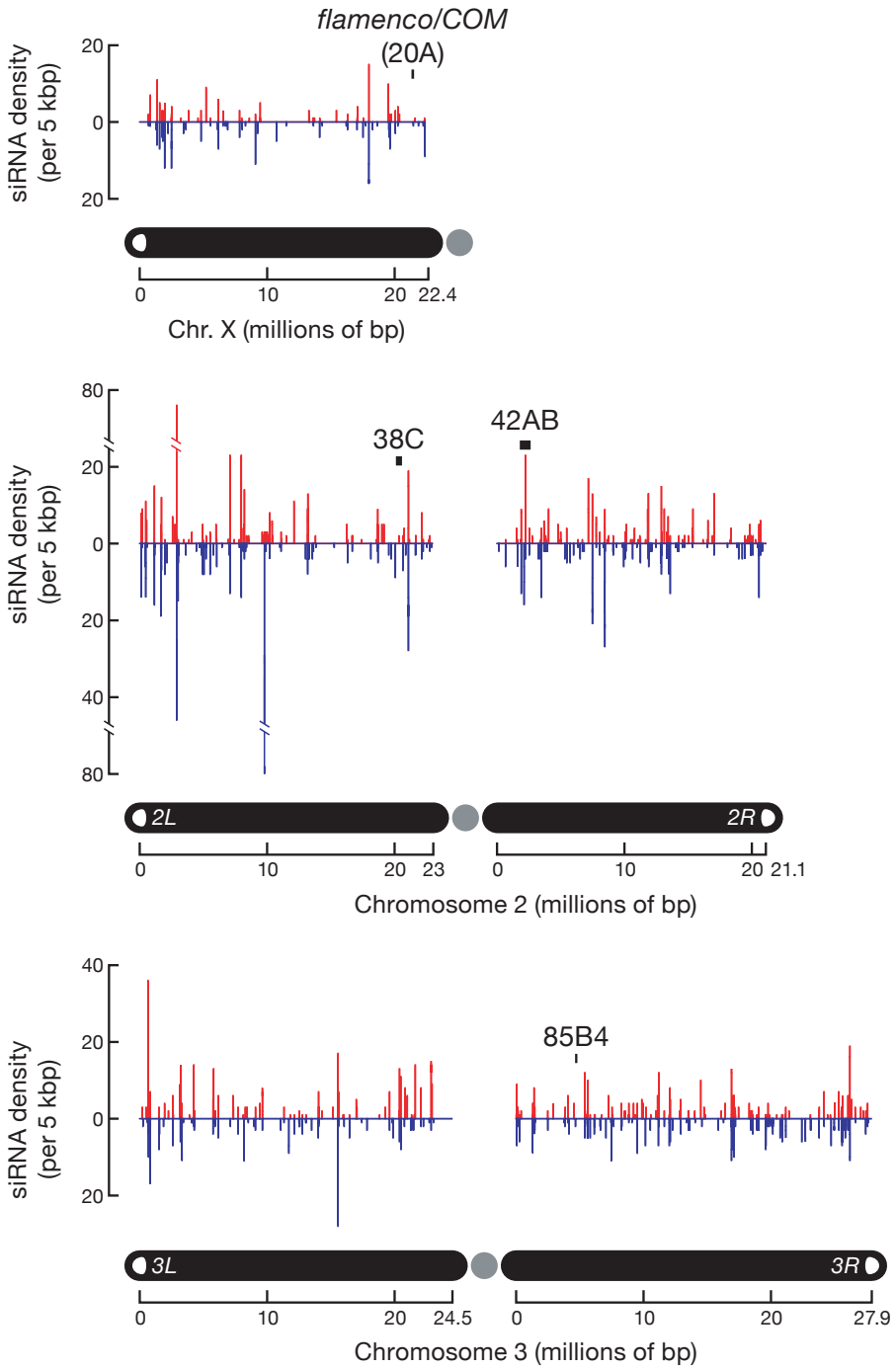
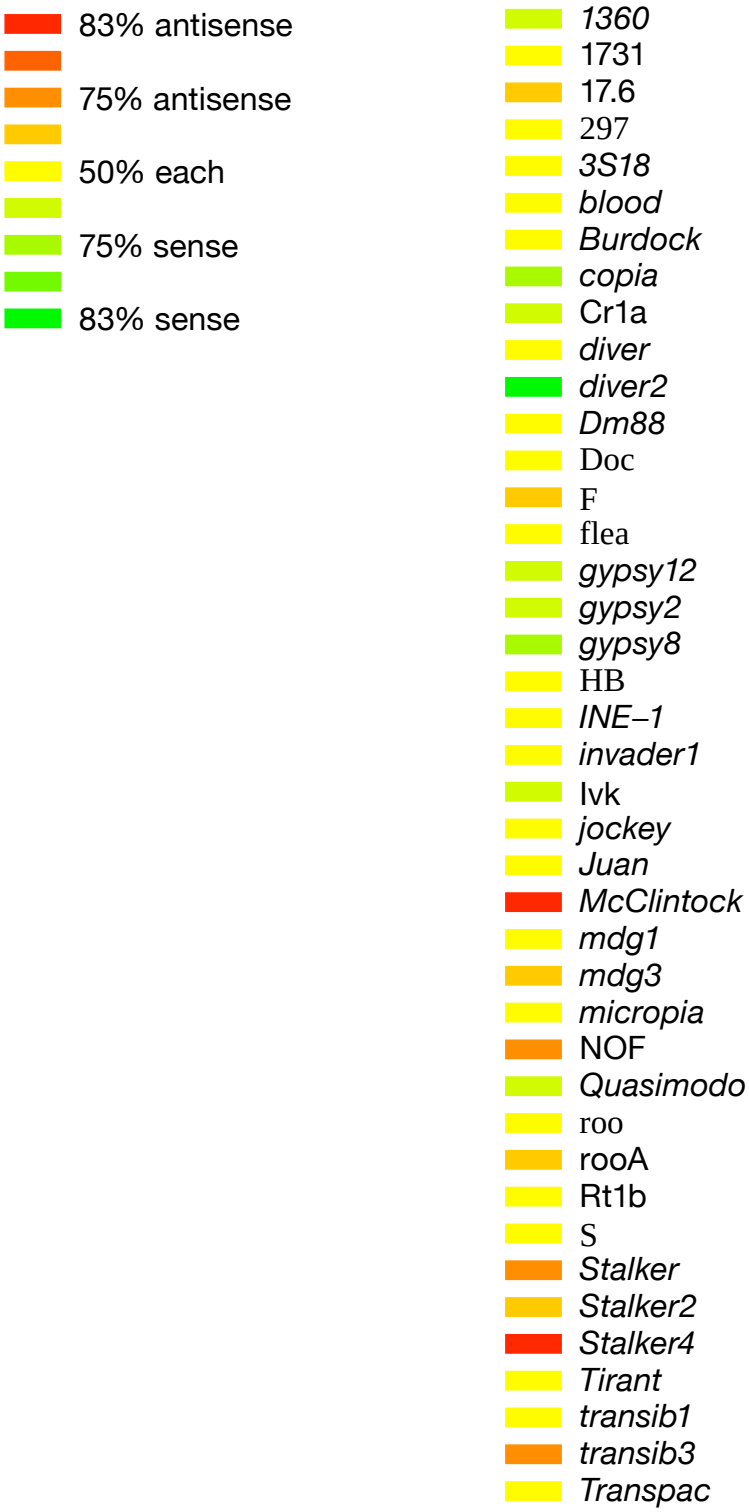

flamenco/COM
(20A)
siRNA density
(per 5 kbp)
20
0
20
0
10
20
22.4
Chr. X (millions of bp)
80
20
0
20
40
80
38C
42AB
2L
2R
0
10
20
23
0
10
20
21.1
Chromosome 2 (millions of bp)
40
20
0
20
85B4
3L
3R
0
10
20
24.5
0
10
20
27.9
Chromosome 3 (millions of bp)

Figure II-S6. In cultured S2 cells, transposon-derived siRNAs generally mapped about equally to sense and antisense orientations. Only transposons with more than fifty 21-nt reads were analyzed.

Figure II-S6.


83% antisense
75% antisense
50% each
75% sense
83% sense
1360
1731
17.6
297
3S18
blood
Burdock
copia
Cr1a
diver
diver2
Dm88
Doc
F
flea
gypsy12
gypsy2
gypsy8
HB
INE–1
invader1
Ivk
jockey
Juan
McClintock
mdg1
mdg3
micropia
NOF
Quasimodo
roo
rooA
Rt1b
S
Stalker
Stalker2
Stalker4
Tirant
transib1
transib3
Transpac

Supplemental Tables

Table II-S1A. mRNA-matching endo-siRNAs in cultured S2 cells. Data are from pyrosequencing of a small RNA library enriched for 3´ terminally modified RNA.

GENE	siRNA	orientation of small RNA	S2 cell reads
5Ptasel	ATATATCGCCCTGTCCCGAGG	sense	2
Aats-gln	TTTGCGTGACCGATGTGCAGG	antisense	1
Ac3	CTCGCGAACAAAGGTTATGTC	antisense	3
Acer	CAGTATTCCGCACCGAAAAGG	sense	1
Ack	TCCTCCTGCGCTCTGTTTTGC	antisense	1
Act42A	CGGCCGTCCACGTCACCAGGG	sense	12
Act42A	ATGGGGTACTTCAGGGTAAGG	antisense	7
Act42A	AATATGGTTTGCTTATGCGTC	antisense	1
Act42A	CTACAACTCAATCATGAAGTG	sense	1
Ada2b	TGTGGCTCTTAATCGAAGGGG	antisense	4
Ada2b	ATTGATTTTCAGTTTTGTAGT	antisense	1
Ada2b	TTGATGAAAATGCCAACGACA	antisense	1
ade2	TGAGTTTTAAAGTTGTTTGAG	antisense	1
ago(archipeligo)	CACCGTTCAAGGTATCCGTGG	antisense	14
AGO2	CCTGACCTTCCTCGATGCTGC	antisense	8
AGO2	GTTGGAAAGCTTATAATGGAG	sense	7
AGO2	TGGCGGACCATCTCAAGGCGG	antisense	7
AGO2	TTAAAAGCCGCCTTGAGATGG	sense	7
AGO2	TGGAATCAATAGAGATGCTCC	antisense	3
AGO2	TGTCCTAAAATGCCACAAACA	antisense	3
AGO2	TTGGAAAGCTTATAATGGAGT	sense	3
AGO2	CTCCATTATAAGCTTTCCAAC	antisense	1
AGO2	TTAATATTCCTAAAAGAAAGG	antisense	1
AnnIX	TAAGGATTTCCTCGTTGGATC	sense	5
AnnIX	AACAGGATGCGAGACTGGGGT	antisense	3
AnnIX	TTTTGCGGAAGATTCATAGCC	sense	1
AnnX	TGGTTGCTCCTGCCGACGAGC	antisense	3
Aos1	TGGTTACTAAATTGGAGGCGC	sense	3
Apc2	AACTATAGGAAAATTGAGACC	antisense	1
Arf79F	CTTATGGGTTGGTGAAATGCG	antisense	2
Arf79F	TCGGCTCGCTTGAAATCAGAG	antisense	1
argos	TATACGAAACCCCATGGATCG	sense	2
Arp5	ATGCGCTCTACAGCTGGAAGC	sense	2
Art1	GTAAGGTGGCCGTGACCGAGC	sense	3
Art4	CTGGCATCTGCCATGGGCTGG	sense	5
Art4	CTTAAATTAGCTTAGGCTTAT	antisense	1
ATPCL	CATGGGCACAGTTGGTGGTGG	antisense	7
ATPCL	TGCTGGCAAAGGAAGCGTGGG	antisense	5
ATPCL	TGTCGAACGAATCAGGAACGT	antisense	1
aux	CTGCTGCTGGTACGTCACTGG	antisense	7
aux	CTGCGAGATGTCCACCATGGT	sense	4
aux	ATGTAAATTCATGTAAAAGTG	sense	2
aux	TGTATGCAGATTTGCTGCGGG	sense	2
aux	CTTTAAAGTTGAAGTATTGGC	sense	1
Bap170	CGTTTCAGGCTTCTCTTGGCC	antisense	3
betaggt-I	TCGTGCTTTGTGCGGCTTCCC	antisense	1
bigmax	TTAACCAGCAGAAACTCAAGC	sense	3
bin3	GTTAGAATCGTCTGTGTCCGC	antisense	5
bin3	CTGCCTAGCCCATGATCCGGC	antisense	4
bin3	TTGGAGTCTGTCTGCAGCTGG	antisense	4
bin3	GTAGAGCGGCGGGGTGCTGGGC	antisense	2
bin3	GTCTGGTAAAGCGGCGAGGCG	antisense	1
Bj1	ACATTATCCTGCGGCGGGGTC	antisense	2
blue	TGCGACTGCGATTTGGTGTGG	antisense	7
bocksbeutel	TCTTTAATTGCTTGTTCTCGC	sense	1
botv	TTACAGTCTGCCATATTGGGG	sense	2
brat	CGGACCAGAATCTCACCAAGG	sense	5
Bruce	CACGTGCCAAGAGATTATGCA	sense	2
Bruce	ACGCTGTAGCAAAACACTAAG	antisense	1
BRWD3	CCTCCTCCTCTTCAATATCGC	antisense	2
BtbVII	TACCGTGAACAACCTAGTCGG	sense	5
btn	TGACCGGACGCCTGGGGAAGG	sense	3

GENE	siRNA	orientation of small RNA	S2 cell reads
Bzd	GAAATCACAGTACCGCCTGGA	sense	2
cact	ATAAAATGCTTGACATTCTGC	antisense	1
Cap-D3	TATGGCCATGTGCCTGCAGGA	sense	1
Cap-H2	CGTCAATCAAAAAGTCATTTG	antisense	2
cbt	CAAGCTGGCGGAACTGATGGG	antisense	9
cbt	CGAGCGATAGCACCGGCGGGC	sense	7
cbt	GTATATTTTCATTTGTTGAGA	antisense	2
Ccn	ATGATGATGATGATGATGATG	sense	1
Cct1	ATAAGTGTGTGTTCGTGGAGC	antisense	4
Cct5	CTGTTCAAATACACTAAAACG	sense	2
Cct5	TGAACAGGGATAGCCCCTTGT	antisense	2
Cct5	CAGGAGAAGTTCACCCAGATG	sense	1
Cct5	TCGTTTTAGTGTATTTGAACA	antisense	1
cg	GATGTGGGGGCGTTCACCTGT	antisense	2
CG10011	TTTGTCGCTGCAATTGCTGTG	antisense	1
CG10151	CAAGGCGGCTGCAGCGGCTGC	sense	4
CG10214	TTGTTAAGCGTGAAGTTAGGC	sense	2
CG10214	AATATAAGCCTAACTTCACGC	antisense	1
CG10214	ATTGTTAAGCGTGAAGTTAGG	sense	1
CG10225	AATGGATAAATGTCTTTTGTC	antisense	1
CG10249	AATACGAAATGGCTTTACTGC	antisense	3
CG10249	TGTTCGTTATTTCTGGTTAGT	antisense	1
CG10274	CTTAAGTCGAATTCACATAGG	antisense	2
CG10274	TTTGCCACAGATGTTACAGGG	antisense	2
CG10274	CCCTCAACTGGTGGCGCGGGT	antisense	1
CG10274	CTTGAACTTCTTCTCCTGGGT	antisense	1
CG10341	CTCGTACTTTCGGGGGCTGGC	antisense	4
CG10365	GGCGGATGTCCTCGTGCGAGT	antisense	6
CG10376	CAGCAGGAAGCACTAAGCGGC	antisense	4
CG10376	AGAAAAAGTGCACAATTACGC	sense	1
CG10376	TTAAGAATGCCATTTACACGC	antisense	1
CG10435	ATATTTATTCTGCTGCTGAGG	antisense	2
CG10435	TAGGGAGCTGGCTGGTCCGGC	sense	2
CG10445	TAAAGTTGCACCTGGAGAAGC	sense	1
CG10462	CAATAGCGCGGACGATCTGGC	sense	4
CG10516	TGGTGTCTGTTTGGATCGAAC	sense	1
CG10576	ATCTGTGGAGCGCAGCGTTGG	antisense	2
CG10669	GCGTTCCTTTGTCGTTCAGCC	sense	2
CG10889	CGAGGGAAAGCCAGCCGCTGG	sense	3
CG10903	TAAGCCGGGCACTTTCGACGG	sense	3
CG10971	CGGAACAACAGCCTTTGGATG	sense	4
CG10971	AACTGGTGAAATCGTTTTGGG	antisense	1
CG1104	TGATTCTGCATTGTCTGCAAA	antisense	1
CG11063	TAGTTTGTTCTGGTTTTGTGC	antisense	1
CG11109	TGGTGGCCCTAGACAATTCGG	sense	5
CG11109	TGTTAATGCAGCGGTATCAGC	sense	3
CG11109	TTGCTCCGCTGTTGGAATGGC	sense	2
CG11109	TTTGCGTCCGTGTCCTGGATC	sense	2
CG11109	AGTGAATTCTCCCACGGGTGC	antisense	1
CG11180	TTGGTAAGCTTCATGGTTTAC	sense	1
CG11198	GTCGACTTCATGCCCACCAAG	sense	2
CG11198	CAAAAGGCTCTGTGTACAAAG	antisense	1
CG11242	GTATGCGGGTCTATTGATTGG	antisense	9
CG11306	CATTGGATCGATGGTCGTGGG	sense	13
CG11306	CATGGTGAACTCCTCGTGGAC	sense	1
CG11377	GTAAGAAGGGCTGGAGCATGG	sense	3
CG11388	CACGCCGGAGTTCAGGCCTGG	antisense	3
CG11388	AGAAGCTGACCCACTCGGAGG	sense	2
CG11448	AGGCGTCGCCCTCTGATGCGG	antisense	2
CG11455	CATCAGTTGCTTCTGCATGCC	antisense	1
CG11526	CCACACCAAATGCCTCGTTGG	sense	9
CG11526	TTTCGCATGCAGGCCGCCCGG	sense	2

CG11620	TAACGATCTCACCTCCGAAGG	antisense	6
CG11777	TAAATCTTAAAACCAGCCAAG	sense	1
CG11790	TGACAGACGGTACATTCGGCC	antisense	6
CG11814	ATCCAGTTCCTGCGGCTGAGG	antisense	4
CG11866	TGACTGCCGGATTCGTGTGAT	antisense	3
CG11872	TATAGCTTCTGTTATTGTGGG	antisense	2
CG11880	ATGTACTGCTTTACGGAGAAG	antisense	2
CG11880	TAGTAGGTGGACGCCCCCGCC	antisense	1
CG11927	ATACAAATGCCAATGGCCGTC	antisense	3
CG11929	TTGTTTCGTGGGATTTGCAGA	antisense	1
CG11943	CCGAGCAGGCGCTGGGCGTGC	sense	2
CG12016	CCACACCAAATGCCTCGTTGG	antisense	9
CG12016	TAGTTATTATGGGTGACATGG	sense	4
CG12030	ACCGACGGTCCGGAGATGTGG	sense	11
CG12030	GCCACGTTCCTGCCGCTCCGC	sense	6
CG12082	AGTCGTTGGCCGGGTCGCTGG	antisense	1
CG12106	CAGCGCCGAATCACTATGGGC	antisense	10
CG12106	AAGCGGGTGCTGTCGTTCTTC	sense	2
CG12106	TGAAGCTGCTTGTGCCGCCGC	antisense	1
CG12118	CAGCGCCGAATCACTATGGGC	sense	10
CG12118	AAGCGGGTGCTGTCGTTCTTC	antisense	2
CG12118	TGAAGCTGCTTGTGCCGCCGC	sense	1
CG12170	GGAAAGTGCACTGTTTTGGTC	sense	3
CG12182	CTATCGATTGCATCTGCAGGC	antisense	6
CG12182	TCAAGGACCTTCTACTGGTGG	sense	5
CG12182	TGGTCGTGGATCCCTTTCCGT	sense	2
CG12262	TTGGGATCCGGATTGGTGCGG	antisense	4
CG12262	GCATCATGACCGCCTTAGAGG	sense	3
CG12299	GTGCACGCACTGCGAGGCGAG	sense	3
CG12341	TTTAAGTTAAGATCTAAGTAT	antisense	1
CG12343	TTTTGGAGGTATCCGCTGTGG	sense	2
CG12393	GGGCTCGGTGTCACGTCGGGC	antisense	6
CG12576	GAGGAAAGCCTGTCAAAGGGG	sense	1
CG12785	TATTAGGCGTTTCCTTTTGGG	antisense	5
CG12936	TTTAAGTTAAGATCTAAGTAT	sense	1
CG1311	CATGTAGAAATTCAGACCGGG	antisense	3
CG13189	GGTTGGCCTCAAAGAGTCTGG	sense	2
CG13220	CTGGCGGGCTTGGGAACTGGC	sense	8
CG13349	GCCTGATGCACTTCTGCTGGA	sense	1
CG13384	CGTTGTGACCTTCGCAGGAGC	sense	2
CG13384	TTGGTCAAAGGTGTCAGGTCC	antisense	1
CG13484	TTGTTATAGTTCTTGCGAGGG	antisense	2
CG13484	AGAAGCCAACGTTTGGATTTT	antisense	1
CG1358	CCGATCCACCGAGGGCGCTGC	sense	8
CG1358	CTGGGCTCCGTGGTATCGGGC	sense	5
CG1358	TTGGTGGCGGTGCTGTTCCGC	antisense	5
CG1358	AGACGGTGGGCAGTGATCTGC	sense	1
CG1358	ATGATGATGATGATGATGATG	antisense	1
CG13601	TCTTTAGTTGTTTGTCTGCGG	antisense	1
CG13762	CGCTGTCCACCTGCACGTCGG	antisense	4
CG13893	TTAAGGTCAACGTTGAGGAGC	sense	1
CG13900	CAGTGCAGCGAAGGTATCGTG	sense	9
CG13900	CCAGGATCTCTGCTCGCCCTC	antisense	7
CG13900	TAAAGAACCTGGTGCTTGTGG	sense	5
CG13900	CTCGCAAATCCGTTGCCTGGA	sense	1
CG13902	GAGATACGGTTCCAGCTGGTC	antisense	2
CG13902	TTGTACATGCCACCCAAATGG	sense	2
CG13924	CTGCTCGTTGCGATTGATGGT	antisense	2
CG14102	GCGACTGCTTCTCAATTTCGG	sense	1
CG14211	CGGCGACGCACATGGAGACGC	sense	6
CG14215	CTGATTTCAATGCAGTGGCGG	antisense	2
CG14230	TCCTCCTTCTTCTCCTCCTCC	antisense	4
CG1434	AGGGTACAATCGATCTGGTGC	sense	1
CG14435	AATAAGTTTGTGTTGCCAGAC	sense	1
CG14476	GTTATGCTGCCATTTGGACGG	sense	1
CG14670	TACTCGAACTCGGTGGCCTGG	antisense	3
CG14782	CAAGGTCGCGTTCTGGTGGGC	sense	7
CG14782	ATGAGCCGCGCTTCTACGGGG	sense	4
CG14786	TGGCCGAGTTTCAGCGACTGG	sense	6
CG14799	ATGATGATGATGATGATGATG	antisense	1
CG14804	TAGCTATGCTCTCCAGTTCGC	sense	3
CG14815	TGAGGTCTGGCAGTTGCTGGG	sense	4
CG14882	AATAGGTTGCTCATTCGTGGG	antisense	3
CG14956	CTGCTAACCGTTCCACCGCGG	sense	3
CG14956	GCAGTAGCAGCAGTGGAGCGG	sense	1
CG14966	TAAGCAGAACGGAATCACAGG	sense	6
CG14967	TACGGATCGAGCGATGCGTGC	antisense	7
CG15011	CGTTGTGGTCTGCACGAAAAG	sense	2
CG15067	TTTAAACTTATGTGGTGGAGG	sense	2
CG15097	TGAAACCCTACGTATTTGCGG	sense	1
CG15099	AACAGCCGGTTTTCATCTCGG	antisense	1
CG1516	TGAGCTGGCTGTGCAGACCGG	sense	5
CG15209	CCATAGGGCTAGCAGGCGGGC	antisense	5
CG15216	CGGGTTCCAGTGCATGGGGGA	sense	2
CG1531	TCCTCTCTCAGCCAGGGCAGT	antisense	4
CG15370	ATGATGATGATGATGATGATG	antisense	1
CG1542	AGTCAGCCGAGAATCGCAAGA	sense	3
CG15438	AAATACTTGGCGTGTCTAGTC	sense	2
CG15482	TCTTAGACTTAAATACATGGC	antisense	1
CG1553	ATTGTCCAGCACGTTGCAGTC	antisense	3
CG1553	TTTAGGTTTATCGTGTTATGA	antisense	2
CG15609	CGAAATAGTATTGGTGGTGGT	sense	6
CG15609	AGGGGTTGCTGTTTCTAGCGC	antisense	5
CG15609	TTTAGCTGCATCTGTCGCCGG	antisense	3
CG15609	CACGACCGTTGGCCGCCACCG	antisense	2
CG15609	TCGCAATGGTTTACGCTGGTC	sense	2
CG15609	TTGGGAGTGTGACATAAATGG	antisense	2
CG15609	TATTTCGGATTAGCCTAAGGA	antisense	1
CG15891	AGCTTGTCCAGTTCCTCCTCC	antisense	1
CG15892	AGCTTGTCCAGTTCCTCCTCC	antisense	1
CG15896	CTCATCGCTGATGGCCACCGG	antisense	6
CG15896	AACTCCACATAGCTTTTGGCC	antisense	1
CG15930	ATGATGATGATGATGATGATG	sense	1
CG1600	GTGCGCTGGTCAAGTCGTCGG	sense	3
CG1621	CTACGGAGCCCATATGCGAGC	sense	4
CG16742	ATCTCGCGCTCGTGGACCTGC	sense	1
CG16903	CGATTTAGTAATGCTAATGTG	sense	2
CG16903	TAGCATTACTAAATCGGTAGA	antisense	1
CG16972	TGGGTGTATCAACTGGTAAGG	antisense	6
CG16989	AACATGCACTTTGAGGGACGC	sense	2
CG17264	TCACTGGCACTGCACTTCTGG	antisense	13
CG17660	AATGTATGTAATCGTAGTTCC	sense	4
CG17660	AAGGGTGTTGCAGTTCAGCAC	sense	2
CG17715	TTAACTCTATACAGTGCCGCT	sense	1
CG17746	AGGTGAAGATGGCGTCGCATC	sense	2
CG17746	ATTGGGAATTCATCGTGCTGG	sense	2
CG17746	TGAGTTCAATCGCGTCAACGG	sense	1
CG17746	TTGGATTGCTTCTGGCAGGGT	antisense	1
CG18107	GTCACTGTCTTTGTGCTTGGT	sense	2
CG1812	TCGCATCTATGCCTGCGGCGG	sense	2
CG1814	GTCACTCTGTGCTCTATTTCG	sense	4
CG18166	CGTGGCTGATCAAGTGCTTGA	antisense	1
CG18259	CCACCAACAGCCTTTCTCCGG	antisense	6
CG18259	GAAAGGAGCAGTTTGGTAAGC	sense	2
CG18259	TGAACAAGCCTTTTTTCAAGC	sense	1
CG18262	CGAAATGTCTGGTTGCAGGCT	antisense	2
CG18273	CGTGGCTGATCAAGTGCTTGA	antisense	1
CG18432	TGGACAGCCCTTTGGCGCGCC	sense	7
CG18542	GGGTACGCCTCACTGGAGTGC	sense	8
CG18809	TAACGTAACGTAGCCGCAACG	sense	7
CG18854	CGGGAAACTATGGATCAAATG	sense	15
CG18854	CATCGCAAGCCAGATTCTTGC	sense	12
CG18854	TCGGTTGAAGCGTTGGCTTTC	sense	6
CG18854	GATCTTGAACATTTCGCCCTC	sense	5
CG18854	GAGGTCCGGTCTGAGCGTGGC	sense	4
CG18854	TAATATAGGGTGGAGCTCAGC	sense	4
CG18854	TTTGGCCCATAGTTTTCCATC	sense	4
CG18854	ATGCTGCTGAAATGGATTCGG	sense	3
CG18854	CTTGAAGCCAGGAATGCCATC	sense	3
CG18854	CTTGGTGATCGCTCGTGCCTC	sense	3
CG18854	ATCACTATCATCATCATCCGA	sense	2
CG18854	ATGCTAATGACTCCGATGTGG	sense	2
CG18854	CAAGCTTTGGAGATGGAGGGC	sense	2
CG18854	CCTTGTAGTGGATTCGGATGA	sense	2
CG18854	CTTGCTTGGCTCTCAGGAATC	sense	2
CG18854	GGATTCAGCTCGGTTAGAAAG	antisense	2
CG18854	TCTGATCTTGAACATTTCGCC	sense	2
CG18854	AACGGATCTCAGGACTGGAGG	antisense	1
CG18854	AAGGGTGGCCAAGATATGTGG	sense	1
CG18854	AGGCGCATGTGCTTTAGTCGC	sense	1
CG18854	ATCCTCTACAACGATTTTTTC	sense	1
CG18854	ATTGGTGGTCAATATGTCGGC	sense	1
CG18854	GATGATTCCCGGGATTCAAGC	sense	1
CG18854	GATTCTTGCTTGGCTCTCAGG	sense	1
CG18854	TATGTTGCCTCCAAGTAGGGC	sense	1
CG18854	TTCGGATGATAATGCTAATGA	sense	1
CG18854	TTTCTTGGTGCCGGCGGTCGG	sense	1
CG1902	TTCCTGATCCTTCTCACTGGC	antisense	2

CG1972	GCAGCACATTCAGTATGTAGG	antisense	7
CG2006	TCTGGATGTCCAACATTTTGT	sense	2
CG2006	TCGTGCATATGCGCATCGCGG	sense	1
CG2034	ATCAAGGCCCTGCCCACCGGC	sense	2
CG2137	AGTTCCCGTATATTGAGGCGG	sense	1
CG2225	AAAGACAATGGAGCCTCGGGG	sense	2
CG2247	GTGTTCTTCGCACTTGTCCAG	sense	1
CG2614	CGGGAACGTCATTGCTGTCGC	antisense	6
CG2698	TTATTTACACCGTAGGAGAGT	sense	1
CG2811	TCAACGCATGGGTATTATTCG	sense	2
CG2926	AGGGATCACGTTGCCGCGAGA	antisense	8
CG2991	CGTGGTGGCCTATCAGCGGGC	sense	2
CG3032	CATTCGAGAGGTGCATGCCCC	sense	2
CG30373	GGAGGAAGATGGCCAGGCTGG	antisense	9
CG30410	TGGGCCATGTTCACGAACAGG	antisense	4
CG31082	CGGAATCAGTGTCTGCATAGG	antisense	6
CG31121	TAACGAAGCCCGACTGTACGG	sense	5
CG31121	ATGATGATGATGATGATGATG	sense	1
CG31158	TTCTAGCTACATGGACATGTC	antisense	1
CG3164	AGAAAGCTTGCGAAATGCGCT	antisense	1
CG3164	TGATTGAGGTGTCCTGCGGGG	sense	1
CG31643	TGAGTTTTAAAGTTGTTTGAG	sense	1
CG3165	TAAAATGAAGTTCATGCTGGA	antisense	3
CG31678	TCAAAATATAAACCCAACAGG	antisense	2
CG31678	TCTTAAGCGCCAGGAATCCGC	sense	1
CG31729	TACGAACAAAGACCGATCCGC	antisense	4
CG3173	ACGAGCTGTTCGAACGACTGC	sense	2
CG31771	ATGATGATGATGATGATGATG	sense	1
CG31793	TAGCTGTGATTGGACCCGTGG	sense	5
CG31812	GAAGGGTGTTCAACGCGTGGC	sense	2
CG31849	CAGGTGGGCTTCATCGGCTGC	sense	4
CG31918	ATGAACGCTTTCCTTTGTGGT	sense	1
CG31919	TGAAGGGACTCGCTTGGTCGC	antisense	8
CG31922	GCTCTTTGCAATTGCGTGTGC	antisense	3
CG31975	AGCAGTGGCTACGTCTCCGGC	sense	4
CG32164	CTGTCTGGTGAAGGATTGCCC	sense	1
CG32164	GTAGAGGACGTCAAGCATGCG	antisense	1
CG32164	TAATGCGCACCGACGACCAGG	sense	1
CG32165	CTGTCTGGTGAAGGATTGCCC	sense	1
CG32165	GTAGAGGACGTCAAGCATGCG	antisense	1
CG32165	TAATGCGCACCGACGACCAGG	sense	1
CG3223	ATATTTATTCTGCTGCTGAGG	sense	2
CG32250	TGGAAGAAGCCGGCGATGTGC	sense	2
CG32409	GAAGTTCGCCAAGGAGAAGGG	sense	1
CG32412	AGGGTGCGGTTGAAATGGACC	antisense	1
CG32425	CTATGTTATGCCATCCGCTGG	antisense	8
CG32495	TTGCCGCCTCCCTCGCGCTGC	antisense	2
CG32694	ATGATGATGATGATGATGATG	sense	1
CG32702	CCAAATGCACTCGGAATGGGG	antisense	5
CG32702	CTTGCATGGATGCTGATCGCA	antisense	1
CG3271	TTTCGCAGGGTTCGACAGTCC	antisense	2
CG3279	GACCTTCTTGTTAATCGATGG	sense	5
CG3279	GAAGAAAGCTGCAGATTACGC	sense	3
CG3279	GGAGAAGGTGGTCTTGTATGG	sense	2
CG32809	TGCGGCAGCCTGGTTTGGTGG	antisense	5
CG32939	GGGTACGCCTCACTGGAGTGC	sense	8
CG3308	GCGTCCGGTTCGGAGAAGTCC	antisense	3
CG3308	TTACGGGTTTAAAGCTGCTGG	antisense	2
CG3308	AATCCGCTTTGTGCCCATTGT	sense	1
CG3308	TTGCACAGGTAGCCCGTCAGG	antisense	1
CG33107	TTCTACCATGGCGGGAATCGC	antisense	1
CG33111	CGGGGGATTGGTGGCAGCTGG	sense	1
CG33249	CGGCAGCCAGCGGCAAGCAGG	sense	4
CG33469	TTGTGAACCATTTTAAGTTGG	antisense	1
CG33470	CCAATGGAGCTAAGAGCGTGG	sense	14
CG33509	GCAGAACCTTTCGGATTTGGG	sense	5
CG33509	TGTGGATGTCATGGAAATGGC	sense	5
CG33509	CATAGGTGTGACCTATATTGG	antisense	4
CG33509	TTCCAAAAGGAATCCTGGCTC	sense	1
CG33510	TTGTTCATCGTGATGCATGGA	sense	2
CG33510	TTAAGGATGAACGACCGGAGG	sense	1
CG33523	GACGATTGCCTCTGCCTGAGT	antisense	2
CG3356	TTACTTGAAGGACTACTCGGG	sense	2
CG3363	ATCGTGATGCCTCCAGGAGGG	antisense	3
CG33649	GTGACCGAAGGCGATTGCCGG	sense	1
CG33932	CTGGCGGAAGGTTATGTCCTC	antisense	3
CG33967	GAAGTTGCTGATGAAGACCGC	sense	3
CG33969	TATGGTGGCGATTACAACCGC	antisense	6
CG33969	CATCAAAGGCATTCTCTTCGC	antisense	4
CG33969	AGCTGGTGACCACGGTTTGTG	antisense	3
CG33969	TTCGAACGACCATACTTCCAC	sense	2
CG33969	ATTGACCTGTCCACGAATTGG	sense	1
CG33969	ATTGCGTTTCTATTGTCAAGT	sense	1
CG33969	CTCGTGTCTAGTTTATTACGG	sense	1
CG33969	TTGATCGGGCCAAAATTCGTT	sense	1
CG33978	TTACTTATAATCACCAAGCGG	sense	1
CG33995	TGAAGGGACTCGCTTGGTCGC	antisense	8
CG3402	TAGCTCACGGCCTTTTTGTGG	antisense	10
CG3402	ATAAGTGGTGCTCGGAGCAGG	sense	3
CG3402	TCCACATAAGTGGTGCTCGGA	sense	1
CG34125	CATGTGTAACCTAAAAAAAGG	antisense	2
CG34125	TTTTTATAACATCAAAGGAGT	antisense	1
CG34126	TGAGAACGTCATGGTTTGGGG	antisense	6
CG34179	GGGCTCGGTGTCACGTCGGGC	sense	6
CG34268	CGTGGTGAAATCTGTTCCGGT	sense	3
CG34335	GGCGGTCGAGTGCCTCACAGT	sense	3
CG34376	ATGGAGGTGGCCATTTGGGTC	sense	4
CG34398	ATGATGATGATGATGATGATG	antisense	1
CG34415	GGAACTGAACTGCACGCCTGG	sense	5
CG34429	CTCCTGCAGGATATCTGGATC	antisense	1
CG34429	TGATCCAGATATCCTGCAGGA	sense	1
CG34430	TCGGTCCTAAAGCATTCACGG	antisense	1
CG3542	TCTCAGGAGTCGGTAATCGGG	antisense	2
CG3542	AGAGCGGGTCTACACCCCTGG	sense	1
CG3605	GTTGGCCGAACATGTGGCGCG	sense	3
CG3683	CAAACAGGCTCGCGGGTCATC	antisense	3
CG3703	CGATGGAGCGCAGTGAGTCGC	antisense	4
CG3703	ATAGGTTGAAACACGGCGAGG	sense	1
CG3703	CAGTGAGTCGCGGAATCCGGT	antisense	1
CG3711	TTTGATTAGGAACCACTTTGG	sense	1
CG3740	TAGCCATCGTAGCGGACGAGC	sense	6
CG3760	TCAACGCATGGGTATTATTCG	antisense	2
CG3764	CGGATGTTTGTCGACAGATGC	antisense	5
CG3792	GAATAACCGAATTGGCAAAGG	antisense	3
CG3814	TAACTTAAGCAATGATAAAGC	antisense	2
CG3831	GATTGGTGTACCTATTTAAGG	sense	1
CG3967	CGATTCTTCATGCCCGTGTGC	antisense	1
CG3973	TGGATCCTTCGAGCGCAATGG	sense	2
CG3973	GAGAGCGTGGAATCTTGCTGA	antisense	1
CG3980	AACCTGAACTGCCAGGATCGG	sense	1
CG40084	TATTGAAAACTGTATTGCTAG	sense	1
CG40228	TGTTCACCCATGTTGCCACGC	sense	1
CG4025	CGATTTAGTAATGCTAATGTG	antisense	2
CG4025	TAGCATTACTAAATCGGTAGA	sense	1
CG40351	GTGTCCTCCAAAGCCGCCTGC	antisense	2
CG40351	TAAACCATTTTGAACAGCACC	antisense	2
CG40351	TGTTAGCCAAATGACGAGGAC	sense	2
CG4061	CGACGGCTCCTACTTGGAGGG	sense	3
CG4061	ATTAAGCATTCACGCCAGAAG	sense	1
CG4068	TTGACTCCAACAAGTTCGCTC	sense	33
CG4068	TGGCGCTTCACAGGCGCTGGA	sense	27
CG4068	GTCCAACTACAGGATACTGGG	sense	15
CG4068	TGACTCCAACAAGTTCGCTCC	sense	9
CG4068	CGGTAGCCTGTAGTTTGACTC	sense	5
CG4068	CTTCCGCTGGCTTTGATTTTC	sense	4
CG4068	TTTGACTCCAACAAGTTCGCT	sense	2
CG40798	AAATGCGAAACTACATTAGAG	sense	2
CG4119	TGGCGCTGCCGTACAAATCGG	antisense	1
CG41322	ATAATACAGGCAAGCGTAAGG	sense	2
CG41421	CGAGAGGAACCGCAGGTACGG	sense	8
CG41484	CAGGAATCCTGTTGGAATCGG	sense	7
CG41484	CAAGGAAAGTGGATTACTCGG	sense	5
CG41484	GGGTACTGTCGGCGGCTCGAC	sense	3
CG41484	ATCCGATTGGTACTCCGCAGC	sense	2
CG41484	ACTTTGAGGCTGCTGTATATC	antisense	1
CG41484	TACATAGCGGTGTATTCGCCG	antisense	1
CG41533	TGCATATCATTCACGCCACGC	antisense	2
CG41584	TATTGAAAACTGTATTGCTAG	sense	1
CG41587	AAATGCGAAACTACATTAGAG	sense	2
CG41589	TATTGAAAACTGTATTGCTAG	sense	1
CG4199	ACATCGGGCATTACCAGCTGG	sense	11
CG4199	AGTACGATATCGAACTGTGGC	sense	1
CG4199	ATTGGGTGTTGAGCTTGGAGC	antisense	1
CG4213	ATGCTCAGTTGGCTTTGCAGC	sense	2
CG4334	CTTCTTGATGACACAGTCTGG	sense	2
CG4582	TTCAATGGTCCTGTCTGGCGG	antisense	2
CG4619	AATAGCACTAAACATAATGGT	sense	1
CG4643	ACACGAGCCCAGCAGCGGAGC	sense	1

CG4643	TGTGGCTCTGCTCGGCTCCGA	sense	1
CG4670	AAGGACGCCGTGCTGGGCTTC	sense	3
CG4670	AGAGTGTCCACAGGGAGCAAC	antisense	2
CG4699	GAGCTTGGCGTGCCTGTCCGC	antisense	2
CG4752	CGATGGCGCCTTAATGGTTGC	sense	4
CG4822	AGCCAATGGTAGAAGCCGTGG	sense	6
CG4822	CGGTTTGACCTCGGTGCTAGA	sense	4
CG4901	TTTAGAGCCAAATGCAAGTGC	antisense	2
CG4963	CGCCGTGGTGCTGGGTGCTGG	sense	5
CG5044	AAGGATCAGAAGCCCCAGTGG	sense	6
CG5062	TGTGTCTGTGTGTGTGTGTGT	antisense	6
CG5104	GAATATAAGCGTGAGTCCCGC	antisense	1
CG5126	TTGGTCCGAGTGCTTCGCTGG	sense	1
CG5130	CAGCGCAAGACTTGGATTTGG	antisense	11
CG5191	TGATAAATGCCATCGTCCAGG	sense	4
CG5362	TCTGAACGGTTTCCACGAAGG	antisense	5
CG5458	TGGGAACTACAGGAAGGGTCG	sense	7
CG5458	TCCACTCGGGTACAGGTTATC	antisense	4
CG5458	TGTCGGCGCCCGAGGAAGAGG	sense	2
CG5508	CTATATGCGGCCTCTGTGCGG	sense	2
CG5508	TTGTTTTCGGTTGTCCTGCCG	antisense	1
CG5510	AACTATAGGAAAATTGAGACC	sense	1
CG5537	GGCATTCACTACGGTCCGAGC	antisense	1
CG5543	ATGAAGCTTTGGGATCTGCGG	sense	2
CG5567	CAGCGGGAGCTTCGTGCGGGC	sense	3
CG5644	TTCGCATGAAGGATCGTCTGG	sense	2
CG5734	TGGCGCGCCTTTTCCTTTTCG	antisense	6
CG5734	CTATGGCATCCGGAACGAGTC	sense	2
CG5734	TCTGCAGGCTGTTCAGTTGGC	antisense	2
CG5734	CGAGGCGCTGGCAGTGGAAAG	sense	1
CG5734	TAGAGATGCCAAGCTAACTGG	sense	1
CG5840	CGATAGGCCAGATCCCTGGGC	antisense	1
CG5853	AATAGCACTAAACATAATGGT	antisense	1
CG5857	ACGAGGTCCTGCAGCTGGCGC	antisense	2
CG5857	TCTTTAGTTGTTTGTCTGCGG	sense	1
CG5871	GAACTGCTCGCGACGCTGGTG	antisense	4
CG5871	TCTAAAGACCCACGGCTCCTG	antisense	4
CG5871	CGGTCTGCCGGCGGCGGAGGC	sense	3
CG5871	TATTTCGACGCAACCTGGAGC	sense	3
CG5899	TTCTGTTGACTGCTCATGGGC	antisense	1
CG5919	CACGCCGTCGCCCGTGTCTGC	antisense	9
CG5919	AATCCGCTTTGTGCCCATTGT	antisense	1
CG5938	TGAAACCACAAAATACTTAGG	antisense	2
CG5986	TAAATATATAATTTCTTGTGG	sense	1
CG6038	TAAACCACACGATCAAGGAGC	sense	7
CG6171	TATATAAAGCCTATTATTCGG	antisense	1
CG6181	CAATGAAGTGTCTTGTTTGCT	antisense	3
CG6181	ATGATTCGCCTACTAAATTGC	sense	2
CG6218	CGTGGCTTCTTTTCATGTATG	sense	7
CG6424	CGAACGATCTATGCGTGGAGG	sense	6
CG6424	AGCTGAAGCGTCTGGTGTCGC	sense	3
CG6424	AGTACGAGTCGCTGTTCGAGG	sense	2
CG6424	CTCTCCACATTATTCGCACGG	sense	2
CG6424	GTGGTCTCTGTGCTGTTGTTC	antisense	1
CG6424	TCTAGGTCTGTATTTGTTTGG	antisense	1
CG6448	GTAACCACAGTGGGCCTAGTC	antisense	2
CG6448	ATCGGTGACGCCTTCGGTTCG	antisense	1
CG6454	TGCCACTTTTCGAGGGACTGG	antisense	1
CG6509	TCCGGATTGTACTGAACGAGC	antisense	2
CG6654	AAGTAAGGATCCGTCTTGGGT	antisense	2
CG6689	CCAAAGACCAAAGAAGAGAGG	sense	5
CG6805	CAGAAGGAACAAGAGAAACGA	sense	1
CG6833	CTTGGCATCCTGGTTTTCACT	antisense	2
CG6876	CAGCAATTCGTGGCGAAGTGG	sense	3
CG6891	GAAAGGAGCAGTTTGGTAAGC	antisense	2
CG6900	GAAAGGAGCAGTTTGGTAAGC	antisense	2
CG6903	TCACTGATGTACTACTTCATC	sense	3
CG6907	CAACTTGAAGGCCAAGTCCGG	antisense	7
CG6907	ATCTGCCGCTGGTGAATTCGG	sense	2
CG6907	TTTGCTCCATCGATTCCTGGT	antisense	1
CG6912	GTGGGCTCTGCCGTTGTTGTT	antisense	1
CG6950	ATGGCCGGTGGAGTGCCCCGC	sense	2
CG6961	CCACCAACAGCCTTTCTCCGG	antisense	6
CG6961	TGAACAAGCCTTTTTTCAAGC	sense	1
CG7011	CAGCAATTCGTGGCGAAGTGG	antisense	3
CG7139	CCACGTTGCTCGGAGACTCGC	antisense	8
CG7139	ACAGTGAACTCAGTAGACACA	sense	2
CG7144	AACCGTTGTGCTTATTCTTGG	antisense	1
CG7177	AAAGGTTGCTGATGTTGGACC	antisense	1

CG7224	ATTTGCATTGCACTTATTGGG	sense	2
CG7224	TTTGCATTGCACTTATTGGGC	sense	2
CG7289	ACTCAAAACCAGTGGCGGAGG	sense	1
CG7324	GTTCGCTGCCCGAGCATCCGG	sense	2
CG7324	CAATTGTCTTAACTTAATGGT	antisense	1
CG7338	CTCCCTTTTGTTTTTCCTGAG	antisense	3
CG7338	CAATTGTCTTAACTTAATGGT	sense	1
CG7376	TGTTTGGCCACAGCAGCAAGG	antisense	3
CG7376	TGCTCGTGGCACTTTAAATGC	antisense	2
CG7379	ACGGCAAGCTGCGCCAGCTGG	sense	3
CG7504	AATCACAGAGGGGCCGCGTGT	antisense	1
CG7518	TGAAACACCGGAAGGAGGAGT	sense	1
CG7519	CAACTTGTTAAACTTCTCGGC	antisense	3
CG7632	ATTTTGAAGGGCTTTGTGGGC	antisense	1
CG7650	CGTAATGGTGAGCTTCTTGGC	antisense	4
CG7650	TCATCGGTAAGCCGCACAAAG	antisense	2
CG7650	GATTTGTGTCGATGGTCAGAC	antisense	1
CG7739	GTTGGAAAGCTTATAATGGAG	antisense	7
CG7739	TGGCGGACCATCTCAAGGCGG	sense	7
CG7739	TTAAAAGCCGCCTTGAGATGG	antisense	7
CG7739	CTAAAGCGCACTTTCGAGGTA	sense	5
CG7739	TGGAATCAATAGAGATGCTCC	sense	3
CG7739	TGTCCTAAAATGCCACAAACA	sense	3
CG7739	TTGGAAAGCTTATAATGGAGT	antisense	3
CG7739	CTCCATTATAAGCTTTCCAAC	sense	1
CG7739	TTAATATTCCTAAAAGAAAGG	sense	1
CG7789	ATGGGTCGCACCATTTGGGGC	sense	5
CG7816	TGAATTTGTTGTACTTTATGG	sense	2
CG7816	TCTAAGATTGTCATTGGCAGT	sense	1
CG7830	TGATCACATTGGTGGCCCTGG	sense	2
CG7912	CTCGGCTGCTGCACACTTGTG	sense	3
CG7988	CTGGTTTCTGGGCGTCATGGT	antisense	1
CG8112	CTTCCGCGGATGGGCCTTGGT	sense	3
CG8112	TTCAACCTGCCCTTTGTGAGT	sense	1
CG8155	TGTGAATCCTCAGAAAGGCGG	antisense	2
CG8199	AGTGGGTCACTTTGTATGTGC	antisense	5
CG8289	TAAGGAGCTGCGCGAATCGCC	sense	4
CG8297	TATGATGAAGTGCCTGCAGCG	sense	4
CG8315	TCATCGCTGGCTGGACTCTGG	sense	4
CG8319	GAACAAAACCTCAAGGCGCAC	sense	1
CG8320	TGCTGGCAAAGGAAGCGTGGG	sense	5
CG8336	GATAAGTCACATCAGCCAAGG	sense	2
CG8336	GATCGTTTATGGCCTCCTCGT	antisense	2
CG8336	AAATGGTGACTGGCTGCCAGC	antisense	1
CG8443	TCGAAGTAGCTACATTGGACT	antisense	2
CG8451	TCATAGGTTCCTGTACTTCGG	antisense	5
CG8478	GTTGTGAACAGATCGCCGGGC	sense	2
CG8478	CTCAGAGACTTCGTTCACTAC	antisense	1
CG8481	CTACGATTTTACTGCGCGCCC	antisense	1
CG8516	GGCGGATGTACTGGCGGTCGG	antisense	5
CG8526	CGCCAATCGACTACGCCAAGG	sense	10
CG8538	CTCGACTGCACCCGTGTCTGG	sense	5
CG8545	ATGATGATGATGATGATGATG	sense	1
CG8594	ATGGAGGCCACTAAGGTGCGC	antisense	6
CG8594	CGTTATAATGCCAGCAATGCG	antisense	1
CG8594	TGTCTGAATTGAAGTTTCTGG	antisense	1
CG8602	CGGTGGCTCGTGCTCGAGGGC	antisense	4
CG8602	AATGGTGGTCAGTGTCGCTGT	antisense	2
CG8668	TGGGGAATTTTTTCTTTCTGG	antisense	2
CG8798	TATAAAACCTATCAACACCCG	antisense	4
CG8798	AGATGTCATGAAGGAGTCGGC	sense	3
CG8798	GTTAACCTGCCGGTTTGGAAG	antisense	3
CG8862	CTGTTCAAATACACTAAAACG	antisense	2
CG8862	TGAACAGGGATAGCCCCTTGT	sense	2
CG8862	TCGTTTTAGTGTATTTGAACA	sense	1
CG8878	TCGTTTCACCCACATTTCATG	sense	2
CG8878	AGAATCTGCGTTCCAAGTCGA	sense	1
CG8950	AGGCTGTCCAGGCATTGGAGC	sense	3
CG8950	GTAAGCACTGCCTCCATTTGC	antisense	3
CG9007	CTTTAGGCATGCTTGTGTTTG	antisense	5
CG9143	CGTAAAGTTCCCAGAGGCGAC	antisense	1
CG9246	TTCTAAAAGCCATGTACCTGG	sense	1
CG9320	CAACAAGGTGGATCTAGTCGA	sense	4
CG9339	CGCGGCAGTGGTCCTCGCTGG	sense	3
CG9346	TCTCAAAAAGGACTCGGGTCC	sense	2
CG9372	GTTAACCTGCCGGTTTGGAAG	sense	3
CG9389	CGTTAACCCCAATGAGTACGG	sense	3
CG9578	CGGCCGAGTCCTTGGGACAGC	antisense	3
CG9629	TTTGAGGATGTACACAATTGG	antisense	3

CG9674	ATGATGATGATGATGATGATG	sense	1
CG9776	CTCGATTCATGTGTTCCTCGG	antisense	1
CG9779	TCGATATTTACTGTGGCATGT	sense	1
CG9780	ATTATCGGAAACCTACGCCGG	sense	2
CG9780	TAATTGGTGGCAGTCTGCATC	antisense	1
CG9795	ACGCCCAATTGTCTTCTTCAT	sense	2
CG9799	TAATAGTGCCATCCCCGTGTC	antisense	4
CG9804	AGGCCGATTCCATGCGTGGTG	antisense	1
CG9804	TTGTGCCCTGCGGAATCGAAG	sense	1
CG9922	TTCCACGAGCACCTCAATTGC	antisense	5
CG9934	CCAGTGCTTAGCATCGTTGCC	sense	4
CG9941	ATGATGATGATGATGATGATG	sense	1
CG9945	TTGGTAAGCTTCATGGTTTAC	antisense	1
cher	GCAACGTGACCGAGGATGCGG	sense	4
cher	CAGGCCGGGTCCGTAGGCAGT	antisense	2
chinmo	ACATGTTGAACGTATGGAACG	sense	3
chinmo	ATACTATTTTATTGTTGCAGG	antisense	1
chinmo	CTGCTGCTGCTGCTCGTAGGT	antisense	1
chinmo	GGCCTCGGCCTGGATATCCGG	antisense	1
CHKov1	TCTCAAATCGAATGTGGACGG	sense	3
Chro	CGATACCTGCGACGACGCCGG	antisense	2
Chro	AAACATGCATTTATCGGGGTC	sense	1
Chro	AACAATAATGCTGCGCCGGGG	sense	1
Cht3	TCTAAAAAGCCGGAGACGACT	sense	1
cnc	CAACATGGCAGCGTGATGCGG	antisense	3
Cog3	CTTGATCTGCGGCGTGCCCTC	antisense	1
colt	CTAGTGTTGCTGTTTTCTGGC	antisense	5
Coq2	CGCTGCAGGAACTCGTGTCGG	sense	4
Cp190	GACCTTGTCCACGCTGGCTGC	antisense	2
Cp190	GCAAGCTCTGCGGTAGCGGGC	antisense	2
CPTI	TAAATCTTAAAACCAGCCAAG	antisense	1
crb	GTTCGCATCGCAGTCGCAGTC	antisense	4
crq	CAATTTAGCCCAGCAAACTGC	sense	6
Csk	CAACCACGACGACCGGCCAGG	sense	11
CstF-64	CCCGAGCTCGTGCCCAGATGC	sense	10
cue	CTGGTTGCACATTGCTGTGGC	antisense	3
cue	GTGTCCAGCCAAGTTTGGAGG	sense	1
CycG	TCCACTTGGCCATCAAGCAGC	sense	4
Cyp28d1	TGTTCGAAGACGCGCTTGGTC	antisense	4
da	ATGCGTTCAGTGGTGCCGGGG	sense	1
Dcr-1	TGCGTGGAACTGCACAGGATC	sense	4
Dcr-1	CTGTCGGTGGTTGTTTACTGA	antisense	1
Ddx1	TTAAACTTCTCGATCTGGTTG	antisense	2
Dg	CCAATAATCCAGGTACTCGGA	sense	2
Dgp-1	ATTGAGCTGCTACAAAAGAGG	sense	1
Dif	TTGGGTGGGTTCATCTGCTGG	antisense	1
DNApol-gamma35	GTGACCGAAGGCGATTGCCGG	sense	1
DNApol-iota	TGGTGACCACCGCTGTCGTGG	antisense	1
Doa	TGGAGTCTCCTTGTTCTGGGC	antisense	7
Dph5	AAGGGGTCGCCCACCACGAGC	antisense	1
drosha	TGTCTTCAAAATCTTCGCTGG	sense	2
Drp1	TGATCGGACTCGTTGAGCAGG	antisense	7
Dyrk3	ATGCTTTGGGAGAATACTCGG	sense	1
Dyrk3	ATTAATTCAAATCAGTGCAAT	sense	1
e(y)1	CTTCTTGCGCGCATGGTTGGC	antisense	1
E2f	GTCAGTGTCGGAACCAGCTGG	antisense	3
Eap	ACAAAGGACGACCTGGAGCTG	sense	6
ed	CGGGTGAGTGTTCAAGATGCC	sense	3
Edem2	CGAGCTGTTTACGACATGGGC	sense	7
edl	ATGGCGGATGATGCACTGATC	antisense	1
EDTP	CGATCCCAACCGCTGATGCAG	antisense	3
eff	CAACGTGTTGCAGGCTGCCGG	antisense	1
egh	AGTGGTCACGGACAAGGCGGT	sense	1
eIF2B-epsilon	ACTAGTGGCCATAAACGCAGC	sense	5
eIF3-S10	GATGACAAGTGGCGGCGTGGC	sense	6
eIF4G	GTTTAGGTCACTTTTAGTTGG	antisense	2
eIF-5A	TCGCTGTAAATTTCTGTTGCA	sense	1
eIF5B	TAATGAACTCTGTTATAATGG	sense	1
Elongin-B	TTAAATATTCACTCTGCTCAC	antisense	2
ETH	TGTCTGTTTCGCTCTTGGTGG	sense	3
ex	GTGGTGGCACCTGAAGGTTGC	antisense	1
Fak56D	TGAGCCTGATGATCATTGAGC	antisense	4
fbl	TATGGTGGCGATTACAACCGC	sense	6
fbl	CATCAAAGGCATTCTCTTCGC	sense	4
fbl	AGCTGGTGACCACGGTTTGTG	sense	3
fbl	ATTGACCTGTCCACGAATTGG	antisense	1
fbl	ATTGCGTTTCTATTGTCAAGT	antisense	1
fbl	CTCGTGTCTAGTTTATTACGG	antisense	1

fbl	TTGATCGGGCCAAAATTCGTT	antisense	1
Fem-1	GAAGGTTGTTTAGCTGGGAGA	antisense	6
Fit1	TCTCGCCAGATGTTTGCACGG	antisense	1
FKBP59	AGATTCCACCAAATGCCACAG	sense	1
fog	TTGTGAAACAGTTGTCGTTGT	sense	1
foi	CTATAGCTTTTGCCGATGCGG	antisense	1
for	CTATGTGGCCTCTCTACTGTC	sense	3
form3	CCATCAGGTCATCCTCCTCGG	antisense	2
frc	CTGTTTATGTTTCTGCCTGCC	sense	1
frtz	CAGCTGGGCGCTCTGATCTGG	antisense	7
fry	ACCAACAAAGCCTGCACCTGG	antisense	6
fry	CTCATCGGGATGGATCAGCGG	antisense	1
Fur2	ATGATGATGATGATGATGATG	antisense	1
fzr2	CGCCGACGCACATAACCCGCT	sense	5
fzy	GGAAAGCTTGTGCTCACGTGC	antisense	5
gammaSnap	ATACTGGTCAAACTAAGGAGG	sense	3
gatA	AAGTCTGGAGCACACAGTGGC	antisense	1
GckIII	TGGCCGAACGTTCCAGTCCGG	sense	5
Gclc	TTTGCCGCCAAATTCTCCGTC	antisense	1
Gclm	TCTGGTACTTCTTCGTTATGG	antisense	2
Gcn2	GAATGCCTTTCTAGTGGAAGC	sense	2
Gdh	TTAATGTTACGGCAACGGAGC	antisense	3
Gdh	GCAATTGCTTCACTTTGACGC	antisense	1
gft	AAAGTCGCGTTTCTTGCCCTC	sense	1
Gmd	CCATGCCACAAGTGCGGATGG	antisense	1
Gp93	CAGGAGACTTCGCAGTTCGCC	sense	3
gry	TATCCGCTGCAAGTGGGCTGG	sense	6
gry	CGTGAAAATGCCCATGGCTTT	sense	1
gry	CTCAACCGCGTTTCGTCCAGG	antisense	1
gry	GTGCAGTTTTTGGCGTCCCGG	antisense	1
GS	TTGCCGCCTCCCTCGCGCTGC	antisense	2
gwl	CAGAGCCCACCAGTCGACGGC	antisense	2
Gycalpha99B	CCAAGGAGCGGAAGAAGAACG	sense	2
hdc	CGCCACAAAATGCTGCAACAC	sense	8
hdc	TGTGGGCGGCCCAAAGTTCGC	sense	4
hdc	GATCGGTGAAGATGTCGTGGC	antisense	3
hdc	AGCCCTCGTTGAGTCCGACGG	sense	2
hdc	ATTCTGCAAGGAGCCGTGGGA	sense	1
hdc	CTGATGCTGCTGGGCATGCGG	antisense	1
hdc	GTCAGCACTTTTCAGAGGGGC	sense	1
Hexo2	AGAAACGATGCGGTTCAGCGG	sense	1
His1:CG33801	TTATTCAAACTAAGGGAAAGG	sense	1
His1:CG33804	TTATTCAAACTAAGGGAAAGG	sense	1
His1:CG33807	TTATTCAAACTAAGGGAAAGG	sense	1
His1:CG33834	TTATTCAAACTAAGGGAAAGG	sense	1
His1:CG33837	TTATTCAAACTAAGGGAAAGG	sense	1
His1:CG33840	TTATTCAAACTAAGGGAAAGG	sense	1
His1:CG33843	TTATTCAAACTAAGGGAAAGG	sense	1
His1:CG33846	TTATTCAAACTAAGGGAAAGG	sense	1
His1:CG33849	TTATTCAAACTAAGGGAAAGG	sense	1
His1:CG33852	TTATTCAAACTAAGGGAAAGG	sense	1
His1:CG33864	TTATTCAAACTAAGGGAAAGG	sense	1
His2A:CG31618	GCGTAGTTTCCCTTCCGGAGC	antisense	4
His2A:CG31618	TCGTTGCGGATGGCCAGTTGC	antisense	3
His2A:CG31618	CAAGAAGACCGAGAAGAAGGC	sense	2
His2A:CG31618	GTTAAACAAGCTGCTCTCCGG	sense	1
His2A:CG33808	GCGTAGTTTCCCTTCCGGAGC	antisense	4
His2A:CG33808	TCGTTGCGGATGGCCAGTTGC	antisense	3
His2A:CG33808	CAAGAAGACCGAGAAGAAGGC	sense	2
His2A:CG33808	GTTAAACAAGCTGCTCTCCGG	sense	1
His2A:CG33814	GCGTAGTTTCCCTTCCGGAGC	antisense	4
His2A:CG33814	TCGTTGCGGATGGCCAGTTGC	antisense	3
His2A:CG33814	CAAGAAGACCGAGAAGAAGGC	sense	2
His2A:CG33814	GTTAAACAAGCTGCTCTCCGG	sense	1
His2A:CG33817	GCGTAGTTTCCCTTCCGGAGC	antisense	4
His2A:CG33817	TCGTTGCGGATGGCCAGTTGC	antisense	3
His2A:CG33817	CAAGAAGACCGAGAAGAAGGC	sense	2
His2A:CG33817	GTTAAACAAGCTGCTCTCCGG	sense	1
His2A:CG33820	GCGTAGTTTCCCTTCCGGAGC	antisense	4
His2A:CG33820	TCGTTGCGGATGGCCAGTTGC	antisense	3
His2A:CG33820	CAAGAAGACCGAGAAGAAGGC	sense	2
His2A:CG33820	GTTAAACAAGCTGCTCTCCGG	sense	1
His2A:CG33823	GCGTAGTTTCCCTTCCGGAGC	antisense	4
His2A:CG33823	TCGTTGCGGATGGCCAGTTGC	antisense	3
His2A:CG33823	CAAGAAGACCGAGAAGAAGGC	sense	2
His2A:CG33823	GTTAAACAAGCTGCTCTCCGG	sense	1
His2A:CG33826	GCGTAGTTTCCCTTCCGGAGC	antisense	4
His2A:CG33826	TCGTTGCGGATGGCCAGTTGC	antisense	3
His2A:CG33826	CAAGAAGACCGAGAAGAAGGC	sense	2

His2A:CG33826	GTTAAACAAGCTGCTCTCCGG	sense	1
His2A:CG33829	GCGTAGTTTCCCTTCCGGAGC	antisense	4
His2A:CG33829	TCGTTGCGGATGGCCAGTTGC	antisense	3
His2A:CG33829	CAAGAAGACCGAGAAGAAGGC	sense	2
His2A:CG33829	GTTAAACAAGCTGCTCTCCGG	sense	1
His2A:CG33832	GCGTAGTTTCCCTTCCGGAGC	antisense	4
His2A:CG33832	TCGTTGCGGATGGCCAGTTGC	antisense	3
His2A:CG33832	CAAGAAGACCGAGAAGAAGGC	sense	2
His2A:CG33832	GTTAAACAAGCTGCTCTCCGG	sense	1
His2A:CG33835	GCGTAGTTTCCCTTCCGGAGC	antisense	4
His2A:CG33835	TCGTTGCGGATGGCCAGTTGC	antisense	3
His2A:CG33835	CAAGAAGACCGAGAAGAAGGC	sense	2
His2A:CG33835	GTTAAACAAGCTGCTCTCCGG	sense	1
His2A:CG33838	GCGTAGTTTCCCTTCCGGAGC	antisense	4
His2A:CG33838	TCGTTGCGGATGGCCAGTTGC	antisense	3
His2A:CG33838	CAAGAAGACCGAGAAGAAGGC	sense	2
His2A:CG33838	GTTAAACAAGCTGCTCTCCGG	sense	1
His2A:CG33841	GCGTAGTTTCCCTTCCGGAGC	antisense	4
His2A:CG33841	TCGTTGCGGATGGCCAGTTGC	antisense	3
His2A:CG33841	CAAGAAGACCGAGAAGAAGGC	sense	2
His2A:CG33841	GTTAAACAAGCTGCTCTCCGG	sense	1
His2A:CG33844	GCGTAGTTTCCCTTCCGGAGC	antisense	4
His2A:CG33844	TCGTTGCGGATGGCCAGTTGC	antisense	3
His2A:CG33844	CAAGAAGACCGAGAAGAAGGC	sense	2
His2A:CG33844	GTTAAACAAGCTGCTCTCCGG	sense	1
His2A:CG33847	GCGTAGTTTCCCTTCCGGAGC	antisense	4
His2A:CG33847	TCGTTGCGGATGGCCAGTTGC	antisense	3
His2A:CG33847	CAAGAAGACCGAGAAGAAGGC	sense	2
His2A:CG33847	GTTAAACAAGCTGCTCTCCGG	sense	1
His2A:CG33850	GCGTAGTTTCCCTTCCGGAGC	antisense	4
His2A:CG33850	TCGTTGCGGATGGCCAGTTGC	antisense	3
His2A:CG33850	CAAGAAGACCGAGAAGAAGGC	sense	2
His2A:CG33850	GTTAAACAAGCTGCTCTCCGG	sense	1
His2A:CG33853	GCGTAGTTTCCCTTCCGGAGC	antisense	4
His2A:CG33853	TCGTTGCGGATGGCCAGTTGC	antisense	3
His2A:CG33853	CAAGAAGACCGAGAAGAAGGC	sense	2
His2A:CG33853	GTTAAACAAGCTGCTCTCCGG	sense	1
His2A:CG33856	GCGTAGTTTCCCTTCCGGAGC	antisense	4
His2A:CG33856	TCGTTGCGGATGGCCAGTTGC	antisense	3
His2A:CG33856	CAAGAAGACCGAGAAGAAGGC	sense	2
His2A:CG33856	GTTAAACAAGCTGCTCTCCGG	sense	1
His2A:CG33859	GCGTAGTTTCCCTTCCGGAGC	antisense	4
His2A:CG33859	TCGTTGCGGATGGCCAGTTGC	antisense	3
His2A:CG33859	CAAGAAGACCGAGAAGAAGGC	sense	2
His2A:CG33859	GTTAAACAAGCTGCTCTCCGG	sense	1
His2A:CG33862	GCGTAGTTTCCCTTCCGGAGC	antisense	4
His2A:CG33862	TCGTTGCGGATGGCCAGTTGC	antisense	3
His2A:CG33862	CAAGAAGACCGAGAAGAAGGC	sense	2
His2A:CG33862	GTTAAACAAGCTGCTCTCCGG	sense	1
His2A:CG33865	GCGTAGTTTCCCTTCCGGAGC	antisense	4
His2A:CG33865	TCGTTGCGGATGGCCAGTTGC	antisense	3
His2A:CG33865	CAAGAAGACCGAGAAGAAGGC	sense	2
His2A:CG33865	GTTAAACAAGCTGCTCTCCGG	sense	1
His2Av	TGTATGTGAGTGTTGGGGAGA	antisense	1
His2B:CG17949	TCGCCTTCGACGAAATTCCGG	antisense	3
His2B:CG17949	CTTGGTGATGTTCTTCTGAGC	antisense	1
His2B:CG17949	GACGAAATTCCGGTGTCAGGA	antisense	1
His2B:CG33868	TCGCCTTCGACGAAATTCCGG	antisense	3
His2B:CG33868	CTTGGTGATGTTCTTCTGAGC	antisense	1
His2B:CG33868	GACGAAATTCCGGTGTCAGGA	antisense	1
His2B:CG33870	TCGCCTTCGACGAAATTCCGG	antisense	3
His2B:CG33870	CTTGGTGATGTTCTTCTGAGC	antisense	1
His2B:CG33870	GACGAAATTCCGGTGTCAGGA	antisense	1
His2B:CG33872	TCGCCTTCGACGAAATTCCGG	antisense	3
His2B:CG33872	CTTGGTGATGTTCTTCTGAGC	antisense	1
His2B:CG33872	GACGAAATTCCGGTGTCAGGA	antisense	1
His2B:CG33874	TCGCCTTCGACGAAATTCCGG	antisense	3
His2B:CG33874	CTTGGTGATGTTCTTCTGAGC	antisense	1
His2B:CG33874	GACGAAATTCCGGTGTCAGGA	antisense	1
His2B:CG33876	TCGCCTTCGACGAAATTCCGG	antisense	3
His2B:CG33876	CTTGGTGATGTTCTTCTGAGC	antisense	1
His2B:CG33876	GACGAAATTCCGGTGTCAGGA	antisense	1
His2B:CG33878	TCGCCTTCGACGAAATTCCGG	antisense	3
His2B:CG33878	CTTGGTGATGTTCTTCTGAGC	antisense	1
His2B:CG33878	GACGAAATTCCGGTGTCAGGA	antisense	1
His2B:CG33880	TCGCCTTCGACGAAATTCCGG	antisense	3
His2B:CG33880	CTTGGTGATGTTCTTCTGAGC	antisense	1
His2B:CG33880	GACGAAATTCCGGTGTCAGGA	antisense	1
His2B:CG33882	TCGCCTTCGACGAAATTCCGG	antisense	3

His2B:CG33882	CTTGGTGATGTTCTTCTGAGC	antisense	1
His2B:CG33882	GACGAAATTCCGGTGTCAGGA	antisense	1
His2B:CG33884	TCGCCTTCGACGAAATTCCGG	antisense	3
His2B:CG33884	CTTGGTGATGTTCTTCTGAGC	antisense	1
His2B:CG33884	GACGAAATTCCGGTGTCAGGA	antisense	1
His2B:CG33886	TCGCCTTCGACGAAATTCCGG	antisense	3
His2B:CG33886	CTTGGTGATGTTCTTCTGAGC	antisense	1
His2B:CG33886	GACGAAATTCCGGTGTCAGGA	antisense	1
His2B:CG33888	TCGCCTTCGACGAAATTCCGG	antisense	3
His2B:CG33888	CTTGGTGATGTTCTTCTGAGC	antisense	1
His2B:CG33888	GACGAAATTCCGGTGTCAGGA	antisense	1
His2B:CG33890	TCGCCTTCGACGAAATTCCGG	antisense	3
His2B:CG33890	CTTGGTGATGTTCTTCTGAGC	antisense	1
His2B:CG33890	GACGAAATTCCGGTGTCAGGA	antisense	1
His2B:CG33892	TCGCCTTCGACGAAATTCCGG	antisense	3
His2B:CG33892	CTTGGTGATGTTCTTCTGAGC	antisense	1
His2B:CG33892	GACGAAATTCCGGTGTCAGGA	antisense	1
His2B:CG33894	TCGCCTTCGACGAAATTCCGG	antisense	3
His2B:CG33894	CTTGGTGATGTTCTTCTGAGC	antisense	1
His2B:CG33894	GACGAAATTCCGGTGTCAGGA	antisense	1
His2B:CG33896	TCGCCTTCGACGAAATTCCGG	antisense	3
His2B:CG33896	CTTGGTGATGTTCTTCTGAGC	antisense	1
His2B:CG33896	GACGAAATTCCGGTGTCAGGA	antisense	1
His2B:CG33898	TCGCCTTCGACGAAATTCCGG	antisense	3
His2B:CG33898	CTTGGTGATGTTCTTCTGAGC	antisense	1
His2B:CG33898	GACGAAATTCCGGTGTCAGGA	antisense	1
His2B:CG33900	TCGCCTTCGACGAAATTCCGG	antisense	3
His2B:CG33900	CTTGGTGATGTTCTTCTGAGC	antisense	1
His2B:CG33900	GACGAAATTCCGGTGTCAGGA	antisense	1
His2B:CG33902	TCGCCTTCGACGAAATTCCGG	antisense	3
His2B:CG33902	CTTGGTGATGTTCTTCTGAGC	antisense	1
His2B:CG33902	GACGAAATTCCGGTGTCAGGA	antisense	1
His2B:CG33904	TCGCCTTCGACGAAATTCCGG	antisense	3
His2B:CG33904	CTTGGTGATGTTCTTCTGAGC	antisense	1
His2B:CG33904	GACGAAATTCCGGTGTCAGGA	antisense	1
His2B:CG33906	TCGCCTTCGACGAAATTCCGG	antisense	3
His2B:CG33906	CTTGGTGATGTTCTTCTGAGC	antisense	1
His2B:CG33906	GACGAAATTCCGGTGTCAGGA	antisense	1
His2B:CG33908	TCGCCTTCGACGAAATTCCGG	antisense	3
His2B:CG33908	CTTGGTGATGTTCTTCTGAGC	antisense	1
His2B:CG33908	GACGAAATTCCGGTGTCAGGA	antisense	1
His2B:CG33910	TCGCCTTCGACGAAATTCCGG	antisense	3
His2B:CG33910	CTTGGTGATGTTCTTCTGAGC	antisense	1
His2B:CG33910	GACGAAATTCCGGTGTCAGGA	antisense	1
His2B:CG40461	TCGCCTTCGACGAAATTCCGG	antisense	3
His2B:CG40461	GACGAAATTCCGGTGTCAGGA	antisense	1
His3:CG31613	GAAGCTCGGTGCTCTTTTGGT	antisense	5
His3:CG31613	AGTAGCCAGTTGTTTGCGTGG	antisense	2
His3:CG31613	CTTTAGTGAAACCCAAATCGG	sense	1
His3:CG31613	GTAGCCAGTTGTTTGCGTGGC	antisense	1
His3:CG31613	TGGTAGCGACGAATTTCACGC	antisense	1
His3:CG33803	GAAGCTCGGTGCTCTTTTGGT	antisense	5
His3:CG33803	AGTAGCCAGTTGTTTGCGTGG	antisense	2
His3:CG33803	CTTTAGTGAAACCCAAATCGG	sense	1
His3:CG33803	GTAGCCAGTTGTTTGCGTGGC	antisense	1
His3:CG33803	TGGTAGCGACGAATTTCACGC	antisense	1
His3:CG33806	GAAGCTCGGTGCTCTTTTGGT	antisense	5
His3:CG33806	AGTAGCCAGTTGTTTGCGTGG	antisense	2
His3:CG33806	CTTTAGTGAAACCCAAATCGG	sense	1
His3:CG33806	GTAGCCAGTTGTTTGCGTGGC	antisense	1
His3:CG33806	TGGTAGCGACGAATTTCACGC	antisense	1
His3:CG33809	GAAGCTCGGTGCTCTTTTGGT	antisense	5
His3:CG33809	AGTAGCCAGTTGTTTGCGTGG	antisense	2
His3:CG33809	CTTTAGTGAAACCCAAATCGG	sense	1
His3:CG33809	GTAGCCAGTTGTTTGCGTGGC	antisense	1
His3:CG33809	TGGTAGCGACGAATTTCACGC	antisense	1
His3:CG33812	GAAGCTCGGTGCTCTTTTGGT	antisense	5
His3:CG33812	AGTAGCCAGTTGTTTGCGTGG	antisense	2
His3:CG33812	CTTTAGTGAAACCCAAATCGG	sense	1
His3:CG33812	GTAGCCAGTTGTTTGCGTGGC	antisense	1
His3:CG33812	TGGTAGCGACGAATTTCACGC	antisense	1
His3:CG33815	GAAGCTCGGTGCTCTTTTGGT	antisense	5
His3:CG33815	AGTAGCCAGTTGTTTGCGTGG	antisense	2
His3:CG33815	CTTTAGTGAAACCCAAATCGG	sense	1
His3:CG33815	GTAGCCAGTTGTTTGCGTGGC	antisense	1
His3:CG33815	TGGTAGCGACGAATTTCACGC	antisense	1
His3:CG33818	GAAGCTCGGTGCTCTTTTGGT	antisense	5
His3:CG33818	AGTAGCCAGTTGTTTGCGTGG	antisense	2
His3:CG33818	CTTTAGTGAAACCCAAATCGG	sense	1

His3:CG33818	GTAGCCAGTTGTTTGCGTGGC	antisense	1
His3:CG33818	TGGTAGCGACGAATTTCACGC	antisense	1
His3:CG33821	GAAGCTCGGTGCTCTTTTGGT	antisense	5
His3:CG33821	AGTAGCCAGTTGTTTGCGTGG	antisense	2
His3:CG33821	CTTTAGTGAAACCCAAATCGG	sense	1
His3:CG33821	GTAGCCAGTTGTTTGCGTGGC	antisense	1
His3:CG33821	TGGTAGCGACGAATTTCACGC	antisense	1
His3:CG33824	GAAGCTCGGTGCTCTTTTGGT	antisense	5
His3:CG33824	AGTAGCCAGTTGTTTGCGTGG	antisense	2
His3:CG33824	CTTTAGTGAAACCCAAATCGG	sense	1
His3:CG33824	GTAGCCAGTTGTTTGCGTGGC	antisense	1
His3:CG33824	TGGTAGCGACGAATTTCACGC	antisense	1
His3:CG33827	GAAGCTCGGTGCTCTTTTGGT	antisense	5
His3:CG33827	AGTAGCCAGTTGTTTGCGTGG	antisense	2
His3:CG33827	CTTTAGTGAAACCCAAATCGG	sense	1
His3:CG33827	GTAGCCAGTTGTTTGCGTGGC	antisense	1
His3:CG33827	TGGTAGCGACGAATTTCACGC	antisense	1
His3:CG33830	GAAGCTCGGTGCTCTTTTGGT	antisense	5
His3:CG33830	AGTAGCCAGTTGTTTGCGTGG	antisense	2
His3:CG33830	CTTTAGTGAAACCCAAATCGG	sense	1
His3:CG33830	GTAGCCAGTTGTTTGCGTGGC	antisense	1
His3:CG33830	TGGTAGCGACGAATTTCACGC	antisense	1
His3:CG33833	GAAGCTCGGTGCTCTTTTGGT	antisense	5
His3:CG33833	AGTAGCCAGTTGTTTGCGTGG	antisense	2
His3:CG33833	CTTTAGTGAAACCCAAATCGG	sense	1
His3:CG33833	GTAGCCAGTTGTTTGCGTGGC	antisense	1
His3:CG33833	TGGTAGCGACGAATTTCACGC	antisense	1
His3:CG33836	GAAGCTCGGTGCTCTTTTGGT	antisense	5
His3:CG33836	AGTAGCCAGTTGTTTGCGTGG	antisense	2
His3:CG33836	CTTTAGTGAAACCCAAATCGG	sense	1
His3:CG33836	GTAGCCAGTTGTTTGCGTGGC	antisense	1
His3:CG33836	TGGTAGCGACGAATTTCACGC	antisense	1
His3:CG33839	GAAGCTCGGTGCTCTTTTGGT	antisense	5
His3:CG33839	AGTAGCCAGTTGTTTGCGTGG	antisense	2
His3:CG33839	CTTTAGTGAAACCCAAATCGG	sense	1
His3:CG33839	GTAGCCAGTTGTTTGCGTGGC	antisense	1
His3:CG33839	TGGTAGCGACGAATTTCACGC	antisense	1
His3:CG33842	GAAGCTCGGTGCTCTTTTGGT	antisense	5
His3:CG33842	AGTAGCCAGTTGTTTGCGTGG	antisense	2
His3:CG33842	CTTTAGTGAAACCCAAATCGG	sense	1
His3:CG33842	GTAGCCAGTTGTTTGCGTGGC	antisense	1
His3:CG33842	TGGTAGCGACGAATTTCACGC	antisense	1
His3:CG33845	GAAGCTCGGTGCTCTTTTGGT	antisense	5
His3:CG33845	AGTAGCCAGTTGTTTGCGTGG	antisense	2
His3:CG33845	CTTTAGTGAAACCCAAATCGG	sense	1
His3:CG33845	GTAGCCAGTTGTTTGCGTGGC	antisense	1
His3:CG33845	TGGTAGCGACGAATTTCACGC	antisense	1
His3:CG33848	GAAGCTCGGTGCTCTTTTGGT	antisense	5
His3:CG33848	AGTAGCCAGTTGTTTGCGTGG	antisense	2
His3:CG33848	CTTTAGTGAAACCCAAATCGG	sense	1
His3:CG33848	GTAGCCAGTTGTTTGCGTGGC	antisense	1
His3:CG33848	TGGTAGCGACGAATTTCACGC	antisense	1
His3:CG33851	GAAGCTCGGTGCTCTTTTGGT	antisense	5
His3:CG33851	AGTAGCCAGTTGTTTGCGTGG	antisense	2
His3:CG33851	CTTTAGTGAAACCCAAATCGG	sense	1
His3:CG33851	GTAGCCAGTTGTTTGCGTGGC	antisense	1
His3:CG33851	TGGTAGCGACGAATTTCACGC	antisense	1
His3:CG33854	GAAGCTCGGTGCTCTTTTGGT	antisense	5
His3:CG33854	AGTAGCCAGTTGTTTGCGTGG	antisense	2
His3:CG33854	CTTTAGTGAAACCCAAATCGG	sense	1
His3:CG33854	GTAGCCAGTTGTTTGCGTGGC	antisense	1
His3:CG33854	TGGTAGCGACGAATTTCACGC	antisense	1
His3:CG33857	GAAGCTCGGTGCTCTTTTGGT	antisense	5
His3:CG33857	AGTAGCCAGTTGTTTGCGTGG	antisense	2
His3:CG33857	CTTTAGTGAAACCCAAATCGG	sense	1
His3:CG33857	GTAGCCAGTTGTTTGCGTGGC	antisense	1
His3:CG33857	TGGTAGCGACGAATTTCACGC	antisense	1
His3:CG33860	GAAGCTCGGTGCTCTTTTGGT	antisense	5
His3:CG33860	AGTAGCCAGTTGTTTGCGTGG	antisense	2
His3:CG33860	CTTTAGTGAAACCCAAATCGG	sense	1
His3:CG33860	GTAGCCAGTTGTTTGCGTGGC	antisense	1
His3:CG33860	TGGTAGCGACGAATTTCACGC	antisense	1
His3:CG33863	GAAGCTCGGTGCTCTTTTGGT	antisense	5
His3:CG33863	AGTAGCCAGTTGTTTGCGTGG	antisense	2
His3:CG33863	CTTTAGTGAAACCCAAATCGG	sense	1
His3:CG33863	GTAGCCAGTTGTTTGCGTGGC	antisense	1
His3:CG33863	TGGTAGCGACGAATTTCACGC	antisense	1
His3:CG33866	GAAGCTCGGTGCTCTTTTGGT	antisense	5
His3:CG33866	AGTAGCCAGTTGTTTGCGTGG	antisense	2

His3:CG33866	CTTTAGTGAAACCCAAATCGG	sense	1
His3:CG33866	GTAGCCAGTTGTTTGCGTGGC	antisense	1
His3:CG33866	TGGTAGCGACGAATTTCACGC	antisense	1
Hmgcr	AATGGATAAATGTCTTTTGTC	sense	1
hoip	AGTTTCCCGGCCAATAGTCGC	sense	2
Hr4	ATGATGATGATGATGATGATG	antisense	1
Hr96	AAACTGCGACATCACTGTGGT	sense	1
Hs3st-A	ATGATGATGATGATGATGATG	sense	1
Hsp70Aa	ATACTCCGGCGCTCTTTTCGC	antisense	7
Hsp70Aa	TACTCCGGCGCTCTTTTCGCG	antisense	3
Hsp70Aa	TCTTTTCGCGAACATTCGAGG	antisense	3
Hsp70Aa	AGATTGTTTAGCTTGTTCAGC	antisense	2
Hsp70Aa	GAGCGCGCCTCGAATGTTCGC	sense	1
Hsp70Aa	TCTATTTATACTCCGGCGCTC	antisense	1
Hsp70Ab	AGATTGTTTAGCTTGTTCAGC	antisense	2
Hsp70Ba	AGATTGTTTAGCTTGTTCAGC	antisense	2
Hsp70Ba	CTTTAACTTGCACTTTACTGC	antisense	1
Hsp70Bb	AGATTGTTTAGCTTGTTCAGC	antisense	2
Hsp70Bb	CTTTAACTTGCACTTTACTGC	antisense	1
Hsp70Bbb	AGATTGTTTAGCTTGTTCAGC	antisense	2
Hsp70Bbb	CTTTAACTTGCACTTTACTGC	antisense	1
Hsp70Bc	AGATTGTTTAGCTTGTTCAGC	antisense	2
Hsp70Bc	CTTTAACTTGCACTTTACTGC	antisense	1
htt	ATGATGATGATGATGATGATG	sense	1
hyd	TGTTTGTAGCCATTGGCTGGC	antisense	1
Iap2	AACTCAAGAGCTTGCGCTGGA	antisense	3
Ice	CAGGGTATCTGCTTCTGCTGG	antisense	6
icln	GTCCTCATCGTGCTCATCGCT	antisense	2
icln	GCATCTACTTCATGCTGGACC	sense	1
igl	TATGAATTAGAAACCAAAGGA	sense	1
Ilp6	AGTCCTGGCCACCTTGTTCGC	sense	3
IM10	CCAATGGAGCTAAGAGCGTGG	sense	14
IP3K1	GAGGTCCGGTCTGAGCGTGGC	antisense	4
IP3K1	TAATATAGGGTGGAGCTCAGC	antisense	4
IP3K1	CTTGGTGATCGCTCGTGCCTC	antisense	3
IP3K1	GGATTCAGCTCGGTTAGAAAG	sense	2
IP3K1	AACGGATCTCAGGACTGGAGG	sense	1
IP3K1	AGGCGCATGTGCTTTAGTCGC	antisense	1
Irbp	TAACTGCGCAGTGTTCTTGCC	sense	6
itp	GAATTAAGTAGCTGCTCGTGC	antisense	1
Itp-r83A	CTATCTTGCGGTTCCTTCAGC	sense	3
Itp-r83A	CGTAGAGGTGGCCTGCAAGGC	sense	1
jet	CCGCACAACGCTTTGTGGAGG	sense	7
jet	TGCACATGCTGCCGCCTCCGG	antisense	5
Jon99Fi	GCAAGGTTCCCTACATCGTGG	sense	5
Khc	GAAGGTTTCCAAGACTGGAGC	sense	2
kis	CTTCATTGCGGGGACAGCTGG	antisense	3
kis	GTGGCTCCTCTGGTCGTGGTC	sense	1
Klp3A	CGTGTAGAGCTATTGTCGGGC	sense	7
Klp3A	TGGCGTGACTCATGGCCCTGG	antisense	2
ksr	TTTAGCCGCATCTTCGCCAGC	antisense	4
kuz	GTTGTTGGGCGCCCGCGTGGT	antisense	2
l(1)G0004	TGGTGGAGCACATGAAGCTGC	sense	4
l(1)G0004	CGGATCTGCAGCTTCATGTGC	antisense	1
l(2)37Cb	CAGGAGGCTCAAGTGGTCGCC	antisense	4
l(2)NC136	CCTTTTGAAGTGGTGTTGGCC	antisense	1
l(2)tid	GAAAGCCTACTACCAGCTGGC	sense	1
l(3)01239	AATTGGTGGTGTGCTGTGCGA	sense	3
l(3)s1921	AAGGAACGCCCACTGAAAGAG	sense	2
l(3)s1921	TTGGTGAGGATGTCTAACGCC	antisense	2
l(3)s1921	GTGAAGAGGGCTCGAAAGCGC	antisense	1
l(3)s1921	TGTTCTCGTACTCGCACATGT	antisense	1
Lac	ATATCCTTGATCTGCTCCTGC	antisense	1
lack	AGCTGGTACGTCTCGTAGGGC	antisense	4
lack	TGTAGAGCGGGTTGGCCCAGC	antisense	4
larp	CGAAGATACCGTTACCGCTGG	antisense	7
ldlCp	AAGGATCAGAAGCCCCAGTGG	antisense	6
ldlCp	CCAAGCTGGGACTCGTTCTGC	antisense	3
ldlCp	CAAATCGGTCATATGCAGCGA	antisense	2
ldlCp	CTTAAGTCCCGACCATCAAGG	sense	2
lid	TTGCTCGCAGAACCCGTTGCC	antisense	5
lig	CTTGCCAGCAGACTTCGACTG	antisense	2
lkb1	ACATTGCATAAACCTCCGCGG	antisense	1
lok	CACATTTGGTAAAAGTAAAGC	antisense	2
lok	TGGTTTGCTGAAGAGTCATTC	antisense	2
lok	GAGGTGTGCCACAGAGCGTGC	antisense	1
Lsd-1	TTGTTAAGCGTGAAGTTAGGC	antisense	2
Lsd-1	AATATAAGCCTAACTTCACGC	sense	1
Lsd-1	ATTGTTAAGCGTGAAGTTAGG	antisense	1

M(2)21AB	AAGGCCGGTCTCTGCAAGCGC	sense	6
M(2)21AB	CAACACACGTCCTATCTTAGC	antisense	6
M(2)21AB	CGGCACCGATCTCTTCCTCGG	antisense	5
M(2)21AB	TACTTGTGTTTTGGAATCGGG	antisense	1
mam	TCGTCTGGCTAATGGAACTGG	antisense	6
Map60	CGGTGGTGTCTGGTTGTCCTC	antisense	5
mask	TTGGTTGCAGGCAGTGCTGGG	antisense	2
MBD-like	CCACGGCCCTGCATATGCTCA	sense	12
Mcm2	CATAGCCGCACTTAACGCAGT	antisense	1
MCPH1	TGCGGTGGCTCTTGTTCATGG	antisense	5
MCPH1	CCAGTCGTTTCCATTCTGTGG	sense	1
Med	CAATCAGCAAATGGGCGGCGG	sense	2
MED15	CACACTCACACTTACGGGCGG	antisense	5
MED21	TAAACCATTTTGAACAGCACC	sense	2
MED24	CACCGCATTCCACCAGCAGCG	antisense	11
Mes-4	ATCGATGCGGGACCGAAGGGC	sense	1
Mi-2	TTGTACAAGGAGGGCCATTGC	sense	5
mib1	AAGGCTGTGCAGACTGTGCGC	antisense	1
milt	ATCATCATGCGGCTAATGCGG	antisense	3
mip130	CTAATCTGCAGCGAAAACCGC	sense	3
mit(1)15	CGAGAGCCTTAAAGATGCGTC	antisense	5
Mitf	TAGATGTGCCACCCCAAGTGC	sense	6
Mitf	ACTTTTTAAACTTCTGCAGGG	sense	1
Mitf	TGATGAAAGCCTTTTTAGAGG	sense	1
Mmp1	CGGCACATGGGAAAGGAGCTC	sense	10
Mmp1	TTTAACGAGCCATATGCAAGG	sense	4
Mmp1	AAATTGTAGCACAGCTGGAGG	antisense	2
Mmp1	TCGTTTTCGTTTTTGTTCAGT	antisense	2
Mmp1	TATTTCAGCGTACATTTATGG	antisense	1
mod(mdg4)	ACGTCTCGCTGGCCGCCGAGG	sense	2
mod(mdg4)	TTAACTGCTGGTCCTCCGGAG	antisense	2
mod(mdg4)	CTTCCTGTGCCGATCGCTGCC	antisense	1
MP1	CTCAAACTACTACATTGCTGG	sense	3
MP1	TCGATATTTACTGTGGCATGT	antisense	1
mRpL18	CTTTTTTCGGGTTTCCATCGC	sense	1
mRpL44	ACTGCGTGTAAAGATTTGGCC	antisense	1
mRpL48	CGACCTGAAGGACGAGCTGGA	sense	5
mRpL48	TTGTTCAGCTGCAGGTTGCGC	antisense	1
mRpS2	ATACAAATGCCAATGGCCGTC	sense	3
mrt	TGAAACCACAAAATACTTAGG	sense	2
msl-1	TTAGGGCTCTACAATGGTGGC	antisense	4
Mst89B	TGGCGGTGTCCTTCGTTTGGC	antisense	6
mt:CoI	TGGGAATGCTATATCAGGAGC	antisense	2
mt:CoI	ATTTTGACTACTACCTCCTGC	sense	1
MTF-1	GGGGTGACTGTGGTGCTGCAG	sense	4
mtTFB1	GTATGGTGAACTTCATGACGG	antisense	2
mus205	TTCGGCCACTGAAGTTAGCGG	antisense	2
mus205	TTTAGGTTTATCGTGTTATGA	sense	2
mus308	AAGTGGCATTGCTGGTCTTTC	antisense	2
mus309	TGTTTTCTTGGACTTCCAGCC	antisense	1
mus81	CGATGGAGCGCAGTGAGTCGC	sense	4
mus81	TGCTCACGCATTCCGACCTGG	sense	2
mus81	CAGTGAGTCGCGGAATCCGGT	sense	1
Mys45A	AACCGTACAAAGTCTATATGA	antisense	3
nAcRbeta-21C	TAATGATGAGACCTCGTATGG	sense	6
ncd	TTCGAACCGTTCATTTTGTGG	antisense	1
NitFhit	GCAAGATGGTGCACAGGCGGG	sense	2
NitFhit	CGGCGGTGCTGCTCAGGAAGC	sense	1
nito	CAAACAACTTCTGGAGCCACT	antisense	1
Nle	AGCGGGCCTTATTTATACATC	sense	1
Nup154	TTGCCAAACCAACTGGACTCG	antisense	1
Nup44A	TGTGTGTATGGGCGAGAGAGC	antisense	2
Nup98	ACGATGAACTGGTGGACCTGG	sense	7
Obp99c	CAACTTCGCCGCGATCGTGCA	sense	3
O-fut1	TGAGGTAGCCATTGGGATCGC	antisense	4
O-fut1	TTGAAGTGGAGCCCCTGAAGG	sense	1
omd	TGATGGTCATCCGCCGCTGGG	antisense	4
opa1-like	TTAGTTAAGCATACTTTGTGC	sense	2
Orc1	TCCAAGCTGGGCGCCGAGCGA	sense	1
osa	AGATTTCTGGCGGTACTTGGT	sense	3
osa	CGCATCTTCCGCCAGCAGAGC	antisense	3
osa	TTCGGCCTGCTCGATGAAGGC	antisense	2
osa	GAATAGGATGCCCGGCATGCC	sense	1
P58IPK	TCTTTAATTGCTTGTTCTCGC	antisense	1
pAbp	TTCGCTTTGTCGTGCGTTGCG	antisense	1
par-1	ACAATGCTGCAGGATCAGCGG	sense	2
par-1	TTGGACACGCTATCCGCTGGC	antisense	1
Pc	ATTGGCTAGTTTTAGTTACGG	sense	1
Pcaf	TGGTTACGCTTTCTCTGCTGG	antisense	2

Pdk	CGACTTCGAGGGCTGCGGCGC	antisense	4
Pect	CTGGATCAGGAAGTAGCTGCT	sense	4
Pect	TACAAGGCTTTTCAATTCGGC	antisense	4
Pect	AGTTTGTCTTGTTTAGTATGT	sense	2
Pect	TCTTAGACTTAAATACATGGC	sense	1
Pen	GAGGACCAGATGTTCAAGCGG	sense	2
pie	CTCAAAGATGCCGCGGTCCTC	antisense	2
PIP82	ATGATGATGATGATGATGATG	antisense	1
pita	ATATTGCATGCGAAAAGTGCA	sense	1
pita	TAAAAGGCCTTCGGTTAAAGG	antisense	1
Pitslre	ACATGTTCTTCACGGCTGGGA	antisense	2
Pka	TCAAGCAGGGAGCAGCTGGCG	sense	8
Pms2	CAATGACGTATACTTAGTGGC	antisense	1
pnt	ATGATGATGATGATGATGATG	antisense	1
Pof	TCGTTTTCGTTTTTGTTCAGT	sense	2
Pof	TATTTCAGCGTACATTTATGG	sense	1
por	ATGAAGCTTATCTCGCTGGGC	sense	1
por	TCATTTACTTTGTTTTTCCGC	sense	1
Pp1-87B	TGTTTGCGTGCGAAAGTGTGG	antisense	4
Pp2C1	GTTGTTTCTTTGGATTTAAAG	antisense	1
ppk13	AGACCCTTTCGACCACGGAGC	antisense	3
ppk13	AAGCTTTAATAAGCAACCAGG	antisense	2
ppk13	CGAGATTGCCTTCTTCCGCGG	sense	2
ppk13	ATCGCATGGAGACTGAGCTGG	sense	1
ppk13	CAGTTACCTCCTCGCCACTGC	antisense	1
ppk13	TCGATTCGCAGACTATTATGG	antisense	1
ppk13	TGCTCCAGTCACCTGTCCCCG	sense	1
Ptp99A	AACAAGAGCGACTATGTGAGC	sense	6
Ptp99A	TAGCTTTGCAGAGTGTGAAGG	antisense	3
Ptp99A	ACACGCACCCGCAACAACATC	sense	1
Pvf2	TGGAGCTGCGTCTGTGGGAGC	sense	3
Pvf2	CTCGACCCTTTTTTTGAGCTC	antisense	1
pyd	AAGGCAAGCAGAGGCAGCTGC	sense	1
pyd	ATATTGTCTGTTAATTGTGCC	sense	1
qkr58E-1	ATCGTCCATCACATATCGAGC	antisense	1
Rab11	TTTGTTGTTGTTCTGCTCGCG	antisense	1
Rab6	TGGGCCAGCGGTTTTCGGGGC	antisense	11
RabX6	GACCTTCTTGTTAATCGATGG	antisense	5
RabX6	GAAGAAAGCTGCAGATTACGC	antisense	3
Rack1	TTTGTCGGTTGCCTTCTCGGC	sense	1
Rbf	CGAGATCGTGGTGCGGCACGG	sense	1
Rbm13	ATGATGATGATGATGATGATG	sense	1
ref(2)P	TTGAGCAGTCTGGGTGTTGGC	antisense	3
Rfabg	ATGGAGTCAAAAATGTGCAAG	sense	2
RfC38	CAAAGGTTGCGTAGGTTCTCC	antisense	1
RhoGAP16F	CAGGAACGCGACAGGGAGAGG	sense	8
RhoGAP16F	GAGCAGAGCCAGCTTGTTGGG	antisense	8
RhoGAP16F	CAGCACCGCACTGTGCCGCGC	sense	7
RhoGAP16F	TTGCTCTGCGTCTTGTGTTCC	antisense	3
RhoGAP16F	TTGTTGGGTCTGGTTTTCAGC	antisense	2
RhoGAP68F	CAGTACGATTTTTGAGTTCGC	antisense	4
RhoGAP68F	TGAAGAGTTTCCTGCGCGATC	sense	1
Ric	CGAAAACGAATCAAATGCGGG	sense	2
r-l	CCGCCTGCTTCGGGCTGCTGG	antisense	6
r-l	CGGACTTCTGCAACTGACGCC	sense	5
r-l	TAAGGATTTCCTCGTTGGATC	antisense	5
r-l	GACCCAGCTTTGCCGTTGAGT	sense	3
r-l	TTTTGCGGAAGATTCATAGCC	antisense	1
Rlip	GTTCGGTCGCCAAGCTAACGG	antisense	2
RpL21	TACGAATTCATTGCCTAAAGG	sense	2
RpL28	AAGTACTGCATACTTTGGGGC	sense	2
RpLP2	TAATAAAAATCAGCAGTGTTT	sense	1
RpLP2	TCCTCCTTCTTGGCCTCCTTC	antisense	1
Rpn2	CGGTCCGCGTGTCTATCTGCC	sense	4
Rpp20	CTGGCGGAAGGTTATGTCCTC	antisense	3
RpS14a	CAATCTTCATGGACGAACGGG	antisense	4
RpS14b	CAATCTTCATGGACGAACGGG	antisense	4
RpS7	CAGATCCTCAAGGATGGCGTC	antisense	6
Rpt4	AAGCCGTCCATCTGGTTGAGC	antisense	2
Rrp4	TGTGGAAGTGCATCTTGCGGC	antisense	2
Rrp42	CGCCGAAGGATCCACCAGGAC	antisense	6
Sas10	GCAAATACCGAAAGGCGCTCA	sense	9
sas-6	TGCGCTGCTCGTTTATTTTGG	antisense	5
sav	GAAGGTCTGCCGGTGGGCTGG	sense	3
sav	TAGATGGGCGACTCCGATCGC	antisense	1
sax	CAGAGCGATGACCAGAAAAGG	antisense	2
sbb	TGGATGCACTCCATCACGCGG	sense	2
ScpX	ATGCTCAGTTGGCTTTGCAGC	sense	2
scra	AATTCAACGCACCACATGGAC	sense	3

scra	CGGTTCGGTGGATTGGGGAGG	antisense	3
scra	ACTCGTCGCAGGTCTCAGCGG	sense	1
scu	CCTGCTTGGCCAGGCGCTCGG	antisense	1
sec15	CGGGAGTACTTCGAGAAGGAC	sense	7
sec15	TTATTTGGAGGATCTGTGGTC	sense	1
sec23	ATAGAATGTGGTCTCGTCGGG	antisense	1
sec31	CTACACCCAGCCACAGGCAGC	sense	5
sec63	CACCAATGTGGTGACCGCCGG	sense	3
sec71	CGTGGCCTTGTCAAACAACGG	sense	2
sens	ATGATGATGATGATGATGATG	sense	1
Sin3A	TCCACCACCAATACCAATGCG	sense	10
Sin3A	TGGGCTTGCCTGGGACTCGTG	antisense	4
Sin3A	ATTAAAGGCGTATTGCTCGGC	antisense	3
Sin3A	CACTATTTATTTCAATGCAGG	antisense	2
Sin3A	TCCAAAAGTCTGCTTGCTGGC	sense	2

skd	TGTGTGTTTGCCCAGCTTGCA	antisense	6
sle	TCCATTGAAAGTTTCTCGAGC	antisense	2
slik	CGCGTTCCTTCTCTTCGTGGT	antisense	1
slmo	GGGCTCGGTGTCACGTCGGGC	sense	6
Snoo	CAGGCGAATTGGAATAGTAGG	sense	5
Snoo	ATACACGTGGAAGCTGAGGGC	antisense	2
Snoo	ATTGGATGCTTCTCCGTGGGC	sense	2
Snoo	CGAGGACACCGGCGGTGGTGA	antisense	2
Snoo	AACTTTTTTATCTTTGCGCTG	antisense	1
Snoo	ATAGTCACCAATGGCAGGAGC	sense	1

Table II-S1B. Summary of mRNA-matching, 21-nt reads from pyrosequencing of a small RNA library enriched for 3´ terminally modified small RNA.

	Total S2 reads			number of unique 21-mers	
Gene	**sense + antisense**	**antisense**	**sense**	**antisense**	**sense**
5Ptasel	2	0	2	0	1
Aats-gln	1	1	0	1	0
Ac3	3	3	0	1	0
Acer	1	0	1	0	1
Ack	1	1	0	1	0
Act42A	21	8	13	2	2
Ada2b	6	6	0	3	0
ade2	1	1	0	1	0
ago (archipeligo)	14	14	0	1	0
AGO2	40	23	17	6	3
AnnIX	9	3	6	1	2
AnnX	3	3	0	1	0
Aos1	3	0	3	0	1
Apc2	1	1	0	1	1
Arf79F	3	3	0	2	0
argos	2	0	2	0	1
Arp5	2	0	2	0	1
Art1	3	0	3	0	1
Art4	6	1	5	1	1
ATPCL	13	13	0	3	0
aux	16	7	9	1	4
Bap170	3	3	0	1	0
betaggt-I	1	1	0	1	0
bigmax	3	0	3	0	1
bin3	16	16	0	5	0
Bj1	2	2	0	1	0
blue	7	7	0	1	0
bocksbeutel	1	0	1	0	1
botv	2	0	2	0	1
brat	5	0	5	0	1
Bruce	3	1	2	1	1
BRWD3	2	2	0	1	0
BtbVII	5	0	5	0	1
btn	2	0	2	0	1
Bzd	2	0	2	0	1
cact	1	0	1	1	0
Cap-D3	1	0	1	0	1
Cap-H2	2	2	0	1	0
cbt	18	11	7	2	1
Ccn	1	0	1	0	1
Cct1	4	4	1	1	0
Cct5	6	3	3	2	2
cg	2	2	0	1	0
CG10011	1	1	0	1	0
CG10151	4	0	4	0	1
CG10214	2	0	3	1	2
CG10225	1	1	9	1	0
CG10249	4	4	0	4	0
CG10274	6	6	0	4	0
CG10341	4	4	0	1	0
CG10365	6	6	0	1	0

	Total S2 reads			number of unique 21-mers	
Gene	**sense + antisense**	**antisense**	**sense**	**antisense**	**sense**
CG10376	6	5	1	2	1
CG10435	4	2	2	1	1
CG10445	1	0	1	0	1
CG10462	4	0	4	0	1
CG10516	1	0	1	0	1
CG10576	2	2	1	1	0
CG10669	2	0	2	0	1
CG10889	3	0	3	0	1
CG10903	3	0	3	0	1
CG10971	5	1	4	1	1
CG1104	1	1	0	1	0
CG11063	1	1	0	1	0
CG11109	13	1	12	1	4
CG11180	1	0	1	0	1
CG11198	3	1	2	1	1
CG11242	9	9	0	1	0
CG11306	14	0	14	0	2
CG11377	3	0	3	0	1
CG11388	5	3	2	1	1
CG11448	2	2	0	1	0
CG11455	1	1	0	1	0
CG11526	11	0	11	0	2
CG11620	6	6	0	1	0
CG11777	1	0	1	0	1
CG11790	6	6	0	1	0
CG11814	4	4	0	1	0
CG11866	3	3	0	1	0
CG11872	2	2	0	1	0
CG11880	3	3	0	2	0
CG11927	3	3	0	1	0
CG11929	1	1	0	1	0
CG11943	2	0	2	0	1
CG12016	13	9	4	1	1
CG12030	17	0	17	0	2
CG12082	1	1	0	1	0
CG12106	13	11	2	2	1
CG12118	13	11	2	1	2
CG12170	3	0	3	0	1
CG12182	13	6	7	1	2
CG12262	7	4	3	1	1
CG12299	3	0	3	0	1
CG12341	1	1	0	1	0
CG12343	2	0	2	0	1
CG12393	6	6	0	1	0
CG12576	1	0	1	0	1
CG12785	5	5	0	1	0
CG12936	1	0	1	0	1
CG1311	3	3	0	1	0
CG13189	2	0	2	0	1
CG13220	8	0	8	0	1
CG13349	1	0	1	0	1

CG13384	3	1	2	1	1
CG13484	3	3	0	2	0
CG1358	20	6	14	2	3
CG13601	1	1	0	1	0
CG13762	4	4	0	1	0
CG13893	1	0	1	0	1
CG13900	22	7	15	1	3
CG13902	4	2	2	1	1
CG13924	2	2	0	1	0
CG14102	1	0	1	0	1
CG14211	6	0	6	0	1
CG14215	2	2	0	1	0
CG14230	4	4	0	1	0
CG1434	1	0	1	0	1
CG14435	1	0	1	0	1
CG14476	1	0	1	0	1
CG14670	3	3	0	1	0
CG14782	11	0	11	0	2
CG14786	6	0	6	0	1
CG14799	1	1	0	1	0
CG14804	3	0	3	0	1
CG14815	4	0	4	0	1
CG14882	3	3	0	1	0
CG14956	4	0	4	0	2
CG14966	6	0	6	0	1
CG14967	7	7	0	1	0
CG15011	2	0	2	0	1
CG15067	2	0	2	0	1
CG15097	1	0	1	0	1
CG15099	1	1	0	1	0
CG1516	5	0	5	0	1
CG15209	5	5	0	1	0
CG15216	2	0	2	0	1
CG1531	4	4	0	1	0
CG15370	1	1	0	1	0
CG1542	3	0	3	0	1
CG15438	2	0	2	0	1
CG15482	1	1	0	1	0
CG1553	5	5	0	2	0
CG15609	21	13	8	5	2
CG15891	1	1	0	1	0
CG15892	1	1	0	1	0
CG15896	7	7	0	2	0
CG15930	1	0	1	0	1
CG1600	3	0	3	0	1
CG1621	4	0	4	0	1
CG16742	1	0	1	0	1
CG16903	3	1	2	1	1
CG16972	6	6	0	1	0
CG16989	2	0	2	0	1
CG17264	13	13	0	1	0
CG17660	6	0	6	0	2
CG17715	1	0	1	0	1
CG17746	6	1	5	1	3
CG18107	2	0	2	0	1
CG1812	2	0	2	0	1
CG1814	4	0	4	0	1
CG18166	1	1	0	1	0
CG18259	9	6	3	1	2
CG18262	2	2	0	1	0
CG18273	1	1	0	1	0
CG18432	7	0	7	0	1
CG18542	8	0	8	0	1
CG18809	7	0	7	0	1
CG18854	83	3	80	2	25
CG1902	2	2	0	1	0
CG1972	7	7	0	1	0
CG2006	3	0	3	0	2
CG2034	2	0	2	0	1
CG2137	1	0	1	0	1
CG2225	2	0	2	0	1
CG2247	1	0	1	0	1
CG2614	6	6	0	1	0
CG2698	1	0	1	0	1
CG2811	2	0	2	0	1
CG2926	8	8	0	1	0
CG2991	2	0	2	0	1
CG3032	2	0	2	0	1
CG30373	9	9	0	1	0
CG30410	4	4	0	1	0
CG31082	6	6	0	1	0
CG31121	6	0	6	0	2
CG31158	1	1	0	1	0
CG3164	2	1	1	1	1
CG31643	1	0	1	0	1
CG3165	3	3	0	1	0
CG31678	3	2	1	1	1
CG31729	4	4	0	1	0
CG3173	2	0	2	0	1
CG31771	1	0	1	0	1
CG31793	5	0	5	0	1
CG31812	2	0	2	0	1
CG31849	4	0	4	0	1
CG31918	1	0	1	0	1
CG31919	8	8	0	1	0
CG31922	3	3	0	1	0
CG31975	4	0	4	0	1
CG32164	3	1	2	1	2
CG32165	3	1	2	1	2
CG3223	2	0	2	0	1
CG32250	2	0	2	0	1
CG32409	1	0	1	0	1
CG32412	1	1	0	1	0
CG32425	8	8	0	1	0
CG32495	2	2	0	1	0
CG32694	1	0	1	0	1
CG32702	6	6	0	2	0
CG3271	2	2	0	1	0
CG3279	10	0	10	0	3
CG32809	5	5	0	1	0
CG32939	8	0	8	0	1
CG3308	7	6	1	3	1
CG33107	1	1	0	1	0
CG33111	1	0	1	0	1
CG33249	4	0	4	0	1
CG33469	1	1	0	1	0
CG33470	14	0	14	0	1
CG33509	15	4	11	1	3
CG33510	3	0	3	0	2
CG33523	2	2	0	1	0
CG3356	2	0	2	0	1
CG3363	3	3	0	1	0
CG33649	1	0	1	0	1
CG33932	3	3	0	1	0
CG33967	3	0	3	0	1
CG33969	19	13	6	3	5
CG33978	1	0	1	0	1
CG33995	8	8	0	1	0
CG3402	14	10	4	1	2
CG34125	3	3	0	2	0
CG34126	6	6	0	1	0
CG34179	6	0	6	0	1
CG34268	3	0	3	0	1
CG34335	3	0	3	0	1
CG34376	4	0	4	0	1
CG34398	1	1	0	1	0
CG34415	5	0	5	0	1
CG34429	2	1	1	1	1
CG34430	1	0	1	1	0
CG3542	3	2	1	1	1
CG3605	3	0	3	0	1
CG3683	3	3	0	1	0
CG3703	6	5	1	2	1
CG3711	1	0	1	0	1
CG3740	6	0	6	0	1
CG3760	2	2	0	1	0
CG3764	5	5	0	1	0
CG3792	3	3	0	1	0
CG3814	2	2	0	1	0
CG3831	1	0	1	0	1
CG3967	1	1	0	1	0
CG3973	3	1	2	1	1
CG3980	1	0	1	0	1
CG40084	1	0	1	0	1
CG40228	1	0	1	0	1
CG4025	2	2	0	1	1
CG40351	6	4	2	2	1
CG4061	4	0	4	0	2

CG4068	95	0	95	0	7
CG40798	2	0	2	0	1
CG4119	1	1	0	1	0
CG41322	2	0	2	0	1
CG41421	8	0	8	0	1
CG41484	19	2	17	2	3
CG41533	2	2	0	1	0
CG41584	1	0	1	0	1
CG41587	2	0	2	0	1
CG41589	1	0	1	0	1
CG4199	13	12	1	1	2
CG4213	2	0	2	0	1
CG4334	2	0	2	0	1
CG4582	2	2	0	1	0
CG4619	1	0	1	0	1
CG4643	2	0	2	0	2
CG4670	5	2	3	1	1
CG4699	2	2	0	1	0
CG4752	4	0	4	0	1
CG4822	10	0	10	0	2
CG4901	2	2	0	1	0
CG4963	5	0	5	0	1
CG5044	6	0	6	0	1
CG5062	6	6	0	1	0
CG5104	1	1	0	1	0
CG5126	1	0	1	0	1
CG5130	11	11	0	1	0
CG5191	4	0	4	0	1
CG5362	5	5	0	1	0
CG5458	13	4	9	1	2
CG5508	3	1	2	1	1
CG5510	1	0	1	0	1
CG5537	1	1	0	1	0
CG5543	2	0	2	0	1
CG5567	3	0	3	0	1
CG5644	2	0	2	0	1
CG5734	12	8	4	2	2
CG5840	1	1	0	1	0
CG5853	1	1	0	1	0
CG5857	3	2	1	1	1
CG5871	14	8	6	2	2
CG5899	1	1	0	1	0
CG5919	10	10	1	2	0
CG5938	2	2	0	1	0
CG5986	1	0	1	0	1
CG6038	7	0	7	0	1
CG6171	1	1	0	1	0
CG6181	5	3	2	1	1
CG6218	7	0	7	0	1
CG6424	15	2	13	2	4
CG6448	3	3	0	2	0
CG6454	1	1	0	1	0
CG6509	2	2	0	1	0
CG6654	2	2	0	1	0
CG6689	5	0	5	0	1
CG6805	1	0	1	0	1
CG6833	2	2	0	1	0
CG6876	3	0	3	0	1
CG6891	2	2	0	1	0
CG6900	2	2	0	1	0
CG6903	3	0	2	0	1
CG6907	10	8	2	2	1
CG6912	1	1	0	1	0
CG6950	2	0	2	0	1
CG6961	7	6	1	1	1
CG7011	3	3	0	1	0
CG7139	10	8	2	1	1
CG7144	1	1	0	1	0
CG7177	1	1	0	1	0
CG7224	4	0	4	0	2
CG7289	1	0	1	0	1
CG7324	3	1	2	1	1
CG7338	4	3	1	1	1
CG7376	5	5	0	2	0
CG7379	3	0	3	0	1
CG7504	2	1	0	1	0
CG7518	1	0	1	0	1
CG7519	3	3	0	1	0
CG7632	1	1	0	1	0

CG7650	7	7	0	3	0
CG7739	37	17	20	3	6
CG7789	5	0	5	0	1
CG7816	3	0	3	0	2
CG7830	2	0	2	0	1
CG7912	3	0	3	0	1
CG7988	1	1	0	1	0
CG8112	4	0	4	0	2
CG8155	2	2	0	1	0
CG8199	5	5	0	1	0
CG8289	4	0	4	0	1
CG8297	4	0	4	0	1
CG8315	4	0	4	0	1
CG8319	1	0	1	0	1
CG8320	5	0	5	0	1
CG8336	5	3	2	2	1
CG8443	2	2	0	1	0
CG8451	5	5	0	1	0
CG8478	3	1	2	1	1
CG8481	1	1	0	1	0
CG8516	5	5	0	1	0
CG8526	10	0	10	0	1
CG8538	5	0	5	0	1
CG8545	1	0	1	0	1
CG8594	8	8	0	3	0
CG8602	6	6	0	2	0
CG8668	2	2	0	1	0
CG8798	10	7	3	2	1
CG8862	5	2	3	1	2
CG8878	3	0	3	0	2
CG8950	6	3	3	1	1
CG9007	5	5	0	1	0
CG9143	1	1	0	1	0
CG9246	1	0	1	0	1
CG9320	4	0	4	0	1
CG9339	3	0	3	0	1
CG9346	2	0	2	0	1
CG9372	3	0	3	0	1
CG9389	3	0	3	0	1
CG9578	3	3	0	1	0
CG9629	3	3	0	1	0
CG9674	1	0	1	0	1
CG9776	1	1	0	1	0
CG9779	1	0	1	0	1
CG9780	3	1	2	1	1
CG9795	2	0	2	0	1
CG9799	4	4	0	1	0
CG9804	2	1	1	1	1
CG9922	5	5	0	1	0
CG9934	4	0	4	0	1
CG9941	1	0	1	0	1
CG9945	1	1	0	1	0
cher	6	2	4	1	1
chinmo	6	3	3	3	1
CHKov1	3	0	3	0	1
Chro	4	2	2	1	2
Cht3	1	0	1	0	1
cnc	3	3	0	1	0
Cog3	1	1	0	1	0
colt	5	5	0	1	0
Coq2	4	0	4	0	1
Cp190	4	4	0	2	0
CPTI	1	1	0	1	0
crb	4	4	0	1	0
crq	6	0	6	0	1
Csk	11	0	11	0	1
CstF-64	10	0	10	0	1
cue	4	3	1	1	1
CycG	4	0	4	0	1
Cyp28d1	4	4	0	1	0
da	1	0	1	0	1
Dcr-1	5	1	4	1	1
Ddx1	2	2	0	1	0
Dg	2	0	2	0	1
Dgp-1	1	0	1	0	1
Dif	1	1	0	1	0
DNApol-gamma35	1	0	1	0	1
DNApol-iota	1	0	1	1	0

Doa	7	7	0	1	0
Dph5	1	1	0	1	0
drosha	2	0	2	0	1
Drp1	7	7	0	1	0
Dyrk3	2	0	2	0	2
e(y)1	1	1	0	1	0
E2f	3	3	0	1	0
Eap	6	0	6	0	1
ed	3	0	3	0	1
Edem2	7	0	7	0	1
edl	1	1	0	1	0
EDTP	3	3	0	1	0
eff	1	1	0	1	0
egh	1	0	1	0	1
eIF2B-epsilon	5	0	5	0	1
eIF3-S10	6	0	6	0	1
eIF4G	2	2	0	1	0
eIF-5A	1	0	1	0	1
eIF5B	1	0	1	0	1
Elongin-B	2	2	0	1	0
ETH	3	0	3	0	1
ex	1	1	0	1	0
Fak56D	4	4	0	1	0
fbl	17	4	13	4	3
Fem-1	6	6	0	1	0
Fit1	1	1	0	1	0
FKBP59	1	0	1	0	1
fog	1	0	1	0	1
foi	1	1	0	1	0
for	3	0	3	0	1
form3	2	2	0	1	0
frc	1	1		0	1
frtz	7	7	0	1	0
fry	7	7	0	2	0
Fur2	1	1	0	1	0
fzr2	5	0	5	0	1
fzy	5	5	0	1	0
gammaSnap	3	0	3	0	1
gatA	1	1	0	1	0
GckIII	5	0	5	0	1
Gclc	1	1	0	1	0
Gclm	2	2	0	1	0
Gcn2	2	0	2	0	1
Gdh	4	4	0	2	0
gft	1	0	1	0	1
Gmd	1	1	0	1	0
Gp93	3	0	3	0	1
gry	9	2	7	2	2
GS	2	2	0	1	0
gwl	2	2	0	1	0
Gycalpha99B	2	0	2	0	1
hdc	20	4	16	2	5
Hexo2	1	0	1	0	1
His1 (11 loci)	1	0	1	0	1
His2A (19 loci)	10	7	3	2	2
His2A:CG31618	10	7	3	2	2
His2Av	1	1	0	1	0
His2B (22 loci)	5	5	0	3	0
His2B:CG17949	5	5	0	3	0
His2B:CG40461	4	4	0	2	0
His3 (23 loci)	10	9	1	4	1
Hmgcr	1	0	1	0	1
hoip	2	0	2	0	1
Hr4	1	1	0	1	0
Hr96	1	0	1	0	1
Hs3st-A	1	0	1	0	1
Hsp70Aa	17	16	1	5	1
Hsp70Ab	2	2	0	1	0
Hsp70Ba	3	3	0	2	0
Hsp70Bb	3	3	0	2	0
Hsp70Bbb	3	3	0	2	0
Hsp70Bc	3	3	0	2	0
htt	1	0	1	0	1
hyd	1	1	0	1	0
Iap2	3	3	0	1	0
Ice	6	6	0	1	0
icln	3	2	1	1	1
igl	1	0	1	0	1
Ilp6	3	0	3	0	1

IM10	14	0	14	0	1
IP3K1	15	12	3	4	2
Irbp	6	0	6	0	1
itp	1	1	0	1	0
Itp-r83A	4	0	4	0	2
jet	12	5	7	1	1
Jon99Fi	5	0	5	0	1
Khc	2	0	2	0	1
kis	4	3	1	1	1
Klp3A	9	2	7	1	1
ksr	4	4	0	1	0
kuz	2	2	0	1	0
l(1)G0004	5	1	4	1	1
l(2)37Cb	4	4	0	1	0
l(2)NC136	1	1	0	1	0
l(2)tid	1	0	1	0	1
l(3)01239	3	0	3	0	1
l(3)s1921	6	4	2	3	1
Lac	1	1	0	1	0
lack	8	8	0	2	0
larp	7	7	0	1	0
ldlCp	13	11	2	3	1
lid	5	5	0	1	0
lig	2	2	0	1	0
lkb1	1	1	0	1	0
lok	5	5	0	3	0
Lsd-1	4	3	1	2	1
M(2)21AB	18	12	6	3	1
mam	6	6	0	1	0
Map60	5	5	0	1	0
mask	2	2	0	1	0
MBD-like	12	0	12	0	1
Mcm2	1	1	0	1	0
MCPH1	6	5	1	1	1
Med	2	0	1	0	1
MED15	5	5	0	1	0
MED21	2	0	2	0	1
MED24	11	11	2	1	0
Mes-4	1	0	1	0	1
Mi-2	5	0	5	0	1
mib1	1	1	0	1	0
milt	3	3	0	1	0
mip130	3	0	3	0	1
mit(1)15	5	5	0	1	0
Mitf	8	0	8	0	3
Mmp1	19	5	14	3	2
mod(mdg4)	5	3	2	2	1
MP1	4	1	3	1	1
mRpL18	1	0	1	0	1
mRpL44	1	1	0	1	0
mRpL48	6	1	5	1	1
mRpS2	3	0	3	0	1
mrt	2	0	2	0	1
msl-1	4	4	0	1	0
Mst89B	6	6	0	1	0
mt:Col	3	2	1	1	1
MTF-1	4	0	4	0	1
mtTFB1	2	2	0	1	0
mus205	4	2	2	1	1
mus308	2	2	0	1	0
mus309	1	1	0	1	0
mus81	7	0	7	0	3
Mys45A	3	3	0	1	0
nAcRbeta-21C	6	0	6	0	1
ncd	1	1	0	1	0
NitFhit	3	0	3	0	2
nito	1	1	0	1	0
Nle	1	0	1	0	1
Nup154	1	1	0	1	0
Nup44A	2	2	0	1	0
Nup98	7	0	7	0	1
Obp99c	3	0	3	0	1
O-fut1	5	4	1	1	1
omd	4	4	0	1	0
opa1-like	2	0	2	0	1
Orc1	1	0	1	0	1
osa	9	5	4	2	2
P58IPK	1	1	0	1	0
pAbp	1	1	0	1	0

par-1	4	1	2	1	1
Pc	1	0	1	0	1
Pcaf	2	2	0	1	0
Pdk	4	4	0	1	0
Pect	11	4	7	1	3
Pen	2	0	2	0	1
pie	2	2	0	1	0
PIP82	1	1	0	1	0
pita	2	1	1	1	1
Pitslre	2	2	0	1	0
Pka	8	0	8	0	1
Pms2	1	1	0	1	0
pnt	1	1	0	1	0
Pof	3	0	3	0	2
por	2	0	2	0	2
Pp1-87B	4	4	0	1	0
Pp2C1	1	1	0	1	0
ppk13	11	7	4	4	3
Ptp99A	10	3	7	1	2
Pvf2	4	1	3	1	1
pyd	2	0	2	0	2
qkr58E-1	1	1	0	1	0
Rab11	1	1	0	1	0
Rab6	11	11	0	1	0
RabX6	8	8	0	2	0
Rack1	1	0	1	0	1
Rbf	1	0	1	0	1
Rbm13	1	0	1	0	1
ref(2)P	3	3	0	1	0
Rfabg	2	0	2	0	1
RfC38	1	1	0	1	0
RhoGAP16F	28	13	15	3	2
RhoGAP68F	5	4	1	1	1
Ric	2	0	2	0	1
r-l	20	12	8	3	2

Rlip	2	2	0	1	0
RpL21	2	0	2	0	1
RpL28	2	0	2	0	1
RpLP2	2	1	1	1	1
Rpn2	4	0	4	0	1
Rpp20	3	3	0	1	0
RpS14a	4	4	0	1	0
RpS14b	4	4	0	1	0
RpS7	6	6	0	1	0
Rpt4	2	2	0	1	0
Rrp4	2	2	0	1	0
Rrp42	6	6	0	1	0
Sas10	9	0	9	0	1
sas-6	5	5	0	1	0
sav	4	1	3	1	1
sax	2	2	0	1	0
sbb	2	0	2	0	1
ScpX	2	0	2	0	1
scra	7	2	4	1	2
scu	1	1	0	1	0
sec15	8	0	8	0	2
sec23	1	1	0	1	0
sec31	5	0	5	0	1
sec63	3	0	3	0	1
sec71	2	0	2	0	1
sens	1	0	1	0	1
Sin3A	21	9	12	3	2
skd	6	6	0	1	0
sle	2	2	0	1	0
slik	1	1	0	1	0
slmo	6	0	6	0	1
Snoo	13	5	8	3	3

Table II-S1C. mRNA-matching endo-siRNAs in wild-type fly heads. Data comprise pyrosequencing and sequencing-by-synthesis of small RNA libraries enriched for 3´ terminally modified RNA.

GENE	siRNA	orientation of small RNA	wild-type reads
5-HT1B	TTGCTGGTGCTGAACTCGGTC	antisense	1
Ack	TCGAATGAAGACCTCGGTTCG	sense	1
Act5C	TGCTTGGAGATCCACATCTGC	antisense	1
Act79B	TGCTTGGAGATCCACATCTGC	antisense	1
Act87E	TGCTTGGAGATCCACATCTGC	antisense	1
ade3	AACATTCAGCGATAATTTGAC	sense	1
ade3	CCAGAAAGGCGACCGGCTCGG	antisense	1
AGO2	TGGAAAGCTTATAATGGAGTT	sense	2
AGO2	AATCAATAGAGATGCTCCTTT	antisense	1
AGO2	CCGCCTGGACAACAACAAGGT	sense	1
AGO2	TAAAATCATAGCTCATCATGG	antisense	1
AGO2	TCTATTAACTCCATTATAAGC	antisense	1
AGO2	TTAAAAGCCGCCTTGAGATGG	sense	1
AGO2	TTGGAAAGCTTATAATGGAGT	sense	1
alc	GTTGTATTTTTATTAGTTATA	sense	1
Ald	AGGACGGTTGCGACTTCGCCA	sense	1
alpha4GT1	TTTCAATAAATGTCTATGTGA	antisense	1
alphaTry	TACTGGAGCTCTGGTGGCGTC	sense	1
alphaTub84B	ACCGGTCTGCAGGGCTTCCTC	sense	1
alphaTub84D	ACCGGTCTGCAGGGCTTCCTC	sense	1
Amy-d	ATCTGATTGATCTCGGCGTGG	sense	1
Amy-d	CACCGCTGGCGCCAGATCTAC	sense	1
Amy-p	ATCTGATTGATCTCGGCGTGG	sense	1
Amy-p	CACCGCTGGCGCCAGATCTAC	sense	1
Ank	CAAGTGTTACTGGAAAATGGT	sense	1
Ank	GTAATATCATAGACGCAGGAC	sense	1
AnnIX	TAAGGATTTCCTCGTTGGATC	sense	1
Apc	GCGGGCTCCAAGGTTGTGGGC	antisense	1

GENE	siRNA	orientation of small RNA	wild-type reads
Apc	TCCTCGACTACGACGCAGTGC	sense	1
Apc	TGATGATGGTGATGATGATGA	antisense	1
Apc2	ATGCCGCTCAGAATGCACTGC	antisense	1
apt	TGATGACGATGATGATGATGA	sense	2
apt	TGATGATGATGACGATGATGA	sense	1
arm	TGCAATAGAACTCGTACATTT	sense	1
Asator	GAAAATTCTGATGATAACGGC	sense	1
asrij	CTTGGAGCCGTCTAAGCCGGG	sense	1
ATPsyn-beta	TGGCTACCGACATGGGTTCTA	sense	1
aux	AGTAGAATGTCCTCGGGCGTT	antisense	1
aux	CCGCAGACTCTGCCCTCGACA	sense	1
aux	TCGACATCGTCTATCAGGACG	sense	1
bel	CGAACGCACCCGACTGGTGGG	sense	1
bel	TGCTGCTGCTGCTGCTGGCCC	antisense	1
betaTry	TACTGGAGCTCTGGTGGCGTC	sense	1
bigmax	CATTTAATGTTTGTACGCAGT	antisense	1
bin3	AATCCTGCTCGCCGGCCGTGT	antisense	1
bin3	TCCTTCACCGCGATATTCGCT	antisense	1
bin3	TGATGATGATGGTGATGATGA	antisense	1
blw	ACTCCGGATGCGCCATGGGAG	sense	1
blw	TAACGTCGGTCTGTCCGTGTC	sense	1
Bruce	CAAACGTATCGAACTGGCGCT	antisense	1
BRWD3	ATCAGTCCTACATTCTATGTG	sense	1
c11.1	TCCAAGTGTGATCCCTTTGCC	sense	1
cal1	GTGTGTCGAAGTTGTGCTTTA	sense	1
CalpB	CGACGATTTCCTGATGTGTGC	sense	1
CalpB	CGATATTGCCAAGTGGCGGGC	sense	1
CalpB	TTCAAGTACCGAGCATCGGTG	sense	1

Ca-P60A	TTATGGTCTAGTACATTGCCA	sense	1
Cap-H2	TTTACCAGAGTGCAAGCTAGT	sense	1
Cas	ATCCTGCTGGTTGGGCTGGGC	antisense	1
cathD	ATCGTGTGGGCTTCGCCGATG	sense	1
Cbl	CACAGCGACTTGTTCTGAGGG	antisense	1
Ccn	ATCATCATCATCATCATCATC	antisense	55
Ccn	TGATGATGATGATGATGATGG	sense	49
Ccn	ATGATGATGATGATGATGATG	sense	28
Ccn	CATCATCATCATCATCATCAT	antisense	6
Ccn	TCATCATCATCATCATCATCT	antisense	4
Ccn	CCATCATCATCATCATCATCA	antisense	2
Ccn	GATGATGATGATGATGATGAT	sense	1
Ccn	GATGATGATGATGATGATGGG	sense	1
Ccn	TCATCATCATCATCATCTTCA	antisense	1
Ccn	TGATGATGATGATGATGGGGA	sense	1
Ccp84Aa	CTCCACCTGTCCCTTGTTGTC	antisense	1
Ccp84Ab	CTCCACCTGTCCCTTGTTGTC	antisense	1
Cct5	TATTTGAACAGGGATAGCCCC	antisense	1
Cdep	CTCTGGCAGTTAATCGAATGC	antisense	1
ced-6	CGGAAATGGCGATGGTAGTGC	sense	1
ced-6	CTCGTTCTTGCCGATTGTCCC	sense	1
Cf2	TTGCCACGCTGGACGGCGGTC	sense	1
CG10011	TATGTATGTATATGCTTCGTT	antisense	1
CG10055	TCCTCATCCTCATCCTCATCC	antisense	1
CG10077	ACTTGCCCAAGATCGTTTGGT	sense	1
CG10147	TCTCGAAACTCGTCGGCGCAC	antisense	1
CG1021	CCAGGGTCTGCGCAACGTGGG	sense	1
CG10214	TCATATTGCCAATAAAGCATT	antisense	1
CG10237	TGGGTCGTCGCATTTTGGTCC	sense	1
CG10249	AGATGAACCTTTTGATTTGAT	antisense	1
CG10274	AATAGATGAATTGATTGTGGC	sense	1
CG10375	TACAAGCACGCCATCTACGTT	sense	1
CG10433	TGCGTCATGGTCGGAGGACTG	sense	1
CG10444	TACACTACAGTGAGCACCGCC	sense	2
CG10479	TTTCACCTATCGTTCCCTTTG	antisense	1
CG10631	TTCGTGTCGTCGTGGTTCAGC	antisense	1
CG10641	CTGGAACTGCTGACGCCGCTG	antisense	1
CG10646	TTTAATATTTGTATAACCTGC	antisense	2
CG10673	GAGAAGCCTGCCTGATCAAGG	sense	1
CG10681	TTTAATATTTGTATAACCTGC	sense	2
CG10713	CTGCTGCTGATGCTGCTGATG	antisense	1
CG10874	TCGATGCCAAACGCCAGTGCC	sense	1
CG10918	GGTGGAGCTGGAGGAGCTGCT	antisense	1
CG10971	GTATTCCGTGTCGATCGCTTT	sense	1
CG11006	GATGATGATGATGATGATGGA	antisense	2
CG11050	TGCTAACCTTTCTGCTCGACG	sense	1
CG11077	TGGAACTGCTACCTCTCCATG	sense	1
CG11122	TCACCTGCTTCGGGATCGGGC	sense	1
CG11122	TCCTCCTCCTCCTCATCCTCC	antisense	1
CG11146	ATGATGATGATGATGATGGTG	antisense	1
CG11146	TGATGATGATGATGGTGATGG	antisense	1
CG1115	CACACATCTCTGCCTGCACTA	sense	1
CG1115	TACTTTCACATATACATATAT	antisense	1
CG1115	TATATCATAGTTTAGTGCAGG	antisense	1
CG11151	ACAACAACAACCATGTCTCTG	sense	2
CG11180	CTAACTAATTAAACTGAACTA	sense	1
CG11188	TCATCATCCTCCTCCTCATCA	antisense	1
CG11198	ATCCTGCACGACTTCTTCTAC	sense	1
CG11198	CCACCACCTCCTTCCGAGGCC	antisense	1
CG11198	CGTTCGATACGATCGTTGGGC	antisense	1
CG11198	GGGGTACCACACCTGACCTGC	antisense	1
CG11198	TCATGCACGCCTTGCGATATC	antisense	1
CG11242	AGTACGAACAGCGAACAGATT	sense	1
CG11284	AAACTGAATTTATTAAACATC	sense	1
CG11490	CTCTGGCCGCAGGCTCTGCAC	antisense	1
CG11498	CTTTTATCAGATCCGATCGCC	antisense	1
CG11498	TCTCGGCGGTCTCGAAGAGGT	antisense	1
CG11501	CGGAATGCTGCACCTCCCGGG	sense	1
CG11526	CACACCAAATGCCTCGTTGGG	sense	1
CG11526	CGTATGTGTGTTTTGTTTCGC	sense	1
CG11526	TTATTCGTATGTGTGTTTTGT	sense	1
CG11534	GATGATGATGATGATGATGGA	sense	2
CG11534	ATCATCATCATCATCATCGTC	antisense	1
CG11710	CTGTACGCCAGCGATGGTCGC	sense	1
CG11771	TACGATGTGTTCGATGCTGAT	sense	1
CG11848	TCACAGTCACCTTCATAGCGT	sense	1
CG11943	CTGCTGCTGTGCATTGTGGAG	sense	1
CG11943	TTACAATCGCGCGTACAAGGT	sense	1
CG11963	GTAGGAAACGTGTTCCTGGAC	antisense	1

CG11963	TGGCTCCCATCGCCGTCCAGC	sense	1
CG11967	TTTTTCGTATTTATAAGTGTG	sense	1
CG11968	TTTTTCGTATTTATAAGTGTG	antisense	1
CG12016	AGTCTTTATTTCGTATTAAGA	sense	1
CG12016	ATTTCACAATCTGGTTACAAG	antisense	1
CG12016	CACACCAAATGCCTCGTTGGG	antisense	1
CG12016	CGTATGTGTGTTTTGTTTCGC	antisense	1
CG12016	CTGGTTACAAGAGCATTCTCC	antisense	1
CG12016	GTGAAATGAAGAAACTCGGTT	sense	1
CG12016	TGCTCTTGTAACCAGATTGTG	sense	1
CG12016	TTATTCGTATGTGTGTTTTGT	antisense	1
CG12017	TCCTCCTCCTCATCATCCTCC	antisense	1
CG12024	AGAAGGTGACCACGGCCAAGC	sense	1
CG12091	TGCCGCCGCCGGGACATGGGC	sense	1
CG12224	TCAACTATGCTCGTTACACCC	sense	1
CG12340	ATCCTGCACCACTGCCCACGT	sense	1
CG12367	TCATGAGAAACTTAACCAGCG	sense	1
CG12393	AACAACTCCACGCTCAGCGAA	antisense	1
CG1244	TCATCATCATCCTCATCATCC	antisense	2
CG1244	TCATCATCATCATCCTCATCA	antisense	1
CG12581	GTTTATGAATAAAGTGTGTGC	sense	1
CG12581	TTATGAATAAAGTGTGTGCGC	sense	1
CG12773	CAGGTTCACCTGTATGCGGGC	sense	1
CG13124	GCTCCGGGGCACCGGCCGAGG	sense	1
CG13130	TCATCATCATCATCATCCTCA	antisense	10
CG13130	ATCATCATCATCATCATCCTC	antisense	5
CG13130	GATGATGATGATGATGATGGG	sense	1
CG13130	TGAGGATGATGATGATGATGA	sense	1
CG13130	TGATGATGATGATGATGGGGA	sense	1
CG13253	TGATGATGATGCTGCTGCTGA	antisense	1
CG1332	GATGTGGCCGATCGACAATTC	sense	1
CG13445	TCATCATCATCATCATCCTCC	antisense	3
CG1358	TGATGATGATGATGATGATGA	antisense	552
CG1358	TCATCATCATCATCATCATCA	sense	197
CG1358	ATCATCATCATCATCATCATC	sense	55
CG1358	ATGATGATGATGATGATGATG	antisense	28
CG1358	CATCATCATCATCATCATCAT	sense	6
CG1358	TCATCATCATCATCATCATCG	sense	6
CG1358	ATCATCATCATCATCATCGTC	sense	1
CG1358	GATGATGATGATGATGATGAT	antisense	1
CG13585	CCTGCGACCAGACGCACTCGC	antisense	1
CG13670	TCCTCCTCCTCATCATCCTCC	antisense	1
CG13907	TCAGGAGTCTACCTCAGCGCC	sense	1
CG14033	CAAAAACATCGTCAATAATGG	antisense	1
CG14235	TATCGCCAGGCTCGTGAGATT	sense	1
CG14342	TGATGATGATGATGATGTTGA	antisense	7
CG14478	TTGCTTCCTCTGCTCGCTTGG	antisense	1
CG14480	TCAACTTTTTATTTGGATTCT	sense	1
CG14561	CATCGATGGCAGTCTGAGCGA	antisense	1
CG14567	CAGTGGCCCCCGTTTTCAACC	sense	1
CG14567	GCCACTGATGTTGCGGGTCTT	antisense	1
CG14646	CTGCATAGGAAACCTGTAGTG	sense	1
CG14799	TGATGATGATGATGATGATGG	antisense	49
CG14799	ATGATGATGATGATGATGATG	antisense	28
CG14799	CATCATCATCATCATCATCAT	sense	6
CG14799	CCATCATCATCATCATCATCA	sense	2
CG14799	TCATCATCATCATCATCATTG	sense	2
CG14799	ATGATGATGATGATGATGGTG	antisense	1
CG1486	CATGTGGCCTTCCAGACGTGC	sense	1
CG14880	AGCACCACCATCGCGCCCGGC	sense	1
CG14906	TGGGTGAGATCCGACTGCGGG	antisense	1
CG14907	TGGGTGAGATCCGACTGCGGG	antisense	1
CG14956	ATCCGCACCACATCTCACTGC	sense	1
CG14967	CTGACCAGGACTTAGCACTGC	sense	1
CG14982	CACCCACACCAGCATCACCCG	sense	1
CG15019	TTAATATGCAATAATAACTCG	antisense	1
CG15067	AGATTGTTTAAACTTATGTGG	sense	1
CG15067	GATAATTTCTTAGTGTTTCAA	antisense	1
CG15099	ACTGACTGGTGTACTTCATGT	antisense	1
CG15099	CATGAAGTACACCAGTCAGTC	sense	1
CG15105	CCACCCAGTTCCGCCAGCAGG	sense	1
CG15118	ACACTCTACCCCCACCCTCCT	sense	1
CG15134	CGCAACATCCGGTGGCCCTAC	sense	1
CG1516	TTGTCGTAGGGCTTTTCCCGC	sense	1
CG15203	CCCTGCTCTTCTCGATGCTCT	sense	1
CG15209	CCGTGGAGCGTACTACTAGCC	antisense	2
CG15209	ATGTCGACGAGCCATTGGGTC	sense	1
CG15240	TGCTGATGATGATGATGATGA	antisense	2
CG15240	ATCATCATCATCATCATCAGC	sense	1

CG15240	TGATGATGATGATGATGATTC	antisense	1
CG15322	CTGCTGCTGATGCTGCTGATG	antisense	1
CG15370	TGATGATGATGATGATGATGA	antisense	552
CG15370	TCATCATCATCATCATCATCA	sense	197
CG15370	ATCATCATCATCATCATCATC	sense	55
CG15370	ATGATGATGATGATGATGATG	antisense	28
CG15370	CATCATCATCATCATCATCAT	sense	6
CG15370	CTGATGATGATGATGATGATG	antisense	4
CG15370	GATGATGATGATGATGATGAT	antisense	1
CG15370	TGATGATGATGATGATGATTT	antisense	1
CG15418	CTGAAGTCGCTTTTAACGATG	sense	1
CG15465	TGATGATGATGATGATGATGC	antisense	35
CG15465	TCATCATCATCATCATCACCA	sense	10
CG15465	GTGATGATGATGATGATGATG	antisense	4
CG15465	ATGGTGATGATGATGATGATG	antisense	1
CG15465	TGGTGATGATGATGATGATGA	antisense	1
CG15482	TATCGGCGCACTGGCCTTAAT	sense	1
CG15529	CTTTTATCAGATCCGATCGCC	sense	1
CG15609	ACAAAATGGTCACCTCAACGC	sense	1
CG15675	TCAAACAATCCGCAAAGCAGA	antisense	1
CG15706	TTCCCTGGCAACCAAATCCTT	sense	1
CG15725	TGATGATGATGATGATGCTGC	antisense	9
CG15725	TGATGATGATGATGCTGCTGA	antisense	4
CG15725	ATGATGATGATGATGCTGCTG	antisense	2
CG15725	CTGATGATGATGATGATGCTG	antisense	1
CG15725	TGATGATGATGCTGCTGATGA	antisense	1
CG15725	TGATGCTGCTGATGATGATGA	antisense	1
CG15771	TGATGATGATGATGATGATGC	antisense	35
CG15771	TGATGATGATGATGATGCTGC	antisense	9
CG15771	ATGATGATGATGATGATGCTG	antisense	5
CG15771	CTGATGATGATGATGATGATG	antisense	4
CG15771	ATGATGATGATGATGCTGCTG	antisense	2
CG15771	TGATGATGATGATGCTGCTGC	antisense	1
CG1578	TGATGATGATAATGATGATGA	sense	1
CG15828	AGGTATCCAGTTTTACTGCTG	sense	1
CG15828	CATCAGATTTCCATCGTAGTG	antisense	1
CG15930	TGATGATGATGATGATGATGA	sense	552
CG15930	TCATCATCATCATCATCATCA	antisense	197
CG15930	ATCATCATCATCATCATCATC	antisense	55
CG15930	ATGATGATGATGATGATGATG	sense	28
CG15930	ATCATCATCATCATCATCATA	antisense	7
CG15930	TCATCATCATCATCATCATAA	antisense	7
CG15930	CATCATCATCATCATCATCAT	antisense	6
CG15930	TTATGATGATGATGATGATGA	sense	2
CG15930	GATGATGATGATGATGATGAT	sense	1
CG15930	TGATGATGATGATGATGATTT	sense	1
CG1599	ATATAAAACTCTACAGTACTC	sense	1
CG1628	AGCACAGCGTCTTGCTCCCGG	antisense	1
CG1637	CAATCGGGCCAGGGATTGGGC	antisense	1
CG1638	ACGGAGCCGAAGTCCAGGGAG	sense	1
CG1662	GGCTCCTCCAGTCGCGCTGCT	antisense	1
CG1665	ATATAAAACTCTACAGTACTC	antisense	1
CG16972	CTACTTCAGTCGATAGTAGCA	antisense	1
CG17065	CACGCACTCGTCCATCGGGGC	antisense	1
CG17108	CAGGGCCGTCACCGATTCCTC	antisense	1
CG17264	TTTCAATAAATGTCTATGTGA	sense	1
CG17528	CTCTTCACTTCAAGCATCCCC	sense	1
CG1753	TCCGCGGGCCTCTCGTATGGC	antisense	1
CG17838	TGATGATGATGACGATGATGA	antisense	1
CG18107	ACATTTTTATTATGATGCTAA	antisense	1
CG18107	CAATCGTCACTGTCTTTGTGC	sense	1
CG18107	CTTTGTGCTTGGTCTTCTGGC	sense	1
CG18107	GTGCTTGGTCTTCTGGCTTTG	sense	1
CG1812	TGCAATAAATTAGGCAGTGTC	antisense	1
CG18135	TCCACCTGCCGATTAAGTCGG	sense	1
CG18208	CTTGTGTGTGCTACGCGTTCT	antisense	1
CG18262	TCCTCCTTCTCCTCCTCCTCC	antisense	1
CG18787	CACTAGCGTATAATGTATATA	sense	1
CG18809	TATCGCCAGGCTCGTGAGATT	sense	1
CG1882	ATAGTTGGTGAGCTGGATGTC	antisense	1
CG1882	CGACCGAATCCGAGGATGTCC	antisense	1
CG18854	CGGGAAACTATGGATCAAATG	sense	37
CG18854	ATCACTATCATCATCATCCGA	sense	29
CG18854	GGTCGACCACGTACCTTGGCC	sense	10
CG18854	TGGCAAAAAATCCTTGTAGTG	sense	10
CG18854	CATCGCAAGCCAGATTCTTGC	sense	8
CG18854	AATCATCCTTTGGCCCATAGT	sense	3
CG18854	GATGATTCCCGGGATTCAAGC	sense	3
CG18854	ACTATCATCATCATCCGAATC	sense	2

CG18854	AGGGTGGCCAAGATATGTGGT	sense	2
CG18854	ATCATCCTTTGGCCCATAGTT	sense	2
CG18854	CTTGCTTGGCTCTCAGGAATC	sense	2
CG18854	TATCATCATCATCCGAATCCT	sense	2
CG18854	TTATTGGTGGTCAATATGTCG	sense	2
CG18854	TTCCATCTGATCTTGAACATT	sense	2
CG18854	ATAAGATTCTTGAAGCCAGGA	sense	1
CG18854	ATCTGATCTTGAACATTTCGC	sense	1
CG18854	ATGATTCCCGGGATTCAAGCA	sense	1
CG18854	ATGCTAATGACTCCGATGTGG	sense	1
CG18854	ATGCTGCTGAAATGGATTCGG	sense	1
CG18854	ATTGAATAAGATTCTTGAAGC	sense	1
CG18854	CAAGATATGTGGTCGACCGAC	sense	1
CG18854	CCACATCGACTGGAATAGTGC	antisense	1
CG18854	CGGCTGCCCATCTTGATGTCC	sense	1
CG18854	CTAATGACTCCGATGTGGACC	sense	1
CG18854	CTATCATCATCATCCGAATCC	sense	1
CG18854	GAAACTATGGATCAAATGATG	sense	1
CG18854	GATCTTGAACATTTCGCCCTC	sense	1
CG18854	GCCTTGACGATCTTAGTCAAT	sense	1
CG18854	GGACCATCGAAGTGCTTGGGC	antisense	1
CG18854	TAATCAAAAAATAACTCAGCA	sense	1
CG18854	TAGTGCATCGCAAGCCAGATT	sense	1
CG18854	TCATCATCCGAATCCTCTACA	sense	1
CG18854	TCATCCGAATCCTCTACAACG	sense	1
CG18854	TCGATTAGTGCATCGCAAGCC	sense	1
CG18854	TCTACAACGATTTTTTCCCCA	sense	1
CG18854	TCTGATCTTGAACATTTCGCC	sense	1
CG18854	TTAATCAAAAAATAACTCAGC	sense	1
CG18854	TTAGTCAATTCGCGCAGCTCC	sense	1
CG18854	TTGAACATTTCGCCCTCCTTG	sense	1
CG18854	TTTCCATCTGATCTTGAACAT	sense	1
CG18854	TTTGGCCCATAGTTTTCCATC	sense	1
CG18870	TTGGCTTAAGACCTACTGACC	antisense	1
CG1893	CAAAATGCCTTGAAAGCTGGC	antisense	1
CG1998	ATCATCATCATCCTCCTCCTC	sense	2
CG1998	TCATCATCATCATCCTCCTCC	sense	2
CG1998	TCATCATCATCCTCCTCCTCC	sense	2
CG1998	TCCTCATCATCATCATCCTCC	sense	2
CG2061	TGGAGCCCGTATCCGATCTCT	sense	1
CG2083	TCCACCAGCCCTGGGAACCGC	sense	1
CG2093	CAGAAATACCCTGTCGGTGTT	sense	1
CG2124	AGCACAGCGTCTTGCTCCCGG	sense	1
CG2165	GTGGCCGTACCTGAGGGGCTT	sense	1
CG2182	TGGGAACGCTCTAGAATCGGC	antisense	1
CG2186	TGATGATGATGATGATGCTGC	antisense	9
CG2186	CTGATGATGATGATGATGCTG	antisense	1
CG2186	TGATGCTGCTGATGATGATGA	antisense	1
CG2186	TGCTGATGATGATGATGATGC	antisense	1
CG2211	TCTTTGCTCGGTCGTAGTATC	antisense	1
CG2233	AAATTACCAGACCATTGACCT	sense	1
CG2233	CATCGGTTCTCCGGCTCCGGT	sense	1
CG2519	TGGAGTGACTATGCTAGTGGC	antisense	1
Cg25C	TTCGAGACCCGGCCCTATGGG	sense	1
CG2604	GCCAGTGGCTGGCTTAAGAGC	sense	1
CG2807	TACTGGAGTTGCTCAAGGCCC	sense	1
CG2989	TCATCATCATCATCATCTTCA	antisense	1
CG30035	CTGTAAGTGTATCTATATGTA	sense	1
CG3011	ATCTGGCTGTCTACACGGGCG	sense	1
CG3011	ATCTTCTTCGAGAGCATGCCG	sense	1
CG31116	TCGCACCGACTTCGGCTTTTG	antisense	1
CG31121	TGATGATGATGATGATGATGA	sense	552
CG31121	TCATCATCATCATCATCATCA	antisense	197
CG31121	ATCATCATCATCATCATCATC	antisense	55
CG31121	ATGATGATGATGATGATGATG	sense	28
CG31121	CATCATCATCATCATCATCAT	antisense	6
CG31121	TCATCATCATCATCATCATCG	antisense	6
CG31121	ATCATCATCATCATCATCGTC	antisense	1
CG31121	ATGACGATGATGATGATGATG	sense	1
CG31121	GATGATGATGATGATGATGAT	sense	1
CG31150	AGTTGTAGTTGCATAAATATA	sense	1
CG31163	TCATCATCATCATCATCAACA	sense	1
CG31284	TCTTCTGCTTCAGCTTCGTGC	sense	2
CG31461	TCATCATCATCATCATCATCG	antisense	6
CG31461	ATCATCATCATCATCATCGTC	antisense	1
CG31461	ATGACGATGATGATGATGATG	sense	1
CG31461	ATGATGATGATGATGATGACG	sense	1
CG31461	TCGTCATCATCATCATCATCA	antisense	1
CG31549	TGGAGTGGCCAGCATGGTTAG	antisense	1

CG31771	ATCATCATCATCATCATCATC	antisense	55
CG31771	TGATGATGATGATGATGATGG	sense	49
CG31771	TCATCATCATCATCATCATCC	antisense	32
CG31771	ATGATGATGATGATGATGATG	sense	28
CG31771	CATCATCATCATCATCATCAT	antisense	6
CG31771	CCATCATCATCATCATCATCA	antisense	2
CG31771	GATGATGATGATGATGATGGA	sense	2
CG31771	GATGATGATGATGATGATGAT	sense	1
CG31790	TGATGATGATGATGATGATGA	antisense	552
CG31790	TCATCATCATCATCATCATCA	sense	197
CG31790	TCATCATCATCATCATCACAG	sense	6
CG31790	TGTGATGATGATGATGATGAT	antisense	5
CG31790	ATCATCATCATCATCATCACA	sense	4
CG31790	GTGATGATGATGATGATGATG	antisense	4
CG31790	CTGTGATGATGATGATGATGA	antisense	2
CG31865	CACTAGCGTATAATGTATATA	antisense	1
CG31866	CACTAGCGTATAATGTATATA	antisense	1
CG32017	ACAAAACATGTACGTGCTTGT	sense	1
CG32048	ATTCCTTTGCCGGGAGTTCGT	sense	1
CG32075	TAAAAACTAGTACTAGATCCA	sense	1
CG32164	TAGTGCAAAATAGGAGTTCTG	antisense	1
CG32165	TAGTGCAAAATAGGAGTTCTG	antisense	1
CG32170	TACGAATTGTTCGGACTGATG	antisense	1
CG32306	AAGACAGACTCGCCGTCCAAG	sense	1
CG32442	GACATCACCTCTGCTCCCTGG	sense	1
CG32521	GTGACGGCAAGGATTGCGGCA	sense	1
CG32667	CATCGCCTCCTTGAAGCCCTG	sense	1
CG32667	CGCCGGAAGGTCGCTGCCTGC	sense	1
CG32676	TGATGGTGATGATGATGATGT	antisense	1
CG32685	ATATTCACCATTTCCCTGGAC	sense	1
CG32694	TGATGATGATGATGATGATGA	sense	552
CG32694	TCATCATCATCATCATCATCA	antisense	197
CG32694	ATGATGATGATGATGATGATG	sense	28
CG32694	CATCATCATCATCATCATCAT	antisense	6
CG32694	TAATGATGATGATGATGATGA	sense	2
CG32694	TCATCATCATCATCATCATTA	antisense	1
CG32694	TCATCATCATCATCATTATCT	antisense	1
CG32694	TGATGATGATGATGATGAAGA	sense	1
CG3270	TTCGCACTCCGGCTCCTCGTC	antisense	1
CG32758	ACTGGCCCACGGCTGCACCGA	sense	1
CG3279	CTTTCTTCCACAGAATATCTC	antisense	1
CG3308	AGTGTGTGTCTGTGTGCGGAC	antisense	1
CG3308	ATAGTGTGTGTCTGTGTGCGG	antisense	1
CG3308	TCCACAAATACGACCCCCATT	sense	1
CG33080	ATACATAAGATGCCTTATCGC	sense	2
CG33080	TACTTAACTAACTATACGCAC	antisense	1
CG33080	TGTTTTTGTGCGTGCGTATAG	sense	1
CG33097	AGTACATCGTGGAGGTGCGGC	sense	1
CG33138	CGCTGTGGGACAGTCGTCTCT	sense	1
CG33144	ATAATTGTATATGTGTTAACT	sense	1
CG3332	TGATGATGATGATGCTGATGC	antisense	4
CG3332	ATGATGATGATGATGCTGATG	antisense	2
CG3337	ACGGTGTTTACTCCGTCCTCT	antisense	2
CG33470	CAGGGTGAGAACTTTGTGGCC	sense	1
CG33472	CTGCTTTTCTATTGATTTGGC	antisense	1
CG33523	TGGTGATGGCTATGGTCGCGC	sense	1
CG3368	CATTTAATGTTTGTACGCAGT	sense	1
CG33969	AGAGAGAAGGCTATTACCGTC	sense	1
CG33969	CAATGGCAATGACTTTGGTCC	antisense	1
CG33981	TGATCTGGCGTTGGGCTCGCT	sense	1
CG34136	TAAAGTTTACGGAAATAAAGG	sense	1
CG34179	AACAACTCCACGCTCAGCGAA	sense	1
CG34260	TGATGATGATGCTGCTGCTGA	sense	1
CG34268	CACCGGAACATGCTGCACCAC	antisense	1
CG34335	GGCGGTCGAGTGCCTCACAGT	sense	1
CG34360	CTGCTGATGATGCTGCTGTTG	antisense	1
CG34360	GTGCAATTGCTGGCAGCAAGA	antisense	1
CG34398	ATCATCATCATCATCATCATC	sense	55
CG34398	TGATGATGATGATGATGATGG	antisense	49
CG34398	ATGATGATGATGATGATGATG	antisense	28
CG34398	CATCATCATCATCATCATCAT	sense	6
CG34398	TCATCATCATCATCATCATCT	sense	4
CG34398	CCATCATCATCATCATCATCA	sense	2
CG34398	GATGATGATGATGATGATGAT	antisense	1
CG34398	GATGATGATGATGATGATGGC	antisense	1
CG34417	CACCACTACCCAAGATCAGGC	sense	1
CG34422	ATCCGCCTTCGCTGCCCGTGG	sense	1
CG3448	CCATCAATGTGTAGACGTGGC	sense	1
CG3523	CACATCCAAGGCGTGCAGGCG	antisense	1

CG3523	TGTACTTCCACGGGCTTTGTG	antisense	1
CG3529	CCATCAATGTGTAGACGTGGC	antisense	1
CG3585	CGAAAATGGGAACACCTGTCC	sense	1
CG3585	CGCAAGATTCGGCTGCAACTG	sense	1
CG3585	GAGCGCAGGTACTTCTTGTGC	antisense	1
CG3597	ACGCTACGACTCCAGATGTGC	antisense	1
CG3597	CGGGTGTGCAACCGATCTCTG	sense	1
CG3829	CTAGCGTGGATTCGGTCCAGC	sense	1
CG4000	AGGACCCGCTGGACTCGGCGC	antisense	1
CG4000	CAGGAGGTGCCGGACAAGCGG	sense	1
CG4000	GGTGGAGCTGGAGGAGCTGCT	sense	1
CG40084	ATACTACCTTCGACATCTTTT	antisense	1
CG40084	TCAGACAGCTTATTTCGTAGG	antisense	1
CG40084	TCCTACTGCCTCGCCTCTTTC	antisense	1
CG40182	CGAGCAGCATTGCCAGCCAAC	antisense	1
CG40271	TCCTACTGCCTCGCCTCTTTC	antisense	1
CG40339	TACAAGTCTGCTTGATGTTGC	antisense	1
CG40351	TGTTCATAAAATCCAAAGTGG	sense	1
CG40351	TTAAAACTATAATTAATTATT	sense	1
CG40351	TTTACAGATCAAATGGGTTTT	antisense	1
CG40451	TTATCGAAGGTGTTGGAATAC	sense	1
CG4068	TTGACTCCAACAAGTTCGCTC	sense	27
CG4068	TGGTAGCCTGTAGTTTGACTC	sense	9
CG4068	CGGTAGCCTGTAGTTTGACTC	sense	8
CG4068	GTCCAACTACAGGATACTGGG	sense	3
CG4068	TTTGACTCCAACAAGTTCGCT	sense	3
CG4068	AAATCTTAACCGCCGGAAGTC	sense	2
CG4068	CTTCCGCTGGCTTTGATTTTC	sense	2
CG4068	TAACCGCCGGAAGTCACTTCC	sense	2
CG4068	TCCAACTACAGGATACTGGGG	sense	2
CG4068	TGGCGCTTCACAGGCGCTGGA	sense	2
CG4068	AATCTTAACCGCCGGAAGTCA	sense	1
CG4068	AGGGACTTGTTTGAGTCCAAC	sense	1
CG4068	AGTCCAACTACAGGATACTGG	sense	1
CG4068	CTGGAAAATCTTAACCGCCGG	sense	1
CG4068	TGACTCCAACAAGTTCGCTCC	sense	1
CG40793	CTAAGAGACGCCTCTGTTGCT	sense	1
CG40798	TTTCTGTTAGCGGTTAACTGC	antisense	1
CG41053	TCCTACTGCCTCGCCTCTTTC	antisense	1
CG41126	ATACTACCTTCGACATCTTTT	antisense	1
CG41126	TCAGACAGCTTATTTCGTAGG	antisense	1
CG41332	TCCTACTGCCTCGCCTCTTTC	antisense	1
CG41484	TCGCCGCACCGTTACCGTTAC	antisense	1
CG41484	TGAGTGGAACTAGTGGGCAAC	sense	1
CG41557	TCAGACAGCTTATTTCGTAGG	antisense	1
CG41560	ATACTACCTTCGACATCTTTT	antisense	1
CG41560	TCAGACAGCTTATTTCGTAGG	antisense	1
CG41573	CTGTTCCCGTTGATTCCCGTT	antisense	1
CG41573	GCTGGTTGACTGTTCCCGTTG	antisense	1
CG41574	TCCTACTGCCTCGCCTCTTTC	antisense	1
CG41579	TCAGACAGCTTATTTCGTAGG	antisense	1
CG41584	ATACTACCTTCGACATCTTTT	antisense	1
CG41587	TTTCTGTTAGCGGTTAACTGC	antisense	1
CG41592	TCCTACTGCCTCGCCTCTTTC	antisense	1
CG4169	AACTCCACGGTCTTGCCAATG	antisense	1
CG4186	TCCTTTCGGTTTTTACTTTGT	antisense	1
CG4278	ATAGCCAGTCGTTGACGCCAC	antisense	1
CG4500	TGGCCATCTGCTGGGCGTCGT	antisense	1
CG4607	GCACCATCGGCTCCACCGACC	sense	1
CG4629	TGCTTCTGCAACCGATTGACC	antisense	1
CG4655	TGATGATGATGATGATGATGC	antisense	35
CG4655	TGATGATGATGATGATGCTGA	antisense	24
CG4655	ATGATGATGATGATGATGCTG	antisense	5
CG4655	TCATCATCATCATCATCAACA	sense	1
CG4658	TCAACCTGATGCACTCCAACT	sense	1
CG4662	ATTTAATCGTGCAATTTGTGT	antisense	1
CG4662	CGAAGAAGTGCAGCTGCAGTG	antisense	1
CG4673	GCTATGCGGTTTCGGCTCAGT	sense	1
CG4688	CCACGGTGATCTGGTGCTGAC	sense	1
CG4699	GTCGACGAAGATCTTTCGGAT	sense	1
CG4756	CTTTTAACGCTGGCCAACTGC	sense	1
CG4756	GTAGCGATAATTGGTATTGGC	antisense	1
CG4769	TCACCGGCGGAGTGGGCGCCC	sense	1
CG4825	TCGAGTGCTGGTGGGATGCCC	sense	1
CG4825	TCGGGCGACGCGCTACTGGAC	sense	1
CG4927	CACAAGATCGATGTGCGCACC	sense	1
CG5044	TAAAGAATATTGCAAAACGGC	antisense	1
CG5270	TGCAAATGGATGCCAGGGCTC	sense	1
CG5273	TTGTCTCCACTCGTCTAAGGG	sense	1

CG5315	TAGCATGTCCTTCGGCGTCCA	antisense	1
CG5455	CTGGTGTCAATGATATTTTGG	sense	1
CG5508	TATATGCGGCCTCTGTGCGGC	sense	1
CG5621	CTGTTTTGGGATCGATTCCAC	antisense	1
CG5644	GAATCCTCTTCCTTCGCCTTC	sense	1
CG5691	TCGCGGGACATGGCCACTCCC	sense	1
CG5728	ATCAGTCCTACATTCTATGTG	antisense	1
CG5734	CACCGAGCAGTTGACCAGGAT	antisense	1
CG5734	TGCCTGCAGTTCGATGAGGCG	sense	1
CG5794	GTGGGACGCCGCCGAGTACTG	sense	1
CG5815	CGGATTAGCCCACGTCGAGAT	antisense	1
CG5871	GCGGAACAGGTCCTTGCGCTG	antisense	1
CG5885	CTTGTCTTTCTTGTAGAAAGC	sense	1
CG5919	AGTGTGTGTCTGTGTGCGGAC	sense	1
CG5919	ATAGTGTGTGTCTGTGTGCGG	sense	1
CG5938	CGTCTATGCCCCAAAGTGCTG	antisense	1
CG5991	TCAGCCAGACTATTGTAGTGC	antisense	1
CG6028	ATCTATTGAAAACTACAAAAT	antisense	1
CG6055	ATCGGGCCTGTCGCAGCCAGC	antisense	1
CG6129	CTACCTCCGCCTTCATGGCCG	antisense	1
CG6201	TGCAGGACTCTTTAAGGACTC	sense	1
CG6218	AAATACTCTATTCTAAGCTCC	antisense	1
CG6218	CTGTCGTGGCTTCTTTTCATG	sense	1
CG6299	CACGGGTGTGAATAGTTTGCC	antisense	1
CG6321	TAAATATGACTTAAAAGGATG	sense	1
CG6404	AAGTTTTGGTAGATGTAATCG	sense	1
CG6459	AGTTTTTATTAGTTGTGTTTT	sense	1
CG6459	ATTGGCAACCGTTTCTATAGT	antisense	1
CG6459	TTTCTATTGTCTGCTGTCGCA	sense	1
CG6498	TACTCGGGTGTGCCGTACACC	antisense	1
CG6503	CTGTTCACGGCTCCTGTCCAC	sense	1
CG6503	GTTTTGCTGCTGTCCGTAATG	sense	1
CG6503	TATATCCCGCAGGACATCCGC	antisense	1
CG6503	TTTGCTGCTGTCCGTAATGGC	sense	1
CG6654	CGCATTTGAGTTGGGTCGTTC	antisense	1
CG6749	CTCAGTCCTGCTCACTTTGTG	sense	1
CG6762	CACGGCGTTGAAACTTGTTTG	antisense	1
CG6770	TTAGCGCCGCATGAAAAGCCA	antisense	1
CG6808	AGCATGTGCCGCACTTGCCGC	sense	1
CG6879	CCGTTGAATGTTGATGGCAGC	antisense	1
CG7156	TTACAAATTTTATTACTTACT	antisense	1
CG7326	TTAAACTATTACTCTACTCTC	antisense	1
CG7376	AATAGATGAATTGATTGTGGC	antisense	1
CG7414	TCCATACGAATTCGGTGGCTG	antisense	1
CG7518	AGCCACCATATGCCCGTTGAC	antisense	1
CG7739	TGGAAAGCTTATAATGGAGTT	antisense	2
CG7739	AATCAATAGAGATGCTCCTTT	sense	1
CG7739	TAAAATCATAGCTCATCATGG	sense	1
CG7739	TCTATTAACTCCATTATAAGC	sense	1
CG7739	TTAAAAGCCGCCTTGAGATGG	antisense	1
CG7739	TTGGAAAGCTTATAATGGAGT	antisense	1
CG7766	CTAGCGGCCGCGAGGTGTTCC	sense	1
CG7781	AATCTCCTCTTAAATGCAATA	sense	1
CG7839	TCATCATCATCATCATCATCG	antisense	6
CG7839	ATCATCATCATCATCATCGTC	antisense	1
CG7839	ATGATGATGATGATGATGACG	sense	1
CG7839	TCGTCATCATCATCATCATCA	antisense	1
CG7839	TGATGATGATGATGACGATGA	sense	1
CG7884	TGCTGCTGCTGCTGCTGGCCC	antisense	1
CG7888	TAAATATGACTTAAAAGGATG	antisense	1
CG7920	CAGGGTGGCGCCGTCCTTGAC	antisense	1
CG7998	AGGTGACCGTTTGCGGTGCCG	sense	1
CG8008	CTTCATTCATCATTTTAATTT	sense	1
CG8008	TCGTAGTAGTGTAAAAGGGTA	sense	1
CG8058	GTTGTATTTTTATTAGTTATA	antisense	1
CG8112	TAAACTGTCTGTACACAGGGC	sense	1
CG8112	TCGATTTGTATTTTAGGTAAA	sense	1
CG8199	AGTGGGTCACTTTGTATGTGC	antisense	1
CG8199	CAAAGTGACCCACTCTGGTGC	sense	1
CG8289	TACAGAAATGATGCCTTACAT	antisense	1
CG8311	CGTAGGAGACACTGCAGCCAG	sense	1
CG8312	TTATCTGAACGGTGTGTGTGC	sense	1
CG8451	AATTTATCACAGACATTATGC	antisense	1
CG8451	CAAACTACCAACAAATTCTGC	antisense	1
CG8455	AATTTATCACAGACATTATGC	sense	1
CG8500	ATTATAATTCGTATGCAACTA	antisense	1
CG8500	TGGTCCGCCGTCCAGCAGTGC	sense	1
CG8500	TTTGGCTGCTGGTCCGCCGTC	sense	1
CG8545	TGATGATGATGATGATGATGA	sense	552

CG8545	TCATCATCATCATCATCATCA	antisense	197
CG8545	ATCATCATCATCATCATCATC	antisense	55
CG8545	ATGATGATGATGATGATGATG	sense	28
CG8545	CATCATCATCATCATCATCAT	antisense	6
CG8545	TCATCATCATCATCATCATCG	antisense	6
CG8545	ATCATCATCATCATCATCGTC	antisense	1
CG8545	GATGATGATGATGATGATGAT	sense	1
CG8549	TCAAATCGCTTGCCACCTTTT	antisense	1
CG8745	TAATCCCCGATATTGTGTGTG	sense	1
CG8798	ATAAAACCTATCAACACCCGC	antisense	1
CG8862	TATTTGAACAGGGATAGCCCC	sense	1
CG9005	ACAACGAATCCCTATGGTTCC	sense	2
CG9062	CCAACTTGTAAAGAGCTCTAT	sense	1
CG9062	TACAGATTCTCCTTGAATGTG	antisense	1
CG9062	TTGAATGTGTTTGTGTTTGTC	antisense	1
CG9132	CTCTCCCTCACTCTCTCTCTC	sense	1
CG9170	TCATCATCATCCTCATCATCA	antisense	3
CG9170	TCCTCATCATCATCCTCATCA	antisense	1
CG9216	AATCTAAGCGTATATATTAAT	sense	1
CG9281	TTTATTTTACCTTTGTCAAGC	antisense	1
CG9311	CCACTCCTGCGCCTCCTTCGA	sense	1
CG9318	ATTCAGTTCGGCAACAGTGGA	sense	1
CG9339	ATACATATATATTTATATAAT	antisense	1
CG9393	TCGGAGTCTAGGAACTTGCGC	antisense	1
CG9425	CACACTGCTGCAGTTCGAGAG	antisense	1
CG9485	CGTCCAGCAGGAAGGCGGGGC	antisense	1
CG9485	GCACATGGTGGACCAGGGCTT	sense	1
CG9512	CCACCTCGATTGAGGGACCCA	sense	1
CG9526	TGGCCACATGTAGTTGGTTGC	antisense	2
CG9619	TAACGTAACGTATCAACACAA	sense	1
CG9629	GCGTCGATGTTGTACCGGAGT	sense	1
CG9666	GCGTCGATGTTGTACCGGAGT	antisense	1
CG9674	TGATGATGATGATGATGATGA	sense	552
CG9674	TCATCATCATCATCATCATCA	antisense	197
CG9674	ATCATCATCATCATCATCATC	antisense	55
CG9674	TGATGATGATGATGATGATGC	sense	35
CG9674	ATGATGATGATGATGATGATG	sense	28
CG9674	ATCATCATCATCATCATCATA	antisense	7
CG9674	CATCATCATCATCATCATCAT	antisense	6
CG9674	GATGATGATGATGATGATGAT	sense	1
CG9674	TAATTGGATCTGCACTGCGTT	antisense	1
CG9779	AAGAAGTGTTGAACTCTGCGC	sense	1
CG9779	TAAAAATCGATATTTACTGTG	sense	1
CG9780	ATCTCCATTCAGCGTAGTGTG	antisense	1
CG9780	ATTGGACGCGGCATACCACTC	sense	1
CG9780	ATTTTTAACACCACCGGTGGC	antisense	1
CG9780	CAGCGCACGCGACGTTTGGCC	sense	1
CG9780	CAGTAAATGCGTTCCTAGGGC	sense	1
CG9780	CATCTTCCAACAGTGCGTAGG	sense	1
CG9780	CATGAAACTGTGCAAATTGTG	antisense	1
CG9780	CGGCGTGCTCAGATTGTGCTG	antisense	1
CG9780	TACTGGAGCGAACATTTGGGC	antisense	1
CG9865	TCAAACAATCCGCAAAGCAGA	sense	1
CG9894	TGTGATGATGATGATGATGAT	sense	5
CG9894	ATCATCATCATCATCATCACA	antisense	4
CG9894	TGATGATGATGATGATAATGA	sense	2
CG9894	GTGATGATGATGATGATGATA	sense	1
CG9906	TCCTCCTTCTCCTCCTCCTCC	antisense	1
CG9914	CTCTCCAGCTCCACCAGCACT	sense	1
CG9915	CACAGAGAGCTCTGACGATGA	sense	1
CG9934	ATCATCATCCTCCTCCTCCTC	antisense	2
CG9934	TCATCATCCTCCTCCTCCTCC	antisense	2
CG9935	ATATCCTTCAAGATTTGCTTT	antisense	1
CG9935	TACCCCAACATCCGCAGTGTC	antisense	1
CG9935	TTCGGTGTACGTTTTCGCGAT	sense	1
CG9941	TGATGATGATGATGATGATGA	sense	552
CG9941	TCATCATCATCATCATCATCA	antisense	197
CG9941	ATCATCATCATCATCATCATC	antisense	55
CG9941	ATGATGATGATGATGATGATG	sense	28
CG9941	CATCATCATCATCATCATCAT	antisense	6
CG9941	ATGATGATGATGATGATGATA	sense	4
CG9941	TAATGATGATGATGATGATGA	sense	2
CG9941	TGATGATGATGATGATAATGA	sense	2
CG9941	GATGATGATGATGATGATGAT	sense	1
CG9941	TCATCATCATCATCATCATTA	antisense	1
CG9945	CTAACTAATTAAACTGAACTA	antisense	1
CG9945	TGGCATCGCCGCACAGCGCAT	sense	1
CG9986	TATGTATGTATATGCTTCGTT	sense	1
Chc	CACAGCTGCTTCTTGCCCTGC	antisense	1

Chc	TTATATTTGAATAAAGAGTGC	antisense	1
cher	AGGGAGCCAGCGCCGGTGGCC	sense	1
cher	GTGTGTCGAAGTTGTGCTTTA	antisense	1
CHES-1-like	TGATGCTGATGCTGATGCTGC	antisense	1
cic	AAGCCGGAGGACGCTGGCTCC	sense	1
Cks30A	TGATGATGATGATGATGTTGT	antisense	4
Cks85A	TCGATTTGTATTTTAGGTAAA	antisense	1
Cp1	ACAAGCACAAGATTGCCAAGC	sense	1
cpo	GATGGCTCTGGCCCCGTTGGG	sense	1
CRMP	GGAGTGGTCTTACCTGGTGGG	antisense	2
crq	AACTTGCAGTTTGCTGGGCTA	antisense	1
Csk	CTACTACCTTTAACTACCTAC	sense	1
CSN8	GGAAGCACCGCCAGCTCGTGG	sense	1
Cyp1	GTTTGATCTTTTGATGTTGGC	sense	1
Cyp28d1	CCGAAGTGATTTCCGATTGTG	sense	1
Cyp6d5	ACGGGAACACCTGATTTGGGC	antisense	1
Cyp6g1	TGGAGCACGAAACCCTTGGGA	antisense	1
Cyp6w1	TGCCCAGGCAGCTGTTTTTCA	sense	1
Cys	GCTGAGTCCAGATCTTCACGG	antisense	1
Cyt-b5-r	CAAAAGGCTGTTGTATTTGGC	antisense	2
Cyt-b5-r	CACCGTCCAGCAGAGGGCCAC	antisense	1
Cyt-b5-r	CATGATGAACTTTGCCGCCTG	sense	1
Cyt-b5-r	CCAGTAGAGGATGTTGGTGTG	antisense	1
Cyt-b5-r	CGTGTGTCCCATGCTCTGTCC	sense	1
Cyt-b5-r	CTCTGCTGGACGGTGATCGTG	sense	1
Cyt-b5-r	TCGGATCGGTAAAGTTTGTGC	sense	1
D2R	CAACAGCTTTGTGAACCCGGT	sense	1
Dcr-1	AACGCGCACCCGTCGCCCATC	antisense	1
Dcr-1	ACTCACTACTGGTCGTGCTGC	sense	1
Dcr-1	AGATGGAGCATGGTTCCGTGC	antisense	1
Dcr-1	CCGCGTGTTTTGCTCTTCGGG	antisense	1
Dcr-1	CCTTGTGGCGAACTGCCGTGT	sense	1
Deaf1	TGATGATGATGATGGTGATGG	antisense	1
Df31	TTACAGCAATACTCTGAAATG	sense	1
Dhc64C	CTTTCAATGCTCAGCATCCAG	sense	1
Dhc64C	GACAGCCGGTTCCACCTGGGC	antisense	1
dik	AGAACAGCCCCTTGGTCGTGC	sense	1
DNApol-iota	CTTGAGGAGTTTATCCTGGGC	antisense	1
dnc	TCGCAGCGGCAGCGGGCGCTG	sense	1
dome	TATTTTGTATTCTTTGCGATG	sense	1
Dot	CTGAAGTCGCTTTTAACGATG	antisense	1
dp	AACTGCCTGATCGTTAACCTC	antisense	1
dp	CCTGCCAACCCTCCCCCTGTG	sense	1
dp	TCGGTGGACAGTCCTCGTTGT	antisense	1
Dpit47	CGGGATTGCCTGGCGACGCTG	sense	1
drosha	CTGTATTGTGAACAGTTTTGC	sense	1
E(Pc)	TTTTTCACACTCTTTGCCGGG	antisense	1
e(y)3	ACGACGTGGAGCTAACTCGCT	antisense	1
ect	TCATCATCATCATCATCTTCT	antisense	2
ect	ATGACGATGATGATGATGATG	sense	1
Edem1	ACTAAACCGCTGTAGATCCAG	antisense	1
Edem2	CTACTTCAGTCGATAGTAGCA	sense	1
Ef1alpha48D	CTACGTGACCATCATTGATGC	sense	1
Ef1alpha48D	GTATGGTGGCTCGGAGGAGTC	antisense	1
Ef2b	ATCGAGGATGTGCCCTCTGGC	sense	1
Ef2b	TGTGTCCAGACCGAAACCGTG	sense	1
EfTuM	CAACTGCAGATGTTCTCCCGC	sense	1
egh	CAATAGACAACGTCTTGGAGC	sense	1
eIF-4a	CATCACCCAGTCGGTAATCTT	sense	1
Eip55E	TTCGCCGCCTTGGATAATGCC	sense	1
epsilonTry	CTTCCGCTCCAGCATTCGCGA	sense	1
Ets97D	TATTATCCATTTCGTATTCGC	sense	1
exba	CCACCTATATAAAACTCAAAA	sense	1
exo70	TCGAGCACCAATATTGTGGGC	sense	1
exo84	CTCTATCTCTTGCTTATTACA	sense	1
fab1	TGATCTGGCGTTGGGCTCGCT	sense	1
faf	CAAGGCGAACTAGATCGGCAG	sense	1
faf	GCGGTAGTGCAACTGGCCTGG	sense	1
fal	TAAATATAAGATGCATTTGTC	sense	1
fat-spondin	GTCAAGTGATAACGCGGGAAA	sense	1
fbl	AGAGAGAAGGCTATTACCGTC	antisense	1
fbl	CAATGGCAATGACTTTGGTCC	sense	1
fh	CCGACCAACGATCGAACGGAC	antisense	1
Fit1	TGGGCGAAAATACGTGGAACC	sense	1
Flo-2	ACTACTTATACAGATCTCTAC	antisense	1
Flo-2	ATGCTATATATACTATATACA	sense	1
for	AACAGAGCTCTGAAACAGAGT	antisense	1
form3	TCAAATATATTTAAGAGTTGG	antisense	1
Fps85D	AGTGATTCATAATTTGTATAT	sense	1

Fs	TGATGCTGATGCTGATGCTGC	antisense	1
Fs	TGCTGATGCTGATGCTGCTGC	antisense	1
fs(2)ltoPP43	TCATCATCATCATCATCCTCA	antisense	10
fs(2)ltoPP43	ATCATCATCATCATCCTCATC	antisense	3
fs(2)ltoPP43	TCATCATCATCATCCTCATCC	antisense	3
fs(2)ltoPP43	ATCATCATCATCCTCATCCTC	antisense	2
fs(2)ltoPP43	TGAGGATGATGATGATGATGA	sense	1
Fur2	TGATGATGATGATGATGATGA	antisense	552
Fur2	TCATCATCATCATCATCATCA	sense	197
Fur2	ATCATCATCATCATCATCATC	sense	55
Fur2	TGATGATGATGATGATGATGG	antisense	49
Fur2	ATGATGATGATGATGATGATG	antisense	28
Fur2	CATCATCATCATCATCATCAT	sense	6
Fur2	TCATCATCATCATCATCATCG	sense	6
Fur2	CCATCATCATCATCATCATCA	sense	2
Fur2	GATGATGATGATGATGATGAT	antisense	1
Fur2	GATGATGATGATGATGATGGG	antisense	1
G9a	CCCTTGGGAGATGTTAAGAGA	sense	1
gammaCop	CCACGTCCTGGTCCGTGCTGC	antisense	1
Gfr	TTGTTGATGGATTGTTTGTGC	antisense	1
Ggamma1	TCCAAAAACGTGTGATTCTTG	antisense	1
Glycogenin	AGAAGACCCTGAAGGACGCCG	sense	1
GlyP	CCACCATGTTGTCCTTGACGG	antisense	1
gro	CTCTATCTCTTGCTTATTACA	antisense	1
gry	CCTACTCCCTCTTTATTGCTA	sense	1
gry	CTGACCAGGACTTAGCACTGC	antisense	1
Gs2	TTAATCCTAAACTACATACAT	sense	1
Gug	TCATCATCATCATCATCCTCA	sense	10
Gug	TGAGGATGATGATGATGATGA	antisense	1
HDAC6	CAAGGAGCACTACGAGTGTCC	sense	2
HDAC6	ACGGACGTTGCAGATGCCGCT	antisense	1
He	CAAATATTTTAAGATTCGTGC	antisense	3
HERC2	CGAGGGAGAGCTGGACCCGGC	sense	1
HERC2	TTGGGGACTGCCCAACGCATC	sense	1
Hexo1	CACACTGGACAACCGCTTCTG	sense	1
His2A:CG31618	CGAGCAGCATTGCCAGCCAAC	antisense	1
His2A:CG33808	CGAGCAGCATTGCCAGCCAAC	antisense	1
His2A:CG33814	CGAGCAGCATTGCCAGCCAAC	antisense	1
His2A:CG33817	CGAGCAGCATTGCCAGCCAAC	antisense	1
His2A:CG33820	CGAGCAGCATTGCCAGCCAAC	antisense	1
His2A:CG33823	CGAGCAGCATTGCCAGCCAAC	antisense	1
His2A:CG33826	CGAGCAGCATTGCCAGCCAAC	antisense	1
His2A:CG33829	CGAGCAGCATTGCCAGCCAAC	antisense	1
His2A:CG33832	CGAGCAGCATTGCCAGCCAAC	antisense	1
His2A:CG33835	CGAGCAGCATTGCCAGCCAAC	antisense	1
His2A:CG33838	CGAGCAGCATTGCCAGCCAAC	antisense	1
His2A:CG33841	CGAGCAGCATTGCCAGCCAAC	antisense	1
His2A:CG33844	CGAGCAGCATTGCCAGCCAAC	antisense	1
His2A:CG33847	CGAGCAGCATTGCCAGCCAAC	antisense	1
His2A:CG33850	CGAGCAGCATTGCCAGCCAAC	antisense	1
His2A:CG33853	CGAGCAGCATTGCCAGCCAAC	antisense	1
His2A:CG33856	CGAGCAGCATTGCCAGCCAAC	antisense	1
His2A:CG33859	CGAGCAGCATTGCCAGCCAAC	antisense	1
His2A:CG33862	CGAGCAGCATTGCCAGCCAAC	antisense	1
His2A:CG33865	CGAGCAGCATTGCCAGCCAAC	antisense	1
His2B:CG17949	TCGCCTTCGACGAAATTCCGG	antisense	2
His2B:CG33868	TCGCCTTCGACGAAATTCCGG	antisense	2
His2B:CG33870	TCGCCTTCGACGAAATTCCGG	antisense	2
His2B:CG33872	TCGCCTTCGACGAAATTCCGG	antisense	2
His2B:CG33874	TCGCCTTCGACGAAATTCCGG	antisense	2
His2B:CG33876	TCGCCTTCGACGAAATTCCGG	antisense	2
His2B:CG33878	TCGCCTTCGACGAAATTCCGG	antisense	2
His2B:CG33880	TCGCCTTCGACGAAATTCCGG	antisense	2
His2B:CG33882	TCGCCTTCGACGAAATTCCGG	antisense	2
His2B:CG33884	TCGCCTTCGACGAAATTCCGG	antisense	2
His2B:CG33886	TCGCCTTCGACGAAATTCCGG	antisense	2
His2B:CG33888	TCGCCTTCGACGAAATTCCGG	antisense	2
His2B:CG33890	TCGCCTTCGACGAAATTCCGG	antisense	2
His2B:CG33892	TCGCCTTCGACGAAATTCCGG	antisense	2
His2B:CG33894	TCGCCTTCGACGAAATTCCGG	antisense	2
His2B:CG33896	TCGCCTTCGACGAAATTCCGG	antisense	2
His2B:CG33898	TCGCCTTCGACGAAATTCCGG	antisense	2
His2B:CG33900	TCGCCTTCGACGAAATTCCGG	antisense	2
His2B:CG33902	TCGCCTTCGACGAAATTCCGG	antisense	2
His2B:CG33904	TCGCCTTCGACGAAATTCCGG	antisense	2
His2B:CG33906	TCGCCTTCGACGAAATTCCGG	antisense	2
His2B:CG33908	TCGCCTTCGACGAAATTCCGG	antisense	2
His2B:CG33910	TCGCCTTCGACGAAATTCCGG	antisense	2
His2B:CG40461	TCGCCTTCGACGAAATTCCGG	antisense	2

His3.3B	TATATGCATATACGTAAGTGT	antisense	1
His3:CG31613	CCGAGCTTCTAATCCGCAAGC	sense	1
His3:CG33803	CCGAGCTTCTAATCCGCAAGC	sense	1
His3:CG33806	CCGAGCTTCTAATCCGCAAGC	sense	1
His3:CG33809	CCGAGCTTCTAATCCGCAAGC	sense	1
His3:CG33812	CCGAGCTTCTAATCCGCAAGC	sense	1
His3:CG33815	CCGAGCTTCTAATCCGCAAGC	sense	1
His3:CG33818	CCGAGCTTCTAATCCGCAAGC	sense	1
His3:CG33821	CCGAGCTTCTAATCCGCAAGC	sense	1
His3:CG33824	CCGAGCTTCTAATCCGCAAGC	sense	1
His3:CG33827	CCGAGCTTCTAATCCGCAAGC	sense	1
His3:CG33830	CCGAGCTTCTAATCCGCAAGC	sense	1
His3:CG33833	CCGAGCTTCTAATCCGCAAGC	sense	1
His3:CG33836	CCGAGCTTCTAATCCGCAAGC	sense	1
His3:CG33839	CCGAGCTTCTAATCCGCAAGC	sense	1
His3:CG33842	CCGAGCTTCTAATCCGCAAGC	sense	1
His3:CG33845	CCGAGCTTCTAATCCGCAAGC	sense	1
His3:CG33848	CCGAGCTTCTAATCCGCAAGC	sense	1
His3:CG33851	CCGAGCTTCTAATCCGCAAGC	sense	1
His3:CG33854	CCGAGCTTCTAATCCGCAAGC	sense	1
His3:CG33857	CCGAGCTTCTAATCCGCAAGC	sense	1
His3:CG33860	CCGAGCTTCTAATCCGCAAGC	sense	1
His3:CG33863	CCGAGCTTCTAATCCGCAAGC	sense	1
His3:CG33866	CCGAGCTTCTAATCCGCAAGC	sense	1
hoe2	GAGCTCACCGAGCACGTTATC	sense	1
Hr4	TGATGATGATGATGATGATGA	antisense	552
Hr4	TCATCATCATCATCATCATCA	sense	197
Hr4	ATCATCATCATCATCATCATC	sense	55
Hr4	ATGATGATGATGATGATGATG	antisense	28
Hr4	CATCATCATCATCATCATCAT	sense	6
Hr4	TAATGATGATGATGATGATGA	antisense	2
Hr4	GATGATGATGATGATGATGAT	antisense	1
Hr4	TCATCATCATCATCATCATTA	sense	1
Hr4	TCATCATCATCATTATCATCA	sense	1
Hr4	TGATGATGATAATGATGATGA	antisense	1
Hr4	TGATGATGATGATGATGAAGA	antisense	1
Hs3st-A	ATCATCATCATCATCATCATC	antisense	55
Hs3st-A	ATGATGATGATGATGATGATG	sense	28
Hs3st-A	TGATGATGATGATGATGATGT	sense	16
Hs3st-A	CATCATCATCATCATCATCAT	antisense	6
Hs3st-A	TCATCATCATCATCATCATCG	antisense	6
Hs3st-A	ATGATGATGATGATGATGTTA	sense	1
Hs3st-A	GATGATGATGATGATGATGAT	sense	1
Hsp70Aa	GTGCAAGTTAAAGTGAATCAA	sense	1
Hsp70Ab	GTGCAAGTTAAAGTGAATCAA	sense	1
Hsp70Ba	GTGCAAGTTAAAGTGAATCAA	sense	1
Hsp70Bb	GTGCAAGTTAAAGTGAATCAA	sense	1
Hsp70Bbb	GTGCAAGTTAAAGTGAATCAA	sense	1
Hsp70Bc	GTGCAAGTTAAAGTGAATCAA	sense	1
Hsp83	TGCTGCAGCAGAACAAGGTCC	sense	1
Hsp83	TGGGTGATCGGGGTTGATCTC	antisense	1
Hsp83	TTCGAGAGCCTGTGCAAGCTG	sense	1
htt	TGATGATGATGATGATGATGA	sense	552
htt	TCATCATCATCATCATCATCA	antisense	197
htt	ATCATCATCATCATCATCATC	antisense	55
htt	TCATCATCATCATCATCATCC	antisense	32
htt	ATGATGATGATGATGATGATG	sense	28
htt	CATCATCATCATCATCATCAT	antisense	6
htt	ATCATCATCATCATCATCCTC	antisense	5
htt	TCATCATCATCATCATCCTCC	antisense	3
htt	ATGATGATGATGATGATGACG	sense	1
htt	GATGATGATGATGATGATGAT	sense	1
htt	TCGTCATCATCATCATCATCA	antisense	1
htt	TGCTGATGCTGATGCTGCTGC	antisense	1
IM10	CAGGGTGAGAACTTTGTGGCC	sense	1
InR	CTGCTGATGATGCTGCTGATG	antisense	1
InR	TGATGATGATGCTGCTGCTGA	antisense	1
IP3K1	CCACATCGACTGGAATAGTGC	sense	1
IP3K1	CGGCTGCCCATCTTGATGTCC	antisense	1
IP3K1	GGACCATCGAAGTGCTTGGGC	sense	1
Irp-1B	TTTTGGGGAGTTCCAGTGGGA	sense	1
JTBR	AGAGTGTATGTACTTAAACTA	antisense	1
Kap-alpha3	TTTCAAAAAGAATGTAGACCT	sense	1
katanin-60	AAAATTTCAGAATCCTCTCTT	antisense	2
katanin-60	CACACATCTCTGCCTGCACTA	antisense	1
katanin-60	TACTTTCACATATACATATAT	sense	1
katanin-60	TATATCATAGTTTAGTGCAGG	sense	1
KrT95D	TTAGGAACCTGCAGGTGTGGC	sense	1
kst	TCGAGTCCTGGCTAGGTGAGC	sense	1

kuk	CTCCGGCTCGCACTTTCGCCC	sense	1
l(1)G0469	GTCACCGGTGTCGGCTGCTGC	antisense	1
l(2)01810	ATCATGGGCTTCTTGGCCATC	sense	1
l(2)gl	ACACGATGCGGTGGACTCCCG	antisense	1
l(2)gl	ACTTTATACCCTTTACCCGTA	sense	1
l(2)gl	CAACACAGACTCTTCATTTTC	antisense	1
l(2)gl	GATAACCCAGAACCGCCGTCG	sense	1
l(2)gl	TAATATCGCCAATAGAGCTGC	antisense	1
l(2)gl	TGATCAGGCGGGTAGCCCTGT	sense	1
l(2)gl	TTCCTCATAAGCCTTCGGCTC	sense	1
l(2)gl	TTCTTCCAACACAGACTCTTC	antisense	1
l(3)02640	TCTTTGCTCGGTCGTAGTATC	sense	1
l(3)04053	CACTGGAACTCCTCCTGTTGC	antisense	1
l(3)73Ah	TACATTCGTGTGTTTTGTACA	antisense	1
l(3)j2D3	TCATCATCCTCCTCCTCCTCC	antisense	2
Lam	TGTTGCTGTTCTTATTGTTGC	antisense	1
ldlCp	CTGAAGCTCTGGCAGATGCGC	antisense	1
ldlCp	TAAAGAATATTGCAAAACGGC	sense	1
Lmpt	TGATGATGATAATGATGATGA	sense	1
lolal	GAAACTTATATGTACGCAAGA	sense	1
Lsd-1	ACGACAACAGAGGTTGCCCAC	sense	2
Lsm11	CAAGCAGTGGAACCTCCTACT	sense	1
lva	TCAATAGCCATGTGGAGCGAG	sense	1
LysS	GACAAGTGGACCTGCATTGCC	sense	1
mask	ATCATCATCATCCTCCTCCTC	antisense	2
mask	TCATCATCATCATCCTCCTCC	antisense	2
mbl	GTAAAACAAAAAACTATCTCT	sense	1
MED21	TGTTCATAAAATCCAAAGTGG	antisense	1
MED21	TTAAAACTATAATTAATTATT	antisense	1
MED21	TTTACAGATCAAATGGGTTTT	sense	1
Mekk1	CAGGAGAGCAATCCACTGGGC	sense	1
Mes-4	CGGCAAACTCAAGGATGATGC	sense	1
milt	TTCACGATCTGCTCATGTTGC	antisense	1
Mio	GGTCACTCAGCAAATGCAGCA	sense	1
Mitf	TGTCTATGTAGTTCTTGGAAG	antisense	1
Mkp3	CTCGGTGGTGCTGGTCCACTG	sense	1
mld	TTTGTTTAGGCTATTCAACTG	antisense	1
Mmp1	TTAAGACAGTCCATTAAACGA	sense	1
mnb	GTACAGACGACTGTGGCGTGT	antisense	1
mnb	TGAGTGAGTTTTCGTGTGTTT	antisense	1
Mnt	TCATCCTCCTCATCATCATCA	sense	1
MP1	AAGAAGTGTTGAACTCTGCGC	antisense	1
MP1	CACTGAGGTTGCTGGTTTCGC	antisense	1
MP1	TAAAAATCGATATTTACTGTG	antisense	1
mRpL1	TTGTTGATGGATTGTTTGTGC	sense	1
mRpL35	ATCTATTGAAAACTACAAAAT	sense	1
mrt	CGTCTATGCCCCAAAGTGCTG	sense	1
msl-3	ATTACGATTAAAGCTTATGCT	sense	1
msl-3	TCAAATATATTTAAGAGTTGG	sense	1
Msp-300	AACCAGTCGGCGCACTGCTTA	antisense	1
Msp-300	TTGATAGCACTTTCATGCGCG	sense	1
mt:ATPase6	TGTGTTTGCTGTATTAAGAAC	sense	1
mt:CoI	ATCCTGGAGCATTAATTGGAG	sense	1
mt:CoI	ATTATAATTGGTGGATTTGGA	sense	1
mt:CoI	CGAGCTGAATTAGGACATCCT	sense	1
mt:CoI	GATTAAAAAGTCATTTCATTA	sense	1
mt:CoI	GGATTTGTTTTTTTATTTACA	sense	1
mt:CoI	TTTTATTTACAGTAGGAGGAT	sense	1
mt:CoII	AATGAATTAATAACTGATGGA	sense	1
mt:CoIII	ACTATTGCAGACTCAATTTAT	sense	1
mt:CoIII	AGAAGGAACATACCAAGGATT	sense	1
mt:CoIII	AGGAGTTACTGTAACTTGAGC	sense	1
mt:CoIII	TTACTATTTTAACTGTATATC	sense	1
mt:Cyt-b	AAGATATTGTAGGATTTATTG	sense	1
mt:Cyt-b	ATAGTGTTAATCATATTTGTC	sense	1
mt:ND1	AATTTTTATAGCTGAATATGC	sense	1
mt:ND3	ATTTTTGATGTAGAGATTGCA	sense	1
mt:ND5	CGGGTTTAACTGTTAGTTATT	sense	1
mt:ND5	TCTTATAATGCTGGTATATTA	sense	1
mts	AAGAACTACAACTACTTAGCT	sense	1
mus309	CTGCCGCCTTGGCTTCCTTGG	antisense	1
Nc73EF	ATGAACAATCCTCCTCCACCA	sense	1
nct	TTTATACCTTCTGCATCTATT	sense	1
nej	ACTACATCTTTCACTGCCATC	sense	1
Ngp	TCAACTTTTTATTTGGATTCT	antisense	1
ninaB	CAACTGCTCGTCGGATGTGTG	sense	1
ninaC	TCGTAATACGACTACCAATAA	sense	1
ninaE	ACAGCAACAACAACAACAAGA	sense	2
Nipped-A	CATCATACTGTTTATCTGGGT	antisense	1

Nrx-1	AGCCGGATCCGGATGTCGCGG	sense	1
Nrx-1	GCCGCCATTCGAATCGGGTAG	sense	1
Ntl	AGAGGCAGGACTCTGGGCCAA	sense	2
Ntl	GATGCGAGCAGAGGCAGGACT	sense	2
Ntl	GAGGCTTGGTGATGCGAGCAG	sense	1
Ntl	GCAGAGGCAGGACTCTGGGCC	sense	1
Ntl	GGCAGGACTCTGGGCCAAACA	sense	1
OstStt3	CTAATCACCTTCGCCATCCTG	sense	1
ovo	TGATGATGATGATGATGATTA	antisense	9
ovo	TCATCATCATCATCATCAACA	sense	1
p47	ACCAAATTGGCCACGGGACTG	antisense	1
Patj	AGAGTGTATGTACTTAAACTA	sense	1
pbl	ATCATCATCATCATCATCGTC	sense	1
pbl	ATGACGATGATGATGATGATG	antisense	1
pbl	ATGATGATGATGATGATGGTG	antisense	1
Pc	TCGGTTTCGCATGGAGTTTTC	antisense	1
Pc	TCTCTTAGCAGTCATTCAAGA	sense	1
PebIII	CTGGTGGACAACGGAAAGTGC	sense	1
Pect	AAACATAAAACTTGAACTCGC	antisense	1
Pect	GCATCCTACTTCCGTTGCTGA	antisense	1
Pect	TATATATTCAGCAACGGAAGT	sense	1
Pepck	AGGAGATGGGAATGCCACGGA	antisense	1
Pgant35A	CCAGCACGGTCTTAATCGAGC	antisense	1
pgant6	CTATCCGCCGGTGGATCCGCC	sense	1
Pgi	CCGCCAAGACCTGGCTCCTGG	sense	1
Pgk	CATCTCGTTGACCTTGTCCAG	antisense	2
Pgm	TAGCTGAAGTTGTCGGCCTCC	antisense	1
Pi3K59F	CTCGGGCCTGTACTCCGAGGA	sense	1
pip	TAAATATAAGATGCATTTGTC	antisense	1
PIP82	TGATGATGATGATGATGATGA	antisense	552
PIP82	TCATCATCATCATCATCATCA	sense	197
PIP82	ATCATCATCATCATCATCATC	sense	55
PIP82	ATGATGATGATGATGATGATG	antisense	28
PIP82	TGATGATGATGATGATGATGT	antisense	16
PIP82	TGATGATGATGATGATGTTGA	antisense	7
PIP82	CATCATCATCATCATCATCAT	sense	6
PIP82	TGTGATGATGATGATGATGAT	antisense	5
PIP82	ATCATCATCATCATCATCACA	sense	4
PIP82	GTGATGATGATGATGATGATG	antisense	4
PIP82	ATGATGATGATGTTGATGATG	antisense	1
PIP82	GATGATGATGATGATGATGAT	antisense	1
Pitslre	TACCGCCGGCGCCCAGCTATG	sense	1
ple	GATTGTTGTATCTATATCATT	antisense	1
plexA	ACGCCATGCTTGCGGAAGAGT	sense	1
plx	CAACTGAAGAGTCCCATGATG	sense	1
pnt	TGATGATGATGATGATGATGA	antisense	552
pnt	TCATCATCATCATCATCATCA	sense	197
pnt	ATCATCATCATCATCATCATC	sense	55
pnt	TGATGATGATGATGATGATGC	antisense	35
pnt	ATGATGATGATGATGATGATG	antisense	28
pnt	TGATGATGATGATGATGCTGA	antisense	24
pnt	TCATCATCATCATCATCACCA	sense	10
pnt	CATCATCATCATCATCATCAT	sense	6
pnt	ATGATGATGATGATGATGCTG	antisense	5
pnt	GTGATGATGATGATGATGATG	antisense	4
pnt	GATGATGATGATGATGATGAT	antisense	1
pnt	TCATCATCATCATCACCACAG	sense	1
pnt	TGGTGATGATGATGATGATGA	antisense	1
poe	GATTGTGCACTGCATTTGTTG	sense	1
Pof	TTAAGACAGTCCATTAAACGA	antisense	1
POSH	TTGGCATGCAACTGGGATTGC	antisense	1
Pp2B-14D	TCACTCTGTTTCAGTATTTGC	antisense	1
Ppt1	AAACTGAATTTATTAAACATC	antisense	1
Prm	ACCTCATCCTGCTGCAGGACG	sense	1
Psf3	AATAAGCCAAAGCGATGTTGC	sense	1
Ptp99A	AACAGATACTAGGACGGACTG	antisense	1
Ptp99A	AACTATAAGTGTAAATCGGCA	sense	1
Ptp99A	CTAGGTTCGCTTAAGTTTGTC	antisense	1
Ptp99A	CTAGTATCTGTTATTCTTTTC	sense	1
Ptp99A	CTCTACACCACCCACACCACA	sense	1
Ptp99A	TACTTTAACTTACACACCCAC	sense	1
Pu	CACGCGTTCAGAGCCCCAATC	sense	1
pyd	ATCGCTACGATTTATAAATGG	antisense	1
PyK	AATGGTGAAGAAGCCACGTCC	sense	1
r2d2	AAATCCAACTGCTGCCCGGCA	sense	1
r2d2	ATGTTAACTTTGTACACGATT	sense	1
r2d2	CATCTATGTTAACTTTGTACA	sense	1
Rab6	TAACTGGAACTGTGGATGGCA	sense	1
RabX6	CTTTCTTCCACAGAATATCTC	sense	1

rad	TTAGCGTTAACGTTATCTAGG	sense	1
Rad23	TATGCAAATAACACCGGAAGC	sense	1
raps	AATTGTAACTATGAGTATTGC	sense	1
Rbm13	TGATGATGATGATGATGATGA	sense	552
Rbm13	TCATCATCATCATCATCATCA	antisense	197
Rbm13	ATCATCATCATCATCATCATC	antisense	55
Rbm13	TCATCATCATCATCATCATCC	antisense	32
Rbm13	ATGATGATGATGATGATGATG	sense	28
Rbm13	CATCATCATCATCATCATCAT	antisense	6
Rbm13	ATCATCATCATCATCATCCTC	antisense	5
Rbm13	TCATCATCATCATCATCCTCC	antisense	3
Rbm13	ATCATCATCATCCTCCTCCTC	antisense	2
Rbm13	TCATCATCATCATCCTCCTCC	antisense	2
Rbm13	ATCATCATCATCATCCTCCTC	antisense	1
Rbm13	ATGATGATGATGATGATGACG	sense	1
Rbm13	GATGATGATGATGATGATGAT	sense	1
Rbm13	TCATCATCATCCTCCTCCTCA	antisense	1
Rbm13	TCGTCATCATCATCATCATCA	antisense	1
Rbp2	TCATCATCATCATCATCCTCC	antisense	3
Rbp2	ATCATCATCCTCCTCCTCCTC	antisense	2
Rbp2	ATCATCATCCTCCTCCTCCTC	antisense	2
Rbp2	TCATCATCATCATCCTCCTCC	antisense	2
Rbp2	TCATCATCATCCTCCTCCTCC	antisense	2
Rbp2	TCATCATCCTCCTCCTCCTCC	antisense	2
Rbp2	ATCATCATCATCATCCTCCTC	antisense	1
Rbp2	TGATGATGATGATGACGATGA	sense	1
ref(2)P	CAGCCATCGCATTCAACGGCG	antisense	1
regucalcin	TCGAGGGCGAAACCTTGGCCG	sense	1
repo	TTACAAATTTTATTACTTACT	sense	1
Rfabg	ACTTAACGCACAGTACGGAGC	sense	1
Rfabg	GGCAACTACTATGACTATTCC	sense	1
Rfabg	GTTTGAATTAAAGTCTCAAAA	sense	1
Rho1	TCGAATTCGTGCTGAGTGTTG	sense	2
RhoGAP71E	GCGAAATGCGATAGGCGAGCG	sense	1
rin	CCACACTCTCAATCGACACAG	sense	1
rin	TACTCTCTACCACCCACCACC	sense	1
r-l	TAAGGATTTCCTCGTTGGATC	antisense	1
rols	GTAGACAGTGCCGCCGCCCGG	antisense	1
RpL19	GATCCCAATGAAATCAACGAG	sense	1
RpL28	TGATCGTGTTGATAAACTTAT	sense	4
RpL31	CCACTCCATTCCGCATTCGCG	sense	1
RpL35A	AACACTTTAAATTTAATTAAA	sense	1
RpL38	ATATTTCTACTGCTAAGGAAT	sense	1
RpL4	GCAGCGTGCGCCGCCTGAACC	sense	1
RpS18	TGGACTCGAAGCTGCGTGACG	sense	1
RpS19a	ACACCGTTGCGCTTGCGTCCG	antisense	1
RpS19a	CGCCCGTTTGGTCGAGAAGCA	sense	1
RpS26	CGCCGTAACGGAGGACGCAAC	sense	1
RpS6	TCCACGAGAGGAGAAATAAAA	sense	1
RpS8	TCCGCAAGAAGCGCAAGTTCG	sense	1
Rpt4	CGGACTTTGTGTGACAGGCCC	antisense	1
rut	TGCTGATGCTGATGCTGCTGC	antisense	1
Rya-r44F	AGTGGATCGCATCGTGGCGAT	sense	1
sano	ATCATCATCATCATCATCATC	antisense	55
sano	TCATCATCATCATCATCATCG	antisense	6
sano	GATGATGATGATGATGATGAT	sense	1
sano	TGATGATGATGATGATGATTC	sense	1
Sap-r	AACCTGCTTTCCCGCCTGATG	sense	1
Sara	GACACTAGCTCTACATTGGGC	sense	1
sdt	TGCTGATGCTGATGCTGCTGC	antisense	1
SelR	CGACGCACGCTTGTTTTCGCC	sense	1
sens	TGATGATGATGATGATGATGA	sense	552
sens	TCATCATCATCATCATCATCA	antisense	197
sens	ATCATCATCATCATCATCATC	antisense	55
sens	ATGATGATGATGATGATGATG	sense	28
sens	CATCATCATCATCATCATCAT	antisense	6
sens	TCATCATCATCATCATCATCT	antisense	4
sens	GATGATGATGATGATGATGAT	sense	1
SerT	ATGGTGATGATGATGATGATG	antisense	1
SerT	GTGATGATGATGATGATGCTG	antisense	1
SerT	TGATGATGATGATGATGCTGT	antisense	1
sesB	CAGTGTTATCGATCGGATGCA	sense	1
sev	TATAAGTTTTATTTGCACCGC	sense	1
sev	TGTCCACTGCCTATTCCTGGC	sense	1
shot	TGATGATGATGATGATGTTGG	antisense	1
sif	CTTTGTGGAGGCTGTTAGCCC	antisense	1
Sk2	TCCATATCGGAGAGCATCTAC	sense	1
slik	GGAAAAGAACGTTTGCTGAGG	sense	1
sls	CACCTACGATTTTGGCTTCGT	sense	1

smg	TGATGCTGATGCTGATGCTGC	antisense	1
smg	TTGCAAAACAAATTGGCCGGT	sense	1
Smg5	CACGCCTTTCTGGTTGCTGCC	sense	1
Smr	TGATGATGATGATGATGATGT	antisense	16
Smr	TGATGATGATGATGATGTTGA	antisense	7
Smr	CTGATGATGATGATGATGATG	antisense	4
Smr	TGATGATGCTGATGATGATGA	antisense	2
Smr	TGCTGATGATGATGATGATGA	antisense	2
Smr	ATCATCATCATCATCATCAGC	sense	1
Smr	ATGATGATGATGTTGATGATG	antisense	1
Smr	ATGATGCTGATGATGATGATG	antisense	1
Smr	CAGAAGCAAGGAACCACCGCC	sense	1
Smr	CTGGAGCCATCGGTGAGATTC	antisense	1
Smr	TGATGCTGATGATGATGATGA	antisense	1
Smr	TGCTGATGATGCTGATGATGA	antisense	1
Sox14	ACTTAAATATCTCTCACATTT	antisense	1
Spase18-21	CATGAAGATCGTGATGATGCC	antisense	1
spen	ATCTCGATCTCGTGCATCTTC	sense	1
spen	CAGCTTCAGCGTCTGCATCCA	sense	1
Sra-1	AAATACTCTATTCTAAGCTCC	sense	1
Sra-1	CTGTCGTGGCTTCTTTTCATG	antisense	1
sta	GGGCAAGACCTGGGAGAAGCT	sense	1
StlP	AGACTACTTTAATGCATATGG	antisense	1
Strn-Mlck	ATGCTGGAGGAGGCGCACGGT	sense	1
sty	TTTGCCACTGTTTTTGTTGTC	antisense	1
svr	AAAGTACGCGAACGCAGCAGC	antisense	1
Syx17	TTAATATGCAATAATAACTCG	sense	1
Syx18	TTTGTTTAGGCTATTCAACTG	sense	1
T48	TATTATCCATTTCGTATTCGC	antisense	1
tafazzin	TACTAATAATGCACACTGATT	antisense	1
tafazzin	TTTGCTTTTGAGCTTGTTGCA	sense	1
Tango7	ATGTGTACTACCACCTGGTCC	sense	1
tara	CGGCGCTGCGAGGTCCACGTC	sense	1
Tcp-1eta	CATCCGCAAGGCCCTGCAGCT	sense	1
TfIIFbeta	TTTACCAGAGTGCAAGCTAGT	antisense	1
th	AGGAGAGCTCTTCGATTGGAG	sense	1
th	CGCAACAGTGGACAGTTGGGC	antisense	1
th	GATGAGAGTGATGTCTGCTGC	antisense	1
th	GTCATGTGGTGGCCTGCGCCA	sense	1
Thiolase	TTTTATACACCTTACAACTAC	sense	1
Thor	AGGAAGGTTGTCATCTCGGAT	sense	2
Tis11	AGGATGCTGCTCGCCCACGGC	sense	1
Tis11	GAATTGATATCAAGGATCGGT	antisense	1
tkv	GTATCTTTATCTGTAACCGTT	sense	1
tkv	TTGTGGTCTGCGAGCAGTAAT	antisense	1
tnc	TTTTCCTCTTCCTCCTCCTCA	antisense	1
tra	TACATTCGTGTGTTTTGTACA	sense	1
trbl	TGGCTCAGAATGCCAATGGGC	sense	1
trk	CTACATGAGCATCGATCCGCC	sense	1

Trxr-1	TGCAGTCCGTACAGAACCACA	sense	1
Tsf1	CGCAACTGGCTACGCGGATGC	sense	1
Tsf1	CGGACCCGCCTGCTCCTGGGC	sense	1
Tsf1	GACAAGTTTGGTGCCCGCGGC	sense	1
Tsf1	TCCTGCTCCACGCTGGTGGTG	antisense	1
Tsp42Er	CCTAAGATTGTTGGCTGGATG	sense	1
Tsp5D	ATAAGCATATCCAGGTCCAAG	antisense	2
tst	TACAAGCACGCCATCTACGTT	antisense	1
TwdlT	TGATGATGATGATGATGATTT	sense	1
UbcD2	AACTATAAACTATTCAACTGC	sense	1
Ubc-E2H	CAAACCAATCAAATCAAAAGC	sense	1
Ubp64E	TAATCTCACCGGTGGACGATG	sense	1
Unc-89	CAGGTAGGTGTAGGTCTTTGG	antisense	1
up	AATCCACACTCTGGGCCCGCC	sense	1
up	TCCTCCTCCTCCTCATCCTCC	antisense	1
Usp36	GCAGTGACAGTAAAGATGTGG	sense	1
Vap-33-1	CATTGAACCAGAACATGAGTT	sense	1
vav	ATTTTTAAATCTATTGTTGCT	antisense	1
Vha100-1	CTCCTGTTTATTAACTATACT	sense	1
Vha26	CATTTCGGTGCTTCGCTGAGC	antisense	1
Vha68-2	ACTTCCCCGAGCTGTCCGTGG	sense	1
Vhl	CCAACTTGTAAAGAGCTCTAT	antisense	1
Vhl	TACAGATTCTCCTTGAATGTG	sense	1
Vhl	TTGAATGTGTTTGTGTTTGTC	sense	1
vir-1	ACCATCACGCCCTCAGCCCGA	sense	1
vn	ATGATGATGATGATGATGGTG	antisense	1
wmd	CGGCGAGCTGATGGAATTGCC	antisense	1
wuho	CGGACTTTGTGTGACAGGCCC	sense	1
Xbp1	ACAGGTGGACACACAGTCGTC	sense	1
yip7	AAAGACCGCTGTTGCCTCCGG	sense	1
Yippee	AAAAGATGGGCTGCTACTCAG	sense	1
yl	GAGCAGAACGGTCACTTTCAC	sense	1
Yp1	AAGTGGATCGTCCAGATGGTC	sense	1
Yp1	AGCTGCGCCGTGTCACCGGTC	sense	1
Yp1	ATCCACACCTCGGTCTACGGC	sense	1
Yp1	GCCGTCGCAGTGGCTCTCCGG	sense	1
Yp1	TGACCGGTCTGGCTCGCGGTG	sense	1
Yp2	ACAATAAAAAACGTTTGCATT	sense	1
Yp2	ACCGATTTCGATCTGCAGGGC	sense	1
Yp2	GCTTCCTGCGCTGCGTTTGCT	antisense	1
Yp3	TGATCGGCCAGGGAATCAGCG	sense	1
Yp3	TGTGGAGACGGCCAAGGCACA	sense	1
zf30C	CTGCGGAACACTTGGTTTTGC	antisense	1
zfh2	CGAAGTCGTTTCTGAAGATGC	sense	1

Table II-S1D. Summary of mRNA-matching, 21-nt reads from pyrosequencing and sequencing-by-synthesis of a small RNA libraries enriched for 3´ terminally modified small RNA from wild-type heads.

	Total S2 reads			number of unique 21-mers	
Gene	**sense + antisense**	**antisense**	**sense**	**antisense**	**sense**
5-HT1B	1	1	0	1	0
Ack	1	0	1	0	1
Act5C	1	1	0	1	0
Act79B	1	1	0	1	0
Act87E	1	1	0	1	0
ade3	1	0	1	0	1
ade3	1	1	0	1	0
AGO2	8	3	5	3	4
alc	1	0	1	0	1
Ald	1	0	1	0	1
alpha4GT1	1	1	0	1	0
alphaTry	1	0	1	0	1

	Total S2 reads			number of unique 21-mers	
Gene	**sense + antisense**	**antisense**	**sense**	**antisense**	**sense**
alphaTub84B	1	0	1	0	1
alphaTub84D	1	0	1	0	1
Amy-d	4	0	4	0	4
Ank	2	0	2	0	2
AnnIX	1	0	1	0	1
Apc	3	2	1	2	1
Apc2	1	1	0	1	0
apt	3	0	3	0	2
arm	1	0	1	0	1
Asator	1	0	1	0	1
asrij	1	0	1	0	1
ATPsyn-beta	1	0	1	0	1

aux	3	1	2	1	2
bel	2	1	1	1	1
betaTry	1	0	1	0	1
bigmax	1	1	0	1	0
bin3	3	3	0	3	0
blw	2	0	2	0	2
Bruce	1	1	0	1	0
BRWD3	1	0	1	0	1
c11.1	1	0	1	0	1
cal1	1	0	1	0	1
CalpB	3	0	3	0	3
Ca-P60A	1	0	1	0	1
Cap-H2	1	0	1	0	1
Cas	1	1	0	1	0
cathD	1	0	1	0	1
Cbl	1	1	0	1	0
Ccn	148	68	80	6	5
Ccp84Aa	1	1	0	1	0
Ccp84Ab	1	1	0	1	0
Cct5	1	1	0	1	0
Cdep	1	1	0	1	0
ced-6	2	0	2	0	2
Cf2	1	0	1	0	1
CG10011	1	1	0	1	0
CG10055	1	1	0	1	0
CG10077	1	0	1	0	1
CG10147	1	1	0	1	0
CG1021	1	0	1	0	1
CG10214	1	1	0	1	0
CG10237	1	0	1	0	1
CG10249	1	1	0	1	0
CG10274	1	0	1	0	1
CG10375	1	0	1	0	1
CG10433	1	0	1	0	1
CG10444	1	0	1	0	1
CG10479	1	1	0	1	0
CG10631	1	1	0	1	0
CG10641	1	1	0	1	0
CG10646	2	2	0	1	0
CG10673	1	0	1	0	1
CG10681	2	0	2	0	1
CG10713	1	1	0	1	0
CG10874	1	0	1	0	1
CG10918	1	1	0	1	0
CG10971	1	0	1	0	1
CG11006	2	2	0	1	0
CG11050	1	0	1	0	1
CG11077	1	0	1	0	1
CG11122	2	1	1	1	1
CG11146	2	2	0	2	0
CG1115	3	2	1	2	1
CG11151	2	0	2	0	1
CG11180	1	0	1	0	1
CG11188	1	1	0	1	0
CG11198	5	4	1	4	1
CG11242	1	0	1	0	1
CG11284	1	0	1	0	1
CG11490	1	1	0	1	0
CG11498	2	2	0	2	0
CG11501	1	0	1	0	1
CG11526	3	0	3	0	3
CG11534	1	1	0	1	0
CG11710	1	0	1	0	1
CG11771	1	0	1	0	1
CG11848	1	0	1	0	1
CG11943	2	0	2	0	2
CG11963	2	1	1	1	1
CG11967	1	0	1	0	1
CG11968	1	1	0	1	0
CG12016	8	5	3	5	3
CG12017	1	1	0	1	0
CG12024	1	0	1	0	1
CG12091	1	0	1	0	1
CG12224	1	0	1	0	1
CG12340	1	0	1	0	1
CG12367	1	0	1	0	1
CG12393	1	1	0	1	0
CG1244	3	3	0	2	0
CG12581	1	0	1	0	1

CG12581	1	0	1	0	1
CG12773	1	0	1	0	1
CG13124	1	0	1	0	1
CG13130	18	15	3	2	3
CG13253	1	1	0	1	0
CG1332	1	0	1	0	1
CG13445	3	3	0	1	0
CG1358	846	581	265	3	5
CG13585	1	1	0	1	0
CG13670	1	1	0	1	0
CG13907	1	0	1	0	1
CG14033	1	1	0	1	0
CG14235	1	0	1	0	1
CG14342	7	7	0	1	0
CG14478	1	1	0	1	0
CG14480	1	0	1	0	1
CG14561	1	1	0	1	0
CG14567	2	1	1	1	1
CG14646	1	0	1	0	1
CG14799	88	78	10	3	3
CG1486	1	0	1	0	1
CG14880	1	0	1	0	1
CG14906	1	1	0	1	0
CG14907	1	1	0	1	0
CG14956	1	0	1	0	1
CG14967	1	0	1	0	1
CG14982	1	0	1	0	1
CG15019	1	1	0	1	0
CG15067	2	1	1	1	1
CG15099	2	1	1	1	1
CG15105	1	0	1	0	1
CG15118	1	0	1	0	1
CG15134	1	0	1	0	1
CG1516	1	0	1	0	1
CG15203	1	0	1	0	1
CG15209	3	2	1	1	1
CG15240	4	2	2	1	2
CG15370	844	586	258	5	3
CG15418	1	0	1	0	1
CG15465	51	41	10	4	1
CG15482	1	0	1	0	1
CG15529	1	0	1	0	1
CG15609	1	0	1	0	1
CG15706	1	0	1	0	1
CG15725	18	18	0	6	0
CG15771	56	56	0	6	0
CG1578	1	0	1	0	1
CG15828	2	1	1	1	1
CG15930	856	272	584	5	5
CG1599	1	0	1	0	1
CG1628	1	1	0	1	0
CG1637	1	1	0	1	0
CG1638	1	0	1	0	1
CG1662	1	1	0	1	0
CG1665	1	1	0	1	0
CG16972	1	1	0	1	0
CG17065	1	1	0	1	0
CG17108	1	1	0	1	0
CG17264	1	0	1	0	1
CG17528	1	0	1	0	1
CG1753	1	1	0	1	0
CG17838	1	1	0	1	0
CG18107	4	1	3	1	3
CG1812	1	1	0	1	0
CG18135	1	0	1	0	1
CG18208	1	1	0	1	0
CG18262	1	1	0	1	0
CG18787	1	0	1	0	1
CG18809	1	0	1	0	1
CG1882	2	2	0	2	0
CG18854	141	139	2	2	39
CG18870	1	1	0	1	0
CG1893	1	1	0	1	0
CG1998	8	0	8	0	4
CG2061	1	0	1	0	1
CG2083	1	0	1	0	1
CG2093	1	0	1	0	1
CG2124	1	0	1	0	1
CG2165	1	0	1	0	1

CG2182	1	1	0	1	0
CG2186	12	12	0	4	0
CG2233	2	0	2	0	2
CG2519	1	1	0	1	0
Cg25C	1	0	1	0	1
CG2604	1	0	1	0	1
CG2807	1	0	1	0	1
CG2989	1	1	0	1	0
CG30035	1	0	1	0	1
CG3011	2	0	2	0	2
CG31116	1	1	0	1	0
CG31121	847	265	582	5	4
CG31150	1	1	0	1	0
CG31163	1	1	0	1	0
CG31284	1	1	0	1	0
CG31461	10	8	2	3	2
CG31549	1	1	0	1	0
CG31771	175	95	80	4	4
CG31790	770	563	207	4	3
CG31865	2	2	0	2	0
CG32017	1	0	1	0	1
CG32048	1	0	1	0	1
CG32075	1	0	1	0	1
CG32164	1	1	0	1	0
CG32165	1	1	0	1	0
CG32170	1	1	0	1	0
CG32306	1	0	1	0	1
CG32442	1	0	1	0	1
CG32521	1	0	1	0	1
CG32667	2	0	2	0	2
CG32676	1	1	0	1	0
CG32685	1	0	1	0	1
CG32694	788	205	583	4	4
CG3270	1	1	0	1	0
CG32758	1	0	1	0	1
CG3279	1	1	0	1	0
CG3308	3	2	1	2	1
CG33080	4	1	3	1	2
CG33080	1	0	1	0	1
CG33097	1	0	1	0	1
CG33138	1	0	1	0	1
CG33144	1	0	1	0	1
CG3332	8	8	0	3	0
CG33470	1	0	1	0	1
CG33472	1	1	0	1	0
CG33523	1	0	1	0	1
CG3368	1	0	1	0	1
CG33969	2	1	1	1	1
CG33981	1	0	1	0	1
CG34136	1	0	1	0	1
CG34179	1	0	1	0	1
CG34260	1	0	1	0	1
CG34268	1	1	0	1	0
CG34335	1	0	1	0	1
CG34360	2	2	0	2	0
CG34398	146	79	67	4	4
CG34417	1	0	1	0	1
CG34422	1	0	1	0	1
CG3448	1	0	1	0	1
CG3523	2	2	0	2	0
CG3529	1	1	0	1	0
CG3585	3	1	2	1	2
CG3597	2	1	1	1	1
CG3829	1	0	1	0	1
CG4000	3	1	2	1	2
CG40084	3	3	0	3	0
CG40182	1	1	0	1	0
CG40271	1	1	0	1	0
CG40339	1	1	0	1	0
CG40351	4	1	3	1	3
CG4068	65	0	65	0	15
CG40793	1	0	1	0	1
CG40798	1	1	0	1	0
CG41053	1	1	0	1	0
CG41126	2	2	0	2	0
CG41332	1	1	0	1	0
CG41484	2	1	1	1	1
CG41557	1	1	0	1	0
CG41560	2	2	0	2	0

CG41573	2	2	0	2	0
CG41574	1	1	0	1	0
CG41579	1	1	0	1	0
CG41584	1	1	0	1	0
CG41587	1	1	0	1	0
CG41592	1	1	0	1	0
CG4169	1	1	0	1	0
CG4186	1	1	0	1	0
CG4278	1	1	0	1	0
CG4500	1	1	0	1	0
CG4607	1	0	1	0	1
CG4629	1	1	0	1	0
CG4655	65	64	1	3	1
CG4658	1	0	1	0	1
CG4662	2	2	0	2	0
CG4673	1	0	1	0	1
CG4688	1	0	1	0	1
CG4699	1	0	1	0	1
CG4756	2	1	1	1	1
CG4756	1	1	0	1	0
CG4769	1	0	1	0	1
CG4825	2	1	1	1	1
CG4927	1	0	1	0	1
CG5044	1	1	0	1	0
CG5270	1	0	1	0	1
CG5273	1	0	1	0	1
CG5315	1	1	0	1	0
CG5455	1	0	1	0	1
CG5508	1	0	1	0	1
CG5621	1	1	0	1	0
CG5644	1	0	1	0	1
CG5691	1	0	1	0	1
CG5728	1	1	0	1	0
CG5734	2	1	1	1	1
CG5794	1	0	1	0	1
CG5815	1	1	0	1	0
CG5871	1	1	0	1	0
CG5885	1	0	1	0	1
CG5919	2	0	2	0	2
CG5938	1	1	0	1	0
CG5991	1	1	0	1	0
CG6028	1	1	0	1	0
CG6055	1	1	0	1	0
CG6129	1	1	0	1	0
CG6201	1	0	1	0	1
CG6218	2	1	1	1	1
CG6299	1	1	0	1	0
CG6321	1	0	1	0	1
CG6404	1	0	1	0	1
CG6459	3	1	2	1	2
CG6498	1	1	0	1	0
CG6503	4	1	3	1	3
CG6654	1	1	0	1	0
CG6749	1	0	1	0	1
CG6762	1	1	0	1	0
CG6770	1	1	0	1	0
CG6808	1	0	1	0	1
CG6879	1	1	0	1	0
CG7156	1	1	0	1	0
CG7326	1	1	0	1	0
CG7376	1	1	0	1	0
CG7414	1	1	0	1	0
CG7518	1	1	0	1	0
CG7739	7	4	3	3	3
CG7766	1	0	1	0	1
CG7781	1	0	1	0	1
CG7839	10	8	2	3	2
CG7884	1	1	0	1	0
CG7888	1	1	0	1	0
CG7920	1	1	0	1	0
CG7998	1	0	1	0	1
CG8008	2	0	2	0	2
CG8058	1	1	0	1	0
CG8112	2	0	2	0	2
CG8199	2	1	1	1	1
CG8289	1	0	1	0	1
CG8311	1	0	1	0	1
CG8312	1	0	1	0	1
CG8451	2	1	1	2	0

CG8455	1	0	1	0	1
CG8500	3	1	2	1	2
CG8545	846	265	582	5	3
CG8549	1	1	0	1	0
CG8745	1	0	1	0	1
CG8798	1	1	0	1	0
CG8862	1	0	1	0	1
CG9005	2	2	0	1	0
CG9062	1	0	1	0	1
CG9062	2	2	0	2	0
CG9132	1	0	1	0	1
CG9170	4	4	0	2	0
CG9216	1	0	1	0	1
CG9281	1	1	0	1	0
CG9311	1	0	1	0	1
CG9318	1	0	1	0	1
CG9339	1	1	0	1	0
CG9393	1	1	0	1	0
CG9425	1	1	0	1	0
CG9485	1	1	0	1	0
CG9485	1	0	1	0	1
CG9512	1	0	1	0	1
CG9526	2	2	0	1	0
CG9619	1	0	1	0	1
CG9629	1	0	1	0	1
CG9666	1	1	0	1	0
CG9674	882	266	616	5	4
CG9779	2	0	2	0	2
CG9780	9	5	4	5	4
CG9865	1	0	1	0	1
CG9894	12	4	8	1	3
CG9906	1	1	0	1	0
CG9914	1	0	1	0	1
CG9915	1	0	1	0	1
CG9934	4	4	0	2	0
CG9935	2	2	0	2	0
CG9935	1	0	1	0	1
CG9941	848	259	589	4	6
CG9945	2	1	1	1	1
CG9986	1	0	1	0	1
Chc	2	2	0	2	0
cher	2	1	1	1	1
CHES-1-like	1	1	0	1	0
cic	1	0	1	0	1
Cks30A	1	1	0	1	0
Cks85A	1	1	0	1	0
Cp1	1	0	1	0	1
cpo	1	0	1	0	1
CRMP	2	2	0	1	0
crq	1	1	0	1	0
Csk	1	0	1	0	1
CSN8	1	0	1	0	1
Cyp1	1	0	1	0	1
Cyp28d1	1	0	1	0	1
Cyp6d5	1	1	0	1	0
Cyp6g1	1	1	0	1	0
Cyp6w1	1	0	1	0	1
Cys	1	1	0	1	0
Cyt-b5-r	8	4	4	3	4
D2R	1	0	1	0	1
Dcr-1	5	3	2	3	2
Deaf1	1	1	0	1	0
Df31	1	0	1	0	1
Dhc64C	2	1	1	1	1
dik	1	0	1	0	1
DNApol-iota	1	1	0	1	0
dnc	1	0	1	0	1
dome	1	0	1	0	1
Dot	1	1	0	1	0
dp	3	2	1	2	1
Dpit47	1	0	1	0	1
drosha	1	0	1	0	1
E(Pc)	1	1	0	1	0
e(y)3	1	1	0	1	0
ect	3	2	1	2	1
Edem1	1	1	0	1	1
Edem2	1	1	0	1	1
Ef1alpha48D	2	1	1	1	1
Ef2b	2	0	2	0	2

EfTuM	1	0	1	0	1
egh	1	0	1	0	1
eIF-4a	1	0	1	0	1
Eip55E	1	0	1	0	1
epsilonTry	1	0	1	0	1
Ets97D	1	0	1	0	1
exba	1	0	1	0	1
exo70	1	0	1	0	1
exo84	1	0	1	0	1
fab1	1	0	1	0	1
faf	3	0	3	0	3
fat-spondin	1	0	1	0	1
fbl	2	1	1	1	1
fh	1	1	0	1	0
Fit1	1	0	1	0	1
Flo-2	2	1	1	1	1
for	1	1	0	1	0
form3	1	1	0	1	0
Fps85D	1	0	1	0	1
Fs	2	2	0	2	0
fs(2)ltoPP43	19	18	1	4	1
Fur2	897	631	266	5	5
G9a	1	0	1	0	1
gammaCop	1	1	0	1	0
Gfr	1	1	0	1	0
Ggamma1	1	1	0	1	0
Glycogenin	1	0	1	0	1
GlyP	1	1	0	1	0
gro	1	1	0	1	0
gry	2	1	1	1	1
Gs2	1	0	1	0	1
Gug	11	1	10	1	1
HDAC6	3	2	1	1	1
He	3	3	0	1	0
HERC2	2	0	2	0	2
Hexo1	1	0	1	0	1
His2A (19 loci)	1	1	0	1	0
His2B:CG17949	2	2	0	1	0
His3 (24 loci)	1	1	0	1	0
His3 (23 loci)	1	0	1	0	1
hoe2	1	0	1	0	1
Hr4	845	585	260	6	5
Hs3st-A	113	67	46	3	4
Hsp70Aa	1	0	1	0	1
Hsp70Ab	1	0	1	0	1
Hsp70Ba	1	0	1	0	1
Hsp70Bb	1	0	1	0	1
Hsp70Bbb	1	0	1	0	1
Hsp70Bc	1	0	1	0	1
Hsp83	3	1	2	1	2
htt	882	300	582	8	4
IM10	1	0	1	0	1
InR	2	2	0	2	0
IP3K1	3	1	2	1	2
Irp-1B	1	0	1	0	1
JTBR	1	1	0	1	0
Kap-alpha3	1	0	1	0	1
katanin-60	5	3	2	2	2
KrT95D	1	0	1	0	1
kst	1	0	1	0	1
kuk	1	0	1	0	1
l(1)G0469	1	1	0	1	0
l(2)01810	1	0	1	0	1
l(2)gl	8	4	4	4	4
l(3)02640	1	0	1	0	1
l(3)04053	1	1	0	1	0
l(3)73Ah	1	1	0	1	0
l(3)j2D3	2	2	0	1	0
Lam	1	1	0	1	0
ldlCp	2	1	1	1	1
Lmpt	1	0	1	0	1
lolal	1	0	1	0	1
Lsd-1	2	0	2	0	2
Lsm11	1	0	1	0	1
lva	1	0	1	0	1
LysS	1	0	1	0	1
mask	4	4	0	2	0
mbl	1	0	1	0	1
MED21	3	2	1	2	1

Mekk1	1	0	1	0	1
Mes-4	1	0	1	0	1
milt	1	1	0	1	0
Mio	1	0	1	0	1
Mitf	1	1	0	1	0
Mkp3	1	0	1	0	1
mld	1	1	0	1	0
Mmp1	1	0	1	0	1
mnb	2	2	0	2	0
Mnt	1	0	1	0	1
MP1	3	3	0	3	0
mRpL1	1	0	1	0	1
mRpL35	1	0	1	0	1
mrt	1	0	1	0	1
msl-3	2	0	2	0	2
Msp-300	2	1	1	1	1
mt:ATPase6	1	0	1	0	1
mt:CoI	6	0	6	0	6
mt:CoII	1	0	1	0	1
mt:CoIII	4	0	4	0	4
mt:Cyt-b	2	0	2	0	2
mt:ND1	1	0	1	0	1
mt:ND3	1	0	1	0	1
mt:ND5	1	0	1	0	1
mt:ND5	1	0	1	0	1
mts	1	0	1	0	1
mus309	1	1	0	1	0
Nc73EF	1	0	1	0	1
nct	1	0	1	0	1
nej	1	0	1	0	1
Ngp	1	1	0	1	0
ninaB	1	0	1	0	1
ninaC	1	0	1	0	1
ninaE	2	0	2	0	1
Nipped-A	1	1	0	1	0
Nrx-1	2	0	2	0	2
Ntl	7	0	7	0	5
OstStt3	1	0	1	0	1
ovo	10	9	1	1	1
p47	1	1	0	1	0
Patj	1	0	1	0	1
pbl	3	2	1	2	1
Pc	2	1	1	1	1
PebIII	1	0	1	0	1
Pect	3	2	1	2	1
Pepck	1	1	0	1	0
Pgant35A	1	1	0	1	0
pgant6	1	0	1	0	1
Pgi	1	0	1	0	1
Pgk	2	2	0	1	0
Pgm	1	1	0	1	0
Pi3K59F	1	0	1	0	1
pip	1	1	0	1	0
PIP82	876	614	262	8	4
Pitslre	1	0	1	0	1
ple	1	1	0	1	0
plexA	1	0	1	0	1
plx	1	0	1	0	1
pnt	919	650	269	8	5
Pof	1	1	0	1	0
POSH	1	1	0	1	0
Pp2B-14D	1	1	0	1	0
Ppt1	1	1	0	1	0
Prm	1	0	1	0	1
Psf3	1	0	1	0	1
Ptp99A	6	2	4	2	4
Pu	1	0	1	0	1
pyd	1	1	0	1	0
PyK	1	0	1	0	1
r2d2	3	0	3	0	3
Rab6	1	0	1	0	1
RabX6	1	0	1	0	1
rad	1	0	1	0	1
Rad23	1	0	1	0	1
raps	1	0	1	0	1
Rbm13	887	303	582	9	4
Rbp2	15	14	1	7	1
Rbp2	1	0	1	0	1
ref(2)P	1	1	0	1	0

regucalcin	1	0	1	0	1
repo	1	0	1	0	1
Rfabg	3	0	3	0	3
Rho1	2	0	2	0	1
RhoGAP71E	1	0	1	0	1
rin	2	0	2	0	2
r-l	1	1	0	1	0
rols	1	1	0	1	0
RpL19	1	0	1	0	1
RpL28	4	0	4	0	1
RpL31	1	0	1	0	1
RpL35A	1	0	1	0	1
RpL38	1	0	1	0	1
RpL4	1	0	1	0	1
RpS18	1	0	1	0	1
RpS19a	2	1	1	1	1
RpS26	1	0	1	0	1
RpS6	1	0	1	0	1
RpS8	1	0	1	0	1
Rpt4	1	1	0	1	0
rut	1	1	0	1	0
Rya-r44F	1	0	1	0	1
sano	63	61	2	2	2
Sap-r	1	0	1	0	1
Sara	1	0	1	0	1
sdt	1	1	0	1	0
SelR	1	0	1	0	1
sens	843	262	581	4	3
SerT	3	3	0	3	0
sesB	1	0	1	0	1
sev	2	0	2	0	2
shot	1	1	0	1	0
sif	1	1	0	1	0
Sk2	1	0	1	0	1
slik	1	0	1	0	1
sls	1	0	1	0	1
smg	2	1	1	1	1
Smr	38	36	2	10	2
Sox14	1	1	0	1	0
Spase18-21	1	1	0	1	0
spen	2	0	2	0	2
Sra-1	2	1	1	1	1
sta	1	0	1	0	1
StlP	1	1	0	1	0
Strn-Mlck	1	0	1	0	1
sty	1	1	0	1	0
svr	1	1	0	1	0
Syx17	1	0	1	0	1
Syx18	1	0	1	0	1
T48	1	1	0	1	0
tafazzin	2	1	1	1	1
Tango7	1	0	1	0	1
tara	1	0	1	0	1
Tcp-1eta	1	0	1	0	1
TfIIFbeta	1	1	0	1	0
th	4	2	2	2	2
Thiolase	1	0	1	0	1
Thor	2	0	2	0	1
Tis11	2	1	1	1	1
tkv	2	1	1	1	1
tnc	1	1	0	1	0
tra	1	0	1	0	1
trbl	1	0	1	0	1
trk	1	0	1	0	1
Trxr-1	1	0	1	0	1
Tsf1	4	1	3	1	3
Tsp42Er	1	0	1	0	1
Tsp5D	2	2	0	1	0
tst	1	1	0	1	0
TwdlT	1	0	1	0	1
UbcD2	1	0	1	0	1
Ubc-E2H	1	0	1	0	1
Ubp64E	1	0	1	0	1
Unc-89	1	1	0	1	0
up	2	1	1	1	1
Usp36	1	0	1	0	1
Vap-33-1	1	0	1	0	1
vav	1	1	0	1	0
Vha100-1	1	0	1	0	1

Vha26	1	1	0	1	0
Vha68-2	1	0	1	0	1
Vhl	3	1	2	1	2
vir-1	1	0	1	0	1
vn	1	1	0	1	0
wmd	1	1	0	1	0
wuho	1	0	1	0	1
Xbp1	1	0	1	0	1
yip7	1	0	1	0	1
Yippee	1	0	1	0	1
yl	1	0	1	0	1
Yp1	5	0	5	0	5
Yp2	3	1	2	1	2
Yp3	2	0	2	0	2
zf30C	1	1	0	1	0
zfh2	1	0	1	0	1

Table II-S2. Endogenous siRNAs map to transposons. Percentages total more than 100, because some siRNAs map to more than one transposon. Red, LTR retrotransposons; green, non-LTR retrotransposons, blue, DNA transposons.

Table II-S2.

S2 cells (36,958 reads excluding pre-miRNA matching)					
Transposon	Sense siRNAs		Antisense siRNAs		Total number of siRNAs
	Number of siRNAs	% of total siRNAs	Number of siRNAs	% of total siRNAs	
297	10,918	29.54	10,833	29.31	21,751
1731	7,887	21.34	6,490	17.56	14,377
mdg1	4,565	12.35	5,156	13.95	7,968
roo	3,101	8.39	4,023	10.89	6,745
Doc	1,794	4.85	1,999	5.41	3,793
blood	1,810	4.90	1,952	5.28	3,762
INE-1	1,194	3.23	1,306	3.53	2,476
diver	1,037	2.81	1,126	3.05	2,163
mdg3	569	1.54	914	2.47	1,483
Cr1a	804	2.18	402	1.09	1,183
jockey	593	1.60	565	1.53	1,158
S	490	1.33	518	1.40	999
Juan	508	1.37	480	1.30	988
copia	615	1.66	246	0.67	861
Tirant	308	0.83	380	1.03	688
17.6	237	0.64	400	1.08	637
Quasimodo	383	1.04	236	0.64	597
3S18	264	0.71	245	0.66	509
transib1	242	0.65	256	0.69	498
F	202	0.55	283	0.77	403
Stalker2	293	0.79	313	0.85	332
gypsy12	200	0.54	125	0.34	325
micropia	161	0.44	163	0.44	324
HB	144	0.39	167	0.45	311
Dm88	142	0.38	142	0.38	284
Stalker4	29	0.08	151	0.41	180
Rt1b	90	0.24	84	0.23	171
flea	63	0.17	59	0.16	122
Transpac	59	0.16	48	0.13	107
Ivk	72	0.19	34	0.09	102
transib3	26	0.07	66	0.18	92
diver2	70	0.19	14	0.04	84
Burdock	45	0.12	35	0.09	80
rooA	24	0.06	50	0.14	74
gypsy2	48	0.13	25	0.07	73
invader1	73	0.20	73	0.20	73
Stalker	17	0.05	52	0.14	69
McClintock	1	0.00	62	0.17	63
NOF	14	0.04	48	0.13	62
gypsy8	44	0.12	13	0.04	57
1360	42	0.11	34	0.09	46
412	8	0.02	32	0.09	40
ninja-Dsim-like	19	0.05	21	0.06	40
jockey2	25	0.07	9	0.02	32
HMS-Beagle	24	0.06	5	0.01	29
Fw2	24	0.06	3	0.01	27
gypsy10	23	0.06	4	0.01	27
gypsy4	11	0.03	16	0.04	27
gypsy6	14	0.04	13	0.04	27

HeT-A	16	0.04	25	0.07	27
FB	23	0.06	10	0.03	23
gypsy	11	0.03	10	0.03	21
opus	21	0.06	17	0.05	21
G	18	0.05	2	0.01	20
G3	7	0.02	9	0.02	16
Rt1c	7	0.02	9	0.02	16
R1-element	2	0.01	12	0.03	14
Tabor	4	0.01	9	0.02	13
gypsy11	0	0.00	12	0.03	12
Fw3	9	0.02	7	0.02	11
Idefix	7	0.02	4	0.01	11
G4	4	0.01	6	0.02	10
Max	6	0.02	8	0.02	10
GATE	2	0.01	7	0.02	9
TART	7	0.02	1	0.00	8
baggins	7	0.02	0	0.00	7
G5A	2	0.01	4	0.01	6
S2	5	0.01	0	0.00	5
looper1	4	0.01	2	0.01	4
gypsy3	2	0.01	1	0.00	3
invader3	2	0.01	1	0.00	3
invader4	2	0.01	1	0.00	3
rover	0	0.00	3	0.01	3
springer	2	0.01	1	0.00	3
frogger	0	0.00	2	0.01	2
accord	1	0.00	0	0.00	1
I	0	0.00	1	0.00	1
invader2	0	0.00	1	0.00	1
pogo	1	0.00	0	0.00	1

intergenic	1,606	4.35	1,406	3.80	2,817
unannotated	N/A	N/A	N/A	N/A	1,715
mRNA not transposon	N/A	N/A	N/A	N/A	1,261
mRNA & transposon	3,247	8.79	3,021	8.17	4,597

Fly Heads (5,600 reads excluding pre-miRNA matching)					
Transposon	Sense siRNAs		Antisense siRNAs		Total number of siRNAs
	Number of siRNAs	% of total siRNAs	Number of siRNAs	% of total siRNAs	
mdg1	533	10.13	540	10.26	720
roo	350	6.65	338	6.42	571
297	185	3.52	189	3.59	374
jockey	72	1.37	112	2.13	184
F	92	1.75	89	1.69	137
Cr1a	64	1.22	58	1.10	119
INE-1	73	1.39	40	0.76	110
Stalker2	65	1.24	61	1.16	100
gypsy12	47	0.89	53	1.01	99
Doc	49	0.93	37	0.70	86
HB	41	0.78	44	0.84	85
Ivk	39	0.74	61	1.16	81
Rt1b	34	0.65	45	0.86	78
Stalker4	40	0.76	31	0.59	71
opus	59	1.12	61	1.16	65
diver2	25	0.48	35	0.67	60
transib3	37	0.70	16	0.30	53
gypsy2	26	0.49	20	0.38	45
blood	17	0.32	26	0.49	43
invader1	41	0.78	41	0.78	43
gypsy6	21	0.40	16	0.30	37
gypsy	19	0.36	16	0.30	35
rooA	9	0.17	23	0.44	32
FB	30	0.57	29	0.55	30
accord2	17	0.32	12	0.23	29
jockey2	20	0.38	10	0.19	26
Stalker	5	0.10	19	0.36	24
NOF	5	0.10	18	0.34	23
gypsy8	14	0.27	4	0.08	18
1360 (hoppel)	7	0.13	13	0.25	16
Max	7	0.13	10	0.19	16
412	3	0.06	11	0.21	14
GATE	7	0.13	7	0.13	14
gypsy3	6	0.11	8	0.15	14
springer	6	0.11	8	0.15	14
Burdock	5	0.10	7	0.13	12
invader3	6	0.11	6	0.11	12
gypsy4	3	0.06	8	0.15	11
Quasimodo	6	0.11	5	0.10	10
R1	5	0.10	5	0.10	10
17.6	6	0.11	3	0.06	9
gypsy10	2	0.04	6	0.11	8
R1-element	3	0.06	5	0.10	8
HMS-Beagle	3	0.06	4	0.08	7
X	7	0.13	5	0.10	7
HeT-A	5	0.10	2	0.04	5
mdg3	2	0.04	3	0.06	5
S	1	0.02	5	0.10	5
copia	3	0.06	1	0.02	4
Tabor	3	0.06	1	0.02	4
Dm88	1	0.02	2	0.04	3
HMS-Beagle2	1	0.02	2	0.04	3

I	2	0.04	1	0.02	3
rover	1	0.02	2	0.04	3
3S18	2	0.04	0	0.00	2
flea	1	0.02	1	0.02	2
G	1	0.02	1	0.02	2
BS	1	0.02	0	0.00	1
Circe	1	0.02	0	0.00	1
Doc2	1	0.02	0	0.00	1
Doc4	1	0.02	0	0.00	1
G2	1	0.02	0	0.00	1
G3	1	0.02	0	0.00	1
G6	1	0.02	0	0.00	1
hopper2	1	0.02	0	0.00	1
invader2	1	0.02	0	0.00	1
invader6	1	0.02	0	0.00	1
looper1	1	0.02	0	0.00	1
McClintock	1	0.02	0	0.00	1
micropia	1	0.02	0	0.00	1
Rt1a	1	0.02	0	0.00	1
Rt1c	1	0.02	0	0.00	1
transib4	1	0.02	0	0.00	1

intergenic	1,727	32.81	1,487	28.25	2,144
unannotated	N/A	N/A	N/A	N/A	1,779
mRNA not transposon	N/A	N/A	N/A	N/A	2,443
mRNA & transposon	2,006	38.12	1,768	33.59	2,441

Table II-S3A. Enodogenous siRNAs from S2 cells were clustered as described by Brennecke et al. (2007), using Drosophila *melanogaster* genome release R5.5 (http://flybase.bio.indiana.edu/).

Table II-S3A.

S2 cells								
Cluster ID	**Chromosome**	**Start**	**End**	**Cluster length (kb)**	**Number of reads mapping uniquely to cluster**	**Number of reads mapping to cluster**	**piRNA cluster I.D. (Brennecke et al., 2007)**	**Cytogenetic location**
1	2L	9,782,623	9,795,136	13	517	517		30C9-30D1
2	3RHet	782,889	796,491	14	159	897		
3	2L	2,898,870	2,913,985	15	128	128		23C2
4	2L	22,752,903	22,780,367	27	65	619		
5	2L	1,655,404	1,717,432	62	50	4845		22A6-22B1
6	2L	21,079,751	21,094,168	14	50	50		39A1
7	3L	15,547,096	15,559,889	13	45	53		71E1
8	U	5,762,659	5,775,688	13	42	106	cluster #10	
9	2L	13,178,621	13,215,680	37	41	41		34A8-34A10
10	3L	645,955	657,148	11	40	40		61C8
11	2L	7,967,622	7,988,787	21	39	4930		28D3
12	2R	8,459,176	8,469,174	10	37	37		49B5-49B6
13	2L	7,073,818	7,084,359	11	36	36		27E1
14	2L	103,176	123,592	20	35	35		21B2
16	3L	3,192,342	3,242,225	50	35	35		63B11-63C1
15	2R	7,478,837	7,491,292	12	35	35		48A3
17	X	17,983,251	17,995,197	12	32	32		16F6
18	2L	447,944	482,135	34	31	31		21C2-21D1
19	2L	1,153,981	1,164,214	10	31	31		21F1
20	3R	26,182,009	26,195,088	13	30	30		99F1-99F2
21	U	9,199,049	9,230,523	31	29	2917		
22	3L	824,291	870,587	46	28	958		61D2
23	2R	2,229,785	2,243,731	14	27	15206	cluster #1	42A15-42A16
24	3R	19,551,888	19,606,927	55	26	26		95B1-95B5
25	3R	16,891,416	16,901,947	11	24	24		93B9-93B10
26	2L	8,195,702	8,225,201	30	23	23		28F1-28F4
27	2R	12,892,015	12,902,518	11	23	23		53E4
28	2R	1,897,692	1,907,690	10	22	22		42A7-42A8
29	3R	5,591,241	5,634,025	43	22	22		85E8-85E10
31	X	1,346,902	1,378,364	31	21	21		2A3-2B1
30	3R	11,159,014	11,191,374	32	21	21		88F1
33	X	1,956,119	1,972,769	17	20	20		2C8-2C10
32	3L	22,860,546	22,871,373	11	20	20		80B1-80B2
34	2L	3,014,084	3,028,983	15	19	19		23C4
39	3R	12,063,559	12,099,214	36	19	19		89B9
35	2L	18,675,544	18,705,832	30	19	19		36F10-37A1
37	3L	20,378,235	20,389,706	11	19	19		77B9-77C1
38	3L	20,473,055	20,511,399	38	19	20		77C4-77C6
36	2R	20,554,111	20,563,855	10	19	19		60D13
40	3L	14,008,142	14,049,713	42	18	18		70C12-70D1
44	U	4,011,232	4,031,670	20	17	96	cluster #7	
43	3R	5,378,423	5,391,934	14	17	17		85D24-85D25
41	2R	7,162,904	7,177,380	14	17	17		47E5-47F1
42	2R	11,876,244	11,887,693	11	17	17		52D9-52D11
46	3R	27,620	57,331	30	16	621		81F6-82A1
48	X	2,491,408	2,503,866	12	16	16		3A6-3A7
45	2R	17,026,606	17,037,555	11	16	16		57C3-57C4
47	3R	17,091,519	17,102,609	11	16	16		93D2
49	3L	3,317,306	3,327,304	10	15	15		63D1
54	U	1,130,212	1,149,062	19	14	11543		
50	3L	4,246,406	4,256,404	10	14	14		64A11-64A12
51	3L	5,798,359	5,808,357	10	14	14		64E11-64E13
53	3R	17,042,277	17,052,592	10	14	14		93C6-93C7
52	3L	21,615,234	21,625,232	10	14	14		78E1
56	3L	2,592,932	2,603,319	10	13	13		62E7
55	2L	4,914,510	4,954,666	40	13	671		25B3-25B4
59	X	6,179,215	6,191,365	12	13	13		5E5-5E6
60	X	9,084,020	9,094,018	10	13	13		8C17-8D1
57	3R	9,500,740	9,510,738	10	13	13		87F7-87F10
58	3R	25,568,106	25,586,472	18	13	14		99B9-99B10
63	3L	1,502,401	1,521,931	20	12	690		62A3
67	X	1,558,009	1,575,576	18	12	12		2B5-2B6
64	3L	9,619,817	9,631,371	12	12	12		67C2-67C3
66	U	9,763,582	9,778,881	15	12	4825		
61	2L	10,200,298	10,210,296	10	12	12		31B1
62	2L	22,127,508	22,137,506	10	12	16		40E4-40E5
65	3R	25,621,570	25,634,567	13	12	12		99C1-99C2
72	3L Het	780,345	791,575	11	11	3535		
68	2L	5,520,595	5,530,593	10	11	11		25E5
70	3L	6,164,197	6,176,924	13	11	11		65A7

73	3R	7,458,172	7,468,170	10	11	11		86E13
71	3L	8,184,373	8,194,371	10	11	11		66B11-66B12
69	2L	12,103,912	12,113,910	10	11	11		33C4
74	3R	12,908,322	12,924,213	16	11	11		89E12-89E13
75	3L	686,750	696,889	10	10	10		61C8
78	3R	1,293,230	1,303,419	10	10	10		83A4-83A5
79	3R	14,481,371	14,491,369	10	10	10		91B8
81	X	19,512,463	19,522,461	10	10	10		18D7-18D8
76	3L	19,596,487	19,606,485	10	10	10		76B9
77	3L	22,056,871	22,066,869	10	10	10		79B2
82	X	22,345,871	22,392,860	47	10	1258		20F2-20F3
80	3R	25,303,931	25,313,929	10	10	10		99B1
83	2L	192,335	202,333	10	9	9		21B4
96	3L Het	248,455	258,453	10	9	249	cluster #15	
84	2L	1,975,628	1,985,626	10	9	9		22B8
91	3L	3,147,736	3,157,734	10	9	9		63B6-63B7
87	2R	4,046,783	4,056,781	10	9	9		44B5-44B8
98	X	5,207,694	5,217,692	10	9	12		4F4-4F5
92	3L	5,748,568	5,758,566	10	9	9		64E5-64E6
93	3L	11,698,935	11,708,933	10	9	9		68D2-68D3
88	2R	11,817,707	11,827,705	10	9	9		52D2-52D3
89	2R	13,120,411	13,131,234	11	9	1004		54B1
90	2R	15,369,295	15,379,293	10	9	9		56D11-56D13
94	3L	15,598,077	15,608,075	10	9	9		71E2-71E3
85	2L	20,059,341	20,069,339	10	9	9		38B1-38B2
86	2L	20,652,091	20,662,089	10	9	9		38D2-38D3
95	3L	22,933,272	22,943,270	10	9	9		80C1
97	3R	26,028,924	26,038,922	10	9	9		99E2
109	X	831,709	841,707	10	8	8		1D2-1D3
108	4	1,218,726	1,228,724	10	8	8		102F8
103	3R	1,403,421	1,413,419	10	8	8		83B2-83B3
110	X	1,809,962	1,819,960	10	8	8		2B15-2B16
100	2R	3,685,613	3,693,468	8	8	8		43E17-43E18
111	X	4,810,933	4,826,291	15	8	7978		4D5-4D7
99	2L	5,041,556	5,051,554	10	8	8		25C1-25C3
104	3R	5,508,105	5,524,570	16	8	22		85E4
101	2R	7,780,030	7,790,028	10	8	8		48C5
105	3R	12,008,921	12,018,919	10	8	8		89B7
106	3R	16,927,742	16,937,740	10	8	8		93B12-93B13
102	3L	20,821,682	20,831,680	10	8	8		77F1
107	3R	21,149,905	21,159,903	10	8	98		96D1
112	2L	146,778	156,776	10	7	7		21B3
113	2L	2,560,743	2,586,937	26	7	4649		22F4-23A1
116	3R	4,058,120	4,068,025	10	7	7		84F4-84F5
115	3L	7,708,825	7,718,481	10	7	7		66A10
120	X	7,838,431	7,844,562	6	7	7		7C9-7D1
117	3R	10,142,716	10,152,714	10	7	13		88B3-88B4
114	2R	13,424,091	13,434,089	10	7	7		54C3
121	X	19,632,553	19,642,551	10	7	10		18E5-18F1
118	3R	24,141,329	24,151,327	10	7	7		98C3
119	3R	24,710,511	24,720,509	10	7	7		98F1-98F2
135	3R	229,824	261,348	32	6	983		82A6-82B1
122	2L	542,059	552,057	10	6	6		21E2
134	3L Het	563,098	600,561	37	6	879		
127	2R	666,812	675,630	9	6	6		41C2
136	3R	1,459,740	1,467,999	8	6	6		83B7
123	2L	4,986,739	4,996,737	10	6	6		25B9-25B10
137	3R	5,805,160	5,815,158	10	6	6		85F4
124	2L	6,043,300	6,053,298	10	6	6		26B3
138	3R	7,231,157	7,241,155	10	6	6		86E4
130	3L	7,316,278	7,325,381	9	6	6		65F4
131	3L	9,078,428	9,086,564	8	6	6		66F5
125	2L	10,389,264	10,399,262	10	6	6		31D11-31E1
139	3R	11,092,595	11,099,452	7	6	6		88E9-88E10
126	2L	16,307,231	16,317,229	10	6	6		35F1
132	3L	16,450,201	16,460,199	10	6	6		73A1
128	2R	16,549,517	16,559,515	10	6	6		57A9-57A10
140	3R	19,016,516	19,022,739	6	6	6		94E5-94E6
145	X	20,061,920	20,069,591	8	6	6		19C1
129	2R	20,663,857	20,673,855	10	6	6		60E1
141	3R	20,869,191	20,879,189	10	6	6		96B17-96B19
133	3L	20,986,085	21,020,979	35	6	969		78A2
142	3R	22,405,317	22,415,315	10	6	6		97C1
143	3R	22,687,952	22,697,947	10	6	6		97D3
144	3R	27,568,150	27,577,970	10	6	6		100D2
173	3R Het	31,087	74,992	44	5	10922		
174	4	551,915	561,913	10	5	5		102C4
146	2L	2,132,930	2,142,928	10	5	5		22D1
161	3R	2,479,803	2,486,260	6	5	5		84A1
147	2L	2,764,639	2,783,911	19	5	4784		23A5-23A6

154	2R	4,779,425	4,789,423	10	5	5		44F3
148	2L	5,986,382	5,996,380	10	5	5		26B2
162	3R	7,039,531	7,049,529	10	5	5		86D8
163	3R	8,839,679	8,847,105	7	5	5		87D8-87D9
175	X	9,452,424	9,462,422	10	5	5		8E7-8E10
164	3R	9,855,890	9,865,888	10	5	5		88A4
155	2R	10,145,197	10,155,195	10	5	5		5.00E+07
176	X	10,731,645	10,741,643	10	5	5		9F4-9F5
149	2L	11,092,188	11,100,393	8	5	5		32D2-32D3
157	3L	12,131,399	12,140,545	9	5	5		68F5-68F6
158	3L	12,759,827	12,768,208	8	5	5		69E2
165	3R	13,512,031	13,521,424	9	5	9		90C1
150	2L	16,249,303	16,259,301	10	5	5		35F1
159	3L	16,984,985	16,994,983	10	5	5		73E1-73E3
166	3R	18,407,620	18,413,515	6	5	5		94B5
167	3R	18,559,695	18,569,011	9	5	5		94C4
151	2L	19,000,410	19,010,408	10	5	5		37B9
152	2L	19,142,575	19,174,358	32	5	4890		37C1-37C6
156	2R	19,833,985	19,843,983	10	5	5		60A13
160	3L	19,874,250	19,884,248	10	5	5		76D3
177	X	20,257,850	20,267,821	10	5	5		19C5-19C6
168	3R	20,704,520	20,714,518	10	5	5		96B2-96B3
153	2L	21,660,567	21,669,076	9	5	5		39E3-39E6
169	3R	23,765,906	23,775,904	10	5	5		98B6
170	3R	25,507,490	25,515,528	8	5	5		99B7
171	3R	25,816,869	25,826,867	10	5	5		99D1
172	3R	26,303,191	26,319,848	17	5	128		99F6

Table II-S3B. siRNAs from fly heads were clustered as described by Brennecke et al. (2007), using *Drosophila melanogaster* genome release R5.5.

Table II-S3B.

WT Heads								
Cluster ID	**Chromosome**	**Start**	**End**	**Cluster length (kb)**	**Number of reads map uniquely to cluster**	**Number of reads mapping to cluster**	**piRNA cluster I.D. (Brennecke et al., 2007)**	**Cytogenetic location**
1	2L	9783876	9795136	11 3	478	478		30C9-30D1
173	3L	886,261	896,260	10.0	14	14		61D3-61D4
174	2L	6,855	17 067	10.2	11	11		21A5
175	X	9,940,973	9,953,050	12.1	10	11		9A5-9B1
176	3R	113,708	123,706	10.0	9	9		82A1
49	3L	3,317,197	3,327,189	10.0	8	8		63D1
111	X	4,811,216	4,826,291	15.1	8	1,771		4D5-4D7
177	2L	7,706,540	7,716,536	10.0	8	8		28C1
7	3L	15,549,041	15,558,952	9.9	8	13		71E1
178	2L	16,784,804	16,794,788	10.0	8	8		36B1
179	3R	6,665	15,118	8.5	7	8		81F6
180	2R	14,267,508	14,277,167	9.7	6	6		55C4
181	X	21,604,591	21,614,589	10.0	6	72	cluster #8	20B1
80	3R	25,305,992	25,315,848	9.9	6	6		99B1
182	3R	1,048,181	1,058,035	9.9	5	5		82F6
183	2L	9,817,453	9,827,451	10.0	5	156		30D1
184	3L	10,687,581	10,697,585	10.0	5	926		67E7

Table II-S3C. piRNA data from Brennecke et al. (2007) were clustered according using *Drosophila melanogaster* genome.

Table II-S3C.

piRNAs (from Brennecke et al., 2007)						
Chromosome	**Start**	**End**	**Cluster length (kb)**	**Number of reads mapping uniquely to cluster**	**Number of reads mapping to cluster**	**Cytogenetic location**
2R	2,140,512	2,389,335	249	1,460	19,441	42A14-42B1
X	21,388,081	21,432,231	44	994	7,351	20A1-20A3
2L	20,143,634	20,232,517	89	445	2,540	38C2-38C3
3L	23,269,813	23,313,601	44	224	1,169	80E3-80F1
4	1,255,371	1,351,506	96	202	5,079	102F8
U	4,010,984	4,077,966	67	162	822	
X	21,501,319	21,548,357	47	122	2,827	20A5-20B1
U	5,743,150	5,797,646	54	115	3,694	
2R	12,713,990	12,723,988	10	109	109	53D11-53D12
X	15,398,513	15,408,511	10	80	80	13C5-13C7
3LHet	2,008,276	2,212,278	204	70	15,385	
3RHet	2,070,375	2,106,781	36	67	1,066	
3LHet	237,482	330,926	93	61	3,703	
U	7,497,140	7,584,470	87	61	8,578	
3R	6,228,871	6,238,915	10	46	46	86B4
X	21,756,108	21,841,785	86	43	3,377	20B3-20C1
4	807,233	867,379	60	41	464	102E1-102E3
2L	20,100,366	20,123,183	23	40	261	38C1-38C2
2L	22,342,790	22,421,219	78	35	3,508	40F7
2L	1	11,667	12	33	20,448	21A5
3LHet	148,660	204,731	56	33	1,701	
3L	24,088,523	24,134,591	46	33	2,113	
3RHet	2,309,480	2,373,211	64	32	2,425	
X	2,061	26,029	24	31	387	1A1
X	11,076,431	11,099,456	23	31	215	10A10-10B1
X	21,580,417	21,687,831	107	31	2,671	20B1
3L	23,449,678	23,478,214	29	30	601	80F6-80F7
3LHet	493,948	685,925	192	29	10,795	
3R	21,467,283	21,482,178	15	29	29	96E6-96E7
2LHet	121,252	266,568	145	27	6,692	
3LHet	285	32,970	33	26	1,164	
2R	742,942	782,203	39	26	2,423	41C4-41C5
U	2,433,298	2,478,920	46	26	1,155	
2L	22,945,885	22,989,803	44	26	1,179	
3L	23,940,894	24,045,838	105	26	6,737	
3R	1,279	23,416	22	25	58	81F6
3RHet	1,607,736	1,674,464	67	25	2,263	
2RHet	1,857,936	1,913,095	55	24	966	
3L	19,845,140	19,864,685	20	23	1,530	76D1-76D3
3RHet	104,786	191,198	86	21	2,505	
3LHet	1,402,112	1,458,965	57	21	1,332	
2L	22,486,772	22,547,558	61	21	3,167	40F7
3L	24,465,528	24,543,475	78	21	1,399	
3RHet	617,618	656,530	39	20	1,197	
2RHet	1,412,742	1,489,780	77	20	824	
3RHet	1,746,563	1,797,611	51	20	3,429	
3RHet	532,053	575,335	43	19	1,169	
3RHet	849,568	921,355	72	19	1,662	
U	2,056,878	2,098,213	41	19	2,277	
U	889,267	1,061,441	172	18	6,217	
3RHet	1,111,034	1,223,916	113	18	3,403	
2R	1,253,143	1,284,240	31	18	1,040	41E5-41E6
2L	19,564,519	19,574,923	10	17	32	37F1-37F2
2L	22,254,319	22,281,479	27	17	701	40F7
3L	23,612,866	23,636,896	24	17	752	80F9

X	8,368,544	8,381,781	13	16	16	7F1
3LHet	770,628	819,852	49	15	2,636	
2R	16,466,415	16,476,583	10	15	15	57A6
3LHet	840,924	895,679	55	14	2,243	
3RHet	1,383,668	1,470,543	87	14	2,624	
3LHet	1,479,139	1,528,684	50	14	4,505	
2R	7,777,083	7,787,544	10	14	14	48C5
2L	8,450,213	8,490,832	41	14	2,178	29C5-29D1
2L	16,693,456	16,703,757	10	14	14	36A10-36A11
X	22,369,187	22,403,875	35	14	1,688	20F3
2R	109,239	149,540	40	13	884	
U	141,712	210,336	69	13	2,023	
3LHet	362,237	394,074	32	13	326	
2R	1,216,294	1,227,635	11	13	39	41E5
3LHet	1,844,970	1,901,261	56	13	2,843	
3L	24,350,206	24,375,909	26	13	819	
3R	27,892,332	27,909,797	17	13	11,215	10E4
U	40,427	117,442	77	12	4,287	
4	1,015,921	1,026,279	10	12	55	102F5
2RHet	2,204,696	2,287,166	82	12	1,695	
2RHet	2,788,079	2,857,172	69	12	2,226	
U	5,625,604	5,649,537	24	12	476	
2L	5,954,935	5,984,574	30	12	12	26A3-26B2
2R	3,316,801	3,331,740	15	11	96	43C1
U	3,519,704	3,551,702	32	11	768	
3R	5,921,675	5,931,673	10	11	11	85F10-85F11
U	9,170,572	9,298,799	128	11	7,794	
2RHet	1,679,952	1,715,467	36	10	730	
2R	9,211,947	9,221,945	10	10	10	50A1-50A3
X	10,164,447	10,174,445	10	10	1,391	9B5-9B6
2R	185,439	225,778	40	9	1,909	

2R	845,724	885,372	40	9	1,552	41C6
2RHet	867,578	909,826	42	9	3,623	
U	5,446,117	5,477,034	31	9	2,751	
2L	7,420,980	7,430,978	10	9	9	27F3-27F4
2R	21,136,534	21,151,342	15	9	1,898	60F5
2L	21,891,204	21,901,202	10	9	9	40B3
3L	24,309,487	24,328,647	19	9	415	
3RHet	9,020	19,018	10	8	71	
3LHet	2,376,347	2,446,273	70	8	2,166	
2RHet	2,878,674	2,939,749	61	8	2,021	
2RHet	2,988,025	3,049,062	61	8	1,036	
U	3,876,652	3,943,760	67	8	1,823	
2L	20,631,611	20,640,251	9	8	18	38D1
XHet	169,257	192,176	23	7	919	
2LHet	302,772	369,442	67	7	3,621	
U	339,589	384,771	45	7	1,130	
3LHet	714,299	741,348	27	7	702	
X	1,371,374	1,381,268	10	7	7	2B1
X	4,017,313	4,027,311	10	7	12	4B1
X	5,201,679	5,211,408	10	7	7	4F4-4F5
U	6,643,127	6,660,684	18	7	1,684	
3R	7,044,221	7,053,379	9	7	7	86D8
3L	8,716,961	8,726,803	10	7	7	66D12
2L	9,891,561	9,901,336	10	7	7	30E1
X	12,660,975	12,670,600	10	7	116	11B16-11C1
X	22,096,745	22,116,991	20	7	129	20D2
3R	27,415,958	27,425,954	10	7	7	100C7
2R	410,424	423,625	13	6	536	41A2
X	652,829	662,184	9	6	6	1C4
4	985,305	995,476	10	6	37	102F4
3LHet	989,120	1,049,237	60	6	4,049	

3RHet	1,252,983	1,338,766	86	6	2,064	
U	1,379,079	1,450,515	71	6	3,660	
2RHet	1,597,675	1,641,429	44	6	2,648	
U	1,962,353	1,972,720	10	6	90	
U	3,109,090	3,139,633	31	6	1,015	
U	5,852,441	6,059,636	207	6	14,423	
U	7,836,590	7,882,892	46	6	6,453	
2L	17,968,726	17,978,719	10	6	6	36E3
3L	23,187,909	23,220,679	33	6	562	80D5-80E1
2RHet	483,216	517,466	34	5	258	
2R	514,257	524,320	10	5	202	41B2
4	609,263	619,456	10	5	1,194	102C6-102D1
3L	825,510	833,058	8	5	5	61D2
2RHet	1,087,648	1,133,687	46	5	1,120	
2RHet	1,339,078	1,389,974	51	5	3,038	
3RHet	2,482,651	2,492,649	10	5	62	
3R	2,909,142	2,918,466	9	5	6	84B2-84B6
X	3,435,144	3,445,036	10	5	7	3D5
U	6,191,840	6,261,703	70	5	6,969	
U	7,020,670	7,055,172	35	5	4,662	
2L	7,825,754	7,830,915	5	5	5	28D1-28D2
3L	10,353,382	10,363,380	10	5	5	67E1-67E2
X	11,787,892	11,794,120	6	5	5	10F4
2L	13,405,034	13,416,223	11	5	646	34B10-34B11
X	19,487,663	19,497,511	10	5	5	18D3
X	21,183,210	21,188,753	6	5	12	19F3-19F4
3L	24,169,238	24,179,236	10	5	95	
3L	24,220,571	24,229,777	9	5	17	

Table II-S4. Endogenous siRNAs matching transposons are depleted in *dcr-2* null mutant fly heads. Percentages total more than 100, because some siRNAs map to more than one transposon. Red, LTR retrotransposons; green, non-LTR retrotransposons, blue, DNA transposons. "Fold decrease" was calculated by normalizing the total number of siRNAs matching the transposon in each genotype to the total number of 18–29 nt RNA reads, excluding pre-miRNA-matching reads, a measure of the small RNA sequencing depth. Some siRNAs match more than one transposon, so the sum of the total number of siRNAs for each transposon is greater than the actual number so 21 nt small RNA reads: 2,524 for *dcr-2*/CyO and 263 for *dcr-2* homozygotes. *p*-value was calculated using Fisher's exact test.

Table II-S4.

	dcr-2^{L811fsX}/CyO (25,822 reads, excluding pre-miRNA-matching reads)					*dcr-2*^{L811fsX} (16,917 reads, excluding pre-miRNA-matching reads)						
Transposon	**Sense siRNAs**		**Antisense siRNAs**		**Total number of siRNAs**	**Transposon**	**Sense siRNAs**		**Antisense siRNAs**		**Total number of siRNAs**	**fold decrease**	***p*-value**
	Number of siRNAs	**% of total siRNAs**	**Number of siRNAs**	**% of total siRNAs**			**Number of siRNAs**	**% of total siRNAs**	**Number of siRNAs**	**% of total siRNAs**			
roo	802	71%	897	80%	1,126	*roo*	56	64%	64	74%	87	**8.5**	0.000
jockey	351	47%	400	53%	751	*jockey*	30	63%	18	38%	48	**10.3**	0.000
mdg1	297	43%	396	58%	687	*mdg1*	35	60%	27	47%	58	**7.8**	0.000
297	288	45%	354	55%	642	*297*	43	55%	35	45%	78	**5.4**	0.000
Cr1a	295	60%	202	41%	490	*Cr1a*	26	58%	19	42%	45	**7.1**	0.000
blood	202	42%	281	58%	483	*blood*	20	54%	17	46%	37	**8.6**	0.000
HB	180	40%	270	60%	450	*HB*	11	44%	14	56%	25	**11.8**	0.000
springer	189	42%	256	58%	445	*springer*	16	64%	9	36%	25	**11.7**	0.000
gypsy3	162	42%	221	58%	383	*gypsy3*	14	64%	8	36%	22	**11.4**	0.000
invader3	162	42%	221	58%	383	*invader3*	14	64%	8	36%	22	**11.4**	0.000
Stalker4	118	32%	246	68%	364	*Stalker4*	11	37%	19	63%	30	**7.9**	0.000
Stalker2	228	68%	265	79%	337	*Stalker2*	7	58%	11	92%	12	**18.4**	0.000
F	115	43%	181	68%	268	*F*	7	41%	13	76%	17	**10.3**	0.000
Doc	85	33%	170	66%	258	*Doc*	6	21%	21	75%	28	**6.0**	0.000
gypsy12	156	62%	97	38%	253	*gypsy12*	8	73%	3	27%	11	**15.1**	0.000
Ivk	186	74%	118	47%	251	*Ivk*	14	88%	3	19%	16	**10.3**	0.000
INE-1	78	38%	133	65%	204	*INE-1*	5	25%	16	80%	20	**6.7**	0.000

Stalker	61	32%	128	68%	189	*Stalker*	3	43%	4	57%	7	**17.7**	0.000
transib3	57	31%	125	69%	182	*transib3*	2	13%	13	87%	15	**7.9**	0.000
rooA	59	33%	122	67%	181	*rooA*	2	25%	6	75%	8	**14.8**	0.000
NOF	58	33%	116	66%	176	*NOF*	2	33%	4	67%	6	**19.2**	0.000
Rt1b	56	42%	77	58%	133	*Rt1b*	4	31%	9	69%	13	**6.7**	0.000
FB	105	81%	112	87%	129	*FB*	2	50%	3	75%	4	**21.1**	0.000
412	39	37%	66	63%	105	*412*	0	0%	1	100%	1	**68.8**	0.000
jockey2	69	68%	36	35%	102	*jockey2*	2	67%	0	0%	3	**22.3**	0.000
gypsy8	44	67%	22	33%	66	*gypsy8*	0	0%	1	100%	1	**43.2**	0.000
opus	39	89%	39	89%	44	*opus*	4	57%	7	100%	7	**4.1**	0.000
diver2	24	57%	18	43%	42	*diver2*	3	33%	6	67%	9	**3.1**	0.001
invader1	31	82%	32	84%	38	*invader1*	2	100%	2	100%	2	**12.4**	0.000
Burdock	13	37%	22	63%	35	*Burdock*	5	100%	0	0%	5	**4.6**	0.000
1360	25	76%	17	52%	33	*1360*	7	70%	4	40%	10	**2.2**	0.029
Quasimodo	20	61%	17	52%	33	*Quasimodo*	5	50%	5	50%	10	**2.2**	0.029
gypsy	13	42%	18	58%	31	*gypsy*	2	50%	2	50%	4	**5.1**	0.000
HMS-Beagle	7	24%	22	76%	29	*HMS-Beagle*	3	43%	4	57%	7	**2.7**	0.016
gypsy2	16	67%	8	33%	24	*gypsy2*	1	33%	2	67%	3	**5.2**	0.002
gypsy6	11	46%	13	54%	24	*gypsy6*	0	0%	2	100%	2	**7.9**	0.000
Transpac	12	55%	10	45%	22	*Transpac*	5	100%	0	0%	5	**2.9**	0.029
gypsy4	7	33%	14	67%	21								
17.6	4	20%	16	80%	20	*17.6*	1	25%	3	75%	4	**3.3**	0.022
GATE	8	47%	12	71%	17	*GATE*	2	50%	2	50%	4	**2.8**	0.072
gypsy10	5	29%	12	71%	17	*gypsy10*	2	50%	2	50%	4	**2.8**	0.072
mdg3	5	29%	12	71%	17	*mdg3*	1	8%	12	92%	13	**0.9**	0.711
Dm88	5	31%	11	69%	16	*Dm88*	1	25%	3	75%	4	**2.6**	0.107

S	7	50%	11	79%	14	*S*	4	57%	3	43%	7	**1.3**	0.659
copia	9	69%	4	31%	13	*copia*	10	83%	2	17%	12	**0.7**	0.417
Max	8	62%	8	62%	13								
accord2	3	30%	7	70%	10	*accord2*	0	0%	1	100%	1	**6.6**	0.059
rover	4	40%	6	60%	10								
flea	6	67%	3	33%	9	*flea*	3	100%	0	0%	3	**2.0**	0.385
I	5	63%	3	38%	8								
HeT-A	4	67%	5	83%	6	*HeT-A*	2	40%	3	60%	5	**0.8**	0.762
R1-element	4	67%	2	33%	6	*R1-element*	3	50%	3	50%	6	**0.7**	0.558
X	5	83%	6	100%	6	*X*	1	100%	1	100%	1	**3.9**	0.256
baggins	1	20%	4	80%	5	*baggins*	0	0%	2	100%	2	**1.6**	0.711
G	2	40%	3	60%	5	*G*	1	17%	5	83%	6	**0.5**	0.362
ninja-Dsim-like	2	40%	3	60%	5								
1731	1	25%	3	75%	4	*1731*	9	56%	7	44%	16	**0.2**	0.000
Idefix	1	25%	3	75%	4	*Idefix*	0	0%	1	100%	1	**2.6**	0.654
Rt1a	0	0%	4	100%	4	*Rt1a*	0	0%	1	100%	1	**2.6**	0.654
Tabor	3	75%	1	25%	4	*Tirant*	3	75%	1	25%	4	**0.7**	0.720
frogger	1	33%	2	67%	3								
Juan	2	67%	1	33%	3	*Juan*	0	0%	4	100%	4	**0.5**	0.446
3S18	1	50%	1	50%	2								
Circe	1	50%	1	50%	2								
diver	1	50%	1	50%	2	*diver*	2	67%	1	33%	3	**0.4**	0.391
Fw2	2	100%	0	0%	2								
Fw3	1	50%	1	50%	2								
G3	1	50%	1	50%	2								
gypsy9	0	0%	2	100%	2								

hopper2	0	0%	2	100%	2
invader4	1	50%	1	50%	2
micropia	0	0%	2	100%	2
Rt1c	1	50%	1	50%	2
transib1	0	0%	2	100%	2
accord	0	0%	1	100%	1
BS3	1	100%	0	0%	1
G2	0	0%	1	100%	1
G4	1	100%	0	0%	1
G5	1	100%	0	0%	1
G6	0	0%	1	100%	1
HMS-Beagle2	0	0%	1	100%	1
invader2	0	0%	1	100%	1
invader6	0	0%	1	100%	1
McClintock	0	0%	1	100%	1
transib4	0	0%	1	100%	1

hopper2	0	0%	3	100%	3	**0.4**	0.391
invader4	0	0%	3	100%	3	**0.4**	0.391
micropia	0	0%	1	100%	1	**1.3**	1.000
transib1	0	0%	2	100%	2	**0.7**	0.651
G2	0	0%	3	100%	3	**0.2**	0.308
invader2	1	50%	1	50%	2	**0.3**	0.567
invader6	0	0%	2	100%	2	**0.3**	0.567
transib4	0	0%	2	100%	2	**0.3**	0.567
pogo	1	100%	0	0%	1		
S2	0	0%	1	100%	1		

Table II-S5. The abundance of miRNA-matching reads was unchanged in *dcr-2*^{L811fsX} heads, compared to their heterozygous siblings. Fold change was calculated by normalizing the total number of miRNAs in each genotype to small RNA sequencing depth, i.e., the total number of 18–29 nt RNA reads (688,323 for *dcr-2* homozygotes; 859,436 for heterozygotes).

Table II-S5.

	Total number of reads		Fold change (homozygotes vs heterozygotes)	*p*-value
miRNA	***dcr2*^{L811fsX}/CyO**	***dcr2*^{L811fsX}**		
miR-14	172,360	101,066	0.73	0.000
miR-276a	141,107	99,817	0.88	0.000
miR-8	84,901	56,233	0.83	0.000
miR-317	47,027	41,865	1.11	0.000
miR-277	40,372	36,318	1.12	0.000
miR-34	34,350	59,032	2.15	0.000
miR-276b	21,520	13,092	0.76	0.000
Bantam	17,977	13,336	0.93	0.000
miR-274	16,754	20,625	1.54	0.169
miR-210	16,142	18,754	1.45	1.2E-11
miR-1	14,885	13,926	1.17	8.9E-135
miR-133	12,532	9,555	0.95	1.6E-296
miR-999	12,065	8,549	0.88	0.000
miR-7	11,707	7,085	0.76	0.000
miR-184	11,679	15,992	1.71	1.8E-14
let-7	11,192	14,617	1.63	3.0E-04
miR-33	10,529	6,842	0.81	0.000
miR-9a	10,101	6,985	0.86	0.000
miR-125	9,397	8,268	1.10	8.1E-121
miR-278	6,942	7,849	1.41	1.5E-09
miR-11	6,562	4,849	0.92	1.5E-172
miR-284	5,486	4,089	0.93	2.2E-140
miR-252	5,188	3,911	0.94	8.9E-128
miR-124	4,181	7,615	2.27	6.2E-89
miR-305	3,398	5,428	1.99	5.4E-30
miR-279	3,395	3,441	1.27	6.1E-18
miR-285	3,198	1,781	0.70	3.5E-173
miR-13a	3,111	1,596	0.64	3.0E-196
miR-996	3,012	1,766	0.73	8.3E-147
miR-987	2,915	2,148	0.92	4.5E-78
miR-981	2,682	2,759	1.28	9.4E-13
miR-932	2,593	1,514	0.73	1.0E-127

miR-307	2,496	2,426	1.21	1.5E-18
miR-12	2,386	1,410	0.74	1.1E-114
miR-927	2,365	1,563	0.83	3.4E-87
miR-306	2,299	2,341	1.27	3.8E-12
miR-282	2,167	2,426	1.40	0.0002
miR-957	1,775	1,998	1.41	1.5E-03
miR-965	1,775	685	0.48	4.2E-170
miR-275	1,647	2,493	1.89	1.1E-09
miR-1000	1,493	1,841	1.54	0.727
miR-79	1,421	1,082	0.95	5.6E-35
miR-304	1,382	770	0.70	1.1E-75
miR-1010	1,300	899	0.86	3.6E-43
miR-263b	1,298	761	0.73	2.9E-64
miR-31a	1,227	1,303	1.33	4.8E-05
miR-970	1,188	1,338	1.41	0.0097
miR-219	980	946	1.21	1.8E-08
miR-1003	917	513	0.70	2.5E-50
miR-315	861	591	0.86	1.2E-29
miR-137	851	729	1.07	8.2E-14
miR-9b	844	473	0.70	2.1E-46
miR-1006	813	533	0.82	1.2E-31
miR-986	714	1,053	1.84	0.0006
miR-316	589	657	1.39	0.049
miR-995	570	690	1.51	0.590
miR-263a	562	872	1.94	5.9E-05
miR-1012	543	377	0.87	1.1E-18
miR-1001	531	392	0.92	2.2E-15
miR-998	526	513	1.22	7.2E-05
miR-1017	497	304	0.76	2.4E-23
miR-9c	478	588	1.54	0.829
miR-993	449	395	1.10	4.0E-07
miR-1009	414	234	0.71	3.2E-23
miR-980	336	276	1.03	2.7E-07
miR-929	335	287	1.07	2.7E-06
miR-190	319	518	2.03	0.0002
miR-2a-2	316	348	1.38	0.109
miR-87	266	139	0.65	1.0E-17
miR-1008	248	200	1.01	4.6E-06
miR-375	243	199	1.02	1.0E-05

miR-100	241	224	1.16	0.0015
miR-988	223	150	0.84	3.1E-09
miR-1004	182	135	0.93	4.2E-06
miR-308	166	281	2.11	0.0019
miR-193	116	126	1.36	0.301
miR-2b-2	87	102	1.46	0.714
miR-2b-1	86	90	1.31	0.255
miR-283	85	92	1.35	0.364
miR-2c	81	117	1.80	0.318
miR-1005	62	40	0.81	0.0013
miR-1007	62	51	1.03	0.029
miR-2a-1	60	63	1.31	0.365
miR-958	58	112	2.41	0.0068
miR-10	53	61	1.44	0.706
miR-971	48	56	1.46	0.768
miR-956	34	78	2.86	0.0030
miR-969	34	36	1.32	0.548
miR-311	30	11	0.46	0.0002
miR-314	21	38	2.26	0.191
miR-3	17	15	1.10	0.375
miR-954	17	27	1.98	0.453
miR-310	16	8	0.62	0.038
miR-312	16	13	1.01	0.266
miR-31b	13	17	1.63	1.000
miR-1016	11	16	1.82	0.847
miR-286	10	1	0.12	0.0035
miR-990	10	24	3.00	0.086
miR-318	7	6	1.07	0.582
miR-92b	7	19	3.39	0.078
miR-960	7	2	0.36	0.088
miR-982	7	5	0.89	0.391
miR-966	6	4	0.83	0.356
miR-991	6	2	0.42	0.151
miR-1013	5	6	1.50	1.000
miR-92a	5	5	1.25	0.759
miR-1011	4	4	1.25	1.000
miR-984	4	5	1.56	1.000
miR-309	3	3	1.25	1.000
miR-313	2	1	0.62	0.589

miR-976	2	4	2.50	0.699
miR-977	2	2	1.25	1.000
miR-303	1	6	7.49	0.141
miR-4	1	2	2.50	1.000
miR-959	1	1	1.25	1.000
miR-961	1	3	3.75	0.634
miR-964	1	2	2.50	1.000
miR-973	1	1	1.25	1.000
miR-989	1	2	2.50	1.000
miR-iab4as	1	4	4.99	0.390

Table II-S6. Primers for quantitative RT-PCR.

Table II-S6.

Detects	Forward primer, reverse primer
Gypsy	CCAGGTCGGGCTGTTATAGG, GAACCGGTGTACTCAAGAGC
297	AAAGGGCGCTCATACAAATG, TGTGCACATAAAATGGTTCG
roo	CGTCTGCAATGTACTGGCTCT, CGGCACTCCACTAACTTCTCC
I-element	TGAAATACGGCATACTGCCCCCA, GCTGATAGGGAGTCGGAGCAGATA
mdg1	CACATGTTCTCATTCCCAACC, TTCGCTTTTTATATTTGCGCTAC
jockey	TGCAGTTGTTTCCCCTAACC, AGTTGGGCAAATGCTAGTGG
INE-1	GGCCATGTCCGTCTGTCC, AGCTAGTGTGAATGCGAACG
blood	TGCCACAGTACCTGATTTCG, GATTCGCCTTTTACGTTTGC
S-element	TGAAAAGCGTCATTCATTCG, TGTTTCTAGCGCACTCAACG
Doc	GGGTGACTATAACGCCAAGC, GCAAAATCGATCAGGTCTGG
1731	AGCAAACGTCTGTTGGAAGG, CGACAGCAAAACAACACTGC
F-element	GCTGGTAGATACCGCTGAGG, GTAGTCGTCCTCCGTTTTCG
412	CACCGGTTTGGTCGAAAG, GGACATGCCTGGTATTTTGG
NOF	AGTTGGACCTGGAATTGTGG, AATGCACACGGAAGAGGAAC
Idefix	AACAAAATCGTGGCAGGAAG, TCCATTTTTCGCGTTTACTG
Het-A	CGCGCGGAACCCATCTTCAGA, CGCCGCAGTCGTTTGGTGAGT
dcr-1	GCTAACGATGGCATCAATCTG, GCTTGGAGCGCAGGTGACTTA
dcr-2	GAGCTGCTCCATCAGTTTCA, TCCCAGTCAAAGCATTTCTGT
ago2	CAAGAAAGGAGGACAGGATAGC, TTGTTGCTGATGCGGTTG

Table II-S7. Sequencing statistics. "Small RNA reads" correspond to genome matching reads after excluding annotated non-coding RNAs. 454, pyrosequencing; Solexa/Illumina, sequencing-by-synthesis. An asterisk indicates data that was pooled as described in the legend to Figure 2. Ambiguous: the reads map to the indicated category and another category or in both orientations within a single category.

Table II-S7.

Genotype	Enriched for modified 3´ ends?	Sequencing method	Genome-matching reads	annotated ncRNAs	Total Small RNA reads	All pre-miRNA-matching reads	Annotated miRNAs only
S2 cells	no	454	81,226	16,921	64,207	56,463	47,599
	yes	454	72,012	5,875	66,056	11,014	7,476
IR-wild-type heads*	no	454	94,772	23,206	71,268	68,596	61,688
	yes	454	30,250	1,526	22,690	11,089	8,740
	no	Illumina	1,245,354	33,429	1,187,572	1,152,293	949,190
	yes	Illumina	33,558	2,219	28,344	10,792	8,849
wild-type male heads*	no	Illumina	387,855	15,671	357,300	347,089	304,740
	yes	Illumina	4,928	422	4,208	3,261	2,856
wild-type female heads*	no	Illumina	916,026	43,081	790,126	754,602	673,105
	yes	Illumina	61,748	2,214	54,495	47,231	41,598
dcr-2^{L811fsX}/CyO heads	no	Illumina	908,508	2,683	859,436	833,614	638,085
dcr-2^{L811fsX} heads	no	Illumina	734,343	7,105	688,323	671,408	549,508
untreated ago2 heads	no	Illumina	749,674	27,908	684,388	649,398	1,094,293
oxidized ago2 heads	yes	Illumina	228,112	871	183,572	73,518	17,327

Table II-S7, continued.

matching coding genes (unambiguous)		matching coding genes (ambiguous)		matching transposons (unambiguous)		matching transposons (ambiguous)		matching only *white* IR trigger		matching *white IR trigger and others*	
sense	antisense	sense	antisense	sense	antisense	sense	antisense				
1,394	670	21,742	20,290	4	15	2,821	2,950				
6,586	4,995	44,670	42,945	23	61	24,442	24,148	sense	antisense	sense	antisense
1,163	927	23,752	22,362	1	12	171	224	708	834	0	12
4,068	5,503	9,197	8,841	9	8	703	792	3,102	5,283	0	19
14,518	8,326	64,396	34,109	26	59	2,873	3,411	7,149	7,229	0	5
1,556	1,522	11,506	9,860	4	4	306	294	904	1,165	0	0
5,078	215	33,053	12,706	2	2	438	753				
114	13	1,150	686	0	0	18	22				
16,312	851	132,279	59,124	20	85	1,751	2,874				
1,258	279	10,398	5,969	5	17	550	664				
5,066	1,458	56,057	32,819	40	107	3,430	4,164				
3,476	876	49,723	33,036	20	52	1,510	1,686				
7,188	1,146	71,620	39,516	103	136	13,106	13,179				
323	93	5,532	4,228	4	12	336	440				

CHAPTER III

Sorting of *Drosophila* small silencing RNAs partitions microRNA* strands into the RNA interference pathway

The following chapter was a collaborative effort. The author conceived the experimental design, performed all experiments and initial small-scale bioinformatic analyses. Jia Xu, Hervé Seitz and Zhiping Weng performed bioinformatic analyses. The author and Phillip Zamore wrote the paper. This chapter appeared in RNA. 2010 Jan;16(1):43-56.

Summary

In flies, small silencing RNAs are sorted between Argonaute1 (Ago1), the central protein component of the microRNA (miRNA) pathway, and Argonaute2 (Ago2), which mediates RNA interference. Extensive double-stranded character—as is found in small interfering RNAs (siRNAs)—directs duplexes into Ago2, whereas central mismatches, like those found in miRNA/miRNA* duplexes, direct duplexes into Ago1. Central to this sorting decision is the affinity of the small RNA duplex for the Dcr-2/R2D2 heterodimer, which loads small RNAs into Ago2. Here, we show that while most *Drosophila* miRNAs are bound to Ago1, miRNA* strands accumulate bound to Ago2. Like siRNA loading, efficient loading of miRNA* strands in Ago2 favors duplexes with a paired central region and requires both Dcr-2 and R2D2. Those miRNA and miRNA* sequences bound to Ago2, like siRNAs diced in vivo from long double-stranded RNA, typically begin with

cytidine, whereas Ago1-bound miRNA and miRNA* disproportionately begin with uridine. Consequently, some pre-miRNA generate two or more isoforms from the same side of the stem that differentially partition between Ago1 and Ago2. Our findings provide the first genome-wide test for the idea that *Drosophila* small RNAs are sorted between Ago1 and Ago2 according to their duplex structure and the identity of their first nucleotide.

Introduction

In animals, microRNAs (miRNAs) regulate the stability and rate of translation of mRNAs, whereas small interfering RNAs (siRNAs) silence transposons, defend against viral pathogens, and regulate mRNA expression[254]. Both miRNAs and siRNAs derive from longer double-stranded RNA (dsRNA) precursors, which are cleaved by RNase III dsRNA-specific endonucleases. miRNA production begins in the nucleus, where long primary miRNAs, transcribed by RNA polymerase II, are converted into ~65 nucleotides (nt) pre-miRNA hairpins by the RNase III ribonuclease, Drosha, aided by a double-stranded RNA-binding domain (dsRBD) partner protein[124-127,255,256]. A minority of pre-miRNAs—mirtrons—correspond to entire introns and are excised from their primary transcripts by the pre-mRNA splicing pathway[138,139]. Pre-miRNAs are then exported to the cytoplasm[129-131], where they are processed by a second RNase III enzyme, Dicer, together with its dsRBD partner protein, into ~22 nt long miRNA/miRNA* duplexes[62-64,133-137,257,258].

siRNA production also requires Dicer, which excises 21-nt siRNA duplexes, comprising a guide and passenger strand, from long dsRNA formed by the base pairing of complementary sense and antisense transcripts, convergently transcribed mRNAs, or by the intra-molecular base pairing of long, self-complementary RNAs. Such endogenous dsRNAs yield endo-siRNAs. Similarly, exogenous dsRNA, introduced experimentally or by viral infection, are converted by Dicer to exo-siRNAs.

In *Drosophila melanogaster*, Dicer-1 (Dcr-1) converts pre-miRNAs into miRNA/miRNA* duplexes; Dicer-2 (Dcr-2) converts long dsRNA into 21-nt siRNA duplexes[22,59,66]. The use of different Dicer proteins to generate miRNAs and siRNAs may minimize competition between the two pathways, so that an RNAi defense to viral infection does not perturb miRNA production.

All small silencing RNAs function bound to Argonaute proteins. Argonaute proteins display nucleotides 2 to 8 of the small RNA guide in a pre-helical geometry that confers on this region special importance in target recognition: the majority of the binding energy for target binding is contributed by this "seed" sequence[159,160,164,165,259-267]. In flies, miRNAs are loaded from miRNA/miRNA* duplexes into Argonaute1 (Ago1), whereas siRNAs are loaded from guide/passenger duplexes into Argonaute2 (Ago2)[58,78-82,241,268]. Two binary choices accompany loading of small RNAs into Argonaute proteins in *Drosophila*: the choice of Ago1 versus Ago2 and the selection of one of the two strands of the duplex as a miRNA or guide strand[73,74].

Although fly miRNAs are overwhelmingly associated with Ago1 and siRNAs with Ago2, small RNA production and Argonaute loading are uncoupled[202,203]. Instead,

miRNA and siRNA duplexes are actively partitioned between Ago1 and Ago2 according to their structure. Extensive double-stranded character directs duplexes, such as siRNAs, into Ago2, which mediates RNAi, whereas bulges and mismatches, like those found in miRNA/miRNA* duplexes, are sorted into Ago1[269]. Central to this sorting decision is the affinity of the small RNA duplex for the Dcr-2/R2D2 heterodimer, which loads small RNAs into Ago2[76,77,202,270]. Central mismatches reduce binding of small RNA duplexes by the Dcr-2/R2D2 heterodimer, antagonizing Ago2 loading and promoting loading into Ago1[202,203,269]. The function of the Dcr-2/R2D2 heterodimer in Ago2 loading is separate and distinct from its role in dicing siRNAs from long dsRNA: Dcr-2 bearing a glycine to arginine substitution (G31R) in its helicase domain cannot dice, but can still load siRNA into Ago2[66].

Increasingly, this simple picture of small RNA strand choice is at odds with the intracellular abundance, processing accuracy, and evolutionary conservation of miRNA* strands. First, some evolutionarily conserved miRNAs are less abundant than their miRNA* strands, which appear to be evolving regulatory functions[188]. Second, miRNA* 5′ ends are far more precisely defined than their 3′ ends, suggesting selective pressure to generate an accurate seed region—implying that they have regulatory targets[188,244,271]. Third, there is mounting evidence that some miRNA*s may have regulatory potential[271-273], and fly miRNA* strands are evolutionarily conserved, albeit not to the same extent as miRNAs[271]. Thus, miRNA* strands may regulate gene expression, rather than serve merely as carriers for loading the miRNA strand. Such a mechanism would make small RNA biogenesis more efficient, with each pre-miRNA producing two different regulatory

small RNAs. Nonetheless, miRNAs are typically far more abundant than their miRNA* counterparts, and regulation by low abundance Ago1-small RNA complexes has not been reported in flies.

Here, we show that while most *Drosophila* miRNAs are bound to Ago1 in vivo, most miRNA* strands accumulate bound to Ago2. Partitioning of miRNAs into Ago1 and Ago2 provide a wide-scale in vivo test for the previously proposed principles for small RNA sorting in flies: miRNAs and miRNA* strands are sorted between the two Argonaute proteins according to the structure of their small RNA duplex, a process that requires both Dcr-2 and R2D2. Like the exo-siRNAs that direct RNAi, miRNA* strands bound to Ago2 typically begin with cytidine, whereas Ago1-bound miRNAs usually begin with uridine. Thus, the identity of the first nucleotide of a small RNA plays a role in its sorting in flies, as previously reported for plants. Finally, miRNA* bound to Ago2 are more abundant than siRNAs that direct RNAi, suggesting that they function to silence target RNAs.

Results

*miRNAs and miRNA*s partition differentially between Ago1 and Ago2*

We used high throughput sequencing of 18–29 nt RNA from fly heads to determine the small RNA profile and distribution of small RNAs between Ago1 and Ago2 in this complex somatic structure (Table S1). Unlike other fly tissues, heads express little if any Piwi-interacting RNA, allowing us to focus on small RNAs bound to Ago1 or Ago2[26]. Of the ~1.6 million genome-matching small RNAs sequenced (excluding annotated non-

coding RNAs such as 2S ribosomal RNA), 90.2% were derived from pre-miRNAs (Fig. 1A). In parallel, we used an Ago1 monoclonal antibody[81] to immunoprecipitate Ago1-associated small RNAs from fly head extracts. Nearly 97% of the > 5.03 million small RNA sequences associated with Ago1 were miRNAs; only 2.2% were miRNA* strands (Fig. 1A).

Ago2-loaded guide strands acquire a 3´ terminal 2´-*O*-methyl modification after their corresponding passenger strand is discarded[83,182]. To enrich for Ago2-loaded small RNAs, we oxidized the 18–29 nt RNAs prior to library preparation, a treatment that excludes from the library most Ago1-loaded small RNAs, which bear 2´,3´ hydroxyl termini, but allows sequencing of Ago2-loaded small RNAs, because their 2´-*O*-methyl modification protects them from reaction with $NaIO_4$[26,244]. In general, the pre-miRNA-derived small RNAs associated with Ago1 correlated well with the total small RNA profile (r = 0.91 for miRNAs; r = 0.70 for miRNA* strands), supporting the view that the majority of small RNAs in fly heads accumulate because they are bound to Ago1. However, a global fit of the sum of the miRNA and miRNA* species detected in the Ago1 immunoprecipitation and the miRNA and miRNA* species detected in the library prepared from oxidized RNA more closely recapitulated the total small RNA profile (r = 0.91 for miRNAs; r = 0.85 for miRNA* strands), suggesting that Ago2-bound miRNA and/or miRNA* species are a significant component of the total pre-miRNA–derived small RNA population.

siRNAs were previously identified as the major class of Ago2-associated endogenous small RNAs in flies[26-29,112,113]. Yet, the population of Ago2-associated small

RNAs contained more miRNA plus miRNA* combined (53.2%) than endo-siRNAs (33.2%) (Fig. 1A). Thus, the identity of the Dicer paralog that generates a small RNA does not determine the Argonaute protein into which it is loaded. Compared to the total small RNA population—where miRNAs represented ~87.5% of all small RNAs, but miRNA* reads were just 2.6%—miRNAs were underrepresented (39.4%) and miRNA* (13.8%) were over-represented among the Ago2-associated small RNA sequences. The abundance of pre-miRNA–derived small RNAs associated with Ago2 calls into question the prevailing view that Ago2 is restricted to the RNAi pathway.

In general, Ago2 was significantly depleted of miRNAs and enriched for miRNA* sequences ($P \leq 2.2 \times 10^{-16}$). Conversely, Ago1 was significantly depleted of miRNA* sequences and enriched for miRNAs ($P \leq 2.2 \times 10^{-16}$). For some of these—especially miRNAs—more of a particular small RNA was present in Ago1 than in Ago2, but more of that small RNA was associated with Ago2 than would be expected by chance. In all, 26 miRNAs and 49 miRNA* were significantly ($P \leq 0.01$) enriched in Ago2, whereas 71 miRNAs and 9 miRNA* were significantly ($P \leq 0.01$) enriched in Ago1 (Fig. 1B). Of the 49 miRNA* enriched in Ago2, 32 had their corresponding miRNA enriched in Ago1, while 15 had their miRNA enriched in Ago2. Among the examples illustrated in Figure 2, the miRNAs *bantam* and miR-308 were enriched in Ago1, whereas *bantam** and miR-308* were enriched in Ago2. Table S2 reports the enrichment or depletion of individual miRNAs and miRNA* species between the two *Drosophila* Argonaute proteins.

Although generally less abundant than miRNAs bound to Ago1, miRNA* isoforms (i.e., all of the species derived from the same side of the stem of a single pre-miRNA *and* sharing a common seed) bound to Ago2 were equally or more abundant than other small RNAs that exert their regulatory functions through Ago2, including the well studied exo-siRNAs derived from an inverted repeat transgene that fully silences the *white* gene via the RNAi pathway[243]. The median abundance for miRNA* isoforms enriched in Ago2 was more than twice that of the median abundance for *white* exo-siRNAs bound to Ago2, and 18 miRNA* were more abundant than the single most abundant *white* exo-siRNA detected in the same fly heads. (These 18 miRNA* are outliers whose abundance was too large to display on the box plot in Fig. 1C.) In fact, the abundance of a single miR-8* isoform alone (2,748 parts per million [ppm]), was nearly two-thirds of the aggregate abundance of *all* antisense *white* exo-siRNAs (4,273 ppm), whose concentration in heads is sufficient to phenocopy a strong loss-of-function *white* mutation. Summing the isoforms of each miRNA*, 25 miRNA* were more abundant than all antisense *white* exo-siRNAs combined.

Figure III-1. miRNA* are loaded in Ago2. (A) Relative abundance of miRNA, miRNA*, and endo-siRNAs among total fly head small RNA, Ago1-bound small RNAs—inferred from co-immunoprecipitation with Ago1, and Ago2-bound small RNAs—inferred from their presence in an oxidized small RNA library. (B) Box plots illustrating the enrichment scores for all miRNA and miRNA associated with Ago1 (i.e., in the Ago1 immunoprecipitate) or Ago2 (i.e., in the oxidized library) and for miRNA and miRNA* that were significantly ($P \leq 0.01$) associated with Ago1 or Ago2. For miRNA* enriched in Ago2, six outliers with enrichment scores greater than 150 are not show: miR-92a* (score = 1206), miR-308* (score = 649), miR-998* (score = 598), miR-315* (score = 514), miR-2a-2* (score = 309), and miR-33* (score = 304). (C) Box plots illustrating the abundance of Ago2-enriched miRNA* and *white* exo-siRNAs in the total RNA library. For miRNA* enriched in Ago2, 18 outliers with abundance greater than 250 ppm are not shown, including miR-8* (2,748 ppm) and miR-34* (1,747 ppm).

Figure III-1.

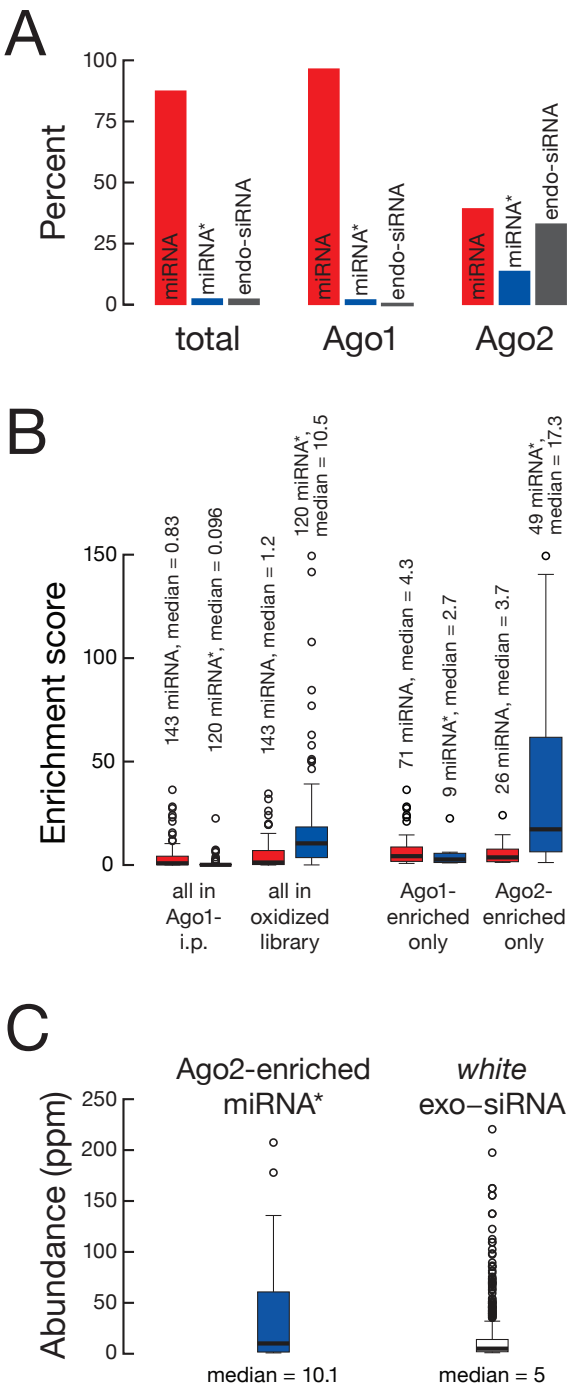

A
Percent
100
75
50
25
0
miRNA
miRNA*
endo-siRNA
total
Ago1
Ago2
B
Enrichment score
150
100
50
0
143 miRNA, median = 0.83
120 miRNA*, median = 0.096
143 miRNA, median = 1.2
120 miRNA*, median = 10.5
71 miRNA, median = 4.3
9 miRNA*, median = 2.7
26 miRNA, median = 3.7
49 miRNA*, median = 17.3
all in Ago1-i.p.
all in oxidized library
Ago1-enriched only
Ago2-enriched only
C
Abundance (ppm)
250
200
150
100
50
0
Ago2-enriched miRNA*
white exo–siRNA
median = 10.1
median = 5

Figure III-2. Exemplary miRNA and miRNA* duplexes. Typical miRNA/miRNA* duplexes load their miRNA strands into Ago1 and their miRNA* strands into Ago2. The examples here correspond to duplexes whose miRNA strand was significantly ($P \leq 0.01$) enriched in Ago1 and whose miRNA* strand was enriched in Ago2. These duplexes present different structures to the Ago1 and Ago2 sorting machinery, as the prospective guide strand occupies a unique position during Argonaute loading. When viewed with miRNA strand as the guide and miRNA* strand as the passenger, the duplex presents a duplex with central bulges, mismatches and G:U wobbles, but when miRNA* strand will become the guide and miRNA strand serves as the passenger, the duplexes present more stably paired central regions. The duplexes are drawn using the guide isoform that was most abundant for the specific Argonaute protein paired to the most abundant passenger sequence detected in the total small RNA library. Red text, seed sequence; shaded bars highlight positions that are significantly different between Ago1- and Ago2-loaded guides (see Fig. 4).

Figure III-2.


Ago1
Ago2
5′ U, unpaired P9–10
bantam
bantam*
bantam*
bantam
5′ C, paired P8–11, P10 G:U
5′ U, P7–P9 G:U, unpaired P10
miR-14
miR-14*
miR-14*
miR-14
5′ G, paired P8–10, unpaired P11
5′ U, unpaired P9–10
miR-282
miR-282*
miR-282*
miR-282
5′ A, paired P8–10, unpaired p11
5′ A, P8–9 G:U, unpaired P11
miR-308
miR-308*
miR-308*
miR-308
5′ C, paired P8,9,11, unpaired P10
5′ C, unpaired P8–9, P10 G:U
miR-929
miR-929*
miR-929*
miR-929
5′ A, paired P8–10, P11 G:U
5′ U, P9 bulge, P11 G:U
miR-996
miR-996*
miR-996*
miR-996
5′ G, paired P8,9,11, P10 G:U

The siRNA-loading machinery sorts miRNA* strands into Ago2

Apart from its function in producing siRNAs, Dcr-2 acts with its double-stranded RNA-binding domain protein partner, R2D2, to both load small RNA duplexes into Ago2 and determine the identity of guide and passenger strands. Thus, both Dcr-2 and R2D2 are required to load Ago2 with siRNAs derived from exogenous dsRNA (exo-siRNAs), such as those derived from a long inverted repeat transcript designed to silence *white* mRNA expression[66,134]. At least one *Drosophila* miRNA, miR-277, which associates equally with Ago1 and Ago2 in cultured S2 cells, requires Dcr-2 and R2D2 to load it into Ago2, even though miR-277 requires Dcr-1 to liberate it from pre-miR-277[203].

Likewise, those miRNA and miRNA* sequences that were enriched in Ago2 required Dcr-2 and R2D2 for their loading (Fig. 3A). The median extent of Ago2 loading of these miRNAs declined 2.7-fold in *dcr-2*^{L811fsX} and 3.3-fold in *r2d2*^{1} heads, compared to wild-type; loading of miRNA* into Ago2 declined 2.1-fold in *dcr-2*^{L811fsX} and 3.1-fold in *r2d2*^{1}. In contrast, the overall abundance of the miRNA or miRNA* sequences that were enriched in Ago1 was unaltered in *dcr-2* or *r2d2* mutant heads.

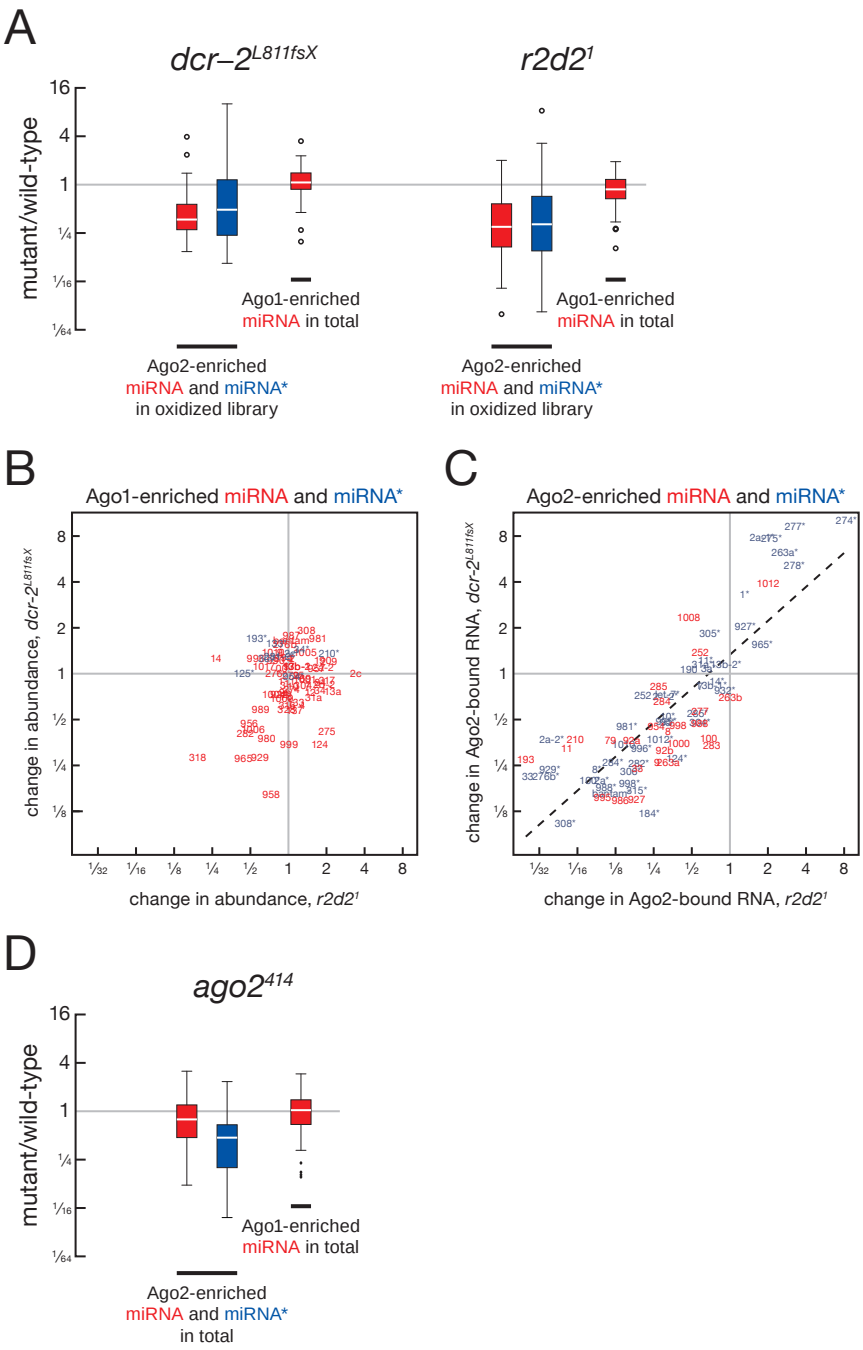
R2D2 is stabilized by its association with Dcr-2[76,203]. Consequently, *dcr-2*^{L811fsX} flies are also deficient in R2D2. For miRNA and miRNA* that were preferentially loaded into Ago2, the effect of the absence of Dcr-2 and R2D2 on Ago2 loading were well correlated (r = 0.828) (Fig. 3B,C). As expected, the abundance of miRNA and miRNA* that were preferentially loaded into Ago1 were largely unchanged in these two mutants.

The median abundance of Ago2-enriched miRNA* sequences in the total RNA library declined ~2.1-fold in the absence of Ago2 (Fig. 3D). In contrast, the median

abundance of miRNA-enriched in Ago1 was unaltered in *ago2*^{414} mutants heads, compared to wild-type (median fold change = 1.0), a significant difference from the Ago2-enriched miRNA* ($P \leq 3.1 \times 10^{-8}$). These data suggest that in the *ago2* mutant, those miRNA* species that normally are loaded into Ago2 become less stable when that Argonaute protein is not available. We envision that these miRNA*/miRNA duplexes, while good substrates for the Ago2-loading machinery, are poor loading substrates for the Ago1-loading machinery. In the absence of Ago2, miRNA*/miRNA duplexes from which the Ago2-enriched miRNA* are normally loaded into Ago2 can no longer be used for this purpose. Instead, they are now used as miRNA/miRNA* duplexes—whose structure typically favors Ago1 loading—to load their miRNA strand into Ago1. The observation that abundance of Ago2-enriched miRNA* sequences declines in *ago2*^{414} heads supports the earlier proposal that the duplex features that promote Ago2-loading are anti-determinants for Ago1 loading[202,269].

Figure III-3. Association of miRNA* with Ago2 relies on the Ago2-loading machinery. (A) Efficient loading into Ago2 of miRNA and miRNA* strands—measured by their abundance in an oxidized small RNA library—was diminished in heads from *dcr-2*^{L811fsX} and *r2d2*^{1} mutants for miRNA and miRNA* normally enriched in Ago2, but the abundance of Ago1-enriched miRNAs was unaltered, as measured in the total small RNA library. Box plots illustrate the fold-change between mutant and wild-type. (B,C) The requirement for Dcr-2 and R2D2 for Ago2 loading was well correlated for miRNA and miRNA* strands preferentially loaded into Ago2. (D) The overall abundance of Ago2-enriched miRNA and miRNA*—measured in the total small RNA library—decline in *ago2* mutant heads. Box plots illustrate the fold-change between mutant and wild-type in total small RNA libraries.

Figure III-3.


A
dcr–2L811fsX
r2d21
mutant/wild-type
16
4
1
¼
1/16
1/64
Ago1-enriched miRNA in total
Ago2-enriched miRNA and miRNA* in oxidized library
B
Ago1-enriched miRNA and miRNA*
change in abundance, dcr-2L811fsX
change in abundance, r2d21
C
Ago2-enriched miRNA and miRNA*
change in Ago2-bound RNA, dcr-2L811fsX
change in Ago2-bound RNA, r2d21
D
ago2414
Ago2-enriched miRNA and miRNA* in total

miRNA/miRNA* duplex structure determines Argonaute loading

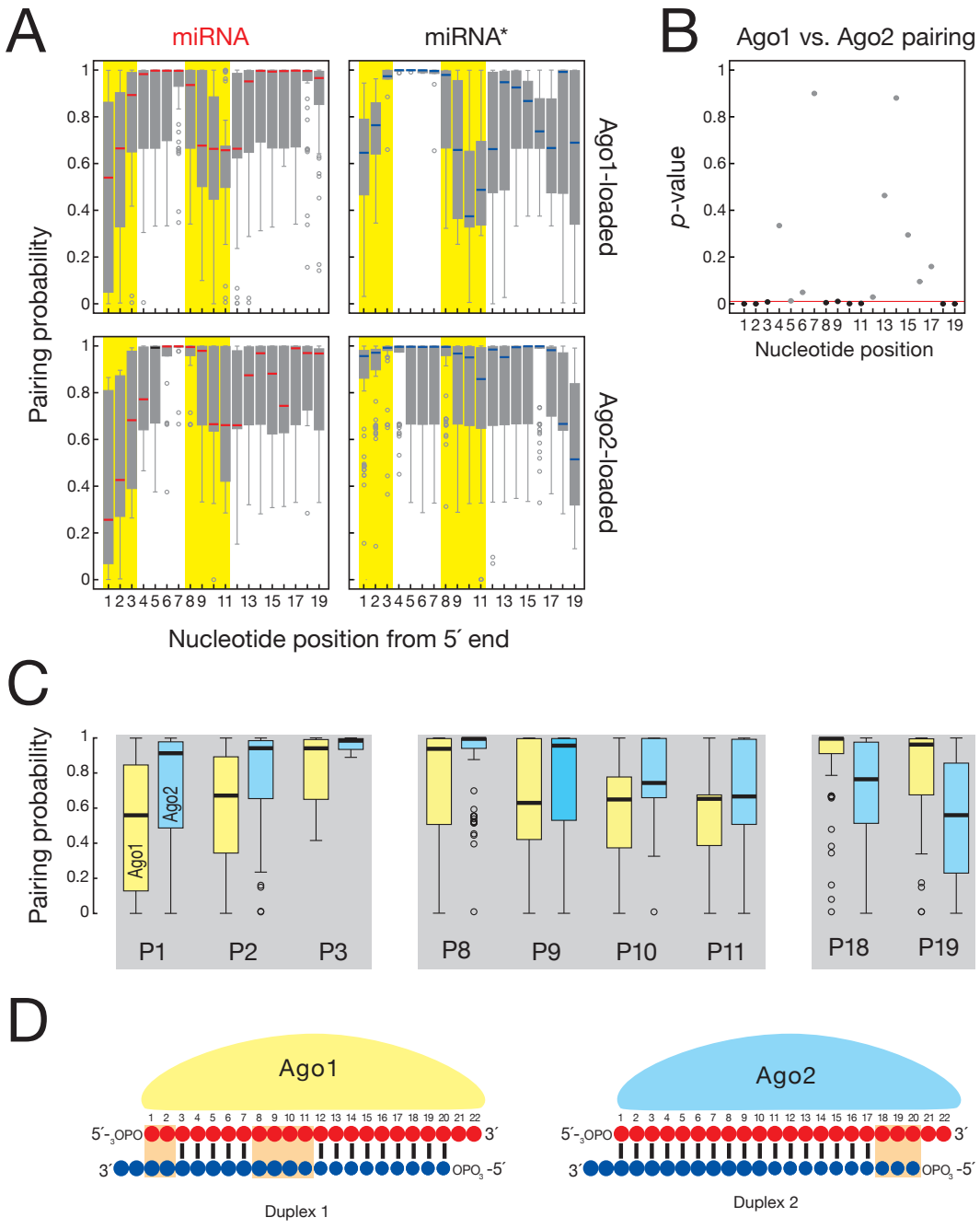
The Dcr-2/R2D2 heterodimer interprets the structure of a small RNA duplex, sorting centrally paired duplexes into Ago2 and leaving duplexes with an unpaired region centered on guide nucleotide 9 to enter the Ago1 loading pathway[202,203]. Each small RNA duplex presents two distinct duplexes to the fly sorting machinery. For example, *bantam*/*bantam** displays mismatches at guide positions 9 and 10 when viewed from the 5´ end of the miRNA, but these positions are paired when viewed from the 5´ end of the miRNA* strand (Fig. 2). That is, the *bantam*/*bantam** and *bantam**/*bantam* duplexes are not equivalent.

To evaluate if miRNA/miRNA* duplexes and miRNA*/miRNA duplexes generally present distinct structures to the *Drosophila* Argonaute loading machineries, we calculated the pairing probability for each nucleotide in each miRNA/miRNA* duplex that loads an Ago1- or Ago2-enriched miRNA or miRNA*/miRNA duplex that loads an Ago1- or Ago2-enriched miRNA* (Fig. 4A). Viewed in this way, two significant ($P < 0.01$) structural differences emerge that distinguish duplexes that load Ago1 from those that load Ago2 (Fig. 4B,C): from the perspective of the loaded strand, Ago1-loading duplexes are more likely to have an unpaired 5´ end and a central unpaired region that spans nucleotide positions 8–11. Conversely, Ago2-loading duplexes more likely have a paired 5′ end and a central region with greater double-stranded character. Ago2-loading duplexes are also more likely to have an unpaired guide 3´ end (Fig. 4D). Remarkably, these differences reflect the “rules” for sorting small RNA duplexes between Ago1 and Ago2 that were inferred previously from biochemical studies[202,269]. Thus, they provide in

vivo validation of the hypothesis that *Drosophila* small RNA duplex structure determines its partitioning between Ago1 and Ago2.

Figure III-4. Pairing profiles of Ago1- and Ago2-loaded small RNA guides. (A) Box plots illustrate the predicted double-stranded character of each nucleotide position, 1–19, for all Ago1- or Ago2-enriched miRNA or miRNA* strands. (B) The Wilcoxon test *P*-value for each comparison was used to identify nucleotide positions that were significantly different between Ago1-enriched miRNA plus miRNA* compared with Ago2-enriched miRNA plus miRNA*. The red line indicates $P = 0.01$. Grey circles, non-significant; black circles, significant. (C) Box plots illustrate the differences in double-stranded character for each position that was significantly different in double-stranded character between Ago1-loaded and Ago2-loaded miRNA plus miRNA* in (B). (D) The data in (A–C) suggest that miRNA duplexes with less stable 5´ ends and central mismatches act as guides for Ago1 and miRNA duplexes with less stable 3´ ends act as guides for Ago2.

Figure III-4.


A
miRNA
miRNA*
Ago1-loaded
Ago2-loaded
Pairing probability
Nucleotide position from 5′ end
B
Ago1 vs. Ago2 pairing
p-value
Nucleotide position
C
Pairing probability
Ago1
Ago2
P1
P2
P3
P8
P9
P10
P11
P18
P19
D
Ago1
Ago2
5′-3OPO
3′
OPO3 -5′
Duplex 1
Duplex 2

The 5´ terminal nucleotide of a small RNA reflects its partitioning between Ago1 and Ago2

Arabidopsis thaliana produces ten distinct AGO proteins, and small RNAs are sorted among them according to their first nucleotide. Of the 187 annotated miRNAs in *Arabidopsis*, ~76% begin with uridine, consistent with the idea that a 5´ U steers a small RNA into plant Ago1[108,204]. *Arabidopsis* Ago2 and Ago4 preferentially load small RNAs that begin with an adenosine, whereas Ago5 favors small RNAs that begin with cytidine108,204. Small RNAs in flies partition between Ago1 and Ago2 according to the structure of the duplex from which they are loaded, yet, as in plants, *Drosophila* miRNAs overwhelmingly begin with U, whereas U is not over-represented as the first nucleotide of siRNAs[26].

We analyzed the sequence composition of Ago1- and Ago2-loaded miRNA and miRNA* strands present in our small RNA libraries from fly heads. To prevent differential rates of transcription or miRNA precursor processing from skewing our analysis, for each set of small RNAs derived from a common precursor, we weighted the sequence bias of each miRNA or miRNA* isoform by its relative abundance, then averaged the sequence bias among all miRNAs or miRNA* strands, weighting each locus equally (Fig. 5).

Our analysis suggests that the first nucleotide of a fly small RNA reflects its sorting between Ago1 and Ago2. miRNAs expressed in fly heads generally began with U (72%) rather than A (15.2%), C (7.6%), or G (5.2%); for miRNAs bound to Ago1, as judged by their co-purification with immunoprecipitated Ago1, 73.5% began with U,

whereas 7.1% began with C. Among the miRNA and miRNA* species that were significantly ($P \leq 0.01$) enriched in Ago1 relative to the total small RNA pool of fly heads, 83.9% began with U; just 3.4% began with C. In contrast, 49% of miRNAs that were enriched in Ago2 began with U; 21.6% began with C and 21.8% began with A, indicating a selection against a 5´ U.

miRNA* strands showed a distinctly different 5´ sequence bias. The miRNA* detected in fly heads typically began with A (28.2%), C (32.1%), or G (22.1%), rather than U (17.6%). In contrast to this overall 5´ sequence bias, those miRNA* that were significantly enriched in Ago1 began either with A (56.3%) or U (29.2%); the population of miRNA* loaded into Ago1 was depleted of miRNA* isoforms that begin with C.

Ago2-loaded miRNA* strands showed the opposite bias: they typically began with C. Nearly 58% of miRNA* strands enriched in Ago2 and detected in the oxidized library began with C, 15.2% began with A, and just 7.7% began with U, a sequence bias significantly different from the composition of nucleotides 2–18 of the same small RNAs ($P \leq 6.7 \times 10^{-10}$, Fisher's exact test) and from the first nucleotide bias of miRNA* overall ($P \leq 6.6 \times 10^{-7}$) and of those miRNA* loaded into Ago1 ($P \leq 0.017$). Overall, 40% of the Ago1-enriched miRNA or miRNA* species began with U, whereas 23% of the Ago2-enriched miRNA or miRNA* species began with C.

Essentially identical sequence biases for both miRNA and miRNA* were present in independent small RNA libraries from male and female heads, in libraries prepared from three distinct genetic backgrounds (Oregon R, *dcr-2*^{L811fsX}/CyO, *or r2d2*^{1}/CyO), in libraries of Ago2-associated small RNAs that were prepared using either oxidation or

oxidation followed by β-elimination, and in libraries processed and sequenced using two different high throughput technologies: pyrosequencing (“454”) or sequencing-by-synthesis (Illumina Genome Analyzer). Together, these data suggest that, in flies, a 5´ terminal U promotes Ago1 loading but discourages association with Ago2, whereas a 5´ terminal C directs a small RNA away from Ago1 and towards Ago2.

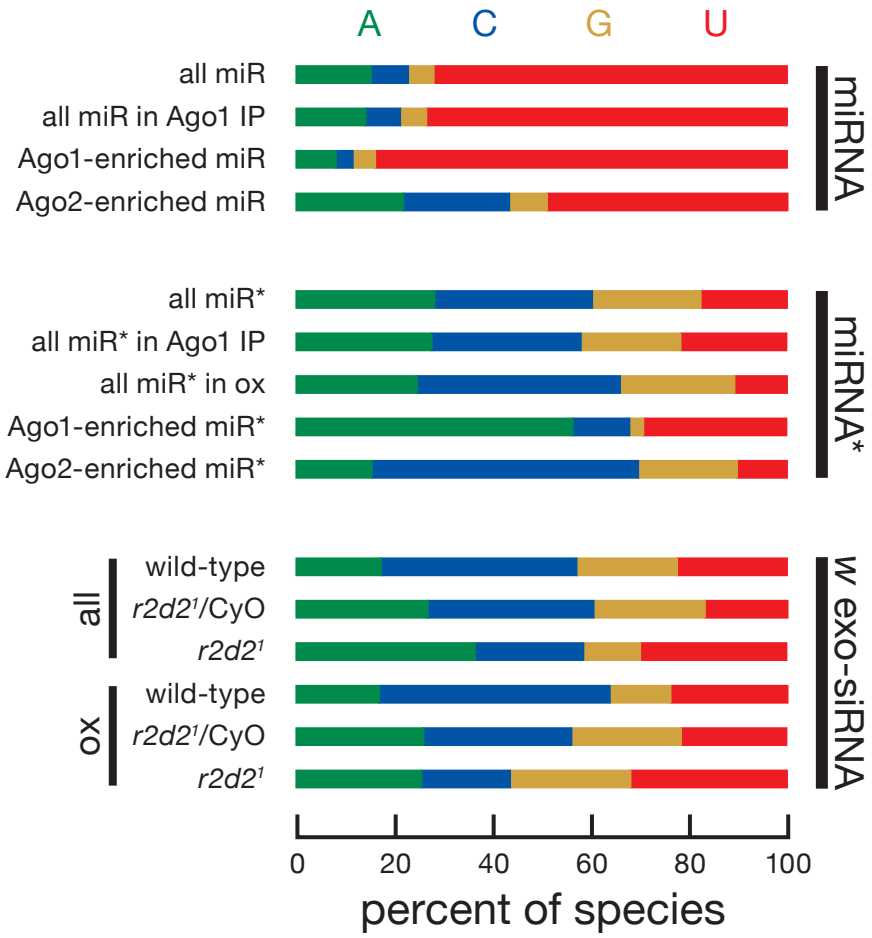
To further test this hypothesis, we analyzed the 5´ nucleotide composition of exo-siRNAs derived from a P-element transgene expressing a long inverted repeat corresponding to exon 3 of the *white* mRNA. We compared the overall population of *white* exo-siRNAs with those *white* exo-siRNAs bound to Ago2, as inferred from their presence in an oxidized small RNA library. Because the *white* exo-siRNA species are transcribed and diced from a common transcript, differences in their steady-state abundance likely reflect, at least in part, their different propensities to load into an Argonaute protein. Supporting this view, *white* exo-siRNAs levels decline >10-fold in vivo in a $dcr\text{-}2^{L811fsX}$, $r2d2^{1}$, or $ago2^{414}$ mutant (T. Du and PDZ, unpublished data). We therefore weighted each first nucleotide according to the abundance of the corresponding *white* exo-siRNA species.

Like Ago2-enriched miRNA*, exo-siRNAs isolated from fly heads typically began with C (39.8%), rather than A (17.3%), G (20.5%), or U (22.4%), a sequence bias significantly different from that of the corresponding strand of the dsRNA from which they are derived ($P \leq 1.8 \times 10^{-9}$). Among the *white* exo-siRNAs in the library prepared from oxidized small RNA—i.e., small RNAs bound to Ago2—47% began with C. Supporting the view that the strong C-bias of exo-siRNAs reflects their association with

Ago2, the 5′ C bias was not observed among the 17-fold lower amount of exo-siRNAs that remained in an *r2d2*^{1} mutant. *r2d2*^{1} mutant flies are defective in loading exo-siRNAs into Ago2 and do not silence *white* expression[134].

Figure III-5. miRNAs and miRNA* show an Argonaute-specific first nucleotide bias. miRNAs and miRNA* associated with Ago1 or Ago2 differ in the bias of their first nucleotide. miRNAs generally begin with uridine; this bias increased for the subset of miRNA that were Ago1-bound (measured in the Ago1 immunoprecipitate library), and increased further for the subset of Ago1-enriched miRNAs (measured in the total small RNA library). In contrast, Ago2-enrcihed miRNAs were depleted of 5´ uridine in the oxidized small RNA library. miRNA* strands generally began with adenosine or cytidine. All miRNA* strands detected in the oxidized library (i.e., loaded in Ago2) or those enriched in Ago2, were significantly more likely to begin with cytidine, whereas those miRNA* enriched in Ago1 were depleted of a 5´ cytidine. A 5´ cytidine bias was also observed for *white* exo-siRNAs and was diminished in *r2d2*^{1}, a mutant defective in Ago2-loading.

Figure III-5.


A
C
G
U
all miR
all miR in Ago1 IP
Ago1-enriched miR
Ago2-enriched miR
miRNA
all miR*
all miR* in Ago1 IP
all miR* in ox
Ago1-enriched miR*
Ago2-enriched miR*
miRNA*
all
wild-type
r2d2¹/CyO
r2d2¹
ox
wild-type
r2d2¹/CyO
r2d2¹
w exo-siRNA
0
20
40
60
80
100
percent of species

To further test the idea that the first nucleotide of a small RNA duplex influences its sorting between Ago1 and Ago2 in flies, we examined the loading of small RNA duplexes in vitro, using a previously described UV cross-linking assay[202]. We synthesized two miRNA duplexes, one corresponding to the authentic *Drosophila let-7* miRNA/miRNA* duplex, which begins with a 5′ U, and a second in which the initial U of *let-7* was changed to a 5′ C (Fig. 6A). In parallel, we also synthesized two siRNA duplexes in which the guide strand was either authentic *let-7* (paired to its reverse complement) or *let-7* bearing a 5′ C instead of a U (Fig. 6A). Each miRNA or guide strand was 5′ ^{32}P-radiolabeled, so that cross-linking identified the proteins, including Ago1 and Ago2, to which it bound when incubated in *Drosophila* embryo lysate.

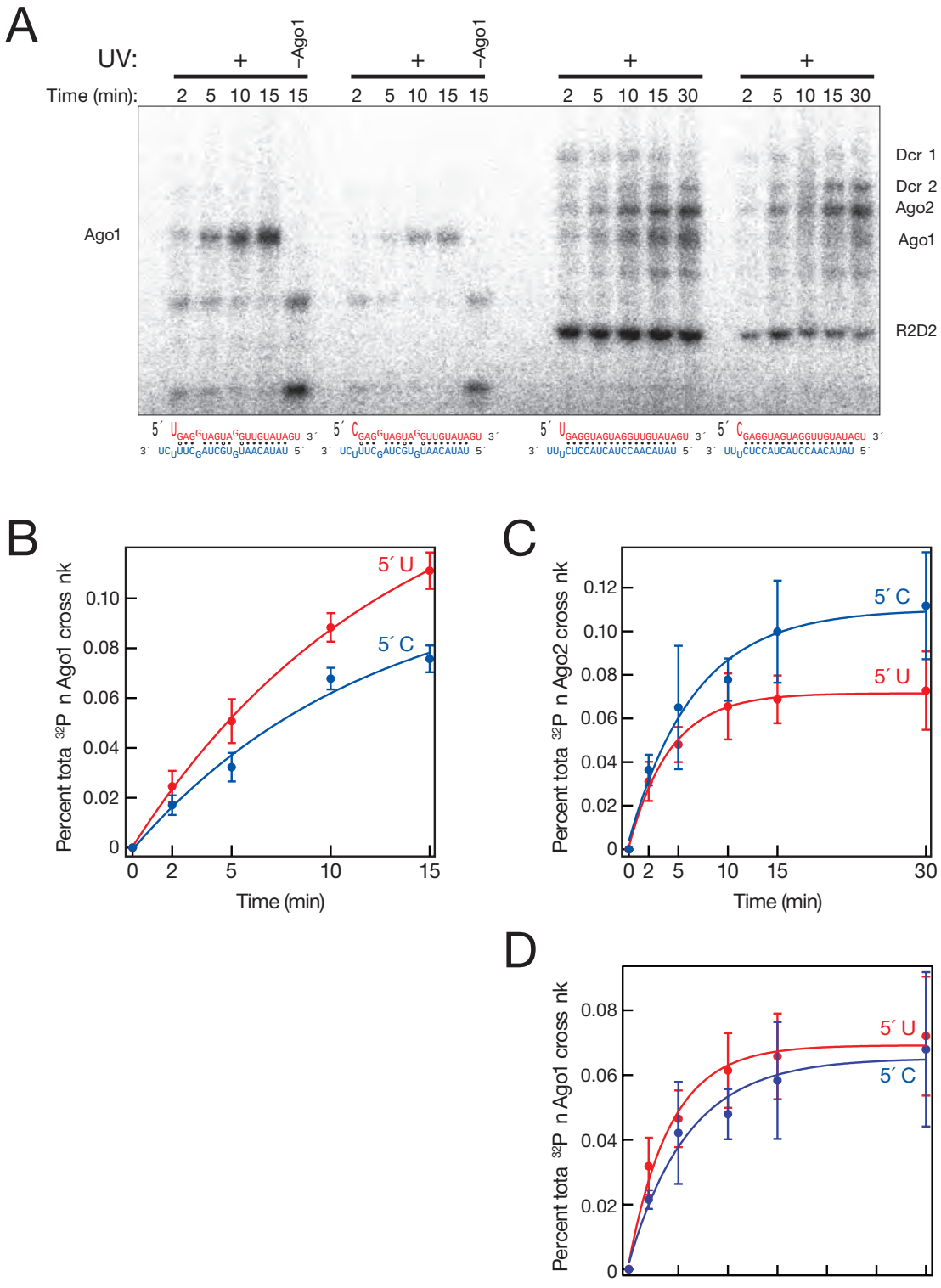
The miRNA/miRNA* duplex containing authentic *let-7* strand—i.e., *let-7* that began with a 5′ U—cross-linked to Ago1 more efficiently than the *let-7* variant that began with a 5′ C (Fig. 6A,B); neither duplex detectably loaded its miRNA strand into Ago2. Moreover, when we performed cross-linking in Ago1 immunodepleted lysate, not only was Ago1 cross-linking absent, but no Ago2 cross-link appeared. We conclude that the structure of a miRNA duplex not only favors Ago1 loading, but actively prevents loading of the miRNA into Ago2. Moreover, the interplay between the structure of the miRNA duplex and its 5′ nucleotide determines its distribution between Ago1 and Ago2.

In contrast to the miRNA/miRNA* duplexes, the siRNA duplexes cross-linked mainly to Ago2, although some Ago1 cross-linking was clearly detected. For the siRNA duplexes, the influence of first nucleotide identity on the efficiency of Ago2 loading was opposite of that observed for the miRNAs: the siRNA duplex whose guide strand began

with a 5´ C loaded more efficiently into Ago2 than the siRNA duplex whose guide began with U (Fig. 6A,C). Together, these in vitro data provide strong support for the hypothesis that the enrichment for a 5´ U among Ago1-loaded miRNAs and for a 5´ C among Ago2-loaded miRNA, miRNA*, and siRNAs reflects a direct role for 5´ nucleotide identity in small RNA sorting between Ago1 and Ago2 in *Drosophila*.

Figure III-6. Ago1 prefers to load miRNAs that begin with a 5´ uridine, while Ago2 prefers siRNAs that begin with a 5´ cytidine. (A) Four small-RNA duplexes were incubated with embryo lysate and then cross-linked with shortwave UV to identify small RNA-bound proteins. Representative data is shown. (B) Kinetic analysis of miRNA association with Ago1, monitored by UV cross-linking. (C) Kinetic analysis of siRNA association with Ago2, monitored by UV cross-linking. (D) Kinetic analysis of siRNA association with Ago1, monitored by UV cross-linking. In B and C, each data point represents the average ± standard deviation for three trials.

Figure III-6.

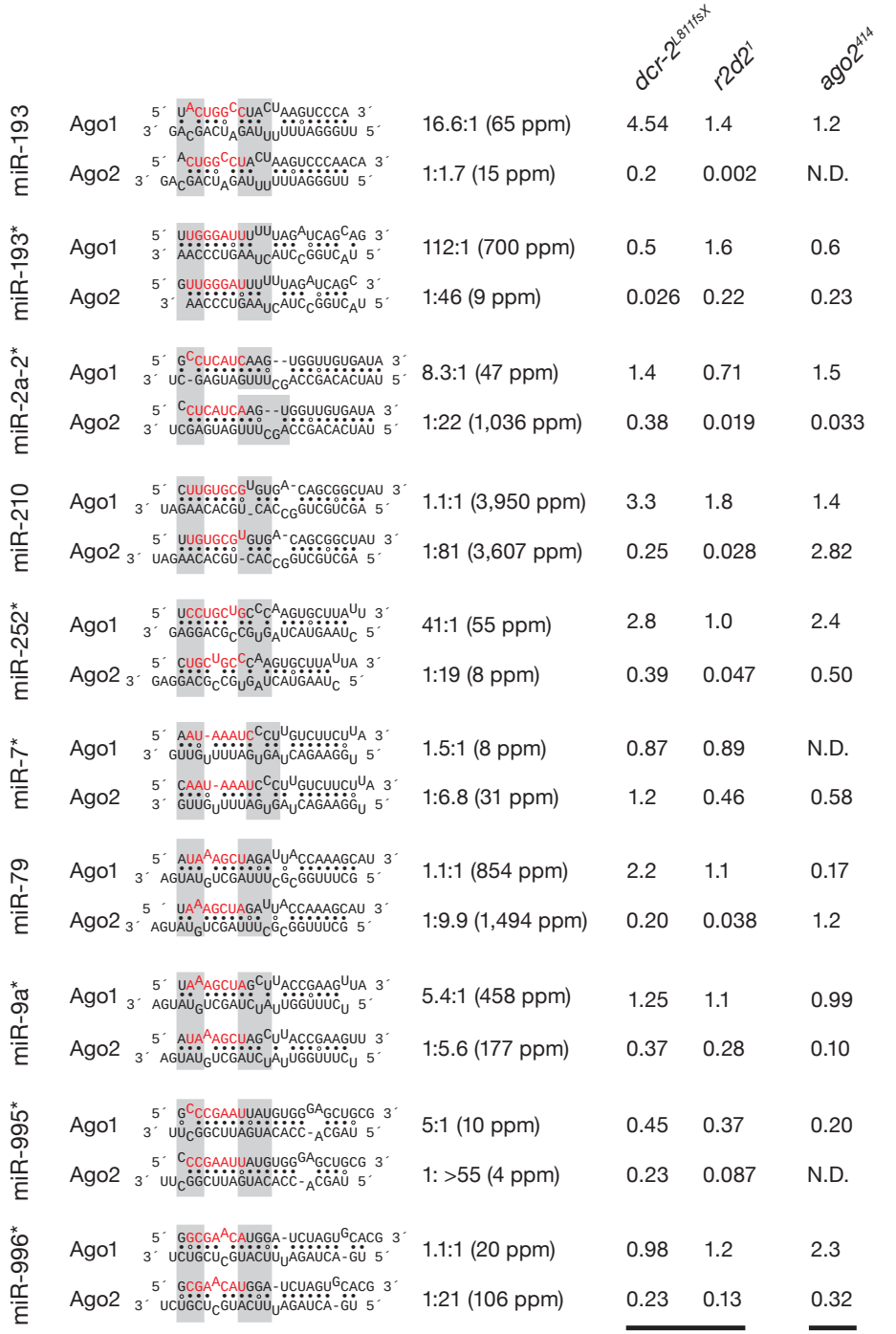

A
UV:
+
–Ago1
Time (min):
2 5 10 15 15
2 5 10 15 15
2 5 10 15 30
2 5 10 15 30
Ago1
Dcr 1
Dcr 2
Ago2
Ago1
R2D2
B
Percent tota 32P n Ago1 cross nk
5´ U
5´ C
Time (min)
C
Percent tota 32P n Ago2 cross nk
5´ C
5´ U
Time (min)
D
Percent tota 32P n Ago1 cross nk
5´ U
5´ C

For some miRNA and miRNA*, distinct isoforms load into Ago1 and Ago2

At least nine *Drosophila* pre-miRNA produce from one side of their stem two small RNAs that partition differentially between Ago1 and Ago2. Such differentially partitioning miRNA or miRNA* isoforms differ at their 5´ ends and therefore present subtly different duplexes to the Argonaute-loading machinery. Moreover, the differentially sorting isoforms have different seed sequences, which would allow them to regulate distinct repertoires of target mRNAs. Figure 7 presents these "seed switching" miRNA and miRNA* isoforms in the context of the duplexes from which they are presumed to be loaded into Ago1 or Ago2. Pre-miR-193 provides a particularly stunning example of such isoform-specific Argonaute loading. This pre-miRNA generates two miR-193 isoforms: one begins with a U and loads into Ago1, whereas a miR-193 isoform that begins at the next nucleotide, an A, loads into Ago2. Pre-miR-193 also generates two miR-193* isoforms. Again, the one that begins with a U loads into Ago1, whereas a less abundant isoform that begins at the G that lies immediately 5´ to the U loads into Ago2. This small collection of seed switching miRNA and miRNA* gives the impression that the sorting of imperfectly paired small RNA duplexes between Ago1 and Ago2 reflects a complex interplay between structural determinants or anti-determinants and first nucleotide preferences and dislikes.

Figure III-7. miRNA and miRNA* can switch seeds between Ago1 and Ago2. Depicted are miRNA/miRNA* duplexes that load distinct isoforms of their miRNA or miRNA* between Ago1 and Ago2, resulting in seed switching between Argonautes. The duplexes are drawn pairing the most abundant guide isoform associated with the particular Argonaute to the most abundant passenger strand isoform in total head small RNA library. Reads in parts per million represent the sum of all isoforms that share the same seed. Ratio reports the relative number of reads for the isoform in Ago1: the number of reads for the isoform in Ago2 as detected within either the library prepared from Ago1 immunoprecipitated small RNAs (Ago1 ratio) or oxidized small RNA (Ago2 ratio). Red text, seed sequence; shaded bars, determinative positions for small RNA sorting between Ago1 and Ago2; N.D., detected in wild-type but not detected in the *ago2*^{414} mutant.

Figure III-7.


dcr-2L811fsX
r2d2l
ago2414
miR-193
Ago1
16.6:1 (65 ppm)
4.54
1.4
1.2
Ago2
1:1.7 (15 ppm)
0.2
0.002
N.D.
miR-193*
Ago1
112:1 (700 ppm)
0.5
1.6
0.6
Ago2
1:46 (9 ppm)
0.026
0.22
0.23
miR-2a-2*
Ago1
8.3:1 (47 ppm)
1.4
0.71
1.5
Ago2
1:22 (1,036 ppm)
0.38
0.019
0.033
miR-210
Ago1
1.1:1 (3,950 ppm)
3.3
1.8
1.4
Ago2
1:81 (3,607 ppm)
0.25
0.028
2.82
miR-252*
Ago1
41:1 (55 ppm)
2.8
1.0
2.4
Ago2
1:19 (8 ppm)
0.39
0.047
0.50
miR-7*
Ago1
1.5:1 (8 ppm)
0.87
0.89
N.D.
Ago2
1:6.8 (31 ppm)
1.2
0.46
0.58
miR-79
Ago1
1.1:1 (854 ppm)
2.2
1.1
0.17
Ago2
1:9.9 (1,494 ppm)
0.20
0.038
1.2
miR-9a*
Ago1
5.4:1 (458 ppm)
1.25
1.1
0.99
Ago2
1:5.6 (177 ppm)
0.37
0.28
0.10
miR-995*
Ago1
5:1 (10 ppm)
0.45
0.37
0.20
Ago2
1: >55 (4 ppm)
0.23
0.087
N.D.
miR-996*
Ago1
1.1:1 (20 ppm)
0.98
1.2
2.3
Ago2
1:21 (106 ppm)
0.23
0.13
0.32
change in Ago2-loaded RNA
change in overall abundance

Discussion

Historically, miRNA were defined as the more abundant of the small RNAs derived from the two sides of a pre-miRNA stem[3-5]. The miRNA* strand has been proposed to be destroyed during Argonaute loading, explaining its considerably lower abundance[73,74]. Yet, high depth sequencing has revealed that many miRNA* species are more abundant than some miRNA species, and miRNA/miRNA* ratios may vary dramatically among developmental stages[271,272].

In fly heads and ovaries, several miRNA* strands are more abundant than their annotated miRNA counterparts (Table 1). In our data sets, miR-92a was more abundant than miR-92a* in ovaries (3,240 ppm miRNA vs. 15 ppm miRNA*), while its miR-92a* was more abundant than miR-92a in heads (24 ppm miRNA vs. 106 ppm miRNA*). Likewise, miR-988 (260 ppm miRNA vs. 300 ppm miRNA* in heads, but 124 ppm miRNA vs. 49 ppm miRNA* in ovaries) and miR-284 (4,993 ppm miRNA vs. 915 ppm miRNA* in heads, 49 ppm miRNA vs. 72 ppm miRNA* in ovaries) showed distinctly different miRNA/miRNA* ratios in ovaries and heads. Such altered ratios may reflect different concentrations of Ago1 and Ago2 or of components of their respective Argonaute-loading machineries in the two organs.

Table III-1. Pre-miRNAs whose miRNA* strands were more abundant than their miRNAs among small RNAs isolated from fly heads and fly ovaries.

Pre-miRNA	miRNA reads (ppm)	miRNA* reads (ppm)
Fly heads		
miR-10	771	1861
miR-1012	219	269
miR-193	211	1,771
miR-281-2	239	390
miR-5	15	17
miR-92a	24	106
miR-988	260	300
Fly ovaries		
miR-10	29	83
miR-1012	19	24
miR-276b	17	45
miR-281-2	240	252
miR-284	49	72

Our analyses show that nearly all miRNA and miRNA* strands sequenced in a total small RNA library correspond to species loaded into Ago1 or Ago2. Ago1 and Ago2 initially bind duplex small RNAs that subsequently separate, leading to one small RNA being retained as a guide and the other being discarded and destroyed. The identity of the destroyed strand, i.e., the passenger strand, can only be loosely inferred from small RNA sequencing data, because the accumulation of both miRNA and miRNA* strands in total libraries reflects their loading as guide RNAs, not their accumulation as discarded passenger strands. We attempted to infer the identity of these passenger strands by searching published high throughput *Drosophila* small RNA libraries—our own and

those of others—for the loop fragments that result from Dicer-1 cleavage of pre-miRNA. Such loop fragments have the potential to reveal the site of Dicer cleavage and therefore might better define the pre-miRNA-derived small RNA that is initially paired to an Argonaute-loaded miRNA or miRNA* strand.

We analyzed 70 independent small RNA libraries comprising > 66 million non-ncRNA, genome-mapping small RNA reads. We detected loop reads for 80 pre-miRNAs. For most of these small RNAs, the loop-based strategy predicted the same base-pairing profile produced by annealing the most abundant miRNA isoform with the most abundant miRNA* isoform present in our total wild-type small RNA library. We conclude that pairing the most abundant passenger strand isoform to the corresponding miRNA or miRNA* is a good approximation of the miRNA/miRNA* or miRNA*/miRNA duplex used as the substrate for Argonaute loading.

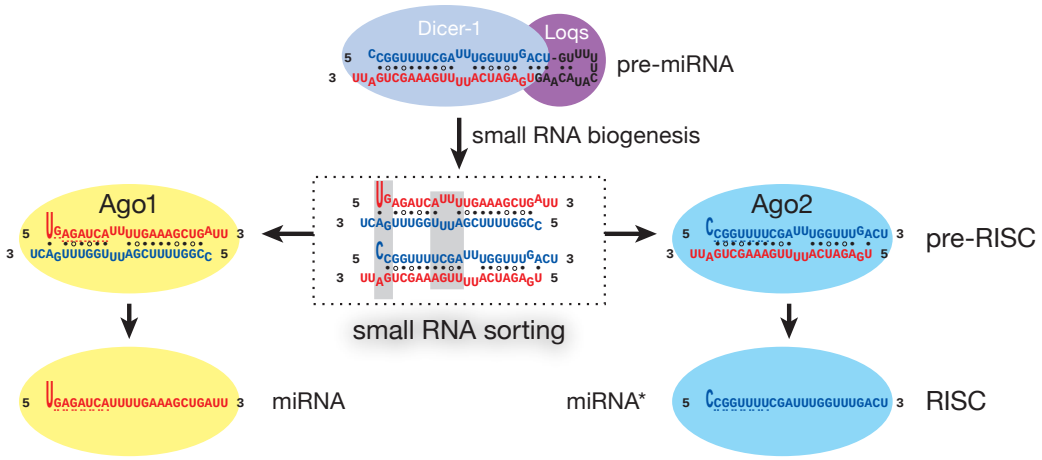
Sorting combines structure and sequence information

In general, miRNAs associate with Ago1 and miRNA* strands associate with Ago2 in *Drosophila*. It is important to note that our data argues strongly against a model in which miRNA* strands bind Ago2 as a consequence of the corresponding miRNA binding Ago1. First, we can identify six miRNAs in which both the miRNA and the miRNA* strand are enriched in Ago1 complexes in fly heads. Second, we find 15 miRNAs for which both the miRNA and the miRNA* strands are enriched in Ago2 complexes. Our data suggest that each miRNA/miRNA* duplex presents two distinct structures to the sorting machinery: One in which the miRNA is the presumptive guide and one in which

the miRNA* assumes that position. Evolution appears to have selected for miRNA/miRNA* duplexes that present sequence and structural features appropriate for loading Ago1 while simultaneously favoring Ago2 loading when the same duplex is viewed from the perspective of the miRNA*. Consequently, miRNA generally load into Ago1, whereas miRNA* load into Ago2, an Argonaute protein previously thought to act only in the RNAi pathway. miRNA* are therefore the first class of *Drosophila* small silencing RNAs produced by Dicer-1, but preferentially loaded into Ago2 (Fig. 8).

Figure III-8. A model for small RNA sorting. Sorting of small RNA into an Argonaute is governed by structure and first nucleotide identity. Consequently, a single miRNA/miRNA* duplex derived from a single pre-miRNA can present two distinct structures to the Argonaute-loading machinery. From one end, the duplex can act as a favorable substrate for loading Ago1, while from the other end, its structure and sequence can favor entry into the RNAi—i.e., the Ago2—pathway.

Figure III-8.


Dicer-1
Loqs
pre-miRNA
small RNA biogenesis
Ago1
Ago2
pre-RISC
small RNA sorting
miRNA
miRNA*
RISC

\

miRNA/miRNA* duplexes that preferentially load Ago1 are typically less stably paired at their 5´ ends and contain central mismatches, bulges, or G:U wobble pairs, whereas miRNA*/miRNA duplex that preferentially load Ago2 possess more stably paired 5´ ends and center, but have less stably paired 3´ ends. In addition to structure, sequence also plays a role in small RNA sorting in flies. Ago1-bound miRNAs begin overwhelmingly with uridine, whereas Ago2-bound miRNA, miRNA*, and siRNA tend to begin with cytidine. Moreover, our in vitro cross-linking experiments show that a 5´ U increased the efficiency of miRNA loading into Ago1, relative to a 5´ C, whereas a 5´ C—in the context of an siRNA duplex—increased the efficiency of Ago2 loading, relative to a 5´ U.

The 5´ terminal nucleotide of a small RNA is anchored in the phosphate-binding pocket of Argonaute proteins and unavailable for base pairing with its RNA target[274,275]. We speculate that the structures of the Ago1 and Ago2 discriminate between U and C by making specific hydrogen-bonding contacts with the edges of the first base of a small RNA guide.

The fate of a miRNA/miRNA* duplex, therefore depends on multiple factors; structure of its duplex, thermodynamic stability of the ends of the duplex and the identity of its 5´ terminal nucleotide. We do not yet know to what extent each factor weighs in the sorting decision.

miRNA loci appear to generate an extraordinary diversity of functional small RNAs. Some miRNA genes are transcribed from both DNA strands, producing two different hairpins from a single genomic locus[276,277]. A few miRNA have been annotated

as producing functional small RNAs—miRNA-5p and miRNA-3p—from a single pre-miRNA, a phenomenon that we suggest may be the rule rather than the exception. Our data argue that the two small RNAs, typically annotated as miRNA and miRNA*, from a single pre-miRNA partition into distinct effector proteins, with the miRNA loading into Ago1 and the miRNA* loading into Ago2. These Ago2-loaded miRNA*s are present at levels comparable to exo-siRNAs. Moreover, Ago2-loaded small RNAs can guide either target cleavage or translational repression[278], suggesting that Ago2-loaded miRNA* function to regulate as yet to be identified target RNAs. Finally, we find that a single arm of a single pre-miRNA hairpin can give rise to several functional RNA isoforms that possess different seed sequence and that associate with different Argonaute proteins that have distinct biological activities. These three layers of functional diversification—multiple small RNAs that partition differently from the two sides of the stem of a single pre-miRNA, different seed isoforms from a single side of a pre-miRNA stem, and distinct partitioning of these RNA seed isoforms—allows a single, compact genomic locus, the miRNA gene, to produce multiple riboregulators, each with a distinct biological activity and target repertoire.

Materials and Methods

General methods

Fly strains were wild-type Oregon R, *dcr-2*^{L811fsX}, *r2d2*^{1}, and *ago2*^{414}. Fly heads were isolated by vigorous shaking of liquid nitrogen-frozen flies in nested, pre-chilled sieves (U.S.A. standard sieve, Humboldt MFG, Chicago, IL), allowing the heads to pass through

the top sieve (No. 25), and collecting them on the bottom sieve (No. 40).

Small RNA sequencing

Total RNA was extracted with the mirVana kit (Ambion, Austin, TX, USA)), then 18-to-30 nt long RNA was gel purified. 2S rRNA was depleted as described [244]. A part of the sample was then oxidized using sodium periodate[83] without β-elimination step. Size-selected RNA derived from at least 68 μg total RNA for oxidation; and size-selected RNA derived from at least 7 μg total RNA for untreated. Library preparation was as described previously[26]. High throughput sequencing was by Genome Analyzer II (Illumina, San Diego, CA, USA).

Preparation of fly head extract

Isolated fly heads were transferred to 1.5 ml micro centrifuge tubes, pre-chilled in liquid-nitrogen, and homogenized using a plastic “pellet pestle” (Kontes, Vineland, NJ, USA) in 1 ml ice-cold Lysis Buffer (100 mM potassium acetate, 30 mM HEPES-KOH at pH 7.4, 2 mM magnesium acetate) containing 5 mM DTT and 1 mg/ml complete “mini” EDTA-free protease inhibitor tablets (Roche Applied Science, Indianapolis, IN, USA) per gram of heads. Lysate was clarified by centrifugation at 14,000 × g for 30 min at 4°C. The supernatant was dispensed into pre-chilled micro centrifuge tubes, flash frozen in liquid nitrogen, and stored at –80°C. Total protein concentration was determined by Bradford assay.

Immunoprecipitation

For small RNA cloning, immunoprecipitation of Ago1 protein was essentially as described[81]. Briefly, 40 µl GammaBind beads (GE Healthcare : #17-0885-01) were washed four times with 1 ml of Lysis Buffer with DTT and protease inhibitors and containing 0.5% v/v NP-40, then incubated with 40 µl monoclonal anti-Ago1 antibody[81] in 1 ml Lysis Buffer at 4°C for 3 h. After washing 5 times with 1 ml of Lysis-IP buffer, the antibody-bound beads were incubated with 910 µl fly head lysate (~4.55 mg total protein) at 4°C for 16 h, and then the supernatant collected and the beads washed 5 times with 1 ml of RIPA buffer (50 mM Tris (pH 8.0), 1.0% v/v NP-40, 150 mM NaCl, 0.5% v/v DOC, 0.1% v/v SDS, 1x Complete-EDTA-free protease inhibitor cocktail tablet). Immunoprecipitation efficiency was confirmed by Western blotting.

UV cross-linking

UV cross-linking was performed in embryo lysates prepared as described[279]. Embryo lysates were immuno-depleted for Ago1 as described above. UV cross-linking was as previously described[202], except that the samples were ~0.5 cm from the UV lamp.

Computational analyses

For each sequence read, the first occurrence of the 6-mer perfectly matching the 5´ end of the 3´ linker was identified. Sequences without a linker match were discarded. The extracted inserts for sequences that contained the 3´ linker were then mapped to the female Drosophila melanogaster genome (Release R5.5). Inserts that matched perfectly

and completely to the genome were collected using either Bowtie[280] or in-house suffix tree-based software, and the corresponding genomic coordinates were determined for downstream functional analysis. Sequences corresponding to pre-miRNA hairpins (miRBase, 13.0) or non-coding RNAs (ncRNAs; Table S3) were identified using the same suffix tree-based software. Gene were retrieved from FlyBase (R5.5). We manually curated mature miRNA*. Mature miRNA annotations were obtained from miRBase (13.0). We allowed sequence reads to differ in 5´ and 3´ ends from mature miRNA or miRNA* for up to 9 nt. Endogenous siRNA (endo-siRNA) were defined as genome mapping 21-mers detected in the oxidized library and that did not map to ncRNA or miRNA hairpins. Exogenous siRNA (exo-siRNA) were 21-mers detected in the oxidized library and that mapped perfectly to the *white* inverted repeat. Except for Fisher's exact test, which requires raw sequence reads, all sequence reads are reported in parts per million (ppm) reads of sequencing depth, with the sequencing depth defined as total number of linker containing, genome-matching reads excluding ncRNAs.

Fisher's exact test was applied to each miRNA or miRNA* to identify those that are enriched in Ago1 or Ago2. Take miR-1 as an example, the 2x2 contingency table includes the following cells: number of reads of miR-1 in detected in the library prepared from the Ago1 immunoprecipitate, number of reads of all other miRNA or miRNA* in this library, number of reads of miR-1 in the library prepared from oxidized small RNA, number of reads of all other miRNA or miRNA* in the oxidized library. p-values ≤ 0.01 were deemed significant. Furthermore, we required a miRNA or miRNA* enriched in an Argonaute protein to be at least 10 ppm in that Ago. Enrichment score (Fig. 1C) was

defined as the number of reads of a particular miRNA or miRNA* in one Argonaute versus the other. A pseudo count (or an informed prior in Bayesian statistics) of 10 ppm was used to control noise arising from extremely low abundance. For example, for miR-1 the enrichment score was [(Number of miR-1 reads in Ago1 + 10) / (total number of all miRNA reads in Ago1)]/ [(Number of miR-1 in Ago2 + 10) / (total number of all miRNA reads in Ago2)]. Similarly, the fold change in a mutant compared with the wild-type, again using miR-1 as an example, was defined as (Number of miR-1 reads in the mutant + 10) / (total number of miR-1 reads in the wild-type), where 10 ppm was the pseudo count.

Pairing probabilities were calculated using RNAcofold (ViennaRNA-1.8.3, http://www.tbi.univie.ac.at/RNA/). For each Argonaute-enriched miRNA or miRNA*, the most abundant isoform for that miRNA or miRNA* was chosen to be the guide strand and the corresponding passenger was taken to be the most abundant isoform of the miRNA* or miRNA from the wild-type untreated experiment (see Supplemental Discussion for empirical support for this approach). Both the guide and passenger were required to pass the aforementioned 10ppm threshold. The probability per position was the sum of the pairing probabilities for that position. Pairing probability for each position was smoothed by the values of the two neighboring nucleotides. For each position, we tested the significance of the difference between all Ago1-enriched miRNA and miRNA* together and all Ago2-enriched miRNA and miRNA* together using the two-sided Wilcoxon ranked-sum test with 0.01 as the threshold for significance.

To compute first nucleotide bias, we used an egalitarian weighting scheme to account for the difference in transcriptional and processing efficiency for different miRNA and miRNA*. The isoforms for a particular miRNA or miRNA* were weighted by their abundance in a data set, then all miRNA and miRNA* were weighted equally. Because *white* exo-siRNAs are produced from the same transcript, we weighted all exo-siRNA sequences by their abundance.

Supplemental Materials

Supplemental Tables

Table III-S1A. Sequencing statistics: reads. Ago2-loaded miRNAs or miRNA* strands were detected by oxidation of small RNAs prior to library construction. Ago1-loaded small RNAs were enriched by immunoprecipitation (I.P.) using a monoclonal antibody specific for Ago1protein. “Total small RNA reads” correspond to genome-matching reads after excluding annotated non-coding RNAs (ncRNAs), such as rRNA, snRNA, snoRNA, or tRNA. Supplemental Table 3 lists the ncRNAs whose sequences were excluded from small RNA reads.

Genotype	Library preparation	Total Reads	Reads perfectly matching genome	Reads matching annotated ncRNAs	Small RNA reads (excluding ncRNAs)	Pre-miRNA–matching reads	Reads excluding ncRNAs & pre-miRNA–matching
Oregon R	untreated	7,307,082	2,072,453	474,124	1,598,329	1,442,072	156,257
Oregon R	oxidized	1,400,012	566,747	6,271	560,476	298,462	262,014
Oregon R	Ago1 I.P.	6,609,187	5,159,876	124,419	5,035,457	4,975,624	59,833
w ; ; ago2^{414}	untreated	1,945,285	530,532	35,802	494,730	474,892	19,838
yw ; dcr-2^{L811fsX}/CyO	untreated	4,315,808	2,425,592	824,657	1,600,935	1,232,982	367,953
yw ; dcr-2^{L811fsX}/CyO	oxidized	1,901,642	540,789	34,956	505,833	222,032	283,801
yw ; dcr-2^{L811fsX}	untreated	2,229,996	1,453,332	157,038	1,296,294	1,251,929	44,365
yw ; dcr-2^{L811fsX}	oxidized	1,208,997	280,052	44,207	235,845	84,412	151,433
r2d2^{1}/CyO *; white-IR*	untreated	6,537,590	1,621,311	239,930	1,381,381	1,251,294	130,087
r2d2^{1}/CyO *; white-IR*	oxidized	6,816,758	735,527	6,112	729,415	266,347	463,068
r2d2^{1} ; white-IR	untreated	7,812,088	3,024,255	635,420	2,388,835	2,165,030	223,805
r2d2^{1} ; white-IR	oxidized	5,557,111	922,828	88,442	834,386	292,038	542,348

Table III-S1B. Sequencing statistics: species. Ago2-loaded miRNAs or miRNA* strands were detected by oxidation of small RNAs prior to library construction. Ago1-loaded small RNAs were enriched by immunoprecipitation (I.P.) using a monoclonal antibody specific for Ago1protein. "Total small RNA species" correspond to genome-matching species after excluding annotated non-coding RNAs (ncRNAs), such as rRNA, snRNA, snoRNA, or tRNA.

Genotype	Library preparation	Total Species	Species perfectly matching genome	Species matching annotated ncRNAs	Small RNA Species (excluding ncRNAs)	Pre-miRNA–matching Species	Species excluding ncRNAs & pre-miRNA–matching
Oregon R	untreated	2,456,441	136,561	46,167	90,394	2,512	87,882
Oregon R	oxidized	259,814	113,249	2,389	110,860	1,641	109,219
Oregon R	Ago1 I.P.	457,656	48,035	22,566	25,469	2,626	22,843
w ; ; ago2^{414}	untreated	647,247	29,375	12,933	16,442	1,384	15,058
yw ; dcr-2^{L811fsX}/CyO	untreated	1,194,213	317,577	55,615	261,962	2,795	259,167
yw ; dcr-2^{L811fsX}/CyO	oxidized	576,310	196,391	10,654	185,737	1,739	183,998
yw ; dcr-2^{L811fsX}	untreated	304,636	60,345	29,135	31,210	2,509	28,701
yw ; dcr-2^{L811fsX}	oxidized	363,108	117,042	8,604	108,438	1,504	106,934
r2d2^{1}/CyO *; white-IR*	untreated	2,212,159	89,505	33,376	56,129	2,019	54,110
r2d2^{1}/CyO *; white-IR*	oxidized	3,114,299	134,995	2,090	132,905	1,282	131,623
r2d2^{1} ; white-IR	untreated	2,272,831	130,214	44,171	86,043	2,233	83,810
r2d2^{1} ; white-IR	oxidized	858,049	167,166	11,329	155,837	1,551	154,286

Table III-S2. miRNA and miRNA* significantly enriched or depleted in Ago1 or Ago2 using Fisher's exact test. Odds ratio was defined as [(the number of reads in Ago1 for an individual miRNA or miRNA*)(the number of reads for every other miRNA or miRNA* in Ago2)]/(the number of reads in Ago2 for that individual miRNA or miRNA*)(the number of reads in Ago1 for every other miRNA or miRNA*]. Enrichment was defined as [(the number of reads in Ago1 for an individual miRNA or miRNA* + 10)(the number of reads for all miRNA or miRNA* in Ago2 + 10)]/(the number of reads in Ago2 for that individual miRNA or miRNA* +10)(the number of reads in Ago1 for all miRNA or miRNA* + 10)].

miRNAs enriched in Ago1

name	***p*-value**	**odds ratio**	**enrich-*ment***
bantam	5.8E-65	10.7	10.3
let-7	4.0E-226	1.7	1.6
miR-1	0	2.2	2.2
miR-2a-1	0	11.1	10.6
miR-2a-2	0	9.7	9.3
miR-2b-1	0	9.4	8.9
miR-2b-2	4.0E-31	9.6	9.0
miR-2c	0	6.6	4.3
miR-7	1.2E-321	7.3	7.1
miR-9a	8.2E-106	5.4	5.2
miR-9b	4.8E-10	12.4	9.1
miR-12	3.8E-88	1.7	1.6
miR-13a	0	3.6	3.4
miR-13b-1	0	8.4	8.0
miR-13b-2	0	8.4	8.0
miR-14	2.8E-132	30.9	27.0
miR-31a	0	10.6	8.6
miR-34	0	22.3	19.9
miR-124	9.6E-41	7.9	7.5
miR-125	0	1.3	1.3
miR-133	1.5E-12	15.4	14.5
miR-137	0	2.8	2.2
miR-184	5.7E-44	3.1	3.0
miR-190	2.3E-09	1.6	1.6
miR-219	0	1.5	1.4
miR-274	4.8E-159	23.1	21.9
miR-275	0	4.7	4.4
miR-276a	0	6.8	6.3
miR-276b	0	7.0	6.7
miR-278	2.9E-252	9.9	9.5
miR-279	2.5E-32	3.3	3.3
miR-281-1	1.7E-32	14.1	5.9
miR-281-2	1.7E-186	14.2	5.9
miR-282	2.4E-04	8.1	7.3
miR-286	1.8E-43	11.1	1.1
miR-304	4.0E-82	3.3	3.0
miR-305	1.2E-77	1.3	1.3
miR-306	0	4.0	3.7
miR-307	0	32.1	26.3
miR-308	3.7E-03	43.1	35.2
miR-311	7.5E-14	8.0	0.8
miR-314	1.2E-50	34.7	3.2
miR-316	0	6.0	4.7
miR-317	1.4E-16	24.9	23.0
miR-318	3.9E-19	Inf	3.8
miR-929	1.8E-57	2.8	2.4
miR-932	6.9E-12	4.0	3.6
miR-956	6.4E-242	29.9	2.8
miR-957	1.4E-07	11.9	10.4
miR-958	8.3E-07	Inf	1.7
miR-965	1.1E-04	1.7	1.5

miR-969	1.9E-09	Inf	1.0
miR-971	5.8E-14	7.9	2.3
miR-980	0	1.9	1.8
miR-981	9.5E-179	15.0	13.6
miR-987	6.5E-04	4.6	4.4
miR-989	6.1E-21	Inf	0.8
miR-990	2.2E-57	52.0	4.8
miR-993	0	10.3	6.8
miR-996	1.3E-68	10.3	9.7
miR-999	5.7E-20	1.8	1.7

miR-1001	1.5E-30	7.4	3.7
miR-1003	2.2E-18	3.6	3.0
miR-1004	1.0E-03	Inf	4.2
miR-1005	4.8E-07	3.8	1.1
miR-1006	1.1E-05	2.2	1.7
miR-1007	6.3E-04	8.5	1.5
miR-1009	2.7E-181	2.0	1.3
miR-1010	7.3E-04	13.3	10.8
miR-1013	8.5E-41	5.9	1.0
miR-1017	5.8E-65	19.6	7.4

*miRNA*s enriched in Ago1*

name	***p*-value**	**odds ratio**	**enrich-*ment***
miR-2c*	1.5E-03	4.3	1.0
miR-34*	0	24.8	21.4
miR-133	2.3E-22	55.4	5.1
miR-125*	1.1E-03	2.8	1.2
miR-193*	1.0E-10	1.7	1.6
miR-210*	5.6E-95	6.6	5.7
miR-281-2*	8.3E-47	10.0	6.2
miR-307	1.1E-58	2.8	2.7
miR-954	1.3E-05	Inf	1.2

miRNAs enriched in Ago2

name	***p*-value**	**odds ratio**	**enrich-*ment***
miR-8	0	1.7	1.7
miR-9c	2.6E-211	4.2	3.6
miR-11	0	4.0	3.4
miR-33	0	8.8	6.2
miR-79	0	6.0	4.7
miR-92a	1.1E-108	23.6	10.6
miR-92b	1.3E-66	7.4	5.7
miR-100	1.4E-41	1.6	1.5
miR-193	0	12.0	7.5
miR-210	0	1.4	1.4
miR-252	0	9.2	6.4
miR-263a	0	6.8	5.2
miR-263b	2.0E-13	1.5	1.5
miR-277	0	1.5	1.5
miR-283	1.2E-106	3.4	3.0
miR-284	1.8E-104	2.4	2.3

miR-8	0	1.7	1.7
miR-9c	2.6E-211	4.2	3.6
miR-11	0	4.0	3.4
miR-33	0	8.8	6.2
miR-79	0	6.0	4.7
miR-92a	1.1E-108	23.6	10.6
miR-92b	1.3E-66	7.4	5.7
miR-100	1.4E-41	1.6	1.5
miR-193	0	12.0	7.5
miR-210	0	1.4	1.4
miR-252	0	9.2	6.4
miR-263a	0	6.8	5.2
miR-263b	2.0E-13	1.5	1.5
miR-277	0	1.5	1.5
miR-283	1.2E-106	3.4	3.0
miR-284	1.8E-104	2.4	2.3

miRNA*s enriched in Ago2

name	*p*-value	odds ratio	enrich-ment
*bantam**	0	28.7	11.3
*let-7**	0	46.5	13.1
miR-1*	1.3E-23	19.1	10.7
miR-2a-1*	5.0E-155	75.0	14.6
miR-2a-2*	0	326.1	17.4
miR-2b-2*	0	28.5	11.2
miR-7*	2.6E-36	12.3	8.2
miR-8*	0	7.2	5.4
miR-9a*	0	9.6	6.5
miR-10*	7.7E-03	1.2	1.2
miR-11*	8.4E-04	2.2	2.7
miR-13a*	2.7E-148	55.8	13.8
miR-13b-1*	1.6E-17	23.3	11.9
miR-13b-2*	8.8E-79	46.7	13.4
miR-14*	0	17.4	9.1
miR-31a*	1.2E-17	3.7	3.6
miR-33*	0	325.0	17.0
miR-92a*	0	4193.2	17.7
miR-100*	0	183.1	16.2
miR-124*	3.9E-198	64.5	14.2
miR-184*	5.3E-96	10.1	6.9
miR-190*	3.0E-50	6.6	5.3
miR-252*	8.9E-03	1.2	1.2
miR-263a*	1.4E-04	4.1	5.5
miR-274*	3.6E-52	16.7	9.4
miR-275*	8.7E-197	18.9	9.6
miR-276a*	0	50.5	13.5
miR-276b*	0	50.5	13.5
miR-277*	6.7E-05	1.4	1.4
miR-278*	4.6E-39	4.4	3.9
miR-282*	0	78.0	14.8
miR-284*	0	85.4	14.9
miR-285*	8.4E-128	11.7	7.5
miR-304*	4.1E-10	5.2	5.3
miR-305*	8.2E-44	2.7	2.5
miR-306*	5.2E-46	4.5	3.9
miR-308*	0	699.6	17.7
miR-315*	0	793.0	17.4
miR-927*	4.2E-04	2.2	2.8
miR-929*	0	10.3	6.8
miR-932*	6.4E-30	3.7	3.4
miR-965*	1.8E-20	5.3	4.8
miR-981*	1.1E-21	7.9	6.5
miR-988*	0	502.4	17.2
miR-995*	3.4E-23	10.4	7.7
miR-996*	0	112.4	15.4
miR-998*	0	1182.8	17.5
miR-1012*	0	16.6	8.9
miR-1010*	9.8E-129	450.2	17.1

Table III-S3. Non-coding RNAs (ncRNAs) excluded prior to small RNA analyses.

Unannotated rRNAs

Name	Sequence
2S rRNA variant	5′-UGCUUGGACUACACAUGGUUGAGGGUUGUA-3′
2S rRNA variant	5′-UGCUUGGACUACAUAUGGUUGAGGGUUGGA-3′
2S rRNA variant	5′-UGCUUGGACUACAUAUGGUUGAGGGUUGUA-3′
5′ extended 5.8S	5′-AAACUCUAAGCGGUGGAU-3′
5′ extended 5.8S	5′-AAAACUCUAAGCGGUGGAU-3′
5′ extended 5.8S	5′-UAAAACUCUAAGCGGUGGAU-3′
5′ extended 5.8S	5′-UAUAAAACUCUAAGCGGUGGAU-3′
5′ extended 5.8S	5′-UUAUAAAACUCUAAGCGGUGGAU-3′

GenBank annotated RNAs

Name	GenBank I.D.	Description	Locus
5.8S rRNA	M21017.1	*D. melanogaster* 18S, 5.8S 2S and 28S rRNA genes, complete, and 18S rRNA gene, 5′ end, clone pDm238	DRORGAB: 2722-2844
18S rRNA	M21017.1	D. *melanogaster* 18S, 5.8S 2S and 28S rRNA genes, complete, and 18S rRNA gene, 5′ end, clone pDm238	DRORGAB: 1-1973

FlyBase annotated RNAs

FlyBase I.D.	Type	Name	Length (nt)
FBtr0111041	snoRNA	snoRNA:Me28S-C3420a-RA	91
FBtr0111042	snoRNA	snoRNA:Me28S-C3420b-RA	91
FBtr0111039	snRNA	snRNA:U11-RA	275
FBtr0077222	snoRNA	snoRNA:Z30-RA	91
FBtr0070292	snoRNA	snoRNA:M-RA	99
FBtr0078834	snRNA	snRNA:U4atac:82E-RA	121
FBtr0084651	snRNA	snRNA:U6:96Ab-RA	107
FBtr0086856	snoRNA	snoRNA:U27:54Eb-RA	72
FBtr0076634	snoRNA	snoRNA:U49:66Da-RA	80
FBtr0074208	snRNA	snRNA:U2:14B-RA	192
FBtr0079659	snRNA	snRNA:U6atac:29B-RA	97
FBtr0084528	snRNA	snRNA:U1:95Ca-RA	164
FBtr0084652	snRNA	snRNA:U6:96Ac-RA	107
FBtr0078028	snRNA	snRNA:U1:21D-RA	172
FBtr0080486	snRNA	snRNA:U2:34ABa-RA	192
FBtr0086347	rRNA	5SrRNA:CR33355-RA	135
FBtr0086362	rRNA	5SrRNA:CR33370-RA	135
FBtr0086372	rRNA	5SrRNA:CR33380-RA	135
FBtr0086373	rRNA	5SrRNA:CR33381-RA	135
FBtr0086374	rRNA	5SrRNA:CR33382-RA	135

FBtr0086375	rRNA	5SrRNA:CR33383-RA	135
FBtr0086380	rRNA	5SrRNA:CR33388-RA	135
FBtr0086381	rRNA	5SrRNA:CR33389-RA	135
FBtr0086390	rRNA	5SrRNA:CR33398-RA	135
FBtr0086391	rRNA	5SrRNA:CR33399-RA	135
FBtr0086393	rRNA	5SrRNA:CR33401-RA	135
FBtr0086441	rRNA	5SrRNA:CR33449-RA	135
FBtr0100848	snRNA	snRNA:U7-RA	71
FBtr0091605	snoRNA	snoRNA:U3:54Ab-RA	173
FBtr0091629	snoRNA	snoRNA:14-RA	108
FBtr0091635	snoRNA	snoRNA:3-RA	16
FBtr0091740	snRNA	snmRNA:430:CR33742-RA	36
FBtr0091741	snRNA	snmRNA:430:CR33743-RA	36
FBtr0091742	snRNA	snmRNA:430:CR33744-RA	36
FBtr0091743	snRNA	snmRNA:430:CR33745-RA	36
FBtr0091744	snRNA	snmRNA:430:CR33746-RA	36
FBtr0091788	snoRNA	snoRNA:644-RA	81
FBtr0091697	snoRNA	snoRNA:165-RA	53
FBtr0091739	snRNA	snmRNA:430:CR33741-RA	36
FBtr0091766	snoRNA	snoRNA:66-RA	137
FBtr0079910	snoRNA	snoRNA:U14:30Eb-RA	81
FBtr0091922	snoRNA	snoRNA:734-RA	133
FBtr0078851	snRNA	snRNA:U1:82Eb-RA	255
FBtr0086421	rRNA	5SrRNA:CR33429-RA	135
FBtr0091798	snoRNA	snoRNA:684-RA	78
FBtr0091781	snoRNA	snoRNA:660-RA	96
FBtr0091789	snRNA	snRNA:U2:34ABc-RA	192
FBtr0091752	snoRNA	snoRNA:708-RA	53
FBtr0091751	snoRNA	snoRNA:328-RA	68
FBtr0091755	snoRNA	snoRNA:50-RA	160
FBtr0091754	snoRNA	snoRNA:586-RA	80
FBtr0091708	snoRNA	snoRNA:755-RA	111
FBtr0091677	snoRNA	snoRNA:72-RA	86
FBtr0086392	rRNA	5SrRNA:CR33400-RA	135
FBtr0091664	snoRNA	snoRNA:229-RA	140
FBtr0086394	rRNA	5SrRNA:CR33402-RA	135
FBtr0086397	rRNA	5SrRNA:CR33405-RA	134
FBtr0086401	rRNA	5SrRNA:CR33409-RA	135
FBtr0086402	rRNA	5SrRNA:CR33410-RA	135
FBtr0086403	rRNA	5SrRNA:CR33411-RA	135
FBtr0086404	rRNA	5SrRNA:CR33412-RA	135
FBtr0086406	rRNA	5SrRNA:CR33414-RA	135
FBtr0086410	rRNA	5SrRNA:CR33418-RA	135
FBtr0086411	rRNA	5SrRNA:CR33419-RA	135
FBtr0086413	rRNA	5SrRNA:CR33421-RA	135
FBtr0086414	rRNA	5SrRNA:CR33422-RA	135
FBtr0091623	snoRNA	snoRNA:825-RA	34
FBtr0091610	snoRNA	snoRNA:203-RA	53
FBtr0086415	rRNA	5SrRNA:CR33423-RA	135
FBtr0086416	rRNA	5SrRNA:CR33424-RA	135

FBtr0086417	rRNA	5SrRNA:CR33425-RA	135
FBtr0086418	rRNA	5SrRNA:CR33426-RA	135
FBtr0091613	snoRNA	snoRNA:461-RA	102
FBtr0091602	snoRNA	snoRNA:783-RA	54
FBtr0086345	rRNA	5SrRNA:CR33353-RA	135
FBtr0086346	rRNA	5SrRNA:CR33354-RA	135
FBtr0086349	rRNA	5SrRNA:CR33357-RA	135
FBtr0086350	rRNA	5SrRNA:CR33358-RA	135
FBtr0086353	rRNA	5SrRNA:CR33361-RA	135
FBtr0086364	rRNA	5SrRNA:CR33372-RA	135
FBtr0086367	rRNA	5SrRNA:CR33375-RA	135
FBtr0086368	rRNA	5SrRNA:CR33376-RA	135
FBtr0086369	rRNA	5SrRNA:CR33377-RA	135
FBtr0086378	rRNA	5SrRNA:CR33386-RA	135
FBtr0086382	rRNA	5SrRNA:CR33390-RA	135
FBtr0086386	rRNA	5SrRNA:CR33394-RA	135
FBtr0086387	rRNA	5SrRNA:CR33395-RA	135
FBtr0086388	rRNA	5SrRNA:CR33396-RA	135
FBtr0086389	rRNA	5SrRNA:CR33397-RA	135
FBtr0080451	snRNA	snRNA:U5:34A-RA	127
FBtr0084650	snRNA	snRNA:U6:96Aa-RA	107
FBtr0075315	snRNA	snRNA:U12:73B-RA	238
FBtr0081489	snRNA	snRNA:U4:39B-RA	143
FBtr0100888	rRNA	mt:lrRNA-RA	1325
FBtr0100890	rRNA	mt:srRNA-RA	786
FBtr0078791	snRNA	snRNA:U4atac:83A-RA	122
FBtr0081560	snoRNA	snoRNA:U85-RA	316
FBtr0084488	snRNA	snRNA:U1:95Cb-RA	164
FBtr0079908	snoRNA	snoRNA:U25:30E-RA	68
FBtr0079909	snoRNA	snoRNA:U14:30Ea-RA	81
FBtr0080770	snRNA	snRNA:U5:35D-RA	126
FBtr0081293	snRNA	snRNA:U2:38ABb-RA	191
FBtr0081292	snRNA	snRNA:U4:38AB-RA	142
FBtr0081294	snRNA	snRNA:U5:38ABb-RA	127
FBtr0081315	snRNA	snRNA:U5:38ABa-RA	127
FBtr0081313	snRNA	snRNA:U2:38ABa-RA	192
FBtr0072259	snoRNA	snoRNA:H1-RA	140
FBtr0086843	snoRNA	snoRNA:U31:54Ea-RA	69
FBtr0086844	snoRNA	snoRNA:U29:54Ea-RA	87
FBtr0086845	snoRNA	snoRNA:U76:54Ea-RA	73
FBtr0086846	snoRNA	snoRNA:U29:54Eb-RA	86
FBtr0086848	snoRNA	snoRNA:U76:54Eb-RA	73
FBtr0086850	snoRNA	snoRNA:U27:54Ea-RA	69
FBtr0086851	snoRNA	snoRNA:snR38:54Ea-RA	77
FBtr0086847	snoRNA	snoRNA:U29:54Ec-RA	88
FBtr0086849	snoRNA	snoRNA:U29:54Ed-RA	87
FBtr0086852	snoRNA	snoRNA:snR38:54Eb-RA	76
FBtr0086853	snoRNA	snoRNA:U31:54Eb-RA	67
FBtr0086854	snoRNA	snoRNA:U31:54Ec-RA	67
FBtr0086855	snoRNA	snoRNA:U31:54Ed-RA	67

FBtr0086857	snoRNA	snoRNA:snR38:54Ec-RA	77
FBtr0086858	snoRNA	snoRNA:U27:54Ec-RA	67
FBtr0088037	snoRNA	snoRNA:Z1-RA	73
FBtr0078576	snoRNA	snoRNA:U21-RA	78
FBtr0073017	snRNA	snRNA:U5:63BC-RA	123
FBtr0076635	snoRNA	snoRNA:U49:66Db-RA	84
FBtr0074249	snRNA	snRNA:U5:14B-RA	110
FBtr0079108	snRNA	snRNA:U4:25F-RA	148
FBtr0077658	snRNA	snRNA:U5:23D-RA	131
FBtr0084487	snRNA	snRNA:U1:95Cc-RA	164
FBtr0080443	snRNA	snRNA:U2:34ABb-RA	192
FBtr0079907	snoRNA	snoRNA:Z5-RA	113
FBtr0086351	rRNA	5SrRNA:CR33359-RA	135
FBtr0086352	rRNA	5SrRNA:CR33360-RA	135
FBtr0086354	rRNA	5SrRNA:CR33362-RA	135
FBtr0086356	rRNA	5SrRNA:CR33364-RA	135
FBtr0086357	rRNA	5SrRNA:CR33365-RA	135
FBtr0086358	rRNA	5SrRNA:CR33366-RA	135
FBtr0086359	rRNA	5SrRNA:CR33367-RA	135
FBtr0086360	rRNA	5SrRNA:CR33368-RA	135
FBtr0086361	rRNA	5SrRNA:CR33369-RA	135
FBtr0086365	rRNA	5SrRNA:CR33373-RA	135
FBtr0086366	rRNA	5SrRNA:CR33374-RA	135
FBtr0086370	rRNA	5SrRNA:CR33378-RA	135
FBtr0086371	rRNA	5SrRNA:CR33379-RA	135
FBtr0086376	rRNA	5SrRNA:CR33384-RA	135
FBtr0086377	rRNA	5SrRNA:CR33385-RA	135
FBtr0077928	snRNA	snRNA:U3:22A-RA	211
FBtr0086379	rRNA	5SrRNA:CR33387-RA	135
FBtr0086383	rRNA	5SrRNA:CR33391-RA	134
FBtr0086384	rRNA	5SrRNA:CR33392-RA	135
FBtr0086385	rRNA	5SrRNA:CR33393-RA	135
FBtr0086395	rRNA	5SrRNA:CR33403-RA	135
FBtr0086396	rRNA	5SrRNA:CR33404-RA	135
FBtr0086398	rRNA	5SrRNA:CR33406-RA	135
FBtr0086399	rRNA	5SrRNA:CR33407-RA	135
FBtr0086400	rRNA	5SrRNA:CR33408-RA	135
FBtr0086405	rRNA	5SrRNA:CR33413-RA	135
FBtr0086407	rRNA	5SrRNA:CR33415-RA	135
FBtr0086409	rRNA	5SrRNA:CR33417-RA	135
FBtr0086412	rRNA	5SrRNA:CR33420-RA	135
FBtr0086419	rRNA	5SrRNA:CR33427-RA	135
FBtr0086420	rRNA	5SrRNA:CR33428-RA	135
FBtr0086431	rRNA	5SrRNA:CR33439-RA	135
FBtr0086422	rRNA	5SrRNA:CR33430-RA	135
FBtr0086423	rRNA	5SrRNA:CR33431-RA	135
FBtr0086424	rRNA	5SrRNA:CR33432-RA	135
FBtr0086425	rRNA	5SrRNA:CR33433-RA	135
FBtr0086426	rRNA	5SrRNA:CR33434-RA	135
FBtr0086427	rRNA	5SrRNA:CR33435-RA	135

FBtr0086428	rRNA	5SrRNA:CR33436-RA	135
FBtr0086429	rRNA	5SrRNA:CR33437-RA	135
FBtr0086430	rRNA	5SrRNA:CR33438-RA	135
FBtr0086432	rRNA	5SrRNA:CR33440-RA	135
FBtr0086433	rRNA	5SrRNA:CR33441-RA	135
FBtr0086434	rRNA	5SrRNA:CR33442-RA	135
FBtr0086435	rRNA	5SrRNA:CR33443-RA	135
FBtr0086436	rRNA	5SrRNA:CR33444-RA	135
FBtr0086437	rRNA	5SrRNA:CR33445-RA	135
FBtr0086438	rRNA	5SrRNA:CR33446-RA	135
FBtr0086439	rRNA	5SrRNA:CR33447-RA	135
FBtr0086440	rRNA	5SrRNA:CR33448-RA	135
FBtr0086442	rRNA	5SrRNA:CR33450-RA	135
FBtr0086443	rRNA	5SrRNA:CR33451-RA	135
FBtr0086444	rRNA	5SrRNA:CR33452-RA	135
FBtr0091611	snoRNA	snoRNA:122-RA	68
FBtr0091614	snoRNA	snoRNA:227-RA	29
FBtr0091615	snoRNA	snoRNA:291-RA	74
FBtr0091640	snoRNA	snoRNA:535-RA	43
FBtr0091641	snoRNA	snoRNA:U3:9B-RA	168
FBtr0091642	snoRNA	snoRNA:284-RA	67
FBtr0091653	snoRNA	snoRNA:737-RA	72
FBtr0091666	snoRNA	snoRNA:269-RA	183
FBtr0091724	snoRNA	snoRNA:700-RA	49
FBtr0091725	snRNA	snmRNA:430:CR33727-RA	36
FBtr0091727	snRNA	snmRNA:430:CR33729-RA	36
FBtr0091729	snRNA	snmRNA:430:CR33731-RA	36
FBtr0091730	snRNA	snmRNA:430:CR33732-RA	36
FBtr0091731	snRNA	snmRNA:430:CR33733-RA	36
FBtr0091732	snRNA	snmRNA:430:CR33734-RA	36
FBtr0091735	snRNA	snmRNA:430:CR33737-RA	36
FBtr0091736	snRNA	snmRNA:430:CR33738-RA	36
FBtr0091737	snRNA	snmRNA:430:CR33739-RA	36
FBtr0091738	snRNA	snmRNA:430:CR33740-RA	36
FBtr0091777	snoRNA	snoRNA:143-RA	123
FBtr0091779	snoRNA	snoRNA:442-RA	46
FBtr0091803	snoRNA	snoRNA:75-RA	38
FBtr0091917	snoRNA	snoRNA:314-RA	64
FBtr0091925	snoRNA	snoRNA:U3:54Aa-RA	173
FBtr0091934	snoRNA	snoRNA:185-RA	55
FBtr0089298	tRNA	tRNA:Y1:22Fa-RA	73
FBtr0071737	tRNA	CR30407-RA	72
FBtr0087665	tRNA	CR30509-RA	74
FBtr0086448	tRNA	tRNA:E4:56Fc-RA	72
FBtr0086449	tRNA	CR30452-RA	72
FBtr0100841	tRNA	CR30505-RA	71
FBtr0087688	tRNA	tRNA:K2:50C-RA	73
FBtr0089302	tRNA	tRNA:L3:49Fb-RA	83
FBtr0086016	tRNA	tRNA:N5:42Ah-RA	74
FBtr0087659	tRNA	tRNA:I:49Fc-RA	74

FBtr0075456	tRNA	CR32153-RA	73
FBtr0073018	tRNA	CR32288-RA	72
FBtr0073019	tRNA	CR32289-RA	72
FBtr0076739	tRNA	CR32363-RA	83
FBtr0074600	tRNA	CR32546-RA	72
FBtr0070911	tRNA	CR32748-RA	73
FBtr0086017	tRNA	tRNA:K2:42Ae-RA	73
FBtr0084597	tRNA	CR31130-RA	84
FBtr0083799	tRNA	tRNA:V3b:92Ba-RA	73
FBtr0083793	tRNA	CR31215-RA	73
FBtr0083544	tRNA	CR31242-RA	72
FBtr0082836	tRNA	CR31331-RA	73
FBtr0083977	tRNA	CR31333-RA	72
FBtr0084714	tRNA	CR31382-RA	72
FBtr0085637	tRNA	CR31383-RA	80
FBtr0084721	tRNA	CR31416-RA	72
FBtr0083792	tRNA	tRNA:V3b:92Bb-RA	73
FBtr0081885	tRNA	tRNA:Y1:85Ab-RA	73
FBtr0083794	tRNA	CR31471-RA	74
FBtr0082919	tRNA	tRNA:S2b:88A-RA	82
FBtr0083975	tRNA	CR31480-RA	72
FBtr0081894	tRNA	tRNA:Y1:85Ad-RA	73
FBtr0081810	tRNA	tRNA:R2:84Fb-RA	73
FBtr0083501	tRNA	tRNA:P:90Ca-RA	72
FBtr0083494	tRNA	CR31569-RA	73
FBtr0081811	tRNA	tRNA:R2:84Fc-RA	73
FBtr0081814	tRNA	tRNA:N5:84F-RA	74
FBtr0081812	tRNA	tRNA:R2:84Fd-RA	73
FBtr0079690	tRNA	CR31603-RA	72
FBtr0077872	tRNA	tRNA:G3:22BCa-RA	71
FBtr0077860	tRNA	CR31669-RA	72
FBtr0079694	tRNA	CR31604-RA	72
FBtr0079692	tRNA	CR31895-RA	73
FBtr0079677	tRNA	CR31896-RA	72
FBtr0077819	tRNA	CR31944-RA	72
FBtr0077812	tRNA	CR31942-RA	73
FBtr0077458	tRNA	CR31963-RA	72
FBtr0089613	tRNA	tRNA:G3:35Ba-RA	71
FBtr0080663	tRNA	tRNA:G3:35Be-RA	71
FBtr0100845	tRNA	CR33536-RA	84
FBtr0072445	tRNA	CR30198-RA	72
FBtr0072447	tRNA	CR30200-RA	72
FBtr0071983	tRNA	CR30201-RA	82
FBtr0071581	tRNA	tRNA:G3:57BCb-RA	71
FBtr0086247	tRNA	CR30211-RA	72
FBtr0086659	tRNA	tRNA:G3:55E-RA	71
FBtr0088703	tRNA	tRNA:L2:44EF-RA	83
FBtr0088787	tRNA	CR30297-RA	74
FBtr0089059	tRNA	CR30298-RA	74
FBtr0086898	tRNA	CR30333-RA	73

FBtr0088086	tRNA	tRNA:M2:48Bb-RA	73
FBtr0077862	tRNA	CR31939-RA	72
FBtr0077861	tRNA	CR31940-RA	72
FBtr0077818	tRNA	CR31943-RA	72
FBtr0077834	tRNA	CR31946-RA	73
FBtr0086332	tRNA	CR30220-RA	72
FBtr0086333	tRNA	CR30449-RA	72
FBtr0100842	tRNA	CR33534-RA	73
FBtr0091521	tRNA	tRNA:R:85Cb-RA	73
FBtr0077821	tRNA	tRNA:G3:22BCb-RA	71
FBtr0100843	tRNA	CR33535-RA	72
FBtr0077613	tRNA	tRNA:S7:23Ea-RA	82
FBtr0083693	tRNA	CR31228-RA	72
FBtr0083285	tRNA	CR31282-RA	74
FBtr0073885	tRNA	tRNA:S7:12Ed-RA	82
FBtr0073872	tRNA	tRNA:S4:12Ee-RA	82
FBtr0073870	tRNA	tRNA:S7:12Eg-RA	82
FBtr0073871	tRNA	tRNA:R:12Ef-RA	73
FBtr0073886	tRNA	tRNA:S774:12Ec-RA	82
FBtr0073865	tRNA	tRNA:S474:12Eh-RA	82
FBtr0073863	tRNA	tRNA:R:12Ed-RA	73
FBtr0073862	tRNA	tRNA:R:12Ec-RA	73
FBtr0073861	tRNA	tRNA:R:12Eb-RA	73
FBtr0073858	tRNA	tRNA:S4:12Ea-RA	82
FBtr0073860	tRNA	tRNA:R:12Ea-RA	73
FBtr0073857	tRNA	tRNA:R:12Ee-RA	73
FBtr0079528	tRNA	tRNA:G3:28D-RA	71
FBtr0077184	tRNA	tRNA:R:19F-RA	73
FBtr0080717	tRNA	tRNA:L:35C-RA	84
FBtr0080609	tRNA	tRNA:Q:34E-RA	72
FBtr0077142	tRNA	tRNA:S7:64D-RA	82
FBtr0077577	tRNA	tRNA:S7:23Eb-RA	82
FBtr0079702	tRNA	CR31892-RA	72
FBtr0079596	tRNA	tRNA:K5:29A-RA	73
FBtr0081809	tRNA	tRNA:R2:84Fa-RA	73
FBtr0081813	tRNA	tRNA:R2:84Fe-RA	73
FBtr0083267	tRNA	tRNA:F2:89BC-RA	73
FBtr0083268	tRNA	tRNA:V4:89BC-RA	73
FBtr0083495	tRNA	tRNA:V4:90C-RA	73
FBtr0084232	tRNA	CR31167-RA	82
FBtr0084482	tRNA	CR31143-RA	83
FBtr0071736	tRNA	CR30406-RA	72
FBtr0089301	tRNA	CR32520-RA	73
FBtr0089300	tRNA	CR32525-RA	73
FBtr0075713	tRNA	tRNA:M3:70Fa-RA	72
FBtr0088145	tRNA	CR30506-RA	72
FBtr0075681	tRNA	tRNA:M3:70Fb-RA	72
FBtr0086334	tRNA	tRNA:K2:56EF-RA	73
FBtr0081923	tRNA	tRNA:R:85Ca-RA	73
FBtr0081660	tRNA	tRNA:K5:84ABa-RA	73

FBtr0081661	tRNA	tRNA:K5:84ABc-RA	73
FBtr0081622	tRNA	tRNA:K5:84ABd-RA	73
FBtr0076083	tRNA	CR32093-RA	72
FBtr0082558	tRNA	CR31356-RA	72
FBtr0081561	tRNA	tRNA:V3b:84Dc-RA	73
FBtr0075831	tRNA	CR32127-RA	83
FBtr0075830	tRNA	CR32126-RA	72
FBtr0075244	tRNA	CR32173-RA	72
FBtr0082616	tRNA	CR31432-RA	72
FBtr0083317	tRNA	CR31497-RA	74
FBtr0081565	tRNA	CR31494-RA	72
FBtr0081610	tRNA	CR31491-RA	72
FBtr0081895	tRNA	tRNA:Y1:85Ae-RA	73
FBtr0075107	tRNA	CR32200-RA	72
FBtr0083545	tRNA	CR31518-RA	72
FBtr0083943	tRNA	CR31506-RA	84
FBtr0083466	tRNA	tRNA:V3b:90BC-RA	73
FBtr0083474	tRNA	CR31579-RA	73
FBtr0083473	tRNA	CR31578-RA	72
FBtr0083475	tRNA	CR31577-RA	73
FBtr0083472	tRNA	CR31573-RA	72
FBtr0083493	tRNA	CR31568-RA	73
FBtr0073020	tRNA	CR32287-RA	72
FBtr0073022	tRNA	CR32286-RA	72
FBtr0073021	tRNA	CR32285-RA	73
FBtr0073121	tRNA	CR32273-RA	73
FBtr0073125	tRNA	CR32272-RA	73
FBtr0084649	tRNA	tRNA:D:96A-RA	72
FBtr0070533	tRNA	CR32493-RA	72
FBtr0078790	tRNA	tRNA:R:83AB-RA	73
FBtr0083496	tRNA	tRNA:T:90Cb-RA	74
FBtr0083497	tRNA	tRNA:T:90Ca-RA	74
FBtr0083492	tRNA	CR31580-RA	73
FBtr0078490	tRNA	CR32449-RA	72
FBtr0078580	tRNA	CR32460-RA	72
FBtr0079693	tRNA	CR31602-RA	72
FBtr0072686	tRNA	tRNA:E4:62Ab-RA	72
FBtr0072687	tRNA	tRNA:E4:62Ac-RA	72
FBtr0072683	tRNA	CR32328-RA	72
FBtr0072688	tRNA	CR32324-RA	72
FBtr0076740	tRNA	CR32362-RA	83
FBtr0079338	tRNA	CR31631-RA	73
FBtr0075380	tRNA	CR32357-RA	73
FBtr0100015	tRNA	tRNA:N5:42Ag-RA	74
FBtr0100881	tRNA	mt:tRNA:T-RA	66
FBtr0100882	tRNA	mt:tRNA:P-RA	61
FBtr0100872	tRNA	mt:tRNA:R-RA	64
FBtr0100873	tRNA	mt:tRNA:N-RA	65
FBtr0100874	tRNA	mt:tRNA:S:AGY-RA	68
FBtr0100875	tRNA	mt:tRNA:E-RA	67

FBtr0100876	tRNA	mt:tRNA:F-RA	65
FBtr0100854	tRNA	mt:tRNA:I-RA	65
FBtr0100855	tRNA	mt:tRNA:Q-RA	69
FBtr0100856	tRNA	mt:tRNA:M-RA	69
FBtr0100887	tRNA	mt:tRNA:L:CUN-RA	65
FBtr0100889	tRNA	mt:tRNA:V-RA	73
FBtr0100878	tRNA	mt:tRNA:H-RA	66
FBtr0100885	tRNA	mt:tRNA:S:UCN-RA	66
FBtr0100869	tRNA	mt:tRNA:G-RA	65
FBtr0100871	tRNA	mt:tRNA:A-RA	65
FBtr0100862	tRNA	mt:tRNA:L:UUR-RA	66
FBtr0100864	tRNA	mt:tRNA:K-RA	71
FBtr0100865	tRNA	mt:tRNA:D-RA	67
FBtr0100858	tRNA	mt:tRNA:W-RA	66
FBtr0100859	tRNA	mt:tRNA:C-RA	62
FBtr0100860	tRNA	mt:tRNA:Y-RA	66
FBtr0081536	tRNA	tRNA:V3b:84Da-RA	73
FBtr0086446	tRNA	tRNA:E4:56Fa-RA	72
FBtr0086445	tRNA	CR30454-RA	72
FBtr0086447	tRNA	tRNA:E4:56Fb-RA	72
FBtr0085982	tRNA	tRNA:N5:42Aa-RA	74
FBtr0085981	tRNA	tRNA:N5:42Ab-RA	74
FBtr0085980	tRNA	tRNA:R2:42Ac-RA	73
FBtr0086006	tRNA	tRNA:R2:42Aa-RA	73
FBtr0086007	tRNA	tRNA:K2:42Ad-RA	73
FBtr0086013	tRNA	tRNA:K2:42Ab-RA	73
FBtr0086014	tRNA	tRNA:K2:42Ac-RA	73
FBtr0086015	tRNA	tRNA:N5:42Ac-RA	74
FBtr0086004	tRNA	tRNA:N5:42Ae-RA	74
FBtr0076505	tRNA	CR32034-RA	74
FBtr0075856	tRNA	tRNA:V4:70BCa-RA	73
FBtr0075847	tRNA	tRNA:V4:70BCb-RA	73
FBtr0075834	tRNA	tRNA:D2:69F-RA	72
FBtr0075829	tRNA	CR32128-RA	72
FBtr0075828	tRNA	CR32129-RA	72
FBtr0072882	tRNA	CR32303-RA	72
FBtr0072840	tRNA	CR32312-RA	72
FBtr0072691	tRNA	tRNA:E4:62Ae-RA	72
FBtr0072690	tRNA	tRNA:E4:62Ad-RA	72
FBtr0072685	tRNA	tRNA:E4:62Aa-RA	72
FBtr0072684	tRNA	CR32329-RA	72
FBtr0072682	tRNA	CR32330-RA	72
FBtr0076744	tRNA	CR32358-RA	83
FBtr0076743	tRNA	CR32359-RA	83
FBtr0076741	tRNA	CR32361-RA	83
FBtr0076742	tRNA	CR32360-RA	83
FBtr0076762	tRNA	CR32370-RA	73
FBtr0077143	tRNA	CR32420-RA	73
FBtr0077158	tRNA	CR32421-RA	73
FBtr0078547	tRNA	CR32456-RA	83

FBtr0072579	tRNA	CR32480-RA	74
FBtr0072580	tRNA	CR32481-RA	74
FBtr0072600	tRNA	tRNA:M-i:61D-RA	72
FBtr0073859	tRNA	tRNA:S4:12Eb-RA	82
FBtr0073866	tRNA	tRNA:S774:12Ef-RA	82
FBtr0070979	tRNA	CR32740-RA	73
FBtr0070603	tRNA	CR32785-RA	72
FBtr0070604	tRNA	tRNA:P:3E-RA	72
FBtr0070001	tRNA	CR32826-RA	72
FBtr0089303	tRNA	tRNA:L3:49Fa-RA	83
FBtr0087713	tRNA	tRNA:I:49Fa-RA	74
FBtr0085636	tRNA	CR31023-RA	80
FBtr0085086	tRNA	CR31070-RA	72
FBtr0085087	tRNA	CR31071-RA	72
FBtr0084233	tRNA	CR31162-RA	82
FBtr0084222	tRNA	CR31165-RA	82
FBtr0084223	tRNA	CR31166-RA	82
FBtr0082544	tRNA	tRNA:K5:87BC-RA	72
FBtr0081706	tRNA	tRNA:M2:83F-RA	73
FBtr0083976	tRNA	CR31334-RA	72
FBtr0082241	tRNA	tRNA:S2b:86A-RA	82
FBtr0081886	tRNA	tRNA:Y1:85Ac-RA	73
FBtr0081884	tRNA	tRNA:Y1:85Aa-RA	73
FBtr0081623	tRNA	tRNA:K5:84ABb-RA	73
FBtr0081558	tRNA	tRNA:V3b:84Dd-RA	73
FBtr0081535	tRNA	tRNA:V3b:84Db-RA	73
FBtr0083468	tRNA	CR31570-RA	73
FBtr0083469	tRNA	tRNA:P:90Cb-RA	72
FBtr0083499	tRNA	tRNA:A:90C-RA	73
FBtr0083491	tRNA	CR31575-RA	73
FBtr0083477	tRNA	CR31576-RA	73
FBtr0082834	tRNA	CR31588-RA	73
FBtr0079820	tRNA	CR31885-RA	74
FBtr0079752	tRNA	CR31888-RA	72
FBtr0079728	tRNA	CR31889-RA	72
FBtr0079729	tRNA	CR31890-RA	72
FBtr0089297	tRNA	tRNA:Y1:28C-RA	73
FBtr0079090	tRNA	CR31914-RA	72
FBtr0079064	tRNA	CR31971-RA	72
FBtr0080666	tRNA	tRNA:G3:35Bb-RA	71
FBtr0080660	tRNA	tRNA:P:35Bb-RA	72
FBtr0080664	tRNA	tRNA:G3:35Bd-RA	71
FBtr0080665	tRNA	tRNA:G3:35Bc-RA	71
FBtr0080646	tRNA	tRNA:P:35Ba-RA	72
FBtr0080647	tRNA	tRNA:P:35Bc-RA	72
FBtr0080644	tRNA	tRNA:P:35Bd-RA	72
FBtr0089299	tRNA	tRNA:Y1:22Fb-RA	73
FBtr0100849	tRNA	tRNA:N5:60C-RA	74
FBtr0100846	tRNA	CR33538-RA	73
FBtr0100847	tRNA	CR33539-RA	72

FBtr0100016	tRNA	tRNA:N5:42Af-RA	74
FBtr0088465	tRNA	tRNA:M3:46A-RA	72
FBtr0088156	tRNA	tRNA:SeC-RA	87
FBtr0086314	tRNA	tRNA:G3:56EFb-RA	71
FBtr0086234	tRNA	CR30155-RA	72
FBtr0072446	tRNA	CR30199-RA	72
FBtr0071962	tRNA	CR30202-RA	82
FBtr0071631	tRNA	CR30206-RA	72
FBtr0071630	tRNA	tRNA:G3:57BCa-RA	71
FBtr0071574	tRNA	CR30208-RA	72
FBtr0071626	tRNA	CR30209-RA	72
FBtr0086326	tRNA	CR30212-RA	73
FBtr0086312	tRNA	tRNA:G3:56EFa-RA	71
FBtr0086330	tRNA	CR30215-RA	74
FBtr0086331	tRNA	CR30218-RA	72
FBtr0086498	tRNA	CR30223-RA	72
FBtr0086526	tRNA	tRNA:S4:56D-RA	82
FBtr0086527	tRNA	CR30225-RA	73
FBtr0086603	tRNA	CR30227-RA	72
FBtr0086611	tRNA	CR30229-RA	82
FBtr0086835	tRNA	CR30231-RA	73
FBtr0086965	tRNA	CR30232-RA	72
FBtr0086908	tRNA	tRNA:H:56E-RA	72
FBtr0087001	tRNA	CR30234-RA	82
FBtr0086988	tRNA	CR30235-RA	82
FBtr0087055	tRNA	tRNA:G3:53E-RA	71
FBtr0087198	tRNA	CR30237-RA	72
FBtr0087197	tRNA	CR30238-RA	72
FBtr0087128	tRNA	CR30239-RA	72
FBtr0087129	tRNA	CR30240-RA	72
FBtr0087425	tRNA	CR30241-RA	82
FBtr0087658	tRNA	tRNA:I:49Fb-RA	74
FBtr0087660	tRNA	tRNA:I:49Fd-RA	74
FBtr0087661	tRNA	tRNA:I:49Fe-RA	74
FBtr0087963	tRNA	CR30249-RA	72
FBtr0087962	tRNA	CR30250-RA	72
FBtr0087961	tRNA	CR30251-RA	72
FBtr0087900	tRNA	tRNA:H:48F-RA	72
FBtr0088087	tRNA	CR30254-RA	73
FBtr0088071	tRNA	tRNA:M2:48Ba-RA	73
FBtr0088131	tRNA	CR30257-RA	72
FBtr0100840	tRNA	CR30260-RA	71
FBtr0089058	tRNA	CR30299-RA	74
FBtr0086190	tRNA	tRNA:K2:42Ea-RA	73
FBtr0086189	tRNA	tRNA:K2:42Eb-RA	73
FBtr0086188	tRNA	tRNA:K2:42Ec-RA	73
FBtr0086187	tRNA	tRNA:K2:42Ed-RA	73
FBtr0085979	tRNA	tRNA:R2:42Ad-RA	73
FBtr0085983	tRNA	tRNA:K2:42Aa-RA	73
FBtr0085984	tRNA	tRNA:I:42A-RA	74

FBtr0085985	tRNA	tRNA:R2:42Ab-RA	73
FBtr0085986	tRNA	CR30316-RA	73
FBtr0085989	tRNA	tRNA:N5:42Ad-RA	74
FBtr0086528	tRNA	CR30326-RA	73

CHAPTER IV

Argonaute loading contributes to the precision of the 5´ ends of both microRNAs and their miRNA* strands in flies

The following chapter was a collaborative effort. The author and Hervé Seitz performed experiments and analyses, respectively, demonstrating the 5´ homogeneity of both miRNAs and miRNA* stands in flies. We also proposed a role for Ago2 loading in purifying 5´ends of miRNA and miRNA* sequences. The author, Hervé Seitz and Phillip Zamore wrote the paper. This chapter appeared in Curr Biol. 2008 Jan 22;18(2):147-51.

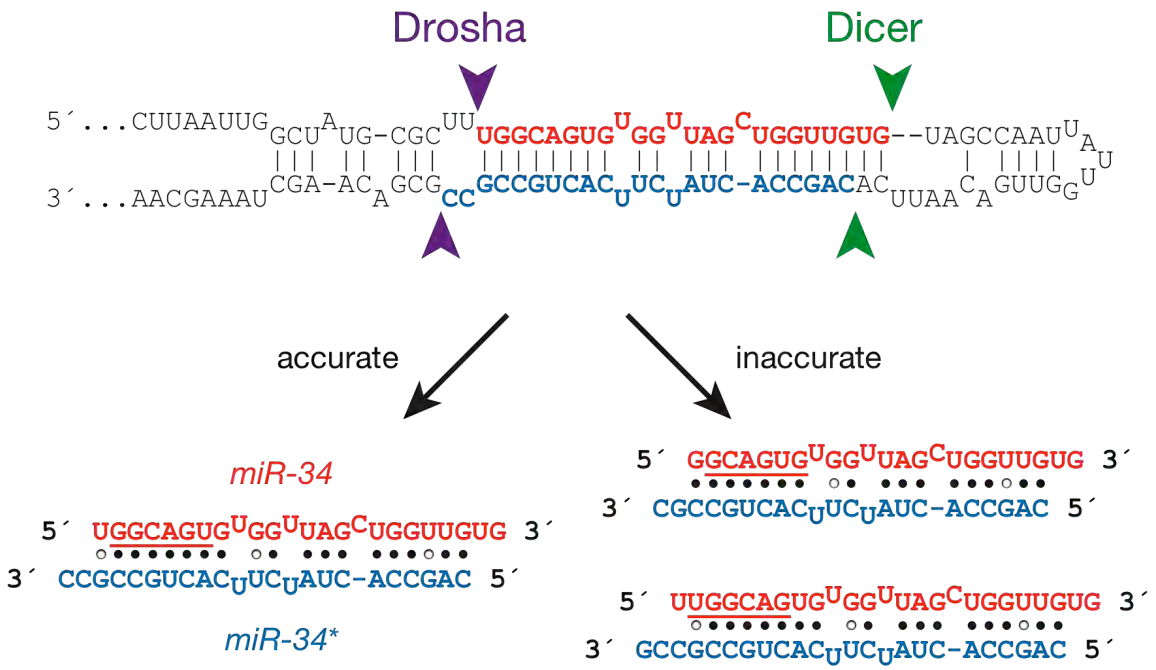
Introduction

MicroRNAs (miRNAs) are short regulatory RNAs that direct repression of their mRNA targets. The miRNA "seed"—nucleotides 2-7—establishes miRNA target specificity by mediating target binding[159,160,165,261,281]. Accurate processing of the miRNA 5´ end is thought to be under strong selective pressure[44,282] because a shift by just one nucleotide in the 5´ end of a miRNA would alter its seed sequence, redefining its repertoire of targets (Fig. 1). Animal miRNAs are produced by the sequential cleavage of partially double-stranded precursor RNAs by the RNase III endonucleases Drosha and Dicer, thereby generating a transitory double-stranded intermediate comprising the miRNA paired to its partially complementary miRNA* strand[283,284]. Here, we report that in flies, the 5´ ends of miRNAs and miRNA* strands are typically more precisely defined than the 3´ ends of

either the miRNA or its miRNA*. Surprisingly, the precision of the 5´ ends of both miRNA and miRNA* sequences increases after Argonaute2 (Ago2) loading. Our data imply that either many miRNA* sequences are under evolutionary pressure to maintain their seed sequences—that is, they have targets—or that secondary constraints such as the sequence requirements for loading small RNAs into functional Argonaute complexes, narrow the range of miRNA and miRNA* 5´ ends that accumulate in flies.

Figure IV-1. Inaccurate processing of the 5´ end of a miRNA alters its seed sequence. miRNA precursors are cleaved by two RNase III enzymes, Drosha and Dicer, liberating a short duplex: in this duplex, the mature miRNA (red) is paired to a partially complementary small RNA, the miRNA* (blue), derived from the opposite arm of the pre-miRNA stem. Inaccurate cleavage of the miRNA 5´ end changes its seed sequence (underlined).

Figure IV-1.



Results

Inaccurate cleavages and non-templated additions cause miRNA heterogeneity

We used high throughput pyrosequencing of 18–30 nt RNAs to identify miRNAs expressed in *Drosophila melanogaster* heads and in cultured *Drosophila* S2 cells. Among the 120,896 miRNA reads (66,377 from fly heads; 54,519 from S2 cells), we observed two sources of heterogeneity for the ends of fly miRNAs: the addition of nucleotides not present in the gene from which the miRNA is transcribed (non-templated nucleotides) and inaccurate or alternative cleavage by Drosha or Dicer. Approximately 5% of the reads for a typical miRNA contained non-templated nucleotides on at least one end (Fig. 2A and Fig. 3), most frequently the addition of single uridines or adenosine to the 3′ end, but longer extensions were also observed, both on the 5′ and 3′ ends (Table S1). Interestingly, longer extensions were also U- and A-rich at the 3′ end, whereas at the 5′ end, the 3′-most non-templated nucleotide was frequently a cytidine, and other added nucleotides were typically uridines. This observation could prove to be useful for the identification of the 5′-elongating enzymatic activity. The non-templated addition of nucleotides, especially uridines, to the 3′ ends of miRNAs has been reported previously in wild-type *Caenorhabditis elegans*[44] and *hen1* mutant *Arabidopsis thaliana*[87]. Overall, the addition of non-templated nucleotides to the 5′ end of miRNAs was rarer (~1%) (Fig. 2A and Table S1).

We also observed a second, more frequent type of heterogeneity: variability in the position of the miRNA 5′ and 3′ ends within the sequence of the miRNA precursors (Fig. 2B). Non-templated nucleotides fortuitously matching the templated sequence are

predicted to occur much less often than the heterogeneity we observe (Table S2). Similar terminal heterogeneity has been noted for the 3′ ends of *C. elegans*[44] and the 5′ and 3′ ends of mouse[285] miRNAs. The aberrant miRNA termini we observe likely reflect imprecision in precursor cleavage by Drosha and Dicer. They are unlikely to correspond to degradation products because we recorded nearly as many miRNA reads that were longer than the dominant species as were shorter (Fig. 4) and because 93% (S2 cells) and 99% (fly heads) of sequences of the fly-specific 30 nt 2S ribosomal RNA (rRNA)—whose termini are expected to be single-stranded—were full-length (Discussion). 3′ degradation was slightly more common than 5′ degradation. We detected 3′ degradation for 1,010 reads versus 5′ degradation for 201 reads among the 33,505 total 2S rRNA reads from S2 cells and fly heads combined; 5 reads corresponded to 2S rRNA trimmed from both ends.

Figure IV-2. Cleavage inaccuracies are more frequent than non-templated additions. (A) The percentage of reads with non-templated 5′ or 3′ extensions was evaluated for each miRNA whose sequence was read at least 100 times. (B) The most abundant 5′ and 3′ ends were identified for each miRNA and all other ends corresponding to the sequence of the primary miRNA transcript were flagged as "alternative". The percentage of reads with alternative ends was then determined for each miRNA read at least 100 times. Note the difference in the y-axis scales in (A) and (B). Box plots follow Tukey's standard conventions: a rectangle encloses all data from the first to the third quartiles, a bold horizontal line reports the median, whiskers connected to the rectangle indicate the largest and smallest non-outlier data, and outliers (values distant from the box by more than 1.5 times the interquartile range) are displayed as open circles.

Figure IV-2.

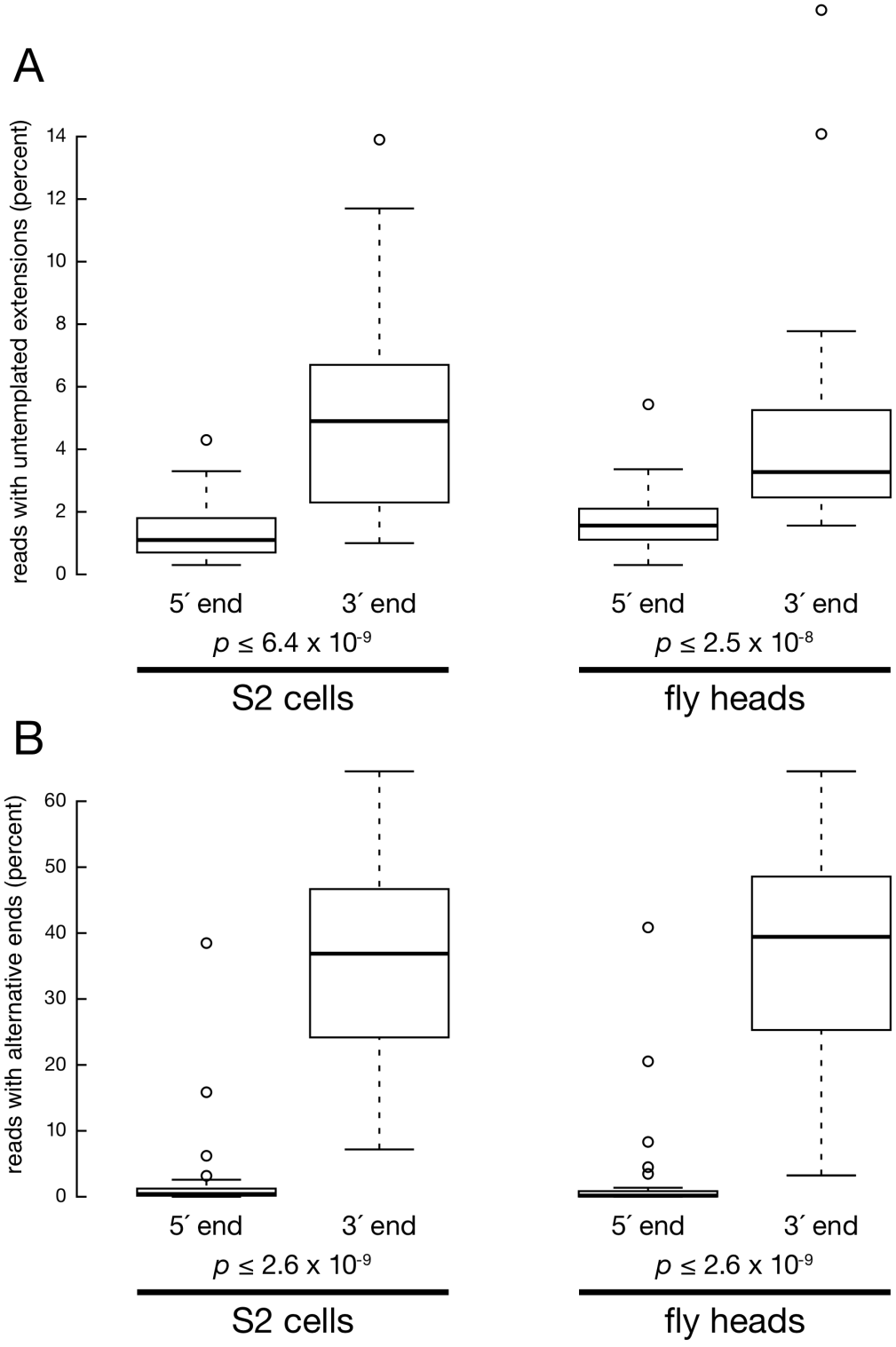

A
reads with untemplated extensions (percent)
0
2
4
6
8
10
12
14
5′ end
3′ end
p ≤ 6.4 x 10-9
S2 cells
5′ end
3′ end
p ≤ 2.5 x 10-8
fly heads
B
reads with alternative ends (percent)
0
10
20
30
40
50
60
5′ end
3′ end
p ≤ 2.6 x 10-9
S2 cells
5′ end
3′ end
p ≤ 2.6 x 10-9
fly heads

Figure IV-3. The abundance of miRNAs with non-templated nucleotides is proportional to the abundance of the miRNA itself.

Figure IV-3.

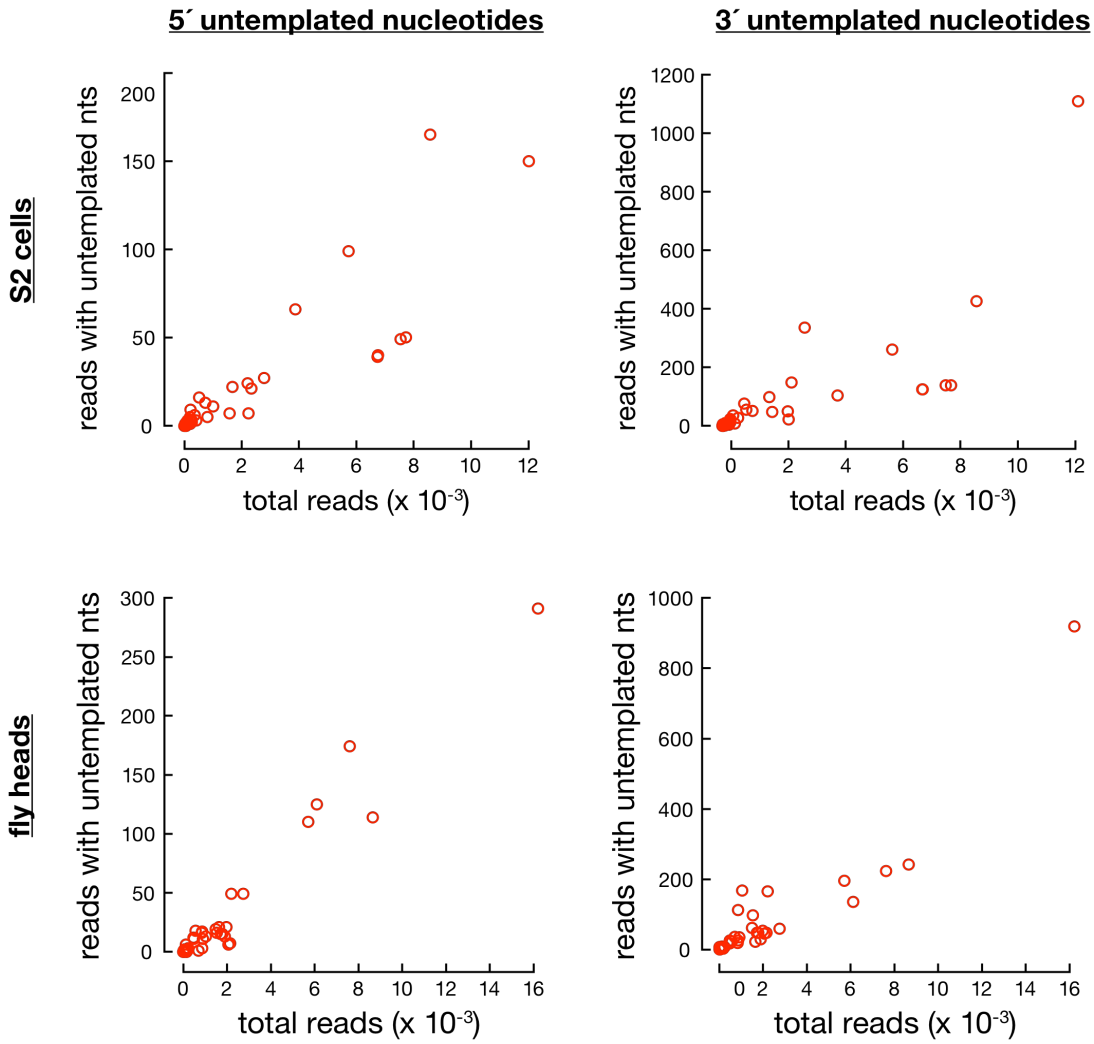
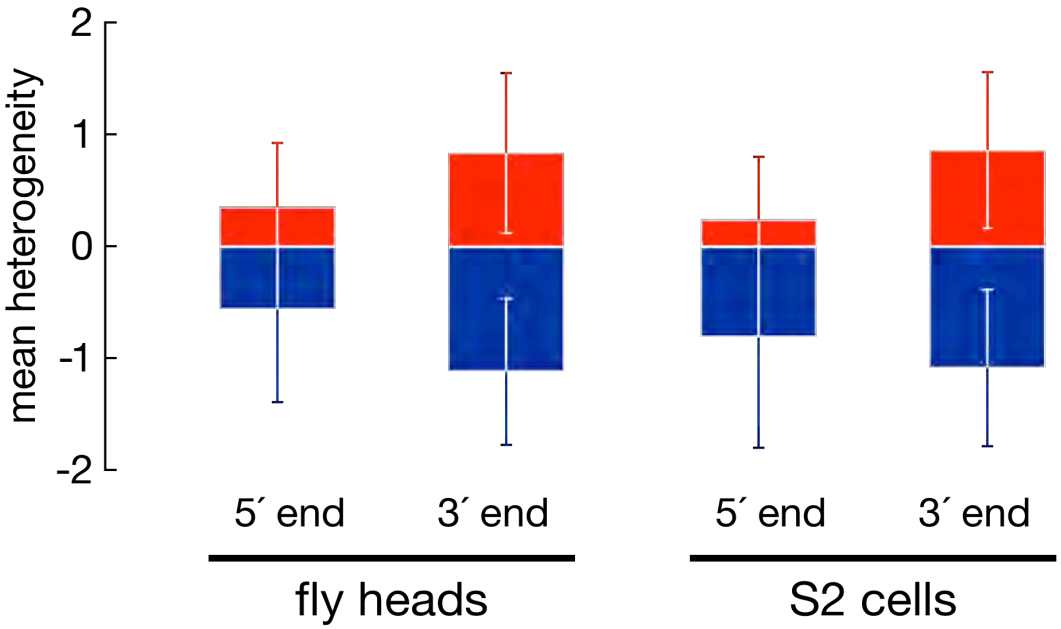

5´ untemplated nucleotides
3´ untemplated nucleotides
S2 cells
fly heads
reads with untemplated nts
total reads (x 10-3)

Figure IV-4. Mean heterogeneity for shorter and longer reads, compared to the most abundant variant for each miRNA. Positive values (red) indicate the reads were longer; negative values (blue) indicate that they were shorter than the most abundant variant for the corresponding miRNA. The bar graphs are essentially symmetrical; the various isoforms do not tend to be shorter than the most abundant one, suggesting that heterogeneity in miRNA ends reflects imprecise processing, rather than degradation. Error bars show standard deviation.

Figure IV-4.

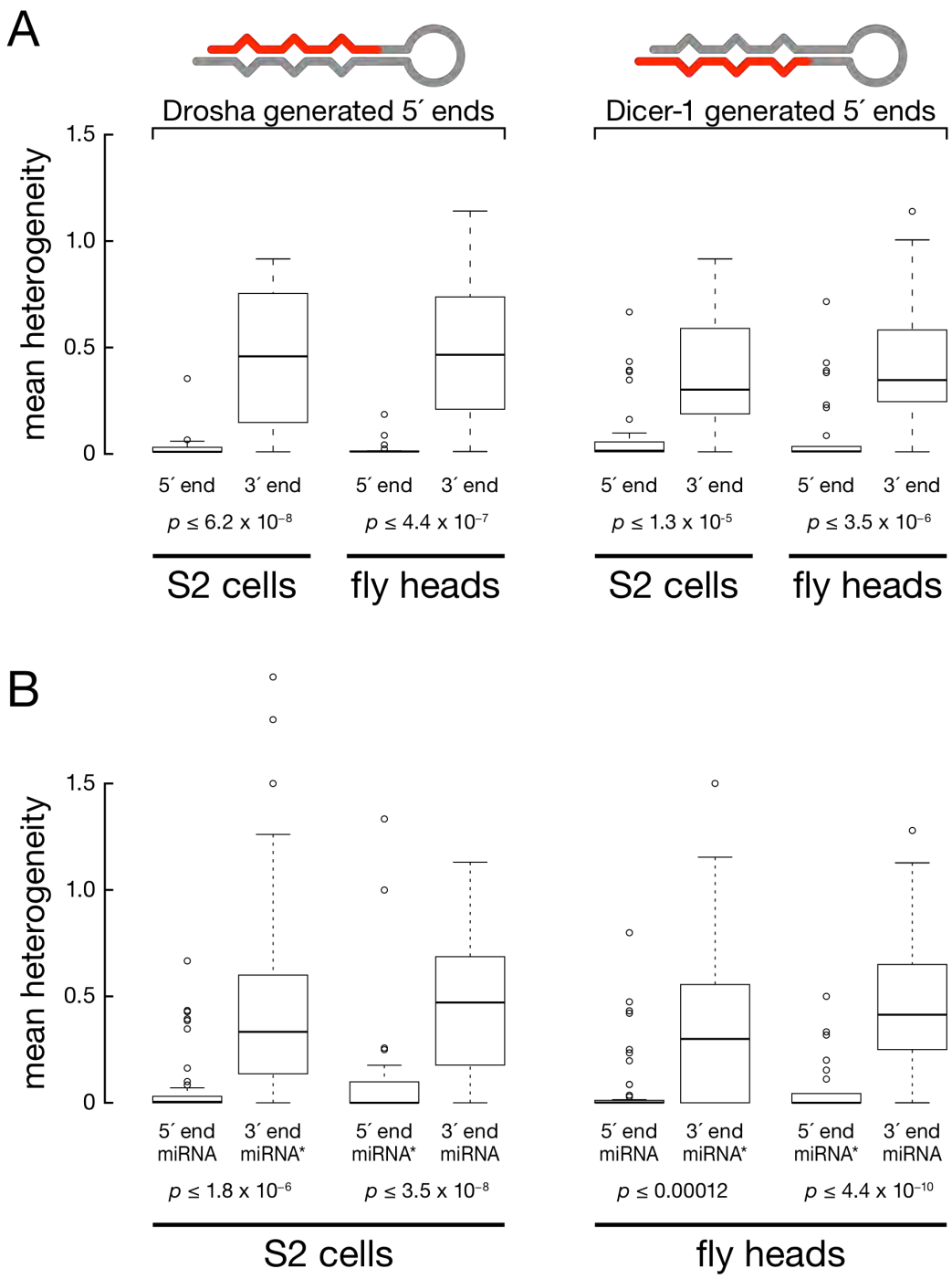

mean heterogeneity
2
1
0
-1
-2
5´ end
3´ end
5´ end
3´ end
fly heads
S2 cells

miRNA and miRNA* have more defined 5´ ends than 3´ ends

The 5´ ends of miRNAs were more precisely defined than their 3´ ends, irrespective of whether the miRNA originated from the 5´ or 3´ arm of the pre-miRNA (Fig. 5A). Thus, the difference in cleavage accuracy between the 5´ and 3´ ends cannot be attributed to an intrinsic difference in fidelity between Drosha and Dcr-1. We expected that the 3´ ends of miRNA* strands would be precisely defined, because they are created by the pair of cuts that generates the 5´ ends of miRNA, and that the 5´ ends of miRNA* strands would be imprecisely determined, because they are created by the pair of cleavages that generates the highly heterogeneous 3´ ends of miRNA. Instead, we found that the 5´ end of a strand (for example, the miRNA) was more accurate than the 3´ end of the adjacent strand (in this example, the miRNA*; Fig. 5B); these two extremities are produced by a pair of cuts catalyzed by the same enzyme.

Figure IV-5. **miRNA and miRNA* 5´ ends are more precisely defined than their 3´ ends.** (A) miRNAs originating from the 5´ (left panels) or 3´ (right panels) arms of their pre-miRNAs were analyzed separately. For each miRNA, the heterogeneity of its termini was calculated as the mean of the absolute values of the distance between the 5´ or 3´ extremity of an individual templated read and the most abundant 5´ or 3´ ends for that miRNA. Sequences read from RNA isolated from fly heads and cultured S2 cells were analyzed separately. (B) Box-plots show the distribution of mean heterogeneity for the 5´ and 3´ ends of miRNA and miRNA* sequences.

Figure IV-5.


A
Drosha generated 5′ ends
Dicer-1 generated 5′ ends
mean heterogeneity
1.5
1.0
0.5
0
5′ end
3′ end
p ≤ 6.2 x 10−8
p ≤ 4.4 x 10−7
p ≤ 1.3 x 10−5
p ≤ 3.5 x 10−6
S2 cells
fly heads
B
mean heterogeneity
5′ end miRNA
3′ end miRNA*
5′ end miRNA*
3′ end miRNA
p ≤ 1.8 x 10−6
p ≤ 3.5 x 10−8
p ≤ 0.00012
p ≤ 4.4 x 10−10
S2 cells
fly heads

Ago2 loading refines 5´ ends of miRNA and miRNA* strands

Current dogma holds that the local sequence or structure of miRNA precursors is under strong selective pressure to generate accurate 5´ ends, because a precise miRNA 5´ end directly establishes the seed sequence and hence the targets of the miRNA. Since we observe that, in flies, the 5´ ends of both the miRNA and the miRNA* are more precisely determined than the 3´ ends of either strand, this explanation implies that miRNA* sequences are under selective pressure to establish a unique seed sequence, implying that they, too, have regulatory targets.

It is also possible that both Drosha and Dcr-1—whose active sites are homologous—may also be intrinsically more precise in 5´ cleavage than in 3´ cutting. Another alternative is that 5´ and 3´ ends might be generated with similar, imperfect accuracy, but subsequent constraints in RISC loading select for those small RNAs that begin with a particular nucleotide or sequence. The subsequent destruction of miRNAs without these 5´ features would increase the apparent accuracy of miRNA 5´ ends while retaining miRNA 3´ heterogeneity. To test this idea, we separately sequenced small RNAs containing modified 3´ termini (Table S3). In flies, the 3´ termini of small RNAs that are loaded into Ago2[84], but not those bound to Argonaute1[182], are 2´-*O*-methylated by *Drosophila* Hen1 as the last step in Ago2-RISC maturation[83]. To sequence small RNAs bearing 2´-*O*-methylated 3´ ends, we treated the total small RNA with $NaIO_4$ followed by β-elimination; this method blocks ligation of adapters to small RNAs bearing 2´,3´ hydroxy termini, preventing them from being sequenced.

To determine whether the greater accuracy of miRNA and miRNA* 5´ versus 3´

ends reflects the constraints of RISC assembly or stability, rather than more accurate 5′ versus 3′ cleavage by Drosha and Dicer, we compared the terminal heterogeneity of miRNA and miRNA* reads from the 3′ modified population to the heterogeneity of the total miRNA and miRNA* population. As a control, we compared the 3′ heterogeneity between the two populations. For both analyses, we only considered miRNA or miRNA* strands displaying some heterogeneity in the total population. For both fly heads and S2 cells, we observed a dramatic increase in the precision of the 5′—but not the 3′—ends of miRNAs and miRNA* strands upon loading into Ago2 (Fig. 6). We also performed the analysis for those small RNAs that both had heterogeneous termini and were specifically enriched in the β-eliminated sequences relative to the non- β-eliminated set. For the 13 small RNAs (4 miRNAs and 9 miRNA*s) meeting these criteria, the 5′ ends in the sub-population of miRNA and miRNA* sequences loaded into Ago2—i.e., those that were 2′-*O*-methylated—were again more precisely defined than the 5′ ends of the same small RNA sequences in the total small RNA population (Fig. 7). We conclude that loading miRNAs into Ago2, and, perhaps into Argonaute proteins in general, imposes a purifying selection on their 5′ ends.

Figure IV-6. Ago2-loading, as evidenced by 3′ terminal 2′-*O*-methylation, refines miRNA and miRNA* 5′ ends. On average, the 5′ ends of the miRNAs and miRNA* strands in the 2′-*O*-methylated populations from both fly heads and S2 cells were more precisely defined than in the total population. We observed no statistically significant increase in the precision of the 3′ ends of the 3′ modified miRNAs and miRNA* strands.

Figure IV-6.

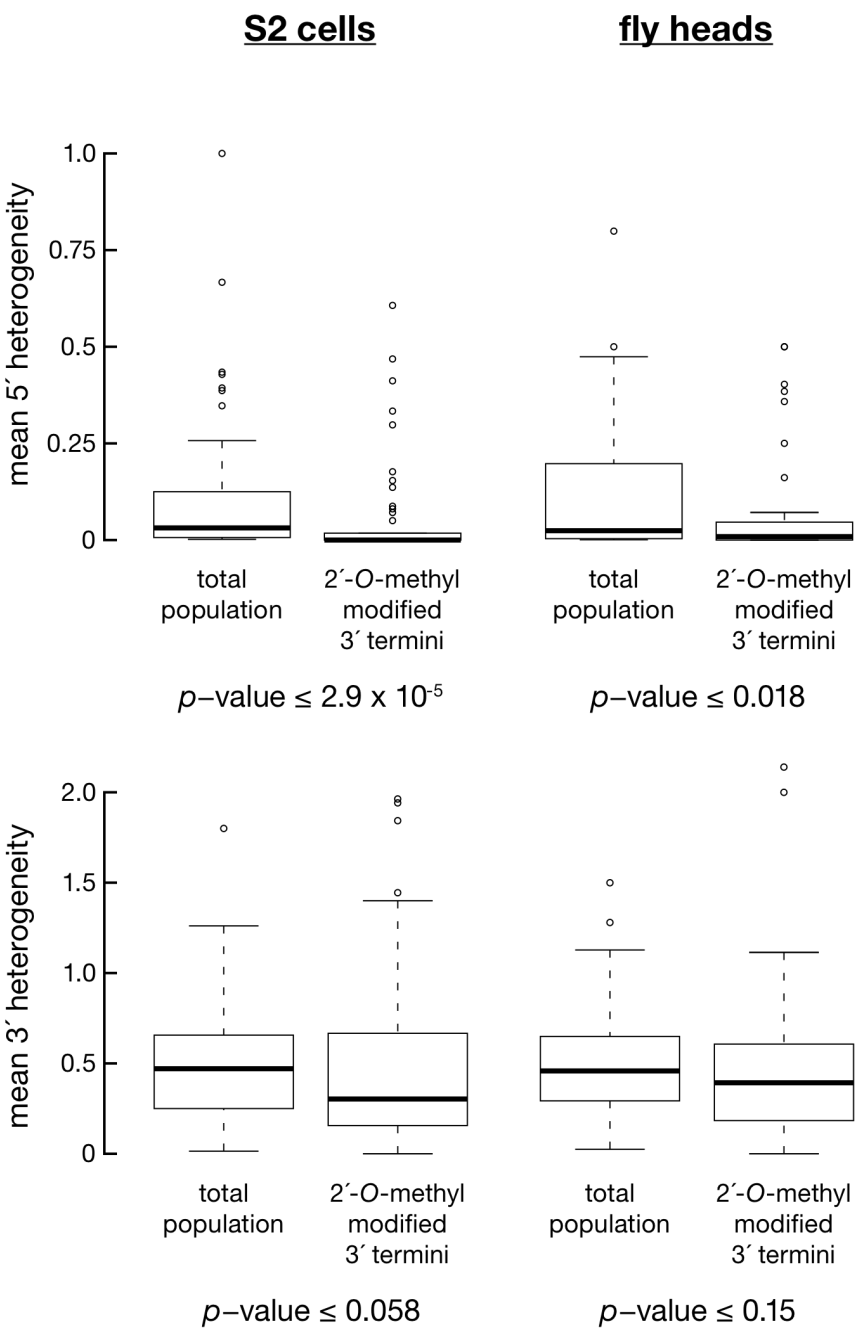
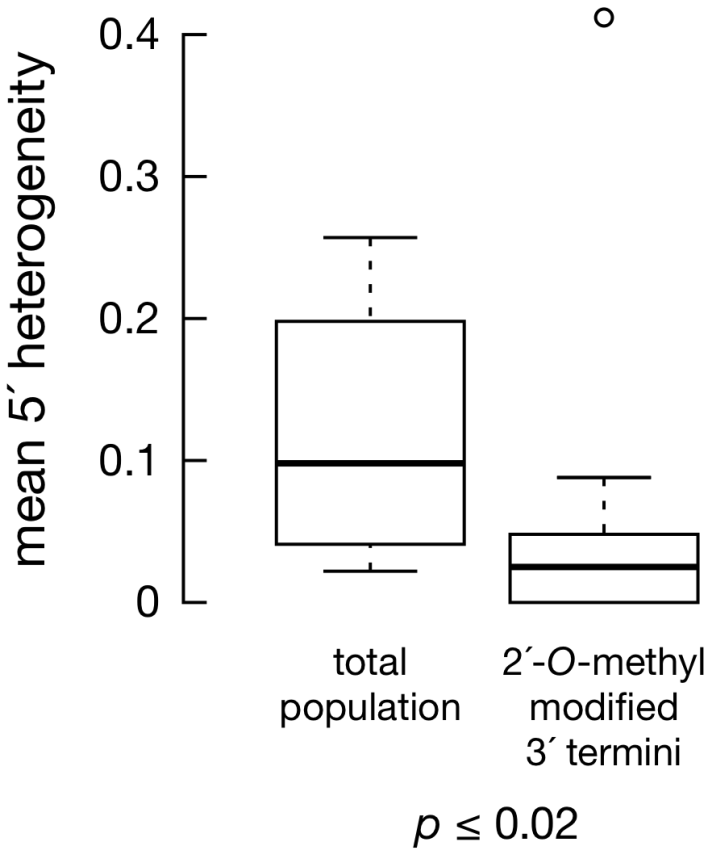

S2 cells
fly heads
mean 5′ heterogeneity
1.0
0.75
0.5
0.25
0
total population
2′-O-methyl modified 3′ termini
p–value ≤ 2.9 x 10-5
p–value ≤ 0.018
mean 3′ heterogeneity
2.0
1.5
1.0
0.5
0
p–value ≤ 0.058
p–value ≤ 0.15

Figure IV-7. Ago2 loading, as evidenced by 3′ terminal 2′ -O-methylation, refines miRNA and miRNA* 5′ ends. Four miRNAs and nine miRNA* species were identified that were both heterogeneous at their 5′ ends (i.e., a mean heterogeneity >0, Table S3) and were enriched among RNAs modified at their 3′ termini. On average, the 5′ ends of these small RNAs were more precisely defined in the 2′ -O-methylated population than in the total population.

Figure IV-7.


mean 5´ heterogeneity
0.4
0.3
0.2
0.1
0
total population
2´-O-methyl modified 3´ termini
p ≤ 0.02

Discussion

Terminal heterogeneity is not a ligation or degradation artifact

A potential explanation for the addition of non-templated nucleotides is that the ligase used to add adapters to each end of the small RNA joined fragments of abundant RNAs or pieces of RNA adapters to the endogenous small RNA. This seems unlikely: the most abundant non-templated trinucleotides are 5′-UAG-3′ added to the 5′ end and 5′-UUU-3′ added to the 3′ end. The only occurence of 5′-UAG-3′ in the adapters we used is in the core of the 5′ adapter (AUC GUA GGC ACC UGA AA); a degradation product of this adapter would likely bear 5′ hydroxyl and 3′ phosphate groups, making it a poor substrate for ligation. 5′-UUU-3′ is indeed the 5′-terminal trinucleotide of the 3′ adapter (UUU AAC CGC GAA UUC CAG) we used for the S2 cells RNAs, but this trinucleotide is absent from the 3′ adapter we used for the RNAs isolated from fly heads (CAC UCG GGC ACC AAG GA), where UUU corresponds to the most common non-templated 3′ trinucleotide. Finally, 5′-UAG-3′ is not the 3′ terminal trinucleotide of any abundant *Drosophila* non-coding RNA (ribosomal RNAs, tRNAs, snoRNAs and snRNAs), making it unlikely to be abundant in a 3′-OH form in our RNA samples. We note that the abundance of small RNA reads containing non-templated nucleotide extensions is proportional to the number of times it was read in the total population, a measure of its relative cellular abundance (Fig. 3).

Moreover, the heterogeneity of templated nucleotides is unlikely to reflect heterogeneous degradation of miRNA and miRNA* extremities by exonucleases, as exemplified by the integrity of detected 2S rRNA sequences. Of 19,811 2S rRNA-

matching reads from fly heads, 19,670 (99 %) corresponded to full-length, 30 nt 2S rRNA; of 13,694 2S rRNA-matching reads from the S2 cells, 12,706 (93 %) were full-length. Additionally, we did not notice any tendency for these heterogeneous reads to be shorter than the most abundant read: shorter and longer reads were detected with similar frequencies (Fig. 4).

Potential 5´ nucleotide purifying mechanisms

Various isoforms of miRNAs and miRNA* sequences, differing in their 5´ or 3´ ends, have been observed to arise from pre-miRNAs[188,244,271,285,286]. The variations result largely from imprecise processing by Drosha or Dicer. Consistent with the need to specify the miRNA seed precisely, the 5´ end of the miRNA strands are more homogenous than their 3´ ends: a change in the 5´ end will alter the identity of the seed region and hence redefine its repertoire of targets. One might reasonably presume the 3´ end of the miRNA* to be more precisely defined than its 5´ end, as it is made by the same pair of cuts that defines the 5´ end of the miRNA. However, the 5´ ends of the miRNA* strands are also more precisely determined than their 3´ ends, regardless of whether they are defined by Drosha or Dicer. Perhaps the miRNA* seed is under selective pressure because miRNA* strands have their own target RNAs. Alternatively, the Drosha and Dicer active sites that cleave the 5´ side of double-stranded RNA may simply be more precise than their 3´ counterparts. Finally, precision in 5´ ends may reflect sequence or structural requirements for loading RISC. Supporting this idea, we show that the 5´ ends

of *Drosophila* miRNA and miRNA* loaded in Ago2 are, on average, more precise than those in the total population.

The mechanism responsible for the homogenization of 5′ ends following Ago2 loading remains to be determined. We can imagine that the efficiency of Argonaute loading is affected by the nature of the 5′ end of a small RNA, much as the stability of its pairing to the other strand influences this process[73]. The 5′ sequence itself may also play a role in RISC assembly, with some miRNA variants loaded more efficiently than others, according to the identity of their 5′ nucleotide(s). Alternatively, some Argonaute complexes might be selectively stabilized after their assembly, for example, by the presence of a target RNA whose binding stabilizes those RISCs containing miRNA isoforms with a complementary seed sequence.

Materials and Methods

General methods

Fly heads were isolated by vigorously shaking liquid nitrogen-frozen flies expressing a long double-stranded hairpin RNA corresponding to *white*[66,243] in nested, pre-chilled sieves (U.S.A. standard sieve, Humboldt MFG Co., Chicago, IL, USA), allowing the heads to pass through the top sieve (No. 25) and collecting them on the bottom sieve (No. 40). S2 cell RNA was prepared from a clonal line containing the stably-integrated GFP transgene (pKF63) and transiently transfected with a double-stranded RNA against GFP[134].

RNA preparation

400 µg total RNA was extracted using the mirVana kit (Ambion), then 18- to 30 nt-long RNAs gel purified. 2S rRNA was depleted by hybridization to immobilized DNA oligonucleotide (5´-biotin-TCA ATG TCG ATA CAA CCC TCA ACC ATA TGT AGT CCA AGC A-3´). 1.6 nmol of the biotinylated oligonucleotide was bound to 32 mg M270 Streptabeads (Dynal, Norway) in 3.2 ml 0.5x SSC for 30 min on ice, then the beads were washed with ice-cold 0.5x SSC, resuspended in 8 ml 0.5x SSC, and incubated 5 min at 65°C. Gel-purified RNAs were diluted with 7 volumes 0.5x SSC to a final volume of 160 µl and denatured at 80°C for 5 min, then added to the bead suspension and incubated 1 h at 50°C. Beads were magnetically captured for 1 min at room temperature, then the 2S rRNA-depleted supernatant collected and precipitated with absolute ethanol. More than 99% of the 2S rRNA was routinely removed without measurably altering miRNA concentration; without the depletion step, nearly all the small RNA reads would correspond to 2S rRNA. Half the sample was then β-eliminated as described [180] and half was subject to the same treatment, except that sodium periodate was omitted.

Amplification and pyrosequencing

Adapters were ligated to the small RNA sample, and the resulting library amplified by PCR as described[4], except that a truncation mutant of RNA ligase 2 [Rnl2(1-249); 287] was used for the 3´ ligation step; T4 RNA ligase (Ambion) was used for 5´ ligation. The 5´ adapter was 5´-dAdTdC dGdTrA rGrGrC rArCrC rUrGrA rArA-3´ (Dharmacon, Lafayette, CO, USA); 3´ ‘preadenylated’ adapters were 5´-rAppdCdA dCdTdC dGdGdG dCdAdC

dCdAdA dGdGdA ddC-3′ for fly head and 5′-rAppdTdT dTdAdA dCdCdG dCdGdA dAdTdT dCdCdA dGddC-3′ for S2 cell RNA (IDT DNA, Coralville, IA, USA). After adapter addition, the RNA was amplified by PCR using DNA primers corresponding to the adapters. This PCR pool was gel purified (4% Metaphor Agarose, Cambrex, East Rutherford, NJ, USA) with Qiaex II (Qiagen, Valencia, CA, USA), then re-amplified by PCR (common 5′ primer, 5′-GCC TCC CTC GCG CCA TCA GAT CGT AGG CAC CTG AAA-3′; 3′ primer for fly heads, 5′-GCC TTG CCA GCC CGC TCA GTC CTT GGT GCC CGA GTG-3′; 3′-primer for S2 cells, 5′-GCC TTG CCA GCC CGC TCA GCT GGA ATT CGC GGT TAA A-3′). The PCR-amplified libraries were pyrosequenced by Roche Applied Science (Branford, CT, USA). Sequence and abundance data are available via the NCBI gene expression omnibus web site (http://www.ncbi.nlm.nih.gob/geo/) using accession number GSE9389.

Computational analyses

Eighteen- to 30-nt long reads were mapped to the *Drosophila melanogaster* genome (FlyBase assemblyR5.1; http://flybase.org/) and to the *D. melanogaster* "stem-loops" (which include the pre-miRNA sequences, usually extended by a few nucleotides) listed in miRBase (http://microrna.sanger.ac.uk/sequences/; version 10.0, August 2007). To identify non-templated microRNA additions, non-genome matching sequences were iteratively trimmed by 1 to 3 nucleotides on either the 5′ or the 3′ end and mapped to stem-loops.

Among stem-loop-matching reads, miRNA-matching and miRNA*-matching

reads were identified, using either the experimentally detected miRNA* sequence (when it was available in the miRBase records) or the product of conceptual dicing of the hairpin [73]. To include reads that showed extremities different from those annotated in miRBase, a distance of as many as 9 nucleotides 5′ or 3′ from the annotated miRNA or miRNA* sequence was tolerated. Statistical calculations were made using the **R** statistical package. *p*-values were calculated using the Wilcoxon test.

Supplemental Materials

Supplemental Tables

Table IV-S1. Addition of non-templated nucleotides to miRNAs in fly heads and in cultured S2 cells. Among pre-miRNA matching reads, some correspond to genomic sequence only if terminal nucleotides are removed. Once trimmed of these non-templated nucleotides, most of these sequences map perfectly to miRNAs; the remaining few percent typically map to miRNA* strands. For each set of pre-miRNA matching reads, the percentage matching the mature miRNA is reported in parentheses. The number of reads matching the pre-miRNA exactly (i.e., miRNA or miRNA*) is in red.

Table IV-S1.

RNA source	End	position of non-templated nucleotides	number of reads matching pre-miRNA (percent matching mature miRNA)	frequency of non-templated nucleotide at position
fly heads	5′	0	65,636 (95%)	NA
		1	500 (88%)	28% A; 62% C; 2% G; 8% U
		2	523 (97%)	14% A; 30% C; 2% G; 55% U
		3	212 (96%)	12% A; 25% C; 1% G; 62% U
	3′	0	65,636 (95%)	NA
		1	2,312 (97%)	32% A; 24% C; 4% G; 40% U
		2	400 (97%)	30% A; 25% C; 5% G; 40% U
		3	181 (90%)	29% A; 26% C; 6% G; 39% U
S2 cells	5′	0	53,683 (94%)	NA
		1	348 (91%)	33% A; 54% C; 4% G; 10% U
		2	284 (96%)	20% A; 31% C; 2% G; 47% U
		3	188 (94%)	16% A; 24% C; 2% G; 58% U
	3′	0	53,683 (94%)	NA
		1	2,629 (97%)	36% A; 12% C; 3% G; 49% U
		2	411 (98%)	36% A; 12% C; 5% G; 47% U
		3	219 (93%)	35% A; 12% C; 6% G; 47% U

Table IV-S2. Templated heterogeneity is unlikely to result from the addition of non-templated nucleotides fortuitously identical to the templated sequence. For each miRNA with at least 10 reads showing heterogeneity to the templated sequence, the observed nucleotide additions at the 5´ and the 3´ ends were compared to the expected distributions of non-templated extensions (assuming the observed nucleotide biases reported Table S1); significance was assessed by the chi-square test. For simplicity, only the first non-templated nucleotide on each end was considered, and we assumed that every non-templated addition followed the average observed nucleotide preferences in Table S1. These simplifications over-estimate the *p*-values and make the test more conservative.

Table IV-S2.

fly heads		S2 cells	
miRNA	***p*-value**	**miRNA**	***p*-value**
bantam	2.70E-134	*bantam*	0
let-7	4.80E-243	*let-7*	7.40E-60
mir-100	2.00E-161	mir-100	1.30E-161
mir-10	3.30E-07	mir-11	8.30E-232
mir-11	0	mir-124	0
mir-124	0	mir-125	3.80E-197
mir-125	0	mir-12	1.40E-18
mir-12	6.30E-54	mir-133	1.9e-318
mir-133	2.40E-276	mir-13a	3.50E-58
mir-13a	6.50E-76	mir-13b-1	0
mir-13b-1	0	mir-13b-2	0
mir-13b-2	0	mir-14	0
mir-14	0	mir-184	0
mir-184	0	mir-1	4.00E-115
mir-1	0	mir-210	0
mir-210	6.20E-213	mir-263a	7.00E-134
mir-263a	1.30E-74	mir-263b	1.10E-29
mir-274	0	mir-274	0
mir-276a	0	mir-275	2.30E-183
mir-276b	0	mir-276a	0
mir-277	0	mir-276b	1.90E-88
mir-278	0	mir-277	0
mir-279	1.70E-66	mir-278	0
mir-281-1	1.10E-12	mir-279	1.70E-83
mir-281-2	1.10E-12	mir-282	8.80E-65
mir-282	1.10E-27	mir-285	0
mir-285	0	mir-2a-1	0
mir-2a-1	0	mir-2a-2	0
mir-2a-2	0	mir-2b-1	0
mir-2b-1	0	mir-2b-2	0
mir-2b-2	0	mir-2c	1.10E-156
mir-2c	0	mir-304	2.90E-67
mir-305	8.30E-114	mir-305	1.30E-88
mir-306	2.80E-07	mir-306	3.00E-12
mir-307	1.80E-295	mir-307	1.30E-98
mir-317	0	mir-316	8.80E-45
mir-31a	2.50E-107	mir-317	0
mir-34	0	mir-31a	1.00E-29
mir-79	1.90E-06	mir-33	0

mir-7	0	mir-34	0
mir-8	0	mir-7	3.00E-56
mir-9a	8.80E-100	mir-8	0
mir-9b	5.90E-11	mir-9a	8.70E-100
		mir-9b	9.00E-35
		mir-9c	4.30E-25

Table IV-S3. 5´ end heterogeneity of miRNA and miRNA* sequences bearing a modified 3´ terminus. miRNAs and miRNA* sequences that were enriched among reads from 3´ terminally modified small RNAs and which were read at least 10 times in that sample were flagged as 2´-*O*-methylated. Mean heterogeneity was calculated as described in the legend to Figure 3. miRNA and miRNA* species used for the analysis in Fig. 7 are highlighted.

Table IV-S1.

fly heads (mean heterogeneity)		
small RNA	total population	3´ terminally modified RNAs
mir-100	0.198	0.046
mir-33	0.000	0.083
mir-100*	0.000	0.000
mir-278*	0.000	0.000
mir-282*	0.032	0.041
mir-284*	0.043	0.000
mir-2a-2*	0.041	0.025
mir-306*	0.000	0.059
mir-308*	0.000	0.000
mir-33*	0.200	0.048
mir-92a*	0.000	0.077

S2 cells (mean heterogeneity)		
small RNA	total population	3´ terminally modified RNAs
mir-283	0.000	0.174
mir-33	0.000	0.003
mir-6	0.000	0.000
mir-9c	0.027	0.000
bantam*	0.049	0.006
mir-100*	0.000	0.000
mir-13b-2*	0.167	0.000
mir-14*	0.250	0.412
mir-184*	0.022	0.000
mir-275*	0.000	0.000
mir-282*	0.098	0.071
mir-284*	0.000	0.000
mir-2a-2*	0.125	0.088
mir-308*	0.000	0.000
mir-33*	0.257	0.000

CHAPTER V

Conclusions and Discussion

R.J. Britten's 1969 proposal that nucleic acid guides could selectively control gene expression by base-pairing with target genes[288], seemed, until the discovery of small silencing RNAs, to be just another elegant strategy ignored by eukaryotes. Then and now, the idea that antisense nucleic acids are uniquely suited to specify regulatory targets is appealing because it is simple. Yet, like nearly all biological mechanisms, small RNA-directed pathways are at once elegantly simple—small RNA guides use sequence complementarity to identify their targets—and shockingly complex, with myriad proteins required to excise small RNA guides from much longer precursors and still more required to carry out small RNA-directed functions. Despite this complexity, the defining features of small silencing RNAs are their short length and their association with members of the Argonaute family of proteins.

Small RNAs predominantly exercise their regulation by base pairing with their target mRNAs, whose expression they repress transcriptionally or post-transcriptionally. It is not known if small RNAs can pair with DNA directly, an appealing model for those small RNAs that direct transcriptional silencing. Three types of small silencing RNAs are common between flies and mammals: microRNAs (miRNAs), small interfering RNAs (siRNAs) and Piwi-interacting RNAs (piRNAs). In the preceding chapters, I have revealed a new class of small RNAs in flies and also attempted to

understand the cross talk and the network of interactions that fine-tune the small RNA pathways, as it will provide vital insight into their regulatory potential.

The new small RNAs: endo-siRNAs

In flies, exogenous sources of dsRNA were considered the sole trigger of a siRNA response and viral defense as its solitary function. Endogenous siRNAs were known to exist only in organisms expressing an RNA-dependent RNA polymerase (RdRP), such as *Arabidopsis*, *Neurospora crassa*, *Schizosaccharomyces pombe* and *C. elegans*[254]. RdRPs transcribe single stranded RNA from an RNA template, producing dsRNA. The genomes of flies and mammals encode no readily recognizable RdRP proteins. Nevertheless, evidence suggested involvement of the siRNA machinery in silencing selfish genetic elements in the fly soma[235-240]. In chapter II, we sequenced small RNAs (18–30 nt) from fly heads and S2 cells, in an attempt to identify potential endo-siRNA candidates. We identified a 21 nt small RNA population which was 2´-*O*-methylated at their 3´ ends, similar to siRNAs derived from exogenous sources. The 21-mers did not exhibit a bias toward beginning with uracil, a characteristic of miRNAs and piRNAs, and were present in both sense and antisense orientations, in contrast to piRNAs. Moreover, the normal accumulation of the 21-mers was dependent on Dcr-2 and Ago2, establishing them as bona fide somatic endo-siRNAs.

Endo-siRNAs have also been cloned from fly gonads[27]. It will be interesting to determine, if these gonadal endo-siRNAs were derived from the germline, or were

present in the somatic cells of the gonads. Alternatively, are endo-siRNAs confined to the soma, similar to piRNAs restrictment to the germline?

Making endo-siRNAs without RdRP

Endo-siRNAs originate from transposons, heterochromatic sequences, intergenic regions and mRNAs; and disruption of the siRNA generating machinery results in enhanced expression of several transposons in the soma. We also observed that endo-siRNAs arise frequently from genomic regions likely to produce convergent transcripts[26]. This provided a strong evidence for the intermolecular production of dsRNA in vivo in metazoans, excluding the use of an RdRP.

Three other groups identified endo-siRNAs in *Drosophila*[27-29,112,113]. All these studies combined, recommend several genomic sources which can act as precursors for endo-siRNAs. These genomic loci—like bidirectionally transcribed loci , complementary overlapping transcripts, and structured loci—can inter- or intra-molecularly base-pair to form dsRNA precursors for endo-siRNA generation. However, the usage of these loci as precursors for endo-siRNAs and the precise dimensions of these precursors, still awaits validation.

Function and biogenesis of endo-siRNAs

Since the discovery of endo-siRNAs, there is much speculation regarding their functions and constant attempts made to ascertain these functions (see Discussion in Chapter II and section on ‘Target prediction for Ago2 bound small RNAs’ below). However, another

interesting question concerns their biogenesis. Are endo-siRNAs constitutively present in the cell, or is their production triggered under certain biotic or abiotic conditions. For example, production of natsiRNAs in plants is triggered in response to stress. NatsiRNAs are generated from a pair of convergently transcribed RNAs: typically, one transcript is expressed constitutively, whereas the complementary RNA is transcribed only when the plant is subject to environmental stress. Futhermore, is endo-siRNA production affected in the event of a viral attack on a fly? Viral infection will overwhelm the siRNA machinery with generation of viral siRNAs, in order to launch a robust RNAi defense. Are the functions mediated by endo-siRNAs dispensable in such a situation?

Possible cross-talk

Intriguingly, we also discovered 24-28nt small RNAs in mutant fly genotypes deficient in endo-siRNAs[26]. 24-28nt small RNAs share many features with germline piRNAs and originate from similar transposon clusters as endo-siRNAs, alluding to the possibility of a locus to act as precursor for both endo-siRNA and piRNA-like small RNA biogenesis[26]. Endo-siRNAs are derived from dsRNA precursors and piRNAs from single-stranded precursors, and it's fascinating to imagine how the same transcript will be directed into different small RNA pathways. May be different isoforms of Polymerse II or an accessory component of the transcription machinery, might channelize the transcripts into either the siRNA or piRNA pathway. Moreover, why are 24-28 nt small RNAs observed in *ago2* mutants? Perhaps, in the absence of endo-siRNAs, piRNAs are produced to resume somatic transposon surveillance. Such a model implies existence of interaction

between the piRNA and endo-siRNA–generating machineries, and is the focus of my ongoing study. Interstingly, the reverse has been shown; overexpression of Aubergine in somatic tissues interferes with proper functioning of RNAi[289]. Are these pathways mutually excusive and do they cross-regulate each other? Hopefully, further research in this field will bring answers to these questions.

The blurring of distinctions (the diminishing line)

A small-scale biochemical approach led to the discovery of small RNA sorting phenomenon in flies. The biogenesis of a small RNA by Dicer and its Argonaute loading are uncoupled events[202,203]. miRNA duplexes produced by Dcr-1 are loaded into Ago2 by the Dcr-2/R2D2 heterodimer (RLC). Our study described in Chapter III, provides the first global in vivo test for small RNA sorting in flies. We performed extensive analyses to validate and expand our knowledge of the factors involved in sorting of small RNAs into distinct Argonaute complexes ensuing their biogenesis. We observed that a miRNA duplex presents distinct structures to the sorting machinery, viewing from either ends. Along with the central region, the edges of the miRNA duplex can determine it's sorting. Ago1 loaded guides were found to be less stably paired at their 5′ ends and bore central mismatches or bulges, whereas Ago2 loaded guides had less stably paired 3′ ends. In addition to structure, we observed a 5′ terminal nucleotide predilection by the different Argonautes, a phenomenon only observed in plants. Ago1 hugely preferred small RNAs that begin with uracil, whereas Ago2 was biased for small RNAs with a 5′ terminal cytosine. This selective advantage bestowed by a 5′ terminal nucleotide is consistent with

previous studies that show the 5′ end of a small RNA anchored in the PIWI domain of *A.fulgidus* Piwi protein[274,275]. Henceforth, the 5′ nucleotide will have to be compatible with the conformation acquired by the Argonaute upon RNA binding. Moreover, the choice of a pyrimidine in each case might reflect the necessity for a less bulky base at the 5′ phosphate binding pocket of the Argonaute, along with the thermodynamically suitable nucleotide at the end of the small RNA duplex.

The fate of a miRNA/miRNA* duplex, therefore depends on multiple factors; structure of its duplex, thermodynamic stability of the ends of the duplex and the 5′ terminal nucleotide. It is intriguing to imagine how these factors cooperate or conflict with each other while making the fateful decision in the event of small RNA duplex sorting. Furthermore, it also implies, though complex, the sorting process must be extremely efficient as it resolves the competition between the two small RNA pathways and maintains them working at their optimum.

Revisiting the definition of miRNA and miRNA* strands

The present definition of miRNA and miRNA* is based on the relative abundance of the two strands, as measured by the number of times it has been sequenced. The strand that is more abundant is referred as the miRNA strand and the other strand as miRNA*[4]. It is also assumed that the miRNA strand is the one with functional relevance and the miRNA* strand will be destroyed following Argonaute loading, hence sequenced less. However, with the advent of high depth sequencing, miRNA* strands are more frequently cloned and sequenced. Many miRNA* species are in fact present much more

abundantly compared to several lowly expressed miRNA loci[271]. Also the ratio of a miRNA: miRNA* can vary dramatically across development, with stages having comparable detectable expression from both strands[271].

We observed that many miRNA* strands are more abundant than their annotated miRNA counterpart[272,290]. Strikingly, we also observed miRNA duplexes that have the annotated miRNA relatively more expressed in one tissue and the miRNA* as the abundant species in a different tissue.

The abundance of a miRNA and miRNA* is also a measure of its association with Ago1, because the total small RNA profile in flies recapitulates the Ago1-bound small RNA library. This might reflect either higher cellular concentration of Ago1 compared to Ago2, or Ago1 is more frequently occupied by a small RNA and only a small fraction of Ago2 associates with a small RNA. However, an assessment of the abundance of a miRNA and miRNA* loaded into Ago2, will vary from the traditional annotation of a miRNA and miRNA* strand. An example is *bantam*, whereas Ago1 preferentially associates with the miRNA strand, Ago2 preferentially binds *bantam**. Therefore, the definition of a miRNA and miRNA* strand may vary across development stages and tissues; and differ across the Argonautes. Moreover, with increasing evidence of evolutionary conservation and target regulation for miRNA* strands, it is more valid to annotate the pre-miRNA derived small RNAs as 3p- miRNA or 5p-miRNA, based on the which arm of the pre-miRNA it is derived form[271-273].

The non-functional star strand?

In chapter IV, we observed stringent 5´ processing for miRNA* strands, which is further refined after Ago2 loading, similar to miRNA strands[244]. Accurate processing of a small RNA 5´ end is vital, as it defines the spectrum of its target RNAs[291]. Also, conservation of miRNA* strands across the 12 Drosophilid species, correlated with their abundance in flies[271]. Mounting evidence for precise processing, evolutionary conservation, intracellular abundance, in addition to Argonaute loading of the miRNA* strands, challenges its definition as a carrier strand and alludes to its role as a regulatory molecule[188,244,271-273]. We showed that the miRNA* strands enriched in Ago2 are present at levels comparable to endo-siRNAs and *white* exo-siRNAs, which phenocopies a loss-of-function *white* mutation. In flies both Ago2 and Ago1 retain endonucleolytic activity, but Ago2 is a far better endonuclease than Ago1 and can catalyze multiple rounds of target cleavage, unlike Ago1[203]. Therefore, we predict even a small amount of miRNA* strands in Ago2 can efficiently regulate their targets. Interestingly, recent evidence indicates that Ago2 can also repress translation of targets, bearing central bulges when paired with the small RNA[278]. Therefore, it is probable for Ago2 loaded miRNA* strands to mediate target regulation, by either translational repression, or cleavage, or both.

Another perplexing observation is the evolutionary conservation profile of all miRNA genes. The miRNA genes exhibit the highest conservation score in the area corresponding to the miRNA strand followed by the miRNA* strand[188,271]. The hairpin loop of the pre-miRNA is not conserved. MiRNAs are known to bind and regulate their targets utilizing the seed sequences present in the first half of the miRNA strand. So why

is the latter half of the miRNA strand conserved? A possible explanation could be the pairing of the latter half of a miRNA strand with the 5′ half of the miRNA* strand which embodies the seed sequence of the miRNA* strand, especially if only the seed is essential for target binding and repression.

Target prediction for Ago2 bound small RNAs

Target prediction algorithms utilize evolutionary conservation of the miRNA target sites, and pairing of the miRNA seed sequence to its target, usually supplemented by beneficial 3´pairing[291]. Genome wide analyses with these target recognition tools led to the identification of many miRNA targets. siRNAs, on the other hand, were considered to guide Ago2 to cleave targets with extensively complementary sequence, in addition to a base-paired seed region. But with emerging knowledge, about the ability of Ago2 to translationally repress targets, complicates target prediction. In order to conduct a genome wide search for potential targets for endo-siRNAs and miRNA/miRNA* strands loaded in Ago2, we will have to specify the constraints of base-pairing required between the guide small RNA and the target mRNA, to elicit target cleavage or translational repression by Ago2. The requisite extent of base-pairing between the small RNA and its target, and the varied requirement for base-pairing at each position across the small RNA-target RNA duplex, for Ago2 to mediate either of the two methods of target regulation, are unspecified. It will be a challenge in the future to lay out the prerequisites to define targets regulated by Ago2, and how they are regulated.

Conclusions

This study not only enforces the functionality of both the strands of a miRNA duplex but also highlights the complex interplay between the small RNA pathways. The miRNA and siRNA pathways are no more distinct end points but form a continuum. This research brings to consensus long-standing conflicts between small RNA biogenesis and evolutionary conservation. It assigns a role beyond viral defense to the siRNA machinery and established miRNA* as a functional entity, elucidating maximal utilization of a Dicer processing event. Interestingly, evolution seems to have selected for miRNA duplexes that present two distinct structures to the sorting machineries. From one end of the duplex, the miRNA strand is favored as guide in Ago1 with an unpaired central region, whereas Ago2 loading is preferred from the perspective of the miRNA* strand.

The highlight of this thesis, however, is unveiling the underlying complexity that interconnects small RNA pathways. Malfunction of small RNAs bear consequences like cancer, infertility, and neurodegeneration. Therefore, cross talk between small RNA pathways creates a dynamic flux leading to a vigilant small RNA-mediated supervision of a multitude of biological processes.

Future Prospects

Despite our growing understanding of the mechanism and function of small RNAs, their evolutionary origins remain obscure. siRNAs are present in all three eukaryotic kingdoms—plants, animals, and fungi—and provide anti-viral defense in at least plants and animals. Thus, the siRNA machinery was present in the last common ancestor of

plants, animals and fungi. In contrast, miRNAs have only been found in land plants, the unicellular green alga, *Chlamydomonas reinhardtii*, and metazoan animals, but not in unicellular choanoflagellates or fungi[1,2,292]. Deep sequencing experiments have found no miRNAs shared by plants and animals, suggesting that miRNA genes, unlike the miRNA protein machinery, arose independently at least twice in evolution. Finally, piRNAs appear to be the youngest major small RNA family, having been found only in metazoan animals[292]. While Dicer proteins have been identified only in eukaryotes, Argonaute proteins can also be found in eubacteria and archea, raising the prospect that small nucleic-acids may have served as guides for proteins at the very dawn of cellular life, and though the machinery might be ancient, the small RNA guides diversified over time to acquire specialized roles.

The history of small silencing RNAs makes predicting the future particularly daunting, as new discoveries have come at a breakneck pace, with each new small RNA mechanism or function forcing a re-evaluation of cherished models and “facts.” Several longstanding but unanswered questions, however, are worth highlighting. First, does RNAi—in the sense of an siRNA-guided defense against external nucleic acid threats such as viruses—exist in mammals? Second, how do miRNAs repress gene expression? Do several parallel mechanisms co-exist in vivo, or will the current, apparently contradictory, models for miRNA-directed translational repression and mRNA decay ultimately be unified in a larger mechanistic scheme? Third, can miRNA regulated genes ever be identified by computation alone, or will computational predictions ultimately give way to high throughput experimental methods for associating individual miRNA species

with their regulatory targets? Will network analysis uncover themes in miRNA-target relationships that reveal why miRNA-regulation is so widespread in animals? Fourth, how are piRNAs made? The feed-forward amplification “ping-pong” model is appealing, but likely underestimates the complexity of piRNA biogenesis mechanisms? We do not yet know how piRNA 3´ ends are generated. Nor do we have a coherent model for how long, antisense transcripts from piRNA clusters are fragmented into piRNAs. Finally, will the increasing number of examples of small RNAs carrying epigenetic information across generations[57,293] ultimately force us to reexamine our Mendelian view of inheritance?

APPENDIX I

Targeted deletion of *loquacious*

The work presented was a collaborative effort. The author generated *loqs* loss of function flies by Flippase mediated targeted recombination of FRT sites leading to deletion of *loqs*. These flies were used in a study led by Tingting Du, to examine the role of Loqs in the siRNA pathway. Tingting Du performed the experiment, demonstrating requirement for Loqs for maximal silencing triggered by a long inverted repeat. The author, Tingting Du and Phillip Zamore, wrote the following text.

Introduction

In most eukaryotes, long double-stranded RNA (dsRNA) triggers the destruction of messenger RNAs with complementary sequences, a phenomenon termed RNA interference (RNAi)[55,294-296]. In *Drosophila*, 'foreign' long dsRNAs, such as those introduced experimentally or produced by viral infection, enter the RNAi pathway when they are processed into ~22 nucleotide, double-stranded small interfering RNAs (siRNAs) by the RNase III endonuclease Dicer-2 (Dcr-2)[21,22,24,59]. (Flies encode two dicer proteins[66,297]). These siRNAs are subsequently loaded into an effector complex—RISC (RNA-induced silencing complex)—containing Argonaute2 (Ago2) by the RISC-loading complex (RLC)[298]. Dcr-2 and its dsRNA-binding protein partner, R2D2, are core components of the RLC[76,270]. They form a stable heterodimer that identifies the siRNA

guide and passenger strands: R2D2 binds to the more stably paired end of the siRNA duplex, thereby positioning Dcr-2 at the less stable end, designating this RNA strand as the future guide[77]. After binding the siRNA, the Dcr-2/R2D2 heterodimer, perhaps together with other RLC components, recruits Ago2 to the double-stranded siRNA[299-301]. The geometry of the siRNA within the Dcr-2/R2D2 heterodimer is preserved when it is passed to Ago2: the 5´ end of the guide siRNA binds the Ago2 5´ phosphate-binding pocket, and the passenger strand assumes the position of a target mRNA.

Ago2 is an RNA-guided, Mg^{2+}-dependent endonuclease[150,268,302-305]. This nuclease activity acts not only in siRNA-guided mRNA cleavage, but also in the maturation of Ago2 to its active form, RISC. Because in immature RISC (pre-RISC) the passenger strand occupies the position of a target RNA, a critical step in RISC assembly is cleavage of the passenger strand by Ago2, a step that facilitates separation of the two siRNA strands[78-82]. Dissociation of the passenger strand leaves Ago2 loaded a single-stranded siRNA guide. Such mature RISC can then find its mRNA targets by nucleobase complementarity to the siRNA guide and destroy them by Ago2-catalyzed endonucleolytic cleavage.

Plants and animals also produce a second class of small regulatory RNAs, microRNAs (miRNAs)[3-5,116,144,283,306-308]. miRNAs are typically transcribed by RNA polymerase II as if they were mRNAs, but are then processed sequentially to generate a ~22 nt small RNA from the initial >1,000 nt transcript, the primary miRNA (pri-miRNA)[120]. In animals, the RNase III enzyme Drosha acts with a dsRNA-binding domain (dsRBD) protein partner, named Pasha in flies, to excise from the pri-miRNA a ~70 nt stem-loop RNA, the pre-miRNA[124-127,309]. Cleavage of the pri-miRNA by Drosha defines either the 5´ end or 3´

end of the mature miRNA, which can reside on either arm of the stem of the pre-miRNA. (A few miRNAs are transcribed directly into pre-miRNAs by RNA polymerase III, at least in human cells[310]).

Pre-miRNAs are converted to miRNAs by Dicer[62-64]. In flies it is Dicer-1 (Dcr-1), together with its dsRBD protein partner, Loquacious, (Loqs), that cleaves pre-miRNA[66,134,135,137]. Dcr-1 cleavage of a pre-miRNA liberates an siRNA-like duplex in which the miRNA is partially paired to a ~22 nt small RNA derived from the other arm of the pre-miRNA stem. This small RNA is the miRNA*[308]. The miRNA strand preferentially assembles into mature RISC, whereas the miRNA* strand is degraded.

It has been proposed that the RNAi and miRNA pathways are separate and parallel, with each using a unique set of proteins to produce small RNAs, to assemble functional RNA-guided enzyme complexes, and to regulate target mRNAs[241]. Such a simple picture likely underestimates the in vivo complexity of these two RNA silencing pathways. First, both *dcr-1* and *loqs* mutants, which are defective in miRNA production, are also impaired in siRNA-directed RNAi[66,134]. Second, Ago2, the Argonaute protein that mediates RNAi in flies, binds at least one endogenous miRNA[203]. Finally, *ago1* and *ago2* interact genetically in embryonic patterning and morphogenesis, suggesting that they function in a common pathway[311].

Results

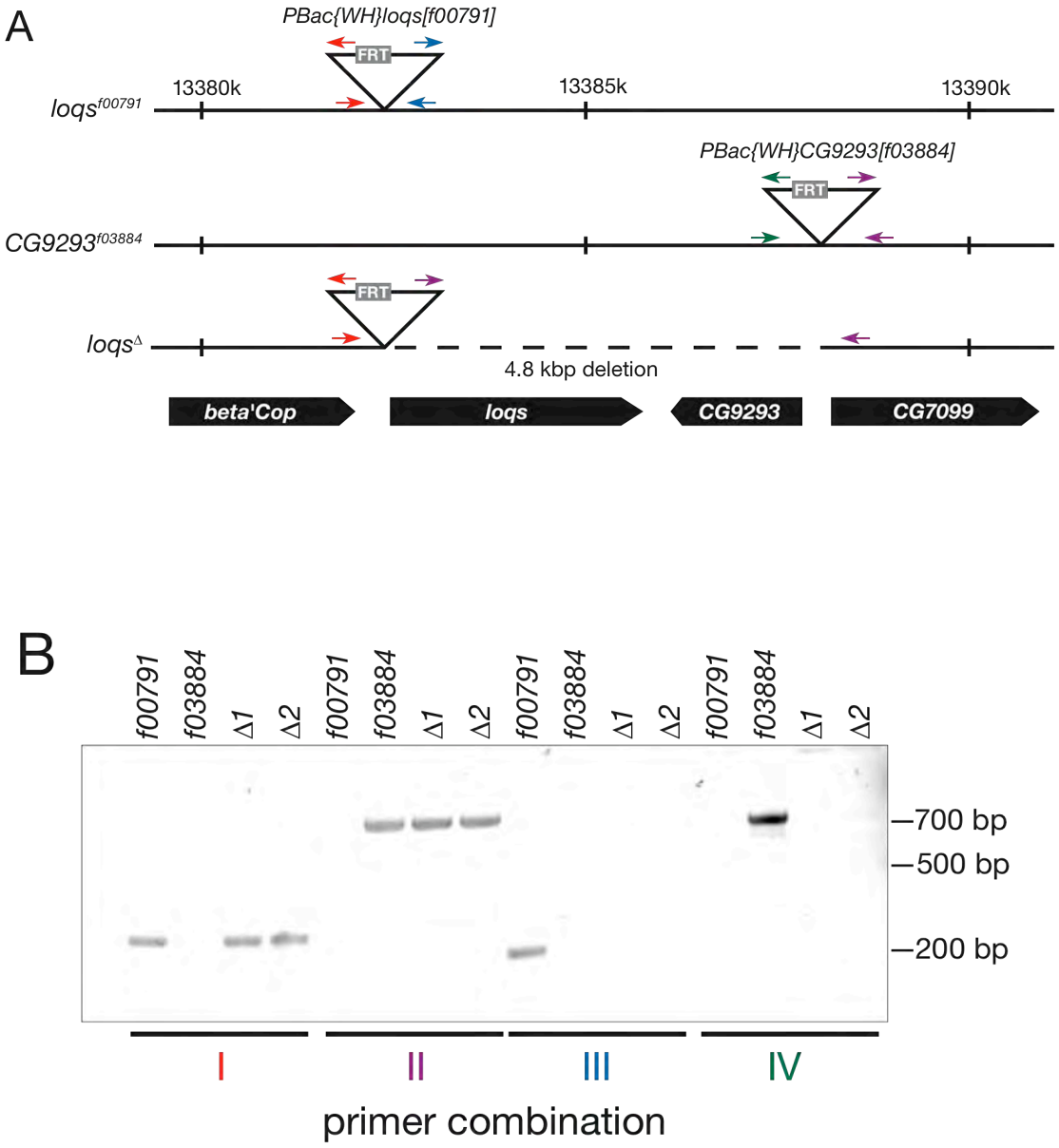
Generation of a loqs deficient allele by Flp-FRT mediated targeted deletion

Tingting Du, in our lab, was interested in examining the molecular function of Dcr-1/Loqs in RNAi pathway. The only *loqs* allele available then was $loqs^{f00791}$, generated by a piggyBac insertion within the first exon and 221 nucleotides upstream of the translational start codon of the *loqs* gene[135,312]. This allele exhibited the strongest phenotype in ovaries; a 40 fold reduction in *loqs* mRNA levels, compared to 5 fold reduction in female somatic tissues[134]. Therefore, consistent with the mutation, the mutant flies were viable but female sterile[134]. As *loqs* mutants were impaired in siRNA mediated silencing, we wanted to determine if the modest effect on silencing was only due to the hypomorphic nature of $loqs^{f00791}$ allele.

Therefore to facilitate analyses of molecular function of Loqs, I created a new allele $loqs^{D1}$, by FLP recombinase-induced mitotic recombination of two, tandem, FRT-bearing piggyBac transposons flanking *loqs* (Fig. 1). This new allele, $loqs^{D1}$, completely deletes *loqs*, as well as an adjacent gene; $loqs^{D1}$ is homozygous lethal.

Figure AI-1. Construction of a *loqs* deletion allele. (A) Strategy for making and identifying a 4.8 kbp deletion that removes the *loqs* gene. The deletion was constructed by FLP recombinase-mediated recombination between the FRT site in *PBac{WH}loqs[f00791]* and the FRT site in *PBac{WH}CG9293[f03884]*. (B) PCR analysis using the four color-coded primer pairs, indicated as arrows in (A), demonstrated that two independent deletion alleles, *loqsΔ1* and *loqsΔ2*, were recovered.

Figure AI-1.


A
PBac{WH}loqs[f00791]
FRT
loqs^f00791
13380k
13385k
13390k
PBac{WH}CG9293[f03884]
CG9293^f03884
loqs^Δ
4.8 kbp deletion
beta'Cop
loqs
CG9293
CG7099
B
f00791
f03884
Δ1
Δ2
—700 bp
—500 bp
—200 bp
I
II
III
IV
primer combination

Loqs is required in vivo for maximal silencing triggered by a long inverted repeat

In flies and other eukaryotes, long inverted repeat (IR) RNAs trigger silencing of complementary mRNAs because they are almost entirely double-stranded. The *Drosophila white* gene, which encodes a protein required for the production and distribution of red eye pigment, can be silenced by a transgene (*GMR-whiteIR*, henceforth, *white*-IR) that expresses in developing eye tissue a 621 nt dsRNA hairpin corresponding to the third exon of *white*[313]. IR-induced silencing of *white* has been proposed to function through the RNAi pathway, because two key components of the pathway, Dcr-2 and R2D2, are required for the process in vivo[66,134].

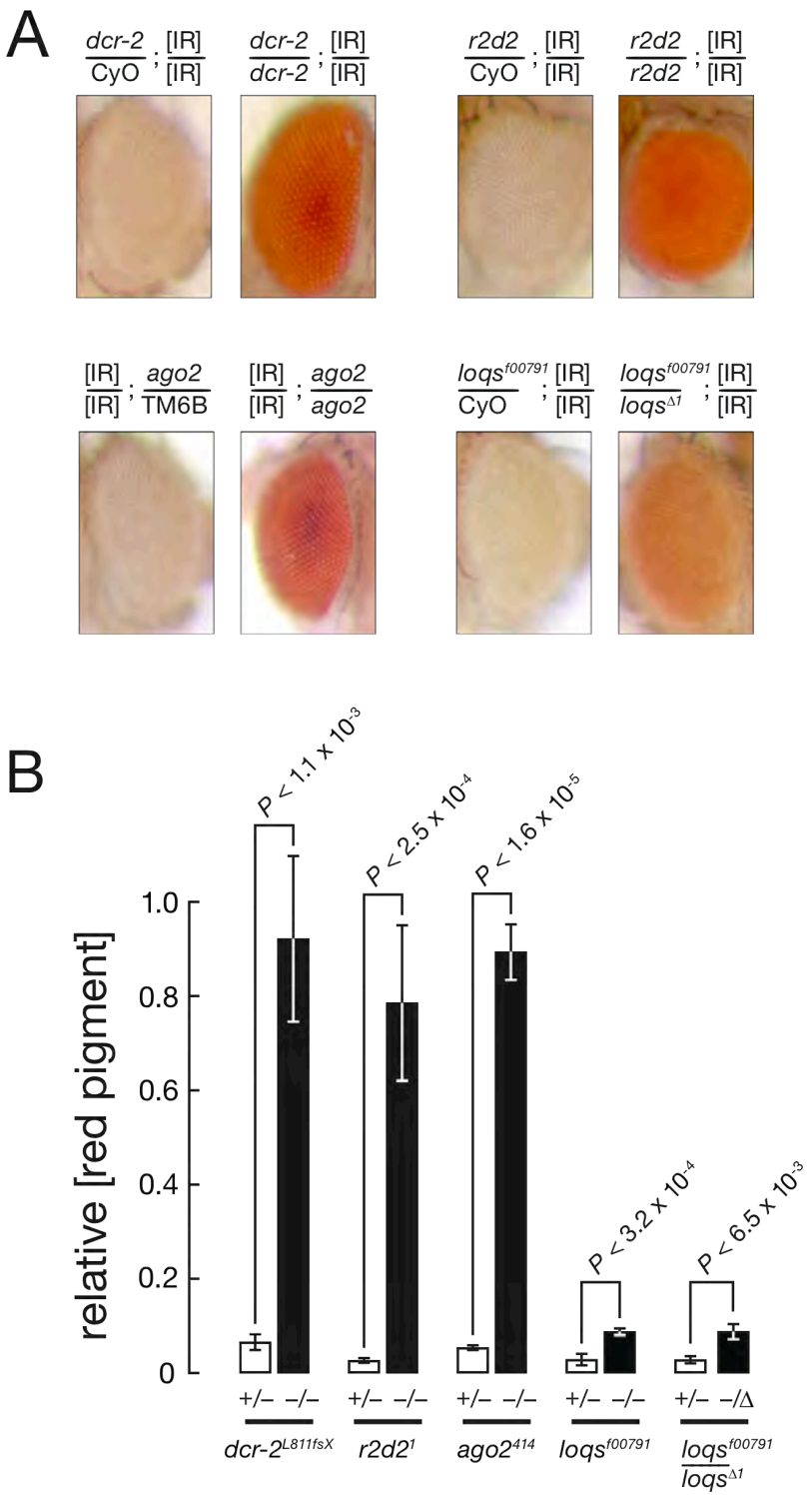
Because *Drosophila* Ago2 is the core of the RISC that mediates RNAi[314], we asked whether Ago2 is required for IR triggered silencing in vivo. In otherwise wild-type flies, two copies of the *white*-IR completely silence *white*, yielding a colorless eye indistinguishable from that of a complete loss-of-function *white* mutant. In contrast, two copies of the *white*-IR failed to silence *white* expression in an $ago2^{414}$ mutant (Fig. 1A); these flies had red—i.e., wild-type—eyes. Thus, all three key components of the somatic RNAi pathway, *dcr-2*, *r2d2*, and *ago2*, are required in vivo in flies for silencing triggered by the *white*-IR.

loqs mutants are partially defective in IR- induced *white* silencing[134]. This defect is reflected by the orange, rather than red, color of the eyes of *loqs* mutant homozygotes expressing two copies of the *white*-IR (Fig. 2A) and can be quantified by measuring the amount of red pigment extracted from fly heads (Fig. 2B). For flies carrying two copies of the *white*-IR, the $loqs^{f00791}$ mutation restored eye pigment to 8 ± 0.6% (average ± standard deviation, n=5) of the concentration in wild-type Oregon R heads. In contrast, strong loss of

function mutations in *r2d2* (*r2d2*^{1}: 73 ± 15%; n=5), *dcr-2* (*dcr-2*^{L811fsX}: 92 ± 18%; n=4) and *ago-2* (*ago-2*^{414} : 89 ± 6%, n=4), essentially eliminated silencing (Fig. 2B). Because *loqs*^{f00791} is a partial loss-of-function allele in the soma, we analyzed the silencing phenotype of *trans*-heterozygous flies bearing one copy of *loqs*^{f00791} and one copy of a *loqs*^{D1}. The loss of *white* silencing was essentially the same in the *loqs*^{f00791} homozygotes and in the *loqs*^{f00791}/*loqs*^{D1} *trans*-heterozygotes (Fig. 2A and B), demonstrating that *loqs*, rather than a second gene fortuitously mutated in the original *loqs*^{f00791} stock, plays a role in robust RNAi in vivo.

Figure AI-2. Loqs facilitates RNAi in vivo. (A) The eye color of heterozygotes was compared to that of homozygotes for the mutant alleles *dcr-2*^{L811fsX}, *r2d2*^{1}, and *ago2*^{414} for age-matched males bearing two copies of the *white*-inverted repeat transgene ([IR]). For *loqs*, flies heterozygous for *loqs*^{f00791} were compared to *loqs*^{f00791}/*loqs*^{D1} trans-heterozygotes. (B) The eye pigment of heterozygotes (+/–) and homozygotes (–/–) for the indicated genotypes, each bearing two copies of *white*-IR transgene, was extracted and its absorbance measured at 480 nm. The graph shows the mean ± standard deviation, relative to wild-type flies lacking the *white*-IR transgene, for four independent measurements. Statistical significance was estimated using a two-sample Student's *t*-test assuming equal variance.

Figure AI-2.


A
dcr-2 / CyO ; [IR] / [IR]
dcr-2 / dcr-2 ; [IR] / [IR]
r2d2 / CyO ; [IR] / [IR]
r2d2 / r2d2 ; [IR] / [IR]
[IR] / [IR] ; ago2 / TM6B
[IR] / [IR] ; ago2 / ago2
loqs^f00791 / CyO ; [IR] / [IR]
loqs^f00791 / loqs^Δ1 ; [IR] / [IR]
B
relative [red pigment]
1.0
0.8
0.6
0.4
0.2
0
P < 1.1 x 10^-3
P < 2.5 x 10^-4
P < 1.6 x 10^-5
P < 3.2 x 10^-4
P < 6.5 x 10^-3
+/– –/–
+/– –/–
+/– –/–
+/– –/–
+/– –/Δ
dcr-2^L811fsX
r2d2^1
ago2^414
loqs^f00791
loqs^f00791 / loqs^Δ1

Discussion

In *Drosophila,* the two best understood RNA silencing pathways are the siRNA-mediated RNAi pathway and the microRNA pathway. These two pathways were originally proposed to be parallel and separate. Increasingly, however, the two pathways appear to be interconnected, with some proteins shared between them. For example Dcr-1 and Loqs, which function together to process pre-miRNA into mature miRNA, are required in vivo for robust RNAi. Loqs is also required for accumulation of endo-siRNAs derived form structured loci[27,28]. Moreover, a recent report suggests that all classes of siRNAs require a sequential action of both Loqs and R2D2; Loqs partners with Dcr-2 in processing these endo-siRNAs, and R2D2 in collaboration with Dcr-2 loads the siRNAs into Ago2- effector complexes[315].

In light of the various speculated roles, assigned to Loqs, the *loqs* deficiency allele provides a useful tool to better dissect the role of Loqs in the various pathways. Additionally, the *loqs* deletion allele, aided in the study lead by Tingting Du (Du *et al,* unpublished manuscript). We showed that the Dcr-1/Loqs complex plays a direct role in the production of Ago2-RISC, the siRNA-programmed RNAi enzyme complex that directs cleavage of target RNAs in response to a dsRNA trigger. Our results suggest that the earliest detectable step in Ago2-RISC assembly is binding of the Dcr-1/Loqs complex to siRNA. Dcr-1/Loqs then transfers the still double-stranded siRNA to the RLC, which contains the Dcr-2/R2D2 heterodimer. Consistent with a role of Dcr-1/Loqs in siRNA loading, accumulation of dsRNA-derived siRNAs in vivo is impaired in *loqs* mutants. Together, our data suggest

considerable functional and genetic overlap between the miRNA and siRNA pathways, with the two sharing components previously thought to be restricted to just a single pathway.

Materials and methods

Fly stocks

The following fly stocks were used: Oregon R, P[*w*-IR]/ P[*w*-IR] (on chromosome 2), P[*w*-IR]/ P[*w*-IR] (on chromosome 3), FRT42D *dcr-2*^{L811fsX}/CyO;P[*w*-IR]/TM6B, *r2d2*^{1}/CyO;P[*w*-IR]/TM6B, P[*w*-IR]/CyO;*ago2*^{414}/TM6B, *loqs*^{f00791}/CyO;P[*w*-IR]/TM6B, *cg9293*^{f03884}/CyO, *loqs*^{f00791}/CyO, *loqs*^{excision}/CyO, P[*w*-IR]; *loqs*^{f00791}/CyO; FRT82B *dcr-1*^{Q1147X}/TM6B.

Quantifying eye color

Red pigment was measured as described[201]. For each genotype, heads were manually dissected from 8 males 3–4 days after eclosion. For each individual measurement, two heads were homogenized in 0.1 ml of 0.01 M HCl dissolved in ethanol. The homogenates were incubated at 4°C overnight, warmed to 50°C for 5 min, and then clarified by centrifugation. The optical density of the supernatant was measured at 480 nm and normalized to that recorded for heads from wild-type Oregon R.

Preparation of lysate from heads

Wild-type or mutant flies were flash frozen in liquid nitrogen. Heads were separated from bodies by vigorous shaking in nested, pre-chilled sieves (U.S.A. standard sieve,

Humboldt MFG Co., Chicago, IL, USA), allowing the heads to pass through the top sieve (No. 25) and collecting them on the bottom sieve (No. 40). Heads were transferred to 0.5 ml microcentrifuge tubes, pre-chilled in liquid-nitrogen, and then homogenized using a plastic “pellet pestle” (Kontes, Vineland, NJ, USA) in 1 ml ice-cold lysis buffer (100 mM potassium acetate, 30 mM HEPES-KOH at pH 7.4, 2 mM magnesium acetate) containing 5 mM DTT and 1 mg/ml complete “mini” EDTA-free protease inhibitor tablets (Roche) per gram of heads. Lysate was clarified by centrifugation at 14,000 × g for 30 min at 4°C. The supernatant was aliquoted into pre-chilled microcentrifuge tubes, flash frozen in liquid nitrogen, and stored at -80°C. For each experiment, siRNA-protein complexes were assembled using equal amounts of total protein for all genotypes.

APPENDIX II

Target-directed destruction of small silencing RNAs

The work presented was a collaborative effort. Stefan Ameres and Michael Horwich demonstrated that extensive complementarity between a target RNA and an Argonaute1-bound miRNA triggers miRNA tailing and destruction. However, Argonaute2-bound small RNAs were immune to this phenomenon. In flies, Argonaute2-bound small RNAs—but not those bound to Argonaute1—bear a 2´-*O*-methyl group at their 3´ ends, added by the methytransferase Hen1. Therefore, we speculated that this modification blocks target-directed degradation for Argonaute2-bound small RNAs. To validate our hypothesis, I performed high-depth sequencing from *hen1* heterozygous and homozygous mutant heads, and found that in flies lacking Hen1, Argonaute2-associated siRNAs are tailed and degraded. Stefan Ameres performed the experiment, demonstrating methylation protected small RNAs from tailing and degradation. Jui-Hung, Jia Xu and Zhiping Weng performed Bioinformatic Analyses. The author, Stefan Ameres and Phillip Zamore, wrote the following text.

Introduction

Small silencing RNAs regulate gene expression, defend against viral infection, and protect the genome from transposons in nearly all eukaryotes[254]. In *Drosophila melanogaster*, conceptually similar but mechanistically different pathways produce

siRNAs and miRNAs. Fly siRNAs guide Argonaute2 (Ago2) to cleave target RNAs with extensive complementarity to the siRNA guide, a process termed RNA interference (RNAi), whereas miRNAs typically act through Argonaute1 (Ago1) to decrease the translation and stability of partially complementary mRNAs[158,241,278,314]. The difference in target complementarity between animal siRNAs and miRNAs stands in contrast to plants, where both siRNAs and miRNAs bind target mRNAs through extensive base pairing across the entire small RNA guide[157].

In flies, a key step in the production of a functional siRNA-Ago2 complex, but not a miRNA-Ago1 complex, is the addition of a 2´-*O*-methyl group to the 3´ end of the small RNA by Hen1[83,84,182], an *S*-adenosylmethionine-dependent methyltransferase first discovered in plants[85]. Plant Hen1 protects siRNAs and miRNAs alike from 3´-terminal uridylation and degradation[85,87,143]. In contrast, terminal 2´-*O*-methylation in flies is a hallmark of small RNAs bound to the RNAi protein Ago2 and is not found on small RNAs—typically miRNAs—bound to Ago1. Here, we report that extensive complementarity between a target RNA and an Argonaute1-bound miRNA triggers miRNA tailing and destruction in vivo and in cell lysates in vitro. The presence of a 3´ terminal 2´-*O*-methyl group blocks such target-dependent small RNA tailing and destruction. We propose that 3´ terminal 2´-*O*-methylation differentiates small RNA guides that extensively base pair to the RNAs they regulate from those small RNAs that bind their targets through only limited complementarity.

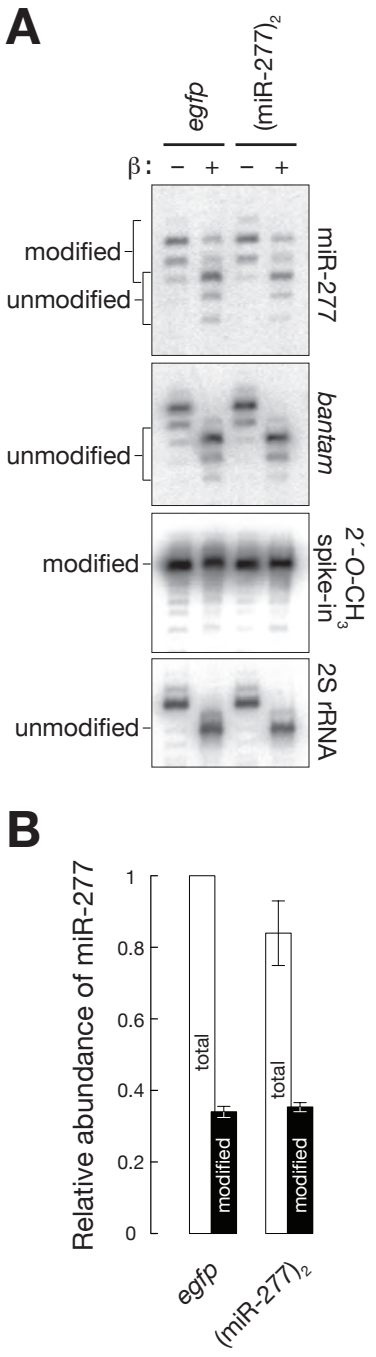
Results

A complementary target RNA directs degradation of Ago1-, but not Ago2-bound miR-277

Drosophila siRNAs and miRNAs partition between Ago1 and Ago2 according to their duplex structure[202,290,316,317]. Consequently, some miRNAs, including miR-277, partition into both Ago1 and Ago2[203]. Expression of *egfp* bearing two sites fully complementary to miR-277 caused a surprisingly small but significant reduction in the abundance of that miRNA ($p < 0.05$) (Fig. 1A and B). We used oxidation with $NaIO_4$ followed by β-elimination to distinguish between Ago1- and Ago2-loaded miR-277[83]. Ago2-loaded miRNAs bear 2′-*O*-methyl modified 3′ termini, making them refractory to oxidation; Ago1-loaded miRNAs bear 2′,3′ hydroxy 3′ termini, and oxidation followed by β-elimination removes their final nucleotide, making them one-nucleotide shorter. Relative to a control reporter, the miR-277–complementary reporter had no effect on Ago2-associated miR-277 (Fig. 1A and B). We conclude that the fully complementary target RNAs decreased the Ago1- but not the Ago2-bound miR-277, consistent with earlier observations that Ago2 but not Ago1 silences an *egfp* reporter with target sites perfectly complementary to miR-277[203].

Figure AII- 1. Methylation protects small RNAs from tailing and degradation. (A) Northern blot analysis of total RNA from clonal S2 cells stably expressing a control *egfp* mRNA (*egfp*) or an *egfp* mRNA bearing two target sites perfectly complementary to miR-277 [$(miR\text{-}277)_2$]. (B) Mean ± standard deviation for at least three biologically independent measurements of miR-277 abundance. “Modified” indicates the population of miR-277 resistant to oxidation and β-elimination.

Figure AII- 1.


A
egfp
(miR-277)2
β: − + − +
modified
unmodified
miR-277
unmodified
bantam
modified
2´-O-CH3 spike-in
unmodified
2S rRNA
B
Relative abundance of miR-277
1
0.8
0.6
0.4
0.2
0
total
modified
total
modified
egfp
(miR-277)2

The methyltransferase, Hen1, is required to stabilize Ago2-bound small RNAs

We sequenced 18 to 29 nt small RNAs from heterozygous or homozygous mutant *hen1*^{f00810} heads (Tables S1 and S2). Consistent with the idea that Hen1 does not act on Ago1-associated small RNAs, the absence of Hen1 in the mutant flies altered neither the abundance nor the length of most miRNAs (Fig. 2A). The most abundant class of Ago2-bound small RNAs are endogenous siRNAs (endo-siRNAs), and they are fully or extensively complementary to cellular or transposon-derived mRNAs[26-28,112]. The length and abundance of perfectly genome-matching endo-siRNA reads derived from transposon sequences was decreased in *hen1*^{f00810} mutant heads, compared to the heads of heterozygous siblings (Fig. 2B, upper panel). In contrast, prefix-matching endo-siRNA reads increased in the *hen1* mutant heads (Fig. 2B lower panel). The sequences of prefix-matching reads correspond to the reference fly genome for their first 15 or more nucleotides, then contain a short tail of 3′ nucleotides not found in genomic sequence. The majority of prefix reads contained a single 3′ uridine tail; the second most abundant nucleotide added was adenine. Longer tails comprising more than one non-genome matching addition were rare, but nearly always corresponded to homopolymeric stretches of uridines (Fig. 2B, lower panel). Such uridine tails are found on small RNAs in plants lacking Hen1 and are believed to tag siRNAs and miRNAs for destruction[87].

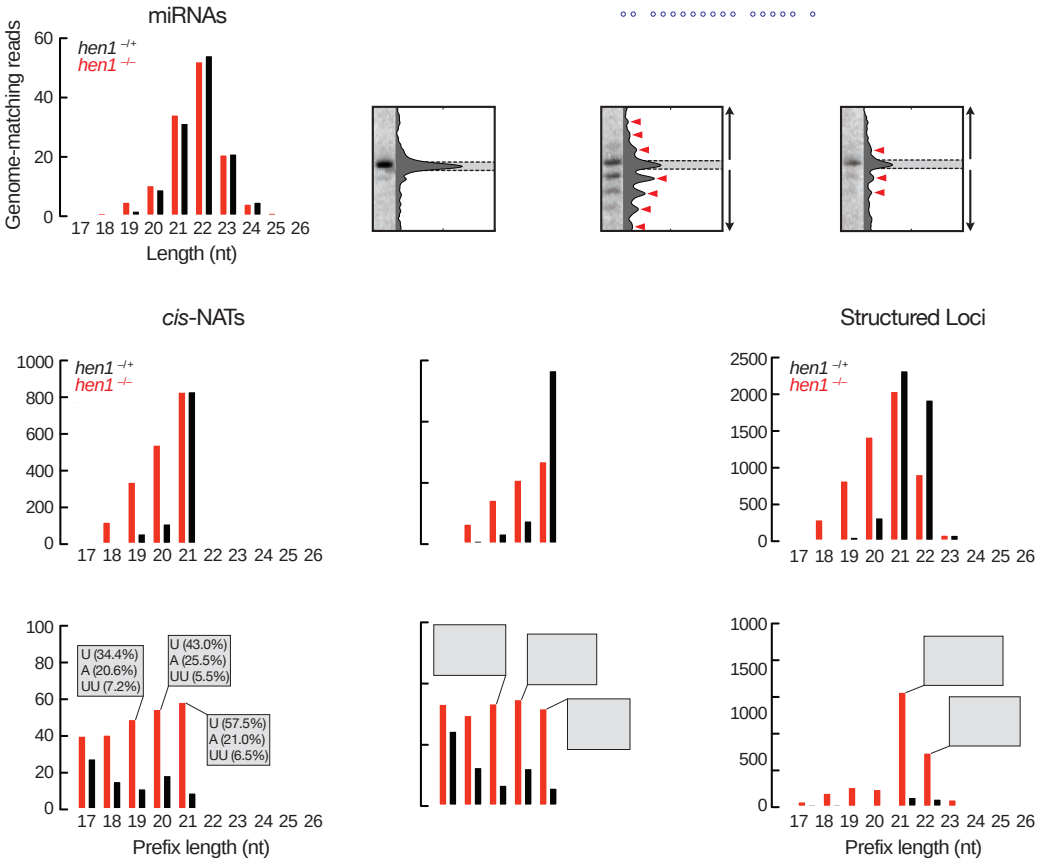
Endogenous siRNAs derived from structured loci (esiRNAs) differ from the classical features of siRNAs in their length (22 instead of 21 nt), their duplex structure (they contain bulges and mismatches instead of perfect pairing throughout the duplex), their biogenesis (they require Loquacious rather than R2D2), and their 5′ nucleotide bias

(they begin with U instead of C)[27-29]. In fact, the known determinants for assembling small RNAs into *Drosophila* Argonaute complexes would predict that some of esiRNAs should preferentially associate with Ago1 rather than Ago2[202,290]. Supporting this view, in UV cross-linking experiments in embryo lysate, the most abundant esiRNA, esi-2.1[27], loads predominantly into Ago1. Although in vivo, endo-siRNAs from structured loci associate to a larger extent with Ago1 than siRNAs derived from transposons or natural antisense transcripts, they accumulate mainly in Ago2[27-29].

In the absence of Hen1, esiRNAs generally become shorter (Fig. 2B). Moreover, analysis of published high throughput sequencing of small RNAs in S2 cells[27], which express both esi-2.1 and its highly complementary mRNA target, *mus308* (Fig. 2C), suggests that Ago1-associated esi-2.1 is subject to ongoing tailing and degradation: Ago1-bound esi-2.1 is shorter (its modal length is 20 rather than 22 nt), more heterogeneous in length, and contains a higher fraction of non-genome matching 3´ nucleotide additions (typically a single uridine) than Ago2-bound esi2.1. esi-2.1 is similarly less abundant, tailed, and degraded in both *hen1*^{f00810} and *ago2*^{414} whole mutant flies, compared to wild-type (Fig. 2C).

Figure AII-2. Small RNA tailing and degradation in vivo. (A) Length distribution of miRNAs from heterozygous (black bars) or homozygous (red bars) *hen1*^{f00810} fly heads. For each individual annotated pre-miRNA, reads were normalized to sequencing depth; reads for each distinct pre-miRNA were weighted equally to eliminate the influence of differences in transcriptional rates. (B) Length distribution of sequence reads perfectly matching the fly reference (top panel) or reads matching only within a 5´ prefix (bottom panel) from heterozygous (black bars) or homozygous (red bars) *hen1*^{f00810} heads. The three classes of endogenous siRNAs are analyzed separately: siRNAs derived from natural antisense transcripts (*cis*-NATs), from transposons, or from structured loci. The most frequent non-genome matching nucleotide additions are indicated in the gray boxes as a percent of all non-genome matching additions for specific prefix lengths. Reads are reported in parts per million. (C) The sequence of the esi-2.1 duplex and its cellular mRNA target, *mus308*. Northern blot analysis was used to detect esi-2.1 in total RNA from whole Oregon R (wild-type) or *hen1*^{f00810} or *ago2*^{414} mutant flies and the observed signal intensities (I, $\log_{10}$ scale) determined for each lane. Tailing and degradation products are marked with red arrowheads.

Figure AII-2.


miRNAs
Genome-matching reads
hen1 −/+
hen1 −/−
Length (nt)
cis-NATs
Structured Loci
U (34.4%)
A (20.6%)
UU (7.2%)
U (43.0%)
A (25.5%)
UU (5.5%)
U (57.5%)
A (21.0%)
UU (6.5%)
Prefix length (nt)

Two other abundant esiRNAs, esi-1.1 and esi-1.2, derive from more siRNA-like duplexes (Fig. 3). Like esi-2.1, esi-1.1 begins with uridine and loads in vitro into both Ago1 and Ago2 (Fig. 3G). In contrast, esi-1.2 starts with cytidine and loads efficiently into Ago2 (Fig. 3L); a 5´ cytidine has been proposed to favor Ago2 loading[26,290]. esi-1.1 and esi1.2 differ in the extent to which they were tailed and degraded in vivo in *hen1*^{f00810} (Fig. 3H and M) and *ago2*^{414} (Fig. 3I and N) mutant fly heads: degradation of esi-1.2, which favors loading into Ago2, was greater in these RNAi pathway mutants than for esi-1.1 (Fig. 4). We speculate that loss of Hen1 and Ago2 produce fundamentally distinct consequences: In a *hen1* mutant, Ago2-bound esiRNAs become tailed and degraded, because they no longer possess their protective, 3´ terminal, 2´-*O*-methyl modification. Thus, for small RNAs such as esi-1.1 and esi-2.1, tailed and shortened species comprise both Ago1- and Ago2-loaded RNAs. In contrast, in an *ago2* mutant, the normally Ago2-bound molecules no longer exist, so the only remaining tailed and shortened species must derive from Ago1-bound molecules.

esi-2.1 is the only structured locus-derived endo-siRNA for which a highly complementary target, *mus308*, has been described. Using quantitative RT-PCR, we were unable to detect *mus308* expression in heads. However, fly heads express mRNAs with sufficient complementarity to esi-1.1, esi-1.2, and esi-2.1 (Fig. 3E, J and O)—based on our in vitro results—to direct tailing and degradation[318].

Figure AII-3. Assembly, genetic requirements and potential destabilizing targets of three abundant structured loci endo-siRNAs. The duplex structures of esi-2.1, esi-1.1 and esi-1.2, three abundant, small RNAs derived from structured loci, predict they will partition differently between Ago1 and Ago2 (A, F and K). When the esi-2.1, esi-1.1 and esi-1.2 duplexes (guide strands were 5´ ^{32}P-radiolabeled) were incubated in *Drosophila* embryo lysate and loading was monitored by UV cross-linking (B, G and L), esi-2.1 and esi-1.1 loaded predominantly Ago1, whereas esi-1.2 loaded Ago2. Analysis of the length distributions of genome-matching (top panel) or prefix only (bottom panel) sequence reads for esi-2.1, esi-1.1 and esi-1.2 from heads of heterozygous or homozygous *hen1*^{f00810} flies (C, H and M) and of wild-type (Oregon R) or homozygous mutant *ago2*^{414} fly heads (D, I and N) revealed tailing and degradation in the mutants. The most frequent non-genome–matching, 3´ nucleotide additions are reported in the gray boxes as a percent of all non-genome–matching additions for each prefix length. Target RNAs that would be predicted from our in vitro results to possess sufficient complementarity to the respective esiRNA (red) to direct small RNA tailing and degradation (E, J and O) are expressed in fly heads according to publicly available data[318].

Figure AII-3.

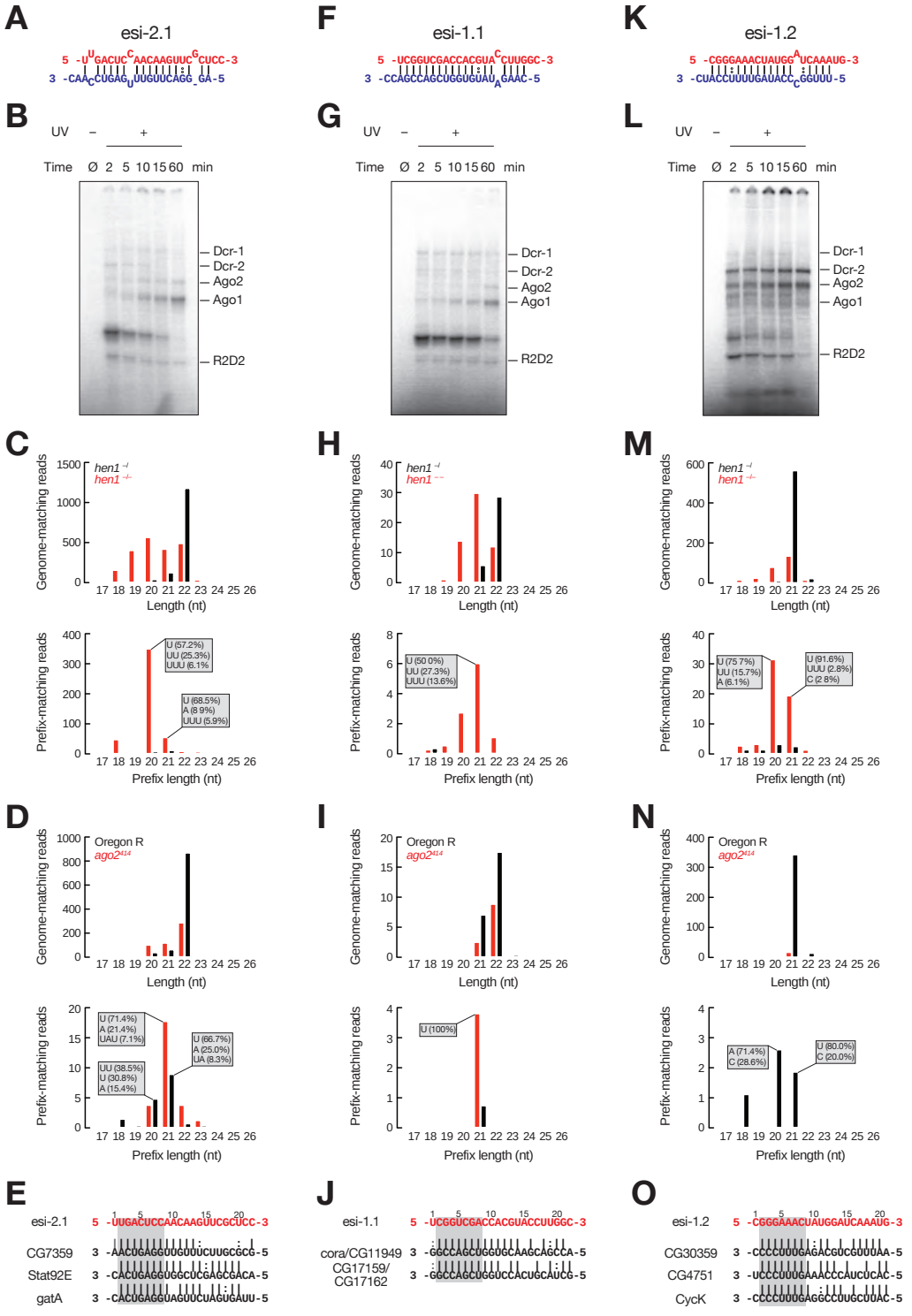
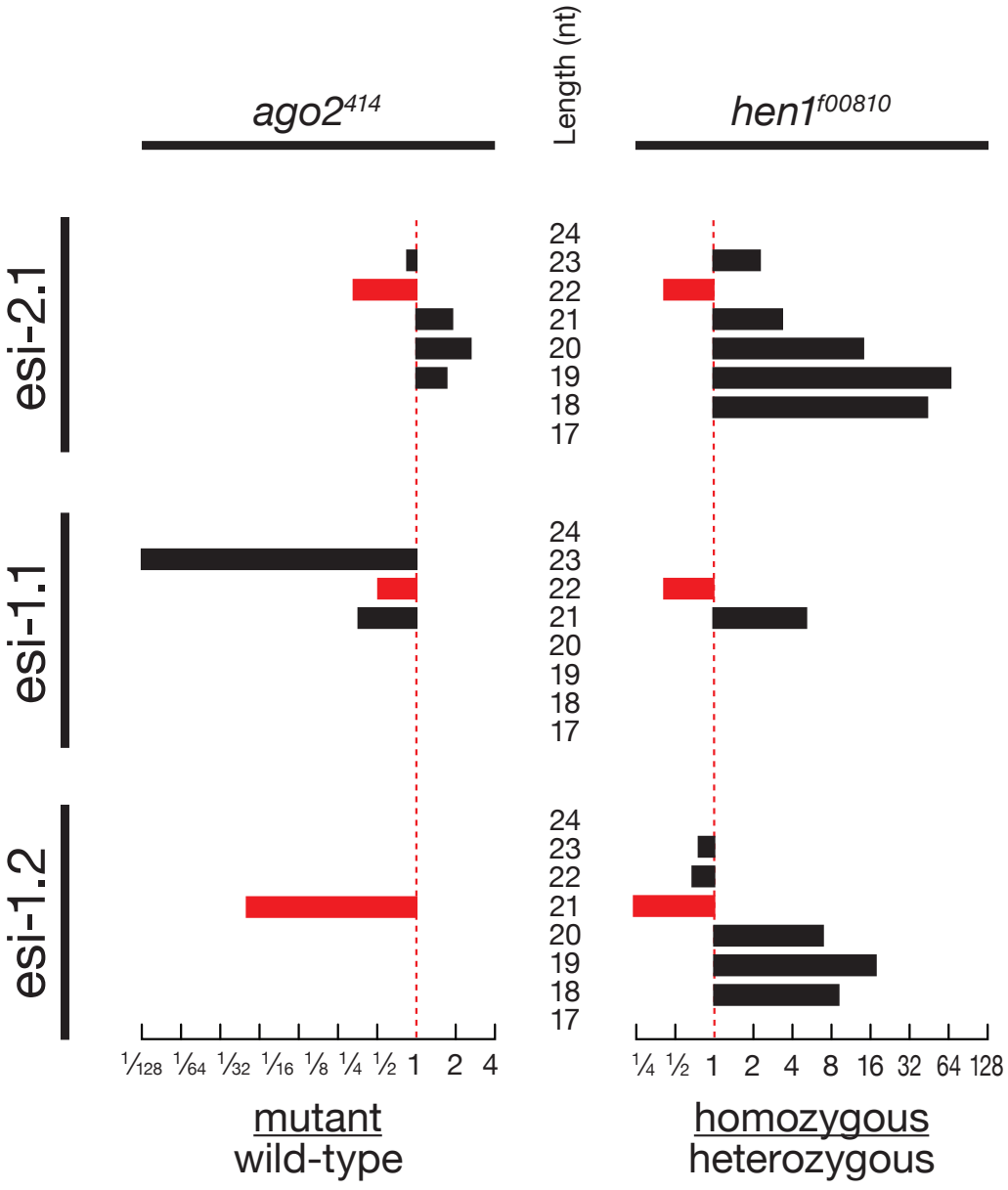

A
esi-2.1
5 -UUGACUCCAACAAGUUCGCUCC-3
3 -CAACCUGAGUUGUUCAGG_GA-5
F
esi-1.1
5 -UCGGUCGACCACGUACCUUGGC-3
3 -CCAGCCAGCUGGUGUAUGAACC-5
K
esi-1.2
5 -CGGGAAACUAUGGAUCAAAUG-3
3 -CUACCUUUUGAUACCCGGUUU-5
B
G
L
UV – +
Time Ø 2 5 10 15 60 min
Dcr-1
Dcr-2
Ago2
Ago1
R2D2
C
H
M
Genome-matching reads
hen1 -/-
hen1 -/-
Length (nt)
Prefix-matching reads
U (57.2%)
UU (25.3%)
UUU (6.1%
U (68.5%)
A (8.9%)
UUU (5.9%)
U (50.0%)
UU (27.3%)
UUU (13.6%)
U (75.7%)
UU (15.7%)
A (6.1%)
U (91.6%)
UUU (2.8%)
C (2.8%)
Prefix length (nt)
D
I
N
Oregon R
ago2414
U (71.4%)
A (21.4%)
UAU (7.1%)
UU (38.5%)
U (30.8%)
A (15.4%)
U (66.7%)
A (25.0%)
UA (8.3%)
U (100%)
A (71.4%)
C (28.6%)
U (80.0%)
C (20.0%)
E
esi-2.1 5 -UUGACUCCAACAAGUUCGCUCC-3
CG7359 3 -AACUGAGGUUGUUCUUGCGCG-5
Stat92E 3 -CACUGAGGUGGCUCGAGCGACA-5
gatA 3 -CACUGAGGUAGUUCUAGUGAUU-5
J
esi-1.1 5 -UCGGUCGACCACGUACCUUGGC-3
cora/CG11949 CG17159/ CG17162 3 -GGCCAGCUGGUGCCAUGCAUCG-5
O
esi-1.2 5 -CGGGAAACUAUGGAUCAAAUG-3
CG30359 3 -CCCCUUUGAGACGUCGUUUAA-5
CG4751 3 -UCCCUUUGAAACCCAUCUCAC-5
CycK 3 -CCCCUUUGAGGCCUUGCUUAC-5

Figure AII-4. Fold-change of esi-2.1, esi-1.1 and esi-1.2 in *hen1*^{f00810} and *ago2*^{414} mutant fly heads. Fold change of esi-2.1, esi-1.1 and esi-1.2 perfect genome matching reads of indicated length in *hen1*^{f00810} fly heads compared to heterozygous siblings (left panel) and *ago2*^{414} fly heads compared to Oregon R (right panel). Bars representing the full length sequence are depicted in red.

Figure AII-4.

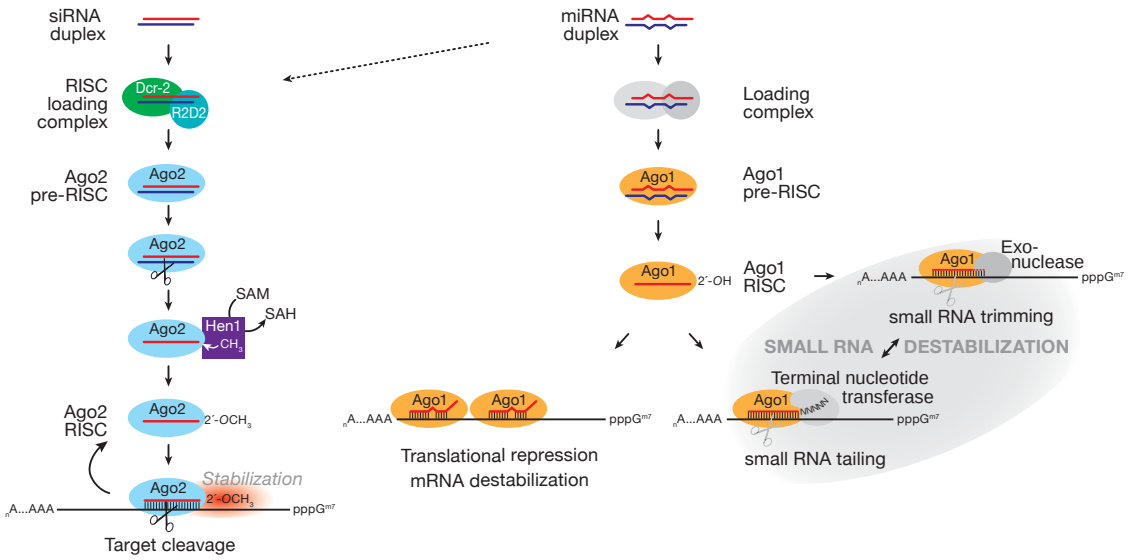

ago2414
Length (nt)
hen1f00810
esi-2.1
esi-1.1
esi-1.2
24
23
22
21
20
19
18
17
1/128 1/64 1/32 1/16 1/8 1/4 1/2 1 2 4
1/4 1/2 1 2 4 8 16 32 64 128
mutant
wild-type
homozygous
heterozygous

A model for small RNA degradation in Drosophila

Our data suggest a model for the influence of target RNA complementarity on small RNA abundance in *Drosophila* (Fig. 5). miRNAs typically direct Ago1 to bind target RNAs and repress their translation and decrease their stability[319]. Such binding is nearly always mediated by complementarity between the miRNA seed sequence and the target, with few additional base pairs tethering the two RNAs together. The presence of transcripts with extensive complementarity to Ago1-bound small RNAs results in small RNA degradation, which our data suggest involves a terminal nucleotide transferase and a 3´-to-5´ exonuclease. In contrast, Ago2-associated small RNAs are 2´-*O*-methyl modified by Hen1 as a final step of Ago2 loading. The methyl group blocks tailing and degradation; in *hen1* mutants, the unmethylated but Ago2-bound small RNAs are subject to target-directed degradation. Our model predicts that differential accumulation of small RNA species between Ago1 and Ago2 reflects not only small RNA sorting, but also the differential stability of Ago1- and Ago2-bound small RNAs in the face of a highly complementary target RNA. Thus, a subgroup of esiRNAs likely load into both Ago1 and Ago2 but accumulate mainly in Ago2 because Ago1-bound esiRNAs are subject to target-directed small RNA degradation.

Figure AII-5. A model for the influence of target RNA complementarity on small RNA stability in *Drosophila*.

Figure AII-5.


siRNA duplex
RISC loading complex
Dcr-2
R2D2
Ago2 pre-RISC
Ago2
SAM
SAH
Hen1
CH3
Ago2 RISC
2′-OCH3
Stabilization
nA...AAA
pppGm7
Target cleavage
miRNA duplex
Loading complex
Ago1
Ago1 pre-RISC
2′-OH
Ago1 RISC
Exo-nuclease
small RNA trimming
SMALL RNA
DESTABILIZATION
Terminal nucleotide transferase
NNNN
small RNA tailing
Translational repression
mRNA destabilization

Discussion

Our data establish that in flies the stability of small RNAs is determined by both the degree of complementarity between the small RNA and its target RNA and the identity of the Argonaute protein to which it is bound: highly complementary targets trigger tailing and degradation of Ago1-associated small RNAs. In contrast, such targets do not induce degradation of Ago2-associated small RNAs. The resistance of Ago2-associated small RNAs to target-directed degradation is thought to reflect the ability of Ago2, but not Ago1, to recruit Hen1 to add a methoxy group to the terminal 2′ carbon of the small RNA guide. Hence, Hen1 and the methoxy group it deposits on the guide RNA lies at the heart of the specialization of the two somatic RNA silencing pathways in flies: RNA methylation by Hen1 enables Ago2 to bind and cleave highly complementary target RNAs; the exclusion of Hen1 from the Ago1-loading pathway restricts Ago1-bound small RNAs to regulate only partially complementary targets. The fact that Ago1-associated small RNAs are sensitive to target-directed tailing and destruction has likely shaped the evolution of miRNA target sites in *Drosophila* and perhaps other animals: most predicted miRNA binding sites in animal 3′ UTRs lack substantial pairing to the small RNA 3′ end[158,291].

Even in *hen1*^{f00810} flies, small RNAs bound to Ago1 are more prone to target RNA-dependent degradation than those bound to Ago2 (Fig. 3). Ago1 is an inefficient ribonuclease whose catalytic rate is limited by the dissociation of its reaction products[203], whereas Ago2 is an efficient multiple-turnover enzyme[165]. The ability of Ago2 to rapidly cleave its RNA targets may limit its susceptibility to target-directed small RNA

degradation. In contrast, Ago1 likely resides on its target RNAs for much more time than Ago2, making Ago1-bound small RNAs good substrates for target-directed tailing and degradation.

Our data also link target RNA-directed small RNA degradation to 3′ uridylation. Uridylation of mRNA and non-coding RNAs has been described in fission yeast and metazoans where it was implicated in general or specific RNA turnover[320,321]. The apparent discrepancy between target RNA-dependent nucleotidyl transfer on small RNAs in vitro, where almost exclusively adenines were added and in vivo, where the most common nucleotide added was uracil, followed by adenine, might be explained by the fact that 3′ nucleotidyl transferases, e.g. terminal uridylyl transferases (TUTases), can use either ATP or UTP in tailing assays[322]. Alternatively distinct enzymes might add U and A to the 3′ end of small RNAs. Also, adenylation of nuclear RNAs is a signal for degradation[323]; the presence of nuclear components in embryo lysate might explain the predominance of A tailing in the lysate.

Uridylation of small RNAs as well as Ago2-cleaved, 5′ target RNA fragments has been linked to RNA turnover[87,324-326]. The molecular basis for the tailing and destruction of Argonaute-bound small RNAs is unknown. The ends of small RNAs are anchored to Argonaute proteins through binding of the 5′ small RNA phosphate to a pocket composed mainly of residues from the Mid domain and binding of the 3′ end of the small RNA to the PAZ domain[274,275,304,327-332]. Access to the 3′ end of the small RNA likely requires dislodging it from the PAZ domain. Recent crystal structures of a eubacterial Argonaute protein confirms earlier suggestions that extensive pairing of the 3′ half of an

siRNA with its target releases its 3′ end from the PAZ domain[298,333]. Our data are consistent with the idea that extensive pairing to a target RNA exposes a small RNA to nucleotidyl transferases and 3′-to-5′ exonuclease enzymes.

Materials and Methods

General Methods

Total RNA from flies, S2 cells or HeLa cells was purified using the MirVana kit (Ambion) or Trizol (Invitrogen). Northern blot analysis[134], β-elimination[83] have been described previously. Stable cultured S2 cell lines were generated as described[134] and transfected using Cellfectin (Invitrogen) according to the manufacturer's instructions; and total RNA was isolated 48 h later. UV cross-linking experiments were performed essentially as described with the sample ~3 cm below the light bulbs[290].

Small RNA library construction and deep sequencing

Library construction and deep sequencing was performed as described[26,244]. Published libraries used in this study were 18–29 nt total RNA libraries from Oregon R and *ago2*^{414} fly heads[26,244,290] and libraries generated from small RNAs immunoprecipitated with Ago1 as well as Ago2 from S2 cell lysates[27].

Supplemental Materials

Supplemental Tables

Table AII-S1. Sequencing statistics: Analysis of genome matching reads. Somatic tissue was prepared by mechanical separation of fly heads from bodies. "Small RNA reads (excluding ncRNAs)" correspond to genome-matching reads after excluding annotated non-coding RNAs (ncRNAs), such as rRNA, snRNA, snoRNA, or tRNA. "Transposon-matching reads" correspond to small RNAs mapped to *Drosophila melanogaster* transposons. "Cis NAT-matching reads" correspond to reads matching to mRNAs[26,27,112]. "Structured loci-matching reads" correspond to reads that map to two distinct loci in the *Drosophila melanogaster* genome (CG18824 and a locus overlapping with CG4068), the transcripts of which fold into long hairpin structures and produce the majority of small RNAs of this class [27,28]. Where reads were normalized to genome matching reads (excluding ncRNAs), they are reported in parts per million (ppm). N.A., not applicable.

Head genotype	Total reads	Reads perfectly matching genome	Reads matching annotated ncRNAs	Small RNA reads (excluding ncRNAs)	Pre-miRNA-matching reads (ppm)	Reads excluding ncRNA and pre-miRNA-matching (ppm)	Transposon-matching reads (ppm)	cis-NAT-matching reads (ppm)	Structured loci-matching reads (ppm)
hen1^{f00810}/CyO	6,413,029	2,310,112	408,102	1,902,013	888,203	111,497	22,139	1,486	6,877
hen1^{f00810}	7,221,663	2,932,242	389,272	2,242,670	871,819	128,181	24,964	2,639	8,049

Table AII-S2. Sequencing statistics: Analysis of 5´ prefix-matching reads. Analysis was as described in Table S1 except that only reads not matching the reference genome across their entire length were considered in order to detect small RNAs bearing 3´ terminal, non-genome matching additions. The analysis employed previously published datasets prepared from Oregon R fly heads[26,244,290], *ago2*^{414} mutant fly heads[26].

Head genotype / Sample	Total reads	Prefixes matching genome	Prefixes excluding internal mm	Prefixes matching annotated ncRNAs	Prefixes (excluding ncRNAs)	Pre-miRNA-matching prefixes	Transposon-matching prefixes	cis-NAT-matching prefixes	Structured loci-matching prefixes
hen1^{f00810}/CyO	6,413,029	1,124,892	937,420	96,733	840,717	108,049	22,860	1,407	477
hen1^{f00810}	7,221,663	1,396,479	1,198,890	106,802	1,092,082	143,823	32,087	2,121	4,374
Oregon R	7,307,082	1,393,383	1,209,692	127,284	1,082,411	84,342	27,820	2,090	344
ago2^{414}	1,942,282	379,709	309,220	16,639	292,911	42,122	9,762	228	82

Bibliography

1. Zhao, T. et al. A complex system of small RNAs in the unicellular green alga Chlamydomonas reinhardtii. *Genes Dev* **21**, 1190-1203 (2007).
2. Molnar, A., Schwach, F., Studholme, D. J., Thuenemann, E. C. & Baulcombe, D. C. miRNAs control gene expression in the single-cell alga Chlamydomonas reinhardtii. *Nature* **447**, 1126-1129 (2007).
3. Lee, R. C. & Ambros, V. An Extensive Class of Small RNAs in Caenorhabditis elegans. *Science* **294**, 862-864 (2001).
4. Lau, N. C., Lim, L. P., Weinstein, E. G. & Bartel, D. P. An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. *Science* **294**, 858-862 (2001).
5. Lagos-Quintana, M., Rauhut, R., Lendeckel, W. & Tuschl, T. Identification of Novel Genes Coding for Small Expressed RNAs. *Science* **294**, 853-858 (2001).
6. Jones-Rhoades, M. W. & Bartel, D. P. Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Mol Cell* **14**, 787-799 (2004).
7. Chan, S. W. et al. RNA silencing genes control de novo DNA methylation. *Science* **303**, 1336 (2004).
8. Tran, R. K. et al. DNA methylation profiling identifies CG methylation clusters in Arabidopsis genes. *Curr Biol* **15**, 154-159 (2005).
9. Zilberman, D., Cao, X. & Jacobsen, S. E. ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. *Science* **299**, 716-719 (2003).
10. Hamilton, A., Voinnet, O., Chappell, L. & Baulcombe, D. Two classes of short interfering RNA in RNA silencing. *EMBO J* **21**, 4671-4679 (2002).
11. Llave, C., Kasschau, K. D., Rector, M. A. & Carrington, J. C. Endogenous and Silencing-Associated Small RNAs in Plants. *Plant Cell* **14**, 1605-1619 (2002).
12. Mette, M. F., Aufsatz, W., van der Winden, J., Matzke, M. A. & Matzke, A. J. Transcriptional silencing and promoter methylation triggered by double- stranded RNA. *EMBO J* **19**, 5194-201. (2000).
13. Xie, Z. et al. Genetic and Functional Diversification of Small RNA Pathways in Plants. *PLoS Biol* **2**, E104 (2004).
14. Vazquez, F. et al. Endogenous trans-acting siRNAs regulate the accumulation of Arabidopsis mRNAs. *Mol Cell* **16**, 69-79 (2004).
15. Peragine, A., Yoshikawa, M., Wu, G., Albrecht, H. L. & Poethig, R. S. SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of trans-acting siRNAs in Arabidopsis. *Genes Dev* **18**, 2368-2379 (2004).
16. Yoshikawa, M., Peragine, A., Park, M. Y. & Poethig, R. S. A pathway for the biogenesis of trans-acting siRNAs in Arabidopsis. *Genes Dev* **19**, 2164-2175 (2005).
17. Williams, L., Carles, C. C., Osmont, K. S. & Fletcher, J. C. A database analysis method identifies an endogenous trans-acting short-interfering RNA that targets the

Arabidopsis ARF2, ARF3, and ARF4 genes. *Proc Natl Acad Sci U S A* **102**, 9703-9708 (2005).
18. Allen, E., Xie, Z., Gustafson, A. M. & Carrington, J. C. microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell* **121**, 207-221 (2005).
19. Katiyar-Agarwal, S. et al. A pathogen-inducible endogenous siRNA in plant immunity. *Proc Natl Acad Sci U S A* **103**, 18002-18007 (2006).
20. Borsani, O., Zhu, J., Verslues, P. E., Sunkar, R. & Zhu, J. K. Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in Arabidopsis. *Cell* **123**, 1279-1291 (2005).
21. Hamilton, A. J. & Baulcombe, D. C. A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* **286**, 950-952 (1999).
22. Zamore, P. D., Tuschl, T., Sharp, P. A. & Bartel, D. P. RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* **101**, 25-33 (2000).
23. Elbashir, S. M. et al. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**, 494-498 (2001).
24. Elbashir, S. M., Lendeckel, W. & Tuschl, T. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev* **15**, 188-200 (2001).
25. Yang, N. & Kazazian, H. H. J. L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. *Nat Struct Mol Biol* **13**, 763-771 (2006).
26. Ghildiyal, M. et al. Endogenous siRNAs derived from transposons and mRNAs in Drosophila somatic cells. *Science* **320**, 1077-1081 (2008).
27. Czech, B. et al. An endogenous small interfering RNA pathway in Drosophila. *Nature* **453**, 798-802 (2008).
28. Okamura, K. et al. The Drosophila hairpin RNA pathway generates endogenous short interfering RNAs. *Nature* **453**, 803-806 (2008).
29. Kawamura, Y. et al. Drosophila endogenous small RNAs bind to Argonaute 2 in somatic cells. *Nature* **453**, 793-797 (2008).
30. Tabara, H. et al. The rde-1 gene, RNA interference, and transposon silencing in C. elegans. *Cell* **99**, 123-132 (1999).
31. Ketting, R. F., Haverkamp, T. H., van Luenen, H. G. & Plasterk, R. H. Mut-7 of C. elegans, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and RNaseD. *Cell* **99**, 133-141 (1999).
32. Sijen, T. & Plasterk, R. H. Transposon silencing in the Caenorhabditis elegans germ line by natural RNAi. *Nature* **426**, 310-314 (2003).
33. Tam, O. H. et al. Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature* (2008).
34. Watanabe, T. et al. Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* **453**, 539-543 (2008).
35. Aravin, A. A. et al. Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the *D. melanogaster* germline. *Curr Biol* **11**, 1017–1027 (2001).

36. Houwing, S. et al. A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. *Cell* **129**, 69-82 (2007).
37. Aravin, A. et al. A novel class of small RNAs bind to MILI protein in mouse testes. *Nature* **442**, 203-207 (2006).
38. Lau, N. C. et al. Characterization of the piRNA complex from rat testes. *Science* **313**, 363-367 (2006).
39. Aravin, A. A., Sachidanandam, R., Girard, A., Fejes-Toth, K. & Hannon, G. J. Developmentally regulated piRNA clusters implicate MILI in transposon control. *Science* **316**, 744-747 (2007).
40. Girard, A., Sachidanandam, R., Hannon, G. J. & Carmell, M. A. A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature* **442**, 199-202 (2006).
41. Grivna, S. T., Beyret, E., Wang, Z. & Lin, H. A novel class of small RNAs in mouse spermatogenic cells. *Genes Dev* **20**, 1709-1714 (2006).
42. Grivna, S. T., Pyhtila, B. & Lin, H. MIWI associates with translational machinery and PIWI-interacting RNAs (piRNAs) in regulating spermatogenesis. *Proc Natl Acad Sci U S A* **103**, 13415-13420 (2006).
43. Brennecke, J. et al. Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell* **128**, 1089-1103 (2007).
44. Ruby, J. G. et al. Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in C. elegans. *Cell* **127**, 1193-1207 (2006).
45. Batista, P. J. et al. PRG-1 and 21U-RNAs Interact to Form the piRNA Complex Required for Fertility in C. elegans. *Mol Cell* (2008).
46. Das, P. P. et al. Piwi and piRNAs Act Upstream of an Endogenous siRNA Pathway to Suppress Tc3 Transposon Mobility in the Caenorhabditis elegans Germline. *Mol Cell* **31**, 79-90 (2008).
47. Wang, G. & Reinke, V. A C. elegans Piwi, PRG-1, Regulates 21U-RNAs during Spermatogenesis. *Curr Biol* **18**, 861-867 (2008).
48. Napoli, C., Lemieux, C. & Jorgensen, R. A. Introduction of a chimeric chalcone synthase gene into Petunia results in reversible co-suppression of homologous genes in trans. *Plant Cell* **2**, 279-289 (1990).
49. van der Krol, A. R., Mur, L. A., Beld, M., Mol, J. N. M. & Stuitji, A. R. Flavonoid genes in petunia: Addition of a limited number of gene copies may lead to a suppression of gene expression. *Plant Cell* **2**, 291-299. (1990).
50. Baulcombe, D. C. RNA as a target and an initiator of post-transcriptional gene silencing in transgenic plants. *Plant Mol Biol* **32**, 79-88 (1996).
51. Lindbo, J. A. & Dougherty, W. G. Untranslatable transcripts of the tobacco etch virus coat protein gene sequence can interfere with tobacco etch virus replication in transgenic plants and protoplasts. *Virology* **189**, 725-733 (1992).
52. Lindbo, J. A. & Dougherty, W. G. Pathogen-derived resistance to a potyvirus: immune and resistant phenotypes in transgenic tobacco expressing altered forms of a potyvirus coat protein nucleotide sequence. *Mol Plant Microbe Interact* **5**, 144-153 (1992).

53. Voinnet, O. & Baulcombe, D. C. Systemic signalling in gene silencing. *Nature* **389**, 553 (1997).
54. Guo, S. & Kemphues, K. J. par-1, a gene required for establishing polarity in C. elegans embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed. *Cell* **81**, 611-20. (1995).
55. Fire, A. et al. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **391**, 806-811 (1998).
56. Voinnet, O. Non-cell autonomous RNA silencing. *FEBS Lett* **579**, 5858-5871 (2005).
57. Grishok, A., Tabara, H. & Mello, C. C. Genetic requirements for inheritance of RNAi in *C. elegans*. *Science* **287**, 2494-2497 (2000).
58. Hammond, S. M., Bernstein, E., Beach, D. & Hannon, G. J. An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. *Nature* **404**, 293-296 (2000).
59. Bernstein, E., Caudy, A. A., Hammond, S. M. & Hannon, G. J. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409**, 363-366 (2001).
60. Elbashir, S. M., Martinez, J., Patkaniowska, A., Lendeckel, W. & Tuschl, T. Functional anatomy of siRNAs for mediating efficient RNAi in Drosophila melanogaster embryo lysate. *EMBO J* **20**, 6877-6888 (2001).
61. Liu, J. et al. A role for the P-body component GW182 in microRNA function. *Nat Cell Biol* **7**, 1161-1166 (2005).
62. Hutvagner, G. et al. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* **293**, 834-838 (2001).
63. 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* **106**, 23-34 (2001).
64. Ketting, R. F. et al. Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. *Genes Dev* **15**, 2654-2659 (2001).
65. Knight, S. W. & Bass, B. L. A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in Caenorhabditis elegans. *Science* **293**, 2269-2271 (2001).
66. Lee, Y. S. et al. Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. *Cell* **117**, 69-81 (2004).
67. Obbard, D. J., Jiggins, F. M., Halligan, D. L. & Little, T. J. Natural selection drives extremely rapid evolution in antiviral RNAi genes. *Curr Biol* **16**, 580-585 (2006).
68. Tabara, H., Yigit, E., Siomi, H. & Mello, C. C. The dsRNA Binding Protein RDE-4 Interacts with RDE-1, DCR-1, and a DexH-Box Helicase to Direct RNAi in C. elegans. *Cell* **109**, 861-871 (2002).
69. Shaham, S. Worming into the cell: viral reproduction in Caenorhabditis elegans. *Proc Natl Acad Sci U S A* **103**, 3955-3956 (2006).
70. Williams, B. R. PKR; a sentinel kinase for cellular stress. *Oncogene* **18**, 6112-6120 (1999).

71. Kunzi, M. S. & Pitha, P. M. Interferon research: a brief history. *Methods Mol Med* **116**, 25-35 (2005).
72. Vilcek, J. Fifty years of interferon research: aiming at a moving target. *Immunity* **25**, 343-348 (2006).
73. Schwarz, D. S. et al. Asymmetry in the assembly of the RNAi enzyme complex. *Cell* **115**, 199-208 (2003).
74. Khvorova, A., Reynolds, A. & Jayasena, S. D. Functional siRNAs and miRNAs exhibit strand bias. *Cell* **115**, 209-216 (2003).
75. Aza-Blanc, P. et al. Identification of modulators of TRAIL-induced apoptosis via RNAi-based phenotypic screening. *Mol Cell* **12**, 627-637 (2003).
76. Liu, Q. et al. R2D2, a Bridge Between the Initiation and Effector Steps of the Drosophila RNAi Pathway. *Science* **301**, 1921-1925 (2003).
77. Tomari, Y., Matranga, C., Haley, B., Martinez, N. & Zamore, P. D. A protein sensor for siRNA asymmetry. *Science* **306**, 1377-1380 (2004).
78. Matranga, C., Tomari, Y., Shin, C., Bartel, D. P. & Zamore, P. D. Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. *Cell* **123**, 607-620 (2005).
79. Kim, K., Lee, Y. S. & Carthew, R. W. Conversion of pre-RISC to holo-RISC by Ago2 during assembly of RNAi complexes. *RNA* **13**, 22-29 (2006).
80. Leuschner, P. J., Ameres, S. L., Kueng, S. & Martinez, J. Cleavage of the siRNA passenger strand during RISC assembly in human cells. *EMBO Rep* **7**, 314-320 (2006).
81. Miyoshi, K., Tsukumo, H., Nagami, T., Siomi, H. & Siomi, M. C. Slicer function of Drosophila Argonautes and its involvement in RISC formation. *Genes Dev* **19**, 2837-2848 (2005).
82. Rand, T. A., Petersen, S., Du, F. & Wang, X. Argonaute2 Cleaves the Anti-Guide Strand of siRNA during RISC Activation. *Cell* **123**, 621-629 (2005).
83. Horwich, M. D. et al. The Drosophila RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. *Curr Biol* **17**, 1265-1272 (2007).
84. Pelisson, A., Sarot, E., Payen-Groschene, G. & Bucheton, A. A novel repeat-associated small interfering RNA-mediated silencing pathway downregulates complementary sense gypsy transcripts in somatic cells of the Drosophila ovary. *J Virol* **81**, 1951-1960 (2007).
85. Yu, B. et al. Methylation as a crucial step in plant microRNA biogenesis. *Science* **307**, 932-935 (2005).
86. Ramachandran, V. & Chen, X. Small RNA metabolism in Arabidopsis. *Trends Plant Sci* **13**, 368-374 (2008).
87. Li, J., Yang, Z., Yu, B., Liu, J. & Chen, X. Methylation protects miRNAs and siRNAs from a 3'-end uridylation activity in Arabidopsis. *Curr Biol* **15**, 1501-1507 (2005).
88. Tolia, N. H. & Joshua-Tor, L. Slicer and the Argonautes. *Nat Chem Biol* **3**, 36-43 (2007).

89. Tang, G., Reinhart, B. J., Bartel, D. P. & Zamore, P. D. A biochemical framework for RNA silencing in plants. *Genes Dev* **17**, 49-63 (2003).
90. Dunoyer, P., Himber, C., Ruiz-Ferrer, V., Alioua, A. & Voinnet, O. Intra- and intercellular RNA interference in Arabidopsis thaliana requires components of the microRNA and heterochromatic silencing pathways. *Nat Genet* **39**, 848-856 (2007).
91. Bouche, N., Lauressergues, D., Gasciolli, V. & Vaucheret, H. An antagonistic function for Arabidopsis DCL2 in development and a new function for DCL4 in generating viral siRNAs. *EMBO J* **25**, 3347-3356 (2006).
92. Deleris, A. et al. Hierarchical action and inhibition of plant Dicer-like proteins in antiviral defense. *Science* **313**, 68-71 (2006).
93. Gasciolli, V., Mallory, A. C., Bartel, D. P. & Vaucheret, H. Partially redundant functions of Arabidopsis DICER-like enzymes and a role for DCL4 in producing trans-acting siRNAs. *Curr Biol* **15**, 1494-1500 (2005).
94. Chan, S. W. Inputs and outputs for chromatin-targeted RNAi. *Trends Plant Sci* **13**, 383-389 (2008).
95. Qu, F. et al. RDR6 has a broad-spectrum but temperature-dependent antiviral defense role in Nicotiana benthamiana. *J Virol* **79**, 15209-15217 (2005).
96. Voinnet, O. Use, tolerance and avoidance of amplified RNA silencing by plants. *Trends Plant Sci* **13**, 317-328 (2008).
97. Gazzani, S., Lawrenson, T., Woodward, C., Headon, D. & Sablowski, R. A link between mRNA turnover and RNA interference in Arabidopsis. *Science* **306**, 1046-1048 (2004).
98. Gy, I. et al. Arabidopsis FIERY1, XRN2, and XRN3 are endogenous RNA silencing suppressors. *Plant Cell* **19**, 3451-3461 (2007).
99. Ambros, V., Lee, R. C., Lavanway, A., Williams, P. T. & Jewell, D. MicroRNAs and other tiny endogenous RNAs in C. elegans. *Curr Biol* **13**, 807-818 (2003).
100. Zheng, X., Zhu, J., Kapoor, A. & Zhu, J. K. Role of Arabidopsis AGO6 in siRNA accumulation, DNA methylation and transcriptional gene silencing. *EMBO J* **26**, 1691-1701 (2007).
101. Boutet, S. et al. Arabidopsis HEN1: a genetic link between endogenous miRNA controlling development and siRNA controlling transgene silencing and virus resistance. *Curr Biol* **13**, 843-848 (2003).
102. El-Shami, M. et al. Reiterated WG/GW motifs form functionally and evolutionarily conserved ARGONAUTE-binding platforms in RNAi-related components. *Genes Dev* **21**, 2539-2544 (2007).
103. Herr, A. J., Jensen, M. B., Dalmay, T. & Baulcombe, D. C. RNA polymerase IV directs silencing of endogenous DNA. *Science* **308**, 118-120 (2005).
104. Kanoh, J., Sadaie, M., Urano, T. & Ishikawa, F. Telomere binding protein Taz1 establishes Swi6 heterochromatin independently of RNAi at telomeres. *Curr Biol* **15**, 1808-1819 (2005).
105. Lee, S. K. et al. Lentiviral delivery of short hairpin RNAs protects CD4 T cells from multiple clades and primary isolates of HIV. *Blood* **106**, 818-826 (2005).
106. Onodera, Y. et al. Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell* **120**, 613-622 (2005).

107. Pontier, D. et al. Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases IV in Arabidopsis. *Genes Dev* **19**, 2030-2040 (2005).
108. Montgomery, T. A. et al. Specificity of ARGONAUTE7-miR390 interaction and dual functionality in TAS3 trans-acting siRNA formation. *Cell* **133**, 128-141 (2008).
109. Axtell, M. J., Jan, C., Rajagopalan, R. & Bartel, D. P. A two-hit trigger for siRNA biogenesis in plants. *Cell* **127**, 565-577 (2006).
110. Zhang, D. & Trudeau, V. L. The XS domain of a plant specific SGS3 protein adopts a unique RNA recognition motif (RRM) fold. *Cell Cycle* **7**, 2268-2270 (2008).
111. Katiyar-Agarwal, S., Gao, S., Vivian-Smith, A. & Jin, H. A novel class of bacteria-induced small RNAs in Arabidopsis. *Genes Dev* **21**, 3123-3134 (2007).
112. Okamura, K., Balla, S., Martin, R., Liu, N. & Lai, E. C. Two distinct mechanisms generate endogenous siRNAs from bidirectional transcription in Drosophila melanogaster. *Nat Struct Mol Biol* **15**, 581-590 (2008).
113. Chung, W. J., Okamura, K., Martin, R. & Lai, E. C. Endogenous RNA interference provides a somatic defense against Drosophila transposons. *Curr Biol* **18**, 795-802 (2008).
114. Sasidharan, R. & Gerstein, M. Genomics: protein fossils live on as RNA. *Nature* **453**, 729-731 (2008).
115. Neilson, J. R. & Sharp, P. A. Small RNA regulators of gene expression. *Cell* **134**, 899-902 (2008).
116. Lee, R. C., Feinbaum, R. L. & Ambros, V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75**, 843-54. (1993).
117. Wightman, B., Burglin, T. R., Gatto, J., Arasu, P. & Ruvkun, G. Negative regulatory sequences in the lin-14 3'-untranslated region are necessary to generate a temporal switch during Caenorhabditis elegans development. *Genes Dev* **5**, 1813-124. (1991).
118. Wightman, B., Ha, I. & Ruvkun, G. Posttranscriptional regulation of the heterochronic gene lin-14 by lin- 4 mediates temporal pattern formation in C. elegans. *Cell* **75**, 855-62. (1993).
119. Griffiths-Jones, S., Saini, H. K., van Dongen, S. & Enright, A. J. miRBase: tools for microRNA genomics. *Nucleic Acids Res* **36**, D154-8 (2008).
120. Lee, Y., Jeon, K., Lee, J. T., Kim, S. & Kim, V. N. MicroRNA maturation: stepwise processing and subcellular localization. *EMBO J* **21**, 4663-4670 (2002).
121. Lee, Y. et al. MicroRNA genes are transcribed by RNA polymerase II. *EMBO J* **23**, 4051-4060 (2004).
122. Cai, X., Hagedorn, C. H. & Cullen, B. R. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA* **10**, 1957-1966 (2004).
123. Birney, E. et al. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447**, 799-816 (2007).

124. Lee, Y. et al. The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**, 415-419 (2003).
125. Denli, A. M., Tops, B. B., Plasterk, R. H., Ketting, R. F. & Hannon, G. J. Processing of primary microRNAs by the Microprocessor complex. *Nature* **432**, 231-235 (2004).
126. Gregory, R. I. et al. The Microprocessor complex mediates the genesis of microRNAs. *Nature* **432**, 235-240 (2004).
127. Han, J. et al. The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev* **18**, 3016-3027 (2004).
128. Landthaler, M., Yalcin, A. & Tuschl, T. The Human DiGeorge Syndrome Critical Region Gene 8 and Its D. melanogaster Homolog Are Required for miRNA Biogenesis. *Curr Biol* **14**, 2162-2167 (2004).
129. Yi, R., Qin, Y., Macara, I. G. & Cullen, B. R. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev* **17**, 3011-3016 (2003).
130. Bohnsack, M. T., Czaplinski, K. & Gorlich, D. Exportin 5 is a Ran GTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA* **10**, 185-191 (2004).
131. Lund, E., Guttinger, S., Calado, A., Dahlberg, J. E. & Kutay, U. Nuclear export of microRNA precursors. *Science* **303**, 95-98 (2004).
132. Zeng, Y. & Cullen, B. R. Structural requirements for pre-microRNA binding and nuclear export by Exportin 5. *Nucleic Acids Res* **32**, 4776-4785 (2004).
133. Chendrimada, T. P. et al. TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature* **436**, 740-744 (2005).
134. Forstemann, K. et al. Normal microRNA maturation and germ-line stem cell maintenance requires Loquacious, a double-stranded RNA-binding domain protein. *PLoS Biol* **3**, e236 (2005).
135. Jiang, F. et al. Dicer-1 and R3D1-L catalyze microRNA maturation in Drosophila. *Genes Dev* **19**, 1674-1679 (2005).
136. Lee, Y. et al. The role of PACT in the RNA silencing pathway. *EMBO J* **25**, 522-532 (2006).
137. Saito, K., Ishizuka, A., Siomi, H. & Siomi, M. C. Processing of pre-microRNAs by the Dicer-1-Loquacious complex in Drosophila cells. *PLoS Biol* **3**, e235 (2005).
138. Okamura, K., Hagen, J. W., Duan, H., Tyler, D. M. & Lai, E. C. The Mirtron Pathway Generates microRNA-Class Regulatory RNAs in Drosophila. *Cell* **130**, 89-100 (2007).
139. Ruby, J. G., Jan, C. H. & Bartel, D. P. Intronic microRNA precursors that bypass Drosha processing. *Nature* **448**, 83-86 (2007).
140. Glazov, E. A. et al. A microRNA catalog of the developing chicken embryo identified by a deep sequencing approach. *Genome Res* **18**, 957-964 (2008).
141. Berezikov, E., Chung, W. J., Willis, J., Cuppen, E. & Lai, E. C. Mammalian mirtron genes. *Mol Cell* **28**, 328-336 (2007).
142. Babiarz, J. E., Ruby, J. G., Wang, Y., Bartel, D. P. & Blelloch, R. Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes Dev* **22**, 2773-2785 (2008).

143. Park, W., Li, J., Song, R., Messing, J. & Chen, X. CARPEL FACTORY, a Dicer Homolog, and HEN1, a Novel Protein, Act in microRNA Metabolism in *Arabidopsis thaliana*. *Current Biology* **12**, 1484–1495 (2002).
144. Reinhart, B. J., Weinstein, E. G., Rhoades, M. W., Bartel, B. & Bartel, D. P. MicroRNAs in plants. *Genes Dev* **16**, 1616-1626 (2002).
145. Papp, I. et al. Evidence for nuclear processing of plant micro RNA and short interfering RNA precursors. *Plant Physiol* **132**, 1382-1390 (2003).
146. Vazquez, F., Gasciolli, V., Crete, P. & Vaucheret, H. The nuclear dsRNA binding protein HYL1 is required for microRNA accumulation and plant development, but not posttranscriptional transgene silencing. *Curr Biol* **14**, 346-351 (2004).
147. Park, M. Y., Wu, G., Gonzalez-Sulser, A., Vaucheret, H. & Poethig, R. S. Nuclear processing and export of microRNAs in Arabidopsis. *Proc Natl Acad Sci U S A* **102**, 3691-3696 (2005).
148. Yang, Z., Ebright, Y. W., Yu, B. & Chen, X. HEN1 recognizes 21-24 nt small RNA duplexes and deposits a methyl group onto the 2' OH of the 3' terminal nucleotide. *Nucleic Acids Res* **34**, 667-675 (2006).
149. Hutvagner, G. & Zamore, P. D. A microRNA in a Multiple-Turnover RNAi Enzyme Complex. *Science* **297**, 2056-2060 (2002).
150. Meister, G. et al. Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol Cell* **15**, 185-197 (2004).
151. Liu, J. et al. Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **305**, 1437-1441 (2004).
152. Mansfield, J. H. et al. MicroRNA-responsive 'sensor' transgenes uncover Hox-like and other developmentally regulated patterns of vertebrate microRNA expression. *Nat Genet* **36**, 1079-1083 (2004).
153. Pfeffer, S. et al. Identification of virus-encoded microRNAs. *Science* **304**, 734-736 (2004).
154. Yekta, S., Shih, I. H. & Bartel, D. P. MicroRNA-directed cleavage of HOXB8 mRNA. *Science* **304**, 594-596 (2004).
155. Davis, E. et al. RNAi-mediated allelic trans-interaction at the imprinted Rtl1/Peg11 locus. *Curr Biol* **15**, 743-749 (2005).
156. Sullivan, C. S., Grundhoff, A. T., Tevethia, S., Pipas, J. M. & Ganem, D. SV40-encoded microRNAs regulate viral gene expression and reduce susceptibility to cytotoxic T cells. *Nature* **435**, 682-686 (2005).
157. Rhoades, M. W. et al. Prediction of plant microRNA targets. *Cell* **110**, 513-520 (2002).
158. Brennecke, J., Stark, A., Russell, R. B. & Cohen, S. M. Principles of microRNA-target recognition. *PLoS Biol* **3**, e85 (2005).
159. Lewis, B. P., Shih, I. H., Jones-Rhoades, M. W., Bartel, D. P. & Burge, C. B. Prediction of mammalian microRNA targets. *Cell* **115**, 787-798 (2003).
160. Lewis, B. P., Burge, C. B. & Bartel, D. P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**, 15-20 (2005).

161. Stark, A., Brennecke, J., Bushati, N., Russell, R. B. & Cohen, S. M. Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3'UTR evolution. *Cell* **123**, 1133-1146 (2005).
162. Xie, X. et al. Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. *Nature* **434**, 338-345 (2005).
163. Lai, E. C. Predicting and validating microRNA targets. *Genome Biol* **5**, 115 (2004).
164. Ameres, S. L., Martinez, J. & Schroeder, R. Molecular basis for target RNA recognition and cleavage by human RISC. *Cell* **130**, 101-112 (2007).
165. Haley, B. & Zamore, P. D. Kinetic analysis of the RNAi enzyme complex. *Nat Struct Mol Biol* **11**, 599-606 (2004).
166. Baek, D. et al. The impact of microRNAs on protein output. *Nature* **455**, 64-71 (2008).
167. Selbach, M. et al. Widespread changes in protein synthesis induced by microRNAs. *Nature* **455**, 58-63 (2008).
168. Kong, Y. W. et al. The mechanism of micro-RNA-mediated translation repression is determined by the promoter of the target gene. *Proc Natl Acad Sci U S A* **105**, 8866-8871 (2008).
169. Brodersen, P. et al. Widespread translational inhibition by plant miRNAs and siRNAs. *Science* **320**, 1185-1190 (2008).
170. Stefani, G. & Slack, F. J. Small non-coding RNAs in animal development. *Nat Rev Mol Cell Biol* **9**, 219-230 (2008).
171. Yekta, S., Tabin, C. J. & Bartel, D. P. MicroRNAs in the Hox network: an apparent link to posterior prevalence. *Nat Rev Genet* **9**, 789-796 (2008).
172. Bartel, D. P. & Chen, C. Z. Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. *Nat Rev Genet* **5**, 396-400 (2004).
173. Bernstein, E. et al. Dicer is essential for mouse development. *Nat Genet* **35**, 215-217 (2003).
174. Wienholds, E. et al. MicroRNA expression in zebrafish embryonic development. *Science* **309**, 310-311 (2005).
175. Giraldez, A. J. et al. MicroRNAs regulate brain morphogenesis in zebrafish. *Science* **308**, 833-838 (2005).
176. Hatfield, S. D. et al. Stem cell division is regulated by the microRNA pathway. *Nature* **435**, 974-978 (2005).
177. Mudhasani, R. et al. Loss of miRNA biogenesis induces p19Arf-p53 signaling and senescence in primary cells. *J Cell Biol* **181**, 1055-1063 (2008).
178. Landgraf, P. et al. A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* **129**, 1401-1414 (2007).
179. Wienholds, E. & Plasterk, R. H. MicroRNA function in animal development. *FEBS Lett* **579**, 5911-5922 (2005).
180. Vagin, V. V. et al. A distinct small RNA pathway silences selfish genetic elements in the germline. *Science* **313**, 320-324 (2006).
181. Saito, K. et al. Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the Drosophila genome. *Genes Dev* **20**, 2214-2222 (2006).

182. Saito, K. et al. Pimet, the Drosophila homolog of HEN1, mediates 2'-O-methylation of Piwi- interacting RNAs at their 3' ends. *Genes Dev* **21**, 1603-1608 (2007).
183. Kirino, Y. & Mourelatos, Z. Mouse Piwi-interacting RNAs are 2'-O-methylated at their 3' termini. *Nat Struct Mol Biol* **14**, 347-348 (2007).
184. Ohara, T. et al. The 3' termini of mouse Piwi-interacting RNAs are 2'-O-methylated. *Nat Struct Mol Biol* **14**, 349-350 (2007).
185. Kato, Y. et al. Role of the Dnmt3 family in de novo methylation of imprinted and repetitive sequences during male germ cell development in the mouse. *Hum Mol Genet* **16**, 2272-2280 (2007).
186. Lees-Murdock, D. J., De Felici, M. & Walsh, C. P. Methylation dynamics of repetitive DNA elements in the mouse germ cell lineage. *Genomics* **82**, 230-237 (2003).
187. Aravin, A. A. et al. A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. *Mol Cell* **31**, 785-799 (2008).
188. Ruby, J. G. et al. Evolution, biogenesis, expression, and target predictions of a substantially expanded set of Drosophila microRNAs. *Genome Res* **17**, 1850-1864 (2007).
189. Nishida, K. M. et al. Gene silencing mechanisms mediated by Aubergine piRNA complexes in Drosophila male gonad. *RNA* **13**, 1911-1922 (2007).
190. Prud'homme, N., Gans, M., Masson, M., Terzian, C. & Bucheton, A. Flamenco, a gene controlling the gypsy retrovirus of Drosophila melanogaster. *Genetics* **139**, 697-711 (1995).
191. Pelisson, A. et al. Gypsy transposition correlates with the production of a retroviral envelope-like protein under the tissue-specific control of the Drosophila flamenco gene. *EMBO J* **13**, 4401-4411 (1994).
192. Desset, S., Buchon, N., Meignin, C., Coiffet, M. & Vaury, C. In Drosophila melanogaster the COM Locus Directs the Somatic Silencing of Two Retrotransposons through both Piwi-Dependent and -Independent Pathways. *PLoS ONE* **3**, e1526 (2008).
193. Desset, S., Meignin, C., Dastugue, B. & Vaury, C. COM, a heterochromatic locus governing the control of independent endogenous retroviruses from Drosophila melanogaster. *Genetics* **164**, 501-509 (2003).
194. Mevel-Ninio, M. T., Pelisson, A., Kinder, J., Campos, A. R. & Bucheton, A. The flamenco locus controls the gypsy and ZAM retroviruses and is required for Drosophila oogenesis. *Genetics* (2007).
195. Gunawardane, L. S. et al. A Slicer-Mediated Mechanism for Repeat-Associated siRNA 5' End Formation in Drosophila. *Science* **315**, 1587-1590 (2007).
196. Cox, D. N. et al. A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. *Genes & Development* **12**, 3715–3727 (1998).
197. Cox, D. N., Chao, A. & Lin, H. piwi encodes a nucleoplasmic factor whose activity modulates the number and division rate of germline stem cells. *Development* **127**, 503-514 (2000).
198. Aravin, A. A. et al. Dissection of a natural RNA silencing process in the Drosophila melanogaster germ line. *Mol Cell Biol* **24**, 6742-6750 (2004).

199. Schupbach, T. & Wieschaus, E. Female sterile mutations on the second chromosome of Drosophila melanogaster. II. Mutations blocking oogenesis or altering egg morphology. *Genetics* **129**, 1119-1136 (1991).
200. Klattenhoff, C. et al. Drosophila rasiRNA pathway mutations disrupt embryonic axis specification through activation of an ATR/Chk2 DNA damage response. *Dev Cell* **12**, 45-55 (2007).
201. Pal-Bhadra, M. et al. Heterochromatic silencing and HP1 localization in Drosophila are dependent on the RNAi machinery. *Science* **303**, 669-672 (2004).
202. Tomari, Y., Du, T. & Zamore, P. D. Sorting of Drosophila small silencing RNAs. *Cell* **130**, 299–308 (2007).
203. Forstemann, K., Horwich, M. D., Wee, L.-M., Tomari, Y. & Zamore, P. D. Drosophila microRNAs are sorted into functionally distinct Argonaute protein complexes after their production by Dicer-1. *Cell* **130**, 287–297 (2007).
204. Mi, S. et al. Sorting of small RNAs into Arabidopsis argonaute complexes is directed by the 5' terminal nucleotide. *Cell* **133**, 116-127 (2008).
205. Vaucheret, H. Plant ARGONAUTES. *Trends Plant Sci* **13**, 350-358 (2008).
206. Tokumaru, S., Suzuki, M., Yamada, H., Nagino, M. & Takahashi, T. let-7 regulates Dicer expression and constitutes a negative feedback loop. *Carcinogenesis* (2008).
207. Forman, J. J., Legesse-Miller, A. & Coller, H. A. A search for conserved sequences in coding regions reveals that the let-7 microRNA targets Dicer within its coding sequence. *Proc Natl Acad Sci U S A* **105**, 14879-14884 (2008).
208. Xie, Z., Kasschau, K. D. & Carrington, J. C. Negative Feedback Regulation of Dicer-Like1 in Arabidopsis by microRNA-Guided mRNA Degradation. *Curr Biol* **13**, 784-789 (2003).
209. Vaucheret, H., Vazquez, F., Crete, P. & Bartel, D. P. The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. *Genes Dev* **18**, 1187-1197 (2004).
210. Vaucheret, H., Mallory, A. C. & Bartel, D. P. AGO1 homeostasis entails coexpression of MIR168 and AGO1 and preferential stabilization of miR168 by AGO1. *Mol Cell* **22**, 129-136 (2006).
211. Sijen, T. et al. On the Role of RNA Amplification in dsRNA-Triggered Gene Silencing. *Cell* **107**, 465-476 (2001).
212. Voinnet, O., Vain, P., Angell, S. & Baulcombe, D. C. Systemic spread of sequence-specific transgene RNA degradation in plants is initiated by localized introduction of ectopic promoterless DNA. *Cell* **95**, 177-187 (1998).
213. Vaistij, F. E., Jones, L. & Baulcombe, D. C. Spreading of RNA targeting and DNA methylation in RNA silencing requires transcription of the target gene and a putative RNA-dependent RNA polymerase. *Plant Cell* **14**, 857-867 (2002).
214. Fusaro, A. F. et al. RNA interference-inducing hairpin RNAs in plants act through the viral defence pathway. *EMBO Rep* **7**, 1168-1175 (2006).
215. Moissiard, G., Parizotto, E. A., Himber, C. & Voinnet, O. Transitivity in Arabidopsis can be primed, requires the redundant action of the antiviral Dicer-like 4 and Dicer-like 2, and is compromised by viral-encoded suppressor proteins. *RNA* **13**, 1268-1278 (2007).

216. Sijen, T., Steiner, F. A., Thijssen, K. L. & Plasterk, R. H. Secondary siRNAs result from unprimed RNA synthesis and form a distinct class. *Science* **315**, 244-247 (2007).
217. Pak, J. & Fire, A. Distinct populations of primary and secondary effectors during RNAi in C. elegans. *Science* **315**, 241-244 (2007).
218. Smardon, A. et al. EGO-1 is related to RNA-directed RNA polymerase and functions in germ-line development and RNA interference in *C. elegans*. *Current Biology* **10**, 169-178 (2000).
219. Aoki, K., Moriguchi, H., Yoshioka, T., Okawa, K. & Tabara, H. In vitro analyses of the production and activity of secondary small interfering RNAs in C. elegans. *EMBO J* (2007).
220. Makeyev, E. V. & Bamford, D. H. Cellular RNA-dependent RNA polymerase involved in posttranscriptional gene silencing has two distinct activity modes. *Mol Cell* **10**, 1417-1427 (2002).
221. Tijsterman, M., Ketting, R. F., Okihara, K. L., Sijen, T. & Plasterk, R. H. RNA helicase MUT-14-dependent gene silencing triggered in C. elegans by short antisense RNAs. *Science* **295**, 694-697 (2002).
222. Caplen, N. J. et al. Rescue of polyglutamine-mediated cytotoxicity by double-stranded RNA-mediated RNA interference. *Hum Mol Genet* **11**, 175-184 (2002).
223. Xia, H., Mao, Q., Paulson, H. L. & Davidson, B. L. siRNA-mediated gene silencing in vitro and in vivo. *Nat Biotechnol* **20**, 1006-1010 (2002).
224. Ding, H. et al. Selective silencing by RNAi of a dominant allele that causes amyotrophic lateral sclerosis. *Aging Cell* **2**, 209-217 (2003).
225. Miller, V. M. et al. Allele-specific silencing of dominant disease genes. *Proc Natl Acad Sci U S A* **100**, 7195-7200 (2003).
226. Martinez, L. A. et al. Synthetic small inhibiting RNAs: efficient tools to inactivate oncogenic mutations and restore p53 pathways. *Proc Natl Acad Sci U S A* **99**, 14849-14854 (2002).
227. Schwarz, D. S., Hutvagner, G., Haley, B. & Zamore, P. D. Evidence that siRNAs function as guides, not primers, in the Drosophila and human RNAi pathways. *Mol Cell* **10**, 537-548 (2002).
228. Nykanen, A., Haley, B. & Zamore, P. D. ATP requirements and small interfering RNA structure in the RNA interference pathway. *Cell* **107**, 309-321 (2001).
229. Martinez, J., Patkaniowska, A., H, H. U., Lührmann, R. & Tuschl, T. Single stranded antisense siRNA guide target RNA cleavage in RNAi. *Cell* **110**, 563-574 (2002).
230. Roignant, J. Y. et al. Absence of transitive and systemic pathways allows cell-specific and isoform-specific RNAi in Drosophila. *RNA* **9**, 299-308 (2003).
231. Rusk, N. & Kiermer, V. Primer: Sequencing--the next generation. *Nat Methods* **5**, 15 (2008).
232. Johnston, R. J. & Hobert, O. A microRNA controlling left/right neuronal asymmetry in Caenorhabditis elegans. *Nature* **426**, 845-849 (2003).
233. Yin, H. & Lin, H. An epigenetic activation role of Piwi and a Piwi-associated piRNA in Drosophila melanogaster. *Nature* **450**, 304-308 (2007).

234. Sunkar, R., Girke, T. & Zhu, J. K. Identification and characterization of endogenous small interfering RNAs from rice. *Nucleic Acids Res* **33**, 4443-4454 (2005).
235. Ishizuka, A., Siomi, M. C. & Siomi, H. A Drosophila fragile X protein interacts with components of RNAi and ribosomal proteins. *Genes Dev* **16**, 2497-2508 (2002).
236. Csink, A. K., Linsk, R. & Birchler, J. A. The Lighten up (Lip) gene of Drosophila melanogaster, a modifier of retroelement expression, position effect variegation and white locus insertion alleles. *Genetics* **138**, 153-163 (1994).
237. Rehwinkel, J. et al. Genome-wide analysis of mRNAs regulated by Drosha and Argonaute proteins in Drosophila melanogaster. *Mol Cell Biol* **26**, 2965-2975 (2006).
238. Haynes, K. A., Caudy, A. A., Collins, L. & Elgin, S. C. Element 1360 and RNAi components contribute to HP1-dependent silencing of a pericentric reporter. *Curr Biol* **16**, 2222-2227 (2006).
239. Pal-Bhadra, M., Bhadra, U. & Birchler, J. A. RNAi Related Mechanisms Affect Both Transcriptional and Posttranscriptional Transgene Silencing in Drosophila. *Mol Cell* **9**, 315-327 (2002).
240. Tchurikov, N. A. & Kretova, O. V. Suffix-specific RNAi leads to silencing of F element in Drosophila melanogaster. *PLoS ONE* **2**, e476 (2007).
241. Okamura, K., Ishizuka, A., Siomi, H. & Siomi, M. C. Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. *Genes Dev* **18**, 1655-1666 (2004).
242. http://flyrnai.org/cgi-bin/RNAi_FAQ_lines.pl.
243. Lee, Y. S. & Carthew, R. W. Making a better RNAi vector for Drosophila: use of intron spacers. *Methods* **30**, 322-329 (2003).
244. Seitz, H., Ghildiyal, M. & Zamore, P. D. Argonaute Loading Improves the 5' Precision of Both MicroRNAs and Their miRNA(*) Strands in Flies. *Curr Biol* **18**, 147-151 (2008).
245. Takaesu, N. T. et al. dSno facilitates baboon signaling in the Drosophila brain by switching the affinity of Medea away from Mad and toward dSmad2. *Genetics* **174**, 1299-1313 (2006).
246. Potter, S. S., Brorein, W. J. J., Dunsmuir, P. & Rubin, G. M. Transposition of elements of the 412, copia and 297 dispersed repeated gene families in Drosophila. *Cell* **17**, 415-427 (1979).
247. Maisonhaute, C., Ogereau, D., Hua-Van, A. & Capy, P. Amplification of the 1731 LTR retrotransposon in Drosophila melanogaster cultured cells: origin of neocopies and impact on the genome. *Gene* **393**, 116-126 (2007).
248. Zhou, R. et al. Comparative analysis of argonaute-dependent small RNA pathways in Drosophila. *Mol Cell* **32**, 592-599 (2008).
249. Lucchetta, E. M., Carthew, R. W. & Ismagilov, R. F. The endo-siRNA pathway is essential for robust development of the Drosophila embryo. *PLoS One* **4**, e7576 (2009).
250. Deshpande, G., Calhoun, G. & Schedl, P. Drosophila argonaute-2 is required early in embryogenesis for the assembly of centric/centromeric heterochromatin, nuclear

division, nuclear migration, and germ-cell formation. *Genes Dev* **19**, 1680-1685 (2005).
251. Li, X., Cassidy, J. J., Reinke, C. A., Fischboeck, S. & Carthew, R. W. A microRNA imparts robustness against environmental fluctuation during development. *Cell* **137**, 273-282 (2009).
252. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402-48. (2001).
253. Peirson, S. N., Butler, J. N. & Foster, R. G. Experimental validation of novel and conventional approaches to quantitative real-time PCR data analysis. *Nucleic Acids Res* **31**, e73 (2003).
254. Ghildiyal, M. & Zamore, P. D. Small silencing RNAs: an expanding universe. *Nat Rev Genet* **10**, 94-108 (2009).
255. Han, J. et al. Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell* **125**, 887-901 (2006).
256. Zeng, Y., Yi, R. & Cullen, B. R. Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha. *EMBO J* **24**, 38-48 (2005).
257. Melo, S. A. et al. A TARBP2 mutation in human cancer impairs microRNA processing and DICER1 function. *Nat Genet* **41**, 365-370 (2009).
258. Maniataki, E. & Mourelatos, Z. A human, ATP-independent, RISC assembly machine fueled by pre-miRNA. *Genes Dev* **19**, 2979-2990 (2005).
259. Grimson, A. et al. MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol Cell* **27**, 91-105 (2007).
260. Grun, D., Wang, Y. L., Langenberger, D., Gunsalus, K. C. & Rajewsky, N. microRNA target predictions across seven Drosophila species and comparison to mammalian targets. *PLoS Comput Biol* **1**, e13 (2005).
261. Krek, A. et al. Combinatorial microRNA target predictions. *Nat Genet* **37**, 495-500 (2005).
262. Rajewsky, N. & Socci, N. D. Computational identification of microRNA targets. *Dev Biol* **267**, 529-535 (2004).
263. Wang, Y. et al. Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. *Nature* **456**, 921-926 (2008).
264. Parker, J. S., Parizotto, E. A., Wang, M., Roe, S. M. & Barford, D. Enhancement of the seed-target recognition step in RNA silencing by a PIWI/MID domain protein. *Mol Cell* **33**, 204-214 (2009).
265. Jackson, A. L. et al. Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity. *RNA* **12**, 1179-1187 (2006).
266. Lim, L. P. et al. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* **433**, 769-773 (2005).
267. Jackson, A. L. et al. Expression profiling reveals off-target gene regulation by RNAi. *Nat Biotechnol* **21**, 635-637 (2003).
268. Rand, T. A., Ginalski, K., Grishin, N. V. & Wang, X. Biochemical identification of Argonaute 2 as the sole protein required for RNA-induced silencing complex activity. *Proc Natl Acad Sci U S A* **101**, 14385-14389 (2004).

269. Kawamata, T., Seitz, H. & Tomari, Y. Structural determinants of miRNAs for RISC loading and slicer-independent unwinding. *Nat Struct Mol Biol* **16**, 953-960 (2009).
270. Tomari, Y. et al. RISC assembly defects in the Drosophila RNAi mutant armitage. *Cell* **116**, 831-841 (2004).
271. Okamura, K. et al. The regulatory activity of microRNA* species has substantial influence on microRNA and 3' UTR evolution. *Nat Struct Mol Biol* **15**, 354-363 (2008).
272. Ro, S., Park, C., Young, D., Sanders, K. M. & Yan, W. Tissue-dependent paired expression of miRNAs. *Nucleic Acids Res* **35**, 5944-5953 (2007).
273. Lin, E. A., Kong, L., Bai, X. H., Luan, Y. & Liu, C. J. miR-199a, a bone morphogenic protein 2-responsive MicroRNA, regulates chondrogenesis via direct targeting to Smad1. *J Biol Chem* **284**, 11326-11335 (2009).
274. Ma, J. B. et al. Structural basis for 5'-end-specific recognition of guide RNA by the A. fulgidus Piwi protein. *Nature* **434**, 666-670 (2005).
275. Parker, J. S., Roe, S. M. & Barford, D. Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex. *Nature* **434**, 663-666 (2005).
276. Stark, A. et al. A single Hox locus in Drosophila produces functional microRNAs from opposite DNA strands. *Genes Dev* **22**, 8-13 (2008).
277. Tyler, D. M. et al. Functionally distinct regulatory RNAs generated by bidirectional transcription and processing of microRNA loci. *Genes Dev* **22**, 26-36 (2008).
278. Iwasaki, S., Kawamata, T. & Tomari, Y. Drosophila argonaute1 and argonaute2 employ distinct mechanisms for translational repression. *Mol Cell* **34**, 58-67 (2009).
279. Tuschl, T., Zamore, P. D., Lehmann, R., Bartel, D. P. & Sharp, P. A. Targeted mRNA degradation by double-stranded RNA in vitro. *Genes Dev* **13**, 3191-3197 (1999).
280. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**, R25 (2009).
281. Lai, E. C. Micro RNAs are complementary to 3' UTR sequence motifs that mediate negative post-transcriptional regulation. *Nat Genet* **30**, 363-364 (2002).
282. Farh, K. K. et al. The widespread impact of mammalian MicroRNAs on mRNA repression and evolution. *Science* **310**, 1817-1821 (2005).
283. Du, T. & Zamore, P. D. microPrimer: the biogenesis and function of microRNA. *Development* **132**, 4645-4652 (2005).
284. Kim, V. N. MicroRNA biogenesis: coordinated cropping and dicing. *Nat Rev Mol Cell Biol* **6**, 376-385 (2005).
285. Wu, H. et al. miRNA Profiling of Naive, Effector and Memory CD8 T Cells. *PLoS ONE* **2**, e1020 (2007).
286. Azuma-Mukai, A. et al. Characterization of endogenous human Argonautes and their miRNA partners in RNA silencing. *Proc Natl Acad Sci U S A* **105**, 7964-7969 (2008).
287. Ho, C. K., Wang, L. K., Lima, C. D. & Shuman, S. Structure and mechanism of RNA ligase. *Structure* **12**, 327-339 (2004).

288. Britten, R. J. & Davidson, E. H. Gene regulation for higher cells: a theory. *Science* **165**, 349-357 (1969).
289. Specchia, V. et al. aubergine Gene Overexpression in Somatic Tissues of auberginesting Mutants Interferes With the RNAi Pathway of a yellow Hairpin dsRNA in Drosophila melanogaster. *Genetics* **178**, 1271-1282 (2008).
290. Ghildiyal, M., Xu, J., Seitz, H., Weng, Z. & Zamore, P. D. Sorting of Drosophila small silencing RNAs partitions microRNA* strands into the RNA interference pathway. *RNA* **16**, 43-56 (2010).
291. Bartel, D. P. MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215-233 (2009).
292. Grimson, A. et al. Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. *Nature* **455**, 1193-1197 (2008).
293. Brennecke, J. et al. An epigenetic role for maternally inherited piRNAs in transposon silencing. *Science* **322**, 1387-1392 (2008).
294. Kennerdell, J. R. & Carthew, R. W. Use of dsRNA-mediated genetic interference to demonstrate that frizzled and frizzled 2 act in the wingless pathway. *Cell* **95**, 1017-1026 (1998).
295. Ngo, H., Tschudi, C., Gull, K. & Ullu, E. Double-stranded RNA induces mRNA degradation in *Trypanosoma brucei*. *Proc Natl Acad Sci U S A* **95**, 14687-14692 (1998).
296. Waterhouse, P. M., Graham, M. W. & Wang, M. B. Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. *Proc Natl Acad Sci U S A* **95**, 13959-13964 (1998).
297. Hoa, N. T., Keene, K. M., Olson, K. E. & Zheng, L. Characterization of RNA interference in an Anopheles gambiae cell line. *Insect Biochem Mol Biol* **33**, 949-957 (2003).
298. Tomari, Y. & Zamore, P. D. Perspective: machines for RNAi. *Genes Dev* **19**, 517-529 (2005).
299. Pham, J. W., Pellino, J. L., Lee, Y. S., Carthew, R. W. & Sontheimer, E. J. A Dicer-2-dependent 80s complex cleaves targeted mRNAs during RNAi in Drosophila. *Cell* **117**, 83-94 (2004).
300. Pham, J. W. & Sontheimer, E. J. Molecular requirements for RNA-induced silencing complex assembly in the Drosophila RNA interference pathway. *J Biol Chem* **280**, 39278-39283 (2005).
301. Preall, J. B., He, Z., Gorra, J. M. & Sontheimer, E. J. Short interfering RNA strand selection is independent of dsRNA processing polarity during RNAi in Drosophila. *Curr Biol* **16**, 530-535 (2006).
302. Schwarz, D. S., Tomari, Y. & Zamore, P. D. The RNA-Induced Silencing Complex Is a Mg(2+)-Dependent Endonuclease. *Curr Biol* **14**, 787-791 (2004).
303. Martinez, J. & Tuschl, T. RISC is a 5' phosphomonoester-producing RNA endonuclease. *Genes Dev* **18**, 975-980 (2004).
304. Song, J. J., Smith, S. K., Hannon, G. J. & Joshua-Tor, L. Crystal structure of Argonaute and its implications for RISC slicer activity. *Science* **305**, 1434-1437 (2004).

305. Rivas, F. V. et al. Purified Argonaute2 and an siRNA form recombinant human RISC. *Nat Struct Mol Biol* **12**, 340-349 (2005).
306. Pasquinelli, A. E. et al. Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature* **408**, 86-89 (2000).
307. Reinhart, B. J. et al. The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature* **403**, 901-96. (2000).
308. Bartel, D. P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281-297 (2004).
309. Yeom, K. H., Lee, Y., Han, J., Suh, M. R. & Kim, V. N. Characterization of DGCR8/Pasha, the essential cofactor for Drosha in primary miRNA processing. *Nucleic Acids Res* **34**, 4622-4629 (2006).
310. Borchert, G. M., Lanier, W. & Davidson, B. L. RNA polymerase III transcribes human microRNAs. *Nat Struct Mol Biol* **13**, 1097-1101 (2006).
311. Meyer, W. J. et al. Overlapping functions of argonaute proteins in patterning and morphogenesis of Drosophila embryos. *PLoS Genet* **2**, e134 (2006).
312. Thibault, S. T. et al. A complementary transposon tool kit for Drosophila melanogaster using P and piggyBac. *Nat Genet* **36**, 283-287 (2004).
313. Kennerdell, J. R. & Carthew, R. W. Heritable gene silencing in Drosophila using double-stranded RNA. *Nat Biotechnol* **18**, 896-898 (2000).
314. Hammond, S. M., Boettcher, S., Caudy, A. A., Kobayashi, R. & Hannon, G. J. Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science* **293**, 1146-1150 (2001).
315. Marques, J. T. et al. Loqs and R2D2 act sequentially in the siRNA pathway in Drosophila. *Nat Struct Mol Biol* (2009).
316. Czech, B. et al. Hierarchical rules for Argonaute loading in Drosophila. *Mol Cell* **36**, 445-456 (2009).
317. Okamura, K., Liu, N. & Lai, E. C. Distinct mechanisms for microRNA strand selection by Drosophila Argonautes. *Mol Cell* **36**, 431-444 (2009).
318. Chintapalli, V. R., Wang, J. & Dow, J. A. Using FlyAtlas to identify better Drosophila melanogaster models of human disease. *Nat Genet* **39**, 715-720 (2007).
319. Brodersen, P. & Voinnet, O. Revisiting the principles of microRNA target recognition and mode of action. *Nat Rev Mol Cell Biol* **10**, 141-148 (2009).
320. Wickens, M. & Kwak, J. E. Molecular biology. A tail tale for U. *Science* **319**, 1344-1345 (2008).
321. Rissland, O. S. & Norbury, C. J. Decapping is preceded by 3' uridylation in a novel pathway of bulk mRNA turnover. *Nat Struct Mol Biol* **16**, 616-623 (2009).
322. Rissland, O. S., Mikulasova, A. & Norbury, C. J. Efficient RNA polyuridylation by noncanonical poly(A) polymerases. *Mol Cell Biol* **27**, 3612-3624 (2007).
323. Houseley, J. & Tollervey, D. The many pathways of RNA degradation. *Cell* **136**, 763-776 (2009).
324. Shen, B. & Goodman, H. M. Uridine addition after microRNA-directed cleavage. *Science* **306**, 997 (2004).

325. Mullen, T. E. & Marzluff, W. F. Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5' to 3' and 3' to 5'. *Genes Dev* **22**, 50-65 (2008).
326. van Wolfswinkel, J. C. et al. CDE-1 affects chromosome segregation through uridylation of CSR-1-bound siRNAs. *Cell* **139**, 135-148 (2009).
327. Song, J. J. et al. The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. *Nat Struct Biol* **10**, 1026-1032 (2003).
328. Yan, K. S. et al. Structure and conserved RNA binding of the PAZ domain. *Nature* **426**, 468-474 (2003).
329. Lingel, A., Simon, B., Izaurralde, E. & Sattler, M. Structure and nucleic-acid binding of the Drosophila Argonaute 2 PAZ domain. *Nature* **426**, 465-469 (2003).
330. Ma, J. B., Ye, K. & Patel, D. J. Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature* **429**, 318-322 (2004).
331. Lingel, A., Simon, B., Izaurralde, E. & Sattler, M. Nucleic acid 3'-end recognition by the Argonaute2 PAZ domain. *Nat Struct Mol Biol* **11**, 576-577 (2004).
332. Parker, J. S., Roe, S. M. & Barford, D. Crystal structure of a PIWI protein suggests mechanisms for siRNA recognition and slicer activity. *EMBO J* **23**, 4727-4737 (2004).
333. Wang, Y. et al. Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. *Nature* **461**, 754-761 (2009).