Detailed Online Methods

Fly stocks

A 3xFLAG-HA tag was inserted at the N-terminus of AGO2-RB into BAC RP98-21A13 via bacterial Red/ET recombination (Gene Bridges GmbH). A 13.9kb AvrII/XhoI fragment of the modified BAC encompassing the AGO2 locus including parts of flanking genes (chr3L coordinates: 15,544,405-15,558,309) was cloned into pCasper4 (XbaI/XhoI). Transgenic flies were generated at Bestgene Inc. Expression of tagged AGO2 in embryos, ovaries and whole flies was verified in multiple lines by western blotting using a monoclonal anti-HA-Peroxidase antibody (1:500; Cat.: 12013819001; Roche). *dcr-2*^{L811Fsx} flies were a kind gift of Richard Carthew [6], *AGO2*^{414} flies were a kind gift of Haruhiko Siomi [30], *loqs*^{f00791} flies were a kind gift of Phil Zamore[1] and *r2d2*^{1} flies were a kind gift of Dean Smith[14]. As wild-type fly stocks, #2057 From Bloomington (Celera sequencing strain) and OregonR flies were used.

Small RNA libraries

Twenty-six 10 cm plates with 50-70% confluent Schneider cells were transfected with pCasper_FLAG-HA-AGO2 using Calcium Phosphate, harvested after 36h and lysed in buffer A (20 mM HEPES pH7.0, 150 mM NaCl, 2.5 mM $MgCl_2$, 0.3% Triton, 30% glycerol) supplemented with 1 mM PMSF and protease inhibitors (Complete; Roche). Cleared extract was split and incubated with rabbit polyclonal AGO1 antibody (1:20; lot: 113754; Abcam) or mouse anti-FLAG® M2-Agarose (1:25; SIGMA) for 4 hours at 4°C. AGO1 antibodies were isolated by adding proteinG beads (1:10; Roche) for 1 hour. Beads were washed 6 times each 10 minutes in buffer B (30 mM HEPES pH7.4, 800 mM NaCl, 2 mM $MgCl_2$; 0.1% NP-40) which contained equal supplements as buffer A. The immunoprecipitation was analyzed by western blotting using anti-AGO1 (1:2000; Abcam) and anti-HA-Peroxidase (1:500; Cat.: 12013819001; Roche). AGO1- and AGO2-associated RNAs were isolated with Phenol/Chloroform and ethanol precipitated. For the ovarian AGO2 IP library ~500 mg ovaries from transgenic FLAG-HA-AGO2 flies were dissected and lysed mechanically in buffer A. AGO2 complexes and associated RNAs were purified as above.

AGO1 and AGO2 bound small RNAs as well as small RNAs from total RNA were cloned as described [12] (detailed protocol available upon request). The following small RNA libraries from total RNA were prepared for this study:

18-28 nt from ovaries of the Celera sequenced strain (Bloom. # 2057)
18-28 nt from ovaries of *dcr-2*^{L811Fsx} homozygous flies
18-28 nt from *loqs*^{f00791} homozygous flies
18-24 nt from testis of OregonR flies.

Libraries were sequenced in house using the Illumina platform. Published libraries used in this study were a 16-26 nt S2 cell total RNA library [15] and Piwi/Aub/AGO3 IP libraries from ovaries [12].

Bioinformatic analysis of small RNA libraries

Small RNA sequences were matched to the *Drosophila* release 5 genome and genomes of *Drosophila* C virus, Flock house virus and Cricket paralysis virus. Only reads matching the fly genome 100% and viral genomes with up to 3 mismatches were used for further analysis. For annotations we used Flybase for protein-coding genes, UCSC for non-coding RNAs and transposons/repeats and the most recent miRNA catalog. [15, 33].

siRNA clusters were extracted by mapping all 20-22nt long RNAs from the AGO2-IP libraries to the genome (only uniquely mapping RNAs were used) and retaining 200 nt windows, which contained at least 3 distinct small RNAs. Windows separated by max. 200 nt were fused and those with more than 40 unique reads were sorted after the density of siRNAs per bp.

For the transposon analysis, 20-22 nt AGO2 bound RNAs from ovaries and S2 cells were mapped onto the Repbase collection of transposons [34] with up to 3 mismatches to construct heatmaps indicating cloning frequency and strand bias of siRNAs. For the latter analysis only siRNAs unambiguously mapped to one strand were considered.

Cleavage site mapping for endo-siRNA targets

Wild-type testes were dissected on ice into 1x PBS. Total RNA was isolated using Trizol (Invitrogen) according to the manufacturer's protocol. 5 µg total RNA were used as starting material. Ligation of an RNA adaptor, reverse transcription using the GeneRacer oligo (dT) primer and 5' RACE-PCR were performed according to the manufacturer's instructions (GeneRacer kit, Invitrogen). 5' RACE-PCR was carried out using the GeneRacer 5' primer (5′-CGACTGGAGCACGAGGACACTGA-3′) and a *mus308* gene-specific reverse primer (5'-TGCTTTGCAGAGTCGAAGCTGATTG-3'), and followed by one round of nested PCR using the GeneRacer 5' nested primer (5′-GGACACTGACATGGACTGAAGGAGTA-3′) and a nested primer specific to *mus308* (5'-CCGCTAGCTCTACCAAACTGGTGAT-3'). PCR products were gel purified and cloned into pCR®4Blunt-TOPO® (Invitrogen). 22 clones were sequenced with T7 (5'-GTAATACGACTCACTATAGGGC-3') and T3 (5'-AATTAACCCTCACTAAAGGG-3') primer, and subjected to further analysis.

dsRNA treatment of Schneider cells

~ $3x10^6$ S2-NP cells were soaked in 1.5mL serum-free Schneider's medium containing 10 ug dsRNAs in 6-well plates, and 3mL serum-containing medium was added 45 minutes later. After 4 days of initial dsRNA treatment, cells were treated with a second round of dsRNAs using the same procedure and were harvested another 4 days later. Total RNA was extracted with Trizol (Invitrogen). Sequences of the primers for generating dsRNAs are listed below.

siRNA reporter constructs

A SalI/BglII fragment from pGL3-Basic (Promega) was cloned to pRmHa-3 using SalI/BamHI (pMT-Firefly-long). The coding region of the *Renilla luciferase* gene was amplified by PCR and cloned into pRmHa-3 using BamHI/EcoRI sites (pMT-Renilla). A pair of oligos containing two perfect binding sites for si1_1,

si1_2 or si2 were annealed and cloned into pMT-Renilla (BamHI/SalI) to generate sensor constructs (Fig. S13).

Transfection was performed in a 384-well plate format. For each well, ~100 ng plasmid DNA (5 ng pMT-Renilla, 20 ng pMT-Renilla-si1_1, 50 ng pMT-si1_2 or 100 ng pMT-si2, 5 ng pMT-Firefly-long, and corresponding amounts of pRmHa-3 serving as carrier DNA) and ~80 ng dsRNA were mixed with 0.8 μl Enhancer in 15 μl EC (Qiagen) and incubated at room temperature for 5 minutes. Then 0.35 μl of Effectene reagent was added and the mixture was immediately dispensed into each well containing dsRNA. After incubation at room temperature for 10 minutes, 40 μl S2-NP cells (10^6 cells/mL) were dispensed into the well. Cells were induced with 200 micromolar $CuSO_4$ 132 hours post transfection and luciferase assays were performed 36 hours later using DualGlo reagents (Promega). For each well, the reporter activity was calculated as the ratio of *Renilla* luciferase to firefly luciferase. Each data point was normalized against the data points where dsRNA against LacZ was transfected. Presented are average results with standard deviation (n=3).

Northern blotting

Total RNA was isolated using Trizol (Invitrogen). 30 ug RNA were separated on a 15% denaturing polyacrylamide gel and transferred onto a Hybond-N+ membrane (Amersham Biosciences) in 1x TBE. The RNA was UV crosslinked to the membrane and pre-hybridized in ULTRAhyb™ buffer (Ambion) for 1 hour. DNA probes complementary to the indicated siRNAs, bantam, and 2S RNA were 5' radiolabeled and added to the hybridization buffer (hybridization over night at 37°C). Membranes were washed 4-6 times in 1x SSC with 0.1% SDS at 37°C and exposed to PhosphoImager screens. Probes were stripped by boiling the membrane twice in 0.2x SSC containing 0.1% SDS in a microwave.

Quantitative real-time PCR

Ovaries and testis from homozygous or heterozygous flies were dissected on ice into 1x PBS. Total RNA of dissected tissues or S2 cells was extracted using Trizol (Invitrogen). RNA was treated with DNase I Amplification Grade (Invitrogen) according to the manufacturer's instructions. cDNA was prepared by reverse transcription using SuperScript™ III Reverse Transcriptase (Invitrogen) and random hexamer primer. qRT-PCR was carried out using SYBR® GREEN PCR Master Mix (Applied Biosystems) and a Chromo4 Real-Time PCR Detector (BioRad). C_T values were calculated within the log-linear phase of the amplification curve using the Opticon Monitor 3.1.32 software (BioRad). Quantification was normalized to the endogenous ribosomal protein *rp49* and relative expression levels were calculated using the following equation: $A = 1.8^{[Ct(ref)-Ct(ref\text{-}control)]-[Ct(sample)-Ct(sample\text{-}control)]}$. Transposon analysis was carried out with four biological replicates (individually shown; error bars indicate technical replicates), siRNA target analysis was carried out with three biological replicates. Oligonucleotide primers used in this study are listed below.

DNA Oligonucleotides

dsRNA generating PCR primers:

T7-Dicer1-F-14	TAATACGACTCACTATAGGGTGCGACAACAATCTGC
T7-Dicer1-R-565	TAATACGACTCACTATAGGGTCAGTTGCTGCAGCTCAC
T7-Dicer2-F-5	TAATACGACTCACTATAGGGAAGATGTGGAAATCAAGCC
T7-Dicer2-R-555	TAATACGACTCACTATAGGGCCACGTTCGTAATTTC
T7-Drosha-F-3356	TAATACGACTCACTATAGGGTGAATCAGGACTGGAACG
T7-Drosha-R-3910	TAATACGACTCACTATAGGGAGCCATCGCTATCACTGC
T7-Exportin5-F-55	TAATACGACTCACTATAGGGATCTAGTCATGAACCCG
T7-Exportin5-R-623	TAATACGACTCACTATAGGGAACGCAGTCACATGCTGC
T7-Ago1-F1225	TAATACGACTCACTATAGGGAACGGACAGACCGTAGAG
T7-Ago1-R1858	TAATACGACTCACTATAGGGTGGCGTACTTACAGAAGC
T7-Ago2-F2211	TAATACGACTCACTATAGGGAGCCACATCGACGAACG
T7-Ago2-R2855	TAATACGACTCACTATAGGGCGAGGATCATCCTTGATC
T7-R2D2-PZ-F	TAATACGACTCACTATAGGGCATACACGGCTTGATGAAGGATTC
T7-R2D2-PZ-R	TAATACGACTCACTATAGGGTTGCTTGTGCTCGCTACTTGC
T7-Pasha-F452	TAATACGACTCACTATAGGGACTTTGAAGTCCTACCCG
T7-Pasha-R1177	TAATACGACTCACTATAGGGCTCCTTGAACTCATAGG
T7-Loqs-F-1	TAATACGACTCACTATAGGGATGGACCAGGAGAATTTCC
T7-Loqs-R-540	TAATACGACTCACTATAGGGAAGGGCGTATCCTTGTC
T7-LacZ-F	TAATACGACTCACTATAGGGCATTATCCGAACCATCC
T7-LacZ-R	TAATACGACTCACTATAGGGCAGAACTGGCGATCGTTCG

siRNA sensor oligos

Bam-esi-1_2-S2-2P-F
GATCCCAACAGTTTATTTACTTGGAGGCAACATAATCAAATGAACTGAGGGTTACTTGGAGG
CAACATAATCAG
Sal-esi-1_2-S2-2P-R
TCGACTGATTATGTTGCCTCCAAGTAACCCTCAGTTCATTTGATTATGTTGCCTCCAAGTAAA
TAAACTGTTGG
Bam-esi-2_1-S2-2P-F
GATCCCAACAGTTTATTGGAGCGAACTTGTTGGAGTCAAAATGAACTGAGGGTGGAGCGAA
CTTGTTGGAGTCAAG
Sal-esi-2_1-S2-2P-R
TCGACTTGACTCCAACAAGTTCGCTCCACCCTCAGTTCATTTTGACTCCAACAAGTTCGCTCC
AATAAACTGTTGG
Bam-esi-1_3-S2-2P-F
GATCCCAACAGTTTATTCATTTGATCCATAGTTTCCCGAATGAACTGAGGGTCATTTGATCCA
TAGTTTCCCGG
Sal-esi-1_3-S2-2P-R
TCGACCGGGAAACTATGGATCAAATGACCCTCAGTTCATTCGGGAAACTATGGATCAAATGA
ATAAACTGTTGG
Bam-esi-1_1_ -S2-2P-F
GATCCCAACAGTTTATTGCCAAGGTACGTGGTCGACCGAAATGAACTGAGGGTGCCAAGGT
ACGTGGTCGACCGAG
Sal-esi-1_1_BC36-S2-2P-R
TCGACTCGGTCGACCACGTACCTTGGCACCCTCAGTTCATTTCGGTCGACCACGTACCTTG
GCAATAAACTGTTGG
Bam- mus308_target-S2-2P-F
GATCCCAACAGTTTATTGGGCGAGCTTGTTGGAGTCAGAATGAACTGAGGGTGGGCGAGCT
TGTTGGAGTCAGG
Sal- mus308_target-S2-2P-R
TCGACCTGACTCCAACAAGCTCGCCCACCCTCAGTTCATTCTGACTCCAACAAGCTCGCCCA
ATAAACTGTTGG

Northern probes:

esi-2.1:	GGAGCGAACTTGTTGGAGTCAA
esi-1.1:	GCCAAGGTACGTGGTCGACCGA
esi-1.2:	CATTTGATCCATAGTTTCCCG
miR-bantam:	AATCAGCTTTCAAAATGATCTCA
2S rRNA	TACAACCCTCAACCATATGTAGTCCAAGCA

Quantitative real-time PCR:

Transposon analysis:

rp49_fwd:	ATGACCATCCGCCCAGCATAC
rp49_rev:	CTGCATGAGCAGGACCTCCAG
GAPDH2_fwd:	TGATGAAATTAAGGCCAAGGTTCAGGA
GAPDH2_rev:	TCGTTGTCGTACCAAGAGATCAGCTTC
actin5c_fwd:	AAGTTGCTGCTCTGGTTGTCG
actin5c_rev:	GCCACACGCAGCTCATTGTAG
BEL1_fwd:	ATTATACAAACGCCCAATTGCCAAAA
BEL1_rev:	TCCGATGAAGCTGCAGACAAATAAGA
BLOOD_fwd:	AGACGTTCATTACAGATCAAGGTACGGA
BLOOD_rev:	AGTTCGTATGGGCAATAGTCATGGACT
DM412_fwd:	AAAGTACGGTCCAATGAAGACG
DM412_rev:	GTGGTGATGAGCTGTTGATGTT
F-element_fwd:	TTGTTGAACAGCATACCACTCC
F-element_rev:	CCAGAGTTGATGAGCCAGTGTA
gypsy6_fwd:	GACAAGGGCATAACCGATACTGTGGA
gypsy6_rev:	AATGATTCTGTTCCGGACTTCCGTCT
MDG1_fwd:	AACAGAAACGCCAGCAACAGC
MDG1_rev:	CGTTCCCATGTCCGTTGTGAT
ROO_fwd:	CGTCTGCAATGTACTGGCTCT
ROO_rev:	CGGCACTCCACTAACTTCTCC
stalker4_fwd:	TTTGGAAGATTACCAAGGCAGTTCGC
stalker4_rev:	GGATCTAACTTATGACCCGATTCGTTCC
ZAM_fwd:	ACTTGACCTGGATACACTCACAAC
ZAM_rev:	GAGTATTACGGCGACTAGGGATAC
FHV_fwd:	CCCTGGAGTCGCTTACTTGAGTGCT
FHV_rev:	ATGGAAGCGTACCTGAAGGAGGACA
DOC_fwd:	TACCTTAAACAAACAAACATGCCCACC
DOC_rev:	TTTGTATGGGTGGTCAGCTTTTCGT
DM297_fwd:	GCCAGTACACACGAACGAAATA
DM297_rev:	AATTGAATTTTGGCAATTTTGG
TABOR_fwd:	GAGCAAGAATTATGCTCGAAGAA
TABOR_rev:	AATTTATGTCCGGTTTCGTTTTT

endo-siRNA target analysis:

mus308_fwd:	AAGGATTAGCGCCAAGCTGGAGGAT
mus308_rev:	ACCACGACCACTGCCACAGAGATTC
CG9203_fwd:	AGCTGGCAGAAAAACCATGACCAGT
CG9203_rev:	CAATTCTTTTGGCGTAGCTTGAGCA

S2 cell knockdown analysis:

S2-Dcr-1_fwd:	ACGCCTTCCATCTCCCAGTTTTACC
S2-Dcr-1_rev:	GCCACCCTGCTTATTCTGACTGCTC

S2-Dcr-2_fwd:	AAACGAGAGATTCGTGCCCAAAACA
S2-Dcr-2_rev:	CTGTCCTTGCTCTTATCGGCCTTGT
S2-Drosha_fwd:	AGATGCCAGAGAACTTCACCATCCA
S2-Drosha_rev:	GAAAGAAGTGAAAAGCTGGGCAGGA
S2-Pasha_fwd:	TGTCAAGGACAAGATAACGGGCAACA
S2-Pasha_rev:	GTTGGGAGATGGCTCCGTCGTCT
S2-AGO1_fwd:	ACTACCACGTTCTGTGGGACGACAA
S2-AGO1_rev:	GAATCGTGCTCCTTCTCCACCAGAT
S2-AGO2_fwd:	AACCCTCAAAAGTAAATCATGGGAAA
S2-AGO2_rev:	ATTTTTGCTGTTGGCCTCCTTG
S2-Loqs_fwd:	GTGTGTGCGTCTGGATTTTGCTGTA
S2-Loqs_rev:	GTTTTCGGGAGGATTCGGTGTGTAT
S2-R2D2_fwd:	GCGAAGACGGAGGGTACGTCTGTAA
S2-R2D2_rev:	AGTCGAATCCTTCATCAAGCCGTGT

Supplementary Tables

Table S1. Endo-siRNA generating loci in ovaries

chromosome	start	stop	annotation	size [nt]	siRNA density [siRNAs/bp]	uniquely mapping siRNAs	uniquely mapping siRNAs (plus)	uniquely mapping siRNAs (minus)
X	1773055	1773076	CG14805/CG14818	21	1.24	102	16	86
3L	15553715	15554131	AGO2/CG7739	416	0.65	1018	683	335
2R	4999020	4999084	tsu/Mys45A	64	0.45	56	13	43
X	6183510	6183583	wuho/Rpt4	73	0.45	69	61	8
2L	20064256	20064339	lok/vls	83	0.39	68	16	52
2L	9788759	9789185	esi1	426	0.36	2955	2	2953
3L	22258908	22259084	Ddx1/CG11523	176	0.35	112	61	51
3L	7970921	7971009	CG8209/CG8038	88	0.33	44	26	18
3R	22736530	22736695	T48/Ets97D	165	0.33	78	45	33
3L	8395167	8395479	CG33057, mkg-p/CG7120	312	0.33	191	105	86
3R	6592104	6592458	Cap-H2/TfIIbeta	354	0.32	214	150	64
2L	9789830	9790265	esi1	435	0.32	8780	0	8780
3R	1053041	1053216	CG1115/katanin-60	175	0.32	82	38	44
2R	6356995	6357092	CG11777/CPT1	97	0.32	43	22	21
3R	20154174	20154376	BRWD3/CG5728	202	0.31	134	70	64
X	8610753	8611055	CG1785/l(1)g0020	302	0.26	202	128	74
2R	17016974	17017205	mago/Magi	231	0.24	106	36	70
2R	3989805	3989943	cul-4/CG11210	138	0.24	52	37	15
2R	13404771	13404916	CG4853/CG11419	145	0.23	51	24	27
X	16683171	16683650	CG4756/Rrp45	479	0.21	161	68	93
2L	14359309	14361319	Pgant35A	2010	0.21	610	301	309
2R	10050284	10050626	O-fut1/Tango7	342	0.20	135	41	94
3L	1600634	1600823	potential overlap	189	0.20	56	21	35
3L	15601809	15602197	potential overlap	388	0.19	105	59	46
2R	7113812	7114223	Cg7737/skf	411	0.18	117	65	52
3R	11163614	11164074	Sra-1/CG6218	460	0.17	129	32	97
2L	12045376	12045575	CG6770; no overlap	199	0.16	55	33	22
X	15599850	15600559	CG8134/CG9281	709	0.15	148	69	79
3R	25689428	25689611	RpS8/CG7816	183	0.15	46	12	34
3R	20449117	20449604	mld/Syx18	487	0.15	129	57	72
3L	12498567	12498831	vih/CG10654	264	0.15	51	26	25
3L	20515289	20515737	CG5104/CG4825	448	0.15	103	48	55
X	2158693	2159323	CG3071/CG2924	630	0.14	197	108	89
3L	691492	691885	RabX6/CG3279	393	0.14	75	36	39
2L	7801526	7802166	r2d2/cdc14	640	0.14	148	78	70
2R	8064757	8065016	CCT5/CG8862	259	0.14	57	25	32
3R	11179363	11179686	CG5038/CG6194	323	0.13	83	35	48
3R	4653189	4653571	pyg/CG8379	382	0.13	76	39	37
3L	259029	259327	potential overlap	298	0.12	41	22	19
3R	18949582	18949926	cenB1A/CG31365	344	0.12	46	21	25
3L	5583103	5583697	CG5146/spo	594	0.12	123	69	54
2R	18060003	18060378	a/CG3045	375	0.11	61	25	36
3L	11682255	11683235	Suq/CG14133	980	0.11	153	80	73
3R	10145711	10146334	pr-set7/CG8538	623	0.10	88	33	55
2R	7168022	7168489	Vh1/CG9062	467	0.10	62	30	32
3L	3211014	3211291	gry/CG14967	277	0.10	43	34	9
3R	23758231	23758796	CG5508/Mes-4	565	0.10	87	47	40
X	12503984	12504356	CG32654/CG1463	372	0.10	47	25	22
3R	12917269	12917631	CG5148/cher	362	0.09	46	28	18
3L	3322130	3322538	CG12016/CG11526	408	0.09	47	24	23

3'UTR overlap or predicted overlap based on EST evidence
structured cluster (note that esi2 is not identified here as it is repetetive, while this table only identifies clusters based on uniquely mapping siRNAs
genes producing siRNAs over their complete length
unclear

Table S2. Endo-siRNA generating loci in S2 cells.

chromosome	start	stop	annotation	size [nt]	siRNA density [siRNAs/bp	uniquely mapping siRNAs	uniquely mapping siRNAs (plus)	uniquely mapping siRNAs (minus)
3L	15550580	15550600	vector	20	1.90	189	120	69
3L	494514	498125	klarsicht cluster A	3611	1.80	81393	38416	42977
X	2683995	2687242	vector	3247	1.56	44169	23057	21112
3L	15549040	15549053	vector	13	1.54	60	56	4
3L	15550847	15558295	vector	7448	1.37	192350	111384	80966
3L	502144	508892	klarsicht cluster B	6748	1.16	93344	44671	48673
X	2690029	2690784	vector	755	0.90	3057	1439	1618
3L	15549735	15550140	vector	405	0.50	916	445	471
3L	15544414	15548374	vector	3960	0.40	4634	2075	2559
2R	2705192	2705225	downstream of koi	33	0.30	50	50	0
2L	9787382	9790285	esi1	2903	0.23	7098	332	6766
3R	1053044	1053208	CG1115/katanin-60	164	0.21	74	26	48
2L	1158668	1159346	tango14	678	0.19	197	100	97
3L	22258937	22259079	Ddx1/CG11523	142	0.18	40	25	15
2L	8687262	8687590	Hnf4/CG9298	328	0.16	85	10	75
3L	625207	627226	bantam precursor	2019	0.14	461	292	169
3L	691194	691891	RabX6/CG3279	697	0.13	161	67	94
3R	18572158	18572408	HP1c/CG17141	250	0.13	52	13	39
X	6183518	6183699	wuho/Rpt4	181	0.12	41	31	10
3L	7970926	7971081	CG8209/CG8038	155	0.12	48	12	36
3L	15601732	15602189	RhoGAP71E/CG7650	457	0.11	74	45	29
X	1371166	1375156	CG14781/CG14782	3990	0.11	683	277	406
2L	10231856	10233707	eIF1delta	1851	0.11	466	295	171
2R	1609634	1622872	apterous	13238	0.10	3423	1142	2281
3R	18883789	18885685	sav no overlap	1896	0.10	292	161	131
X	16683236	16683641	CG4756/Rrp45	405	0.10	64	21	43
3R	12916970	12917884	CG5148/cher	914	0.10	146	94	52
2L	8431103	8431648	grk/Akap200	545	0.10	108	33	75
3R	18545073	18546819	wge no overlap	1746	0.09	227	120	107
X	1772815	1773164	CG14805/CG14818	349	0.09	210	23	187
2L	21086874	21089157	ppk13; whole gene	2283	0.09	312	128	184
3L	20515308	20515699	CG5104/CG4825	391	0.08	47	30	17
3L	3321963	3322631	CG12016/CG11562	668	0.08	89	43	46
2R	10049715	10050536	O-Fut1/Tango7	821	0.08	201	48	153
2R	6986288	6988174	luna no overlap	1886	0.08	216	105	111
3R	17096019	17098513	CG5919/CG3308	2494	0.08	285	141	144
3R	6591969	6592520	Cap-H2/TfIIFbeta	551	0.07	61	50	11
2L	475050	475851	MED-15/POTENTIALLY CBT	801	0.07	87	36	51
2R	1595804	1604222	apterous	8418	0.07	852	414	438
2R	6418923	6420361	lola	1438	0.07	149	69	80
2R	1623127	1626081	upstream of ap	2954	0.07	310	149	161
2R	10658051	10658501	BEAF32 5'UTR sense/antisense	450	0.07	43	6	37
3R	21086061	21086445	CG11857/CG10425	384	0.07	42	22	20
2R	12083949	12084373	CG8446 5'UTR	424	0.07	44	11	33
3L	650261	652829	bantam precursor	2568	0.06	219	137	82
X	1959462	1963465	CG4199	4003	0.06	351	166	185
2L	160250	161426	spen intron	1176	0.06	102	56	46
3R	27572676	27573235	antisense gene in sense gene	559	0.06	48	27	21
3R	50391	54775	auxillin/CG18143	4384	0.06	369	145	224

siRNAs derived from the transfected plasmid which contains AGO2, flanking seqeunces and white
3'UTR overlap or predicted overlap based on EST evidence
structured cluster
other

Supplementary Figures

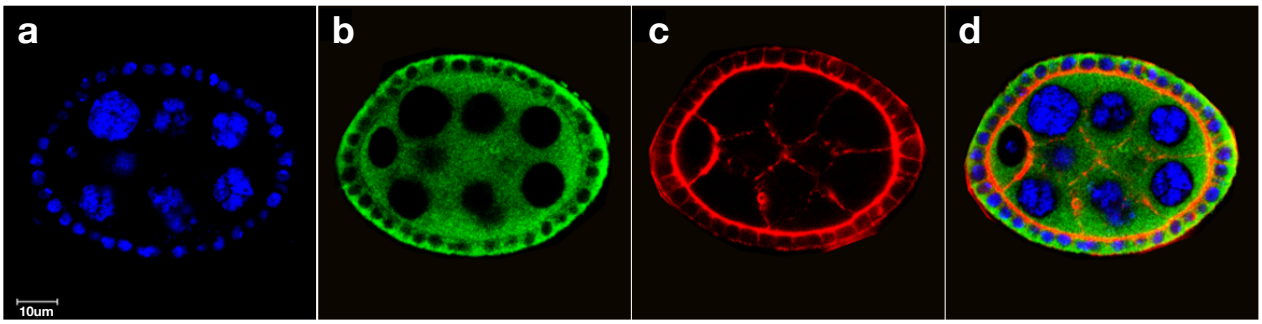


Figure S1. AGO2 localization in a developing egg chamber. Shown is immunofluorescence staining for Flag-HA tagged AGO2 expressed from its native promoter in a developing egg chamber. **(a)** DNA stained with TOPRO3; **(b)** AGO2 stained with anti HA; **(c)** actin stained with Phalloidin; **(d)** overlay; AGO2 is detected uniformly in germline and somatic cells of the ovary.

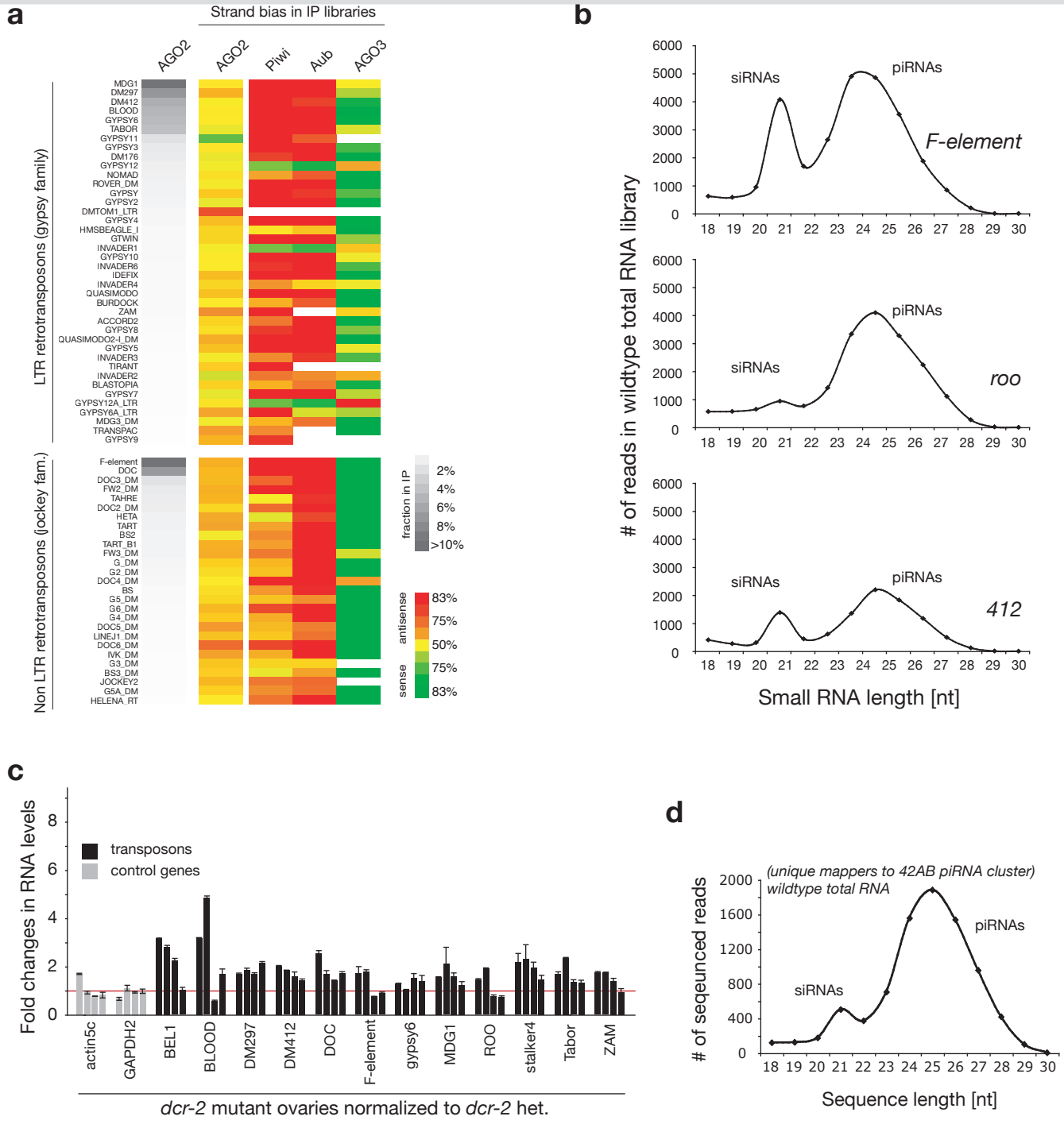
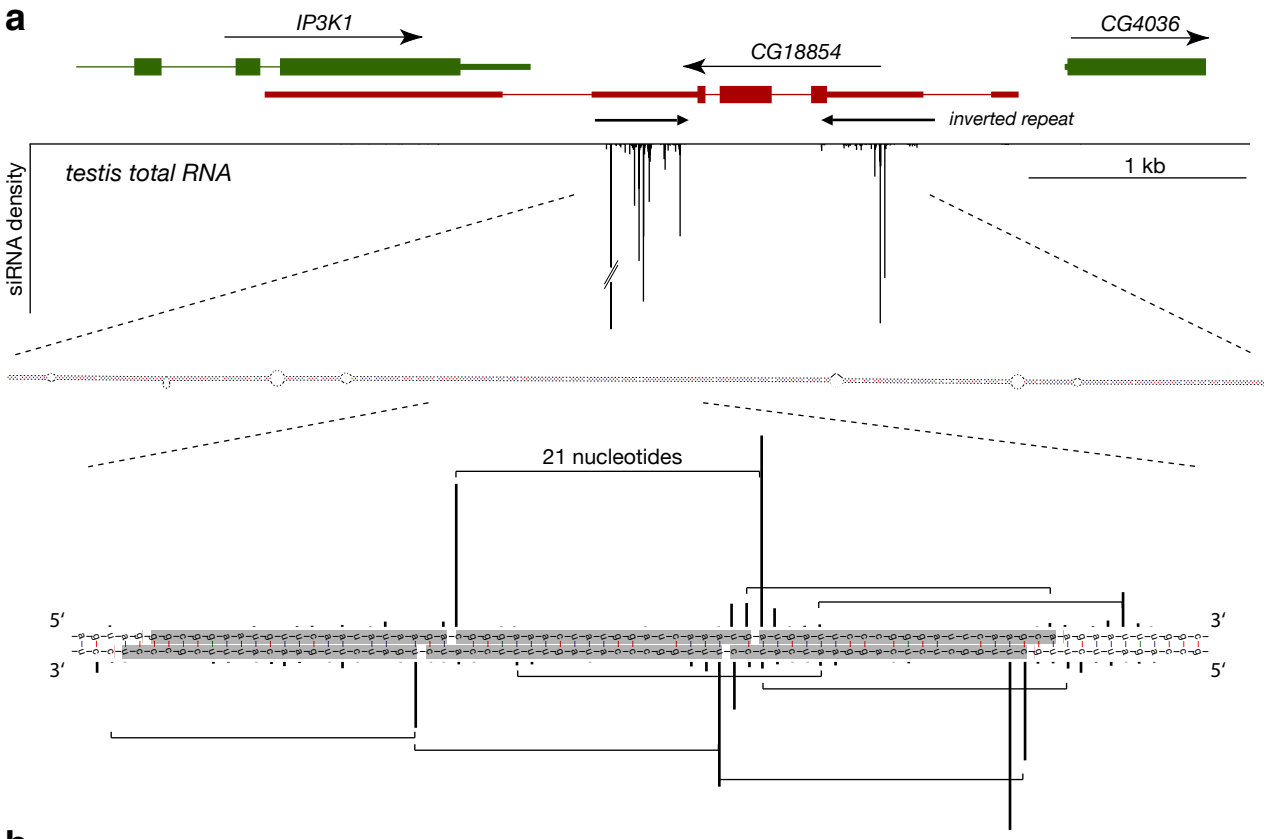


Figure S2. Transposon-derived siRNAs are not strand biased. **(a)** A heatmap was generated indicating cloning frequency (grey) and strand bias (red antisense and green sense) of AGO2 bound siRNAs from ovaries. To the right, the strand bias of Piwi, Aubergine and AGO3 bound piRNAs is provided for comparison (piRNA sequences as published in ref.12). Shown are elements from the gypsy and the jockey family of retrotransposons. Similar results were obtained from other transposon families. **(b)** Length profiles of 18-29 nt small RNAs obtained from total RNA libraries from wild-type ovaries mapping with up to three mismatches to the indicated transposons. **(c)** Steady-state RNA levels of twelve transposons and two control genes (normalized to rp49) are shown in ovaries mutant for *dcr-2* as compared to *dcr-2* heterozygotes (four independent biological replicates with error bars indicating technical variation). **(d)** Length profile of 18-29 nt small RNAs obtained from a wildtype ovarian total RNA library mapping exclusively to the 42AB piRNA cluster.



b

```
position  400                    position  408                    position  743
CG8289 5'      A                 CG8289 5'                C     CG8289 5'         G           G
           CGUU UAGAGGAUUCGGAUGA              AGGAUUCGGAUGAUGAUG UA              GCUAAGG ACGUGGUCGAC G
           GCAA AUCUCCUAAGCCUACU              UCCUAAGCCUACUACUAC AU              CGGUUCC UGCACCAGCUG C
siRNA  3'      C                 siRNA  3'                     U     siRNA  3'          A           G

position  772                    position  842                    position  851
CG8289 5'                A     CG8289 5'                C     CG8289 5'          C           C
           CACCAAGAAAGCCGACG UUCA             GCCACCAAGGAGGGCGA AUG              GAGGGCGA AUGUUUAAGAUC
           GUGGUUCUUUCGGUUGC AAGU             CGGUGGUUCCUCCCGCU UAC              CUCCCGCU UACAAGUUCUAG
siRNA  3'                G     siRNA  3'                     U     siRNA  3'           U           U

position  860                    position  884                    position  893
CG8289 5' C           C          CG8289 5'  G     A               CG8289 5'
           AUGUUUAAGAUC GAUGGAA               UAUGG CCGAAGGAUGAUU              AAGGAUGAUUCCUGGGAGCCGA
           UACAAGUUCUAG CUACCUU               AUACC GGUUUCCUACUAA              UUCCUACUAAGGACUCUCGGUU
siRNA  3' U            U       siRNA  3' UG     C               siRNA  3'

position  898                    position  903                    position  915
CG8289 5'                        CG8289 5'                        CG8289 5'              G
           UGAUUCCUGGGAGCCGAGUAA              CCUGGGAGCCGAGUAAGAAUCU             GUAAGAAUCUGGC UGCGAUG
           ACUAAGGACUCUCGGUUCGUU              GGACUCUCGGUUCGUUCUUAGA             CGUUCUUAGACCG ACGCUAC
siRNA  3'                        siRNA  3'                        siRNA  3'               A

position  920
CG8289 5'          G           U
           AAUCUGGC UGCGAUGCGCU
           UUAGACCG ACGCUACGUGA
siRNA  3'          A           U
```

Figure S3. esi-1 produces abundant phased siRNAs. **(a)** CG18854 on chromosome 2L encodes a prominent structured siRNA locus. Shown is the annotated Flybase gene structure indicating the CDS, and 5' and 3' UTRs and introns. Genes on the genomic plus strand are in green, with those on the minus strand in red. The density of siRNAs across the region is depicted below. All siRNAs derive from the genomic minus strand. The structure of the complete ~400 bp fold-back and a close-up are shown below including the 5' positions of cloned siRNAs from the total testis RNA library (vertical bars indicating cloning frequency). Brackets indicate phased 21 nt intervals. **(b)** Multiple abundant siRNAs from esi-1 share significant complementarity to the protein-coding gene, CG8289. Shown are the duplex structures and the start position within the CG8289 mRNA. Sites within CG8289 were identified using the RNAhybrid website (http://bibiserv.techfak.uni-bielefeld.de/rnahybrid/submission.html; ref. 31).

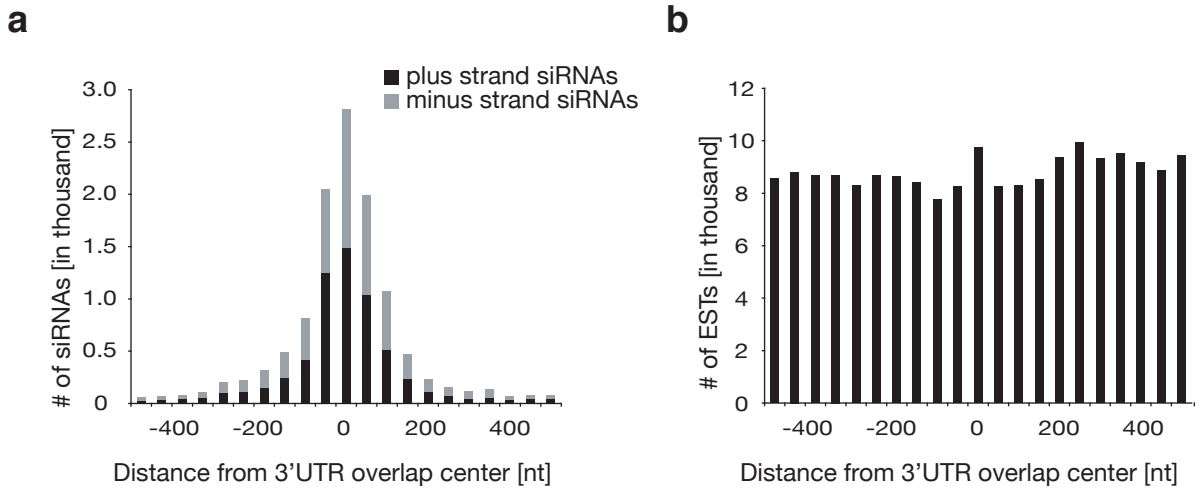


Figure S4. siRNA generation is generally confined to 3'UTR overlaps of convergently transcribed genes. **(a)** Shown are the total numbers of siRNAs (from the ovarian AGO2 IP library) mapping to all pairs of convergently transcribed genes with annotated transcript overlaps in the genome (bioinformatically extracted). Zero depicts the center of the annotated overlap and each bar represents the number of siRNAs in consecutive 50 nt windows upstream and downstream from this center. **(b)** To exclude that siRNA density is a sheer reflection of the 2 fold elevated rate of transcription in overlap regions, we plotted the number of annotated ESTs in the identical windows as for (a).

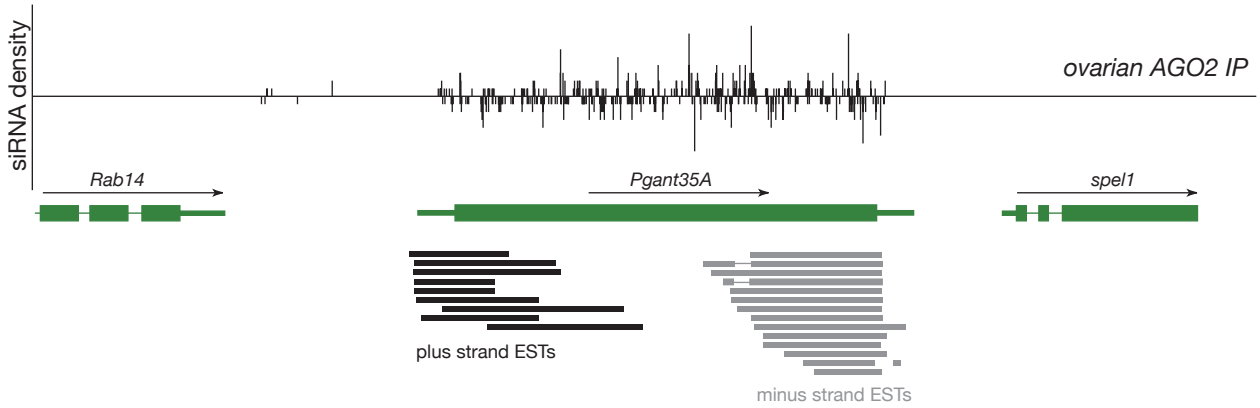


Figure S5. The Pgant35A gene produces siRNAs from both strands over the entire gene. Shown is the Pgant35A locus with flanking genes and the siRNA density from the ovarian AGO2 IP library over the entire locus above. Multiple ESTs annotated to Pgant35A support extensive sense/antisense transcription at this locus. No corresponding siRNAs were detected in the S2 cell AGO2 IP library.

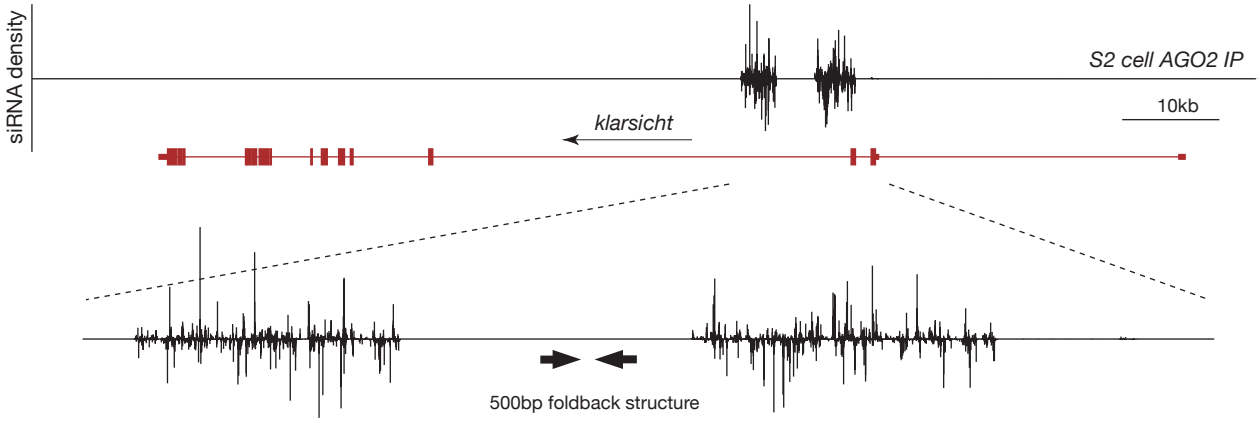


Figure S6. Two bi-directional siRNA clusters residing within the klarsicht locus. siRNAs from these two neighboring clusters account for ~16% of all S2 cell endo-siRNAs. One cluster partially overlaps a coding exon of klarsicht, likely enabling a regulatory interaction. The two clusters share no sequence relationship to each other or to any other sequence in the *Drosophila* genome or the 'nr' database at NCBI. Interestingly, two annotated transcripts residing between the siRNA clusters are nearly perfect reverse complements allowing for a dsRNA fold-back structure.
However, no siRNAs mapping to this structure were identified.

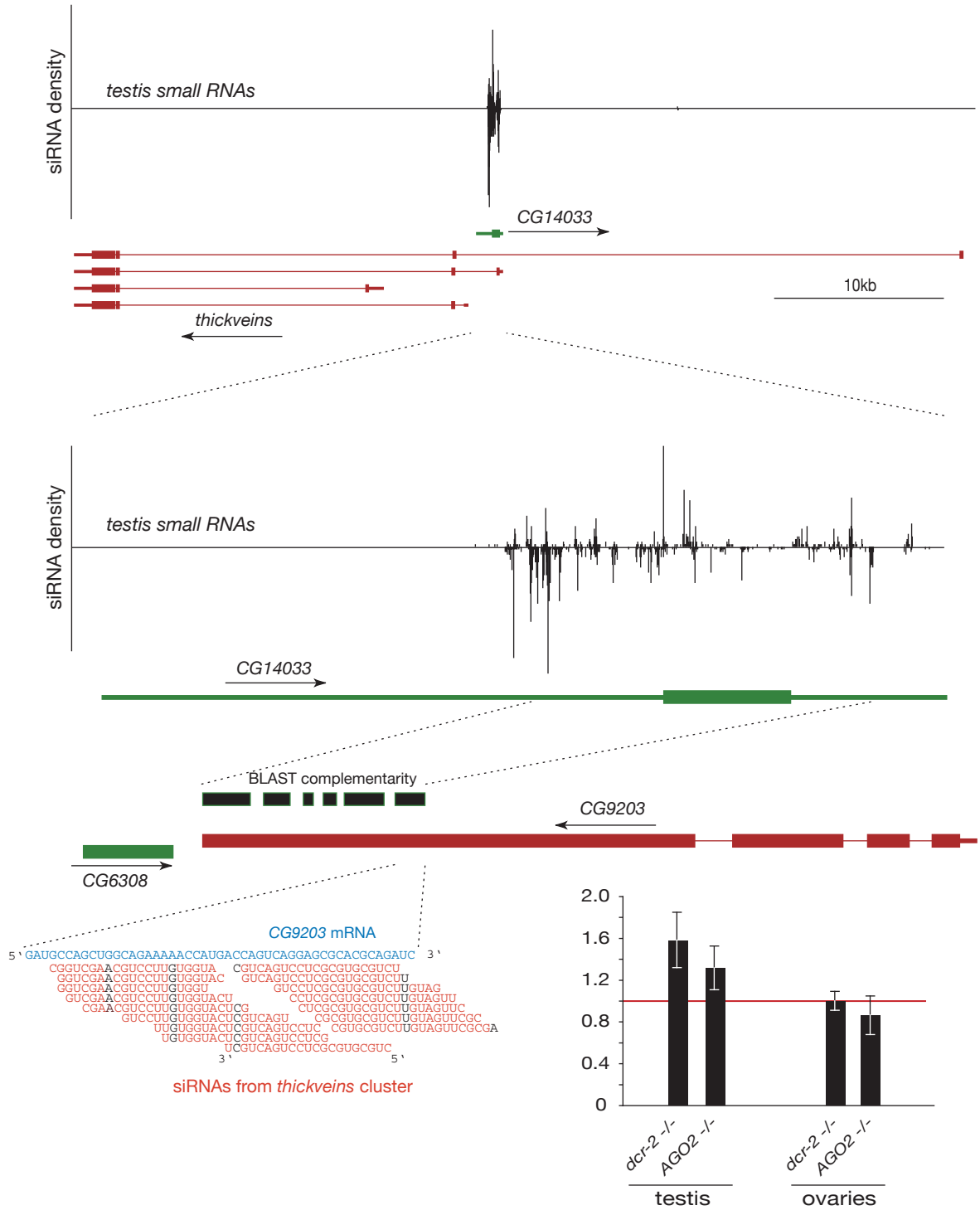


Figure S7. An endo siRNA locus resides within the thickveins gene. Shown is the siRNA density across the entire thickveins locus with annotated transcripts shown in red (genomic minus) and green (genomic plus). siRNAs are exclusively produced from an area which overlaps the annotated gene, CG14033. A magnification of the locus is shown below. A BLAST homology search reveals that the only portion of CG14033 with a significant match in the genome is confined to the area, which gives rise to endo-siRNAs. The BLAST complementarity to CG9203 is shown as black boxes. The detail below depicts one part of the sequence complementarity between the CG9203 target locus and cloned siRNAs from the thickveins cluster. The bar graph indicates CG9203 mRNA levels in testis and ovaries mutant for *AGO2* or *dcr-2* compared to their respective heterozygotes, with standard deviation indicated by the error bars (n=3). Note that siRNAs from the thickveins cluster are only detected in testis.

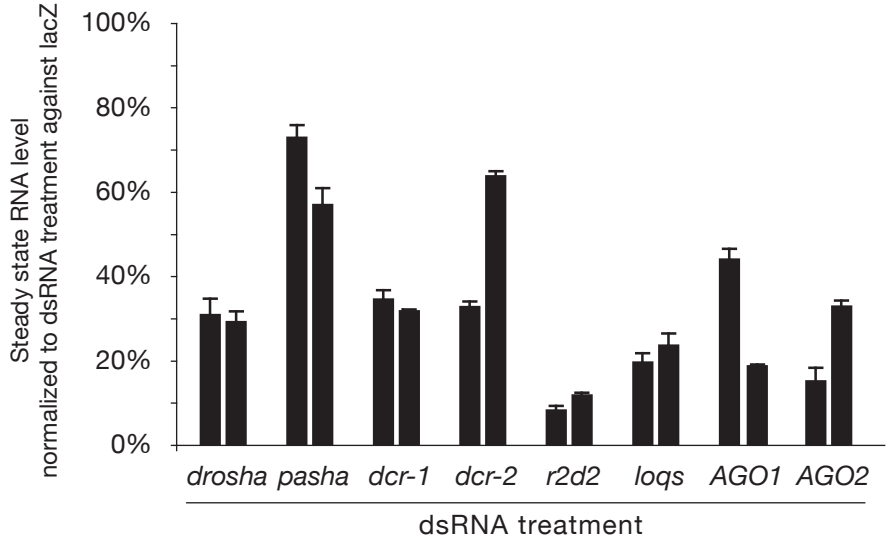


Figure S8. Knockdown efficiency by RNA interference of genes analyzed in Figure 4. Shown are steady state RNA levels (normalized to rp49) of the indicated genes after eight days of dsRNA treatment measured by qRT-PCR. Two independent experiments are shown (error bars indicate deviation within technical replicates). qRT-PCR primers match outside the regions used for dsRNA synthesis.

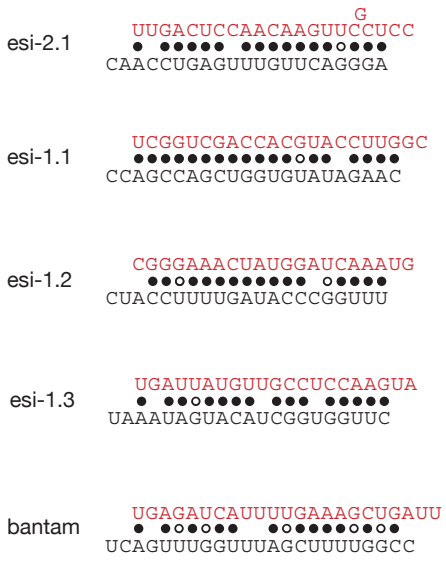


Figure S9. Duplex structures for the siRNAs analyzed in Figure 4. Shown is the siRNA sequence detected by northern blot in red and the corresponding passenger strand in black as extracted from the respective fold-back structure. Watson Crick base pairs are shown as solid, GU base pairs as open dots. Note that esi-1.1 most resembles a perfect siRNA duplex, potentially explaining its partial dependence on R2D2 (see Figure 4B).

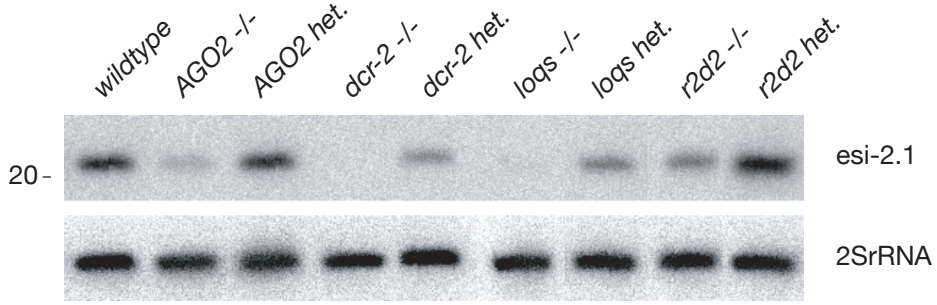


Figure S10. Genetic requirements for esi-2.1 biogenesis and stability in flies. Northern analysis of the most abundant endo-siRNA derived from esi-2. Total RNA from wild-type, mutant and heterozygote female flies (as indicated) was electrophoresed alongside RNA size markers (indicated to the left in nt). The membrane was re-probed for 2S rRNA levels to control for equal loading. Note, that *dcr-2* and *loqs* mutants show a complete or nearly complete loss of this siRNA, while *AGO2* mutants retain low levels. A measurable but less pronounced decrease is observed in *r2d2* mutants. Also note that heterozygotes for *dcr-2* and *loqs* already exhibit substantial decreases in siRNA levels.

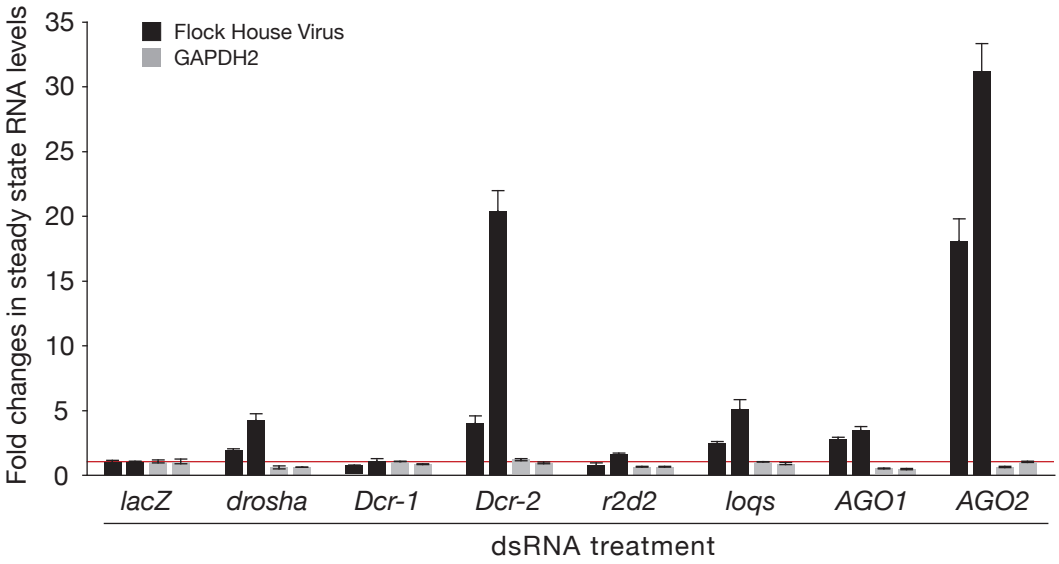


Figure S11. RNAi pathway genes and Loqs are required to control the persistent flock house virus infection in S2 cells. Shown are steady state RNA levels (normalized to rp49) of flock house virus compared to GAPDH2 in S2 cells treated with dsRNA against the indicated genes. Two independent experiments are shown and error bars indicate standard deviation within technical replicates. Values were normalized to the lacZ control. The red line indicates no change.

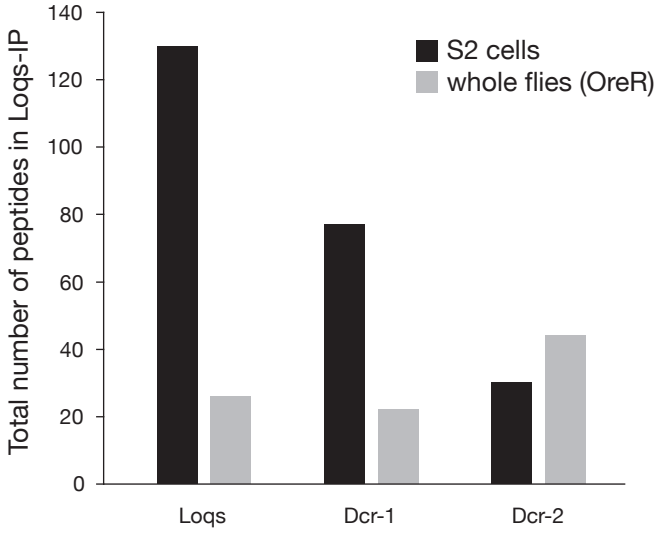


Figure S12. Loqs interacts with Dcr-1 and Dcr-2. Shown are counts of identified peptides (spectral counts) in Loqs immunoprecipitates from S2 cells and flies analyzed by MudPit (ref. 32). A polyclonal antibody against Loqs was used for immunoprecipitations (ref.1).

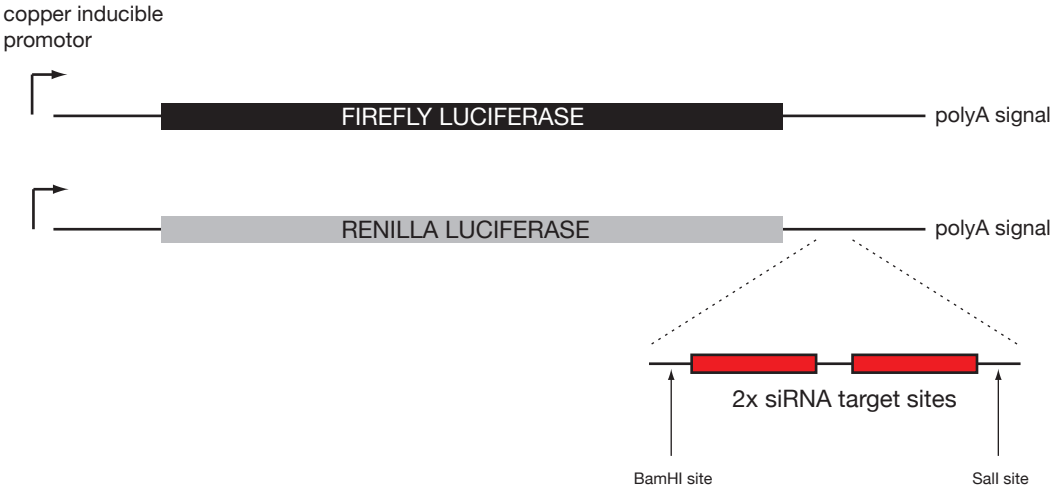


Figure S13. Schematic of Renilla/Firefly luciferase reporters used for Figure 4. Indicated are cartoons of the control reporter expressing firefly luciferase and the experimantal reporter expressing Renilla luciferase. Both reporters contain CuSO4 inducible promotors and the Renilla reporter harbors two target sites for a specific endo-siRNA.

Literature Cited

31. Rehmsmeier, M., Steffen, P., Hochsmann, M. & Giegerich, R. Fast and effective prediction of microRNA/target duplexes. Rna 10, 1507-17 (2004).
32. Wolters, D. A., Washburn, M. P. & Yates, J. R., 3rd. An automated multidimensional protein identification technology for shotgun proteomics. Anal Chem 73, 5683-90 (2001).
33. Stark, A. et al. Systematic discovery and characterization of fly microRNAs using 12 Drosophila genomes. Genome Res 17, 1865-79 (2007).
34. Jurka, J. et al. Repbase Update, a database of eukaryotic repetitive elements. Cytogenet Genome Res 110, 462-7 (2005).