Supplementary information for

Chromatin-associated RNAi components contribute to transcriptional regulation in *Drosophila.*

Filippo M. Cernilogar , Maria Cristina Onorati, Greg O. Kothe, A. Maxwell Burroughs, Krishna Mohan Parsi, Achim Breiling, Federica lo Sardo, Alka Saxena, Keita Miyoshi, Haruhiko Siomi, Mikiko C. Siomi, Piero Carninci, David S. Gilmour, Davide F.V. Corona and Valerio Orlando.

This file includes:

Figures S1 to S14

Tables S1 to S4,S6

Supplementary discussion

Supplementary references

Other supplementary information for this manuscript includes the following:

Table S5

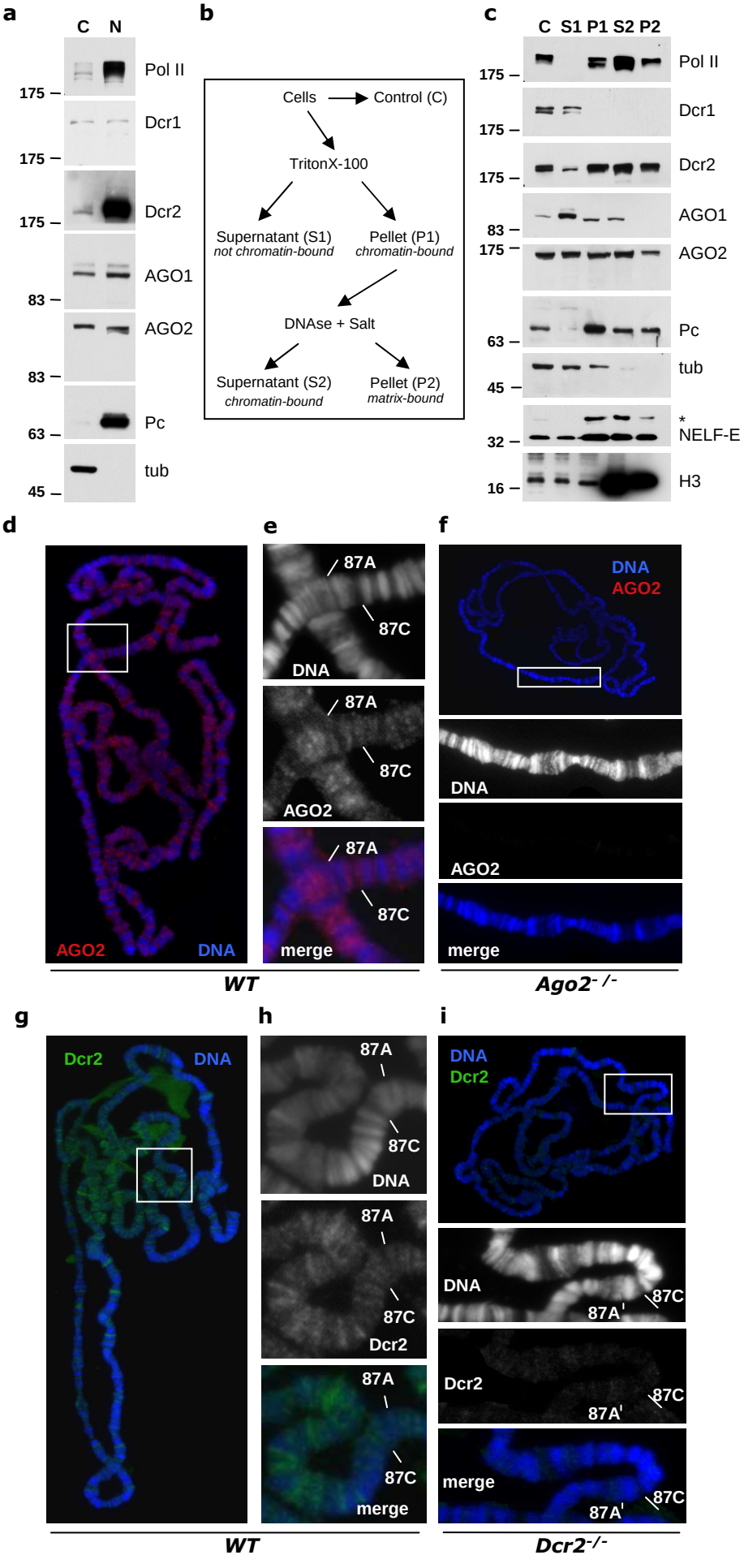
a
C N
Pol II
Dcr1
Dcr2
AGO1
AGO2
Pc
tub
175
175
175
83
83
63
45
b
Cells
Control (C)
TritonX-100
Supernatant (S1)
not chromatin-bound
Pellet (P1)
chromatin-bound
DNAse + Salt
Supernatant (S2)
chromatin-bound
Pellet (P2)
matrix-bound
c
C S1 P1 S2 P2
Pol II
Dcr1
Dcr2
AGO1
AGO2
Pc
tub
* NELF-E
H3
175
175
175
83
175
63
45
32
16
d
AGO2
DNA
e
87A
87C
DNA
AGO2
merge
f
DNA
AGO2
merge
WT
Ago2-/-
g
Dcr2
DNA
h
87A
87C
DNA
Dcr2
merge
i
DNA
Dcr2
87A'
merge
WT
Dcr2-/-

Figure S1: RNAi components are associated to chromatin. a) Equivalent amounts of cytoplasmic (C) and nuclear (N) extracts from *Drosophila* S2 cells were analyzed by western blot for the presence of the indicated proteins: RNA polymerase II (Pol II), Dicer-1 (Dcr1), Dicer-2 (Dcr2), Argonaute1 (AGO1), Argonaute2 (AGO2), Polycomb (Pc) and Tubulin (tub). **b**) Scheme of the procedure used to fractionate S2 cells extracts (adapted from ref.11). Insoluble fractions were dissolved in RIPA buffer (see methods). Chromatin associated proteins should be found in fractions P1 and S2. **c**) Equivalent amounts of the resulting samples were analyzed by western blot for the presence of the indicated proteins. Tubulin (tub) serves as a chromatin unbound marker; Pol II, Polycomb (Pc), NELF-E (∗ uncharacterised band) are chromatin bound markers; histone H3 (H3) is a tightly chromatin bound marker that is fully released only after high salt and DNAse treatment. Three independent biological samples have been analyzed. Shown are representative pictures. **d-f**) Immunolocalization of the AGO2 protein (red) on WT (**d,e**) and homozygous *Ago2414* (**f**) mutant chromosomes. DNA is stained in blue. **e**) Higher magnification of the boxed area from **d**. **f**) below the whole chromosome is shown a higher magnification of the boxed area. **g-i**) Immunolocalization of the Dcr2 protein (green) on WT (**g,h**) and homozygous *Dcr2L811fsX* (**i**) mutant chromosomes. **h**) Higher magnification of the boxed area from **g**. **i**) below the whole chromosome is shown a higher magnification of the boxed area. Single signals are shown in black and white.

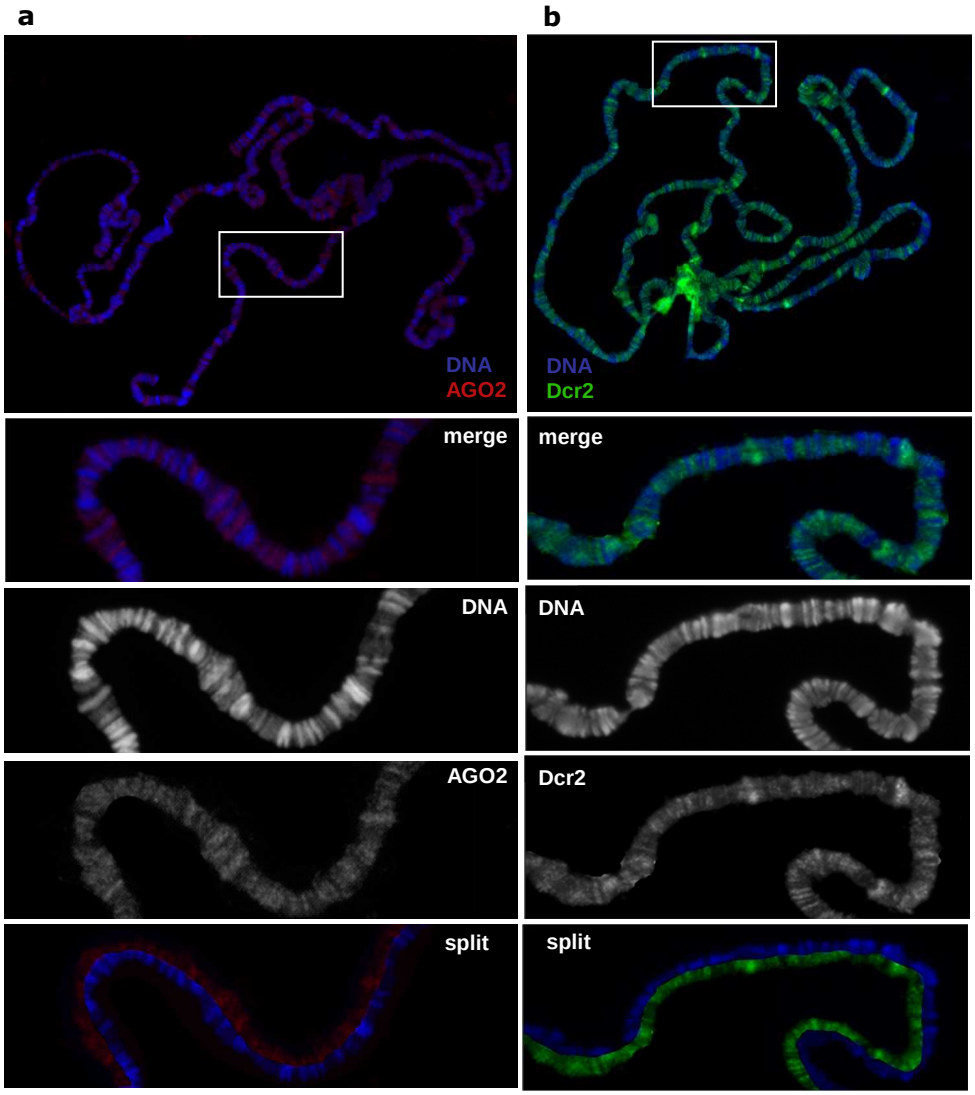

Figure S2: AGO2 and Dcr2 associate with many euchromatic sites on polytene chromosomes. Immunolocalization of AGO2 (**a**) and Dcr2 (**b**) in WT chromosomes . Blue is DNA staining. Below is shown a higher magnification of the boxed area with single signals in black and white. In order to highlight AGO2 and Dcr2 chromatin binding, the cytoplasmic and nucleoplasmic background signals were removed from polytene staining (see Materials and Methods).

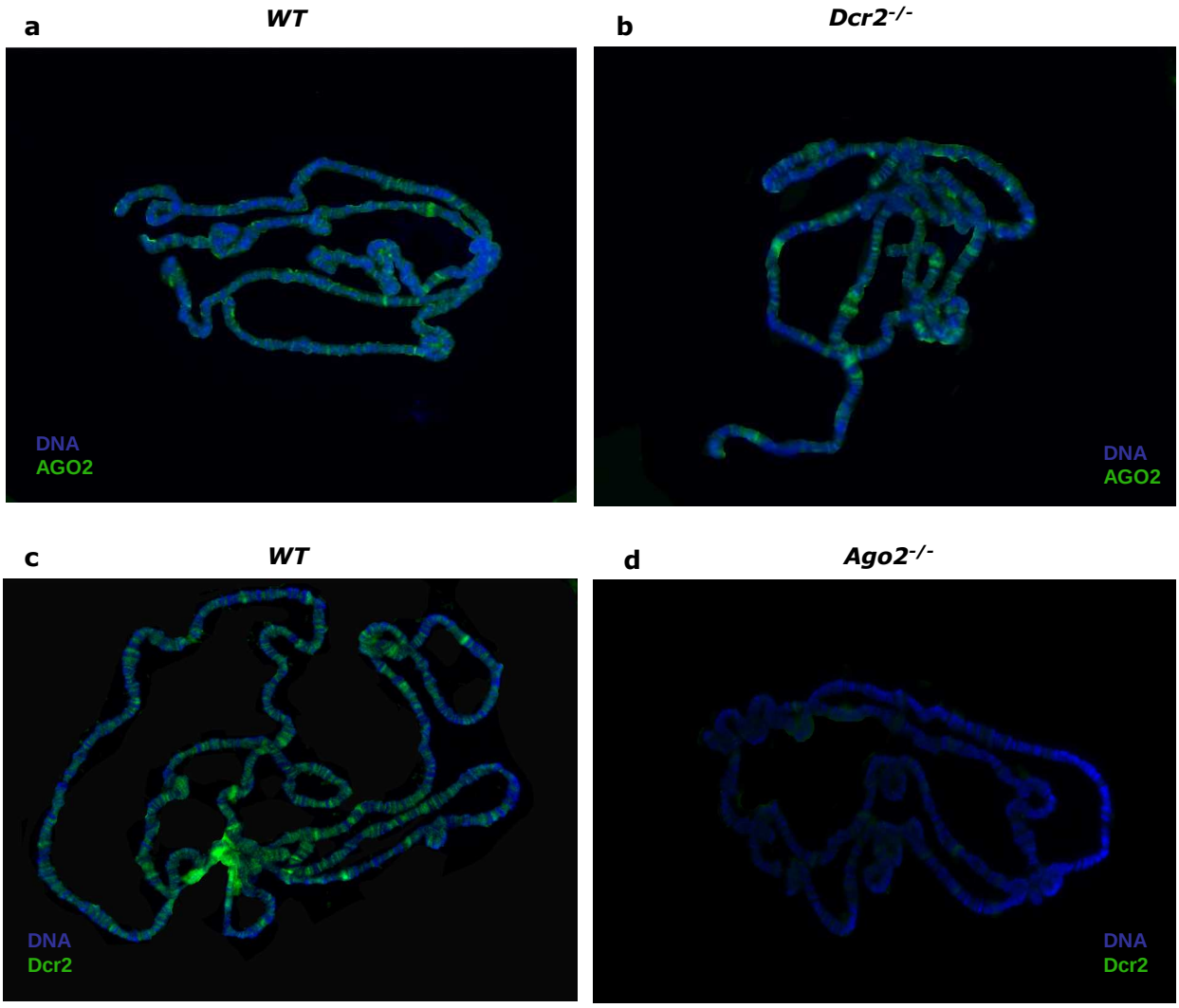

Figure S3: Chromatin binding of Dcr2 requires AGO2 but not vice versa.

a,b) Immunolocalization of AGO2 in WT (**a**) and homozygous *Dcr2*^{L811fsX} (**b**) mutant chromosomes . **c,d**) Immunolocalization of Dcr2 in WT (**c**) and homozygous *Ago2*^{414} (**b**) mutant chromosomes. Blue is DNA staining. In order to highlight AGO2 and Dcr2 chromatin binding, the cytoplasmic and nucleoplasmic background signals were removed from polytene staining (see Materials and Methods).

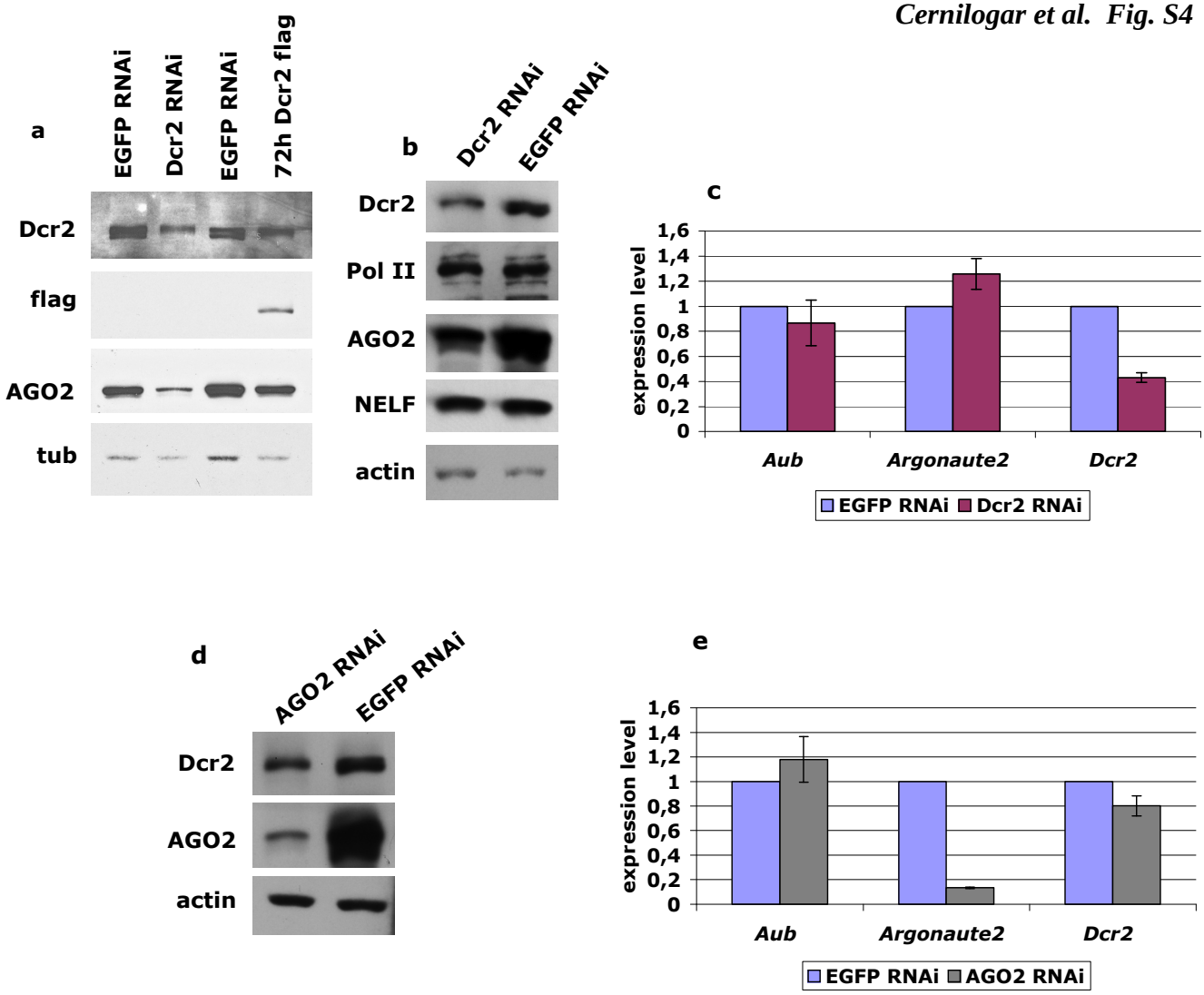

Figure S4: Protein and transcript levels in Dcr2, AGO2 depleted cells and after the expression of the Dcr2-flag transgene. a) Western blot showing the Dcr2, Dcr2-flag, AGO-2 and tubulin (loading control) protein levels in the same samples analyzed in fig. 1c. **b,c**) The samples analyzed were S2 cells treated with EGFP dsRNA (control) or Dcr2 dsRNA. **b**) Western blot with the indicated antibodies. Actin is the loading control. **c**) Quantitative RT-PCR of *Aubergine* (*Aub*; negative control), *Dcr2* and *AGO2* transcripts. n=3, bars represent the mean± standard deviation. **d,e**) The samples analyzed were S2 cells treated with EGFP dsRNA (control) or AGO2 dsRNA. **d**) Western blot with the indicated antibodies. **e**) Quantitative RT-PCR. n=3, bars represent the mean± standard deviation.

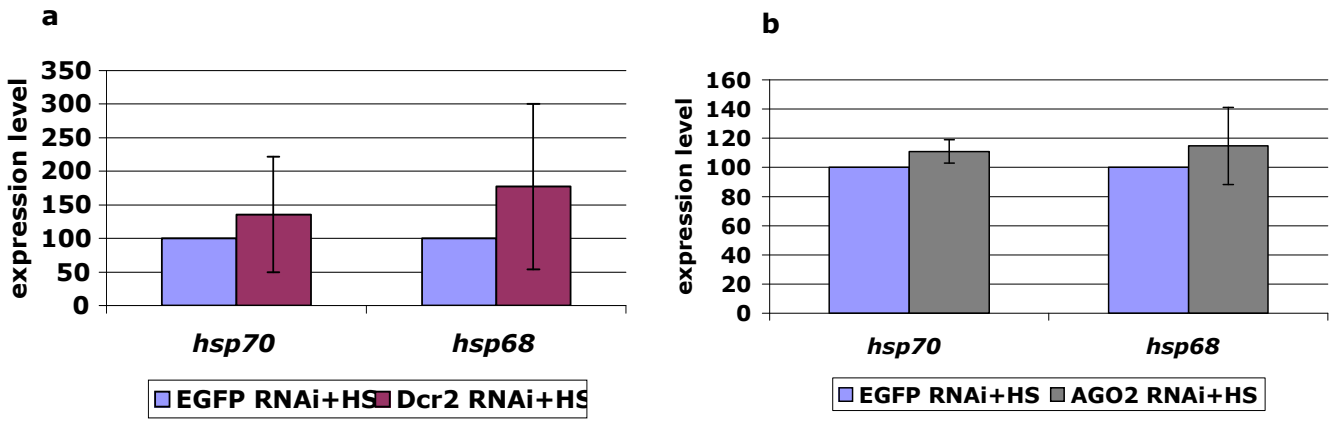

Figure S5: ***hsp70 and hsp68*** **transcript levels in S2 cells following heat shock activation.** Heat shock activates transcription of heat shock genes. In response to heat shock, the heat shock factor (HSF) rapidly associates with the promoter region. Quantitative RT-PCR analysis in S2 cells treated with EGFP dsRNA (control) or Dcr2 dsRNA (**a**) or AGO2 dsRNA (**b**) and exposed to heat shock (HS). n=3 independent RT-PCR experiments. Bars represent the mean±standard deviation.

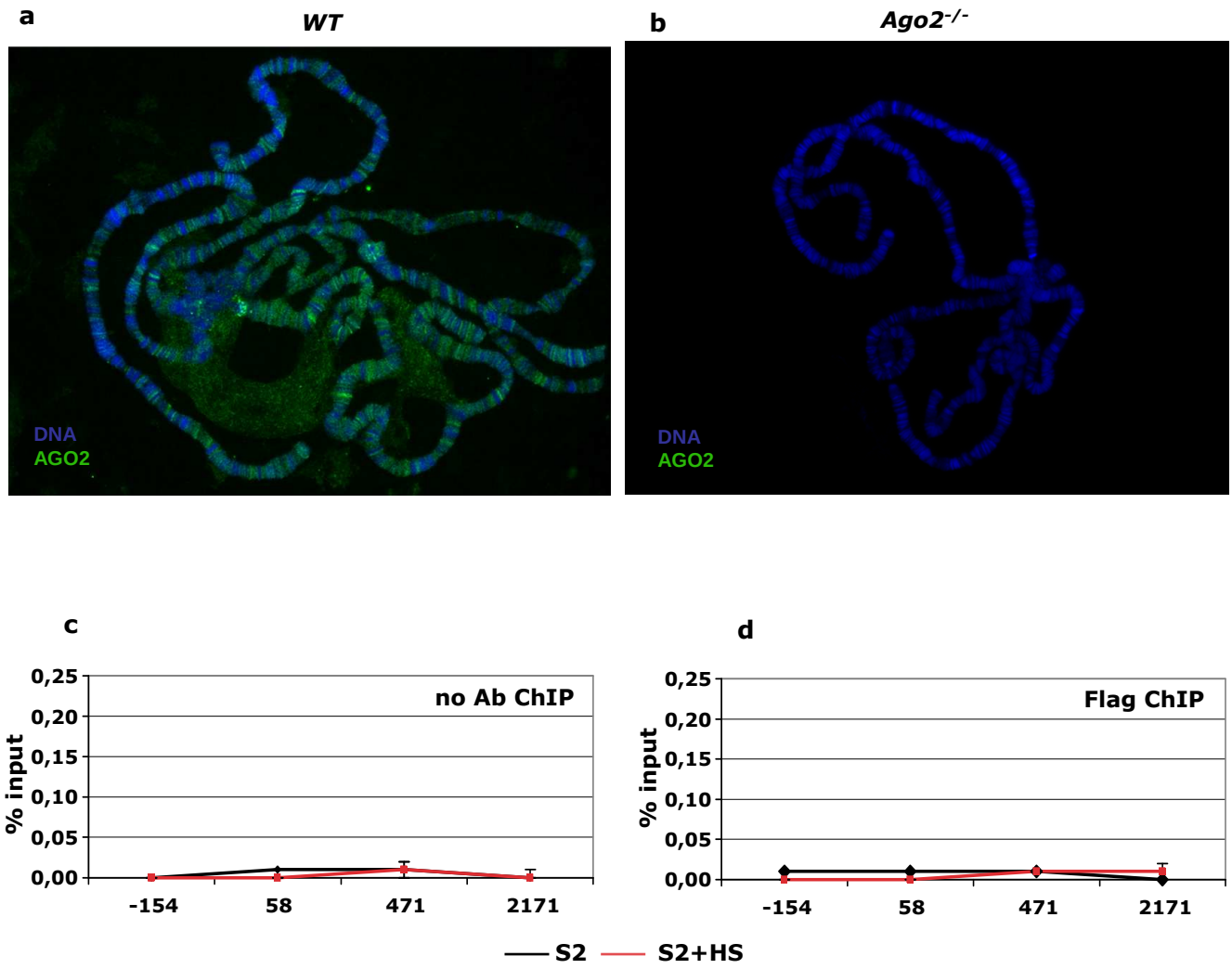

Figure S6: Controls for the ChIP and for the specificity of the anti-AGO2 (9D6) antibody. The anti-AGO2 (9D6) ChIP grade antibody recognizes specifically the AGO2 protein on polytene chromosomes**.** Immunolocalization of AGO2 (green) in WT (**a**) and homozygous *Ago2*^{414} (**b**) mutant chromosomes . Blue is DNA staining. (**c,d**) As negative control for the ChIP chromatin from S2 cells or S2 cells after exposure to heat shock (HS) was immunoprecipitated with no antibody (**c)** or anti-Flag antibody (**d)**, that has no target in S2 cells. n=3, bars represent the mean±standard deviation.

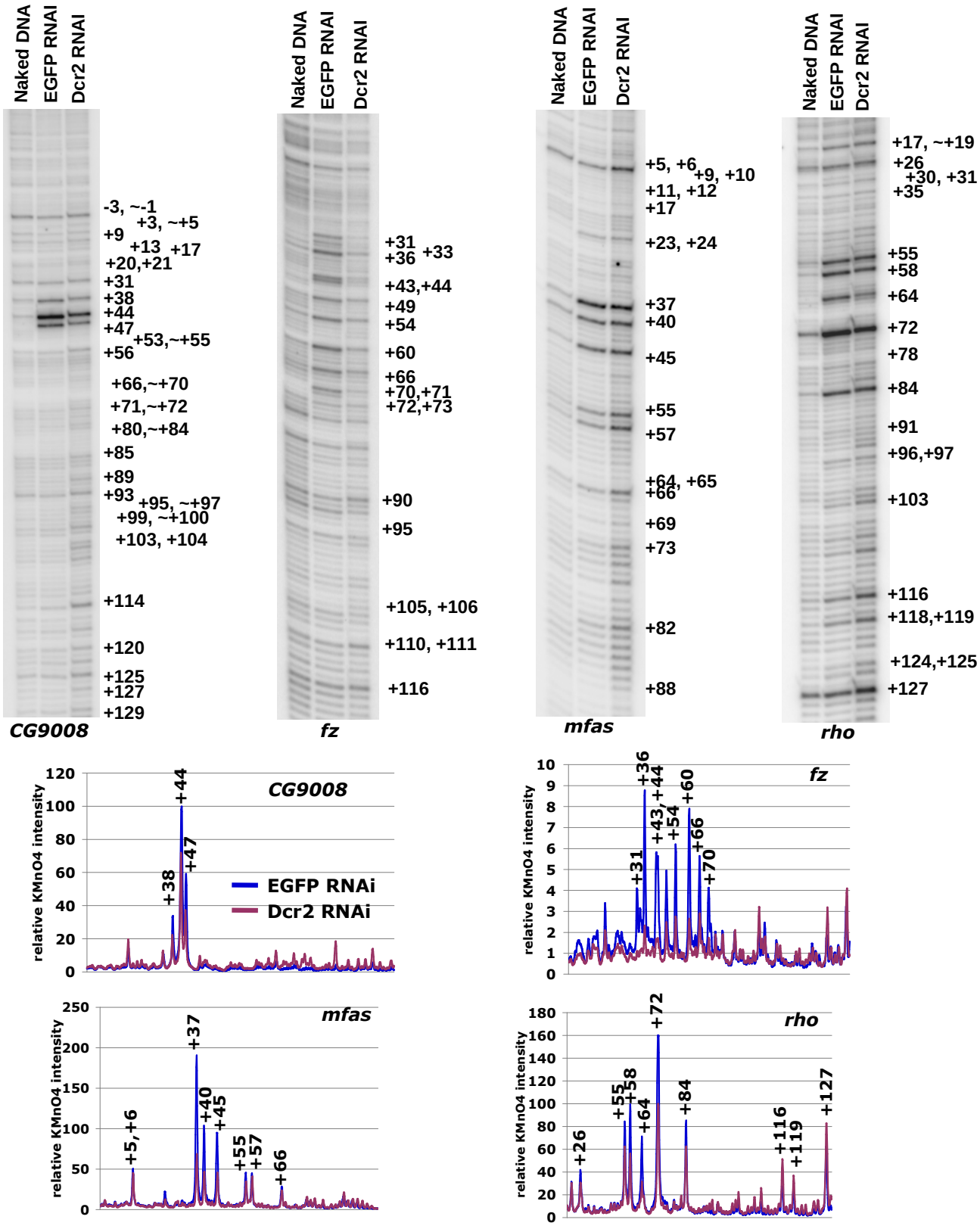

Figure S7: Transcription of non-heat shock genes is perturbed in Dcr2 depleted cells. Upper part shows the permanganate footprinting analysis of the indicated genes. Each of these panels includes the permanganate reactivity of naked DNA as a control (Naked DNA) and the permanganate reactivity of thymine residues observed in S2 cells treated with EGFP dsRNA (control) or Dcr2 dsRNA. Hyper-reactive thymines (T) are indicated. Lower part shows the quantification of the permanganate reactivity. Some hyper-reactive T are indicated.

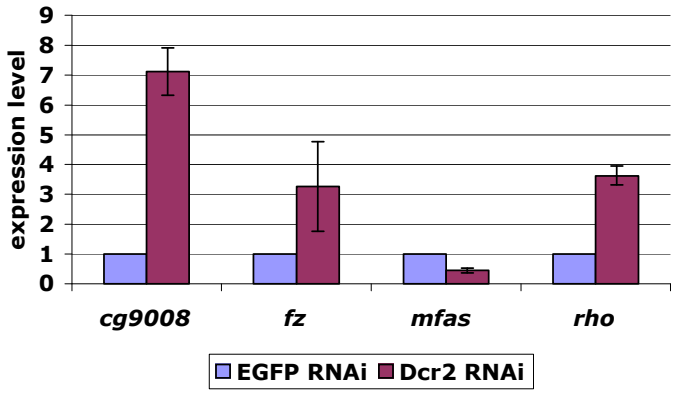

Figure S8: Quantitative RT-PCR analysis of the indicated genes in EGFP and Dcr2 RNAi samples. n=3 independent RT-PCR experiments. Bars represent the mean±standard deviation.

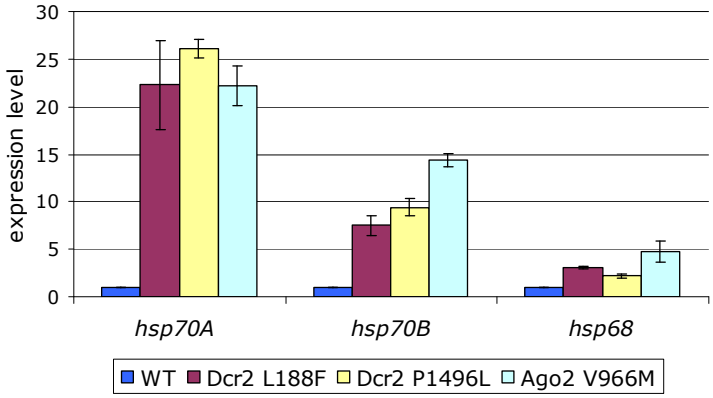

Figure S9:. **Missense mutations in Dcr2 or AGO2 change the expression levels of heat-shock genes.** Quantitative RT-PCR of *hsp70A, hsp70B* and *hsp68* transcript levels in *WT*, homozygous *Ago2V966M*, homozygous *Dcr2L188F*, homozygous *Dcr2P1496L* larvae. *n*=3 independent experiments. Bars represent the mean±standard error of the mean.

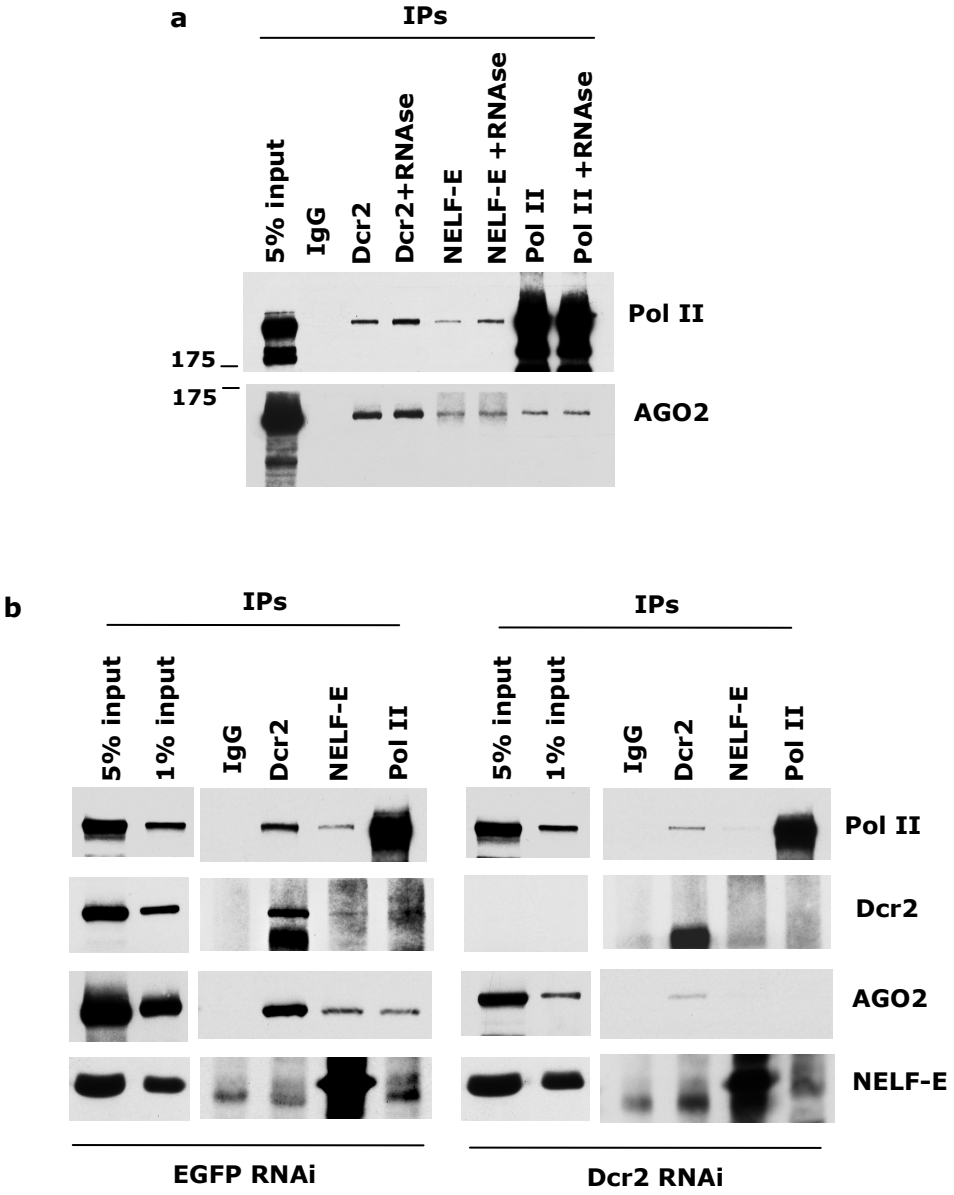

Figure S10:. **Figure 5: Dcr2 and the RNAi effector protein AGO2 associate with Pol II and NELF. a)** Nuclear extracts from S2 cells were immunoprecipitated and analyzed by western blot for the presence of Pol II and AGO2. Where indicated the samples have been treated with a RNase cocktail (RNase-A and T1). **b**) Nuclear extracts from S2 cells treated with EGFP dsRNA (control) or Dcr2 dsRNA were immunoprecipitated and analyzed by western blot for the presence of Pol II, Dcr2, AGO2, NELF-E. We observed also the reduction of AGO2 levels upon Dcr2 depletion.

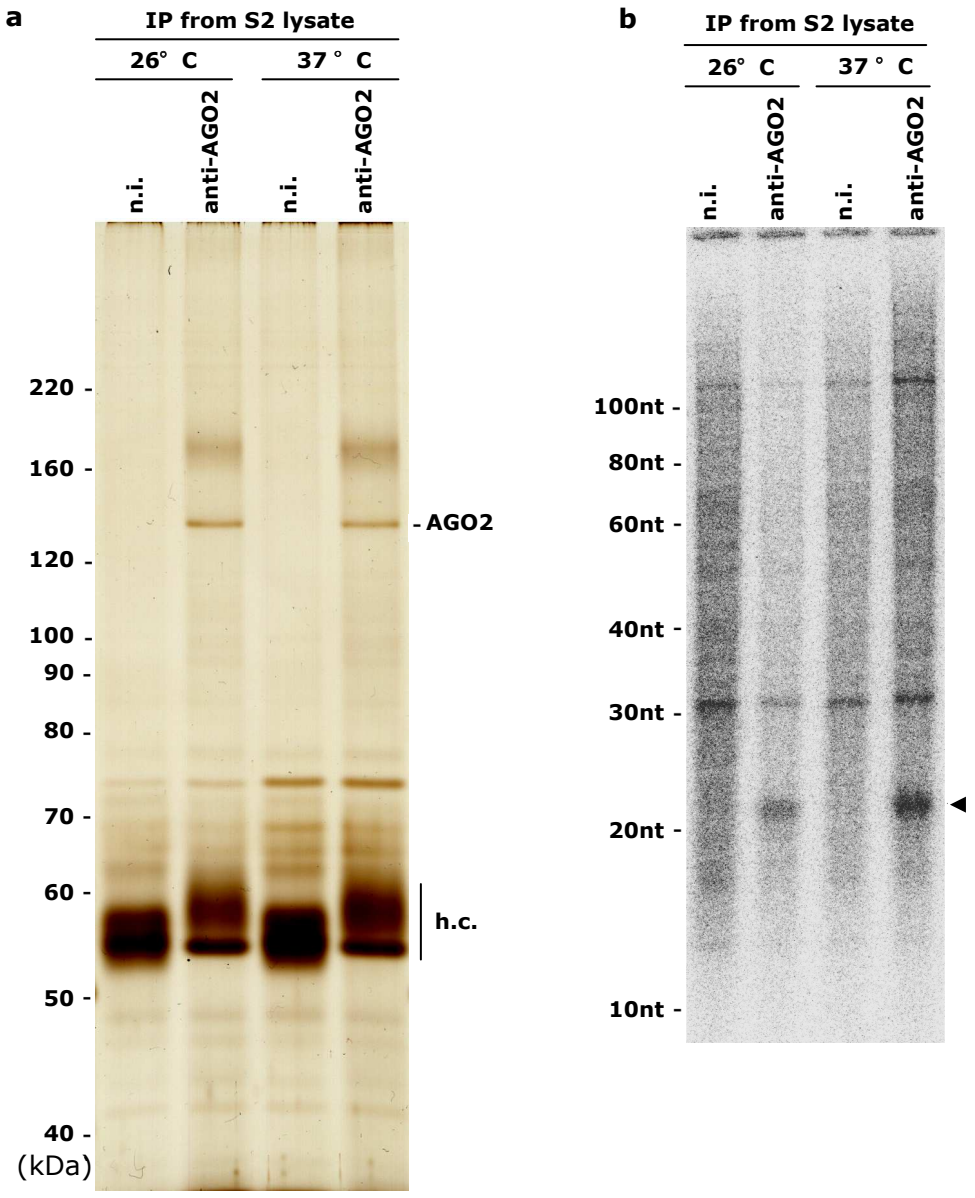

Figure S11: Purification of AGO2-associated small RNAs. (**a**) S2 cells were lysed before and after heat-shock and AGO2 was immunopurified from the lysates using anti-Ago2 antibody. Immunoprecipitated AGO2 was visualized by silver staining. n.i.: non-immune IgG (negative control), h.c.: heavy chains of antibodies. (**b**) Small RNAs co-immunopurified with AGO2 in (**a**) were visualized by [^{32}P]ATP-labeling.

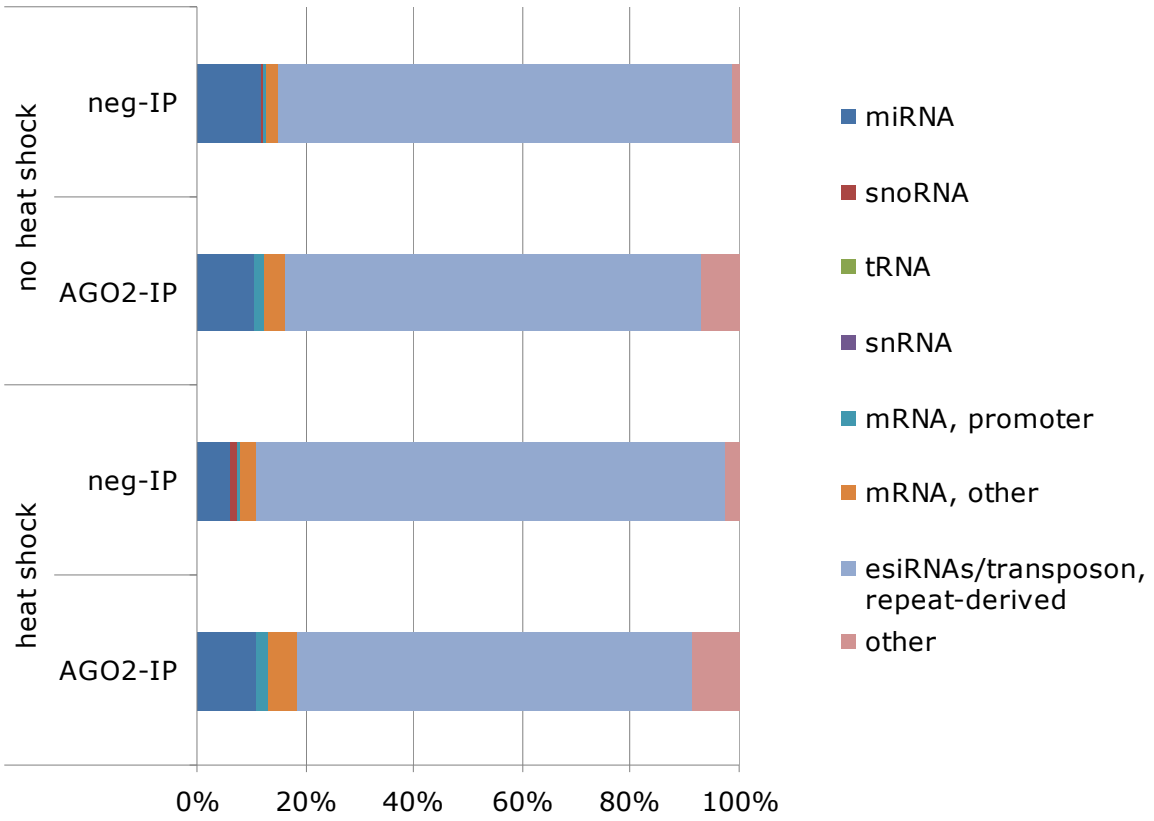

Figure S12: Comparisons of genome locations to which tags mapped. All mapped tags were assigned to various genome annotations derived from FlyBase and the UCSC genome browser.

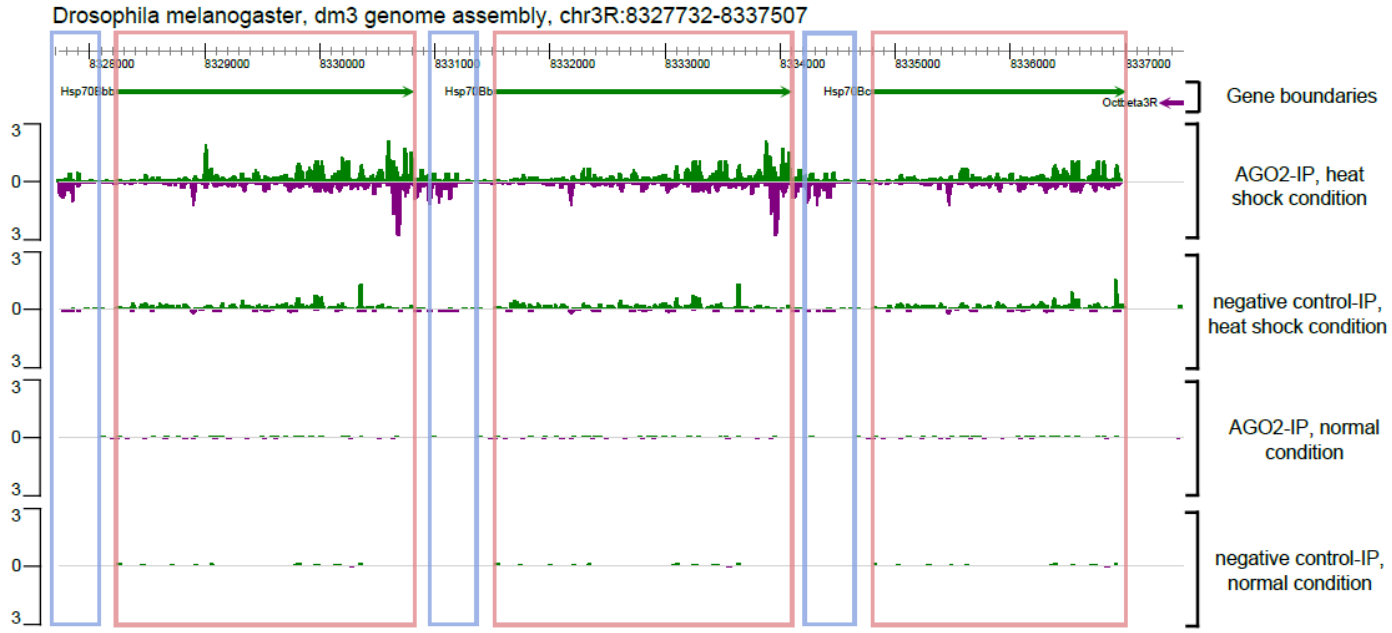

Figure S13: Gene browser view of the *Hsp70Bbb, Hsp70Bb, and Hsp70Bc* gene loci. Normalized tag counts mapping to the displayed FlyBase gene loci are visualized for each condition; condition labels are given to the right. Sense and antisense tags are depicted in green and purple, respectively. Tags are displayed by summing over the total counts at a location along the entire length of the tag. The same scale is used for each library; scales are provided to the left. Regions containing the bulk of the tags found in the promoter regions are outlined in blue. Regions corresponding to the boundaries of the genes are outlined in red. While an increase in tags is observed in the 3' region of the *Hsp70Bbb and Hsp70Bb* loci, this is not a consistently observed trend (see global analysis in Fig. S14a) Genome visualization provided by the ZENBU Omics Data Integration and Visualization System (Severin J, http://fantom.gsc.riken.jp/zenbu/).

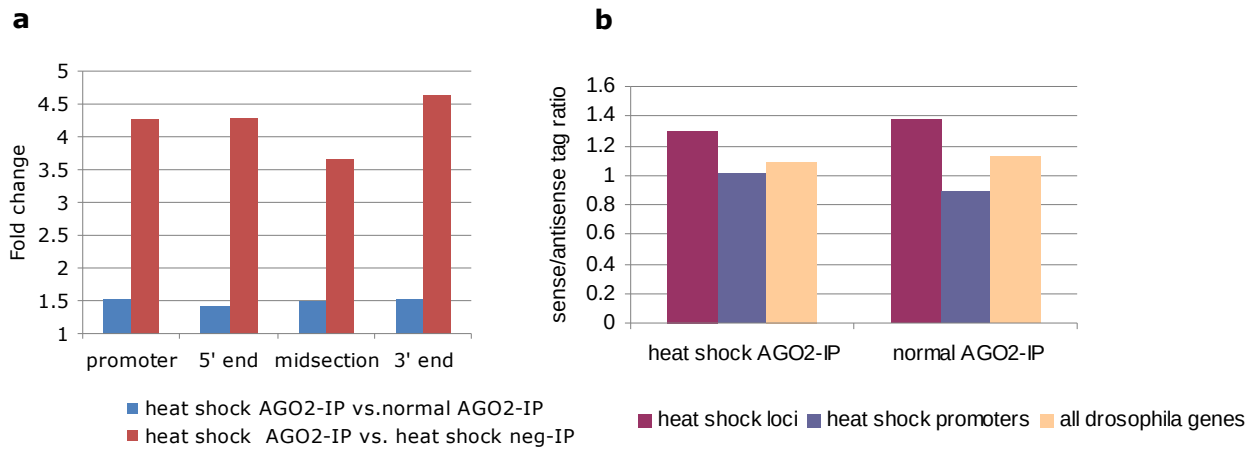

Figure S14:. **Features of AGO2-associated small RNAs. a)** Comparison of short RNA tags mapping to different locations of transcribed regions. To ascertain whether short RNAs derived from transcribed regions were biased in their origins for specific spatial locations, all transcripts (defined by Flybase) were divided into four segments: a fixed promoter length (-500/+50) centered around the TSS and three equivalently-sized sections normalized to the length of a given transcript yielding a 5' region, a midsection region, and a 3' region. Total tag counts were summed across all segments and relative fold enrichment (*y*-axis) was compared across the libraries listed at the bottom of the figure. **b)**Sense/antisense tag ratio for the indicated short RNA IP-seq libraries; ratios calculated by summing all tag counts across the indicated gene definitions in both the sense and antisense directions (see Methods for more information).

a

Mutant's name	Molecular lesion	Additional information
Dcr2^{L811fsx} (null)	Premature stop codon	Full length protein:NO Dicing activity:NO RNAi activity:severely reduced
Dcr2^{L188F}	Amino acid substitution in the DExH helicase domain	Full length protein:YES Dicing activity:YES RNAi activity: reduced
Dcr2^{P1496L}	Amino acid substitution in the RNAse III domain	Full length protein:YES Dicing activity:NO RNAi activity: reduced
Ago2^{414} (null)	Deletion of exons 1 and 2	Protein:NO Slicing activity:NO RNAi activity:severely reduced
Ago2^{V966M}	Amino acid substitution in the PIWI domain	Full length protein:YES Slicing activity:NO RNAi activity: reduced

b

Genotype	Chromosomes analyzed	Puffed	Not puffed	% Puffed	% Not puffed
w^{1118}	18	3	15	16.60%	83.40%
Dcr2^{L811fsx} *(null)*	35	34	1	97.15%	2.85%
Dcr2^{P1496L}	24	8	16	33.30%	66.70%
Dcr2^{L188F}	22	8	14	36.30%	63.70%
Ago2^{414} *(null)*	36	24	12	66.66%	33.34%
Ago2^{V966M}	24	12	12	50%	50%

Table S1: a) mutants analyzed in this study. **b**) results of the *hsp70* DNA-FISH. The relative puffing ratio existing between 87C and 87A was calculated by quantitative densitometric analysis. Any deviation of 33% from the normal average relative puffing ratio existing in wild type at the 87C/87A loci was considered an increase in puffing at the 87C relative to 87A in the *Ago2* and *Dcr2* mutants analyzed.

Genotype	Tot Bands	Not puffed bands	Puffed bands
w^{1118}	38	25	13
	32	24	8
	30	20	10
	39	29	10
	38	26	12
	38	27	11
$Ago2^{V966M}$	104	92	12
	102	89	13
	117	103	14
	98	85	13
	117	102	15
$Ago2^{414}$*(null)*	105	93	12
	108	95	13
	112	99	13
	112	101	11
	103	91	12
$Dcr2^{L811fsx}$*(null)*	87	75	12
	91	78	13
	101	87	14
	97	84	13
	91	79	12
$Dcr2^{L188F}$	83	71	12
	88	77	11
	87	76	11
	94	81	13
	95	83	12
$Dcr2^{P1496L}$	102	89	13
	95	83	12
	90	80	10
	103	90	13
	97	86	11

Table S2: number of bands, corresponding to active Pol II, on polytene chromosomes after heat shock. Five/six chromosomes for each genotype have been analyzed**.**

condition	library	total extracted	total mapping	mapping%	map exactly	map 1 mismatch	map 2 mismatches
heatshock	**AGO2-IP**	12801981	12357974	96.5	9900075	1830382	627517
	IgG-IP	4545467	4097076	90.1	3633206	314027	149843
no heatshock	**AGO2-IP**	11931648	11500101	96.4	9757820	1298258	444023
	IgG-IP	4578303	4303460	94.0	3994882	191910	116668

Table S3: A summary of the basic statistics for the four libraries is given .

	Raw tag counts				**tags per million (tpm)**			
	heat shock		**no heat shock**		**heat shock**		**no heat shock**	
	AGO2-IP	**neg-IP**	**AGO2-IP**	**neg-IP**	**AGO2-IP**	**neg-IP**	**AGO2-IP**	**neg-IP**
CG14207	0.5	0.0	0.0	0.0	0.1	0.1	0.1	0.1
Hsc70-2	0.0	0.1	0.0	0.0	0.1	0.1	0.1	0.1
Hsc70-4	2.0	0.0	2.0	0.0	0.2	0.1	0.2	0.0
Hsc70-5	0.0	0.0	1.0	0.0	0.1	0.1	0.1	0.1
Hsf	0.0	0.0	1.0	0.0	0.1	0.1	0.1	0.1
Hsp22	1.0	0.0	0.0	0.1	0.1	0.1	0.1	0.1
Hsp23	21.0	0.0	0.0	0.0	1.7	0.1	0.1	0.1
Hsp27	4.0	0.0	0.0	0.0	0.3	0.1	0.1	0.1
Hsp60	0.0	1.0	0.0	0.0	0.1	0.2	0.1	0.1
Hsp60B	0.0	0.3	0.0	0.0	0.1	0.1	0.1	0.1
Hsp67Ba	3.0	0.0	1.5	0.0	0.2	0.1	0.1	0.1
Hsp67Bc	0.0	0.0	0.0	0.1	0.1	0.1	0.1	0.1
Hsp68	1.0	0.0	7.0	0.0	0.1	0.1	0.6	0.1
Hsp70Aa	13.8	0.4	1.0	0.0	1.1	0.1	0.1	0.1
Hsp70Ab	13.2	1.7	0.0	0.0	1.1	0.4	0.1	0.1
Hsp70Ba	604.0	15.0	0.6	0.7	48.9	3.7	0.1	0.2
Hsp70Bb	159.1	3.9	1.7	0.4	12.9	0.9	0.1	0.1
Hsp70Bbb	167.1	3.9	1.7	0.4	13.5	0.9	0.1	0.1
Hsp70Bc	159.1	3.9	1.7	0.4	12.9	0.9	0.1	0.1
Hsp83	0.0	0.2	0.0	0.0	0.1	0.1	0.1	0.1
Hsromega	2.0	0.0	0.0	0.0	0.2	0.1	0.1	0.1

Table S4: Short RNA raw (left) and normalized (right) counts mapping to promoter regions -500/+50 around transcriptional start sites. Tags counts are strandles; derived from the sense or anti-sense direction of the loci. Tags are normalized by calculating parts per million (tags per million or tpm). Loci for which no tags were recovered are not shown in the table. As an additional normalization strategy, tags per million miRNA counts were also tabulated for each library (Table S5) because miRNA percentages were largely unaffected by heat shock treatment (Figure S12); fold enrichments calculated across conditions with these values were consistent with tpm normalization (data not shown).

heat shock locus	CAGE expression, sense	CAGE expression, antisense
CG14207	31034.0	0.0
Hsc70-1	18.0	17.0
Hsc70-2	5.0	31.0
Hsc70-3	15729.0	52.0
Hsc70-4	21561.0	1.0
Hsc70-5	905.0	12.0
Hsf	3183.0	44.0
Hsp22	44.0	5.0
Hsp23	660.0	4.0
Hsp26	275.0	8.0
Hsp27	16881.0	35.5
Hsp60	1255.0	19.0
Hsp60B	2.0	7.0
Hsp67Ba	0.0	7.0
Hsp67Bc	218.0	2.0
Hsp68	155.0	32.0
Hsp70Aa	102.3	102.0
Hsp70Ab	55.5	80.0
Hsp70Ba	44.8	3.0
Hsp70Bb	46.8	46.3
Hsp70Bbb	25.4	46.3
Hsp70Bc	250.4	50.3
Hsp83	15548.0	42.0
Hsromega	236.0	0.0

Table S6: Sense and antisense transcriptional start site activity as measured by CAGE expression at the 5’ and 3’ ends of heat shock loci (see Methods). CAGE data generated for embryonic tissue as part of the Drosophila modENCODE project. Rows highlighted in orange correspond to loci with roughly equivalent sense and antisense expression according to CAGE; additional loci with unusually high relative antisense expression or lowly-expressed equivalent sense and antisense expression are highlighted in light red.

Supplementary discussion

Classic studies of RNAi function in heterochromatin formation have been suggesting a role of chromatin associated RNAi in gene silencing[1]. However, heterochromatin formation requires transcription, thus even heterochromatin gene silencing may involve co-transcriptional mechanisms, perhaps via non coding RNA. Our data revealed an unexpected preferential association of RNAi components with transcriptionally active loci and loss of function alleles determine defects in RNA Pol II pausing and transcriptional up-regulation, particularly at inducible heat shock genes, suggesting that RNAi function is required for correct RNA Pol II function.

The precise role of RNAi and short RNAs in Pol II regulation remains to be elucidated. However, RNA silencing mechanisms are known to act on aberrant transcripts. It is unclear how aberrant transcripts are recognized but the mechanism seems to involve competition between different RNA processing pathways, including exosome activity[2] . The notion of aberrant RNA remains to be fully understood. The cell may consider as aberrant either a transcript not present normally in the cell or also illegitimate levels of transcription. The latter may require a fine tuned co-transcriptional mechanisms that may involve Pol II. Our data show that chromatin associated RNAi components affect Pol II, particularly after heat shock and that loss of these components impair Pol II in heat shock stress response. It is tempting to speculate that in *Drosophila*, in particular under stress conditions, aberrant RNAs are produced and play a role in initiation of RNAi-mediated silencing. Indeed, our findings have intriguing parallelisms with a recent report regarding the role in fission yeast of RNAi in heterochromatin formation and transcriptome. In that work, Moazed and co-workers proposed a model for RNAi as a surveillance mechanism in which small RNAs called priRNA, derived from degradation of abundant transcripts, would bind AGO1 to target transcripts that result from bidirectional transcription of DNA repeats as well as other mRNAs[2] . A similar model for Dicer function was proposed, in the same system, in gene silencing control of some euchromatic sites and transposable elements[3]. Our data indicate that AGO2-associated short RNA sequences match sense and antisense transcripts at heat shock loci, suggesting that RNAi operates on both. Of note, we found a strong bias for antisense, particularly after heat shock as well as throughout the *Drosophila* genome[4]. Thus, we propose that in higher eukaryotes chromatin associated RNAi activities may operate co-transcriptionally to control the transcriptome via sense and antisense transcripts.

Supplementary references

1. Grewal, S.I., Elgin, S.C. Transcription and RNA interference in the formation of heterochromatin. *Nature*. **447**, 399-406 (2007).

2. Halic, M., Moazed, D. Dicer-independent primal RNAs trigger RNAi and heterochromatin formation. *Cell*. **140**, 504-16 (2010).

3. Woolcock, K.J., Gaidatzis, D., Punga, T., Bühler, M. Dicer associates with chromatin to repress genome activity in Schizosaccharomyces pombe. *Nat Struct Mol Biol.* **18**, 94-9 (2011).

4. Hoskins, R.A. et al. Genome-wide analysis of promoter architecture in Drosophila melanogaster. *Genome Res.* **21**,182-192 (2011).