Supplemental information

The HP1 homolog Rhino anchors a nuclear complex that suppresses piRNA precursor splicing

Zhao Zhang, Jie Wang, Nadine Schultz, Fan Zhang, Swapnil S. Parhad, Shikui Tu, Thom Vreven, Zhiping Weng and William E. Theurkauf

SUPPLEMENTAL FIGURE LEGENDS

Figure S1. Rhi does not bind with flamenco cluster. Related to Figure 1.

(A) piRNA and Rhi ChIP-Seq signal across the *flamenco* (*flam*) cluster and cluster 32, a dual-strand piRNA cluster located ~60 kb downstream of flamenco. Rhi binds to dual-strand cluster 32, but does not associate with *flam.*

(B) Scatter plots showing ChIP-Seq enrichment (*y-axis*) relative to reduction in piRNA production in $rhi^{2/KG}$ mutants (*x-axis*). Each point represents a distinct piRNA cluster. Pre-I.S. is Pre-Immune Serum. The major germline piRNA cluster at 42AB and the major somatic cluster (*flam*) are indicated in green.

Figure S2. Rhi binding may not depend on the piRNA production. Related to Figure 1.

(A) Small RNA-Seq and Rhi ChIP-Seq signal across the right arm of chromosome 2 in control (*armi/CyO* and w^1) and *armi* mutants. The 42AB cluster is indicated. Mutations in *armi* nearly eliminate piRNAs mapping to 42AB (Malone et al., 2009), but do not reduce Rhi binding.

(B) Scatter plot showing Rhi ChIP-Seq enrichment in $armi^{1/72.1}$ mutants relative to w^1 controls. Each dot represents a distinct piRNA cluster.

(C) A biological replicate of Panel B.

Figure S3. Loss of Rhi, Cuff or UAP56 does not affect total cluster and gene transcript steady state level. Related to Figure 2. Scatter plots comparing normalized RNA-Seq reads uniquely mapping to clusters and genes in *rhino*$^{2/KG}$, *cuff*^{wm25} and *uap56*$^{sz/28}$ ovaries and corresponding controls.

Figure S4. Mutations in *rhi, cuff* and *uap56* increase splicing at the *sox102F* locus. Related to Figure 2. The experimentally determined intron-exon structure of the somatic *sox102F* transcript is shown in blue (top). Note that the gene is transcribed on the minus strand. Next, Rhi ChIP-Seq signal (red) is superimposed on the input control (dark blue). The following 6 tracks (black) show RNA-Seq signal in Ore. R and *w*^{1118} controls, and in the indicated *qin, cuff, rhi* and *uap56* mutants. The *qin* mutation, which disrupts expression of a Tudor domain protein the localizes to nuage (Zhang et al., 2012), does not increase splicing. By contrast, mutations in *cuff, rhi* and *uap56,* which encode nuclear proteins that localize to clusters, increase in splicing. piRNA expression for each genotype is indicated in the green tracks. The "mappability" of the locus, reflecting the extent of unique sequence, is shown at the bottom of the figure.

Figure S5. Rhi suppresses cluster splicing and only unspliced cluster transcripts are directed for piRNA production. Related to Figure 2 and Figure 3.

(A) Expression of eGFP::Rhi in *rhi* mutants restores insufficient splicing for the Sox102F locus. Data are mean ± standard deviation for 3 independent biological samples.

(B) Spliced transcript abundance from the Sox102F locus, relative to w^{1118}. qRT-PCR data were first normalized to *rp49*. Data are mean ± standard deviation for 3 independent biological samples.

(C) Un-spliced transcript abundance from the Sox102F locus, relative to w^{1118}. qRT-PCR data were first normalized to *rp49*. Data are mean ± standard deviation for 3 independent biological samples.

(D) piRNAs production from Sox102F splicing junction sites in w^{1118} ovaries. All cloned piRNAs are from the unspliced transcripts, none of them are mapped to the spliced junctions. DNA sequences are in black. Splicing donor sites are in orange, and green means splicing accepter sites. Blue stands for sense piRNAs, and antisense piRNAs are in red.

Figure S6. Tethering LacI::Rhi, but not LacI , silences GFP protein expression. Related to Figure 5. Top. Western blot showing LacI (green) and a tubulin control (red) expression in ovaries carrying inducible LacI or LacI::Rhi fusion protein genes and an EGFP reporter, in the absence of Gal4 induction (-Gal4) or the presence of Gal4 induction (+Gal4). Bottom. Parallel blot for EGFP (green) and tubulin control (red). Biological triplicate data are shown. EGFP expression is only silenced when LacI::Rhi is expressed.

Figure S7. Rhi binding and expression of complementary transcripts are needed to induce piRNA production. Related to Figure 6. Length distribution of the small RNAs mapping to the GFP constructs. Un-oxidized RNAs were used to generate libraries for sequencing. Blue, sense piRNAs; red, antisense.

Table S2. High throughput sequencing statistics. Linked to Figures 1-3.

Table S2a. ChIP-Seq samples.

Ovary genotype	Sample type	Total reads	Genome mapping reads	Uniq genome mapping reads	Uniq piRNA cluster-mapping reads	Sequencing type
Oregon R	Input	19,690,683	19,207,681	16,953,468	185,114	Single-end 36
	Rhi-ChIP	19,188,196	16,157,556	12,789,291	648,108	Single-end 36
	Pre-immune-serum ChIP	17,631,445	3,743,201	3,350,336	36,292	Single-end 36
w^{1}	Input-rep1	45,732,160	42,864,288	35,176,654	559,429	Single-end 50
	Rhi-ChIP-rep1	152,966,819	34,182,612	28,787,037	730,674	Single-end 50
	Input-rep2	21,118,300	19,687,920	17,038,394	222,400	Single-end 50
	Rhi-ChIP-rep2	30,298,852	25,665,877	20,886,439	757,626	Single-end 50
armi$^{1/72.1}$	Input-rep1	130,342,282	125,148,797	106,350,666	1,732,302	Single-end 50
	Rhi-ChIP-rep1	111,639,082	64,610,367	53,827,732	1,667,665	Single-end 50
	Input-rep2	22,516,588	21,506,254	18,476,720	322,285	Single-end 50
	Rhi-ChIP-rep2	13,795,364	12,220,648	8,839,898	753,164	Single-end 50

Table S2b. Small RNA-Seq samples.

Small RNA sequencing statistics: analysis of genome matching sequences by reads. “Reads excluding ncRNA” 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.

EGFP Construct Age Oxidization?	LacI/ LacI::Rhi	Total reads	Reads perfectly matching genome	ncRNA reads	Reads excluding ncRNA	miRNA-matching reads	Reads excluding ncRNA & miRNA	23–29 nt small RNA reads			
								Total	Transposon-matching reads		
									Total	Sense	Antisense
LacO-Vasp-EGFP (OVG) 2-4 days Oxidized	LacI-rep1	16,249,348	14,119,765	359,519	13,760,246	121,789	13,638,457	12,358,000	9,594,293	2,682,118	7,169,470
	LacI::Rhi-rep1	11,068,854	9,353,884	266,473	9,087,411	122,909	8,964,502	8,282,858	6,029,337	1,739,662	4,486,354
	LacI-rep2	153,157,842	127,489,611	3,255,606	124,234,005	1,083,941	123,150,064	111,624,430	86,889,772	24,302,289	64,877,410
	LacI::Rhi-rep2	177,409,973	151,678,815	4,367,112	147,311,703	1,929,109	145,382,594	134,368,774	97,914,490	28,401,266	72,641,185
LacO-Vasp-EGFP/EGFP-Vasp-LacO (OVG/GVO) 2-4 days Oxidized	LacI-rep1	46,380,606	40,244,084	822,685	39,421,399	156,822	39,264,577	35,219,730	27,366,611	8,755,773	19,149,566
	LacI::Rhi-rep1	48,735,470	41,682,162	708,211	40,973,951	180,312	40,793,639	37,315,761	28,477,289	9,319,330	19,801,667
	LacI-rep2	38,513,048	33,520,351	679,720	32,840,631	124,096	32,716,535	29,414,978	22,867,873	7,356,176	15,959,693
	LacI::Rhi-rep2	35,355,856	30,303,103	519,668	29,783,435	143,046	29,640,389	26,982,099	20,612,548	6,767,095	14,310,916
LacO-Vasp-EGFP/EGFP-	LacI-rep1	45,272,747	38,325,461	778,436	37,547,025	92,432	37,454,593	35,312,835	27,500,652	8,911,593	19,178,922

Vasp-LacO (OVG/GVO) 14-16 days Oxidized	LacI::Rhi-rep1	43,792,021	37,701,250	800,990	36,900,260	181,314	36,718,946	33,148,612	25,666,838	8,417,116	17,785,189
	LacI-rep2	32,665,847	27,627,049	552,691	27,074,358	63,953	27,010,405	25,459,224	19,823,005	6,466,333	13,782,482
	LacI::Rhi-rep2	41,996,269	35,966,320	747,541	35,218,779	166,390	35,052,389	31,613,148	24,477,416	8,064,835	16,927,858
OVG 2-4 days Un-oxidized	LacI	16,698,738	14,572,404	957,746	13,614,658	7,480,048	6,134,610	5,258,640	4,007,781	1,240,435	2,876,079
	LacI::Rhi	19,658,683	17,218,302	1,471,495	15,746,807	9,037,230	6,709,577	5,979,746	4,371,854	1,352,805	3,164,875
OVG/GVO 2-4 days Un-oxidized	LacI	27,945,613	22,347,290	1,418,311	20,928,979	7,108,237	13,820,742	12,355,771	9,507,077	3,147,024	6,582,276
	LacI::Rhi	25,523,627	20,291,554	2,269,452	18,022,102	5,693,260	12,328,842	11,056,947	8,338,308	2,778,530	5,798,249

Table S2c. RNA-Seq samples.

Ovary genotype	Total reads	Genome-mapping reads	Uniquely mapping reads	Gene mapping reads	Transposon mapping reads	piRNA cluster-mapping reads	Sequencing type
Ore.R.	97,529,610	89,452,521	83,347,362	76,141,253	261,801	128,970	Paired-end 100
w^{1}-rep1	106,023,094	100,624,999	89,461,051	81,431,340	313,153	89,316	Paired-end 100
$rhi^{2/KG}$-rep1	101,764,128	90,224,659	78,162,381	66,926,117	5,206,641	387,375	Paired-end 100
w^{1}-rep2	71,051,534	67,979,397	64,136,391	58,685,981	979,334	129,752	Paired-end 100
$rhi^{2/KG}$-rep2	78,918,984	75,340,847	67,572,190	57,609,542	5,485,277	282,868	Paired-end 100
$cuff^{wm25}$	106,448,028	97,728,390	84,346,886	70,421,237	8,392,499	631,809	Paired-end 100
$uap56^{28/sz}$	108,230,898	100,104,694	78,541,631	70,768,111	2,724,098	141,180	Paired-end 100

Table S3. Synthetic DNA oligonucleotides used in this study (5′-to-3′). Linked to Figure 3.

For quantitative RT-PCR to measure splicing ratio

sox102F-RT	TGT CCA TGA CCA TTT CCT TG
sox102F-Left	TGC AGG TAC AGG GCC TAG TT
sox102F-Right-Spliced	CTT CTA AAA AGT CAT GGG AGA GTG
sox102F-Right-Unspliced	CTT TAA TTT GTT CTA GGG GAG AGT G
42AB-RT	CTG GAA AGG CGC TCC ACT AC
42AB-Left	GCA GTT GCC GTC TCT CCT T
42AB-Right-Spliced	TGG GTC AAA GTG CAG CAG TTT T
42AB-Right-Unspliced	CGG GAA TAT AAT CGC AGC AGTT TT
rp49-RT	CGG GAA TAT AAT CGC AGC AGT TTT
rp49-Left	CCG CTT CAA GGG ACA GTA TCT G
rp49-Right	ATC TCG CCG CAG TAA ACG C
EGFP-RT	TGC TCA GGT AGT GGT TGT CG
EGFP-Right	GAA CTT CAG GGT CAG CTT GC
EGFP-Left-Spliced	ATA TGG TGA GCA AGG GCG A
EGFP-Left-Unspliced	CTC ATC CAC AGG TGA GCA AG

For ChIP-quantitative PCR

42AB	CGT CCC AGC CTA CCT AGT CA; ACT TCC CGG TGA AGA CTC CT

VT-q1	GCG ATA GCA CAA TGG GAA AT; GGC TTG ACA AAC GTA AAA CGA
VT-q2	CAT TTG ATG TGT TAG TGG AAA ACG; GGC AAG CTG TCG ACT TGT G
tubIn	GGC AAG CTG TCG ACT TGT G; AAC AGC TCC TCG CCC TTG
GFP3'	CGA CAA CCA CTA CCT GAG CA; ATC AGC TCG GGA TCT GAG TC
GT3	AAC AGC TCC TCG CCC TTG; CCC ATC GAG CGT TGA AGT
mocs	TCA CTG CGG ATG GAA ACA TA; GGG GAG AGA GTG TGG TGT GT
suUR	TAG CTC GTT GTC CTC GGA GT; CAC CTC AGA ATC GTT GAG CA

Table S4. Published fly alleles used in this study.

$rhi^{2/KG}$	(Klattenhoff et al., 2009)
$cuff^{wm25}$	(Pane et al., 2011)
$uap56^{28/sz}$	(Zhang et al., 2012a)
qin^{kumo}/Df	Zhang et al., *EMBO J.* In press; (Anand and Kai, 2012)
$armi^{1/72.1}$	(Cook et al., 2004)

Table S5. Published high-throughput sequencing data used in this study.

$rhi^{2/KG}$-small RNA	SRP002060; (Klattenhoff et al., 2009)
$armi^{1/72.1}$-small RNA	GSE15186; (Malone et al., 2009)
$uap56^{28/sz}$-small RNA	GSE35638; (Zhang et al., 2012a)
$cuff^{wm25}$-small RNA	(Pane et al., 2011)
Ore.R.-small RNA	SRP000458; (Li et al., 2009)
W^{1118}-RNA-Seq	SRP024291; Zhang et al.; *EMBO J.* In press

qin^{kumo}/*Df*-RNA-Seq	SRP024291; Zhang et al.; *EMBO J.* In press

Table S6. Antibody information.

anti-Rhi	Custom made, guinea pig 1943; 25 µl/ChIP
anti-GFP	Clontech, cat# 632460; Western blotting: 1:2000 dilution
anti-LacI	Rockland, cat# 600-401-B04; Western blotting: 1:2000 dilution
anti-LacI	US Biological, cat# L0899; Immunostaining: 1:500 dilution
anti-Tubulin	DSHB, cat# 12G10; Western blotting: 1:5000 dilution

SUPPLEMENTAL EXPERIMENT PROCEDURES

Transgenic flies for tethering Rhi to the GFP locus.

Transgenes expressing LacI or LacI::Rhi fusion were made as follows: the 1.1 kb lacI binding domain from lacI-HP1 in pCas-hs-act, provided by L. L. Wallrath (Li et al., 2003), was PCR amplified (Forward primer: AAA GAA TTC GCC ATG GTG AAA CCA GTA ACG; Reverse primer: AAA GGA TCC AAC CTT CCT CTT CAT C), and the 1.4 kb *rhi* coding sequence from the full length rhino cDNA clone RE36324 (Klattenhoff et al., 2009) was PCR amplified (Forward primer: AAA GGA TCC GTT ATG TCT CGC AAC CAT CAG; Reverse primer: AAA CGC GCC GCT TTA CTT GGG CAC AAT GAT). Rhi PCR product was digested with BamHI/NotI and cloned into pBstII KS+ to generate pBst-Rhi. The LacI PCR product was digested with EcoRI and BamHI and cloned into pBst-Rhi to generate pBst-lacI::Rhi. The entire insert in pBst-lacI::Rhi was cut out with KpnI/NotI and cloned into the transformation vector pUASp to generate pUASp-lacI::Rhi. pUASp-lacI::Rhi was cut with BamHI to remove the *rhi* cDNA and re-circularized to yield pUASp-lacI. All intermediates and final plasmid clones were verified by direct sequencing. The final pUASp-lacI and pUASp-lacI::Rhi constructs were used to make germline transgenes using standard protocols.

To generate transgenic flies carrying the LacO-EGFP, XbaI was used to partially digest pSV2-dhfr-8.32 provided by A. S. Belmont (Robinett et al., 1996). A 1.2 kb XbaI fragment corresponding to 32 repeats of the 36bp Lac operon was cloned into pBstII KS+ to produce pBst-32mer. This clone was used to provide the LacO repeats that were subsequently cloned into unique restriction sites upstream of a truncated *vasa* promoter driving expression of the 84B alpha tublin 5'UTR and first intron followed by EGFP-NLS and the tublin-3'UTR. The repeats were excised from pBst with SmaI/NotI, the ends were polished with Klenow Large

Fragment and the DNA was cloned into either the 5' Not/blunted site or the 3' Bam/blunted site of the EGFP reporter construct. The following primers were used to amplify the partial *vasa* promoter from vasp-EGFP (Forward: AAA GGA TCC ATA TGA ATG AAT CAC TTA GG; Reverse: AAA GGA TCC GTG GAA TTT CCC ATT GTG C). This product was cut with BamHI and cloned into the unique BamHI site at the 3' end of the reporter construct in the anti-sense direction to get the GFP-vaspAS-LacO construct. Due to the instability of the lacO repeats, all clones containing these repeat sequences were transformed into Max Efficiency Stbl2 chemically competent cells (Invitrogen cat#10268-019). Otherwise, DH5α electrocompetent cells (home made) were used. All reporter constructs contain the attB site and were integrated onto the attP2 site located at chromosome 3L-68A4.

Transgenic flies for rescuing the splicing.

For *rhi* transgene rescue, *rhi* cDNA was cloned into *rhi*-promoter-pPGW vector having attB site. The transgenic flies were produced by site specific insertion at chromosomal location 3L-68A4 by PhiC31 mediated transformation. Thus, N terminal eGFP tagged Rhi was expressed under endogenous *rhi* promoter. Expression of eGFP::Rhi in *rhi* mutant background leads to a reduction of high splicing efficiency observed in *rhi* mutants and the spliced/unspliced ratio is comparable to wild type control, indicating that the splicing phenotype is specific for *rhi* and not due to second site mutation.

ChIP-Seq and data analysis.

For Rhi ChIP-Seq, 1 ml ovaries were first crosslinked with 2% formaldehyde for 10 minutes in Robb's medium (100 mM HEPES pH 7.4, 55 mM sodium acetate, 40 mM potassium acetate, 100 mM sucrose, 10 mM glucose, 1.2 mM $MgCl_2$, 1 mM $CaCl_2$). Then the reaction was quenched by adding Glycine to 120 mM and

incubating for 5 minutes with rotation. The ovaries were then washed twice with TBS (50 mM Tris-HCl pH 7.5, 150 mM NaCl), and twice with ChIP lysis buffer (50 mM Hepes/KOH pH 7.5, 140 mM NaCl, 1% [v/v] Triton X-100, 0.1% [w/v] Na-Deoxycholate, 0.1% [w/v] SDS). Washed ovaries were sonicated for 4 ×15 minutes in sonication buffer (1% SDS, 10 mM EDTA, 50 mM Tris-HCl pH 8.0, 1 Proteinase Inhibitor tablet freshly added) with a Bioruptor Standard (diagenode cat# B01010001). The ovary lysate was centrifuged at 14000 rpm at 4˚C for 15 minutes, 200 µl supernatant was saved as the input control and the remaining supernatant was diluted 7 fold with dilution buffer (20 mM Tris-HCl, 167 mM NaCl, 1.2 mM EDTA, 0.01% [w/v] SDS, 1% [v/v] Triton X-100, 1 Proteinase Inhibitor tablet freshly added). For each ChIP-Seq library, 25 µl anti-Rhi antibody or Pre-Immune Serum(custom made, guinea pig 1943) was conjugated to 400 µl Dynabeads Protein A (Life technologies, cat # 10001D). The diluted supernatant added to the conjugated beads and incubated was at 4˚C overnight. The beads were then wash two times each with 1 ml Wash buffer A (20 mM Tris-HCl pH 8.0, 2 mM EDTA, 0.1% [w/v] SDS, 1%[v/v] Triton X-100, 150 mM NaCl), 1 ml Wash buffer B (20 mM Tris-HCl pH 8.0, 2 mM EDTA, 0.1% [w/v] SDS, 1%[v/v] Triton X-100, 500 mM NaCl), 1 ml Wash buffer C (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 1% [v/v] NP-40, 1% [w/v] Na-deoxycolate, 0.25 M LiCl) and 1 ml Wash buffer D (10 mM Tris-HCl pH 8.0, 1 mM EDTA). The beads were then resuspended with 200 µl sonication buffer, an the saved input sample was thawed and processed in parallel with the ChIP sample, as follows: Crosslinking was reversed by adding 2 µl 5 M NaCl to the beads/input and incubating at 65˚C for 6 hours. Then 200 µl Tris-HCl buffer (10 mM Tris-HCl pH 8.5) was added to each sample. To remove RNA, 6 µl 30 mg/ml RNaseA was added and incubated at 37˚C for 2 hours. To digest protein, 20 µl 20 mg/ml Proteinase K was added and the mixture was incubated at 55 ˚C for 2 hours. Finally, phenol:chloroform extraction was

used to purify the immunoprecipitated DNA, which was dissolved in 34 µl water. The sequencing library was constructed by sequentially performing end-repair, A-tailing, Y-shaped adapter ligation and PCR amplification as described (Zhang et al., 2012b). The libraries from Oregon R. ovaries were prepared with illumina Paired End DNA oligos and sequenced by illumina GAII. The libraries for w^{1}-rep1 and $armi^{1/72.1}$-rep1 were made in parallel with illumina Paired End DNA oligos and sequenced by illumina HiSeq. The libraries for w^{1}-rep2 and $armi^{1/72.1}$-rep2 were made in parallel with illumina Multiplexing oligos and sequenced by HiSeq.

The sequencing reads were mapped to the *Drosophila melanogaster* genome (FlyBase r5.45/dm3) using bwa-0.6.1 (Li and Durbin, 2009). All libraries were normalized to sequencing depth, using total genome mapping reads. For each library, bigwig files were generated for UCSC browser visualization. To calculate the Rhi binding enrichment for the piRNA cluster regions, only the reads that uniquely mapped to one genome position were used. The mean ppm value over each cluster was calculated by bigWigAverageOverBed, with a pseudo count of 0.01.

SUPPLEMENTAL REFERENCES

Anand, A., and Kai, T. (2012). The tudor domain protein kumo is required to assemble the nuage and to generate germline piRNAs in Drosophila. The EMBO journal *31*, 870-882.

Cook, H. A., Koppetsch, B. S., Wu, J., and Theurkauf, W. E. (2004). The Drosophila SDE3 homolog armitage is required for oskar mRNA silencing and embryonic axis specification. Cell *116*, 817-829.

Klattenhoff, C., Xi, H., Li, C., Lee, S., Xu, J., Khurana, J. S., Zhang, F., Schultz, N., Koppetsch, B. S., Nowosielska, A., Seitz, H., Zamore, P. D., Weng, Z., and Theurkauf, W. E. (2009). The Drosophila HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. Cell *138*, 1137-1149.

Li, C., Vagin, V. V., Lee, S., Xu, J., Ma, S., Xi, H., Seitz, H., Horwich, M. D., Syrzycka, M., Honda, B. M., Kittler, E. L. W., Zapp, M. L., Klattenhoff, C., Schulz, N., Theurkauf, W. E., Weng, Z., and Zamore, P. D. (2009). Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. Cell *137*, 509-521.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754-1760.

Li, Y., Danzer, J. R., Alvarez, P., Belmont, A. S., and Wallrath, L. L. (2003). Effects of tethering HP1 to euchromatic regions of the Drosophila genome. Development *130*, 1817-1824.

Malone, C. D., Brennecke, J., Dus, M., Stark, A., McCombie, W. R., Sachidanandam, R., and Hannon, G. J. (2009). Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. Cell *137*, 522-535.

Pane, A., Jiang, P., Zhao, D. Y., Singh, M., and Schüpbach, T. (2011). The Cutoff protein regulates piRNA cluster expression and piRNA production in the Drosophila germline. The EMBO journal *30*, 4601-4615.

Robinett, C. C., Straight, A., Li, G., Willhelm, C., Sudlow, G., Murray, A., and Belmont, A. S. (1996). In vivo localization of DNA sequences and visualization of large-scale chromatin organization using lac operator/repressor recognition. J Cell Biol *135*, 1685-1700.

Zhang, F., Wang, J., Xu, J., Zhang, Z., Koppetsch, B. S., Schultz, N., Vreven, T., Meignin, C., Davis, I., Zamore, P. D., Weng, Z., and Theurkauf, W. E. (2012a). UAP56 Couples piRNA Clusters to the Perinuclear Transposon Silencing Machinery. Cell *151*, 871-884.

Zhang, Z., Theurkauf, W. E., Weng, Z., and Zamore, P. D. (2012b). Strand-specific libraries for high throughput RNA sequencing (RNA-Seq) prepared without poly(A) selection. Silence *3*, 9.