Supplementary figures and legends for Chen et al. 2021

Figure S1. Analysis pipeline of gonad small RNAs.

Flow chart showing step-wise isolation of piRNAs from total small RNAs and subsequent mappings to different annotations (repeats, protein-coding genes and genome).

Figure S2. Profiles of complex satellite- and TE-mapping piRNAs in two sexes.

(A) Coverage plots of piRNAs over *Rsp* (top), *HETRP* (middle) and *SAR* (bottom), in testis (left) and ovary (right).

(B) Distributions of 3'-to-5' piRNA distance on the same strand for each of the three complex satellites in testis and ovary. Z_1 score marks the statistical significance of a putative enrichment of 1nt (P<0.05 for z>1.96). Significant ones are marked pink with the rest in gray.

(C) Examples of complementary pairs of *Rsp*-mapping piRNAs. Note that in ovary (red) they show an enrichment for 10-nt overlap, i.e., ping-pong signature, but in testis (blue) they show near perfect-complementarity with no evidence for ping-pong signature.

(D) Top 10 TEs targeted by the most abundant antisense piRNAs in testis (left) and ovary (right). Heights of slices correspond to relative abundance in each sex, and the sum of top 10 TEs is then scaled to the same height between sexes. Each TE family is given a unique color, and the same TE family is connected by a line to help visualize distinct rank-orders between sexes. Names of TE families are shown following the same order, though not directly next to respective slices.

Figure S3. TE levels in piRNA pathway mutants and curation of TEs regulated by Rhi in at least one sex.

(A) Bar graphs showing TE levels in piRNA pathway mutant (*rhi*) testes (orange) and ovaries (blue). TEs that have >10 RPKM in either sex is shown at the top, with the rest at the bottom.

(B) Top 10 most expressed TE families in piRNA pathway mutant testis (left) and ovary (right). $rhi^{-/-}$ was used, where piRNA production from genome-wide dual-strand piRNA clusters collapses. Slice heights and colors were depicted as described in Supplementary Fig. S2D, though the same TE can be marked by a different color.

(C) Table reporting manual curation of 36 confidently affected TE families by *rhi*$^{-/-}$. Silencing potential is TRUE when there are normally >100 RPM antisense piRNAs and they show >2-fold reduction in *rhi* mutants. TEs are deemed de-repressed when having >3-fold up-regulation. Note a few unexpected cases where TE de-repression is not accompanied by piRNA loss, the ovary ones of which were described before (Klattenhoff et al. 2009).

Figure S4. An algorithm that includes local repeats in piRNA cluster definition and analysis.

(A) Flow chart showing steps of the new algorithm that includes local repeats in piRNA cluster definition and analysis. See also methods.

(B) Histogram showing the distribution of "max distances" defined in (A) to identify a meaningful cutoff (2Mb) for distinguishing local from non-local repeats. See also methods.

Figure S5. Characterization of *pira* homology in *D. mel,* and the phasing pattern of *pira*-derived sense piRNAs.

(A) Homology between two *D. melanogaster* paralogs: *velo* and *CG12717/pira*. The homologous regions are marked using BLAT and they share 75% nucleotide sequence identity.

(B) Alignment of duplicated, partial copies of *CG12717* at *petrel* on *D. melanogaster* Y to its *CG12717* gene (left), and their genomic coordinates (right). Note that there are two small regions of *CG12717* absent on Y. RNA *in situ* HCR was targeted against the ORF region unique to *CG12717* gene.

(C) Distribution of 3'-to-5' piRNA distance on the coding strand of *CG12717/pira* in testis. Only genome-unique sense piRNAs with zero mismatch are used. Z_1=3.2 is equivalent to P< 0.001.

Table S1. Genome-wide piRNA clusters in testis and ovary as well as major piRNA clusters defined in this study.

Figure S1. Chen et al.

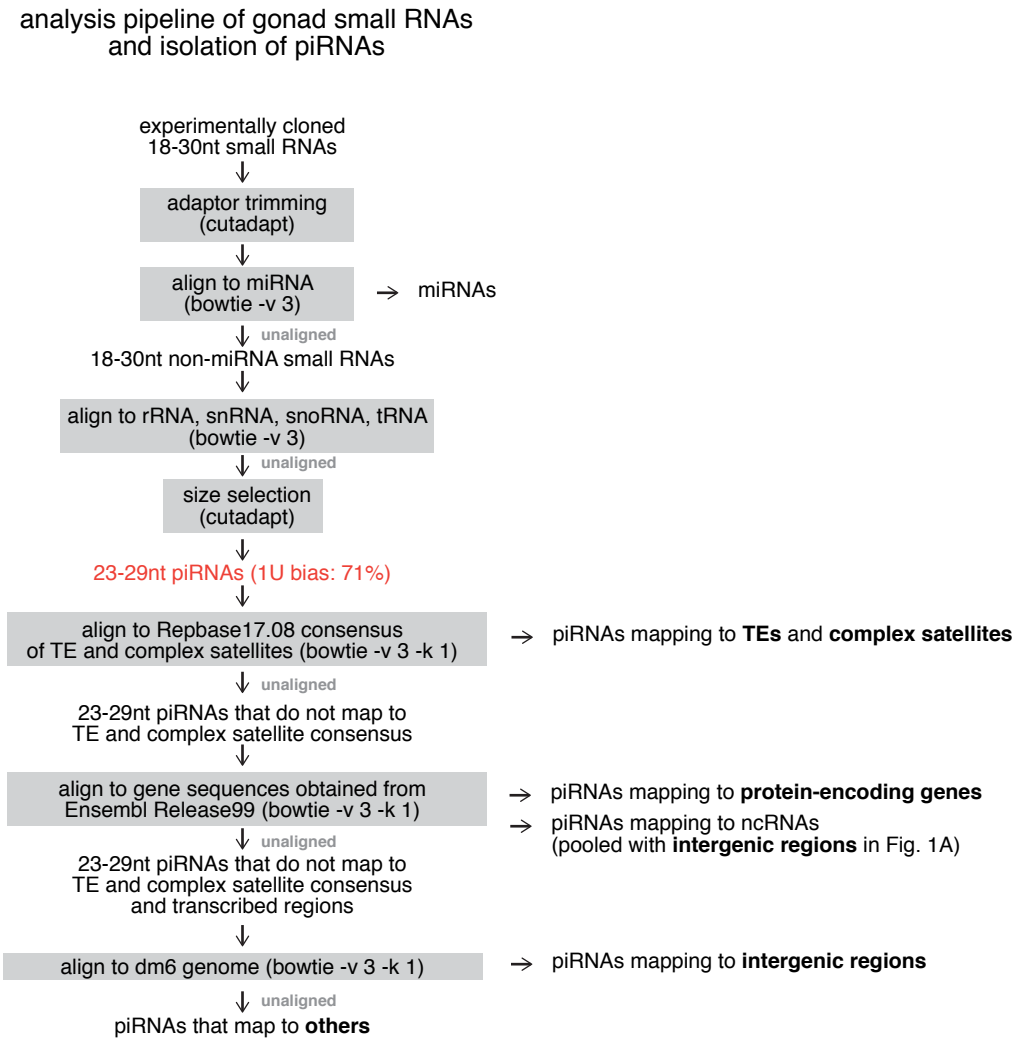
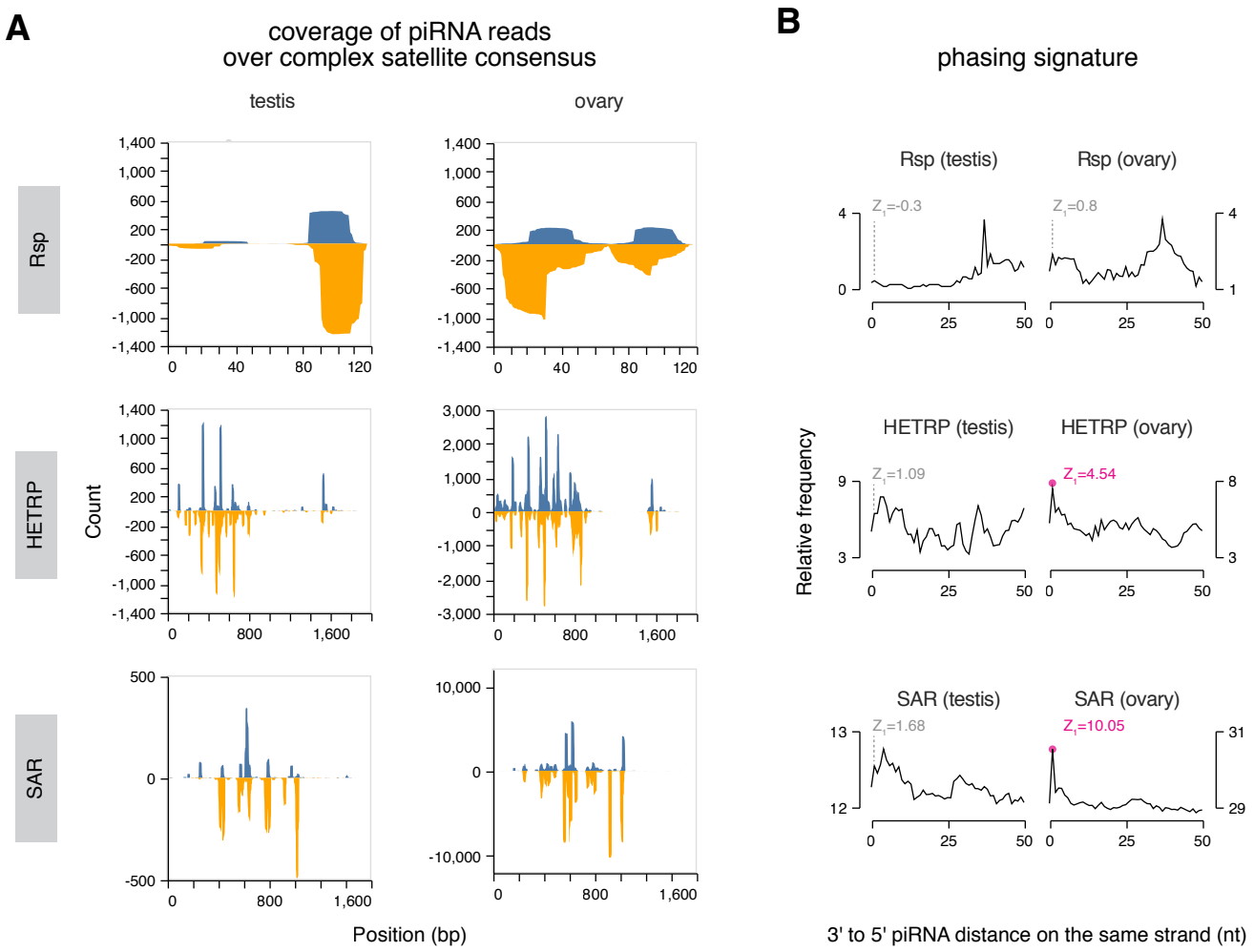


Figure S2. Chen et al.



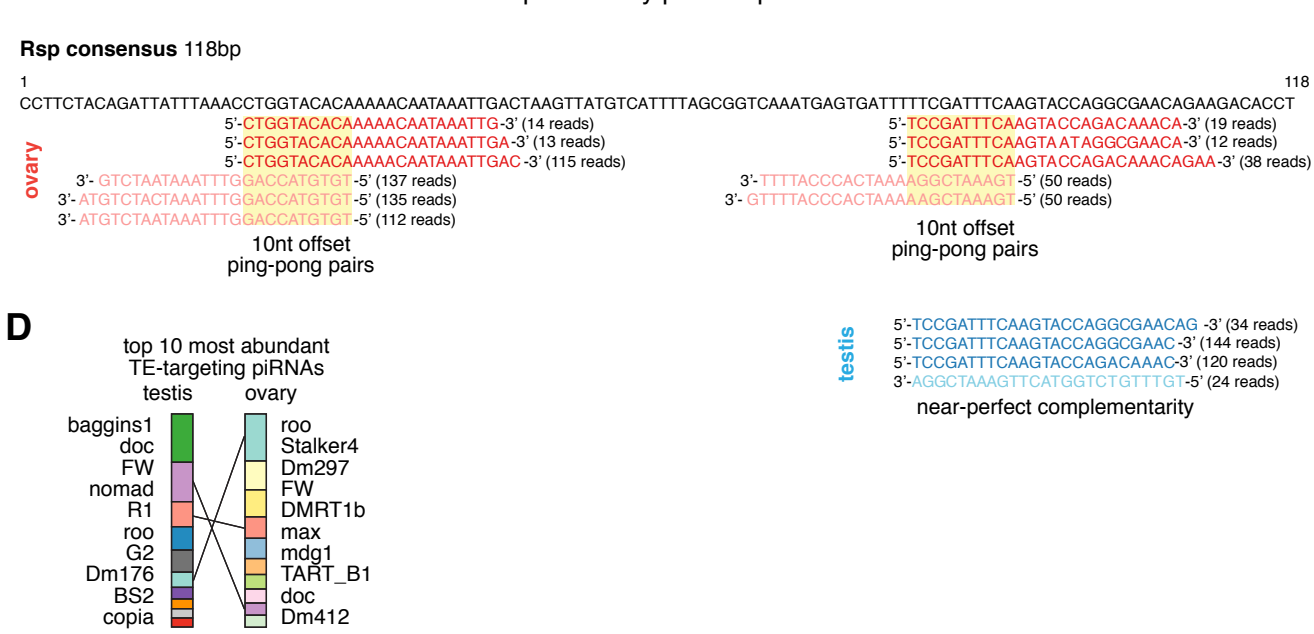
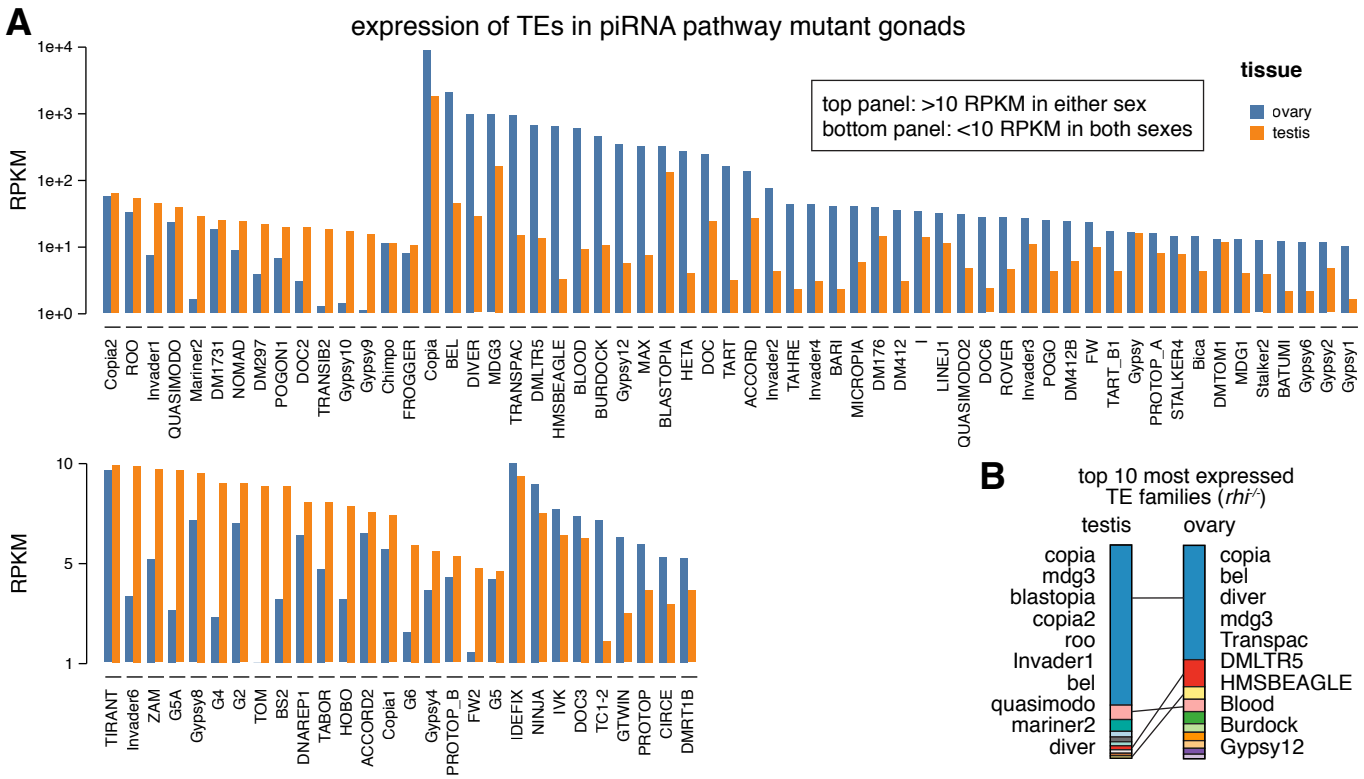


Figure S3. Chen et al.



C

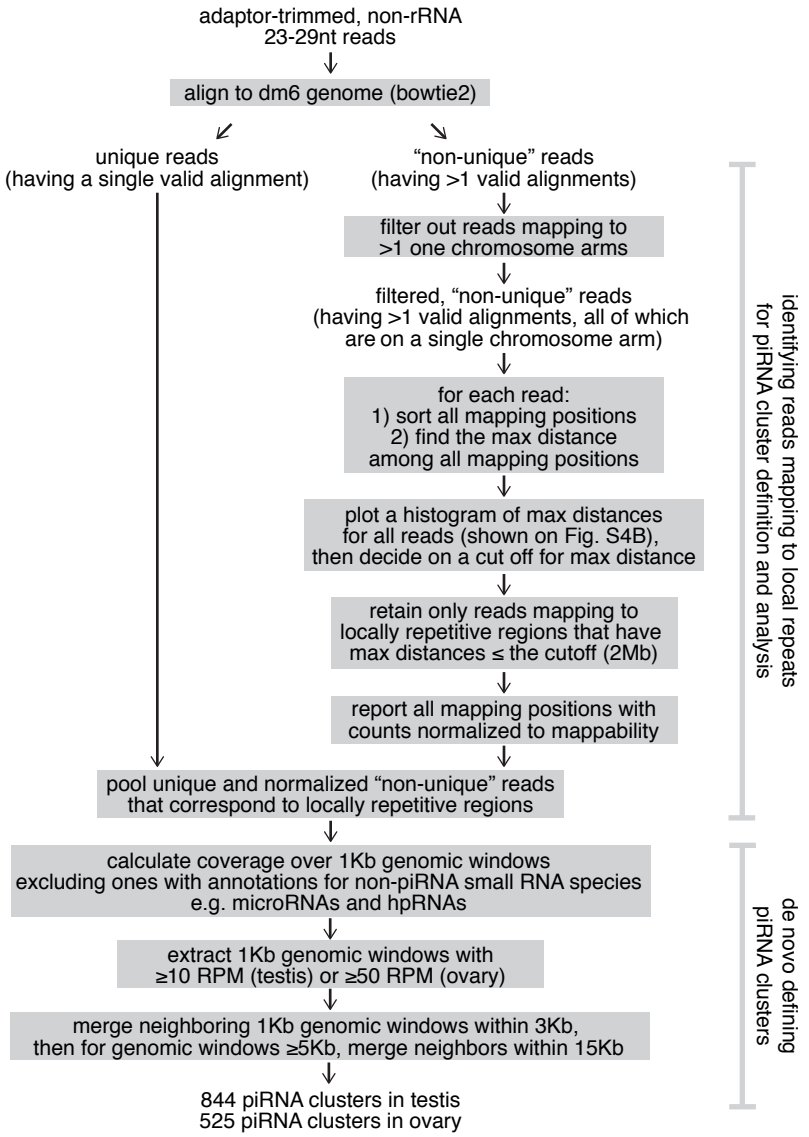
profiles of 36 TEs regulated by Rhi in at least one sex

TE_merge	ovary_TE_fc	ovary_silencing potential	ovary_pi	ovary_pi_FC	testis_TE_fc	testis_silencing potential	testis_pi	testis_pi_FC	Derepressed in	CAN EXPLAIN	Note
MDG3	382.2	TRUE	2098.3	4.3	33.2	TRUE	2944.7	8	Both sex	YES	
ACCORD	109.6	TRUE	1282.2	15.4	14.4	TRUE	496.8	9.6	Both sex	YES	
Copia	114.4	TRUE	1110.3	4.4	12.5	TRUE	5141.1	26.4	Both sex	YES	
Invader3	9.7	TRUE	731.9	20.4	12.1	TRUE	4857.3	81.8	Both sex	YES	
I	22.2	TRUE	2632.6	9.3	7.9	TRUE	2586.3	19.6	Both sex	YES	
DIVER	70.7	TRUE	1563.2	9.2	3.6	TRUE	1985.5	2.9	Both sex	YES	
MICROPIA	100	TRUE	1357.9	25.7	3.5	TRUE	2531.6	124	Both sex	YES	
BS2	2.9	TRUE	7253.5	39.6	7.8	TRUE	5274.9	105.7	Testis	YES	testis biased TE
TRANSIB2	1.4	TRUE	1062.1	20.1	4.2	TRUE	3090.7	161.3	Testis	YES	testis biased TE
DOC2	1.9	TRUE	2090.2	13	3	TRUE	725.3	3.6	Testis	YES	testis biased TE
DM297	0.9	FALSE	27748.2	0.3	4.8	FALSE	3626.7	0.7	Testis	?	TE derepression w/o tissue-wide piRNA loss
TOM	1	FALSE	137.2	0.7	3.8	FALSE	28.7	1.3	Testis	?	TE derepression w/o tissue-wide piRNA loss
ROO	1.6	FALSE	48306.1	0.5	3.1	FALSE	8654.7	0.8	Testis	?	TE derepression w/o tissue-wide piRNA loss
Gypsy12	389.1	TRUE	2051.2	29.3	2.5	TRUE	854.4	3.6	Ovary	YES	ovary biased TE
BLASTOPIA	109.2	TRUE	2319.5	7.7	2.1	TRUE	3376	16.5	Ovary	YES	ovary biased TE
MAX	93.7	TRUE	16354	38.9	1.2	TRUE	1505.1	20.5	Ovary	YES	ovary biased TE
Copia2	42.8	TRUE	1856.3	25.6	1.8	TRUE	1648.4	23	Ovary	YES	ovary biased TE
DM1731	40.3	TRUE	1047.1	23.9	2.1	TRUE	713.9	3.4	Ovary	YES	ovary biased TE
DMLTR5	27	TRUE	108.3	2.1	1.5	TRUE	377.9	12.8	Ovary	YES	ovary biased TE
Invader1	5.2	TRUE	2021.1	32.1	1.6	TRUE	872.7	41.5	Ovary	YES	ovary biased TE
BURDOCK	317	TRUE	3755.7	47.1	2	FALSE	69.9	2	Ovary	YES	ovary biased TE
HETA	238.5	TRUE	5725.4	30.9	0.4	FALSE	53.3	0.3	Ovary	YES	ovary biased TE
TRANSPAC	192.6	TRUE	751.1	6.3	0.9	FALSE	43.6	2	Ovary	YES	ovary biased TE
BEL	130.6	TRUE	2245.7	8.7	0.6	FALSE	145.4	0.5	Ovary	YES	ovary biased TE
TART	80.1	TRUE	4354.2	2.6	0.3	FALSE	2756.7	0.6	Ovary	YES	ovary biased TE
HMSBEAGLE	69.2	TRUE	2461	4.2	2.3	FALSE	1871.8	1.7	Ovary	YES	ovary biased TE
TART_B1	23.2	TRUE	14384.2	15	0.8	FALSE	2709.2	1.7	Ovary	YES	ovary biased TE
LINEJ1	20.5	TRUE	2184.4	17	1.9	FALSE	40.4	1	Ovary	YES	ovary biased TE
Invader2	19.6	TRUE	2859.8	7.1	2.1	FALSE	4728.2	1.2	Ovary	YES	ovary biased TE
TAHRE	18.9	TRUE	5589	53.1	0.5	FALSE	74.4	0.5	Ovary	YES	ovary biased TE
DOC	9.3	TRUE	12492.9	14.8	2.6	FALSE	22434.2	2	Ovary	YES	ovary biased TE
QUASIMODO	8.2	TRUE	5956.2	2.2	2.1	FALSE	2290	1.4	Ovary	YES	ovary biased TE
FW	6.7	TRUE	22035	6.5	2.2	FALSE	14250.2	1.4	Ovary	YES	ovary biased TE
DOC6	4.4	TRUE	2051.2	9.3	2.5	FALSE	69.3	1.8	Ovary	YES	ovary biased TE
QUASIMODO2	28.6	FALSE	1169.4	1.3	1.9	TRUE	351.7	5.4	Ovary	?	TE derepression w/o tissue-wide piRNA loss
BLOOD	50.9	FALSE	5419.9	1	1.5	FALSE	3524.8	0.9	Ovary	?	TE derepression w/o tissue-wide piRNA loss

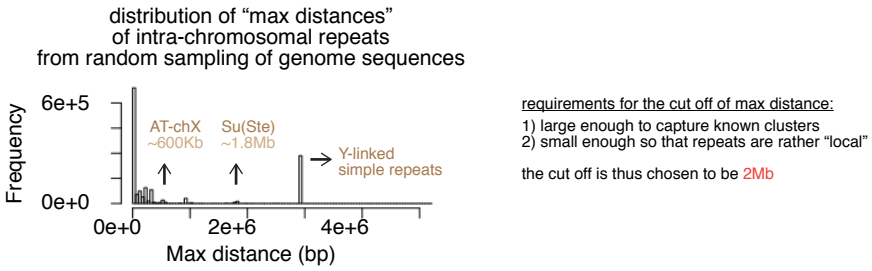
"silencing potential" is TRUE when 1) there is >100 RPM piRNAs normally AND 2) there is >2-fold piRNA loss in rhi mutants.

Figure S4. Chen et al.

A



B



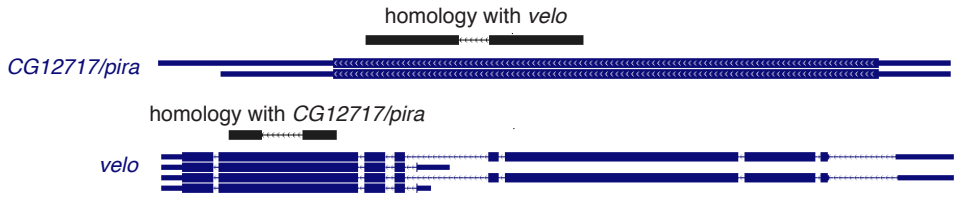
requirements for the cut off of max distance:

1) large enough to capture known clusters
2) small enough so that repeats are rather "local"

the cut off is thus chosen to be 2Mb

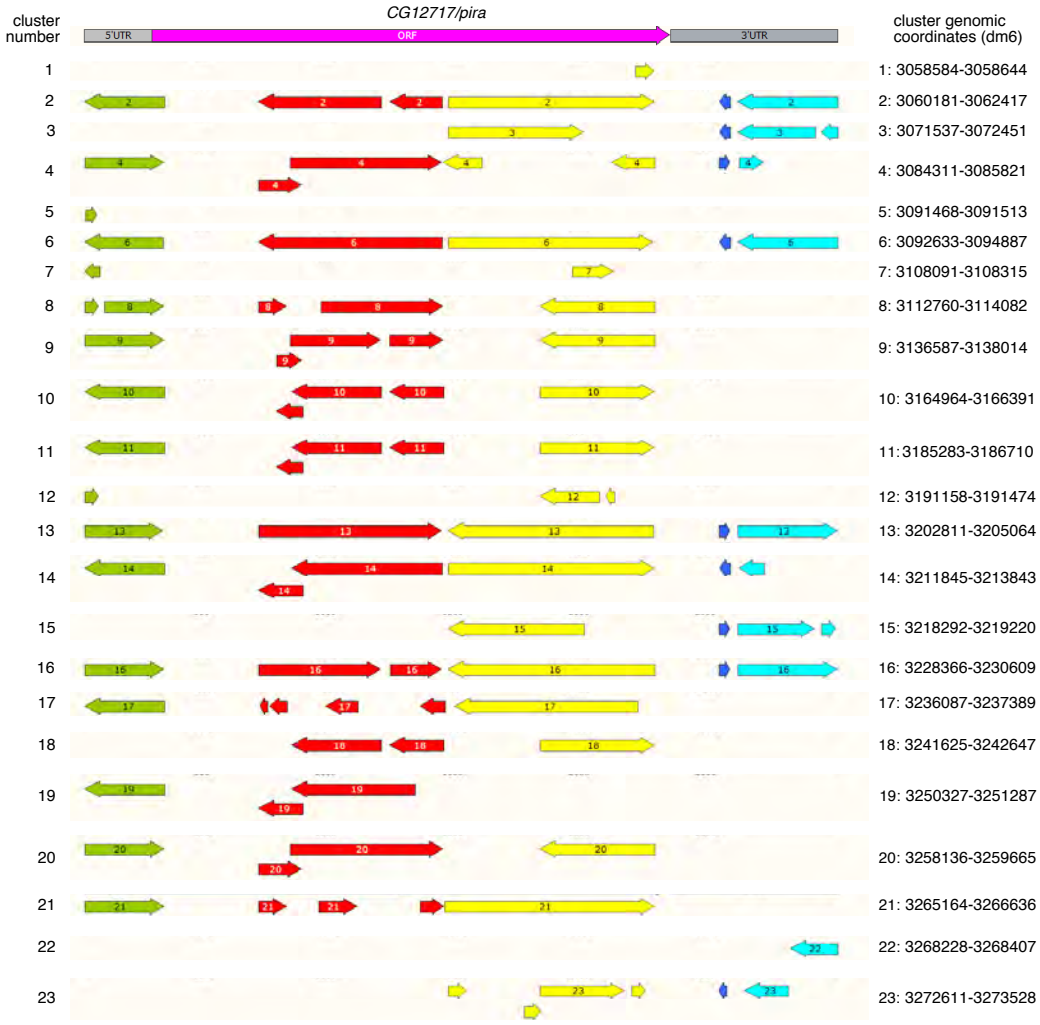
Figure S5. Chen et al.

A



B

alignment of *CG12717/pira*-related fragments at *petrel* locus along the *CG12717/pira* gene



C