Supplementary information

Materials and Methods
Reference
Supplementary Figures S1~S5

Materials and Methods

Pull-down assay

AGO3-2-sDMA (without R72-sDMA) and Aub-sDMA (without R72-sDMA) peptides were synthesized by Operon.

Western blot analysis

Western blot analysis was performed as described previously (Miyoshi *et al*, 2005). Anti-Spn-E monoclonal antibody was used at 1:2,000 dilution. Culture supernatant of hybridomas producing anti-Tud monoclonal antibody was used without dilution as a primary antibody.

RNAi

RNAi in S2 cells was performed as described (Ishizuka *et al*, 2002). The primers utilized to produce dsRNAs were;

T7-spnE3501: 5'-TAATACGACTCACTATAGGGAGTGGTCTAGTCAAAAGCTG-3' and

T7-spnE2848:

5'-TAATACGACTCACTATAGGGAAACGCGATAGTTATTTCCA-3'.

TnT

The cDNA encoding the ORF of dTdrd1 (nucleotide number; 4-4809) was amplified by PCR using primers;

BamHI-dTdrd1:

5'-GGGGGGATCCGCTATCTTCGTCCAGCTAACCACAGATAAT-3' and

SacI-dTdrd1:

5'-GGGGGAGCTCTTATGATTTTTTCAGGAACATCCAATTTTC-3'. The resultant DNA fragment was inserted into pET-28a (Novagen) to produce pET-dTdrd1. ^{35}S-methonine labeled dTdrd1 protein was produced using the TNT Coupled Reticulocyte Lysate Systems (Promega), according to the manufacturer's instructions.

Sucrose gradient

Ovary lysates prepared as described in the Materials and Methods section were resolved on a linear sucrose gradient (5 to 30%). Besides sucrose, Gradient buffer contained 30 mM HEPES-KOH (pH 7.3), 100 mM KOAc, 4 mM MgOAc, 1 mM DTT, 2 mg/mL Pepstatin, 2 mg/mL Leupeptin, and 0.5% Aprotinin. The gradient was centrifuged at 4°C in a Beckman MLS-50 rotor at 23,000 rpm for 15 hr and 10 min. Following centrifugation, fractions were collected (200 μl each) and proteins in the fractions were TCA-precipitated and subjected to western blot analyses using anti-Tudor, anti-AGO3 and anti-Aub antibodies.

RT-PCR

Total RNAs were isolated from the immunopurified complexes by phenol: chloroform extraction and precipitated with ethanol. Reverse transcription (RT) was performed using Accuscript RT (Stratagene) according to the manufacturer's protocol. Primers used were;

roo-primer-1761: 5'-AATAAAATTGAATTTTTATGGCATAAAATA-3'

roo1960-rev: 5'-GGTAACTCCTCCGCCTTAAC-3'.

PCR was performed using KOD plus (TOYOBO) with the primers for RT and primers indicated below;

roo2060-rev: 5'-ATCCTCTCTCACATGTATGA-3'

roo2160-rev: 5'-CATAATTTTCTCCAGAGCCG-3'

roo2260-rev: 5'-TCTAAATTTAAATTTGGCAT-3'

roo-primer-1791: 5'-GATAACCTGATTGAACAGGTGAATAGTCGT-3'.

Deep-sequencing

Small RNA libraries were prepared as described previously (Morin *et al*, 2008). Deep-sequencing was performed on an Illumina Genome Analyzer at Hokkaido system

science. For bioinformatics analyses (Figure S5), we randomly picked up 30,000 read sequences from each of *tud* heterozygous and homozygous mutant libraries.

Cloning frequency and strand bias

Short reads ranging from 20 to 30 nt were extracted from original sequencing results of *tud* heterozygous and homozygous mutants, respectively. The extracted short reads were mapped to the fruit fly genome sequence (dm3) without allowing any base mismatches. We obtained 27,197 and 28,721 perfectly matched reads from *tud* heterozygous and homozygous mutants, respectively. By comparing the mapped reads to RepeatMasker tracks of the UCSC Genome Browser (http://genome.ucsc.edu/), 19,738 and 21,716 reads were found to be transcripts of transposons. The transposon-derived short reads were annotated and assigned to transposon families. Per family population was computed for each transposon family. The population was normalized to a percentile over the total transposon-derived reads. We visualized the percentile population of transposon families as a heatmap image. For visual simplicity, minor families which appear less than 1% of the total population are not shown. For each transposon family, a ratio between sense and antisense reads was computed using the following equation: (sense reads − antisense reads)/total reads, which ranges from -1.0 to 1.0. The ratio represents strand bias of transposon-derived short reads for each transposon family. We visualized the ratio per transposon family as a heatmap image by using a two-color microarray-like color scheme. An antisense-biased family is rendered as a red-colored box whereas a sense-biased family is rendered green.

Population pie-chart

The mapped short reads were annotated by comparison to the following data tracks of the UCSC Genome Browser, where various gene classes are represented: flyBaseNoncoding (for tRNA, rRNA, snoRNA, snRNA, miRNA, and other ncRNA), rmsk, and simpleRepeat. After the annotation, the population of each gene class was computed and visualized as a pie-chart.

Reference

Morin RD, O'Connor MD, Griffith M, Kuchenbauer F, Delaney A, Prabhu AL, Zhao Y, McDonald H, Zeng T, Hirst M, Eaves CJ, Marra MA. 2008. Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome Res.* **18**: 610-621
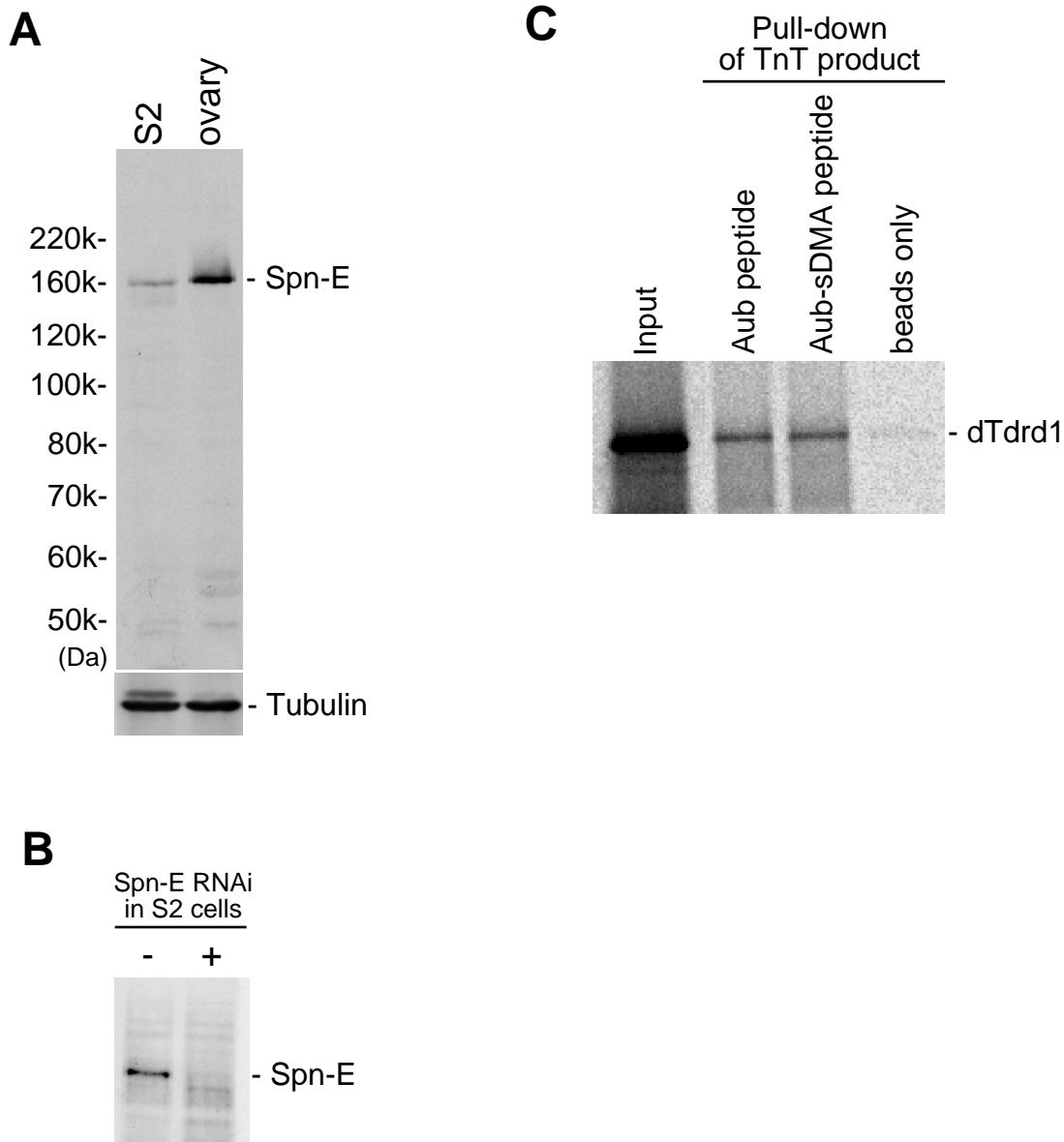
A

S2   ovary

220k-
160k- - Spn-E
120k-
100k-
80k-
70k-
60k-
50k-
(Da)

- Tubulin

C

Pull-down
of TnT product

Input   Aub peptide   Aub-sDMA peptide   beads only

- dTdrd1

B

Spn-E RNAi
in S2 cells

-   +

- Spn-E

Figure S1. (**A**) Western blot analysis of S2 cells and ovary lysates using anti-Spn-E antibodies raised in our laboratory. Tubulin was detected as an internal control. (**B**) Western blot analysis of S2 cells, with and without RNAi treatment against the *spn-E* gene, shows the specificity of the anti-Spn-E antibodies utilized in (A). (**C**) dTdrd1 does not show a specific interaction with the peptides. 35S-labeled dTdrd1 was produced in rabbit reticulocyte lysates and utilized in pull-down assays using Aub and Aub-sDMA peptides (Figure 2A). The signals were slightly higher in both Aub and Aub-sDMA peptide lanes compared with the negative control lane (beads only), indicating that dTdrd1 might be able to bind weakly with the peptides; however, the sDMA-dependency was not observed. Input shows 10% of dTdrd1 used for each pull-down reaction.
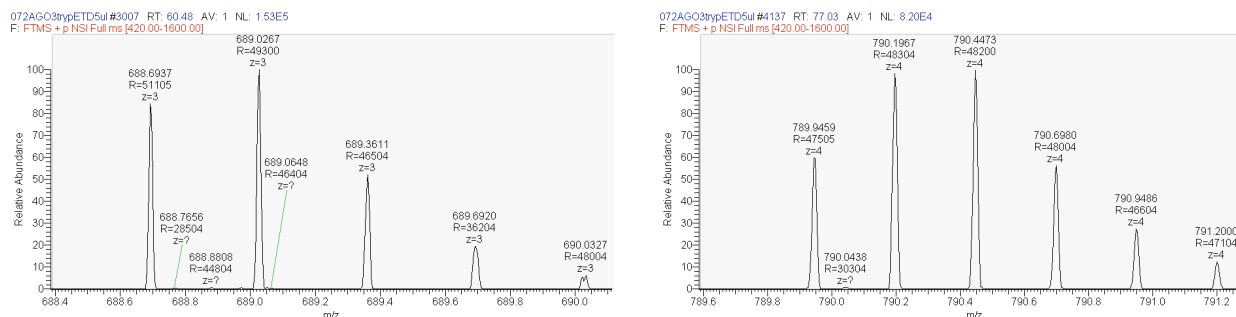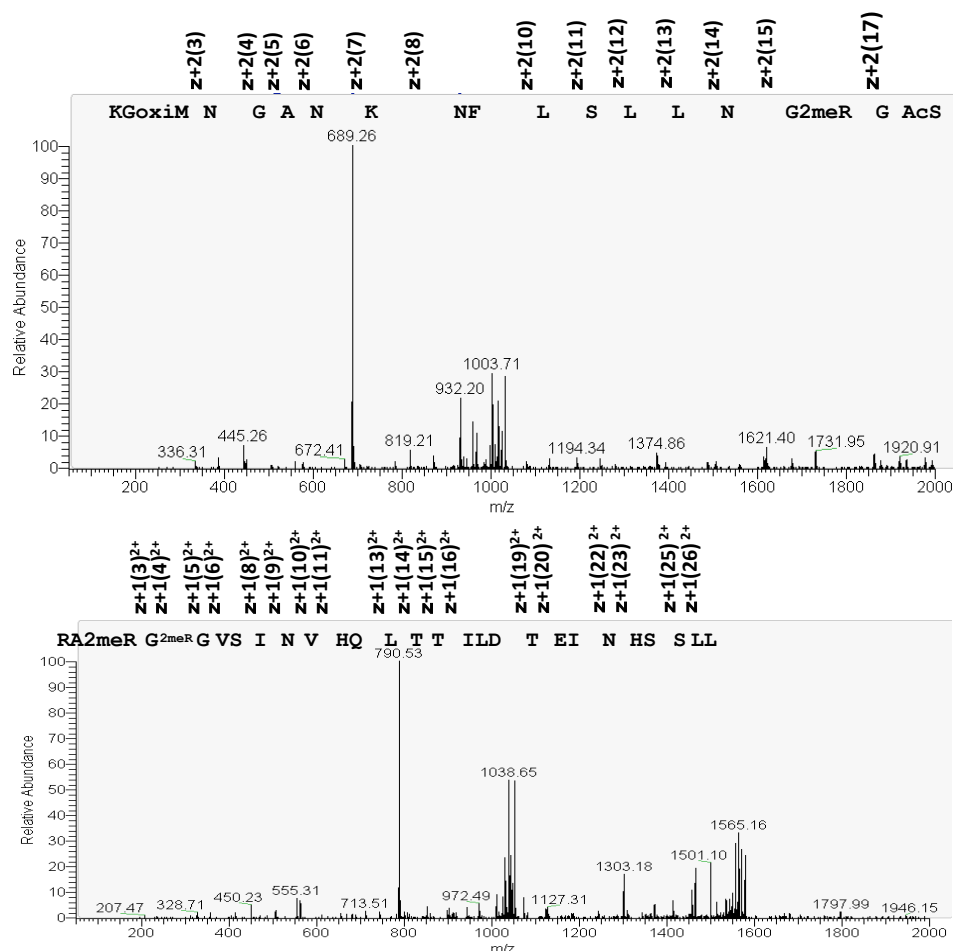
Figure S2. Analysis of AGO3 modifications by LC-MS/MS. (**A**) AGO3 peptide was identified by MASCOT database searching. Letters with yellow background are search hits. Those with green background indicate amino acids with modifications; of those, M, R, and C indicate methionine with oxidation, arginine with dimethylation, and cysteine with carbamidomethylation, respectively. The N-terminal serine (S2) was found to be N-terminally acetylated. (**B**) MS spectrum of the AGO3 peptide; amino acid sequences from S2 to K20 (S2-K20; m/z = 688.6937, theoretical value = 688.6938) and from L45 to R72 (L45-R72; m/z = 789.9459, theoretical value = 789.9460). (**C**) ETD MS/MS spectrum of the AGO3 peptides S2-K20 and L45-R72 with Mascot Ion scores of 82 and 90, respectively. 2meR indicates an arginine with dimethylation. Ac indicates an N-terminal acetylation. z-series ions assigned to the amino acid sequences are indicated.
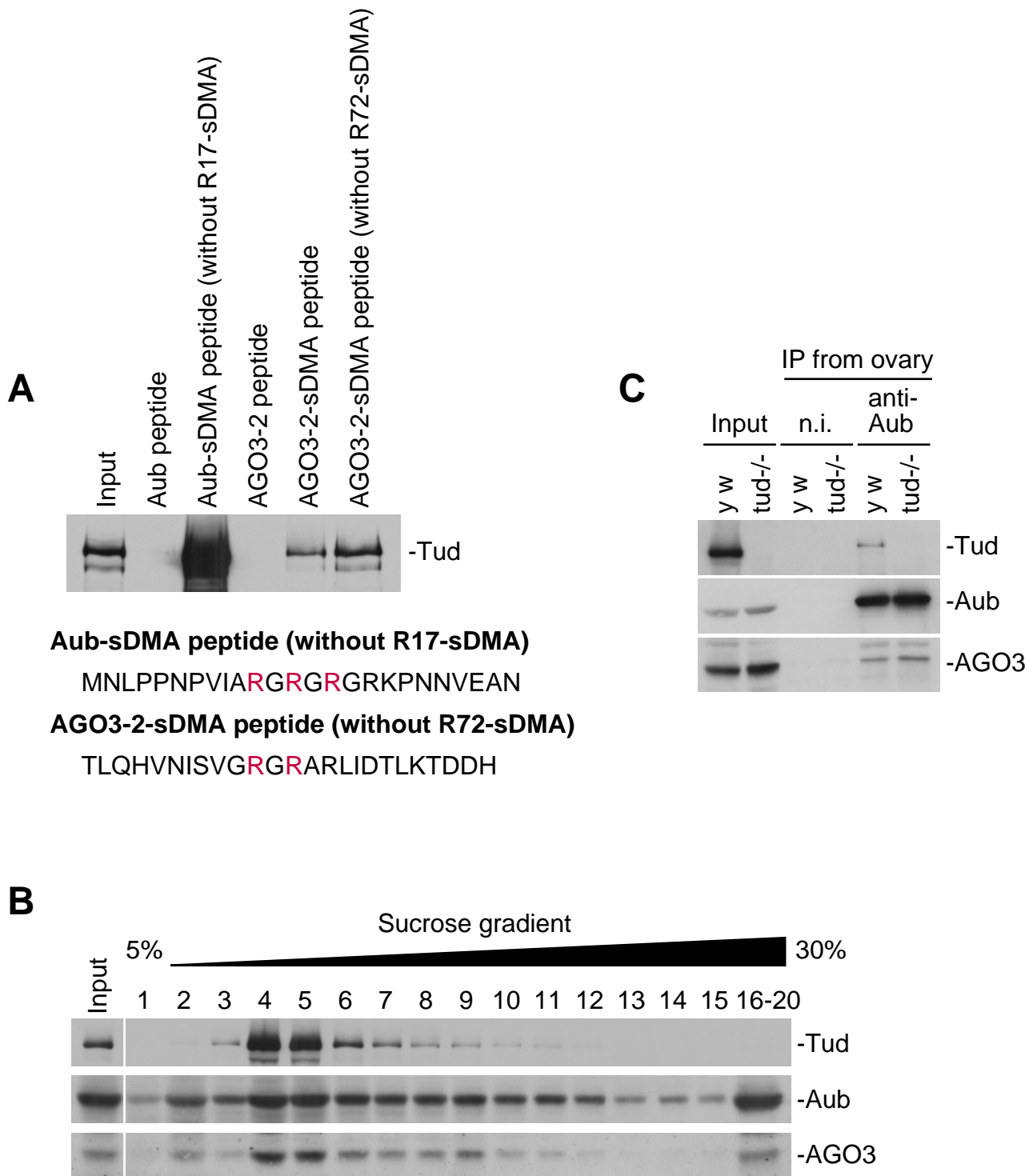
A

Aub-sDMA peptide (without R17-sDMA)

MNLPPNPVIARGRGRGRKPNNVEAN

AGO3-2-sDMA peptide (without R72-sDMA)

TLQHVNISVGRGRARLIDTLKTDDH

Figure S3. (**A**) Pull-down assays were performed as in Figure 2A, from ovary lysates using Aub-sDMA (without R17-sDMA) and AGO3-2-sDMA (without R72-sDMA) peptides. The sequences of the peptides are shown below. Rs in red indicate Rs that were sDMA-modified. The protein pools obtained were probed with anti-Tud antibodies. (**B**)Linear gradient sedimentation experiments were performed using wt ovary lysates. Western blot analyses show that Tud, Aub and AGO3 co-sediment even in relatively heavier fractions. (**C**) Immunoprecipitation was performed from ovary lysates of wt and *tud* mutants using anti-Aub antibodies. Non-immune IgG (n.i.) was used as a negative control. The immunopurified complexes were probed with anti-Tud, anti-Aub and anti-AGO3 antibodies. Tud was observed only in the Aub complex obtained from wt ovaries. It is noted that Aub is able to interact with AGO3 even in *tud* mutants.
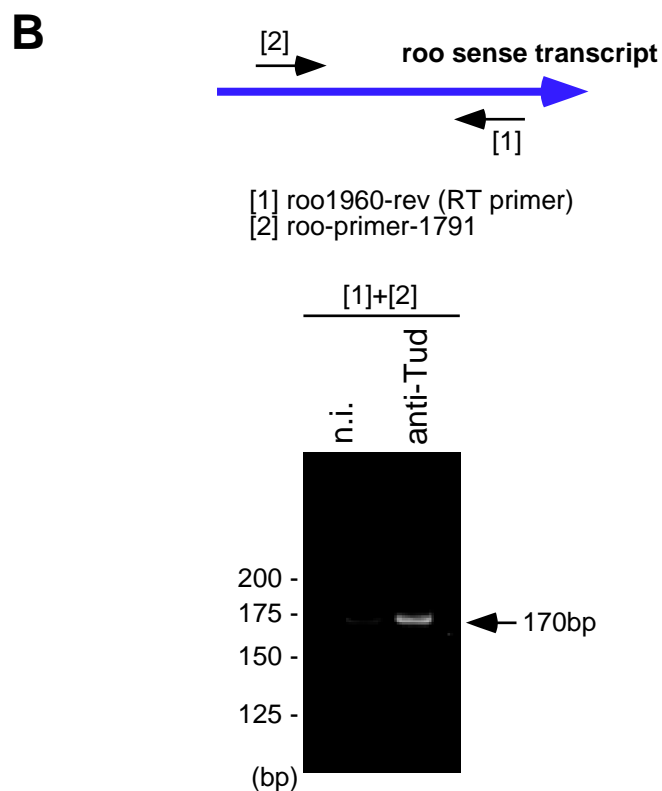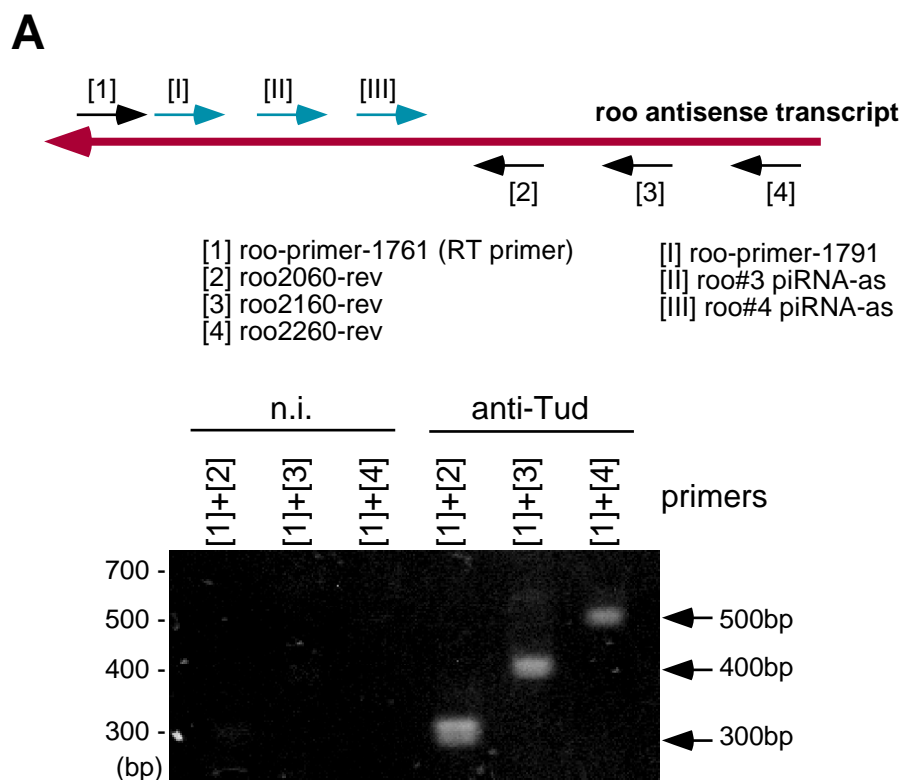
A

[1] roo-primer-1761 (RT primer)
[2] roo2060-rev
[3] roo2160-rev
[4] roo2260-rev

[I] roo-primer-1791
[II] roo#3 piRNA-as
[III] roo#4 piRNA-as

B

[1] roo1960-rev (RT primer)
[2] roo-primer-1791

Figure S4. (**A**) RT-PCR showed that the anti-Tud immunoprecipitates (Figure 4B) contain part of the *roo* antisense transcript. Non-immune IgG (n.i.) was used as a negative control. The primer used for the RT (reverse transcription) reaction was roo-primer-1761 (RT primer). roo-primer-1761 was also utilized for PCR in combination with either roo2060-rev, roo2160-rev, or roo2260-rev primer. The positions of oligos used for northern blot analysis in Figure 4B (roo-primer-1791, roo#3 piRNA-as and roo#4 piRNA-as) are also indicated in this figure. (**B**) RT-PCR shows that the anti-Tud immunoprecipitates contain sense *roo* piRNA precursor-like molecules.
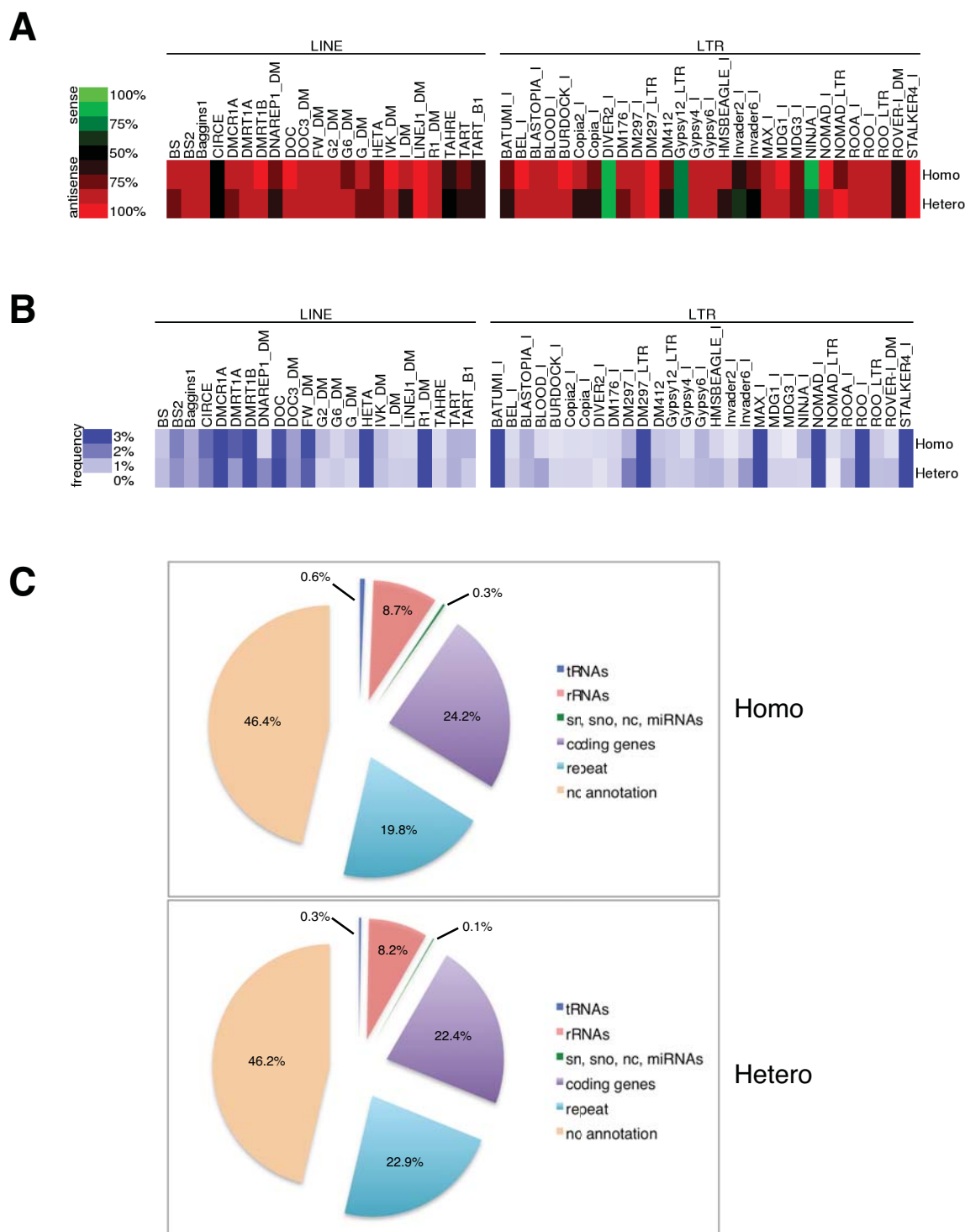
Figure S5. (**A**) Strand bias shows a comparison of strand preferences between transposon-derived short reads from *tud* heterozygous and homozygous mutants. The figure uses a color scheme analogous to that of two-color microarray data presentation. A red-colored box represents the corresponding transposon family that expresses more antisense transcripts than sense transcripts. A green-colored box represents the reverse situation. (**B**) Cloning frequency shows population differences between transposon-derived short reads from *tud* heterozygous and homozygous mutants. The population of each transposon family is normalized as a percentile over total reads. A darker box represents a greater percentile. (**C**) The pie-charts show the contents of short reads from *tud* heterozygous and homozygous mutants. It should be noted that transposon-derived piRNAs are omitted from the charts.