PIWI-piRNA pathway function in *Hydra* somatic stem cells

Celina E. Juliano, Adrian Reich, Na Liu, Jessica Götzfried, Mei Zhong, Selen Uman, Robert A. Reenan, Gary M. Wessel, Robert E. Steele, and Haifan Lin

SUPPLEMENTARY INFORMATION

Table of Contents

I. Supplemental Materials and Methods

Hydra strains and culturing conditions

Hydra magnipapillata strain 105 and *Hydra vulgaris* strain AEP were cultured at 18°C by standard procedures (1). All experiments described were performed using the 105 strain unless otherwise noted. Transgenic AEP *Hydra* expressing GFP in either the ectoderm or endoderm were used for immunofluorescent labeling and transgenic animals expressing DsRed2 in the ectoderm and GFP in the endoderm were used for FACS. These transgenic animals were made as previously described (2, 3). sf-1 strain *Hydra* were cultured at 25°C for 4 days to remove the interstitial cell lineage and then processed for immunoblot and immunofluorescence analysis as described below (4).

Hywi and Hyli identification and antibody generation

Hywi and *Hyli* were identified by BLAST analysis of the *Hydra magnipapillata* 105 genome (http://hydrazome.metazome.net/cgi-bin/gbrowse/hydra/) (5). Full-length cDNA *hywi* and *hyli* sequences, including UTRs, were obtained using the First Choice RLM-RACE Kit (Life Technologies; Carlsbad, CA) and deposited into GenBank (Hywi, KF411461; Hyli, KF411462). Using PAUP (Phylogenetic Analysis Using Parsimony), an unrooted neighbor-joining phylogram

was made from full-length *piwi* coding sequences; bootstrap replicate values are from 1000 iterations (6). The Hywi N-terminus (amino acids 1-227), the Hywi mid-domain (amino acids 444-583), the Hyli N-terminus (amino acids 1-270), and the Hyli mid-domain (amino acids 479-622) were cloned into the Gateway expression vector pDEST17, which has an N-terminal 6xHis tag (Life Technologies; Carlsbad, CA). Recombinant protein was expressed in BL21-AI bacterial cells (Life Technologies; Carlsbad, CA), purified on Ni-NTA resin (Qiagen; Valencia, CA), and further purified by SDS-PAGE separation and electro-elution. Purified proteins were used to raise antisera in rabbits (Hywi) or guinea pigs (Hyli) (Cocalico Biologicals Inc.; Reamstown, PA). The Hyli N-terminus antibody was affinity purified for immune-EM and immunoprecipitation. Hyli N-terminus recombinant protein was immobilized on an Affi-gel 10 column per the manufacturer's instructions (Bio-Rad; Hercules, CA). Heat-inactivated antiserum was passed over the antigen-immobilized column and bound antibodies were eluted with 1 ml 100 mM glycine (pH 2.5) into 50 ul 1 M Tris (pH 9.5) and used directly for immuno-EM and IP.

Immunoblot and immunofluorescence analysis

For immunoblot analysis, protein extracts were made by removing the culture medium and adding 1X SDS-PAGE loading buffer with 5 mM DTT to *Hydra*. Approximately 10 µl of buffer was used per *Hydra* polyp. To obtain equal loading, the Pierce™ BCA Protein Assay Kit (Thermo Scientific; Rockford, IL) was used for quantitation of total protein. Samples were vortexed, heated at 100°C for 10 minutes, and spun at 15K RPM for 1 minute. Samples were loaded onto 4-15% Mini-PROTEAN TGX Precast Gels (Biorad; Hercules, CA). After transfer to nitrocellulose (Biorad; Hercules, CA) proteins were exposed to primary antibodies overnight at 4°C in TBST + 3% dry milk. Primary antibody dilutions were as follows: Hywi N-terminal serum 1:20,000; Hywi Mid-domain serum 1:2000; Hyli N-terminal serum 1:2000; Hyli Mid-domain serum 1:2000; α-Alpha-Tubulin 1:10,000 (12G10 from the Hybridoma bank), GAPDH 1:1,000 (Sigma 9545; St. Louis, MO); Histone H3 1:1,000 (Abcam 1791; Cambridge, MA). HRP-conjugated secondary antibodies (Jackson ImmunoResearch; West Grove, PA) were diluted 1:10,000 and incubated in blocking buffer for 1 hour at room temperature and visualized by standard ECL detection (ThermoScientific; Rockford, IL).

Whole mount immunofluorescence was performed as previously described (7). Briefly, *Hydra* were relaxed in 2% urethane in *Hydra* medium, fixed in 4% PFA in *Hydra* medium, washed with PBS, and permeabilized with 0.5% Triton X-100 in PBS. Samples were incubated with blocking solution (1% BSA; 10% normal goat serum; 0.1% Triton X-100 in PBS) for one hour. Primary antibodies were diluted in blocking solution and incubations were done overnight

at 4°C. Antibody dilutions were as follows: Hywi N-terminal serum 1:1,000; Hyli N-terminal serum 1:1,000; C41 monoclonal 1:2 (8); GFP 1:500 (Roche Cat #11814460001; Indianapolis, IN); dsRed2 1:50 (Santa Cruz # sc-81595; Santa Cruz, CA). Alexa Fluor-conjugated secondary antibodies were diluted 1:500 in blocking buffer and incubations were done for 1 hour at room temperature (Invitrogen; Carlsbad, CA). For labeling of cells from dissociated *Hydra* polyps, macerations were done as previously described (9). Slides were dried for at least 3 hours and labeling was then performed using the same steps described for whole-mount labeling. Whole-mount *Hydra* preparations were imaged on a Leica TCS SP5 confocal microscope (Leica Microsystems, Bannockburn, IL) and single cells were imaged either on the Leica TCS SP5 or on a Zeiss AxioImager Z1 microscope (Carl Zeiss, Inc.; Thornwood, NY). DNA was labeled with 1:000 Hoechst 33342 diluted in PBS (Life Technologies; Carlsbad CA).

Immuno-electron microscopy

Samples were fixed in 4% paraformaldehyde/0.1% gluteraldehyde in PBS for 15 minutes, followed by 4% PFA in PBS for 1 hour. Samples were cryoprotected in 2.3 M sucrose overnight at 4°C. The samples were rapidly frozen onto aluminum pins in liquid nitrogen. The frozen block was trimmed on a Leica Cryo-EMUC6UltraCut (Leica Microsystems, Bannockburn, IL) and 60 nm sections were collected as previously described (10). The frozen sections were thawed and placed on a nickel formvar /carbon-coated grid floated in a dish of PBS ready for immunolabeling. For immunolabeling, samples on grids were placed section side down on drops of 0.1 M ammonium chloride to quench unreacted aldehyde groups, then blocked for nonspecific binding with 1% fish skin gelatin in PBS. The grids were incubated with primary antibodies for 30 minutes, Hywi serum 1:150 or purified Hyli antibody 1:50. For Hyli, a rabbit anti-guinea pig bridging serum (Jackson ImmunoResearch; West Grove, PA) was used. Rinsed grids were placed on Protein A-gold 10 nm (UtrechtUMC) for 30 minutes. All grids were rinsed in PBS, fixed using 1% gluteraldehyde, then rinsed and transferred to a UA/methylcellulose drop for 10 minutes. Samples were viewed using a FEI Tencai Biotwin Transmission Electron Microscope (FEI; Hillsboro, Oregon) at 80Kv. Images were captured using Morada CCD and iTEM (Olympus) software.

Nuclear-Cytoplasmic Fractionation

Approximately 100 AEP *Hydra* polyps were dissociated into single cells as previously described and filtered through a 100 µm cell strainer (11). $3x10^5$ cells were collected and fractionation was performed following the cell culture protocol of the ProteoExtract Subcellular

Proteome Extraction Kit #539790 (EMD Millipore; Darmstadt, Germany). Nuclear and cytoplasmic fractions were analyzed by immunoblot analysis as described above.

Fluorescence Activated Cell Sorting (FACS)

Transgenic AEP *Hydra* polyps expressing DsRed2 in the ectoderm and GFP in the endoderm were used for all FACS experiments (12). For small RNA sequencing the animals were dissociated using Pronase E (Sigma; St. Louis, MO) as previously described (13). However, this method did not work for immunoblotting because proteins were severely degraded. For immunoblotting *Hydra* were dissociated with 0.25% Trypsin-EDTA solution (Life Technologies; Carlsbad, CA). Approximately 120 *Hydra* were divided evenly into 3 wells of a 24-well plate, the *Hydra* medium was removed and 1 ml of trypsin solution was added. *Hydra* were incubated twice at 37°C for 5 minutes and pipetted up and down after each incubation to dissociate cells. Cells were moved to a 15 ml conical tube and volume was increased to 5 ml with dissociation medium (11). Trypsin was neutralized with fetal bovine serum. For both Pronase E and trypsin dissociation procedures, cells were filtered through a 100 µm filter and washed twice with dissociation medium. Cells were collected after each wash by spinning at 200xg. Cells were sorted on a FACSAria Cell Sorter with 100 µm nozzle (see Figure S4) (BD Biosciences; San Jose, CA).

Immunoprecipitation and piRNA sequencing

For immunoprecipitations (IPs), approximately 100 *Hydra magnipapillata* strain105 polyps were homogenized in 1 ml MCB buffer [50 mM HEPES, pH7.5; 150 mM KOAc; 2 mM Mg(OAc)2; 10% glycerol; 0.1% TritonX-100; 0.1% NP-40; 1 mM DTT] and complete protease inhibitor cocktail (Roche; Indianapolis, IN). The protein concentration of the resulting protein extract was ~1 mg/ml. The total protein extract was pre-cleared by incubating with 50 mg of protein A sepharose CL-4B (GE Healthcare; Piscataway, NJ) for 1 hour at 4°C on a rotator. Protein extracts were incubated with antibody overnight at 4°C on a rotator. For Hywi IP, 5 µl of N-terminal antibody serum or 5 µl of pre-bleed serum was added to ~250 mg of protein extract. For Hyli IP, 4 µg of affinity-purified N-terminal antibody, 5 µl of N-terminal serum, or 5 µl of pre-bleed serum was added to ~250 mg of protein extract. Protein A beads (60 µl of a 2X slurry) were added to each IP and incubated for 1 hour at 4°C on a rotator. Beads were washed 5 times with MCB buffer. For immunoblotting, 40 µl of SDS Sample buffer plus 5 mM DTT was added to beads after removal of the last wash. Samples were vortexed and incubated at 100°C

for 5 minutes. Beads were removed by centrifugation and the resulting supernatant was used for immunoblotting.

For isolation of Hywi- and Hyli-bound piRNAs, RNase OUT was added to the lysate (Life Technologies; Carlsbad, CA). After IP, 300 µl of Trizol-LS was added directly to the beads (Life Technologies; Carlsbad, CA) and RNA was isolated according to manufacturer's protocol. RNA pellet was re-suspended in 10 µL of nuclease-free water: 4 µL was used for 5'-end labeling with [γ-^{32}P] ATP by polynucleotide kinase and 6 µL was used for piRNA sequencing. For total small RNA sequencing 10 µg of total RNA was used as starting material. Small RNA libraries were prepared using Illumina small RNA Preparation Kit v1.5 (Illumina; San Diego, CA) following the manufacturer's protocol. In brief, RNA was electrophoresed in a 15% TBE urea gel and small RNAs were eluted from the gel. Adaptors were ligated to the 3' and 5' ends followed by reverse transcription and low cycle PCR. Libraries were gel-purified and sequenced using the Genome Analyzer II (Illumina; San Diego, CA).

β-elimination and small RNA northern blot

β-elimination reactions and small RNA northern blots were performed as previously described (14, 15). 20 µg of total RNA from *Hydra* AEP in 13.5 µl of water was combined with 4 µl of 5x borate buffer (148 mM borax, 148 mM boric acid, pH 8.6) (Polysciences; Warrington, PA) and 2.5 µl of freshly prepared 200 mM $NaIO_4$. The reaction was incubated for 15 minutes in the dark at room temperature and then 2 µl of 100% glycerol was added to quench unreacted $NaIO_4$. Reactions were then incubated for an additional 15 minutes at room temperature. The reactions were dried in a SpeedVac evaporator for 75 minutes at room temperature. Pellets were resuspended in 50 µl 1x borate buffer with 50mM NaOH (pH 9.5) and incubated at 45°C for 90 minutes. 2 µl of glycogen was added and RNA was EtOH precipitated. Pellets were collected by centrifugation, washed with 70% EtOH and 95% EtOH respectively, and air-dried pellets were resuspended in 15 µl of water. β-elimination reactions and untreated RNA samples were electrophoresed for 5 hours at 200V in 15% polyacrylamide midi-gels with 8M Urea and 0.5X TBE as buffer. RNA was transferred from the gels onto Hybond N+ membrane (GE Healthcare Life Sciences; Pittsburgh, PA) in 0.5X TBE at 100 mA for 2.5 hours. The membranes were then washed two times for 10 minutes with 2X SSC and then UV cross-linked using the UV Stratalinker 1800 (Stratagene; Santa Clara, CA) on the "Auto crosslink" setting. The membranes were incubated for 1 hour at 42°C in hybridization buffer (0.2M $NaHPO_4$ pH 7.2, 1mM EDTA, 1% BSA, 7%SDS). Probes for hybridization were an oligo complementary to miR2030 (probe sequence: CAAATTTATTTTTGCGCTCTCA) (16) and an oligo complementary

to an abundant transposon-derived piRNA (probe sequence: AATCCAAACGCCAGGAATTCGATCACC). 10 pmol of each oligo was 5'-end labeled with [ɣ-^{32}P] ATP by polynucleotide kinase for 1 hour at 37°C. Labeled oligos were purified using oligo quick spin columns (Roche; Indianapolis, IN). The entire labeled oligo sample was added to the hybridization buffer for incubation overnight at 42°C with rotation. Membranes were washed four times for 10 minutes with 2xSSC/0.1%SDS at 50°C with rotation. Finally, the membranes were wrapped in Saran Wrap, exposed to phosphor plates for 6 hours, and imaged on the Typhoon Trio Variable Mode Imager (GE Healthcare Life Sciences; Pittsburgh, PA).

Sequencing of lineage-specific small RNAs

After FACS, RNA was isolated from each lineage by Trizol extraction (Life Technologies; Carlsbad, CA). Approximately 2 µg of RNA was collected from each of the epithelial lineages, ~7 µg of RNA was collected from the interstitial lineage. This RNA was used to generate small RNA libraries using the TruSeq Small RNA Sample Prep Kit (Illumina; San Diego, CA) according to the manufacturer's protocol. Briefly, the libraries were generated by ligation of specific 5' and 3' adapters to the RNA and ligated products were reverse transcribed and amplified by PCR. PCR products were separated by polyacrylamide gel electrophoresis (PAGE), and products corresponding to adaptor ligated 18 – 35 nt long RNAs (~150 bp) were used for pooled gel purification. The libraries were sequenced using the HiSeq™ 2000 (Illumina; San Diego, CA). The quality of the RNA and the corresponding cDNA was analyzed by the Agilent 2100 Bioanalyzer (Agilent Technologies; Santa Clara, CA). For the interstitial lineage there were 15,302,191 trimmed reads less than 23 nucleotides and 23,713,174 trimmed reads 23 nucleotides or greater. For the ectodermal lineage there were 14,873,803 trimmed reads less than 23 nucleotides and 12,715,400 trimmed reads 23 nucleotides or greater. For the endodermal lineage there were 26,964,740 trimmed reads less than 23 nucleotides and 21,073,941 trimmed reads 23 nucleotides or greater.

Bioinformatic analysis and genomic mapping of small RNAs

For analysis of PIWI-bound piRNAs, three small RNA libraries were sequenced: 1. Total small RNAs (total), 2. Hywi-bound piRNAs (Hywi), and 3. Hyli-bound piRNAs (Hyli). Linker sequences were successfully trimmed from more than 90% of the sequences. 18-32nt RNAs were selected and mapped to the *Hydra* genome for further analysis. From the total library, ~10 million sequences were mapped to the genome, from the Hywi library ~4.7 million were mapped, and from the Hwyli library ~6.3 million were mapped. Approximately 50% of the

sequenced small RNAs were mapped to the genome without ambiguity. First, the small RNAs that mapped to tRNAs were annotated. Because *Hydra* tRNA annotation is not yet complete, up to 2 mismatches were allowed including insertions/deletions. For miRNA annotation, we extended 30 nucleotides on both ends of the small RNA sequences obtained by the 454 sequencing carried out as part of the *Hydra* genome project (17). For transposon/repeat annotation, we referred to the annotation information obtained via RepeatMasker as part of the *Hydra* genome project (17). For the gene annotation, we adopted the Berkeley group's annotation (17). To test whether the ping-pong mechanism of piRNA biogenesis functions in *Hydra*, we searched for the ping-pong signature among these libraries by determining whether the piRNAs have sequence partners with a 10 nucleotide off-set. The partners were defined as piRNAs whose 5' 10 nucleotides are reverse compliments, as previously described (18). The same methods were used to test for the presence of the ping-pong signature in the lineage-specific small RNA sequencing data (26 to 34 nucleotides): 1. Interstitial, 2. Ectoderm, and 3. Endoderm.

Assembly of the *Hydra* transcriptome and mapping of piRNAs and lineage-specific small RNAs

Total RNA was isolated from 2-day starved *Hydra magnipapillata* strain 105. A cDNA library was constructed using the Illumina TruSeq RNA Sample Preparation Kit using a slightly modified procedure. The total RNA was only sheared for 1.5 minutes and prior to the final PCR enrichment, the library was run on a LabChip XT DNA 750 (PerkinElmer; Waltham, Massachusetts) to select for 500bp fragments (+/-5%). The final library was sequenced on the Illumina HiSeq 2000 (Illumina; San Diego, CA) yielding 43.6 million 100 base pair paired-end reads. The transcriptome was assembled using a previously described pipeline (19). Approximately 27,000 sequences were initially assembled and then curated to 9,986 sequences with BLAST hits to the Swiss-Prot database (E value of $1xe^{-5}$ or less), thus allowing for identification of open reading frames and transcript orientation. To identify transposon sequences, all known *Hydra* transposons were first downloaded from RepBase (current version August, 5th 2013). One sequence was removed (hAT-64_HM|hAT|Hydra_magnipapillata) due to ambiguous characters, yielding 565 transposons. We compared the *de novo* transcriptome against the *Hydra* transposon database by BLAST analysis (tblastx -evalue .00001) and any transcripts with a BLAST hit were flagged as transposons. Small RNAs were trimmed of adapter sequence using CutAdapt (version 1.0 "-a ATCTCGTATGCCGTCTTCTGCTTG -m 10 --too-

short-output --untrimmed-output", though subsequent analysis ignored small RNAs shorter than 23nt) and were mapped to the transcriptome allowing up to a 3 base-pair mismatch using Bowtie 1 (version 0.12.9) with the following settings: "-a --best --strata -v 3". 1.4 million small RNAs at least 23 base pairs long from the total small RNA library were mapped to the transcriptome, 63% of these mapped to transposon sequences. Hywi-bound piRNAs and Hyli-bound piRNAs at least 23 base pairs long were also mapped to the transcriptome: 725,000 Hywi-bound piRNAs mapped and 963,000 Hyli-bound piRNAs mapped. For the lineage-specific small RNA sequences, adapter sequences were trimmed, using CutAdapt with the following settings: "-a TGGAATTCTCGGGTGCCAAGGC -m 15". The trimmed reads were mapped against the transcriptome with Bowtie 1 (version 0.12.9) with the same settings as above and yielded: 1,247,033 interstitial, 488,903 ectodermal, and 848,964 endodermal small RNAs 23 nucleotides or greater. Transcripts with at least 10 times more mapped piRNAs from a specific lineage as compared to other lineages (normalized to the size of the libraries) were considered putative lineage-specific targets for that particular lineage. Gene ontology analysis of putative PIWI-piRNA pathway targets was done using DAVID (20), employing the Swiss-Prot identification from the closest BLAST hit. All small RNA short reads, RNA-seq short reads and the assembled transcriptome were deposited under NCBI BioProject PRJNA213706.

Real-time quantitative PCR to test *hywi* knockdown levels

RNA was isolated from F1 *Hydra* hatchlings expressing the *hywi* RNAi-1 transgene and from wild type F1 siblings by Trizol extraction (Life Technologies; Carlsbad, CA). Further purification of the RNA and on-column DNAse-treatment was performed using the RNeasy Mini Kit (Qiagen; Valencia, CA) according to the manufacturer's protocol. Reverse transcription was performed using the High Capacity cDNA Reverse Transcription Kit (Life Technologies; Carlsbad, CA) according to the manufacturer's protocol. Real-time quantitative PCR was carried out using iQ SYBR Green 2x Supermix on a CFX96™ thermal cycler (Bio-rad; Hercules, CA). Primer pairs were as follows: Actin (Forward: AAGCTCAGAGCAAACGTGGT; Reverse: GGACAGGGTGTTCTTCTGGA); GAPDH (Forward: GACAACCATTCATGCCACAA; Reverse: ACAGCTTTTGCAGCTCCAGT); Hywi-1 (Forward: CCACAACCTCCTGTTGGAGT; Reverse: TGAGCAGTTTGCTGAGGTTG); Hywi-2(Forward: ACCCAAGGACCAATCCTTTT; Reverse: AAATTTTTCGCACGCATCTC); Hyli (Forward: GCCCTGGAAACACCTATGAA; Reverse: GGATGAGTGCCCATTCACTT).

RNAi plasmid description and construction

The RNAi plasmids were designed in an operon configuration such that the upstream hairpin and the downstream DsRed2 gene are transcribed together from an actin gene promoter, with the bicistronic primary transcript then being separated into a hairpin RNA and a DsRed2 mRNA by trans-splicing (2) (See Fig 4*A*). The RNA hairpin consists of a ~500 base pair fragment from the target gene cloned in an inverted orientation around an actin intron sequence (483 base pairs). Downstream of the hairpin cassette is the DsRed2 gene followed by the actin 3'UTR (500 base pairs). In between the hairpin cassette and the DsRed2 gene is the RFC140/*flp* intergenic sequence, which contains an acceptor for trans-spliceed leader addition (2). Thus, the RNA hairpin and DsRed2 transcript are arranged in an operon configuration and are spliced apart after transcription.

The operon plasmid pHyVec11 (2) was modified by the insertion of a GFP hairpin (nucleotides 1-552) separated by the intron from a *Hydra* actin gene (21) in the upstream position of the operon. This plasmid, named pHyVec12 (*gfp* RNAi), was then used to construct the *hywi* RNAi-1 and *hywi* RNAi-2 plasmids. pHyVec12 was cut with NheI and BamHI to remove the GFP hairpin and the actin intron. *Hywi* forward and reverse sequences (379-899 or 1557-2093), surrounding the actin intron sequence, were then inserted into the plasmid using the Cold Fusion Cloning Kit (System Biosciences; Mountain View, CA).

Generation of transgenic *Hydra*.

The generation of transgenic *Hydra* was performed as previously described (2, 3). *Hywi* RNAi and *gfp* RNAi plasmids were prepared by Maxiprep (Qiagen, Valencia, CA) and eluted in RNase-free water. Plasmid DNA was injected at a final concentration of 1 mg/mL using an IM-9B Narishige microinjector (Narishige; East Meadow, NY) under a Zeiss dissecting scope (Carl Zeiss, Inc.; Thornwood, NY).

References

1. Lenhoff HM & Brown RD (1970) Mass culture of hydra: an improved method and its application to other aquatic invertebrates. *Laboratory animals* 4(1):139-154.
2. Dana CE, Glauber KM, Chan TA, Bridge DM, & Steele RE (2012) Incorporation of a horizontally transferred gene into an operon during cnidarian evolution. *PLoS One* 7(2):e31643.

3. Wittlieb J, Khalturin K, Lohmann JU, Anton-Erxleben F, & Bosch TC (2006) Transgenic Hydra allow in vivo tracking of individual stem cells during morphogenesis. *Proc Natl Acad Sci U S A* 103(16):6208-6211.
4. Marcum BA, Fuijsawa T, & Sugiyama T (1980) A mutant strain (sf-1) containing temperature-sensitive interstitial cells. *Developmental and Cellular Biology of Coelenterates*, eds Tardent P & Tardent R (Elsevier, Amsterdam), pp 429-434.
5. Altschul SF, Gish W, Miller W, Myers EW, & Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215(3):403-410.
6. Swofford DL (2002) PAUP*: Phylogenetic Analysis Using Parsimony (*and other methods). Version 4. (Sinauer Associates, Sunderland, MA).
7. Munder S*, et al.* (2010) Notch signalling defines critical boundary during budding in Hydra. *Dev Biol* 344(1):331-345.
8. David CN, Fujisawa T, & Bosch TC (1991) Interstitial stem cell proliferation in hydra: evidence for strain-specific regulatory signals. *Dev Biol* 148(2):501-507.
9. David CN (1973) A Quantitative Method for Maceration of Hydra Tissue. *Wilhelm Roux' Archiv* 171:259-268.
10. Tokuyasu KT (1973) A technique for ultracryotomy of cell suspensions and tissues. *J Cell Biol* 57(2):551-565.
11. Gierer A*, et al.* (1972) Regeneration of hydra from reaggregated cells. *Nature: New biology* 239(91):98-101.
12. Glauber KM*, et al.* (2013) A small molecule screen identifies a novel compound that induces a homeotic transformation in *Hydra*. *in press at Development*.
13. Hemmrich G*, et al.* (2012) Molecular signatures of the three stem cell lineages in Hydra and the emergence of stem cell function at the base of multicellularity. *Mol Biol Evol*.
14. Vagin VV*, et al.* (2006) A distinct small RNA pathway silences selfish genetic elements in the germline. *Science* 313(5785):320-324.
15. Watanabe T, Totoki Y, Sasaki H, Minami N, & Imai H (2007) Analysis of small RNA profiles during development. *Methods in enzymology* 427:155-169.
16. Krishna S*, et al.* (2013) Deep sequencing reveals unique small RNA repertoire that is regulated during head regeneration in Hydra magnipapillata. *Nucleic Acids Res* 41(1):599-616.
17. Chapman JA*, et al.* (2010) The dynamic genome of Hydra. *Nature* 464(7288):592-596.
18. Brennecke J*, et al.* (2007) Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell* 128(6):1089-1103.

19. Howison M, Sinnott-Armstrong N, & Dunn CW (2012) BioLite, a lightweight bioinformatics framework with automated tracking of diagnostics and provenance. *Proceedings of the 4th USENIX Workshop on the Theory and Practice of Provenance (TaPP '12)*.
20. Dennis G, Jr.*, et al.* (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* 4(5):P3.
21. Fisher DA & Bode HR (1989) Nucleotide sequence of an actin-encoding gene from Hydra attenuata: structural characteristics and evolutionary implications. *Gene* 84(1):55-64.
22. Slautterback DB & Fawcett DW (1959) The development of the cnidoblasts of Hydra; an electron microscope study of cell differentiation. *The Journal of biophysical and biochemical cytology* 5(3):441-452.
23. David CN & Gierer A (1974) Cell cycle kinetics and development of Hydra attenuata. III. Nerve and nematocyte differentiation. *J Cell Sci* 16(2):359-375.
24. Horwich MD*, et al.* (2007) The Drosophila RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. *Curr Biol* 17(14):1265-1272.
25. Kirino Y & Mourelatos Z (2007) Mouse Piwi-interacting RNAs are 2'-O-methylated at their 3' termini. *Nat Struct Mol Biol* 14(4):347-348.
26. Ohara T, Sakaguchi Y, Suzuki T, Ueda H, & Miyauchi K (2007) The 3' termini of mouse Piwi-interacting RNAs are 2'-O-methylated. *Nat Struct Mol Biol* 14(4):349-350.
27. Saito K, Sakaguchi Y, Suzuki T, Siomi H, & Siomi MC (2007) Pimet, the Drosophila homolog of HEN1, mediates 2'-O-methylation of Piwi- interacting RNAs at their 3' ends. *Genes Dev* 21(13):1603-1608.
28. Mortazavi A, Williams BA, McCue K, Schaeffer L, & Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* 5(7):621-628.

II. Supplementary Figure Legends

Supplemental Figure 1: Generation and characterization of antibodies against Hywi and Hyli.

(A) *Hydra* is a member of the phylum Cnidaria, which is the sister group to the bilaterians. (B) An unrooted neighbor-joining phylogram demonstrates that *hywi* and *hyli* cluster with *piwi* family

genes and that *Hy-ago1* and *Hy-ago2* cluster with *ago* family genes. *Hywi* clusters with *miwi*, *ziwi*, and *seawi* (red lines) and *hyli* clusters with *mili*, *zili*, and *seali* (green lines). Numbers indicate bootstrap replicate values from 1000 iterations. Abbreviations are as follows: Mm (*Mus musculus*), Dr (*Danio rerio*, i.e. zebrafish), Sp (*Strongylocentrotus purpuratus*, i.e. sea urchin), Sm (*Schmidtea mediterranea*, i.e. planarian), Dm (*Drosophila Melanogaster*). (C) *Hywi* and *hyli* have conserved ARGONAUTE family domain structures. Sequence comparisons between the protein domains of Hywi and Hyli show that the N-terminal and Mid-domains have lower sequence identity (23% and 27%) as compared to the PAZ and PIWI domains (48% and 58%). (D,E) Polyclonal antibodies were raised against the N-terminal and Mid-domains of Hywi (in rabbit) and Hyli (in guinea pig). (F) The 100kDa protein immunoprecipitated with the Hywi N-terminus (N) antibody is recognized by the antibody raised against the Mid-domain (MD) of Hywi, but not by the N-terminal Hyli antibody. (G) The 100 kDa protein immunoprecipitated with the Hyli N-terminus antibody is recognized by the antibody raised against the Mid-domain of Hyli, but not by the N-terminal Hywi antibody.

Supplemental Figure 2: Hyli protein is expressed in interstitial stem cells and mitotically active epithelial stem/progenitor cells.

(A-C) Hywi (red) and Hyli (green) antibodies stain the same population of cells throughout the body column. Green nematocyte (e.g. asterisk) labeling in panel B is non-specific labeling from the secondary antibody. (D-F) Hyli (red) is expressed in the C41-positive (green) I-cells. (G-J) Confocal images of Hyli accumulation in perinuclear granules in ectodermal (G,H) and endodermal (I,J) epithelial cells. Staining was done on transgenic *Hydra* that express GFP in either the ectodermal (H) or endodermal (J) lineages. DNA is labeled with Hoechst 33342.

Supplemental Figure 3: *Hydra* PIWI proteins are expressed in developing nematoblasts.

Interstitial stem cells in the process of differentiating into nematocytes, the specialized stinging cells of cnidarians, undergo four divisions with incomplete cytokinesis that give rise to 2-, 4-, 8- and 16-cell nematoblast nests that are distributed throughout the body column. In the final step of differentiation, the cells break apart from the nests and migrate to the tentacles (22, 23). (A-L) Staining of dissociated *Hydra* cells show that Hywi (red) and Hyli (green) proteins are expressed in 4-cell (A-C), 8-cell (D-F), and 16-cell (G-I) nematoblast nests, but not in differentiated nematocytes (J-L). DNA is labeled with Hoechst 33342.

Supplemental Figure 4: FACS isolation of ectodermal and endodermal cells.

(A) Transgenic *Hydra* expressing GFP in the endoderm and DsRed2 in the ectoderm (see Fig. 2*B*) (12) were dissociated into single cells which were then sorted by FACS. Double-negative (DN) cells were also collected as the interstitial lineage cell population (B) Imaging of cells after a sort demonstrates that the GFP-positive cells were successfully separated away from the total cell population.

Supplemental Figure 5: Isolation, deep-sequencing, and mapping of Hywi and Hyli bound piRNAs to the *Hydra* genome and beta elimination assay

(A) Both total RNA and RNA extracted from Hywi and Hyli immunoprecipitates was 5' end-labeled with [γ-^{32}P] ATP. Immunoprecipitated RNA contains piRNAs just below the 30 nucleotide marker (arrows). (B,C) Mapping of *Hydra* piRNAs to the genome: "Total" is small RNAs selected by size and "Hywi" and "Hyli" are the piRNAs isolated from Hywi and Hyli immunoprecipitates respectively. (B) piRNAs that map uniquely to the *Hydra* genome (~50%) showed no enrichment for transposon/repeat sequences (the *Hydra* genome is 57% transposon/repeat sequences) (17). (C) Analyzing all piRNAs that map to the genome, even those that map in multiple places, also shows no significant enrichment for transposon/repeat sequences. In this analysis, piRNAs that map more than once are weighted. (D) Analysis of nucleotide distribution across the length of Hywi-bound and Hyli-bound piRNAs demonstrates that Hywi-bound piRNAs have a strong preference for uridine at the 5' position and Hyli-bound piRNAs have a strong preference for adenine at the 10th position. (E) Northern blot analysis of total RNA subjected to β-elimination or control total RNA. Anti-sense probes detect *Hydra* miR2030 (16) or an abundant *Hydra* transposon-derived Hyli-associated piRNA with the following sequence: GGTGATCGAATTCCTGGCGTTTGGATT. The piRNA, but not the miRNA, is protected from nucleotide loss due to β-elimination, thus indicating that the piRNA is 2'-O-methylated at the 3' end similar to piRNAs in *Drosophila* and mice (24-27).

Supplemental Figure 6: Analysis of lineage-specific small RNAs. (A) Size distribution of lineage specific small RNAs. (B) Analysis of nucleotide distribution across the length of small RNAs between 26 and 34 nucleotides long from each lineage (interstitial, ectoderm, and endoderm) demonstrates a strong preference for uridine at the 5' position for small RNAs from all three lineages. (C) Small RNAs between 26 and 34 nucleotides long isolated from the interstitial lineage have a higher frequency of complementary overlap 10 bases from their 5' end compared to small RNAs of the same length isolated from the ectodermal or endodermal lineages. (D-E) In order to compare the numbers of small RNAs mapping to the same transcript

across the three lineage datasets and to compare piRNA mapping numbers between different transcripts, the mapped reads were normalized similarly to RPKM (28) (here piPKM) and log transformed. (D) The interstitial lineage has many more small RNAs mapping to transposons [average log(piPKM) in the interstitial lineage is 0.48] compared with the epithelial lineages [average 0.11, -0.02 in the ectoderm and endoderm lineages respectively]. (E) Non transposon transcripts have many fewer small RNAs mapping to them and there is no lineage specific difference [average log(piPKM) is -0.74, -0.51, -0.60 in interstitial, ectoderm and endoderm lineages respectively].

Supplemental Figure 7: Transmission of the *hywi* RNAi-1 transgene through the germline and knockdown of *hywi* in the epithelial cells of F1 hatchlings.
(A-C) A transgenic line was established that uniformly expresses DsRed2 under control of an actin promoter in all three lineages. This was accomplished by sexual transmission of the transgene. Double labeling with antibodies against Hywi and DsRed2 demonstrates that the transgene is expressed in the epithelial cells, but not in the interstitial stem cells (arrow in panel C). (D) Stable lines were created expressing the *hywi* RNAi-1 or *hywi* RNAi-2 transgene (Fig. 4A,B) in the interstitial lineage under the control of the actin promoter, which can be observed by DsRed2 expression in the differentiated cells of the lineage. (E,F) In one line *hywi* RNAi-1 is transmitted through the germline. (G-I) The resulting F1 *Hydra* hatchlings do not express Hywi in the epithelial cells (H,I), but Hywi protein is detected in nontransgenic F1 siblings (G). (J-L) Hywi protein is still detected in the interstitial stem cells of *hywi* knockdown hatchlings because the actin promoter is not active in these cells. Transgenic cells are identified by labeling with an antibody against DsRed2. DNA is labeled with Hoechst 33342. (M) *hywi* mRNA levels were tested by qRT-PCR at several time points after hatching and eating in *hywi* knockdown F1 *Hydra* as compared to wild type F1 sibling controls (normalized to actin).

III. Supplemental Table Legends

Supplemental Table 1: piRNA mapping to the *Hydra* transcriptome.
The numbers correspond to the bar graph in Figure 3E. For transposon transcripts, the majority of Hywi-bound piRNAs are mapped in the anti-sense orientation (yellow boxes) and the majority of Hyli-bound piRNAs are mapped in the sense orientation (grey boxes). The majority of both Hywi- and Hyli-bound piRNAs that map to non-transposon transcripts map in the sense orientation (yellow boxes).

Supplemental Table 2: Gene ontology analysis of transcripts with greater than 10 Hywi-bound piRNAs mapped.

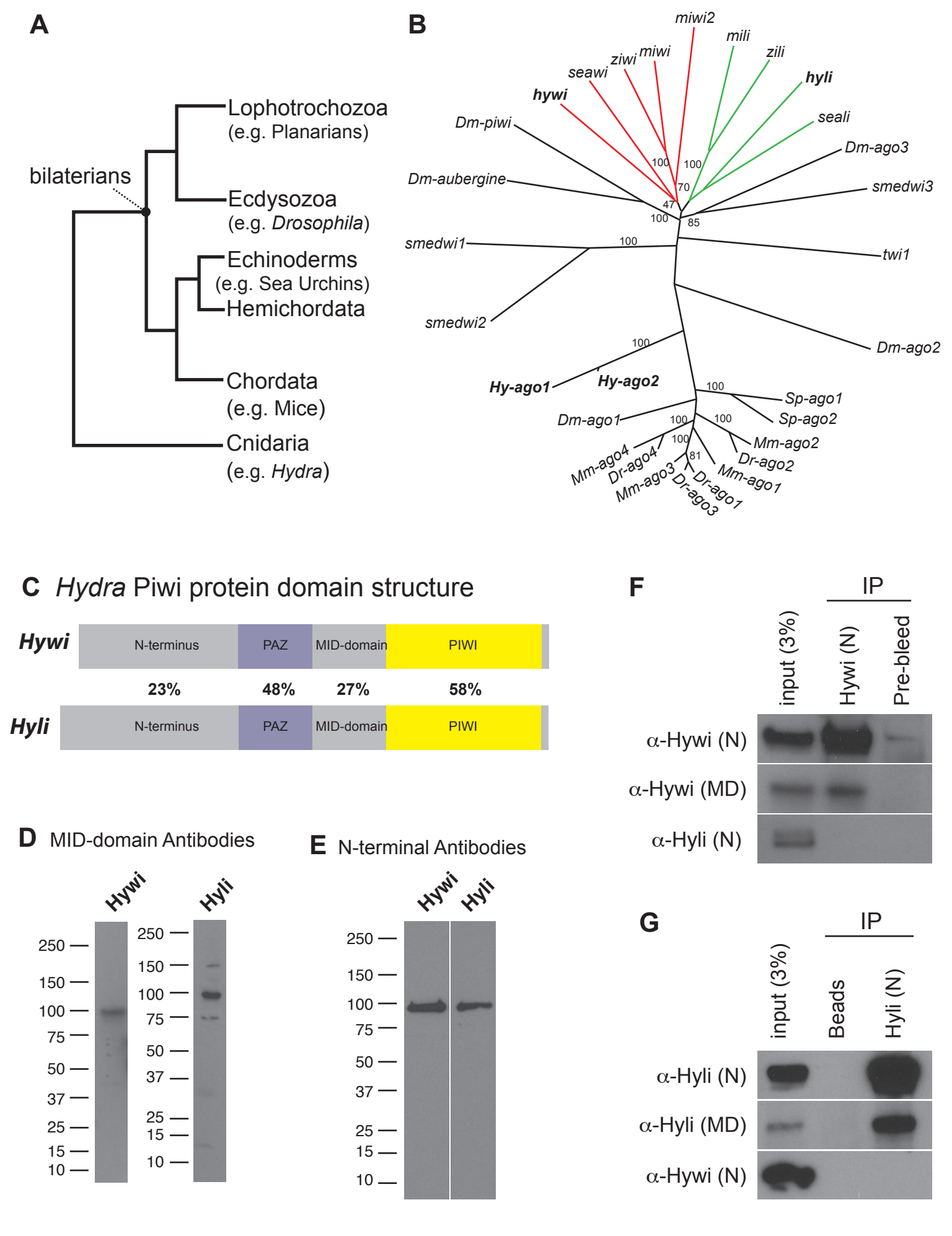
Categories highlighted in green are enriched only for transcripts with Hywi-bound mapping piRNAs and categories highlighted in purple are enriched for transcripts with both Hywi- and Hyli-bound mapping piRNAs.

Supplemental Table 3: Gene ontology analysis of transcripts with greater than 10 Hyli-bound piRNAs mapped.

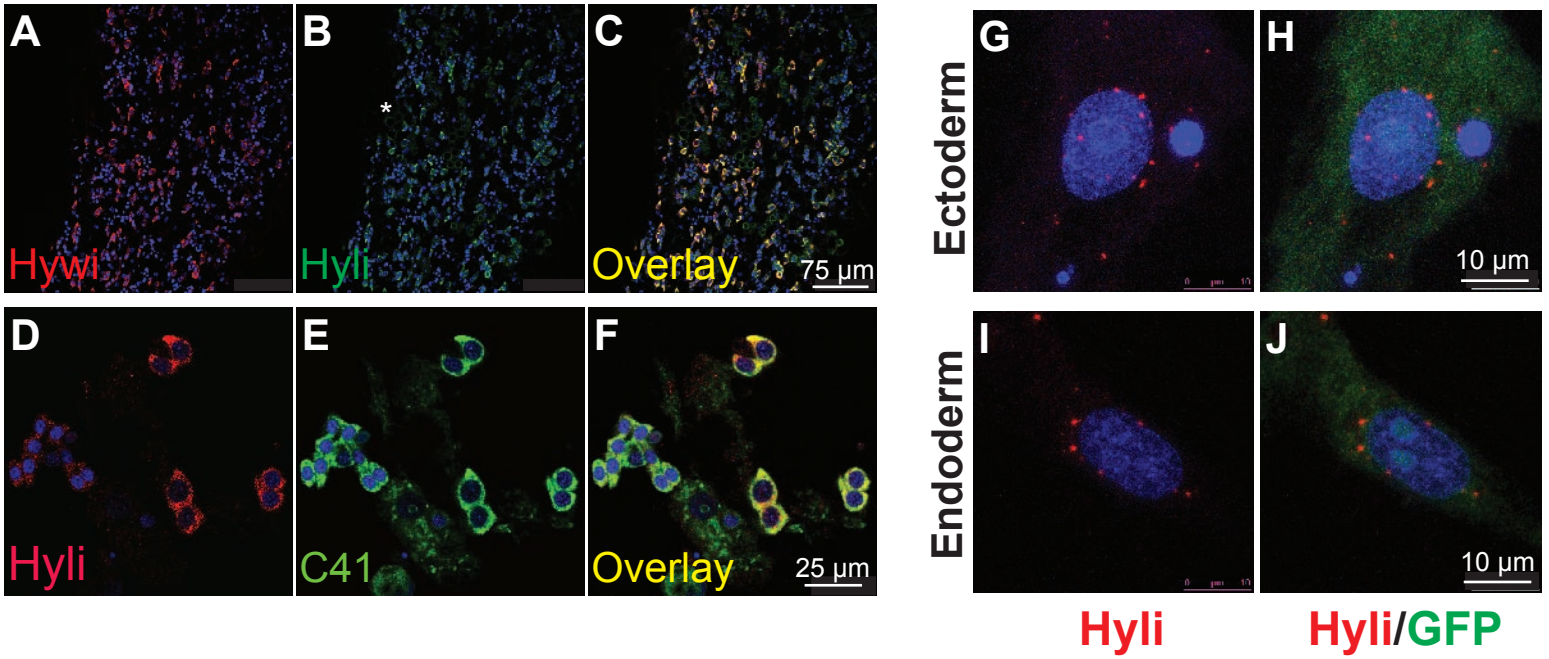
Categories highlighted in blue are enriched only for transcripts with Hyli-bound mapping piRNAs and categories highlighted in purple are enriched for transcripts with both Hywi- and Hyli-bound mapping piRNAs.

Supplemental Table 4: Gene ontology analysis of putative lineage-specific targets of the PIWI-piRNA pathway.

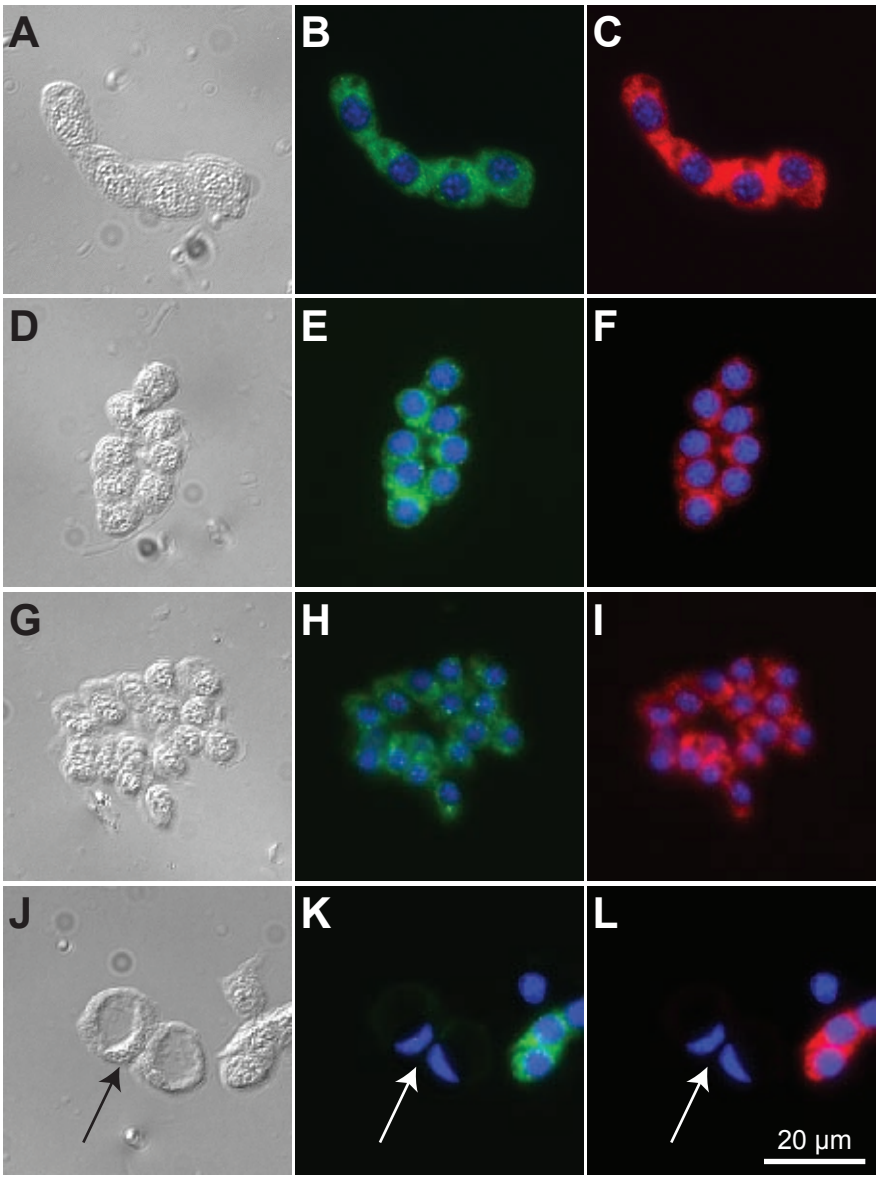
Lineage-specific small RNAs 23 nucleotides or greater in length were mapped to the *Hydra* transcriptome. Transcripts with at least 10 times more small RNAs mapping from a specific lineage were identified as putative targets specific to that lineage. The category “epithelium” is a combination of both ectodermal and endodermal small RNAs as compared to interstitial small RNAs, thus identifying putative mRNAs that are targeted in both epithelial layers, but not in the interstitial lineage.



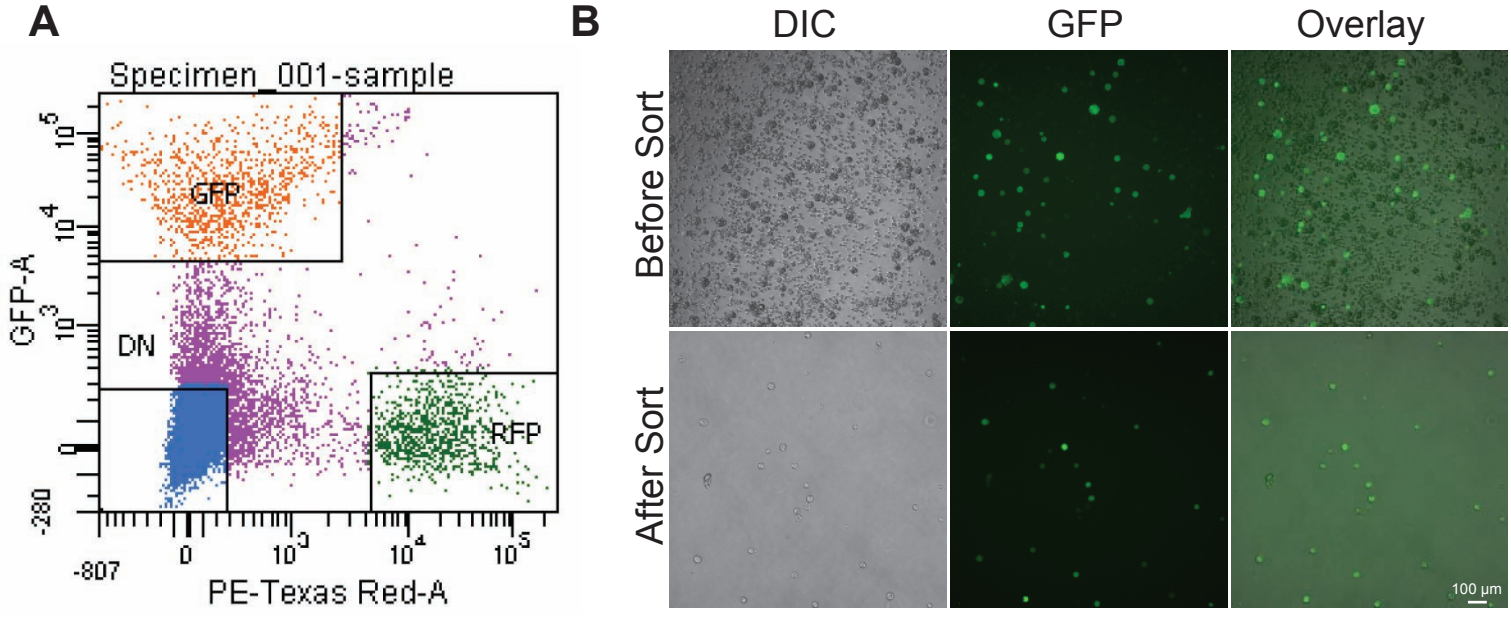
Supplemental Figure 1



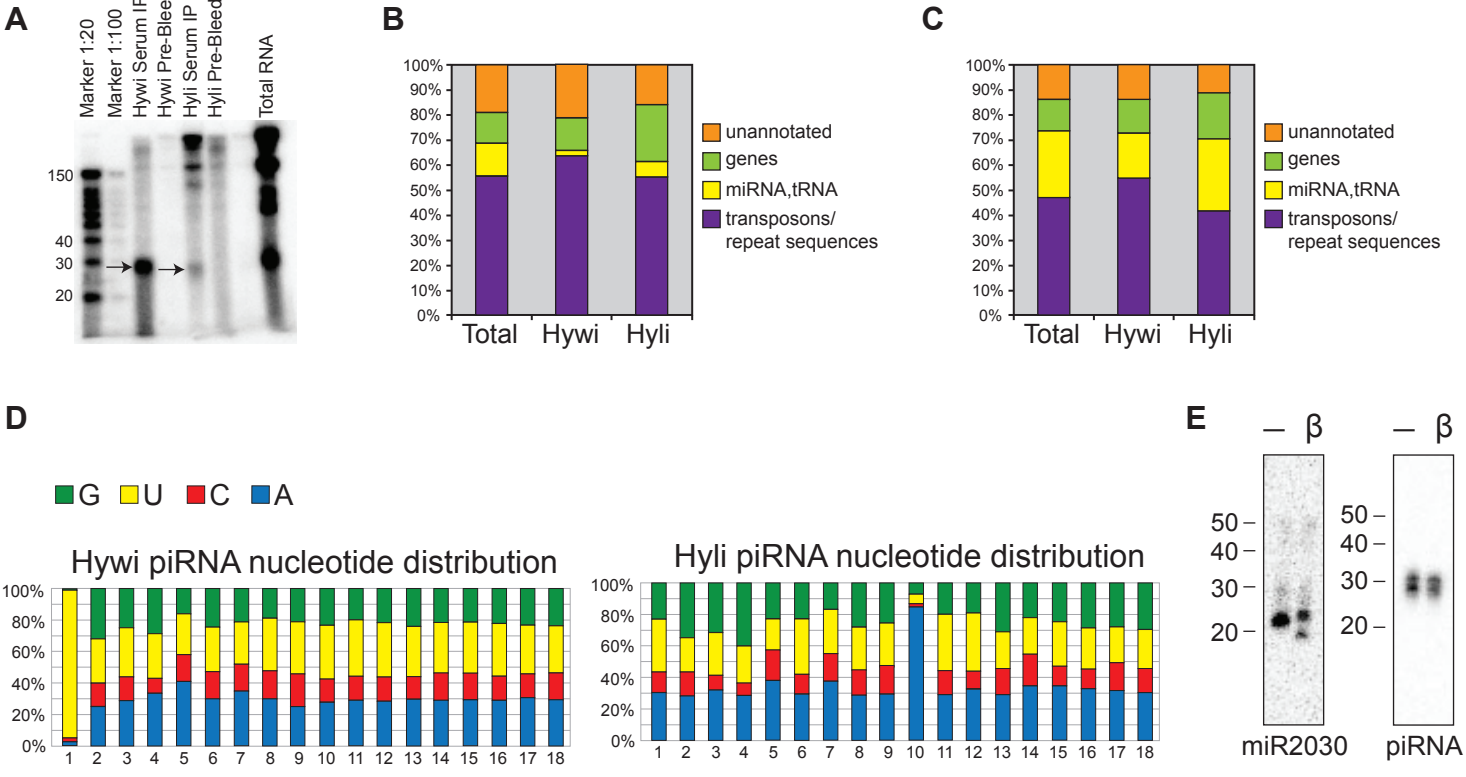
Supplemental Figure 2



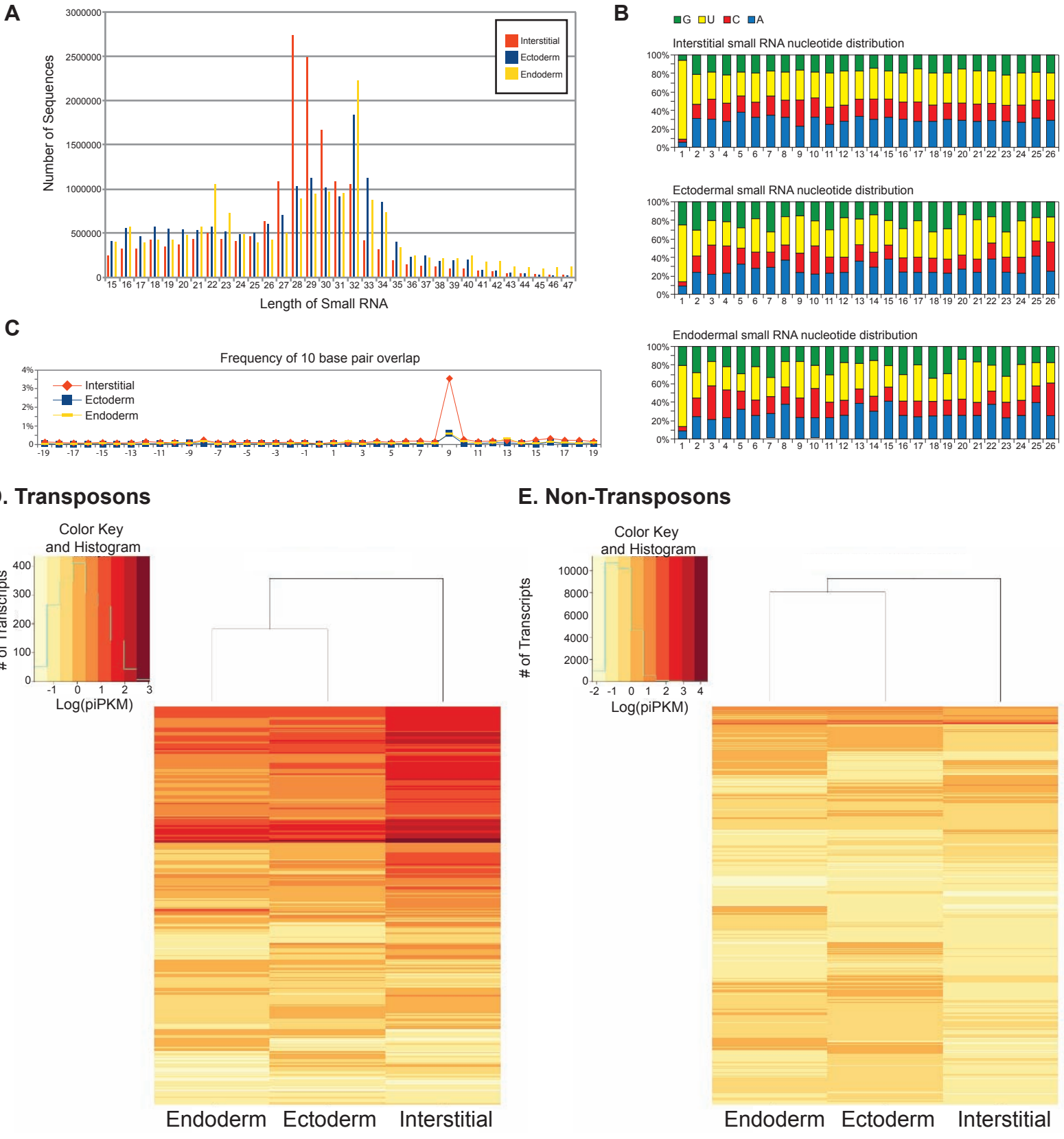
Supplemental Figure 3



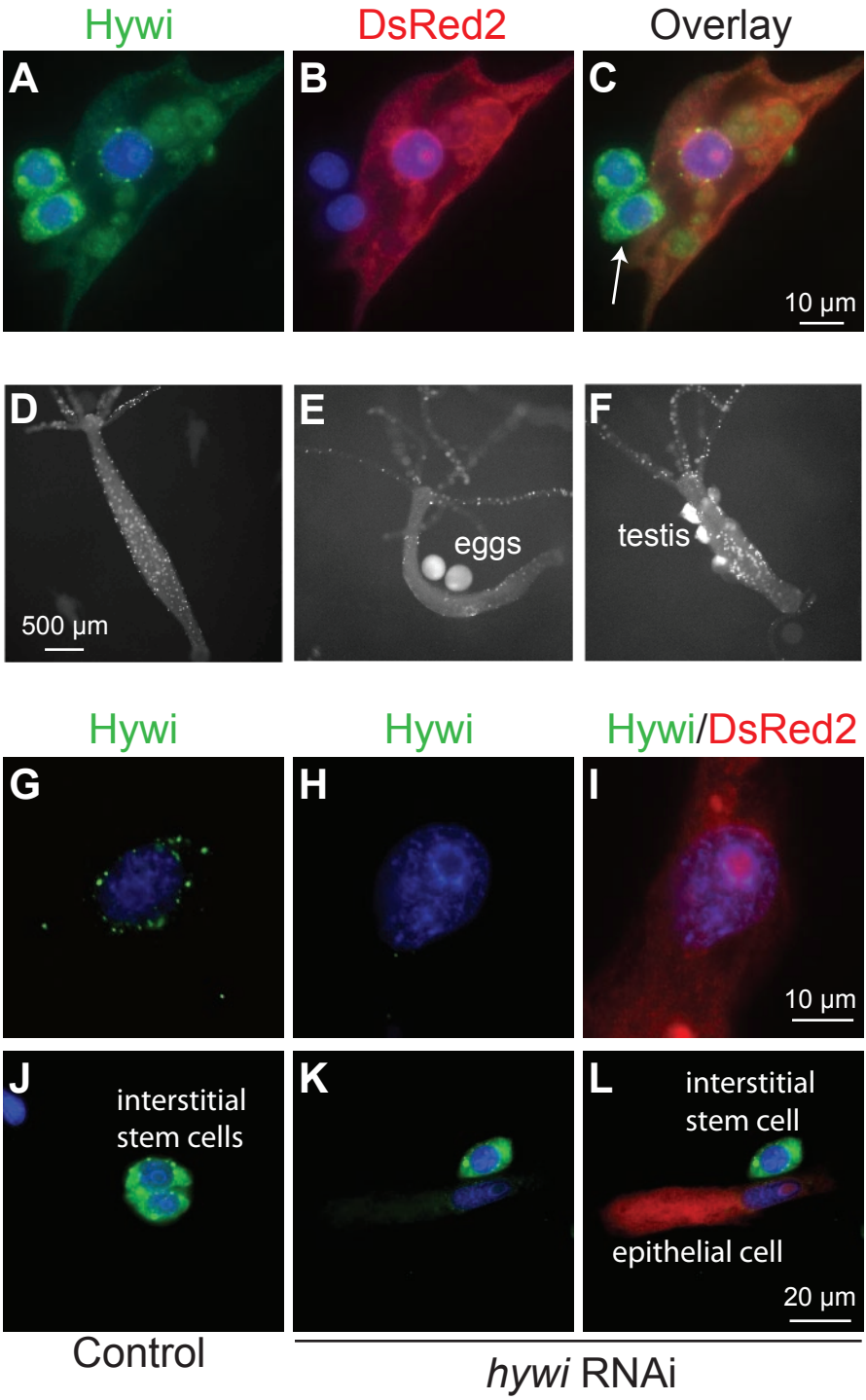
Supplemental Figure 4

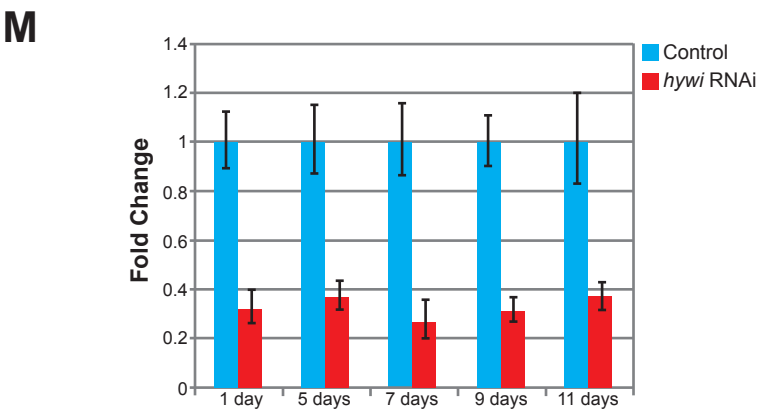


Supplemental Figure 5

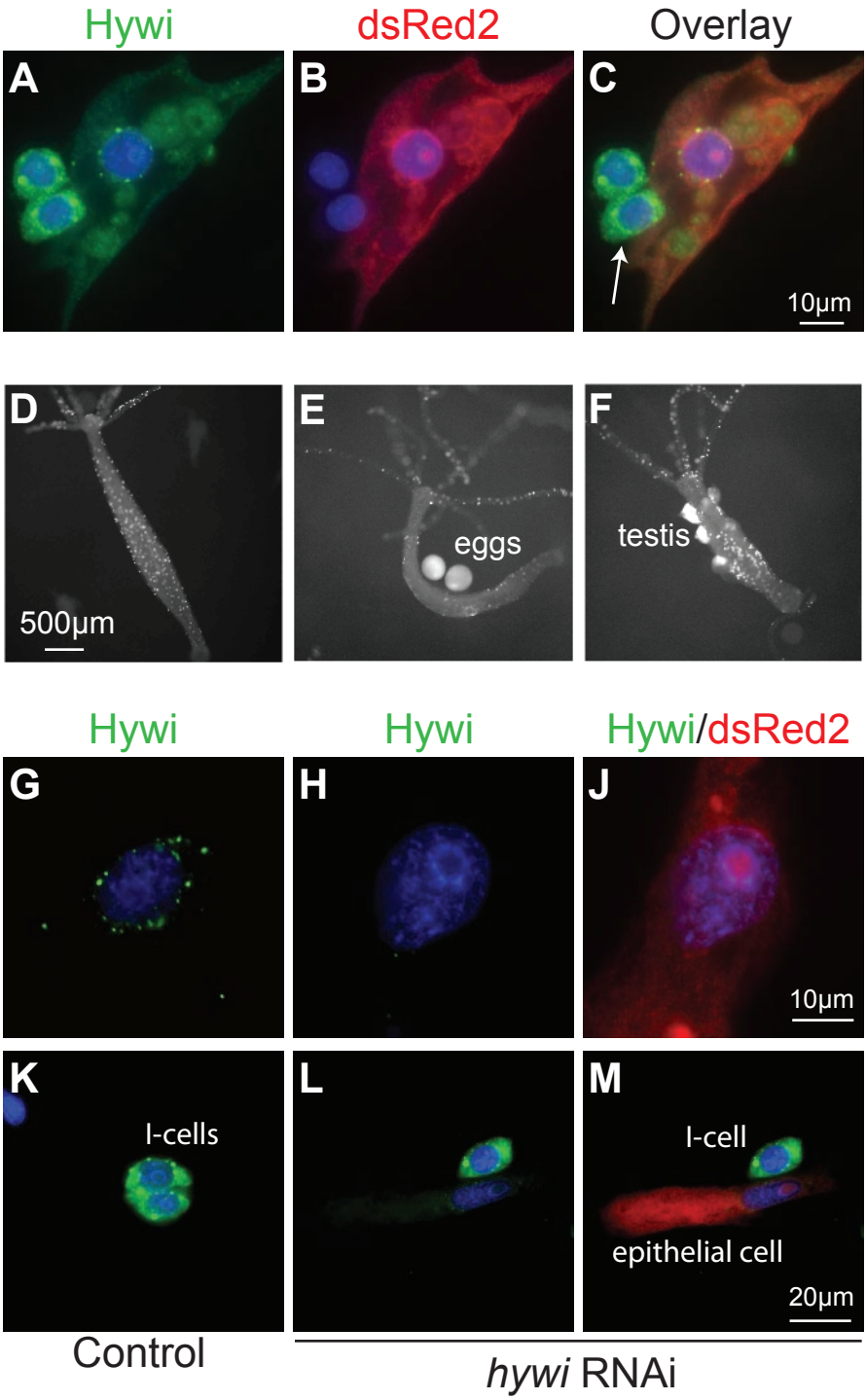


Supplemental Figure 6

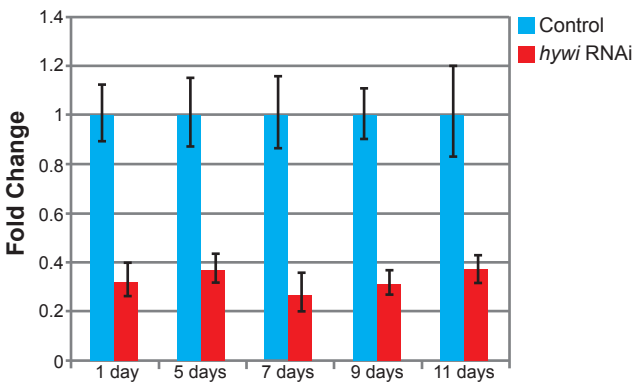




Supplemental Figure 7



N



Supplemental Figure 8

Supplemental Table1

	5'UTR		ORF		3'UTR	
Average Transcript	10%		76%		14%	
Hywi piRNAs Transposons	23%		55%		22%	
	6%	17%	14%	41%	6%	16%
Hyli piRNAs Transposons	22%		49%		29%	
	15%	7%	41%	8%	20%	9%
Hywi piRNAs non-Transposons	11%		66%		23%	
	4%	7%	59%	7%	10%	13%
Hyli piRNAs non-Transposons	18%		65%		17%	
	15%	3%	61%	4%	9%	8%

Supplemental Table 2: Gene ontology analysis of transcripts with greater than 10 Hywi-bound piRNAs mapped.

Category	Term	Count	Fold Enrichment	P-Value
Biological Process	Gastrulation With Mouth Forming First	2	25.4	7.70E-02
Biological Process	Nucleosome Assembly	5	6.7	5.60E-03
Biological Process	Neuropeptide Signaling Pathway	12	6.6	1.10E-06
Biological Process	Chromatin Assembly	5	6.3	6.80E-03
Biological Process	Translational Elongation	9	6.3	5.80E-05
Biological Process	Nucleosome Organization	5	6	8.20E-03
Biological Process	Protein-DNA Complex Assembly	5	5.5	1.10E-02
Biological Process	DNA packaging	6	4.8	7.50E-03
Biological Process	Cell Growth	4	4.6	5.30E-02
Biological Process	Epithelial Cell Differentiation	4	4.4	5.90E-02
Biological Process	Induction of Apoptosis by Extracellular Signals	4	3.9	8.02E-02
Biological Process	Chromatin Assembly or Disassembly	5	3.7	4.30E-02
Biological Process	Cellular Respiration	7	3	2.90E-02
Biological Process	Translation	30	3	1.70E-07
Biological Process	ATP biosynthetic process	6	2.9	5.20E-02
Biological Process	Purine Nucleoside Triphosphate Biosynthetic Process	6	2.7	7.20E-02
Biological Process	Purine Ribonucleoside Triphosphate Biosynthetic Process	6	2.7	7.20E-02
Biological Process	Ribonucleoside Triphosphate Biosynthetic Process	6	2.7	7.20E-02
Biological Process	ATP Metabolic Process	6	2.6	8.10E-02
Biological Process	di-, tri-valent inorganic cation transport	8	2.6	3.10E-02
Biological Process	Nucleoside Triphosphate Biosynthetic Process	6	2.6	7.60E-02
Biological Process	Sensory Perception of Light Stimulus	6	2.6	7.60E-02
Biological Process	Visual Perception	6	2.6	7.60E-02
Biological Process	Electron Transport Chain	9	2.5	2.40E-02
Biological Process	Heart Development	7	2.2	9.30E-02
Biological Process	Chromatin Organization	13	2.1	2.20E-02
Biological Process	Chromosome Organization	15	1.8	3.70E-02
Biological Process	Generation of Precursor Metabolites and Energy	13	1.8	5.00E-02
Biological Process	GPCR Signaing	16	1.8	2.70E-02
Biological Process	Regulation of Cell Proliferation	12	1.7	7.90E-02
Biological Process	Cation Transport	15	1.6	7.70E-02
Biological Process	Regulation of Biological Quality	29	1.6	1.30E-02
Biological Process	Anatomical Structure Development	41	1.4	2.90E-02
Biological Process	Cell Surface Receptor Linked Signal Transduction	25	1.4	6.30E-02
Biological Process	Cellular Protein Metabolic Process	66	1.4	4.30E-03
Biological Process	Protein Metabolic Process	73	1.3	1.20E-02
Biological Process	Signal Transduction	38	1.3	4.70E-02
Biological Process	System Development	35	1.3	9.10E-02
Biological Process	Establishment of Localization	60	1.2	4.40E-02
Biological Process	Localization	65	1.2	4.20E-02
Biological Process	Transport	59	1.2	5.50E-02
Molecular Function	Sodium:Potassium-Exchange ATPase Activity	3	24.8	4.70E-03
Molecular Function	rRNA binding	4	4.5	5.60E-02
Molecular Function	Structural Constituent of Ribosome	27	4.5	9.40E-11
Molecular Function	Cytochrome-C Oxidase Activity	4	4.1	7.00E-02
Molecular Function	Heme-Copper Terminal Oxidase Activity	4	4.1	7.00E-02
Molecular Function	Oxidoreductase Activity, Acting on Heme Group of Donors	4	4.1	7.00E-02
Molecular Function	Oxidoreductase Activity, Acting on Heme Group of Donors, Oxygen as Receptor	4	4.1	7.00E-02
Molecular Function	Structural Molecule Activity	34	3.4	4.80E-10
Molecular Function	Inorganic Cation Transmembrane Transporter Activity	12	3.2	1.10E-03
Molecular Function	Monovalent Inorganic Cation Transmembrane Transporter Activity	8	3.1	1.30E-02
Molecular Function	Cysteine-Type Endopeptidase Activity	5	2.8	9.90E-02
Molecular Function	Hydrogen Ion Transmembrane Transporter Activity	6	2.5	9.20E-02
Molecular Function	Cation Transmembrane Transporter Activity	19	1.9	9.70E-03
Molecular Function	G-protein Coupled Receptor Activity	14	1.9	2.80E-02
Molecular Function	GTP binding	14	1.9	3.70E-02
Molecular Function	Guanyl Nucleotide Binding	14	1.8	4.00E-02
Molecular Function	Guanyl Ribonucleotide Binding	14	1.8	4.00E-02
Molecular Function	Calcium Ion Binding	27	1.7	8.50E-03
Molecular Function	Ion Transmembrane Transporter Activity	21	1.7	1.80E-02
Molecular Function	Transmembrane Receptor Activity	19	1.7	3.00E-02
Molecular Function	Substrate-Secific Transporter Activity	26	1.6	2.20E-02
Molecular Function	Substrate-Specific Transmembrane Transporter Activity	22	1.6	3.80E-02
Molecular Function	Transporter Activity	32	1.6	9.70E-03

Molecular Function	Transmembrane Transporter Activity	24	1.5	3.70E-02
Molecular Function	Receptor Activity	24	1.4	7.70E-01
Cellular Component	Kinesin Complex	5	10.4	9.40E-04
Cellular Component	Nucleosome	5	9.6	1.30E-03
Cellular Component	Protein-DNA Complex	6	6.5	1.80E-03
Cellular Component	Cytosolic Large Ribosomal Subunit	5	5	1.60E-02
Cellular Component	Cytosolic Ribosome	8	4.5	1.60E-03
Cellular Component	Ribosome	28	4.1	4.10E-10
Cellular Component	Respiratory Chain	8	3.6	6.40E-03
Cellular Component	Cytosolic Part	9	3.3	5.80E-03
Cellular Component	Large Ribosomal Subunit	6	3.3	3.50E-02
Cellular Component	Ribosomal Subunit	10	3.3	3.00E-03
Cellular Component	Clatherin-Coated Vesicle	7	2.5	5.60E-02
Cellular Component	Ribonucleoprotein Complex	32	2.3	2.00E-05
Cellular Component	Extracellular Space	8	2.2	7.00E-02
Cellular Component	Extracellular Region Part	14	1.9	3.60E-02
Cellular Component	Chromosome	15	1.8	3.30E-02
Cellular Component	Cytoplasmic Membrane-Bounded Vesicles	13	1.7	8.30E-02
Cellular Component	Vesicle	17	1.7	4.80E-02
Cellular Component	Cytoplasmic Vesicle	16	1.6	6.70E-02
Cellular Component	Cytosol	28	1.5	2.40E-02
Cellular Component	Intracellular Non-Membrane-Bounded Organelle	66	1.5	2.00E-04
Cellular Component	Non-Membrane-Bounded Organelle	66	1.5	2.00E-04
Cellular Component	Plasma Membrane	59	1.4	2.20E-03
Cellular Component	Cytoplasmic Part	114	1.3	7.10E-04
Cellular Component	Intrinsic to Membranes	81	1.2	8.10E-02
Cellular Component	Macromolecular Complex	66	1.2	6.00E-02
Cellular Component	Cytoplasm	149	1.1	1.40E-02

Supplemental Table 3: Gene ontology analysis of transcripts with greater than 10 Hyli-bound piRNAs mapped.

Category	Term	Count	Fold Enrichment	P-Value
Biological Process	De Novo Posttranslational Protein Folding	3	10.1	3.10E-02
Biological Process	De Novo Protein Folding	3	10.1	3.10E-02
Biological Process	Translational Elongation	18	8.4	2.90E-12
Biological Process	Mitotic Spindle Elongation	7	7.8	1.30E-04
Biological Process	Spindle Elongation	7	7.8	1.30E-04
Biological Process	Ribosomal Small Subunit Biogenesis	3	7.2	6.10E-02
Biological Process	Nucleosome Assembly	6	5.3	4.20E-03
Biological Process	Chromatin Assembly	6	5.0	5.30E-03
Biological Process	Mitotic Spindle Organization	7	4.9	2.30E-03
Biological Process	Dentrite Morphogenesis	4	4.8	4.60E-02
Biological Process	Nucleosome Organization	6	4.8	6.60E-03
Biological Process	Imaginal Disc Development	5	4.7	1.90E-02
Biological Process	Protein-DNA complex Assembly	6	4.4	9.90E-03
Biological Process	ATP Synthesis Coupled Proton Transport	7	4.1	6.20E-03
Biological Process	Energy Coupled Proton Transport, Down Electrochemical Gradient	7	4.1	6.20E-03
Biological Process	Spindle Organization	9	4.0	1.40E-03
Biological Process	Instar Larval or Pupal Morphogenesis	4	3.9	7.60E-02
Biological Process	Translation	59	3.9	3.40E-20
Biological Process	DNA Packaging	7	3.7	1.00E-02
Biological Process	Metamorphosis	4	3.7	8.70E-02
Biological Process	Translational Initation	6	3.7	2.00E-02
Biological Process	Ion Transmembrane Transport	7	3.6	1.20E-02
Biological Process	Proton Transport	7	3.6	1.20E-02
Biological Process	Hydrogen Transport	7	3.5	1.40E-02
Biological Process	Oxidative Phosphorylation	10	3.2	3.10E-03
Biological Process	Negative Regulation of Transport	6	3.1	4.30E-02
Biological Process	Aerobic Respiration	5	3.0	8.10E-02
Biological Process	Chromatin Assembly or Dissasembly	6	3.0	4.80E-02
Biological Process	ATP biosynthetic process	8	2.6	3.30E-02
Biological Process	Nucleoside Triphosphate Biosynthetic Process	9	2.6	2.00E-02
Biological Process	Purine Nucleoside Triphosphate Biosynthetic Process	9	2.6	1.80E-02
Biological Process	Purine Ribonucleoside Triphosphate Biosynthetic Process	9	2.6	1.80E-02
Biological Process	Ribonucleoside Triphosphate Biosynthetic Process	9	2.6	1.80E-02
Biological Process	Cellular Respiration	9	2.5	2.40E-02
Biological Process	Ribonucleotide Biosynthetic Process	10	2.4	2.40E-02
Biological Process	ATP Metabolic Process	8	2.3	5.90E-02
Biological Process	Purine Nucleoside Triphosphate Metabolic Process	9	2.3	4.40E-02
Biological Process	Purine Ribonucleoside Triphosphate Metabolic Process	9	2.3	3.70E-02
Biological Process	Purine Ribonucleotide Biosynthetic Process	9	2.3	4.40E-02
Biological Process	Ribonucleoside Triphosphate Metabolic Process	9	2.3	3.70E-02
Biological Process	Electron Transport Chain	12	2.2	1.70E-02
Biological Process	Generation of Precursor Metabolic and Energy	24	2.2	3.60E-04
Biological Process	Microtubule Cytoskeleton Organization	10	2.2	3.80E-02
Biological Process	Nucleotide Biosynthetic Process	13	2.2	1.40E-02
Biological Process	Purine Nucleotide Biosynthetic Process	10	2.2	3.80E-02
Biological Process	Nucleobase, Nucleoside and Nucleic Acid Biosynthetic Process	13	2.1	1.70E-02
Biological Process	Nucleobase, Nucleoside and Nucleotide Biosynthetic Process	13	2.1	1.70E-02
Biological Process	Nucleoside Triphosphate Metabolic Process	9	2.1	6.60E-02
Biological Process	Ribonucleotide Metabolic Process	10	2.1	5.00E-02
Biological Process	Purine Ribonucleotide Metabolic Process	9	2	8.50E-02
Biological Process	Chromatin Organization	17	1.8	2.60E-02
Biological Process	Cytoskeleton Organization	17	1.8	2.70E-02
Biological Process	Microtubule-Based Process	16	1.8	2.90E-02
Biological Process	Cellular Macromolecular Complex Assembly	13	1.7	8.80E-02
Biological Process	Cellular Macromolecular Complex Subunit Organization	14	1.6	8.50E-02
Biological Process	Chromosome Organization	20	1.6	4.40E-02
Biological Process	Nitrogen Compound Biosynthetic Process	18	1.6	5.10E-02
Biological Process	Cellular Protein Metabolic Process	111	1.5	9.30E-07
Biological Process	Mitotic Cell Cycle	17	1.5	9.80E-02
Biological Process	Protein Metabolic Process	121	1.4	1.50E-05
Biological Process	Cellular Biosynthetic Process	107	1.3	2.90E-03
Biological Process	Cellular Component Biogenesis	33	1.3	7.40E-02
Biological Process	Cellular Macromolecule Biosynthetic Process	85	1.3	3.60E-03
Biological Process	Gene Expression	93	1.3	1.60E-03
Biological Process	Macromolecule Biosynthetic Process	85	1.3	3.90E-03
Biological Process	Organelle Organization	45	1.3	5.30E-02
Biological Process	Biosynthetic Process	107	1.2	7.40E-03
Biological Process	Cellular Metabolic Process	196	1.1	7.30E-02
Molecular Function	Structural Constituent of Ribosome	48	5.5	3.30E-23
Molecular Function	Translation Elongation Factor Activity	5	4.7	1.90E-02
Molecular Function	rRNA Binding	6	4.6	8.10E-03
Molecular Function	Cytochrome-C Oxidase Activity	6	4.2	1.20E-02
Molecular Function	Heme-Copper Terminal Oxidase Activity	6	4.2	1.20E-02

Molecular Function	Oxidoreductase Activity, Acting on Heme Group of Donors	6	4.2	1.20E-02
Molecular Function	Oxidoreductase Activity, Acting on Heme Group of Donors, Oxygen as Acceptor	6	4.2	1.20E-02
Molecular Function	Antioxidant Activity	4	4	7.60E-02
Molecular Function	Monovalent Inorganic Cation Transmembrane Transporter Activity	15	4.0	1.30E-05
Molecular Function	Hydrogen Ion Transmembrane Transporter Activity	14	3.9	3.60E-05
Molecular Function	Structural Molecule Activity	57	3.9	1.10E-19
Molecular Function	Histone Methyltransferase Activity	6	3.1	4.30E-02
Molecular Function	Inorganic CationTransmembrane Transporter Activity	16	2.9	2.90E-04
Molecular Function	N-methyltransferase Activity	6	2.6	7.80E-02
Molecular Function	Microtubule Motor Activity	7	2.3	7.90E-02
Molecular Function	Motor Activity	10	2.2	4.00E-02
Molecular Function	Cation Transmembrane Transporter Activity	24	1.6	1.90E-02
Molecular Function	Ion Transmembrane Transporter Activity	25	1.4	8.50E-02
Cellular Component	Mitochondrial Proton-Transporting ATP Synthase Complex, Coupling Factor F(o)	3	12.5	2.00E-02
Cellular Component	Polytene Chromosome	3	12.5	2.00E-02
Cellular Component	Cytosolic Large Ribosomal Subunit	11	7.4	7.30E-07
Cellular Component	Mitochondrial Proton-Transporting ATP Synthase Complex	3	7.2	6.10E-02
Cellular Component	Proton-Transporting V-Type ATPase Complex	3	7.2	6.10E-02
Cellular Component	Kinesin Complex	5	7.0	4.20E-03
Cellular Component	Cytosolic Ribosome	17	6.5	1.60E-09
Cellular Component	Cytosolic Small Ribosomal Subunit	5	5.6	1.00E-02
Cellular Component	Lipid Particle	5	5.6	1.00E-02
Cellular Component	Proton-Transporting ATP synthesis complex, Coupling Factor F(o)	5	5.6	1.00E-02
Cellular Component	Nuclesome	4	5.1	3.80E-02
Cellular Component	Ribosome	50	4.9	8.10E-22
Cellular Component	Small Ribosomal Subunit	8	4.5	1.60E-03
Cellular Component	Large Ribosomal Subunit	12	4.4	5.80E-05
Cellular Component	Proton-Transporting Two-Sector ATPase Complex, Proton-Transporting Domain	5	4.4	2.40E-02
Cellular Component	Ribosomal Subunit	20	4.4	5.60E-08
Cellular Component	Contractile Fiber	6	3.9	1.70E-02
Cellular Component	Cytosolic Part	16	3.9	9.40E-06
Cellular Component	Proton-Transport Two-Sector ATPase complex	7	3.8	8.90E-03
Cellular Component	Contractile Fiber Part	5	3.6	4.50E-02
Cellular Component	Protein-DNA Complex	5	3.6	4.50E-02
Cellular Component	Proton-Transporting ATP Synthase Complex	5	3.5	5.10E-02
Cellular Component	Microtubule Associated Complex	11	3.2	2.00E-03
Cellular Component	Ribonucleoprotein Complex	60	2.8	8.40E-14
Cellular Component	Mitochondrial Membrane Part	7	2.5	5.40E-02
Cellular Component	Respiratory Chain	8	2.4	4.70E-02
Cellular Component	Mitochondrial Inner Membrane	21	1.9	8.50E-03
Cellular Component	Organelle Inner Membrane	21	1.8	1.50E-02
Cellular Component	Cytosol	46	1.7	4.60E-04
Cellular Component	Intracellular Non-Membrane-Bounded Organelle	108	1.7	8.50E-09
Cellular Component	Non-Membrane-Bounded-Organelle	108	1.7	8.50E-09
Cellular Component	Miochondrial Envelope	23	1.6	2.60E-02
Cellular Component	Mitochondrial Membrane	22	1.6	2.90E-02
Cellular Component	Envelope	30	1.5	2.70E-02
Cellular Component	Macromolecular Complex	123	1.5	3.20E-07
Cellular Component	Organelle Envelopes	28	1.4	6.20E-01
Cellular Component	Cytoplasmic Part	173	1.3	9.50E-06
Cellular Component	Mitochondrion	51	1.3	2.50E-02
Cellular Component	Cytoplasm	235	1.2	1.70E-05
Cellular Component	Intracellular	297	1.1	1.60E-02
Cellular Component	Intracellular Organelle	256	1.1	2.80E-03
Cellular Component	Intracellular Part	295	1.1	1.80E-02
Cellular Component	Organelle	257	1.1	2.20E-03

Supplemental Table 4: Gene ontology analysis of putative lineage-specific targets

Endoderm

Category	Term	Count	Fold Enrichment	P-Value
Biological Process	proteolysis	6	2.9	4.10E-02
Cellular Component	collagen	4	196.7	5.10E-07
Cellular Component	proteinaceous extracellular matrix	7	16.9	1.80E-06
Cellular Component	extracellular matrix	7	16.6	2.10E-06
Cellular Component	extracellular region part	8	10.5	4.70E-06
Cellular Component	extracellular region	10	6.5	6.00E-06
Cellular Component	extracellular matrix part	5	30	1.60E-05
Cellular Component	collagen type I	2	245.9	7.80E-03
Cellular Component	fibrillar collagen	2	163.9	1.20E-02
Cellular Component	cell surface	3	8.4	4.50E-02
Cellular Component	extracellular space	3	8.1	4.80E-02
Molecular Function	extracellular matrix structural constituent	4	62.8	2.60E-05
Molecular Function	peptidase activity, acting on L-amino acid peptides	6	5.7	2.50E-03
Molecular Function	peptidase activity	6	5.5	3.00E-03
Molecular Function	endopeptidase activity	4	5.6	2.90E-02
Molecular Function	structural molecule activity	4	4.4	5.50E-02
Molecular Function	SMAD binding	2	31.4	5.90E-02
Molecular Function	metallopeptidase activity	3	6.9	6.40E-02

Ectoderm

Category	Term	Count	Fold Enrichment	P-Value
Biological Process	signal transduction	11	1.8	7.00E-02
Biological Process	cell surface receptor linked signal transduction	9	2.4	2.90E-02
Biological Process	biological adhesion	5	3.3	5.90E-02
Biological Process	cell adhesion	5	3.3	5.90E-02
Biological Process	G-protein coupled receptor protein signaling pathway	7	3.7	1.00E-02
Biological Process	neuropeptide signaling pathway	4	10.2	6.50E-03
Biological Process	negative regulation of angiogenesis	2	21.2	8.90E-02
Cellular Component	membrane	31	1.5	2.10E-03
Cellular Component	membrane part	27	1.5	1.10E-02
Cellular Component	intrinsic to membrane	24	1.7	4.20E-03
Cellular Component	integral to membrane	23	1.7	6.30E-03
Cellular Component	plasma membrane part	10	2	5.30E-02
Cellular Component	plasma membrane	19	2.3	4.90E-04
Cellular Component	extracellular region	12	3.9	1.40E-04
Cellular Component	intrinsic to plasma membrane	7	3.9	8.00E-03
Cellular Component	integral to plasma membrane	7	4	6.90E-03
Cellular Component	extracellular region part	7	4.6	3.50E-03
Cellular Component	extracellular matrix	6	7.1	1.30E-03
Cellular Component	proteinaceous extracellular matrix	6	7.2	1.20E-03
Cellular Component	extracellular matrix part	4	12	4.10E-03
Cellular Component	fibril	2	61.5	3.10E-02
Cellular Component	microfibril	2	82	2.40E-02
Molecular Function	metal ion binding	22	1.5	3.30E-02
Molecular Function	cation binding	22	1.5	3.80E-02
Molecular Function	ion binding	22	1.5	4.20E-02
Molecular Function	molecular transducer activity	13	2.7	2.00E-03
Molecular Function	signal transducer activity	13	2.7	2.00E-03
Molecular Function	receptor activity	13	3.5	2.00E-04
Molecular Function	carbohydrate binding	4	4.2	6.70E-02
Molecular Function	calcium ion binding	15	4.4	3.20E-06
Molecular Function	transmembrane receptor activity	11	4.6	9.20E-05
Molecular Function	G-protein coupled receptor activity	8	5.1	7.20E-04
Molecular Function	pattern binding	3	7.3	6.00E-02

Molecular Function	polysaccharide binding	3	7.3	6.00E-02
Molecular Function	peptide receptor activity, G-protein coupled	3	9.1	4.10E-02
Molecular Function	peptide receptor activity	3	9.1	4.10E-02
Molecular Function	extracellular matrix structural constituent	4	27.1	3.60E-04

Epithelium

Category	Term	Count	Fold Enrichment	P-Value
Biological Process	positive regulation of DNA binding	3	24	6.40E-03
Biological Process	amine metabolic process	7	3.7	9.90E-03
Biological Process	positive regulation of binding	3	18.3	1.10E-02
Biological Process	transport	20	1.7	1.30E-02
Biological Process	establishment of localization	20	1.7	1.40E-02
Biological Process	localization	21	1.6	1.70E-02
Biological Process	oxidation reduction	9	2.6	1.70E-02
Biological Process	regulation of DNA binding	3	12.5	2.30E-02
Biological Process	cholesterol metabolic process	3	10.7	3.00E-02
Biological Process	Ras protein signal transduction	3	10	3.40E-02
Biological Process	gas transport	2	51.9	3.70E-02
Biological Process	sterol metabolic process	3	9.4	3.80E-02
Biological Process	response to chemical stimulus	8	2.4	3.90E-02
Biological Process	biogenic amine metabolic process	3	8	5.20E-02
Biological Process	electron transport chain	4	4.6	5.30E-02
Biological Process	cellular amine metabolic process	5	3.3	6.00E-02
Biological Process	regulation of binding	3	6.9	6.70E-02
Biological Process	positive regulation of NF-kappaB transcription factor activity	2	25.9	7.30E-02
Biological Process	cellular amino acid and derivative metabolic process	5	2.9	8.80E-02
Biological Process	generation of precursor metabolites and energy	5	2.9	8.90E-02
Biological Process	regulation of multicellular organismal process	6	2.5	9.00E-02
Biological Process	Rho protein signal transduction	2	20.8	9.10E-02
Biological Process	positive regulation of transcription factor activity	2	20.8	9.10E-02
Biological Process	alcohol metabolic process	5	2.9	9.10E-02
Biological Process	amine biosynthetic process	3	5.8	9.20E-02
Biological Process	steroid metabolic process	3	5.6	9.80E-02
Biological Process	nitrogen compound biosynthetic process	5	2.8	9.90E-02
Cellular Component	extracellular region	9	2.4	2.80E-02
Cellular Component	extracellular region part	6	3.2	3.40E-02
Cellular Component	collagen	2	40.7	4.70E-02
Cellular Component	vacuole	5	3.5	5.20E-02
Cellular Component	external side of plasma membrane	3	7.8	5.40E-02
Cellular Component	proteinaceous extracellular matrix	4	4	7.50E-02
Cellular Component	extracellular matrix	4	3.9	7.90E-02
Cellular Component	cytoplasmic part	28	1.3	9.00E-02
Molecular Function	cofactor binding	6	3.9	1.70E-02
Molecular Function	active transmembrane transporter activity	6	3.7	2.00E-02
Molecular Function	endopeptidase activity	6	3.2	3.80E-02
Molecular Function	calcium ion binding	9	2.3	3.90E-02
Molecular Function	secondary active transmembrane transporter activity	4	5.1	4.10E-02
Molecular Function	vitamin binding	4	4.7	5.10E-02
Molecular Function	peptidase activity, acting on L-amino acid peptides	7	2.5	5.60E-02
Molecular Function	inorganic anion transmembrane transporter activity	2	33.2	5.80E-02
Molecular Function	peptidase activity	7	2.4	6.50E-02
Molecular Function	oxidoreductase activity	8	2.2	6.70E-02
Molecular Function	cysteine-type endopeptidase activity	3	6.8	7.00E-02
Molecular Function	oxidoreductase activity	3	6.5	7.50E-02
Molecular Function	cytoskeletal protein binding	5	3	8.10E-02
Molecular Function	coenzyme binding	4	3.9	8.10E-02
Molecular Function	cation binding	23	1.3	1.00E-01

Interstitial cells

Category	Term	Count	Fold Enrichment	P-Value
Biological Process	DNA integration	3	33	3.40E-03
Biological Process	M phase	6	5	5.40E-03
Biological Process	response to stress	9	2.9	7.60E-03
Biological Process	cell cycle phase	6	4.5	8.40E-03
Biological Process	protein processing	3	15	1.60E-02
Biological Process	protein maturation	3	14.5	1.70E-02
Biological Process	cell division	5	4.7	1.90E-02
Biological Process	cellular response to stress	6	3.7	2.00E-02
Biological Process	cell cycle process	6	3.6	2.10E-02
Biological Process	cell cycle	7	3	2.30E-02
Biological Process	mitosis	4	5.3	3.60E-02
Biological Process	nuclear division	4	5.3	3.60E-02
Biological Process	M phase of mitotic cell cycle	4	5.2	3.70E-02
Biological Process	regulation of biological process	19	1.5	4.10E-02
Biological Process	organelle fission	4	5	4.20E-02
Biological Process	positive regulation of biological process	8	2.3	4.30E-02
Biological Process	regulation of response to stress	3	8.2	4.80E-02
Biological Process	cellular response to stimulus	6	2.9	4.90E-02
Biological Process	DNA recombination	3	8.1	5.00E-02
Biological Process	signal transduction	9	2.1	5.20E-02
Biological Process	MAPKKK cascade	3	7.7	5.40E-02
Biological Process	positive regulation of catalytic activity	4	4.3	6.00E-02
Biological Process	regulation of multicellular organismal process	5	3.2	6.00E-02
Biological Process	response to stimulus	10	1.8	6.80E-02
Biological Process	biological regulation	19	1.4	7.40E-02
Biological Process	positive regulation of molecular function	4	3.9	7.50E-02
Biological Process	intracellular signaling cascade	6	2.4	9.00E-02
Biological Process	DNA metabolic process	5	2.8	9.30E-02
Biological Process	regulation of localization	4	3.5	9.50E-02
Biological Process	regulation of response to stimulus	3	5.5	9.80E-02
Biological Process	mitotic cell cycle	4	3.5	1.00E-01
Cellular Component	extracellular region	7	3	2.20E-02
Molecular Function	histone-lysine N-methyltransferase activity	3	15.7	1.50E-02
Molecular Function	protein-lysine N-methyltransferase activity	3	15.7	1.50E-02
Molecular Function	lysine N-methyltransferase activity	3	15.7	1.50E-02
Molecular Function	histone methyltransferase activity	3	13.3	2.00E-02
Molecular Function	N-methyltransferase activity	3	11.3	2.70E-02
Molecular Function	DNA binding	10	2.2	3.00E-02
Molecular Function	protein methyltransferase activity	3	10.5	3.10E-02
Molecular Function	S-adenosylmethionine-dependent methyltransferase activity	3	7.5	5.80E-02
Molecular Function	magnesium ion binding	5	3.1	7.30E-02