Nucleic Acids Research, 2024, **52**, 2590–2608
https://doi.org/10.1093/nar/gkad1212
Advance access publication date: 24 December 2023
RNA and RNA-protein complexes

piRNA generation is associated with the pioneer round of translation in stem cells

Sudheesh Allikka Parambil [1,2,3], Danyan Li [1,2,3], Michael Zelko [1,2,3], Axel Poulet [1,2,3] and Josien C. van Wolfswinkel [1,2,3,*]

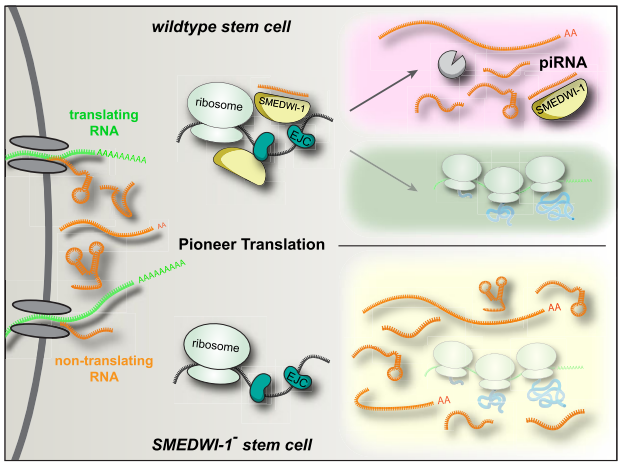
[1]Department of Molecular Cellular and Developmental Biology, Yale University, New Haven, CT 06511, USA
[2]Yale Stem Cell Center, Yale School of Medicine, New Haven, CT 06511, USA
[3]Center for RNA science and medicine, Yale School of Medicine, New Haven. CT 06511, USA

[*]To whom correspondence should be addressed. Tel: +1 203 432 3520; Email: josien.van.wolfswinkel@yale.edu
Present addresses:
Danyan Li, Department of Basic Medical Sciences, School of Medicine, Tsinghua University, Beijing, China.
Michael Zelko, Cell Molecular Biology Graduate Program, University of Chicago, Chicago, IL 60637, USA.

Abstract

Much insight has been gained on how stem cells maintain genomic integrity, but less attention has been paid to how they maintain their transcriptome. Here, we report that the PIWI protein SMEDWI-1 plays a role in the filtering of dysfunctional transcripts from the transcriptome of planarian stem cells. SMEDWI-1 accomplishes this through association with the ribosomes during the pioneer round of translation, and processing of poorly translated transcripts into piRNAs. This results in the removal of such transcripts from the cytoplasmic pool and at the same time creates a dynamic pool of small RNAs for post-transcriptional surveillance through the piRNA pathway. Loss of SMEDWI-1 results in elevated levels of several non-coding transcripts, including rRNAs, snRNAs and pseudogene mRNAs, while reducing levels of several coding transcripts. In the absence of SMEDWI-1, stem cell colonies are delayed in their expansion and a higher fraction of descendants exit the stem cell state, indicating that this transcriptomic sanitation mediated by SMEDWI-1 is essential to maintain stem cell health. This study presents a new model for the function of PIWI proteins in stem cell maintenance, that complements their role in transposon repression, and proposes a new biogenesis pathway for piRNAs in stem cells.

Graphical abstract



Introduction

PIWI proteins were initially discovered as factors involved in stem cell self-renewal in the *Drosophila* ovaries (1,2), and were subsequently identified widely in other animal germlines. The somatic expression of PIWI proteins is more limited, but they were detected in several potent stem cell types including hematopoietic stem cells and embryonic stem cells (3,4). Interestingly, expression of PIWI genes was also found to be a prime characteristic of the adult somatic stem cells of highly regenerative animals such as planarians, acoels, cnidarians, and sponges (5–8), suggesting that they may well have a conserved function in adult stem cell biology and regeneration. While much progress has been made on the elucidation of the biogenesis of piRNAs and their

Received: April 13, 2023. **Revised:** December 5, 2023. **Editorial Decision:** December 6, 2023. **Accepted:** December 8, 2023

role in transposon silencing, their involvement in stem cell maintenance has remained unclear.

Over the past decades, planarians have developed into a powerful model for regenerative biology and stem cell regulation (9). Their large population of stem cells (neoblasts) grants them the ability to recover from any injury and to replace any lost part of their body with newly generated structures (10). Early studies of the stem cells of *Schmidtea mediterranea* identified the transcript of the PIWI gene *smedwi-1* as highly specific to the stem cell population (8), and this transcript has been routinely used as a stem cell marker ever since (11–16).

Subsequent studies have shown that the SMEDWI-1 protein has a slightly wider expression range than the mRNA, and is detectable in both stem cells and early stem cell progeny, but is rapidly lost during cell differentiation (15). Notwithstanding this tightly restricted expression pattern, prior studies have not identified an essential role for SMEDWI-1 (8,17) and instead have instilled the notion that SMEDWI-1 is dispensable for overall planarian stem cell function with the exception of subtle effects on the localization of histone transcripts (18).

In contrast to this perception, we here report that SMEDWI-1 has an important function in the continuous sanitation of the planarian stem cell transcriptome. We find that SMEDWI-1 interacts with the ribosome to identify poorly coding transcripts, and mediates the degradation of aberrant transcripts while producing SMEDWI-1 bound piRNA. In the absence of SMEDWI-1, such non-coding and poorly adenylated transcripts accumulate. This negatively impacts the stability of the stem cells and their ability to respond to challenges such as expansion after irradiation.

Overall, we offer a model that introduces a new role of PIWI proteins in stem cell systems and we identify a new method of piRNA biogenesis that provides a highly flexible way to initiate PIWI-mediated transcriptomic surveillance. This study emphasizes the far-reaching role of PIWI proteins in the regulation of adult stem cells and in gene regulation.

Materials and methods

Animal husbandry

Schmidtea mediterranea asexual clonal strain ClW4 and sexual strain S2 were maintained as previously described (19). Briefly, animals were cultured in 1× Montjuic salts at 20°C, fed homogenized beef liver paste every 1–2 weeks, and expanded through continuous cycles of amputation or fissioning and regeneration. Animals were starved 1–2 weeks prior to experiments. To facilitate certain RNAi experiments, planarian water was supplemented with Gentamicin (50 μg/ml) to inhibit bacterial growth. Unless indicated, experiments were performed using asexual animals which do not contain a germline, and express *smedwi-1* only in their adult somatic stem cells (neoblasts).

RNAi, cordycepin and sublethal irradiation

Regions of planarian genes 0.5–2 kb in length were amplified from complementary DNA (cDNA) using sequence specific primers (Supplementary Table S1) with adaptor sequences. The PCR product was cloned into the pGEM-T vector (Promega) and verified by Sanger sequencing. Both RNA strands were synthesized *in vitro* from PCR-generated forward and reverse templates with flanking T7 promoters (TAATACGACTCACTATAGG), and annealed by incubation at 37°C for 30 min. The transcribed ssRNA as well as the final dsRNA product were verified by gel electrophoresis.

Animals were starved 1–2 weeks prior to RNAi experiments. RNAi food was prepared by mixing 2 μl of generic food coloring, 2 μl of dsRNA and 50 μl of homogenized beef liver (20) and fed to animals in 3-day intervals. *smedwi-2(RNAi), smedwi-1 smedwi-2(double RNAi)*, animals were fed twice (on day 0 and day 3), while *smedwi-1(RNAi), zucchini(RNAi)* and controls were fed four times (on day 0, day 2, day 7 and day 9), unless noted otherwise. DsRNA matching *C. elegans* gene *unc-22* was used as a negative control.

For disruption of RNA synthesis, animals were incubated in a final concentration of 100 μM cordycepin in planarian water for 24 h, after which they were harvested.

For sublethal irradiation animals were placed in 6cm dishes with 6ml of water and exposed to 1500 Rads of gamma irradiation after Thoreaux filter in a MultiRad350 (Precission Inc.) in exposure control mode. Animals were subsequently maintained in water supplemented with 50 μg/ml Gentamicin sulfate (VWR).

Whole-mount fluorescent *in situ* hybridization and immunofluorescence

Fixations and whole-mount *in situ* hybridizations (ISH) and immunofluorescence were performed as previously described (21), with alterations described in (22). Briefly, formaldehyde fixed animals were bleached using formamide bleach solution and treated with proteinase K (2 μg/ml) in PBSTx. For FISH, following overnight hybridization at 56°C, samples were washed sequentially in pre-hyb solution, 1:1 pre-hyb-2× SSC, 2× SSC and 0.2× SSC at 56°C. Probes were detected with anti-DIG-POD (Roche 11207733910), anti-Fl-POD (Roche 11426346910) or anti-DNP-HRP (Perkin Elmer PF1129). After tyramide development (22), peroxidase was inactivated by incubation in 1% sodium azide. Specimens were counterstained with DAPI (Sigma). For immunofluorescence animals were blocked and incubated with primary antibody overnight, followed by incubation with goat anti-Rabbit IgG HRP Conjugate (Life Technologies) or goat anti-Mouse IgG HRP Conjugate (Life Technologies). Primary antibodies used were rabbit anti-phospho-Histone3[Ser10] (Millipore, clone 63-1C-8) 1:750, and mouse anti-BrdU (BD Biosciences). Signals were developed using Tyramide SuperBoost™ Kits (Invitrogen).

Neoblast isolation and staining

Neoblasts in G2/M phase (X1), or G0/G1 phase (X2), and differentiated cells (Xins) were isolated by Fluorescence-Activated Cell Sorting based on DNA content (Hoechst fluorescence) as reported by (23), following the procedures described previously (13).

For staining of the isolated cells, cell suspensions isolated by FACS were collected in CMFB and centrifuged at ∼300 g for 5 min at 4°C. Cells were washed in CMF, spotted onto poly-D-lysine coated coverslips (BD Biosciences), allowed to settle for ∼30 min, and fixed in 4% PFA (in PBS) for 20 min at room temperature. For SYTO RNAselect staining, fixation instead was performed for 10 min at −20°C in ice-cold methanol. Controls and treatment were always spotted on the same cover slip, and went through all staining steps using the same

solutions in the same well. IF and FISH labelings were carried out similarly to the whole-mount protocol, with wash steps and antibody incubations shortened to 10 min and 1 hour, respectively.

Microscopy and image analysis

Images were taken on a Zeiss LSM800 Confocal Microscope. Control and RNAi animals were imaged with the same magnification, laser intensity and gain, at comparable anatomical position. Cell counting and quantification of fluorescence intensity were performed in Fiji (24), or by using PIQ (Planaria Image Quantification; https://gitlab.com/vanwolfswinkel/PIQ).

The quantification of EdU-positive neoblasts was performed in PIQ using 2D images. We used the 'neigh' segmentation method (adapted from NODeJ (25)) with separate settings for detection of EdU signal (Gaussian: 2.0, Factor: 2.0, NeighborhoodSize: 5, MaxVolumeObjectFilter: 50.0, MinVolumeObjectFilter: 10.0) and for the neoblasts (Gaussian: 1.0, Factor: 1.0, NeighborhoodSize: 2, MaxVolumeObjectFilter: 100.0, MinVolumeObjectFilter: 5.0).

For analysis of the isolated neoblasts, PIQ was used to compute cell size and intensity in 3D. The 'neigh' segmentation method was used with the following parameters: Gaussian: 1.0, Factor: 3.0, NeighborhoodSize: 7, MaxVolumeObjectFilter: 300.0, MinVolumeObjectFilter: 30.0.

qPCR analysis

Total RNA was isolated by TRIzol (Invitrogen) and quantified by Qubit. To distinguish between changes in global RNA level or in level of polyadenylation, two cDNA preparations were synthesized from the same RNA samples: one primed by hexamers and one primed by oligo dT. For several non-coding RNAs a mix of dedicated RT primers was added to the reaction; sequences are listed in Supplementary Table S1. For all cDNA preparations ProtoScriptII (NEB) was used according to the manufacturer instructions, using 1 μg RNA as starting material in a 20 μl reaction. cDNA was diluted 1:5 in MilliQ water and 1 μl was applied to a 10 μl qPCR reaction using EvaGreen master mix (Biotium). Primers are listed in Supplementary Table S1. RT and qPCR reactions of samples and controls were run in parallel in the same plates. qPCRs were run on a QuantStudio 3 instrument (ABI) with the following program: 95°C, 20 s; 40 cycles of 95°C, 5 s; 60°C, 20 s; followed by a melting curve analysis.

RNA-seq library generation

For mRNA-seq libraries of bulk neoblasts and differentiated tissues, large animals were fed control *unc-22* or *smedwi-1* dsRNA in liver for at least 3 weeks. RNA was extracted from isolated tissues using TRIzol Reagent (Invitrogen), and libraries were generated using TruSeq RNA Library Prep Kit v2 (Illumina) following manufacturer's instructions.

For generation of RiboMinus libraries, a dedicated planarian probeset (siTOOLs Biotech) was used to deplete RNA samples from rRNA sequences following the manufacturer's instructions. Remaining RNA was used for library synthesis using the TruSeq RNA Library Prep Kit v2 (Illumina), starting after the polyA selection steps.

Expression and purification of SMEDWI-1 protein

The N-terminal domain of SMEDWI-1 protein (aa 1–157) was cloned into pET28 in front of a GST tag. Constructs were verified by Sanger sequencing, and transformed into *E. coli* strain BL21 for expression. SMEDWI-GST protein was purified using Pierce Glutathione Agarose (Thermo Scientific) following manufacturer's instruction. The GST tag was cleaved using PreScission Protease (GE Healthcare), and removed by passage over the glutathione column.

SMEDWI-1 antibody generation

Protein antibody against SMEDWI-1 was generated by Cocalico Biologicals (Stevens, PA). The serum was purified using antigens immobilized on CNBr columns, as described by (26). To eliminate cross-reactivity, antibody was negatively selected on column with *in vitro* expressed SMEDWI-2 N-terminal domain.

As a separate approach, peptide antibody was generated by immunization of rabbits with a peptide derived from the PAZ domain of SMEDWI-1, followed by affinity purification of the antibody from serum. Western blots and immunofluorescent stainings of SMEDWI-1 have been verified with both antibodies. Verification of these antibodies is shown in Supplementary Figure S6.

Immunoprecipitation (IP) of SMEDWI-1

Starting material consisted of 10 million total cells. Cells were lysed in RIPA lysis buffer supplemented with 5% glycerol, 0.3% Western Blot Buffer (Roche), and protease inhibitor (Roche)). RNase treatment was performed by incubation for 30 min at 37°C with 50 μg RNaseA per ml lysate. 10% of the lysate was taken as input. Depending on the lysate, 0.25–1.25 μg of anti-SMEDWI-1 protein antibody and 5–25 μl of Protein A magnetic beads (NEB) were added to the sample. Immunoprecipitation was performed following the NEB protocol. Small RNA was extracted using TRIzol.

Protein co-IP was performed by incubation with antibodies overnight at 4°C followed by incubation with Dynabeads Protein A/G (Invitrogen) and nutated for 2 h at 4 °C. The magnetic beads were washed four times with IP buffer and the proteins were eluted from the beads with SDS-PAGE sample buffer.

Small RNA library generation

Following TRIzol extraction, IP or input small RNAs were further cleaned up either by eliminating longer RNAs using AmPure (Beckman Coulter), or by selecting for piRNAs with 3′ modifications using NaIO4 mediated oxidation (27). Small RNA libraries were generated following a protocol adapted from the Zamore lab (28). Briefly, end-blocked, 5′ adenylated 3′ adaptor was ligated using T4 RNA Ligase 2, truncated KQ (NEB). Excess RT primer was added to anneal with the remaining 3′ adaptor, and 5′ adaptor was ligated using T4 RNA Ligase 1 (NEB). Reverse transcription was performed using AMV reverse transcriptase (NEB). PCR amplified libraries 144–158 bp in size were selected by PAGE.

Processing of mRNA-seq data

mRNA libraries were sequenced on HiSeq 2500 or NovaSeq (Illumina). Reads were mapped against *Schmidtea mediterranea* transcriptomes WIX1 (13), dd_Smed_v6 (29)

or unigene (30) using Bowtie2 (31), and further processed with SAMtools (32). For transposon expression analysis, reads were mapped against the *Schmidtea mediterranea* SMESG.1 genome (33) using STAR (34). Genomic locations of transposon fragments were identified using RepeatMasker (35), based on consensus sequences in Repbase libraries (36).

To investigate upregulation of transposons, RNAseq reads were mapped to individual transposon copies and ambiguously mapped reads were assigned to the most probable source using Telescope (37). Transposon copies with a minimum average expression of 1 RPM per sample were filtered for differential expression analysis using DESeq2 (38).

Processing of small RNA-seq data

Small RNA libraries were sequenced on NovaSeq, and 10–25M reads were generated for each sample. Following adaptor trimming by Cutadapt (39), reads were mapped against the *Schmidtea mediterranea* SMESG.1 genome using Bowtie (40), allowing for 2 mismatches and up to 20 mapping locations. The reads are then counted strand-sensitively toward exons of transcripts or transposon copies using BEDTools.

To investigate targets of SMEDWI-1 associated piRNAs, adaptor trimmed reads were mapped sequentially against the following genomic sequences: *Schmidtea mediterranea* rRNA, tRNA, other structural RNA, miRNA, transposons, the transcriptome, and the genome. At each category, mapping with up to 2 mismatches allowed was attempted using Bowtie, mapped reads were excluded, and the rest of the library was used for mapping in the next category. Reads that were not accepted in any of the categories are marked as mapping to 'none'.

To investigate enrichment of SMEDWI-1 associated piRNAs over input, piRNAs from input or pull-down libraries were mapped to individual transcripts and transposon copies, and the ratio between input and pull down piRNAs was calculated per transcript.

To investigate the relationship between piRNA count and transposon/transcript expression, piRNA were mapped to transposon consensus or transcripts, allowing up to 2 mismatches and randomly assigning multi-mappers to one of their best matching locations. Changes in transposon/transcript expression were plotted against the counts of SMEDWI-1 associated piRNAs.

Mass spectrometry

Samples were loaded onto a 10% acrylamide gel and separated at 160 V for 10 min. The portion of the gel with sample was diced and subjected to in-gel trypsin digestion. Peptides were extracted with 5% formic acid/50% ACN, concentrated with vacuum centrifugation, and desalted using a C-18 ziptip (Millipore). Peptide separation was performed on an EASY-nLC 1200 system (Thermo Scientific) using an analytical PicoFrit column (New Objective) (75 μm × 50 cm length) self-packed with ReproSil-Pur 120A C18-Q 1.9 μm (Dr Maisch GmbH). An Orbitrap Fusion Lumos Tribrid mass spectrometer (Thermo Scientific) with a NanoFlex ion source was coupled to the LC platform. The DIA-MS method consisted of a MS1 survey scan and 40 MS2 scans of variable windows. The default peptide charge state was set to 2. Both of MS1 and MS2 spectra were recorded in profile mode.

All MS/MS samples were analyzed using Mascot (Matrix Science, London, UK; version 2.8.0 in Proteome Discoverer 2.5.0.400) on a custom generated planarian proteome database assuming only trypsin digestion. Mascot was searched with a fragment ion mass tolerance of 0.020 Da and a parent ion tolerance of 10.0 PPM. Deamidation of asparagine and glutamine, oxidation of methionine and carbamidomethylation of cysteine were specified in Mascot as variable modifications.

Scaffold (version Scaffold_5.2.2, Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 95.0% probability. Protein identifications were accepted if they could be established at greater than 99.0% probability and contained at least 2 identified peptides.

Polysome profiling

Polysome profiling was carried out as described in (41). Briefly, worms were treated with cycloheximide (100 μg/ml) for 24 hours and lysates were prepared in a pre-chilled polysome extraction buffer (20 mM Tris–HCl pH 7.5, 150 mM NaCl, 1.5 mM $MgCl_2$, 0.6% Triton X-100, 0.5 mM DTT, 1× protease inhibitor, cycloheximide (100 μg/ml)). Cleared lysate was layered on the top of a 10–50% sucrose density gradient prepared in the polysome extraction buffer and subjected to ultracentrifugation at 40 000 RPM for 16 h at 4°C (Beckman SW 41 Ti rotor). Fractions were collected by upward displacement and the absorbance was monitored at 254 nm. For further analysis, RNA was purified using TRIzol extraction, and protein was isolated by TCA precipitation.

For immunoprecipitations from monosome and polysome fractions, protein content was concentrated by ultracentrifugation at 100 000 g for 2 h at 4°C (Beckman 50 Ti rotor). Pellets were solubilized in Urea IP buffer (1 × TBS, 1% SDS, 4 M urea, 10 mM EDTA, and Complete protease inhibitor cocktail) by heating at 42°C for 30 min. Insoluble material was removed by centrifugation at 15 000 g for 10 minutes, and excess urea and SDS were removed by dialysis using Amicon Ultra-15 centrifugal filters (Millipore). Immunoprecipitation was performed as described above.

SDS-PAGE and western blotting

Individual 1–3mm sized animals were homogenized in protein loading buffer (60 mM Tris–Cl pH6.8, 5% glycerol, 1% SDS and 2.5% β-mercaptoethanol) and separated on 8–10% denaturing polyacrylamide gel. Samples were transferred to PVDF membrane, blocked, and incubated with the primary antibody followed by secondary antibody, in PBSTw (1× PBS with 0.1% Tween-20) containing 1.5% milk. The following commercial antibodies were used: mouse anti-α-tubulin (Millipore) at 1:10 000; rabbit anti-RPL10 at 1:1000 (Invitrogen); rabbit anti-RPS17 at 1:1000 (Invitrogen); rabbit anti-eIF4A3 at 1:300 (Abcam); goat anti-Rabbit IgG HRP conjugate at 1:5000 and goat anti-Mouse IgG HRP Conjugate (Life Technologies) at 1:2500.

RNA gels

For detection of piRNAs in *zucchini(RNAi)* samples, total RNA was run on a 12% denaturing acrylamide gel and stained with SYBR gold for 30 min at room temperature. Analysis of RNA from *smedwi-1(RNAi)* and control neoblasts was performed by radioactive labeling. RNA was end-labeled with 32P, separated on a 12% denaturing acrylamide gel,

vacuum dried, and imaged using a phosphorimager screen on a GE Typhoon FLA 9000 gel imager.

RNase protection assay was performed using an internally labeled ssRNA probe purified by size selection from gel. The probe was hybridized with the RNA sample overnight at 46°C, followed by digestion with RNaseA (40ug per sample), RNaseT1 (1U), and RNase One (20U) for 45 min at 40°C, followed by 30 min at 37°C. RNases were inactivated by Proteinase K digestion. Protected RNA was purified by TRIzol extraction, separated on a 12% denaturing acrylamide gel, vacuum dried, and imaged using a phosphorimager screen on a GE Typhoon FLA 9000 gel imager.

Quantification and statistical analysis

Levels of significance were calculated with two-tailed Student's *t* test, using the Prism software package. Analysis of genome-wide data was carried out as described above.

Results

Depletion of SMEDWI-1 results in delayed neoblast expansion

Previous studies had reported that *smedwi-1(RNAi)* treatment did not result in a detectable phenotype (8,17). We first wanted to verify that homeostatic *smedwi-1(RNAi)* animals showed no detectable changes in the composition of the stem cell population. Indeed, based on qPCR for several neoblast-specific genes we did not detect a reduction in neoblast levels upon loss of SMEDWI-1 (Supplementary Figure S1A). We also determined the fraction of neoblasts in S/G2/M phase (X1 cells) by FACS analysis based on nuclear DNA content, and found that *smedwi-1(RNAi)* animals do not show a reduction in the fraction of cycling cells (Figure 1A, Supplementary Figure S1B). Both these observations were confirmed by staining of intact animals with a probe against the stem cell marker *bruli* and with antibodies against the mitotic marker phosphorylated Histone 3 (H3P) (Figure 1B). To gain a broader perspective on the state of the neoblasts we also isolated homeostatic neoblasts for RNA sequencing, and found no major changes in the expression of common neoblasts markers or in markers for any of the major neoblast classes (Figure 1C).

To identify more subtle effects on stem cell function, we subjected *smedwi-1(RNAi)* and control animals to sublethal irradiation (Figure 1D). This procedure strongly reduces the number of neoblasts, followed by expansion of the surviving stem cells. The dynamics of the stem cell expansion process can reveal alterations in stem cell dynamics that would remain undetected under homeostatic conditions (42). Upon irradiation, the stem cell number in *smedwi-1(RNAi)* animals was reduced to a similar extent as in control animals (Figure 1E). However, 10 days after irradiation, the number of stem cells was significantly lower in the *smedwi-1(RNAi)* animals, indicating that neoblast expansion was delayed (Figure 1E,F). This slower expansion of the neoblast population could be caused by a higher rate of cell death, a higher rate of cell differentiation, a reduced fraction of stem cells that is actively cycling, or a slower cell cycle rate. Based on TUNEL staining we found that the rate of apoptosis was not significantly altered upon loss of SMEDWI-1 (Supplementary Figure S1C), arguing against changes in the rate of cell death. Similarly, the proportion of the stem cell population residing in M-phase of the cell cycle, as determined by phosphorylation of histone H3 (H3P), was not reduced in *smedwi-1(RNAi)* animals compared to controls (Figure 1B), suggesting no reduction in the fraction of stem cells that is actively cycling. In contrast, the fraction of cells in S-phase as determined by the active incorporation of the thymidine analog EdU into newly synthesized DNA was increased in *smedwi-1(RNAi)* animals compared to controls (Figure 1G). This suggests that *smedwi-1(RNAi)* neoblasts spend a larger fraction of their cycle in S-phase compared to control stem cells. Further, the ratio between early epidermal progeny cells (as marked by *prog-1* and *prog-2*) and neoblasts (as marked by *bruli* and *h2b*) was increased in the absence of SMEDWI-1 (Figure 1H), suggesting increased cell differentiation and reduced neoblast self-renewal in the *smedwi-1(RNAi)* animals.

Together, these data show that *smedwi-1(RNAi)* neoblasts have impaired ability to recover from the stress applied by irradiation, and indicate that this is due to compromised neoblast self-renewal and reduced ability to maintain the neoblast state.

SMEDWI-1 functions as a back-up layer of transposon defense in the stem cells

To gain insight into how these changes in stem cell dynamics relate to the loss of SMEDWI-1, we focused on the bound small RNAs (piRNAs) that target SMEDWI proteins to specific transcript sequences. PIWI proteins are best known for their role in transposon silencing in animal germlines and have been found to play similar roles in the stem cells of regenerative animals (43,44). Indeed, planarian PIWI protein genes *smedwi-2* and *smedwi-3* were previously reported to bind piRNAs related to transposon sequences and function in transposon repression in various tissues and contexts (45,46). SMEDWI-1-bound piRNAs overlapped significantly with SMEDWI-2-bound piRNAs. They had a strong 5′U bias, and most of them mapped to transposons or repeat-associated regions of the genome (Figure 2A). They were distinguished from SMEDWI-2-bound small RNAs by a difference in length preference (32nt for SMEDWI-1 versus 33nt for SMEDWI-2) and a significantly higher prevalence of sequences matching structural RNA (Supplementary Figure S2A-C).

We previously reported that SMEDWI-2 plays an important role in the silencing of transposons during the process of cell differentiation (46). To determine whether SMEDWI-1 also functions in the repression of transposons, we analyzed the levels of transposon transcripts in the RNAseq libraries from *smedwi-1(RNAi)* neoblasts and differentiated tissues. In *smedwi-1* neoblasts around 50 transposons were altered, of which half were upregulated up to 3-fold (Figure 2B). Upregulated transposons were strongly biased towards DNA elements, and most had very low baseline expression, rendering the 3-fold increased levels still barely detectable. In differentiated cells similarly around 50 transposons changed in expression, half of which were upregulated, and most were lowly expressed DNA elements that had modest fold changes. Compared to the (increased) transposon levels upon knockdown of SMEDWI-2, the expression of transposon transcripts in the absence of SMEDWI-1 is indistinguishable from the background levels in control samples (Supplementary Figure S2D). In summary no major changes in

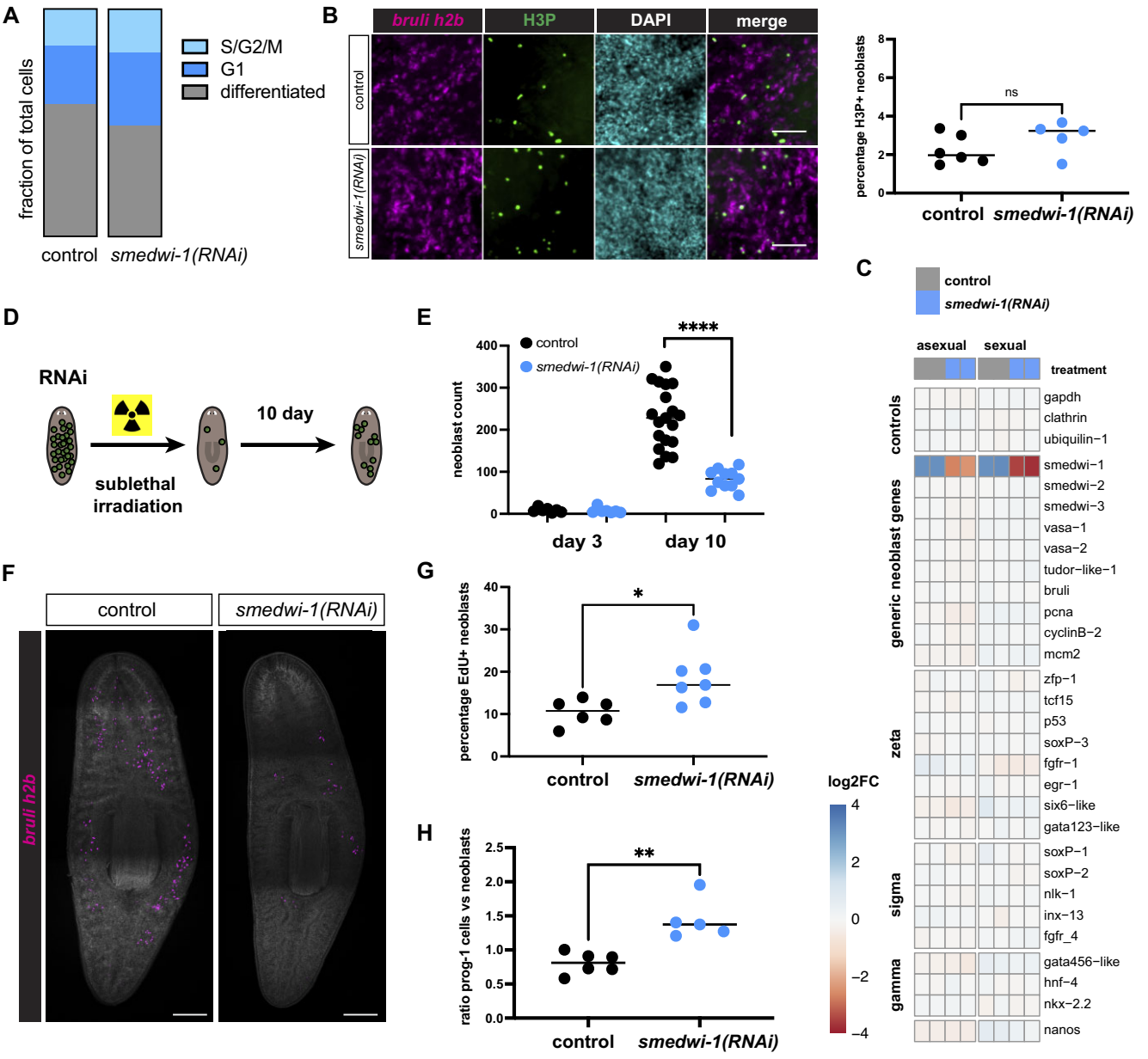


Figure 1. Loss of SMEDWI-1 leads to delayed neoblast expansion. (**A**) Cell composition of control and *smedwi-1(RNAi)* animals was analyzed by flow cytometry based on staining with Hoechst (Supplementary Figure S1B). (**B**) Representative images of control and *smedwi-1(RNAi)* animals stained with RNA probe against neoblast genes *bruli* and *h2b* to reveal neoblasts, and antibody against phosphorylated histone 3 (H3P) to show cells in M-phase. Scale bar 50 μm. Right: quantification of the percentage of mitotic neoblasts in control and *smedwi-1(RNAi)* animals. Statistical significance was determined using a *t*-test: ns: not significant. (**C**) Heatmap of RNA expression levels of control genes and neoblast genes in control and *smedwi-1(RNAi)* neoblasts as determined by RNAseq analysis. Shown are log2 fold changes relative to the averaged expression. (**D**) Workflow schematic of the sublethal irradiation experiment. (**E**) Number of neoblasts at d3 and d10 after irradiation at 1500 Rads in individual control and *smedwi-1(RNAi)* animals. Statistical significance was determined using a t-test: **** $P < 0.0001$. (**F**) Representative images of neoblasts (as labeled with RNA probe against neoblast genes *bruli* and *h2b)* on d10 after 1500 Rads. Scale bar 100μm. (**G**) Quantification of the number of cells in S-phase as determined by EdU incorporation, normalized to the total number of neoblasts quantified by FISH. Statistical significance was determined using a *t*-test: * $P < 0.05$. (**H**) Quantification of the number of early progeny cells normalized to the number of neoblasts, both quantified by FISH. Statistical significance was determined using a *t*-test: ** $P < 0.01$.

transposon regulation were detected in the *smedwi-1(RNAi)* animals.

The lack of transposon deregulation in *smedwi-1(RNAi)* animals could be explained by the continued presence of the nuclear PIWI protein SMEDWI-2. We had previously reported that SMEDWI-2 coordinates the silencing of repetitive regions at the chromatin level, and thus very low levels of transposon transcripts are generated while SMEDWI-2 is active (46). We previously reported that in the absence of SMEDWI-2 protein, transposon transcripts are readily detected in the differentiated cells, but not in the stem cells (46) (Figure 2C), and proposed that one of the other PIWI proteins likely constitutes an additional mechanism for transposon control specifically in the stem cells. To test whether SMEDWI-1 could act as this stem cell-specific backup, we evaluated the level of transposon transcript in the stem cells in the

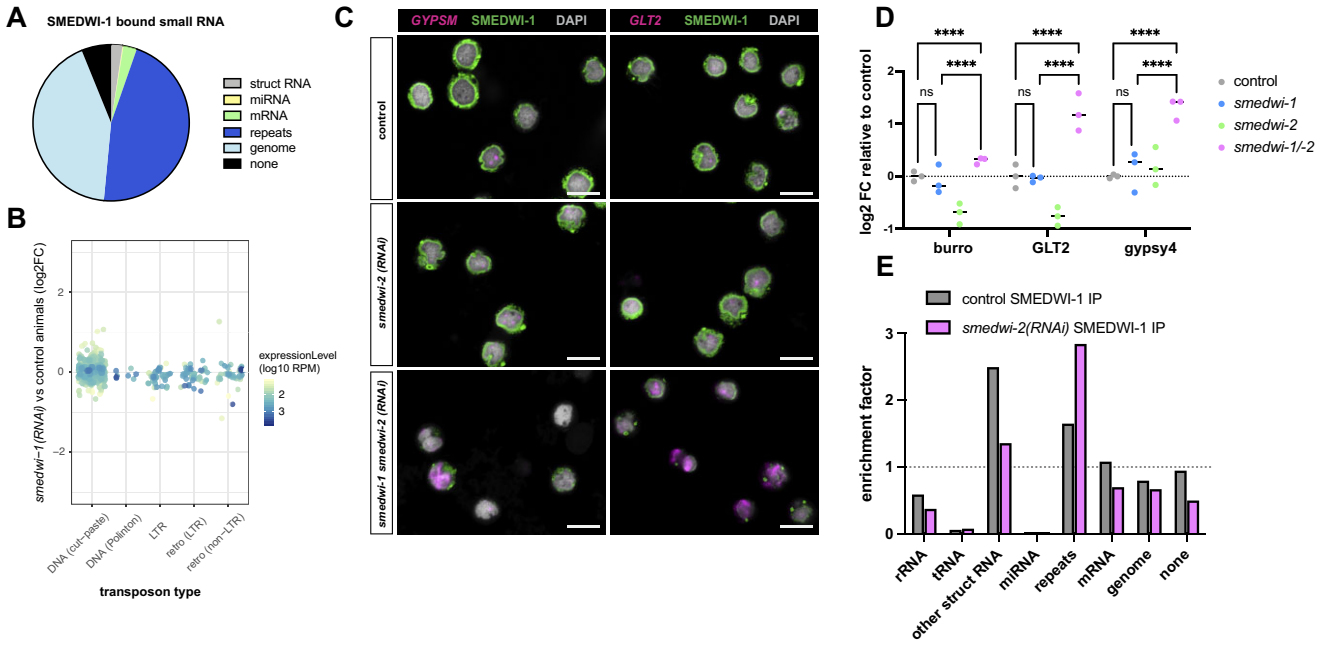


Figure 2. SMEDWI-1 provides a backup defense against transposons in the neoblasts. (**A**) Relative proportions of small RNA reads mapping to various genetic elements in the SMEDWI-1 bound piRNAs. The reads mapping to the genome are largely located in repetitive regions associated with transposons (see (46)). (**B**) Effect of loss of SMEDWI-1 on transposon levels in whole animals. Transposons are classified by type and their baseline expression level is shown by coloring. No major changes in transposon levels are detected (see also Supplementary Figure S2D). (**C**) Stainings of isolated neoblasts from control animals, *smedwi-2(RNAi)* animals, and *smedwi-1/-2(RNAi)* animals. Shown are SMEDWI-1 protein (green) and transposon transcripts *GypSM* and *GLT2* (magenta). Only upon loss of both SMEDWI-2 and SMEDWI-1 are transposons desilenced in the neoblasts. Scale bar 10 μm. (**D**) qPCR of transposon transcripts in neoblasts from control animals, *smedwi-1(RNAi)* animals, *smedwi-2(RNAi)* animals, and *smedwi-1/-2(RNAi)* animals. Transposon levels are significantly elevated in the *smedwi-1/-2(RNAi)* neoblasts, but not in the *smedwi-1 (RNAi)* or *smedwi-2 (RNAi)* neoblasts. Statistical significance was determined using a *t*-test: **** $P < 0.0001$. (**E**) Enrichment of small RNAs in SMEDWI-1 immunoprecipitation compared to total small RNA input in control samples and *smedwi-2 (RNAi)* samples. Upon loss of SMEDWI-2, repeat elements take a more prominent place in the SMEDWI-1 bound small RNAs, whereas contributions of all other RNAs are reduced.

absence of one or both PIWI proteins. We found that if both SMEDWI-2 and SMEDWI-1 are depleted, transposon transcripts are clearly detected in a subset of neoblasts by Fluorescent In Situ Hybridization (FISH) (Figure 2C). This indicates that SMEDWI-1 functions as a secondary filter on the transcripts from SMEDWI-2 targeted transposons. This notion was confirmed by qPCR: whereas loss of SMEDWI-2 or SMEDWI-1 alone had no effect on the levels of several transposons, the loss of both SMEDWI-1 and SMEDWI-2 in the same samples resulted in a significant increase in transposon levels in the stem cells (Figure 2D). Further, the fraction of repeat-targeting piRNAs bound to SMEDWI-1 was increased in the absence of SMEDWI-2 (Figure 2E), indicating that SMEDWI-1 responds to fight the increased transcriptional load of transposons when chromatin silencing through SMEDWI-2 is lost. The functional defect of the *smedwi-1(RNAi)* neoblasts as described in Figure 1 (Figure 1E–H) however is observed in the presence of SMEDWI-2 protein, and thus under maintained chromatin-mediated silencing of the transposons. Deregulation of transposon sequences therefore is unlikely to be the cause of this observed *smedwi-1(RNAi)* phenotype.

Loss of SMEDWI-1 results in an increase in structural RNA molecules

The finding that SMEDWI-1 functions as a (backup) filter for transposon transcripts in the stem cells motivated us to look more closely at the regulation of other transcripts that match SMEDWI-1-bound piRNAs. After transposon-related sequences, the largest contribution to the SMEDWI-1 piRNAs was from structural non-coding RNA (ncRNA), in particular rRNA and snRNA (Figure 3A). While rRNA sequences were not proportionally enriched in the SMEDWI-1-bound piRNAs relative to the total small RNA (Figure 2E, Supplementary Figure S3A), rRNA-related small RNAs are highly abundant in planarian cells, and thus they still made up a considerable fraction of the read library. Furthermore, rRNA-related piRNAs were more abundant in SMEDWI-1 immunoprecipitations (IPs) than in precipitations of the related PIWI protein SMEDWI-2, suggesting at least a specific preference for the SMEDWI-1 protein. In particular small RNAs matching the 5.8S rRNA were highly represented in the SMEDWI-1 IP. Small RNAs from snRNA are quite rare in the total piRNA population, and were depleted from SMEDWI-2 IPs, but were significantly enriched in the SMEDWI-1 IPs (Figure 2E, Supplementary Figure S3B).

rRNAs and snRNA are not polyadenylated, and thus should not appear in oligo-dT selected RNA sequencing data. However, we found significant numbers of reads in our libraries, and the fraction of reads from these transcripts was significantly higher in *smedwi-1* RNAi samples compared to controls (Figure 3B). Transcripts for all three annotated rRNAs (18S, 5.8S and 28S), and several snRNAs (U2, U4 and U6) were elevated upon loss of SMEDWI-1. To independently confirm whether the levels of rRNA and snRNA were altered in *smedwi-1(RNAi)* samples, we performed qPCR using random hexamers for cDNA synthesis (Figure 3C), and FISH

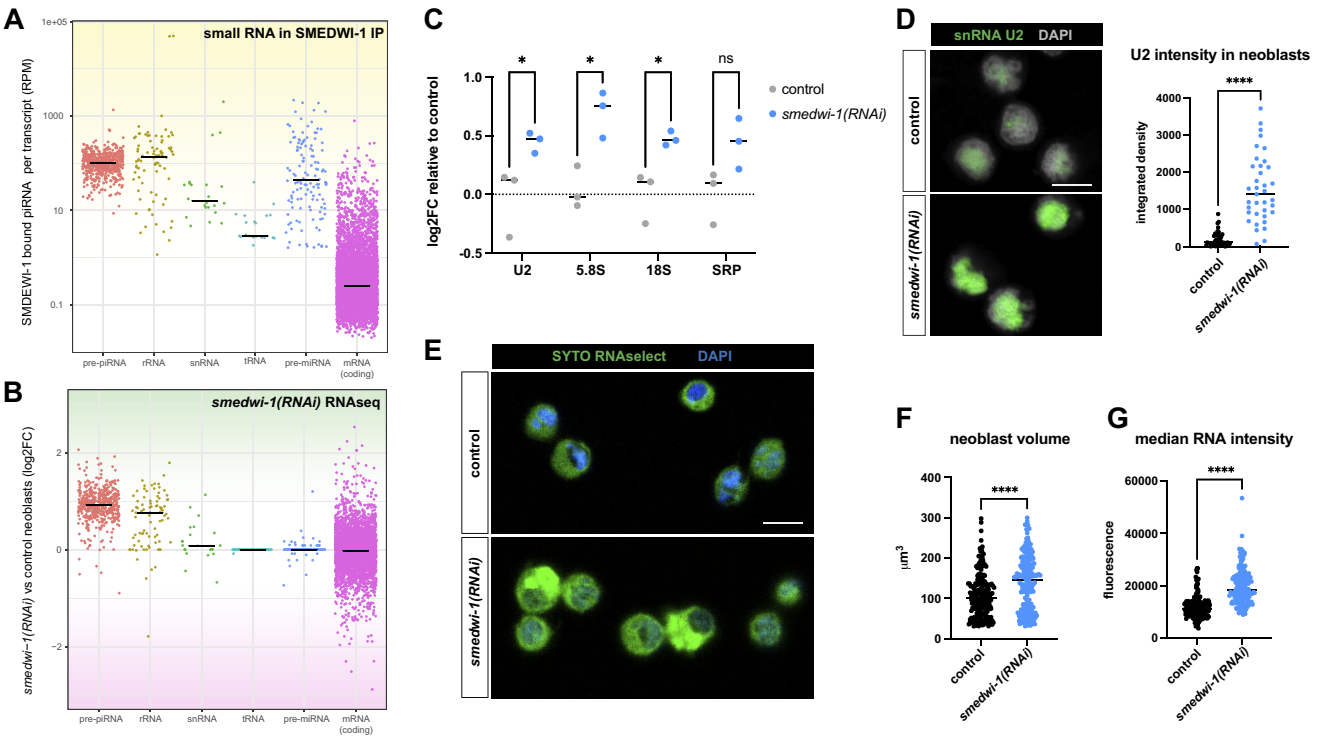


Figure 3. SMEDWI-1 regulates levels of structural RNAs. (**A**) Number of SMEDWI-1-bound small RNAs mapping to individual elements in each of the RNA classes. Pre-piRNAs are defined as genomic regions that harbor a significant enrichment of piRNAs (piRNA clusters). Several structural RNAs have large numbers of small RNAs bound to SMEDWI-1. (**B**) Changes in neoblast RNA levels of individual elements upon loss of SMEDWI-1 as determined by RNAseq analysis. Several structural RNAs show increases in RNA abundance in the RNA libraries upon loss of SMEDWI-1. (**C**) qPCR on neoblast cDNA generated with random hexamers confirms increased levels of several structural RNAs upon loss of SMEDWI-1. Statistical significance was determined using a *t*-test: * $P < 0.05$; ns: not significant. (**D**) Stainings of isolated neoblasts from control animals and *smedwi-1(RNAi)* animals labeled with RNA probe antisense to snRNA U2 (left), and quantification of the signal intensity (right). The U2 signal is significantly increased in *smedwi-1(RNAi)* neoblasts. Scale bar 10μm. Statistical significance was determined using a *t*-test: **** $P < 0.0001$. (**E**) Stainings of isolated neoblasts from control animals and *smedwi-1(RNAi)* animals labeled with RNA-specific dye SYTO RNA-select. Scale bar 10 μm. (F, G) Quantification of E, showing a significant increase in volume (**F**) and RNA intensity (**G**) in *smedwi-1(RNAi)* neoblasts. Statistical significance was determined using a *t*-test: **** $P < 0.0001$.

(Figure 3D). Both methods confirmed an increase in the overall levels of these transcripts in the *smedwi-1(RNAi)* samples, indicating that in wildtype stem cells, SMEDWI-1 functions to reduce the levels of these structured and non-polyadenylated RNAs. It is important to note that the piRNAs matching these transcripts are largely of the sense orientation (Supplementary Figure S3A,B) and thus do not target the transcripts by base-pairing, but rather reflect fragments of these structural non-coding RNAs.

Structural RNA constitutes a major fraction of the total cellular RNA content. The finding of increased levels of several of these RNAs suggests that *smedwi-1(RNAi)* neoblasts may contain more cellular RNA than control neoblasts. To quantify cellular RNA content, we isolated neoblasts from control animals and *smedwi-1(RNAi)* animals, and measured fluorescence using the RNA-specific dye SYTO RNAselect (Figure 3E). We found that *smedwi-1(RNAi)* neoblasts are significantly larger and significantly brighter, indicating that they have a higher total RNA content than control neoblasts (Figure 3F,G).

We next asked whether the increased levels of structural RNAs detected in the *smedwi-1(RNAi)* neoblasts by qPCR and FISH reflected intact mature sequences, or rather might reflect processing intermediates or degradation products of these molecules. We performed northern blot analysis for RNA from control neoblasts and *smedwi-1(RNAi)* neoblasts and probed with a 5′ fragment of the 28S rRNA to determine the sizes of RNA species related to this ribosomal RNA (Supplementary Figure S4A). We did not find a notable difference in the size of the labeled bands, suggesting that rRNA is largely intact, and that a significant fraction of the accumulating structural RNA is full-length.

SMEDWI-1 controls non-polyadenylated mRNAs

mRNA-related reads made up only a small fraction of the piRNAs, and most mRNAs with matching piRNAs were not affected by the loss of *smedwi-1* (Figure 4A). A small number of these transcripts however had both piRNA coverage and clear transcriptional changes in the absence of SMEDWI-1 suggesting that they could be directly regulated. Remarkably, most of these transcripts corresponded to histones. In a previous study, the cellular localization of histone transcripts was found to be deregulated in the absence of SMEDWI-1 protein (18). We noticed however that only a subset of histone transcripts was upregulated in the absence of SMEDWI-1, while others remained unaffected (Figure 4A).

Histone transcripts come in two major types: transcripts for histone variants are typically polyadenylated just like other mRNAs, whereas transcripts for replication-dependent canonical histones terminate in a 3′ stem-loop structure and thus are distinct from other mRNAs. *S. mediterranea*

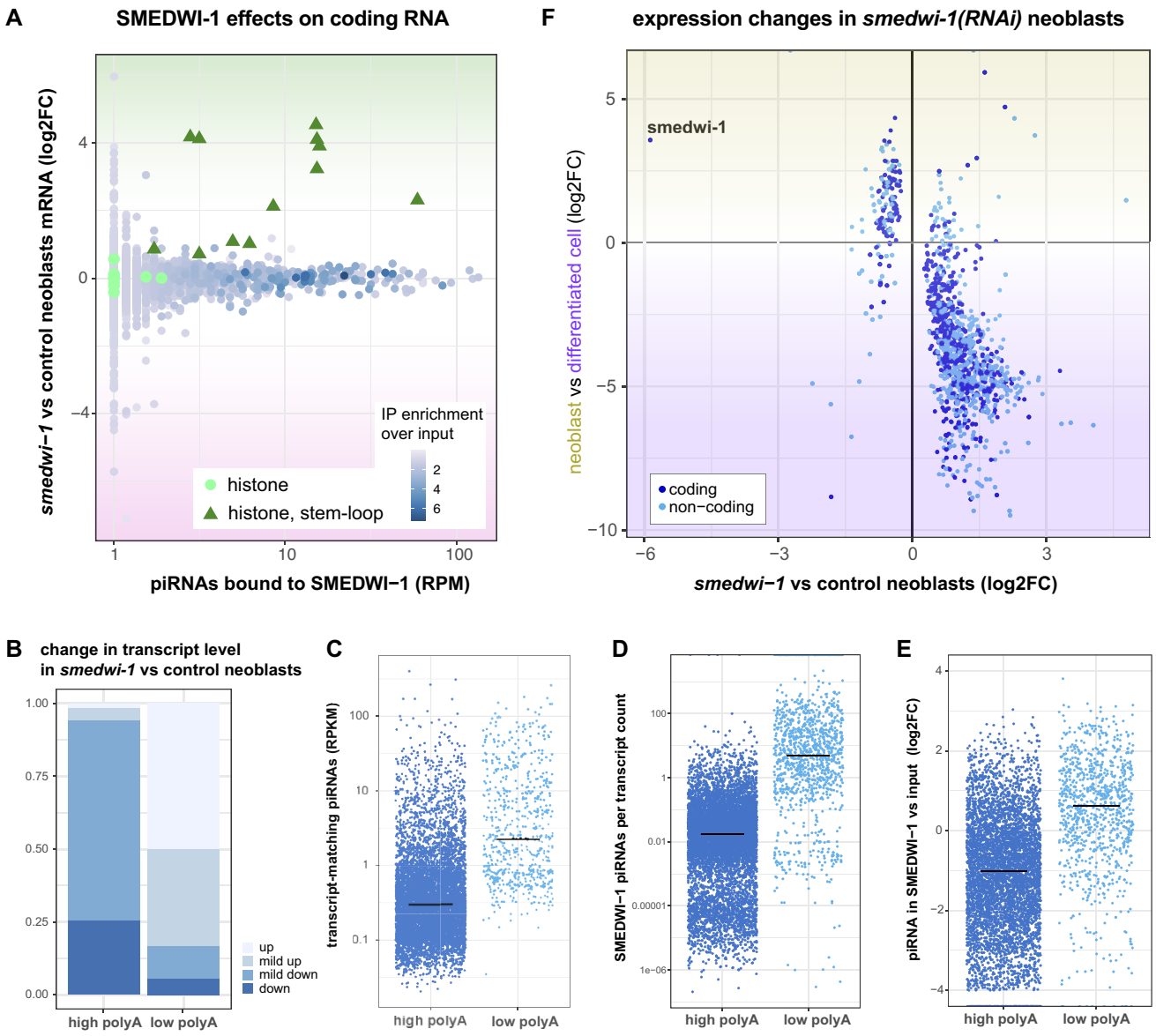


Figure 4. SMEDWI-1 restricts levels of non-polyadenylated mRNAs. (**A**) Scatterplot of mRNA sequences showing the SMEDWI-1-bound piRNA density at the mRNA versus the fold change of the mRNA in neoblasts upon loss of SMEDWI-1. Coloring indicates the fold enrichment of the piRNAs in SMEDWI-1 over input. The majority of mRNAs that have matching piRNAs bound to SMEDWI-1 are not significantly affected in level upon loss of SMEDWI-1. Conversely most mRNAs that are altered in level upon loss of SMEDWI-1 do not have SMEDWI-1-bound piRNAs. The major exception is a subset of the histone mRNAs (green). This subset consists of the canonical histones that are characterized by a terminal stem-loop (triangles). (**B**) Bar graph showing the fraction of transcripts that is up or down in neoblasts upon loss of SMEDWI-1, separated by polyadenylation level. Transcripts with high polyadenylation are more likely to be reduced in levels, whereas transcripts with low polyadenylation tend to be increased. Mild change corresponds to $\log_2$ fold change of 0.7; full change corresponds to log2 fold change of 1.5. (**C**) Graph showing the number of piRNAs bound to SMEDWI-1 for each mRNA, separated by polyadenylation level. Transcripts with low polyadenylation tend to have higher numbers of piRNAs. (**D**) Graph showing the number of piRNAs bound to SMEDWI-1 for each mRNA normalized to the expression level of the mRNA, separated by polyadenylation level. Transcripts with low polyadenylation tend to have higher numbers of piRNAs per transcript. (**E**) Graph showing the enrichment of piRNAs in SMEDWI-1 compared the level of the piRNAs in the input for each mRNA, separated by polyadenylation level. Small RNAs related to transcripts with low polyadenylation tend to be enriched in SMEDWI-1, whereas small RNAs related to transcripts with high polyadenylation are depleted. (**F**) Scatterplot of mRNA sequences showing the fold change of the mRNA in neoblast samples upon loss of SMEDWI-1 versus the neoblast enrichment of the transcript as determined by RNAseq analysis. Transcripts that are overrepresented in the *smedwi-1(RNAi)* neoblasts (right half) compared to controls are transcripts that are typically depleted from neoblasts (purple shade), whereas the few downregulated transcripts (left half) largely correspond to neoblast genes (yellow shade).

encodes several copies of each of the five histones and only a subset of them encodes a 3′ stem-loop structure. We evaluated the presence of a stem-loop on each of the histone genes and generated matching RiboMinus RNAseq libraries to determine polyA-independent levels of the various histone transcripts. Indeed, the histone transcripts that were overrepresented in *smedwi-1(RNAi)* neoblasts largely overlapped with the ones that have a terminal stem-loop (Figure 4A) and were underrepresented in the oligo-dT-selected libraries (Supplementary Figure S4B). The histone transcripts that are regulated by SMEDWI-1 thus are the non-polyadenylated transcripts that are distinct from regular mRNAs.

The trend observed among the structural ncRNAs and the histone transcripts suggested that SMEDWI-1 may have a more general role in the regulation of non-adenylated transcripts. To explore this further, we defined sets of high-polyA transcripts and low-polyA transcripts in the planarian transcriptome, based on their relative abundance in the oligo dT-based RNAseq versus the RiboMinus-based RNAseq. Next, we analysed the foldchange of the transcripts in these classes upon loss of SMEDWI-1. Of the high-polyA transcripts that had significant changes in expression, over 90% were reduced in abundance in *smedwi-1(RNAi)* libraries, whereas over 80% of altered low-polyA transcripts were increased (Figure 4B). Further we found that the low-polyA transcripts had significantly higher density of SMEDWI-1 bound piRNAs than the high-polyA transcripts (Figure 4C), and this effect was even more pronounced when considering that the full-length low-polyA transcripts tend to be less abundant than the high-polyA transcripts (Figure 4D). Furthermore, we found that piRNAs from low-polyA transcripts were specifically enriched in the SMEDWI-1 IP compared to their presence in total small RNA pool, whereas piRNAs from high-polyA mRNAs were depleted (Figure 4E, Supplementary Figure S3E,F). Together, these data indicate that SMEDWI-1 preferentially binds piRNAs matching low-polyA transcripts, and that the outcome of SMEDWI-1 activity is a reduction in the levels of atypical transcripts that lack prominent polyA tails.

Interestingly, the transcripts that were most affected by this selection mechanism differed between the stem cells and the early differentiating cells. In the stem cells, typical stem cell genes were not affected (Figure 1C) but loss of SMEDWI-1 did reduce levels of some other stem cell-specific transcripts, and increased levels of low abundant and largely non-coding transcripts that are normally enriched in differentiated cells (Figure 4F, Supplementary Figure S4D,E). In the differentiating cells, the inverse pattern was detected, involving an increase of some stem cell-specific transcripts and a decrease of differentiation-related transcripts (Supplementary Figure S4C,D). This indicates that the misregulation of the low-adenylated and structural RNAs leads to a broader destabilization of the stem cell transcriptome.

In accordance with this notion, we found that the overall size profile of the neoblast RNA changes in the absence of SMEDWI-1. While the RNA of control neoblasts contained several well-defined prominent bands, reflecting abundant RNA species, the RNA of *smedwi-1(RNAi)* neoblasts instead showed a smear indicating the disruption of the RNA species in these cells (Supplementary Figure S4F).

SMEDWI-1 piRNAs rely in part on the endonuclease Zucchini

An intriguing question is how these SMEDWI-1-bound piRNAs matching low-polyA mRNAs and structural RNAs are formed. Studies in *Drosophila* and mouse have proposed that the majority of piRNAs are generated from piRNA clusters, which are genomic loci that are rich in transposon fragments and poor in genic content (27,47–49). Primary transcripts from these loci are processed into smaller fragments by the endonuclease Zucchini and bound by a PIWI protein, after which they are trimmed to mature piRNA length and methylated on their 3′ end (50–58). Most PIWI proteins preferentially bind RNA fragments with a Uridine at the 5′ position and these primary piRNAs therefore have a strong 5′U bias (59). Targeting of complementary transcripts by base-pairing to these PIWI-piRNA complexes can lead to a piRNA amplification cycle known as the ping-pong cycle, which sequentially produces an antisense piRNA that has a bias for an Adenine at position 10 and reproduces the original piRNA (47).

The most parsimonious explanation for the origin of the low-polyA mRNA piRNAs is that they are generated directly from the mRNA transcripts. This is supported by several findings: 1. The SMEDWI-1-bound piRNAs matching these transcripts are overwhelmingly of the sense orientation (Supplementary Figure S3), indicating that they must derive from transcripts that contain (fragments of) the exact same sequence as the mRNA. 2. The piRNAs have a strong 5′U bias in the absence of enrichment of Adenine on the 10th position (Figure 5A, Supplementary Figure S3) and no antisense piRNAs with a ping-pong signature are identified for these transcripts. Therefore, most of them are expected to be ‘primary piRNAs’ that are generated independently of the PIWI-mediated ping-pong amplification loop. 3. The piRNAs do not map to other regions in the genome (Figure 5B, Supplementary Figure S5A), indicating that they don’t derive from piRNA clusters and thus must be generated from the matching transcripts themselves.

In mouse and *Drosophila*, the nuclease Zucchini is involved in the generation of new 5′ ends from longer transcripts that are then processed into mature piRNAs (53–58). Zucchini is crucial for piRNA synthesis in the *Drosophila* soma, and in its absence primary piRNA transcripts accumulated. Zucchini further was found to be required for the stability of the PIWI protein to which its piRNAs are primarily bound (60,61). *Schmidtea* encodes only a single *zucchini* homolog, and we hypothesized that it may be involved in the production of the sense piRNAs bound to SMEDWI-1.

We used RNAi-mediated knockdown of Zucchini to investigate the effect on the stem cell piRNAs and SMEDWI-1 protein. Loss of Zucchini resulted in a strong reduction of the overall piRNA levels (Figure 5C), indicating that the role of Zucchini in piRNA production is conserved in *Schmidtea*. Zucchini knockdown also reduced the level of SMEDWI-1 protein without a significant effect on the level of *smedwi-1* transcript levels (Figure 5D-G, Supplementary Figure S5B,C), indicating that the reduction in piRNAs likely affects the loading and stability of the SMEDWI-1 protein, similar to what was found in other model systems. We further found that the RNA-dense granules marked by the Y12 antibody staining, which are thought to reflect the sites of piRNA processing (nuage), are

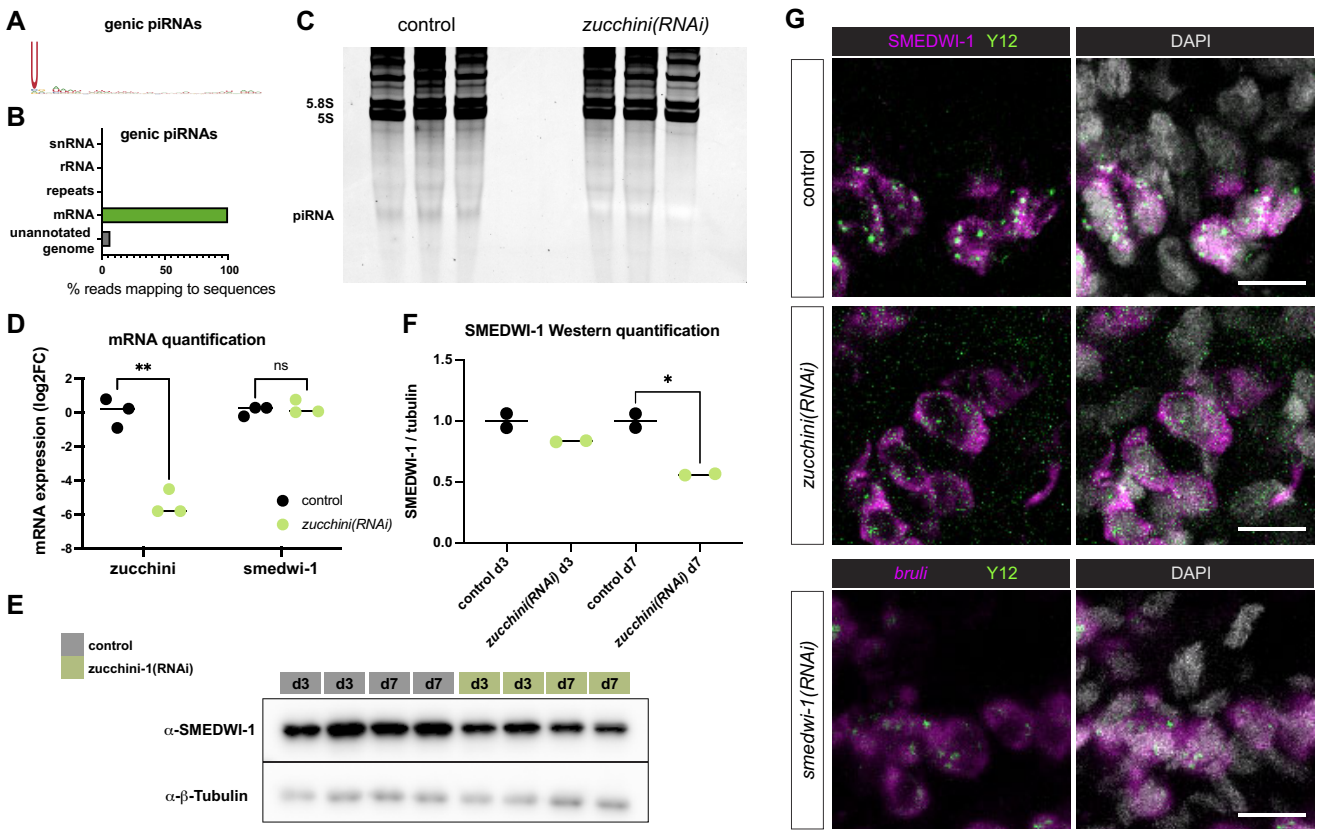


Figure 5. SMEDWI-1 function intersects with nuclease Zucchini. (**A**) Sequence logo of the SMEDWI-1 bound piRNAs related to mRNAs shows a bias for a 5′ Uridine. See also Supplementary Figure S3. (**B**) The majority of piRNAs related to mRNAs does not map to other loci in the genome than the annotated mRNA region. See also Supplementary Figure S5A. (**C**) Total RNA from 3 independent sets of control animals and *zucchini(RNAi)* animals shows a reduction in piRNAs upon loss of Zucchini. (**D**) qPCR showing that in *zucchini(RNAi)* animals, *smedwi-1* transcript levels remain stable. Statistical significance was determined using a *t*-test: ** $P < 0.01$; ns: not significant. (**E**) Western blot showing the reduction of SMEDWI-1 protein in *zucchini(RNAi)* samples. (**F**) Quantification of E. Statistical significance was determined using a *t*-test: * $P < 0.05$. (**G**) Immunostainings of whole mounts from control animals and *zucchini(RNAi)* animals, and *smedwi-1(RNAi)* animals with the nuage-labeling antibody Y12 (green) and SMEDWI-1 protein (magenta) or *bruli* RNA (magenta). In the absence of Zucchini or SMEDWI-1, nuage structures are reduced. Scale bar 10 μm.

reduced in the absence of Zucchini as well as in the absence of SMEDWI-1 (Figure 5G, Supplementary Figure S5C). However, in the zucchini(RNAi) samples we did not observe accumulation of the transcripts that matched SMEDWI-1 bound piRNAs, suggesting that these transcripts may be subject to another degradation mechanism in the absence of Zucchini.

Based on these findings we conclude that Zucchini plays an important role in the generation of piRNAs in planarian stem cells, and is likely involved in the production of SMEDWI-1 bound piRNAs. However, the activity of Zucchini by itself does not provide an explanation for how the previously identified low-polyA transcripts are selected for the generation of piRNAs while high-polyA transcripts are left largely unperturbed.

SMEDWI-1 associates with monosomes and factors of the pioneer round of translation

To identify other factors involved in the mechanisms surrounding SMEDWI-1, we precipitated SMEDWI-1 protein (Supplementary Figure S6A,B) and identified its interactors by mass spectrometry. After removal of common cytoskeletal contaminants, we found that we retrieved almost all subunits of the planarian ribosomal complexes (Figure 6A, Supplementary Tables S2 and S3). To verify the interaction of SMEDWI-1 with the ribosomes, we used the antibody Y10B, which recognizes rRNA, and probed for SMEDWI-1 by Western blot (Figure 6B). In addition, we used antibodies that recognize the small ribosomal subunit protein RPS17 or the large ribosomal subunit protein RPL10 to precipitate ribosomes, and again probed for SMEDWI-1 (Figure 6C). In each of the three IPs we found a strong enrichment of SMEDWI-1 in the IP over the input, whereas SMEDWI-1 was not enriched in IPs with unrelated antibody (Supplementary Figure S6C), confirming the specific interaction between SMEDWI-1 and ribosomes. Conversely, we also found strong enrichment of the ribosomal protein RPS17 in the SMEDWI-1 IP, even after treatment of the lysates with RNase to remove indirect interactions through bridging RNAs (Figure 6D, Supplementary Figure S6D).

Ribosomal proteins can be part of pre-initiation ribosomal subunits, monosomes, or polysomes. To determine which of these fractions of ribosomal proteins interact with SMEDWI-1, we used a sucrose gradient to separate ribosomal complexes by size (Figure 6E). SMEDWI-1 was not found in the polysome fractions, but did show in the fractions corresponding to the monosomes, indicating that SMEDWI-1 is not associated with ribosomes that are involved in efficient translation. To confirm the interaction of SMEDWI-1 with

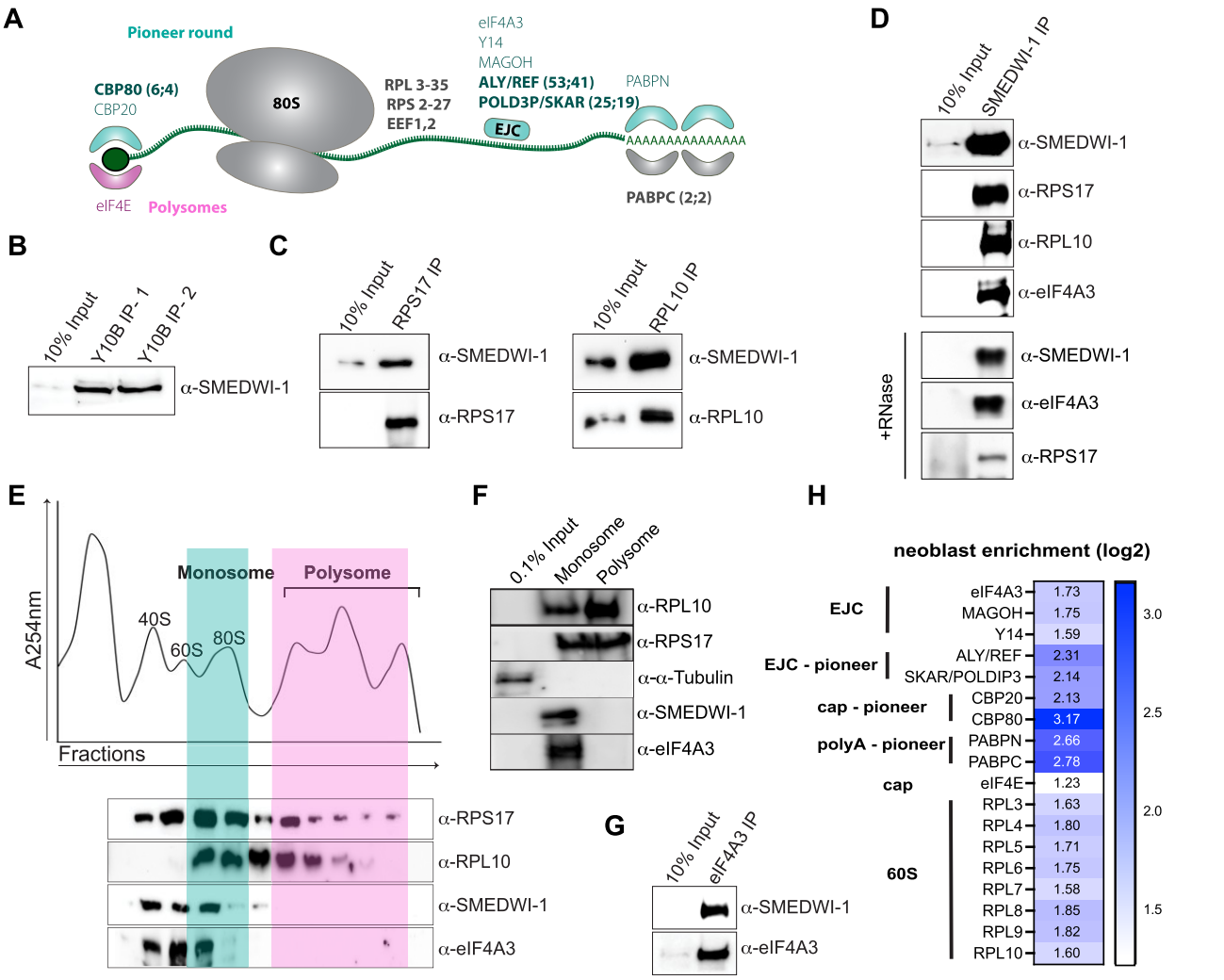


Figure 6. SMEDWI-1 interacts with ribosomes during the pioneer round of translation. (**A**) Schematic indicating protein factors involved in translation during the pioneer round (teal) and during efficient translation on ribosomes (magenta). Factors that are in common between both stages are indicated in grey. Factors that were recovered in the SMEDWI-1 mass spectrometry are indicated in bold; numbers indicate the number of peptides recovered. No peptides to these proteins were recovered in negative controls. For ribosomal proteins, peptide counts ranged from 55 to 2 (see Supplementary Table S2). (**B**) Western blot showing co-precipitation of SMEDWI-1 upon pull down with the rRNA antibody Y10B. Replicate experiments are shown next to the input. (**C**) Western blots showing co-precipitation of SMEDWI-1 with ribosomal proteins RPS17 and RPL10. Precipitated complex is shown next to the input. (**D**) Western blots showing co-precipitation of ribosomal proteins and EJC protein eIF4A3 with SMEDWI-1. Precipitated complex is shown next to the input. eIF4A3 and RPS17 still co-precipitate with SMEDWI-1 after treatment with RNAseA. (**E**) Polysome profile with Western blots of the fractions shows the presence of RPS17 and RPL10 in both monosome and polysome fractions. SMEDWI-1 and eIF4A3 are present only in the monosome fractions. (**F**) Western blot showing co-precipitation of RPS17 with RPL10 in both monosomal and polysomal fractions, and co-precipitation of SMEDWI-1 and eIF4A3 with RPL10 specifically in the monosomal fractions. (**G**) Western blot showing co-precipitation of SMEDWI-1 with Exon Junction Complex protein eIF4A3. Precipitated complex is shown next to the input. (**H**) Transcript enrichment (log_2) of translation-related genes in neoblasts over differentiated cells. All translation-related genes show higher expression in neoblasts than in differentiated cells, but pioneer-related genes have even higher transcript enrichment levels.

the ribosomes in monosomal state, we isolated monosomes and polysomes from the sucrose gradient and performed IP for the large ribosomal subunit protein RPL10. This IP recovered the small ribosomal subunit protein RPS17 in both the monosomal and the polysomal fraction, and conversely did not co-purify the unrelated cytoskeletal protein α-tubulin (Figure 6F, Supplementary Figure S6E), indicating that the IP from ribosomal fractions can recover true interactions. SMEDWI-1 was not detected in the IP from the polysomal fractions, but clearly co-precipitated with RPL10 in the monosomal fraction (Figure 6F). This confirms that SMEDWI-1 associates with ribosomes, specifically in the monosomal state.

We examined the mass spectrometry data for other translation-related proteins and found a highly suggestive subset (Figure 6A). The mass spectrometry recovered peptides matching the homolog of CBP80, which is part of the cap-binding complex of newly synthesized RNAs that is replaced by eIF4E after the first round of translation, also known as the pioneer round (62). We further found strong enrichment of the homologs of proteins ALY/REF and SKAR/POLDIP3, both of which associate with the Exon-Junction Complex (EJC). The EJC is loaded on newly synthesized RNAs co-transcriptionally and is also typically removed during the pioneer round of translation (63). The mRNA export factor ALY/REF and translation activating

factor SKAR/POLDIP3 associate with the EJC and are similarly removed during pioneer translation. One of the core proteins of the Exon Junction Complex is the RNA binding protein eIF4A3, and this protein additionally assists in the unwinding of RNA during pioneer translation (64). We confirmed that eIF4A3 is found in the monosome fraction, and is not associated with the polysomes, similar to SMEDWI-1 (Figure 6E). Further we found that pull-down of SMEDWI-1 co-precipitates eIF4A3 and vice versa, confirming the association of SMEDWI-1 with the pioneer round of translation (Figure 6D,G).

The EJC was previously found to be required for stem cell maintenance in planarians without major effects on differentiated cells (65). In agreement with this, we found that genes associated with the pioneer round of translation generally showed increased expression in the planarian neoblasts compared to the differentiated tissues. Whereas mRNAs corresponding to ribosomal proteins and the EJC itself were all moderately enriched in neoblasts compared to differentiated cells, the factors associated with pioneer translation were even higher expressed (Figure 6H). Together, these findings indicate that pioneer translation is actively used in the planarian stem cells, and that SMEDWI-1 associates with RNAs going through this process.

Transcripts that generate SMEDWI-1-associated piRNAs are enriched in monosomes

During the pioneer round of translation, newly synthesized transcripts are tested for the presence of features that are required for efficient translation, and this therefore could be a time at which coding high-polyA transcripts are distinguished from low-polyA transcripts. To examine whether the monosomal pioneer complex could play a role in passing a subset of transcripts to SMEDWI-1, we inspected the distribution of transcripts between the monosomal fraction and the polysomal fraction.

We first confirmed that transcripts of several well-polyadenylated control genes that are not associated with production of piRNAs were enriched in the polysomal fraction (Figure 7A). We then tested the two types of histone transcripts. We found that transcripts of canonical histones (which terminate in a stem-loop structure instead of a polyA tail) were enriched in the monosomes, whereas the transcripts of histone variants were more equally distributed between monosomes and polysomes (Figure 7A). This matches our previous finding that stem-loop histone transcripts are regulated by SMEDWI-1, whereas the polyA-tailed histone transcripts are not. Finally, we tested the distribution of several poorly adenylated transcripts that give rise to piRNAs. We found that these piRNA-generating transcripts were enriched in the monosomal fraction, and the same was true for the transcript of an active Ty3-like transposon that generates a large number of piRNAs (Figure 7B).

Disruption of mature RNA production leads to increased piRNAs

Our data suggest that transcripts that are atypical or have features of non-translated RNAs can be diverted to piRNA production and binding to SMEDWI-1. We therefore wondered whether disrupted RNA synthesis would lead to increased piRNA production from transcripts that are otherwise not processed in this manner. To test this, we treated animals with the Adenosine analog cordycepin to disrupt RNA elongation, and analyzed changes in piRNAs bound to SMEDWI-1 compared to an untreated sample. At low concentrations, cordycepin inhibits polyadenylation, but at higher concentrations it also disrupts general RNA synthesis which primarily affects rRNA transcripts and processing (66). We indeed found a highly significant increase in rRNA-related piRNAs bound to SMEDWI-1 upon treatment with cordycepin (Figure 7C). Furthermore, we found increased production of piRNAs related to polyadenylated transcripts, and among the piRNAs derived from mRNA-related sequences we detected a stronger increase of piRNAs from high-polyA transcripts compared to the change in piRNAs from low-polyA transcripts (Figure 7D), confirming our model.

Finally, we asked whether an exogenous non-coding transcript would be recognized by this mechanism and processed into piRNAs. To test this, we injected animals with a non-coding RNA construct and analyzed the RNA species derived from this construct by RNase protection assay. In the injected animals we indeed detected ∼30nt long fragments derived from this sequence that were absent from mock treated animals (Figure 7E).

Taken together our data indicate that atypical transcripts that are not well polyadenylated are enriched in the monosomes, which retains them in proximity to SMEDWI-1. We propose that these transcripts are degraded by multiple mechanisms, including by the piRNA-related endonuclease Zucchini. We propose that SMEDWI-1 is present at the monosomes to bind the fragments of the degraded RNAs and stabilizes 5′U fragments to form piRNAs derived from these sequences. This association of SMEDWI-1 with the pioneer round of translation thus allows this PIWI protein to pick up a wide variety of sequence fragments for piRNA production and for transcriptional monitoring.

Discussion

PIWI proteins are a consistent presence in the stem cells of regenerative organisms and are essential for these organism's regenerative abilities, suggesting that they have a conserved role in the functioning of these stem cells. While suppression of repetitive elements is a prime function of PIWI proteins throughout germline and stem cell systems, there have been several indications that their role in stem cell maintenance can be separated from that in transposon regulation (44,46,67). Here we show that in the stem cells of *S. mediterranea*, the quintessential stem cell marker SMEDWI-1 functions as a backup mechanism for transposon control—in line with the functions identified for other PIWI proteins—but has an additional role in monitoring and filtering the transcriptional products of the cell to restrict the levels of non-coding or poorly adenylated RNAs (Figure 7F). We propose that this streamlines the gene expression in these cells and that it endows the SMEDWI-1 expressing stem cells with improved stability under stress.

It has long been recognized that sequences matching structural RNAs are found in small RNA libraries of a wide range of organisms (68–75), however it often remained unclear whether these are true small RNAs that are processed by small RNA generating pathways, or whether they are degradation products that happen to be of similar size as regulatory small RNAs. Similarly, small RNAs derived from

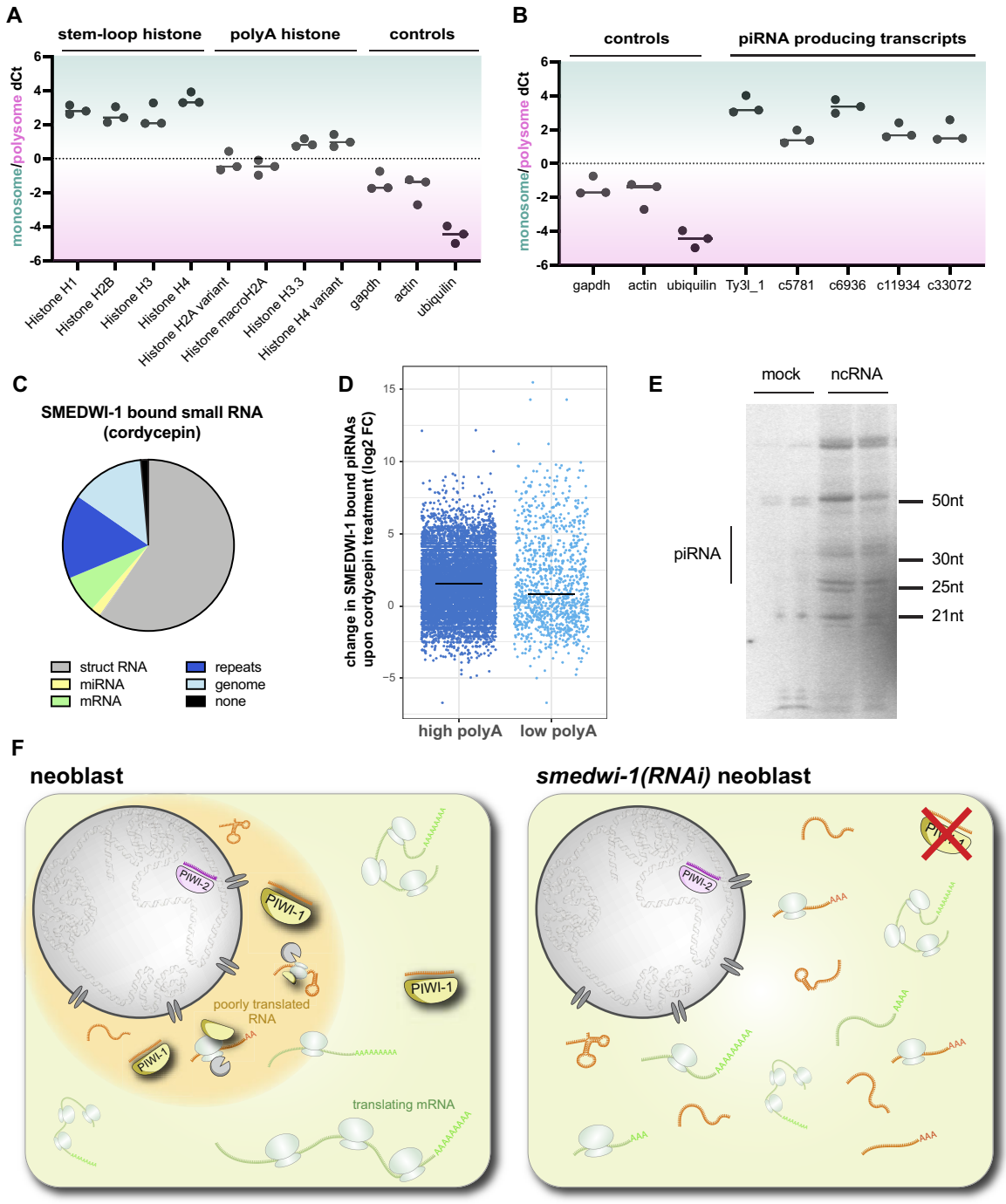


Figure 7. Poorly translated transcripts give rise to SMEDWI-1 bound piRNAs. (A, B) Relative amount of histone transcript associated with monosomes compared to the amount associated with polysomes as determined by qPCR. Shown are $\log_2$ fold changes (delta Ct values) for three independent experiments. Highly translated control transcripts are enriched in the polysomal fraction. (**A**) Stem-loop histone transcripts are enriched in monosomes, whereas the related polyA histone transcripts are more equally distributed between the two ribosomal states. (**B**) The transcript for a TY3-related transposon that generates many piRNAs is enriched in the monosomes. Transcripts that give rise to SMEDWI-1 bound piRNAs are also enriched in the monosomal fraction. Controls are repeated from panel A for clarity. (**C**) Relative proportions of small RNA reads mapping to various genetic elements in the SMEDWI-1 bound piRNAs after treatment with cordycepin (compare to Figure 2A). Small RNAs related to structural RNA are strongly increased. (**D**) Comparison of the foldchange in SMEDWI-1 bound piRNAs derived from low polyA transcripts and high polyA transcrtipts upon treatment with cordycepin. High-polyA transcripts that are normally rarely processed into piRNAs are more prominently processed upon disruption of RNA synthesis and maturation by cordycepin. (**E**) RNase protection assay showing that introduction an exogenous poorly translated transcript leads to the production of small RNAs of the typical size of piRNAs derived from this sequence. (**F**) Model for the function of SMEDWI-1. In wildtype neoblasts (left), SMEDWI-1 associates with monosomes during the pioneer round of translation, where it filters transcripts as they are exported from the nucleus, probably while they are still in the perinuclear region. Transcripts with high levels of secondary structure or low terminal adenylation (orange) are degraded and fragments are bound to SMEDWI-1. Only well-translating transcripts (green) are retained and can form polysomes. This mechanism keeps the rest of the cytoplasm free from aberrant transcripts. SMEDWI-1 protein loaded with piRNAs derived from the degraded transcripts continues to scan the well-translating transcripts to induce piRNA-mediated silencing if significant base-pairing is achieved. In the absence of SMEDWI-1 (right), atypical transcripts persist and the transcriptome of the neoblast is disorganized, leading to inefficiencies in gene expression and in neoblast function.

histone mRNA have been detected in various organisms including mosquito (76), *C. elegans* (77), and planarians (18), but their specificity and function have remained unclear. Here we find by immunoprecipitation that small RNAs overlapping snRNA, rRNA and histone mRNA sequences are bound to SMEDWI-1, have a 5′ U bias, preferential length of just over 30nt, and are methylated – all of which are hallmarks of planarian piRNAs.

These histone-related and structural RNA-related piRNAs are over 95% of the sense orientation and do not match other loci in the genome (Supplementary Figure S3 and S5A). This poses conundrums for their function and their biogenesis. Typically, the function of piRNAs is to recruit the PIWI protein to specific sequences by base-pairing, thereby resulting in degradation of the complementary target RNA or bringing in additional proteins to modify chromatin organization. Being of the sense orientation, this group of piRNAs can not target snRNAs, rRNAs, or histone mRNAs by base pairing. They do not show significant base pairing with other coding transcripts or genomic regions, making it unlikely that they directly target mRNAs or transposons. They could be used as a scanning mechanism to detect and degrade antisense transcripts related to structured RNAs. Both snRNAs and rRNAs are encoded in genomic clusters that may at low frequency produce antisense transcripts, however no accumulation of antisense transcripts from snRNA or rRNA loci was detected in the absence of SMEDWI-1. snRNAs and rRNAs also have complex structures with hairpins, and some sense piRNAs therefore could target the other strand of a hairpin by imperfect base-pairing. However, the fact that these small RNAs are not enriched in regions with structural elements also renders this explanation unlikely (Supplementary Figure S3A,B,D).

Alternatively, these piRNAs could have a regulatory role related to other small RNAs rather than to other long transcripts. Recently in *C. elegans*, PIWI proteins were found to restrict the amplification of antisense small RNAs matching to rRNA, thereby protecting the rRNA from fragmentation (78). We did not find increased levels of antisense reads matching rRNAs in the absence of SMEDWI-1, indicating that this mechanism is not likely to be at play in *Schmidtea*. Additionally, the RdRP-mediated amplification mechanism of small RNAs as found in *C. elegans* is not present in most other animal models, and the need to restrict potentially deranging amplification may thus be a nematode-specific problem.

In contrast, we propose that these sense small RNAs are the result of a rather a-specific degradation mechanism and that the majority does not have a sequence-based function per se. We propose that these piRNAs are the consequence of the way planarian stem cells accomplish (part of) their piRNA synthesis. Studies on the biogenesis of piRNAs have focused on the tightly regulated generation of specific piRNA sequences from specific piRNA precursor transcripts, as exemplified by the 21U RNAs in C. *elegans* and the piRNA clusters in *Drosophila*. Although genomic piRNA clusters can be identified in planarians, we find that a significant fraction of the planarian piRNAs does not derive from these cluster sequences, and we propose that their biogenesis involves the recognition of aberrant transcripts that could come from a wide variety of sources. They could derive from transposon-related sequences, but they could also come from pseudogenes, or from mislocalized structural RNAs. Indeed, we identified 2′O-methylated SMEDWI-1-bound piRNAs from each of these types of transcripts without presence of these sequences in genomic piRNA clusters (Supplementary Figure S5A).

A major question then is how the piRNA machinery recognizes these aberrant transcripts and distinguishes them from normal coding mRNAs. Our finding of SMEDWI-1 in association with monosomes and several proteins that are specific to the pioneer round of translation provides an appealing explanation. The pioneer round of translation provides a natural opportunity to separate the proper coding mRNAs that should be passed on to polysomes, from transcripts that look abnormal and should be removed from circulation (Figure 7F). The pioneer round of translation has been reported to identify a wide range of dysfunctional transcripts, such as transcripts that lack coding potential, contain a premature stop, lack essential binding proteins, or have excessive secondary structure (79). The proximity of SMEDWI-1 may ensure that when these transcripts are degraded, fragments can be bound by SMEDWI-1 and directly enter the piRNA pathway.

In support of this model, we find higher levels of piRNAs generated from mRNAs with short polyA tails than from mRNAs with long polyA tails (Figure 4C-E), and polyadenylated non-coding transcripts have higher levels of piRNAs per unit of full-length transcript than coding mRNAs (Supplementary Figure S4E). In the case of canonical histones, we propose that the effect may relate to the cell cycle-based regulation of histone translation. Replication-dependent histone transcripts are translated primarily during S-phase, and this requires the Stem Loop Binding Protein (SLBP). Outside of S-phase the absence of SLBP may make them targets for piRNA production. Some support for this model comes from Mosquito (76), where it was reported that the peak of histone piRNAs is a few hours after the peak of canonical histone expression and when most cells have arrived in G2/M phase, suggesting that this happens when translation of these canonical histone transcripts is already reduced. Furthermore, we find that disruption of RNA synthesis leads to increased production of piRNAs from transcripts related to rRNAs and coding polyadenylated mRNAs (Figure 7C, D). The disruption of RNA elongation by cordycepin leads to truncated versions of transcripts as well as reduced polyadenylation. According to our model, these aberrant transcripts will be unable to proceed to polysomes for efficient translation, and instead will be processed into SMEDWI-1 bound piRNAs, and this is exactly what we find.

The outcome of this proposed collaboration between pioneer translation and SMEDWI-1 is 2-fold. In the first place it produces piRNA from odd-looking potentially dangerous transcripts and allows SMEDWI-1 to scan the cytoplasm for antisense transcripts related to these RNAs and target them for degradation. This operational manner of defining what transcripts need to be monitored allows for maximum flexibility and for the immediate identification of new threats for which no established piRNA clusters exist. In the second place, this mechanism functions as a clean-up strategy for aberrant RNAs. Such aberrant transcripts may be mis-expressed, have splicing defects, or have other dysfunctional properties, and it thus benefits the cell to remove them from the translational pool. How mechanistically SMEDWI-1 assists in this second part of the outcome is not clear from our data. SMEDWI-1 may play a role in stabilization of the complexes involved in pioneer translation, or it may assist in recruiting a nuclease, possibly Zucchini, to

degrade rejected transcripts. Further studies are necessary to elucidate that part of the mechanism.

The processing of piRNAs is associated with perinuclear RNA-protein-rich domains known as nuage (47,80,81). Nuage may associate with nuclear pores and thus is positioned in an ideal location to filter transcripts that exit the nucleus. In our experiments, nuage (as labeled with the antibody Y12) was less pronounced in the absence of SMEDWI-1 or of the piRNA-producing endonuclease Zucchini, indicating that piRNA processing may indeed be affected by the absence of these proteins. However, how these components interact spatially remains unclear from our data. Nuage is at the nuclear membrane (81), Zucchini is associated with mitochondria (56), and SMEDWI-1 protein is present broadly in the neoblast cytoplasm. In *Drosophila*, the helicase Armitage was proposed to transport pre-piRNAs between nuage and the mitochondria (82). *S. mediterranea* encodes a homolog of Armitage, so it is possible that this function is conserved, but how the pioneer round of translation fits in this spatial arrangement currently remains unknown.

In the absence of SMEDWI-1, structural RNAs accumulated in the neoblasts, and in wildtype cells we found abundant SMEDWI-1-bound piRNAs from rRNA and snRNA. Structural RNAs are not translated and are not known to pass through pioneer translation, but part of the quality control of snRNA and rRNA does take place in the cytoplasm, and the degradation pathways may intersect with those of abnormal mRNAs (83–85). The cleared rRNAs do not appear to be of aberrant length, but they could be incorrectly folded, incorrectly modified, or not correctly bound by ribosomal proteins, targeting them for degradation. It is possible that stem cells have particularly stringent quality control of their ribosome biogenesis and that minor imperfections trigger rRNA degradation in these cells. Further studies will be required to characterize the quality control of rRNA in neoblasts to determine why large amounts of rRNA are routinely degraded.

In this study, we have focused on the SMEDWI-1 bound piRNAs that are not related to transposons, but it is important to emphasize that this concerns a minority of the reads in the SMEDWI-1 libraries. Most piRNAs bound to SMEDWI-1 and SMEDWI-2 are related to transposons, and none of our findings contest the view that silencing of transposable elements is the primary role of the piRNA pathway. In fact, transcripts derived from transposon loci will often look aberrant. Many genomic transposon copies are incomplete, leading to transcripts that are truncated and poorly coding. Such transcripts would be prime candidates for degradation and processing into SMEDWI-1 bound piRNAs by the mechanism proposed above. In support of this notion, we found that the transcript of a piRNA producing transposon was enriched in monosomes over polysomes, similar to the mRNA-like piRNA-producing transcripts (Figure 7B). In contrast to structural RNA and mRNA-like transcripts, antisense transcripts of transposons are relatively prominent. The transposon-derived piRNAs therefore are likely to find targets and amplify by ping-pong amplification or Zucchini-mediated processing of the recognized targets into new piRNAs.

While the changes in the coding neoblast transcriptome upon loss of SMEDWI-1 are subtle, we find that this correlates with a significant delay in neoblast expansion (Figure 1). This may in part be due to changes in cell cycle, as we found evidence of an extended S-phase and a higher overall RNA content of the neoblasts. The subtle deregulation of many transcripts may also lead to inefficiencies in the general functioning of the neoblasts and in their ability to respond to challenges. This may additionally make them more likely to exit the stem cell state, as indicated by our findings. Inversely, this phenomenon may benefit stem cells with high SMEDWI-1 levels, giving them a competitive advantage over low-expressing cells, and thereby continuously maintaining high SMEDWI-1 levels in the stem cell population. We hypothesize that this stabilizing effect of the PIWI protein on the transcriptome may be a conserved feature in potent stem cell systems, and that this may explain the conserved presence of PIWI proteins in these cells.

Data availability

Sequencing data generated over the course of this study have been deposited in the SRA under accession number PRJNA905109. Proteomics data have been deposited in the massIVE database, under accession code MSV000091709.

Supplementary data

Supplementary Data are available at NAR Online.

Acknowledgements

We are grateful to Joan Steitz for sharing the Y10B rRNA antibody, and to the Keck DNA Sequencing Facility at Yale for assistance with sequencing. We also thank the MS & Proteomics Resource at Yale University for providing the necessary mass spectrometers and accompanying biotechnology tools funded in part by the NIH Office of the Director (S10OD02365101, S10OD019967, S10OD018034 and S10OD030363). We thank Karla Neugebauer and members of the Van Wolfswinkel Lab for discussion and comments.

Author contributions: S.A.P., D.L., M.Z. and J.C.vW. designed the studies. S.A.P., D.L., M.Z. and J.C.vW. performed the experiments. D.L., A.P. and J.vW. performed the computational analyses. J.C.vW. supervised the study. J.Cv.W. wrote the manuscript with input from the other authors.

Funding

National Institutes of Health [R35GM128619, R01AG078926]; Vallee Foundation. Funding for open access charge: NIH [R35GM128619, R01AG078926].

Conflict of interest statement

None declared.

References

1. Cox,D.N., Chao,A., Baker,J., Chang,L., Qiao,D. and Lin,H. (1998) A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. *Genes Dev.*, **12**, 3715–3727.
2. Cox,D.N., Chao,A. and Lin,H. (2000) piwi encodes a nucleoplasmic factor whose activity modulates the number and division rate of germline stem cells. *Development*, **127**, 503–514.
3. Cheng,E.C., Kang,D., Wang,Z. and Lin,H. (2014) PIWI proteins are dispensable for mouse somatic development and

reprogramming of fibroblasts into pluripotent stem cells. *PLoS One*, **9**, e97821.
4. Sharma,A.K., Nelson,M.C., Brandt,J.E., Wessman,M., Mahmud,N., Weller,K.P. and Hoffman,R. (2001) Human CD34(+) stem cells express the hiwi gene, a human homologue of the Drosophila gene piwi. *Blood*, **97**, 426–434.
5. Srivastava,M., Mazza-Curll,K.L., van Wolfswinkel,J.C. and Reddien,P.W. (2014) Whole-body acoel regeneration is controlled by wnt and bmp-admp signaling. *Curr. Biol.*, **24**, 1107–1113.
6. Juliano,C.E., Reich,A., Liu,N., Gotzfried,J., Zhong,M., Uman,S., Reenan,R.A., Wessel,G.M., Steele,R.E. and Lin,H. (2014) PIWI proteins and PIWI-interacting RNAs function in hydra somatic stem cells. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, 337–342.
7. Funayama,N., Nakatsukasa,M., Mohri,K., Masuda,Y. and Agata,K. (2010) Piwi expression in archeocytes and choanocytes in demosponges: insights into the stem cell system in demosponges. *Evol. Dev.*, **12**, 275–287.
8. Reddien,P.W., Oviedo,N.J., Jennings,J.R., Jenkin,J.C. and Sanchez Alvarado,A. (2005) SMEDWI-2 is a PIWI-like protein that regulates planarian stem cells. *Science*, **310**, 1327–1330.
9. Reddien,P.W. and Sanchez Alvarado,A. (2004) Fundamentals of planarian regeneration. *Annu. Rev. Cell Dev. Biol.*, **20**, 725–757.
10. Newmark,P.A. and Sanchez Alvarado,A. (2000) Bromodeoxyuridine specifically labels the regenerative stem cells of planarians. *Dev. Biol.*, **220**, 142–153.
11. Fincher,C.T., Wurtzel,O., de Hoog,T., Kravarik,K.M. and Reddien,P.W. (2018) Cell type transcriptome atlas for the planarian Schmidtea mediterranea. *Science*, **360**, 112.
12. Raz,A.A., Wurtzel,O. and Reddien,P.W. (2021) Planarian stem cells specify fate yet retain potency during the cell cycle. *Cell Stem Cell*, **28**, 1307–1322.
13. van Wolfswinkel,J.C., Wagner,D.E. and Reddien,P.W. (2014) Single-cell analysis reveals functionally distinct classes within the planarian stem cell compartment. *Cell Stem Cell*, **15**, 326–339.
14. Wagner,D.E., Wang,I.E. and Reddien,P.W. (2011) Clonogenic neoblasts are pluripotent adult stem cells that underlie planarian regeneration. *Science*, **332**, 811–816.
15. Guo,T., Peters,A.H. and Newmark,P.A. (2006) A Bruno-like gene is required for stem cell maintenance in planarians. *Dev. Cell*, **11**, 159–169.
16. Zeng,A., Li,H., Guo,L., Gao,X., McKinney,S., Wang,Y., Yu,Z., Park,J., Semerad,C., Ross,E., *et al.* (2018) Prospectively isolated tetraspanin(+) neoblasts are adult pluripotent stem cells underlying planaria regeneration. *Cell*, **173**, 1593–1608.
17. Palakodeti,D., Smielewska,M., Lu,Y.C., Yeo,G.W. and Graveley,B.R. (2008) The PIWI proteins SMEDWI-2 and SMEDWI-3 are required for stem cell function and piRNA expression in planarians. *RNA*, **14**, 1174–1186.
18. Rouhana,L., Weiss,J.A., King,R.S. and Newmark,P.A. (2014) PIWI homologs mediate histone H4 mRNA localization to planarian chromatoid bodies. *Development*, **141**, 2592–2601.
19. Newmark,P.A. and Sánchez Alvarado,A. (2000) Bromodeoxyuridine specifically labels the regenerative stem cells of planarians. *Dev. Biol.*, **220**, 142–153.
20. Rouhana,L., Weiss,J.A., Forsthoefel,D.J., Lee,H., King,R.S., Inoue,T., Shibata,N., Agata,K. and Newmark,P.A. (2013) RNA interference by feeding in vitro-synthesized double-stranded RNA to planarians: methodology and dynamics. *Dev. Dyn.*, **242**, 718–730.
21. Pearson,B.J., Eisenhoffer,G.T., Gurley,K.A., Rink,J.C., Miller,D.E. and Sánchez Alvarado,A. (2009) Formaldehyde-based whole-mount *in situ* hybridization method for planarians. *Dev. Dyn.*, **238**, 443–450.
22. King,R.S. and Newmark,P.A. (2013) In situ hybridization protocol for enhanced detection of gene expression in the planarian Schmidtea mediterranea. *BMC Dev. Biol.*, **13**, 8.
23. Hayashi,T., Asami,M., Higuchi,S., Shibata,N. and Agata,K. (2006) Isolation of planarian X-ray-sensitive stem cells by fluorescence-activated cell sorting. *Dev. Growth Differ.*, **48**, 371–380.
24. Schindelin,J., Arganda-Carreras,I., Frise,E., Kaynig,V., Longair,M., Pietzsch,T., Preibisch,S., Rueden,C., Saalfeld,S., Schmid,B., *et al.* (2012) Fiji: an open-source platform for biological-image analysis. *Nat. Methods*, **9**, 676–682.
25. Dubos,T., Poulet,A., Thomson,G., Pery,E., Chausse,F., Tatout,C., Desset,S., van Wolfswinkel,J.C. and Jacob,Y. (2022) NODeJ: an ImageJ plugin for 3D segmentation of nuclear objects. *BMC Bioinf.*, **23**, 216.
26. Kavran,J.M. and Leahy,D.J. (2014) Coupling antibody to cyanogen bromide-activated sepharose. *Methods Enzymol.*, **541**, 27–34.
27. Vagin,V.V., Sigova,A., Li,C., Seitz,H., Gvozdev,V. and Zamore,P.D. (2006) A distinct small RNA pathway silences selfish genetic elements in the germline. *Science*, **313**, 320–324.
28. Li,C., Vagin,V.V., Lee,S., Xu,J., Ma,S., Xi,H., Seitz,H., Horwich,M.D., Syrzycka,M., Honda,B.M., *et al.* (2009) Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. *Cell*, **137**, 509–521.
29. Liu,S.Y., Selck,C., Friedrich,B., Lutz,R., Vila-Farre,M., Dahl,A., Brandl,H., Lakshmanaperumal,N., Henry,I. and Rink,J.C. (2013) Reactivating head regrowth in a regeneration-deficient planarian species. *Nature*, **500**, 81–84.
30. Robb,S.M., Gotting,K., Ross,E. and Sanchez Alvarado,A. (2015) SmedGD 2.0: the *Schmidtea mediterranea* genome database. *Genesis*, **53**, 535–546.
31. Langmead,B. and Salzberg,S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods*, **9**, 357–359.
32. Li,H., Handsaker,B., Wysoker,A., Fennell,T., Ruan,J., Homer,N., Marth,G., Abecasis,G., Durbin,R. and Genome Project Data Processing, S.Genome Project Data Processing, S. (2009) The sequence alignment/map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.
33. Grohme,M.A., Schloissnig,S., Rozanski,A., Pippel,M., Young,G.R., Winkler,S., Brandl,H., Henry,I., Dahl,A., Powell,S., *et al.* (2018) The genome of *Schmidtea mediterranea* and the evolution of core cellular mechanisms. *Nature*, **554**, 56–61.
34. Dobin,A., Davis,C.A., Schlesinger,F., Drenkow,J., Zaleski,C., Jha,S., Batut,P., Chaisson,M. and Gingeras,T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.
35. Smit,A., Hubley,R. and Green,P. (2013-2015) RepeatMasker Open-4.0.
36. Bao,W., Kojima,K.K. and Kohany,O. (2015) Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mobile DNA*, **6**, 11.
37. Bendall,M.L., de Mulder,M., Iniguez,L.P., Lecanda-Sanchez,A., Perez-Losada,M., Ostrowski,M.A., Jones,R.B., Mulder,L.C.F., Reyes-Teran,G., Crandall,K.A., *et al.* (2019) Telescope: characterization of the retrotranscriptome by accurate estimation of transposable element expression. *PLoS Comput. Biol.*, **15**, e1006453.
38. Love,M.I., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
39. Martin,M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, **2011**, 17,
40. Langmead,B. (2010) Aligning short sequencing reads with Bowtie. *Curr. Protoc. Bioinformatics*, **Chapter 11**, Unit 11,17.
41. Panda,A.C., Martindale,J.L. and Gorospe,M. (2017) Polysome fractionation to analyze mRNA distribution profiles. *Bio Protoc*, **7**, e2126.
42. Wagner,D.E., Ho,J.J. and Reddien,P.W. (2012) Genetic regulators of a pluripotent adult stem cell system in planarians identified by RNAi and clonal analysis. *Cell Stem Cell*, **10**, 299–311.
43. Teefy,B.B., Siebert,S., Cazet,J.F., Lin,H. and Juliano,C.E. (2020) PIWI-piRNA pathway-mediated transposable element repression in hydra somatic stem cells. *RNA*, **26**, 550–563.

44. Zhou,X., Battistoni,G., El Demerdash,O., Gurtowski,J., Wunderer,J., Falciatori,I., Ladurner,P., Schatz,M.C., Hannon,G.J. and Wasik,K.A. (2015) Dual functions of Macpiwi1 in transposon silencing and stem cell maintenance in the flatworm Macrostomum lignano. *RNA*, **21**, 1885–1897.
45. Kim,I.V., Duncan,E.M., Ross,E.J., Gorbovytska,V., Nowotarski,S.H., Elliott,S.A., Sanchez Alvarado,A. and Kuhn,C.D. (2019) Planarians recruit piRNAs for mRNA turnover in adult stem cells. *Genes Dev.*, **33**, 1575–1590.
46. Li,D., Taylor,D.H. and van Wolfswinkel,J.C. (2021) PIWI-mediated control of tissue-specific transposons is essential for somatic cell differentiation. *Cell Rep.*, **37**, 109776.
47. Brennecke,J., Aravin,A.A., Stark,A., Dus,M., Kellis,M., Sachidanandam,R. and Hannon,G.J. (2007) Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell*, **128**, 1089–1103.
48. Aravin,A.A., Sachidanandam,R., Girard,A., Fejes-Toth,K. and Hannon,G.J. (2007) Developmentally regulated piRNA clusters implicate MILI in transposon control. *Science*, **316**, 744–747.
49. Li,X.Z., Roy,C.K., Dong,X., Bolcun-Filas,E., Wang,J., Han,B.W., Xu,J., Moore,M.J., Schimenti,J.C., Weng,Z., *et al.* (2013) An ancient transcription factor initiates the burst of piRNA production during early meiosis in mouse testes. *Mol. Cell*, **50**, 67–81.
50. Feltzin,V.L., Khaladkar,M., Abe,M., Parisi,M., Hendriks,G.J., Kim,J. and Bonini,N.M. (2015) The exonuclease Nibbler regulates age-associated traits and modulates piRNA length in Drosophila. *Aging Cell*, **14**, 443–452.
51. Horwich,M.D., Li,C., Matranga,C., Vagin,V., Farley,G., Wang,P. and Zamore,P.D. (2007) The drosophila RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. *Curr. Biol.*, **17**, 1265–1272.
52. Saito,K., Sakaguchi,Y., Suzuki,T., Suzuki,T., Siomi,H. and Siomi,M.C. (2007) Pimet, the Drosophila homolog of HEN1, mediates 2'-O-methylation of Piwi- interacting RNAs at their 3' ends. *Genes Dev.*, **21**, 1603–1608.
53. Nishimasu,H., Ishizu,H., Saito,K., Fukuhara,S., Kamatani,M.K., Bonnefond,L., Matsumoto,N., Nishizawa,T., Nakanaga,K., Aoki,J., *et al.* (2012) Structure and function of Zucchini endoribonuclease in piRNA biogenesis. *Nature*, **491**, 284–287.
54. Gainetdinov,I., Colpan,C., Arif,A., Cecchini,K. and Zamore,P.D. (2018) A single mechanism of biogenesis, initiated and directed by PIWI proteins, explains piRNA production in most animals. *Mol. Cell*, **71**, 775–790.
55. Ipsaro,J.J., Haase,A.D., Knott,S.R., Joshua-Tor,L. and Hannon,G.J. (2012) The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis. *Nature*, **491**, 279–283.
56. Watanabe,T., Chuma,S., Yamamoto,Y., Kuramochi-Miyagawa,S., Totoki,Y., Toyoda,A., Hoki,Y., Fujiyama,A., Shibata,T., Sado,T., *et al.* (2011) MITOPLD is a mitochondrial protein essential for nuage formation and piRNA biogenesis in the mouse germline. *Dev. Cell*, **20**, 364–375.
57. Mohn,F., Handler,D. and Brennecke,J. (2015) Noncoding RNA. PiRNA-guided slicing specifies transcripts for zucchini-dependent, phased piRNA biogenesis. *Science*, **348**, 812–817.
58. Han,B.W., Wang,W., Li,C., Weng,Z. and Zamore,P.D. (2015) Noncoding RNA. PiRNA-guided transposon cleavage initiates zucchini-dependent, phased piRNA production. *Science*, **348**, 817–821.
59. Cora,E., Pandey,R.R., Xiol,J., Taylor,J., Sachidanandam,R., McCarthy,A.A. and Pillai,R.S. (2014) The MID-PIWI module of Piwi proteins specifies nucleotide- and strand-biases of piRNAs. *RNA*, **20**, 773–781.
60. Haase,A.D., Fenoglio,S., Muerdter,F., Guzzardo,P.M., Czech,B., Pappin,D.J., Chen,C., Gordon,A. and Hannon,G.J. (2010) Probing the initiation and effector phases of the somatic piRNA pathway in Drosophila. *Genes Dev.*, **24**, 2499–2504.
61. Olivieri,D., Sykora,M.M., Sachidanandam,R., Mechtler,K. and Brennecke,J. (2010) An in vivo RNAi assay identifies major genetic and cellular requirements for primary piRNA biogenesis in Drosophila. *EMBO J.*, **29**, 3301–3317.
62. Ishigaki,Y., Li,X., Serin,G. and Maquat,L.E. (2001) Evidence for a pioneer round of mRNA translation: mRNAs subject to nonsense-mediated decay in mammalian cells are bound by CBP80 and CBP20. *Cell*, **106**, 607–617.
63. Lejeune,F., Ishigaki,Y., Li,X. and Maquat,L.E. (2002) The exon junction complex is detected on CBP80-bound but not eIF4E-bound mRNA in mammalian cells: dynamics of mRNP remodeling. *EMBO J.*, **21**, 3536–3545.
64. Choe,J., Ryu,I., Park,O.H., Park,J., Cho,H., Yoo,J.S., Chi,S.W., Kim,M.K., Song,H.K. and Kim,Y.K. (2014) eIF4AIII enhances translation of nuclear cap-binding complex-bound mRNAs by promoting disruption of secondary structures in 5'UTR. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, E4577–E4586.
65. Kimball,C., Powers,K., Dustin,J., Poirier,V. and Pellettieri,J. (2020) The exon junction complex is required for stem and progenitor cell maintenance in planarians. *Dev. Biol.*, **457**, 119–127.
66. Siev,M., Weinberg,R. and Penman,S. (1969) The selective interruption of nucleolar RNA synthesis in HeLa cells by cordycepin. *J. Cell Biol.*, **41**, 510–520.
67. Klenov,M.S., Sokolova,O.A., Yakushev,E.Y., Stolyarenko,A.D., Mikhaleva,E.A., Lavrov,S.A. and Gvozdev,V.A. (2011) Separation of stem cell maintenance and transposon silencing functions of Piwi protein. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 18760–18765.
68. Xie,Z., Johansen,L.K., Gustafson,A.M., Kasschau,K.D., Lellis,A.D., Zilberman,D., Jacobsen,S.E. and Carrington,J.C. (2004) Genetic and functional diversification of small RNA pathways in plants. *PLoS Biol.*, **2**, E104.
69. Cecere,G. and Cogoni,C. (2009) Quelling targets the rDNA locus and functions in rDNA copy number control. *BMC Microbiol.*, **9**, 44.
70. Cam,H.P., Sugiyama,T., Chen,E.S., Chen,X., FitzGerald,P.C. and Grewal,S.I. (2005) Comprehensive analysis of heterochromatin- and RNAi-mediated epigenetic control of the fission yeast genome. *Nat. Genet.*, **37**, 809–819.
71. Chak,L.L., Mohammed,J., Lai,E.C., Tucker-Kellogg,G. and Okamura,K. (2015) A deeply conserved, noncanonical miRNA hosted by ribosomal DNA. *RNA*, **21**, 375–384.
72. Zhou,X., Feng,X., Mao,H., Li,M., Xu,F., Hu,K. and Guang,S. (2017) RdRP-synthesized antisense ribosomal siRNAs silence pre-rRNA via the nuclear RNAi pathway. *Nat. Struct. Mol. Biol.*, **24**, 258–269.
73. Ender,C., Krek,A., Friedlander,M.R., Beitzinger,M., Weinmann,L., Chen,W., Pfeffer,S., Rajewsky,N. and Meister,G. (2008) A human snoRNA with microRNA-like functions. *Mol. Cell*, **32**, 519–528.
74. Haussecker,D., Huang,Y., Lau,A., Parameswaran,P., Fire,A.Z. and Kay,M.A. (2010) Human tRNA-derived small RNAs in the global regulation of RNA silencing. *RNA*, **16**, 673–695.
75. Cherlin,T., Magee,R., Jing,Y., Pliatsika,V., Loher,P. and Rigoutsos,I. (2020) Ribosomal RNA fragmentation into short RNAs (rRFs) is modulated in a sex- and population of origin-specific manner. *BMC Biol.*, **18**, 38.
76. Girardi,E., Miesen,P., Pennings,B., Frangeul,L., Saleh,M.C. and van Rij,R.P. (2017) Histone-derived piRNA biogenesis depends on the ping-pong partners Piwi5 and Ago3 in Aedes aegypti. *Nucleic Acids Res.*, **45**, 4881–4892.
77. Barucci,G., Cornes,E., Singh,M., Li,B., Ugolini,M., Samolygo,A., Didier,C., Dingli,F., Loew,D., Quarato,P., *et al.* (2020) Small-RNA-mediated transgenerational silencing of histone genes impairs fertility in piRNA mutants. *Nat. Cell Biol.*, **22**, 235–245.
78. Wahba,L., Hansen,L. and Fire,A.Z. (2021) An essential role for the piRNA pathway in regulating the ribosomal RNA pool in C. elegans. *Dev. Cell*, **56**, 2295–2312.
79. Wolin,S.L. and Maquat,L.E. (2019) Cellular RNA surveillance in health and disease. *Science*, **366**, 822–827.
80. Lim,A.K. and Kai,T. (2007) Unique germ-line organelle, nuage, functions to repress selfish genetic elements in Drosophila melanogaster. *Proc. Natl. Acad. Sci. USA*, **104**, 6714–6719.

81. Voronina,E., Seydoux,G., Sassone-Corsi,P. and Nagamori,I. (2011) RNA granules in germ cells. *Cold Spring Harb. Perspect. Biol.*, **3**, a002774.
82. Ge,D.T., Wang,W., Tipping,C., Gainetdinov,I., Weng,Z. and Zamore,P.D. (2019) The RNA-binding ATPase, armitage, couples piRNA amplification in nuage to phased piRNA production on mitochondria. *Mol. Cell*, **74**, 982–995.
83. Cole,S.E., LaRiviere,F.J., Merrikh,C.N. and Moore,M.J. (2009) A convergence of rRNA and mRNA quality control pathways revealed by mechanistic analysis of nonfunctional rRNA decay. *Mol. Cell*, **34**, 440–450.
84. LaRiviere,F.J., Cole,S.E., Ferullo,D.J. and Moore,M.J. (2006) A late-acting quality control process for mature eukaryotic rRNAs. *Mol. Cell*, **24**, 619–626.
85. Becker,D., Hirsch,A.G., Bender,L., Lingner,T., Salinas,G. and Krebber,H. (2019) Nuclear pre-snRNA export is an essential quality assurance mechanism for functional spliceosomes. *Cell Rep.*, **27**, 3199–3214.

Received: April 13, 2023. **Revised:** December 5, 2023. **Editorial Decision:** December 6, 2023. **Accepted:** December 8, 2023

Supplementary Figure Legends

Supplemental Figure 1. Ad Figure 1
A. qPCR of individual neoblast genes in control samples and *smedwi-1(RNAi)* samples.
B. Examples of FACS data related to Figure 1A. Shown are the far red (horizontal axis) and blue (vertical axis) emission signals from Hoechst 33342 on samples of control animals (top) and *smedwi-1(RNAi)* animals (bottom).
C. Quantification of TUNEL assay to label apoptotic cells in control animals and *smedwi-1(RNAi)* animals.

Supplemental Figure 2. Ad Figure 2.
A. Relative proportions of small RNA reads mapping to various genetic elements in the total RNA input, SMEDWI-1 bound piRNAs (β-eliminated – repeated from Figure 2A for clarity), and SMEDWI-2 bound piRNAs (β-eliminated).
B. Size distribution of piRNAs bound to SMEDWI-1 and to SMEDWI-2.
C. Comparison of the relative contributions of the piRNAs bound to SMEDWI-1 and SMEDWI-2.
D. Heatmap comparing the expression levels (10log RPM) of transposable elements in whole animals in control samples, *smedwi-1(RNAi)* samples, and *smedwi-2(RNAi)* samples. Each column reflects an independent sequencing library.

Supplemental Figure 3. Ad Figure 3.
Size distribution, directionality, sequence logos, and genome browser views of small RNAs, separated by the type of locus they match. Size distributions and directionality are shown for input (left panel), SMEDWI-1 IP (middle panel) and β-eliminated SMEDWI-1 IP (right panel). Sense piRNAs are shown in blue, antisense in red. Sequence logos for the piRNAs are shown above the size plots for sense (blue shade) and antisense (red shade) piRNAs. On the right a genome browser view shows the distribution of the piRNAs over the related genomic region.
A. rRNA-matching reads have a wide size range in the input, but only a fraction binds to SMEDWI-1. SMEDWI-1-bound sequences with methylation have a 32nt size preference and a 5`U bias, suggesting they are indistinguishable from other piRNAs. These piRNAs are largely of the sense orientation and derive from the entire ribosomal transcript.
B. snRNA-matching reads consist of a few dominating sequences. The sequences bound to SMEDWI-1 are of the sense orientation and have a 5`U bias.
C. Repeat-matching piRNAs bound to SMEDWI-1 are more commonly of the antisense orientation. Both sense and antisense piRNAs have a 32nt preferred length and a 5`U bias.
D. Histone-matching piRNAs bound to SMEDWI-1 are of the sense orientation, have a 32nt preferred length, and have a 5`U bias. They derive from the entire histone transcript.
E. SMEDWI-1 bound piRNAs that match high polyA transcripts are rare – especially if they would be normalized for the high abundance of these transcripts. In the input a wide range of sizes is detected, suggesting that other (degradation) products of these transcripts exist. They are depleted from SMEDWI-1. The reads that remain in the SMEDWI-1 IPs are predominantly sense, and have a 5`U bias. They are very sparse in browser views.
F. SMEDWI-1 bound piRNAs that match low polyA transcripts in contrast are enriched in SMEDWI-1. They have a 32nt preferred length already in the input, indicating that there is no significant population of other degradation products present in the cells. The orientation of these

transcripts is sometimes difficult to determine as they are often non-coding, and the transcript levels are low. However, in genome browser views the piRNAs are predominantly from one of the strands.

Supplemental Figure 4. Ad Figure 4.
A. Northern blot of RNA from control samples and *smedwi-1(RNAi)* samples shows no evidence of abnormal rRNA products or prominent degradation fragments upon loss of SMEDWI-1. The probe used matches the 5' end of the mature 28S transcript. SYBR gold staining of the same gel is shown on the right.
B. Heatmap showing the fold change (2log) of each histone transcript upon loss of SMEDWI-1, together with the baseline expression (arbitrary bins) in the RNAseq libraries. In general, the transcripts with the highest fold changes are the ones with low baseline abundance in the sequencing libraries.
C. Scatterplot of mRNA sequences showing the fold change of the mRNA in differentiated cell samples upon loss of SMEDWI-1 versus the neoblast enrichment of the transcript as determined by RNAseq analysis. Transcripts that are overrepresented in the *smedwi-1(RNAi)* differentiated cells are largely coding transcripts that are typically enriched in neoblasts (yellow half), whereas the downregulated transcripts largely correspond to genes expressed more in differentiated cells (purple half).
D. Relative contributions of coding and non-coding transcripts to the upregulated transcripts in *smedwi-1* neoblasts (yellow shade) and differentiated cells (purple shade). Transcripts were divided by general expression level, where transcripts with less that 30 RPKM were classified as low expressed. In the neoblasts most transcripts that are increased in abundance are non-coding. This is not the case in the differentiated cells.
E. Non-coding transcripts on average give rise to double the numbers of piRNAs per transcript copy compared to coding transcripts.
F. End-labeled total neoblast RNA shows several prominent bands corresponding to structural RNAs in control neoblasts, whereas in in *smedwi-1(RNAi)* neoblasts many products of varying sizes are detected.

Supplemental Figure 5. Ad Supplementary Figure 3.
A. Computational analysis of the potential sources of piRNAs classified as matching various genomic elements. For each of the piRNA classes, reads were collected and mapped to the genome allowing any hits with no more than 2 mismatches to determine from which loci these piRNAs could have originated. piRNAs annotated as related to genes predominantly map to the mRNA transcripts that they were associated with. A small fraction (less than 5%) in addition matches unannotated parts of the genome. Often these are unannotated pseudogenes that resemble the original transcript. piRNAs related to repeats at low frequency could derive from mRNA or unannotated regions, but the majority can only have derived from the repeats themselves. Around half of the piRNAs that match rRNA could also have originated from repeat regions. This is caused by the frequent inclusion of rRNA fragments in repeat loci. The other half of rRNA-matching piRNAs can only derive from the rRNAs themselves.
snRNA-matching piRNAs can derive from loci annotated as snRNA and from unannotated loci. This is largely due to the incomplete annotation of snRNA loci in the planarian genome.

B. Immunostainings of whole mounts from control animals and *zucchini(RNAi)* animals with SMEDWI-1 antibody (magenta). In the absence of Zucchini, SMEDWI-1 levels are reduced, and remaining protein is localized in small concentrations in the cell. Scale bar 10μm.
C. Immunostainings of whole mounts from control animals, *zucchini(RNAi)* animals, and *smedwi-1(RNAi)* animals with the nuage-labeling antibody Y12 (green) and SMEDWI-1 protein (magenta). In the absence of Zucchini or SMEDWI-1, nuage structures are reduced. Scale bar 50μm.

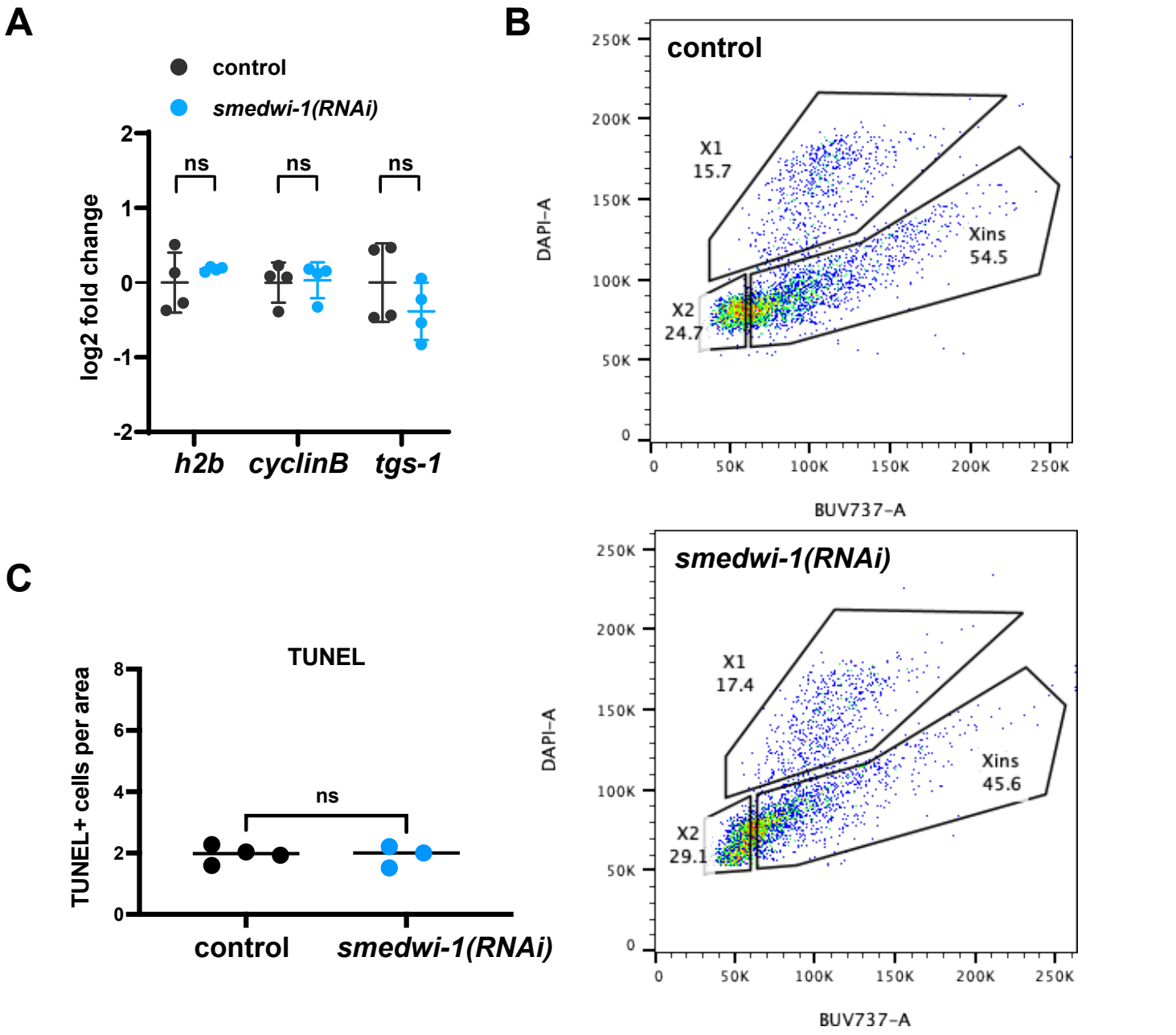
Supplemental Figure 6. Ad Materials.
Verification of antibodies and IPs.
A. Western blot showing total planarian lysate labeled with the antibody raised against a SMEDWI-1 peptide (left) or the antibody raised against the N-terminal domain of SMEDWI-1 (right). In both cases the blot shows a single band at the expected size.
B. Western blot of control lysate, lysate from irradiated animals to deplete the neoblasts, and animals treated with RNAi for *smedwi-1*. Each blot is stained with antibody for β-tubulin as a loading control, and with the SMEDWI-1 antibody (peptide antibody top; protein antibody bottom). The band recognized by the antibody is lost upon elimination of the neoblasts, or upon knockdown of *smedwi-1* confirming the specificity of the antibodies.
C. Control IPs. SMEDWI-1 protein is not aspecifically co-precipitated with antibody against rabbit IgG (left), or β-tubulin (right), indicating that the coprecipitation of SMEDWI-1 with other antibodies is specific.
D. Verification of the RNaseA treatment. RNaseA treatment results in a strong reduction of polysomes, indicating that larger RNA complexes have been dissolved.
E. Verification of immunoprecipitation from urea precipitated protein samples. Immunoprecipitation of β-tubulin results in the co-precipitation of α-tubulin, both in native samples and urea precipitated samples. SMEDWI-1 protein is still not aspecifically co-precipitated in urea precipitated samples, indicating that co-immunoprecipitation found in the urea precipitated samples is specific.

Supplementary Table 1
Primer sequences used in this study.

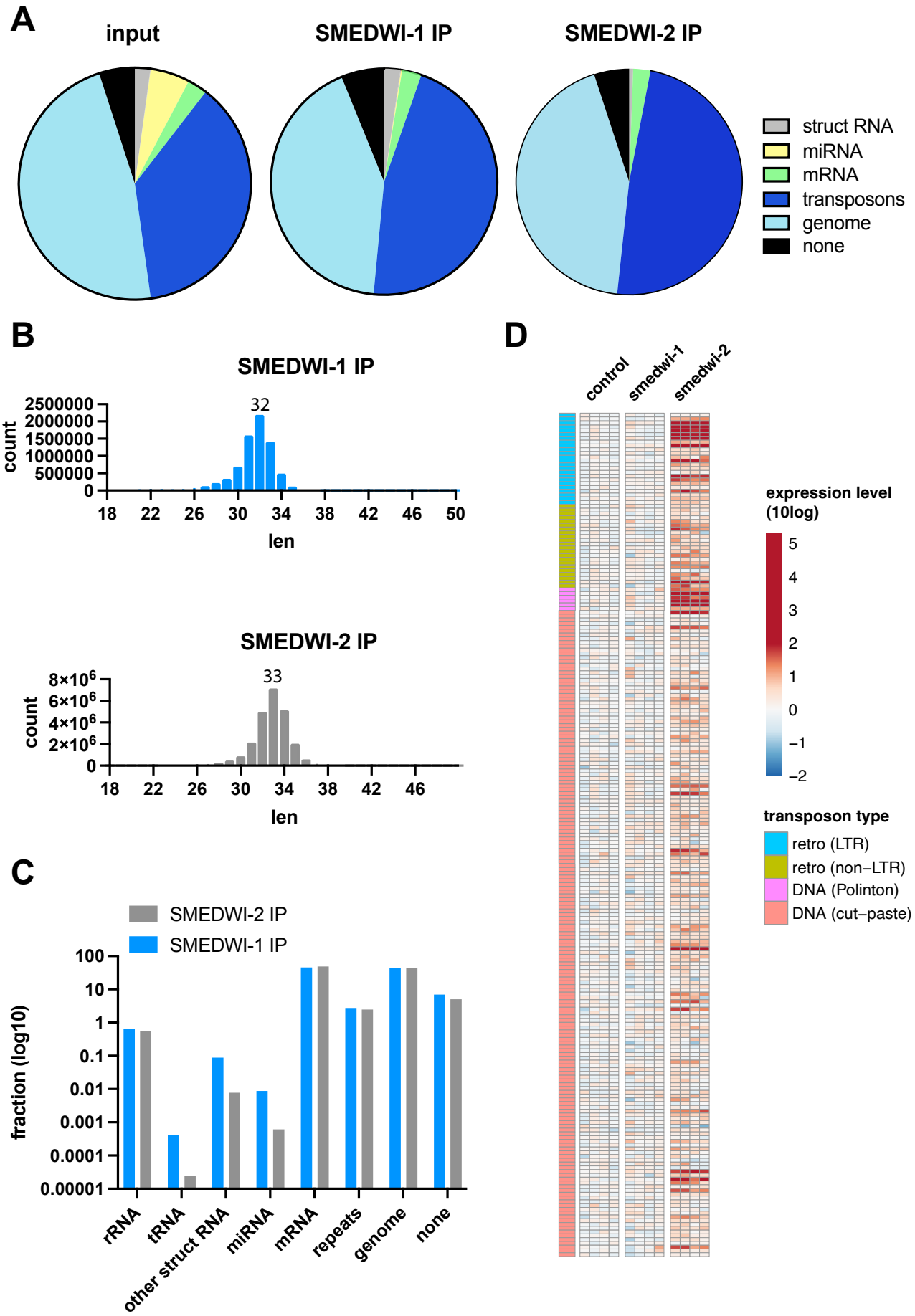
Supplementary Table 2
Peptide counts for translation-related proteins identified in SMEDWI-1 mass spectrometry.

Supplementary Table 3
Annotated table of mass spectrometry identifications from SMEDWI-1 immunoprecipitation.

Suppl Fig 1

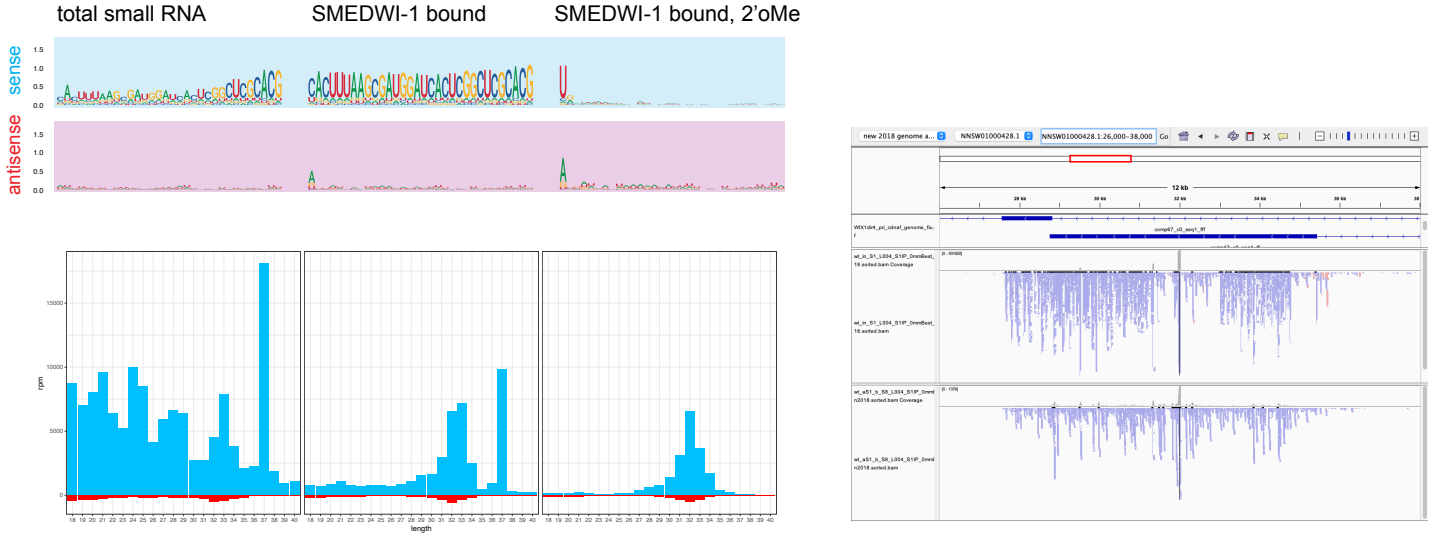

A
control
smedwi-1(RNAi)
log2 fold change
2
1
0
-1
-2
ns
ns
ns
h2b
cyclinB
tgs-1
B
control
DAPI-A
BUV737-A
X1
15.7
X2
24.7
Xins
54.5
250K
200K
150K
100K
50K
0
smedwi-1(RNAi)
X1
17.4
X2
29.1
Xins
45.6
C
TUNEL
TUNEL+ cells per area
8
6
4
2
0
ns
control
smedwi-1(RNAi)

Suppl Fig 2

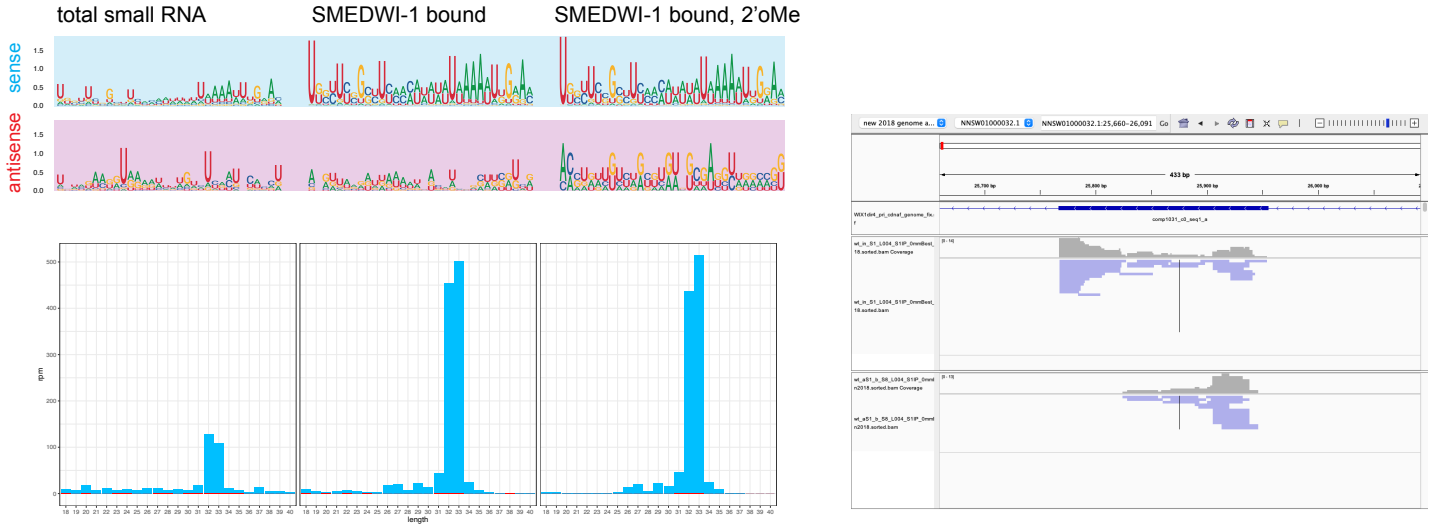

A
input
SMEDWI-1 IP
SMEDWI-2 IP
struct RNA
miRNA
mRNA
transposons
genome
none
B
SMEDWI-1 IP
count
len
32
SMEDWI-2 IP
33
C
SMEDWI-2 IP
SMEDWI-1 IP
fraction (log10)
rRNA
tRNA
other struct RNA
miRNA
mRNA
repeats
genome
none
D
control
smedwi-1
smedwi-2
expression level (10log)
transposon type
retro (LTR)
retro (non–LTR)
DNA (Polinton)
DNA (cut–paste)

Suppl Fig 3

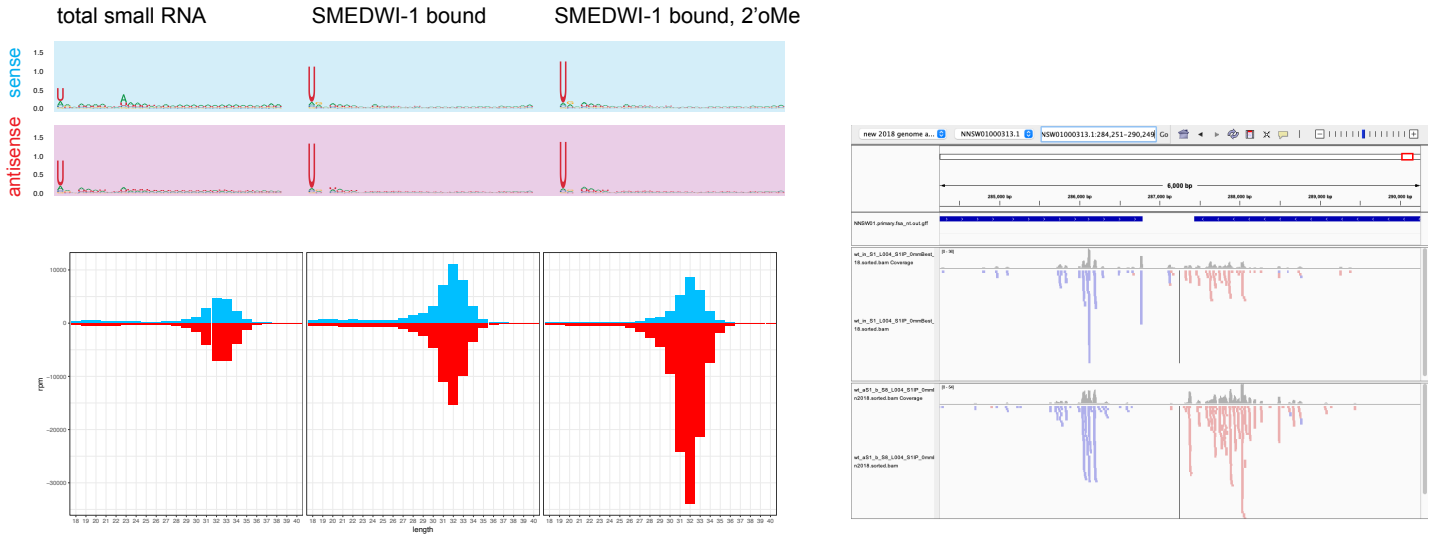
A rRNA-matching reads



B snRNA-matching reads

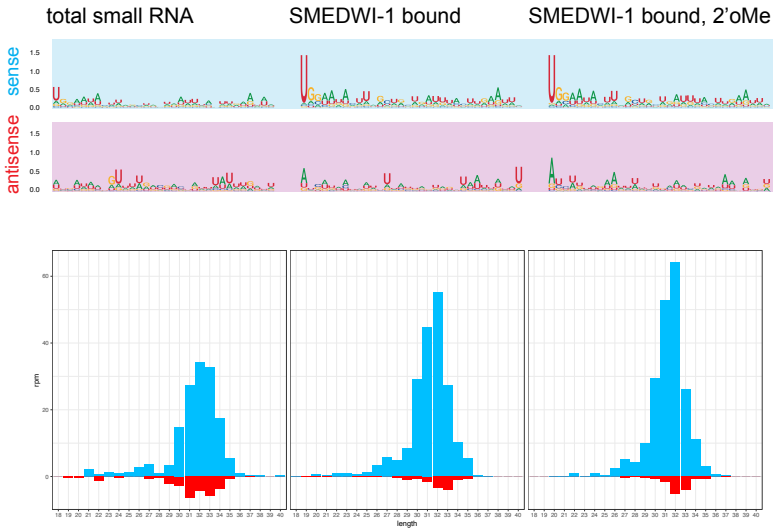


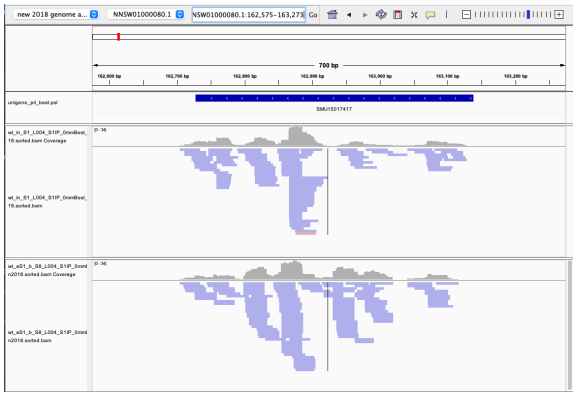
C repeat-matching reads



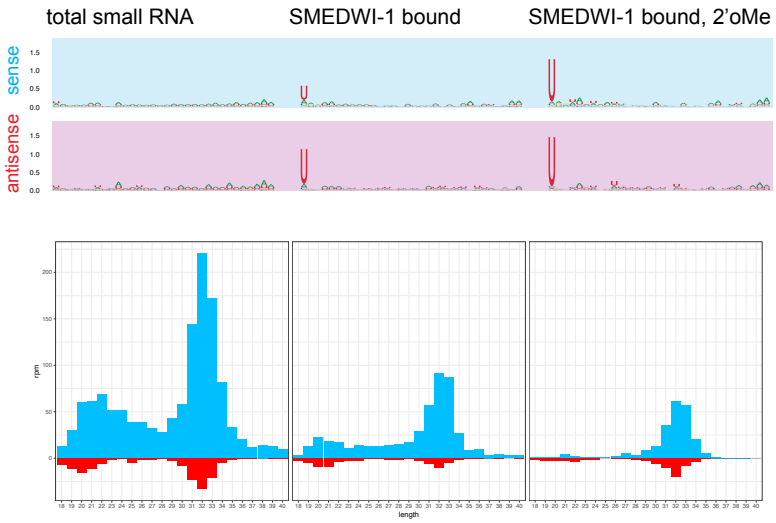
Suppl Fig 3 continued

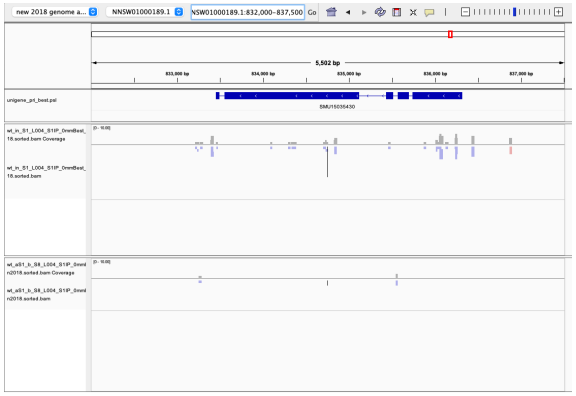
D histone mRNA-matching reads



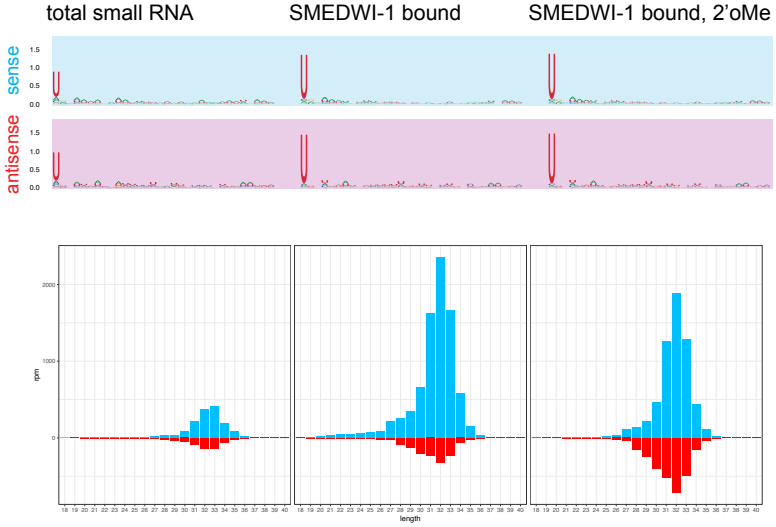


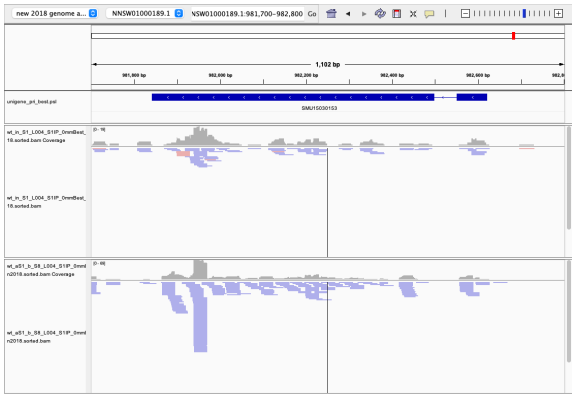
E high polyA mRNA-matching reads



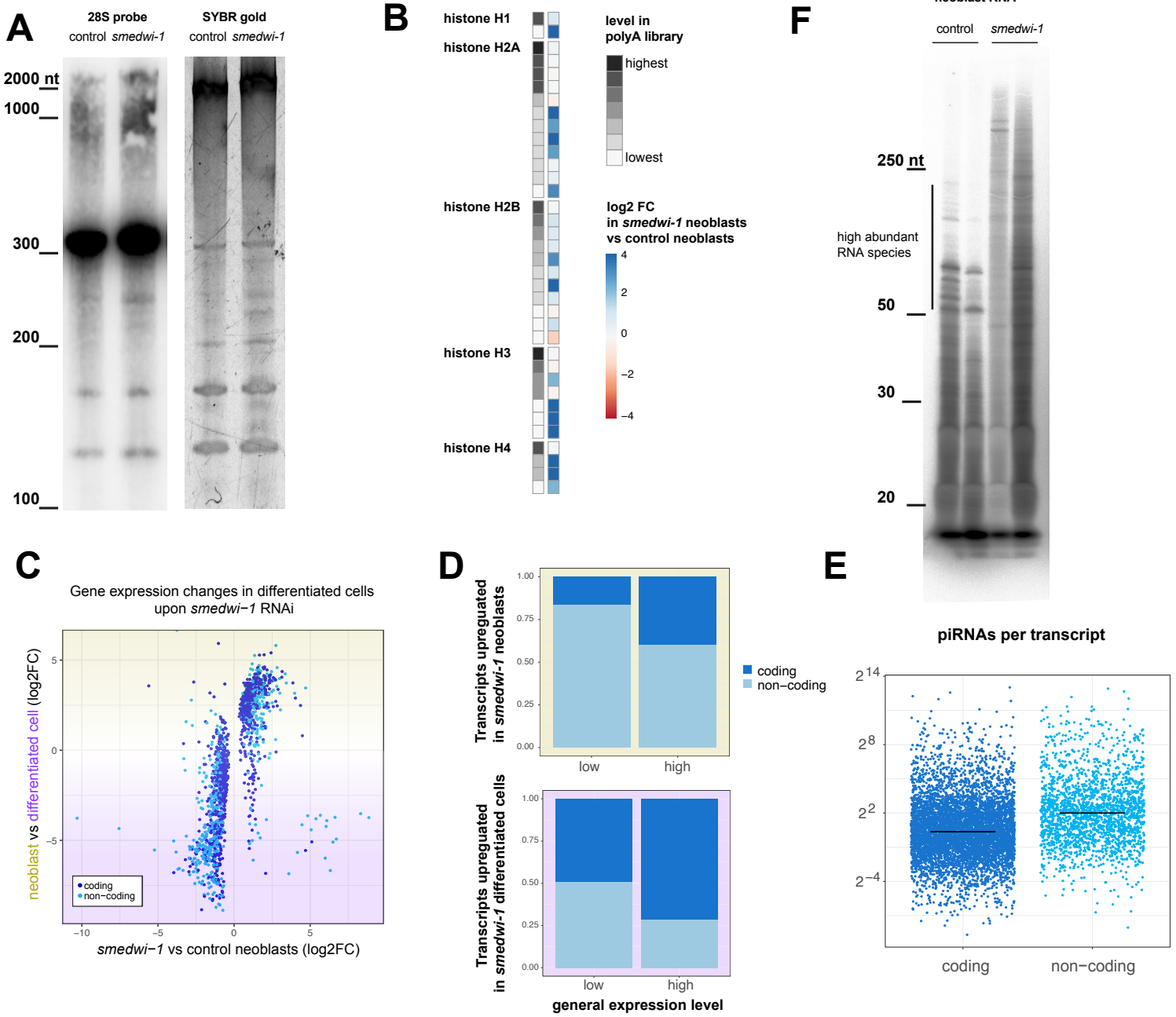


F low polyA mRNA-matching reads

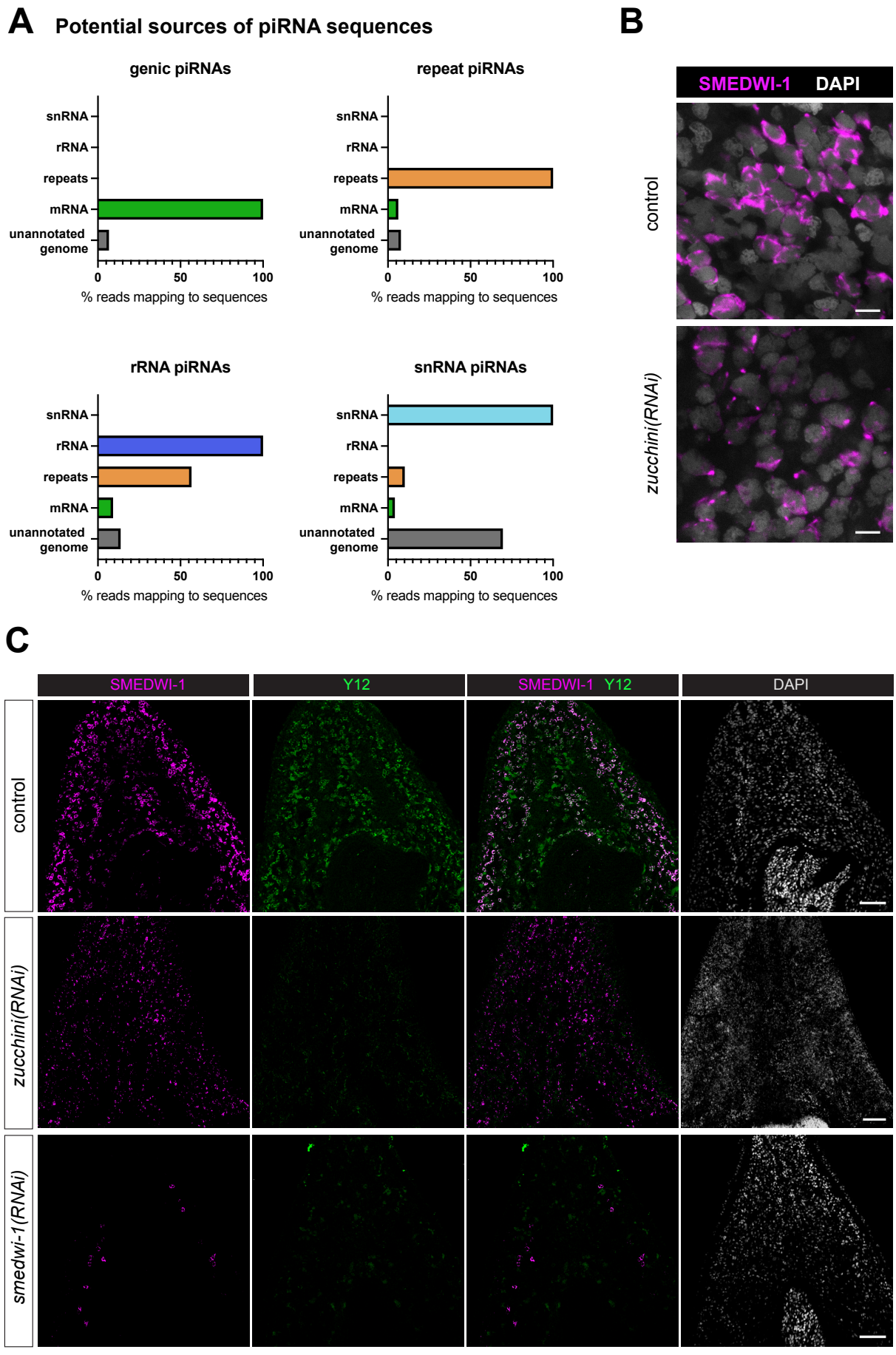




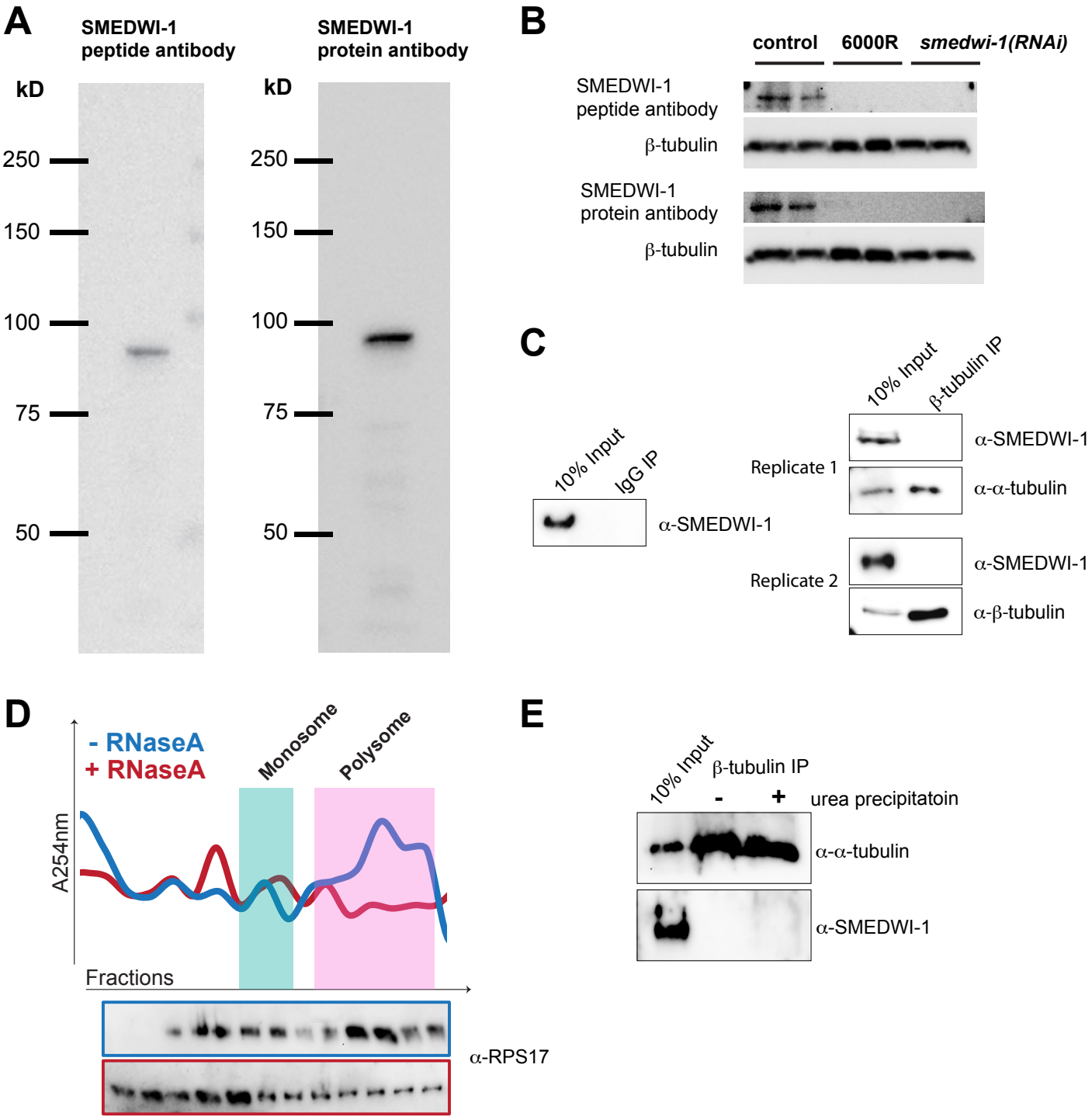
Suppl Fig 4


A
28S probe
control smedwi-1
SYBR gold
control smedwi-1
2000 nt
1000
300
200
100
B
histone H1
histone H2A
histone H2B
histone H3
histone H4
level in polyA library
highest
lowest
log2 FC in smedwi-1 neoblasts vs control neoblasts
4
2
0
−2
−4
F
neoblast RNA
control
smedwi-1
250 nt
high abundant RNA species
50
30
20
C
Gene expression changes in differentiated cells upon smedwi−1 RNAi
neoblast vs differentiated cell (log2FC)
smedwi−1 vs control neoblasts (log2FC)
coding
non−coding
D
Transcripts upregulated in smedwi-1 neoblasts
Transcripts upregulated in smedwi-1 differentiated cells
low
high
general expression level
coding
non−coding
E
piRNAs per transcript
coding
non−coding

Suppl Fig 5


A
Potential sources of piRNA sequences
genic piRNAs
repeat piRNAs
rRNA piRNAs
snRNA piRNAs
snRNA
rRNA
repeats
mRNA
unannotated genome
0
50
100
% reads mapping to sequences
B
SMEDWI-1
DAPI
control
zucchini(RNAi)
C
SMEDWI-1
Y12
SMEDWI-1 Y12
DAPI
control
zucchini(RNAi)
smedwi-1(RNAi)

Suppl Fig 6


A
SMEDWI-1
peptide antibody
SMEDWI-1
protein antibody
kD
250
150
100
75
50
B
control
6000R
smedwi-1(RNAi)
SMEDWI-1
peptide antibody
β-tubulin
SMEDWI-1
protein antibody
β-tubulin
C
10% Input
IgG IP
α-SMEDWI-1
10% Input
β-tubulin IP
Replicate 1
α-SMEDWI-1
α-α-tubulin
Replicate 2
α-SMEDWI-1
α-β-tubulin
D
- RNaseA
+ RNaseA
Monosome
Polysome
A254nm
Fractions
α-RPS17
E
10% Input
β-tubulin IP
-
+
urea precipitatoin
α-α-tubulin
α-SMEDWI-1