nature structural & molecular biology

Loss of the Mili-interacting Tudor domain–containing protein-1 activates transposons and alters the Mili-associated small RNA profile

Michael Reuter[1], Shinichiro Chuma[2], Takashi Tanaka[2], Thomas Franz[3], Alexander Stark[4] & Ramesh S Pillai[1]

Piwi proteins and their associated Piwi-interacting RNAs (piRNAs) are implicated in transposon silencing in the mouse germ line. There is currently little information on additional proteins in the murine Piwi complex and how they might regulate the entry of transcripts that accumulate as piRNAs in the Piwi ribonucleoprotein (piRNP). We isolated Mili-containing complexes from adult mouse testes and identified Tudor domain–containing protein-1 (Tdrd1) as a factor specifically associated with the Mili piRNP throughout spermatogenesis. Complex formation is promoted by the recognition of symmetrically dimethylated arginines at the N terminus of Mili by the tudor domains of Tdrd1. Similar to a *Mili* mutant, mice lacking *Tdrd1* show derepression of L1 transposons accompanied by a loss of DNA methylation at their regulatory elements and delocalization of Miwi2 from the nucleus to the cytoplasm. Finally, we show that Mili piRNPs devoid of Tdrd1 accept the entry of abundant cellular transcripts into the piRNA pathway and accumulate piRNAs with a profile that is drastically different from that of the wild type. Our data suggest that Tdrd1 ensures the entry of correct transcripts into the normal piRNA pool.

Small RNAs function in gene-silencing pathways by guiding members of the Argonaute (Ago) family of RNA binding proteins to their targets. The ~21-nucleotide (nt) microRNAs (miRNAs) bound by the Ago subfamily members participate in post-transcriptional regulation of gene expression[1]. In several animal species, ~24–31-nt germ line–specific small RNAs, called piRNAs, associate with the Piwi subfamily members of the Ago family[2,3]. Piwi proteins and piRNAs in *Drosophila melanogaster* and mice are implicated in transposon silencing[2,4].

The mouse male germ line expresses three Piwi proteins: Miwi2, Mili and Miwi. The expression of Mili is detected early in embryonic development and persists into the postnatal stages[5]. In adults, Mili expression peaks in cells in the pachytene stage of meiosis (pachytene spermatocytes), with reduced levels seen in haploid postmeiotic round spermatids. Mili expression overlaps with that of Miwi2 during embryonic stages, as Miwi2 is not detected beyond 3 d after birth (3 dpp; days postpartum). The functional implication of this overlapping expression for biogenesis of piRNAs has been described[6]. The expression of the third Piwi protein, Miwi, overlaps with Mili in pachytene spermatocytes and round spermatids[7]. Thus, Mili is a central protein in the piRNA pathway and is expressed throughout spermatogenesis up to the round spermatids stage.

The small RNAs entering Mili have been analyzed extensively, and they show dynamic changes throughout development[6]. During embryonic stages at 16.5 dpc (days post coitum), Mili piRNAs are enriched (45%) in transposon-derived small RNAs[6], with exonic sequences forming only a tiny (4%) fraction of the bound pool. At postnatal stages, including 10 dpp[8] and in adults[9], Mili starts accumulating increasing amounts of small RNAs derived from cellular transcripts. It is not clear if this change reflects a programmed shift in the Piwi-associated small RNA constituents or a random sampling of the prevailing cellular transcripts; nor is it known whether there are mechanisms in place to prevent entry of abundant cellular transcripts into this small RNA pathway.

The functional significance of all RNAs entering Mili is presently unknown. However, the transposon-derived small RNAs entering Mili are believed to guide DNA methylation, thus silencing these elements[8,10]. A *Mili* mutant derepresses both LINE-1 (L1; long interspersed nuclear elements) and intracisternal A particle (IAP) retrotransposons and shows a loss of DNA methylation in their regulatory regions. Another phenotype of this mutant is the mislocalization of Miwi2 from its normal nuclear location to the cytoplasm. The mislocalized Miwi2 was shown to be unloaded, pointing to a functional relationship between the two Piwi proteins in piRNA loading[6].

In an effort to understand Mili function, we aimed to identify additional proteins in the Mili complex. In this study, we describe the characterization of Tdrd1 as a Mili-associated factor and demonstrate its functional link to the piRNA pathway.

[1]European Molecular Biology Laboratory, Grenoble Outstation, France. [2]Institute for Frontier Medical Sciences, Kyoto University, Japan. [3]European Molecular Biology Laboratory, Heidelberg, Germany. [4]Institute of Molecular Pathology, Vienna, Austria. Correspondence should be addressed to R.S.P. (pillai@embl.fr).

Received 14 November 2008; accepted 6 May 2009; published online 24 May 2009; doi:10.1038/nsmb.1615

RESULTS

Tdrd1 is a Mili-associated protein

To characterize Mili-interacting proteins, we immunopurified Mili complexes (**Supplementary Fig. 1a** online) using a specific monoclonal antibody (**Supplementary Fig. 1b**) and identified individual polypeptides by MS. Tdrd1 (ref. 11), a 129-kDa chromatoid body component, was present in three independent purifications. It is composed of an N-terminal myeloid–nervy–DEAF-1 (MYND) domain followed by four tandem tudor domains (**Fig. 1a**). We confirmed the presence of Tdrd1 in Mili complexes isolated from extracts from adult testes by western blot analysis. This interaction is further validated by the reciprocal detection of Mili in isolated Tdrd1 complexes (**Fig. 1a**). Miwi, also present in these extracts (**Supplementary Fig. 1c**), does not interact with Tdrd1 (**Fig. 1a**). Similarly, we find that Tdrd1 is also not associated with endogenous Miwi2 (see below). To further investigate the specificity of the interaction with Mili, we coexpressed the indicated hemagglutinin (HA)-tagged Argonaute proteins with Myc-tagged Tdrd1 in human embryonic kidney (HEK) 293T cell cultures. Anti-HA immunoprecipitation and western blot analysis with the anti-Myc antibody revealed association of Myc-Tdrd1 with only HA-Mili (**Fig. 1a**, below).

We obtained further evidence for Tdrd1's association with Mili by co-immunoprecipitation of Tdrd1-associated RNAs. To this end, Piwi proteins and Tdrd1 were immunoprecipitated from adult testes extracts and the associated RNAs were revealed by 5′ end labeling. As expected, Mili and Miwi bind ~26-nt and ~30-nt RNAs, respectively. Confirming its presence in a Mili complex, Tdrd1 most clearly co-immunoprecipitates ~26-nt RNAs (**Fig. 1b**). We consistently did not observe any 30-nt RNA species with Tdrd1 (see below). Treatment of isolated endogenous Mili complexes with RNases did not affect retention of Tdrd1, suggesting that the interaction is RNA independent (**Fig. 1c**). Taken together, these results confirm Tdrd1 as a previously uncharacterized Mili-specific interacting protein.

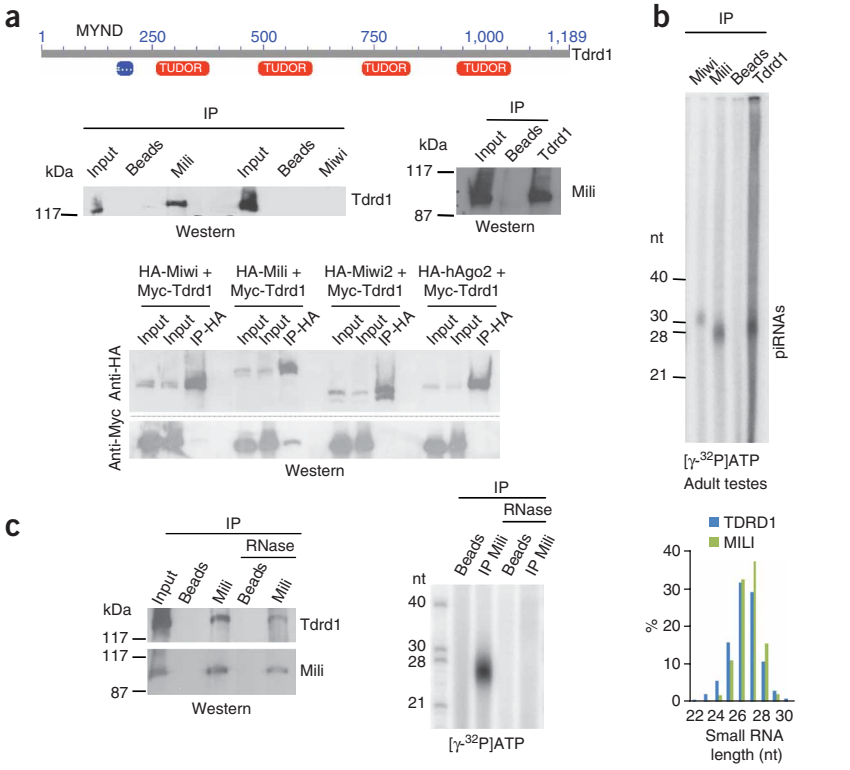

Figure 1 Tdrd1 is a Mili-specific interacting factor. (**a**) Above, known protein domains of Tdrd1 are shown. Middle, Tdrd1 interacts with Mili, but not Miwi, in adult mouse testes extracts. Reciprocally, Mili is detected in an anti-Tdrd1 immunoprecipitate (IP). Below, Myc-Tdrd1 specifically interacts with HA-Mili in HEK 293T cell cultures. Extracts of cells coexpressing the indicated tagged proteins were subjected to anti-HA immunoprecipitation (IP-HA) and analyzed with anti-Myc and anti-HA antibodies by western blot analysis. (**b**) Indicated proteins were immunoprecipitated from adult testes extracts and the associated RNAs were subjected to 5′ end labeling. The histogram shows size distribution (by percentage) of RNAs deep-sequenced from the indicated small RNA libraries (**Fig. 3**) prepared from adult mouse testes. (**c**) The Mili–Tdrd1 interaction is RNA independent, as it resists treatment with RNases. A fraction of these complexes was used for extraction of associated RNAs and detected by 5′ end labeling ([γ-^{32}P]-ATP, right). Size markers in nucleotides or kilodaltons are indicated.

Tdrd1 tudor domains recognize arginine dimethylation in Mili

We mapped the interaction between Mili and Tdrd1 by a series of coexpression and co-immunoprecipitation studies in HEK 293T cells using tagged deletion constructs. Co-immunoprecipitation of Myc-Tdrd1 with HA-Mili was lost upon deletion of the Mili N terminus (HA-Mili$^{\Delta N\text{-term}}$; lanes 6 and 9, **Fig. 2a**). In fact, a fragment comprising the Mili N terminus alone (HA-Mili$^{N\text{-term}}$) can associate with Myc-Tdrd1 (lane 7, **Fig. 2a**). Analysis of Tdrd1 indicates that deletion (Myc-Tdrd1$^{4\times Tudor}$) of its N-terminal MYND domain, a reported protein-protein interaction interface[12], did not affect its association with HA-Mili (lanes 8 and 9, **Fig. 2b**). Notably, the N terminus of Mili interacts with Myc-Tdrd1$^{4\times Tudor}$ (lanes 10 and 11, **Fig. 2b**). These results suggest that the N terminus of Mili interacts with the tudor domains of Tdrd1.

An examination of the primary sequence of the Mili polypeptide reveals the presence of several RG and GR dipeptides at the N terminus (**Supplementary Fig. 2a** online). Arginines in a similar context are known to be symmetrically dimethylated in Sm proteins, which form part of the splicing snRNPs[13,14]. We confirmed the presence of a dimethylation mark on Arg74 of endogenous Mili by MS (**Fig. 2a**). Although it is likely that further RG and GR dipeptides are also modified, they were not detected in this analysis. The symmetrically dimethylated arginine (sDMA)-glycine dipeptides of Sm proteins are recognized by the anti-Sm mAb Y12 (ref. 14), which has been used to demonstrate the existence of similar modifications on other proteins[15,16]. Using the Y12 antibody, a band corresponding to the full-length HA-Mili polypeptide, but not its deletion fragments, is clearly recognized by western blot analysis (**Fig. 2a**, below). The failure to detect HA-Mili$^{N\text{-term}}$ with the Y12 antibody (**Supplementary Fig. 2b**) may be explained by the fact that HA-Mili$^{N\text{-term}}$ may be a poor substrate for the protein arginine methyltransferase in HEK 293T cells responsible for this modification. The finding that HA-Mili$^{N\text{-term}}$ still interacts with Myc-Tdrd1 might indicate that some interaction can still occur in the absence of dimethylation, as has previously been observed for Sm protein interactions[17]. Taken together,

we conclude that the N terminus of Mili, which interacts with Tdrd1's tudor domains, is modified by sDMA.

Next, to investigate whether the endogenous Mili complexes can be recognized by the Y12 antibody, we performed immunoprecipitations from adult mouse testes and HEK 293T cell extracts using the Y12 mAb. The Y12 antibody precipitated longer RNA species from both extracts, some of which were likely to be the Sm protein–associated small nuclear RNAs (snRNAs). Notably, small RNAs similar in size to those associated with Mili and Miwi are precipitated with the Y12 antibody from adult mouse testes extracts (**Fig. 2c**). An examination of the Miwi protein sequence reveals the presence of N-terminal RG and GR dipeptides (**Supplementary Fig. 2a**), suggesting that these might also be modified. We confirmed the presence of both Mili and Miwi in the Y12 immunoprecipitate by western blot analysis (**Fig. 2c**). The overrepresentation of 30-nt RNAs in the Y12 immunoprecipitate is explained by the fact that they constitute the majority of piRNAs in total RNA populations[9]. Similar precipitations with the Y12 antibody using early-stage testes extracts revealed the clear presence of both Mili- and Miwi-associated piRNAs (**Supplementary Fig. 2c**). These results suggest that endogenous Piwi proteins carry arginine dimethylation marks on their N-terminal RG and GR dipeptides.

To investigate the importance of arginine dimethylation of Mili for its interaction with Tdrd1, we incubated transfected cell cultures with the methyltransferase inhibitor 5′-deoxy-5′-(methylthio)adenosine (MTA)[18] (**Fig. 2d**, above). After an 18-h treatment, we harvested the cells and performed immunoprecipitations similar to those shown in **Figure 2a**. MTA treatment reduced co-immunoprecipitation of Myc-Tdrd1 with HA-Mili when compared to cells treated with the solvent

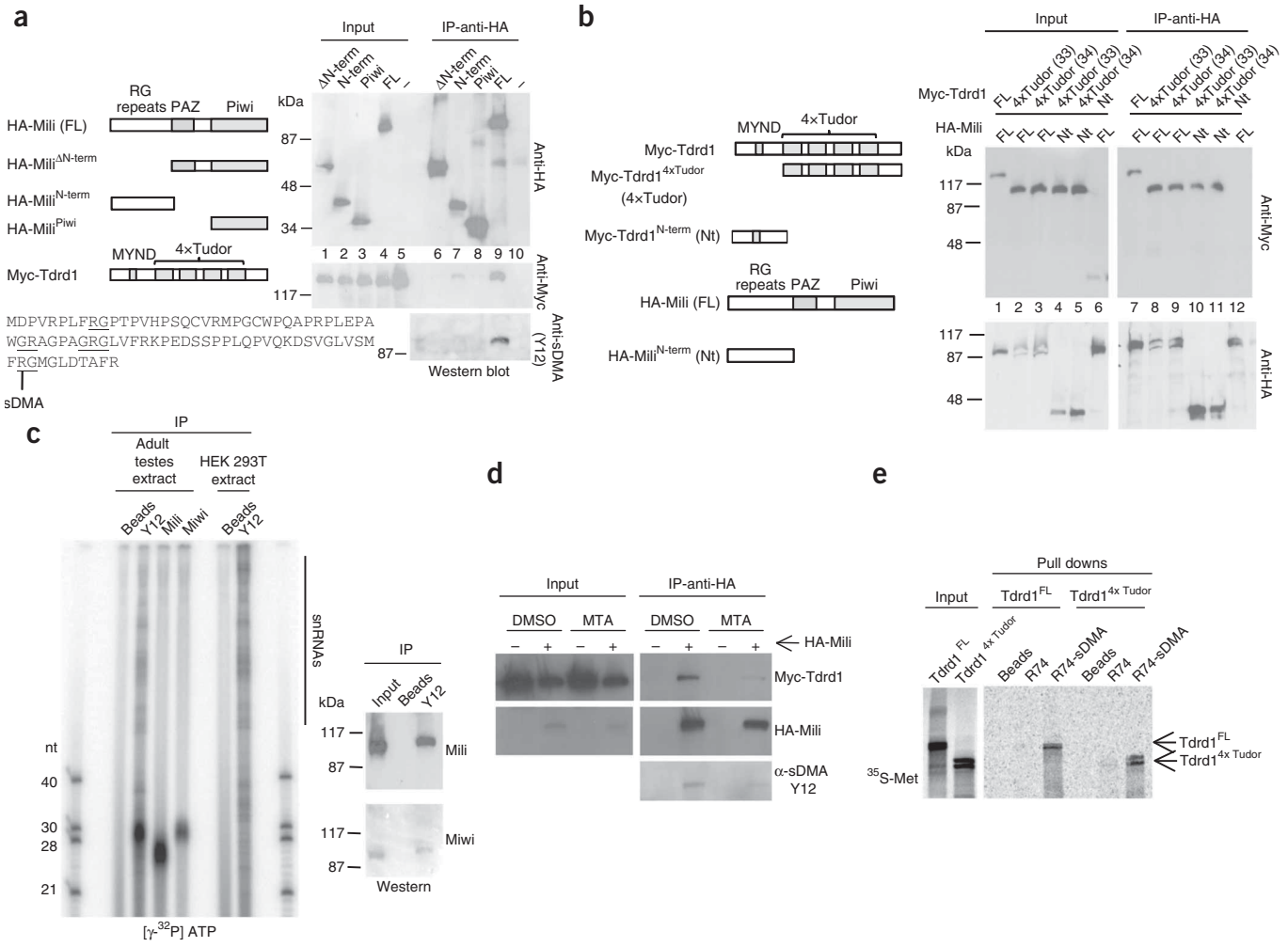

Figure 2 Tudor domains of Tdrd1 interact with the Mili N terminus, which carries symmetrical dimethyl modifications on RG dipeptides. (**a**) HA-Mili or the deleted versions shown were coexpressed with Myc-Tdrd1 in HEK 293T cells and subjected to analysis as in **Figure 1a**. The N terminus of Mili is sufficient to mediate interaction with Myc-Tdrd1 (lane 7). The blot was stripped and reprobed sequentially with anti-Myc and anti-symmetrical dimethyl (sDMA) Y12 antibodies (below). Part of the Mili sequence (amino acids 1–88) is shown with the RG and GR dipeptides underlined and the dimethyl arginine (Arg74) identified by MS indicated (labeled sDMA). Cartoons show deletion fragments of tagged proteins used and the protein domains therein. (**b**) Myc-Tdrd1 or deletions were coexpressed with HA-Mili or deletions and analyzed as above. Two different constructs (clones #33 and #34 differ by 19 amino acids in the C terminus) expressing the 4×Tudor domains were used. (**c**) Immunoprecipitations (IP) with anti-sDMA Y12 antibody using extracts from HEK 293T cells or adult mouse testes. Control IPs were performed with anti-Miwi or anti-Mili antibodies or beads alone. Associated RNAs were analyzed by 5′ end labeling and proteins by western blot analysis. (**d**) Methylation of HA-Mili (α-sDMA Y12) is reduced upon treatment of transfected HEK 293T cells with the methyltransferase inhibitor, MTA. This leads to reduced interaction of Myc-Tdrd1 with HA-Mili (**e**) *In vitro*–translated full-length Tdrd1 ($Tdrd1^{FL}$) and $Tdrd1^{4\times Tudor}$ (^{35}S-Met) interact with a synthetic 12-mer Mili peptide based on the amino acid sequence surrounding Arg74 only when it carries a sDMA mark. Size markers in nucleotides or kilodaltons are indicated.

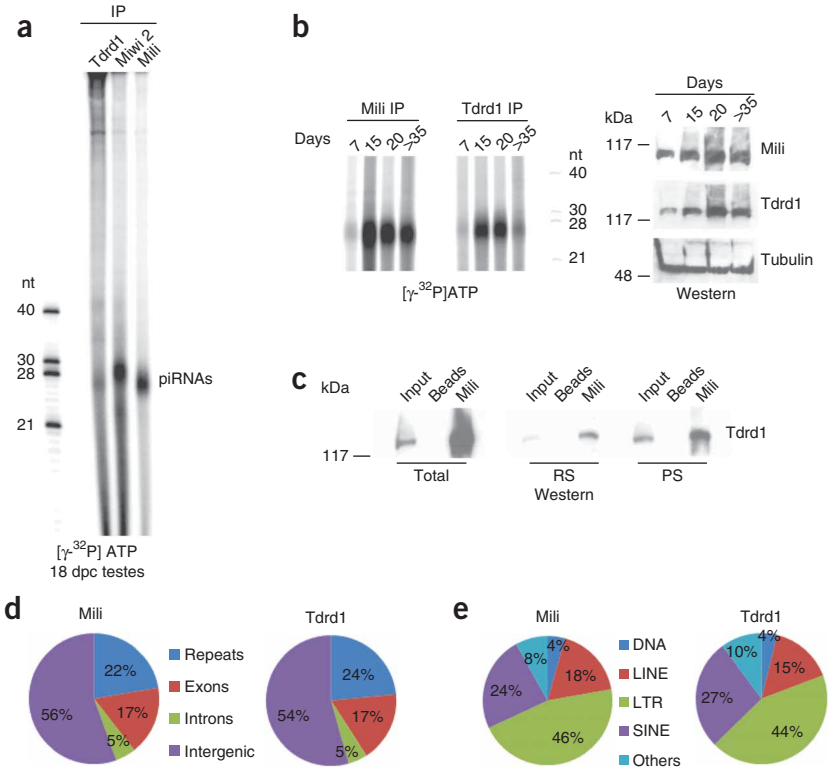

Figure 3 Tdrd1 associates with the Mili small RNAs throughout spermatogenesis. (**a**) Immunoprecipitation (IP) of Tdrd1 and Piwi proteins from testes extracts from 18-dpc embryos, and labeling of associated small RNAs. Tdrd1-associated small RNAs have the same migration as those associated with Mili. (**b**) Analysis of piRNAs ([$\gamma-^{32}$P]ATP) associated with Mili and Tdrd1 in extracts of testes isolated from animals of the indicated age (in days after birth). Western blotting to detect indicated proteins in these extracts is shown. (**c**) IP of Mili from extracts prepared with purified murine germ cells and western blotting for indicated proteins. Total, total testis extract; RS, round spermatids; PS, pachytene spermatocytes. (**d**,**e**) Genome annotation (**d**) and representation of major transposon classes (**e**) in small RNA libraries prepared from Mili- or Tdrd1-associated RNAs. Size markers in nucleotides or kilodaltons are indicated.

DMSO (**Fig. 2d**, above). This was accompanied by reduced methylation of HA-Mili, as detected by the Y12 antibody (**Fig. 2d**, below). MTA treatment did not affect either the total levels of the tagged proteins in the extracts or that of HA-Mili in the immunoprecipitations. This suggests that Mili methylation promotes its interaction with Tdrd1. To confirm this further, we incubated *in vitro*–translated Tdrd1 with biotinylated synthetic peptides corresponding to amino acids 69–80 of Mili, where Arg74 was either unmodified or carries a symmetrical dimethyl mark (Arg74-sDMA). Precipitation with streptavidin beads revealed the presence of Tdrd1 or Tdrd1$^{4\times Tudor}$ only when the methylated peptide was used and not with the unmodified peptide or beads alone (**Fig. 2e**). Using a similar experiment, we confirmed our earlier observation (**Fig. 2b**) that the N-terminal part of Tdrd1 with the MYND domain is not involved in this interaction (**Supplementary Fig. 2d**). Taken together, these results indicate that the tudor domains of Tdrd1 interact with the symmetrically dimethylated N terminus of Mili.

Tdrd1 is present in the Mili piRNP throughout development

As for Mili, the expression of Tdrd1 is detected early in development[11] and carefully traces the pattern that is reported for Mili into the adult stage. We wished to ascertain their complex formation by probing for the association of Tdrd1 with piRNAs during early spermatogenic stages. We prepared extracts from embryonic testes isolated from 18-dpc embryos and carried out immunoprecipitations with the indicated antibodies (**Fig. 3a**). As expected[6,10], small RNAs of distinct sizes co-precipitated with Miwi2 and Mili. Notably, the prominent Tdrd1-associated RNAs are similar in size to that associated with Mili, suggesting that Tdrd1 is part of the Mili complex at this embryonic stage. This interaction persists into postnatal stages, as demonstrated by co-precipitation of ~26-nt small RNAs with Tdrd1 in testes extracts from 7-dpp to adult animals (**Fig. 3b**). Although weaker in early stages, peak RNA association was observed in testes extracts from 15-dpp and 20-dpp mice, the stages when pachytene spermatocytes and round spermatids, respectively, appear in the seminiferous tubules. We examined complex formation of Tdrd1 with Mili directly in these cells by co-immunoprecipitation and western blot analysis of extracts prepared from purified germ cells (**Fig. 3c** and **Supplementary Fig. 2e**). These experiments suggest that the Tdrd1–Mili piRNP complex functions from primordial germ cells (PGCs) through to haploid postmeiotic germ cells.

To investigate whether Tdrd1 forms an integral part of the Mili complex or interacts with only a subset of the total piRNP pool, we prepared Mili- and Tdrd1-associated small RNA libraries from adult testes and deeply sequenced them using Solexa technology (**Supplementary Methods** online). A total of more than 6 million reads were obtained, mapping perfectly to the mouse genome. Consistent with its Mili-associated status, Tdrd1 reads peaked in the 26–27-nt size range (**Fig. 1b**). Similarly to other piRNAs[9], the reads from both libraries show an expected bias (~70%) for a uracil residue at the position of the 5′ nucleotide. Approximately ~0.5 million unique sequences were contained in each library, attesting to the complexity typical of published piRNA pools[6,8,9]. Although most were cloned only once, some reads were represented thousands of times with no detectable bias for the nature of the genomic region giving rise to it (**Supplementary Fig. 3a** online).

Genome annotation of the Mili library revealed that most reads derive from intergenic (~56%) unannotated regions, with repeats (22%) and genic (22%) regions forming the rest (**Fig. 3d**). This pattern is similar to that described for Mili-associated pachytene piRNAs[9]. We found a similar genomic representation of reads in the Tdrd1 library. Detailed analysis of the repeat-associated piRNAs revealed that the long terminal repeat (LTR, ~43%) retrotransposons dominate both libraries, whereas the short interspersed elements (SINEs, ~23%) and LINEs (~15%) form the remaining two major classes (**Fig. 3e**). Read-representation for different transposon families was also identical for the libraries (**Supplementary Fig. 3b** and **Supplementary Table 1** online). Indeed, at least 33% of the reads in the two libraries were identical, and 80% of the reads showed shared genomic origins (**Supplementary Fig. 3c**,**d**). Strong clustering,

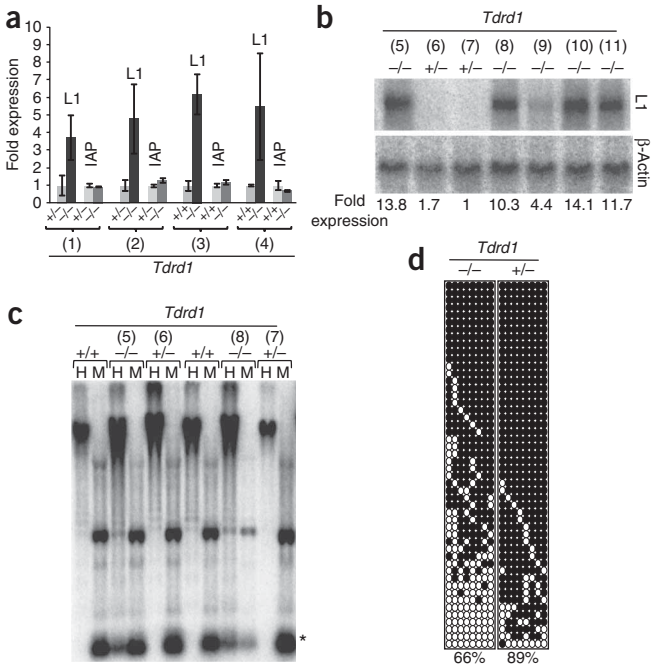

Figure 4 Activation of L1 transposons in *Tdrd1* mutant testes is coupled to loss of DNA methylation of the 5′ regulatory region of the L1 element. (**a**) qRT-PCR analysis of L1 expression in testes from wild-type (+/+), heterozygous (+/−) or homozygous (−/−) 15-dpp *Tdrd1* donor mice. Error bars represent s.d. from three experiments. DNA methylation analysis of testes genomic DNA for the same animals (number in parentheses) is given in **Supplementary Figure 4a**. (**b**) Northern blot analysis of L1 transcripts with total RNA from testes of mice with the indicated genetic backgrounds. The blot was stripped and reprobed for actin mRNA. The L1 transcript levels were normalized to that of actin and indicated below as fold expression of L1. (**c**) Methylation-sensitive Southern blot analysis using an L1 probe. Testes genomic DNA isolated from the same animals (numbers in parentheses) used for RNA analysis in **b** was digested with methylation-sensitive (HpaII, H) and methylation-insensitive (MspI, M) enzymes. The asterisk indicates the position of the methylation-sensitive (H) restriction product appearing in the homozygous mutants. (**d**) Bisulfite-treatment of genomic DNA obtained from FACS-purified germ cells of testes from 15-dpp *Tdrd1*$^{+/-}$ and *Tdrd1*$^{-/-}$ mutants, followed by sequencing of a specific L1 element[21] and determination of the methylation status of several CpGs. The results from 50 sequences obtained for each genotype are shown. Filled circles indicate methylated CpG and unfilled circles represent unmethylated CpG. The percentage of total methylated CpGs is given below.

a feature identified with published pachytene piRNAs[9], was also observed for reads in the two libraries (**Supplementary Fig. 3d** and **Supplementary Tables 2** and **3** online). Within a given cluster, the reads from the libraries arise from the same strand. These results identify the Tdrd1-containing Mili complex as being associated with most of the defined Mili-associated piRNAs[9].

Loss of DNA methylation and L1 activation in *tdrd1* mice

Loss of *Mili* leads to activation of L1 and IAP LTR retrotransposon elements in the murine germ line[8,10]. A *Tdrd1* mutant[19] shows spermatogenic arrest at the pachytene stage, similar to that seen in a *Mili* mutant[5]. Owing to the heterogeneous penetrance of the phenotype in the *Tdrd1* mutant, the germ cells in some seminiferous tubules proceed to the round spermatids stage[19]. To probe potential functions of the Mili–Tdrd1 complex, we examined whether Tdrd1 might also be required for transposon silencing. Quantitative reverse trancription–PCR (qRT-PCR) analyses of transposons in the testes of 15-dpp *Tdrd1* mutants revealed that L1 levels are elevated by fourfold to ninefold when compared to levels in heterozygous littermates (shown for four separate pairs in **Fig. 4a**). However, in contrast to the *Mili* mutant[8,10], we did not observe any changes in the expression of IAP LTR transcripts. We reached a similar conclusion using northern blot analysis of testes total RNA extracts from several *Tdrd1* mutant animals. L1 transcript levels were upregulated by 4-fold to 14-fold in the homozygous mutants, compared to their heterozygous littermates (**Fig. 4b**). Analysis with an IAP probe did not yield any signal, suggesting lack of activation (data not shown).

Activation of transposons in the *Mili* mutant is linked to loss of DNA methylation on these elements[8,10]. To examine the DNA methylation status of L1 in the *Tdrd1* mutant testes, we performed methylation-sensitive Southern blot analysis. We isolated genomic DNA from testes from the same animals used above for the RNA analyses in **Figure 4a**,**b** and digested it with a methylation-sensitive restriction enzyme, HpaII. This analysis revealed that L1 elements were demethylated in the *Tdrd1* mutants as compared to those in wild-type and heterozygous animals (**Fig. 4c** and **Supplementary Fig. 4a** online). To directly examine the methylation status of CpGs, we used flourescence-activated cell sorting (FACS) to purify testicular germ cells from 15-dpp animals, extracted the DNA and performed bisulfite sequencing of a 230-bp fragment of an L1 5′ end[20]. Whereas 89% of all CpGs were modified in the heterozygous animals, only 66% were modified in the homozygous *Tdrd1* mutants (**Fig. 4d**). The 23% difference in CpG methylation observed in the *Tdrd1* mutant is modest when compared to the 30% and 56% differences seen over two L1 regions in a *Mili* mutant[10]. A likely explanation is that, because Mili is central to the piRNA pathway, it has a more profound effect on DNA methylation. Also, some Mili functions might still be retained in the *Tdrd1* mutant. Nevertheless, these results point to a role for the Tdrd1–Mili complex in L1 silencing by DNA methylation.

Miwi2 is mislocalized in *tdrd1* mice

As Mili is implicated in *de novo* DNA methylation[10] of transposons during a tight developmental window in the embryonic stage, we examined testes isolated from 18-dpc *Tdrd1* mutant embryos for a potential impact of Tdrd1 depletion on L1 silencing. We reasoned that any transcriptional activation of L1 should lead to accumulation of the L1-encoded protein products—L1 ORF1p and L1 ORF2p. Consistently, examination by immunofluorescence revealed a large increase in the L1 ORF1p[21] (**Fig. 5a**). It accumulated uniformly in the cytoplasm of the germ cells, with enrichment in several perinuclear granules, whereas the surrounding somatic cells were devoid of any signal (**Fig. 5b**). The fetal germ cells in the heterozygous animals show only weak background levels of the protein. This result indicates activation of L1 elements in the *Tdrd1* mutant fetal testes during this crucial window when *de novo* DNA methylation is set.

In the fetal germ cells, Miwi2 is nuclear[6], whereas Mili and Tdrd1 are both cytoplasmic[11]. Mili is proposed to interact functionally with Miwi2 to generate new piRNAs in these cells. In fact, in *Mili* mutants, Miwi2 becomes cytoplasmic and is devoid of piRNAs[6]. To investigate whether there is a similar effect on Miwi2 localization in the *Tdrd1* mutant, we analyzed the two Piwi proteins by immunoflourescence. In the heterozygous animal, Mili was detected present diffusely in the cytoplasm of fetal germ cells with enrichment in a few granules. This pattern was unchanged in the *Tdrd1* mutant (**Fig. 5c**). As expected, Miwi2 was mainly nuclear, with one or two cytoplasmic granules in

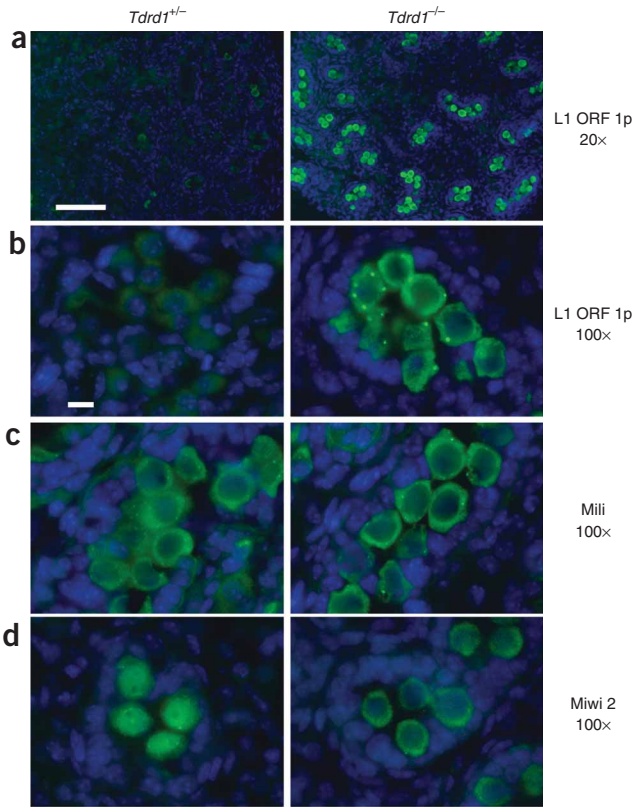

Figure 5 Activation of L1 and mislocalization of Miwi2 in the fetal germ cells of *Tdrd1* mutants. (**a**) Embryonic testes isolated from 18-dpc embryos of the indicated *Tdrd1* mutant backgrounds were examined by immunofluoresence for L1 ORF1p with a rabbit polyclonal antibody (green). Massive accumulation of the L1-encoded protein is seen in the knockout mutant germ cells. The magnification used is indicated on the right. (**b**) A 100× magnification to show cytoplasmic granules containing L1 ORF1p. (**c**) Mili localization is unaffected in the *Tdrd1* mutant. Mili is detected (green) in the cytoplasm with a few perinuclear granules. (**d**) Miwi2 localization depends on Tdrd1. Miwi2 was detected with a rabbit polyclonal antibody to the protein[6] and found to localize to the nucleus in the $Tdrd1^{+/-}$ mutant. In *Tdrd1* homozygous mutant fetal germ cells, Miwi2 is mislocalized from the nucleus to the cytoplasm. Staining for DNA with Hoechst is shown in blue. Scale bar represents 100 μm for the 20× and 10 μm for the 100× magnification.

the heterozygous animal. In a pattern markedly similar to that of a *Mili* mutant, Miwi2 became uniformly cytosolic in the *Tdrd1* mutant (**Fig. 5d**). Taken together, these results suggest that Tdrd1 is required for Mili's role in ensuring correct nuclear localization of Miwi2 to mediate *de novo* DNA methylation.

Mili-associated small RNA profile is altered in *tdrd1* mice

To understand how Tdrd1 functions in the Mili complex, we examined the abundance and subcellular localization of Mili protein in testes from *Tdrd1* mutant animals. Mili protein levels were unaffected in the homozygous mutant when compared to levels in the heterozygous animals (**Fig. 6a**). Similarly, subcellular localization was also unaffected in the mutants (**Fig. 5c** and **Supplementary Fig. 5** online). Next, we examined whether the entry of piRNAs into Mili is affected in the absence of Tdrd1. We recovered Mili complexes from testes of 15-dpp (before the onset of spermatogenic arrest in mutants) heterozygous and homozygous *Tdrd1* mutants and used end labeling to detect the associated small RNAs. We found that the overall presence of small RNAs in Mili was unchanged (**Fig. 6b**).

To characterize the nature of Mili piRNAs in the mutant, we deep sequenced Mili-associated small RNAs isolated from the homozygous ($Tdrd1^{-/-}$) mutant testes. We obtained more than 4 million uniquely mapping reads, comprising around 1.8 million unique sequences. Sequence annotations revealed a strong overrepresentation of sense-strand exonic reads in the $Tdrd1^{-/-}$ (48% exonic) library. This result was markedly different from what was observed for the wild-type Mili library (16% exonic) (**Fig. 6c**) and other published Mili pools[8,9]. The percentage of reads corresponding to sense intronic reads was also

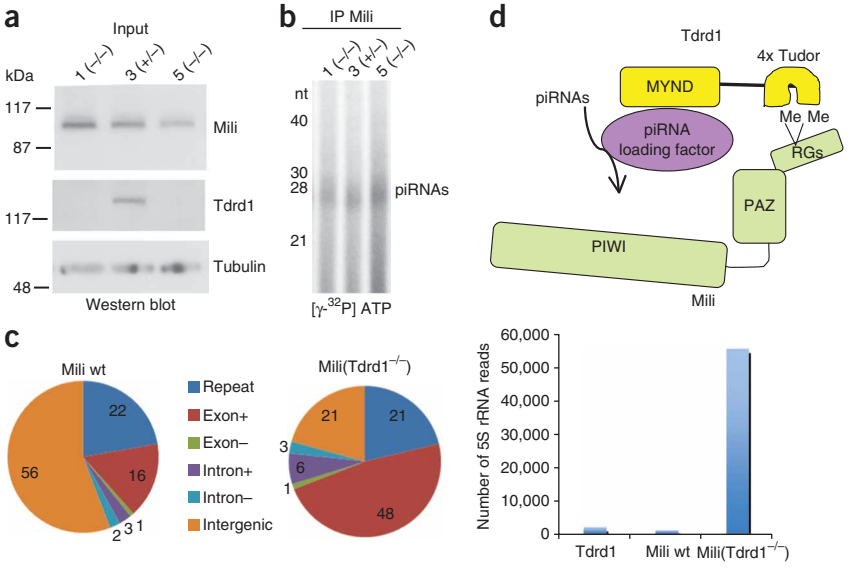

Figure 6 Loss of *Tdrd1* leads to overrepresentation of exonic reads in Mili. (**a**) Mili protein levels are unaffected in *Tdrd1* mutants. Western blot analysis of Mili, Tdrd1 and a tubulin control in extracts prepared from testes of 15-dpp mice with the indicated *Tdrd1* backgrounds. (**b**) The small RNA association of Mili is unchanged in *Tdrd1* mutants. Immunoprecipitation (IP) of Mili and labeling of associated RNAs from the testes of the same animals used in **a**. (**c**) Sequence annotations of the Mili pool of small RNAs isolated from the testes of wild-type (Mili-wt) and 15-dpp *Tdrd1* homozygous (–/–) mutants. Note the increased contribution from the exonic reads to the Mili pool in the $Tdrd1^{-/-}$. The histogram shows the number of reads corresponding to 5S rRNA in the indicated libraries. The differences cannot be accounted for by the depth of sequencing (total number of reads) or the number of reads that match to the genome, both of which differ at most twofold between the libraries. 'Tdrd1' indicates Tdrd1-associated reads from adult testes; 'Mili-wt' library is from adult testes. (**d**) A model for Tdrd1–Mili complex formation and function in the piRNA pathway. Recognition of the dimethylated arginines (RGs) in the Mili N teminus by the tudor domains of Tdrd1 mediates complex formation. The MYND domain of Tdrd1 can then recruit a putative piRNA biogenesis or loading factor to facilitate proper piRNA assembly.

higher in the mutant library. There were no changes in the antisense exonic and intronic reads, suggesting an increased contribution to the Mili pool in the *Tdrd1*$^{-/-}$ mutant from cellular transcripts. We did not observe any change in the representation of transposon families in the mutant. Sequence comparison revealed that only 19% of the *Tdrd1*$^{-/-}$ sequences are found in the wild-type Mili pool. Notably, cluster analysis revealed that there is absolutely no overlap (0.68%) in the genomic regions that give rise to reads in the two libraries. The above analyses suggest that, in the absence of Tdrd1, cellular transcripts contribute disproportionally to the Mili pool of piRNAs. This is exemplified by the presence of substantially more reads corresponding to cellular mRNAs and ribosomal RNAs (**Fig. 6c** and **Supplementary Fig. 4b**). In fact, the top genomic clusters giving rise to reads in the mutant Mili pool include genomic regions containing transposons sequences, mRNAs and rRNAs (**Supplementary Table 4** online).

DISCUSSION

Our purification of the Mili complexes identifies Tdrd1 as a previously uncharacterized Mili-specific component. Tdrd1 associates with the Mili piRNP throughout germ cell development from embryonic stages to the adult. The N terminus of Mili is modified by symmetrical dimethyl groups on RG and GR dipeptides and interacts with tudor domains of Tdrd1 to mediate this complex formation. Loss of *Tdrd1* activates L1 elements during the developmental window when Mili is required to silence transposons by *de novo* DNA methylation[10,22]. Our qRT-PCR and northern analyses of this mutant show that the activation of L1 transcripts is coupled to a loss of DNA methylation of its regulatory elements. We also show that Tdrd1 is required for Mili's role in ensuring proper nuclear localization of Miwi2. Finally, we show that in the absence of Tdrd1, Mili accumulates increasing amounts of cellular transcripts and has a distinct small RNA profile.

Our analysis of the *Tdrd1* mutant links Tdrd1 to two reported phenotypes in a *Mili* mutant. First, activation of L1 transposons is coupled to loss of DNA methylation. Second, Tdrd1 is required for proper subcellular localization of Miwi2 in the fetal germ cells. However, what is unexpected is the lack of an impact on IAP LTR silencing in the *Tdrd1* mutant. In this context, it is tempting to note that the *Tdrd1* mutant phenotype is similar to that of the the *Miwi2* mutation: robust L1 derepression, with no or very little upregulation of IAP LTR elements[10,20]. On the other hand, the *Mili* mutant shows robust activation of both L1 and IAP LTR elements[8,10]. When considered together with the mislocalization of Miwi2, it can be suggested that the loss of *Tdrd1* separates the central role of Mili in silencing both transposons.

Mili-associated RNAs show dynamic changes in their profile throughout development. Compared to embryonic stages, where 45% of the Mili pool is composed of repeat-associated piRNAs, the cellular transcripts contribute more in reported postnatal Mili piRNA pools[6,8]. This suggests a continual pressure to include small RNAs derived from such transcripts in the Mili piRNP complex. The presence of exonic reads is believed to be a consequence of random sampling of the prevailing transcriptome. The strong overrepresentation of exonic reads in the Mili pool from the *Tdrd1* mutant suggests that Tdrd1 might have a role in limiting the entry of cellular transcripts into the piRNA pathway. Loss of *Tdrd1* leads to a shift in piRNA clusters to one that is dominated by transposons and cellular transcripts. Such a Mili complex might fail to participate effectively in the RNA-driven piRNA-biogenesis pathway[2] and show failings in its transposon-silencing function.

Recognition of post-translational modifications by specific protein modules has an important role in various cellular processes. Arginine methylation is known to regulate protein-protein interactions or subcellular localization of complex components[15,17,24]. Our demonstration that the methylation status of Mili affects its interaction with Tdrd1 suggests that Tdrd1's tudor domains may sample these modifications on the Mili N terminus to facilitate complex formation. Similar symmetrical dimethylation of the RG repeats in Sm proteins is known to increase their affinity for the snRNP assembly factor SMN[13,14]. The tudor domain of SMN and those in other proteins have been described to be methylated ligand binding modules[25,26]. Consistently, a fragment consisting of the four tandem tudor domains of Tdrd1 (Tdrd1$^{4\times Tudor}$) specifically interacts with a Mili peptide dependent on the presence of a symmetrical dimethylation mark. The absence of a change in subcellular localization of Mili (both in fetal and postnatal stages) in the *Tdrd1* mutant mice suggests that this complex formation has other roles. The functional importance of such a complex is highlighted by the changed profile of Mili-associated small RNAs in the *Tdrd1* mutant. In a purely speculative model, complex formation with Tdrd1 allows the recruitment of a putative piRNA biogenesis or loading factor that ensures the correct entry of transcripts and piRNAs into Mili (**Fig. 6d**).

While this manuscript was under review, the identification of Tdrd1 as a Mili-interacting factor was reported[27]. This study mapped the interaction domain on Tdrd1 to a region encompassing the N-terminal MYND domain together with the first two tudor domains of Tdrd1; in contrast, we find that the tudor domains alone are sufficient for this interaction. That report also found that the total piRNA levels in a Tdrd1 mutant are unchanged. Consistently, we find that the loss of Tdrd1 does not affect abundance of small RNAs in Mili but rather the identity of its constituents, which in turn has an impact on Mili-mediated transposon silencing. At the same time, another study[28] identified sDMA modification in Piwi proteins from diverse species. They identified the fly protein arginine methyltransferase 5 (PRMT5)[29,30] as the enzyme responsible for this modification. The absence of PRMT5 in flies leads to instability of Piwi proteins and reduced piRNA levels. It would be interesting to see whether the unmethylated fly Piwi proteins also accumulate piRNAs with an altered profile. Homologs of Tdrd1 are also present in other species, and their analysis should yield information about whether Tdrd1 has a conserved function. Notably, a defect in a fly tudor domain protein, Krimper, results in loss of transposon silencing and reduced piRNA levels[31]. It remains to be seen whether it has any role similar to that of Tdrd1 in mice. To conclude, our identification of Tdrd1 as a Mili-interacting factor and of Piwi protein methylation should now direct further investigation into a link between the other tudor proteins in the mouse germ line and the piRNA pathway.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/nsmb/.

Accession numbers. Gene Expression Omnibus: Small RNA sequencing data used in this study are deposited under the accession number GSE16023.

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

ACKNOWLEDGMENTS

We thank A. Aravin and G. Hannon (Cold Spring Harbour Laboratory), J. Martinez (Institute of Molecular Biotechnology, Vienna), S. Martin (University of Colorado School of Medicine), A. Bortvin (Carnegie Institution of Washington) and D. Schuemperli and M. Ruepp (University of Bern) for reagents. We thank the European Molecular Biology Laboratory (EMBL)

Mononclonal Antibody Facility, Protein Expression and Gene Core facilities for antibody production and Solexa sequencing. We are grateful to D. O'Carroll, A. Verdel and W. Filipowicz and members of the Pillai group for excellent discussions and reading of the manuscript. Research in the Pillai group is supported by EMBL and a grant from Region Rhone Alp (CIBLE 2008).

AUTHOR CONTRIBUTIONS

M.R. performed most of the experiments described in this study; S.C. provided immunostaining data and *Tdrd1* mutant mice; T.T. purified germ cells by FACS; T.F. provided MS analysis; A.S. performed Solexa sequence analysis; R.S.P. designed the research and wrote the paper.

Published online at http://www.nature.com/nsmb/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Filipowicz, W., Bhattacharyya, S.N. & Sonenberg, N. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat. Rev. Genet.* **9**, 102–114 (2008).
2. Aravin, A.A., Hannon, G.J. & Brennecke, J. The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science* **318**, 761–764 (2007).
3. Seto, A.G., Kingston, R.E. & Lau, N.C. The coming of age for Piwi proteins. *Mol. Cell* **26**, 603–609 (2007).
4. Brennecke, J. *et al.* Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* **128**, 1089–1103 (2007).
5. Kuramochi-Miyagawa, S. *et al.* Mili, a mammalian member of *piwi* family gene, is essential for spermatogenesis. *Development* **131**, 839–849 (2004).
6. Aravin, A.A. *et al.* A piRNA pathway primed by individual transposons is linked to *de novo* DNA methylation in mice. *Mol. Cell* **31**, 785–799 (2008).
7. Deng, W. & Lin, H. *miwi*, a murine homolog of *piwi*, encodes a cytoplasmic protein essential for spermatogenesis. *Dev. Cell* **2**, 819–830 (2002).
8. Aravin, A.A., Sachidanandam, R., Girard, A., Fejes-Toth, K. & Hannon, G.J. Developmentally regulated piRNA clusters implicate MILI in transposon control. *Science* **316**, 744–747 (2007).
9. Aravin, A. *et al.* A novel class of small RNAs bind to MILI protein in mouse testes. *Nature* **442**, 203–207 (2006).
10. Kuramochi-Miyagawa, S. *et al.* DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. *Genes Dev.* **22**, 908–917 (2008).
11. Chuma, S. *et al.* Mouse tudor repeat-1 (MTR-1) is a novel component of chromatoid bodies/nuages in male germ cells and forms a complex with snRNPs. *Mech. Dev.* **120**, 979–990 (2003).
12. Liu, Y. *et al.* Structural basis for recognition of SMRT/N-CoR by the MYND domain and its contribution to AML1/ETO's activity. *Cancer Cell* **11**, 483–497 (2007).
13. Friesen, W.J., Massenet, S., Paushkin, S., Wyce, A. & Dreyfuss, G. SMN, the product of the spinal muscular atrophy gene, binds preferentially to dimethylarginine-containing protein targets. *Mol. Cell* **7**, 1111–1117 (2001).
14. Brahms, H. *et al.* The C-terminal RG dipeptide repeats of the spliceosomal Sm proteins D1 and D3 contain symmetrical dimethylarginines, which form a major B-cell epitope for anti-Sm autoantibodies. *J. Biol. Chem.* **275**, 17122–17129 (2000).
15. Hebert, M.D., Shpargel, K.B., Ospina, J.K., Tucker, K.E. & Matera, A.G. Coilin methylation regulates nuclear body formation. *Dev. Cell* **3**, 329–337 (2002).
16. Boisvert, F.M. *et al.* Symmetrical dimethylarginine methylation is required for the localization of SMN in Cajal bodies and pre-mRNA splicing. *J. Cell Biol.* **159**, 957–969 (2002).
17. Meister, G. & Fischer, U. Assisted RNP assembly: SMN and PRMT5 complexes cooperate in the formation of spliceosomal UsnRNPs. *EMBO J.* **21**, 5853–5863 (2002).
18. Gonsalvez, G.B. *et al.* Two distinct arginine methyltransferases are required for biogenesis of Sm-class ribonucleoproteins. *J. Cell Biol.* **178**, 733–740 (2007).
19. Chuma, S. *et al.* Tdrd1/Mtr-1, a *tudor*-related gene, is essential for male germ-cell differentiation and nuage/germinal granule formation in mice. *Proc. Natl. Acad. Sci. USA* **103**, 15894–15899 (2006).
20. Lane, N. *et al.* Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. *Genesis* **35**, 88–93 (2003).
21. Martin, S.L. & Branciforte, D. Synchronous expression of LINE-1 RNA and protein in mouse embryonal carcinoma cells. *Mol. Cell. Biol.* **13**, 5383–5392 (1993).
22. Aravin, A.A. & Bourc'his, D. Small RNA guides for *de novo* DNA methylation in mammalian germ cells. *Genes Dev.* **22**, 970–975 (2008).
23. Carmell, M.A. *et al.* MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. *Dev. Cell* **12**, 503–514 (2007).
24. Zhao, Q. *et al.* PRMT5-mediated methylation of histone H4R3 recruits DNMT3A, coupling histone and DNA methylation in gene silencing. *Nat. Struct. Mol. Biol.* **16**, 304–311 (2009).
25. Botuyan, M.V. *et al.* Structural basis for the methylation state-specific recognition of histone H4–K20 by 53BP1 and Crb2 in DNA repair. *Cell* **127**, 1361–1373 (2006).
26. Selenko, P. *et al.* SMN tudor domain structure and its interaction with the Sm proteins. *Nat. Struct. Biol.* **8**, 27–31 (2001).
27. Wang, J., Saxe, J.P., Tanaka, T., Chuma, S. & Lin, H. Mili interacts with tudor domain-containing protein 1 in regulating spermatogenesis. *Curr. Biol.* **19**, 640–644 (2009).
28. Kirino, Y. *et al.* Arginine methylation of Piwi proteins catalysed by dPRMT5 is required for Ago3 and Aub stability. *Nat. Cell Biol.* **11**, 652–658 (2009).
29. Anne, J., Ollo, R., Ephrussi, A. & Mechler, B.M. Arginine methyltransferase Capsuleen is essential for methylation of spliceosomal Sm proteins and germ cell formation in *Drosophila*. *Development* **134**, 137–146 (2007).
30. Gonsalvez, G.B., Rajendra, T.K., Tian, L. & Matera, A.G. The Sm-protein methyltransferase, dart5, is essential for germ-cell specification and maintenance. *Curr. Biol.* **16**, 1077–1089 (2006).
31. Lim, A.K. & Kai, T. Unique germ-line organelle, nuage, functions to repress selfish genetic elements in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **104**, 6714–6719 (2007).

ONLINE METHODS

Antibodies and complex purifications. For anti-Miwi polyclonal rabbit antibodies, we produced a GST-fusion of a fragment consisting of amino acids 1–200 of Miwi in *E. coli* for use as an antigen. The anti-Mili monoclonal antibody (13E-3) recognizes the peptide corresponding to amino acids 107–122 of Mili. All antibodies were affinity-purified, and the quality of antibodies was verified by western blotting or immunofluorescence or both (**Supplementary Fig. 1b,c**). Anti-Tdrd1 (ref. 11), anti-L1 ORF1p[22], anti-Sm Y12, which recognizes symmetrical dimethyl arginines[14], and anti-Miwi2 antibodies[6] are described. The anti–β-tubulin (Abcam) and anti-HA (Santa Cruz) antibodies were purchased. All information on plasmid constructs used in this study is available upon request.

We chemically cross-linked anti-Mili antibodies to protein G–Sepharose beads using dimethyl pimelimidate (DMP; Sigma D-8388) and used these to purify Mili complexes from adult mouse testes extracts (50 mM Tris, pH 8, 150 mM NaCl, 5 mM $MgCl_2$, 1mM DTT, 0.5% sodium deoxycholate (Sigma), 1% Triton X-100, 1 tablet of complete protease inhibitor (Roche) per 5 ml, 50 μg ml^{-1} yeast tRNA, 2 mM Vanadyl ribonucleoside complex (Sigma)). After five washes (10 mM Tris, pH 8, 150 mM NaCl, 0.05% (v/v) NP-40), the retained proteins were eluted by 0.1 M glycine, pH 3, neutralized with Tris pH 8, and precipitated before being resolved by 10% SDS-PAGE. After staining with Coomassie blue, we identified all the visible bands by LC/MS/MS (Proteomics Core Facility, EMBL Heidelberg).

When required, we extracted RNA from the immunoprecipitations by proteinase K treatment and phenol extraction. After dephosphorylation by shrimp alkaline phosphatase (Roche) and inactivation by heat treatment, we 5′ end labeled the RNAs with T4 polynucleotide kinase (Fermentas) and [γ-^{32}P]ATP. For western blot analysis of the complex components, the blots were stripped and reprobed successively with respective antibodies.

Mammalian cell transfections and immunoprecipitations. We transiently transfected HEK 293T cells with 3 μg each of the required tagged protein expression constructs using Lipofectamine Plus (Invitrogen). When required, total plasmid amounts used for transfections were kept the same with the use of empty vector. Cells were harvested after 48 h and lysed for immunoprecipitations. For anti-HA immunoprecipitations, we used anti-HA affinity matrix beads (Roche). Where mentioned, transfected cells were treated with methyltransferase inhibitor 5′-deoxy-5′-(methylthio)adenosine (MTA; Sigma) at a final concentration of 750 μM or the solvent DMSO alone. In these experiments, we collected cells after 18 h for analysis.

Mass spectrometry. We digested the protein mixture in solution with trypsin (Roche Applied Science) at 37 °C for at least 4 h. The ratio of trypsin and protein (w/w) was 1:50. The digested proteins were analyzed by an ion-trap MS HCT ultra PTM Discovery System (Bruker Daltonics) coupled with a Nano-LC·2D HPLC system (Eksigent).

We loaded the digest into a CapRod monolithic C18 pre-column 40 mm × 0.1 mm, washed with Phase A (2% acetonitrile and 0.1% formic acid in water) (both from Merck). The peptides were separated by another CapRod monolithic C18 column of 100 μm inner diameter and 15 cm length. The linear elution gradient started at 10% B (99.9% acetonitrile, 0.1% formic acid) and ended at 70% B over a span of 20 min with a constant flow rate of 300nl min^{-1}. We used the data-dependent mode to do MS/MS. Two precursor ions per screen were selected over *m*/*z* range from 400 to 1200. Fragmentation was performed subsequently for 60 ms from 300 to 2000 *m*/*z*. We processed the raw data using DataAnalysis (Version 3.1, Bruker) and submitted the extracted MS/MS data to the MASCOT (version 2.103, Matrix Science) in-house server via Biotools (version3.0, Brucker Daltonics).

We identified the proteins by searching the peptide lists against SwissProt or NCBI databases. The following parameters were used for searching for post-translational modifications: taxonomy mammalia; Enzyme: trypsin; Max Missed Cleavages: 2; Variable modifications: Oxidation (M); dimethylation (R); Peptide Mass Tolerance: ± 0.5 Da: Fragment Mass Tolerance: ± 0.5 Da.

Biotinylated peptide binding assays. Synthetic biotinylated (C-term) peptides corresponding to Mili amino acids 69–80 were synthesized (Peptide Specialty Laboratories) with the Arg74 either unmodified or modified with a symmetrical dimethyl mark (sDMA). Proteins were labeled with ^{35}S-methionine (Perkin Elmer) during *in vitro* translation in rabbit reticulocyte lysates (Promega). We incubated equal amounts of the labeled protein with 20 μg each of the modified or unmodified peptide for 1 h at 4 °C. We recovered bound protein-peptide complexes by addition of 50 μl of Magnetic Streptavidin beads (Promega) per reaction. The beads were pre-blocked with HEK 293T cell lysate to reduce background. Subsequent to 45 min incubation, we washed the beads four times with wash buffer (1× PBS with 0.05% (v/v) NP-40) and the proteins were resolved by 10% SDS-PAGE. Dried gels were exposed to a Phosphor Storage screen and scanned (Typhoon scanner; GE Health).

Mouse strains and purification of germ cells. Testes from mice of different ages were harvested and extracts were normalized by protein estimation. We used equal total protein amounts in the experiment shown in **Figure 3b**. The *Tdrd1* mutant is described[19]. Preparation of purified cell populations by sedimentation and staining procedure was as described[32].

For tissue sections, testes from 18-dpc fetuses were extracted from heterozygous and homozygous animals. Adult testes were isolated from wild-type and *Tdrd1* homozygous mutant mice. Testes were fixed in 2% (v/v) paraformaldehyde in PBS and cryoembedded in OCT compound (Sakura). Staining with anti-Miwi, anti-Miwi2, anti-L1 ORF1p or anti-Mili (used at 1:10) was performed on 10 μm cryosections and detected with Alexa488–conjugated secondary antibodies (Invitrogen). Nuclei were counterstained with a Hoechst dye (Sigma). Images were taken with a fluorescence microscope mounted with a CCD.

Quantitative RT-PCR analysis, Southern and bisulfite methylation analysis. We extracted total RNA from testes isolated from mice of the indicated age with Trizol (Invitrogen). Quantitative real-time PCR primers and analysis were as described[8,20]. We extracted whole-testis DNA from 15-dpp mice and carried our Southern blot analysis with an L1 probe as described[33].

Genomic DNA was isolated from FACS-purified male germ cells (spermatogonia to early spermatocytes) of 15-dpp animals. Bisulfite sequencing of the L1-sequence was as described[21].

Small RNA libraries and bioinformatics. We prepared small RNA libraries with the Digital Gene Expression for Small RNA kit (Illumina) and deep-sequenced (Solexa, Illumina) at EMBL Gene Core facilty, Heidelberg.

The raw sequencing reads from Mili (3 million), Mili (*Tdrd1*$^{-/-}$) (5 million) and Tdrd1 (4.5 million) were mapped to the mouse genome (July 2007, mm9 assembly), and only perfectly mapping reads were considered for further analysis: Mili (2.7 million), Tdrd1 (3.5 million) and Mili (Tdrd1$^{-/-}$) (4 million).

For the library comparison, we counted the number of reads that were identical between both libraries. To account for essentially random processing of primary piRNA transcripts, we also counted the number of reads in each library that seemed to have an identical genomic origin. For this, we considered two reads to be identical when they mapped to the genome with an offset of at most 10 nt (**Supplementary Fig. 3d**). We compared this number to three controls, in which all the genomic coordinates for reads from one of the two libraries was shifted by a constant 1,000 nt (control 1), a random offset between 0 nt and 1,000 nt (control 2), or a random offset between –500 nt and +500 nt (control 3). Controls 2 and 3 are exceedingly stringent as they place the control reads in the vicinity of the actual reads (with the identical position being allowed).

To map the reads to transposon consensus sequences from Repbase[34], we allowed up to three differences or mismatches. Strand orientations (**Supplementary Fig. 3e,f**) were calculated by mapping the repeat reads in the libraries to consensus sequences for different transposon families. Strand orientations for genic reads were calculated by aligning to known or predicted genes in the databases. Functional annotation of sequences into known genic regions, repeat classes and other known small RNAs were done by intersecting the matching positions in the genome with the respective genome annotation from UCSC genome browser[35].

For cluster analysis, the reads in the libraries were used to scan the genome in 2-kb tiles (in 10-nt increments), which we refined to the reads at the window start and end. We selected the top 1% of windows with the highest

doi:10.1038/nsmb.1615

read density and collapsed neighboring windows (within 500 nt) into larger non-overlapping clusters that we again ranked based on their read density (**Supplementary Tables 2** and **3**).

32. Pivot-Pajot, C. *et al*. Acetylation-dependent chromatin reorganization by BRDT, a testis-specific bromodomain-containing protein. *Mol. Cell. Biol.* **23**, 5354–5365 (2003).
33. Soper, S.F. *et al*. Mouse maelstrom, a component of nuage, is essential for spermatogenesis and transposon repression in meiosis. *Dev. Cell* **15**, 285–297 (2008).
34. Jurka, J. Repbase update: a database and an electronic journal of repetitive elements. *Trends Genet.* **16**, 418–420 (2000).
35. Kent, W.J. *et al*. The human genome browser at UCSC. *Genome Res.* **12**, 996–1006 (2002).