Functional Analysis of Tudor-domain-containing proteins in the zebrafish germline

For Mark and my parents

The research described in this thesis was performed at the Hubrecht Institute for Developmental Biology and Stem Cell Research, part of the Royal Acadamy of the Arts and Sciences (KNAW), within the framework of the Graduate School of Cancer, Genomics and Developmental Biology (CGDB) in Utrecht, The Netherlands.

ISBN: 978-90-393-5700-2

Cover: Confocal image of immunostaining against Tdrd6 in 1dpf zebrafish PGCs.

Layout and Illustrations:
Hsin-Yi Huang and Mark van Eekelen

Printing: Off Page, Amsterdam

Functional Analysis of Tudor-domain-containing proteins in the zebrafish germline

Een functionele analyse van Tudor-domein-bevattende eiwitten in zebravis kiemcellen
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof.dr. G.J. van der Zwaan, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op dinsdag 17 januari 2012 des ochtends te 10.30 uur

door

Hsin-Yi Huang

geboren op 15 maart 1981
te Hsin-Chu City, Taiwan R.O.C.

Promotor:
Prof.dr. R.F. Ketting

Table of contents

Chapter 1

Introduction

1

1

Germ cells carry the genetic make-up of the next generation, therefore it is essential to maintain the genome integrity of the germ cells. Several mechanisms are used to protect the genome from internal and external threats, such as free radicals, UV radiation, viral infection, and transposable elements (TEs). One of the defense mechanism against TEs in germ cells is the PIWI-piRNA pathway. In this introduction, I will discuss germ cells, a germ cell-specific structure called germ plasm and the relation between the germ plasm and the PIWI-piRNA pathway in controlling TEs.

Germline specification, germ plasm and different forms of germ plasm

Germline specification and germ plasm

All sexually reproducing animals combine the genetic information from two organisms to create a new one. This genetic information is stored in all cell types but only one specific cell type, the germ cell, can transmit the information to the next generation. In most studied animal species, germ cells are specified and segregated from somatic cell lineages during the initial stages of embryogenesis. There are two main ways to specify germ cells from somatic cells. In *C. elegans* (Strome et al., 1994; Wolf et al., 1983), *Drosophila* (Williamson and Lehmann, 1996), *Xenopus* (Ikenishi, 1998) and zebrafish (*Danio rerio*) (Raz, 2003), the “preformation” mode is used. The germ cell fates in these species are determined through localization of maternally inherited cytoplasmic determinants, called germ plasm. Only the cells inheriting germ plasm become PGCs. Alternatively, in mammals, germ cells are specified later in embryogenesis, from among a population of pluripotent epiblast cells, depending on inter-cellular inductions from surrounding somatic cells during gastrulation (Hayashi et al., 2002; McLaren, 2003). This is called the “inductive” mode of germ cell specification. Germ plasm components from species that use the preformation mode, such as Vasa and PIWI proteins, are not detectable in newly specified PGCs and do not appear until the PGCs enter the genital ridge (de Sousa Lopes and Roelen, 2010). The inductive mode is regulated by the bone morphogenetic protein (BMP) signaling pathway (de Sousa Lopes et al., 2004; Hayashi et al., 2002).

Germ granules/Nuage

Germ plasm is characterized by special electron-dense cytoplasmic organelles called germ granules or nuage. Different terms are used to describe germ granules in different species. They are called P granules in *C. elegans* and nematodes; polar granules in *Drosophila* and insects; and nuage in zebrafish. Here I use the term nuage, since zebrafish is the model organism used in this thesis. Nuage was first identified in electron micrographs as an

amorphous electron–dense cloud, which appears as discrete cytoplasmic inclusions, surrounding the nurse cell nuclei in *Drosophila* (Allis et al., 1979; Mahowald, 1971). It is often found associated with the nuclear membrane, in close proximity to nuclear pores and mitochondria. The association of nuage with nuclear pores is proposed to have a role in receiving RNA molecules from the nucleus. The nuage is strongly associated with germline fate in all organisms, including the species that specify germ cells by inductive signals. It has been observed not only in PGCs, but also in oogonia, oocytes, spermatogonia, spermatocytes and spermatids, and has been documented in at least eight animal phyla (Eddy, 1974), which suggests that it has an important and conserved role in germ line development. The presence of nuage is obvious in early embryogenesis of worms, flies, frogs and zebrafish. In mammals, fetal nuage-like structures have been described in mouse gonocytes (Aravin et al., 2009), and the nuage becomes more obvious at later stages of germ cell differentiation, such as the intermitochondrial cement (IMC) in the spermatogonia, early spermatocytes and oocytes; the chromatoid body (CB) in the round spermatids (Chuma et al., 2009). Nuage is enriched in RNA and RNA binding proteins, such as Piwi proteins, RNA helicases and nuclease proteins. These proteins and RNAs are either specifically localized to nuage or protected from degradation in the nuage. Accumulating evidence supports the role for nuage in RNA metabolism and storage (Chuma et al., 2009).

In zebrafish, primordial germ cells (PGCs) are determined by germ plasm, which is formed in the distal ends of the cleavage furrows in the 4-cell stage embryos. Germ plasm contains mRNAs from germline-specific genes, such as *vasa*, *nanos*, *dazl*, as well as some maternally provided piRNA pathway components, like Ziwi (Houwing et al., 2007) and Tdrd6 (this thesis). At the 32-cell stage, the four aggregates of the germ plasm segregate to four cells referred to as promptive primordial germ cells (pPGCs) (Braat et al., 1999). The PGC number during the blastula stages has been estimated to be five (Lin et al., 1992), implying PGC number does not rise substantially through the cell division of early embryos until the 4K-cell stage. From the dome stage, PGCs start to migrate to the future gonad area while proliferating at the same time. At 24hpf, the migration of PGCs is complete and there are between 25 to 50 PGCs at the future gonad area ((Raz, 2003) and Fig.1, schematic picture). Nuage structure is observed as a dynamic subcellular structure during PGC migration (Brangwynne et al., 2009; Strasser et al., 2008) and can be clearly seen in 24hpf zebrafish PGCs. It forms massive aggregates and often resides at the place where the nuclear envelope is engulfed (curved toward the nucleus) (Fig.2A). We speculate that this shape of the nuclear envelope is caused by clustering of nuclear pores to adapt to the presence of nuage, which may have a function in receiving mRNA molecules after they are transcribed in the nucleus. As the

development of PGCs proceeds, nuage structure decreases in size. From 3 wkpf to 5 wkpf, using immunohistochemistry against Zili, Tdrd1 or Tdrd6, three nuage components, nuage can be observed as distinct tiny circular structures surrounding the nucleus of juvenile oocytes (Fig.2B). These circular structures are similar to Medaka Tdrd1 described in the developing gonad of both male and female germ cells (Aoki et al., 2008). It is likely that there are sub-structural compartments for different nuage components. Immuno-electron-microscopic analysis of the localization of different nuage components will be able to give further insight.

In adulthood, granular nuage is only obvious in the early stages of male and female germ cells, such as oogonia and stage I – II oocytes in the female gonads (Fig.2C) and spermatogonia and early spermatocytes in the male gonads (Fig.2D). During spermatogenesis in mice, nuage is first observed in spermatogonia, followed by the appearance of IMC in spermatocytes in meiotic prophase (Yokota, 2008). Following the meiotic divisions, a prominent CB represents a special form of nuage at the surface of the haploid nucleus of round spermatids (Parvinen and Jokelainen, 1974). A structure similar to the CB has not yet been identified during zebrafish spermatogenesis.

PIWI-piRNA pathway components are enriched in the nuage structures of many species. For instance, in *Drosophila*, germinal granules are rich in PIWI pathway proteins, Aub and Ago3 (Brennecke et al., 2007); in mice, IMC is abundant in MILI, TDRD1, TDRD6 and GASZ; CB is rich in MILI, TDRD1, TDRD9 and MVH (Mammalian Vasa Homologue), which are essential components for piRNA biogenesis (Chuma et al., 2006; Hosokawa et al., 2007). Loss of Tdrd1 leads to less electron-density or absence of nuage structures in both mice and zebrafish (Chuma et al., 2006; Huang et al., 2011b). Nuage components also fail to localize properly in *mili* mutants (Aravin et al., 2009). This suggests that nuage are the subcellular location for the PIWI-piRNA machinery and that the pathway is generally required for nuage formation. The summaries of the expression time points of each PIWI-piRNA pathway protein that have been studied in zebrafish are illustrated in Fig.3.

Balbiani body

The embryonic germ plasm originates from the oocytes, where different components required for germ plasm formation aggregate into a large electron-dense structure, called the Balbiani body. The Balbiani body is another form of the germ plasm in the female germline, however it has a structure distinct from that of nuage. It is a subcellular cluster of organelles, such as mitochondria, ER and Golgi, and proteins/RNAs in the oocytes. In some species, such as *Xenopus* and zebrafish, nuage are also gathered in the Balbiani body (Kloc et al., 2004). It has been observed that the nuage

is part of the Balbiani bodies in zebrafish (Marlow and Mullins, 2008). In mice, Balbiani body was only recently observed in a short period of early oogenesis, from just before primordial follicle formation to briefly in young primordial follicles (Pepling et al., 2007). So far, there is no clear association between nuage and Balbiani body found in mice. PIWI proteins were not detected in the Balbiani bodies. Similarly, the Balbiani body component, TRAL, is absent from the nuage (Chuma et al., 2009; Pepling et al., 2007).

Although the functions of the Balbiani body remain elusive, it has been speculated to be involved in the selection of healthy mitochondria and the transport of specific mRNAs in the oocytes (Kloc et al., 2004). The quality of mitochondria in the oocytes was recently reported to influence the success rate of fertilization (Bentov et al., 2011). In zebrafish and *Xenopus* oocytes, it has been shown that the Balbiani body functions in transporting germ plasm mRNAs (Kosaka et al., 2007). Balbiani bodies can be first observed in the oocytes of 5-week-post-fertilization zebrafish larvae, where they locate to one side of the ooplasm and associate with the nucleus in early stage oocytes. As the oocyte develops, the Balbiani body moves toward the vegetal cortex of the oocytes and dissociates into small islands (Fig.3B, schematic pictures for Tdrd6a in oogenesis). Interestingly, we notice that nuage and the Balbiani body share some common components, such as Vasa, Ziwi (Houwing et al., 2007) and Tdrd6a (this thesis, Chapter 3) that are known to be involved in the PIWI-piRNA pathway. However, the link between the nuage and the Balbiani body in zebrafish remains enigmatic and awaits further investigation.

Transposable elements (TEs)

TEs are repeated DNA sequences that are able to move and replicate within genomes. Various classes of TEs comprise a large fraction of most eukaryotic genomes, such as 45% in the human genome and about 20% in the *Drosophila* genome (Lander et al., 2001; Petrov et al., 2011). Besides being considered as “selfish” and “parasitic”, TEs also play a role in driving genome evolution: Cryptic TEs, which are epigenetically silenced for a long time and have thus lost the ability to mobilize, are important structural components of the centromeres and telomeres of difference species (Slotkin and Martienssen, 2007), and LINE elements have been reported to play an important role in X chromosome inactivation in mammals by promoting heterochromatin formation (Chow et al., 2010). Nevertheless, they are highly mutagenic and can alter the genome in many different ways, such as insertion into protein-coding genes. which causes chromosome breakage and genome rearrangement. Therefore tight control of TEs is essential to maintain genome integrity.

TEs can be broadly divided into two groups: retro-transposons (class I) and DNA transposons (class II). Class I TEs require a reverse-transcription step and move in the genome via an RNA intermediate. They undergo

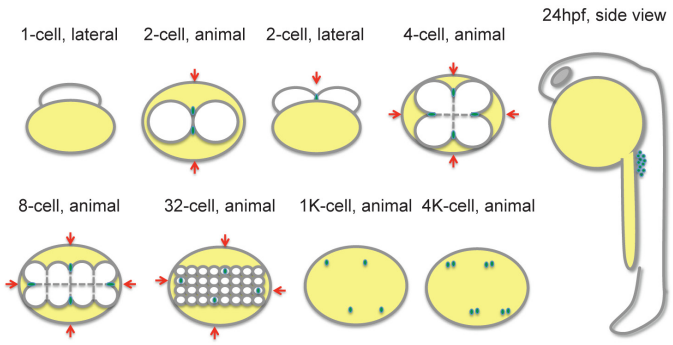


Figure 1. Zebrafish PGC specification during early embryogenesis. Zebrafish germ cells are defined by the presence of the germ plasm. Here *vasa* RNA (green patches) is a marker of germ plasm. *Vasa* RNA can be detected at the cleavage planes at 2- to 8-cell stages (indicated with red arrows). At 32-cell stage, *vasa* RNA aggregates into four cells, defining the pPGCs. The germ cell number does not increase until 4K-cell stage. Later, PGCs start proliferation while migrating to the future gonad area. At 24hpf, most PGCs localize to the beginning of the yolk extension where the gonad develops (Adapted from Raz, 2003). White area indicates the embryo and the yellow area indicates the yolk.

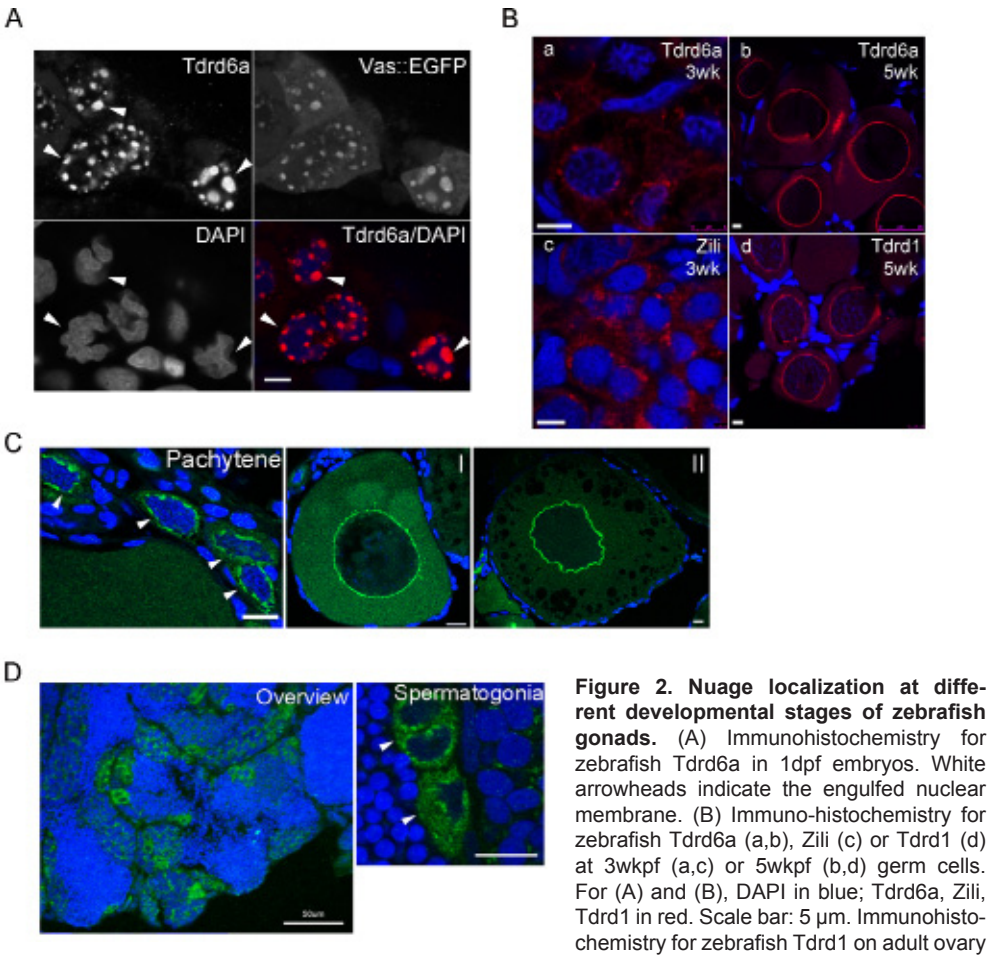


Figure 2. Nuage localization at different developmental stages of zebrafish gonads. (A) Immunohistochemistry for zebrafish Tdrd6a in 1dpf embryos. White arrowheads indicate the engulfed nuclear membrane. (B) Immuno-histochemistry for zebrafish Tdrd6a (a,b), Zili (c) or Tdrd1 (d) at 3wkpf (a,c) or 5wkpf (b,d) germ cells. For (A) and (B), DAPI in blue; Tdrd6a, Zili, Tdrd1 in red. Scale bar: 5 µm. Immunohistochemistry for zebrafish Tdrd1 on adult ovary (C) and testis (D). Pachytene, I, II: stages of oocytes. Tdrd1 in green, DAPI in blue. Scale bar: 10 µm, otherwise indicated in the figure.

duplicative transposition, which means that after each transposition the total number of the elements increases. Therefore class I TEs are likely to enlarge the genome size. Class I TEs can be further divided into two types, LTR and non-LTR retro-transposons, depending on the presence or absence of direct long terminal repeats (LTR) at the end of the element. LTR retrotransposons encode Gag and Pol proteins, which form virus-like particles, where reverse-transcription of the mRNAs and production of the DNA intermediates take place. LTR retrotransposons also encode Integrase proteins that further integrate the DNA intermediates into the genome. Non-LTR retrotransposons can be further divided into two sub-groups: long interspersed nuclear elements (LINEs) and short interspersed nuclear elements (SINEs). LINEs encode ORF1 and ORF2 proteins, which contain the reverse transcriptase domain, while SINEs do not encode reverse-transcriptases and use enzymes from LINEs or other transposons for their transposition (Siomi et al., 2011; Slotkin and Martienssen, 2007).

The vast majority of class II TEs mobilize through a cut-and-paste mechanism. DNA transposons encode transposase proteins that recognize the inverted/directed repeat sequence at the end of a DNA transposon and excise the element out of the donor position and move it into a new acceptor position. Some transposases recognize specific target sequences and some transposases can randomly insert DNA transposons into a new site in the genome. The DNA sequence is cut by the transposases in a staggered manner. After the DNA transposon is inserted into the acceptor site, the host DNA repair machinery repairs the single-stranded gaps in the acceptor site as well as the double-stranded breaks at the donor site (Miskey et al., 2005).

A list of zebrafish transposons that are relevant for this thesis is presented in Table1.

Guardian of the germline genome: PIWI-piRNA pathway

The only function of the germ line is to transmit genetic information between generations, whereas the soma mediates germline transmission by ensuring reproductive fitness of the organism. Therefore, maintaining the genome fitness of both the germline and the soma is crucial. TE activity threatens the genome integrity of both germ cells and somatic cells. To control selfish TEs and viruses, organisms have developed several defense systems to protect their genome via Argonaute family proteins. In animals, the Argonaute superfamily is divided into two clades, AGO and PIWI (P-element-induced wimpy testis) (Farazi et al., 2008) (Fig.4), based on their sequence and expression pattern. Argonaute family proteins function together with small RNAs (sRNAs), mostly ~20-30 nucleotide in length, and together form so called RNA-induced silencing complexes (RISCs). Three characteristic domains in Argonaute proteins are important for RISC function: the PAZ domain binds to the 3' end of the sRNA; the MID domain forms a binding

pocket for the 5' phosphate group of the sRNA; the PIWI domain is an RNase H-like domain harboring conserved residues (DDH) for the endonucleolytic activity (from now on referred to as Slicer activity). Argonaute proteins possess a DDH motif and have the potential to perform endonucleolytic cleavage of its target at the position opposite to the 10th and the 11th nucleotide of the guiding sRNA.

Silencing sRNAs in animals can be divided into three classes: microRNAs (miRNAs), small interfering RNAs (siRNAs) and PIWI-interacting RNAs (piRNAs). These sRNAs possess a 5' phosphate group and 3' hydroxyl termini, which in some cases are modified by, such as, methylation. Different sRNA groups often associate with specific members of the Argonaute family (Czech and Hannon, 2011; Siomi et al., 2011). AGO proteins are mostly ubiquitously expressed throughout development and are associated with miRNAs and/or siRNAs. miRNA-guided RISCs (miRISCs) are important in gene regulation during development, while siRNA-guided RISCs (siRISCs) control the exogenous viral infection and endogenous TE activities. siRISCs regulate TE by degrading TE transcripts (post-transcriptional silencing) or promoting heterochromatin formation at TE genomic regions (transcriptional silencing) (Obbard et al., 2009). Mutations in *ago* genes lead to developmental defects; for instance, *ago2*-null mice shows embryonic lethality (Liu et al., 2004).

PIWI proteins use piRNAs as a guide for substrate recognition and target cleavage. PIWI and piRNA form the piRNA-induced silencing complex (piRISC) that is known to suppress the activity of TEs and maintain genome integrity of the germline (Aravin et al., 2007). PIWI as well as piRNA expression are mainly restricted to germline cells. Contrary to *ago* mutants, *piwi* mutants show defects in gametogenesis but exhibit normal development. Loss of PIWI proteins causes infertility in *Drosophila*, mice and zebrafish. Mutant animals also show de-repression of transposon activities, which is thought to be the reason for the defect in gametogenesis (Carmell et al., 2007; Houwing et al., 2008; Houwing et al., 2007; Kuramochi-Miyagawa et al., 2008; Li et al., 2009; Vagin et al., 2006). Deletion of piRNA clusters (the genomic source of piRNAs, explained below), such as the *flamenco* locus in *Drosophila*, also results in de-repression of particular transposons (Desset et al., 2003; Mevel-Ninio et al., 2007; Prud'homme et al., 1995). These studies show that both Piwi proteins and piRNAs are required for controlling the activity of TEs.

In most organisms, multiple PIWI paralogs are present. The ones that have been more intensively studied up to now are Aub/AGO3/Piwi in *Drosophila*, MIWI/MIWI2/MILI in mice, Ziwi/Zili in zebrafish and SMEDWI-1/SMEDWI-2/SMEDWI-3 in planarian. These Piwi proteins are speculated to have Slicer activity, which involves the silencing of transposon mRNAs via cleavage and is also essential for the biogenesis of piRNAs. Below, I will compare different aspects of the PIWI-piRNA pathway among various

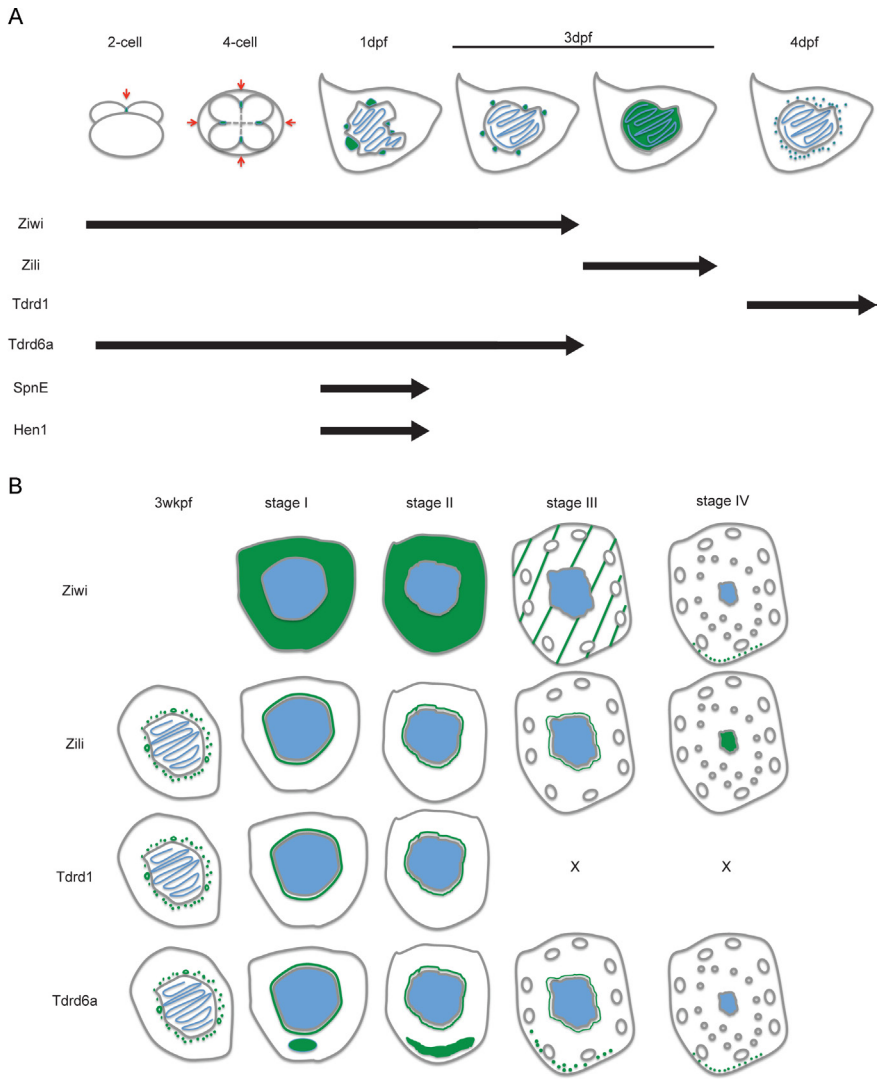


Figure 3. Schematic expression patterns of known zebrafish PIWI-piRNA pathway components. (A) Expression patterns of Ziwi, Zili, Tdrd1, Tdrd6a, SpnE and Hen1 during early embryogenesis. Ziwi and Tdrd6a are maternally provided and can be detected from 2-cell stage at the cleavage plane. At 1dpf, both Ziwi and Tdrd6a display prominent granular structures (big nuage) around the nucleus. Small nuage structures are also seen. At later PGC stages, both Ziwi and Tdrd6 are seen as perinuclear granules with more homogeneous size distribution. Zili can be detected in the nucleus of 3dpf PGCs and the nuclear localization remains till 7dpf. Tdrd1 can be detected from 4dpf as fine perinuclear granules. Whole embryo lateral and dorsal view at the 2- and 4-cell stage, respectively. Single PGC is shown for 1dpf to 4dpf. (B) Expression patterns of Ziwi, Zili, Tdrd1 and Tdrd6a in oogenesis. Ziwi has dispersed cytoplasmic expression during oogenesis. At stage III oocytes, Ziwi signal is diluted by the size of the cell (shown as green lines). At stage IV, Ziwi is detectable as small granules along the cortex of the oocyte. Zili, Tdrd1 and Tdrd6 have similar perinuclear expression patterns: at 3wkpf, they are seen as fine granules or circular structures around the nucleus (show as green dots and circles). During the development of the oocytes, Zili, Tdrd1 and Tdrd6 show fine granule structures, which form a rim around the nucleus. At stage IV oocyte, Zili translocates into the nucleus, while Tdrd1 and Tdrd6 are not detectable at perinuclear area. Tdrd6 is also localized at a distinct structure, the Balbiani body. Balbiani body localization of Tdrd6 can

be detected through out the oogenesis. Each stage is represented by a single germ cell. x indicates no signal detectable. For both (A) and (B), green shows the localization of each protein, blue indicate DAPI positive area (DNA or nuclei).

species, of which *Drosophila* is the best characterized.

PIWI-piRNA pathway in Drosophila

In *Drosophila*, the three PIWI proteins show various sub-cellular localization. Piwi is localized in the nuclei of both germline cells and somatic follicle cells in the ovaries, but only in the somatic niche cells that maintain the germ stem cells in the testes. Aub and AGO3 are detected mainly in the nuage, where piRNA production takes place, in both male and female germ cells, but are absent from gonadal somatic cells (Saito and Siomi, 2010). In addition to the localization differences, different PIWI proteins also associate with distinct groups of piRNAs.

Characteristics and Biogenesis of piRNAs

The genomic loci of piRNAs are conserved but their sequences are poorly conserved even between close species (Aravin et al., 2006). Large portions of the piRNAs have only been cloned once showing the high complexity of the piRNA population (Siomi et al., 2011). However, they map to relatively restricted genomic regions, the piRNA clusters, which can range from several to hundreds of kilobases in length. piRNA clusters are characterized by a high density of transposon fragments and a dearth of protein coding genes, and are located in pericentromeric and subtelomeric heterochromatin regions (Aravin et al., 2006; Brennecke et al., 2007; Girard et al., 2006). Most of the transposon contents in the piRNA clusters are in the form of nested, truncated or damaged copies that are not likely to express or mobilize. Nevertheless, there are some piRNAs mapped to protein-coding genes, such as suppressor of *stellate [su(ste)]* and *traffic jam (tj)*, which down-regulate the protein coding *stellate (ste)* and *fasciclin III (fasIII)* genes, respectively (Nishida et al., 2007; Robine et al., 2009; Saito et al., 2009; Vagin et al., 2006).

piRNA clusters in *Drosophila* can be divided into two subgroups: dual-strand clusters and uni-strand clusters. The dual-strand clusters produce piRNAs from both genomic strands, such as *42AB* on *Drosophila* chromosome 2 (Klattenhoff et al., 2009). On the other hand, uni-strand clusters, such as the *flamenco* locus on *Drosophila* chromosome X, are only transcribed form one direction. piRNAs derived from uni-strand clusters are usually anti-sense to the TE mRNAs (Brennecke et al., 2007; Malone et al., 2009). Based on the study in *Drosophila* showing that P element insertion at the 5' end of *flamenco* disrupts the piRNA production from this cluster. This suggests that piRNAs are derived from long single-stranded precursor RNA transcripts. Therefore unlike miRNAs and siRNAs, piRNAs can be generated

in a Dicer-independent manner (Houwing et al., 2007; Vagin et al., 2006).

In general, piRNAs differ in many respects from siRNAs and miRNAs. Mature piRNAs are 26-30 nucleotides (nt) long, while miRNAs and siRNAs are between 20-24 nt. piRNAs have 2' methoxy, 3' hydroxy termini, which stabilize piRNAs in vivo by limiting further 3' end processing (Houwing et al., 2007; Kirino and Mourelatos, 2007; Ohara et al., 2007; Vagin et al., 2006). This methylation modification, which is absent in the animal siRNAs and miRNAs, is done by the methyltransferase Hen1. In *Drosophila*, it has been reported that RNA methyltransferase *Dm*Hen1 (Pimet) carries out the methylation reaction. However, *hen1* mutants are viable and fertile, suggesting that this modification is not essential for piRNA function (Horwich et al., 2007; Saito et al., 2007). piRNAs possess neither specific sequence motifs nor significant secondary structures, however, many piRNAs have either a uridine residue at the 1st position (1U) or an adenine residue at the 10th position (10A) counting from their 5' ends (Aravin et al., 2006; Brennecke et al., 2007; Girard et al., 2006; Gunawardane et al., 2007; Houwing et al., 2008; Houwing et al., 2007).

Besides the source of genomic encoded piRNAs, it was found in *Drosophila* by a phenomenon called hybrid dysgenesis that some piRNAs are maternally inherited (Brennecke et al., 2008). Hybrid dysgenesis happens only when one strain of males carrying certain transposons crosses with a different strain of naïve females that do not have the elements, and leads to sterility of the progeny, while the reciprocal cross does not have this effect. From small RNA analysis in the ovaries and early embryos with yet active zygotic genome, Brennecke et al (2008) found the piRNA profiles are similar between the mother and the offspring. Taken together, it suggests that maternally inherited piRNAs are the factors that protect the progenies from the paternal transposons. Further more, the population of these maternal piRNAs corresponds to the maternally deposited PIWI proteins (Brennecke et al., 2008).

Bioinformatic analyses of piRNAs that associate with PIWI proteins in *Drosophila* have led to two models of piRNA biogenesis: the primary processing pathway (Fig.5A) and the secondary amplification pathway, also called the "Ping-Pong" cycle (Brennecke et al., 2007; Gunawardane et al., 2007) (Fig.5B). The primary pathway and the maternally inherited piRISCs provide an initial pool of piRNAs to target various TEs. Then the secondary pathway further shapes the piRNA populations according to the abundance of the active TEs.

It is noted that different PIWI proteins have distinct length preferences for their bound piRNAs: ~23, ~24, ~25 for Ago3, Aub and PIWI, respectively (Siomi et al., 2011). This suggests that PIWI proteins may function as rulers to define the size of the mature piRNAs. In addition to the length preference, piRNAs of different PIWI proteins also display diverse properties, such as

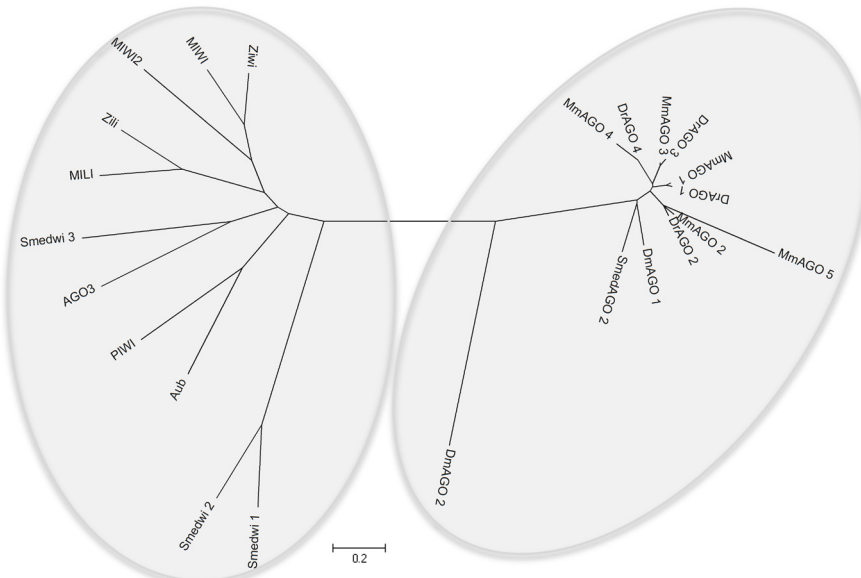


Figure 4. Phylogenetic tree of Argonaute protein family. AGO and PIWI clade of proteins from mouse, zebrafish, *Drosophila* and planarian, *Schmidtea mediterranea*, are aligned in the tree. Phylogenetic analyses were conducted in MEGA4 (Tamura et al., 2007).

strand bias and sequence motif preference: Aub and Piwi preferentially bind to piRNAs with 1U preference that are anti-sense to transposon mRNAs, while AGO3 binds to sense-oriented piRNAs with 10A bias. These properties have led to the model of secondary amplification of piRNAs.

The secondary pathway is a Slicer-dependent, feed-forward amplification cycle, which was first proposed in *Drosophila* by Brennecke et al. (2007) and Gunawardane et al. (2007), to produce piRNAs that target active TEs. A large fraction of *Drosophila* piRNAs are generated by the secondary pathway. Maternally inherited piRISCs and primary piRNAs are the initial driving force of the amplification loop (Brennecke et al., 2008). The initiation of the amplification cycle requires the presence of target transcripts that are complementary to the primary piRNAs. These target transcripts can be mRNAs of active TEs or transcripts of a dual-strand piRNA cluster, such as the *42AB* locus. A primary piRNA, mostly anti-sense to the TE transcript, is loaded onto Aub, which possesses Slicer activity to cleave the complementary transcript 10 nucleotides away from the 5' end of the primary piRNA. This event produces the 5' end of a future-secondary piRNA, which has a 10A bias from its 5' end to pair with the 1U at the primary piRNA and is sense-oriented to the TE transcript. It also results in 10 nucleotides of complementary sequences between the primary and secondary piRNAs. The future-secondary piRNA can then be loaded onto AGO3, where it is processed by the following events to become a mature piRNA: it is trimmed at the 3' end

to a specific length by unknown nucleases and then 2'O methylated by Hen1 to protect the piRNAs from further 3' end processing (Horwich et al., 2007). The mature secondary piRNA further guides AGO3 to target and cleave a complementary sequence, further producing a new piRNA that reflects the original primary piRNA with the 1U preference. The cycle can then be repeated with the loading of a new primary piRN onto Aub. Therefore the amplification cycle is dubbed as "Ping-Pong." The 1U at the primary piRNA, 10A at the secondary piRNA and the 10 nucleotide overlapping between the primary and secondary piRNAs are considered together as the "Ping-Pong signature," which defines whether certain piRNAs are produced via the Ping-Pong cycle (Brennecke et al., 2007; Gunawardane et al., 2007).

piRNAs associated with AGO3 and Aub possess a strong Ping-Pong signature, however, Piwi bound piRNAs display a weak signature. This suggests that Piwi is probably not involved in the amplification cycle. Later it was discovered that Piwi bound piRNAs are produced via the primary processing pathway from uni-strand clusters in the germline and ovarian follicle cells (Malone et al., 2009). The two key players, Aub and AGO3, are only expressed in the germline, thus implying that the Ping-Pong cycle is a germline-specific event (Brennecke et al., 2007) (Fig.5A-B).

The primary processing pathway is involved in the cleavage of long single-stranded precursor RNAs from piRNA clusters, such as the *flamenco* locus, to generate new primary piRNAs. In the somatic follicle cells of the *Drosophila* ovary, where only Piwi but not Aub or AGO3 are expressed, long piRNA precursors are processed into piRNA intermediates via a yet-to-be determined pathway. The piRNA intermediates are further processed into mature primary piRNAs in a specific cytoplasmic granule named the Yb-body. FS(1)YB/Yb is the core component of Yb-bodies and together with ARMI and Piwi is essential for the processing of piRNA intermediates (Haase et al., 2010; Li et al., 2009; Malone et al., 2009; Olivieri et al., 2010; Qi et al., 2011; Saito et al., 2010). Another factor involved in the primary process is ZUC/MITOPLD, a phospholipase or putative nuclease localizing to the outer mitochondrial membrane. It is essential for the maturation of primary piRNAs (Huang et al., 2011a; Olivieri et al., 2010; Saito et al., 2010; Watanabe et al., 2011). Mature primary piRNA then can be loaded onto Piwi, a step that is important for the nuclear localization of Piwi (Saito et al., 2010). The 5' end of the primary piRNA has the 1U bias. This bias can result from preferential cleavage at the uridine residues on a long precursor RNA. It is also possible that the cleavage is totally unbiased, only the PIWI proteins in the pathway selectively associate with piRNAs with 1U at their 5' end.

PIWI-piRNA pathway and heterochromatin

Since Piwi protein displays nuclear localization and the piRNA

clusters are localized in heterochromatic regions, much effort has been made to understand the relationship between heterochromatin and the PIWI-piRNA pathway. Heterochromatin protein 1, HP1, is a conserved component of heterochromatin. Both yeast-two-hybrid analysis and immunoprecipitation assays using *Drosophila* embryonic nuclear lysate have shown that Piwi interacts with *Drosophila* HP1 homolog, HP1a/Rhino (Brower-Toland et al., 2007). Later, Rhino was reported to promote production of piRNAs from the germline dual-strand clusters, but not the uni-strand clusters such as the *flamenco* locus, by association with the dual-strand cluster regions. Loss of Rhino disrupts nuage structure and Aub as well as AGO3 localization to the nuage. *rhi* mutants also show significant reduction of putative RNA precursors, suggesting Rhino promotes the transcription of the dual-strand cluster piRNAs that further drive the Ping-Pong amplification cycle (Klattenhoff et al., 2009). Recently, it was shown that heterochromatin formation is required for piRNA production. The methyltransferase dSETDB1 deposits H3K9me3 heterochromatic marks on the piRNA clusters of the germline and somatic tissue of the gonads. These repressive marks can be recognized by Rhino and are essential for piRNA transcription of both dual-strand and uni-strand clusters (Rangan et al., 2011). Another report showed that heterochromatin formation at the somatic piRNA clusters is independent of PIWI-piRNA pathway. In this report, increased HP1a chromatin association at piRNA clusters was described in Piwi-pathway mutants (Moshkovich and Lei, 2010). However, this study used *Drosophila* heads as a source of material and it is still not convincingly shown that the PIWI-piRNA pathway is truly active in that tissue. Furthermore, depletion of Piwi in ovarian somatic cells (OSCs), where Piwi produces piRNAs from the uni-stranded *flamenco* cluster via the primary processing pathway does not affect HP1a chromatin association (Saito et al., 2010). It suggests that Piwi does not direct HP1a recruitment in the (gonadal) somatic cells. Taken together these results suggest that, although Rhino/HP1a can directly interact with Piwi, there are likely alternative, dSETDB1-dependent, Rhino/HP1a-independent mechanisms. These mechanisms can recruit heterochromatin at the gonadal somatic piRNA clusters and uni-strand clusters in the germline, to promote transcription of piRNA precursor transcripts.

PIWI-piRNA pathway in mice

In mice, MILI expression starts in PGCs at E12.5 and continues until the round spermatid stage (Kuramochi-Miyagawa et al., 2004). MILI is present in the nuage in both male and female PGCs, and also in the nuage in oocytes and CBs in round spermatids (Aravin et al., 2008). MIWI2 is transiently expressed from E15.5 and declined soon after birth in male but not female embryonic gonocytes. MIWI2 is located both in the cytoplasm and the nucleus (Aravin et al., 2008; Kuramochi-Miyagawa et al., 2008).

Table1. Transposons in the zebrafish genome

Class I transposons				Class II transposons	
Types	**Superfamily**	**Family**	**Rank**	**Family**	**Rank**
LTR		Gypsy	1	hAT	5
		DIRS	2	DNA	7
		ERV1	3	Enspm	9
		BEL	8	Polinton	10
		Copia	12	Mariner/Tc1	11
Non-LTR	LINE	CR1	4	Harbinger	13
	LINE	L1	6	TDR	14
	LINE	I	17	Helitron	15
	LINE	RTE	20	Kolobok	16
	SINE		18	Looper	19
				PiggyBac	21
				ISL2EU	22
				MuDR	23

Table 1. Zebrafish transposons. A list of zebrafish transposons that are targeted by piRNAs. Categorized according to their types/families. The rank refelcets the abundance of the total piRNAs mapped to these transposon families in the small RNA library made from wild-type adult ovary with Tupfel long fin (TL) background (also see Chapter 2). Rank 1: highest piRNA abundance, rank 23: lowest piRNA abundance.

Even though both MILI and MIWI2 are located in the cytoplasm, it is often found that they are at distinct nuage-like foci. Upon associating with TDRD1, MILI is located in smaller granular structures, called “pi-bodies,” with MVH and GASZ. Cytoplasmic MIWI2 is found in company with TDRD9 and MAEL in larger granular structures, which are often in close proximity to the MILI granules (Aravin et al., 2008). MIWI2 granules contain known components of the processing body (p-body), such as GW182, DDX6, DCP1a, in the somatic cells, therefore they are referred as “piP-bodies.” It was further found that the pi-body is essential for piP-body formation (Aravin et al., 2009). MIWI is only expressed between the pachytene stage of meiosis and the haploid round spermatid stage. It is localized in the cytoplasm of spermatocytes and CBs in the round spermatids (Deng and Lin, 2002; Kuramochi-Miyagawa et al., 2001). Similar to the case in *Drosophila*, the three different mouse PIWIs also exhibit size preferences for bound piRNAs: ~26, ~28, ~30 nt for MILI, MIWI2 and MIWI, respectively.

In mammals, piRNAs can be divided into two populations: the pre-pachytene piRNAs and the pachytene piRNAs. Pachytene piRNAs were the first piRNAs identified (Aravin et al., 2006; Girard et al., 2006). They are highly expressed in pachytene stage spermatocytes and gradually decline until the haploid round spermatid stage. Pachytene piRNAs interact with MILI and MIWI. Most of the pachytene piRNAs originate from uni-strand cluster regions. At some pachytene piRNA loci a bi-directional promoter region

has been suggested to drive expression of long RNA precursors (Fig.5C). Pachytene piRNAs are relatively low in repeat sequences suggesting they do not function in TE suppression. MIWI-null mutants display spermatogenic arrest at the beginning of round spermatids stage (Deng and Lin, 2002), implying MIWI and associated pachytene piRNAs may play a role in meiosis. However, the actual function of pachytene piRNAs is still unclear (Aravin et al., 2006; Girard et al., 2006).

Pre-pachytene piRNAs are present in germ cells before meiosis. They are associated with MILI and MIWI2. Pre-pachytene piRNAs originate from different genomic loci compared to pachytene piRNAs (Aravin et al., 2007). Most of the pre-pachytene piRNAs are made from piRNA clusters similar to the ones in *Drosophila*, containing multi-typed-repeat-derived sequences. Pre-pachytene piRNAs also map to protein coding genes (28%), mostly in the 3' untranslated regions (3'UTRs) and in those cases they originate from the sense transcript. Therefore, these pre-pachytene piRNAs are not likely to guide post-transcriptional gene silencing (Aravin et al., 2007). However, loss of MILI and MIWI2 lead to failure of *de novo* methylation of TEs in mouse germline and consequent TE upregulation. The effect on methylation status of TEs is likely via translocation of piRNA-loaded MIWI2 into the nucleus (Aravin et al., 2008).

In mice, MIWI2 and MILI are the main players in the Ping-Pong cycle, while MIWI is not involved. Contrary to the orientation of piRNAs in *Drosophila*, mouse primary piRNAs are mostly sense while secondary piRNAs are anti-sense-oriented. MILI preferentially binds primary piRNAs, which display a strong bias for a U at position one (1U). MIWI2-associated piRNAs are secondary piRNAs. Consequently, they are anti-sense-oriented and display a bias for adenosine at position ten (10A). In the absence of MILI, mice fail to generate primary piRNAs and MIWI2 remains unloaded (Aravin et al., 2008) (Fig.5C). However, it is possible that MILI alone is able to conduct amplification cycle since MILI can accept both primary and secondary piRNAs at certain periods of the pre-pachytene stage (Aravin et al., 2007).

PIWI-piRNA pathway in zebrafish

During early embryogenesis in zebrafish, Ziwi is detected in the nuage and as embryogenesis continues, Ziwi gradually disperses into the cytoplasm (Houwing et al., 2007). Zili is detected in the nucleus between 3dpf to 7dpf of embryogenesis. Later during development, Zili moves into perinuclear nuage until the stage of the maturing oocyte, where Zili translocates back into the nucleus (Houwing et al., 2008). piRNA clusters in zebrafish, similar to in *Drosophila*, are largely mapped to repeat sequences. However, the transcription of piRNAs often only happens from one genomic strand within a given region of the piRNA cluster (Houwing et al., 2007).

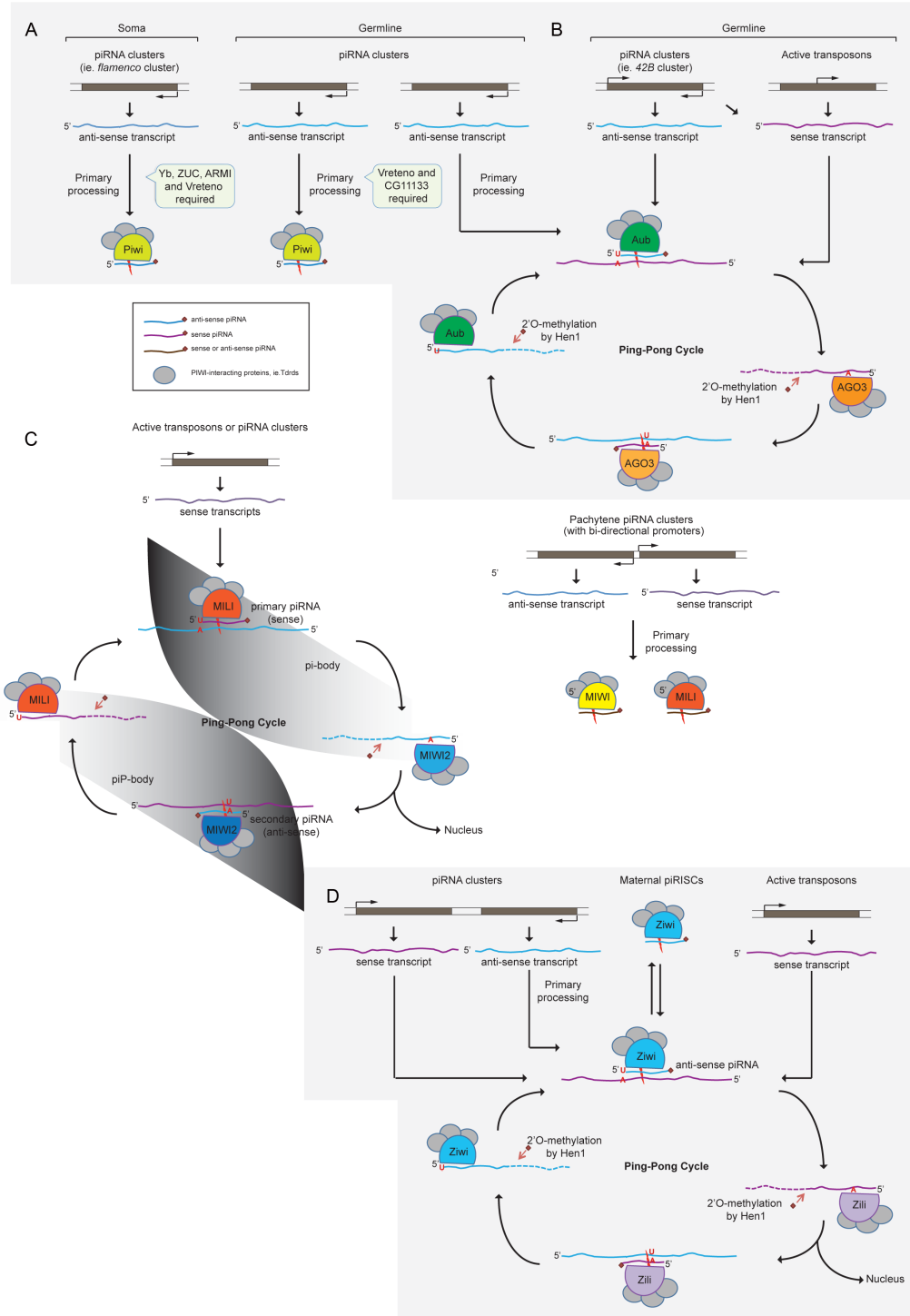


Figure 5. Schematic piRNA biogenesis in *Drosophila*, mice and zebrafish. (A) Primary processing pathway in somatic follicle cells and germ cells in the *Drosophila* ovary. (B) Ping-Pong amplification cy-

cle in the *Drosophila* germline. (C) Left: Ping-Pong amplification cycle during mouse spermatogenesis. Right: Primary processing of pachytene piRNAs. (D) Ping-Pong amplification cycle in zebrafish germline development.

In the Ping-Pong cycle in zebrafish germline, Ziwi and Zili share conserved molecular features with their mouse homologs. Both MILI and Zili preferentially bind sense piRNAs, while MIWI2 and Ziwi can bind both sense and anti-sense piRNAs. However, an important difference between the two systems is that Ziwi and Zili have reversed roles compared to their mouse homologs: Ziwi interacts with primary piRNAs whereas Zili associates with secondary piRNAs with strong enrichment of 10A bias (Houwing et al., 2008) (Fig.5D). Interestingly, both secondary piRNA-accepting PIWIs, MIWI2 and Zili, are able to shuttle between the nucleus and the cytoplasm, implying they may have conserved roles in chromatin remodeling.

In zebrafish, Ziwi, piRNAs (Houwing et al., 2007) and Tdrd6 (this thesis, Chapter 3), important components of the piRNA pathway, are also delivered maternally. So we can speculate that the maternal inherited piRISCs provide the first defense for the progeny to defeat active TEs in the early embryogenesis in zebrafish as well. piRNA profiling during different developmental stages has never been demonstrated in any system. In Chapter 2 of this thesis, we analyze and compare zebrafish piRNA profiles from various developmental stages. Our results also confirm the relationship between the maternal and early embryonic piRNAs in a vertebrate.

PIWI-piRNA pathway in Stem cells

PIWI proteins have been implicated in stem cell maintenance. *Drosophila* Piwi is essential for self-renewal in both male and female germline stem cells (Cox et al., 1998; Lin and Spradling, 1997). Mouse MIWI2 is important for germline stem cell maintenance (Carmell et al., 2007). PIWI-encoding genes are also found in flatworms that are known for their regenerative capacity. In the planarian *Schmidtea mediterranea*, three out of the ten *piwi* family genes, *smedwi-1*, *smedwi-2* and *smedwi-3*, are better characterized (Palakodeti et al., 2008; Reddien et al., 2005). PIWI proteins encoded by these three genes are unexceptionally expressed in the neoblasts, which are the dividing adult stem cells in planarians, and possess conserved catalytic residues that are required for Slicer activity (Palakodeti et al., 2008; Reddien et al., 2005). In planarians, piRNA-like small RNAs are also identified, although their direct interactions with PIWIs have not been demonstrated. Using RNAi assays, it was shown that both SMEDWI-2 and SMEDWI-3, but not SMEDWI-1, are required for neoblast function, regeneration and homeostasis, as well as the synthesis and/or expression of planarian piRNAs (Palakodeti et al., 2008).

PIWI arginine methylation and Tudor proteins

Several mass spectrometry analyses of purified PIWI proteins revealed that PIWI proteins, but not AGO, possess methylated arginine residues in their N-termini (Chen et al., 2009; Huang et al., 2011b; Kirino et al., 2009; Nishida et al., 2009; Reuter et al., 2009; Shoji et al., 2009; Vagin et al., 2009b). Arginine residues can be either mono (MMA) or di-methylated (DMA), while di-methylation can be symmetric (sDMA) or asymmetric (aDMA) (Fig.6A). The enzymes that are responsible for arginine methylation, called protein arginine N-methyl-transferases (PRMTs), are categorized into two groups. Type I (PRMT1, PRMT3, PRMT4 and PRMT6) and type II (PRMT5 and PRMT7) both can catalyse methylation of arginine residues to a mono-methylated intermediate, but type I PRMTs generate asymmetrical dimethylated arginine (aDMA) residues, while type II PRMTs produce symmetrical dimethylated arginine (sDMA) residues (Boisvert et al., 2005; Gary and Clarke, 1998; Lee et al., 2005; McBride and Silver, 2001). PRMT5 and its co-factor, WDR77, were co-immunoprecipitated with all three mouse PIWI complexes (Vagin et al., 2009a). In *Drosophila*, loss of PRMT5 (also known as Dart5 or Capsuleen) function phenocopies *aub* mutants, including de-repression of TEs and delocalization of Aub from the nuage in the mutant ovaries. Furthermore, PIWI proteins lost the sDMA modifications in *prmt5* mutants, suggesting that PRMT5 is the factor that mediates the sDMA modification (Harris and Macdonald, 2001; Kirino et al., 2009; Nishida et al., 2009). The interactions between PIWI proteins and PRMT5, Western blotting experiments using sDMA- or aDMA-specific antibodies, as well as mass spectrometry data coherently indicate that the methylation sites on the N-terminal tail of the PIWI proteins are symmetrically dimethylated arginines (sDMAs). These sDMAs are found within distinct motifs with the sequences of GRG and ARG/GRA (G, glycine; R, arginine; A, alanine).

At the same time, proteomic analysis of PIWI-associated proteins identified many tudor-domain-containing proteins (Tdrds) in various organisms (Chen et al., 2009; Huang et al., 2011b; Kirino et al., 2009; Kirino et al., 2010b; Nishida et al., 2009; Shoji et al., 2009; Vagin et al., 2009b; Vasileva et al., 2009; Wang et al., 2009) and this thesis). Tudor domains are commonly found in proteins in a wide range of organisms (Maurer-Stroh et al., 2003; Ponting, 1997; Talbot et al., 1998). They contain about 60 amino acids that form a beta-barrel-like core structure composed of beta sheets forming a hydrophobic pocket. The pocket comprises a cluster of conserved aromatic residues that constitutes a protein-protein interaction surface, which, at least in some cases, binds selectively to methylated lysine or methylated arginine residues (Jin et al., 2009; Selenko et al., 2001; Sprangers et al., 2003).

According to the distinct functions and sequences, Tdrd proteins are sub-divided into four groups (Jin et al., 2009) (Fig.6B). The first group is represented by the JMJD (JuMonJi Domain protein) family of histone demethylases, which bind to H3K4me and are involved in epigenetics

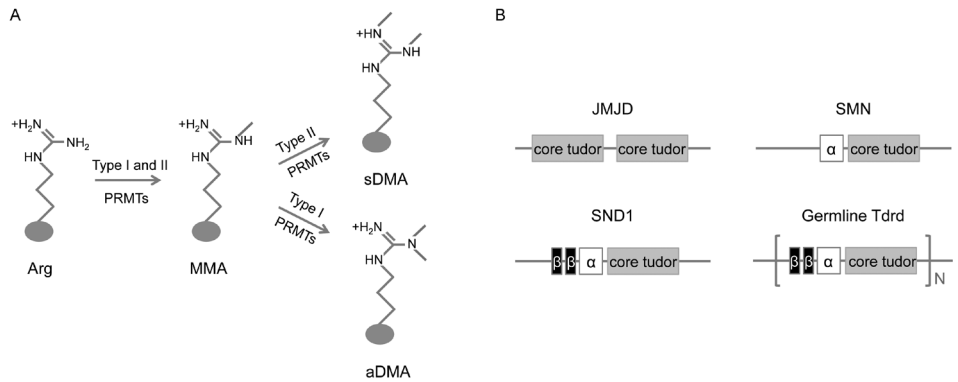


Figure 6. Tdrd proteins and arginine methylation. (A) Schematic presentation of arginine methylation. MMA: mono-methylated arginine; sDMA: symmetrically di-methylated arginine; aDMA: asymmetrically di-methylated arginine. (B) Schematic presentation of the four types of Tdrd proteins mentioned in the main text. Gray box: core tudor; white box: α-helix; black box: β-sheet. N is an integer that is equal to or greater than one (Adapted from Jin et al., 2009).

through chromatin regulation. JMJD Tdrds are characterized by two core tudor domains in tandem. The second group is represented by the survival motor neuron (SMN) protein family. SMN Tdrds have an α-helix N-terminal to the core tudor domain and are involved in mRNA splicing. The third group is represented by SND1 protein (staphylococcal nuclease and tudor domain containing 1), a component of the somatic RISC (Caudy et al., 2003). SND1 has five conserved tandem staphylococcal nuclease (SN)-like domains and the tudor domain is inserted after the 2nd β-strand of the 5th SN domain and creates the SN-tudor structure. It was suggested that the SN-tudor structure brings tudor domains into the pathways involving small RNA processing. The last one is the germline group, whose function remains unclear, but many studies have proposed their functions in the PIWI-piRNA pathway. Both SND1 and germline Tdrds have the basic structure of two β-strands and an α-helix preceding the tudor core. However, a feature that distinguishes the germline Tdrds from SND1 Tdrds is that many of the germline Tdrd members form tandem tudor repeats. These tandem tudor repeats have been hypothesized to simultaneously interact with several proteins harboring methylated arginines and to play a role in facilitating the generation of piRNAs (Jin et al., 2009).

Several germline Tdrd proteins have been associated with the PIWI-piRNA pathway in different animal models, such as Tdrd1, Tdrd2/Tdrkh, Tdrd4/RNF17, Tdrd5, Tdrd6, Tdrd7, Tdrd9, Tdrd12, Yb, Vreteno, PAPI, Tejas and Krimp (Chen et al., 2009; Handler et al., 2011; Huang et al., 2011b; Kirino et al., 2010b; Kojima et al., 2009; Lim and Kai, 2007; Liu et al., 2011; Patil and Kai, 2010; Saito et al., 2010; Shoji et al., 2009; Vagin et al., 2009a; Vasileva et al., 2009; Wang et al., 2009; Yabuta et al., 2011; Zamparini et al., 2011). Many studies have shown that interactions between various Tdrds and PIWI proteins are essential for nuage formation. Many

of the Tdrd-PIWI interactions are RNA-independent and can be reproduced under heterologous settings. Treatment with MTA [59-deoxy-59-(methylthio)-adenosine], a general methyl-transferase inhibitor, in cell culture has shown that sDMAs on the PIWI proteins are essential for the PIWI-Tdrds interaction (Vagin et al., 2009b). Below I will describe the findings of some germline Tdrds involved in the PIWI-piRNA pathway. The germline Tdrds that are identified to date are summarized in Table 2.

***Drosophila* Tudor** is the founder of the germline Tdrds. It contains 11 tudor domains and is required for the assembly of the germ plasm (Thomson and Lasko, 2005). Tudor associates with Aub and AGO3, and the Tudor-Aub interaction is important for their localization and germ plasm formation (Kirino et al., 2009; Nishida et al., 2009). Further analysis shows that the 7th to the 11th tudor domains of Tudor are sufficient for Tudor-Aub interaction and germ cell formation (Arkov et al., 2006; Kirino et al., 2010b). Later a mutation study on the 7th -11th tudor domains of Tudor revealed that the 8th domain is dispensible for the function, while the rest of the domains are required. Liu H et al. (2010) also identified a sequence motif, YR-D(F/Y)GN, which appears to be conserved between the essential tudor domains. Based on the structural analysis of the 11th tudor domain, the aromatic residues that form the sDMA binding pocket were also identified (Liu et al., 2010a; Liu et al., 2010b). It will be very interesting to further investigate whether these four tudor domains interact with Aub independently or synergistically.

Tdrd1 contains an N-terminus MYND (myeloid-Nervy-DEAF-1) zinc finger domain and four tudor domains. The MYND domain has two putative non-DNA-binding zinc fingers and is involved in protein-protein interactions. It is often found in nuclear proteins and has been implicated in transcriptional repression (Liu et al., 2007). TDRD1 is essential for spermatogenesis in mice and germline development in zebrafish (Chuma et al., 2006; Huang et al., 2011b). In mice, TDRD1 interacts with MILI and MIWI2 (Vagin et al., 2009b; Wang et al., 2009). The TDRD1-MILI interaction is mediated through the MYND domain and the first two tudor domains (Wang et al., 2009). MILI proteins fail to form cytoplasmic granules when expressed in a heterologous cell culture system, but are recruited to the granules if co-expressed with TDRD1 (Vagin et al., 2009). Treatment with MTA abolishes the association of MILI with TDRD1 suggesting that the MILI-TDRD1 interaction depends on sDMA modification (Reuter et al., 2009; Vagin et al., 2009b). Loss of TDRD1 also leads to impaired Ping-Pong cycle in mice, as well as transposon de-repression, particularly LINE-1 (Reuter et al., 2009). In zebrafish, Tdrd1 interacts strongly with Zili and Ziwi, and it is also important for nuage localization of Zili. Different from the mouse study, we found that all four tudor domains can interact with Zili *in vitro*. We showed that sDMA modification is not sufficient to promote Tdrd1-Zili association. There may be some sequence specificity in the Tdrd1-Zili interaction. Loss of zebrafish

Tdrd1, similar to mice, leads to severe reduction of piRNA population and upregulation of transposons. In zebrafish testis, using immunoprecipitation assays we identified piRNAs and Tdrd1-associated transcripts (TATs), which are likely piRNA intermediates. It suggests that Tdrd1 acts as a molecular scaffold for PIWI proteins, piRNA and TATs and therefore facilitates Ping-Pong amplification (Huang et al., 2011b).

Tdrd2/Tdrkh contains one tudor domain and two tandem KH domains. KH domains are nucleic acid recognition motifs associated with transcriptional and translational regulation. KH domains bind RNA or single-stranded DNA molecules. Multiple copies of KH domains within a protein have been proposed to achieve better affinity and specificity of nucleic acid binding (Valverde et al., 2008). In mice, TDRKH is highly expressed in the testis and to a lesser extent in the brain. From the temporal expression of TDRKH, it was suggested to play a role in prospermatogonia differentiation and adult meiosis (Chen et al., 2009). Tdrkh colocalizes with MVH and interacts with MIWI in an RNA-independent manner. Chen et al (2009) performed mass spectrometry analysis and predicted three sDMA cluster candidates at the N-terminus of MIWI. Mutation studies on either the putative sDMA sites on MIWI or the conserved consensus residues in the tudor domain show that the TDRKH-MIWI interaction is sDMA-dependent. However, the molecular function of TDRKH remains elusive.

TDRD4/RNF17 in mice has two isoforms, called RNF17S and RNF17L. Both contain one RING domain and either three (RNF17S) or five (RNF17L) tudor domains (Pan et al., 2005). The RING domain is a zinc-binding motif and is often found in proteins with ubiquitin E3 ligase activity (Joazeiro and Weissman, 2000). RNF17 is testis-specific and was reported by electron microscopy studies to form cytoplasmic nuage-like RNF17 granules via self-self interaction. RNF17 granules do not colocalize with TDRD1 and TDRD6, two CB and IMC components, indicating that RNF17 granules are distinct nuage-like structures. Animals lacking *rnf17* are viable but male sterile, due to complete arrest at the round spermatids stage. Interestingly, mouse mutants of *miwi*, *tdrd5* and *tdrd6* show similar spermiogenic phenotypes (Deng and Lin, 2002; Vasileva et al., 2009; Yabuta et al., 2011), which is also similar to the phenotype of the mutant for CREM, cAMP-responsive element modulator, a key regulator of spermiogenesis (Blendy et al., 1996; Nantel and Sassone-Corsi, 1996). Loss of either *rnf17*, *tdrd5* or *miwi* does not affect CREM expression, but down-regulation of putative CREM target genes, *prm1*, *prm2*, *tnp1*, as well as a testis-specific coactivator of CREM, *act,* was observed (Deng and Lin, 2002; Pan et al., 2005; Yabuta et al., 2011). RNF17 was identified in MIWI immunoprecipitates while TDRD5 colocalizes with MIWI in male germ cells (Vagin et al., 2009b; Yabuta et al., 2011). Taken together, it was suggested that RNF17 and TDRD5 may work together with MIWI to stabilize mRNAs of ACT and other CREM target

Table2. Summary of germline Tdrds in *Drosophila*, mice and zebrafish.

Tdrd protein	Species	Motif	With YR-D(F/Y)GN?	Function	Reference
Tudor	*Drosophila*	tudor x 11	2th, 5th-7th, 9th-11th	Germline development piRNA biogenesis	Schupbach and Wieschau, 1986; Nishida et al., 2009; Kirino et al., 2010; Liu H et al., 2010; Liu K et al., 2010
Tdrd1	Mouse	tudor x 4, MYND	2nd-4th	Spermatogenesis piRNA biogenesis	Chuma et al., 2006; Reuter et al., 2009
	Zebrafish	tudor x 4, MYND	1st-4th	Germline development piRNA biogenesis	Huang et al., 2011
Tdrd2/Tdrkh	Mouse	tudor, KH x 2	None	Implicated in spermatogenesis	Chen et al., 2009
Tdrd4/RNF17	Mouse	tudor x 5, BBC	2nd	Implicated in spermatogenesis	Pan et al., 2005
Tdrd5	Mouse	tudor, LOTUS x 3	Yes	Spermatogenesis, piRNA biogenesis	Yabuta et al., 2011
Tdrd6	Mouse	tudor x 7	4th-6th	Spermatogenesis	Hosokawa et al., 2007; Vasileva et al., 2009; Tanaka et al., 2011
	Zebrafish	tudor x 7	4th, 6th, 7th	Germ plasm assembly piRNA biogenesis	this thesis, Chapter 4
Tdrd7/TRAP	Mouse	tudor x 3, LOTUS x 3	None	Spermatogenesis transposon silencing	Tanaka et al., 2011
Tdrd8/Stk31	Mouse	tudor, polC, PKc-like	Yes	Impicated in spermatogenesis	Chen et al., 2009
Tdrd9/Spn-E	*Drosophila*	tudor, DExD/H, HA2, HELICc	None	Germ plasm assembly piRNA biogenesis, histon methylation	Pal-Bhadra et al., 2004; Lim and Kai 2007
	Mouse	tudor, DExD/H, HA2, HELICc	None	Spermatogenesis, piRNA biogenesis	Shoij et al., 2009
	Zebrafish	tudor, DExD/H, HA2 HELICc	Yes	Germline development piRNA biogenesis	This thesis, Chapter 5
Tdrd12	Mouse	tudor	None	Unknown	
Yb (mouse Tdrd12 homolog)	*Drosophila*	tudor	None	Somatic primary piRNA biogenesis	Saito et al., 2009; Olivieri et al., 2009
CG11133 (Brother of Yb)	*Drosophila*	tudor	None	Germline primary piRNA biogenesis	Handler et al., 2011
CG31755 (Sister of Yb)	*Drosophila*	tudor	None	Somatic/germline primary piRNA biogenesis	Handler et al., 2011
Krimper	*Drosophila*	tudor	Yes	piRNA biogenesis	Lim and Kai, 2007
Ovarian tumor	*Drosophila*	tudor, OUT	None	RNA localization in nurse cell, chromosom disp	Goodrich et al., 2004
Vreteno	*Drosophila*	tudor x 2	None	Somatic and germline primary piRNA biogenes	Handler et al., 2011; Zampirini et al., 2011

Table 2. Germline Tdrd proteins involved in PIWI-piRNA pathway in mice, *Drosophila* and zebrafish. A list of Tdrd proteins that have been reported or speculated with functions in the germline and/or PIWI-piRNA pathway is presented. The protein name, species of origin, motifs residing in the protein, possible function and the references are included in the table. Tudor domains containing the consensus sequence (YR-D(F/Y)GN are also indicated in the table. MYND: myeloid, Nervy, and DEAF-1 Zinc finger; KH: K homology RNA binding domain; RING: RING zinc binding motif; polC: DNA polymerase III PolC; PKc: protein kinases catalytic domain; DExD/H: DEAD-like helicase superfamily; HA2: helicase associated domain; HELICc: helicase superfamily C-terminal domain; OUT: ovarian tumor-like cysteine protease (Adopted from Siomi et al., 2010).

genes by forming a complex with them, to regulate the expression of these spermiogenic genes (Deng and Lin, 2002; Yabuta et al., 2011). Besides the spermiogenic phenotype, TDRD5 was also found to be important in the assembly of IMC and CB. Loss of *tdrd5* leads to mis-localization of IMC/CB components together with retrotransposon upregulation, particularly LINE-1, due to demethylation at the promoter regions (Yabuta et al., 2011).

Tdrd5 is composed of a tudor domain and three LOTUS domains at its very N-terminus. The LOTUS (Limkain, Oskar and Tudor-containing proteins 5 and 7) domain was recently identified via bioinformatic analysis (Callebaut and Mornon, 2010). Although its actual function remains unknown, the presence of LOTUS in germline-specific proteins, such as Oskar, Tdrd5 and Tdrd7, suggests its possible role in the germline (Callebaut and Mornon, 2010). Tdrd5 and Tdrd7 are the only two Tdrds possessing the LOTUS domain. In our mass spectrometry analysis using immunoprecipitation in adult zebrafish gonads, Tdrd1 brought down Tdrd5, while Tdrd6 brought down Tdrd7. It is possible that Tdrd1/Tdrd5 and Tdrd6/Tdrd7 represent functional pairs in the PIWI-piRNA pathway in zebrafish (This thesis, Chapter 3 and 4).

Tdrd6 contains seven tudor domains (Hosokawa et al., 2007) and is the closest vertebrate homolog of *Drosophila* Tudor. Most of the studies regarding Tdrd6 have been done in mice and TDRD6 was reported to interact with both MILI and MIWI (Chen et al., 2009; Kirino et al., 2010b; Vagin et al., 2009b; Vasileva et al., 2009). TDRD6 is important in spermiogenesis and chromatoid body formation, together with TDRD1 and TDRD7 (Hosokawa et al., 2007; Tanaka et al., 2011; Vasileva et al., 2009). So far, mouse TDRD6 seems not to be involved in transposon silencing (Tanaka et al., 2011; Vasileva et al., 2009). In zebrafish, *tdrd6a*, a paralog of three zebrafish tdrd6 genes, is involved in germ plasm assembly and transposon silencing. We also found that Tdrd6a may be part of an mRNA processing/metabolism machinery in the zebrafish oocytes. The role of zebrafish Tdrd6a is further described in Chapter 4 in this thesis.

Tdrd7 has three tudor domains and three N-terminal LOTUS domains. Like many germline Tdrds in mice, TDRD7/TRAP is essential for male fertility. It forms a complex with TDRD6 and TDRD1, and is a CB component down stream of MVH (Hosokawa et al., 2007). Recently, Tanaka et al. (2011)

showed that CB is a dynamic structure with a changing composition during spermatogenesis. TDRD7 is essential for the fusion of CBs with p-bodies, and may affect the translational profile during spermiogenesis. It was also reported that TDRD7 suppresses retrotransposons in a piRNA biogenesis-independent manner (Tanaka et al., 2011). Beside its germline function, TDRD7 forms cytoplasmic RNP granules in the lens fiber cells and is critical for normal lens development (Lachke et al., 2011). This brings up the possibility that more germline Tdrd proteins may be involved in RNP formation and organogenesis other than gonads. In zebrafish PGCs, Tdrd7 is localized in the germ plasm and is essential for determining the normal nuage morphology and number (Strasser et al., 2008). The identification of Tdrd7 in immuno-complexes of Zili and Tdrd6 also confirm the property of RNP formation of zebrafish Tdrd7 (Huang et al., 2011b and this thesis, Chapter 4).

Tdrd9 is composed of only one tudor domain and an ATPase/DExD/H-type RNA helicase domain. ATPase/DExD/H RNA helicase belongs to the DEAD-box protein superfamily. This type of helicase can move along nucleic acids and some of them can unwind RNA secondary structures in an ATP-dependent manner. The ATPase activity of the helicase is dependent on or stimulated by the association with RNAs. Members of this type of RNA helicase are involved in various RNA processing and translational regulations. (Cordin et al., 2006). In *Drosophila*, SpnE, the homolog of Tdrd9, is crucial for Ping-Pong amplification of piRNAs, transposon silencing and proper nuage localization of Vasa, MAEL and Tejas (Findley et al., 2003; Gillespie and Berg, 1995; Gonzalez-Reyes et al., 1997; Patil and Kai, 2010). In mice, TDRD9 is strongly associated with MIWI2. Loss of Tdrd9 results in changes of piRNA population as well as transposon upregulation in mutants, particularly of LINE-1, but has no affect on PIWI protein localization (Shoji et al., 2009). The zebrafish homolog of Tdrd9 is also involved in transposon repression and piRNA amplification. We show that the tudor domain of Tdrd9 can interact with Zili *in vitro* and the DExD/H helicase domain is essential for Tdrd9 function in the germline (This thesis, Chapter 5). Interestingly, the tudor domain of *Drosophila* SpnE possesses amino acid changes in the consensus sequence required for sDMA binding. It has been shown that the tudor domain of SpnE is not responsible for the SpnE-Aub interaction. In mice, TDRD9 tudor domain also contains changes in the consensus sequence, while the consensus sequence is intact in zebrafish Tdrd9. This strengthens the idea that certain sequences besides the consensus sequence are essential for Tdrd9-PIWI interaction.

Vreteno and three orthologs of murine **Tdrd12**, Yb, CG11133 (Brother of Yb) and CG31755 (Sister of Yb), are the Tdrds recently reported to be essential for the somatic and/or germline primary processing pathway of piRNAs in *Drosophila* (Handler et al., 2011; Olivieri et al., 2010; Saito

et al., 2010; Zamparini et al., 2011). None of them contain the complete consensus sequences for sDMA binding. Among these four Tdrds, Vreteno and CG31755 are present both in the somatic gonads and germline cells, Yb is only in the somatic gonads, and CG11133 is only in the germline. In the somatic follicle cells Yb is required for mature piRNA loading onto Piwi prior to the nuclear localization of Piwi (Olivieri et al., 2010; Saito et al., 2010). Vreteno accumulates in Yb bodies and interacts with Piwi and Yb. Loss of Vreteno leads to loss of Piwi, a phenotype of disrupted primary processing pathway, and de-repression of *gypsy*, *ZAM* and *Idefix*, which are TEs regulated by Piwi (Handler et al., 2011; Malone et al., 2009; Zamparini et al., 2011). On the other hand, in the germline CG11133 and Vreteno are localized in the nuage and are crucial for transposon silencing. CG31755 is located in different perinuclear foci from nuage. It functions together with CG11133 to stabilize Piwi localization and plays a role in the primary piRNA biogenesis in the germline (Handler et al., 2011; Zamparini et al., 2011).

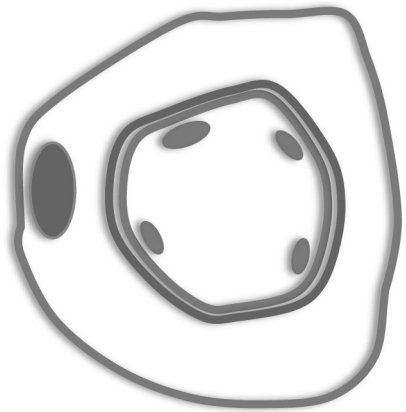
Another three nuage Tdrds, **Krimp**, **Tejas** and **PAPI**, are also found to be crucial for PIWI localization and transposon silencing in *Drosophila* (Lim and Kai, 2007; Liu et al., 2011; Patil and Kai, 2010). Tejas is essential for piRNA production in the germline but not the somatic gonad. The conserved N-terminal tejas domain is essential for Tejas function while the tudor domain is dispensable. PAPI interacts with all three *Drosophila* PIWIs and at least in the case of AGO3 is via sDMAs. It should be noted that the tudor domain of PAPI is required but not sufficient for PAPI-AGO3 interaction (Liu et al., 2011).

To summarize the above, Tdrd-PIWI interactions in most cases are crucial for proper intracellular localization of the PIWI-piRNA machinery and its silencing mechanism. Many of the Tdrds mentioned above were confirmed to possess the conserved aromatic cage, which is known to bind sDMAs, in at least one of their tudor domains (Table 2). Although the physical interactions of these Tdrds with PIWI proteins may not yet all be proven, the presence of the conserved aromatic cages suggests that most tdrd proteins will interact with sDMAs found on PIWIs and other PIWI-piRNA pathway components, such as Vasa (Chen et al., 2009; Huang et al., 2011b; Kirino et al., 2010a; Liu et al., 2010a; Liu et al., 2010b; Liu et al., 2011; Reuter et al., 2009; Shoji et al., 2009; Vagin et al., 2009b; Wang et al., 2009). Vasa has a conserved and essential role in germline formation and is often found in association with PIWI-piRNA pathway components in the nuage (Siomi et al., 2011). Mouse, *Xenopus* and *Drosophila* Vasa, was recently reported to possess sDMA and aDMA. Mouse Vasa (MVH) was also found to interact with mouse TDRD1, TDRD6 as well as MILI and MIWI (Kirino et al., 2010a). It was therefore suggested that arginine methylation creates a specific code essential for germline development (Kirino et al., 2010a). At the same time, several Tdrds, such as *Drosophila* SpnE, Vreteno, Tejas, either have

mutations in the aromatic cage or were found to interact with PIWI in a sDMA-independent manner, suggesting the presence of alternative mechanisms for Tdrd-PIWI interaction (Nishida et al., 2009; Patil and Kai, 2010; Handler et al., 2011). Other protein domains in Tdrds, for example, KH RNA binding domain (Tdrkh), Tejas domain (Tejas), MYND zinc finger domain (Tdrd1), DExH RNA helicase domain (Tdrd9), are also likely to be involved in protein-protein or protein-RNA interaction in the piRISCs. It is known that many nuage components mutually depend on each other for their localization to nuage. It is therefore suggestive that through the multiple tudor domains and other RNA binding domains present in the Tdrds, macromolecules composed of PIWI-piRNA pathway partners then form various nuage and nuage-like RNP structures where essential RNA processing reactions, including piRNA biogenesis, take place.

In this thesis, I will present novel findings regarding germline Tdrd proteins in the zebrafish system. The main focus will be on their function in gametogenesis and their relation to the PIWI-piRNA pathway.

Chapter 2



piRNA profiles during germline development in zebrafish

2

Huang HY, Kaaij LJT, Berezikov E, Ketting RF.

2

INTRODUCTION

Zebrafish germline development

In zebrafish, germ cell fate is determined by the inheritance of the germ plasm. Germ plasm is a maternally provided cytoplasm composed of RNA and RNA-binding proteins. It is first formed at the distal end of the cleavage furrows at the 2- and 4-cell stage embryos (Yoon et al., 1997). At the 32-cell stage, 4 cells have obtained the germ plasm aggregates and become promptive primordial germ cells (pPGCs) (Braat et al., 1999). These cells undergo migration and proliferation and become primordial germ cells (PGCs). At 24 hour post fertilization (hpf), PGCs end up at the future gonad area, numbering between 25 to 50 cells (Raz, 2003). Zebrafish do not have sex chromosomes as many other species do. Sex determination of zebrafish depends on several factors, such as the environment and the population density. It has been proposed that the number of germ cells present in the early gonads also affects the sex determination in zebrafish (Siegfried and Nusslein-Volhard, 2008). Zebrafish are considered juvenile hermaphrodites. All fish first develop an immature ovary, called juvenile ovary, during juvenile development. Later in development, the gonads differentiate into either ovaries or testes (Takahashi, 1974). The timing of the gonadal developmental stages varies between different laboratory settings, but the sequence of the stages remains the same. In our laboratory, we observe that PGCs have differentiated into juvenile oocytes around three weeks post fertilization (wkpf). At 5wkpf, while most of the fish retain the juvenile ovaries, some possess gonads undergoing transformation into testes. The animals having higher number of juvenile oocytes are speculated to become future-females. Other animals are speculated to become males based on the following phenotypes: degeneration of the juvenile oocytes, higher numbers of gonadal somatic cells and onset of spermatogenesis (Siegfried, 2010). The transition from juvenile hermaphrodites to mature females or males can take up to 8 weeks. After 12wkpf, most of the animals are sexually mature.

Transposable elements (TEs)

During the whole process of germline development, it is essential to maintain genome integrity in the germ cells. Transposable elements (TEs), DNA sequences that can move from one place of the genome to another, are a notorious threat to the genome. Based on the mechanisms for their transposition, TEs can be divided into two main classes: Class I and Class II transposons. Class I transposons are also called retrotransposons. They are characterized by the “copy and paste” mechanism using RNA intermediates that are later reverse-transcribed back to DNA during transposition (Weiner et al., 1986). Class I elements are composed of two types of retrotransposons: LTR and non-LTR. The main difference of these two types is the presence or absence of the long-terminal repeat (LTR) sequences at both ends of the

elements. LTR retrotransposons can be further categorized into four main super-families: Gypsy, Copia, BEL and DIRS, based on their sequence similarities and structural features (de al Chaux et al., 2011). Non-LTR retrotransposons are divided into long interspersed nuclear elements (LINEs) and short interspersed nuclear elements (SINEs). LINEs have two ORFs, which encode all the proteins necessary for the retro-transposition, are autonomous retro-elements. On the other hand, SINEs use the machinery from LINEs and function non-autonomously. Class II elements utilize a different mechanism to mobilize: cutting DNA segments out of one place and then pasting it into another place of the genome. Most of the class II elements encode a transposase, the enzyme that cleaves the transposon out of its original site, as well as producing a breakage at the target site for the transposon insertion. This process can be dependent or independent of the flanking sequences of the transposons. Because of the transposition activities of the TEs, they are deleterious to genome integrity.

The dynamic expression of PIWI proteins

During development of the germline, there is a particular group of proteins essential to maintain this process and ensure the genome integrity in the germ cells. They are called the PIWI proteins, a subclass of the Argonaute protein family that is predominantly expressed in the germline (Farazi et al., 2008). PIWI proteins and their interacting small RNAs, piRNAs, are crucial for the development and function of the germline in various species (Siomi et al., 2011). The main function of the PIWI-piRNA pathway, which has been identified so far, is defense against selfish TE activity in the germline.

There are two PIWI proteins encoded in the zebrafish genome, Ziwi and Zili. Ziwi protein is maternally provided and can be first clearly detected at the cleavage furrow at the 2- and 4-cell stage embryos (Houwing et al., 2007). At 24hpf, Ziwi is localized in the prominent perinuclear structures termed nuage (Houwing et al., 2007). Nuage is a form of germ plasm and can be detected throughout the germline development (Eddy, 1974). During the development, Ziwi protein gradually diffuses in the cytoplasm and this cytoplasmic localization remains until adulthood in the germ cells of both sexes (Eddy, 1974). Zili expression starts at 3dpf and can be detected in the nucleus of the PGCs. Zili remains in the nucleus until 7dpf and then translocates into perinuclear nuage structures. During oogenesis, Zili appears as perinuclear nuage granules before the maturing oocyte stage, when Zili is again located in the nucleus. During spermatogenesis, Zili granules are detected in the spermatogonia and spermatocytes (Houwing et al., 2008).

Ziwi and Zili are proposed to conduct the biogenesis of piRNAs and destroy active TE mRNAs in the zebrafish germline via a feed-forward amplification cycle (Houwing et al., 2008). The cycle begins with maternally inherited Ziwi together with its associated piRNAs, referred as primary

piRNAs. piRNAs produced from piRNA clusters can also be the input of the amplification cycle. Ziwi, guided by piRNA, recognizes active TE mRNAs with complementary sequences and cleaves the mRNAs. This event will produce the 5' end of a new piRNA, the secondary piRNA, which can be loaded onto Zili. The 3' end of the secondary piRNA will be further processed to a specific size and protected with 2' O-methylation by the Hen1 protein (Kamminga et al., 2010). The mature secondary piRNA then further guides Zili to target complementary TE mRNAs and produces a new piRNA that resembles the first primary piRNA and can load onto Ziwi. This process is therefore dubbed as the Ping-Pong cycle (Brennecke et al., 2007; Gunawardane et al., 2007). Ziwi is preferentially loaded with primary piRNAs with an uridine at its 1st 5' end nucleotide (1U), whereas Zili prefers secondary piRNAs with an adenine as the 10th nucleotide from their 5' end (10A), corresponding to the 1U on Ziwi piRNAs. Ziwi and Zili piRNAs often have complementary sequences of 10 nucleotides from their 5' ends. The above features of the piRNAs taken together are considered the Ping-Pong signatures (Brennecke et al., 2007; Gunawardane et al., 2007; Houwing et al., 2008).

In mice, piRNA profiles of type A spermatogonia (mitotic stage), pachytene spermatocytes (meiotic stage) and round spermatids (post-meiotic stage) were recently analyzed and compared by Gan et al (2011). However, there was no report yet regarding piRNAs during the process of the germline development. To have better insights of the piRNA functions during the gonadal development in zebrafish, we studied the piRNA profiles choosing the following time points during development: 3dpf (when Zili starts to express and locates in the nucleus), 7dpf (when Zili leaves the nucleus and enters the nuage), 3wkpf (gonads appear as juvenile ovaries), 5wkpf (in the process of sex differentiation), and adulthood (both ovaries (TL-F/ TU-F) and testes (TL-M/ TU-M)). It is to be noted that at 5wk, all the gonads were pooled without separating animals that are male- or female-like. Although desirable, this is technically difficult since the animal/gonad size is very small and hard to distinguish while fast sample processing to prevent RNA degradation is necessary.

In this study, we demonstrate that piRNA population is not constant during gonadal development. Distinct levels of piRNAs at different developmental stages may reflect the activity of the TEs from which the piRNAs are derived.

RESULTS

Vasa is a germ cell marker throughout the zebrafish germline (Yoon et al., 1997). A fish line expressing enhanced green fluorescent protein (EGFP) under the vasa promoter (vas::EGFP) has been reported (Krovel and Olsen, 2002). We took advantage of this transgenic line and used animals with the vas::EGFP transgene to isolate germ cells. In total we have 5 developmental

time points and 8 small RNA libraries. Most of the samples are obtained from a wild-type fish line with TL (Tupfel Long fin) background (3dpf, 7dpf, 3wkpf, 5wkpf, TL-F and TL-M), except one set of ovary and testis samples (TU-F and TU-M), which are from wild-type animals with TU (Tubingen) background. TU adult libraries are only compared with TL adult libraries to study the piRNA profile difference between different fish strains.

Even though we do find piRNAs derived from genes and other parts of the genome, we have so far not observed effects of Piwi pathway mutations on the expression of these genes. Therefore, in the analyses presented here we have only focused on transposon-derived piRNAs (referred to as repeat piRNAs), as thus far, transposon silencing represents the only consistent activity of piRNA pathways in various organisms.

Composition of the libraries

The composition of the 8 small RNA libraries is present in Table1 and Fig.S1A. Owing to different preparation methods for our small libraries, such as FACS purified germ cells vs. total gonads, cloning the small RNAs with polyA tailing vs. 3' end adaptor ligation, as well as the differentiation of the somatic tissue in the gonads, prevent a proper comparison between different time points. Therefore using miRNA as a standard to normalize piRNA reads in all the libraries cannot be applied. However, some pairs of the libraries came from the same preparation, 3dpf vs. 7dpf, TL-F vs. TL-M and TU-F vs. TU-M; we therefore compare them by miRNA normalization. When we look at only the repeat piRNAs, it is clear that 7dpf PGCs have more piRNAs than 3dpf PGCs (Fig.S1B). TL-F samples clearly have more piRNAs than TL-M samples (Fig.S1C), while TU-F and TU-M samples have similar piRNA abundance (Fig.S1D).

Ping-Pong amplified piRNAs are more dominant in PGC stages

To check whether piRNAs from different stages all display similar Ping-Pong signature, we checked the identity of the most 5' end nucleotide as well as the overlapping between two opposite strands of piRNAs within each

Table1. Library composition

Libraries	miRNA	repeat piRNA	genic piRNA	other piRNA	snRNA	snoRNA	tRNA	rRNA	Total
3dpf	159981	2354637	11549	1095465	1139	3160	218892	238365	4083188
7dpf	107980	5873337	25988	3247588	2169	1904	145486	265988	9670440
3wkpf	4221083	7515215	82346	4769450	4904	6236	186723	1226726	18012683
5wkpf	226060	611458	20745	447999	4644	6625	243597	36567	1597695
TL-F	1189723	3697820.5	24726.5	2157992	3596	19043.5	48588.5	82195	7223685
TL-M	1887906	1288259	12878	1715978	7328.5	3201	20409	21760.5	4957720
TU-F	777228	12632648	116748	6785442	14331	3394	124783	339487	20794061
TU-M	608837	9864163	131685	10897333	3296	876	241817	358412	22106419

Table1. Small RNA composition of all libraries. Reads annotated from Illumina sequencing. It is to be noted that 3dpf, 5wkpf and TL-M libraries contain relatively low reads. 3dpf, 7dpf: libraries made from 3dpf and 7dpf PGCs. 3wkpf, 5wkpf: libraries made from isolated gonads at 3wkpf and 5wkpf. TL-F, TL-M: libraries made from isolated adult ovaries (F) and testes (M) with TL background. TU-F, TU-M: libraries made from isolated adult ovaries (F) and testes (M) with TU background.

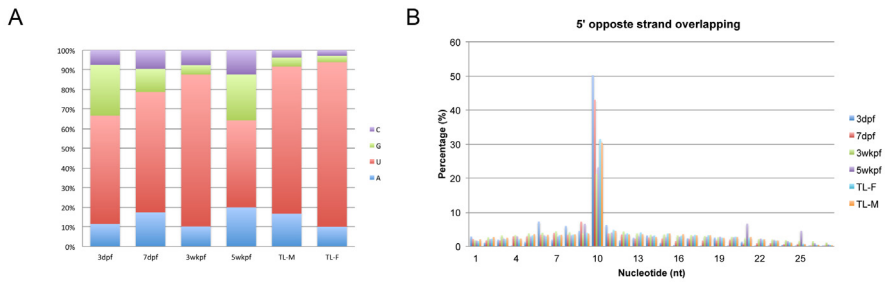


Figure 1. Ping-Pong signature in the piRNAs of all developmental stages. (A) The percentages of each nucleotide (A, U, C or G) at the most 5' end of mapped piRNAs are plotted in the graph. (B) The number of overlapped nucleotides between two opposite strand piRNAs within a library (x-axis) is plotted against the amount of corresponding incidents in percentage (y-axis). Different libraries are indicated in different colors.

library. In all stages, the most 5' end nucleotide of the piRNAs is largely present as a uridine residue (Fig.1A). This can be the result of dominance of Ziwi-bound piRNAs in all developmental stages. The 10-nucleotide overlapping between two opposite-stranded piRNAs within one library is also consistent in all stages (Fig.1B). However, interestingly, piRNAs from 3dpf and 7dpf PGCs show higher percentage (15-20% more than in the adults; 20~25% more than in the juveniles) of piRNA reads with 10-nucleotide overlapping. It suggests that piRNAs from these stages are either mainly produced via Ping-Pong cycle to target active TE transcripts or the maternally provided piRNAs are mostly produced by Ping-Pong cycle. Of course, both of the cases can hold true at the same time. Either way, it suggests that controlling active TEs is a very active process during early germline development.

PGC piRNAs are shorter in length

In an early report, it has been shown that zebrafish piRNAs from adults are 26-28 nucleotides (nt) long (Houwing et al., 2007), we wondered whether the size of the piRNAs remains the same throughout the development. We plotted the percentage of perfectly mapped reads against the length in each developmental stage. Owing to the size range of the small RNAs (18 – 50 nt) that are considered, we observed two main peaks for each stage as we expected: the miRNA peak (between 21-22nt) and the piRNA peak (between 26-30nt) (Fig.S2). Despite the fact that piRNA peaks at all stages display a Gaussian distribution, the distribution maxes out at different lengths at different stages. The peaked length at 3dpf, 7dpf, 3wk, 5wk, adult female (TL-F) and adult male (TL-M) are 26-27nt, 27-28nt, 28-29nt, 28nt, 28-29nt and 28nt respectively (Fig.2). Our data shows that piRNAs are 1 to 2 nt shorter at the PGC stages for yet unknown reasons.

A number of explanations may exist for this interesting observation. First, piRNAs may be trimmed to a shorter length in the PGCs because the Piwi proteins may adopt different conformations due to different interacting

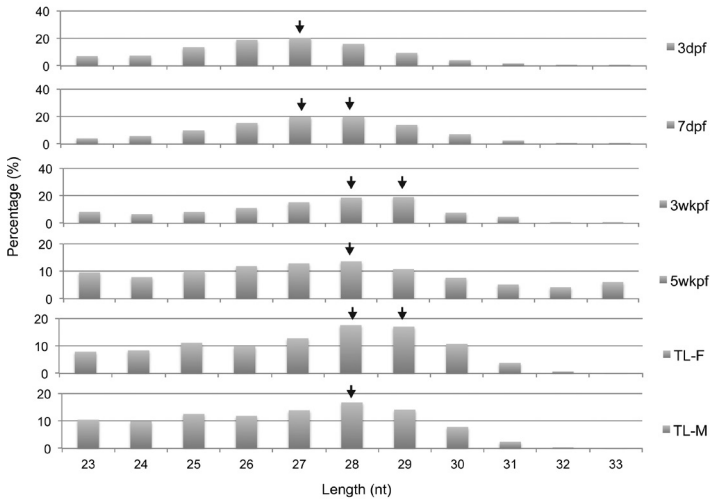


Figure 2. piRNAs are shorter in PGC stages. The percentage of mapped piRNA reads (y-axis) is plotted against the length of the piRNA in nucleotides (x-axis) in bar diagrams. The stages of the germ cell source are indicated above each sub diagrams. Black arrows indicate the highest peaks at each developmental stage.

partners during development. PIWI proteins in zebrafish display dynamic expressions: at 3dpf and 7dpf, Ziwi is present in the perinuclear nuage and Zili is in the nucleus; from 3wkpf till adulthood, Zili is detected in the perinuclear nuage structures while Ziwi is dispersed in the cytoplasm. This may lead to or be the consequence of different interacting partners that are also involved in the piRNA biogenesis. Remarkably, the length distribution we observed has a correlation with the expression pattern of the PIWI proteins in zebrafish. The piRNAs are shorter only when Zili is located in the nucleus. Second, in the timespan between the production of piRNAs in the oocytes and our detection in the PGCs, they may have become shorter, especially in 3dpf PGCs where piRNAs bound by maternally inherited Ziwi may still contribute to a big fraction of the piRNA population. This is not unreasonable, as we know that piRNAs need to be stabilized by methylation to prevent shortening at their 3' end (Kamminga et al., 2010). Even though the maternally provided piRNAs are methylated, the existing trimming activities may over longer times result in progressive shortening of piRNAs. Identification of the molecules involved in piRNA processing, especially at the PGC stages, will help us to better understand how the length differences originate and the functions of the piRNAs at different developmental stages.

Ziwi bound piRNAs are dominant throughout development

It has previously been shown that in the adult gonads, Ziwi-bound piRNAs (mostly anti-sense oriented) are more prominent than Zili-bound piRNAs (mostly sense oriented). At the same time, ovarian piRNAs are much more anti-sense oriented than testis piRNAs (Houwing et al., 2007). When we took together other developmental stages, we noted that overall piRNAs from different stages also display anti-sense bias. Consistent with the observation of prominent 1U bias of piRNAs in all stages, it shows that Ziwi

piRNAs are likely dominant during the whole gonadal development. Among all stages, piRNAs from 3dpf display the strongest strand bias (Fig.3A). A possible reason of the strong strand-bias of the 3dpf piRNAs may be due to Ziwi but not Zili being maternally provided while the zygotic Zili expression only starts at 3dpf. It is plausible that a large portion of the piRNAs at 3dpf is still inherited from the mother together with the maternal Ziwi proteins. In fact, we have observed maternal Ziwi protein in animals that were two weeks old (Houwing, unpublished data).

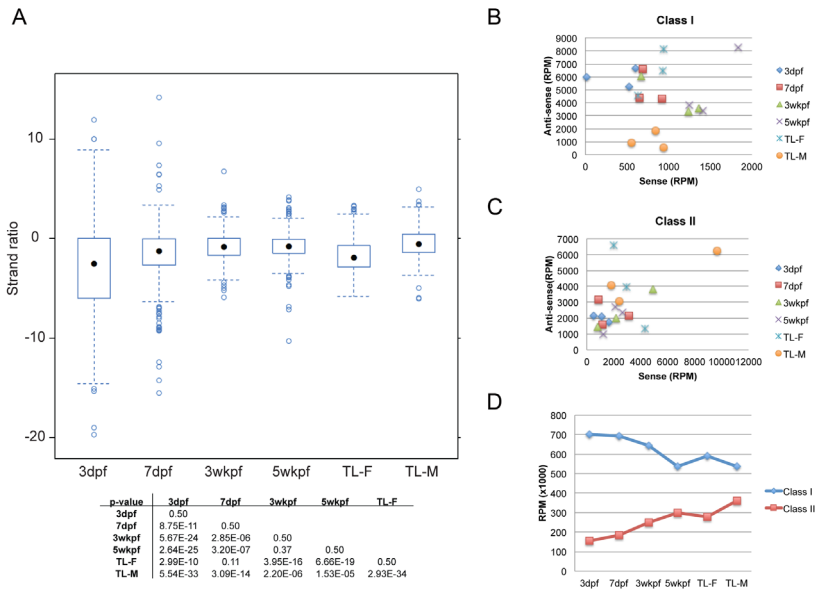
The strand bias gradually decreases during the development until 5wkpf stage. This holds true when we look at individual elements in class I but not class II transposons (Fig.3B-C), indicating that the strand bias is mainly caused by class I transposons. Since anti-sense piRNAs can target active TE mRNAs with complementary sequence, it implies class I elements are likely more actively transcribed during early development.

Retrotransposons are more active in early development

We decided to have a more detailed look at the population of the repeat piRNAs. To compare the bioinformatic analyses we obtained from different libraries, we normalized the total number of piRNA reads mapped to the repeat sequences in each library to one million reads. It is to be noted that prior to normalization some libraries only have 10-50% of total mapped reads compared to other libraries (Table1). Therefore the libraries with relatively low reads may display lower complexity.

Overall, class I element-derived piRNAs are much more abundant than class II element-derived piRNAs throughout the developmental stages we have examined so far (Fig.3D). This is consistent with earlier findings that most of the piRNAs are linked to retrotransposons (Houwing et al., 2007), and the observation that class II transposons either have been inactivated or tightly controlled in the vertebrate genome (Levin and Moran, 2011). Since piRNA populations have been proposed to react on transcripts of active transposon on a feed-forward amplification loop (Brennecke et al., 2007), our data strengthen the idea that vast majority of the class II elements are likely inactivated in the genome of zebrafish germline.

We noticed the population of piRNAs mapped to class I elements is the highest at 3dpf and it gradually goes down through 5wkpf (Fig.3D). Taken together with the observation that class I but not class II element-derived piRNAs have strong anti-sense strand bias at 3dpf, suggests that these class I element derived-piRNAs actually target active retrotransposon transcripts. It is possible that during the development or differentiation of the germ cells, retrotransposons gradually come under control and less retrotransposon-derived piRNAs are needed in the germline. This would also result in the balance between sense and anti-sense piRNA production, which is consistent with the observation that the strand bias of piRNAs is



p-value	3dpf	7dpf	3wkpf	5wkpf	TL-F
3dpf	0.50				
7dpf	8.75E-11	0.50			
3wkpf	5.67E-24	2.85E-06	0.50		
5wkpf	2.64E-25	3.20E-07	0.37	0.50	
TL-F	2.99E-10	0.11	3.95E-16	6.66E-19	0.50
TL-M	5.54E-33	3.09E-14	2.20E-06	1.53E-05	2.93E-34

Figure 3. Dynamic piRNA population during the development of germline. (A) Boxplot representation of strand bias of all the mapped piRNAs at different developmental stages. Y-axis indicates the ratio between the reads of sense and anti-sense stranded piRNAs derived from the same element. The more positive the value, the more sense stranded piRNA are. Each circle represents piRNAs mapped to an individual transposon element. P-values are calculated using student t-test. Three most abundant class I (B) and class II (C) transposons are selected. Each data point on the scattered plot indicates piRNAs mapped to one of the selected transposons. Different shapes of the data points indicate different developmental stages. The anti-sense reads (y-axis) and sense reads (x-axis) of the piRNAs mapped to each transposon are plotted against each other. The closer to the left-upper corner, the more anti-sense strand biased of the total piRNAs mapped to this transposon. (D) Total reads of piRNAs mapped to class I and class II transposons during the development are plotted in blue and red, respectively. For (B)-(D), RPM stands for reads per million.

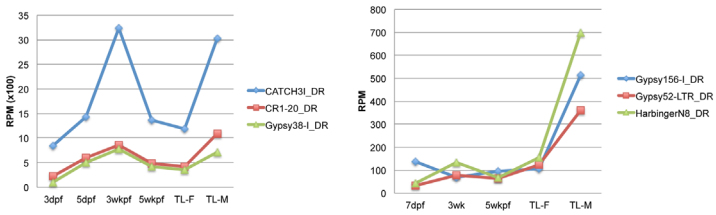


Figure 5 (Right). LoTSE-derived piRNAs are enriched at 3wkpf and adult testis stage. piRNA reads mapped to three representative LoTSEs, CATCH3I_DR, CR1-20_DR, Gypsy38-I_DR, are plotted. Y-axis shows the total reads of piRNAs mapped to these LoTSEs. RPM: reads per million.

Figure 4 (Left). 3dpf-absent element-derived piRNAs are enriched in adult testis. piRNA reads mapped to three representative 3dpf-absent elements, Gypsy156-I-DR, Gypsy52-LTR_DR and HarbingerN8_DR, are plotted. Y-axis shows the total reads of piRNAs mapped to these elements. RPM: reads per million.

gradually lost during the development. In adult ovaries, we detected more retrotransposon-derived piRNAs than in the testes. This may be explained by the need of maternally contributed piRNAs to secure early development of the germ cell lineage. Nevertheless, we cannot rule out a possibility that class II elements become more active during later development and that triggers a drop of class I element derived piRNAs.

When we considered the population of piRNAs derived from different transposon families, we observed that hAT, DNA and EnSpm families contribute to the major class II transposon-derived piRNA population. Gypsy and DIRS, two LTR type transposons, are the most abundant class I piRNA sources. Among all families of transposons, hAT (class II) and SINE (class I) derived piRNAs are most enriched at 5wkpf (Fig.S3A). Unlike most class I element-derived piRNAs, LINE- derived piRNAs are enriched in the adult testes (Fig.S3B). A recent study has reported that LINE elements are typically DNA methylated and (surprisingly) not particularly enriched in neither repressive (H3K9me3) nor activating (H3K14ac) marks on the chromatins of zebrafish sperm (Wu et al., 2011). On the other hand, Lehmann and colleagues (2011) showed that repressive marks such as H3K9me3, which initiate heterochromatin formation, are essential for the piRNA production in Drosophila (Rangan et al., 2011). Taken together the above, it suggests that piRNA production from the heterochromatic region is not very active in zebrafish sperm. Therefore LINE-derived piRNAs we detected in testes, if produced from piRNA clusters, are mainly contributed by spermatogenic cells before the spermatozoa stage.

Germ cells lack certain piRNAs at 3dpf

Even though piRNAs derived from LTR type transposons are the most abundant at 3dpf, we noticed that piRNAs derived from certain individual elements are completely absent at this stage. We refer to these elements as 3dpf-absent elements (Table S1). Most of them belong to the Gypsy and BEL families. piRNAs from the 3dpf-absent elements in all other developmental stages display similar strand bias as other elements, suggesting they can be bound by both Ziwi or Zili. It means that the absence of these piRNAs is not likely due to lack of (maternal) Zili. To investigate further the 3dpf-absent element, we checked their genomic origin by blasting the annotated sequences from RepBase in the UCSC genome browser. Six of the 3dpf-absent elements were manually examined and only the loci with piRNA reads are considered. In total, 216 out of 271 genomic loci displayed piRNA reads. These 216 loci are not clustered in specific regions (TableS2), therefore absence of piRNAs is not likely due to lack of transcription in certain genomic regions. Among these 216 loci, 105 are annotated within intronic regions (TableS2). Interestingly, there are often two, or in one case three, copies of the same element in tandem within one gene (marked in gray in TableS2). Intronic transposons can be co-transcribed with the embedded genes. However, whether the absence of 3dpf-absent element-derived piRNAs is due to lack of co-transcription of these genes at 3dpf stage awaits further investigation. Furthermore, we perceived that piRNAs derived from some of the 3dpf-absent elements (10 out of 16, 62.5%) are enriched in the adult testes relative to all other developmental stages, including ovaries (Fig.4).

It appears that embryos do not obtain these piRNAs from the mothers and these piRNAs are likely to be important during spermatogenesis.

LoTSEs reflect their abundance at 3wkpf

In one of the PIWI-piRNA pathway mutants, *tdrd1*^{fh244}, we have detected that piRNAs derived from certain transposon in the 3wkpf gonads dramatically decreased in numbers (at least 15 fold less than in the wild-type germ cells) and piRNAs from some other transposon are little effected by loss of Tdrd1 (reads ratio between the mutant and the wild-type is 1.5 to 1 fold) (Huang et al., 2011), hereby, we refer to these elements as “loss-of-Tdrd1 sensitive elements” (LoTSEs) and “loss-of-Tdrd1 resistant elements” (LoTREs, Table2), respectively. LoTSEs are all class I transposons while LoTREs are mostly class II transposons. We noticed that many of the LoTSE-derived piRNAs are relatively enriched at 3wkpf compared to other developmental stages (19 out of 23, 82.6%), while only 6.4% (3 out of 47) the LoTRE-derived piRNAs are enriched at 3wkpf. Since Tdrd1 functions to facilitate the piRNA amplification cycle (Huang et al., 2011), the relative high abundance of piRNAs derived from these elements may explain the strong reduction of their population in the *tdrd1* mutants. Interestingly, we discerned that there is a positive correlation between the abundance of certain LoTSE-derived piRNAs at 3wkpf and adult testes stages (15 out of 23, 65.2%). If the piRNAs derived from these LoTSEs are enriched at 3wkpf, they are also more abundant in the adult testes (Fig.5). In Drosophila and zebrafish, piRNAs display sex-specific genomic organizations (Li et al., 2009; Malone et al., 2009; Zhou et al., 2010). It was also suggested that piRNAs are involved in sexual development in zebrafish (Zhou et al., 2010). It is plausible that these LoTSE-derived piRNAs are involved in the transition from juvenile oocytes to the onset of the spermatogenesis in the 3wkpf gonads. Another possibility is that these transposons are more expressed in juvenile oocytes and spermatogenic cells and that the differential expression of the LoTSE-derived piRNAs simply reflects cell identities in the tissues used for library preparation.

piRNA populations differ between fish strains

When we compare the male and female libraries within different fish strains, we notice that TL adult pair and TU adult pair, although from the same stage, show distinct patterns. TL-F clearly has more piRNAs than TL-M (Fig.S1C), while TU-F and TU-M seems to have similar level of repeat piRNAs (Fig.S1D). Furthermore, the composition of class I and II derived-piRNAs seems to be different between strains. Both testis libraries have similar class I vs. class II ratio (about 1.5 fold), while TU-F displays higher ratio than TL-F (Fig.6A). It indicates that transposon activities may vary between different strains, which was observed in different Drosophila strains

Table2. Identities of LoTSEs and LoTREs

LoTSEs		LoTREs	
Element	Ratio	Element	Ratio
Gypsy42-I_DR	33.521	DNA-8-14_DR	1.507
Gypsy-116-I_DR*	28.800	DNA-2-8_DR	1.492
BEL8-I_DR*	23.678	DNA25TWA2_DR	1.489
Gypsy40-I_DR*	21.884	HATN6_DR	1.481
Gypsy38-I_DR*	21.507	EnSpm-N6_DR	1.478
Gypsy148-I_DR*	21.465	HATN8_DR	1.476
CATCH3I_DR*	21.382	HARBINGERN6_DR	1.463
Gypsy58-I_DR*	20.869	HATN10_DR	1.463
Gypsy124-I_DR*	20.041	Tc1N1_DR	1.435
Gypsy46-I_DR*	18.855	SINE3-1a	1.429
Gypsy36-I_DR*	17.764	LOOPERN6_DR	1.416
L1-10_DR*	17.757	MuDR-N1_DR	1.398
L1-4_DR	17.693	TDR8	1.397
Gypsy117-I_DR*	17.638	TDR7	1.395
Gypsy45-I_DR*	16.997	HARBINGERN2_DR	1.376
L1-7_DR*	16.952	LOOPERN4_DR	1.333
SUSHIIDR1	15.727	DNA-TA-2_DR	1.320
Gypsy-32-I_DR*	15.621	hAT-N30_DR	1.320
Gypsy62-I_DR	15.544	TDR15	1.314
CR1-20_DR*	15.479	TDR3	1.307
Gypsy-117-I_DR*	15.369	TDR13	1.306
Gypsy59-I_DR*	15.011	HATN3_DR	1.284
Gypsy164-I_DR*	15.000	DANA	1.272
		TDR5	1.267
		hAT-N66_DR	1.266
		DNA-X-4_DR	1.261
		HE1_DR1*	1.236
		ANGEL	1.229
		DNA9NNN1_DR	1.228
		LOOPERN1_DR*	1.219
		TDR20	1.218
		TDR18	1.195
		TC1DR2	1.183
		Kolobok-N1_DR	1.179
		DNA25TWA1_DR	1.172
		HARBINGER3N_DR	1.146
		TDR8B	1.145
		TDR23	1.143
		DNA15TA1_DR	1.140
		DNA11TA1_DR	1.137
		DNA2-2_DR*	1.127
		HATN12_DR	1.108
		hAT-N39_DR	1.066
		TDR2	1.041
		TC1DR3	1.039
		DNAX-1_DR	1.031
		Gypsy10-LTR_DR	0.908

Table2. Identities of LoTSEs and LoTREs. The name of each element and the ratio of the amount of piRNAs, which mapped to each element, in 3-week-old wild-type and *tdrd1* mutant gonads are listed. Asterisk indicates the enrichment of this transposon-derived piRNAs at 3wkpf.

as well (Brennecke et al., 2007). To understand which transposon families contribute to the difference between TU and TL, we plot the repeat piRNA composition based on their derived transposon class and family. Among class I transposons, the Gypsy family is ~10% higher in TU than in TL libraries (Fig.6B); L2 and I families are 5-9% higher in TL than in TU libraries; whereas other class I families seem to display similar percentages (less than 5% difference) between libraries. The composition of class II transposons is almost identical between TU and TL pairs (Fig.6C), suggesting these two fish strains have different levels of Gypsy transposon activities.

DISCUSSION AND CONCLUSION

The differences between the piRNAs at various developmental stages, suggests that piRNAs together with PIWI proteins regulate and play distinct roles during different stages of the germline development. According to our analyses, piRNAs from 3dpf stage have very distinct properties compared to piRNAs from later developmental stages. They are shorter, higher in retrotransposon content, from both the data of piRNA origin and the Ping-Pong signature, and are highly anti-sense oriented. It is therefore of great interest to investigate whether the main function of these 3dpf piRNAs is directly targeting active retrotransposons or if they are for other unknown purposes. Especially during 3dpf to 7dpf, Zili is present in the nucleus of the PGCs, which gives rise to the possibility that PGC piRNAs may be involved in chromatin modifications to regulate gene expression or TE transcription for normal development of the germline. piRNAs produced by certain transposon families are enriched in distinct stages, such as hAT and SINE in 5wkpf and LINE in adult testis.

One explanation is that these families of transposons are more active during these specific gonadal stages. Another explanation is that specific piRNA clusters are activated at these stages. The fact that none of these piRNAs display extreme strand ratios at the indicated stages, suggests that the later explanation is more in favor. Nevertheless, recent findings about SINEs functions in eukaryotic cells reveal that some SINEs regulate mRNA transcription: In some cases they block the Pol II machinery under heat shock stress. In others they serve as a boundary element and change the chromatin status to promote transcription in a developmental stage-specific manner during organogenesis (reviewed in Ponicsan et al., 2010). Therefore we cannot rule out the possibility that SINEs play similar roles during germline development.

In the analyses of comparing piRNA profiles between two different fish strains, we observed that all the elements detected are present in both TU and TL libraries. A phenomenon, called hybrid dysgenesis, describes crosses between two different strains of animals and lead to sterility of the off-springs. It was later explained by Brennecke et al (2008) that it is due to lack of certain maternal piRNAs, which can target paternal inherited transposons. We do not observe sterility defects in the progenies of crosses between TU and TL animals, which is consistent with the above results.

Recently Han and colleagues (2011) showed that piRNAs from three different stages of germ cells during mouse spermatogenesis have distinct properties and origins: the percentage of piRNA population increases from mitotic to post-meiotic stage; piRNA length is shorter in mitotic and meiotic stage germ cells than in post-meiotic stage. Another recent study on retrotransposon-derived piRNA profiles in developing mouse testis also showed that during the development of the testis, piRNA origin shifts from active retrotransposon transcripts to intronic retrotransposons that co-transcribed with mRNAs (Mourier, 2011). In our analyses, all types of spermatogenic cells were pooled together therefore we were not able to verify whether the above observations are also conserved in zebrafish. Further studies on the differences of piRNA profiles from purified cells with specific differentiation stages will bring more answers and insights of the possible divergent as well as conserved functions of the PIWI-piRNA pathway during the gonadal development.

MATERIAL AND METHODS

Zebrafish strains and genetics

Zebrafish were kept under standard conditions. Wildtype fish with either TL or TU in combination with vas::EGFP transgene (Krovel and Olsen 2004) background are used.

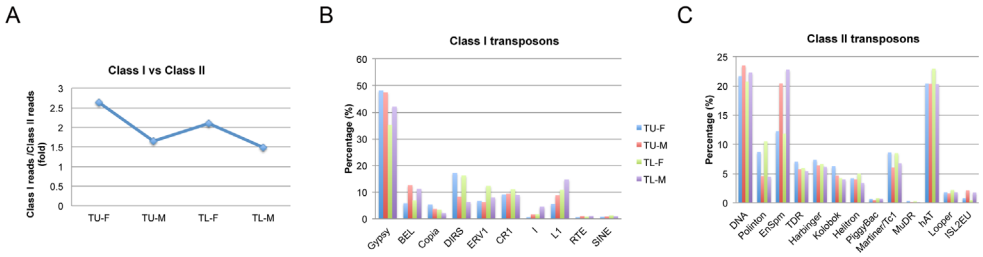


Figure 6. Distinct piRNA compositions between TU and TL fish strains. (A) The ratio between the reads of piRNAs mapped to class I and class II transposons within individual library is plotted. Percentages of piRNAs mapped to various transposon families of class I (B) and class II (C) within one library are plotted.

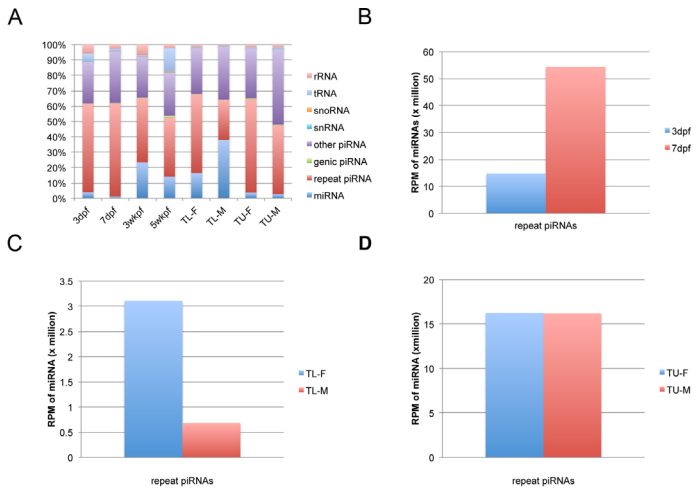


Figure S1. Small RNA composition changes during development. (A) Small RNA types considered are plotted according to their percentage in each library. Repeat piRNA and other piRNA reads in 3dpf and 7dpf (B), TL-F and TL-M (C) and TU-F and TU-M (D) libraries are compared after normalizing by setting miRNA reads in each library to 1 million.

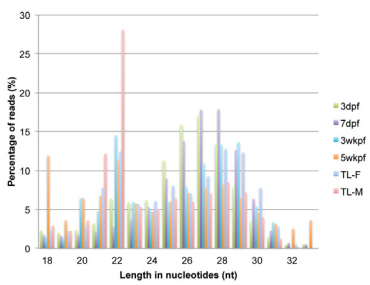


Figure S2. Length distribution of small RNAs in all libraries. The percentage of mapped small RNA reads (y-axis) is plotted against the length of the small RNA in nucleotides (x-axis) in bar diagrams. Each developmental stage is indicated in different colors.

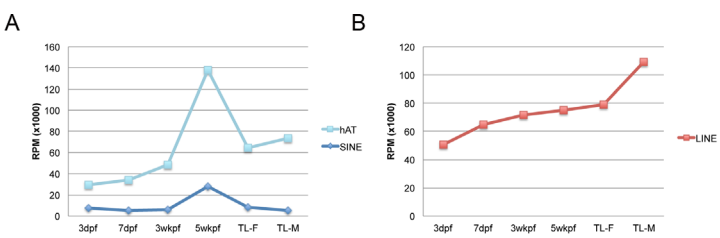


Figure S3. Differential expression of piRNAs during development. piRNAs mapped to certain transposon families are illustrated (A) hAT and SINE families. (B) LINE family. RPM: reads per million.

TableS1. 3dpf-absent elements	
Class I	Gypsy156-I_DR*#
	Gypsy67-I_DR*#
	BEL-38-LTR_DR#
	Gypsy36-LTR_DR#
	Gypsy-14-LTR_DR*#
	Gypsy52-LTR_DR*#
	BEL9-LTR_DR*
	CR1-15_DR*
	Gypsy102-LTR_DR
	BEL18-LTR_DR*
	BEL4-LTR_DR
	Gypsy-19-I_DR*
	Copia4-LTR_DR
Class II	HarbingerN8_DR*

TableS1. Identities of 3dpf-absent elements. The names of the elements are listed according to their belonging class (class I or class II). Asterisk indicates that piRNAs mapped to this element are enriched at adult testis stage. # indicates the genomic origin of the elements have been checked and described in TableS2.

Table S2. Genomic loci of some 3dpf-absent elements

Name	Identity	Genomic location		Intronic?
		Start	End	
Gypsy156-I_DR	100%	28831840	28838250	No
	96.20%	60601111	60601136	within hspa12b
	100%	30336158	30340193	within zgc:174708
Gypsy67-I_DR	88.10%	9256884	9257314	within arhgef9
	87.30%	9338233	9338663	within arhgef9
	80.50%	12334022	12334377	within zgc:165409
	80.30%	12512098	12512442	within zgc:165409
	72.10%	40652964	40653289	within psmd1
BEL-38-LTR_DR	98.10%	50802207	50802518	within zgc:173617
	97.50%	6857910	6858218	within ech1
	97.50%	6864770	6865078	within ech1
	97.10%	9497752	9498060	within NPAS3
	97.10%	9501104	9501412	within NPAS3
	96.50%	44473709	44474015	within atg10
Gypsy36-LTR_DR	97.20%	63603535	63603920	within SAT2
	96.90%	63609277	63609661	within SAT2
	94.30%	50757606	50757981	within si:dkey-51d8.1
	94.70%	50751260	50751634	within si:dkey-51d8.1
	95.70%	14610080	14610463	within C20H1orf9
	97.20%	49098925	49105651	within si:ch211-246b8.1
	95.40%	49105277	49105512	within si:ch211-246b8.1
Gypsy52-LTR_DR	100%	21700465	21701937	within AL773596.3
	100%	21707412	21708884	within AL773596.3
	98.10%	16504208	16505670	within si:dkeyp-116h7.2
	98.10%	16497269	16498731	within si:dkeyp-116h7.2
	97.50%	21598201	21599644	within drd3
	97.40%	21605121	21606564	within drd3
Gypsy-14-LTR_DI	97.70%	60173183	60174443	within cyyr1
	99.00%	1764767	1766062	within si:ch1073-109g21.1

TableS2. Genomic origins of some 3dpf-absent elements. The genomic locations, whether they are located in the intronic region of genes, and the associated gene names are indicated in the table.

Isolation of adult gonads

Adult animals were sacrificed in icy water. Gonads of these animals were isolated under standard dissection microscope and immediately stored in Trizol-LS (Invitrogen) for RNA isolation according to the manufacturer's protocol.

Isolation of 3wkpf and 5wkpf gonads

3wkpf and 5wkpf animals with vas::EGFP transgene were sacrificed in icy water and kept on ice prior dissection. Gonads of these animals were isolated under M2Bio GFP microscope and immediately stored in Trizol-LS (Invitrogen) for RNA isolation according to the manufacturer's protocol.

FACS of 3dpf and 7dpf PGCs

3dpf and 7dpf embryos with vas::EGFP transgene were collected sacrificed in icy water. Whole embryos were treated with TrypLE Express (GIBCO) for 60 minutes at 28 °C to dissociate the cells. Fetal calf serum was added to 20% to stop the reaction. Cells were spun down at 1500rpm for 1 minute and washed once with PBS. Resuspended cells were kept in PBS on ice before sorting. The cell suspension was further filtered and sorted using a MoFlo Fluorescence Activated Cell Sorting machine (Dakocytomatino). Total RNA was isolated from sorted high GFP cells.

Library preparation

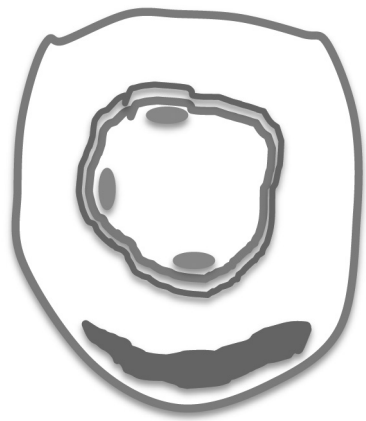
For 3dpf, 7dpf, 3wkpf, TL-F, TL-M, TU-F and TU-M libraries: Total RNA was size selected (18-32 nt); subsequently both 3' and 5' adapters were ligated. The ligated RNA was reverse transcribed into cDNA. This cDNA was amplified with 16-18 cycles of PCR and purified using the Macherey & Nagel NucleoSpin Extract II kit. The cDNA was eventually sent for a 50 cycle sequencing run on an Illumina/Solexa platform.

For 5wkpf library: Total RNA was poly(A)-tailed using poly(A) polymerase followed by ligation of a RNA adaptor to the 5'-phosphate of the RNAs. First-strand cDNA synthesis was performed using an oligo(dT)-linker primer and M-MLV-RNase H- reverse transcriptase. The cDNA was PCR-amplified for 16 cycles according to the instructions of Illumina/Solexa.

Sequence analysis

Bioinformatic analysis was performed as previously described (Huang et al., 2011).

Chapter 3



Tdrd1 associates with piRNA targets and Piwi proteins in zebrafish

3

Adopted from:
Huang HY, Houwing S#, Kaaij LJT#, Meppelink A, Redl S, Gauci S, Vos H, Draper BW, Moens CB, Burgering BM, Ladurner P, Krijgsveld J, Berezikov E and Ketting RF.
EMBO J. 2011 Jul 8;30(16):3298-308.

# contributed equally

Tdrd1 associates with piRNA targets and Piwi proteins in zebrafish

Hsin-Yi Huang[1], Saskia Houwing[1,2,#], Lucas J.T. Kaaij[1,#], Amanda Meppelink[1], Stefan Redl[3], Sharon Gauci[4], Harmjan Vos[5], Bruce W. Draper[6], Cecilia B. Moens[7], Boudewijn M. Burgering[5], Peter Ladurner[3], Jeroen Krijgsveld[4,8], Eugene Berezikov[1] and René F. Ketting[1,*]

[1]Hubrecht Institute-KNAW & University Medical Centre Utrecht, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands

[2]Present address: Kimmel Center for Biology and Medicine at the Skirball Institute,
NYU School of Medicine, 540 First Avenue, New York, NY 10016, USA

[3]Institute of Zoology, Technikerstrasse 25, A-6020 Innsbruck, Austria

[4]Biomolecular Mass Spectrometry and Proteomics Group, Bijvoet Centre for Biomolecular Research, Utrecht Institute for Pharmaceutical Sciences and Netherlands Proteomics Center, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands

[5]Department of Physiological Chemistry, University Medical Center Utrecht, Utrecht, The Netherlands

[6]Molecular and Cellular Biology, University of California, Davis, One Shields Avenue, Davis, CA 95616, USA

[7]Howard Hughes Medical Institute and Fred Hutchinson Cancer Research Center, P.O. Box 19024, Seattle, WA 98109, USA

[8]Present address: EMBL, Genome Biology Unit, Meyerhofstrasse 1, 69117, Heidelberg, Germany.

[#] Contributed equally to this work

[*] To whom correspondence should be addressed. E-mail: r.ketting@hubrecht.eu

ABSTRACT

Piwi proteins function in an RNAi-like pathway that silences transposons. Piwi-associated RNAs, also known as piRNAs, act as a guide to identify Piwi targets. The tudor-domain-containing protein Tdrd1 has been linked to this pathway but its function has thus far remained unclear. We show that zebrafish Tdrd1 is required for efficient Piwi-pathway activity and proper nuage formation. Furthermore, we find that Tdrd1 binds both zebrafish Piwi proteins, Ziwi and Zili, and reveal sequence specificity in the interaction between Tdrd1 tudor domains and symmetrically dimethylated arginines (sDMAs) in Zili. Finally, we show that Tdrd1 associates with long RNA molecules. These Tdrd1 associated transcripts (TATs) likely represent cleaved Piwi pathway targets and may serve as piRNA biogenesis intermediates. Altogether, our data suggest that Tdrd1 acts as a molecular scaffold for Piwi-pathway components, comprising Piwi proteins, their piRNA cofactors and their targets, and provides specificity through interactions between specific sDMAs in Piwi proteins and Tdrd1.

Key words
Tdrd1, Piwi, zebrafish, piRNA, piRNA target, nuage

INTRODUCTION

In gonads of animals, the activity of a specialized RNAi-like pathway is essential for reproduction. This pathway is characterized by a subclass of the Argonaute protein family known as the Piwi subfamily (Carmell et al, 2002). It is required for efficient transposon silencing, germ cell differentiation and meiosis (Aravin et al, 2001; Brennecke et al, 2007; Carmell et al, 2007; Cox et al, 1998; Cox et al, 2000; Deng & Lin, 2002; Ghildiyal & Zamore, 2009; Houwing et al, 2008; Houwing et al, 2007; Kuramochi-Miyagawa et al, 2004). The Piwi pathway operates in distinct subcellular compartments, with nuage -an electron dense, germ cell-specific structure associated with mitochondria- functioning as a major site of Piwi protein activity (Aravin et al, 2009; Harris & Macdonald, 2001; Li et al, 2009; Lim & Kai, 2007; Lim et al, 2009; Malone et al, 2009; Pane et al, 2007; Patil & Kai, 2010; Shoji et al, 2009). Nuage is known to associate with clusters of nuclear pores (Pitt et al, 2000), suggesting that nuage has a function in the trafficking of molecules between the nucleus and the cytoplasm. Interestingly, nuage-like structures named Yb bodies have been described in *Drosophila* somatic nurse cells surrounding the oocyte (Szakmary et al, 2009) and these bodies have been implicated in Piwi-pathway activity as well, in particular the loading of *Drosophila* Piwi with piRNA and its subsequent transit into the nucleus (Olivieri et al, 2010; Qi et al, 2011; Saito et al, 2010).

Small RNAs named piRNAs, and sometimes rasiRNAs in *Drosophila*,

guide Piwi proteins to their targets (Aravin et al, 2006; Girard et al, 2006; Lau et al, 2006; Saito et al, 2006). Two types of piRNAs have been identified: primary and secondary (Li et al, 2009; Malone et al, 2009). Primary piRNAs are characterized by uracil at their 5' end. They are likely derived from long single stranded transcripts through cleavage by an unknown nuclease, after which the 5'-end of the 3'-cleavage product is bound by a Piwi protein followed by 3'-end processing (Horwich et al, 2007; Houwing et al, 2007; Kirino & Mourelatos, 2007; Saito et al, 2007; Vagin et al, 2006). Primary piRNAs can trigger the generation of secondary piRNAs through cleavage of single stranded transcripts by the Piwi proteins themselves (Brennecke et al, 2007; Gunawardane et al, 2007; Li et al, 2009). This step defines the 5' end of a novel piRNA, which is then again presumably followed by 3' end trimming. Consequently, secondary piRNAs have opposite polarity compared to primary piRNAs and are characterized by adenosine at position ten. In turn, if complementary target RNA molecules are present, secondary piRNAs may trigger generation of additional primary-like piRNAs. Because of the reciprocal interactions involved, this mode of piRNA biogenesis has been dubbed 'ping-pong' (Brennecke et al, 2007).

Additional components of these mechanisms have been identified. Many of these interactors (Tdrd proteins) contain so-called Tudor domains (Chen et al, 2009; Kirino et al, 2009a; Liu et al, 2011; Nishida et al, 2009; Patil & Kai, 2010; Reuter et al, 2009; Vagin et al, 2009; Vasileva et al, 2009; Wang et al, 2009; Yabuta et al, 2011). It was shown for tudor domains of the p100 and Tdrd3 proteins that they can bind peptides containing symmetrically dimethylated arginine (sDMA) (Cote & Richard, 2005) and subsequent structural studies have revealed the physical basis behind these interactions (Friberg et al, 2009; Liu et al, 2010a; Liu et al, 2010b). Indeed, sDMAs have been identified in the N-terminal tails of Piwi proteins and in other Piwi-pathway components like Vasa (Kirino et al, 2010), and these mediate interaction with tudor-domain containing proteins (Chen et al, 2009; Kirino et al, 2009a; Nishida et al, 2009; Reuter et al, 2009; Vagin et al, 2009). Interestingly, multiple sDMA sites in Piwi proteins have been described and many Tdrd proteins contain multiple Tudor domains that potentially can bind multiple peptides simultaneously (Liu et al, 2010a). It is thus far, however, not clear whether tudor-Piwi interactions display sequence specificity. Finally, Nishida *et al* (2009) reported on long RNA molecules associated with the *Drosophila* multi-tudor domain protein Tud and suggested that these may be piRNA precursors, but a detailed molecular analysis supporting this notion was not provided.

Mutants in Tdrd protein encoding genes have been described in both mice and *Drosophila*. In mice, mutants lacking the tudor-domain-containing proteins Tdrd1, Tdrd5, Tdrd6 and Tdrd9 result in male sterility (Chuma et al, 2006; Shoji et al, 2009; Vasileva et al, 2009; Wang et al, 2009; Yabuta et al,

2011). In addition, relatively weak effects on piRNA production in *tdrd1* and *tdrd9* mutants have been reported (Reuter et al, 2009; Shoji et al, 2009; Vagin et al, 2009) demonstrating a functional role of these proteins in the Piwi-pathway. In *Drosophila*, mutations in Piwi-pathway associated *tdrd* genes have defects in fertility, transposon silencing and subcellular localization of Piwi-pathway components, again revealing a functional relevance of these Piwi-interacting proteins (Aravin et al, 2001; Kennerdell et al, 2002; Kirino et al, 2009b; Liu et al, 2011; Nishida et al, 2009; Patil & Kai, 2010). However, it is still far from clear what most Tdrd proteins do to support Piwi protein activity.

Zebrafish germ cells express two Piwi proteins, Ziwi and Zili (Houwing et al, 2008; Houwing et al, 2007). Ziwi predominantly binds primary(-like) piRNAs, which are anti-sense with respect to targeted transposons, whereas Zili mostly associates with sense, secondary piRNAs. Both Piwi proteins are required for normal germ cell development. Here we demonstrate that zebrafish Tdrd1 is an interactor of both Ziwi and Zili. It is required for nuage formation and *tdrd1* mutation results in mild delocalization of Zili. Other Piwi-pathway-related defects such as transposon desilencing and germ cell defects are observed as well in *tdrd1* mutants, consistent with results obtained in mice. We then go on to reveal sequence specificity in the Tudor-domain-Zili interaction and identify specific arginines that, when symmetrically dimethylated, bind to Tdrd1. Finally, we describe Tdrd1-associated long RNA molecules that display all the characteristics of -thus far undetected- piRNA biogenesis intermediates. Taken together, we demonstrate that sequence specificity is a relevant factor in assessing Tdrd-Piwi interactions and that Tdrd1 interacts with multiple RNA and protein components of the Piwi-pathway. This, for the first time, provides direct support for the idea that Tdrd proteins may function as molecular scaffolds in this pathway.

RESULTS

Zili interacting proteins

We performed mass spectrometry on Zili immuno-precipitates (IPs) (TableS1). Apart from Zili, we mainly enriched for tudor domain-containing proteins, in line with previously published findings regarding Piwi protein interactors (Chen et al, 2009; Kirino et al, 2009b; Nishida et al, 2009; Reuter et al, 2009; Siomi et al, 2010; Vagin et al, 2009; Wang et al, 2009). We focused our analysis on Tdrd1, a strongly conserved protein with four Tudor domains and an N-terminal MYND domain that has been implicated in nuage formation, germ-cell differentiation and Piwi pathway activity in mice (Chuma et al, 2006; Hosokawa et al, 2007; Reuter et al, 2009; Vagin et al, 2009; Wang et al, 2009). We raised an antibody against zebrafish Tdrd1 that recognizes a gonad-specific protein of roughly the expected molecular weight of 130kDa (Fig 1A). Immunostainings using this antibody reveal germ

cell specific signals in wild-type but not in *tdrd1* mutant animals (see below). Mass spectrometry on Tdrd1 IPs identified both Ziwi and Zili but also Vasa and the tudor-domain-containing proteins Tdrd4 and Tdrd5 (Table S2).

Tdrd1 expression

Using Western blotting analysis we can detect Tdrd1 expression in testis and immature oocytes (Fig. 1A and S1A-C). Tdrd1 signal in total ovary, mostly consisting of large, mature oocytes is often not detectable (Fig. 1A), most likely due to the fact that maturing oocytes accumulate high amounts of yolk proteins, diluting the Tdrd1 signal. Consistent with this, we only detect strong Tdrd1 signals in oocytes of stage II and younger in immuno-histochemistry experiments (Fig. 1B). During development, Tdrd1 protein becomes detectable in the primordial germ cells (PGCs) at four days post fertilization (4dpf; Fig. 1C and S1D). At this stage, Tdrd1 is cytoplasmic, displaying fine granular distribution (Fig. 1B) distinct from the nuclear localization of Zili at 4dpf (Houwing et al, 2008). At 5dpf Tdrd1 starts to assemble into larger perinuclear granules (Fig. S1D), much like we have described for Zili (Houwing et al, 2008). At later time points, Tdrd1 and Zili display very similar sub-cellular localization, most prominently in nuage-like structures in early stages of both oocyte and sperm development (Fig. 1B,D and S1C). Using immunogold electron microscopy (EM) we confirmed that indeed both Zili and Tdrd1 localize to electron-dense structures (Fig. S1E and S2A). This localization profile closely resembles that described for mouse Tdrd1 (Aravin et al, 2009; Chuma et al, 2006).

Loss of Tdrd1 affects nuage and Zili localization

We obtained a *tdrd1* mutant allele, *fh224*, harboring a premature stop codon truncating Tdrd1 between the MYND- and the first Tudor domain. Tdrd1 immunohistochemistry on *tdrd1* mutant cells shows no detectable staining (Fig.1B), indicating that *fh224* represents a strong loss of function allele, perhaps a null allele. *Tdrd1* homozygous mutants are viable but loose their germ cells during development (Fig. 2A and S2C), similar to the *ziwi* and *zili* loss-of-function phenotypes (Houwing et al, 2008; Houwing et al, 2007). However, in contrast to *ziwi* and *zili* mutants, *tdrd1* mutants contain differentiating germ cells at three weeks of age (Fig. 2A and S2C). This indicates that the *tdrd1* mutant phenotype is relatively weak. The zebrafish genome (version Zv9) only encodes one *tdrd1* gene ruling out redundancy through gene duplication. Although we cannot exclude the possibility that other Tdrd proteins can partially substitute for Tdrd1 we prefer the interpretation that Tdrd1 function is not absolutely required for Piwi-pathway activity but rather potentiates it.

Next, we determined the sub-cellular localization of Tdrd1 and Zili in each other's absence. Zili remains nuclear in *tdrd1* mutant PGCs and

Tdrd1 remains present in fine cytoplasmic granules in absence of Zili (Fig. 2B,C). At later stages, when most Zili is found in nuage, we observe that Zili immunostaining becomes less focused around the nucleus and spreads into the cytoplasm in *tdrd1* mutants (Fig. 2D and FigS2D-E). However, perinuclear Zili-positive granules are still present. We also used transmission EM to analyze nuage in more detail in wild-type and *tdrd1* mutant oocytes. In general, *tdrd1* mutant oocytes contain only few nuage-like patches compared to wild-type, and these patches are much less electron dense (Fig. 2E). These structures are no longer tightly associated with mitochondria but their overall association with the nuclear envelope is less affected (Fig. S3). Immunogold labelling showed that Zili is no longer detectable in the structures that persist in absence of Tdrd1 (Fig. S2B). This implies that the peri-nuclear localization of Piwi proteins observed in immuno-fluorescence experiments does not reflect nuage as defined by EM and that other mechanisms are in place to steer Piwi proteins towards peri-nuclear regions.

We noted in the TEM experiments that nuage in oocytes associates with clusters of nuclear pores. This resembles findings in *C. elegans*, where P-granules have been found associated with nuclear pore clusters (Pitt et al, 2000). Interestingly, in *tdrd1* mutant oocytes we find significantly more solitary pores and less extended aggregates (Fig. 2F), suggesting that Tdrd1 is one of the factors driving the clustering of nuclear pores in zebrafish oocytes.

Tdrd1 affects piRNA populations

Mouse Tdrd1 has been described to have effects on ping-pong efficiency (Vagin et al, 2009) and to filter RNA molecules entering the ping-pong cycle (Reuter et al, 2009). We probed the effect of zebrafish Tdrd1 on piRNAs to see if these effects can be seen in the zebrafish. We made small RNA libraries from three-week-old wild-type and mutant gonads, a stage in which *tdrd1* mutant gonad morphology is relatively normal, and analyzed these using deep sequencing (TableS3). First, assuming miRNA levels remain unchanged, a mild decrease in piRNA abundance was observed (Fig. S4A). In contrast to what has been described in mice (Reuter et al, 2009), we do not observe an increase in gene-derived, or other RNA species-derived piRNAs (TableS3). We then looked at transposon-derived piRNAs specifically, and noted an especially strong effect of *tdrd1* mutation on retro-transposon piRNAs (Fig. 3A). This may derive from the fact that retrotransposons, while representing a minority of all transposable elements in the zebrafish genome, produce most piRNAs (Houwing et al, 2007), indicating that the majority of Piwi activity is directed against retro-transposons. For some elements we observed a change in strand bias of the piRNAs (Fig.S4B), and in those cases the class that is mostly bound by Ziwi (Houwing et al, 2008) was reduced. Next, we quantified transcript levels of five individual transposons using q-RT-PCR on RNA isolated from FACS

purified germ cells. This revealed that most elements tested are upregulated in the *tdrd1* mutant (Fig. 3B), showing that in *tdrd1* mutants transposon silencing activity is reduced. The extent of transcript upregulation and piRNA decrease for these elements match rather well (Fig. S5), although we cannot make general statements regarding such correlation based on five individual elements. Finally, we analyzed Piwi-cleavage activity by looking at the so-called ping-pong signature, the characteristic signal of 10-nucleotide sense-antisense 5'-overlaps between individual piRNAs. We observed a very low ping-pong signature in *tdrd1* mutants (Fig. S4C). However, when we correct for the amount of overlapping reads in total, a wild-type signature is evident (Fig. S4D). Taken together, these data suggest that in absence of Tdrd1, Piwi proteins can still cleave their targets, but do so less frequently.

Tdrd1-Zili binding characteristics

Co-IP experiments using zebrafish Tdrd1-specific antibodies confirmed that Tdrd1 binds, directly or indirectly, to Zili (Fig. 4A). In mouse, Tdrd1 has been suggested to bind exclusively to Mili (Reuter et al, 2009) but also to be less specific in its choice for specific Piwi paralogs, i.e. binding Mili, Miwi and Miwi2 (Chen et al, 2009; Vagin et al, 2009). In zebrafish, Tdrd1 clearly interacts with both Zili and Ziwi (Fig. 4A), although little Tdrd1 is detected in Ziwi IPs. The latter may be due to Tdrd1 shielding the antibody-epitope on Ziwi or to an abundance of non-Tdrd1-bound Ziwi.

Binding of Tudor domains to Piwi proteins has been shown to require sDMA residues in the N-terminal tails of the Piwi proteins (Kirino et al, 2009a; Nishida et al, 2009; Reuter et al, 2009; Vagin et al, 2009). Using arginine methylation specific antibodies we found that also Zili contains this type of arginine modification (Fig. 4B). Using mass spectrometry on immuno-purified Zili protein, we identified four methylation sites in the N-terminal region of Zili (R68, 163, 221 and 228) (Fig. S6 and data not shown). We tested whether these four sites can mediate Tdrd1 binding by performing pull-down experiments in testis extracts, using synthetic biotinylated peptides as bait. We found that Tdrd1 binds most strongly to peptides containing dimethylated R221 or R228 (Fig. 4C). Peptides with unmethylated arginines (Fig. 4C) and of unrelated sequence but containing sDMA (Fig. 4C) do not interact with Tdrd1. These interactions are RNA-independent as they are performed in the presence of RNaseA. We also tested which of the Tdrd1 tudor domains can bind Zili through GST-pull-down experiments using testis extracts. This revealed that the most C-terminal tudor-domain of Tdrd1 most strongly associates with Zili (Fig. 4D). This differs from observations on Tdrd1-Mili interactions in a heterologous setting that implicated the N-terminal half of Tdrd1 in Mili binding (Wang et al, 2009).

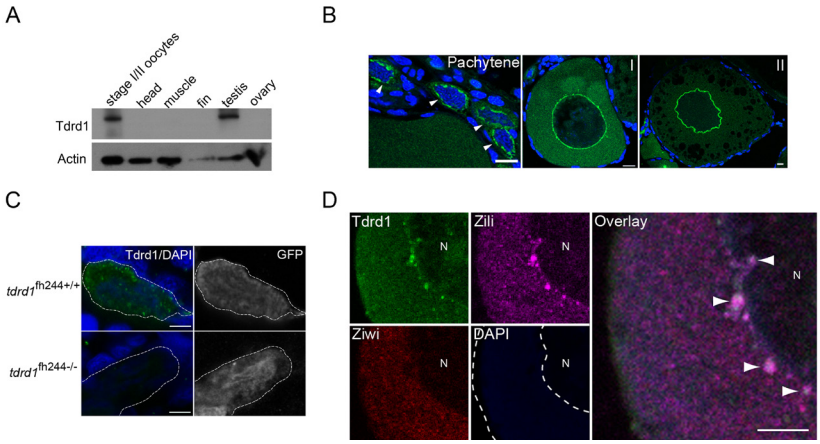


Figure 1. Tdrd1 expression in zebrafish. (**A**) Western blot analysis of Tdrd1 in diverse tissues. (**B**) Immuno-histochemistry on ovary. Arrow-heads indicate pachytene stage oogonia. Scale bars are 10µm. I, II: oocyte stage. (**C**) Immunohisto-chemistry on PGCs in 4 days post fertilisation (dpf) embryos. Vas::EGFP expression marks PGCs. Green: Tdrd1. Blue: DAPI. Scale bars are 10µm. Dashed line outlines a single PGC. (**D**) Co-staining of Ziwi, Zili and Tdrd1 in stage I oocyte. N: nucleus. Arrow-heads: sites of colocalization. Scale bar 5µm. Green: Tdrd1. Red: Ziwi. Magenta: Zili.

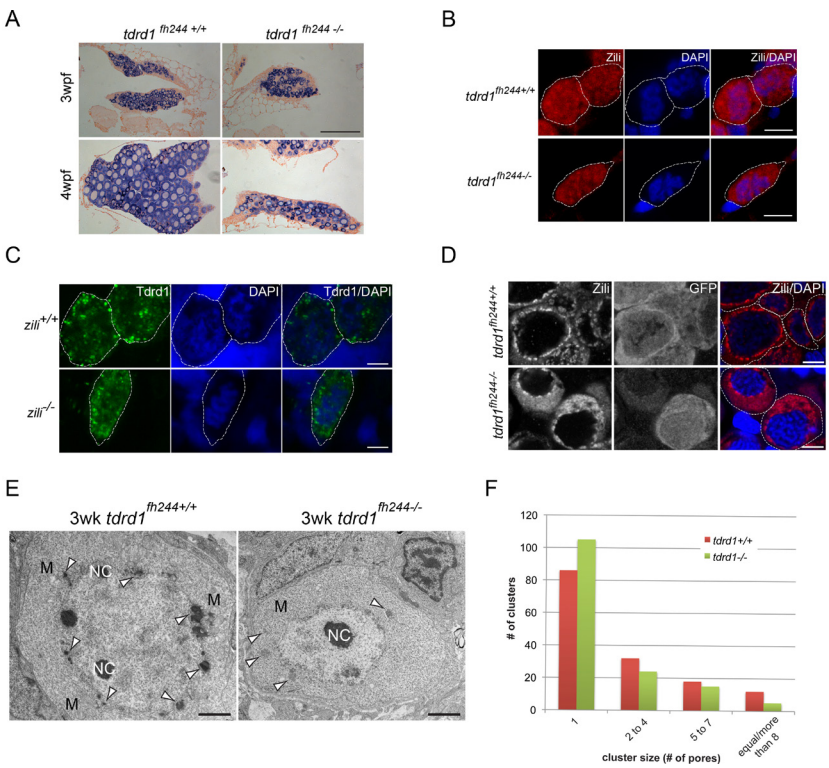


Figure 2. *tdrd1* mutant analysis. (**A**) *Vasa* ISH (purple) on gonads from wild-type and *tdrd1* mutant females at 3 and 4 weeks post fertilization (wpf). Scale-bar: 100µm. (**B**) Zili immuno-histochemistry (red) in wild-type and *tdrd1* PGCs (5dpf). Blue: DAPI. Dashed line outlines individual PGCs. (**C**) Tdrd1 immuno-histochemistry (green) in wild-type and *zili* mutant PGCs (5dpf). Blue: DAPI. Dashed line outlines individual PGCs. (**D**) Zili staining on wild-type and *tdrd1* oocytes at 3 wpf. Scale-bars B-D: 5µm. (**E**) Transmission

electron microscopy on 3 wpf ovaries from wild-type and *tdrd1* mutant animals. Mitochondria (M), nuage (arrow-heads) and nucleoli (NC) are labeled. Scale-bar: 1µm. (**F**) Cluster sizes of nuclear pores were determined from TEM images in wild-type and *tdrd1* mutant ovaries of 3wpf. Y-axis shows how often each cluster size was found. X-axis shows cluster size. ($p<0.001$; Chi-squared test)

Tdrd1 associated piRNAs

We next analyzed the RNA molecules associated with Tdrd1 in zebrafish testis. As can be expected, we found RNA species that resemble mature piRNAs in Tdrd1 IPs (Fig 5A). We size-selected these small RNAs and sequenced them. As comparison we sequenced testis-derived piRNAs from Ziwi and Zili IPs and piRNAs from total testis extracts. The observed length profiles from Ziwi and Zili IPs correlate well with our previously published results (Houwing et al, 2008). Interestingly, the Tdrd1 bound piRNA length profile is intermediate between the Ziwi and Zili profiles (Fig. 5B), suggesting roughly equimolar amounts of Ziwi- and Zili-bound piRNAs in Tdrd1 IPs. We further analyzed transposon derived piRNA sequences in terms of their strand biases. As described before (Houwing et al, 2008) for many repeats (287 out of 866) Ziwi and Zili enrich for anti-sense and sense piRNAs respectively, but we also detect a set of transposons in which this bias is inverted (97 out of 866). The Tdrd1-bound piRNAs from these two sets of transposons (producing 50% of all transposon piRNAs) display intermediate strand biases, consistent with a mix of Ziwi and Zili-bound piRNAs (Fig. 5C) in Tdrd1 IPs.

One potential function for Tdrd1 could be to select specific transcripts for entry into the piRNA amplification cycle. Such a function has been proposed for mouse Tdrd1 (Reuter et al, 2009). We therefor checked whether the piRNAs retrieved by Tdrd1 IP display the same transposon profile as we find in the total Ziwi and Zili IPs (Fig. 5D). This revealed a very strong correlation between piRNA obtained from both Tdrd1- and Ziwi IPs and Tdrd1- and Zili IPs, strongly suggesting that Tdrd1 does not play a significant role in target discrimination, at least not between different transposon substrates.

Tdrd1 associated transcripts

Next to piRNAs we also detect long RNA molecules in Tdrd1 IPs, further referred to as Tdrd1-associated-transcripts, or TATs) (Fig. 5A). We cloned these longer species from Tdrd1 IP samples without size selection or enzymatic pre-treatments. They likely lack 2'-OMe modifications since we cloned almost exclusively long RNA molecules from the Tdrd1 IP while piRNAs are present at similar levels (Fig. 5A). Since we do not treat the RNA with either phosphatases or kinases, the majority of TATs likely contain 5'P and 3'OH termini. After cloning and sequencing the first 46 nucleotides from their 5' end, we find that loci producing TATs tend to also produce piRNAs, and abundances of both RNA types correlate well (Fig. 6A and S7A). Overall, TATs, like the Tdrd1-associated piRNAs, display a roughly 1:1 ratio of sense versus antisense polarities (Fig.S7B). Interestingly, at the level of individual

transposons we find that the strand ratios of Tdrd1 associated piRNAs anti-correlate with the strand ratios found in TATs (Fig. 6B), suggesting a guide-target-RNA interaction between the two populations.

To follow up this observation, we asked whether TATs displayed a ping-pong signal when compared with the transposon-derived piRNAs obtained from Tdrd1 IPs. A signal of 10 bases overlap was indeed observed (Fig. 6C). Many overlaps also covered the complete piRNA length, possibly reflecting uncleaved targets or the fact that our libraries are not saturated (TableS3). We also probed the distances between TAT and piRNA 5' ends. This revealed that most TATs contain a 5' end that is identical to that of a mature piRNA (Fig. 6D). Together, these data suggest that TATs represent cleaved piRNA target fragments that may be converted into mature piRNAs through 3' end trimming activities.

DISCUSSION

Collectively, our data provide a model for the molecular function of the vertebrate Piwi-pathway component Tdrd1. We show that Tdrd1 binds both RNA and protein components of the Piwi pathway and that in absence of Tdrd1 the Piwi pathway operates with a much lower efficiency. This results in partial piRNA target activation and defects in germ cell development. In addition, we identify RNA transcripts that display hallmarks of piRNA targets and/or cleavage products. Below we further discuss both protein and RNA interactions of Tdrd1.

Tdrd1 and nuage formation

We describe strong defects in nuage formation in *tdrd1* mutants. Electron-dense structures are largely absent and those that are present do not associate with the nuclear membrane. Yet, like has been observed in mice (Vagin et al, 2009), perinuclear foci of Piwi proteins can still be observed in *tdrd1* mutants. How can this be resolved? One potential explanation for these findings may be found in the fact that in *Drosophila* the Piwi protein Aubergine, but also Vasa and the arginine-protein-methyltransferase-5 cofactor Valois genetically act upstream of the tudor-domain protein Tud (Cavey et al, 2005; Ephrussi & Lehmann, 1992; Thomson & Lasko, 2005). This could imply that Piwi-pathway activity triggers the onset of nuage formation and that the action of Tdrd proteins is subsequently required to boost this process into formation of nuage as we can detect it in wild-type germ cells. The Piwi positive granules detected in absence of Tdrd1 may thus represent immature nuage with insufficient RNA and protein content to allow detection in TEM experiments.

Specificity in the interactions between Tdrd1 and Zili

Tdrd1 has four tudor-domains. We find that the most C-terminal

domain binds Zili with the highest affinity. This differs from a previous study on mouse Tdrd1 in which the N-terminal part of Tdrd1 was implicated in Mili binding (Wang et al, 2009). This may be due to differences in the mouse and zebrafish Piwi-pathways or to the fact that Wang *et al* investigated the Mili-Tdrd1 interaction in 293T cells, which could lead to sDMA formation on Mili residues that mediate binding to the N-terminal Tdrd1 tudor domains. Related to this hypothesis we find that the Tdrd1-Zili interaction indeed displays sequence specificity. We identify two closely positioned sDMA sites in Zili that associate strongly with Tdrd1, while other sDMA containing peptides display a much weaker interaction. Furthermore, we have not been able to detect another tudor-domain-containing protein, Tdrd6, coming down in our peptide-pull-down assays (Huang and Ketting, unpublished data). With the caveat that our experiments can report indirect interactions, these results suggest that sequence specificity in the interactions between tudor domains and sDMA residues in Piwi proteins may be a general phenomenon. In case of Tdrd1-Zili interactions, the most C-terminal tudor domain of Tdrd1 may interact with sDMAs at positions 221 and 228 in Zili. Due to the poor sequence conservation in the N-terminal regions of Piwi proteins it is difficult to directly relate these residues to homologous arginines in Zili homologues. The available structures of Tudor domains in complex with sDMA containing peptides (Liu et al, 2010a; Liu et al, 2010b) have revealed rather broad groves in which peptides are bound, suggesting that the specificity here observed may not result from sequence specific interactions between the domain. Rather it may be regional electrostatic and hydrophobic interactions that may determine how well a specific peptide will interact with a specific tudor domain.

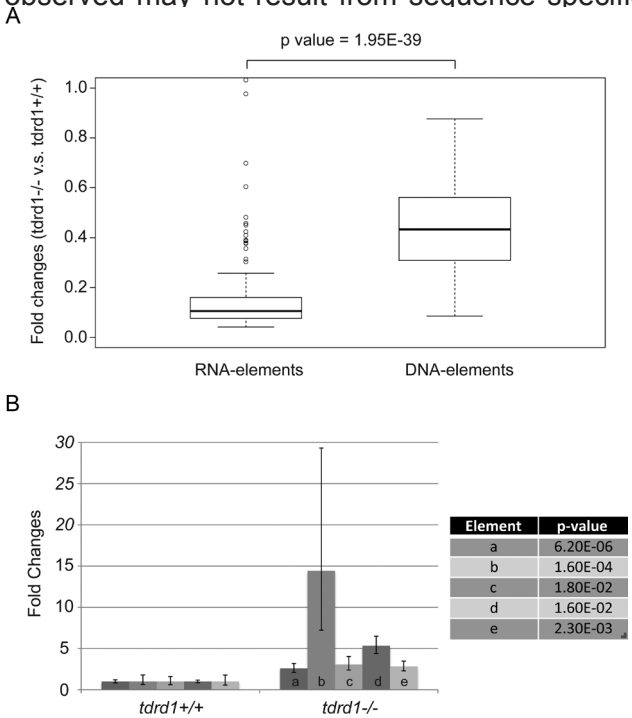


Figure 3. Effect of tdrd1 mutation on transposons and piRNAs. (**A**) Box plot displaying the effects of *tdrd1* mutation on piRNA levels derived from transposable elements that transpose either via an RNA- or a DNA intermediate. Ratios are based on all transposon derived reads that were normalized between the wild-type and *tdrd1* libraries. (**B**) q-RT-PCR analysis on five individual transposons, using RNA from FACS isolated germ cells from wild-type and *tdrd1* mutant gonads. a-e: EnSpmN1, I-1, GypsyDR2, Ngaro1, Polinton1.

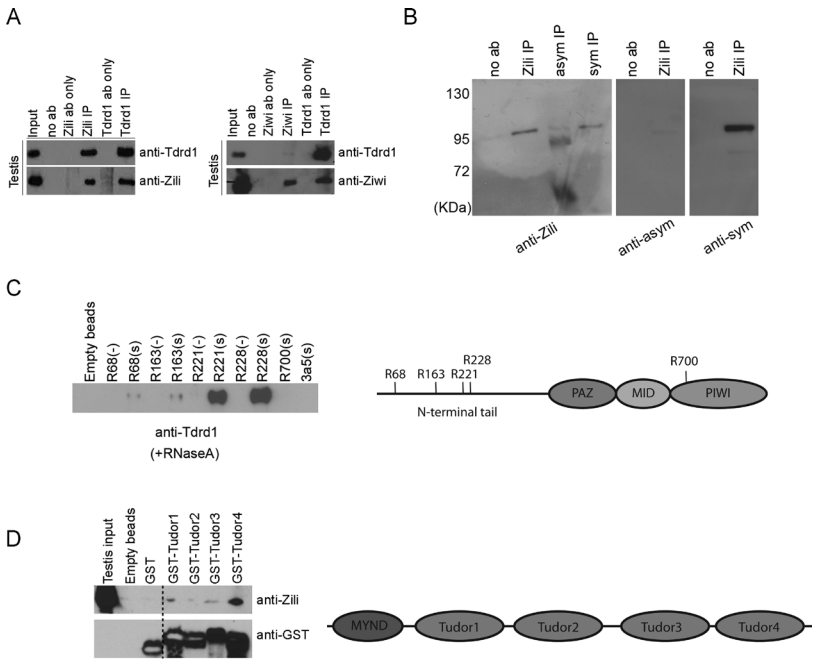


Figure 4. Specific arginine methylation on Zili mediates Tdrd1 binding. (**A**) Western blots on IP material from testis extracts using antibodies directed against Ziwi, Zili and Tdrd1. (**B**) IP-western blot analysis on testis extracts using antibodies recognizing Zili, and symmetrically and asymmetrically dimethylated arginines. (**C**) Pull-down experiments using biotinylated peptides spanning Zili residues R68, R163, R221, R228 and R700. A histone-H3-derived peptide containing sDMA (3a5) was also used as a control. Pull-down sample was analyzed through Western blotting. (-): unmethylated arginine; (s) sDMA. (**D**) Tdrd1 tudor domains were individually expressed as GST-fusion proteins and used as baits in testis extracts. Bound material was probed for Zili and GST using Western blotting.

TATs: piRNA biogenesis intermediates?

We have identified long RNA molecules associated with Tdrd1 (TATs) and their characteristics suggest they may be directly derived from piRNA targets. Such a target RNA association function for Tudor domain proteins may be conserved, as for example *Drosophila* Tudor has been found associated with long RNA transcripts, although these have not been analyzed in detail (Nishida et al, 2009). In absence of a biochemical system for piRNA biogenesis it is difficult to prove that TATs are genuine piRNA biogenesis intermediates. However, they should at least be viewed as good candidates for the following reasons. First, by definition, TATs are found in association with Tdrd1, suggesting they may be localized in nuage, a site where ping-pong amplification of piRNAs is believed to occur (Aravin et al, 2009; Harris & Macdonald, 2001; Li et al, 2009; Lim & Kai, 2007; Lim et al, 2009; Malone et al, 2009; Pane et al, 2007; Patil & Kai, 2010; Shoji et al, 2009). Direct testing of this hypothesis will require high-resolution RNA *in situ* hybridization techniques. Second, TATs tend to be of opposite polarity compared to the piRNAs found in the same Tdrd1 environment. And third, TAT 5' ends match

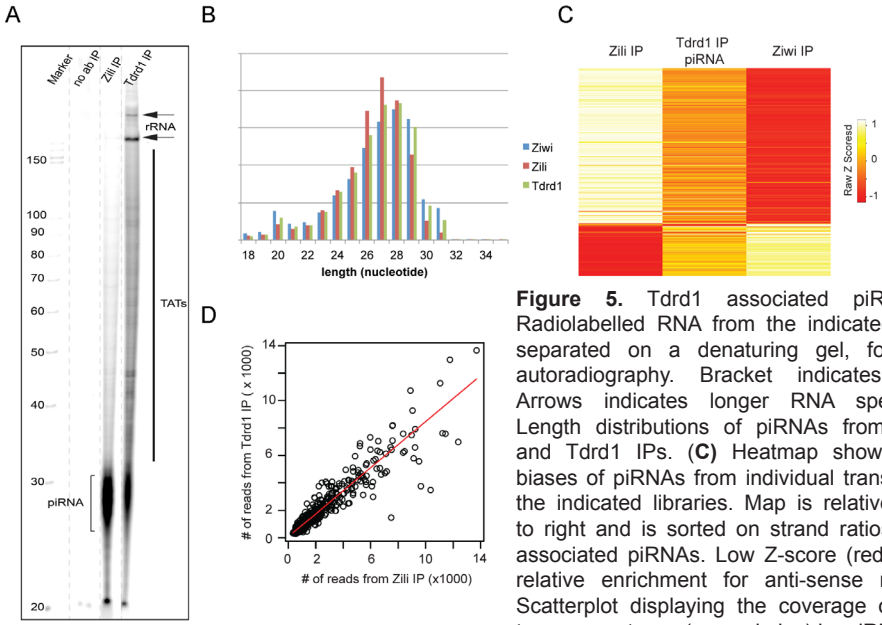


Figure 5. Tdrd1 associated piRNAs. (**A**) Radiolabelled RNA from the indicated IPs was separated on a denaturing gel, followed by autoradiography. Bracket indicates piRNAs. Arrows indicates longer RNA species. (**B**) Length distributions of piRNAs from Ziwi, Zili and Tdrd1 IPs. (**C**) Heatmap showing strand biases of piRNAs from individual transposons in the indicated libraries. Map is relative from left to right and is sorted on strand ratios of Tdrd1 associated piRNAs. Low Z-score (red) indicates relative enrichment for anti-sense reads. (**D**) Scatterplot displaying the coverage of different transposon types (open circles) by piRNAs from a Zili IP (X-axis) and a Tdrd1 IP (Y-axis). Correlation coefficient R=0.93.

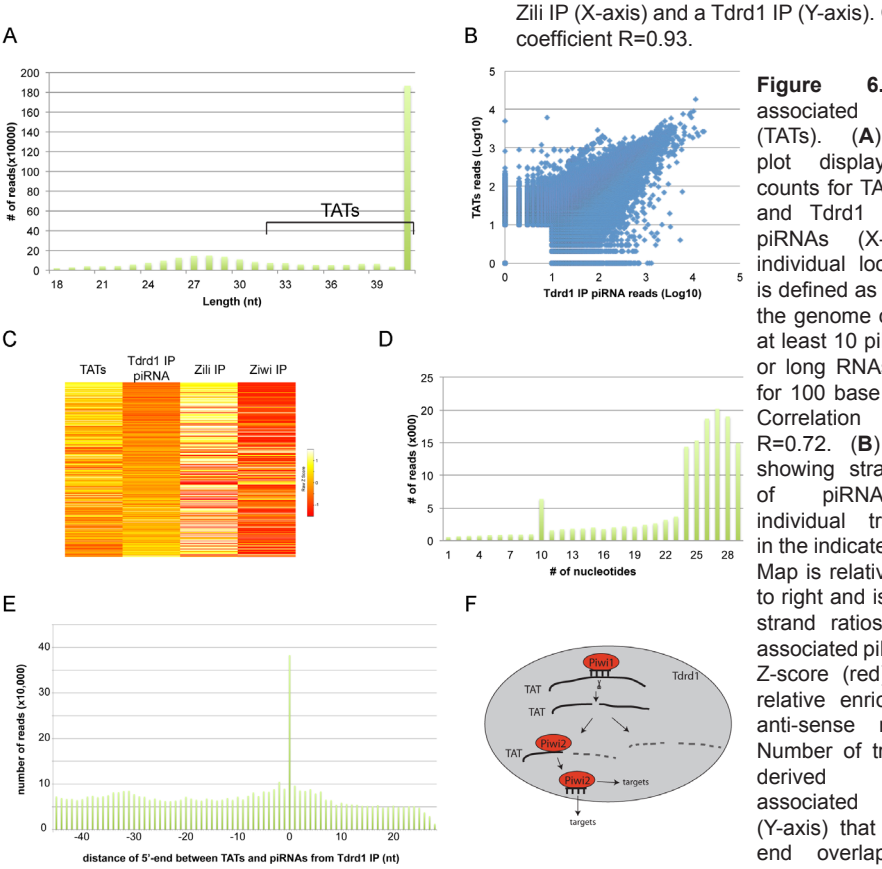


Figure 6. Tdrd1 associated transcripts (TATs). (**A**) Scatter-plot displaying read counts for TATs (Y-axis) and Tdrd1 associated piRNAs (X-axis) for individual loci. A locus is defined as a region of the genome covered by at least 10 piRNAs and/or long RNAs, allowing for 100 base pair gaps. Correlation coefficient R=0.72. (**B**) Heatmap showing strand biases of piRNAs from individual transposons in the indicated libraries. Map is relative from left to right and is sorted on strand ratios of Tdrd1 associated piRNAs. Low Z-score (red) indicates relative enrichment for anti-sense reads. **C)** Number of transposon-derived Tdrd1-associated piRNAs (Y-axis) that have a 5' end overlap (X-axis)

with TATs from the opposite strand. (**D**) Bar diagram displaying the distance between the 5' ends of TATs and Tdrd1-associated piRNAs from the same genomic strands. X-axis reflects the most frequent distance of 5' ends between any TAT species compared to all overlapping piRNAs. Y-axis displays how often each distance is observed.

those of mature piRNAs and they display a ping-pong signal with mature piRNAs, suggesting that TATs are generated through piRNA-mediated cleavage. However, since we find a clear anti-correlation between the strand ratios of long RNAs and piRNAs it is clear that not all TATs can be converted into piRNAs, at least not into Tdrd1-associated piRNAs. We currently entertain two mutually non-exclusive explanations for this observation. First, the majority of TATs may be quickly degraded, preventing further participation in the ping-pong cycle. Second, piRNAs generated in the context of Tdrd1 may not remain associated with Tdrd1. The latter option is consistent with significant, non-Tdrd1 associated Piwi pools that have been observed in many studies. For example, in the adult ovary and testis of zebrafish most Ziwi is not found in peri-nuclear nuage but rather in a diffuse distribution in the cytoplasm (Houwing et al, 2008). In *Xenopus* oocytes Xiwi has been found associated with the Balbiani body (Lau et al, 2009), a structure that we do not find to contain Tdrd1 in the zebrafish. Finally, in mice Miwi2 associates with piP-bodies that lack Tdrd1 (Aravin et al, 2009). These observations strongly suggest that Tdrd1 is not the only scaffold for the Piwi-pathway and that additional sites of piRNA accumulation await further characterisation.

Further research into TATs will be required to fully understand their properties and their relevance to the Piwi pathway. For example, although we did not sequence their complete lengths, many TATs seem to be of rather discrete size (around 170 bases). Currently we have no good explanation for this observation, but perhaps initially a more heterogeneous population is generated that is then trimmed to the observed size? This would suggest that the 5' parts of TATs are being shielded from further degradation. Could nuage be responsible for this shielding? Consistent with this idea, we have previously described that Hen1, an enzyme involved in piRNA maturation, is indeed found in nuage (Kamminga et al, 2010). It will also be of interest to see how the RNA molecules here described relate to the potential piRNA intermediates (named piR-ILs) described by Saito *et al.* (Saito et al, 2010). TATs are longer and likely have 5' phosphate ends that match to 5' ends of mature piRNAs, while piR-ILs lack 5' phosphate groups and do not align well with mature piRNAs. Since piR-ILs have been detected in the context of the primary piRNA biogenesis pathway it could be that piR-ILs are specific to primary piRNA biogenesis and that the TATs described here relate specifically to the ping-pong mechanism. Furthermore, what fraction of the long Tdrd1-associated transcripts will actually be converted into a piRNA and does Tdrd1 interact with long RNAs in the absence of Piwi proteins? Unfortunately, the extremely small size of the gonads at the time when *ziwi*,

zili and *tdrd1* mutants still contain relatively healthy germ cells prevents biochemical analysis of TATs in these mutant backgrounds. Perhaps such experiments will be feasible in mice. Finally, do other Tdrd proteins provide similar environments for specific steps of the Piwi cycle? Questions like these will need to be answered to arrive at a more solid understanding of piRNA biogenesis and function.

MATERIALS AND METHODS

Detailed procedures can be found in the supplemental material.

Zebrafish strains and genetics

Zebrafish were kept under standard conditions. The *tdrd1*$^{fh244+/-}$ mutant allele zebrafish was generated by TILLING and obtained from Zebrafish International Resource Center (ZIRC). Animals carrying *tdrd1*$^{fh244+/-}$ were out crossed against wild-type fish (TL) or vas::EGFP transgenic fish (Krovel & Olsen, 2004) and subsequently in-crossed to obtain *tdrd1*$^{fh244-/-}$ offspring. Experiments involving zebrafish were done with permission from the Animal Experiments Committee of the KNAW.

Tdrd1 antibody

Tdrd1 antibodies were raised in rabbits against the synthetic peptide H2N-RRPATGPSSLSPRGPC-CONH2. Antisera were affinity purified (Eurogentec).

Peptide pull down

Peptide pull downs were done with biotinylated synthetic peptides at a concentration of 2mM. Peptides were incubated with 500µl lysate from 3 testes for 2 hours at 4 ^{o}C while rotating. After washing beads were eluted using SDS-PAAGE loading buffer.

Sequencing

RNA was poly(A)-tailed using poly(A) polymerase followed by ligation of a RNA adaptor to the 5'-phosphate of the small RNAs. First-strand cDNA synthesis was performed using an oligo(dT)-linker primer and M-MLV-RNase H- reverse transcriptase. After amplification the cDNA was sent for a 50-cycle sequencing run on an Illumina/Solexa platform. The four-base adaptor sequences were trimmed from generated data and reads were mapped to the Zv8 zebrafish genome assembly.

ACKNOWLEDGEMENTS

We thank Willi Salvenmoser, Jeroen Korving and Marc van de Wetering for technical assistance, Richard van Schaik and Marc Timmers for assistance in peptide synthesis, Erma van Asselt and Mark Reijnen for fish care, members of the Ketting lab for discussions and suggestions, and the Hubrecht Imaging

Center (HIC) for supporting microscopy experiments. TILLING was done with the support of NIH grant HG002995 to C.B.M. This work was further supported by grants to RFK: a VIDI (864.05.010) and ECHO (700.57.006) grant from the Netherlands Organization for Scientific Research, an ERC Starting Grant from the Ideas Program of the European Union Seventh Framework Program (202819) and the European Union Sixth Framework Program Integrated Project SIROCCO (LSHG-CT-2006-037900).

REFERENCES

Aravin A, Gaidatzis D, Pfeffer S, Lagos-Quintana M, Landgraf P, Iovino N, Morris P, Brownstein MJ, Kuramochi-Miyagawa S, Nakano T, Chien M, Russo JJ, Ju J, Sheridan R, Sander C, Zavolan M, Tuschl T (2006) A novel class of small RNAs bind to MILI protein in mouse testes. Nature 442: 203-207

Aravin AA, Naumova NM, Tulin AV, Vagin VV, Rozovsky YM, Gvozdev VA (2001) Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the D. melanogaster germline. Current biology : CB 11: 1017-1027

Aravin AA, van der Heijden GW, Castaneda J, Vagin VV, Hannon GJ, Bortvin A (2009) Cytoplasmic compartmentalization of the fetal piRNA pathway in mice. PLoS Genet 5: e1000764

Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ (2007) Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. Cell 128: 1089-1103

Carmell MA, Girard A, van de Kant HJ, Bourc'his D, Bestor TH, de Rooij DG, Hannon GJ (2007) MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. Dev Cell 12: 503-514

Carmell MA, Xuan Z, Zhang MQ, Hannon GJ (2002) The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. Genes Dev 16: 2733-2742

Cavey M, Hijal S, Zhang X, Suter B (2005) Drosophila valois encodes a divergent WD protein that is required for Vasa localization and Oskar protein accumulation. Development 132: 459-468

Chen C, Jin J, James DA, Adams-Cioaba MA, Park JG, Guo Y, Tenaglia E, Xu C, Gish G, Min J, Pawson T (2009) Mouse Piwi interactome identifies binding mechanism of Tdrkh Tudor domain to arginine methylated Miwi. Proc Natl Acad Sci U S A 106: 20336-20341

Chuma S, Hosokawa M, Kitamura K, Kasai S, Fujioka M, Hiyoshi M, Takamune K, Noce T, Nakatsuji N (2006) 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 U S A 103: 15894-15899

Cote J, Richard S (2005) Tudor domains bind symmetrical dimethylated arginines. The Journal of biological chemistry 280: 28476-28483

Cox DN, Chao A, Baker J, Chang L, Qiao D, 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

Cox DN, Chao A, Lin H (2000) piwi encodes a nucleoplasmic factor whose activity modulates the number and division rate of germline stem cells. Development 127: 503-514

Deng W, Lin H (2002) miwi, a murine homolog of piwi, encodes a cytoplasmic protein essential for spermatogenesis. Dev Cell 2: 819-830

Ephrussi A, Lehmann R (1992) Induction of germ cell formation by oskar. Nature 358: 387-392

Friberg A, Corsini L, Mourao A, Sattler M (2009) Structure and ligand binding of the extended Tudor domain of D. melanogaster Tudor-SN. J Mol Biol 387: 921-934
Ghildiyal M, Zamore PD (2009) Small silencing RNAs: an expanding universe. Nat Rev Genet 10: 94-108
Girard A, Sachidanandam R, Hannon GJ, Carmell MA (2006) A germline-specific class of small RNAs binds mammalian Piwi proteins. Nature 442: 199-202
Gunawardane LS, Saito K, Nishida KM, Miyoshi K, Kawamura Y, Nagami T, Siomi H, Siomi MC (2007) A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila. Science 315: 1587-1590
Harris AN, Macdonald PM (2001) Aubergine encodes a Drosophila polar granule component required for pole cell formation and related to eIF2C. Development 128: 2823-2832
Horwich MD, Li C, Matranga C, Vagin V, Farley G, Wang P, Zamore PD (2007) The Drosophila RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. Curr Biol 17: 1265-1272
Hosokawa M, Shoji M, Kitamura K, Tanaka T, Noce T, Chuma S, Nakatsuji N (2007) Tudor-related proteins TDRD1/MTR-1, TDRD6 and TDRD7/TRAP: domain composition, intracellular localization, and function in male germ cells in mice. Dev Biol 301: 38-52
Houwing S, Berezikov E, Ketting RF (2008) Zili is required for germ cell differentiation and meiosis in zebrafish. EMBO J 27: 2702-2711
Houwing S, Kamminga LM, Berezikov E, Cronembold D, Girard A, van den Elst H, Filippov DV, Blaser H, Raz E, Moens CB, Plasterk RH, Hannon GJ, Draper BW, Ketting RF (2007) A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. Cell 129: 69-82
Kamminga LM, Luteijn MJ, den Broeder MJ, Redl S, Kaaij LJ, Roovers EF, Ladurner P, Berezikov E, Ketting RF (2010) Hen1 is required for oocyte development and piRNA stability in zebrafish. EMBO J 29: 3688-3700
Kennerdell JR, Yamaguchi S, Carthew RW (2002) RNAi is activated during Drosophila oocyte maturation in a manner dependent on aubergine and spindle-E. Genes & development 16: 1884-1889
Kirino Y, Kim N, de Planell-Saguer M, Khandros E, Chiorean S, Klein PS, Rigoutsos I, Jongens TA, Mourelatos Z (2009a) Arginine methylation of Piwi proteins catalysed by dPRMT5 is required for Ago3 and Aub stability. Nat Cell Biol 11: 652-658
Kirino Y, Mourelatos Z (2007) Mouse Piwi-interacting RNAs are 2'-O-methylated at their 3' termini. Nat Struct Mol Biol 14: 347-348
Kirino Y, Vourekas A, Kim N, de Lima Alves F, Rappsilber J, Klein PS, Jongens TA, Mourelatos Z (2010) Arginine methylation of vasa protein is conserved across phyla. J Biol Chem 285: 8148-8154
Kirino Y, Vourekas A, Sayed N, de Lima Alves F, Thomson T, Lasko P, Rappsilber J, Jongens TA, Mourelatos Z (2009b) Arginine methylation of Aubergine mediates Tudor binding and germ plasm localization. RNA 16: 70-78
Krovel AV, Olsen LC (2004) Sexual dimorphic expression pattern of a splice variant of zebrafish vasa during gonadal development. Dev Biol 271: 190-197
Kuramochi-Miyagawa S, Kimura T, Ijiri TW, Isobe T, Asada N, Fujita Y, Ikawa M, Iwai N, Okabe M, Deng W, Lin H, Matsuda Y, Nakano T (2004) Mili, a mammalian member of piwi family gene, is essential for spermatogenesis. Development 131: 839-849
Lau NC, Ohsumi T, Borowsky M, Kingston RE, Blower MD (2009) Systematic and single cell analysis of Xenopus Piwi-interacting RNAs and Xiwi. EMBO J 28: 2945-2958
Lau NC, Seto AG, Kim J, Kuramochi-Miyagawa S, Nakano T, Bartel DP, Kingston RE (2006) Characterization of the piRNA complex from rat testes. Science 313: 363-367
Li C, Vagin VV, Lee S, Xu J, Ma S, Xi H, Seitz H, Horwich MD, Syrzycka M, Honda BM, Kittler EL, Zapp ML, Klattenhoff C, Schulz N, Theurkauf WE, Weng Z, Zamore PD (2009) Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. Cell 137:

509-521
Lim AK, Kai T (2007) Unique germ-line organelle, nuage, functions to repress selfish genetic elements in Drosophila melanogaster. Proc Natl Acad Sci U S A 104: 6714-6719
Lim AK, Tao L, Kai T (2009) piRNAs mediate posttranscriptional retroelement silencing and localization to pi-bodies in the Drosophila germline. J Cell Biol 186: 333-342
Liu H, Wang JY, Huang Y, Li Z, Gong W, Lehmann R, Xu RM (2010a) Structural basis for methylarginine-dependent recognition of Aubergine by Tudor. Genes & development 24: 1876-1881
Liu K, Chen C, Guo Y, Lam R, Bian C, Xu C, Zhao DY, Jin J, MacKenzie F, Pawson T, Min J (2010b) Structural basis for recognition of arginine methylated Piwi proteins by the extended Tudor domain. Proceedings of the National Academy of Sciences of the United States of America 107: 18398-18403
Liu L, Qi H, Wang J, Lin H (2011) PAPI, a novel Tudor-domain protein, complexes with AGO3, Me31b and Tral in the nuage to silence transposition. Development DOI10.1242/dev.059287
Malone CD, Brennecke J, Dus M, Stark A, McCombie WR, Sachidanandam R, Hannon GJ (2009) Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. Cell 137: 522-535
Nishida KM, Okada TN, Kawamura T, Mituyama T, Kawamura Y, Inagaki S, Huang H, Chen D, Kodama T, Siomi H, Siomi MC (2009) Functional involvement of Tudor and dPRMT5 in the piRNA processing pathway in Drosophila germlines. EMBO J 28: 3820-3831
Olivieri D, Sykora MM, Sachidanandam R, Mechtler K, 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
Pane A, Wehr K, Schupbach T (2007) zucchini and squash encode two putative nucleases required for rasiRNA production in the Drosophila germline. Developmental cell 12: 851-862
Patil VS, Kai T (2010) Repression of Retroelements in Drosophila Germline via piRNA Pathway by the Tudor Domain Protein Tejas. Curr Biol 20: 724-730
Pitt JN, Schisa JA, Priess JR (2000) P granules in the germ cells of Caenorhabditis elegans adults are associated with clusters of nuclear pores and contain RNA. Dev Biol 219: 315-333
Qi H, Watanabe T, Ku HY, Liu N, Zhong M, Lin H (2011) The Yb body, a major site for Piwi-associated RNA biogenesis and a gateway for Piwi expression and transport to the nucleus in somatic cells. The Journal of biological chemistry 286: 3789-3797
Reuter M, Chuma S, Tanaka T, Franz T, Stark A, Pillai RS (2009) Loss of the Mili-interacting Tudor domain-containing protein-1 activates transposons and alters the Mili-associated small RNA profile. Nat Struct Mol Biol 16: 639-646
Saito K, Ishizu H, Komai M, Kotani H, Kawamura Y, Nishida KM, Siomi H, Siomi MC (2010) Roles for the Yb body components Armitage and Yb in primary piRNA biogenesis in Drosophila. Genes Dev 24: 2493-2498
Saito K, Nishida KM, Mori T, Kawamura Y, Miyoshi K, Nagami T, Siomi H, Siomi MC (2006) Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the Drosophila genome. Genes Dev 20: 2214-2222
Saito K, Sakaguchi Y, Suzuki T, Siomi H, Siomi MC (2007) Pimet, the Drosophila homolog of HEN1, mediates 2'-O-methylation of Piwi- interacting RNAs at their 3' ends. Genes Dev 21: 1603-1608
Shoji M, Tanaka T, Hosokawa M, Reuter M, Stark A, Kato Y, Kondoh G, Okawa K, Chujo T, Suzuki T, Hata K, Martin SL, Noce T, Kuramochi-Miyagawa S, Nakano T, Sasaki H, Pillai RS, Nakatsuji N, Chuma S (2009) The TDRD9-MIWI2 complex is essential for piRNA-mediated retrotransposon silencing in the mouse male germline. Dev Cell 17: 775-787
Siomi MC, Mannen T, Siomi H (2010) How does the royal family of Tudor rule the PIWI-interacting RNA pathway? Genes Dev 24: 636-646
Szakmary A, Reedy M, Qi H, Lin H (2009) The Yb protein defines a novel organelle and regulates male germline stem cell self-renewal in Drosophila melanogaster. The Journal of

cell biology 185: 613-627
Thomson T, Lasko P (2005) Tudor and its domains: germ cell for
Vagin VV, Sigova A, Li C, Seitz H, Gvozdev V, Zamore PD (2006) A distinct small RNA pathway silences selfish genetic elements in the germline. Science 313: 320-324
Vagin VV, Wohlschlegel J, Qu J, Jonsson Z, Huang X, Chuma S, Girard A, Sachidanandam R, Hannon GJ, Aravin AA (2009) Proteomic analysis of murine Piwi proteins reveals a role for arginine methylation in specifying interaction with Tudor family members. Genes Dev 23: 1749-1762
Vasileva A, Tiedau D, Firooznia A, Muller-Reichert T, Jessberger R (2009) Tdrd6 is required for spermiogenesis, chromatoid body architecture, and regulation of miRNA expression. Curr Biol 19: 630-639
Wang J, Saxe JP, Tanaka T, Chuma S, Lin H (2009) Mili interacts with tudor domain-containing protein 1 in regulating spermatogenesis. Curr Biol 19: 640-644
Yabuta Y, Ohta H, Abe T, Kurimoto K, Chuma S, Saitou M (2011) TDRD5 is required for retrotransposon silencing, chromatoid body assembly, and spermiogenesis in mice. The Journal of cell biology 192: 781-795

SUPPLEMENTAL MATERIAL

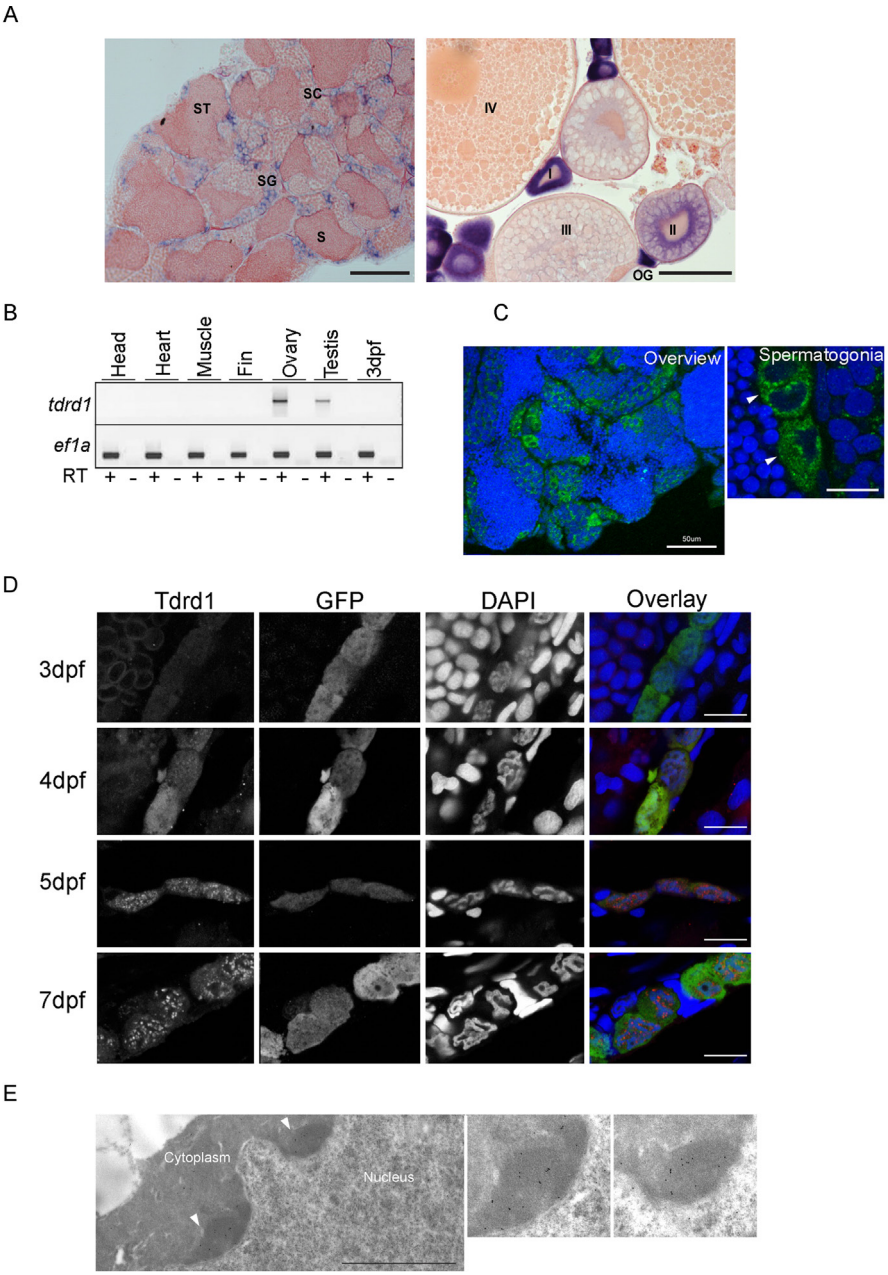


Figure S1. Tdrd1 expression data. (A) Transcripts from tdrd1 were visualized by *in situ* hybridisation (purple) on adult testis (left panel) and ovary (right panel). Sections were counterstained with H&E. Scale bars are 100µm. (B) RT-PCR analysis for tdrd1 expression on indicated samples. (C) Immunohistochemistry for Tdrd1 on testis. Scale bars are 10µm in the right panel and 50µm in the left panel. (D) Immunohistochemistry for Tdrd1 (red) and GFP (green) on embryos between 3 and 7dpf, containing a vasa::EGFP transgene marking the PGCs. Scale bars are 10µm. (E) Tdrd1 immunogold labelling on wild-type adult ovaries. Scale bar is 2µm. White arrow-heads indicate nuage.

MATERIALS AND METHODS

Zebrafish strains and genetics

Zebrafish were kept under standard conditions. The *tdrd1*$^{fh244+/-}$ mutant allele zebrafish was generated by TILLING and obtained from Zebrafish International Resource Center (ZIRC). Animals carrying *tdrd1*$^{fh244+/-}$ were out crossed against wild-type fish (TL) or vas::EGFP transgenic fish (Krovel and Olsen 2004) and subsequently in-crossed to obtain *tdrd1*$^{fh244-/-}$ offspring. For genotyping, the DNA was extracted from caudal fin tissue amputated from anesthetized fish. The primers used to amplify and re-sequence the allele are: fh244_03: 5'-GAA AAA CCT AAG GAG TCA AAA GCT G-3' and fh244_04: 5'-GGCAGAGTGTCTATGCTTGGATAAC-3'. The lesion induces a truncation after amino acid E175. This residue precedes the epitope used for immunization.

Western blot and immuno staining analysis

Western blot and immuno stainings were done as described before (Houwing et al. 2007). Tdrd1 antibodies were raised in rabbits with the synthetic peptide H_2N-RRP ATG PSS LSP RGP C- $CONH_2$. Antisera were subsequently purified against the synthetic peptide (Eurogentec). Embryos were fixed in 4 % PFA at RT for 3hours. Mouse anti-GFP B-2 (Santa Cruz) was used.

Immuno precipitation

Immuno precipitations were done with Ziwi(Girard et al. 2006), Zili(Houwing et al. 2008), Tdrd1 affinity purified antibodies and SYM11 (Millipore) and ASYM24 (Millipore). Tissues in IP lysis buffer were sonicated for 2 minutes at 4 ^{o}C and centrifuged for 10 minutes at 16000 rpm at 4 ^{o}C. Supernatant was used for IP. One IP contains 30 µl Dynabeads (Invitrogen), 3 or 6 testes, and Ziwi antibody 1:200, Zili antibody 1:250 or Tdrd1 antibody 1:100 in a total volume of 500 µl. IP lysis buffer and IP wash buffer are described before (Houwing et al. 2008), except the NaCl concentration in IP wash buffer was raised to 500 mM. Beads, antibody and lysate were incubated for 3 hrs at 4 ^{o}C while rotating. Beads were used for Trizon LS (Invitrogen) RNA isolation or eluted with crack buffer for Western blot or Mass Spectrum analysis. RNA was dephosphorylated and 5' labelled with ^{32}P and run on 12 % denaturing polyacrylamide gels. Gels were exposed to phosphor-imager screens, which were scanned on a BAS-2500 imager (Fuji).

Reverse Transcriptase PCR

cDNA synthesis and RT-PCR were performed as described before (Houwing et al. 2007). Primers used for amplifying Tdrd1 fragments are: Tdrd1-F1: 5'-GTA TAA AGT CTC TAA CAG ACA CCT GG -3' and Tdrd1-R1: 5'-CAG GTA ATA CGA CTC ACT ATA GGG GGG GAA TTT CAG TGG AG-3'; Tdrd1-F2: 5'- GTA AAG AGC ACT TCC CTG TA -3' and Tdrd1-R2: 5'-CAG GTA ATA

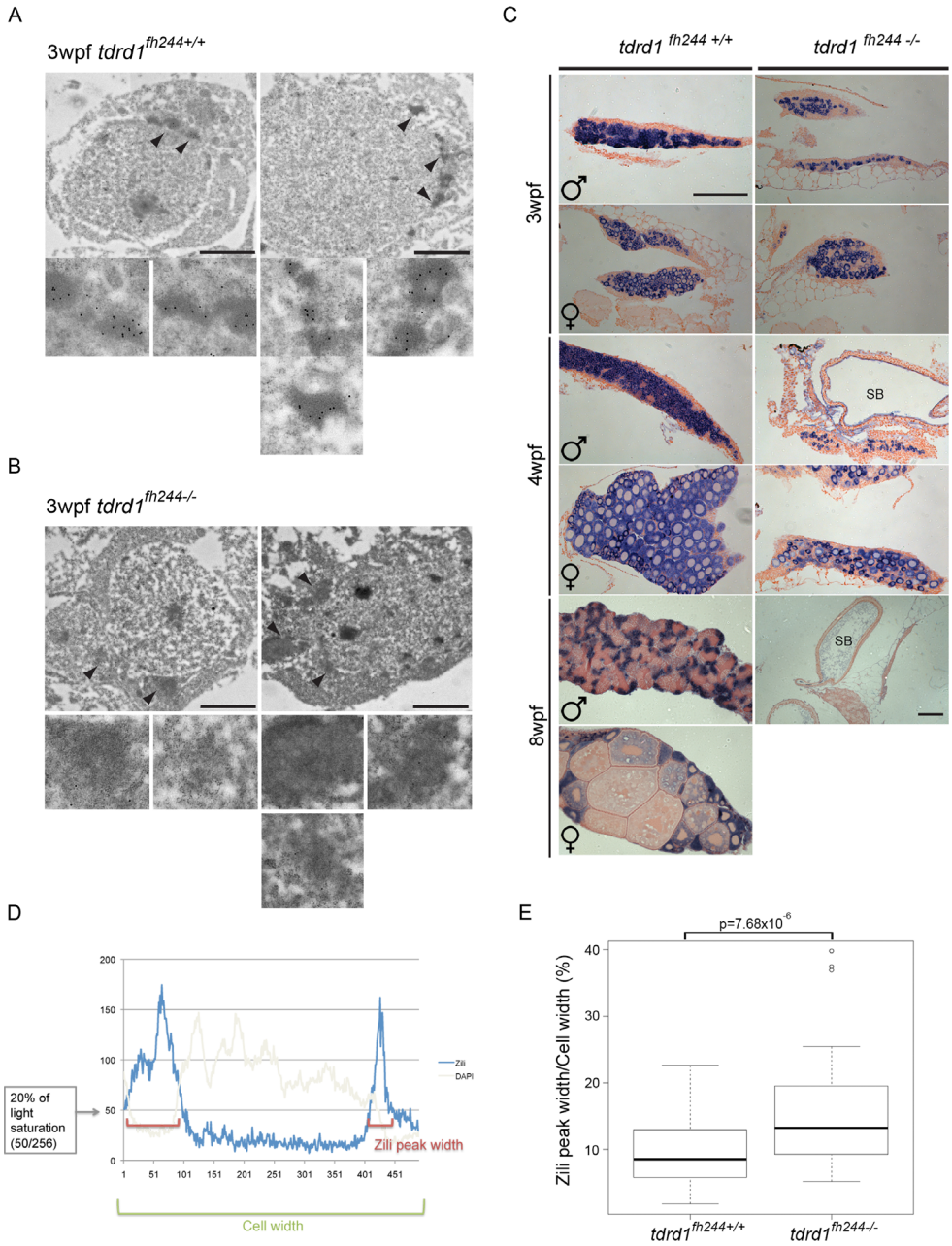


Figure S2. tdrd1 mutant morphology analysis. (A) Zili immunogold labelling on three-week-old wild-type ovaries. (B) Zili immunogold labelling on three-week-old *tdrd1* mutant ovaries. Scale bars are 2µm in A and B. Black arrow-heads indicate nuage. (C) Vasa in situ hybridization (purple) on gonads from wild-type and tdrd1 mutant animals at different developmental stages. Sex was judged by the presence or absence of juvenile oocytes. Wpf: weeks post fertlization. SB: swim-bladder (gonads are close to the swimbladder and in some preparations the SB is visible). Scale: 100µm. (D, E) Quantification method and the quantification result of Zili immunohistochemistry on wild-type and tdrd1 oocytes at 3 weeks. The width of the Zili peak was determined by making a line scan through the cell, always taking the longest axis. At 20%

CGA CTC ACT ATA GGG TGG TTC CTC GTG CTT-3'.

Q-PCR

Juvenile gonads (3wpf) were isolated from incrosses of three-week-old *tdrd1*$^{fh244+/-}$ fish with vas::EGFP transgene. Gonads were treated wih TrypLE Express (GIBCO) for 30 minutes at 28 ^{o}C to dissociate the cells. Fetal calf serum was added to 20% to stop the reaction. Cells were spun down at 1500rpm for 1 minute and washed once with PBS. Resuspended cells were kept in PBS on ice before sorting. The cell suspension was further filtered and sorted using a MoFlo Fluorescence Activated Cell Sorting machine (Dakocytomatino). Total RNA was isolated from sorted high GFP cells. Q-PCR was done as described before (Houwing et al. 2008). Sequences of primers for I-1 element: I-1-F: 5'- GAG TTG AGG ATT GGA AGA TCA TT-3' and I-1-R: 5'- CTT GAG TTT TAG CAA TAG AAA ATC TAA ATC-3'.

***In Situ* Hybridizations**

In situ hybridization was performed as described before (Houwing et al. 2007). Two *tdrd1* fragments were PCR amplified using the same primers as for RT-PCR with an T7 polymerase promoter added to the reverse primer. The PCR product was used as a template for probe synthesis with T7 polymerase.

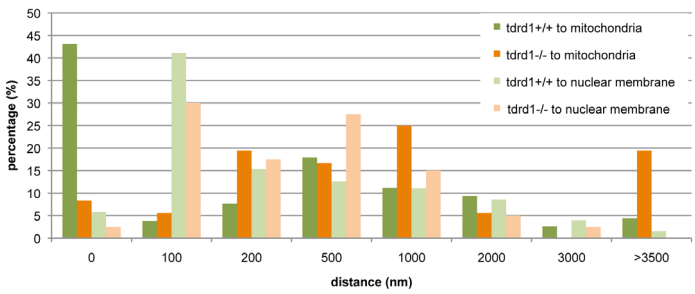
Peptide pull down

Peptide pull downs were done with biotinylated synthetic peptides with concentration of 2 mM. Peptide sequences: R68(s): H_2N-GEM PVR FGR(s) GIT QSI AAK(BiotinC6) -$CONH_2$;R68(-): H_2N-GEM PVR FGR GIT QSI AAK(BiotinC6)-$CONH_2$; R163(s): H_2N-GSS LVS MFR(s) GLG IEP GK(BiotinC6)-$CONH_2$; R163(-): H_2N-GSS LVS MFR GLG IEP GK(BiotinC6)-$CONH_2$; R221(s): H_2N-EES ISF LGR(s) GFT GFG RAK(BiotinC6)-$CONH_2$; R221(-): H_2N-EES ISF LGR GFT GFG RAK(BiotinC6)-$CONH_2$; R228(s): H_2N-GRG FTG FGR(s) AAM PHM TVK(BiotinC6)-$CONH_2$; R228(-): H_2N-GRG FTG FGR AAM PHM TVK(BiotinC6)-$CONH_2$; R700(s): H_2N-EEL VTT FSR(s) VAG PMG MRK(BiotinC6)-$CONH_2$; 3a5(s): H_2N-AR(s)T KQT ARK STG GKA PRG GK(BiotinC6)-$CONH_2$. 1µl of peptide was incubated with 30 µl Dynabeads M-280 Streptavidin in 500 µl IP lysis buffer for 1 hour at 4 ^{o}C while rotating. Supernatant was removed. The beads were incubated with 500 µl lysate from 3 testes for 2 hours at 4 ^{o}C while rotating and washed extensivelty. Lysate and washing buffers are identical to those used in IPs. Beads were eluted with crack buffer for Western blot analysis.

GST-tudor protein production and pull down

PCR products for individual Tdrd1 tudor domains were cloned into pGEX-KG vector containing GST tag at the 5' end of the multiple cloning site. GST-fusion tudor constructs were transformed into BL21 bacteria and induced

of light saturation the width of the Zili peak was **determined. This was** divided by the cell width (which is unaffected by tdrd1 mutation). A Chi-squared-test was used to calculate the indicated p-value.



	tdrd1+/+	*tdrd1-/-*
Total number of cells	51	50
Total number of nuage	343	43
Average number of nuages per cell	6.725	0.86

Figure S3. Distance measurements between nuage and mitochondria and the nuclear membrane. The distance of nuage patches to mitochondria the nuclear membrane was measured on TEM images obtained from wild-type and *tdrd1* mutant samples. Results are displayed in a bin plot. Numbers of observations are given in the table below. Significance was assessed through Chi-test analysis of wild-type versus mutant patterns: p=1.6*10-5 for nuage-mitochondria distance. p=0.42 for nuage-nuclear envelope distance.

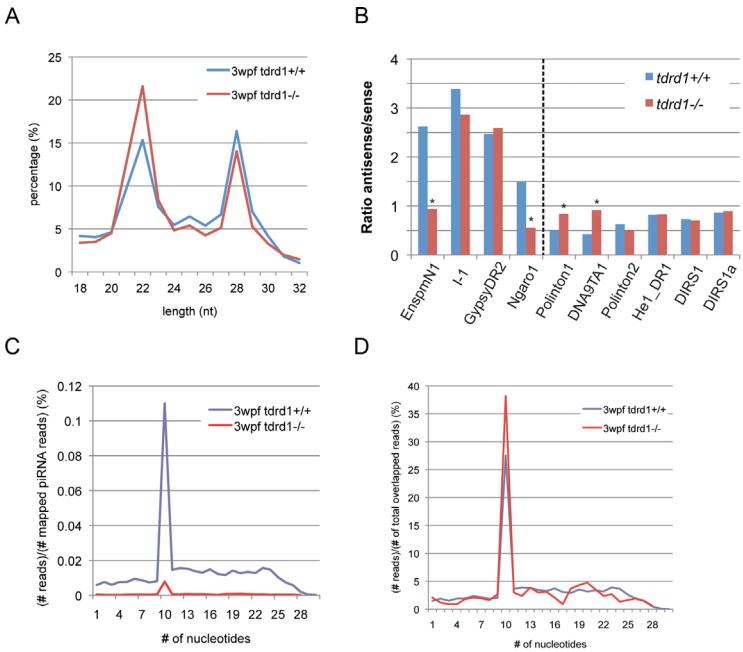


Figure S4. tdrd1 mutant piRNA analysis. (A) Length profile of cloned small RNAs from wild-type and tdrd1 mutant gonads isolated from 3-week-old animals. (B, C) Ping-pong signal among wild-type and tdrd1 mutant piRNAs. In (B) the overlap length is plotted against the ratio between number of reads given a particular length overlap and the total mapped piRNAs. This reveals relatively little overlapping piRNA species in the tdrd1 mutant library. In (C) the overlap length is plotted against the ratio between number of reads given a particular length overlap and the total amount of overlapping piRNAs. This corrects for the low amount of overlaps in the tdrd1 mutant data, and reveals a functional ping-pong mechanism in the tdrd1 mutant, although much less efficient. (D) Anti-sense/sense piRNA ratios for indicated elements. Asterisks indicate significant differences (p<10-15; Chi-squared test). Wild-type: blue. Tdrd1: red. Elements left of the vertical line have antisense enrichment in Ziwi, elements to the right have sense enrichment in Ziwi (Houwing et al. 2008).

with 1 mM IPTG for 3 hours at 30 °C. Bacteria pellets were resuspended in lysis buffer (150 mM NaCl, 50 mM Tris pH 7.5, 1 mM DTT, 1 % NP40, and protease inhibitor (Complete Mini EDTA-free, Roche)) and sonicated 1 minute at 4 °C. Fusion proteins were purified with glutathione agarose beads (Sigma-Aldrich) from the supernatant.

LC-MS/MS

Proteins were digested in-gel with trypsin, and peptides were extracted for subsequent identification by LC-MS/MS. Nanoflow liquid chromatography was performed on an Agilent 1100HPCL binary solvent delivery system (Agilent Technologies, Waldbronn, Germany) with a thermostated wellplate autosampler coupled to an LTQ-Orbitrap mass spectrometer (Thermo Electron, Bremen, Germany). Peptides were trapped at 5 µl/min in 100 % A (0.1 M acetic cid in water) on a trapping column (30 mm x 100 µm packed in-house with Aqua C18, Phenomenex, Torrance, CA) for 10 minutes. After flow-splitting down to around 100 nl/min, peptides were transferred to the analytical column (200 mm x 50 µm packed in-house with Reprosil-Pur C18 AQ, Dr. Maisch GmbH, Ammerbuch, Germany) and eluted with a gradient of 0-40 % B (80 % Acetonitrile/0,1 M Acetic Acid) in a 60-minute gradient. Nano spray was achieved using a coated fused silica emitter (New Objective, Cambridge, MA) (o.d., 360 µm; i.d., 20 µm, tip i.d. 10 µm).

The LTQ-Orbitrap mass spectrometer was operated in data-dependent mode, automatically switching between MS and MS/MS. The two most intense peaks were selected for collision-induced dissociation (CID) in the linear ion trap at normalized collision energy of 35 %. Full scan MS spectra were acquired with a resolution of 60,000 at 400 m/z after accumulation to a target value of 500000.

All MS/MS spectra were converted to DTA files using Bioworks 3.3 (Thermo, San Jose) and searched against the IPI Zebrafish protein database (release 3.35) using MAS-COT (Version 2.2.01, Matrix Science, London, UK). Cysteine carbamidomethylation and methionine oxidation were used as fixed and variable modification, respectively. In selected cases, mono-and di-methylation of Arginine, or mono, di, and tri-methylation of Lysine was used as variable modifications. A peptide mass tolerance of 5 ppm and fragment mass tolerance of 0.8 Da were selected, using Trypsin as the proteolytic enzyme allowing one missed cleavage. All data were loaded into Scaffold (Version 02.01.00, Proteome-Software, Portland, OR) to probabilistically validate peptide and protein identifications. Peptide and protein identifications were accepted when reaching a minimum of 95 % probability, with a minimum of two identified peptides per protein.

Deep sequencing

Long RNA was poly(A)-tailed using poly(A) polymerase followed by ligation of

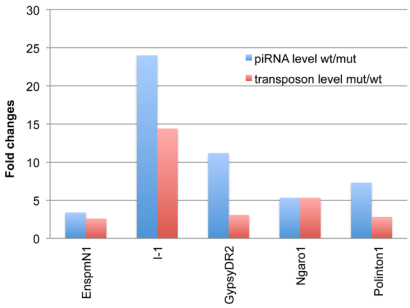


Figure S5. Comparison between changes in transcript levels and changes in piRNA abundance in wild-type versus tdrd1 mutant animals.

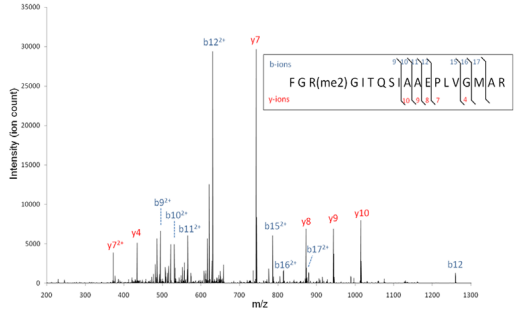


Figure S6. Arginine dimethylation on Zili. Mass spectrum identifying a Zili peptide spanning R68 in a dimethylated form. Fragments (b- and y-ions) arising from a triply-charged precursor peptide (668.2819 m/z) are indicated in the spectrum and in the identified peptide.

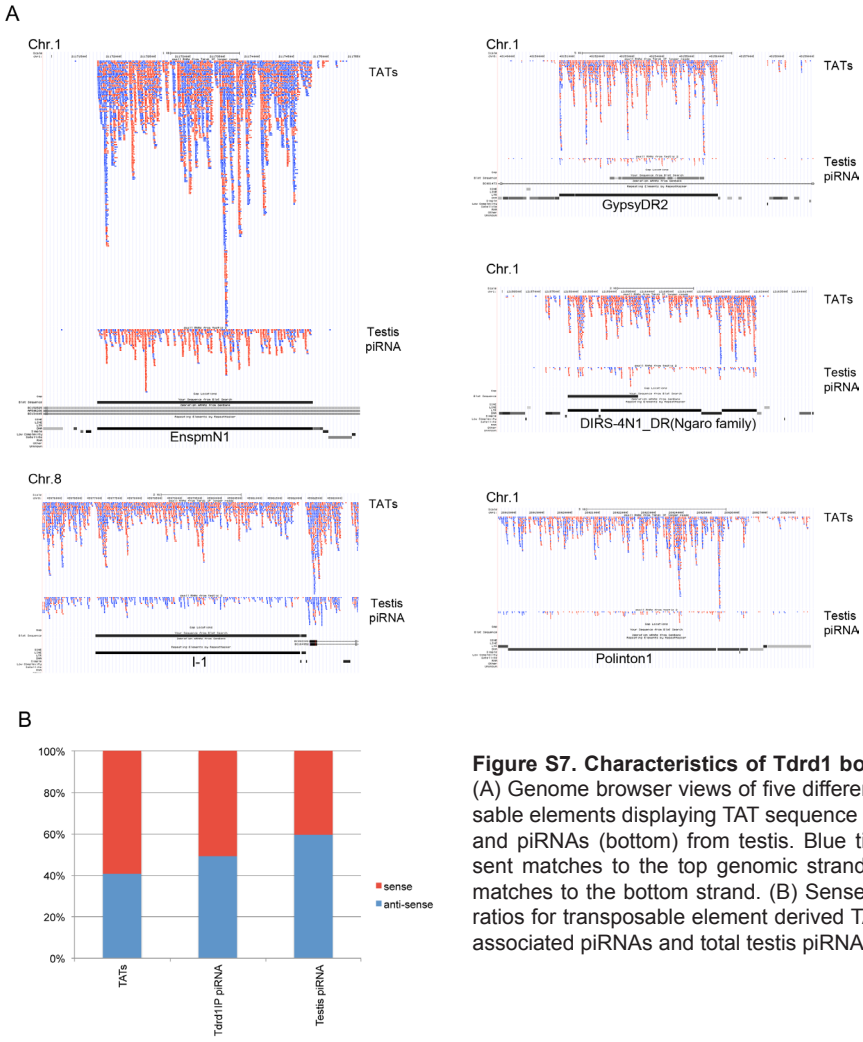


Figure S7. Characteristics of Tdrd1 bound RNA. (A) Genome browser views of five different transposable elements displaying TAT sequence tags (tops) and piRNAs (bottom) from testis. Blue ticks represent matches to the top genomic strand, red ticks matches to the bottom strand. (B) Sense-antisense ratios for transposable element derived TATs, Tdrd1 associated piRNAs and total testis piRNAs.

a RNA adaptor to the 5'-phosphate of the RNAs. First-strand cDNA synthesis was performed using an oligo(dT)-linker primer and M-MLV-RNase H- reverse transcriptase. The cDNA was PCR-amplified for 16 cycles according to the instructions of Illumina/Solexa. cDNA was finally purified using the Macherey & Nagel NucleoSpin Extract II kit and was sent for a 50 cycle sequencing run on an Illumina/Solexa platform. piRNAs were cloned using both 5' and 3' adaptor ligations, and sequenced with 36 cycles.

Sequence analysis

The four-base adaptor sequences were trimmed from generated data using custom scripts. Resulting inserts were mapped to the zebrafish genome (Zv8 assembly) using the megablast program (Zhang et al. 2000), allowing mismatches in the reads starting from nucleotide 19 and considering the longest possible matches as true mapping positions. The mismatched 3' ends of the reads were trimmed, keeping track of the identity of trimmed nucleotides for later counting. Genomic annotations of mapped reads were retrieved from the Ensembl database (release 56) using Perl API provided by Ensembl (Hubbard et al. 2007). Read counts for the different annotated classes are listed in Table S3. Sequences will be made available at GEO upon publication. In the analyses transposon-derived piRNAs are defined as reads with a length between 24 and 29 nucleotides. Heatmap was constructed with basic R programming. The transposons used were filtered on the presence of a strandbias between Ziwi and Zili piRNA populations. The map was sorted on Tdrd1 associated piRNAs. Coloring indicates relative strand ratios comparing the four different piRNA pools.

TEM

Trunks of 3 week old zebrafish were fixed with half-strength Karnovsky fixative (pH 7.4) (Karnovski 1965) and postfixed with 1% OsO4 in 0.05M cacodylate buffer (pH 7.4), dehydrated in acetone series and embedded in Spurr's low viscosity resin (Spurr 1969). Semi-thin, 1-µm-thick sections were cut with a histo Butler diamond knife (Diatome) and mounted on glass slides. After drying, the sections were stained with methylen blue Azur II mixture according to Richardson (Richardson et al. 1960) for light microscopy. Ultrathin sections (70nm in thickness) were placed on copper grids, stained with uranyl acetate followed by lead citrate, and examined with an electron microscope (ZEISS Libra 120 energy filter electron microscope). Image acquisition and analysis were performed using a 2 k Vario Speed SSCCD camera (Droendle), the iTEM software (TEM imaging platform, Olympus) and Adobe Photoshop.

Immunogold EM

Ovaries and testes of wild type zebrafish were dissected and fixed in ice

cold 4% PFA in 0.1 M PBS pH7.4 for 3 hours. Three-week-old *tdrd1*$^{fh244+/-}$ incrossed fish were genotyped and decapitated before fixation as described above. After fixation, samples were washed 3 times for 30 minutes in PBS. Samples were further dehydrated in a standard alcohol series and embedded in LR White medium resin (London Resin Company Ltd). Polymerization was performed at 60°C for 48 hours in an oxygen free atmosphere. Semi-thin sections were cut on a Leica 2040 autocut and ultra-thin sections (90 nm) on a Leica ultramicrotome UCT.

Ultra-thin sections were mounted on pioloform coated gold slot grids, humidified, washed in PBS and blocking was performed using 1% BSA in 0.1 M PBS and 1 % BSA + 0.1% FG in 0.1 M PBS for 25 min each. The grids were incubated with primary antibody in 0.1 M PBS with 1% BSA and 0.1% FG at 4°C over night. After washing they were incubated in goat anti-rabbit secondary antibody, conjugated to 10 nm or 20 nm gold particles, at a dilution of 1:100 for 1 hour at room temperature. After washing and postfixation with 2.5 % GA in 0.1 M PBS, they were counterstained with saturated uranyl acetate and lead citrate. Samples were examined with a Zeiss Libra 120 Energy Filter Transmission Electron Microscope. Zili and Tdrd1 antibodies used for immunogold were affinity purified. Concentrations used for immunogold stainings are: Zili 1:2000, Tdrd1 1:400.

REFERENCES

Girard A, Sachidanandam R, Hannon GJ, and Carmell MA. 2006. A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature* **442**(7099): 199-202.

Houwing S, Berezikov E, and Ketting RF. 2008. Zili is required for germ cell differentiation and meiosis in zebrafish. *EMBO J* **27**(20): 2702-2711.

Houwing S, Kamminga LM, Berezikov E, Cronembold D, Girard A, van den Elst H, Filippov DV, Blaser H, Raz E, Moens CB et al. 2007. A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. *Cell* **129**(1): 69-82.

Hubbard TJ, Aken BL, Beal K, Ballester B, Caccamo M, Chen Y, Clarke L, Coates G, Cunningham F, Cutts T et al. 2007. Ensembl 2007. *Nucleic Acids Res* **35**(Database issue): D610-617.

Karnovski MJ. 1965. A formaldehyde - glutaraldehyde fixative of high osmolality for use in electron microscopy. *JCell Biol* **27**: 137A-138A.

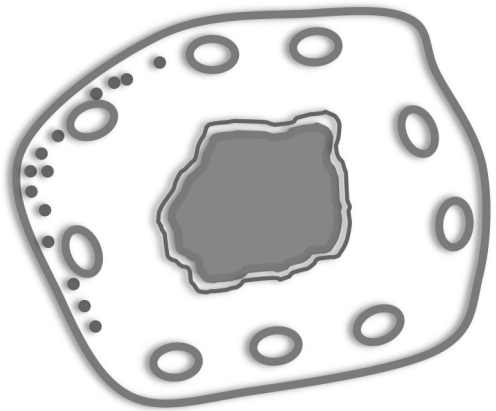
Krovel AV and Olsen LC. 2004. Sexual dimorphic expression pattern of a splice variant of zebrafish vasa during gonadal development. *Dev Biol* **271**(1): 190-197.

Richardson KC, Jarret L, and Finke EH. 1960. Embedding in epoxy resins for ultrathin sectioning in electron microscopy. *Stain Technology* **35**: 313-323.

Spurr AR. 1969. A low-viscosity epoxy resin embedding medium for electron microscopy. *J Ultrastruct Res* **26**: 31-45.

Zhang Z, Schwartz S, Wagner L, and Miller W. 2000. A greedy algorithm for aligning DNA sequences. *J Comput Biol* **7**(1-2): 203-214.

Chapter 4



Tdrd6a is required for normal primordial germ cell formation and accumulation of maternally inherited piRNAs in zebrafish

4

Huang HY, Kaaij LJT, Krijgsveld J, Dosch R, Berezikov E, Ketting RF.

Manuscript in preparation.

INTRODUCTION

Germ plasm is a specialized region of the cytoplasm in early embryos that can impose a germ cell fate on embryonic cells (Blaser et al., 2005; Hashimoto et al., 2004; Koprunner et al., 2001; Yoon et al., 1997). In different species, germ plasm is known with different terms, such as P granules in *C. elegans* and other nematodes, pole plasm in *Drosophila* and other insects and germinal granules in *Xenopus*. The embryonic germ plasm is inherited from the oocyte, where the different components required for germ plasm function aggregate into a large electron dense structure named the Balbiani body. The Balbiani body is an evolutionarily conserved structure in oocytes. It is quite electron dense and is composed of RNA, proteins and organelles, such as mitochondria, ER and Golgi (Guraya, 1979; Kloc et al., 2004; Pepling et al., 2007). The Balbiani body translocates towards the vegetal side of the cortex during the development of the oocytes (Kloc et al., 2004). In zebrafish, the Balbiani body first appears as a prominent aggregate of organelles associated with the nucleus, at the vegetal side; during further development of the oocytes, it moves toward the oocyte cortex at the vegetal side and breaks down into small islands (Marlow and Mullins, 2008).

An interesting gene regarding Balbiani body formation in zebrafish is *bucky ball (buc). Buc*$^{-/-}$ mutants were first found in a maternal-effect screen and the mutant eggs lose its polarity and fail to precede further cell divisions (Dosch et al., 2004). Buc is essential for the set-up of animal-vegetal polarity and Balbiani body assembly in zebrafish oocytes (Marlow and Mullins, 2008). Interestingly, Buc has also been reported to be involved in germ plasm assembly in the embryos (Bontems et al., 2009), where its overexpression has been shown to trigger the formation of additional primordial germ cells (PGCs). Dosch and colleagues suggested that excess Buc induces extra PGCs by recruiting unknown germ plasm components that are already present in the early embryos (Bontems et al., 2009).

All stages of both male and female germ cells are characterized by germ plasm-like structures. This is often referred to as nuage (Eddy, 1974). Nuage is also associated with mitochondria and is mostly associated with the nuclear membrane, where it is often found in the close proximity of the nuclear pores (Pitt et al., 2000; Updike et al., 2011). One of the driving forces behind nuage formation is the PIWI-piRNA pathway, a small RNA-driven silencing pathway. Piwi proteins belong to a clade of Argonaute protein family that is primarily expressed in the germlines of various animal species. Piwi proteins, together with their bound small RNA co-factors called Piwi-interacting small RNAs (piRNAs), form a piRNA-induced silencing complex (piRISC) (Siomi et al., 2011) that suppresses activities of transposable elements (TEs) and maintains genome integrity in the germline. piRNAs are 26-30 nucleotide long and are derived from long single-stranded precursor RNA transcripts from specific genomic regions named piRNA clusters (Aravin et al., 2006;

Brennecke et al., 2007; Girard et al., 2006). Two pathways for piRNA biogenesis have been described: the primary processing pathway and the secondary, also named the Ping-Pong amplification cycle. In *Drosphila* and the zebrafish, the primary processing pathway is involved in the production of the piRNAs from the piRNA clusters. These clusters are loaded with TE copies and are transcribed to produce long single-stranded precursor transcripts that are mostly anti-sense to the TE mRNA transcripts (Brennecke et al., 2007; Houwing et al., 2007). The nucleases involved in processing these transcripts into piRNAs are not known. Interestingly, in mammals a particular set of primary piRNAs is expressed around pachytene stage of meiosis, and these so-called pachytene piRNAs are in fact depleted from transposon sequences. Primary piRNAs are characterized by a strong bias for a uridine (U) residue at the most 5' position.

The secondary pathway involves amplification of piRNA populations, usually by pairs of Piwi paralogs. Piwi proteins loaded with primary piRNAs can cleave target RNAs with complementary sequences, after which the 3' cleavage product can be converted into a novel piRNA. The 5'U bias of primary piRNAs is thus translated into a preference for A at position 10 in secondary piRNAs, since Piwi proteins cleave their target RNA between bases 10 and 11 counted from the 5' end of the guiding piRNA. Secondary piRNAs are generally associated with a different Piwi paralog than primary piRNAs (Aravin et al., 2009; Brennecke et al., 2007; Gunawardane et al., 2007; Houwing et al., 2008; Houwing et al., 2007). For example, the zebrafish genome encodes two Piwi paralogs, Ziwi and Zili. Ziwi binds primary piRNAs that are of overall anti-sense polarity with regard to their transposon targets, while Zili binds secondary piRNAs that are mainly of sense polarity (Houwing et al., 2008).

Both Ziwi and Zili show nuage localization, although at different developmental stages (Houwing et al., 2008; Houwing et al., 2007). Other well-known Piwi-related components of nuage are tudor-domain-containing (Tdrd) proteins, a protein family with at least one conserved tudor domain. The tudor domain is an evolutionarily conserved motif of around 60 amino acids. Some of the Tdrd proteins have been shown to recognize and bind symmetrically dimethylated arginines (sDMAs) present in Piwi proteins, Vasa and perhaps other PIWI-pathway components (Chen et al., 2009; Handler et al., 2011; Huang et al., 2011; Kirino et al., 2010a; Kirino et al., 2010b; Kojima et al., 2009; Lim and Kai, 2007; Liu et al., 2010a; Liu et al., 2010b; Liu et al., 2011; Nishida et al., 2009; Patil and Kai, 2010; Saito et al., 2010; Shoji et al., 2009; Vagin et al., 2009a; Vagin et al., 2009b; Vasileva et al., 2009; Wang et al., 2009; Yabuta et al., 2011; Zamparini et al., 2011). Effects of various Tdrd proteins on PIWI-pathway activity have been described. Studies on Tdrd1 in both mice and zebrafish have revealed that Tdrd1 likely acts in the secondary piRNA pathway, recruiting multiple PIWI pathway components including

different PIWI proteins and their targets (Hosokawa et al., 2007; Huang et al., 2011; Reuter et al., 2009; Wang et al., 2009). TDRD9 has also been shown to act in the secondary piRNA pathway in mice ((Shoji et al., 2009) and this thesis, Chapter 5), but distinct from Tdrd1. In contrast, TDRD6 and TDRD7 do not seem to affect PIWI pathway activity in mice (Tanaka et al., 2011; Vagin et al., 2009b; Vasileva et al., 2009), although both proteins have been shown to physically interact with PIWI proteins. *Drosophila* Tudor, the closest homolog of Tdrd6 in flies, is required for the assembly of the germ plasm (Thomson and Lasko, 2005) and the Tudor-Aub interaction is important for their localization and germ plasm formation (Kirino et al., 2009; Nishida et al., 2009). Maternally provided Piwi is also essential for the maintenance of germ plasm as well as the formation of PGCs in Drosophila (Megosh et al., 2006).

A

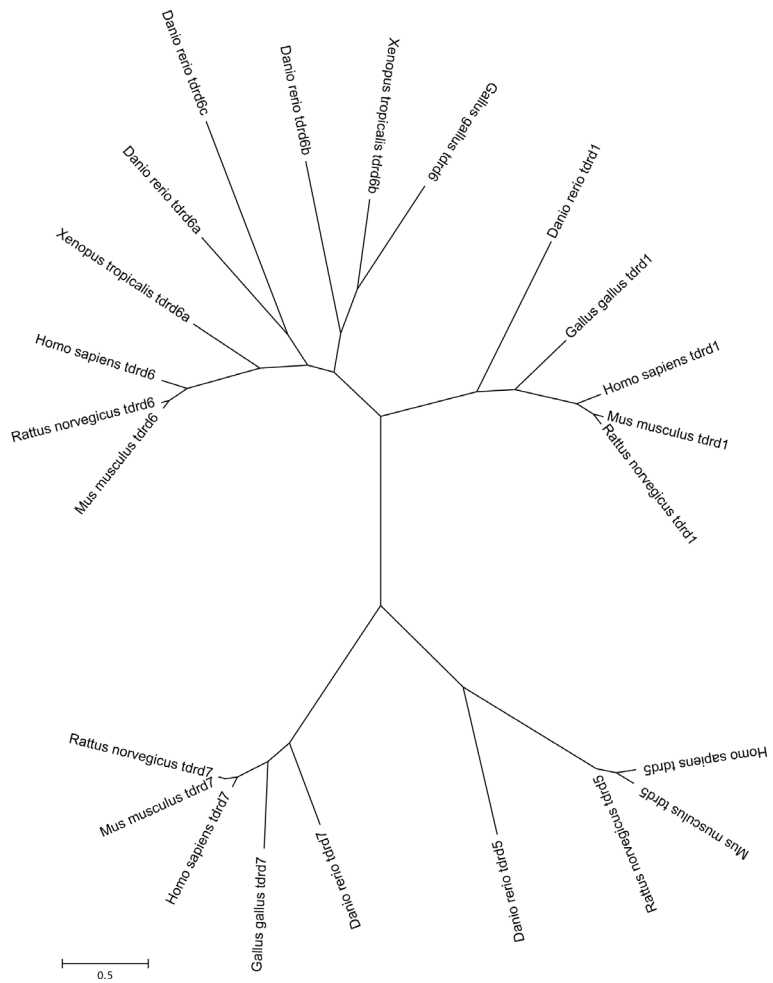


Figure 1. Tdrd6 is conserved across species. (A) Phylogenetic tree of Tdrd1, Tdrd5, Tdrd6 and Tdrd7 proteins in zebrafish, Xenopus, chicken, mouse, rat and human.

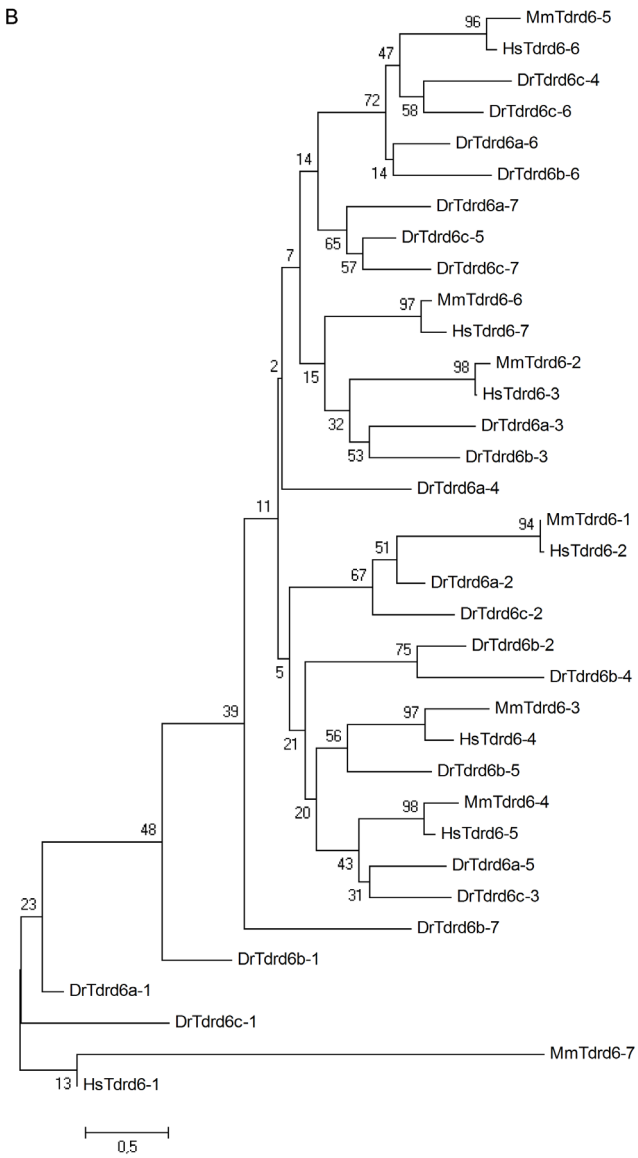


Figure 1. Tdrd6 is conserved across species. ((B) Phylogenetic tree of single tudor domain of Tdrd6 proteins in zebrafish, mouse and human. Tdrd6-1: 1st tudor domain of Tdrd6, Tdrd6-2: 2nd tudor domain of Tdrd6, and so on.

In this study, we show that zebrafish Tudor-domain-containing protein 6a, Tdrd6a, is localized in the nuage of germ cells at most stages of development, but also in the Balbiani body in oocytes. Loss of Tdrd6a leads to transposon transcript up-regulation in the adult male germline. It also causes compromised piRNA function in both male and female gonads. Tdrd6a interacts almost exclusively with Ziwi that is in turn loaded with antisense piRNAs, and mutation of Tdrd6a leads to a specific depletion of anti-sense piRNAs. These data show that Tdrd6a plays a role in the accumulation of anti-sense piRNA pools that are in turn required for proper transposon silencing. Furthermore, maternally provided Tdrd6a plays an important role in recruiting germ plasm components for proper PGC formation in zebrafish embryos. We suggest a model where Buc serves to organize germ plasm by recruiting, amongst others, Tdrd6a, which is in turn required to specify germ cell fate during embryogenesis.

RESULTS AND DISCUSSION

The zebrafish genome encodes three Tdrd6 paralogs

We identified three Tdrd6 homologs in the zebrafish genome by

blasting mouse (ENSMUSG00000040140) and human (ENSG00000180113) Tdrd6 proteins. We named the corresponding genes *tdrd6a (*ENSDARG00000070052), *tdrd6b (*ENSDARG00000014039) and *tdrd6c* (ENSDARG00000089954). *Tdrd6a* and *tdrd6b* are neighboring genes on chromosome 20, while *tdrd6c* is located on chromosome 17. In mammals and chicken we identified only one Tdrd6 homolog, but we did identify two homologs in *Xenopus tropicalis*. (Fig.1A) According to our *in silico* analysis, it seems that the *tdrd6* gene is duplicated specifically in zebrafish and *Xenopus*. All three zebrafish Tdrd6 paralogs contain seven tudor domains that are each marked by sequence characteristics that are evolutionary conserved (Fig. 1B). Tudor domains have been reported to interact with symmetrically dimethylated arginines (sDMA), for instance the sDMAs at the N-terminus of PIWI proteins (Siomi et al., 2010). Indeed, many Tdrd6a peptides, and a few Tdrd6b derived peptides were identified in a Zili immunoprecipitation-Mass Spectrometry experiment (Huang et al., 2011).

Tdrd6a expression in adult gonads

To study the expression of the Tdrd6a protein, we raised a poly-clonal antibody that specifically recognizes zebrafish Tdrd6a. The Tdrd6a antibody was raised against the very C-terminal part of the protein. Using both Western blotting and immunohistochemistry we show that Tdrd6a is strongly expressed in the adult gonads (Fig.2A-B, S1A). In the testis, Tdrd6a is mainly detected in the early stages of spermatogenesis (Fig.2A). Tdrd6a displays big circular structures around the nucleus of spermatogonium, resembling what we described before for the zebrafish Tdrd1 protein (Huang et al., 2011) and what has been described for Medaka Tdrd1 (Aoki et al., 2008). Given these independent observations, it is likely that there are actually sub-structural compartments within nuage in fish. In the ovary, Tdrd6a also appears as nuage-like structures around the nuclei of oogonia and pachytene stage oocytes (Fig.2B). At stage I to III oocytes, Tdrd6a forms smaller granular structures with perinuclear localization (Fig.2B), very similar to zebrafish Tdrd1 in the ovary (Huang et al., 2011), with the notion that Tdrd6a remains detectable up to stage III while Tdrd1 expression quickly fades after stage II.

Apart from the perinuclear nuage localization, we often observe Tdrd6a in an additional prominent structure in stage I oocytes. This structure resembles the Balbiani body. Consistent with the possibility that the Tdrd6a positive structure we find in stage I oocytes is the Balbiani body, we find Tdrd6a present at cortical locations in late stage II and stage III oocytes (Fig.2B). To further test whether this Tdrd6a-positive structure is indeed the Balbiani body, we performed immunohistochemistry experiments probing Tdrd6a localization in *buc*$^{+/+}$ and *buc*$^{-/-}$ ovaries. We did not observe any Balbiani body-like staining in the *buc*$^{-/-}$ mutants, while the perinuclear staining remained

unaffected. This confirms that the Tdrd6a-containing structure is indeed the Balbiani body (Fig.2C, S1B). This data also illustrates that perinuclear nuage in zebrafish oocytes is independent from the Balbiani body, although the Balbiani body could still depend on the presence of perinuclear nuage.

At this point we do not know the function of Tdrd6a in the Balbiani body. Through co-immunoprecipitation experiments and Mass Spectrometry analysis, we have not been able to detect any interaction between Buc and Tdrd6a (Fig.S2). It is therefore likely that Tdrd6a and Buc do not interact directly or are present in a joint macromolecular complex. Many mRNA components of the Balbiani body are related to the germ cell fate; however, the function and the protein components of the Balbiani body are largely unknown (Kloc et al., 2004). It has been proposed that the Balbiani body transports germline-specific mRNAs and proteins to the vegetal pole of the oocyte (Cox and Spradling, 2003). However, the mechanism that keeps the mRNA and protein components stay in the Balbiani body remains elusive. It is plausible that Tdrd6a, with multiple tudor domains, may function as an anchor for mRNAs and proteins to reside in the Balbiani body.

Tdrd6a interacting proteins

To better understand the function of Tdrd6a, we performed immunoprecipitation (IP) using affinity-purified antibody against Tdrd6a in wild-type adult ovaries and testes, respectively, followed by mass spectrometry analysis. In this experiment, IPs using *tdrd6a* mutant ovaries and testes (see below) served as negative control. As may be expected, many PIWI-piRNA pathway components are found, but proteins with by far the highest peptide coverage were Ziwi and Tdrd7. The strong interaction between Tdrd6a and Ziwi was also confirmed by co-IP experiment (Fig. S3A). Other PIWI-pathway factors, like Tdrd1, Vasa, Rnf17/Tdrd4, Tdrd5 and Zili were recovered at much reduced frequencies or not recovered at all, suggesting that the Tdrd6a interactions with the PIWI-pathway are primarily via Ziwi and Tdrd7 (Table1), and are quite distinct from those of, for example, Tdrd1.

Interestingly, we also found a lot of Tdrd6a interacting proteins in the ovaries that are known to be involved in mRNA translational regulation and processing (Table1). These include polyA-binding protein cytoplasmic 1 (PABPC1) that has been shown to bind GW182 (Tritschler et al., 2010), a miRNA pathway component, the Exon-Junction-Complex components Btz, eIFA3 and Y14, RNA helicases DDX19A and DDX43. Finally, we found components of the CPEB and Tral RNP complexes, including CPEB and eIF4E-T/CUP, with high coverage, as well as PAT1 and eIF4E1b. The CPEB complex has been shown to be involved in protein synthesis inhibition in Xenopus oocytes (Minshall et al., 2007). The Tral (also called LSM14A or Rap55) RNP complex, sharing some components with the CPEB complex, including

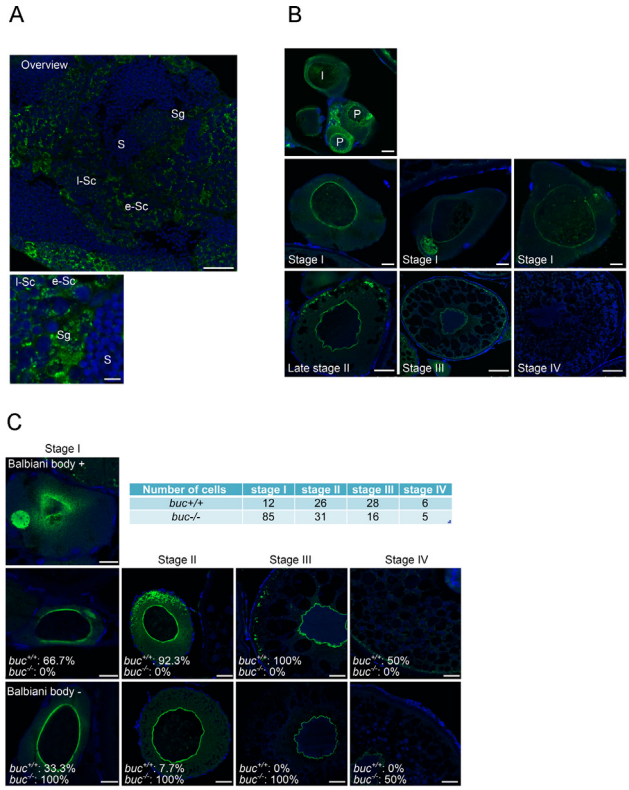


Figure 2. Tdrd6a expression in zebrafish adult gonads. (A) Immunohistochemistry for zebrafish Tdrd6a in adult testis. Sg: spermatogonia; e-Sc: early spermatocyte; l-Sc: late spermatocyte; S: spermatids. Scale bar: upper panel: 25 µm; lower panel: 5 µm. (B) Immunohistochemistry for zebrafish Tdrd6a in adult ovary. P: pachytene stage oocytes. I: stage I oocyte. Scale bar: pachytene and stage I: 10 µm; stage II: 25 µm; stage III and IV: 50 µm. (C) Immunohistochemistry for Tdrd6a in *buc*$^{+/+}$ or *buc*$^{-/-}$ ovaries. Scale bar: stage I: 10 µm; stage II-IV: 25 µm.

Tdrd6a interactors Me31B/DDX6, CPEB, Ybx1 and eIf4E1b and eIF4E-T/CUP, is involved in mRNA localization and translation regulation in Drosophila oogenesis (Wilhelm et al., 2000; Nakamura et al., 2001), and Tral itself has been found in the Balbiani body in mouse and Drosophila oocytes (Pepling et al., 2007). On a mechanistic level, eIF4E-T/CUP was recently reported to reduce the expression of target mRNA by promoting de-adenylation while inhibiting de-capping of the target mRNA at the same time (Igreja and Izaurralde, 2011). These factors of the CPEB and the Tral RNP complexes were not recovered from Tdrd6a IPs using testes lysate. Therefore, we speculate that these interactions arise from the embedding of Tdrd6a in the Balbiani body. The nature of the identified factors from the ovary data strongly suggest that Tdrd6a plays a role in the recruitment or maintenance of mRNA molecules in the Balbiani Body in a translationally inactive state. Given the identification of the complete EJC complex, these mRNAs may be in a state even preceding their first round of translation (Giorgi and Moore, 2007).

Tdrd6a mutants are viable and fertile

To study the loss of function of *tdrd6a*, we isolated a zebrafish allele with a stop mutation right after the first tudor domain of *tdrd6a* (*tdrd6a*^{Q180X}). At different developmental stages, we could not detect Tdrd6a protein in the homozygous mutant germline (Fig.3A, S4A and data not shown), confirming the specificity of the Tdrd6a antibody. *Tdrd6a* homozygous mutants are viable

and fertile. Unlike what we observe in, for example, Hen1 mutant families (Kamminga et al., 2010), we do not observe a significant bias towards male development in *tdrd6a* mutant families, suggesting that *tdrd6a* mutant germ cells are relatively healthy. To check for maternal rescue of a potential *tdrd6a* associated sterility phenotype we also analyzed maternal zygotic mutant animals. These are also viable and fertile, demonstrating that *tdrd6a* is non-essential for viability and fertility.

Tdrd6a expression in the embryo

Using immunohistochemistry on whole mount zebrafish embryos, Tdrd6a protein can be detected at the distal end of the cleavage planes of 4-cell stage embryos (Fig.3A), where the germ plasm is first formed in zebrafish embryos (Blaser et al., 2005; Hashimoto et al., 2004; Koprunner et al., 2001; Mishima et al., 2006; Weidinger et al., 2003; Yoon et al., 1997). At this stage Ziwi, Bruno-like and *vasa* mRNA also localize to these regions (Houwing et al., 2007; Yoon et al., 1997). Later in development, Tdrd6a is specifically found in PGCs where it localizes to perinuclear nuage-like structures.

Taking advantage of the fact that *tdrd6a* mutant animals are fertile, we are able to detect the time point when the zygotic Tdrd6a protein starts to be expressed. Zygotic Tdrd6a can be detected from 3 day-post-fertilization (dpf) onwards. Although the mutant animals are fertile, we notice that maternal Tdrd6a protein is contributing to form bigger nuage structures in 3 dpf PGCs (Fig.3B, S4B). Interestingly, zygotic Tdrd6a is more diffusely located throughout the cytoplasm and the nuage structures seem to be smaller compared to those in PGCs from wild-type mothers. A reasonable explanation is that maternal Tdrd6a is important for gathering nuage components in the embryo. When maternal Tdrd6a is absent, this process may be disturbed resulting in smaller nuage structures. Consistent with this idea we found that Ziwi protein is rather diffusely expressed in 1dpf PGCs of maternal-zygotic *tdrd6a* mutant embryos while Ziwi forms discrete granules around the nucleus of the PGCs in wild-type embryos (Fig.3C). Finally, we analyzed Tdrd1 expression in the absence or presence of maternal Tdrd6a. Tdrd1 protein normally appears as fine granules in the 7dpf PGCs. We measured the average intensity of Tdrd1 signals in the Vas::EGFP-positive area, which labels the PGCs, in the wild-type embryos or embryos lacking maternal Tdrd6a. The average intensity of Tdrd1 signal is further normalized with the average intensity of the Vas::EGFP signal. In embryos lacking maternal Tdrd6a, we observe an overall lower intensity of Tdrd1 signals in the PGCs compare to wildtype PGCs (N= 43, p-value = 1.58E-16), while the subcellular distribution of Tdrd1 seems to be not affected (Fig.3D-E).

It has been shown that interactions between PIWI and Tdrd proteins are important for the subcellular localization and the silencing function of the

Table2. Tdrd6a IP in adult ovary and testis

Accession		a total peptides	b total peptides	a-b	b/a	c total peptide	d total peptide	c-d	d/c
IPI00515878	Tdrd6a	578	0	578	0	113	0	113	0
IPI00499125	Tdrd7	69	0	69	0	50	0	50	0
IPI00851496	eIF4Enif1 / eIF4E-T/CUP	57	0	57	0	0	0	0	
IPI00500092	Btz/MLN51	42	0	42	0	10	0	10	0
IPI00501800	PAT1	38	0	38	0	0	0	0	
IPI00481285	eIF4A3	37	0	37	0	0	0	0	
IPI00485280	PABPC1a	37	0	37	0	52	0	52	0
IPI00866683	Zorba/CPEB	25	0	25	0	0	0	0	
IPI00504207	Actin	19	0	19	0	19	0	19	0
IPI00499534	DDX19A	14	0	14	0	11	0	11	0
IPI00734993	ATP synthase beta subunit	11	0	11	0	0	0	0	
IPI00503192	a-tubulin 1	9	0	9	0	0	0	0	
IPI00886666	DDX43	9	0	9	0	6	0	6	0
IPI00492673	Rnf17	8	0	8	0	0	0	0	
IPI00481706	Tdrd1	7	0	7	0	8	0	8	0
IPI00495861	Vasa	7	0	7	0	3	3	0	1
IPI00760290	Sugar binding	7	0	7	0	0	0	0	
IPI00769676	Esrp2	7	0	7	0	0	0	0	
IPI00806344	Microtubule binding	6	0	6	0	0	0	0	
IPI00505312	RBM8A / Y14	5	0	5	0	2	0	2	0
IPI00551974	Dynll2a	5	0	5	0	0	0	0	
IPI00852189	PABPN1	5	0	5	0	0	0	0	
IPI00495103	Tardbp	4	0	4	0	0	0	0	
IPI00497265	Zili	4	0	4	0	2	0	2	0
IPI00506057	HSP90	4	0	4	0	0	0	0	
IPI00508000	Dynll1	4	0	4	0	0	0	0	
IPI00483853	Glud1b	3	0	3	0	0	0	0	
IPI00484090	Sugar binding	3	0	3	0	0	0	0	
IPI00493643	Pentaxin domain protein	3	0	3	0	0	0	0	
IPI00497603	eIF4E1b	3	0	3	0	0	0	0	
IPI00633761	Uba52	3	0	3	0	0	0	0	
IPI00817780	Rbpms2	3	0	3	0	0	0	0	
IPI00831961	eIF4E-like	3	0	3	0	0	0	0	
IPI00501372	Ziwi	109	3	106	0.03	16	2	14	0.13
IPI00570066	Zar1	59	3	56	0.05	0	0	0	
IPI00482002	Elav-like 2	62	5	57	0.08	0	0	0	
IPI00863260	PABPC1	98	11	87	0.11	0	0	0	
IPI00500251	LSM14 B like	90	14	76	0.16			0	
IPI00837236	Ybx1/Yps	63	10	53	0.16	0	0	0	
IPI00505165	DDX6/Me31B	35	7	28	0.2	0	0	0	
IPI00934329	b-tubulin 2	10	2	8	0.2	8	27	-19	3.38
IPI00497892	igf2bp3 / Imp3	50	11	27	0.22	0	0	0	
IPI00633205	Stau2	31	7	24	0.23	0	0	0	
IPI00932942	HSC70	48	15	33	0.31	14	7	7	0.5
IPI00845047	Sugar binding	11	4	7	0.36	0	0	0	
IPI00507473	Sugar binding	7	3	4	0.43	0	0	0	
IPI00851883	b-actin	8	4	4	0.5	0	0	0	
IPI00505166	Mad2l1	7	4	3	0.57	0	0	0	
IPI00493464	PABPC1b	0	0	0		6	0	6	0
IPI00800089	Histon H4	0	0	0		5	0	5	0
IPI00804253	rbm45	0	0	0		2	0	2	0
IPI00491975	atp5a1	0	0	0		2	0	2	0

**Note: a: wild-type ovary; b: tdrd6a mutant ovary; c: wild-type testis; d: tdrd6a mutant testis

Table 1. Tdrd6a IP in adult ovary and testis. (a) and (c) indicate the reads are from Tdrd6a IP using wild-type ovary and testis lysate, respectively. (b) and (d) indicate the reads are from Tdrd6a IP using *tdrd6a*^{Q180X} mutant ovary and testis lysate, respectively.

PIWI-piRNA pathway (Shoji et al., 2009; Vagin et al., 2009b). In accordance with these earlier general findings, but unlike observations regarding PIWI protein localization in the testis of *tdrd6* mutant mice, our data shows that loss of Tdrd6a leads to de-localization of Ziwi and reduced Tdrd1 expression. These data may indicate that germ plasm formation in the zygote is affected in absence of Tdrd6a. We therefore further analyzed germ cell specification in absence or presence of Tdrd6a.

Tdrd6a is required for normal germ cell specification

Given the above results on germ plasm formation in *tdrd6a* MZ

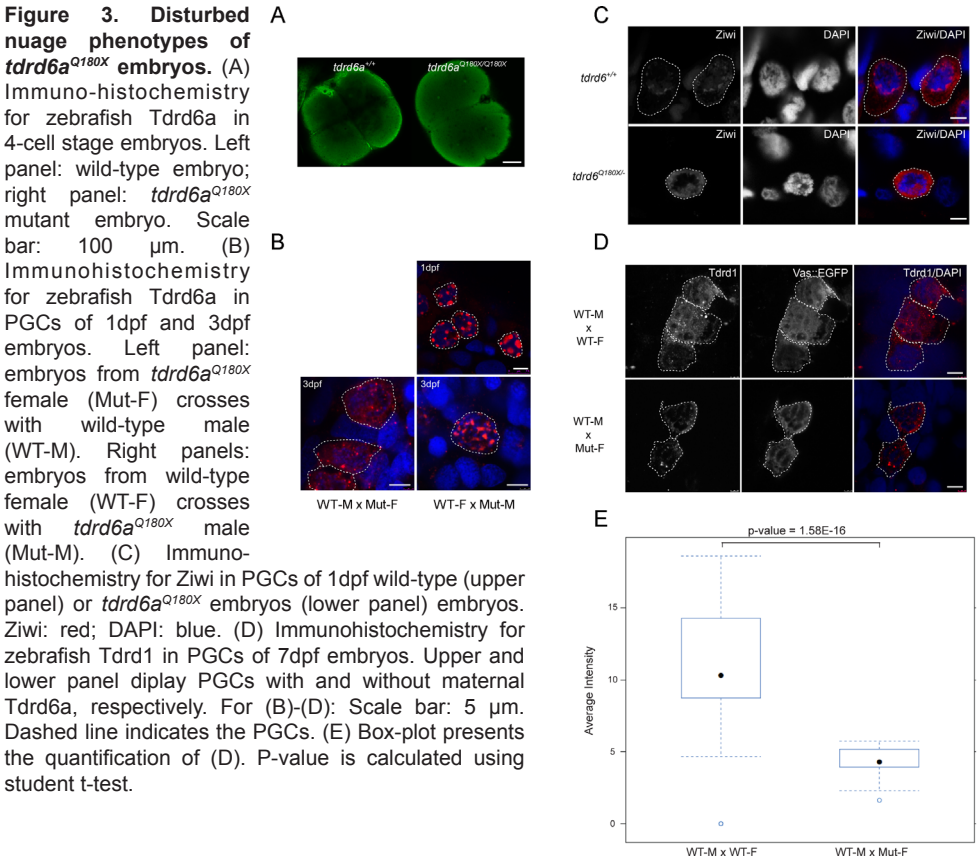


Figure 3. Disturbed nuage phenotypes of *tdrd6a*^{Q180X} embryos. (A) Immuno-histochemistry for zebrafish Tdrd6a in 4-cell stage embryos. Left panel: wild-type embryo; right panel: *tdrd6a*^{Q180X} mutant embryo. Scale bar: 100 µm. (B) Immunohistochemistry for zebrafish Tdrd6a in PGCs of 1dpf and 3dpf embryos. Left panel: embryos from *tdrd6a*^{Q180X} female (Mut-F) crosses with wild-type male (WT-M). Right panels: embryos from wild-type female (WT-F) crosses with *tdrd6a*^{Q180X} male (Mut-M). (C) Immuno-histochemistry for Ziwi in PGCs of 1dpf wild-type (upper panel) or *tdrd6a*^{Q180X} embryos (lower panel) embryos. Ziwi: red; DAPI: blue. (D) Immunohistochemistry for zebrafish Tdrd1 in PGCs of 7dpf embryos. Upper and lower panel diplay PGCs with and without maternal Tdrd6a, respectively. For (B)-(D): Scale bar: 5 µm. Dashed line indicates the PGCs. (E) Box-plot presents the quantification of (D). P-value is calculated using student t-test.

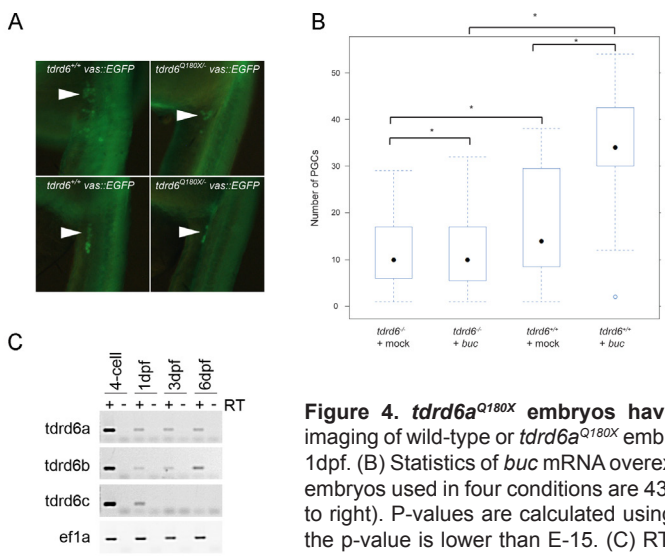
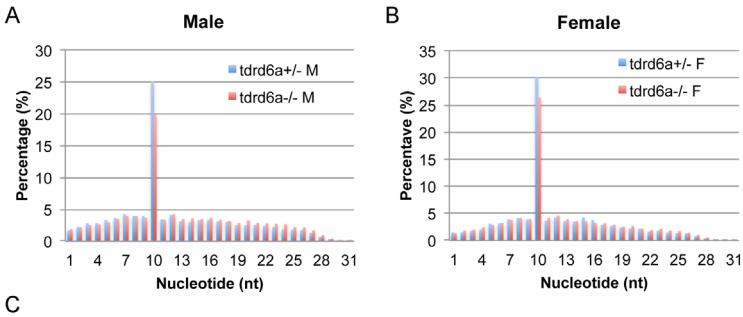


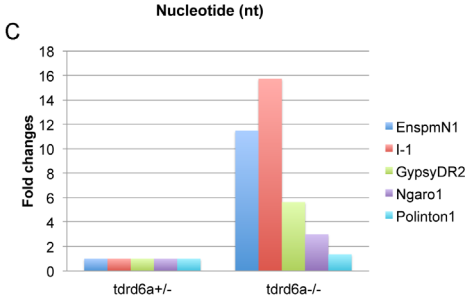
Figure 4. *tdrd6a*^{Q180X} embryos have less PGCs at 1dpf. (A) Live imaging of wild-type or *tdrd6a*^{Q180X} embryos with VAS::EGFP transgene at 1dpf. (B) Statistics of *buc* mRNA overexpression experiment. Numbers of embryos used in four conditions are 43, 39, 39, 37, respectively (from left to right). P-values are calculated using student t-test. Asterisk indicates the p-value is lower than E-15. (C) RT-PCR for three Tdrd6 paralogs in zebrafish. Total RNA samples are isolated from 4-cell, 1dpf, 3dpf and 6dpf stage embryos. *ef1a* is used as loading control.

embryos we turned our attention to PGC formation in these embryos. Interestingly, despite the fact that PGC number in zebrafish embryos displays a wide range from 25 to 50 cells per embryo at 1dpf ((Yoon et al., 1997) and Fig.4A), we observed that *tdrd6a* maternal-zygotic mutant (MZ) embryos tend to have a lower number of PGCs at 1dpf (Fig.4A, 4B). This is consistent with a defect in germ plasm formation. It has been reported that overexpression of *bucky ball* (*buc*) mRNA can induce extra PGCs in the zebrafish embryos, potentially by recruiting germ cell determinants to ectopic germ plasm in somatic cells (Bontems et al., 2009). To test whether *buc* mRNA can rescue the PGC phenotype of *tdrd6a* MZ embryos, we injected *buc* mRNA into 1-cell stage embryos and counted the PGCs at 1dpf. As reported before, wild-type embryos show at least a 30% increase in PGC numbers upon overexpression of Buc (Fig.4B). Interestingly, however, *tdrd6a MZ* embryos do not respond to Buc overexpression at all (Fig.4B).

It has been proposed that Buc functions in recruiting germ plasm components in early embryos and that the aggregation of these components may promote their stability during later embryogenesis (Bontems et al., 2009). Without Tdrd6a, Buc seems not to be able to recruit or stabilize germ plasm components as in wild-type embryos. However, Tdrd6a is not essential for germ cell specification. This may be due to the presence of the *tdrd6b* and *tdrd6c* genes we described above. To probe their expression, we isolated total RNA from different stages of zebrafish embryos, and used reverse transcription-PCR with gene-specific primers to detect *tdrd6a, b and c* transcripts. Transcripts of all three *tdrd6* genes can be detected from 4-cell stage zebrafish embryos onwards (Fig.4C), suggesting that all three RNA transcripts are maternally provided. Therefore it is not unlikely that functional redundancy between these three *tdrd6*-like genes may explain the non-essential nature of the loss of Tdrd6a. We speculate that Tdrd6a acts downstream of Buc, and that either germ plasm components need to first associate with Tdrd6a before Buc can recruit them, or that Tdrd6a is essential to stabilize the aggregate of germ plasm components in order to impose germ cell fate upon the blastomeres of the embryo.

It is likely that the PGC phenotype of *tdrd6a* mutants arises as follows. In the early zebrafish embryo, only cells possessing the germ plasm will obtain the PGC fate (Raz, 2003). Loss of Tdrd6a leads to smaller germ plasm structures, and consequently fewer blastomeres can inherit sufficient germ plasm to be specified as germ cells, leading to embryos with fewer PGCs. From our experience, one single PGC can propagate and populate a whole gonad (Kamminga & Wittkopp, unpublished data). Therefore, even though *tdrd6a* mutant embryos start out with fewer PGCs, the PGCs that are specified can still divide and populate healthy male or female gonads.

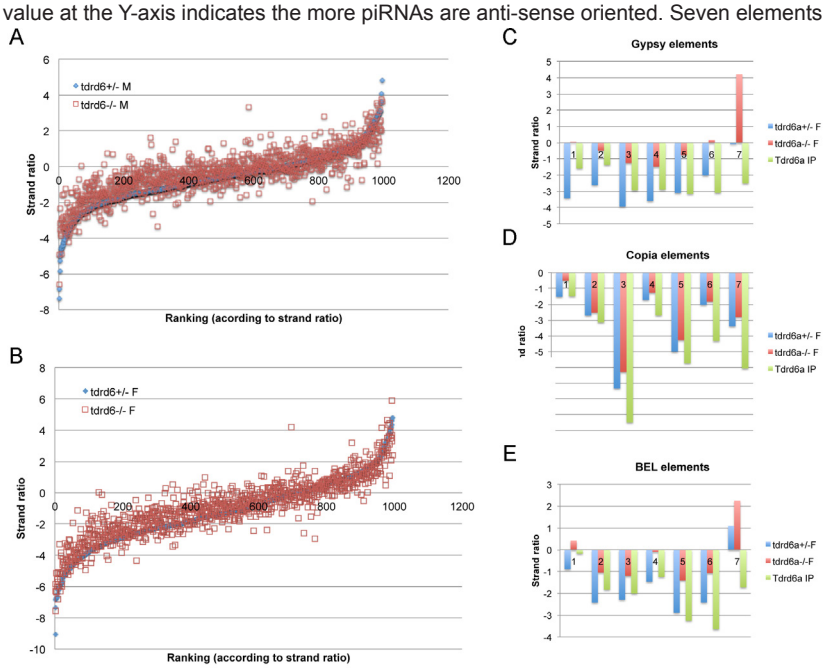




Element	EnspmN1	I-1	GypsyDR2	Ngaro1	Polinton
p-value	1.40E-08	1.35E-04	2.54E-10	2.18E-07	8.35E-04

Figure 5. piRNA function is compromised in *tdrd6a*^{Q180X} males. (A) Ping-pong signature for transposon-derived piRNAs in the small RNA libraries made from *tdrd6a*^{Q180X} adult males and females. (B) Quantitative-RT-PCR experiment shows relative expression of transposon transcripts in *tdrd6a*$^{Q180X/+}$ and *tdrd6a*^{Q180X} adult male gonads. Error bars show SD for three biological replicates. P-values are calculated using student t-test.

Figure 6. Tdrd6a functions as anti-sense piRNA reservoir. For (A) and (B), strand ratio all the transpaoson-derived piRNAs. Each dot indicates the strand ratio of all piRNAs mapped to one certain transposon. (A) Comparing strand ratios of piRNAs in *tdrd6a*$^{Q180X/+}$ adult female (tdrd6a+/- F, blue) and *tdrd6a*^{Q180X} adult female (tdrd6a-/- F, red). (B) Comparing strand ratios of piRNAs in *tdrd6a*$^{Q180X/+}$ adult male (tdrd6a+/- M, blue) and *tdrd6a*^{Q180X} adult male (tdrd6a-/- M, red). The ranking is according to the strand ratio. Dot with more negative strand ratio is closer to the y-axis. For (C)- (E), strand ratio of transposon-derived piRNAs in the following libraries: *tdrd6a*$^{Q180X/+}$ adult female (tdrd6a+/- F, blue), *tdrd6a*^{Q180X} adult female (tdrd6a-/- F, red), and Tdrd6a IP using wild-type adult ovary (Tdrd6a IP, green). The more negative value at the Y-axis indicates the more piRNAs are anti-sense oriented. Seven elements belong to certain



retrotransposon family are presented with bar diagram. (C) Elements in the Gypsy family. 1: Gypsy149-LTR; 2: Gypsy60-I; 3: Gypsy75-LTR; 4: Gypsy-21-I; 5: Gypsy77-I; 6: Gypsy150-I; 7:Gypsy9-LTR. (D) Elements in the Copia family. 1: Copia1-I; 2: Copia2-I; 3: Copia2-LTR; 4: Copia4-LTR; 5: Copia6-LTR; 6: Copia-6-I; 7: Copia-8-LTR. (E) Elements in the BEL family. 1: BEL4-I; 2: BEL32-I; 3: BEL15-I; 4: BEL12-I; 5: CATCH2I; 6: BEL11-I; 7: BEL12-LTR.

piRNA function is compromised in the *tdrd6a*$^{Q180X/Q180X}$ males

To probe the role of Tdrd6a in the PIWI-piRNA pathway, we prepared piRNA libraries from *tdrd6a* mutant and wild-type adult ovary and testis. We observe mild reduction of the total piRNA numbers (~7%) in the mutant female but not male gonads (TableS1). However, the "Ping-Pong signature," which shows the efficiency of the piRNA amplification, is reduced in the gonads of both sexes (Fig.5A-B). Many zebrafish piRNAs map to transposon loci (Houwing et al., 2007) and the PIWI-piRNA pathway in general is known to function in suppression of transposable elements in the germline. To check the transcript level of a number of transposons in the adult gonad, we performed quantitative-PCR using transposon specific primer sets. Two out of five transposons that were tested show more than five-fold up-regulation in *tdrd6a* mutant testes compared to *tdrd6a* heterozygous testes (Fig. 5B). Interestingly, it is often speculated that transposon up-regulation is the main reason of sterility in PIWI-piRNA pathway mutants (Siomi et al., 2011). However, we find similar transposon transcript upregulation in *tdrd6a* mutant males as we do in, for example, *tdrd1* mutant males, and yet *tdrd6a* males are fertile while *tdrd1* males lose all their germ cells. This suggests that de-repression of transposons may not necessarily lead to sterility in the zebrafish, although we cannot exclude the possibility that the fertility defects are caused by activation of transposons that we have not been testing in our qRT-PCR assays.

Recent studies showed that loss of TDRD6 in mouse, lead to spermiogenic defects (Vasileva et al., 2009) but no retrotransposon de-repression (Tanaka et al., 2011). Tanaka et al. speculated the reason is due to the phenotype in *tdrd6*$^{-/-}$, which is established in later stages of development, beyond the time point when the TE control is important. In the zebrafish, Tdrd6a is first detectable in spermatogonia, while in mice it is first seen in the early spermatocytes (Vasileva et al., 2009). This may explain why zebrafish Tdrd6a is more involved in retrotransposon silencing as shown by our data.

Tdrd6a functions as an anti-sense piRNA-reservoir in ovaries

To probe for the molecular function of Tdrd6a, we performed Tdrd6a IPs using wild-type adult ovaries and prepared a piRNA library from them. We have observed before that ovary piRNAs tend to be strongly anti-sense orientated, caused by an abundance of Ziwi-bound piRNAs (Houwing et al., 2007). We observed that Tdrd6a-associated piRNAs also display very strong anti-sense strand bias, consistent with our findings that Tdrd6a more strongly

associates with Ziwi than with Zili. Loss of Tdrd6a leads to loss of strand bias of piRNAs, where the strand bias shifts towards sense orientation. This effect is stronger in the ovaries than in the testes (Fig.6A-B). When individual TEs were examined, we noticed that some retrotransposons, like *gypsy*, *copia* and *bel* elements show an interesting pattern when heterozygous ovary, mutant ovary and Tdrd6a IP piRNA profiles were compared (Fig.6C-E). The piRNAs from mutant ovaries show a general shift towards sense piRNAs while the Tdrd6a IP reveals strong anti-sense enrichment. As we do not observe defects in transposon transcript regulation in ovaries, these findings are more consistent with the idea that Tdrd6a functions as a reservoir for anti-sense piRNAs in ovaries, to be deposited in the embryo, rather than as an actual silencing platform within the ovaries themselves.

CONCLUSION REMARKS

Tdrd6 has been studied before in mice, and a role of TDRD6 in chromatoid body assembly during spermatogenesis has been shown (Hosokawa et al., 2007; Tanaka et al., 2011; Vasileva et al., 2009). In zebrafish sperm, no chromatoid body has been described yet. However, our data do implicate fish Tdrd6a in structures similar to it. We show Tdrd6a is present in perinuclear nuage structures in adult germ cells, where it does not seem to be essential for its formation, although we noted small effects on perinuclear nuage in oocytes. We also noted that the interaction between Tdrd6a and Tdrd1 or Zili are relatively low (Table2 and Fig.S3B) suggesting that the granular peri-nuclear structures we detected with Tdrd1 and Tdrd6a antibodies are possibly distinct sub-compartments of nuage, containing either Zili/Ziwi-Tdrd1 or Ziwi-Tdrd6a. Given that one of the most prominent Tdrd6a interactors in both testis and ovary, PABPC1, is a known p-body component, this organization may be similar to the pi-body/piP-body compartmentalization that has been proposed in the mouse (Aravin et al., 2009). In that case, the Tdrd6a/Ziwi structures would be the zebrafish piP-bodies, while Tdrd1, as in mice, would define the pi-bodies.

Apart from perinuclear nuage, we describe for the first time that Tdrd6a is important in the assembly of the germ plasm in the oocytes and early embryos, and as such also affects germ cell formation in the embryos. This may relate to the function we describe for Tdrd6a in allowing the accumulation of Ziwi-piRNA complexes. Perhaps, in absence of Tdrd6a, less blastomeres can inherit sufficient Ziwi-piRNA complexes to allow the specification, or maintenance of primordial germ cells. Apart from that it is also possible that Tdrd6a plays a role in the deposition of mRNAs like that of nos1 and vasa, and that this function relates to the PGC phenotype. Whether piRNAs would be involved in this function is at this point unclear, and will be important to address in the future.

MATERIALS AND METHODS

Zebrafish strains and genetics

Zebrafish were kept under standard conditions (ref). The *tdrd6a*$^{Q180X/+}$ mutant allele zebrafish was derived from ENU mutagenized libraries using target-selected mutagenesis as described {Wienholds, 2004 #422};. Animals carrying *tdrd6a*$^{Q180X/+}$ were out crossed against wild-type fish (TL)

TableS1. Library composition

Libraries	miRNA	repeats	genic piRNA	other piRNA	snRNA	snoRNA	tRNA	rRNA	total
tdrd6+/-F	256890	13039160	431484	5770243	2558	4346	5290	8338	19518309
tdrd6-/-F	572137	11248756	485353	6524290	4539	4844	9906	19718	18869543
tdrd6+/-M	262660	7228921	843496	10196192	4047	1955	12790	8530	18558591
tdrd6-/-M	424967	7124612	822033	9516784	4816	3237	11969	11293	17919711

Table S1. Composition of the small RNA libraries. RNA samples were prepared from: *tdrd6*$^{+/-}$ F: *tdrd6a*$^{Q180X/+}$ adult female; *tdrd6a*$^{-/-}$ F: *tdrd6a*^{Q180X} adult female; *tdrd6a*$^{+/-}$ M: *tdrd6a*$^{Q180X/+}$ adult male; *tdrd6a*$^{-/-}$ M: *tdrd6a*^{Q180X} adult male.

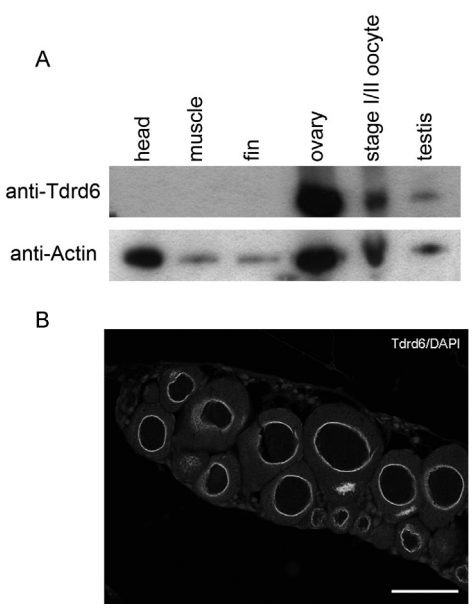


Figure S1 (upper panel). Tdrd6a expression. (A) Western blotting against Tdrd6a in various adult tissues. Actin is used as loading control. (B) Immunohistochemistry for Tdrd6a in 5wkpf gonads. Balbiani body is clearly seen in some stage I oocytes. Green: Tdrd6a; blue: DAPI. Scale bar = 50 µm.

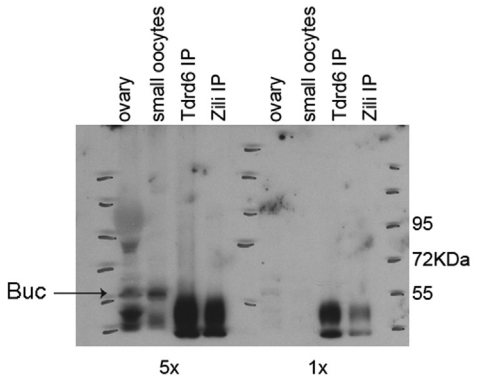


Figure S2. Tdrd6a and Buc does not interact directly with each other. Western blotting against Buc in Tdrd6a and Zili IPs. 5x and 1x indicates the relative amount of protein loaded in the left and right part of the gel, respectively.

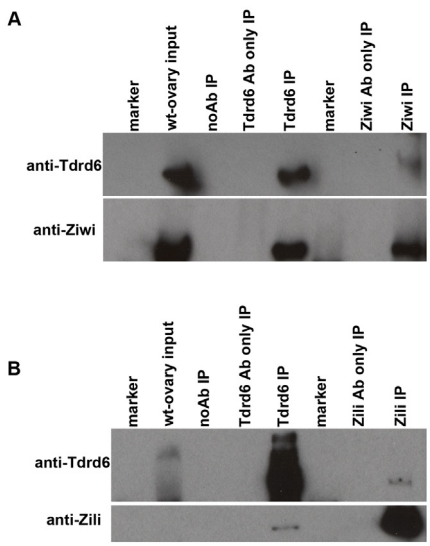


Figure S3. Tdrd6a interacts with both Ziwi and Zili. (A) Co-immunoprecipitation of Tdrd6a and Ziwi using wild-type adult ovaries. (B) Co-immunoprecipitation of Tdrd6a and Ziii using wild-type adult ovaries.

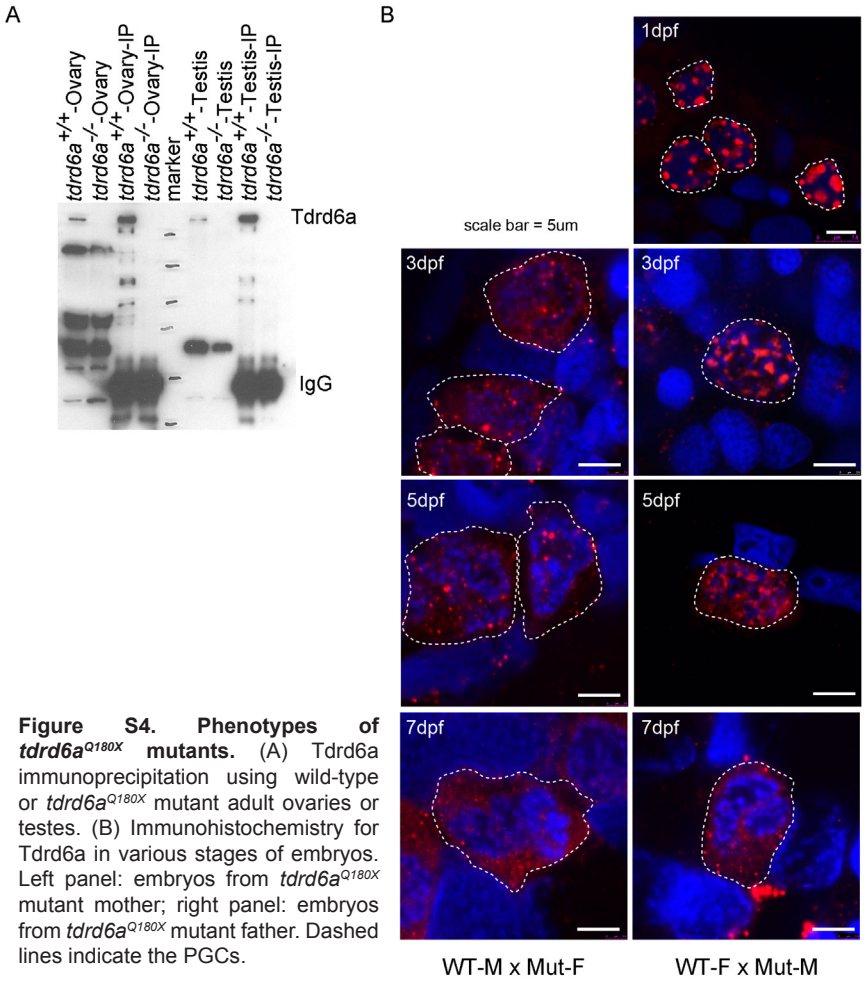


Figure S4. Phenotypes of *tdrd6a*^{Q180X} mutants. (A) Tdrd6a immunoprecipitation using wild-type or *tdrd6a*^{Q180X} mutant adult ovaries or testes. (B) Immunohistochemistry for Tdrd6a in various stages of embryos. Left panel: embryos from *tdrd6a*^{Q180X} mutant mother; right panel: embryos from *tdrd6a*^{Q180X} mutant father. Dashed lines indicate the PGCs.

or vas::EGFP transgenic fish (Krovel and Olsen 2004) and subsequently in-crossed to obtain *tdrd6a*$^{Q180X/Q180X}$ offspring. For genotyping, the DNA was extracted from caudal fin tissue amputated from anesthetized fish. The primers used to amplify and re-sequence the allele are: Tdrd6a_ALC: 5'-GAA GGT GAC CAA GTT CAT GCT TGA CAT TCC TTG TCT GTC AAG GC -3', Tdrd6a_ALT: 5'- GAA GGT CGG AGT CAA CGG ATT CTT GAC ATT CCT TGT CTG TCA AGG T -3', and Tdrd6a_C2: 5'- CAC TGT ACA ATT TCT TTG CAA AGC CCA TT -3'. The lesion induces a truncation after amino acid Q180. This residue precedes the epitope used for immunization.

Western blot and immuno staining analysis

Western blot and immuno stainings were done as described before (Houwing et al. 2007). Tdrd6a antibodies were raised in rabbits with the synthetic peptide H_2N-QAV VHE PES EKE KRD- CONH$_2$. Antisera were subsequently

purified against the synthetic peptide (Eurogentec). Embryos were fixed in 4 % PFA at RT for 3hours. Mouse anti-GFP B-2 (Santa Cruz) was used at 1 to 500.

Immuno-precipitation

Immuno precipitations were done as described before (Huang et al., 2011) with Ziwi(Girard et al. 2006), Zili(Houwing et al. 2008) and Tdrd6a affinity purified antibodies. Tdrd6a antibody was used at 1 to 100.

Reverse Transcriptase PCR

cDNA synthesis and RT-PCR were performed as described before (Houwing et al. 2007). Primers used for amplifying Tdrd1 fragments are: Tdrd6a-F1: 5'-GGA AGG AGG CGC GAA AGT A -3' and Tdrd6a-R1: 5'- ACA CAT GAT GGG TTG AGA TTT AC -3'; Tdrd6a-F2: 5'- TAA CGA GAC CTA CTC TGA CCT TAA A -3'; Tdrd6a-R2: 5'- GCA GCT GTC GAA GGT TAA AAC TAT CAA CA -3'; Tdrd6b-F1: 5'- GGA CAA AAG AGG AAT CAA AGC TC -3'; Tdrd6b-R1: 5'- TTG AGG ACA AAC AGT TGG G -3'; Tdrd6b-F2: 5'- GGA GGG AGT TCT TGG AGT T -3'; Tdrd6b-R2: 5'- CTT TCC TCT GCA ACT CTC -3'; Tdrd6c-F1: 5'- AAT CAC TTT GGA CGA CAA GCG AAG T -3'; Tdrd6c-R1: 5'- TTG GAT GGA GGG ACT TCT TGG CAA A -3'; Tdrd6c-F2: 5'- TGA CAT ATG GCA ATA TGC ATG G -3'and Tdrd6c-R2: 5'- TTT TTG TTA CGG CGT CAG TA -3'.

Q-PCR

Adult gonads were isolated from incrosses of adult *tdrd6a*$^{Q180X/+}$ fish. Total RNA was isolated from the whole gonads. Q-PCR was done as described before (Houwing et al. 2008; Huang et al., 2011).

***buc* mRNA overexpression and PGC counting**

buc mRNA solution was prepared from a pCS2+-buc clone (a generous gift from Dr. Dosch) according to the instruction. 1-cell stage embryos were collected from incrosses of adult *tdrd6a*$^{Q180X/+}$ fish or adult wild-type fish with vas::EGFP transgene. Solutions containing 170ng/µl of buc mRNA and 10% phenol red were injected into collected1-cell stage embryos using glass micropipettes. Mock injections only use 10% phenol red solution. Injected embryos were maintained in the 28 ^{o}C incubator till they reach 24hpf stage. The embryos were further placed at room temperature while the procedure of PGC counting. The embryos were sedated with MS222 and further transferred to methylcellulose solution for immobilization purpose. PGCs clusters at both sides of the yolk extension were counted under Leica GFP microscopes.

LC-MS/MS

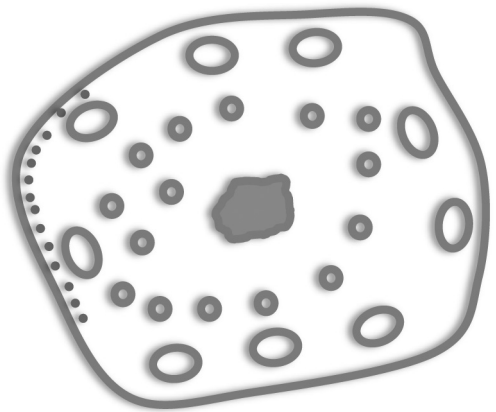
Sample preparation and LC-MS/MS settings were performed as described

before (Huang et al., 2011).

Small RNA libraries and sequencing
Library preparation was performed as described before (Huang et al., 2011).

Sequence analysis
Bioinformatic analysis was performed as previously described (Huang et al., 2011).

Chapter 5



Tdrd9 affects piRNA accumulation and is required for germ cell maintenance in adult male zebrafish

5

Houwing S[#], Huang HY[#], Siegfried KR, Kaaij LJT, Meppelink A, Redl S, Wiebrands K, Waaijers S, Luteijn MJ, Ladurner P, Nüsslein-Volhard C, Berezikov E, Ketting RF.
Submitted.

# contributed equally

Tdrd9 affects piRNA accumulation and is required for germ cell maintenance in adult male zebrafish.

Saskia Houwing[1,2,3], Hsin-Yi Huang[1,3], Kellee R. Siegfried[4,5], Lucas J.T. Kaaij[1], Amanda Meppelink[1], Stefan Redl[6], Kay Wiebrands[1], Selma Waaijers[1], Maartje J. Luteijn[1], Peter Ladurner[6], Christiane Nüsslein-Volhard[4], Eugene Berezikov[1], René F. Ketting[1,7]

[1] Hubrecht Institute-KNAW & University Medical Centre Utrecht, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands
[2] present address: Kimmel Center for Biology and Medicine at the Skirball Institute,
NYU School of Medicine, 540 First Avenue, New York, NY 10016, USA
[3] Contributed equally
[4] Max Planck Institute for Developmental Biology, Department of Genetics Spemannstrasse 35, 72076 Tübingen, Germany
[5] Present address: Childrens Hospital Boston, Orthopedic Research Laboratories, 300 Longwood Ave, Enders 260, Boston, MA02115, USA
[6]Institute of Zoology, Technikerstrasse 25, A-6020 Innsbruck, Austria
[7] corresponding author: r.ketting@hubrecht.eu

ABSTRACT

The Piwi pathway represents a small RNA mediated mechanism that is mostly aimed at the defense of genomes against transposable elements in germ cells. Apart from the Piwi proteins themselves, directly binding the small RNA cofactors known as piRNAs, several additional proteins have been implicated in this process. Among them is the tudor-domain-containing protein Tdrd9, also known as SpindleE in the fruit fly. Apart from its tudor domain, this factor belongs to the family of DExH-box RNA helicases. A precise function of Tdrd9 in the Piwi mechanism has not been described. Here, we present studies on the zebrafish homolog of Tdrd9. We describe the expression of tdrd9 during development and show that the Tdrd9 tudor domain has an intrinsic property to localize to nuage, a germ cell specific subcellular aggregate known to be involved in the Piwi pathway. Furthermore, we present two tdrd9 mutant alleles, both resulting in strong germ cell developmental defects. These defects are further studied at morphological the level, revealing a role for Tdrd9 and its RNA helicase domain in the progression of meiosis into pachytene stages and in the maintenance of germ cells in the adult. Finally, experiments aimed at the effects of Tdrd9 on piRNAs suggest a function for Tdrd9 in the Piwi pathway downstream of the piRNA biogenesis process known as ping-pong amplification. Summarizing, our work confirms the involvement of Tdrd9 in the vertebrate Piwi pathway and provides novel clues to the molecular function of Tdrd9 in this still poorly understood pathway.

INTRODUCTION

Small RNA (sRNA) based regulatory mechanisms play a major role in animal germ cells. One of these mechanisms is better known as the Piwi-pathway, characterized by a specific subfamily of Argonaute proteins, named Piwi proteins, and small RNAs that are better known as piRNAs (Piwi interacting RNAs) (Ghildiyal, 2009; Ketting, 2011). In most animals the Piwi-pathway is involved in silencing transposable elements (TEs) (Malone, 2009; Siomi, 2011). TEs are genomic elements that can induce DNA damage by excision and/or re-integration into the genome, and as such representing immediate threats to the integrity of the hosting genome. Consequently, if the Piwi pathway does not work properly germ cells suffer from an increased transposition load. This has been best established in the fruit fly *Drosophila melanogaster* (Brennecke, 2007; Brennecke, 2008) but there are strong indications that this occurs in other animal species as well (Aravin, 2007; Houwing, 2007). Perhaps related to the activation of transposable elements, germ cells in Piwi mutant animals have been found to display a variety of phenotypes, including increased levels of apoptosis, meiotic failure and stem cell maintenance defects (Carmell, 2007; Kuramochi-Miyagawa, 2004;Deng,

2002; Houwing, 2008; Houwing, 2007).

At a molecular level, the Piwi pathway has been rather well characterized, and it has become clear that two types of piRNA biogenesis exist. Both of these start with a single stranded RNA (ssRNA) molecule and are independent of the Dicer enzyme (Houwing, 2007; Vagin, 2006), a major player in the generation of other sRNA species like miRNAs. This ssRNA is subject to endo-nucleolytic activity, specifying the 5' end of the piRNA to-be-formed. It is this endo-nucleolytic step that distinguishes the two modes of piRNA biogenesis. In the so-called primary piRNA biogenesis pathway an as yet unknown factor induces the cleavage of transcripts that originate from clusters that host many transposon-fragments, the so-called piRNA clusters (Brennecke, 2007; Li, 2009; Malone, 2009). These clusters can be regarded as records of transposition activity in the past and serve as a template to initiate TE defense. In the secondary pathway, the endo-nucleolytic cleavage is performed by a Piwi protein, and the targeted ssRNA can either be piRNA cluster transcripts or TE derived mRNAs. The Piwi protein accepting the new piRNA generally is another Piwi paralog than the one generating the 5'-end cleavage (Brennecke, 2007; Gunawardane, 2007). In both the primary and the secondary pathway, the newly exposed 5' end is bound by a Piwi protein, after which the 3' end is trimmed to fit the hosting Piwi protein. Finally, the piRNA is methylated at its 3'-end by a methyl-transferase enzyme named Hen1. This serves to stabilize piRNAs in response to target recognition (Saito, 2007; Horwich, 2007; Kirino, 2007; Kamminga, 2010; Kurth, 2009). In the secondary pathway, a newly formed piRNA can again trigger the maturation of another piRNA molecule that will be identical in sequence to the original piRNA. Thus, the secondary piRNA biogenesis pathway leads to the amplification of piRNA levels in response to available target RNA; this process is often referred to as ping-pong. This constellation ensures a constant and up-to-date response to TE mediated threats. Interestingly, in mammals the primary piRNA pathway is also employed outside the context of TE silencing (Aravin, 2006; Girard, 2006). These piRNAs are expressed at the onset of pachytene stage and play a role in meiotic progression in male germ cells. Their molecular function remains unknown.

Most animals express at least two different Piwi paralogs that together mediate TE silencing through the ping-pong mechanism. One of these, usually the one accepting both primary and secondary piRNAs, remains in the cytoplasm, where it is associated with an electron-dense material known as nuage (reviewed in Ketting, 2011). Although the precise functions of this structure are not well known, it has become clear that many Piwi-pathway related reactions take place within nuage and that in fact the Piwi pathway itself is one of the driving forces making nuage. The other Piwi paralog, accepting only secondary piRNAs, can translocate to the nucleus upon being loaded with a piRNA (Aravin, 2008). The precise function of nuclear Piwi

proteins has remained unclear sofar, although they have been correlated with *de novo* DNA methylation in mouse primordial germ cells (Aravin, 2008; Kuramochi-Miyagawa, 2008). Many other proteins participate in the Piwi-pathway (Reuter, 2009; Vagin, 2009; Malone, 2009; Kuramochi-Miyagawa, 2010; Liu, 2011; Handler, 2011; Zamparini, 2011; Chen, 2009; Huang, 2011). Among these are many proteins that contain one or more tudor domains (Siomi, 2010). Tudor domains can bind to methylated arginines or lysines and some of the tudor-domain-containing proteins (Tdrd's) functioning in the Piwi-pathway have been shown to bind to symmetrically dimethylated arginines present in the N-terminal regions of the Piwi proteins.

One of these Tdrd proteins is Tdrd9, a protein containing a DExH-box RNA helicase domain in the N-terminal part and one tudor domain in its C-terminal part. The *Drosophila* homolog of Tdrd9, also named Spindle-E or Homeless, has been shown to be crucial for ping-pong amplification of piRNAs and transposon silencing, and for the proper localization of the Piwi proteins Aub and Ago3 (Malone, 2009). In mice, this protein has been shown to bind the Piwi proteins MIWI2 and MILI and to be required for proper transposon silencing (Shoji, 2009; Vagin, 2009). However, TDRD9 mutation does not have an impact on the gross subcellular localization of either MIWI2 or MILI (Shoji, 2009). The functional requirement of either the tudor domain or the DExH box of Tdrd9 has not been demonstrated in any system.

In zebrafish, two Piwi proteins are encoded by the genome (Houwing, 2007), of which Ziwi binds both primary and secondary piRNAs while Zili binds secondary piRNAs (Houwing, 2008). Both Ziwi and Zili are present in the female and male germ cells and proper Piwi pathway activity has been shown to be required for normal TE transcript down-regulation. Like in other systems, Piwi proteins have been found to associate with Tdrd proteins, including Tdrd9 (Huang, 2011). Here we describe the impact of Tdrd9 on the zebrafish Piwi pathway, both on the phenotypic as well as on the molecular level, revealing a requirement for Tdrd9 in the amplification of piRNA levels without affecting the ping-pong signature of the remaining piRNAs. In addition, we show the Tdrd9 Tudor domain can localize to nuage and we describe a requirement for the DExH box domain for proper Tdrd9 function.

RESULTS

Tdrd9 is expressed in germ cells of the zebrafish

We performed RT-PCR to determine in what organs zebrafish *tdrd9* (ENSDARG00000013; Figure 1A) is expressed. Consistent with data from both mouse and *Drosophila*, we find gonad specific expression of *tdrd9*. Using *in situ* hybridization (ISH) we further defined the expression of *tdrd9* in the germ cells of the gonad, as well as in the primordial germ cells (PGCs) in the embryo (Figure 1B,C). Expression of *tdrd9* in the PGCs starts around three days post fertilization (dpf).

Disruption of *tdrd9* results in sterility

In an attempt to identify mutant alleles of Piwi pathway components we re-sequenced a set of ten genes in a collection of 36 mutants that all display a male sterility phenotype (Saito, 2011413)(Table S1). We identified one allele of *tdrd9*, disrupting the splice donor site of the first intron (*lb025*) (Figure 2A). In parallel, using a TILLING approach (Wienholds, 2004), we identified a missense mutation in *tdrd9*, changing the threonine of the highly conserved, DExH box associated, Walker A motif (GSGKT) into an isoleucine (*hu3177;* Figure 2A). The Walker A motif has been shown to be involved in ATP binding (Hishida, 1999), suggesting that this allele of *tdrd9* will produce crippled Tdrd9 protein. While in *tdrd9(hu3177)* animals *tdrd9* mRNA is readily detectable by *in situ* hybridization (ISH), the splice mutation results in loss of detectable Tdrd9 mRNA (Figure 2B), suggesting that *tdrd9(lb025)* triggers nonsense-mediated mRNA decay (Le Hir, 2008).

When homozygous, both alleles result in sterility due to an absence of germ cells, resulting in strict male development (Slanchev, 2005; Siegfried, 2008). Germ cell loss initiates around seven weeks of development (Figure 2C-E), accompanied by low levels of apoptosis (not shown). Similar phenotypes are observed in trans-heterozygous animals, showing that the phenotypes are caused by the lesions in *tdrd9*. This strengthens the idea that the ATPase function of the helicase domain is indeed important for Tdrd9 function *in vivo*.

Tdrd9 is required for germ cell maintenance

While most *tdrd9* mutant adults are sterile, we consistently observe a small fraction (approximately 10%) of the adult males to be sub-fertile. Oocytes fertilized by these males develop normally and homozygous *tdrd9* mutants derived from these males develop indistinguishably from *tdrd9* mutants obtained from heterozygous incrosses. Interestingly, when we analyze *tdrd9* mutant testes from 20 week- old adults by *vasa* ISH to visualize the germ cells, about 10-40% of these testes contain germ cells that are mostly spermatogonia-like in appearance (Figure 2C). Even at 30 weeks of age, some individuals maintained germ cells, although in these cases only spermatozoa-like cells were observed in the tubules (Figure 2D). These data strongly suggest that loss of *tdrd9* permits germ cell development at low frequency. However, in these cases germ cells seem to all differentiate into mature sperm, eventually depleting the germ cell population. Consistent with this notion, the sub-fertile animals cannot maintain fertility over longer periods (not shown).

Tdrd9 has impact on nuage formation

In *tdrd9* mutants, germ cells can be readily detected at three weeks of age.

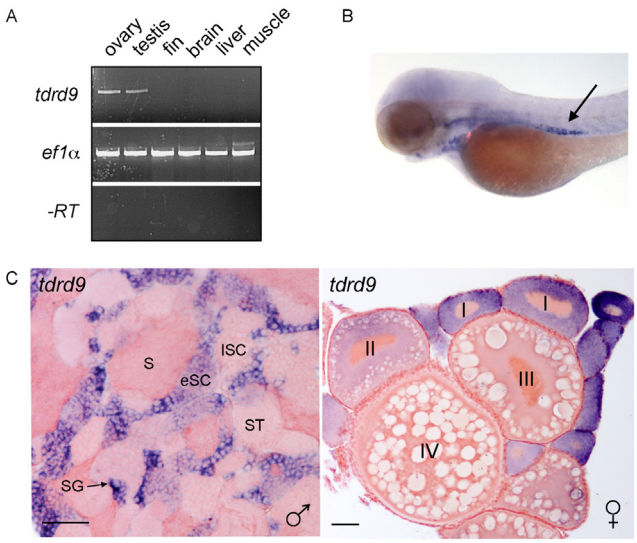


Figure 1
(A) Expression of *tdrd9* in different tissues as indicated, using RT-PCR. *Ef1-alpha* was used as control gene. (B) Whole mount *tdrd9* ISH of a 3dpf embryo. The black arrows indicates the location of the PGCs. (C) ISH for *tdrd9* on sections from testis and ovary. SG: spermatogonia; eSC: easly spermatocyes; lSC: late spermatocytes; ST: spermatids; S: sperm. I-IV: progressing stages of oogenesis. Scale bars: 50µm.

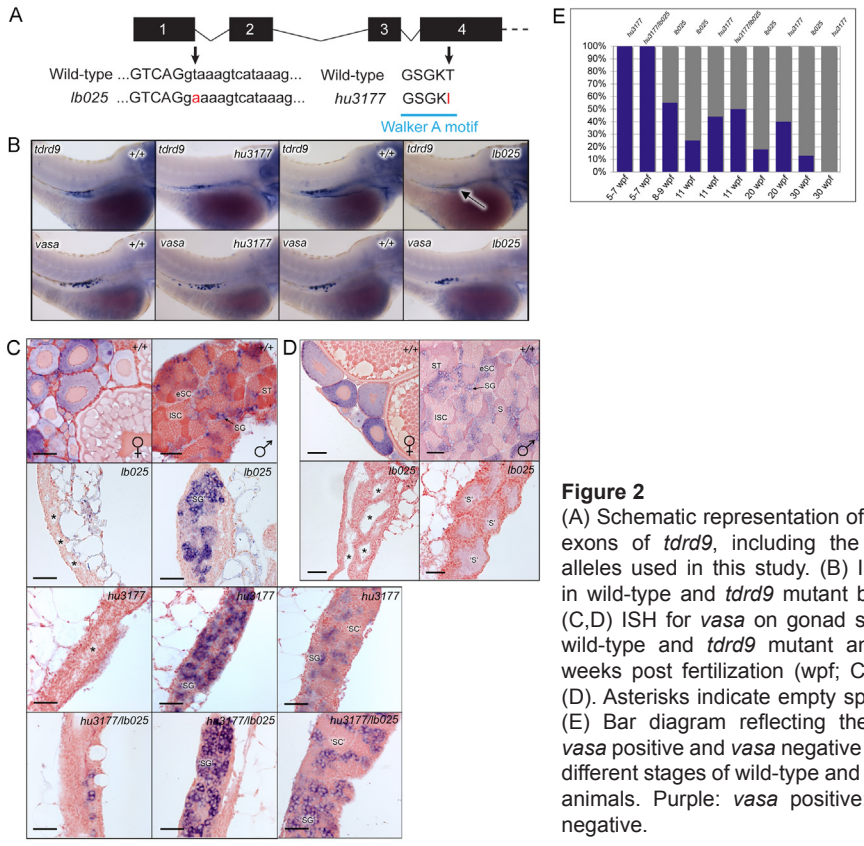


Figure 2
(A) Schematic representation of the first four exons of *tdrd9*, including the two mutant alleles used in this study. (B) ISH for *tdrd9* in wild-type and *tdrd9* mutant backgrounds. (C,D) ISH for *vasa* on gonad sections from wild-type and *tdrd9* mutant animals at 11 weeks post fertilization (wpf; C) and 30wpf (D). Asterisks indicate empty sperm tubules. (E) Bar diagram reflecting the number of *vasa* positive and *vasa* negative gonads from different stages of wild-type and *tdrd9* mutant animals. Purple: *vasa* positive; grey: *vasa* negative.

This allows us to determine the effect of Tdrd9 on the sub-cellular localization of Piwi pathway components. As can be seen in Figures 3A, Tdrd9 is not required for the global nuage-like localization of either Zili or Tdrd1, suggesting

that nuage formation is not disrupted and that Tdrd9 may act downstream of Tdrd1, as also suggested for TDRD9 in mice (Shoji, 2009; Vagin, 2009).

To look in more detail, we also used transmission electron microscopy (TEM) to analyze *tdrd9* mutant oocytes (Figure 3B). First, many wild-type looking patches of nuage are readily detectable in *tdrd9* mutants. This is consistent with the above-described Tdrd1 localization. However, next to these normally looking nuage structures we detect less-electron-dense structures in 38% (31/82) of the *tdrd9* mutant oocytes (Figure 3B), while these were not observed among eighty-six wild-type oocytes. At first sight, these nuage-like structures resemble those observed in *tdrd1* mutant oocytes (Huang, 2011). In addition, while in wild-type sections we found 39% (32/82) of the oocytes to contained associated synaptonemal complexes (Figure 3C), among the eighty-two *tdrd9* mutant oocytes analyzed none displayed such complexes (Figure 3C). This suggests *tdrd9* mutant oocytes are not able to enter or complete the zygotene/pachytene stage of meiosis, a phenotype also observed in mice (Shoji, 2009).

The Tdrd9 tudor domain interacts with Zili and localizes to nuage

Since tudor domains have been shown to mediate interaction with Piwi proteins we have studied the tudor domain of Tdrd9 in isolation. First, we expressed a GST-fused Tdrd9 tudor domain in *E. coli*. As the tudor domains of Tdrd proteins acting in the Piwi pathway often contain N-terminal extensions (Jin, 2009), we made two versions of this domain, one containing a predicted N-terminal beta-sheet and one without. We next incubated these purified domains, coupled to GST-beads, in testis extract and asked whether we could retrieve Zili. As shown in Figure 4A, Zili can be found on the Tdrd9 tudor domain containing the N-terminal beta-sheet, but not on the version lacking this N-terminal extension. Unfortunately, blotting for Ziwi revealed high a-specific binding of Ziwi, preventing interrogation of the interaction of this tudor domain with Ziwi. Nevertheless, these results indicate that Tdrd9 and Piwi proteins may interact via the tudor domain of Tdrd9.

Next, we also asked how this tudor domain localizes *in vivo*. For this we injected mRNA encoding GFP-tagged tudor domain, including the N-terminal extension, into one-cell-stage embryos. The mRNA included a *nos1* 3'UTR to restrict expression to the PGCs (Koprunner, 2001). This resulted in a punctate localization of the Tdrd9-tudor domain around the nucleus, strongly suggestive of nuage (Figure 4B). We further checked this by immuno-histochemistry, co-staining for GFP and Ziwi. Although the Tudor-GFP association with nuage appears to be rather unstable, given the high cytosolic signals in these experiments, this revealed co-localization of GFP and Ziwi (Figure 4C), confirming that the Tdrd9 tudor domain can localize to nuage.

Transposon transcripts are upregulated in *tdrd9* mutant gonads

We performed RT-qPCR to evaluate the levels of TE transcripts in *tdrd9* mutant gonads. We used gonads from five-week old animals as in these animals the global morphology has not yet been affected a lot. As can be seen in Figure 5, and consistent with our previously published data on Piwi pathway components, we can detect clear up-regulation of TE derived transcripts. This confirms the role for *tdrd9* in TE regulation.

Loss of Tdrd9 results in a global reduction of piRNA levels.

In order to study the impact of zebrafish *tdrd9* on piRNAs we derived small RNA libraries from wild-type and *tdrd9(lb025)* mutant gonads at three weeks of development (Table S2). In this analysis we have focused on the transposon derived piRNAs and have used miRNAs as an internal standard to determine relative piRNA levels. Reads derived from structural RNAs, like rRNA and tRNA are considered background and have not been further considered.

At a global level transposon-derived piRNA levels decrease approximately 1.5-fold upon loss of Tdrd9 (Figure 6A). Loss of piRNAs especially impacts on TEs replicating through an RNA intermediate (Figure 6B), the so-called retro-transposons. DNA-based TEs are much less affected. This is consistent with our previous findings (Huang, 2011; Kamminga, 2010), further strengthening the differential impact of Piwi pathway mutations on retro-TEs compared to DNA-TEs in zebrafish. For both types of transposons *tdrd9* has an equal impact on sense and anti-sense polarity piRNAs (Figure 6C), suggesting that *tdrd9* affects both Piwi proteins, Ziwi and Zili, similarly. Surprisingly, the ping-pong signal, defined as the percentage of sense/anti-sense read pairs that display an overlap of ten nucleotides, is still present in *tdrd9* mutant germ cells, and may even be stronger than in wild-type germ cells (Figure 6D).

DISCUSSION

In this study we have shown that zebrafish *tdrd9* functions in the Piwi-pathway and that it is essential for the development and maintenance of mature germ cells. The phenotypes associated with loss of Tdrd9 function are milder than those observed in mutants lacking Piwi proteins, suggesting that *tdrd9* mutation permits residual Piwi-pathway activity, although not enough to allow normal germ cell maturation. While consistent with previously reported data on Tdrd9 in mice and flies, our data add to the understanding of this protein in the Piwi-mediated RNA pathway in the following ways.

Tdrd9 helicase function

We provide strong indications that the ATPase function of Tdrd9 is required for its function: a mutant allele disrupting the highly conserved

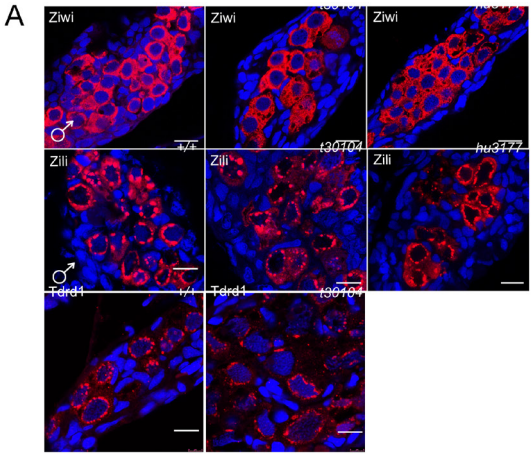


Figure 3 (Left panel)
(A) Immunohistochemistry for Ziwi, Zili and Tdrd1 on wild-type and *tdrd9* mutant gonads at five weeks of development. Blue: DAPI; Red: Ziwi, Zili or Tdrd1, as indicated. Scale bars: 10µm. (B) Transmission EM on wild-type and *tdrd9* mutant oocytes. *: normal looking nuage. **: nuage-like structure with lower electron-density, seen in 31/82 *tdrd9* mutant oocytes and in 0/86 wild-type oocytes. M: mitochondria. (C) Transmission EM on wild-type and *tdrd9* mutant oocytes. The black arrows in the wild-type panel indicate a synaptonemal complex. Such structures have been seen in 32/86 wild-type oocytes and in 0/82 *tdrd9* mutant oocytes.

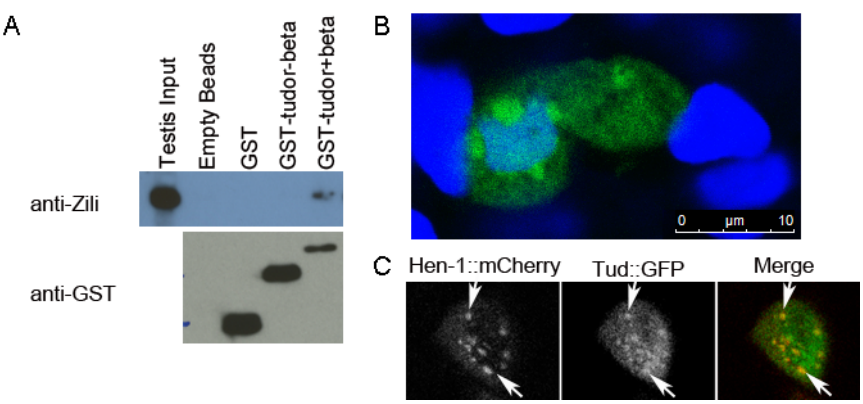


Figure 4 (Lower panel)
(A) GST-pulldown assay with GST-Tdrd9-tudor in testis extracts, followed by western blotting for Zili and GST. (B) Tdrd9-tudor-GFP in 24hpf PGC. Blue: DAPI, Green: Tdrd9-tudor-GFP. (C) Colocalization of Ziwi and Tdrd9-tudor-GFP. Blue: DAPI; Green: Tdrd9-tudor-GFP; Red: Ziwi. White arrows indicate patches of nuage.

Walker A motif induces a phenotype that is indistinguishable from that induced by a splice site mutation triggering strong mRNA destabilization. DExH box domains have been reported to be required for a number of possible functions. As their name suggests, these proteins can have RNA helicase activity, an activity that unwinds double stranded RNA structures. In addition, however, it has also been suggested that these modules act on protein-protein or protein-RNA interactions, where the energy release by ATP hydrolysis is used to remodel RNP complexes (Pan, 2010; Schwer, 2001). Both functions can be well envisaged in the Piwi-pathway. The enzyme Dicer, involved in dsRNA processing during si- and miRNA biogenesis, also contains a DExH motif. This motif in Dicer has been shown to play a role in the processing of thermodynamically unstable hairpin RNA duplexes (Soifer, 2008), suggesting it may have a true RNA helicase function. While the Piwi-pathway is believed to consume single stranded RNAs, RNA helicase activities may still be required to unwind double stranded structures to allow the activity of proteins acting on single stranded RNA. Remodeling interactions between proteins and RNA also seems a well-suited function for Tdrd9. Given that Tdrd9 affects both Ziwi and Zili-bound piRNAs it may play a role in restructuring RNPs containing both Ziwi and Zili, for example when bound by Tdrd1 (Huang, 2011).

Tdrd9 localization

We demonstrate that the tudor domain of Tdrd9 has affinity for nuage and can interact with Zili. Unfortunately, in the absence of functional antibodies and rescuing transgenes, despite several attempts, we are at present unable to further probe the subcellular localization of Tdrd9 and the interactions between Tdrd9 and Piwi proteins. In mouse, TDRD9 has been shown to interact with MIWI2 in P-body-like granules that have been named piP-bodies (Aravin, 2009; Shoji, 2009; Vagin, 2009), distinct from structures that harbor TDRD1 and MILI. It is certainly possible that this is true in zebrafish as well, but our observations that both Ziwi (Houwing, 2007) and the Tdrd9 tudor domain display nuage-like localization suggest that these bodies would be in very close contact with nuage. Interestingly, in *tdrd9* mutant oocytes we observe less electron-dense structures that could be homologous to the piP-bodies observed in mice. The fact we do not observe these in wild-type oocytes may indicate that normally the interaction between these bodies and nuage may be so close that it is hard to see by TEM. Alternatively, these bodies could be aggregates containing TE transcripts or proteins, as has been demonstrated in mice (Soper, 2008). Similar bodies have also been observed in *Tdrd9* mutant germ cells in mice (Shoji, 2009).

Mouse TDRD9 has been shown to be a nucleo-cytoplasmic protein, whose nuclear localization depends on MILI and TDRD1. While we have not been able to directly address this issue, our data do indicate Tdrd9 interacts

with Zili, and *tdrd9* mRNA is expressed at the time when Zili is nuclear (between three and five days post fertilisation). Our data are therefore consistent with the TDRD9 localization data in mice. However, we only observe defects in *tdrd9* mutant germ cells at time-points when Zili is no longer nuclear. This may indicate maternal rescue, most likely through maternally provided Tdrd9 protein as we do not detect *tdrd9* mRNA before 3dpf. Since we do observe a phenotype in germ cells later in development, when the bulk of Zili protein is found peri-nuclear, we conclude that Tdrd9 in zebrafish likely has a function outside the nucleus as well.

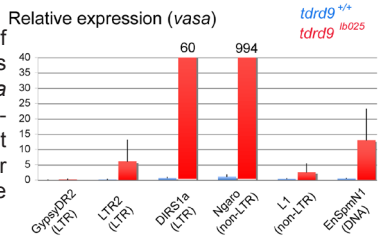


Figure 5
Relative expression of transposon transcripts compared to *vasa* mRNA in *tdrd9* wild-type and *hu3177* mutant gonads at 5wpf. Error bars show SD for three biological replicates.

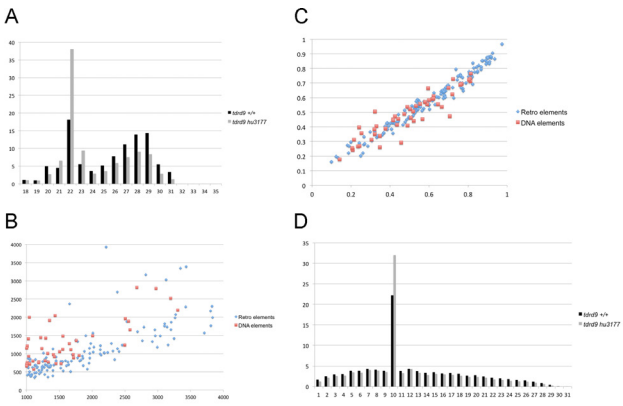


Figure 6
(A) Length distribution profiles of small RNAs cloned from wild-type and *tdrd9* mutant gonads at three weeks of development. (B) Scatter plot displaying read-abundance for different transposons in wild-type and *tdrd9* mutant gonads. (C) Scatter plot displaying the fraction of antisense polarity piRNAs for different transposons. (D) Ping-pong signatures for all transposon derived piRNAs combined derived from the piRNA cloned from wild-type and *tdrd9* mutant gonads.

Tdrd9 is required for piRNA amplification, but not for ping-pong

Our data show a clear effect of Tdrd9 on the accumulation of piRNAs. The piRNAs targeting most TEs are significantly reduced in level. Especially the retro-transposon derived piRNAs are strongly affected, like we have previously described for *hen1* (Kamminga, 2010) and *tdrd1* (Huang, 2011) mutants as well. Surprisingly, we find no effect of Tdrd9 on the molecular signature left by the typical piRNA amplification mechanism know as ping-pong (Brennecke, 2007). This is in contrast to observations in *Drosophila* (Malone, 2009), where Spindle E, the fruit-fly Tdrd9 homolog, leads to a total collapse of the ping-pong signature. This may indicate that both our zebrafish *tdrd9* alleles are hypomorphs, leaving significant Tdrd9 activity intact to support the observed piRNA amplification. We cannot rule this out, but we have two arguments to think differently. First, the splice-site allele of *tdrd9* leaves very little RNA that we can detect by RT-PCR. In fact, we have not

been able to amplify alternatively spliced mRNA molecules from this mutant background, while we did amplify non-spliced versions (data not shown). Second, piRNA levels are in fact affected, suggesting that our alleles are strong enough to leave footprints at the molecular level. We these reasons we also entertain an alternative explanation, and that is that this discrepancy may reflect a much stronger dependence of the fly on SpnE, compared to the need of Tdrd9 in zebrafish. However, in absence of a clearly defined function for Tdrd9/SpnE it remains difficult to speculate what the basis for such a difference could be.

Whatever the underlying reasons, our finding remains that while piRNA levels decline, the ping-pong signature remains fully intact. This is in fact consistent with the suggestion that Tdrd9 in both fish and mice functions downstream of Tdrd1 (this work and Shoji, 2009). Since Tdrd1 has been proposed to be a molecular platform for piRNA amplification through ping-pong mechanisms (Huang, 2011), a function of Tdrd9 downstream of Tdrd1 would imply no effect of Tdrd9 on ping-pong efficiencies. Even though it is at present not possible to place Tdrd9 at a discrete step in the Piwi pathway, a function in releasing Piwi proteins from the ping-pong environment would be consistent with all the obtained data so far. In that scenario, the detection of electron-weak nuage patches may reflect sites that are awaiting the arrival of newly created Piwi:piRNA complexes. Testing such a hypothesis would require, for example, the identification of proteins resident to these electron-weak patches.

The *tdrd9* mutant phenotype

The *tdrd9* mutant zebrafish do not display the same phenotype as either *zili* or *ziwi* mutants. In *tdrd9* mutants germ cells do develop and differentiate, although they have a meiotic defect and cannot be maintained, while in *piwi* mutant animals germ cells do not reach these stages (Houwing, 2008; Houwing, 2007). This may be because of partial maternal rescue of the *tdrd9* mutant embryos. Since we do not observe *tdrd9* mRNA before 3dpf, this would have to be through maternal inheritance of Tdrd9 protein, something that has been observed for Ziwi (Houwing, 2007). In absence of Tdrd9 antibodies this, however, cannot be tested. Alternatively, the phenotype triggered by loss of *tdrd9* may simply be less strong, as we described previously for *tdrd1* (Huang, 2011), a gene for which we have not been able to see either maternal mRNA or protein.

Interestingly, some *tdrd9* mutant individuals maintain germ cells until adulthood and can produce functional gametes at low frequency. Progeny resulting from these *tdrd9* mutant gametes develop normally, indicating that no major defects have occurred during the development of these specific gametes. However, the individuals making these gametes cannot maintain them, strongly suggesting Tdrd9 is required for the maintenance of a germ

cell population, possibly through a defect in spermatogonial stem cells preventing self-renewing divisions. This is reminiscent of the germ cell loss phenotype found in *Miwi2* mutant mice (Carmell, 2007), suggesting that this function for the Piwi pathway is conserved.

MATERIALS AND METHODS

Zebrafish strains and genotyping

Zebrafish were kept under standard conditions (Westerfield, 1993). *tdrd9(hu3177)* mutant allele was derived from ENU mutagenized libraries using target-selected mutagenesis as described (Wienholds, 2004); *tdrd9(lb025)* mutant allele was identified in a forward genetic screen for gonadogenesis mutants based on the absence of a germ line in adult mutant fish. Founders with mutant alleles *tdrd9(hu3177)* and *tdrd9(lb025)* were out crossed against wild-type fish (TL) or vas::EGFP transgenic fish (Krovel, 2004) and subsequently incrossed against each other to obtain homozygous offspring. For genotyping, the DNA was extracted from caudal fin tissue amputated from anesthetized fish. The primers used to amplify and re-sequence the alleles are: *tdrd9*_hu3177_F 5'- GAT TTC ACT ATA TCT CAT GTC -3', *tdrd9*_hu3177_R 5'- TGC AAA TTA ATT CCC ATT CC -3', *tdrd9*_lb025 F 5'- GCA TCA GGC CAT CAC GTC AGC -3' and *tdrd9*_lb025 R 5'- GTG CAG GTG TTA AAC CGC ATA AG -3'.

Reverse Transcriptase PCR

cDNA synthesis and RT-PCR were performed as described before (Houwing, 2007). Primers used to amplify *tdrd9* cDNA are Forward: 5'- CAT CCT GCA GGT GGT CAT TGC -3'. Reverse: 5'- CCC ACA TCG ATC ACC TCA GTA ATG -3';
elongation factor 1-alpha (*ef1-a;pha*) primers are: Forward: 5'- GGC CAC GTC GAC TCC GGA AAG TCC -3'. Reverse: 5'- CTC AAA ACG AGC CTG GCT GTA AGG -3'. Q-PCR was done as described before (Houwing, 2008).

Western blot and Immunohistochemistry

Western blot and immuno stainings were done as described before (Houwing, 2007). Zili and Ziwi antibodies were used as described before (Houwing, 2008;Houwing, 2007). Mouse anti-GFP B-2 (Santa Cruz) was used at 1:500. Mouse anti-GST was used at 1:5000.

In situ hybridization

In situ hybridization was performed as described before (Houwing, 2007). Full-length *tdrd9* cDNA (accession number EU888.1) was used for probe synthesis.

TEM

Sample preparation and imaging were performed as described before (Huang, 2011).

GST-pulldown

PCR products for Tdrd9 tudor domain with or without predicted N-terminal beta sheet were cloned into pGEX-KG vector containing GST tag at the 5' end of the multiple cloning site. Primers used to amplify the respective Tdrd9 fragments are: Forward, short N-terminus: 5'- cgggatccGAGAAACAGTGTCAATTAACAGC -3'. Forward, extended: 5'- cgggatccGATTTCGTCATCAACATTACTGAG -3. Reverse, both clones: 5'- cggaattcCTGCGCTTGAAAGGCTGG -3. The pull down assay was performed as described before (Huang, 2011).

Microinjection of mRNA into embryos

The Tdrd9 tudor domain, with extended N-terminal beta-sheet, as described in the GST-pull-down section, was cloned into the GATEWAY system (Invitrogen), and recombined with a vector harboring the Sp6 promoter and GFP-nos3'UTR sequences, yielding a Sp6-Tudor-GFP-nos3'UTR construct. mRNAs were transcribed *in vitro* using the Sp6 RNA polymerase following manufacturers instructions. Solutions containing 150 – 200 ng/µl mRNA and 10% phenol red were injected into 1-cell stage wild-type embryos using glass micropipettes. Injected embryos were maintained in the 28 ^{o}C incubator till they reach 24hpf stage.

Small RNA libraries and sequencing

Library preparation was essentially as described before (Huang, 2011). In short: Total or RNA isolated from immune precipitations was size selected on a 15% denaturing poly-acrylamide gel (18-32 nt). Following this, 3' and 5' adapters were ligated, after which the ligated RNA was reverse transcribed into cDNA. This cDNA was amplified with 16-18 cycles of PCR and sequenced on a Solexa platform.

Small RNA analysis

Bioinformatic analysis was performed as previously described (Huang, 2011).

ACKNOWLEDGEMENTS

We thank Marjo den Broeder for technical assistance. We thank Drs Cuppen and Plasterk (Hubrecht Institute) and Dr Stemple (Welcome Trust Sanger Institute) for providing a *tdrd9* mutant allele, which was generated as part of the ZF-MODELS Integrated Project in the 6th Framework Programme (Contract No. LSHG-CT-2003-503496) funded by the European Commission. This work was supported by grants to RFK: a VIDI (864.05.010) and ECHO (700.57.006) grant from the Netherlands Organization for Scientific Research, an ERC Starting Grant from the Ideas Program of the European

Union Seventh Framework Program (202819) and the European Union Sixth Framework Program Integrated Project SIROCCO (LSHG-CT-2006-037900).

5

Chapter 6

General Discussion

6

The PIWI-piRNA pathway is a defense mechanism against transposable elements (TEs) to maintain the genome integrity of germ cells. It mainly functions in the germline and, in some cases, somatic cells in the gonads. Many germline tudor-domain containing (Tdrd) proteins are interacting partners of PIWIs and are also important in germline maintenance and function. Drosophila Tudor was the first Tdrd protein identified and is essential for germ plasm formation (Thomson and Lasko, 2005). Later more Tdrds were found to play a role in spermatogenesis as well as nuage formation in mice (Siomi et al., 2010). Recently, the finding of sDMA sites at the N-terminal end of various PIWI proteins links the Tdrd-PIWI interactions, and the tudor-sDMA coding seems to be germline-specific. Here I will summarize the findings presented in this thesis and discuss the relationships between Tdrd1, Tdrd6a, Tdrd9 and the PIWI-piRNA pathway in zebrafish.

Tdrds and Nuage/Germ plasm assembly

The nuage is a conserved electron-dense structure often found in the perinuclear region or associated with mitochondria in germ cells (Eddy, 1974). It is generally considered as the location for the PIWI-piRNA pathway. Tdrds and the PIWI-piRNA pathway have both been described to be involved in nuage assembly (Reviewed in Siomi et al., 2010; Siomi et al., 2011). The PIWI-Tdrd interaction is, in some cases, essential for nuage localization of PIWIs (Siomi et al., 2010; this thesis). In mice, TDRD1, TDRD6 and TDRD7 are essential for the assembly of nuage (Hosokawa et al., 2007; Tanaka et al., 2011). Nuage localization of MILI is dependent on interacting with TDRD1 (Vagin et al., 2009). In zebrafish, Tdrd1 is essential for nuage formation in the oocytes. We observed that nuage is less in number and electron density as well as mis-localized in *tdrd1* mutants. One of the nuage components, Zili, lost nuage localization but its perinuclear localization is little affected in *tdrd1* mutants. This suggests that perinuclear localization of Zili is independent of Tdrd1. (Huang et al., 2011). In tdrd9 mutants some of the nuage appeared less electron dense in the juvenile oocytes, but Ziwi and Zili localization seem not to be affected during gametogenesis (Chapter 5), implying that Tdrd9 is downstream of Ziwi and Zili in nuage assembly and tat the “missing” electron density may represent the Tdrd9-containing sub compartment of nuage. Loss of Tdrd6a also does not impact PIWI and Tdrd1 localization during gametogenesis, but it does disrupt the germ plasm assembly in the early PGCs where Tdrd1 and Ziwi expression are affected. Tdrd6a probably acts downstream of Buc, a protein important for Balbiani body and germ plasm assembly in zebrafish. Components of the germ plasm require the presence of Tdrd6a in order to be recruited by Buc for proper germ plasm formation (Chapter 4). There are three paralogs of *tdrd6* encoded in the zebrafish genome. Since no defects were detected after PGC stage in tdrd6a mutants,

the three paralogs may function partially redundant, and act in nuage/germ plasm assembly at different developmental time points (Chapter 4).

Tdrds in transposon regulation and piRNA biogenesis

Many Tdrds have been reported to be involved in piRNA biogenesis, either the primary processing or the secondary Ping-Pong amplification cycle. In mice and zebrafish, the secondary pathway is better understood. In many cases, Tdrds are involved in transposon regulation, which is the main function known so far of the Ping-Pong cycle. In mice, loss of TDRD1, TDRD5, TDRD7 or TDRD9 but not TDRD6 leads to an increase of transposon transcripts accompanied by defective spermatogenic phenotypes, and consequently sterile animals (Chuma et al., 2006; Reuter et al., 2009; Shoji et al., 2009; Vagin et al., 2009; Vasileva et al., 2009; Wang et al., 2009; Yabuta et al., 2011). In zebrafish, *tdrd1*, *tdrd6a* and *tdrd9* mutants all display transposon upregulation ((Huang et al., 2011), Chapter 4 and 5), suggesting the conserved roles of Tdrds in controlling transposons in the germline. Loss of zebrafish *tdrd6a*, unlike loss of mouse TDRD6, leads to transposon derepression. This is likely due to the earlier expression of zebrafish Tdrd6a compared to mouse TDRD6 during gametogenesis.

In zebrafish, *tdrd1* and *tdrd9* mutants start to lose germ cells at 4 and 7 week-post-fertilization (wkpf) ((Huang et al., 2011) and Chapter 5), respectively, which are relatively mild phenotypes compared to PIWI mutants (Houwing et al., 2008; Houwing et al., 2007). This suggests that *tdrd9* and *tdrd1* mutants allow a certain degree of PIWI-piRNA activity but not enough to maintain the germline. In both *tdrd1* and *tdrd9* mutants, class I element (retro-element)-derived piRNA populations are largely reduced, while class II element-derived piRNAs are not significantly affected, a finding consistent with earlier reports (Houwing et al., 2008; Houwing et al., 2007) and our findings in Chapter 2 that retroelements are the main target of the Ping-Pong cycle. Furthermore, the sensitivity of retro-element-derived piRNAs in *tdrd1* mutants is correlated with their abundance at 3wkpf, the stage that the piRNA populations were examined. In many cases, the highly sensitive piRNAs are very abundant at 3wkpf (Chapter 2), confirming the role of Tdrd1 in facilitating the Ping-Pong cycle (Huang et al., 2011). We also identified Tdrd1-associated transcripts (TATs), which are good candidates for piRNA biogenesis intermediates. The association of Tdrd1 with both PIWI proteins, mature piRNAs and TATs indicates Tdrd1 functions as a molecular platform for the Ping-Pong cycle (Huang et al., 2011).

Although tdrd9 mutants display a global reduction of piRNA levels, the Ping-Pong signature of the remaining piRNAs is not affected. This suggests that Tdrd9 is not required for the Ping-Pong cycle. Instead, a suggested function of Tdrd9 is to remodel RNPs containing Ziwi, Zili, and Tdrd1 and release PIWI proteins from the RNP complex in the Ping-Pong cycle via its

DExH domain (Chapter 5).

Transposon up-regulation is generally thought to be the cause of sterility in the PIWI-piRNA pathway mutants. However, although loss of Tdrd6a leads to
upregulation of transposons in the adult testes, both male and female mutants are fertile. There are two possible explanations. First, there could be other yet unknown factors underlying the fertility defects besides only transposon upregulation. For example, the less-electron dense nuage found in *tdrd1* and *tdrd9* mutants suggests disrupted nuage assembly. It is possible that conserved nuage structures have other functions besides hosting the PIWI-piRNA pathway in the germline. Second, there may be other highly derepressed transposons, which are not examined in our Q-PCR assays, causing the sterility phenotypes in *tdrd1* and *tdrd9* but not *tdrd6a* mutants.

Based on the small reduction of Ping-Pong signature of piRNAs in *tdrd6a* mutants, Tdrd6a is likely playing a minor role in the amplification cycle. Loss of Tdrd6a results in loss of anti-sense strand bias of piRNAs, therefore Tdrd6a is likely to act as an anti-sense piRNA reservoir for Ziwi. Both Ziwi and Tdrd6a are maternally provided, and Tdrd6a may actually facilitate the deposition of maternal piRISCs, containing Ziwi and mostly anti-sense piRNAs, to defend the early germ cell lineage from retrotransposon activities (Chapter 2 and 4).

Tdrds in compartmentalization of nuage

In mice, MILI and MIWI2 are the main players of the Ping-Pong cycle, while they reside in different compartments of nuage-like structures together with their co-factors. MILI together with TDRD1, MVH and GASZ localize in the pi-body, while MIWI2 together with TDRD9, MAEL and several known P-body components localize in the piP-body (Aravin et al., 2009). In zebrafish, it is likely that similar compartments of nuage are also present. There are several lines of evidence to support this hypothesis. First, using immunohistochemistry assays, both Tdrd1 and Tdrd6a display circular perinuclear granules during early gametogenesis, which suggests there may be different compartments in the granular structure of nuage. Second, according to our mass spectrometry analyses of immunoprecipitates, Tdrd1 interacts with Ziwi, Zili and Tdrd5, while Tdrd6a interacts preferentially with Ziwi and Tdrd7. Tdrd5 and Tdrd7 are the only two known tudor proteins with three LOTUS domains. Although the functions of LOTUS domains are unclear, Tdrd5 and Tdrd7 may have conserved function through these domains. It is therefore likely that the perinuclear structures containing Tdrd1 and Tdrd6a are actually distinct compartments of nuage. In the Tdrd6a IP mass spectrometry analysis, many known components of P-bodies, such as PABPC1, el4FE-T and PAT1 were also identified. Thus, Zili/Ziwi/Tdrd1/Tdrd5-containing granules may resemble the pi-body, which are mainly responsible

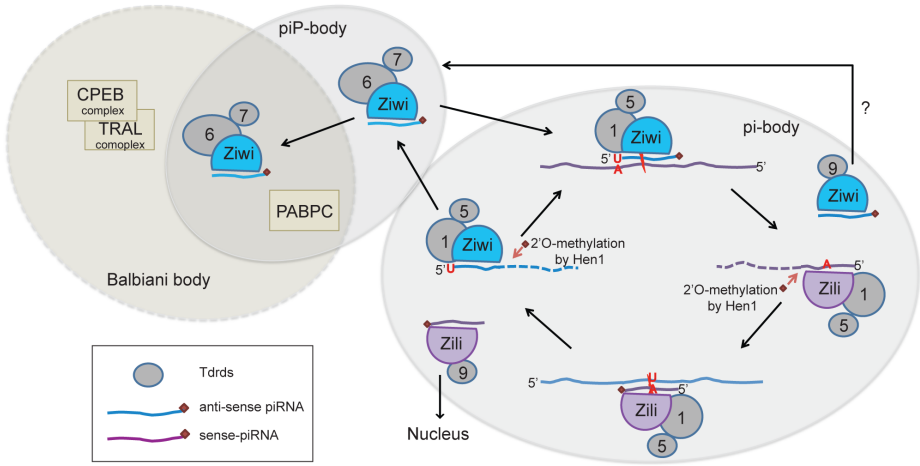


Figure 1. Schematic picture of zebrafish Tdrds in PIWI-piRNA pathway.
Ziwi/Tdrd6 are providing maternal Ziwi-associated piRNA-induced silencing complexes (piRISCs), which is loaded with anti-sense piRNAs. The maternal piRISCs and piRNA transcribed from piRNA clusters are the driving force of the Ping-Pong cycle. Ping-Pong cycle is probably happened in the pi-body composed of Ziwi/Zili/Tdrd1/Tdrd5 and Tdrd9. Tdrd1 facilitates the Ping-Pong cycle while Tdrd9 helps PIWI proteins to release from the Ziwi/Zili/Tdrd1 RNP. Released PIWI proteins may recycle in the pi-body and target other TE transcripts. Or they may be transported to the nucleus or piP-body (as indicated with the arrows) for yet unknown functions. piP-body contains Ziwi/Tdrd6/Tdrd7 and PABPC, which is a GW182 interactor, and may have a function in mRNA degradation in the germline. In the ovary, piP-body is possibly embedded in the Balbiani body and function together with CPEB and TRAL complexes in translational regulation. Numbers in the gray circles are 1: Tdrd1; 5: Tdrd5; 6:Tdrd6; 7:Tdrd7; 9:Tdrd9.

for performing Ping-Pong amplification cycle. Whereas Ziwi/Tdrd6/Tdrd7 are present in the piP-body-like granules and are involved in mRNA degradation. These two structures were not distinguishable until the piP-body-like less electron-dense nuage was observed in the tdrd9 mutants, suggesting that in the wild-type situation these two structures are closely associated with each other. It also implies that Tdrd9 is likely present in the pi-body together with Ziwi/Zili/Tdrd1/Tdrd5. A model concerning the involvement of Tdrds in the compartmentalization of nuage and Ping-Pong cycle is illustrated in Fig1.

Beyond transposon silencing

Besides transposon regulation in the germline, PIWI pathway components have been reported to be involved in other processes during germline development. MIWI, TDRD5 and RNF17 may function together to stabilize the mRNA of ACT and other CREM target genes by forming a complex with them, to regulate the expression of these spermiogenic genes (Deng and Lin, 2002; Yabuta et al., 2011). In Chapter 4, we demonstrated that Tdrd6a interacts with many proteins involved in RNA translational regulation, such as components of CPEB, TRAL as well as the EJC complexes, particularly in the adult ovary but not testis. It has been reported that nuage

is embedded in the Balbiani body in *Xenopus* and zebrafish (Kloc et al., 2004; Marlow and Mullins, 2008). *Xenopus* Xiwi as well as zebrafish Tdrd6a protein and Ziwi mRNA are all located in the Balbiani body ((Lau et al., 2009); Houwing, unpublished data; this thesis, Chapter 4). It is therefore possible that Tdrd6a together with Ziwi are actually regulating mRNA expression within the Balbiani body. The fact that many components in the CPEB, TRAL complexes are also present in the P-body, we cannot rule out the possibility that Ziwi/Tdrd6a-containing piP-body is actually embedded in the Balbiani body (Chapter 4).

Tandem tudor repeats and arginine methylation

Although mass spectrometry analyses have identified several (clusters of) arginine methylation sites on various PIWI proteins, it has been shown that their interacting Tdrds only recognize specific arginine methylation sites. For instance, in mice, TDRKH only recognizes the first RG-rich cluster of MIWI (Chen et al., 2009) and in *Drosophila*, Tudor only recognizes the second RG-rich cluster of AGO3 (Nishida et al., 2009). In zebrafish, we found that Tdrd1 differentially interacts with four sDMA sites on the N-terminus of Zili. Among these four sites, two of them are close to each other and display much stronger interaction with Tdrd1 relative to the other two sDMAs (Huang et al., 2011). All the above indicate that there is some specificity between the tudor-sDMA interaction, either determined by the surrounding sequences or the regional electrostatic/hydrophobic interactions. Multiple sDMA sites or RG-rich clusters on PIWI proteins are therefore likely to be recognized by different Tdrd proteins. Arginine methylation on different locations within a PIWI protein may represent the markers of distinct spatial and/or temporal modes of the protein. Although it is technically difficult to isolate PIWI proteins from different cellular or sub-cellular compartments, such studies would help to understand many phenomena that we are not able to explain at the moment.

Closer look at the piRNAs during germline development

So far most of our understanding of the PIWI-piRNA pathway in zebrafish, particularly piRNA biogenesis, is based on findings in the adult system. In

Chapter 2, we illustrated that piRNA profiles are actually not constant during development of the germline. piRNAs appeared shorter at PGC stages. Based on the analyses regarding strand bias, Ping-Pong signature and the piRNA population from different classes of transposons, it is clear that piRNAs are highly active in targeting active transposon transcripts during the early germline development. However, the differential expression of certain transposon-derived piRNAs during development remains an open question. To investigate the underlying reasons will bring more insights into other

possible functions of piRNAs during various stages of development.

Appendix

Appendix

References

Allis, C.D., Underwood, E.M., Caulton, J.H., and Mahowald, A.P. (1979). Pole cells of Drosophila melanogaster in culture. Normal metabolism, ultrastructure, and functional capabilities. Dev Biol *69*, 451-465.

Ameres, S.L., Horwich, M.D., Hung, J.H., Xu, J., Ghildiyal, M., Weng, Z., and Zamore, P.D. (2010). Target RNA-directed trimming and tailing of small silencing RNAs. Science *in press*.

Aoki, Y., Nagao, I., Saito, D., Ebe, Y., Kinjo, M., and Tanaka, M. (2008). Temporal and spatial localization of three germline-specific proteins in medaka. Dev Dyn *237*, 800-807.

Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., Morris, P., Brownstein, M.J., Kuramochi-Miyagawa, S., Nakano, T*., et al.* (2006). A novel class of small RNAs bind to MILI protein in mouse testes. Nature *442*, 203-207.

Aravin, A.A., and Hannon, G.J. (2008). Small RNA silencing pathways in germ and stem cells. Cold Spring Harb Symp Quant Biol *73*, 283-290.

Aravin, A.A., Hannon, G.J., and Brennecke, J. (2007a). The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. Science *318*, 761-764.

Aravin, A.A., Naumova, N.M., Tulin, A.V., Vagin, V.V., Rozovsky, Y.M., and Gvozdev, V.A. (2001). Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the D. melanogaster germline. Current biology : CB *11*, 1017-1027.

Aravin, A.A., Sachidanandam, R., Bourc'his, D., Schaefer, C., Pezic, D., Toth, K.F., Bestor, T., and Hannon, G.J. (2008). A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. Mol Cell *31*, 785-799.

Aravin, A.A., Sachidanandam, R., Girard, A., Fejes-Toth, K., and Hannon, G.J. (2007b). Developmentally regulated piRNA clusters implicate MILI in transposon control. Science *316*, 744-747.

Aravin, A.A., van der Heijden, G.W., Castaneda, J., Vagin, V.V., Hannon, G.J., and Bortvin, A. (2009). Cytoplasmic compartmentalization of the fetal piRNA pathway in mice. PLoS Genet *5*, e1000764.

Arkov, A.L., Wang, J.Y., Ramos, A., and Lehmann, R. (2006). The role of Tudor domains in germline development and polar granule architecture. Development *133*, 4053-4062.

Bentov, Y., Yavorska, T., Esfandiari, N., Jurisicova, A., and Casper, R.F. (2011). The contribution of mitochondrial function to reproductive aging. J Assist Reprod Genet *28*, 773-783.

Blaser, H., Eisenbeiss, S., Neumann, M., Reichman-Fried, M., Thisse, B., Thisse, C., and Raz, E. (2005). Transition from non-motile behaviour to directed migration during early PGC development in zebrafish. J Cell Sci *118*, 4027-4038.

Blendy, J.A., Kaestner, K.H., Weinbauer, G.F., Nieschlag, E., and Schutz, G. (1996). Severe impairment of spermatogenesis in mice lacking the CREM gene. Nature *380*, 162-165.

Boisvert, F.M., Dery, U., Masson, J.Y., and Richard, S. (2005). Arginine methylation of MRE11 by PRMT1 is required for DNA damage checkpoint control. Genes Dev *19*, 671-676.

Bontems, F., Stein, A., Marlow, F., Lyautey, J., Gupta, T., Mullins, M.C., and Dosch, R. (2009). Bucky ball organizes germ plasm assembly in zebrafish. Curr Biol *19*, 414-422.

Braat, A.K., Speksnijder, J.E., and Zivkovic, D. (1999). Germ line development in fishes. Int J Dev Biol *43*, 745-760.

Brangwynne, C.P., Eckmann, C.R., Courson, D.S., Rybarska, A., Hoege, C., Gharakhani, J., Julicher, F., and Hyman, A.A. (2009). Germline P granules are liquid droplets that localize by controlled dissolution/condensation. Science *324*, 1729-1732.

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.

Brennecke, J., Malone, C.D., Aravin, A.A., Sachidanandam, R., Stark, A., and Hannon, G.J. (2008). An epigenetic role for maternally inherited piRNAs in transposon silencing. Science *322*, 1387-1392.

Brower-Toland, B., Findley, S.D., Jiang, L., Liu, L., Yin, H., Dus, M., Zhou, P., Elgin, S.C., and Lin, H. (2007). Drosophila PIWI associates with chromatin and interacts directly with HP1a. Genes Dev *21*, 2300-2311.

Callebaut, I., and Mornon, J.P. (2010). LOTUS, a new domain associated with small RNA pathways in the germline. Bioinformatics *26*, 1140-1144.

Carmell, M.A., Girard, A., van de Kant, H.J., Bourc'his, D., Bestor, T.H., de Rooij, D.G., and Hannon, G.J. (2007). MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. Dev Cell *12*, 503-514.

Carmell, M.A., Xuan, Z., Zhang, M.Q., and Hannon, G.J. (2002). The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. Genes Dev *16*, 2733-2742.

Caudy, A.A., Ketting, R.F., Hammond, S.M., Denli, A.M., Bathoorn, A.M., Tops, B.B., Silva, J.M., Myers, M.M., Hannon, G.J., and Plasterk, R.H. (2003). A micrococcal nuclease homologue in RNAi effector complexes. Nature *425*, 411-414.

Cavey, M., Hijal, S., Zhang, X., and Suter, B. (2005). Drosophila valois encodes a divergent WD protein that is required for Vasa localization and Oskar protein accumulation. Development *132*, 459-468.

Chen, C., Jin, J., James, D.A., Adams-Cioaba, M.A., Park, J.G., Guo, Y., Tenaglia, E., Xu, C., Gish, G., Min, J.*, et al.* (2009a). Mouse Piwi interactome identifies binding mechanism of Tdrkh Tudor domain to arginine methylated Miwi. Proc Natl Acad Sci U S A *106*, 20336-20341.

Chen, D., Yu, Y., Lin, H., Huang, P., Shan, Z., and Wang, Y. (2010). Ultraviolet-blue to near-infrared downconversion of Nd(3+)-Yb(3+) couple. Optics letters *35*, 220-222.

Chen, X., Liu, J., Cheng, Y., and Jia, D. (2002). HEN1 functions pleiotropically in Arabidopsis development and acts in C function in the flower. Development *129*, 1085-1094.

Chen, Y.R., Solter, L.F., Chien, T.Y., Jiang, M.H., Lin, H.F., Fan, H.S., Lo, C.F., and Wang, C.H. (2009b). Characterization of a new insect cell line (NTU-YB) derived from the common grass yellow butterfly, Eurema hecabe (Linnaeus) (Pieridae: Lepidoptera) and its susceptibility to microsporidia. Journal of invertebrate pathology *102*, 256-262.

Chow, J.C., Ciaudo, C., Fazzari, M.J., Mise, N., Servant, N., Glass, J.L., Attreed, M., Avner, P., Wutz, A., Barillot, E.*, et al.* (2010). LINE-1 activity in facultative heterochromatin formation during X chromosome inactivation. Cell *141*, 956-969.

Chuma, S., Hiyoshi, M., Yamamoto, A., Hosokawa, M., Takamune, K., and Nakatsuji, N. (2003). 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.

Chuma, S., Hosokawa, M., Kitamura, K., Kasai, S., Fujioka, M., Hiyoshi, M., Takamune, K., Noce, T., and Nakatsuji, N. (2006). 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 U S A *103*, 15894-15899.

Chuma, S., Hosokawa, M., Tanaka, T., and Nakatsuji, N. (2009). Ultrastructural characterization of spermatogenesis and its evolutionary conservation in the germline: germinal granules in mammals. Mol Cell Endocrinol *306*, 17-23.

Cook, H.A., Koppetsch, B.S., Wu, J., and Theurkauf, W.E. (2004). The Drosophila SDE3 homolog armitage is required for oskar mRNA silencing and embryonic axis specification. Cell *116*, 817-829.

Cordin, O., Banroques, J., Tanner, N.K., and Linder, P. (2006). The DEAD-box protein family of RNA helicases. Gene *367*, 17-37.

Cote, J., and Richard, S. (2005). Tudor domains bind symmetrical dimethylated arginines. The Journal of biological chemistry *280*, 28476-28483.

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.

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.

Cox, R.T., and Spradling, A.C. (2003). A Balbiani body and the fusome mediate mitochondrial inheritance during Drosophila oogenesis. Development *130*, 1579-1590.

Czech, B., and Hannon, G.J. (2011). Small RNA sorting: matchmaking for Argonautes. Nat Rev Genet *12*, 19-31.

de Sousa Lopes, S.M., and Roelen, B.A. (2010). An overview on the diversity of cellular organelles during the germ cell cycle. Histol Histopathol *25*, 267-276.

de Sousa Lopes, S.M., Roelen, B.A., Monteiro, R.M., Emmens, R., Lin, H.Y., Li, E., Lawson, K.A., and Mummery, C.L. (2004). BMP signaling mediated by ALK2 in the visceral endoderm is necessary for the generation of primordial germ cells in the mouse embryo. Genes Dev *18*, 1838-1849.

Deng, W., and Lin, H. (2002). miwi, a murine homolog of piwi, encodes a cytoplasmic protein essential for spermatogenesis. Dev Cell *2*, 819-830.

Dennis, R.A., Wang, J., Huang, K., Helsley, A., and Robinson, A.G. (2006). The Central Codebook (CCB) at the David Geffen School of Medicine at UCLA. AMIA Annu Symp Proc, 908.

Desset, S., Meignin, C., Dastugue, B., and Vaury, C. (2003). COM, a heterochromatic locus governing the control of independent endogenous retroviruses from Drosophila melanogaster. Genetics *164*, 501-509.

Dosch, R., Wagner, D.S., Mintzer, K.A., Runke, G., Wiemelt, A.P., and Mullins, M.C. (2004). Maternal control of vertebrate development before the midblastula transition: mutants from the zebrafish I. Dev Cell *6*, 771-780.

Eddy, E.M. (1974). Fine structural observations on the form and distribution of nuage in germ cells of the rat. Anat Rec *178*, 731-757.

Elbashir, S.M., Lendeckel, W., and Tuschl, T. (2001). RNA interference is mediated by 21- and 22-nucleotide RNAs. Genes Dev *15*, 188-200.

Ephrussi, A., and Lehmann, R. (1992). Induction of germ cell formation by oskar. Nature *358*, 387-392.

Farazi, T.A., Juranek, S.A., and Tuschl, T. (2008). The growing catalog of small RNAs and their association with distinct Argonaute/Piwi family members. Development *135*, 1201-1214.

Fenske, M., Maack, G., Schafers, C., and Segner, H. (2005). An environmentally relevant concentration of estrogen induces arrest of male gonad development in zebrafish, Danio rerio. Environ Toxicol Chem *24*, 1088-1098.

Feschotte, C., and Pritham, E.J. (2007). DNA transposons and the evolution of eukaryotic genomes. Annu Rev Genet *41*, 331-368.

Findley, S.D., Tamanaha, M., Clegg, N.J., and Ruohola-Baker, H. (2003). Maelstrom, a Drosophila spindle-class gene, encodes a protein that colocalizes with Vasa and RDE1/AGO1 homolog, Aubergine, in nuage. Development *130*, 859-871.

Flicek, P., Aken, B.L., Beal, K., Ballester, B., Caccamo, M., Chen, Y., Clarke, L., Coates, G., Cunningham, F., Cutts, T*., et al.* (2008). Ensembl 2008. Nucleic Acids Res *36*, D707-714.

Forstemann, K., Tomari, Y., Du, T., Vagin, V.V., Denli, A.M., Bratu, D.P., Klattenhoff, C., Theurkauf, W.E., and Zamore, P.D. (2005). Normal microRNA maturation and germ-line stem cell maintenance requires Loquacious, a double-stranded RNA-binding domain protein.

PLoS Biol *3*, e236.

Friberg, A., Corsini, L., Mourao, A., and Sattler, M. (2009). Structure and ligand binding of the extended Tudor domain of D. melanogaster Tudor-SN. Journal of molecular biology *387*, 921-934.

Gary, J.D., and Clarke, S. (1998). RNA and protein interactions modulated by protein arginine methylation. Prog Nucleic Acid Res Mol Biol *61*, 65-131.

Ghildiyal, M., Seitz, H., Horwich, M.D., Li, C., Du, T., Lee, S., Xu, J., Kittler, E.L., Zapp, M.L., Weng, Z.*, et al.* (2008). Endogenous siRNAs derived from transposons and mRNAs in Drosophila somatic cells. Science *320*, 1077-1081.

Ghildiyal, M., and Zamore, P.D. (2009). Small silencing RNAs: an expanding universe. Nat Rev Genet *10*, 94-108.

Gillespie, D.E., and Berg, C.A. (1995). Homeless is required for RNA localization in Drosophila oogenesis and encodes a new member of the DE-H family of RNA-dependent ATPases. Genes Dev *9*, 2495-2508.

Giorgi, C., and Moore, M.J. (2007). The nuclear nurture and cytoplasmic nature of localized mRNPs. Semin Cell Dev Biol *18*, 186-193.

Girard, A., and Hannon, G.J. (2007). [The role of piRNAs in mouse spermatogenesis.]. J Soc Biol *201*, 411-418.

Girard, A., and Hannon, G.J. (2008). Conserved themes in small-RNA-mediated transposon control. Trends Cell Biol *18*, 136-148.

Girard, A., Sachidanandam, R., Hannon, G.J., and Carmell, M.A. (2006). A germline-specific class of small RNAs binds mammalian Piwi proteins. Nature *442*, 199-202.

Gonzalez-Reyes, A., Elliott, H., and St Johnston, D. (1997). Oocyte determination and the origin of polarity in Drosophila: the role of the spindle genes. Development *124*, 4927-4937.

Grivna, S.T., Pyhtila, B., and Lin, H. (2006). MIWI associates with translational machinery and PIWI-interacting RNAs (piRNAs) in regulating spermatogenesis. Proc Natl Acad Sci U S A *103*, 13415-13420.

Gunawardane, L.S., Saito, K., Nishida, K.M., Miyoshi, K., Kawamura, Y., Nagami, T., Siomi, H., and Siomi, M.C. (2007). A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila. Science *315*, 1587-1590.

Guraya, S.S. (1979). Recent advances in the morphology, cytochemistry, and function of Balbiani's vitelline body in animal oocytes. Int Rev Cytol *59*, 249-321.

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.

Handler, D., Olivieri, D., Novatchkova, M., Gruber, F.S., Meixner, K., Mechtler, K., Stark, A., Sachidanandam, R., and Brennecke, J. (2011). A systematic analysis of Drosophila TUDOR domain-containing proteins identifies Vreteno and the Tdrd12 family as essential primary piRNA pathway factors. EMBO J.

Harris, A.N., and Macdonald, P.M. (2001). Aubergine encodes a Drosophila polar granule component required for pole cell formation and related to eIF2C. Development *128*, 2823-2832.

Hashimoto, Y., Maegawa, S., Nagai, T., Yamaha, E., Suzuki, H., Yasuda, K., and Inoue, K. (2004). Localized maternal factors are required for zebrafish germ cell formation. Dev Biol *268*, 152-161.

Hayashi, K., Kobayashi, T., Umino, T., Goitsuka, R., Matsui, Y., and Kitamura, D. (2002). SMAD1 signaling is critical for initial commitment of germ cell lineage from mouse epiblast. Mech Dev *118*, 99-109.

Hishida, T., Iwasaki, H., Yagi, T., and Shinagawa, H. (1999). Role of walker motif A of RuvB protein in promoting branch migration of holliday junctions. Walker motif a mutations affect Atp binding, Atp hydrolyzing, and DNA binding activities of Ruvb. J Biol Chem *274*,

25335-25342.

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.

Hosokawa, M., Shoji, M., Kitamura, K., Tanaka, T., Noce, T., Chuma, S., and Nakatsuji, N. (2007). Tudor-related proteins TDRD1/MTR-1, TDRD6 and TDRD7/TRAP: domain composition, intracellular localization, and function in male germ cells in mice. Dev Biol *301*, 38-52.

Houwing, S., Berezikov, E., and Ketting, R.F. (2008). Zili is required for germ cell differentiation and meiosis in zebrafish. EMBO J *27*, 2702-2711.

Houwing, S., Kamminga, L.M., Berezikov, E., Cronembold, D., Girard, A., van den Elst, H., Filippov, D.V., Blaser, H., Raz, E., Moens, C.B.*, et al.* (2007). A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. Cell *129*, 69-82.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc *4*, 44-57.

Huang da, W., Sherman, B.T., Tan, Q., Kir, J., Liu, D., Bryant, D., Guo, Y., Stephens, R., Baseler, M.W., Lane, H.C.*, et al.* (2007). DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. Nucleic Acids Res *35*, W169-175.

Huang, H., Gao, Q., Peng, X., Choi, S.Y., Sarma, K., Ren, H., Morris, A.J., and Frohman, M.A. (2011a). piRNA-associated germline nuage formation and spermatogenesis require MitoPLD profusogenic mitochondrial-surface lipid signaling. Dev Cell *20*, 376-387.

Huang, H.Y., Houwing, S., Kaaij, L.J., Meppelink, A., Redl, S., Gauci, S., Vos, H., Draper, B.W., Moens, C.B., Burgering, B.M.*, et al.* (2011b). Tdrd1 acts as a molecular scaffold for Piwi proteins and piRNA targets in zebrafish. EMBO J *30*, 3298-3308.

Huang, Y., Ji, L., Huang, Q., Vassylyev, D.G., Chen, X., and Ma, J.B. (2009). Structural insights into mechanisms of the small RNA methyltransferase HEN1. Nature *461*, 823-827.

Hubbard, T.J., Aken, B.L., Ayling, S., Ballester, B., Beal, K., Bragin, E., Brent, S., Chen, Y., Clapham, P., Clarke, L.*, et al.* (2009). Ensembl 2009. Nucleic Acids Res *37*, D690-697.

Hubbard, T.J., Aken, B.L., Beal, K., Ballester, B., Caccamo, M., Chen, Y., Clarke, L., Coates, G., Cunningham, F., Cutts, T.*, et al.* (2007). Ensembl 2007. Nucleic Acids Res *35*, D610-617.

Hutvagner, G., and Simard, M.J. (2008). Argonaute proteins: key players in RNA silencing. Nat Rev Mol Cell Biol *9*, 22-32.

Ibrahim, F., Rymarquis, L.A., Kim, E.J., Becker, J., Balassa, E., Green, P.J., and Cerutti, H. (2010). Uridylation of mature miRNAs and siRNAs by the MUT68 nucleotidyltransferase promotes their degradation in Chlamydomonas. Proc Natl Acad Sci U S A *107*, 3906-3911.

Igreja, C., and Izaurralde, E. (2011). CUP promotes deadenylation and inhibits decapping of mRNA targets. Genes Dev *25*, 1955-1967.

Ikenishi, K. (1998). Germ plasm in Caenorhabditis elegans, Drosophila and Xenopus. Dev Growth Differ *40*, 1-10.

Iwasaki, S., Kobayashi, M., Yoda, M., Sakaguchi, Y., Katsuma, S., Suzuki, T., and Tomari, Y. (2010). Hsc70/Hsp90 Chaperone Machinery Mediates ATP-Dependent RISC Loading of Small RNA Duplexes. Mol Cell *39*, 292-299.

Jin, J., Xie, X., Chen, C., Park, J.G., Stark, C., James, D.A., Olhovsky, M., Linding, R., Mao, Y., and Pawson, T. (2009). Eukaryotic protein domains as functional units of cellular evolution. Sci Signal *2*, ra76.

Joazeiro, C.A., and Weissman, A.M. (2000). RING finger proteins: mediators of ubiquitin ligase activity. Cell *102*, 549-552.

Johnston, M., Geoffroy, M.C., Sobala, A., Hay, R., and Hutvagner, G. (2010). HSP90 Protein Stabilizes Unloaded Argonaute Complexes and Microscopic P-bodies in Human Cells.

Mol Biol Cell.

Kamminga, L.M., Luteijn, M.J., den Broeder, M.J., Redl, S., Kaaij, L.J., Roovers, E.F., Ladurner, P., Berezikov, E., and Ketting, R.F. (2010). Hen1 is required for oocyte development and piRNA stability in zebrafish. EMBO J *29*, 3688-3700.

Karnovski, M.J. (1965). A formaldehyde - glutaraldehyde fixative of high osmolality for use in electron microscopy. JCell Biol *27*, 137A-138A.

Katoh, T., Sakaguchi, Y., Miyauchi, K., Suzuki, T., Kashiwabara, S., Baba, T., and Suzuki, T. (2009). Selective stabilization of mammalian microRNAs by 3' adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2. Genes Dev *23*, 433-438.

Kawaoka, S., Hayashi, N., Suzuki, Y., Abe, H., Sugano, S., Tomari, Y., Shimada, T., and Katsuma, S. (2009). The Bombyx ovary-derived cell line endogenously expresses PIWI/PIWI-interacting RNA complexes. RNA *15*, 1258-1264.

Kedde, M., Strasser, M.J., Boldajipour, B., Oude Vrielink, J.A., Slanchev, K., le Sage, C., Nagel, R., Voorhoeve, P.M., van Duijse, J., Orom, U.A.*, et al.* (2007). RNA-binding protein Dnd1 inhibits microRNA access to target mRNA. Cell *131*, 1273-1286.

Kennerdell, J.R., Yamaguchi, S., and Carthew, R.W. (2002). RNAi is activated during Drosophila oocyte maturation in a manner dependent on aubergine and spindle-E. Genes & development *16*, 1884-1889.

Ketting, R.F. (2011). The many faces of RNAi. Dev Cell *20*, 148-161.

Kimmel, C.B., Ballard, W.W., Kimmel, S.R., Ullmann, B., and Schilling, T.F. (1995). Stages of embryonic development of the zebrafish. Dev Dyn *203*, 253-310.

King, F.J., and Lin, H. (1999). Somatic signaling mediated by fs(1)Yb isKing, F.J., Szakmary, A., Cox, D.N., and Lin, H. (2001). Yb modulates the divisions of both germline and somatic stem cells through piwi- and hh-mediated mechanisms in the Drosophila ovary. Mol Cell *7*, 497-508.

Kirino, Y., Kim, N., de Planell-Saguer, M., Khandros, E., Chiorean, S., Klein, P.S., Rigoutsos, I., Jongens, T.A., and Mourelatos, Z. (2009a). Arginine methylation of Piwi proteins catalysed by dPRMT5 is required for Ago3 and Aub stability. Nat Cell Biol *11*, 652-658.

Kirino, Y., and Mourelatos, Z. (2007a). The mouse homolog of HEN1 is a potential methylase for Piwi-interacting RNAs. RNA *13*, 1397-1401.

Kirino, Y., and Mourelatos, Z. (2007b). Mouse Piwi-interacting RNAs are 2'-O-methylated at their 3' termini. Nat Struct Mol Biol *14*, 347-348.

Kirino, Y., and Mourelatos, Z. (2008). Site-specific crosslinking of human microRNPs to RNA targets. RNA *14*, 2254-2259.

Kirino, Y., Vourekas, A., Kim, N., de Lima Alves, F., Rappsilber, J., Klein, P.S., Jongens, T.A., and Mourelatos, Z. (2010a). Arginine methylation of vasa protein is conserved across phyla. J Biol Chem *285*, 8148-8154.

Kirino, Y., Vourekas, A., Sayed, N., de Lima Alves, F., Thomson, T., Lasko, P., Rappsilber, J., Jongens, T.A., and Mourelatos, Z. (2009b). Arginine methylation of Aubergine mediates Tudor binding and germ plasm localization. RNA *16*, 70-78.

Kirino, Y., Vourekas, A., Sayed, N., de Lima Alves, F., Thomson, T., Lasko, P., Rappsilber, J., Jongens, T.A., and Mourelatos, Z. (2010b). Arginine methylation of Aubergine mediates Tudor binding and germ plasm localization. RNA *16*, 70-78.

Klattenhoff, C., Xi, H., Li, C., Lee, S., Xu, J., Khurana, J.S., Zhang, F., Schultz, N., Koppetsch, B.S., Nowosielska, A.*, et al.* (2009). The Drosophila HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. Cell *138*, 1137-1149.

Klenov, M.S., Lavrov, S.A., Stolyarenko, A.D., Ryazansky, S.S., Aravin, A.A., Tuschl, T., and Gvozdev, V.A. (2007). Repeat-associated siRNAs cause chromatin silencing of retrotransposons in the Drosophila melanogaster germline. Nucleic Acids Res *35*, 5430-5438.

Kloc, M., Bilinski, S., and Etkin, L.D. (2004). The Balbiani body and germ cell determinants: 150 years later. Curr Top Dev Biol *59*, 1-36.

Kojima, K., Kuramochi-Miyagawa, S., Chuma, S., Tanaka, T., Nakatsuji, N., Kimura, T., and Nakano, T. (2009). Associations between PIWI proteins and TDRD1/MTR-1 are critical for integrated subcellular localization in murine male germ cells. Genes Cells *14*, 1155-1165.

Koprunner, M., Thisse, C., Thisse, B., and Raz, E. (2001). A zebrafish nanos-related gene is essential for the development of primordial germ cells. Genes Dev *15*, 2877-2885.

Kosaka, K., Kawakami, K., Sakamoto, H., and Inoue, K. (2007). Spatiotemporal localization of germ plasm RNAs during zebrafish oogenesis. Mech Dev *124*, 279-289.

Kotaja, N., Bhattacharyya, S.N., Jaskiewicz, L., Kimmins, S., Parvinen, M., Filipowicz, W., and Sassone-Corsi, P. (2006a). The chromatoid body of male germ cells: similarity with processing bodies and presence of Dicer and microRNA pathway components. Proc Natl Acad Sci U S A *103*, 2647-2652.

Kotaja, N., Lin, H., Parvinen, M., and Sassone-Corsi, P. (2006b). Interplay of PIWI/Argonaute protein MIWI and kinesin KIF17b in chromatoid bodies of male germ cells. J Cell Sci *119*, 2819-2825.

Krovel, A.V., and Olsen, L.C. (2002). Expression of a vas::EGFP transgene in primordial germ cells of the zebrafish. Mech Dev *116*, 141-150.

Krovel, A.V., and Olsen, L.C. (2004). Sexual dimorphic expression pattern of a splice variant of zebrafish vasa during gonadal development. Dev Biol *271*, 190-197.

Kuramochi-Miyagawa, S., Kimura, T., Ijiri, T.W., Isobe, T., Asada, N., Fujita, Y., Ikawa, M., Iwai, N., Okabe, M., Deng, W*., et al.* (2004). Mili, a mammalian member of piwi family gene, is essential for spermatogenesis. Development *131*, 839-849.

Kuramochi-Miyagawa, S., Kimura, T., Yomogida, K., Kuroiwa, A., Tadokoro, Y., Fujita, Y., Sato, M., Matsuda, Y., and Nakano, T. (2001). Two mouse piwi-related genes: miwi and mili. Mech Dev *108*, 121-133.

Kuramochi-Miyagawa, S., Watanabe, T., Gotoh, K., Takamatsu, K., Chuma, S., Kojima-Kita, K., Shiromoto, Y., Asada, N., Toyoda, A., Fujiyama, A*., et al.* (2010). MVH in piRNA processing and gene silencing of retrotransposons. Genes Dev *24*, 887-892.

Kuramochi-Miyagawa, S., Watanabe, T., Gotoh, K., Totoki, Y., Toyoda, A., Ikawa, M., Asada, N., Kojima, K., Yamaguchi, Y., Ijiri, T.W*., et al.* (2008). DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. Genes Dev *22*, 908-917.

Kurth, H.M., and Mochizuki, K. (2009a). 2'-O-methylation stabilizes Piwi-associated small RNAs and ensures DNA elimination in Tetrahymena. RNA *15*, 675-685.

Kurth, H.M., and Mochizuki, K. (2009b). Non-coding RNA: a bridge between small RNA and DNA. RNA Biol *6*, 138-140.

Lachke, S.A., Alkuraya, F.S., Kneeland, S.C., Ohn, T., Aboukhalil, A., Howell, G.R., Saadi, I., Cavallesco, R., Yue, Y., Tsai, A.C*., et al.* (2011). Mutations in the RNA granule component TDRD7 cause cataract and glaucoma. Science *331*, 1571-1576.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W*., et al.* (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921.

Lau, N.C., Ohsumi, T., Borowsky, M., Kingston, R.E., and Blower, M.D. (2009a). Systematic and single cell analysis of Xenopus Piwi-interacting RNAs and Xiwi. EMBO J *28*, 2945-2958.

Lau, N.C., Robine, N., Martin, R., Chung, W.J., Niki, Y., Berezikov, E., and Lai, E.C. (2009b). Abundant primary piRNAs, endo-siRNAs, and microRNAs in a Drosophila ovary cell line. Genome Res *19*, 1776-1785.

Lau, N.C., Seto, A.G., Kim, J., Kuramochi-Miyagawa, S., Nakano, T., Bartel, D.P., and Kingston, R.E. (2006). Characterization of the piRNA complex from rat testes. Science *313*, 363-367.

Le Hir, H., and Seraphin, B. (2008). EJCs at the heart of translational control. Cell *133*, 213-216.

Lee, J.H., Cook, J.R., Yang, Z.H., Mirochnitchenko, O., Gunderson, S.I., Felix, A.M., Herth, N., Hoffmann, R., and Pestka, S. (2005). PRMT7, a new protein arginine methyltransferase that synthesizes symmetric dimethylarginine. J Biol Chem *280*, 3656-3664.

Leu, D., and Draper, B.W. (2010). The ziwi promoter drives germ cell-specific gene expression in zebrafish. Submitted.

Levin, H.L., and Moran, J.V. (2011). Dynamic interactions between transposable elements and their hosts. Nat Rev Genet *12*, 615-627.

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.

Li, J., Yang, Z., Yu, B., Liu, J., and Chen, X. (2005). Methylation protects miRNAs and siRNAs from a 3'-end uridylation activity in Arabidopsis. Curr Biol *15*, 1501-1507.

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 U S A *104*, 6714-6719.

Lim, A.K., Tao, L., and Kai, T. (2009). piRNAs mediate posttranscriptional retroelement silencing and localization to pi-bodies in the Drosophila germline. J Cell Biol *186*, 333-342.

Lin, H., and Spradling, A.C. (1997). A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary. Development *124*, 2463-2476.

Lin, H., and Yin, H. (2008). A novel epigenetic mechanism in Drosophila somatic cells mediated by Piwi and piRNAs. Cold Spring Harb Symp Quant Biol *73*, 273-281.

Lin, S., Long, W., Chen, J., and Hopkins, N. (1992). Production of germ-line chimeras in zebrafish by cell transplants from genetically pigmented to albino embryos. Proc Natl Acad Sci U S A *89*, 4519-4523.

Liu, H., Wang, J.Y., Huang, Y., Li, Z., Gong, W., Lehmann, R., and Xu, R.M. (2010a). Structural basis for methylarginine-dependent recognition of Aubergine by Tudor. Genes Dev *24*, 1876-1881.

Liu, J., Carmell, M.A., Rivas, F.V., Marsden, C.G., Thomson, J.M., Song, J.J., Hammond, S.M., Joshua-Tor, L., and Hannon, G.J. (2004). Argonaute2 is the catalytic engine of mammalian RNAi. Science *305*, 1437-1441.

Liu, K., Chen, C., Guo, Y., Lam, R., Bian, C., Xu, C., Zhao, D.Y., Jin, J., MacKenzie, F., Pawson, T.*, et al.* (2010b). Structural basis for recognition of arginine methylated Piwi proteins by the extended Tudor domain. Proc Natl Acad Sci U S A *107*, 18398-18403.

Liu, L., Qi, H., Wang, J., and Lin, H. (2011). PAPI, a novel TUDOR-domain protein, complexes with AGO3, ME31B and TRAL in the nuage to silence transposition. Development *138*, 1863-1873.

Liu, Y., Chen, W., Gaudet, J., Cheney, M.D., Roudaia, L., Cierpicki, T., Klet, R.C., Hartman, K., Laue, T.M., Speck, N.A.*, et al.* (2007). 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.

Maack, G., and Segner, H. (2003). Morpholigical development of the gonads in zebrafish. J Fish Biol *62*, 895-906.

Mahowald, A.P. (1971). Polar granules of Drosophila. 3. The continuity of polar granules during the life cycle of Drosophila. J Exp Zool *176*, 329-343.

Malone, C.D., Brennecke, J., Dus, M., Stark, A., McCombie, W.R., Sachidanandam, R., and Hannon, G.J. (2009). Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. Cell *137*, 522-535.

Malone, C.D., and Hannon, G.J. (2009a). Molecular evolution of piRNA and transposon control pathways in Drosophila. Cold Spring Harb Symp Quant Biol *74*, 225-234.

Malone, C.D., and Hannon, G.J. (2009b). Small RNAs as guardians of the genome. Cell *136*, 656-668.

Marlow, F.L., and Mullins, M.C. (2008). Bucky ball functions in Balbiani body assembly and animal-vegetal polarity in the oocyte and follicle cell layer in zebrafish. Dev Biol *321*, 40-50.

Martinez, J., Patkaniowska, A., Urlaub, H., Luhrmann, R., and Tuschl, T. (2002). Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. Cell *110*, 563-574.

Martinez, J., and Tuschl, T. (2004). RISC is a 5' phosphomonoester-producing RNA endonuclease. Genes Dev *18*, 975-980.

Matthews, J.M., Bhati, M., Lehtomaki, E., Mansfield, R.E., Cubeddu, L., and Mackay, J.P. (2009). It takes two to tango: the structure and function of LIM, RING, PHD and MYND domains. Current pharmaceutical design *15*, 3681-3696.

Maurer-Stroh, S., Dickens, N.J., Hughes-Davies, L., Kouzarides, T., Eisenhaber, F., and Ponting, C.P. (2003). The Tudor domain 'Royal Family': Tudor, plant Agenet, Chromo, PWWP and MBT domains. Trends Biochem Sci *28*, 69-74.

McBride, A.E., and Silver, P.A. (2001). State of the arg: protein methylation at arginine comes of age. Cell *106*, 5-8.

McLaren, A. (2003). Primordial germ cells in the mouse. Dev Biol *262*, 1-15.

Megosh, H.B., Cox, D.N., Campbell, C., and Lin, H. (2006). The role of PIWI and the miRNA machinery in Drosophila germline determination. Curr Biol *16*, 1884-1894.

Mevel-Ninio, M., Pelisson, A., Kinder, J., Campos, A.R., and Bucheton, A. (2007). The flamenco locus controls the gypsy and ZAM retroviruses and is required for Drosophila oogenesis. Genetics *175*, 1615-1624.

Minshall, N., Reiter, M.H., Weil, D., and Standart, N. (2007). CPEB interacts with an ovary-specific eIF4E and 4E-T in early Xenopus oocytes. J Biol Chem *282*, 37389-37401.

Mishima, Y., Giraldez, A.J., Takeda, Y., Fujiwara, T., Sakamoto, H., Schier, A.F., and Inoue, K. (2006). Differential regulation of germline mRNAs in soma and germ cells by zebrafish miR-430. Curr Biol *16*, 2135-2142.

Moshkovich, N., and Lei, E.P. (2010). HP1 recruitment in the absence of argonaute proteins in Drosophila. PLoS Genet *6*, e1000880.

Nakamura, A., Amikura, R., Hanyu, K., and Kobayashi, S. (2001). Me31B silences translation of oocyte-localizing RNAs through the formation of cytoplasmic RNP complex during Drosophila oogenesis. Development *128*, 3233-3242.

Nantel, F., and Sassone-Corsi, P. (1996). CREM: a transcriptional master switch during the spermatogenesis differentiation program. Front Biosci *1*, d266-269.

Nishida, K.M., Okada, T.N., Kawamura, T., Mituyama, T., Kawamura, Y., Inagaki, S., Huang, H., Chen, D., Kodama, T., Siomi, H.*, et al.* (2009). Functional involvement of Tudor and dPRMT5 in the piRNA processing pathway in Drosophila germlines. EMBO J *28*, 3820-3831.

Nishida, K.M., Saito, K., Mori, T., Kawamura, Y., Nagami-Okada, T., Inagaki, S., Siomi, H., and Siomi, M.C. (2007). Gene silencing mechanisms mediated by Aubergine piRNA complexes in Drosophila male gonad. RNA *13*, 1911-1922.

Noto, T., Kurth, H.M., Kataoka, K., Aronica, L., DeSouza, L.V., Siu, K.W., Pearlman, R.E., Gorovsky, M.A., and Mochizuki, K. (2010). The Tetrahymena argonaute-binding protein Giw1p directs a mature argonaute-siRNA complex to the nucleus. Cell *140*, 692-703.

Obbard, D.J., Gordon, K.H., Buck, A.H., and Jiggins, F.M. (2009). The evolution of RNAi as a defence against viruses and transposable elements. Philos Trans R Soc Lond B Biol Sci *364*, 99-115.

Ohara, T., Sakaguchi, Y., Suzuki, T., Ueda, H., and Miyauchi, K. (2007a). The 3' termini of mouse Piwi-interacting RNAs are 2'-O-methylated. Nat Struct Mol Biol *14*, 349-350.

Ohara, T., Sakaguchi, Y., Suzuki, T., Ueda, H., Miyauchi, K., and Suzuki, T. (2007b). The 3' termini of mouse Piwi-interacting RNAs are 2'-O-methylated. Nat Struct Mol Biol *14*, 349-350.

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.

Olson, A.J., Brennecke, J., Aravin, A.A., Hannon, G.J., and Sachidanandam, R. (2008). Analysis of large-scale sequencing of small RNAs. Pac Symp Biocomput, 126-136.

Pal-Bhadra, M., Leibovitch, B.A., Gandhi, S.G., Rao, M., Bhadra, U., Birchler, J.A., and Elgin, S.C. (2004). Heterochromatic silencing and HP1 localization in Drosophila are dependent on the RNAi machinery. Science *303*, 669-672.

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.

Pan, B.Y., Dou, S.X., Yang, Y., Xu, Y.N., Bugnard, E., Ding, X.Y., Zhang, L., Wang, P.Y., Li, M., and Xi, X.G. (2010). Mutual inhibition of RecQ molecules in DNA unwinding. J Biol Chem *285*, 15884-15893.

Pan, J., Goodheart, M., Chuma, S., Nakatsuji, N., Page, D.C., and Wang, P.J. (2005). RNF17, a component of the mammalian germ cell nuage, is essential for spermiogenesis. Development *132*, 4029-4039.

Pane, A., Wehr, K., and Schupbach, T. (2007). zucchini and squash encode two putative nucleases required for rasiRNA production in the Drosophila germline. Developmental cell *12*, 851-862.

Park, W., Li, J., Song, R., Messing, J., and Chen, X. (2002). CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in Arabidopsis thaliana. Curr Biol *12*, 1484-1495.

Parvinen, M., and Jokelainen, P.T. (1974). Rapid movements of the chromatoid body in living early spermatids of the rat. Biol Reprod *11*, 85-92.

Patil, V.S., and Kai, T. (2010). Repression of retroelements in Drosophila germline via piRNA pathway by the tudor domain protein tejas. Curr Biol *20*, 724-730.

Pek, J.W., and Kai, T. (2011). A role for vasa in regulating mitotic chromosome condensation in Drosophila. Curr Biol *21*, 39-44.

Pek, J.W., Lim, A.K., and Kai, T. (2009). Drosophila maelstrom ensures proper germline stem cell lineage differentiation by repressing microRNA-7. Dev Cell *17*, 417-424.

Pepling, M.E., Wilhelm, J.E., O'Hara, A.L., Gephardt, G.W., and Spradling, A.C. (2007). Mouse oocytes within germ cell cysts and primordial follicles contain a Balbiani body. Proc Natl Acad Sci U S A *104*, 187-192.

Petrov, D.A., Fiston-Lavier, A.S., Lipatov, M., Lenkov, K., and Gonzalez, J. (2011). Population genomics of transposable elements in Drosophila melanogaster. Mol Biol Evol *28*, 1633-1644.

Pitt, J.N., Schisa, J.A., and Priess, J.R. (2000). P granules in the germ cells of Caenorhabditis elegans adults are associated with clusters of nuclear pores and contain RNA. Dev Biol *219*, 315-333.

Ponting, C.P. (1997). Tudor domains in proteins that interact with RNA. Trends Biochem Sci *22*, 51-52.

Prud'homme, N., Gans, M., Masson, M., Terzian, C., and Bucheton, A. (1995). Flamenco, a gene controlling the gypsy retrovirus of Drosophila melanogaster. Genetics *139*, 697-711.

Qi, H., Watanabe, T., Ku, H.Y., Liu, N., Zhong, M., and Lin, H. (2011). The Yb body, a major site for Piwi-associated RNA biogenesis and a gateway for Piwi expression and transport to the nucleus in somatic cells. J Biol Chem *286*, 3789-3797.

Qiao, D., Zeeman, A.M., Deng, W., Looijenga, L.H., and Lin, H. (2002). Molecular characterization of hiwi, a human member of the piwi gene family whose overexpression is correlated to seminomas. Oncogene *21*, 3988-3999.

Rangan, P., Malone, C.D., Navarro, C., Newbold, S.P., Hayes, P.S., Sachidanandam, R., Hannon, G.J., and Lehmann, R. (2011). piRNA Production Requires Heterochromatin

Formation in Drosophila. Curr Biol *21*, 1373-1379.

Raz, E. (2003). Primordial germ-cell development: the zebrafish perspective. Nat Rev Genet *4*, 690-700.

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.

Reuter, M., Chuma, S., Tanaka, T., Franz, T., Stark, A., and Pillai, R.S. (2009). Loss of the Mili-interacting Tudor domain-containing protein-1 activates transposons and alters the Mili-associated small RNA profile. Nat Struct Mol Biol *16*, 639-646.

Richardson, K.C., Jarret, L., and Finke, E.H. (1960). Embedding in epoxy resins for ultrathin sectioning in electron microscopy. Stain Technology *35*, 313-323.

Robine, N., Lau, N.C., Balla, S., Jin, Z., Okamura, K., Kuramochi-Miyagawa, S., Blower, M.D., and Lai, E.C. (2009). A broadly conserved pathway generates 3'UTR-directed primary piRNAs. Curr Biol *19*, 2066-2076.

Saito, K., Inagaki, S., Mituyama, T., Kawamura, Y., Ono, Y., Sakota, E., Kotani, H., Asai, K., Siomi, H., and Siomi, M.C. (2009). A regulatory circuit for piwi by the large Maf gene traffic jam in Drosophila. Nature *461*, 1296-1299.

Saito, K., Ishizu, H., Komai, M., Kotani, H., Kawamura, Y., Nishida, K.M., Siomi, H., and Siomi, M.C. (2010). Roles for the Yb body components Armitage and Yb in primary piRNA biogenesis in Drosophila. Genes Dev *24*, 2493-2498.

Saito, K., Nishida, K.M., Mori, T., Kawamura, Y., Miyoshi, K., Nagami, T., Siomi, H., and Siomi, M.C. (2006). Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the Drosophila genome. Genes Dev *20*, 2214-2222.

Saito, K., Sakaguchi, Y., 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.

Saito, K., and Siomi, M.C. (2010). Small RNA-mediated quiescence of transposable elements in animals. Dev Cell *19*, 687-697.

Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol *4*, 406-425.

Schafers, C., Teigeler, M., Wenzel, A., Maack, G., Fenske, M., and Segner, H. (2007). Concentration- and time-dependent effects of the synthetic estrogen, 17alpha-ethinylestradiol, on reproductive capabilities of the zebrafish, Danio rerio. J Toxicol Environ Health A *70*, 768-779.

Schwer, B. (2001). A new twist on RNA helicases: DExH/D box proteins as RNPases. Nat Struct Biol *8*, 113-116.

Selenko, P., Sprangers, R., Stier, G., Buhler, D., Fischer, U., and Sattler, M. (2001). SMN tudor domain structure and its interaction with the Sm proteins. Nat Struct Biol *8*, 27-31.

Seto, A.G., Kingston, R.E., and Lau, N.C. (2007). The coming of age for Piwi proteins. Mol Cell *26*, 603-609.

Shoji, M., Tanaka, T., Hosokawa, M., Reuter, M., Stark, A., Kato, Y., Kondoh, G., Okawa, K., Chujo, T., Suzuki, T.*, et al.* (2009). The TDRD9-MIWI2 complex isSiegfried, K.R. (2010). In search of determinants: gene expression during gonadal sex differentiation. J Fish Biol *76*, 1879-1902.

Siegfried, K.R., and Nusslein-Volhard, C. (2008). Germ line control of female sex determination in zebrafish. Dev Biol *324*, 277-287.

Siomi, M.C., Mannen, T., and Siomi, H. (2010). How does the royal family of Tudor rule the PIWI-interacting RNA pathway? Genes Dev *24*, 636-646.

Siomi, M.C., Sato, K., Pezic, D., and Aravin, A.A. (2011). PIWI-interacting small RNAs: the vanguard of genome defence. Nat Rev Mol Cell Biol *12*, 246-258.

Slanchev, K., Stebler, J., de la Cueva-Mendez, G., and Raz, E. (2005). Development without germ cells: the role of the germ line in zebrafish sex differentiation. Proc Natl Acad Sci

U S A *102*, 4074-4079.

Slotkin, R.K., and Martienssen, R. (2007). Transposable elements and the epigenetic regulation of the genome. Nat Rev Genet *8*, 272-285.

Smulders-Srinivasan, T.K., and Lin, H. (2003). Screens for piwi suppressors in Drosophila identify dosage-dependent regulators of germline stem cell division. Genetics *165*, 1971-1991.

Soifer, H.S., Sano, M., Sakurai, K., Chomchan, P., Saetrom, P., Sherman, M.A., Collingwood, M.A., Behlke, M.A., and Rossi, J.J. (2008). A role for the Dicer helicase domain in the processing of thermodynamically unstable hairpin RNAs. Nucleic Acids Res *36*, 6511-6522.

Soper, S.F., van der Heijden, G.W., Hardiman, T.C., Goodheart, M., Martin, S.L., de Boer, P., and Bortvin, A. (2008). Mouse maelstrom, a component of nuage, is essential for spermatogenesis and transposon repression in meiosis. Dev Cell *15*, 285-297.

Sprangers, R., Groves, M.R., Sinning, I., and Sattler, M. (2003). High-resolution X-ray and NMR structures of the SMN Tudor domain: conformational variation in the binding site for symmetrically dimethylated arginine residues. J Mol Biol *327*, 507-520.

Spurr, A.R. (1969). A low-viscosity epoxy resin embedding medium for electron microscopy. J Ultrastruct Res *26*, 31-45.

Strasser, M.J., Mackenzie, N.C., Dumstrei, K., Nakkrasae, L.I., Stebler, J., and Raz, E. (2008). Control over the morphology and segregation of Zebrafish germ cell granules during embryonic development. BMC Dev Biol *8*, 58.

Strome, S., Garvin, C., Paulsen, J., Capowski, E., Martin, P., and Beanan, M. (1994). Specification and development of the germline in Caenorhabditis elegans. Ciba Found Symp *182*, 31-45; discussion 45-57.

Szakmary, A., Cox, D.N., Wang, Z., and Lin, H. (2005). Regulatory relationship among piwi, pumilio, and bag-of-marbles in Drosophila germline stem cell self-renewal and differentiation. Curr Biol *15*, 171-178.

Szakmary, A., Reedy, M., Qi, H., and Lin, H. (2009). The Yb protein defines a novel organelle and regulates male germline stem cell self-renewal in Drosophila melanogaster. The Journal of cell biology *185*, 613-627.

Takahashi, H. (1974). Juvenile Hermaphroditism in the Zebrafish, Brachyodanio rerio. Bulletin of the Faculty of Fisheries Hokkaido University *28*, 57-65.

Talbot, K., Miguel-Aliaga, I., Mohaghegh, P., Ponting, C.P., and Davies, K.E. (1998). Characterization of a gene encoding survival motor neuron (SMN)-related protein, a constituent of the spliceosome complex. Hum Mol Genet *7*, 2149-2156.

Tam, O.H., Aravin, A.A., Stein, P., Girard, A., Murchison, E.P., Cheloufi, S., Hodges, E., Anger, M., Sachidanandam, R., Schultz, R.M.*, et al.* (2008). Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. Nature *453*, 534-538.

Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol *24*, 1596-1599.

Tanaka, T., Hosokawa, M., Vagin, V.V., Reuter, M., Hayashi, E., Mochizuki, A.L., Kitamura, K., Yamanaka, H., Kondoh, G., Okawa, K.*, et al.* (2011). Tudor domain containing 7 (Tdrd7) is essential for dynamic ribonucleoprotein (RNP) remodeling of chromatoid bodies during spermatogenesis. Proc Natl Acad Sci U S A *108*, 10579-10584.

Thomson, T., and Lasko, P. (2004). Drosophila tudor is essential for polar granule assembly and pole cell specification, but not for posterior patterning. Genesis *40*, 164-170.

Thomson, T., and Lasko, P. (2005). Tudor and its domains: germ cell formation from a Tudor perspective. Cell Res *15*, 281-291.

Thomson, T., and Lin, H. (2009). The biogenesis and function of PIWI proteins and piRNAs: progress and prospect. Annu Rev Cell Dev Biol *25*, 355-376.

Thomson, T., Liu, N., Arkov, A., Lehmann, R., and Lasko, P. (2008). Isolation of new polar granule components in Drosophila reveals P body and ER associated proteins.

Mechanisms of development *125*, 865-873.

Tkaczuk, K.L., Obarska, A., and Bujnicki, J.M. (2006). Molecular phylogenetics and comparative modeling of HEN1, a methyltransferase involved in plant microRNA biogenesis. BMC Evol Biol *6*, 6.

Tritschler, F., Huntzinger, E., and Izaurralde, E. (2010). Role of GW182 proteins and PABPC1 in the miRNA pathway: a sense of deja vu. Nat Rev Mol Cell Biol *11*, 379-384.

Uchida, D., Yamashita, M., Kitano, T., and Iguchi, T. (2002). Oocyte apoptosis during the transition from ovary-like tissue to testes during sex differentiation of juvenile zebrafish. J Exp Biol *205*, 711-718.

Uchida, D., Yamashita, M., Kitano, T., and Iguchi, T. (2004). An aromatase inhibitor or high water temperature induce oocyte apoptosis and depletion of P450 aromatase activity in the gonads of genetic female zebrafish during sex-reversal. Comp Biochem Physiol A Mol Integr Physiol *137*, 11-20.

Unhavaithaya, Y., Hao, Y., Beyret, E., Yin, H., Kuramochi-Miyagawa, S., Nakano, T., and Lin, H. (2009). MILI, a PIWI-interacting RNA-binding protein, is required for germ line stem cell self-renewal and appears to positively regulate translation. J Biol Chem *284*, 6507-6519.

Updike, D.L., Hachey, S.J., Kreher, J., and Strome, S. (2011). P granules extend the nuclear pore complex environment in the C. elegans germ line. J Cell Biol *192*, 939-948.

Vagin, V.V., Hannon, G.J., and Aravin, A.A. (2009a). Arginine methylation as a molecular signature of the Piwi small RNA pathway. Cell Cycle *8*, 4003-4004.

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.

Vagin, V.V., Wohlschlegel, J., Qu, J., Jonsson, Z., Huang, X., Chuma, S., Girard, A., Sachidanandam, R., Hannon, G.J., and Aravin, A.A. (2009b). Proteomic analysis of murine Piwi proteins reveals a role for arginine methylation in specifying interaction with Tudor family members. Genes Dev *23*, 1749-1762.

Valverde, R., Edwards, L., and Regan, L. (2008). Structure and function of KH domains. FEBS J *275*, 2712-2726.

van Wolfswinkel, J.C., Claycomb, J.M., Batista, P.J., Mello, C.C., Berezikov, E., and Ketting, R.F. (2009). CDE-1 affects chromosome segregation through uridylation of CSR-1-bound siRNAs. Cell *139*, 135-148.

Vasileva, A., Tiedau, D., Firooznia, A., Muller-Reichert, T., and Jessberger, R. (2009). Tdrd6 is required for spermiogenesis, chromatoid body architecture, and regulation of miRNA expression. Curr Biol *19*, 630-639.

Voronina, E., and Seydoux, G. (2010). The C. elegans homolog of nucleoporin Nup98 is required for the integrity and function of germline P granules. Development *137*, 1441-1450.

Wang, J., Saxe, J.P., Tanaka, T., Chuma, S., and Lin, H. (2009). Mili interacts with tudor domain-containing protein 1 in regulating spermatogenesis. Curr Biol *19*, 640-644.

Wang, Y., Juranek, S., Li, H., Sheng, G., Tuschl, T., and Patel, D.J. (2008). Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. Nature *456*, 921-926.

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.

Weidinger, G., Stebler, J., Slanchev, K., Dumstrei, K., Wise, C., Lovell-Badge, R., Thisse, C., Thisse, B., and Raz, E. (2003). dead end, a novel vertebrate germ plasm component, is required for zebrafish primordial germ cell migration and survival. Curr Biol *13*, 1429-1434.

Weiner, J.H., Cammack, R., Cole, S.T., Condon, C., Honore, N., Lemire, B.D., and Shaw, G. (1986). A mutant of Escherichia coli fumarate reductase decoupled from electron transport. Proc Natl Acad Sci U S A *83*, 2056-2060.

Westerfield (1993). The Zebrafish Book.

Wienholds, E., and Plasterk, R.H. (2004). Target-selected gene inactivation in zebrafish. Methods Cell Biol *77*, 69-90.

Wienholds, E., Schulte-Merker, S., Walderich, B., and Plasterk, R.H. (2002). Target-selected inactivation of the zebrafish rag1 gene. Science *297*, 99-102.

Wienholds, E., van Eeden, F., Kosters, M., Mudde, J., Plasterk, R.H., and Cuppen, E. (2003). Efficient target-selected mutagenesis in zebrafish. Genome Res *13*, 2700-2707.

Wilhelm, J.E., Mansfield, J., Hom-Booher, N., Wang, S., Turck, C.W., Hazelrigg, T., and Vale, R.D. (2000). Isolation of a ribonucleoprotein complex involved in mRNA localization in Drosophila oocytes. J Cell Biol *148*, 427-440.

Williamson, A., and Lehmann, R. (1996). Germ cell development in Drosophila. Annu Rev Cell Dev Biol *12*, 365-391.

Wolf, N., Priess, J., and Hirsh, D. (1983). Segregation of germline granules in early embryos of Caenorhabditis elegans: an electron microscopic analysis. J Embryol Exp Morphol *73*, 297-306.

Wu, S.F., Zhang, H., and Cairns, B.R. (2011). Genes for embryo development are packaged in blocks of multivalent chromatin in zebrafish sperm. Genome Res *21*, 578-589.

Yabuta, Y., Ohta, H., Abe, T., Kurimoto, K., Chuma, S., and Saitou, M. (2011). TDRD5 is required for retrotransposon silencing, chromatoid body assembly, and spermiogenesis in mice. J Cell Biol *192*, 781-795.

Yang, Z., Ebright, Y.W., Yu, B., and Chen, X. (2006). HEN1 recognizes 21-24 nt small RNA duplexes and deposits a methyl group onto the 2' OH of the 3' terminal nucleotide. Nucleic Acids Res *34*, 667-675.

Yin, H., and Lin, H. (2007). An epigenetic activation role of Piwi and a Piwi-associated piRNA in Drosophila melanogaster. Nature *450*, 304-308.

Yokota, S. (2008). Historical survey on chromatoid body research. Acta Histochem Cytochem *41*, 65-82.

Yoon, C., Kawakami, K., and Hopkins, N. (1997). Zebrafish vasa homologue RNA is localized to the cleavage planes of 2- and 4-cell-stage embryos and is expressed in the primordial germ cells. Development *124*, 3157-3165.

Yu, B., Yang, Z., Li, J., Minakhina, S., Yang, M., Padgett, R.W., Steward, R., and Chen, X. (2005). Methylation as a crucial step in plant microRNA biogenesis. Science *307*, 932-935.

Zamparini, A.L., Davis, M.Y., Malone, C.D., Vieira, E., Zavadil, J., Sachidanandam, R., Hannon, G.J., and Lehmann, R. (2011). Vreteno, a gonad-specific protein, is essential for germline development and primary piRNA biogenesis in Drosophila. Development *138*, 4039-4050.

Zhang, Z., Schwartz, S., Wagner, L., and Miller, W. (2000). A greedy algorithm for aligning DNA sequences. J Comput Biol *7*, 203-214.

Zhou, X., Zuo, Z., Zhou, F., Zhao, W., Sakaguchi, Y., Suzuki, T., Cheng, H., and Zhou, R. (2010). Profiling sex-specific piRNAs in zebrafish. Genetics *186*, 1175-1185.

Summary for general public

On average, about 15% of couples have problems to conceive a child. So it is important to study the biological processes that influence fertility. Genetic factors are largely unknown in infertility issues. We use the zebrafish as a study model, because zebrafish and human genes are very similar. We found when certain genes are missing in zebrafish, they become sterile. We would like to find out how these genes affect the fertility. In the future, based on what we learn from this study, we may have better solutions to deal with or prevent fertility problems in human beings. In this summary, you will read about my research on genes that are important in fertility in animals.

Infertility

Infertility is defined as after trying for a year, a couple can't conceive a child. Worldwide, it is estimated that 1 in 7 couples have this problem. There are many possible factors that cause infertility. In general, a third is caused by men, a third is by women and the rest are either unknown or from both genders. The most commonly known reasons in both genders are age, stress, environmental factors (exposure to toxins), genetic factors, or having another disease (such as diabetes or a sexually transmitted diseases), etc. The main focus of this study is on the genetic factors that are important in fertility.

Protect our babies

A new life starts when a sperm cell fertilizes an egg. Having changes in the genes of a fertilized egg can lead to problems in the health or the development of the offspring. Most of the animals, including human and zebrafish, have several defense systems to maintain the consistency of the genes before they pass on the genes to the next generation. If the parents' body can not repair or prevent the changes/damages in the genes of the sperm or the egg cells, these broken sperm/egg cells will be destroyed. This could then lead to infertility.

Transposons/Jumping genes

The unwanted changes in the genes can be caused by virus infection or, by some genes that are present in our DNA, called jumping genes. Jumping genes are officially named transposons. They are a kind of "junk DNA" in the cells. They have the ability to multiply themselves

and jump from one place to another in the DNA. When the jumping genes jump into the middle of an important gene, it can damage the function of this gene.

Piwi, the defender

So how do the animal body defend the important genes in the sperm and egg cells? Here comes the *piwi* genes! The *piwi* genes carry the information to make specific proteins, called Piwi. Piwi proteins are only present in the special animal cells that will later become sperm and egg cells. Piwi proteins are like the army guarding the integrity of the genes in the sperm and egg cells. What we think is that the Piwi proteins form a defense system with some other partner proteins, such as Tudor-containing proteins, to target and destroy jumping genes (transposons). In this way, Piwi proteins prevent unwanted changes in the genes caused by jumping genes in the (future) sperm and egg cells.

Using the zebrafish as a study model

We use the zebrafish to perform our research. Zebrafish and humans share about 70% of identity of their genes. As you may know, all life forms in the world are derived from a common ancestor. The common ancestor for human and fish lived about 400 million years ago. We and other species have inherited many genes from this common ancestor. Therefore we can learn more about the genes in the human body by studying them in other species, for example the zebrafish. Zebrafish has the advantage of easy maintenance. At the same time, their genes are relatively easy to manipulate. For example, making mutants without a specific gene or animals with fluorescent-colored organs. Moreover, the zebrafish has been used as a research model for many human diseases, for instance, genetic disorders, cancers, immune response-related diseases, wound healing problems, etc.

Strategy of the research

The strategy we use to study the *piwi* genes in the zebrafish is by removing these genes. We then see what happens to the zebrafish without *piwi* genes (referred to as mutant) by observing the shape changes of their testis and ovaries. We found the mutants become sterile. They lose all the sperm and egg cells. We also check the activity of some jumping genes. In some cases, the jumping genes are

40 times more active in the mutants compare to the normal fish. We think the reason why these fish don't have normal sperm/egg cells is because the jumping genes damage the sperm/egg cells. The "broken" sperms and eggs are either not able to mature or eventually destroyed by the animal body. This leads to the infertility of the mutant zebrafish.

What we can learn from zebrafish?

The human genome has been sequenced since 2003, however, we still don't know what most of the genes do in our bodies. Because of the high similarity between human and zebrafish genes (70%), the zebrafish becomes a nice tool to study genes of interest. In conclusion: *piwi* genes are present in human and zebrafish. The zebrafish *piwi* genes are important to maintain the fertility of the animal. We speculate that the human *piwi* genes may have similar function as what we observed in zebrafish. To learn from zebrafish how *piwi* genes work to protect the sperm and egg cells may give us more insight how the defense system works in human body. Understanding the reasons behind the infertility caused by loss of the Piwi defense system may help us to deal with some genetically related infertility problems in the future.

Algemene samenvatting

Gemiddeld heeft ongeveer 15% van alle echtparen problemen zwanger te geraken. Het is daarom van belang de biologische processen te onderzoeken welke vruchtbaarheid beïnvloeden. Genetische factoren zijn grotendeels onbekend aangaande vruchtbaarheid. Wij gebruiken de zebravis als een modelsysteem, omdat zebravis en humane genen erg op elkaar lijken. We vonden dat als bepaalde genen afwezig zijn in de zebravis, deze onvruchtbaar wordt. Wij willen graag weten hoe deze genen vruchtbaarheid beïnvloeden. Gebaseerd op wat we hiermee leren kunnen er wellicht in de toekomst betere behandelmethoden beschikbaar zijn voor onvruchtbaarheid in mensen. In deze samenvatting beschrijf ik mijn onderzoek naar genen die belangrijk zijn voor vruchtbaarheid in mensen.

Onvruchtbaarheid

Onvruchtbaarheid is gedefiniëerd als een echtpaar dat na een jaar van proberen nog steeds niet zwanger is geworden. Er wordt geschat dat wereldwijd 1 op de 7 echtparen met dit probleem te maken heeft. Er zijn vele mogelijke oorzaken van onvruchtbaarheid. Over het algemeen ligt in een derde van de gevallen de oorzaak bij de man, bij een derde de oorzaak bij de vrouw, en bij een derde is de oorzaak ofwel onbekend of ligt bij beide. De meest veelvoorkomende gemeenschappelijke oorzaken zijn leeftijd, stress, externe factoren (blootstelling aan gifstoffen), genetische factoren, ziektes (zoals diabetes of seksueel overdraagbare aandoeningen), enz. In deze studie concentreren we voornamelijk op de genetische factoren die belangrijk zijn in vruchtbaarheid.

Bescherm onze babies

Een nieuw leven begint wanneer een spermacel een eicel bevrucht. Veranderingen in de genen van een bevruchte eicel kunnen tot gezondheids- of ontwikkelingsproblemen leiden in de nakomelingen. De meeste dieren, inclusief mensen en zebravissen, hebben verschillende verdedigingssystemen om de structurele integriteit van de genen te handhaven, voordat deze aan de volgende generatie overgedragen worden. Als het lichaam van de ouders de schade aan de genen niet kan repareren of voorkomen in spermacellen of eicellen, worden deze beschadigde spermacellen of eicellen vernietigd. Dit kan

tot onvruchtbaarheid leiden.

Transposons/springende genen
Ongewenste veranderingen aan genen kunnen veroorzaakt worden door virus infecties of door sommige genen aanwezig in ons DNA, genaamd springende genen. Springende genen heten officiëel transposons. Ze zijn een soort “junk DNA” in de cellen. Ze hebben het vermogen zichzelf te kopiëren, en om van de ene naar de andere plaats te springen binnen het DNA. Wanneer springende genen midden in een belangrijk gen springen, kan het de functionaliteit van dit gen aantasten.

Piwi, de verdediging
Hoe beschermt het lichaam dan belangrijke genen in spermacellen en eicellen? Hier komen we bij de *piwi* genen terecht. De *piwi* genen dragen erfelijke informatie om specifieke eiwitten te maken, genaamd Piwi. Piwi eiwitten bevinden zich slechts in speciale dierlijke cellen welke later ofwel spermacel of eicel worden. Piwi eiwitten zijn als een leger dat de integriteit van de genen bewaakt. We denken dat Piwi eiwitten een verdedigings systeem vormt samen met andere eiwitten zoals Tudor-domein bevattende eiwitten, om springende genen (transposons) onschadelijk te maken. Op deze manier kunnnen Piwi eiwitten ongewenste veranderingen aan de genen veroorzaakt door springende genen voorkomen.

De zebravis als modelsysteem
Wij gebruiken de zebravis om ons onderzoek te doen. Zebravissen en mensen delen ongeveer 70% van de indentiteit van hun genen. Zoals bekend stammen alle levensvormen of van dezelfde gemeenschappelijke voorouder. De gemeenschappelijke voorouder van zebravissen en mensen leefde ongeveer 400 miljoen jaar geleden.Wij en andere diersoorten hebben vele genen geërfd van gemeenschappelijke voorouders. We kunnen daarom veel leren over de genen in het menselijk lichaam door deze genen in andere diersoorten te bestuderen, zoals bijvoorbeeld de zebravis. Zebravissen hebben als voordeel dat ze makkelijk te onderhouden zijn. Tegelijkertijd zijn hun genen redelijk makkelijk te manipuleren. Het is bijvoorbeeld makkelijk mutanten zonder een bepaald gen te maken, of dieren met fluorescerende organen. Tenslotte is de zebravis veelvuldig gebruikt in

onderzoek naar humane ziektes, bijvoorbeeld genetische afwijkingen, kanker, immuun gerelateerde afwijkingen, wond genezings problemen, enz.

Onderzoekstrategie
De strategie die we hanteren om *piwi* genen te onderzoeken in zebravissen is om deze genen te verwijderen. We kijken dan wat er gebeurt met de zebravissen zonder *piwi* genen (mutanten) door afwijkingen aan de testes of eierstokken waar te nemen. We bevonden dat mutanten onvruchtbaar worden. Ze veriezen alle sperma- en eicellen. Als we de activiteit van de springende genen bekijken blijken deze in sommige gevallen 40 keer zo actief te zijn als normaal. We denken dat deze vissen geen sperma- en eicellen hebben omdat deze beschadigd zijn door de springende genen. De beschadigde sperma- en eicellen kunnen ofwel niet volgroeien of worden door het lichaam afgebroken. Dit leidt tot onvruchtbaarheid van de mutante zebravissen.

Wat kunnen we leren van de zebravis?
Het humane genoom is sinds 2003 gesequenced, echter weten we nog steeds niet wat het grootste deel van deze genen doen in ons lichaam. Vanwege de grote overeenkomst tussen humane- en zebravis genen, is de zebravis een welkom gereedschap om interessante genen te bestuderen. Tot slot: *piwi* genen bevinden zich in mensen en zebravissen. De zebravis *piwi* genen zijn belangrijk om de vruchtbaarheid van de dieren te waarborgen. We speculeren dat humane *piwi* genen een vergelijkbare rol spelen als geobserveerd in de zebravis. De wetenschap hoe *piwi* genen werken in het beschermen van sperma- en eicellen kan ons meer inzicht verschaffen hoe het verdedigingssysteem in het menselijke lichaam werkt. Het begrijpen van onvruchtbaarheid veroozaakt door verlies van Piwi functionaliteit kan er mogelijk toe leiden dat enkele genetische vormen van onvruchtbaarheid in de toekomst behandeld kunnen worden.

Acknowledgement/Dankwoord

After four and half years of working in the lab as a PhD student, I thought writing up the thesis was the most difficult part. But I just realized that it is actually not as easy as I thought to write the acknowledgement. I hope I didn't miss anyone who I would like to show my appreciation for his or her support and help during the past years in my second home country, The Netherlands.

First of all, I would like to thank Rene. I am grateful for your guidance and support, especially during the last two years and while I was writing the thesis in Toronto. I enjoyed the discussions we had at work. As I often told you, I always feel much better about my work after talking to you.It is very likely you are going to have a big change in your life, hope you and your family will all enjoy the new place.

Thanks to all the former/present Kettingers. It was great to work with you all! I learned a lot from many of you, and enjoyed the time both inside and outside the lab. I will miss the dinners, movies, BBQs and the bowling evenings we had together. Especially, I want to thank Lucas, who is the co-author on all my chapters! I was very lucky to know you and work with you. Without your generous help and support, I would have a much harder time in the last two years. I also enjoyed the time when we gassiped about whatever, and had your favorite beans together ;-) And thank you so much for helping with arranging the defense and related issues when I am too far away to do it myself. Good luck with your PhD, I am convinced that you will do great. Josien, Leonie, Nadine and Marjo thank you for giving me advises and comfort when I feel down in the lab. I will so miss your warm hugs! Amanda, Anisa, Bruno, Elke, Maartje, Selma and Petra, good luck with your PhD and research too! --- No matter where you are now or where you will be, wish you all enjoy what you are doing.

Monika and Cesca, thank you for your friendship and support. It is always so much fun to chat with you in the small microscope room on the 2nd floor. Wish both of you a lot of success and enjoy life after your PhD. Lin, it was so nice to have you. I enjoyed your company very much! I heard you have some good news, wish everything goes

well as you planned. Vivian, Vincent and Ade, it was a pity that I really get to know you so late. Thank you so much for making our most important day so beautiful! Andrei, Ana, thank you for the good times we had together in Utrecht and NYC. Your little girl is coming soon, wish everything goes well and we look forward to see you all again. --- And please don't forget, if you have chance to drop by Toronto, you are all very welcome at our place!

Ira, thank you for all your kind help since I was a master student in Hubrecht, and all the help you did for me to arrange the defense and paper work. Thanks people from de Rooij group, you gave me a lot of help and tips with WB and MS. Marc and Robert, thank you for helping with the FACS. People from the Korswagen group, thank you for all the funny stories when I needed to work in the microscope room-dissecting gonads from my little 3-week-old fishes.

Thanks to many people in the institute, without you the life of PhD students would be much harder: Henk, Pim, Ewart for the screening and sequencing facility; Bert, Mark, Marc, Erma, and Rob for taking care of the fishes; Jeroen for helping with making sections; Jules and Thea, thank you for helping arrange all kinds of things we encounter outside the lab space.

I also would like to thank people I meet in Toronto, especially Adrian and Don for helping me with the R program and checking the grammar of the thesis, respectively. Your generous help makes this book possible.

Hans and Kyra, thank you for treating me as family. I am so lucky to meet you when I was all alone studying in the Netherlands. You are my first bonding with this country. Wish you both lots of happiness and good health for many years. Tony, thank you for your kindness and generosity. Both Mark and I enjoyed very much all the events/dinners we had with you. We both miss to cook for the ING BBQ. And I must say, international students in Utrecht are lucky to have a passionate person like you to lead the ING.

Chris and Liesbeth, words are not enough to show my gratitude and appreciation to your love, support and advices. I am so lucky to have parents-in-law like you. Your home is always the best place to be when

I needed to be re-charged before going back to the stressful PhD student life. Hsin-Ho, Susan and Paul, thank you for making me feel loved. It always gives me more energy to go on.

My dear mom and dad, thank you so much for all the love, support and guidance during these years when I was away from home. Even though physically we are so far apart, your voices on the phone always comfort me. Hopefully we will be able to see each other more often after we move to Toronto. ☺

Mark, you are the most important person behind the scene during my whole PhD study. You went through all the frustration as well as happiness with me. The past few months was not easy for me to write under the condition of new environment, no friends and no job. You are probably the only one who read my whole thesis more than three times. ;-) Honestly speaking, without you I don't think the thesis will ever be done. Thank you so much for your endless support and love. We are lucky to meet a lot of nice people in Toronto, and I think we both enjoy our work and life here. Every step in life is great because of having you besides me. Love you~

Curriculum Vitae

Hsin-Yi Huang was born on the 15th of March, 1981 in Hsinchu, Taiwan. In 2003, she received her B.Sc. diploma with chemistry (major) and life science (minor) at National Tsing-Hua University (NTHU) in Taiwan. She worked as a research/ administrative assistant in the lab of Prof. Hsing-Jang Liu in NTHU for a year and after that she started her study in the program of Biomolecular Sciences in Utrecht University, The Netherlands, and received her M.Sc. in Aug 2006. In Nov 2006, she started as a PhD student in the group of Prof. Dr. Rene Ketting at the Hubrecht Institute. Her research project is focusing on the role of Tudor-domain-containing proteins in the PIWI-piRNA pathway in the zebrafish germline. The results of the project during her PhD study are described in this thesis. From Oct 2011, Hsin-Yi started to work as a postdoctoral fellow in the group of Dr. Aaron Schimmer at the Ontario Cancer Institute in the Princess Margaret Hospital in Toronto, Canada, where she will work on developing new strategies for pancreatic cancer treatment.

List of publications

Huang HY, Houwing S#, Kaaij LJT#, Meppelink A, Redl S, Gauci S, Vos H, Draper BW, Moens CB, Burgering BM, Ladurner P, Krijgsveld J, Berezikov E and Ketting RF. **Tdrd1 associates with piRNA targets and Piwi proteins in zebrafish.** *EMBO J.* 2011 Jul 8;30(16):3298-308.

Houwing S#, Huang HY#, Siegfried KR, Kaaij LJT, Meppelink A, Redl S, Wiebrands K, Waaijers S, Luteijn MJ, Ladurner P, Nüsslein-Volhard C, Berezikov E, Ketting RF. **Tdrd9 affects piRNA accumulation and is required for germ cell maintenance in adult male zebrafish.** *Submitted.*

Huang HY, Kaaij LJT, Krijgsveld J, Dosch R, Berezikov E, Ketting RF. **Tdrd6a is required for normal primordial germ cell formation and accumulation of maternally inherited piRNAs in zebrafish.** *Manuscript in preparation.*

# = contributed equally