

**Germ plasm condensates and the PIWI pathway: maternal
and zygotic elements that
shape the germline**

Elke Francisca Roovers

**Germ plasm condensates and the PIWI pathway: maternal
and zygotic elements that
shape the germline**

**Kiemplasma condensaten en de PIWI-route: maternale en
zygotische elementen die de kiembaan vormgeven**
(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. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties in het openbaar
te verdedigen op woensdag 24 april 2019 des middags te 4.15 uur

door

Elke Francisca Roovers
geboren op 21 juni 1986
te Goes

Cover: Zebrafish germ plasm marked by Buc-RtoK-eGFP in a stage II oocyte

Layout & printing: Off Page, Amsterdam

ISBN: 978-94-6182-948-1

Copyright © by Elke Roovers, all rights reserved. No part of this thesis may be reproduced in any form or
by any means without prior written consent of the author.

Promotoren: Prof. dr. R.F. Ketting
Prof. dr. N. Geijsen

TABLE OF CONTENTS

Outline and scope of this thesis	9
Chapter 1 Introduction	11
Chapter 2 Zebrafish re-shape piRNA populations in the absence of TdrKH	45
Chapter 3 PIWI proteins and piRNAs in mammalian oocytes and early embryos	67
Chapter 4 Tdrd6a regulates the aggregation of Buc into functional subcellular compartments that drive germ cell specification	101
Chapter 5 Bucky ball stabilizes late vegetal pathway components and oocyte polarity in the absence of a Balbiani body	145
Chapter 6 Discussion	171
Addendum Nederlandse samenvatting	189
Dankwoord/Acknowledgements	195
Curriculum vitae	199
List of publications	200

Voor mijn ouders

OUTLINE AND SCOPE OF THIS THESIS

Sexually reproducing organisms produce germ cells in order to pass on their genetic information to the next generation. These cells require specialization during early development and in addition, the protection by the PIWI pathway: a germline-specific small RNA pathway that silences transposable elements through the recognition of their transcripts by PIWI-interacting RNAs (piRNAs) bound to a PIWI protein. In this thesis, several aspects of these processes are studied in order to broaden our knowledge regarding these matters.

Chapter 1 provides a general overview of germline development and the PIWI pathway in different model organisms. Furthermore, I will discuss a process called liquid-liquid phase separation, which plays a role in these processes as well, but also in many other biological phenomena and, when out of control, various disease phenotypes including many neurological illnesses.

In **chapter 2**, we investigate one aspect of piRNA biogenesis in more detail: its 3' end processing. This step is catalyzed by a 3'-to-5' exonuclease, 'Trimmer,' which requires recruitment to the precursor by a Tudor domain-containing protein TdrKH. In some organisms such as mouse, TdrKH is therefore an essential factor for fertility of the animal. Using CRISPR-Cas9, we obtained *tdrkh* mutant zebrafish and observe that, even though piRNAs are longer, likely reflecting 3' end processing defects, we still see that part of the population remains fertile. We will speculate about the possible reasons for this, including the shift to a different arm of the pathway.

In mammals, the PIWI pathway has been studied mostly in the male germline. The reason for this is the absence of an essential PIWI pathway in the female germline of the mouse. It has therefore been assumed that the PIWI pathway is dispensable in the female mammalian germline. The recent discovery of an overly active Dicer variant in the murine ovary and the evolutionary loss of a PIWI paralog, PIWIL3, prompted us to explore a potential role for PIWIL3 and piRNAs in the female germline of other mammalian species. Our results, described in **chapter 3**, suggest that the PIWI pathway does play a role in female germline of other mammals, including humans.

In **chapter 4 and 5** we investigate properties of germ cell-specific subcellular aggregates formed through liquid-liquid phase separation. The protein Bucky ball (Buc) contains a prion-like domain and is an essential component of the Balbiani body, an example of such an aggregate in the oocyte. This structure is highly dynamic throughout oogenesis and embryogenesis. One factor that modulates Buc-containing structures is Tdrd6a, which increases the mobility of these structures, which is ultimately required for its normal function in germ cell specification. This is described in **chapter 4**.

In **chapter 5** we follow up on a mutated version of Buc, Buc-RtoK, which in the absence of a wild-type copy of Buc, fails to form a Balbiani body. We find that, without the presence of a Balbiani body, Buc-RtoK can still phase-separate during the late vegetal pathway. Furthermore, we identify novel candidates that co-localize and affect Buc-structures in BmN4 cells.

In **chapter 6**, the findings described in this thesis are discussed more generally.

Chapter

Introduction

1

In all sexually reproducing organisms, germ cells or gametes play a central role in the intergenerational inheritance of genomic information. These cells are separated from the soma during early development in order to remain capable to give rise to an entire organism and all its cell types. For this reason germ cells have also been coined ‘the ultimate stem cell.’ They face many challenges during development. First of all, germ cells have to be specified and set aside from somatic lineages during early development. This means they have to be able to ignore signals from surrounding tissues and prevent differentiation into one somatic program or another, while they migrate through developing somatic areas that provide many tissue-specific developmental cues. Still, germ cells manage to keep their own identity and potency. Secondly, since two fused germ cells alone will provide the genetic information of an entire organism – including its germ cells and therefore ultimately the entire descending line – it is imperative that these cells minimize genomic damage. Nature has evolved ways to efficiently regulate these processes.

I will discuss how germ cell specification is coordinated in different organisms and how a germ cell-specific small RNA pathway – the PIWI-piRNA pathway – protects the germline genome from detrimental effects by transposable elements. Lastly, I will discuss how the phenomenon of compartmentalization through liquid-liquid phase separation in germ cell biology can help us to understand disease phenotypes that occur through aberrant protein aggregation in conditions like Alzheimer’s disease and Parkinson’s disease.

GERMLINE DEVELOPMENT

Germline specification – two modes, one goal

Germ cells of animals segregate from the soma during early development, which requires certain cells to become specified as ‘primordial germ cells’ (PGCs), the precursors of sperm and oocytes (Extavour, 2003; Lawson et al., 1999; Raz, 2003; Williamson and Lehmann, 1996; Ying et al., 2000; Yoon et al., 1997). Upon their specification, PGCs migrate towards the somatic gonads, where they will proliferate and further differentiate. Though this concept is conserved, two different mechanisms exist to achieve this. PGCs are either induced by signaling from the surrounding tissue, or are formed through maternal inheritance of PGC-specifying material, called germ plasm. The inductive mode is assumed to be the ancestral mechanism and is most common, while inheritance of germ plasm instead is believed to have evolved independently several times (Extavour, 2003; Whittle and Extavour, 2017). It is unknown why these two modes exist. It has been hypothesized that in vertebrates, germ plasm allows for more plasticity and accelerates evolution, since animal clades using the inherited mode would display increased speciation rates, compared to clades that have induced germlines (Crother et al., 2016; Evans et al., 2014). However, other studies showed this needs to be defined more precisely and ascribe enhanced speciation to *timing* of PGC development: before or after gastrulation (Johnson and Alberio, 2015; Whittle and Extavour, 2016). When PGC induction takes place only after gastrulation, evolution of somatic lineages are inhibited since they need to remain standby in case they receive inductive impulses, whereas early PGC specification allows evolution to occur to a greater extent, since PGCs already segregated (Johnson and Alberio, 2015).

1

In mammals, germ cells are actively induced by extra-embryonic signaling to the epiblast. In mouse, BMP4 and BMP8b expression in the extraembryonic ectoderm can induce proximal cells in the embryonic epiblast to become PGCs around E6.25 (Lawson et al., 1999; Ying et al., 2000, 2001). Upon BMP signaling Blimp1, a transcriptional repressor, is produced and interacts with the methyltransferase Prmt5. Together, Blimp1/Prmt5 translocate to the nucleus where Prmt5 methylates and epigenetically silences Blimp1 targets (Ancelin et al., 2006; Ohinata et al., 2005). Around E7.25, approximately 40 PGCs have been formed just below the developing precursor tissue of the umbilical cord, the allantois, and migrate towards the genital ridges until ~E10.5 (Molyneaux et al., 2001). Blimp1 and downstream signaling restrict these cells to the PGC cell fate. It is required for migration and proliferation of PGCs, while suppressing Hox genes and other somatic factors (Kurimoto et al., 2008; Ohinata et al., 2005; Saitou et al., 2002).

Other vertebrates and most insects specify germ cells through the maternal inheritance of cytoplasmic membrane-less structures, which contain organelles and germ-cell specifying transcripts and proteins. These structures are capable of inducing germ cell fate in zygotic cells that receive this material. *C. elegans* embryos receive so-called 'P-granules' from the mother, which specify the P1 blastomere as early as the 2-cell stage (Updike and Strome, 2010). This is

followed by 3 asymmetric divisions until the PGC 'P4' is specified (Figure 1) (Strome and Wood, 1982). Around the 100-cell stage, P4 forms Z2 and Z3 through a symmetric division, which do not divide further until the L4 stage after the somatic gonad has formed, upon which both Z2 and Z3 will proliferate and populate one gonad arm each (Figure 1). *Drosophila* embryos inherit structures similar to P-granules, called 'pole plasm', which localizes posteriorly in the egg (Figure 1). Upon fertilization, a syncytium forms through nuclear divisions, and as soon as nuclei reach the pole plasm, they bud off the posterior into 'pole buds.' Two divisions later, these buds cellularize to form the pole cells, the equivalent of PGCs in *Drosophila*. Just as in *C. elegans*, the cell cycle of PGCs occurs asynchronously and slower than the somatic cells (Su et al., 1998). During migration, pole cells are arrested in G2, however, upon reaching the gonad, mitotic cues can initiate proliferation again (Su et al., 1998). In zebrafish, the inherited 'germ plasm' originates from a large membrane-less structure in the oocyte called the 'Balbiani body' (Figure 1). Upon fertilization, the germ plasm localizes to the cleavage planes during the first divisions after fertilization and is taken up by 4 distinct cells by the 32-cell stage (Yoon et al., 1997). These 'presumptive PGCs' divide asymmetrically up to the 1000-cell stage. During these divisions, the germ plasm, marked by *vasa*, localizes asymmetrically to one centrosome and ends up in only one of the daughter cells (Braat et al., 1999; Knaut et al., 2000). After the 1000-cell stage, they segregate from the soma and are now *bona fide* PGCs (Figure 1) (Raz, 2003; Yoon et al., 1997). From this point onwards, the PGCs undergo symmetric divisions while migrating towards the gonadal ridge, where they effectively all can be found 1 day post fertilization (dpf) (Figure 1). In zebrafish, the amount of correctly specified and localized PGCs at 1dpf can vary widely per individual and is known to be a factor influencing sexual development, since severely depleted PGCs result in male development rather than female development (Tzung et al., 2015). Upon arrival, PGCs do not proliferate until around 9-10dpf, which is when the somatic gonad has been established as well (Braat et al., 1999; Leerberg et al., 2017; Tzung et al., 2015). These cells are not quiescent during this time window, but start expression of zygotic members of the piRNA pathway, a small RNA pathway that protects the germline from genetic parasites (see later) (Houwing et al., 2008; Huang et al., 2011). It has therefore been hypothesized that these factors complement maternally inherited factors in order to set the zygotic defense system in place during this time.

The two different modes of germ cell specification require different mechanisms and timing. For instance, animals that use the inductive mode require signaling to permissive cells during a certain time window during embryogenesis. Animals that require inherited germ plasm, incorporate this *before* the zygotic program has started. This means that in this case, zygotic PGC specification has to be organized much earlier: prior to fertilization during oogenesis. This occurs in parallel with other events that are vital for oocyte maturation and requires the accumulation of germ plasm factors in subcellular heritable assemblies. These structures contain proteins, organelles and translationally repressed mRNAs, and require specific and timely accumulation and localization in oocytes. In the next paragraphs, germ plasm formation and Balbiani body function will be discussed further.

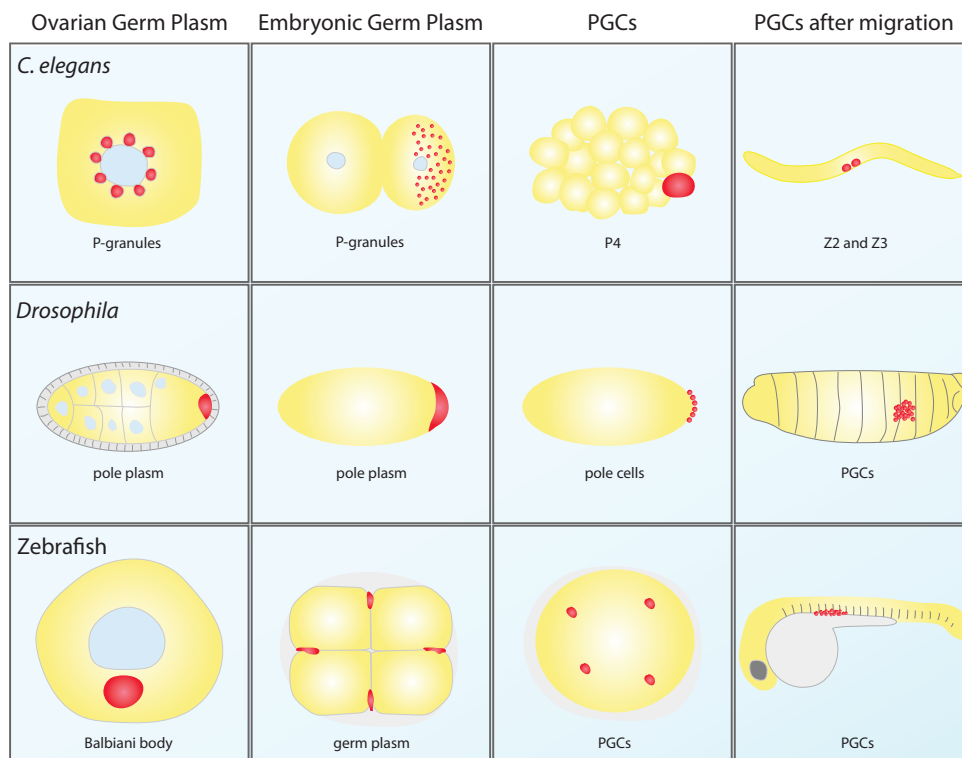


Figure 1. Schematic overview of PGC formation in animals with an inherited germline. Germ plasm inheritance and localization followed by PGC formation in *C. elegans*, *Drosophila* and zebrafish. Germ plasm and the location of the PGCs are indicated in red.

1

1 Zebrafish germ plasm formation in the oocyte – a two-step process

The onset of germ plasm formation in zebrafish is entangled with two other events, namely meiotic bouquet formation and oocyte polarization (Elkouby et al., 2016). During the zygotene stage of meiosis I, telomeres cluster at one side of the nucleus, which is called the chromosomal bouquet (Elkouby et al., 2016). Bouquet formation is a highly conserved process and involves interactions of the cytoskeleton, through the nuclear envelope, with telomeres. It has been hypothesized that this is required for successful meiotic recombination and is found from yeast to plants to humans (Scherthan, 2001). Zebrafish oocytes also become polarized during this time. The first axis to be established is the Animal-Vegetal (AV) axis. The earliest structure to become apparent during symmetry breaking is the centrosome, which is found at the end of cytoplasmic bridges between oogonia, marking the orientation of the last mitotic division from its sister oogonium (Elkouby et al., 2016). This marks the vegetal pole of the oocyte and ultimately the embryo. Microtubules from the centrosome recruit the chromosomal bouquet, which is therefore oriented vegetally as well. This sets the stage for germ plasm formation, which involves recruitment of factors required for PGC specification, like mRNA and proteins, and stabilization of these factors into heritable, subcellular granules. Localization occurs during two stages, first described in *Xenopus*, but zebrafish show many similarities (Kloc and Etkin, 1995; Kosaka et al., 2007; Yisraeli et al., 1990).

First, a nuclear cleft forms vegetally and a protein called Bucky Ball (Buc) accumulates there and surrounds the centrosome (Elkouby et al., 2016; Heim et al., 2014). The centrosome is inactivated and disappears soon after. However, Buc vegetal localization is stabilized and forms the Balbiani body (Figure 2). At this moment, various transcripts, including *buc* mRNA, also localize there. Other transcripts that localize to this area, such as *vasa*, *nanos3* and *dazl*, are all factors that are required for fertility and localize to the PGCs in the zygote (Knaut et al., 2000; Koprunner et al., 2001; Kosaka et al., 2007; Takeda et al., 2009). This pathway is called the ‘early vegetal pathway’ or METRO pathway (**M**essenger **T**ransport **O**rganizer) (Bontems et al., 2009; Draper et al., 2007; King et al., 1999; Kirino et al., 2010a; Kloc and Etkin, 1995; Knaut et al., 2000; Kosaka et al., 2007). At the end of oogenesis stage I, this structure will move towards the vegetal pole and dissociates into germinal granules (Figure 2). It is largely unknown how this is coordinated, but the cytoskeleton plays a role in this reorganization. One protein that is required for dissociation is the microtubule-actin crosslinking factor 1 (Macf1 or Magellan, Mgn), which can interact with actin filaments, microtubules and intermediate filaments (Escobar-Aguirre et al., 2017). This protein is required for dissociation of the Balbiani body towards the end of oogenesis stage I and typical peripheral localization of microtubules in the oocyte (Escobar-Aguirre et al., 2017; Gupta et al., 2010).

The second localization step, or ‘late vegetal pathway,’ is initiated at the end of stage I and continues during stage II and III. Germinal granules first move towards the vegetal pole and are then distributed more cortically, from the vegetal pole towards the animal pole (Figure 2). Certain factors that are non-localized at first, are recruited to these germinal granules during the late pathway, including *mago nashi* and *bruno-like (brul)* (Kosaka et al., 2007). Furthermore,

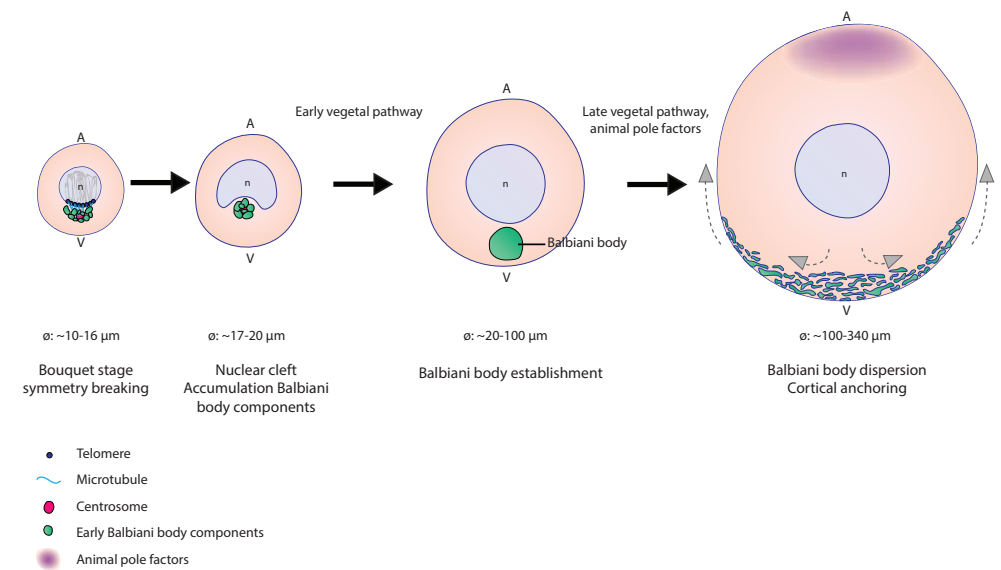


Figure 2. Schematic overview of germ plasm formation during zebrafish oogenesis. Before the onset of the mRNA localization of the early vegetal pathway, the bouquet forms and indicates the first axis – the AV-axis – of the oocyte. Next, Balbiani body components accumulate at the same side in a nuclear cleft and are stabilized into the Balbiani body. During the late vegetal pathway, the Balbiani body moves towards the vegetal pole and disperses into smaller fragments. These fragments dock to the oocyte cortex and spread along the cortex, from the vegetal pole towards the animal pole.

the PIWI protein Ziwi, which initially localizes throughout the cytoplasm, is incorporated in the germ plasm and displays germinal granule localization at these stages as well (Houwing et al., 2007). Conversely, *nanos3* and *buc* de-localize from germinal granules and become diffusely cytoplasmic or move towards the animal pole, respectively (Bontems et al., 2009; Draper et al., 2007).

After fertilization, the germinal granules accumulate again into larger germ plasm structures during the first divisions, before they are taken up by the presumptive PGCs (Riemer et al., 2015). By now, many transcripts that localize to the germ plasm are known, and it has been demonstrated that each factor plays an essential role in PGC specification, maintenance or migration, since loss or knock-down of single components readily results in PGC defects (Koprunner et al., 2001; Slaidina and Lehmann, 2017; Tzung et al., 2015; Weidinger et al., 2003). Furthermore, in *Drosophila*, PGC transcripts initially accumulate in germ granules in a homotypic fashion before they form heterotypic granules (Little et al., 2015; Trcek et al., 2015). Even though certain ratios between transcript clusters exist, they remain spatially separated within the granule (Niepielko et al., 2018; Trcek et al., 2015). Also in zebrafish germ plasm transcripts accumulate in certain fixed ratios, while transcript smFISH signals do not overlap, and signals from individual transcripts display enrichment in specific areas of the germ plasm (see chapter 4 of this thesis). Finally, the presence of a certain transcript in a germ granule

1 stimulates accumulation of transcripts of the same kind, which led to the idea that transcript clustering could also be *intrinsically* driven through RNA-RNA-interactions (Little et al., 2015; Niepielko et al., 2018; Trcek et al., 2015). Together, it has been established that germ plasm transcripts are not merely localized, but that timing and concentration of individual transcripts is tightly orchestrated as well, potentially even on a more intrinsically structural level.

The Balbiani body: zebrafish versus the rest

As mentioned earlier, the Balbiani body is a large, membrane-less compartment present in early stage oocytes (Cox, 2003; Heasman et al., 1984; Kloc et al., 2004; Pepling et al., 2007; De Smedt et al., 2000). It contains high concentrations of mitochondria and other organelles, and has a high electron density relative to the cytosol. In zebrafish, the Balbiani body is a key assembly for the establishment of germ plasm. Interestingly, regardless of the mode of germline formation, the Balbiani body is conserved from insects to mammals, and thus also present in animals that have an induced germline (Cox, 2003; Heasman et al., 1984; Kloc et al., 2004; Pepling et al., 2007; De Smedt et al., 2000). The Balbiani body is either formed by accumulating material from multiple sister oocytes into one, like in *Drosophila* and mouse, or present in all oocytes, like in zebrafish and *Xenopus*. What they all have in common is that in order for an oocyte to survive, a Balbiani body is required (see below).

The Balbiani body of zebrafish is not only involved in the accumulation of germ plasm factors, but also in oocyte patterning and viability of the embryo. When Buc is absent, the Balbiani body does not form, even though the upstream symmetry breaking events, the vegetal centrosome and bouquet, still occur (Bontems et al., 2009; Elkouby et al., 2016; Marlow and Mullins, 2008). This ultimately causes loss of polarization, so that multiple sperm entry points or ‘micropyles’ will form, instead of a single one at the animal pole, resulting in polyspermy (Marlow and Mullins, 2008). It therefore seems that Buc and the Balbiani body are required to stabilize and maintain the AV-axis. Buc protein localization occurs before its transcript localizes, which could point in the direction of RNA-binding proteins (RBPs) that are recruited by Buc and bind *buc* (Heim et al., 2014). Indeed, some RBPs are known to localize to the Balbiani body, like Rbpms2a/b, that are also required for normal Balbiani body formation (Heim et al., 2014; Kaufman et al., 2018). Furthermore, some RBPs including Rbpms2a/b and *Dazl* have also been demonstrated to associate with early pathway transcripts (Heim et al., 2014). Interestingly, Buc phenotypes cannot be rescued completely by a Buc transgene with absent/incomplete introns and UTRs, suggesting there might be additional roles for non-coding sequences and/or splicing (Heim et al., 2014). The mechanism of Balbiani body and germ plasm formation in *Xenopus* is very similar to zebrafish, including an upstream key factor, Buc-homolog X Velo, and transcript localization, including *Xcat-2* (*nanos* homolog) and *Xdazl*, during two vegetal pathways (Forristall et al., 1995; Houston et al., 1998; Kloc and Etkin, 1995; Zhou and King, 1996). Also similar accessory proteins are involved, including RNA-binding protein Hermes (Rbpms2), and the presence of sequence motifs in the 3’UTRs of localizing mRNAs, required for interaction and recruitment (Aguero et al., 2016; King et al., 1999; Kloc and Etkin, 1995).

1 In *Drosophila* and mouse, the Balbiani body is not present in all oocytes. Just as zebrafish, murine and *Drosophila* oocytes initially coexist in interconnected ‘cysts.’ This way, oocytes can communicate and traffic components such as RNA and organelles between each other. In both *Drosophila* and mouse, the Balbiani body ‘selects’ certain oocytes for survival. *Drosophila* egg chambers typically exist of 16 connected cells. The most posterior cell that forms the eventual oocyte receives organelles and transcripts from the 15 other ‘nurse’ cells through ring canals along microtubules where they accumulate in the posterior oocyte in the Balbiani body (Cox, 2003). In *Drosophila*, the germ plasm also requires a factor functionally analogous to Buc and Xvelo in order to be nucleated and stabilized, called *oskar* (Ephrussi and Lehmann, 1992; Ephrussi et al., 1991). *Oskar* mRNA only partially overlaps with the *Drosophila* Balbiani body, prior to its dissociation after oogenesis stage 7 when microtubules reorganize, whereas *oskar* remains at the oocyte posterior, where germ plasm is formed. This is in contrast to *Xenopus* and zebrafish in which mitochondria remain associated with germ plasm transcripts during the late pathway as well (Cox, 2003; Kloc and Etkin, 1995).

During mouse oogenesis, at E14.5, each germ cell in the cyst contains a single centrosome. Next, ring canal transport makes it possible for a subset of germ cells to accumulate multiple (up to 4 at birth) (Lei and Spradling, 2016). The same germ cells accumulate other organelles like Golgi and mitochondria this way, which together form the Balbiani body, and will eventually survive at the costs of other cyst-members (Lei and Spradling, 2016). Thus far, no nucleating factor has been identified for the murine Balbiani body and besides organelles, it is unclear whether other components, such as mRNA, accumulate as well.

Even though the Balbiani body is conserved, it can have different functions between species. However, some functions are commonly found in all species studied. The Balbiani body is rich in mitochondria and therefore potentially functions in mitochondrial selection. It has been suggested that only mitochondria bearing few mtDNA mutations are incorporated in order to provide the progeny with healthy mitochondria (Marlow, 2017; Zhou et al., 2010). Furthermore, it has been shown that unfavorable mitochondrial mutations are selectively eliminated during oogenesis, also referred to as the ‘mitochondrial bottleneck,’ in order to maintain mitochondrial fitness (Bergstrom and Pritchard, 1998; Fan et al., 2008; Hill et al., 2014; Marlow, 2017). Another feature different species have in common is that all oocytes that can become fertilized and are compatible with zygotic development are marked by a Balbiani body; in zebrafish and *Xenopus*, virtually all oocytes have a Balbiani body and can get fertilized, whereas in other species, substantial oocyte selection takes place during oogenesis and only successful oocytes have one. Species such as zebrafish and *Xenopus* could have modified and specialized this structure further and combined it with germ plasm function, so that in addition to organelle inheritance, factors required for PGC formation are incorporated as well.

PROTECTING THE GERMLINE – SMALL RNAs GUARD THE GENOME

Transposable elements – friend or foe?

Germ cells contain the genetic information for all generations that follow. Without these cells, spread of information of the individual is terminated and consequently, there is strong evolutionary pressure to preserve and actively protect these cells. One major threat to germ cells are ‘transposable elements’ (TEs) or transposons. The phenomenon of transposing or ‘jumping’ DNA was first discovered by Barbara McClintock, who found that certain ‘controlling units’ were capable of influencing gene expression in maize (McClintock, 1951). She discovered that these elements could move from one region in the genome to another, contrasting views at the time that genes can change only little at once. Later, it was appreciated that these elements are genomic parasites that insert their genome into that of the host, in order to thrive. TEs show similarities to viruses, including dependency on the host, encoding replication or DNA-integration machineries or the ability to use the machinery of others, and some even display production of capsid proteins. TEs are roughly divided into two types, class I and II transposons, based on the nature of their replication intermediate. TEs of class I are retrotransposons that replicate using an RNA intermediate, and can insert itself again after reverse transcription into DNA. They are subdivided in those that contain long terminal repeats (LTR) and elements and those that do not (non-LTRs). LTRs show similarities to retroviruses and encode typical viral proteins like *gag* and *pol* and replicate in the cytoplasm, whereas non-LTRs replicate in the nucleus through reverse transcription (Boeke and Voytas, 2002; Han, 2010). Class II transposons are DNA transposons that move around without gaining copies, unlike retrotransposons, however, the genome usually still bears marks (‘transposon footprints’), left by imperfect DNA repair (Feschotte and Pritham, 2007; Kazazian, 2004). The majority of elements lack crucial components to transpose and are rather dormant remnants of ancient manifestations. However, there are still active elements present in most genomes. In mammals for instance, the non-LTR element L1 is still functionally independent and makes up about 14% of the human genome (Brouha et al., 2003; Sassaman et al., 1997). Still, TEs are usually suppressed through epigenetic silencing, even though transposition of certain, mostly evolutionarily young, retrotransposons in the soma but also the germline still occurs (Baillie et al., 2011; Brouha et al., 2003; Iskow et al., 2010; Sassaman et al., 1997; Smit, 1999).

The character of uncontrolled transposition causes a potential threat to the host genome. Transposons can insert themselves in exons thereby causing alterations or transcript termination in their coding region. Furthermore, homologies between non-allelic loci can cause aberrant recombination and translocations, which can cause severe genomic rearrangements. Transposons can also alter gene expression by jumping into promoters, introns or more distant regulatory regions, or function as promoters themselves, either sense or antisense, causing abnormal expression or the creation of pseudogenes. Additionally, by nature, transposition induces DNA double-stranded breaks (DSBs) leading to more general genomic instability. An example of damaging effects is the TE insertion that causes hemophilia (Kazazian et al., 1988).

However, the effects do not have to be detrimental per definition and can accelerate favorable mutations as well, or have a more structural function in the genome as discussed below.

One famous example is the color adaptation of British peppered moths. Due to industrialization and coal pollution, the pale camouflage of peppered moths made them highly visible to predators. They rapidly adapted to an advantageous black appearance, due to natural selection of a 22kb insertion of a DNA element in the first intron of the *cortex* gene, that regulates development of the wing (Hof et al., 2016). Furthermore, since TEs ‘attract’ silencing mechanisms, it is thought that they play a role in genomic areas that are rich in constitutive heterochromatin, such as centromeres and telomeres. These regions, which serve important structural functions in the genome, are demarcated by TE-rich regions, which are thought to contribute to maintenance of silencing of these areas. This way, *Drosophila*, lacking telomerase, even entirely depends on coordinated insertion of retrotransposons that make up their telomeres (Levis et al., 1993; Pardue and DeBaryshe, 2003). Lastly, some insertions have major biological consequences, while it is unclear whether the insertion itself is causal, or if it is rather an accelerated adaptation to preceding mutations that were detrimental. This is the case for instance in the female germline of Muridae, which comprises rat and mouse species. These animals have lost a paralog of the PIWI clade, PIWIL3, which is conserved in other mammals. PIWI proteins are Argonaute proteins that play a central role in a germline-specific small RNA pathway, the PIWI pathway (see below). In contrast to other animals studied, in Muridae, only the male germline relies on a functional PIWI pathway, whereas in other species, this pathway is of importance in both germlines, though particularly in the ovary (Aravin et al., 2006; Carmell et al., 2007; Cheng et al., 2014; Houwing et al., 2007; Kamminga et al., 2010; Tam et al., 2008; Watanabe et al., 2008). Instead, Muridae express an overly active Dicer variant in the ovary (Dicer^o), driven by a TE insertion in intron 6, that produces large amounts of siRNAs, whereas male gonads display more typical dependency on piRNAs (Flemr et al., 2013).

In conclusion, TEs are rather unpredictable in their mobility throughout the genome, acting like an accelerating mutational force that can either be harmful, since adaptations cannot keep up, or can have unexpected positive effects that benefit evolution and diversity. Hence, TEs are also called ‘drivers of genome evolution’ (Kazazian, 2004).

Germline protection by the PIWI pathway

With the current techniques available, we now know that roughly 50% of the mammalian genome exists of sequences derived from TEs (Kazazian, 2004). If such a vast amount of DNA would continuously transpose, this would have extensive detrimental consequences for the genome. When harmful transposition events occur in the soma, the effects are carried by this cell and its potential descendants only. If this occurs in the germline, this could have consequences that reach over generations, since germ cells will make up half of the genome of the progeny. In order to protect germ cells from these invasive elements, a small RNA pathway is active in the animal germline. At the core of this mechanism are specialized Argonaute proteins of the PIWI family, of which most organisms encode two or more paralogs, and their PIWI-interacting RNAs

1 (piRNAs) (Aravin et al., 2006; Brennecke et al., 2007; Gainetdinov et al., 2018; Houwing, 2009). These piRNAs direct PIWI-mediated cleavage of transposon transcripts, or can induce their transcriptional silencing by targeting genomic regions that encode them, however, to date, only for mouse prenatal testis and *Drosophila* ovary functionality of a nuclear piRNA pathway has been demonstrated experimentally (Aravin et al., 2008, 2009; Brennecke et al., 2007; Le Thomas et al., 2013; Vagin et al., 2006). Unlike other small RNA pathways, like miRNAs or siRNAs, piRNAs are derived from single-stranded precursors rather than double-stranded, the latter making it a suitable Dicer substrate. Instead, piRNA are produced independently of Dicer and use specialized alternative biogenesis mechanisms (Houwing et al., 2007; Vagin et al., 2006). Generally, two pathways are being distinguished, the primary and secondary pathway, which have been studied in most detail in *Drosophila*. Below, these pathways will be discussed.

In the primary pathway large regions of the genome are being transcribed as long piRNA-precursors that contain numerous TE fragments. These genomic loci are called 'piRNA clusters' and it has been hypothesized that transposons eventually will be 'trapped' in one of those regions, inherent to their mobile nature (Figure 3) (Brennecke et al., 2007; Yamanaka et al., 2014). The more recent its invasion, the more intact a transposon sequence resides in a cluster and thus the more precise it can be targeted. This way, active elements *specifically* are incorporated in piRNA clusters and provide a genetic memory for this defense mechanism. Therefore, the piRNA pathway has strong analogies to an adaptive immune system: it has a target-specific response to an active threat. These regions can be predicted bioinformatically, demonstrating that this mechanism is likely to be general, even though relative dependency on the primary pathway can differ per species (Gainetdinov et al., 2018; Rosenkranz and Zischler, 2012).

Precursor transcripts are exported out of the nucleus to a conserved peri-nuclear electron-dense structure called 'nuage' (French for 'cloud'), typically found in animal germ cells (Eddy, 1976; Mahowald, 1971; Zhang et al., 2012). Cytoplasmic piRNA pathway components, including PIWI proteins typically localize there and nuage requires an active piRNA pathway in order to be stable, making it the principal location for further processing and piRNA pathway action (Aravin et al., 2009; Chuma et al., 2006; Findley, 2003; Houwing et al., 2007, 2008; Huang et al., 2011; Senti and Brennecke, 2010). Since nuage is positioned around the nucleus, all exported transcripts can be 'scanned' and degraded if required, for instance in the case of TE transcripts.

In order to initiate precursor transcript processing, a 'trigger' event takes place catalyzed by the PIWI protein Ago3, bound to a (sense) piRNA, which recognizes the antisense precursor transcript and cleaves it between position 10-11 of the piRNA target site (Figure 3). Through endonuclease activity of Zucchini (Zuc, Pld6) long piRNA precursors are processed into smaller fragments, creating the 3' end of a so-called 'pre-pre-piRNA' and the 5' end of the remaining, unprocessed transcript, the 'pre-pre-piRNA' (Figure 3) (Han et al., 2015; Malone et al., 2009; Mohn et al., 2015; Nishimasu et al., 2012; Watanabe et al., 2011). The first, 'responder' piRNA is bound by the PIWI protein Aubergine (Aub), and the following 'trail' piRNAs by a third PIWI protein, Piwi (Figure 3). This way, Zuc creates piRNAs in a phased manner (this process is therefore referred to as 'phasing'), cutting the transcript at a certain distance from the 1 position

1 from the pre-pre-piRNA, where it will cleave before a Uracil of the next piRNA and the next and so on (Figure 3).

This mechanism creates 3' ends that are generally not very well defined, besides a preference for U at the '+1' position. Therefore, piRNA lengths can vary widely and since PIWI proteins typically interact with piRNAs of a certain length distribution, pre-piRNAs often exceed this (Aravin et al., 2006; Brennecke et al., 2007; Gainetdinov et al., 2018; Girard et al., 2006; Houwing

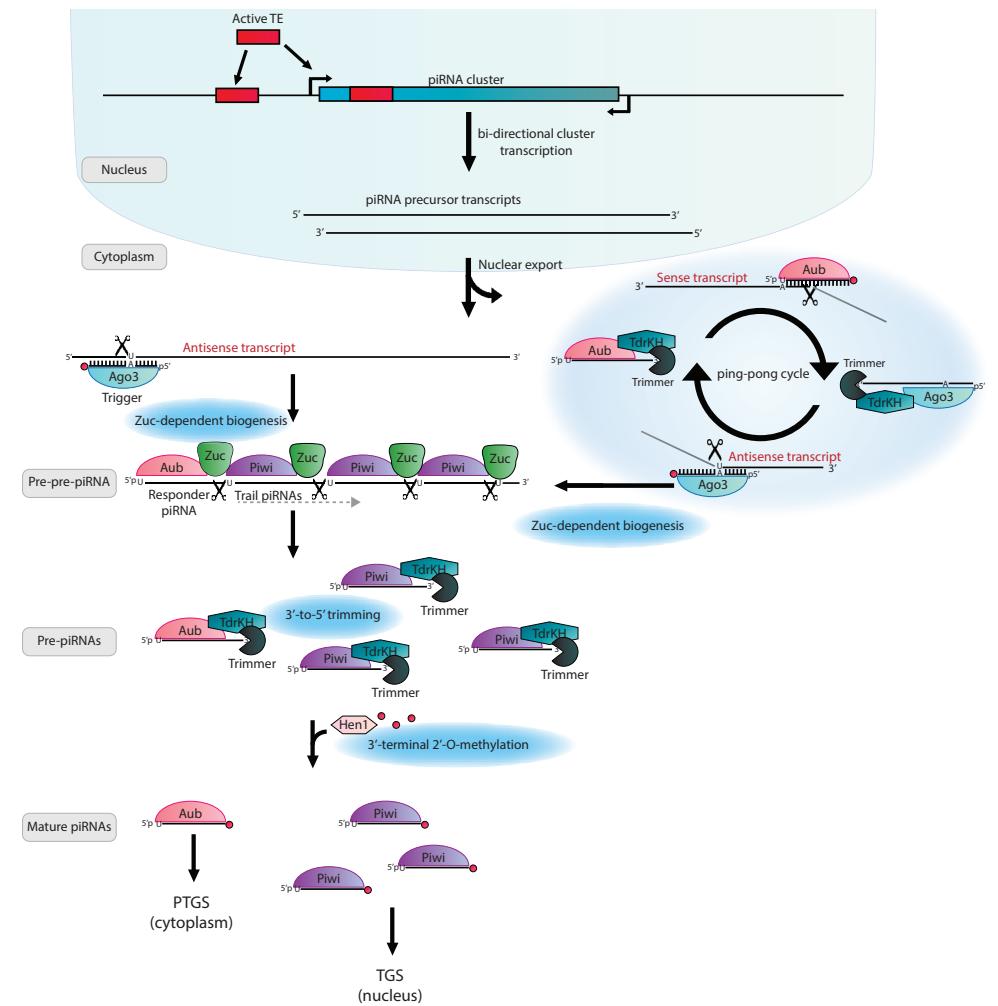


Figure 3. The current model of piRNA biogenesis. An active TE that ends up in a piRNA cluster is transcribed as a part of a large piRNA cluster. Upon export, cluster transcripts can be processed into smaller, phased, pre-piRNAs, through Zucchini-mediated cleavage, after Ago3 has triggered this event. PiRNA maturation involves trimming and methylation by Hen1. The ping-pong amplification loop targets transposons in a feed-forward loop, and expands piRNAs against a certain sequence each cleavage event. At the same time, the ping-pong cycle can provide pre-pre-piRNAs to the Zuc-dependent arm of piRNA biogenesis.

1

et al., 2008; Huang et al., 2011). This is why many pre-piRNAs undergo an additional processing step by PNLDC1 or ‘Trimmer,’ an enzyme with 3’-to-5’ exonuclease activity (Ding et al., 2017; Izumi et al., 2016; Nishimura et al., 2018; Zhang et al., 2017b). Trimming requires the Tudor domain-containing protein TdrKH (a.k.a. Tdrd2 or *Partner of Piwis*; Papi) which interacts with PIWI proteins and recruits PNLDC1/Trimmer, which in turn trims the piRNA until its mature length (Honda et al., 2013; Liu et al., 2011; Saxe et al., 2013)(Figure 3). *Drosophila* has no copy of PNLDC1 and there, at least one other 3’-to-5’ exonuclease is involved in piRNA trimming, Nibbler (Exd3, MUT-7), an enzyme that is also involved in miRNA processing (Feltzin et al., 2015; Hayashi et al., 2016; Wang et al., 2016). Finally, the methyltransferase HEN1 methylates and stabilizes the piRNAs at their 3’ ends (Kamminga et al., 2010; Saito et al., 2007). Mature Aub-piRNA complexes stay in the cytoplasm in order to establish post-transcriptional gene silencing (PTGS), whereas Piwi-piRNA complexes go to the nucleus and can silence transposons at the transcriptional level (TGS) (Figure 3).

The secondary pathway, or ‘ping-pong’ cycle, relies on the synergy between two PIWI paralogs, in the case of *Drosophila* Ago3 and Aub, that are often found to preferentially bind piRNAs that are either sense (Ago3) or antisense (Aub) to TE sequences. Ago3 bound to a sense piRNA cleaves an antisense transcript, derived from a piRNA cluster for instance, followed by binding of the cleavage product by Aub (Figure 3). Slicing activity of Ago3 therefore provides the precursor piRNA for the other PIWI protein, Aub. Upon loading, Aub in turn can target sense TE sequences (Figure 3). This cleavage product is handed over again to Ago3 which can cleave antisense targets again, closing the cycle. This way, the silencing signal is specifically amplified when targets are highly expressed, i.e. posing the highest threat. piRNAs derived from the ping-pong cycle undergo further maturation steps as well, including trimming and HEN1-mediated methylation.

The primary and secondary pathway are not strictly separated, but rather entangled with each other. Phasing initiation requires target cleavage by Ago3, in order for Aub and Piwi to continue (Figure 3). Consequently, bioinformatic analysis will identify both a ping-pong pair, typical for the secondary pathway, as well as a phasing signature as found in the primary pathway on cluster transcripts. Therefore, a ‘unified model’ has recently been proposed, integrating both pathways (Gainetdinov et al., 2018). Furthermore, the historical division of primary versus secondary has become blurred by now, since the secondary pathway also prompts the primary pathway. For this reason, novel views rather distinguish piRNA biogenesis pathways on their dependence or independence on Zuc-mediated cleavage (Czech et al., 2018).

piRNAs in mammals

In mammals, four PIWI paralogs exist, PIWIL1-4, of which PIWIL1, 2 and 4 are present in mouse as well. Interestingly, they do not act simultaneously but have different expression patterns and serve two distinct pathways involving piRNAs: the pre-pachytene and pachytene pathway. Mili (PIWIL2) starts to become expressed in PGCs that have arrived in the gonadal ridge at E12.5 and remains expressed until adulthood, up to the round spermatid stage (Aravin

et al., 2008). Miwi2 (PIWIL4) on the other hand, has a more narrow window of expression between E15.5 up until 4 days post partum (Aravin et al., 2008). During fetal development, Miwi2 and Mili interact with piRNAs that map to TE sequences and Miwi2 can translocate to the nucleus and establishes epigenetic silencing of target sequences (Aravin et al., 2008; Kojima-Kita et al., 2016; Kuramochi-Miyagawa et al., 2008). Loss of either Mili or Miwi2 therefore results in failure to establish *de novo* DNA methylation of genomic TE loci, which will become highly expressed, resulting in DNA damage and germ cell loss (Carmell et al., 2007; Kuramochi-Miyagawa et al., 2004). Interestingly, expression of Miwi2 is restricted to the male gonad and therefore, silencing of TEs by these fetal or ‘pre-pachytene’ piRNAs and mutant phenotypes occur in the male germline only (Aravin et al., 2008; Carmell et al., 2007).

Another class of piRNAs exists in mammals, which do not target TEs, but are rather depleted of repetitive sequences. They have been studied mostly in the murine male germline, where they are expressed postnatally at the onset of meiosis prior to the pachytene stage. The PIWI protein that functions in this postnatal pathway, in addition to Mili, is Miwi (PIWIL1), which also becomes expressed around this time (Deng and Lin, 2002; Grivna et al., 2006a; Li et al., 2013). These so-called pachytene piRNAs are transcribed from specific clusters, different from the ones transcribed for pre-pachytene piRNA production. These clusters are driven by the A-MYB transcription factor, just as the PIWI pathway factors that are required for the pachytene pathway to occur, including Miwi (Li et al., 2013). Pachytene piRNAs are functionally distinct from the TE-targeting pre-pachytene piRNAs, and are required for successful meiosis. This potentially occurs through post-transcriptional regulation of spermatogenic genes by Miwi, since mRNA transcriptome defects occur in *miwi* mutants (Gou et al., 2014; Grivna et al., 2006b; Vourekas et al., 2012). Furthermore, even though pachytene piRNAs generally do not target TEs, upregulation of the L1 TE and DNA damage has been observed as well in catalytically dead *miwi* mutants, indicating a more classic piRNA function for this pathway as well (Reuter et al., 2011).

Tudor domain-containing proteins in germ cells

The piRNA pathway cannot take place without a plethora of accessory proteins that assist PIWI-piRNA complexes to be formed and to function adequately. One family of proteins that is highly represented in essentially all steps of the piRNA pathway are Tudor domain-containing proteins (Tdrds). A genetic screen in *Drosophila* for maternal factors required for fertility of the progeny identified *tudor* (*tud*) (Boswell and Mahowald, 1985). Tud is a protein that was found to consist of 11 analogous domains that were consequently called ‘Tudor domains,’ named after the ‘grandchildlessness’ of the English ‘House of Tudor’ which went extinct after lack of a successor in the family line (Boswell and Mahowald, 1985; Callebaut and Mornon, 1997; Ponting, 1997). This conserved motif comprises ~60 amino acids that form a barrel-like fold consisting of four to five anti-parallel β -sheets, creating an aromatic, hydrophobic cage (Sprangers et al., 2003). The Tudor domain is involved in protein-protein interactions, often mediated by symmetrically dimethylated arginine (sDMA), and sometimes lysine, residues (Côté and Richard, 2005; Kim et al., 2006; Sprangers et al., 2003).

1

Table 1. Tdrd proteins in mouse, *Drosophila*, zebrafish and silkworm

Name	Alias	Tudor domains	Additional domain	Function	Reference
Tudor		11	-	<i>Drosophila</i> : Polar granule assembly, germ cell formation	(Arkov et al., 2006; Boswell and Mahowald, 1985a)
Tdrd1		4	1 x MYND	Mouse: male germ cell development, piRNA biogenesis, nuage formation Zebrafish: germ cell development, nuage stabilization, molecular scaffold for Zwi and Zili	(Chuma et al., 2006b; Reuter et al., 2009) (Huang et al., 2011)
Tdrd2	Partner of Pkwis (PaPi)/TdrKH	1	2 x KH	<i>Drosophila</i> : trimming Mouse: trimming, recruitment of PNLDc1 Silkworm: trimming, recruitment of PNLDc1 Zebrafish: trimming	(Liu et al., 2011) (Chen et al., 2009; Ding et al., 2017; Saxe et al., 2013) (Honda et al., 2013; Izumi et al., 2016) This thesis
Tdrd3		1	1 x DUF, 1 x UBA	Mouse: suppression of R-loop formation through TOP3 β recruitment <i>Drosophila</i> : stress granule localization, synapse formation, bridging TOP3 β with FMRP	(Yang et al., 2014) (Xu et al., 2013)
Tdrd4	Kumo/Qin/Rnf17	5	BBC	Mouse: fertility, suppression of promiscuous PIWI protein activity	(Pan, 2005; Wasik et al., 2015)
Tdrd5	Tejas	1	3 x LOTUS	<i>Drosophila</i> : piRNA biogenesis, heterotypic ping-pong Silkworm: piRNA biogenesis <i>Drosophila</i> : transposon silencing, piRNA production, localization of piRNA pathway factors (together with Tdrd7/Tapas) Mouse: pachytene piRNA biogenesis, chromatoid body assembly, piRNA-precursor binding	(Anand and Kai, 2012; Zhang et al., 2011) (Nishida et al., 2015) (Patil et al., 2014) (Ding et al., 2018)
Tdrd6		7	-	Mouse: chromatoid body architecture, spliceosome assembly Zebrafish: normal Balbiani body and germ plasm formation	(Akpınar et al., 2017; Vasileva et al., 2009) This thesis

Table 1. Continued

Name	Alias	Tudor domains	Additional domain	Function	Reference
Tdrd7	Tapas	4	3 x LOTUS	<i>Drosophila</i> : transposon silencing, piRNA production, localization of piRNA pathway factors (together with Tdrd5/Tejas) Mouse: spermatogenesis, lens development, post-transcriptional control of lens mRNAs	(Patil and Kai, 2010; Patil et al., 2014) (Lachke et al., 2011)
Tdrd8	Stk31	1	poIC, PKC-like	Mouse: spermatogenesis	(Bao et al., 2013; Zhou et al., 2014)
Tdrd9	Spindle-E (Spn-E)	1	1 x HA2, 1 x HELICc	Mouse: transposon silencing, fertility <i>Drosophila</i> : piRNA biogenesis, transposon silencing Silkworm: primary piRNA biogenesis	(Shoji et al., 2009) (Malone et al., 2009) (Nishida et al., 2015)
Tdrd12	Yb	2	1 x DEXDc, 1 x ACD	<i>Drosophila</i> : Somatic piRNA biogenesis, recognition and phased processing of piRNA precursors in follicle cells Mouse: Transposon silencing, secondary piRNA biogenesis, fertility, Exd1 recruitment and MIWI2 loading Zebrafish: germ cell development	(Pandey et al., 2013, 2018) (Dai et al., 2017)

KH: K-Homology; **ACD**: α -crystallin domain; **DEXDc**: DEAD-like helicase superfamily; **DUF**: domain of unknown function, **SNC**: Staphylococcal nuclease homologs; **MYND**: MYND finger; **UBA**: ubiquitin-associated domain; **BBC**: B-Box C-terminal domain; **RRM**: RNA Recognition Motif; **HA2**: helicase-associated domain; **PKC-like**: protein kinases.

1

Besides the Tudor domain itself, the plant-specific Agenet domain, the chromo domain, the PWWP domain and the Agenet-like domain share an ancestral core of three beta strands and together form the ‘Royal’ domain superfamily (Maurer-Stroh et al., 2003). Tdrds are widespread in germ cell biology and particularly important in germ cell-specific structures like nuage. Since Tdrds are implicated in protein-protein interactions, it has been suggested that they can act as adapter proteins or platforms that dock and locally accumulate piRNA pathway components and stabilize the nuage structure, so that all components stay concentrated in order to act efficiently.

Germline Tdrds often contain multiple Tudor domains in tandem, together with additional domains that are involved in RNA interactions such as helicases, KH domains and LOTUS domains (Pek et al., 2012; Siomi et al., 2010) (Table 1). Furthermore, germline Tdrds often contain an additional N-terminal α -helix and two β -strands, named the Maternal Tudor or extended Tudor domain (eTUD) and it is thought that the extension stimulates Tdrd-sDMA interactions (Bateman, 2004; Liu et al., 2010). PIWI proteins typically have a relatively disordered N-terminal domain, containing arginine-glycine(RG)-motifs which can be recognized by Tdrds, often in a symmetrically dimethylated state of the arginine residue (Huang et al., 2011; Kirino et al., 2009; Nishida et al., 2009; Reuter et al., 2009; Vagin et al., 2009). Besides, another conserved and essential piRNA pathway factor Vasa contains sDMAs as well and can interact with Tdrds, indicating this type of interaction could be a more general feature of the piRNA pathway (Kirino et al., 2010a). On the other hand, some Tdrd-PIWI interactions were shown to be independent of sDMAs, since Tdrds either bear mutations in the conserved aromatic residues or, by mutating the arginine residues, were shown to be independent of sDMAs (Handler et al., 2011, 2013; Patil and Kai, 2010; Zhang et al., 2017a). Most germline Tdrds contain at least one Tudor domain with an intact conserved aromatic cage, which could indicate that the other Tudor domains have adapted their binding specificity for other types of interactions, potentially evolving Tdrd function. Interactions between Tdrds and PIWI pathway factors have been shown to be essential for PIWI pathway function and fertility (Kirino et al., 2009, 2010b; Liu et al., 2010). Animal germlines express a wide variety of Tdrds and by now have been implicated in essentially all steps of the PIWI pathway, from biogenesis, to amplification, to inheritance (Chuma et al., 2006; Honda et al., 2013; Huang et al., 2011; Kirino et al., 2009; Nishida et al., 2009; Saxe et al., 2013; Shoji et al., 2009).

GERM CELLS AS A MODEL FOR PRION BIOLOGY

Over the last years, there has been increasing interest in the biology of membrane-less compartments. At the core of this phenomenon, called ‘liquid-liquid phase separation’ (LLPS), are proteins that have long been largely overlooked since they often lack typical protein domain structures and are rather *disordered*, meaning that the three-dimensional structure of a protein can be fluctuating more than is typically the case. Protein structures can in general be dynamic, however, proteins that trigger LLPS often contain stretches of low complexity that

are unstructured in solution, which nevertheless can adopt definite, potentially well-defined structures in a different – *separated* from the cytosolic – state.

Prion domains in biology – a widespread phenomenon

Prions were discovered only after a time during which mysterious infectious diseases occurred caused by factors termed ‘slow viruses’ (Lantos, 1992; Sigurdsson, 1954). The infective agents had not been identified, however, while it did spread from one individual to the other, including cross-species transmission, through the consumption of infected brain tissue. Examples are Creutzfeldt-Jacob disease (CJD), scrapie and bovine spongiform encephalopathy (BSE) (Lantos, 1992). Stanley Prusiner found that what these diseases had in common was that they were actually not caused by viruses, but rather by aberrantly folded versions of the ‘Prion protein’ (PrP) (Prusiner, 1982, 1998). He coined the term ‘prion,’ based on ‘proteinaceous’ and ‘infectious.’ Even though these proteins do not contain genetic material, they are still capable of spreading information by stimulating the healthy form of the protein (PrP^C) to refold and adopt the conformation of the disease-causing PrP^{Sc} by acting as a dominant template (Prusiner, 1998).

Later, more proteins with prion properties were identified in yeast, including Ure2p and Sup35 that are known to have prion states [URE3] and [PSI] (Wickner, 1994). These proteins however, require chaperone activity of Hsp104 in order to break them down and convert them into a propagating prion-like state (Parsell et al., 1994; Wegrzyn et al., 2001). These were indications that prions could serve normal cellular function as well, for instance following environmental stresses. By now, many proteins have been identified which bear ‘prion-like domains’ (PLDs) – domains that are similar to yeast prion domains – which went unnoticed before. Generally, these domains are of low complexity (LC), overrepresenting certain amino acids, and disregarded in the search of ordered protein regions that fold into functional motifs in the more classical sense. Typically, PLDs are rich in polar, uncharged residues and PLD-containing proteins have been shown to be able to phase-separate into droplets *in vitro*, which requires a certain concentration and mixing with crowding agents like dextran or PEG (Boeynaems et al., 2017; Boke et al., 2016; Franzmann et al., 2018; Patel et al., 2015; Woodruff et al., 2017). This suggests that *in vivo*, a concentration threshold is required to phase-separate. Furthermore, molecular seeds like Poly(ADP ribose) (or ‘PAR’), RNA, or environmental factors like temperature or pH can influence phase-separating properties as well (Franzmann et al., 2018; Langdon et al., 2018; Molliex et al., 2015; Patel et al., 2015; Smith et al., 2016; Zhang et al., 2015).

It has now been established that PLDs are widespread and found in proteins involved in numerous cellular processes, ranging from cell division, to stress responses, to heterochromatin formation and many more (Brangwynne et al., 2011; Franzmann et al., 2018; Larson et al., 2017; Lemke, 2016; Molliex et al., 2015; Woodruff et al., 2017). Incorporating such domains in proteins can therefore support the cell to temporarily store proteins and transcripts, like stress granules, or preserve certain conformations more long-term, like heterochromatin. On the other hand, PLDs also bring about the risk of irreversible aggregation, which ultimately

1

1 results in deposition of aberrant, non-functional protein aggregates that can be cytotoxic and disease-causing. Alzheimer's disease, Parkinson's disease and Amyotrophic Lateral Sclerosis (ALS) are all examples of illnesses caused by detrimental depositions of misfolded, irreversibly aggregated proteins. The causes of such diseases can be to a certain extent genetic, usually in combination with aging. The latter implicating the *duration* of aggregation contributes to its irreversibility. In *in vitro* studies, droplet 'aging' has been demonstrated, which can ultimately turn into irreversible 'solids,' stimulated further by typical disease mutations (Patel et al., 2015; Zhang et al., 2015). Considering aging as a common factor in these diseases, it can be speculated that, up to a certain point, these protein condensates are kept in check as healthy and functional phase-separated structures. However, upon turnover towards a cytotoxic version, possibly because these structures themselves 'age' as well, control mechanisms fall short. This, combined with additional deleterious effects by disease-associated mutations, can result in life-threatening illnesses. It is therefore important to study not only how these condensates assemble, but also how they are controlled and broken down again, in order to learn how disease phenotypes can be treated or even prevented.

Phase-separated structures in germ cell biology

Over the last decade, people have started to study phase-separated structures in many different healthy contexts as well. One of the first realizations that a commonly studied structure was actually an example of LLPS, was the finding that P-granules in the *C. elegans* gonad behave like liquid droplets *in vivo* (Brangwynne et al., 2009). Many structures related to P-granules exist across species. Moreover, of all protein coding genes in humans, ~1% contains a predicted PLD and this presence is biased towards RNA-binding proteins (RBPs), of which ~12% contains a PLD (King et al., 2012). Germ cells of animals provide a broad variety of phase-separated granules that are rich in RNA and RBPs. Next to P-granules, examples are nuage, chromatoid bodies, mutator foci and novel variants that are still being discovered (Brangwynne et al., 2009; Chuma et al., 2006; Phillips et al., 2012; Senti and Brennecke, 2010; Updike and Strome, 2010; Wan et al., 2018). Moreover, the above-discussed conserved germ cell-specific Balbiani body has been shown to display amyloid-like features in a healthy, wild-type context (Boke et al., 2016). Interestingly, this is characteristic for pathological variants of Amyloid- β and Tau, two prion proteins that can cause various forms of dementia in a similar way as CJD is caused, through cytotoxic depositions, even though transmissibility of Amyloid- β and Tau prions between people is still under investigation (Abbott, 2015; Daviglus et al., 2011; Goedert and Spillantini, 2006; Goudsmit et al., 1980). Besides the Balbiani body, many other physiological, functional structures adopt a fibrillary structure similar to amyloids (amyloid-like). Examples are insect silks, Pmel1 in melanosomes (organelles that synthesize pigment) and certain fungal adhesion molecules (Fowler et al., 2006; Garcia et al., 2011; Ramsook et al., 2010; Slotta et al., 2007). Incorporating amyloid-like fibers in a structure therefore seems to be a more general mechanism in biological systems in order to provide additional rigidity to a structure.

1 Even though the Balbiani body has amyloid-like properties, this structure remains dynamic throughout oogenesis and transforms from a single, large assembly into fragments during oogenesis, and back into larger assemblies upon fertilization. Furthermore, it contains several key components that can modulate phase separation, including PLD-containing proteins (Buc), mRNA (early and late pathway components) and Tdrds (Tdrd6a, see chapter 4). The same is true for germ plasm-related structures in other species, like P-granules in *C. elegans* and polar granules in *Drosophila*.

P-granules consist of a 'soft' core, made up by PGL-1 and PGL-3, that are predicted to have RNA-binding capacity through their C-terminal RGG-repeats (Aoki et al., 2016; Kawasaki et al., 1998; Seydoux, 2018). The RGG-repeat, and therefore potentially RNA, is required to stabilize this condensate and recruit other P-granule components, including MEG-3 and MEG-4. These MEG proteins, that assemble to PGL-containing granules later during oogenesis, provide a sturdy, outer layer ('cage-like' scaffold), required for P-granule stability and localization (Wang et al., 2014). In the fertilized egg, competition for mRNA with an anterior RNA-binding protein MEX-5, restricts P-granules to the embryonic posterior (Smith et al., 2016).

In *Drosophila* embryos, germ plasm granules marked by Oskar, Vasa and Tudor contain transcripts that accumulate more and more towards the posterior pole. These transcripts seed their own accumulation in a homotypic fashion, followed by combination of mRNA-containing aggregates into heterotypic granules posteriorly, resulting in an organized deposition of transcript ratios into germ cells (Little et al., 2015; Niepielko et al., 2018; Trcek et al., 2015).

Just as Buc, other germ plasm core factors including PGL and MEG proteins and Oskar and Tudor contain low complexity domains and PLDs. And similar to zebrafish, granule size, composition and localization are subject to continuous change throughout oogenesis and embryogenesis, illustrative for a highly coordinated process. It is unknown whether P-granules or polar granules, or a certain sub-compartment within these structures, is amyloid-like too, but the concept of providing sturdiness to these kind of heritable structures seems to be a common theme. Therefore, germ cells can be exploited to explore diverse aspects of phase-separated structures – including assembly, disassembly and misaggregation – in an *in vivo* setting using genetic tools, live cell imaging and biochemical approaches.

OUTSTANDING QUESTIONS

One of the major limitations of knowledge regarding the PIWI pathway is that it has been studied extensively, though only in few organisms. In *Drosophila*, it has been studied in such depth that stands in vast contrast with other species, even though many crucial factors are not conserved, like the Tdrds Krimper and Vreteno and factors of the nuclear pathway (Handler et al., 2011; Klattenhoff et al., 2009; Mohn et al., 2014; Sato et al., 2015; Webster et al., 2015). Furthermore, the principal mammalian model organism, mouse, has lost PIWI paralog PIWIL3, which is present in most other mammals, including humans. This indicates that even though certain conserved concepts have been revealed, much still remains to be discovered regarding piRNA pathway biogenesis and function. In this thesis, we aim to broaden our knowledge by

1 exploring two aspects of the piRNA pathway, using zebrafish and mammalian systems, other than mice and rats. In chapter 2, we study the function of the Tdrd TdrKH in zebrafish, which is required in mouse and silkworm to recruit the 3'-to-5' exonuclease PNLDC1. Studies in mouse, silkworm and *Drosophila* have shown that TdrKH is required for trimming, but to which extent seems to vary widely, just as the outcoming phenotypes regarding fertility. We investigate TdrKH function in zebrafish, in order to distinguish species-specific effects from general concepts regarding piRNA trimming.

In chapter 3, we explore potential piRNA-pathway function in the female mammalian germline. In mouse and rat, the piRNA pathway is only essential in males, whereas in other species like *Drosophila*, silkworm and zebrafish, the piRNA-pathway plays an important role, particularly in ovary. Is loss of PIWIL3 related to loss of ovarian piRNA function? Since most other mammals still have this copy, this discrepancy has raised the suspicion that PIWIL3 could play an ancestral role in protecting the female mammalian germline. We investigate gonadal material from cows, macaques and humans of different stages and provide compelling evidence of the presence of an active PIWI pathway in the female mammalian germline and early embryogenesis, which has been disregarded before.

Next, we investigate aspects of phase separation in zebrafish germ cells and embryos and the role of Tdrds and sDMAs in modulating these structures. Tdrds and RG-motifs are widespread in phase-separated structures, although their precise roles are still poorly understood, particularly in *in vivo* settings. We will study phenotypes in zebrafish oogenesis and embryogenesis caused by mutations in a Tdrd, Tdrd6a (chapter 4), and in RG motifs in Buc (chapter 5), in order to address these issues. Together, our findings provide contributions to the general understanding of phase-separated structures, using an *in vivo* model, and how Tdrds and sDMAs could provide a control mechanism over assembly, disassembly and content of phase-separated structures.

REFERENCES

- Abbott, A. (2015). News In Focus: Autopsies reveal signs of Alzheimer's in growth-hormone patients. *Nature* 525, 165–166.
- Aguero, T., Zhou, Y., Kloc, M., Chang, P., Houliston, E., and King, M. (2016). Hermes (Rbpm5) is a Critical Component of RNP Complexes that Sequester Germline RNAs during Oogenesis. *J. Dev. Biol.*
- Akpınar, M., Lesche, M., Fanourgakis, G., Fu, J., Anastasiadis, K., Dahl, A., and Jessberger, R. (2017). TDRD6 mediates early steps of spliceosome maturation in primary spermatocytes. *PLoS Genet.* 13.
- Anand, A., and Kai, T. (2012). The tudor domain protein Kumo is required to assemble the nuage and to generate germline piRNAs in *Drosophila*. *EMBO J.*
- Ancelin, K., Lange, U.C., Hajkova, P., Schneider, R., Bannister, A.J., Kouzarides, T., and Surani, M.A. (2006). Blimp1 associates with Prmt5 and directs histone arginine methylation in mouse germ cells. *Nat. Cell Biol.* 8, 623–630.
- Aoki, S.T., Kershner, A.M., Bingman, C.A., Wickens, M., and Kimble, J. (2016). PGL germ granule assembly protein is a base-specific, single-stranded RNase. *Proc. Natl. Acad. Sci.*
- 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., 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., 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.
- Baillie, J.K., Barnett, M.W., Upton, K.R., Gerhardt, D.J., Richmond, T.A., De Sapio, F., Brennan, P., Rizzu, P., Smith, S., Fell, M., et al. (2011). Somatic retrotransposition alters the genetic landscape of the human brain. *Nature* 479, 534–537.
- Bao, J., Yuan, S., Maestas, A., Bhetwal, B.P., Schuster, A., and Yan, W. (2013). Stk31 is dispensable for embryonic development and spermatogenesis in mice. *Mol. Reprod. Dev.*
- Bateman, A. (2004). The Pfam protein families database. *Nucleic Acids Res.*
- Bergstrom, C.T., and Pritchard, J. (1998). Germline bottlenecks and the evolutionary maintenance of mitochondrial genomes. *Genetics.*
- Boeke, J.D., and Voytas, D.F. (2002). Ty1 and Ty5 of *Saccharomyces cerevisiae*. In *Mobile DNA II*, p.
- Boeynaems, S., Bogaert, E., Kovacs, D., Konijnenberg, A., Timmerman, E., Volkov, A., Guharoy, M., De Decker, M., Jaspers, T., Ryan, V.H., et al. (2017). Phase Separation of C9orf72 Dipeptide Repeats Perturbs Stress Granule Dynamics. *Mol. Cell.*
- Boke, E., Ruer, M., Wühr, M., Coughlin, M., Lemaitre, R., Gygi, S.P., Alberti, S., Drechsel, D., Hyman, A.A., and Mitchison, T.J. (2016). Amyloid-like Self-Assembly of a Cellular Compartment. *Cell* 166, 637–650.
- 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.
- Boswell, R.E., and Mahowald, A.P. (1985). tudor, a gene required for assembly of the germ plasm in *Drosophila melanogaster*. *Cell* 43, 97–104.
- Braat, A.K., Zandbergen, T., Van De Water, S., Goos, H.J.T.H., and Zivkovic, D. (1999). Characterization of zebrafish primordial germ cells: Morphology and early distribution of vasa RNA. *Dev. Dyn.*
- 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

- 1 by Controlled Dissolution/Condensation. *Science* (80-.). 324, 1729–1732.
22. Brangwynne, C.P., Mitchison, T.J., and Hyman, A.A. (2011). Active liquid-like behavior of nucleoli determines their size and shape in *Xenopus laevis* oocytes. *Proc. Natl. Acad. Sci.* 108, 4334–4339.
23. 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.
24. Brouha, B., Schustak, J., Badge, R.M., Lutz-Prigge, S., Farley, A.H., Moran, J. V., and Kazazian, H.H. (2003). Hot L1s account for the bulk of retrotransposition in the human population. *Proc. Natl. Acad. Sci.*
25. Callebaut, I., and Mornon, J.P. (1997). The human EBNA-2 coactivator p100: multidomain organization and relationship to the staphylococcal nuclease fold and to the tudor protein involved in *Drosophila melanogaster* development. *Biochem. J.* 321 (*Pt 1*), 125–132.
26. 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.
27. 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. (2009). Mouse Piwi interactome identifies binding mechanism of Tdrkh Tudor domain to arginine methylated Miwi. *Proc. Natl. Acad. Sci.*
28. Cheng, E.C., Kang, D., Wang, Z., and Lin, H. (2014). PIWI proteins are dispensable for mouse somatic development and reprogramming of fibroblasts into pluripotent stem cells. *PLoS One*.
29. 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.
30. Côté, J., and Richard, S. (2005). Tudor domains bind symmetrical dimethylated arginines. *J. Biol. Chem.*
31. Cox, R.T. (2003). A Balbiani body and the fusome mediate mitochondrial inheritance during *Drosophila* oogenesis. *Development*.
32. Crother, B.I., White, M.E., and Johnson, A.D. (2016). Diversification and Germ-Line Determination Revisited: Linking Developmental Mechanism with Species Richness. *Front. Ecol. Evol.* 4, 26.
33. Czech, B., Munafò, M., Ciabrelli, F., Eastwood, E.L., Fabry, M.H., Kneuss, E., and Hannon, G.J. (2018). PiRNA-guided genome defense: From biogenesis to silencing. *Annu. Rev. Genet.*
34. Dai, X., Shu, Y., Lou, Q., Tian, Q., Zhai, G., Song, J., Lu, S., Yu, H., He, J., and Yin, Z. (2017). Tdrd12 is essential for germ cell development and maintenance in zebrafish. *Int. J. Mol. Sci.*
35. Daviglus, M.L., Plassman, B.L., Pirezada, A., Bell, C.C., Bowen, P.E., Burke, J.R., Connolly, E.S., Dunbar-Jacob, J.M., Granieri, E.C., McGarry, K., et al. (2011). Risk factors and preventive interventions for alzheimer disease: State of the science. *Arch. Neurol.*
36. Deng, W., and Lin, H. (2002). miwi, a murine homolog of piwi, encodes a cytoplasmic protein essential for spermatogenesis. *Dev Cell* 2, 819–830.
37. Ding, D., Liu, J., Dong, K., Midic, U., Hess, R.A., Xie, H., Demireva, E.Y., and Chen, C. (2017). PNLDC1 is essential for piRNA 3' end trimming and transposon silencing during spermatogenesis in mice. *Nat. Commun.* 8.
38. Ding, D., Liu, J., Midic, U., Wu, Y., Dong, K., Melnick, A., Latham, K.E., and Chen, C. (2018). TDRD5 binds piRNA precursors and selectively enhances pachytene piRNA processing in mice. *Nat. Commun.*
39. Draper, B.W., McCallum, C.M., and Moens, C.B. (2007). nanos1 is required to maintain oocyte production in adult zebrafish. *Dev. Biol.* 305, 589–598.
40. Eddy, E.M. (1976). Germ Plasm and the Differentiation of the Germ Cell Line. *Int. Rev. Cytol.*
41. Elkouby, Y.M., Jamieson-Lucy, A., and Mullins, M.C. (2016). Oocyte Polarization Is Coupled to the Chromosomal Bouquet, a Conserved Polarized Nuclear Configuration in Meiosis. *PLoS Biol.* 14.
42. Ephrussi, A., and Lehmann, R. (1992). Induction of germ cell formation by oskar. *Nature* 358, 387–392.
43. Ephrussi, A., Dickinson, L.K., and Lehmann, R. (1991). Oskar organizes the germ plasm and directs localization of the posterior determinant nanos. *Cell* 66, 37–50.
44. Escobar-Aguirre, M., Zhang, H., Jamieson-Lucy, A., and Mullins, M.C. (2017). Microtubule-actin crosslinking factor 1 (Macf1) domain function in Balbiani body dissociation and nuclear positioning. *PLoS Genet.* 13.
45. Evans, T., Wade, C.M., Chapman, F.A., Johnson, A.D., and Loose, M. (2014). Acquisition of germ plasm accelerates vertebrate evolution. *Science* (80-.). 344, 200–203.
46. Extavour, C.G. (2003). Mechanisms of germ cell specification across the metazoans: epigenesis and preformation. *Development* 130, 5869–5884.
47. Fan, W., Waymire, K.G., Narula, N., Li, P., Rocher, C., Coskun, P.E., Vannan, M.A., Narula, J., MacGregor, G.R., and Wallace, D.C. (2008). A mouse model of mitochondrial disease reveals germline selection against severe mtDNA mutations. *Science* (80-.).
48. Feltzin, V.L., Khaladkar, M., Abe, M., Parisi, M., Hendriks, G.J., Kim, J., and Bonini, N.M. (2015). The exonuclease Nibbler regulates age-associated traits and modulates piRNA length in *Drosophila*. *Aging Cell* 14, 443–452.
49. Feschotte, C., and Pritham, E.J. (2007). DNA Transposons and the Evolution of Eukaryotic Genomes. *Annu. Rev. Genet.*
50. Findley, S.D. (2003). Maelstrom, a *Drosophila* spindle-class gene, encodes a protein that colocalizes with Vasa and RDE1/AGO1 homolog, Aubergine, in nuage. *Development*.
51. Flehr, M., Malik, R., Franke, V., Nejezinska, J., Sedlacek, R., Vlahoviček, K., and Svoboda, P. (2013). A retrotransposon-driven Dicer isoform directs endogenous siRNA production in mouse oocytes. *Cell* 155, 807–816.
52. Forristall, C., Pondel, M., Chen, L., and King, M.L. (1995). Patterns of localization and cytoskeletal association of two vegetally localized RNAs, Vg1 and Xcat-2. *Development* 121, 201–208.
53. Fowler, D.M., Koulov, A. V., Alory-Jost, C., Marks, M.S., Balch, W.E., and Kelly, J.W. (2006). Functional amyloid formation within mammalian tissue. *PLoS Biol.*
54. Franzmann, T.M., Jahnel, M., Pozniakovskiy, A., Mahamid, J., Holehouse, A.S., Nüske, E., Richter, D., Baumeister, W., Grill, S.W., Pappu, R. V., et al. (2018). Phase separation of a yeast prion protein promotes cellular fitness. *Science* (80-.). 359.
55. Gainetdinov, I., Colpan, C., Arif, A., Cecchini, K., and Zamore, P.D. (2018). A Single Mechanism of Biogenesis, Initiated and Directed by PIWI Proteins, Explains piRNA Production in Most Animals. *Mol. Cell*.
56. Garcia, M.C., Lee, J.T., Ramsook, C.B., Alsteens, D., Dufrêne, Y.F., and Lipke, P.N. (2011). A role for amyloid in cell aggregation and biofilm formation. *PLoS One*.
57. Girard, A., Sachidanandam, R., Hannon, G.J., and Carmell, M.A. (2006). A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature*.
58. Goedert, M., and Spillantini, M.G. (2006). A century of Alzheimer's disease. *Science* (80-.).
59. Gou, L.T., Dai, P., Yang, J.H., Xue, Y., Hu, Y.P., Zhou, Y., Kang, J.Y., Wang, X., Li, H., Hua, M.M., et al. (2014). Pachytene piRNAs instruct massive mRNA elimination during late spermiogenesis. *Cell Res.*
60. Goudsmit, J., Morrow, C.H., Asher, D.M., Yanagihara, R.T., Masters, C.L., Gibbs, C.J., and Gajdusek, D.C. (1980). Evidence for and against the transmissibility of Alzheimer disease. *Neurology*.
61. Grivna, S.T., Beyret, E., Wang, Z., and Lin, H. (2006a). A novel class of small RNAs in mouse spermatogenic cells. *Genes Dev.*
62. Grivna, S.T., Pyhtila, B., and Lin, H. (2006b). MIWI associates with translational machinery and PIWI-interacting RNAs (piRNAs) in regulating spermatogenesis. *Proc. Natl. Acad. Sci.*

63. Gupta, T., Marlow, F.L., Ferriola, D., Mackiewicz, K., Dapprich, J., Monos, D., and Mullins, M.C. (2010). Microtubule actin crosslinking factor 1 regulates the balbiani body and animal-vegetal polarity of the zebrafish oocyte. *PLoS Genet.* 6.
64. Han, J.S. (2010). Non-long terminal repeat (non-LTR) retrotransposons: Mechanisms, recent developments, and unanswered questions. *Mob. DNA.*
65. Han, B.W., Wang, W., Li, C., Weng, Z., and Zamore, P.D. (2015). PiRNA-guided transposon cleavage initiates Zucchini-dependent, phased piRNA production. *Science (80-)*. 348, 817–821.
66. 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.*
67. Handler, D., Meixner, K., Pizka, M., Lauss, K., Schmied, C., Gruber, F.S., and Brennecke, J. (2013). The genetic makeup of the Drosophila piRNA pathway. *Mol. Cell* 50, 762–777.
68. Hayashi, R., Schnabl, J., Handler, D., Mohn, F., Ameres, S.L., and Brennecke, J. (2016). Genetic and mechanistic diversity of piRNA 3'-end formation. *Nature.*
69. Heasman, J., Quarmby, J., and Wylie, C.C. (1984). The mitochondrial cloud of Xenopus oocytes: The source of germinal granule material. *Dev. Biol.*
70. Heim, A.E., Hartung, O., Rothhamel, S., Ferreira, E., Jenny, A., and Marlow, F.L. (2014). Oocyte polarity requires a Bucky ball-dependent feedback amplification loop. *Development* 141, 842–854.
71. Hill, J.H., Chen, Z., and Xu, H. (2014). Selective propagation of functional mitochondrial DNA during oogenesis restricts the transmission of a deleterious mitochondrial variant. *Nat. Genet.*
72. Hof, A.E.V. t., Campagne, P., Rigden, D.J., Yung, C.J., Lingley, J., Quail, M.A., Hall, N., Darby, A.C., and Saccheri, I.J. (2016). The industrial melanism mutation in British peppered moths is a transposable element. *Nature* 534, 102–105.
73. Honda, S., Kirino, Y.Y., Maragkakis, M., Alexiou, P., Ohtaki, A., Murali, R., Mourelatos, Z., and Kirino, Y.Y. (2013). Mitochondrial protein BmPAPI modulates the length of mature piRNAs. *RNA* 19, 1405–1418.
74. Houston, D.W., Zhang, J., Maines, J.Z., Wasserman, S.A., and King, M.L. (1998). A Xenopus DAZ-like gene encodes an RNA component of germ plasm and is a functional homologue of Drosophila boule. *Development* 125, 171–180.
75. Houwing, S. (2009). Piwi-piRNA complexes in the zebrafish germline.
76. 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.
77. 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.
78. Huang, H.-Y., Houwing, S., Kaaij, L.J.T., Meppelink, A., Redl, S., Gauci, S., Vos, H., Draper, B.W., Moens, C.B., Burgering, B.M., et al. (2011). Tdrd1 acts as a molecular scaffold for Piwi proteins and piRNA targets in zebrafish. *EMBO J.* 30, 3298–3308.
79. Iskow, R.C., McCabe, M.T., Mills, R.E., Torene, S., Pittard, W.S., Neuwald, A.F., Van Meir, E.G., Vertino, P.M., and Devine, S.E. (2010). Natural mutagenesis of human genomes by endogenous retrotransposons. *Cell* 141, 1253–1261.
80. Izumi, N., Shoji, K., Sakaguchi, Y., Honda, S., Kirino, Y., Suzuki, T., Katsuma, S., and Tomari, Y. (2016). Identification and Functional Analysis of the Pre-piRNA 3' Trimmer in Silkworms. *Cell.*
81. Johnson, A.D., and Alberio, R. (2015). Primordial germ cells: the first cell lineage or the last cells standing? *Development.*
82. Kamminga, L.M., Luteijn, M.J., den Broeder, M.J., Redl, S., Kaaij, L.J.T., 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.
83. Kaufman, O.H., Lee, K.A., Martin, M., Rothhämel, S., and Marlow, F.L. (2018). rbpms2 functions in Balbiani body architecture and ovary fate. *PLoS Genet.*
84. Kawasaki, I., Shim, Y.H., Kirchner, J., Kaminker, J., Wood, W.B., and Strome, S. (1998). PGL-1, a predicted RNA-binding component of germ granules, is essential for fertility in *C. elegans*. *Cell.*
85. Kazazian, H.H. (2004). Mobile Elements: Drivers of Genome Evolution. *Science (80-)*. 303, 1626–1632.
86. Kazazian, H.H., Wong, C., Youssoufian, H., Scott, A.F., Phillips, D.G., and Antonarakis, S.E. (1988). Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. *Nature.*
87. Kim, J., Daniel, J., Espejo, A., Lake, A., Krishna, M., Xia, L., Zhang, Y., and Bedford, M.T. (2006). Tudor, MBT and chromo domains gauge the degree of lysine methylation. *EMBO Rep.*
88. King, M. Lou, Zhou, Y., and Bubunenko, M. (1999). Polarizing genetic information in the egg: RNA localization in the frog oocyte. *BioEssays* 21, 546–557.
89. King, O.D., Gitler, A.D., and Shorter, J. (2012). The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease. *Brain Res.*
90. Kirino, Y., Kim, N., de Planell-Saguer, M., Khandros, E., Chiorean, S., Klein, P.S., Rigoutsos, I., Jongens, T.A., and Mourelatos, Z. (2009). Arginine methylation of Piwi proteins catalysed by dPRMT5 is required for Ago3 and Aub stability. *Nat. Cell Biol.* 11, 652–658.
91. 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.
92. 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.
93. 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.
94. Kloc, M., and Etkin, L.D. (1995). Two distinct pathways for the localization of RNAs at the vegetal cortex in Xenopus oocytes. *Development.*
95. 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.
96. Knaut, H., Pelegri, F., Bohmann, K., Schwarz, H., Nusslein-Volhard, C., and Nüsslein-Volhard, C. (2000). Zebrafish vasa RNA but not its protein is a component of the germ plasm and segregates asymmetrically before germline specification. *J Cell Biol* 149, 875–888.
97. Kojima-Kita, K., Kuramochi-Miyagawa, S., Nagamori, I., Ogonuki, N., Ogura, A., Hasuwa, H., Akazawa, T., Inoue, N., and Nakano, T. (2016). MIWI2 as an Effector of DNA Methylation and Gene Silencing in Embryonic Male Germ Cells. *Cell Rep.*
98. Kopranner, 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.
99. Kosaka, K., Kawakami, K., Sakamoto, H., and Inoue, K. (2007). Spatiotemporal localization of germ plasm RNAs during zebrafish oogenesis. *Mech Dev* 124, 279–289.
100. 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.
101. 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.
102. Kurimoto, K., Yabuta, Y., Ohinata, Y., Shigetani, M., Yamanaka, K., and Saitou, M.

- 1 (2008). Complex genome-wide transcription dynamics orchestrated by Blimp1 for the specification of the germ cell lineage in mice. *Genes Dev.*
103. 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.-H., et al. (2011). Mutations in the RNA granule component TDRD7 cause cataract and glaucoma. *Science* 331, 1571–1576.
104. Langdon, E.M., Qiu, Y., Niaki, A.G., McLaughlin, G.A., Weidmann, C.A., Gerbich, T.M., Smith, J.A., Crutchley, J.M., Termini, C.M., Weeks, K.M., et al. (2018). mRNA structure determines specificity of a polyQ-driven phase separation. *Science* (80-).
105. Lantos, P.L. (1992). From slow virus to prion: a review of transmissible spongiform encephalopathies. *Histopathology.*
106. Larson, A., Elnatan, D., Keenen, M., Trnka, M., Johnston, J., Burlingame, A., Agard, D., Redding, S., and Narlikar, G. (2017). Liquid droplet formation by HP1 α suggests a role for phase separation in heterochromatin. *Nature* 547, 236–240.
107. Lawson, K.A., Dunn, N.R., Roelen, B.A., Zeinstra, L.M., Davis, A.M., Wright, C. V., Korving, J.P., and Hogan, B.L. (1999). Bmp4 is required for the generation of primordial germ cells in the mouse embryo. *Genes Dev.* 13, 424–436.
108. Leerberg, D.M., Sano, K., and Draper, B.W. (2017). Fibroblast growth factor signaling is required for early somatic gonad development in zebrafish. *PLoS Genet.*
109. Lei, L., and Spradling, A.C. (2016). Mouse oocytes differentiate through organelle enrichment from sister cyst germ cells. *Science* (80-). 352, 95–99.
110. Lemke, E.A. (2016). The Multiple Faces of Disordered Nucleoporins. *J. Mol. Biol.*
111. Levis, R.W., Ganesan, R., Houtchens, K., Tolar, L.A., and Sheen, F. miin (1993). Transposons in place of telomeric repeats at a *Drosophila* telomere. *Cell.*
112. Li, X.Z., Roy, C.K., Dong, X., Bolcun-Filas, E., Wang, J., Han, B.W., Xu, J., Moore, M.J., Schimenti, J.C., Weng, Z., et al. (2013). An Ancient Transcription Factor Initiates the Burst of piRNA Production during Early Meiosis in Mouse Testes. *Mol. Cell.*
113. Little, S.C., Sinsimer, K.S., Lee, J.J., Wieschaus, E.F., and Gavis, E.R. (2015). Independent and coordinate trafficking of single *Drosophila* germ plasm mRNAs. *Nat. Cell Biol.* 17, 558–568.
114. Liu, K., Chen, C., Guo, Y., Lam, R., Bian, C., Xu, C., Zhao, D.Y., Jin, J., MacKenzie, F., Pawson, T., et al. (2010). Structural basis for recognition of arginine methylated Piwi proteins by the extended Tudor domain. *Proc. Natl. Acad. Sci.*
115. 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.
116. Mahowald, A.P. (1971). Polar granules of *Drosophila*. III. The continuity of polar granules during the life cycle of *Drosophila*. *J. Exp. Zool.*
117. 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.
118. Marlow, F.L. (2017). Mitochondrial matters: Mitochondrial bottlenecks, self-assembling structures, and entrapment in the female germline. *Stem Cell Res.*
119. 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.
120. 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.
121. McClintock, B. (1951). Chromosome organization and genic expression. *Cold Spring Harb. Symp. Quant. Biol.* 16, 13–47.
122. Mohn, F., Sienski, G., Handler, D., and Brennecke, J. (2014). The Rhino-Deadlock-Cutoff Complex Licenses Noncanonical Transcription of Dual-Strand piRNA Clusters in *Drosophila*. *Cell* 157, 1364–1379.
123. Mohn, F., Handler, D., and Brennecke, J. (2015). PiRNA-guided slicing specifies transcripts for Zucchini-dependent, phased piRNA biogenesis. *Science* (80-). 348, 812–817.
124. Molliex, A., Temirov, J., Lee, J., Coughlin, M., Kanagaraj, A.P., Kim, H.J., Mittag, T., and Taylor, J.P. (2015). Phase Separation by Low Complexity Domains Promotes Stress Granule Assembly and Drives Pathological Fibrillization. *Cell.*
125. Molyneaux, K.A., Stallock, J., Schaible, K., and Wylie, C. (2001). Time-lapse analysis of living mouse germ cell migration. *Dev. Biol.*
126. Niepielko, M.G., Eagle, W.V.I., and Gavis, E.R. (2018). Stochastic Seeding Coupled with mRNA Self-Recruitment Generates Heterogeneous *Drosophila* Germ Granules. *Curr. Biol.*
127. 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.
128. Nishida, K.M., Iwasaki, Y.W., Murota, Y., Nagao, A., Mannen, T., Kato, Y., Siomi, H., and Siomi, M.C. (2015). Respective functions of two distinct siwi complexes assembled during PIWI-interacting RNA biogenesis in *bombix* germ cells. *Cell Rep.*
129. Nishimasu, H., Ishizu, H., Saito, K., Fukuhara, S., Kamatani, M.K., Bonnefond, L., Matsumoto, N., Nishizawa, T., Nakanaga, K., Aoki, J., et al. (2012). Structure and function of Zucchini endoribonuclease in piRNA biogenesis. *Nature* 491, 284–287.
130. Nishimura, T., Nagamori, I., Nakatani, T., Izumi, N., Tomari, Y., Kuramochi-Miyagawa, S., and Nakano, T. (2018). PNLDCl, mouse pre-piRNA Trimmer, is required for meiotic and post-meiotic male germ cell development. *EMBO Rep.* e44957.
131. Ohinata, Y., Payer, B., O’Carroll, D., Ancelin, K., Ono, Y., Sano, M., Barton, S.C., Obukhanych, T., Nussenzweig, M., Tarakhovskiy, A., et al. (2005). Blimp1 is a critical determinant of the germ cell lineage in mice. *Nature* 436, 207–213.
132. Pan, J. (2005). RNF17, a component of the mammalian germ cell nuage, is essential for spermiogenesis. *Development.*
133. Pandey, R.R., Tokuzawa, Y., Yang, Z., Hayashi, E., Ichisaka, T., Kajita, S., Asano, Y., Kunieda, T., Sachidanandam, R., Chuma, S., et al. (2013). Tudor domain containing 12 (TDRD12) is essential for secondary PIWI interacting RNA biogenesis in mice. *Proc. Natl. Acad. Sci.*
134. Pandey, R.R., Homolka, D., Chen, K.M., Sachidanandam, R., Fauvarque, M.O., and Pillai, R.S. (2017). Recruitment of Armitage and Yb to a transcript triggers its phased processing into primary piRNAs in *Drosophila* ovaries. *PLoS Genet.*
135. Pandey, R.R., Homolka, D., Olotu, O., Sachidanandam, R., Kotaja, N., and Pillai, R.S. (2018). Exonuclease Domain-Containing 1 Enhances MIWI2 piRNA Biogenesis via Its Interaction with TDRD12. *Cell Rep.*
136. Pardue, M.-L., and DeBaryshe, P.G. (2003). Retrotransposons Provide an Evolutionarily Robust Non-Telomerase Mechanism to Maintain Telomeres. *Annu. Rev. Genet.*
137. Parsell, D.A., Kowal, A.S., Singer, M.A., and Lindquist, S. (1994). Protein disaggregation mediated by heat-shock protein Hsp104. *Nature.*
138. Patel, A., Lee, H.O., Jawerth, L., Maharana, S., Jahnel, M., Hein, M.Y., Stoynev, S., Mahamid, J., Saha, S., Franzmann, T.M., et al. (2015). A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* 162, 1066–1077.
139. Patil, V.S., and Kai, T. (2010). Repression of Retroelements in *Drosophila* Germline via piRNA Pathway by the Tudor Domain Protein Tejas. *Curr. Biol.*
140. Patil, V.S., Anand, A., Chakrabarti, A., and Kai, T. (2014). The Tudor domain protein Tapas, a homolog of the vertebrate Tdrd7, functions in the piRNA pathway to regulate retrotransposons in germline of *Drosophila melanogaster*. *BMC Biol.*
141. Pek, J.W., Anand, A., and Kai, T. (2012). Tudor domain proteins in development. *Development* 139, 2255–2266.
142. Pepling, M.E., Wilhelm, J.E., O’Hara, A.L., Gephardt, G.W., and Spradling, A.C. (2007).

- 1 Mouse oocytes within germ cell cysts and primordial follicles contain a Balbiani body. *Proc. Natl. Acad. Sci.*
143. Phillips, C.M., Montgomery, T.A., Breen, P.C., and Ruvkun, G. (2012). MUT-16 promotes formation of perinuclear Mutator foci required for RNA silencing in the *C. elegans* germline. *Genes Dev.* 26, 1433–1444.
144. Ponting, C.P. (1997). Tudor domains in proteins that interact with RNA. *Trends Biochem. Sci.* 22, 51–52.
145. Prusiner, S.B. (1982). Novel proteinaceous infectious particles cause scrapie. *Science* (80-).
146. Prusiner, S.B. (1998). Prions. *Proc. Natl. Acad. Sci. U. S. A.* 95, 13363.
147. Ramsook, C.B., Tan, C., Garcia, M.C., Fung, R., Soybelman, G., Henry, R., Litewka, A., O'Meally, S., Otoo, H.N., Khalaf, R.A., et al. (2010). Yeast cell adhesion molecules have functional amyloid-forming sequences. *Eukaryot. Cell.*
148. Raz, E. (2003). Primordial germ-cell development: the zebrafish perspective. *Nat. Rev. Genet.* 4, 690–700.
149. 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.*
150. Reuter, M., Berninger, P., Chuma, S., Shah, H., Hosokawa, M., Funaya, C., Antony, C., Sachidanandam, R., and Pillai, R.S. (2011). Miwi catalysis is required for piRNA amplification-independent LINE1 transposon silencing. *Nature.*
151. Riemer, S., Bontems, F., Krishnakumar, P., Gömann, J., and Dosch, R. (2015). A functional Bucky ball-GFP transgene visualizes germ plasm in living zebrafish. *Gene Expr. Patterns* 18, 44–52.
152. Rosenkranz, D., and Zischler, H. (2012). proTRAC - a software for probabilistic piRNA cluster detection, visualization and analysis. *BMC Bioinformatics.*
153. Saito, K., Sakaguchi, Y., Suzuki, T., Suzuki, T., Siomi, H., and Siomi, M.C. (2007). Pimet, the *Drosophila* homolog of HEN1, mediates 2'-O-methylation of Piwi-interacting RNAs at their 3' ends. *Genes Dev.* 21, 1603–1608.
154. Saitou, M., Barton, S.C., and Surani, M.A. (2002). A molecular programme for the specification of germ cell fate in mice. *Nature.*
155. Sassaman, D.M., Dombroski, B.A., Moran, J. V., Kimberland, M.L., Naas, T.P., DeBerardinis, R.J., Gabriel, A., Swergold, G.D., and Kazazian, H.H. (1997). Many human L1 elements are capable of retrotransposition. *Nat. Genet.*
156. Sato, K., Iwasaki, Y.W., Shibuya, A., Carninci, P., Tsuchizawa, Y., Ishizu, H., Siomi, M.C., and Siomi, H. (2015). Krimper Enforces an Antisense Bias on piRNA Pools by Binding AGO3 in the *Drosophila* Germline. *Mol. Cell.*
157. Saxe, J.P., Chen, M., Zhao, H., and Lin, H. (2013). Tdrkh is essential for spermatogenesis and participates in primary piRNA biogenesis in the germline. *EMBO J.* 32, 1869–1885.
158. Scherthan, H. (2001). A bouquet makes ends meet. *Nat. Rev. Mol. Cell Biol.*
159. Senti, K.-A.A., and Brennecke, J. (2010). The piRNA pathway: a fly's perspective on the guardian of the genome. *Trends Genet.* 26, 499–509.
160. Seydoux, G. (2018). The P Granules of *C. elegans*: A Genetic Model for the Study of RNA-Protein Condensates. *J. Mol. Biol.*
161. 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 Is Essential for piRNA-Mediated Retrotransposon Silencing in the Mouse Male Germline. *Dev. Cell* 17, 775–787.
162. Sigurdsson, B. (1954). Rida, A chronic encephalitis of Sheep With General Remarks on Infections Which Develop Slowly and Some of Their Special Characteristics. *Br. Vet. J.*
163. 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.
164. Slaidina, M., and Lehmann, R. (2017). Quantitative Differences in a Single Maternal Factor Determine Survival Probabilities among *Drosophila* Germ Cells. *Curr. Biol.* 27, 291–297.
165. Slotta, U., Hess, S., Spieß, K., Stromer, T., Serpell, L., and Scheibel, T. (2007). Spider silk and amyloid fibrils: A structural comparison. *Macromol. Biosci.*
166. De Smedt, V., Szöllösi, D., and Kloc, M. (2000). The Balbiani body: Asymmetry in the mammalian oocyte. *Genesis.*
167. Smit, A.F. (1999). Interspersed repeats and other mementos of transposable elements in mammalian genomes. *Curr. Opin. Genet. Dev.*
168. Smith, J., Calidas, D., Schmidt, H., Lu, T., Rasoloson, D., and Seydoux, G. (2016). Spatial patterning of P granules by RNA-induced phase separation of the intrinsically-disordered protein MEG-3. *Elife* 5.
169. 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.
170. Strome, S., and Wood, W.B. (1982). Immunofluorescence visualization of germline-specific cytoplasmic granules in embryos, larvae, and adults of *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. U. S. A.* 79, 1558–1562.
171. Su, T.T., Campbell, S.D., and O'Farrell, P.H. (1998). The cell cycle program in germ cells of the *Drosophila* embryo. *Dev. Biol.* 196, 160–170.
172. Takeda, Y., Mishima, Y., Fujiwara, T., Sakamoto, H., and Inoue, K. (2009). DAZL relieves miRNA-mediated repression of germline mRNAs by controlling poly(A) tail length in zebrafish. *PLoS One.*
173. 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.*
174. Le Thomas, A., Rogers, A.K., Webster, A., Marinov, G.K., Liao, S.E., Perkins, E.M., Hur, J.K., Aravin, A.A., and Toth, K.F. (2013). Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. *Genes Dev* 27, 390–399.
175. Trcek, T., Grosch, M., York, A., Shroff, H., Lionnet, T., and Lehmann, R. (2015). *Drosophila* germ granules are structured and contain homotypic mRNA clusters. *Nat. Commun.* 6, 7962.
176. Tzung, K.W., Goto, R., Saju, J.M., Sreenivasan, R., Saito, T., Arai, K., Yamaha, E., Hossain, M.S., Calvert, M.E.K., and Orbán, L. (2015). Early depletion of primordial germ cells in zebrafish promotes testis formation. *Stem Cell Reports* 4, 61–73.
177. Updike, D., and Strome, S. (2010). P granule assembly and function in *Caenorhabditis elegans* germ cells. In *Journal of Andrology*, pp. 53–60.
178. 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* (80-). 313, 320–324.
179. Vagin, V. V., Wohlschlegel, J., Qu, J., Jonsson, Z., Huang, X., Chuma, S., Girard, A., Sachidanandam, R., Hannon, G.J., and Aravin, A.A. (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.
180. 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.
181. Vourekas, A., Zheng, Q., Alexiou, P., Maragkakis, M., Kirino, Y., Gregory, B.D., and Mourelatos, Z. (2012). Mili and Miwi target RNA repertoire reveals piRNA biogenesis and function of Miwi in spermiogenesis. *Nat. Struct. Mol. Biol.*
182. Wan, G., Fields, B.D., Spracklin, G., Shukla, A., Phillips, C.M., and Kennedy, S. (2018). Spatiotemporal regulation of liquid-like condensates in epigenetic inheritance. *Nature.*
183. Wang, H., Ma, Z., Niu, K., Xiao, Y., Wu, X., Pan, C., Zhao, Y., Wang, K., Zhang, Y., and Liu, N. (2016). Antagonistic roles of Nibbler and Hen1 in modulating piRNA 3' ends in *Drosophila*. *Development* 143, 530–539.
184. Wang, J.T., Smith, J., Chen, B.C., Schmidt, H., Rasoloson, D., Paix, A., Lambrus, B.G., Calidas, D., Betzig, E., and Seydoux, G. (2014). Regulation of RNA granule dynamics by phosphorylation of serine-rich, intrinsically disordered proteins in *C. elegans*. *Elife* 3.
185. Wasik, K.A., Tam, O.H., Knott, S.R., Falcatori, I., Hammell, M., Vagin, V. V., and Hannon, G.J. (2015). RNF17 blocks

- 1 promiscuous activity of PIWI proteins in mouse testes. *Genes Dev.*
186. Watanabe, T., Totoki, Y., Toyoda, A., Kaneda, M., Kuramochi-Miyagawa, S., Obata, Y., Chiba, H., Kohara, Y., Kono, T., Nakano, T., et al. (2008). Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature.*
187. 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.*
188. Webster, A., Li, S., Hur, J.K., Wachsmuth, M., Bois, J.S., Perkins, E.M., Patel, D.J., and Aravin, A.A. (2015). Aub and Ago3 Are Recruited to Nuage through Two Mechanisms to Form a Ping-Pong Complex Assembled by Krimper. *Mol. Cell.*
189. Wegrzyn, R.D., Bapat, K., Newnam, G.P., Zink, A.D., and Chernoff, Y.O. (2001). Mechanism of prion loss after Hsp104 inactivation in yeast. *Mol. Cell. Biol.*
190. 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.
191. Whittle, C.A., and Extavour, C.G. (2016). Refuting the hypothesis that the acquisition of germ plasm accelerates animal evolution. *Nat. Commun.* *7*.
192. Whittle, C.A., and Extavour, C.G. (2017). Causes and evolutionary consequences of primordial germ-cell specification mode in metazoans. *Proc. Natl. Acad. Sci.* *114*, 5784–5791.
193. Wickner, R.B. (1994). [URE3] as an altered URE2 protein: Evidence for a prion analog in *Saccharomyces cerevisiae*. *Science* (80-.).
194. Williamson, A., and Lehmann, R. (1996). Germ cell development in *Drosophila*. *Annu. Rev. Cell Dev. Biol.* *12*, 365–391.
195. Woodruff, J.B., Ferreira Gomes, B., Widlund, P.O., Mahamid, J., Honigsmann, A., and Hyman, A.A. (2017). The Centrosome Is a Selective Condensate that Nucleates Microtubules by Concentrating Tubulin. *Cell* *169*, 1066–1077.e10.
196. Xu, D., Shen, W., Guo, R., Xue, Y., Peng, W., Sima, J., Yang, J., Sharov, A., Srikantan, S., Yang, J., et al. (2013). Top3 β is an RNA topoisomerase that works with fragile X syndrome protein to promote synapse formation. *Nat. Neurosci.*
197. Yamanaka, S., Siomi, M.C., and Siomi, H. (2014). PiRNA clusters and open chromatin structure. *Mob. DNA.*
198. Yang, Y., McBride, K.M., Hensley, S., Lu, Y., Chedin, F., and Bedford, M.T. (2014). Arginine Methylation Facilitates the Recruitment of TOP3B to Chromatin to Prevent R Loop Accumulation. *Mol. Cell.*
199. Ying, Y., Liu, X.M., Marble, a, Lawson, K. a, and Zhao, G.Q. (2000). Requirement of Bmp8b for the Generation of Primordial Germ Cells in the Mouse. *Mol. Endocrinol.* *14*, 1053–1063.
200. Ying, Y., Qi, X., and Zhao, G.-Q. (2001). Induction of primordial germ cells from murine epiblasts by synergistic action of BMP4 and BMP8B signaling pathways. *Proc. Natl. Acad. Sci.* *98*, 7858–7862.
201. Yisraeli, J.K., Sokol, S., and Melton, D. a (1990). A two-step model for the localization of maternal mRNA in *Xenopus* oocytes: involvement of microtubules and microfilaments in the translocation and anchoring of Vg1 mRNA. *Development* *108*, 289–298.
202. 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.
203. Zhang, F., Wang, J., Xu, J., Zhang, Z., Koppetsch, B.S., Schultz, N., Vreven, T., Meignin, C., Davis, I., Zamore, P.D., et al. (2012). UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery. *Cell* *151*, 871–884.
204. Zhang, H., Elbaum-Garfinkle, S., Langdon, E.M., Taylor, N., Occhipinti, P., Bridges, A.A., Brangwynne, C.P., and Gladfelter, A.S. (2015). RNA Controls PolyQ Protein Phase Transitions. *Mol. Cell.*
205. Zhang, H., Liu, K., Izumi, N., Huang, H., Ding, D., Ni, Z., Sidhu, S.S., Chen, C., Tomari, Y., and Min, J. (2017a). Structural basis for arginine methylation-independent recognition of PIWIL1 by TDRD2. *Proc. Natl. Acad. Sci.*
206. Zhang, Y., Guo, R., Cui, Y., Zhu, Z., Zhang, Y., Wu, H., Zheng, B., Yue, Q., Bai, S., Zeng, W., et al. (2017b). An essential role for PNLDC1 in piRNA 3' end trimming and male fertility in mice. *Cell Res.* *27*, 1392–1396.
207. Zhang, Z., Xu, J., Koppetsch, B.S., Wang, J., Tipping, C., Ma, S., Weng, Z., Theurkauf, W.E., and Zamore, P.D. (2011). Heterotypic piRNA Ping-Pong Requires Qin, a Protein with Both E3 Ligase and Tudor Domains. *Mol. Cell.*
208. Zhou, Y., and King, M.L. (1996). Localization of Xcat-2 RNA, a putative germ plasm component, to the mitochondrial cloud in *Xenopus* stage I oocytes. *Development.*
209. Zhou, J., Leu, N.A., Eckardt, S., McLaughlin, K.J., and Wang, P.J. (2014). STK31/TDRD8, a germ cell-specific factor, is dispensable for reproduction in mice. *PLoS One.*
210. Zhou, R.R., Wang, B., Wang, J., Schatten, H., and Zhang, Y.Z. (2010). Is the mitochondrial cloud the selection machinery for preferentially transmitting wild-type mtDNA between generations? Rewinding müller's ratchet efficiently. *Curr. Genet.*

Chapter

Zebrafish re-shape piRNA populations
in the absence of TdrKH

2

ABSTRACT

In animals, piRNAs represent an essential class of small RNAs that PIWI proteins use to target transposons in order to silence them and protect the germline genome from deleterious events. Defects in the piRNA-pathway commonly result in fertility problems, which stresses the importance of their appropriate production. Over the last years, much has been discovered regarding their biogenesis. One aspect is trimming of 3' ends since many piRNAs are initially too long for PIWI proteins to exert their function. TdrKH is a conserved Tudor domain-containing protein that is required for this process. It recruits the exonuclease PNLDC1 to PIWI proteins bound to pre-piRNAs, in order to trim them to their mature, PIWI-compatible length. The zebrafish does not encode a direct PNLDC1 homolog, but it does contain a gene encoding TdrKH. We found that, as in mice and flies, TdrKH protein is associated with mitochondria. Loss of TdrKH results in fertility defects, but this defect is rescued in about half of the population. We find that in these fertile *tdrkh* mutant animals piRNAs are lower in number, and are produced from fewer sites. In addition, they are longer at their 3' ends, suggesting a 3' trimming defect. Finally, *tdrkh* mutant piRNAs display a stronger ping-pong signal compared to wild-type piRNAs, indicating a relative increase in PIWI-induced 5' end generation. This study reveals a role for TdrKH in the zebrafish piRNA pathway, in particular in 3' end trimming, and highlights that a compromised piRNA defense mechanism can still be compatible with effective germ cell production.

INTRODUCTION

Transposable elements (TEs) are genomic parasites that can cause mutations through transposition events in the genome. Animal germ cells are protected from detrimental effects by TEs by a small RNA pathway called the piRNA pathway. Central to this pathway are PIWI proteins that are guided by PIWI-interacting RNAs (piRNAs) to target and cleave transcripts of active TEs and in some species, silence them transcriptionally as well. The importance of a functional piRNA pathway is illustrated by the fact that defects generally result in sterility (Brennecke et al., 2007; Carmell et al., 2007; Houwing et al., 2007, 2008; Kuramochi-Miyagawa et al., 2004).

During the primary piRNA pathway, transcription of so-called piRNA clusters results in the production of long precursor transcripts, that are processed further into smaller precursor piRNAs (pre-piRNAs) by the conserved endonuclease Zucchini (Zuc, a.k.a. Pld6) (Brennecke et al., 2007; Ipsaro et al., 2012). Zuc is expected to cleave before Uracils *in vivo*, however the incentive for this specificity remains unknown (Han et al., 2015). The cleavage products, thus starting predominantly with U and a 5' monophosphate, are loaded into PIWI proteins. Many of these piRNA intermediates (pre-piRNAs) undergo 3'-to-5' trimming down to their mature length between ~23-30 nucleotides (nt). Lastly, 3' ends are methylated by Hen1 for stabilization, upon which they can find and cleave a target (Kamminga et al., 2010; Saito et al., 2007). Hence, in this pathway, the 5' ends of the piRNAs are defined by the endonuclease Zuc, while the 3' ends are determined by an exonucleolytic activity.

The secondary piRNA pathway, called the ping-pong cycle, results from the interplay between two PIWI paralogs that target and cleave transcripts in an alternating fashion. This way, target cleavage by one PIWI results in a newly exposed 5'-monophosphate end that can be loaded into another PIWI protein. In other words, the 5' ends of these pre-piRNAs are not set by Zuc, but by a PIWI protein. These pre-piRNAs undergo further maturation steps similarly to the primary pathway, including 3' end trimming and methylation. piRNAs derived from opposite strands through this pathway thus typically bear a 10 nt overlap with each other, referred to as the 'ping-pong signature' (Brennecke et al., 2007; Gunawardane et al., 2007). Since the resulting mature PIWI-piRNA complexes can cleave further complementary targets again, the secondary pathway ensures that the silencing signal of active elements, which provide many targets, will be amplified.

These two main pathways are not functioning independently of each other, but rather cooperatively. For instance, the primary pathway requires the secondary pathway to provide a 'trigger' event: endonucleolytic cleavage of a target by a PIWI protein (Gainetdinov et al., 2018; Han et al., 2015; Mohn et al., 2015). This trigger initiates downstream Zuc cleavages, which produce a series of piRNAs from one transcript (Han et al., 2015; Mohn et al., 2015). Such a series of piRNAs is referred to as 'phased'. In absence of 3' end trimming, phased piRNAs are characterized by directly adjacent 3' and 5' ends.

It has recently been demonstrated that the conserved PARN-like 3'-to-5' exonuclease PNLDC1 or 'Trimmer' is responsible for pre-piRNA 3' end trimming in mice and the silkworm

Elke F. Roovers^{1,*}, António M. de Jesus Domingues^{1,*}, Stefan Redl¹, Hanna Braun², Falk Butter², René F. Ketting¹

¹ Biology of Non-coding RNA Group, Institute of Molecular Biology, Ackermannweg 4, 55128, Mainz, Germany

² Quantitative Proteomics Group, Institute of Molecular Biology, Ackermannweg 4, 55128, Mainz, Germany

*These authors contributed equally

Bombyx mori (Ding et al., 2017; Izumi et al., 2016; Zhang et al., 2017). In *Drosophila*, PARN-like nucleases are absent and at least one other 3'-to-5' exonuclease, Nibbler (MUT-7/Exd3), can trim piRNA 3' ends as well, in addition to its role in miRNA biogenesis (Hayashi et al., 2016; Wang et al., 2016). Nibbler acts predominantly on Ago3-bound piRNAs derived by PIWI slicing in the ping-pong cycle, whereas Zuc-derived piRNAs remain unaffected by mutations in Nibbler (Hayashi et al., 2016). Interestingly, Nibbler is absent in mouse, whereas silkworm encodes both Nibbler and PNLDC1 (Hayashi et al., 2016).

In order to act on the pre-piRNA, Trimmer has to be recruited to the PIWI-pre-piRNA complex. An essential protein for this step of piRNA maturation is the Tudor domain-containing protein TdrKH (a.k.a. Tdr2 or *Partner of Piwis*, PAPI) (Honda et al., 2013; Saxe et al., 2013). Tudor domains are protein-protein interaction domains that recognize symmetrically dimethylated arginine (sDMA) residues that are commonly present on the N-terminal domain of PIWI proteins (Huang et al., 2011; Kirino et al., 2010; Nishida et al., 2009; Vagin et al., 2009). In addition to its Tudor domain, TdrKH contains two KH domains. First identified in human hnRNP K, the K-Homology (KH) domain is widespread, has affinity for single-stranded nucleic acids and is required by TdrKH to interact with RNA (Nishida et al., 2018; Siomi et al., 1993; Valverde et al., 2008). TdrKH is a mitochondrion-associated protein that interacts with PIWI proteins in mice (Miwi and Miwi2), flies (Ago3, Aub and Piwi) and silkworm (Siwi) (Izumi et al., 2016; Liu et al., 2011; Saxe et al., 2013). In the absence of TdrKH, piRNAs are longer at their 3' ends while they are still methylated by Hen1 (Honda et al., 2013; Izumi et al., 2016; Saxe et al., 2013). How much longer piRNAs are without TdrKH/PAPI differs per species, varying from 1-2 nt in *Drosophila* and silkworm, to 3-10 nt in mouse, suggesting that different pre-piRNA lengths and/or alternative trimming activities exist in parallel (Han et al., 2015; Izumi et al., 2016; Saxe et al., 2013). It is thought that TdrKH and Trimmer/PNLDC1 predominantly act on primary pathway products of Zuc. However, in *Drosophila*, piRNA trimming is also affected in *papi* mutants, even though it does not have a PNLDC1 homolog, indicating it can either recruit Nibbler or a different 3'-to-5' exonuclease as well, or this reflects another unknown function of TdrKH (Han et al., 2015; Hayashi et al., 2016).

Zebrafish encode two PIWI proteins, Ziwi and Zili, of which Ziwi predominantly binds antisense piRNAs of ~29nt while Zili prefers sense, ~27nt long piRNAs (Houwing et al., 2007, 2008). Furthermore, in addition to copies of Zuc and TdrKH, zebrafish also encode Nibbler (Exd3) and PARN-like exonucleases, but it has lost PNLDC1. Given this diversity between species, we isolated zebrafish *tdrkh* mutant alleles in order to investigate its effect on piRNA biogenesis and fertility. *Tdrkh* mutant fish mainly develop into males that are fertile in about 50% of the cases. We found that piRNAs are generally longer than in wild-type and display an increase of phasing signatures, indicating trimming is affected. Interestingly, ping-pong signals increase, suggesting adaptation from Zuc-produced towards PIWI-induced 5' end formation. This potentially reflects a switch towards TdrKH-independent piRNA biogenesis. These results demonstrate that the piRNA pathway is extremely flexible and can quickly, within one generation, shift between alternative mechanisms for piRNA biogenesis, in order to ensure piRNA functionality and fertility.

RESULTS

TdrKH is required for normal fertility in zebrafish

We used CRISPR-Cas9 mutagenesis to target the *tdrkh* locus in the first KH domain in order to create small indels in the coding region (Figure 1A). We obtained 2 alleles; one carrying a deletion of 8 nt and one carrying a deletion of 25 nt (hereafter called $\Delta 8$ and $\Delta 25$, respectively) (Figure 1A). F0 founders were outcrossed and heterozygous progeny was selected via sequencing. These were outcrossed again and F2 fish were incrossed to obtain *tdrkh* homozygous mutants in the F3. There was a strong tendency for mutant animals to develop as males. Occasionally (less than 1 in 20), *tdrkh* mutants did develop as females, which is uncommon for piRNA pathway mutants. Strong mutants such as *ziwi* and *zili* mutant animals develop exclusively as sterile males. Other, milder mutants like *hen1* develop as males that are always fertile, yet mutants never develop as females (Houwing et al., 2007, 2008; Kamminga et al., 2010).

We next asked whether *tdrkh* mutants showed fertility defects. Interestingly, about 50% of the *tdrkh* mutant males were infertile, whereas the other half was fertile, like their heterozygous siblings (Figure 1B). The few identified mutant females were fertile as well (data not shown). In fertile mutants, Ziwi and Zili were both present at high levels, whereas in infertile mutants, almost no Ziwi and Zili was detected. Interestingly, the relative amounts of Zili mildly increased in the fertile mutants and was found to be ~1.2 times the amount that was present in heterozygous siblings when normalized to Actin (Figure 1C). All these observations were made in both alleles, including heteroallelic $\Delta 8/\Delta 25$ fish, indicating no allele-specific effects.

Zebrafish TdrKH localizes to mitochondria and interacts with Ziwi

We made a transgene producing a TdrKH-GFP fusion protein driven by a Ziwi promoter. This transgene is expressed in juvenile ovaries up to stage I oocytes (Figure 2A). TdrKH has a threadlike arrangement in stage I oocytes (Figure 2A, arrows) and more aggregated in clusters during earlier stages, mirroring a typical mitochondrial localization pattern (Figure 2A, arrowheads) (Zhang et al., 2008).

Next, we performed immunoprecipitation on the TdrKH-GFP fusion protein, followed by label-free quantitative mass-spectrometry (IP/LFQ-MS). TdrKH was significantly enriched in these experiments, and was found to interact with Ziwi (Figure 2B). We could confirm TdrKH interaction with Ziwi on Western blot (Figure 2C). Zili was neither identified in our MS experiment, nor detected on Western blot.

Zebrafish lacking TdrKH have longer piRNAs with stronger ping-pong signal

In mice and insects, *tdrkh* mutations led to longer piRNAs, since it is required to recruit PNLDC1/Trimmer (Han et al., 2015; Honda et al., 2013; Saxe et al., 2013). Since zebrafish do not encode PNLDC1/Trimmer, we wondered if TdrKH still has an effect on piRNA length. We therefore sequenced small RNAs derived from *tdrkh*^{+/-} and *tdrkh*^{-/-} testes and ovaries from fertile animals and investigated whether piRNA lengths were affected. We observed that piRNAs were

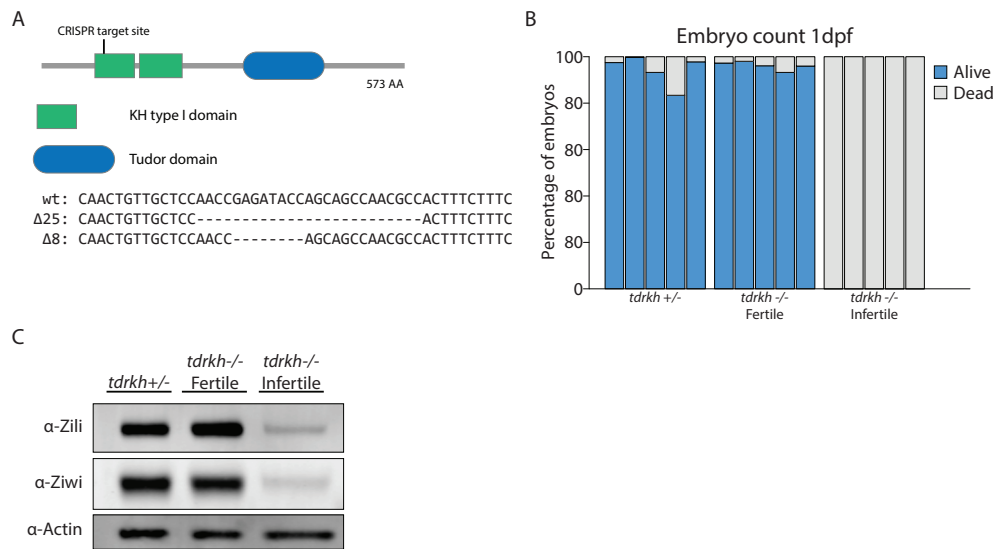


Figure 1. TdrKH is required for normal fertility and sex ratios in zebrafish. (A) Schematic overview of TdrKH. The first KH domain was targeted for mutagenesis. (B) Barplot indicating percentages of living and dead offspring of different *tdrkh* individuals. About 50% of *tdrkh* mutants is infertile, regardless of the allele. The average of three eggclays was taken per individual and embryo viability was assessed 1 dpf. (C) Western blot of heterozygous and mutant testes (fertile and infertile) probed against Ziwi and Zili. Fertile mutants have both Ziwi and Zili at high amounts, whereas infertile mutants lack most of the protein. When normalized to Actin, the proportion of Ziwi and Zili shifted in fertile mutants compared to heterozygous at a ratio of 1 : 1.2 : 0.1 for Zili, whereas Ziwi was at 1 : 0.9 : 0.5 (*tdrkh*^{+/-} : *tdrkh*^{-/-} fertile : *tdrkh*^{-/-} infertile). Testes of animals were taken out after mating, to reduce differences caused by staging.

indeed longer: ~2 nt in ovary and ~3 nt in testis (Figure 3A). In addition, the piRNA length profile peaks for the mutant are less defined compared to those from heterozygous animals, illustrating an increase in length variability. In addition, particularly in ovary, a shoulder of somewhat shorter piRNAs, ~27 nt, is observed in the mutant piRNA profile (Figure 3A, arrows), suggesting that two distinct piRNA populations may be present in the mutant. The transposon-targeting repertoire and 5'U bias of the piRNAs did not change significantly, although LINE elements appeared to be stronger affected than others, in both ovary and testis (Figure 3B and 3C). Interestingly, the ping-pong signatures have stronger Z-scores in *tdrkh* mutants than in the heterozygous situation, indicating that the ping-pong cycle is more prominent in *tdrkh* mutants (Figure 3D).

Ziwi- and Zili-bound piRNAs are differentially affected in *tdrkh* mutants

We next addressed whether the overall increase in piRNA length described above represented a change in the Ziwi- or Zili-bound piRNA repertoire specifically. We therefore performed Ziwi and Zili RNA immunoprecipitations (RIPs) in *tdrkh*^{+/-} and *tdrkh*^{-/-} ovary and testis, followed by

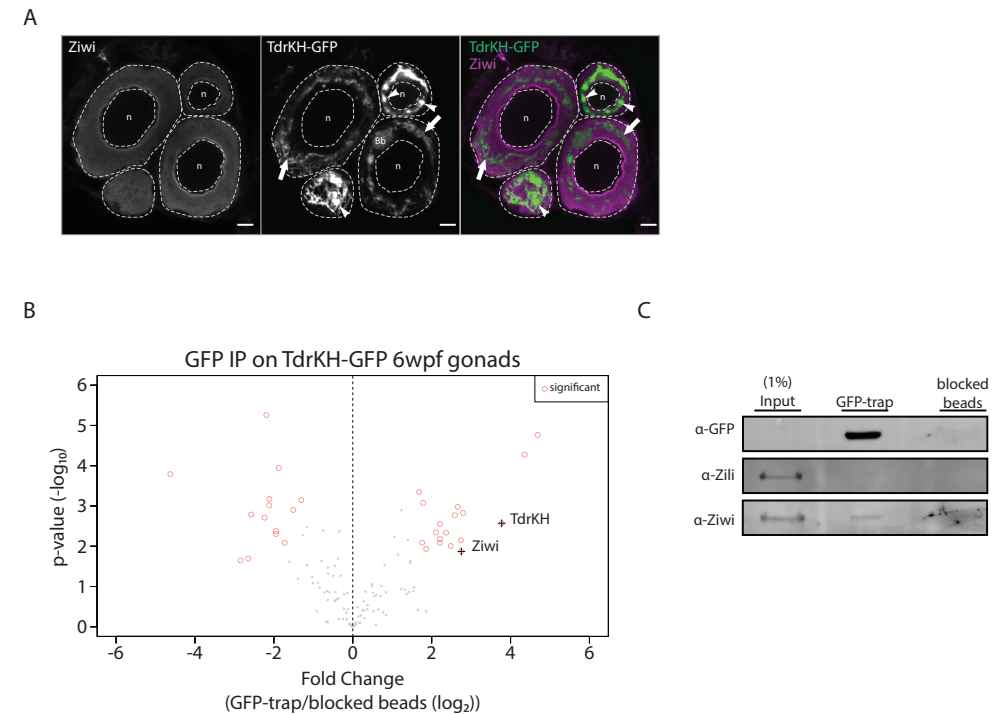


Figure 2. Zebrafish TdrKH interacts with Ziwi and Zili. (A) Confocal image of IHC against Ziwi in 6wpf juvenile ovaries positive for TdrKH-GFP. Ziwi shows perinuclear nuage localization, whereas TdrKH displays typical mitochondrial localization and localizes from more punctate (arrowhead) to more thread-like (arrow) in later stage I oocytes. n = nucleus. Scalebar indicates 10 μ m. (B) Volcanoplot of LFQ-MS on pull-downs of GFP nanobody-coupled beads compared to blocked beads. TdrKH itself is strongly enriched, as well as Ziwi, whereas Zili was not identified in the MS. (C) Western blot confirmation of GFP pull-down of TdrKH, which interacts with Ziwi.

sequencing. We confirmed that in *tdrkh*^{+/-}, Ziwi predominantly interacts with antisense piRNAs with an average length of ~29 nt, whereas piRNA lengths from its ping-pong partner Zili peak at ~27 nt and are mostly sense (Figure 4A). In *tdrkh*^{-/-} gonads, however, Ziwi IPs contained piRNAs that appeared to derive from two sub-populations, one with a peak at ~31 nt and one with a peak at ~27 nt (Figure 4A, arrows). This mirrors our observations in the total RNA samples (Figure 3). On the other hand, we found that the length distribution of Zili-interacting piRNAs still displays a single, but rather broad peak in *tdrkh*^{-/-}, with a clear bias towards longer piRNA species (Figure 4B). Especially in the ovary, the antisense reads as found in the Zili RIPs resemble the Ziwi-bound profile. We therefore speculate that, since Ziwi and Zili interact, these reads likely come from Ziwi contamination in the Zili RIP (Houwing et al., 2008). We conclude that in zebrafish lacking TdrKH both Ziwi- and Zili-bound piRNAs are longer than in wild-type, and that specifically for Ziwi, two piRNA sub-populations are apparent: one shorter than in heterozygous animals, and one longer.

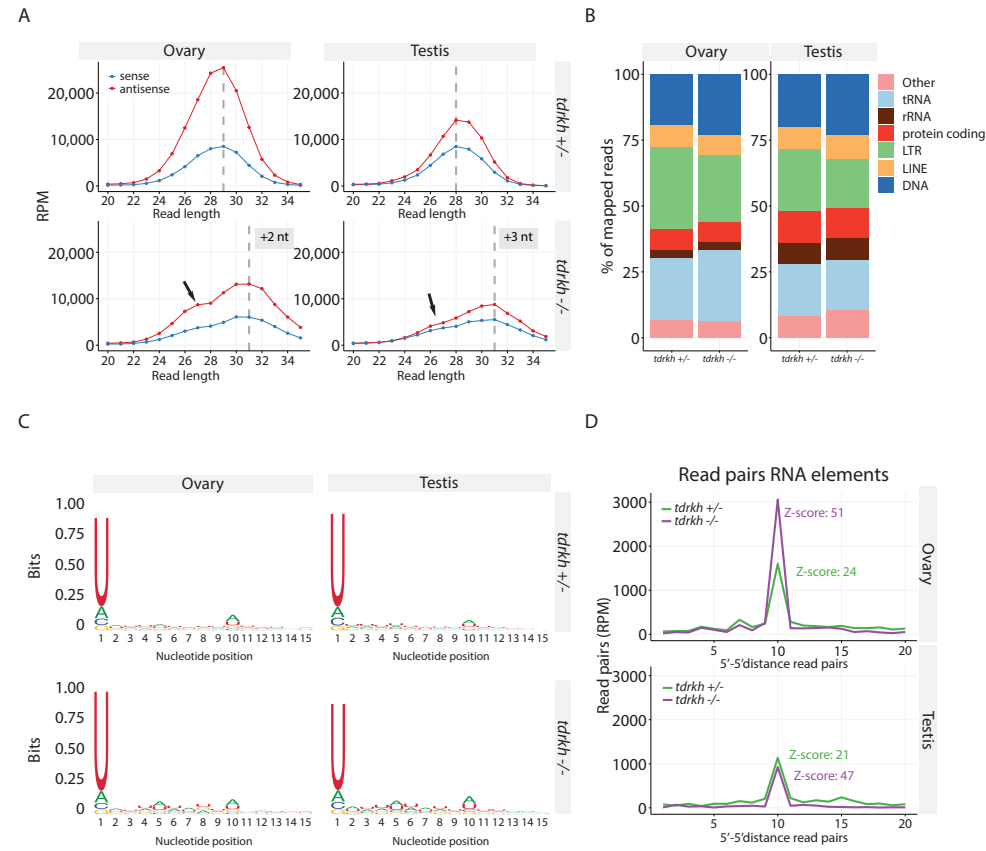


Figure 3. *Tdrkh* mutants display longer piRNAs and stronger ping-pong. (A) Length profiles for piRNAs mapped against RNA elements from *tdrkh*^{+/-} and *tdrkh*^{-/-} ovary and testis. Vertical dashed lines indicate peak maxima. Arrow indicated second maximum in *tdrkh*^{-/-} ovary and testis. (B) Transposon targeting repertoire of *tdrkh*^{+/-} and *tdrkh*^{-/-} piRNAs of ovary and testis do not display dramatic changes. (C) Sequence logos of piRNAs from *tdrkh*^{+/-} and *tdrkh*^{-/-} ovary and testis show a strong 5'U bias, characteristic for piRNAs. (D) Ping-pong signatures from *tdrkh*^{+/-} and *tdrkh*^{-/-} ovary and testis. Mutant samples show a higher Z-score compared to heterozygous samples, indicative of an increase in read pairs that overlap exactly 10 nucleotides in the absence of *tdrkh*.

TdrKH is required for 3' end trimming in zebrafish

In mouse, piRNAs from wild-type and both *tdrkh* and *pnlcd1* mutants could be matched based on their 5' end position, revealing that the increase in length resulted from a lack of 3' end trimming (Ding et al., 2018; Saxe et al., 2013). The finding that the first nucleotide of the piRNA directly downstream of a non-trimmed mouse piRNA (the +1 position) has a strong bias for U in *tdrkh* mutants, led to the idea that TdrKH/trimming follows Zuc-mediated processing and that Zuc produces phased piRNAs (Han et al., 2015; Mohn et al., 2015). Again, given the absence of zebrafish PNLDC1/Trimmer, we looked at these features in our datasets as well. First, a clear +1U bias is already detectable in wild-type piRNAs, indicating that a significant fraction of

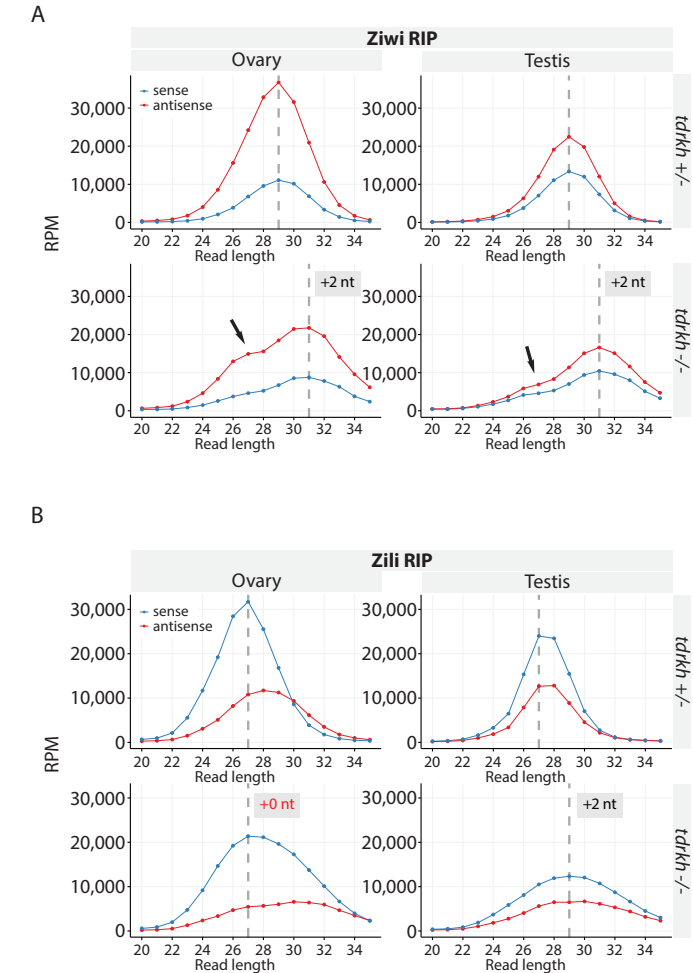


Figure 4. Ziwi- and Zili-bound piRNAs are affected to different extents without TdrKH. (A) Length profiles for Ziwi-bound piRNAs mapped against RNA elements from *tdrkh*^{+/-} and *tdrkh*^{-/-} ovary and testis. A typical antisense bias is seen, also in *tdrkh*^{-/-} even though mutants display an increase in length of ~2nt. The extra maximum is found in Ziwi-RIPs of *tdrkh*^{-/-} ovary just as in input samples, while a mild shoulder is observed in *tdrkh*^{-/-} testis (arrows). (B) Length profiles for Zili-bound piRNAs mapped against RNA elements from *tdrkh*^{+/-} and *tdrkh*^{-/-} ovary and testis, displaying a typical sense bias. While Zili-RIPs in *tdrkh*^{-/-} testis show an increase in length of ~2nt, Zili RIPs from *tdrkh*^{-/-} ovary peak at 27nt, similar to *tdrkh*^{+/-} even though the peaks are much wider.

piRNAs in wild-type animals are not trimmed. The +1U-bias increases mildly in *tdrkh* mutant gonads (Figure 5A). The effect is certainly not as strong as the effect in mouse, in which a +1U signal is completely absent in the wild-type and becomes very prominent in the mutant, but rather similar to the increase observed in the *papi* mutant in *Drosophila* (Han et al., 2015). The juxtaposition of the 5' and the 3' ends of two neighbouring piRNAs (3'-to-5' distance) on the same genomic strand also became more pronounced in zebrafish *tdrkh* mutants, confirming

that the fraction of non-trimmed Zuc products increases (Figure 5B). These results confirm that in zebrafish, phasing is detectable in wild-type animals (Han et al., 2015), and that it becomes more apparent in absence of TdrKH.

To address whether TdrKH affects 3' end trimming, we first identified those piRNA 5' ends that were cloned from both heterozygous and mutant gonads. When the length profiles of these piRNA sub-populations are plotted, it becomes apparent that in the *tdrkh* mutants, piRNAs are longer at their 3' ends (Figure 5C). When we isolated piRNA species with identical 5' ends from wild type and mutant datasets, we noticed that the majority of the 5' positions is in fact unique per genotype and that only a fraction is shared (Supplemental Figure 2). Nevertheless, if we divide all shared and unique 5' positions over the total amount of reads of each 5' position (regardless of their read length as long as they were between 24 and 35 nt) we find that most piRNAs are actually derived from the shared positions (Figure 5D). This implies that these positions represent robust piRNA sites, which still produce stable intermediates in the *tdrkh* mutant condition (Figure 5D). In addition, we found that overall, there were fewer 5' positions in the mutant from which all piRNAs were produced, compared to heterozygous. This indicates that 5' end positions are lost, and that in the *tdrkh* mutants, the piRNA repertoire may be less complex (Figure 5D).

DISCUSSION

Over the last years, much has been revealed about the function and biogenesis of piRNAs, particularly in *Drosophila*, mice and silkworm. Regardless of many similarities and conserved factors, much variation can be found between the different animals, for example to which extent they rely on ping-pong and/or phasing, but also which 3'-to-5' exonucleases are involved in piRNA trimming (Ding et al., 2017; Gainetdinov et al., 2018; Hayashi et al., 2016; Izumi et al., 2016; Zhang et al., 2017). The recent identification of piRNA 3' processing factors TdrKH and PNLDC1 prompted us to investigate the situation in zebrafish. We analyzed TdrKH function and investigated fertility and piRNA features in *tdrkh* mutant animals. Below, we further discuss our findings.

Zebrafish TdrKH is required for piRNA 3'end trimming

It has been demonstrated that TdrKH mediates recruitment of PNLDC1/Trimmer to the PIWI-pre-piRNA complex in order to process piRNAs at their 3' end (Ding et al., 2017; Izumi et al., 2016; Zhang et al., 2017). Without either TdrKH or PNLDC1, piRNAs retain additional nucleotides at their 3' ends specifically. This could hinder PIWI-protein function, since this may affect how well the piRNA can bind to the PIWI protein, but it will also influence the kinetics of target RNA recognition and release. We found that TdrKH is required for 3' end trimming in zebrafish, confirming previous findings, however, it remains unknown whether this depends on recruitment of PARN-like nucleases by TdrKH in zebrafish as well (Honda et al., 2013; Saxe et al., 2013).

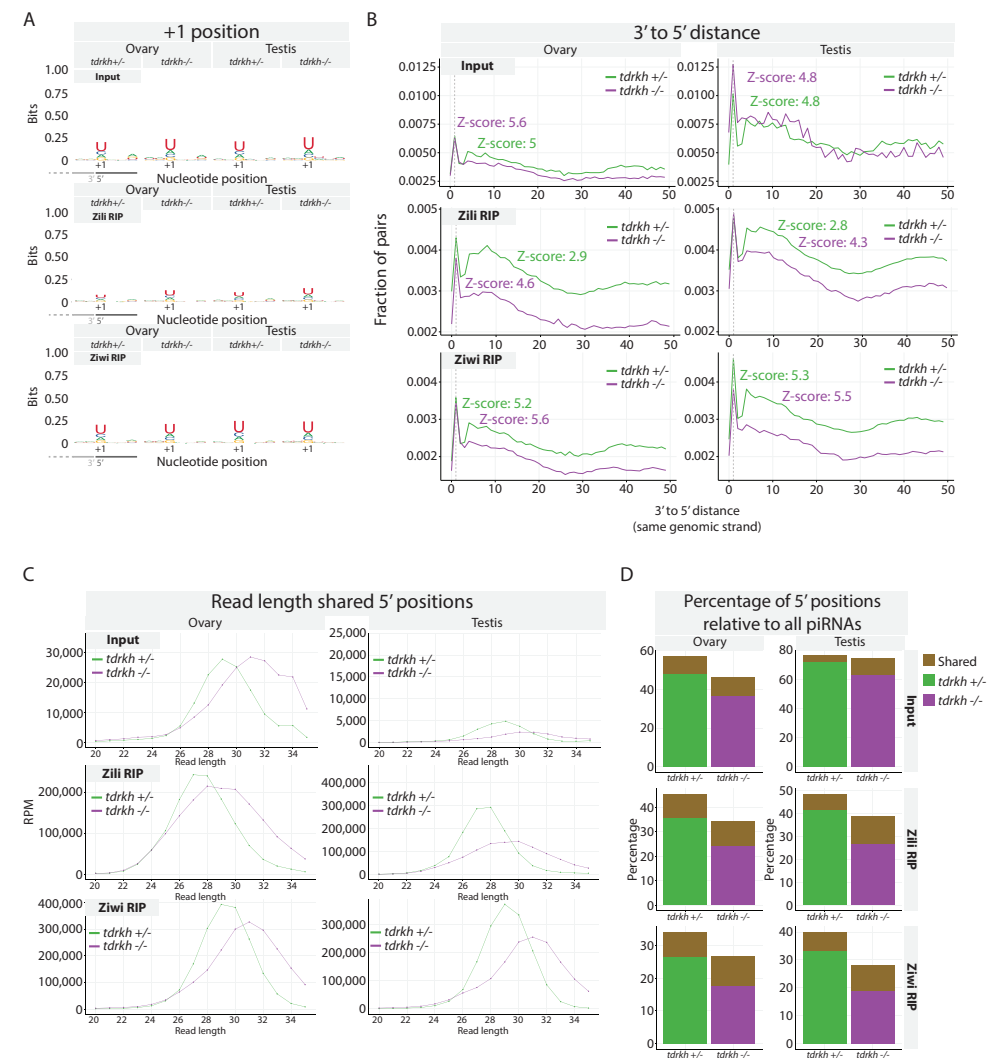


Figure 5. Dominant 5' positions shift to trimming-independent populations in *tdrkh* mutants. (A) Sequence logos displaying the nucleotide bias at the +1 position. There is a mild +1U-bias in heterozygous samples, which slightly increases in *tdrkh* mutants. (B) 3'-to-5' distance of adjacent piRNAs from the same genomic strand. There is a mild increase in Z-score in the mutant, indicating there is more phasing, due to loss of 3' end processing. (C) Length profiles of the shared 5' positions between *tdrkh* heterozygous and mutant. The piRNAs in the mutant with the same 5' end as heterozygous are longer. (D) Percentage of the individual 5' piRNA positions identified in both genotypes, of which some are shared between the genotypes (brown) and some are only found in heterozygous (green) or mutant (purple), divided by all piRNAs that share this 5' position. In heterozygous, the percentage of positions relative to the total number of piRNAs is consistently higher, suggesting a higher diversity compared to the mutant in which all piRNAs map to fewer positions.

It has been predicted that zebrafish pre-piRNAs, *i.e.* prior to trimming, are ~35 nt in length (Gainetdinov et al., 2018), indicating a requirement for 3'-to-5' exonuclease activity in order to trim ~6-8 nt. This was based on measuring 5'-5' end distance of neighbouring piRNA species from the same strand. Here, we only observe a shift of ~2 nt. How can this be reconciled? First, it could be that pre-piRNAs of ~35 nt are simply too long to be stable in the mutant, or that pre-piRNAs with a somewhat shorter length are more effectively used in the ping-pong cycle. Both could lead to a severe under-representation of the predicted 35 nt pre-piRNAs, and may also explain the loss of piRNA species, as defined by their 5' ends, in *tdrkh* mutants. Second, other exonucleases could act on the pre-piRNAs independently of TdrKH. In addition to PARN-like exonucleases, zebrafish have a copy of Nibbler/Exd3, which, in *Drosophila*, functions in the 3'-to-5' trimming of piRNAs that are derived from the ping-pong cycle. It is thought that TdrKH recruits PNLDC1/PARN, but it is unclear whether TdrKH can be responsible for Nibbler recruitment as well. *Drosophila* piRNAs are longer in *papi* mutants, even though PARN exonucleases are absent (Han et al., 2015; Hayashi et al., 2016). In *Drosophila*, *papi* therefore plays role in facilitating trimming as well, although it remains unknown whether this is caused by assisting trimming by Nibbler, or through an unknown function of TdrKH.

Tdrkh mutants adapt piRNA populations

In mouse, loss of *tdrkh* results in longer piRNAs, resembling Zuc products, which require trimming in order to become compatible with PIWI function. Furthermore, the phasing signal in the piRNA population increases without TdrKH (Han et al., 2015; Mohn et al., 2015). In *Drosophila*, Zuc-dependent piRNAs require only minimal trimming and only a slight increase in phasing was demonstrated in *papi* mutants (Han et al., 2015). This variation suggests that different species rely on *tdrkh* to different extents and likely have distinct compensation mechanisms. We analyzed how piRNAs changed in *tdrkh* mutants in zebrafish. First of all, phasing was stronger in *tdrkh* mutants, indicating that piRNA processing upon Zuc-dependent cleavage is affected. When we looked at piRNAs with a matching 5' position in heterozygous compared to mutant, we found that these common piRNAs were indeed longer, implying that the increase in length resulted from a reduction in trimming of specific piRNAs. We also identified fewer 5' positions in the mutant. This could be explained by the fact that Zuc-dependent processing results in a variety of pre-piRNA lengths. This way, certain 5' positions could provide Zuc products that are too long in order to be stabilized by a PIWI protein and will therefore be lost.

Interestingly, piRNA populations in both ovary and testis have an increased ping-pong signature, indicating piRNA production in *tdrkh* mutants relies relatively more on the ping-pong pathway. In other words, PIWI-generated 5' ends become more prominent compared to Zuc-generated 5' ends. This is paralleled by the fact that in the *tdrkh* mutants Zili protein levels are mildly increased, suggesting that Zili partakes more in piRNA production. In *Drosophila*, Nibbler (Exd3 in zebrafish) acts predominantly on the Zili-homolog Ago3. It could therefore be that Exd3 in zebrafish mostly trims Zili-bound piRNAs. If this would be TdrKH-independent, it could explain the relative stabilization of Zili-piRNA complexes, and a more substantial role for

the ping-pong cycle in *tdrkh* mutants. These ideas will need to be tested further experimentally, by creating *exd3* mutant zebrafish.

Since the ping-pong signal becomes more pronounced in *tdrkh* mutants, it seems that piRNA biogenesis shifts from Zuc-dependent biogenesis, which relies more on TdrKH in other species, to Zuc-independent. Furthermore, the relative decline in Zuc-dependent piRNA-production could also explain why the increase of the phasing score is so mild in *tdrkh* mutants. Lastly, a prominent shoulder in the Zili-bound piRNA length distribution suggests the existence of a distinct Zili-bound piRNA population in *tdrkh* mutants. It will be interesting to see whether these piRNAs represent a TdrKH-independent/PIWI-slicing dependent population that is stabilized preferably and perhaps explains the increase in ping-pong signal in the *tdrkh* mutant. For this, however, deeper sequencing of the RIP-input material will be necessary.

TdrKH is required for normal sex development and fertility in zebrafish

In *tdrkh* mutant zebrafish there was a strong bias towards male development. In zebrafish, certain amounts of germ cells are required to maintain an ovary and drive female development (Slanchev et al., 2005). When zebrafish proceed male development instead, oocytes in the juvenile ovary degenerate and undergo apoptosis (Uchida et al., 2002). When the piRNA pathway is impaired, germ cell apoptosis increases strongly (Houwing et al., 2007), possibly due to an increase in transposon activity. Consequently, in null mutants for *ziwi*, *zili* and *tdrd1*, all germ cells are lost (Houwing et al., 2007, 2008; Huang et al., 2011). Since the piRNA pathway in *tdrkh* mutants is compromised, it is likely that many, but not all germ cells are lost, and that consequently most animals will end up forming a testis and that some even develop into fertile males. In relatively rare cases, enough germ cells may be preserved such that even female development can be supported. Whether this depends on the stochastic inheritance of a relatively large population of maternal piRNAs, or to a more effective switch to TdrKH-independent piRNA biogenesis is currently not clear, and will require the analysis of piRNA populations of many individual animals at earlier stages of development.

MATERIALS AND METHODS

Zebrafish

Zebrafish strains were housed at the Institute of Molecular Biology in Mainz and bred and maintained under standard conditions (26-28°C room and water temperature and lighting conditions in cycles of 14:10 hours light:dark) as described by (Westerfield, 1995). Larvae < 5 days post fertilization were kept in E3 medium (5 mM NaCl, 0.17 mM KCl, 0.33 mM CaCl₂, 0.33 mM MgSO₄) at 28°C. CRISPR mutants and the transgenic line were obtained as described below. All experiments were conducted according to the European animal welfare law and approved and licensed by the ministry of Rhineland-Palatinate.

Obtaining CRISPR mutants

A CRISPR target site was found using ZiFiT targeter (<http://zifit.partners.org/ZiFiT/>) and corresponding CRISPR oligos (see below) were mixed, briefly heated to 95°C and cooled down, followed by ligation of the resulting duplex into pDR274 (Joung lab) that was pre-digested with BsaI. The resulting plasmid was then digested with DraI and transcribed with HiScribe T7 RNA polymerase (NEB). After cleaning up with the NucleoSpin RNA clean up kit (Macherey-Nagel), the obtained sgRNA was injected at a concentration of 20ng/μl together with 30ng/μl *cas9* mRNA in freshly laid zebrafish embryos. Finclips were analyzed and F0 fish containing mutations, confirmed by sequencing with *tdrkh_seq_fwd* and *tdrkh_seq_rev*, were outcrossed to obtain heterozygous carriers of *tdrkhΔ8* and *tdrkhΔ25* that were outcrossed again before they were crossed in to get a homozygous mutant line.

Transgenic line

The CDS of *tdrkh* was amplified using *ggggacaagttgtacaaaaagcaggctATGGATGTGG CAGTCCCTAAATCC* and *ggggaccactttgtacaagaaagctgggtgGTTGAACTCCTCATCTATAG TGTCATCCAC* and cloned into pDonr221 using Gateway BP Clonase II (ThermoFisher). Next, the *tdrkh* middle entry clone was recombined with the 5' entry clone containing the *ziwi* promoter and the 3' entry clone containing *gfp-pA* into the destination vector Tol2CG2, using Gateway LR Clonase II Plus (ThermoFisher) (Kwan et al., 2007; Leu and Draper, 2010). Embryos positive for *cmlc2:gfp* were raised in order to create a line. Gonads of positive 5wpf fish were dissected and lysed for IP/MS.

Immunoprecipitation

For adult Ziwi and Zili IPs, gonads were taken up in 700μL IP/lysis buffer (25 mM Tris pH 7.5, 150 mM NaCl, 1.5 mM MgCl₂, 1% Triton-X100, 1mM DTT, protease inhibitor), mashed and sonicated. Next, they were spun down at 12,000 x g at 4°C. The supernatant was collected and 1:500 Rb-α-Zili and 1:200 Rt-α-Ziwi was added and incubated with the samples for 1.5 hours at 4°C while rotating. 30μL washed Dynabeads were added and the samples were incubated for another 45 minutes at 4°C while rotating. This was followed by 3 washes with wash buffer (25 mM Tris pH 7.5, 300 mM NaCl, 1.5 mM MgCl₂, 1mM DTT) upon which the beads were taken up in TRIzol. RNA extractions were performed according to the TRIzol protocol.

For IPs on 5wpf juvenile ovaries, 5 dissected ovaries were pooled and taken up in 700μL IP/lysis buffer and were mashed and sonicated. After spinning at 12,000 x g at 4°C, the supernatant

Target site	GGCGTTGGCTGCTGGTATCT
<i>tdrkh_crispr_oligo_1</i>	TAGGCGTTGGCTGCTGGTATCT
<i>tdrkh_crispr_oligo_2</i>	AAACAGATACCAGCAGCCAACG
<i>tdrkh_seq_fwd</i>	GGAAGACAGTGAAATATCATGCACTTTG
<i>tdrkh_seq_rev</i>	GTCTCAGTGTTTTGTGATCTTTGTTCTC

was collected. 25μL washed slurry of GFP-trap beads (in house production) was added to each sample and the samples were incubated for 1 hour at 4°C while rotating. After 3 washes with wash buffer, the samples were taken up in NuPAGE LDS sample buffer (Thermo Fisher) for downstream MS analysis (see below).

Libraries

NGS library prep was performed with NEBNext's Small RNA Library Prep Kit for Illumina following instructions of manual, with a modification of the adaptors, for which custom made random barcodes for both 3' SR Adaptor and 5' SR Adaptor were used (HISS Diagnostics GmbH, 5' rApprrnnrnrnAGATCGGAAGAGCACACGTCT-NH2-3', and 5' rGrUrUrCrArGrArGrUrUrCrUrArCrArGrUrCrCrGrArCrGrArUrCrnnrnrn-3', respectively).

Libraries were size selected for the 135 to 175bp range using a LabChip GX (Caliper) in a DNA 300 chip. Size-selected libraries were profiled in a High Sensitivity DNA on a 2100 Bioanalyzer (Agilent technologies) and quantified using the Qubit dsDNA HS Assay Kit, in a Qubit 2.0 Fluorometer (Life technologies). All 12 samples were pooled in equimolar ratio and sequenced on 1 lane on a HiSeq 2000, SR for 1x 51 cycles plus 7 cycles for the index read.

Bioinformatic analyses

Raw reads were quality controlled with FastQC before adapter trimming with `cutadapt (v1.9.1) -a AGATCGGAAGAGCACACGTCT -O 8 -m 21 -M 51``, followed by removal of sequences with low-quality calls using `fastq_quality_filter (-q 20 -p 100 -Q 33)` from the FASTX-Toolkit (v0.0.14). PCR duplicates were removed making use of UMIs added during library preparation by collapsing reads with the same sequence, including UMIs, using a combination of unix command-line programs. Processed reads were mapped to the Zv9 *Danio rerio* genome assembly with `bowtie (v0.12.8) -q --sam --phred33-quals --tryhard --best --strata --chunkmbs 256 -v 0 -m 1 --trim5 4 --trim3 4``, that is, allowing only for uniquely mapping reads without mismatches. For subsequent analysis, and when specified, only reads mapping to RNA repeat elements (LINE, SINE, LTR) were used. Repeat element annotations were retrieved from the UCSC genome browser track RepeatMasker (Zv9).

Ping-pong signal was determined by calculating the base pair distance between the 5' of piRNAs overlapping in opposite strands (Brennecke et al., 2007). The Z-score was calculated as in Wasik et al., 2015, and is defined as $Z = (P10 - M) / SD$, where P10 is the number of reads pairs with 10 bp overlap, and M and S are the mean and standard deviation, respectively, of the number of read pairs at with 1-9 and 11-30 overlap (Wasik et al., 2015).

To determine the degree of phasing we used the 3'-5' piRNA distance (d) as described in (Han et al., 2015). Briefly, for each piRNA 3' position (k) the number of piRNAs (a) is compared to the number (b) of 5' piRNAs ends located at k + n, and the score for this piRNA (k) is $\min(a, b)$. This is done for n up to 50, and the final score for the 3'-5' distance (d) is the sum of the scores at n. The Z-score is calculated as described above, $Z = (P1 - M) / SD$, with the difference that now the background signal is calculated using position 2-50.

Nucleotide bias of piRNAs determined by summarizing the number of times a base is present in any given piRNA position. For the downstream U bias, the downstream annotated genomic bases, in the same mapping strand of the piRNA, were used (Han et al., 2015; Mohn et al., 2015).

Mass spectrometry

Immunoprecipitates were resuspended in 1x NuPAGE LDS Sample Buffer (Thermo) supplemented with 0.1 M DTT (Sigma) and heated for 10 minutes at 70 °C. The respective samples were run on a 4%-12% NuPAGE Bis-Tris gradient gel (Thermo) in 1x MOPS buffer (Thermo) at 180 V for 10 min. The samples were processed by in-gel digestion as essentially previously described (Shevchenko et al., 2007) and desalted using C18 StageTips (Rappsilber et al., 2007).

The digested peptides were separated by an EASY-nLC1000 uHPLC system with a C18 column (25 cm long, 75 µm inner diameter, packed in-house with ReproSil-Pur C18-AQ 1.9 µm (Dr. Maisch GmbH)) and sprayed via ESI online into a Q Exactive Plus mass spectrometer (Thermo). A 2h MS run consisting of a gradient from 2% to 40% acetonitrile and a washout of 95% acetonitrile at a flow rate of 200 nl/min was used. The instrument performed a top10 data-dependent acquisition with up to 10 HCD fragmentations per MS full scan (70k resolution, 300-1650 m/z).

Raw files were processed using MaxQuant v.1.5.2.8 (Cox and Mann, 2008) searching against a Uniprot/Trembl *Danio rerio* (www.uniprot.org) fasta file. Carbamidomethylation was set as fixed modification while methionine oxidation, protein N-acetylation, phosphorylation (STY) and lysine/arginine dimethylation were considered as variable modifications. The false discovery rate for peptide and protein identifications was set to 0.01. LFQ quantitation was performed with standard settings and the match between runs option activated.

The resulting proteome data was filtered for at least two peptide identifications per protein group, one unique and one razor. Protein Groups identified by site, known contaminants and reverse hits were disregarded. LFQ values were then log2 transformed. The mean LFQ intensities of each experimental condition (measured in quadruplicates) as well as the p-values (Welch t-test) comparing these conditions were calculated and used for the determination of enriched proteins.

AUTHOR CONTRIBUTIONS

EFR, AMJD and RFK designed the study. EFR and RFK wrote the manuscript. EFR performed most experiments. AMJD performed all bioinformatics analyses. SR made the transgenic line and performed dissections. HB performed and analysed the mass-spectrometry experiment. FB and RFK supervised the project.

ACKNOWLEDGMENTS

We would like to thank Yasmin el Sherif and Svenja Hellmann for technical support. We also acknowledge the IMB Genomics core facility, the Protein production core facility, the Microscopy core facility and the Media Lab.

REFERENCES

- 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.
- Carmell, M.A., Girard, A., van de Kant, H.J.G., 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.
- Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 26, 1367–1372.
- Ding, D., Liu, J., Dong, K., Midic, U., Hess, R.A., Xie, H., Demireva, E.Y., and Chen, C. (2017). PNLDC1 is essential for piRNA 3' end trimming and transposon silencing during spermatogenesis in mice. *Nat. Commun.* 8.
- Ding, D., Liu, J., Midic, U., Wu, Y., Dong, K., Melnick, A., Latham, K.E., and Chen, C. (2018). TDRD5 binds piRNA precursors and selectively enhances pachytene piRNA processing in mice. *Nat. Commun.*
- Gainetdinov, I., Colpan, C., Arif, A., Cecchini, K., and Zamore, P.D. (2018). A Single Mechanism of Biogenesis, Initiated and Directed by PIWI Proteins, Explains piRNA Production in Most Animals. *Mol. Cell*
- 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* (80-.). 348, 817–821.
- Han, B.W., Wang, W., Li, C., Weng, Z., and Zamore, P.D. (2015). PiRNA-guided transposon cleavage initiates Zucchini-dependent, phased piRNA production. *Science* (80-.). 348, 817–821.
- Hayashi, R., Schnabl, J., Handler, D., Mohn, F., Ameres, S.L., and Brennecke, J. (2016). Genetic and mechanistic diversity of piRNA 3'-end formation. *Nature*.
- Honda, S., Kirino, Y.Y., Maragkakis, M., Alexiou, P., Ohtaki, A., Murali, R., Mourelatos, Z., and Kirino, Y.Y. (2013). Mitochondrial protein BmPAPI modulates the length of mature piRNAs. *RNA* 19, 1405–1418.
- 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.
- 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.
- Huang, H.-Y., Houwing, S., Kaaij, L.J.T., Meppelink, A., Redl, S., Gauci, S., Vos, H., Draper, B.W., Moens, C.B., Burgering, B.M., et al. (2011). Tdrd1 acts as a molecular scaffold for Piwi proteins and piRNA targets in zebrafish. *EMBO J.* 30, 3298–3308.
- Ipsaro, J.J., Haase, A.D., Knott, S.R., Joshua-Tor, L., and Hannon, G.J. (2012). The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis. *Nature* 491, 279–283.
- Izumi, N., Shoji, K., Sakaguchi, Y., Honda, S., Kirino, Y., Suzuki, T., Katsuma, S., and Tomari, Y. (2016). Identification and Functional Analysis of the Pre-piRNA 3' Trimmer in Silkworms. *Cell*.
- Kamminga, L.M., Luteijn, M.J., den Broeder, M.J., Redl, S., Kaaij, L.J.T., 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.
- Kirino, Y., Vourekas, A., Sayed, N., de Lima Alves, F., Thomson, T., Lasko, P., Rappsilber, J., Jongens, T.A., and Mourelatos, Z. (2010). Arginine methylation of Aubergine mediates Tudor binding and germ plasm localization. *RNA* 16, 70–78.
- 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.

19. Kwan, K.M., Fujimoto, E., Grabher, C., Mangum, B.D., Hardy, M.E., Campbell, D.S., Parant, J.M., Yost, H.J., Kanki, J.P., and Chien, C. Bin (2007). The Tol2kit: A multisite gateway-based construction Kit for Tol2 transposon transgenesis constructs. *Dev. Dyn.* 236, 3088–3099.
20. Leu, D.H., and Draper, B.W. (2010). The ziwi promoter drives germline-specific gene expression in zebrafish. *Dev. Dyn.* 239, 2714–2721.
21. 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.
22. Mohn, F., Handler, D., and Brennecke, J. (2015). PiRNA-guided slicing specifies transcripts for Zucchini-dependent, phased piRNA biogenesis. *Science* (80-.). 348, 812–817.
23. 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.
24. Nishida, K.M., Sakakibara, K., Iwasaki, Y.W., Yamada, H., Murakami, R., Murota, Y., Kawamura, T., Kodama, T., Siomi, H., and Siomi, M.C. (2018). Hierarchical roles of mitochondrial Papi and Zucchini in *Bombyx* germline piRNA biogenesis. *Nature*.
25. Rappsilber, J., Mann, M., and Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* 2, 1896–1906.
26. Saito, K., Sakaguchi, Y., Suzuki, T., Suzuki, T., Siomi, H., and Siomi, M.C. (2007). Pimet, the *Drosophila* homolog of HEN1, mediates 2'-O-methylation of Piwi-interacting RNAs at their 3' ends. *Genes Dev.* 21, 1603–1608.
27. Saxe, J.P., Chen, M., Zhao, H., and Lin, H. (2013). Tdrkh is essential for spermatogenesis and participates in primary piRNA biogenesis in the germline. *EMBO J.* 32, 1869–1885.
28. Shevchenko, A., Tomas, H., Havli, J., Olsen, J. V., and Mann, M. (2007). In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* 1, 2856–2860.
29. Siomi, H., Matunis, M.J., Michael, W.M., and Dreyfuss, G. (1993). The Pre-Messenger RNA Binding K-Protein Contains a Novel Evolutionarily Conserved Motif. *Nucleic Acids Res* 21, 1193–1198.
30. 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.* 102, 4074–4079.
31. 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.
32. Vagin, V. V., Wohlschlegel, J., Qu, J., Jonsson, Z., Huang, X., Chuma, S., Girard, A., Sachidanandam, R., Hannon, G.J., and Aravin, A.A. (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.
33. Valverde, R., Edwards, L., and Regan, L. (2008). Structure and function of KH domains. *FEBS J.*
34. Wang, H., Ma, Z., Niu, K., Xiao, Y., Wu, X., Pan, C., Zhao, Y., Wang, K., Zhang, Y., and Liu, N. (2016). Antagonistic roles of Nibbler and Hen1 in modulating piRNA 3' ends in *Drosophila*. *Development* 143, 530–539.
35. Wasik, K.A., Tam, O.H., Knott, S.R., Falcatori, I., Hammell, M., Vagin, V. V., and Hannon, G.J. (2015). RNF17 blocks promiscuous activity of PIWI proteins in mouse testes. *Genes Dev.*
36. Westerfield, M. (1995). *The Zebrafish Book. A Guide for the Laboratory Use of Zebrafish (Danio rerio)*, 3rd Edition. Eugene, OR, Univ. Oregon Press 385.
37. Zhang, Y., Guo, R., Cui, Y., Zhu, Z., Zhang, Y., Wu, H., Zheng, B., Yue, Q., Bai, S., Zeng, W., et al. (2017). An essential role for PNLDC1 in piRNA 3' end trimming and male fertility in mice. *Cell Res.* 27, 1392–1396.
38. Zhang, Y.Z., Ouyang, Y.C., Hou, Y., Schatten, H., Chen, D.Y., and Sun, Q.Y. (2008). Mitochondrial behavior during oogenesis in zebrafish: A confocal microscopy analysis. *Dev. Growth Differ.*

SUPPLEMENTARY FIGURES

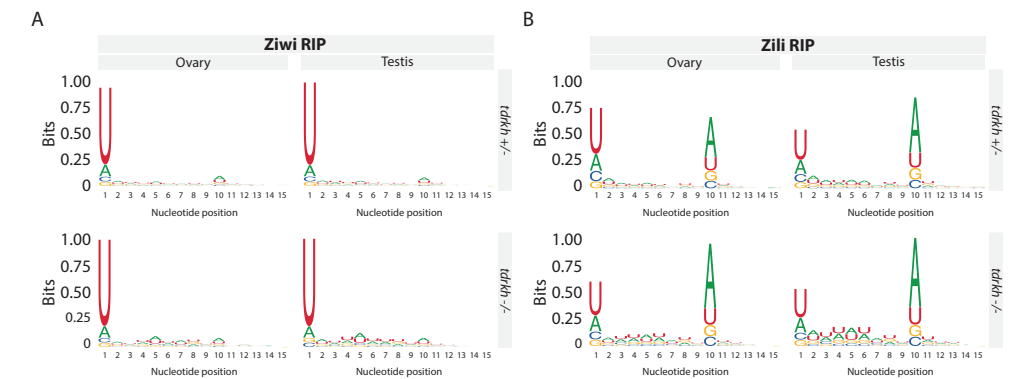


Figure S1, Related to Figure 4. (A, B) Sequence logos of piRNAs from Ziwi (A) and Zili (B) RIPs on *tdrkh*^{+/+} and *tdrkh*^{-/-} ovary and testis. A strong 5'U bias is found especially in Ziwi RIPs, whereas Zili RIPs display a strong A at position 10 characteristic for ping-pong.

2

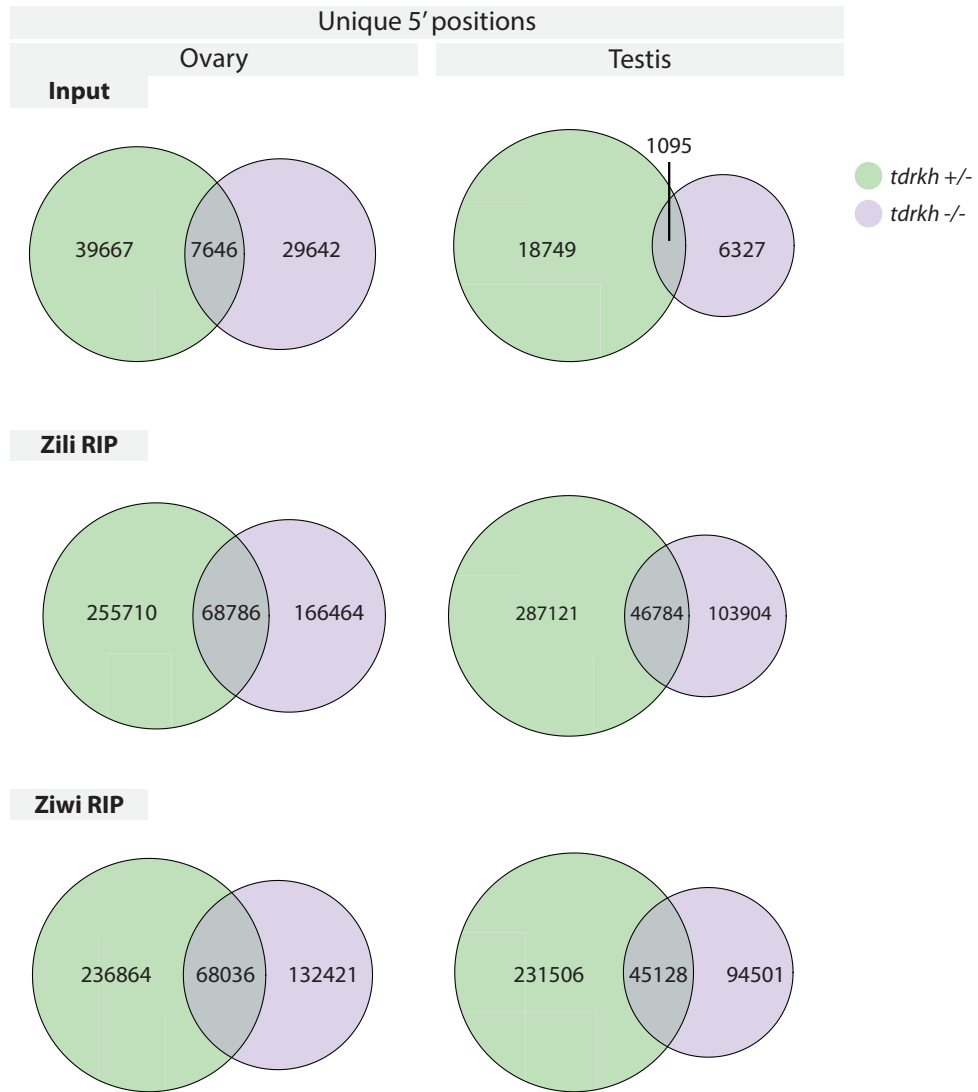


Figure S2, Related to Figure 5. Venn diagrams displaying the number of unique and the shared 5' positions. Even though only few positions are shared, these positions still produce the majority of all piRNAs (see Figure 5D).

Chapter

PIWI proteins and piRNAs in
mammalian oocytes and
early embryos

3

ABSTRACT

Germ cells of most animals critically depend on piRNAs and Piwi proteins. Surprisingly, piRNAs in mouse oocytes are relatively rare and dispensable. We present compelling evidence for strong Piwi-piRNA expression in oocytes of other mammals. Human fetal oocytes express PIWIL2 and transposon-enriched piRNAs. Oocytes in adult human ovary express PIWIL1 and PIWIL2, while those in bovine ovary only express PIWIL1. In human, macaque and bovine ovaries we find piRNAs that resemble testis-borne pachytene piRNAs. Isolated bovine follicular oocytes were shown to contain abundant, relatively short piRNAs that preferentially target transposable elements. Using label-free quantitative proteome analysis we show that these maturing oocytes strongly and specifically express the thus-far uncharacterized PIWIL3 protein, alongside other, known piRNA-pathway components. A piRNA pool is still present in early bovine embryos, revealing a potential impact of piRNAs on mammalian embryogenesis. Our results reveal unexpected, highly dynamic piRNA pathways in mammalian oocytes and early embryos.

Elke F. Roovers^{1,†}, David Rosenkranz^{2,†}, Mahdi Mahdipour^{3,†}, Chung-Ting Han⁴, Nannan He⁵, Susana M. Chuva de Sousa Lopes⁵, Lucette A.J. van der Westerlaken⁶, Hans Zischler², Falk Butter⁷, Bernard A.J. Roelen^{3,8}, René F. Ketting^{1,8}

¹Biology of Non-coding RNA Group, Institute of Molecular Biology (IMB), Ackermannweg 4, 55128 Mainz, Germany.

²Johannes Gutenberg University Mainz, Institute of Anthropology, Anselm-Franz-von-Bentzel-Weg 7, 55128 Mainz, Germany.

³Department of Farm Animal Health, Faculty of Veterinary Medicine, Utrecht University, 3584 CM Utrecht, The Netherlands.

⁴Genomics Core Facility, Institute of Molecular Biology (IMB), Ackermannweg 4, 55128 Mainz, Germany.

⁵Department of Anatomy & Embryology, Leiden University Medical Center, Leiden, The Netherlands.

⁶Department of Gynaecology, Leiden University Medical Center, Leiden, The Netherlands.

⁷Quantitative Proteomics Group, Institute of Molecular Biology (IMB), Ackermannweg 4, 55128 Mainz, Germany.

⁸Co-senior author

[†]Contributed equally

Adapted from: Roovers, E.F.*, Rosenkranz, D.*, Mahdipour, M.*, Han, C.-T.T., He, N., de Sousa Lopes, S.M.C., van der Westerlaken, L.A.J., Zischler, H., Butter, F., Roelen, B.A.J., and Ketting, R.F. (2015). Piwi proteins and piRNAs in Mammalian Oocytes and early embryos. *Cell Rep.* 10, 2070–2083.

INTRODUCTION

piRNAs represent a predominantly germ cell-specific class of small RNAs that interact with a specific class of Argonaute proteins, known as Piwi proteins (Ghildiyal and Zamore, 2009; Malone and Hannon, 2009). Both Piwi proteins and piRNAs are essential for germ cell development and function (Ketting, 2011). In their absence, germ cells can arrest during proliferation or meiosis, or they can undergo apoptosis. The molecular mechanisms behind these phenotypes have not been properly described in all cases, but commonly, defects in the silencing of repetitive sequences in the genome, such as transposable elements, are important causative factors (Saito and Siomi, 2010). In line with this, piRNA populations are often enriched for transposon-derived sequences. One notable exception is the piRNA population that is expressed during spermatogenesis in mammals. These so-called pachytene piRNAs, named after the fact that they start to be expressed at the onset of pachytene stage during meiosis, are depleted of transposon-derived sequences. Therefore, this group of piRNAs could be important for additional regulatory developmental mechanisms (Gou et al., 2014; Weick and Miska, 2014), even though they can still play a role in the control of transposon activity (Di Giacomo et al., 2013).

piRNAs such as the pachytene piRNAs are derived from so-called piRNA-precursor transcripts that in turn come from large clusters that are often bi-directionally transcribed (Aravin et al., 2006; Girard et al., 2006; Lau et al., 2006). Among various mammals, the sequences of these clusters are not conserved, but their location within the genome is. The precursor transcripts are processed by a nuclease named Zucchini (Ipsaro et al., 2012; Voigt et al., 2012), after which a Piwi protein binds to the 5' end of one of the generated fragments. This process results in piRNAs that are characterized by a strong bias for uracil at their 5' ends, most likely because of a 5' nucleotide preference of the Piwi protein involved (Kawaoka et al., 2011). Biogenesis is finalized by trimming of the 3' end of the piRNA intermediate (Kawaoka et al., 2011) and 2'O-methylation of the most 3' base of the piRNA (Houwing et al., 2007; Vagin et al., 2006). This modification is placed by the enzyme Hen1, and serves to inhibit the addition of 3' non-templated nucleotides to the 3' ends of piRNAs (Ameres et al., 2010; Horwich et al., 2007; Kamminga et al., 2010; Saito et al., 2007). These non-templated nucleotides include mostly uridine and adenosine and are related to piRNA turnover.

In addition to this Zucchini-driven pathway, also referred to as primary biogenesis, a secondary piRNA biogenesis mechanism exists. This has been described best in *Drosophila* oocytes, where two Piwi paralogs, Aub and Ago3, participate in this mechanism (Brennecke et al., 2007; Gunawardane et al., 2007), but operates in vertebrates as well (Aravin et al., 2008; Aravin et al., 2007; Houwing et al., 2008; Houwing et al., 2007). First, Aub is loaded through the primary mechanism described above. Then, when Aub cleaves a target transcript, Ago3 is loaded with the 5' cleavage fragment generated by Aub.

Reversely, when Ago3 cleaves a target RNA, an empty Aub molecule can be loaded with the generated 5' cleavage fragment. This results in a very characteristic relationship, also known as the ping-pong signal, between the piRNAs bound by the two different Piwi paralogs: they

are derived from complementary transcripts and overlap precisely ten nucleotides at their 5' ends (Brennecke et al., 2007; Gunawardane et al., 2007). This derives from the fact that Argonaute proteins cleave their target transcripts precisely between nucleotides ten and eleven from the 5' end of the bound small RNA. Consequently, Aub-bound piRNAs are marked by a U at position one, while Ago3-bound piRNAs have a characteristic adenosine enrichment at position ten.

Piwi proteins can also induce transcriptional silencing. In *Drosophila* this is done by Piwi (Le Thomas et al., 2013; Rozhkov et al., 2013; Sienski et al., 2012), a PIWI paralog that is exclusively loaded through primary biogenesis. In mouse embryonic, male germ cells a nuclear Piwi pathway has been shown to be driven by MIWI2 (PIWIL4) (Aravin et al., 2008; Kuramochi-Miyagawa et al., 2008). MIWI2 is specifically loaded through a secondary piRNA biogenesis mechanism, in which the Piwi-paralog MILI (PIWIL2) acts to initiate the loading of MIWI2 (Aravin et al., 2008).

In model organisms like *Drosophila* and zebrafish, PIWI proteins and piRNAs are essential during both spermatogenesis and oogenesis (Ketting, 2011). Strikingly, none of the PIWI proteins affect female fertility (Carmell et al., 2007; Deng and Lin, 2002; Kuramochi-Miyagawa et al., 2004). Mouse oocytes do express PIWI proteins (Aravin et al., 2008; Ding et al., 2013; Lim et al., 2013), but only relatively minor amounts of piRNAs have been detected (Tam et al., 2008; Watanabe et al., 2008). In contrast, the siRNA-generating enzyme DICER is essential for progression of mouse oocytes through meiosis (Murchison et al., 2007), paralleled by relatively high levels of siRNAs in oocytes (Tam et al., 2008; Watanabe et al., 2008). Interestingly, Dicer expression in mouse ovary is driven by a Muridae-specific retro-transposon insertion into the Dicer locus (Flemer et al., 2013). Furthermore, the mouse and rat genomes lack one of the four Piwi paralogs that is found in most other mammals, including primates and humans, PIWIL3.

We reasoned that both the lack of PIWIL3 from the mouse and rat genomes as well as the Muridae-specific Dicer locus may be indications that the small-RNA-status of mouse oocytes is not representative for mammals in general. We therefore profiled small RNAs from bovine, macaque and human ovaries and from isolated, maturing bovine oocytes. We also used label-free quantitative mass spectrometry, to analyze the proteome of bovine oocytes. These experiments uncovered very dynamic PIWI pathways in the female germline of these mammals, including oocyte-specific expression of PIWIL3, while DICER-driven pathways are only poorly represented.

RESULTS

PIWIL1 expression in bovine ovary

We first tested whether Piwi proteins can be detected in adult mammalian ovarian tissue by performing immunofluorescence (IF) experiments on bovine (*Bos taurus*) ovarian tissue, using testis as a positive control (Figure S1A). We observed PIWIL1 expression oocytes starting at primordial stage (Figure 1A). In addition, we could also detect PIWIL1 transcripts using RT-PCR in both tissues (not shown). While PIWIL1 and VASA (DDX4) displayed the typical

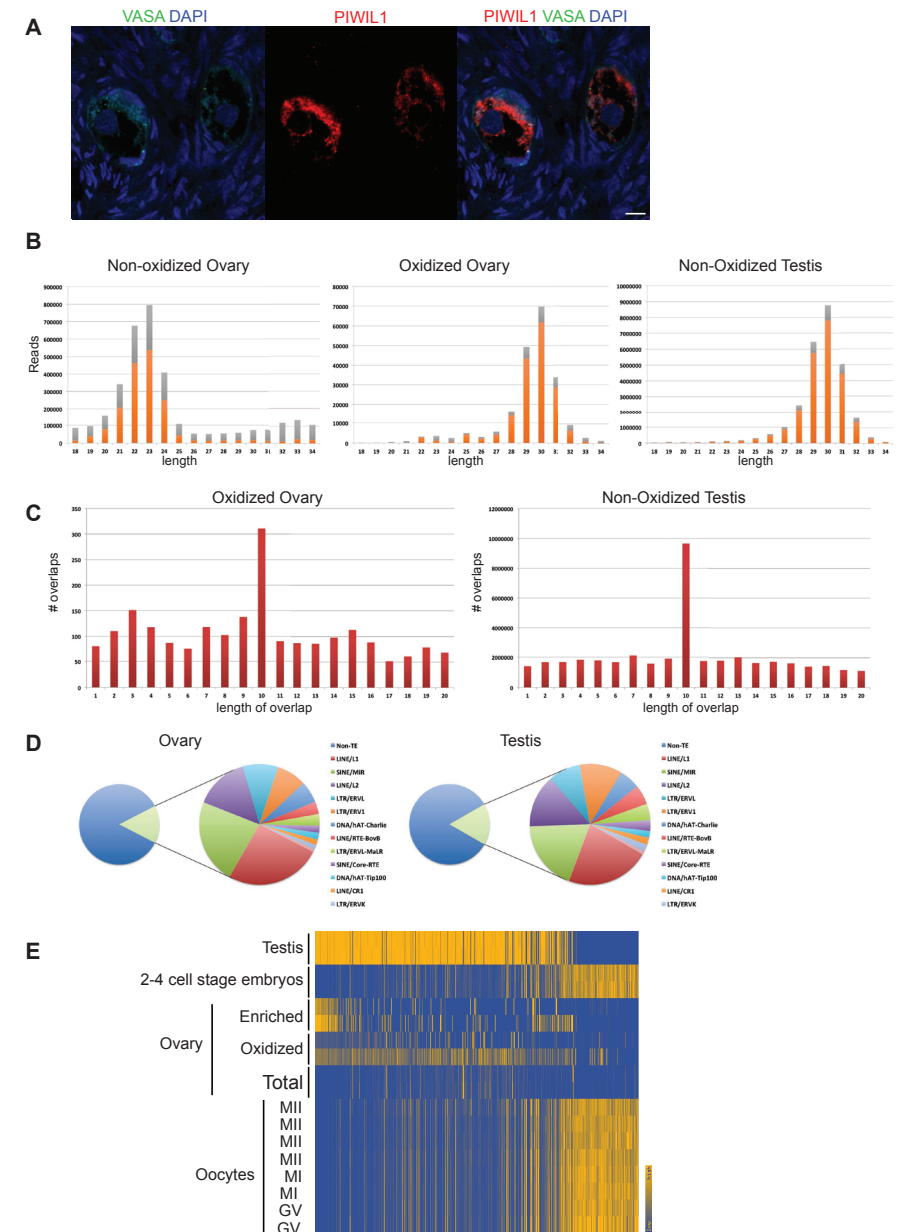


Figure 1. Bovine ovarian piRNAs. (A) Confocal image of PIWIL1 and VASA staining on bovine ovarian section. Scale bar is 10 μ m. (B) Length profiles of the indicated libraries. Only non-annotated reads are depicted. Orange: reads starting with a U. Grey: reads not starting with a U. (C) Overlaps of 5' ends of reads that are mapped to opposite strands of the same locus (ping-pong signal). (D) Diagram depicting the transposon content of bovine ovarian and testicular piRNA populations. (E) Heatmap comparing piRNA clusters called from ovarian, testicular and oocyte-derived (see Figure 5) piRNAs. The map is ordered according to piRNA abundance in testis clusters, and reflects relative piRNA density between the various samples (highest score at each locus was set to 1 and the lowest to 0). 'Enriched' refers to reads from loci that are at least 5-fold enriched following oxidation.

granular structure in testis (Figure S1A), such structures were not obvious in oocytes. PIWIL2 expression was not detected in adult bovine ovary (not shown).

Bovine ovarian piRNAs

To follow up on the detection of PIWIL1 in bovine adult ovary, we made small RNA libraries from bovine ovary total RNA. As a comparison, we also generated small RNA libraries from testis. In both replicates of the testis a strong peak of ~30-nucleotide piRNAs could be observed, and relatively few miRNA reads (Figure 1B, S1B, S2), consistent with previously published data on testis piRNAs in mouse, rat and human (Aravin et al., 2006; Girard et al., 2006; Lau et al., 2006). In the ovary, however, no such population could be detected (Figure 1B, S1B). Given that the germ-cell-count in ovarian tissue is much lower than in testis, we reasoned that a piRNA population might be hidden by somatic small RNAs. To circumvent this, we made additional libraries of the same ovarian RNA samples following treatment with sodium-periodate (NaIO₄). This makes most small RNA species non-accessible for cloning, except when they are protected from oxidation at their 3' end, as piRNAs are, through 2'O-methylation (Vagin et al., 2006). Plotting only those reads from loci that are at least 5-fold enriched through this procedure, a clear piRNA-like population was detected (Figure 1B, S1B). Interestingly, the length-distribution of these RNAs, and their 5'-bias for uracil is virtually indistinguishable from those obtained from testis (Figure 1B). The piRNAs from testis, as well as the NaIO₄-enriched population from ovary show a pronounced ping-pong signal (Figure 1C). Since the vast majority of these piRNAs is presumably bound by PIWIL1, this signature likely results from homotypic ping-pong interaction, and may not play an important role in piRNA amplification. Nevertheless, it is a strong indication that these RNAs are bona fide piRNAs. The transposon-repertoires of the ovarian and testicular piRNA pools are very similar (Figure 1D, S1C, S1D), with relatively low transposon coverage among the total small RNA pool (14% for ovary, 18% for testis). Finally, we defined piRNA clusters, using previously described software (Rosenkranz and Zischler, 2012). In testis, 564 clusters were called, representing 0.2% of the genome. More than 90% of all sequenced piRNAs can come from these loci. In ovary, 121 clusters were called (0.05% of the genome), and these can produce 84% of all piRNAs sequenced. The majority of ovarian clusters called are also represented among testis piRNA clusters (Figure 1E). Interestingly, approximately 90% of the ovarian piRNAs (defined as 5-fold enriched following oxidation) can come from testicular piRNA cluster loci. We conclude that bovine ovary expresses PIWIL1 and piRNAs that are very similar to piRNAs expressed during pachytene stage of spermatogenesis.

Ovary-derived piRNAs from macaque

We then turned to primates and analyzed adult macaque (*Macaca fascicularis*) ovarian and testicular tissue, using the same NaIO₄-treatment-strategy as for bovine ovary. We now considered 2-fold enrichment, since 5-fold enrichment reduced the read numbers dramatically. Using these settings we could clearly detect piRNA-like populations in ovarian samples of two macaque individuals (Figure 2A, S3A). Again, the length of these piRNAs is similar between

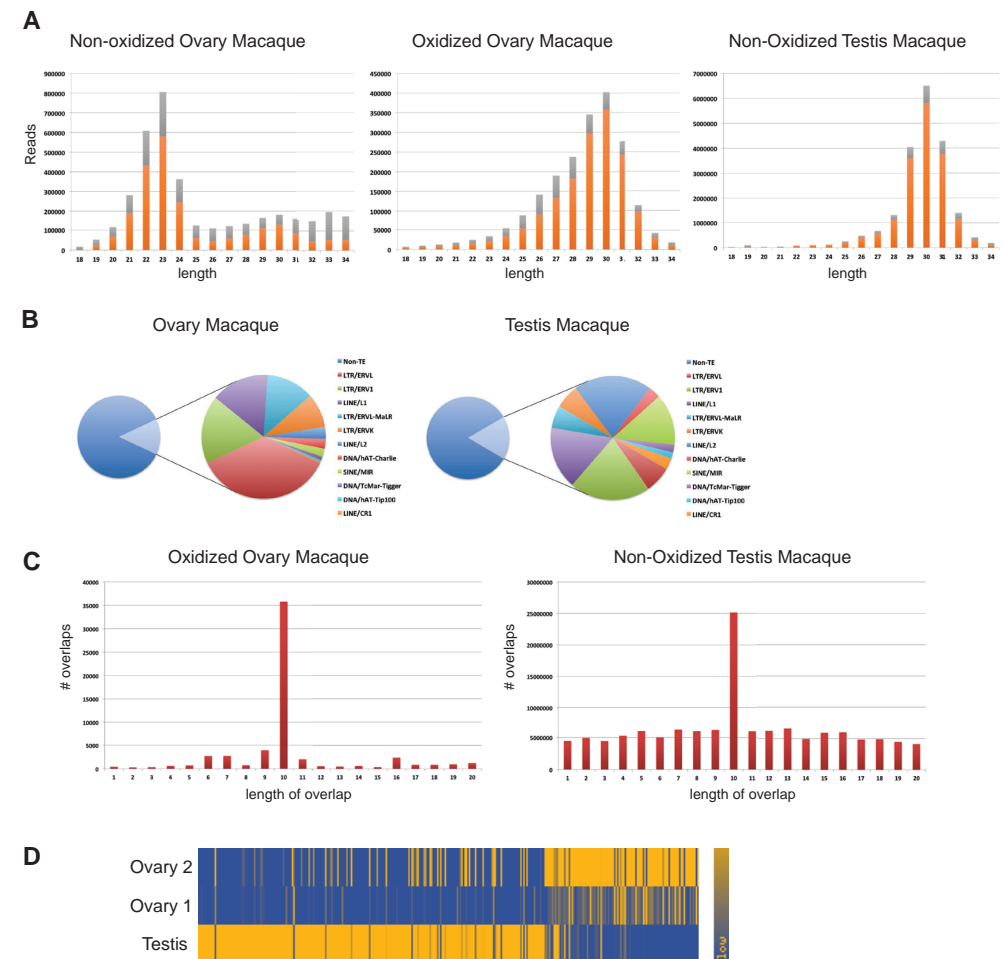


Figure 2. Macaque ovarian piRNAs. (A) Length profiles of the indicated libraries. Only non-annotated reads are depicted. Orange: reads starting with a U. Grey: reads not starting with a U. (B) Diagram depicting the transposon content of macaque ovarian and testicular piRNA populations. (C) Overlaps of 5' ends of reads that are mapped to opposite strands of the same locus. (D) Heatmap comparing piRNA clusters called from ovarian and testicular piRNAs. The map is ordered according to piRNA density in testicular clusters, and reflects relative piRNA density between the various samples (highest score at each locus was set to 1 and the lowest to 0).

testicular and ovarian samples (Figure 2A), although the ovarian piRNA population seems to be a bit broader in size. Possibly, a second, somewhat smaller piRNA population also exists next to the 30-nucleotide PIWIL1 piRNAs (also see next section).

With regard to transposons, the ovarian and testicular piRNA pools are similar, with the exception of ERVL and L2 (Figure 2B, S3B, S3C). Overall, transposon coverage is again low for both ovary (14%) and testis (17%). As among bovine ovarian and testicular piRNAs, a clear ping-pong signal for macaque ovarian and testicular piRNAs is evident (Figure 2C).

Finally, there is again a significant overlap between the piRNA clusters called for ovary piRNAs (264 clusters, explaining 94% of the piRNAs) and those called for testis-derived piRNAs (605 clusters, explaining 92% of the sequenced piRNAs) (Figure 2D).

That said, the resemblance between ovarian and testicular clusters is less convincing for macaque than for bovine piRNAs, and accordingly, only about 10% of ovarian piRNAs can come from testicular piRNA clusters. Based on this data we conclude that also the macaque ovary expresses piRNAs in a manner that resembles pachytene piRNAs in testis, even though both tissues clearly produce unique piRNAs.

Expression of PIWIL1 and piRNAs in human ovary

Next, we analyzed PIWIL1 expression in human adult ovaries using IF. This revealed clear expression of PIWIL1 in oocytes that were marked by expression of VASA (Figure 3A). The staining pattern is very similar to the staining of PIWIL1 in bovine ovary. In addition to PIWIL1, and different from what we observed in bovine ovary, oocytes in human ovary also stained positive for PIWIL2 (Figure 3B). No obvious granular structures were observed for PIWIL1, PIWIL2 or VASA.

We then profiled the small RNAs of two human ovarian cortex samples from two different patients with reproductive age (Figure 3C, S4A). Following NaIO₄-mediated oxidation of the RNA, and considering sequences from loci that were at least 5-fold enriched, we could detect clear piRNA-like small RNAs, with characteristics that are very similar to those obtained from macaque ovary.

Again, the size profile of this piRNA pool suggests the existence of two piRNA pools, one with a size of ~30 nucleotides and one that is somewhat smaller. This would be consistent with a PIWIL1- and a PIWIL2-bound piRNA pool in human ovary. In comparison, we only detected PIWIL1 in bovine ovary, and the size profile for bovine ovarian piRNAs is much sharper. While we cannot prove it, our data would be consistent with a more dominant, PIWIL1-bound piRNA pool in bovine ovary compared to human and macaque ovary.

We could define only 34 human ovarian piRNA clusters that explain only 5% of the sequences that were enriched through oxidation. Possibly, the relatively low number of reads that was significantly enriched following oxidation prevents efficient cluster calling for these human samples. The oxidation-resistant RNA pool revealed a clear ping-pong signal (Figure 3D), and only 18% of these RNAs is derived from transposable elements (Figure 3E, S4B, C). Overall, these results strongly suggest that adult human oocytes contain PIWIL1- and PIWIL2-bound piRNAs.

Human fetal oocytes express PIWIL2

We also addressed the Piwi status of oocytes during development using IF. For this, we analyzed both first and second trimester ovary samples for PIWIL1 and PIWIL2. We did not detect VASA, and we did not see specific PIWIL1/2 staining in OCT4 positive cells (germ cells) in first trimester ovaries (data not shown). However, we detect strong staining for PIWIL2 in

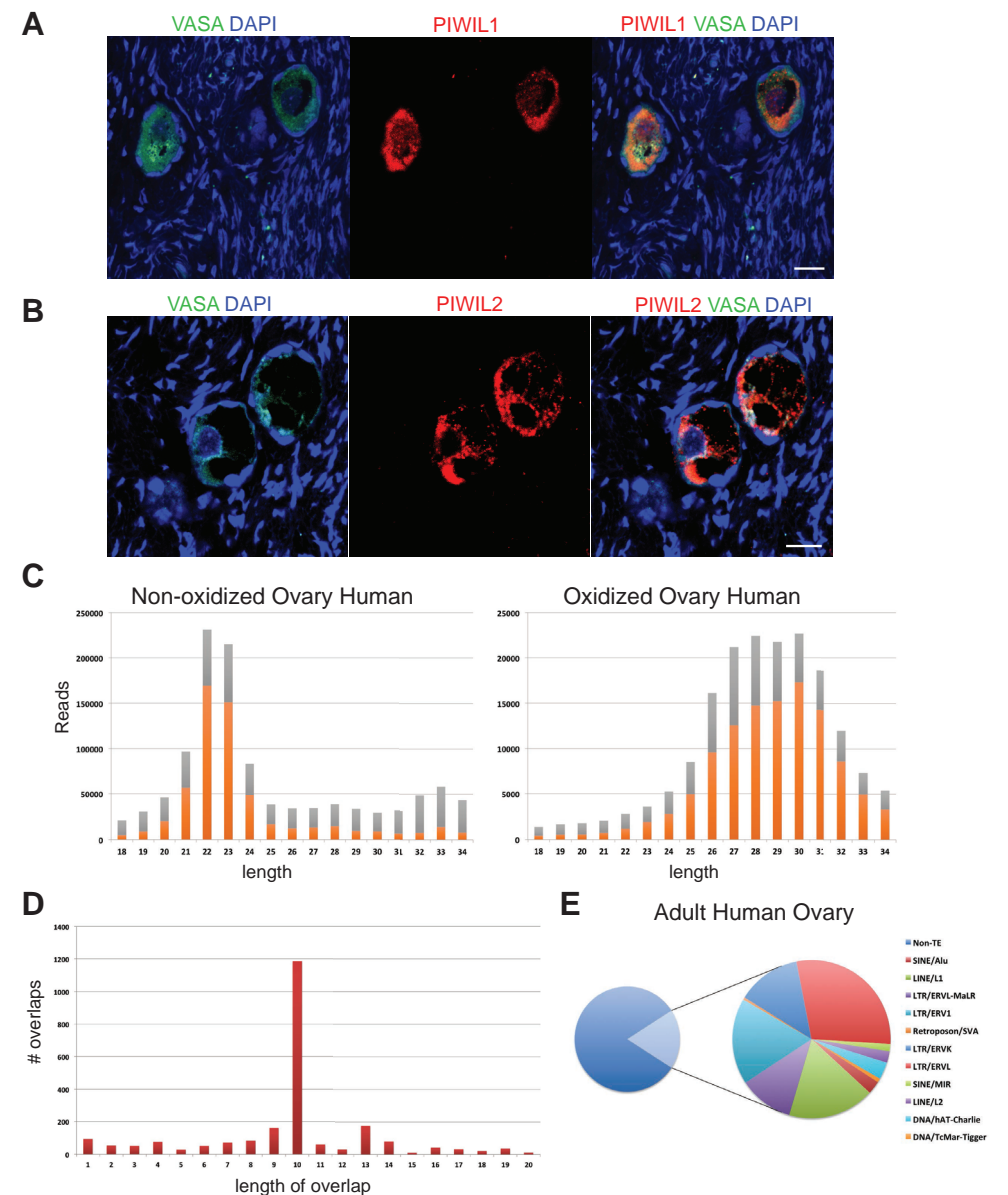


Figure 3. Human adult ovarian piRNAs. (A) Confocal image of PIWIL1 and VASA staining on human adult ovarian sections. Scale bar is 15 μ m. (B) Confocal image of PIWIL2 and VASA staining on human adult ovarian sections. Scale bar is 15 μ m. (C) Length profiles of the indicated libraries. Only non-annotated reads are depicted. Orange: reads starting with a U. Grey: reads not starting with a U. (D) Overlaps of 5' ends of reads that are mapped to opposite strands of the same locus. (E) Diagram depicting the transposon content of human ovarian piRNAs.

second trimester ovaries, especially in those germ cells that are no longer OCT4 positive, but VASA positive (Figure 4A). Strikingly, both VASA and PIWIL2 localize to aggregates close to the nucleus, but are both absent from the nucleus itself. This is a pattern that has been observed in mouse gonocytes as well. In fact, VASA is known to promote the ping-pong cycle (Xiol et al., 2014). In mouse gonocytes, PIWIL4 acts as the secondary Piwi protein, downstream of PIWIL2 (Aravin et al., 2008), but could not detect convincing, germ cell specific staining with available PIWIL4 antibodies. Consistent with this observation, published RNA expression data (Gkoutela et al., 2013) show relatively low expression of *Piwil4* in fetal human oocytes (Figure S5D).

piRNAs from human fetal ovaries

We also sequenced small RNA populations from human fetal ovaries. Two independent samples of first trimester ovaries (8.5 and 10.5 weeks of gestation) did not yield convincing piRNA populations, whether we treated the RNA with NaIO_4 or not (Figure S5A). This is consistent with the fact that during the first trimester oocytes typically are still negative for VASA and PIWIL2 (data not shown). Strikingly, in two independent second trimester ovaries (19 and 19.5 weeks of gestation), when VASA and PIWIL2 are strongly expressed (Figure 4A, B), we detect convincing piRNA-like profiles (Figure 4C, S5B), displaying a prominent ping-pong signal (Figure 4D). Roughly 50% of the oxidation-resistant RNAs can be explained by ~500 clusters. With on average 27 nucleotides, these piRNAs are significantly smaller than the 30-nucleotide piRNAs detected in the human adult ovary, and their transposon-targeting capacity (30%) is higher than in the adult ovary (18%) (Figure 4E, S4C, S5C). In mouse gonocytes, the overall piRNA population peaks at 27 nucleotides as well, with a similar shoulder towards somewhat longer piRNAs (Aravin et al., 2008). For these mouse piRNAs it was established that the overall population is in fact a merge of two piRNA populations, bound by PIWIL2 (binding primary piRNAs of ~26 nucleotides) and PIWIL4 (binding secondary piRNAs of ~28 nucleotides) (Aravin et al., 2008; Aravin et al., 2007). In absence of PIWIL2/4 IP data, we checked whether the sense and antisense piRNAs from the two most abundantly represented transposons (ALU and L1) displayed a significant length difference. We could, however, not detect a difference that was consistent between the two fetal ovarian samples. We did detect a modest, but consistent increase in 1U bias for anti-sense piRNAs, and an increase in 10A bias for sense piRNAs for both transposon types in both samples, while this was absent for other abundantly represented transposons (Figure 4F). These results suggest that fetal human ovaries express piRNAs that are mostly bound by PIWIL2. A clean signature hinting at a PIWIL2-PIWIL4 interaction could not be distilled from our data.

piRNAs from isolated bovine oocytes from antral follicles

Are piRNAs still present in more mature oocytes? We tested that by generating and sequencing small RNA libraries from three independent sets of bovine germinal vesicle (GV) stage oocytes, collected from antral follicles and stripped from their accompanying cumulus cells. Less than

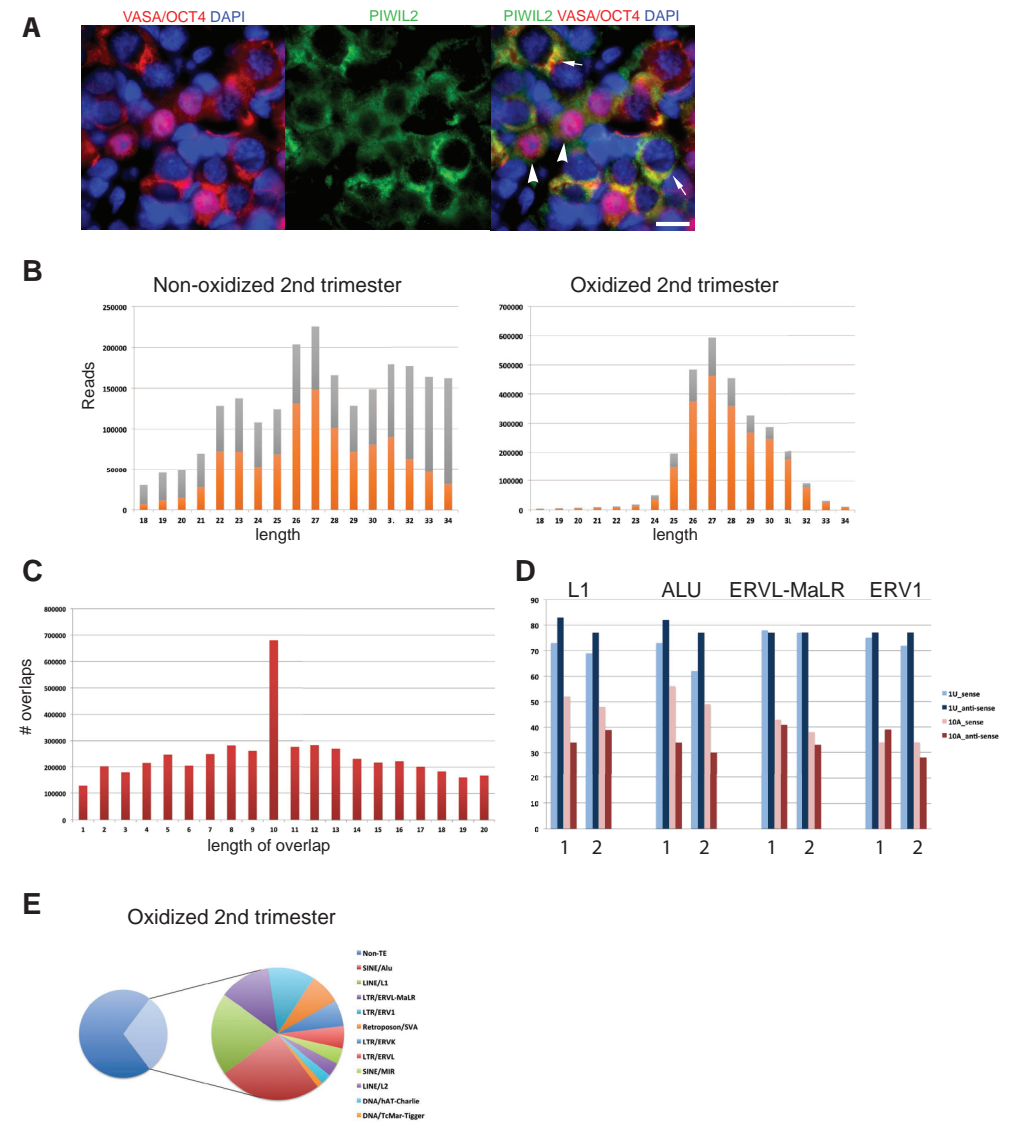


Figure 4. Human fetal ovarian piRNAs. (A) Wide-field image of PIWIL2 (green), VASA (red, cytoplasm) and OCT4 (red, nucleus) staining on human 2nd trimester ovary. Arrows indicate VASA positive cells and arrowheads indicate OCT4 positive cells. Note the PIWIL2 and VASA aggregates in the OCT4 negative cells. Scale bar is 10 μm . (B) Length profiles of the indicated libraries. Only non-annotated reads are depicted. Orange: reads starting with a U. Grey: reads not starting with a U. (C) Overlaps of 5' ends of reads that are mapped to opposite strands of the same locus. (D) Diagram depicting the transposon content of 2nd trimester human ovarian piRNAs. (E) 1U and 10A bias among piRNA reads derived from the indicated repetitive elements. Reads were separated according to sense or anti-sense polarity. Labels '1' and '2' refer to both biological replicates.

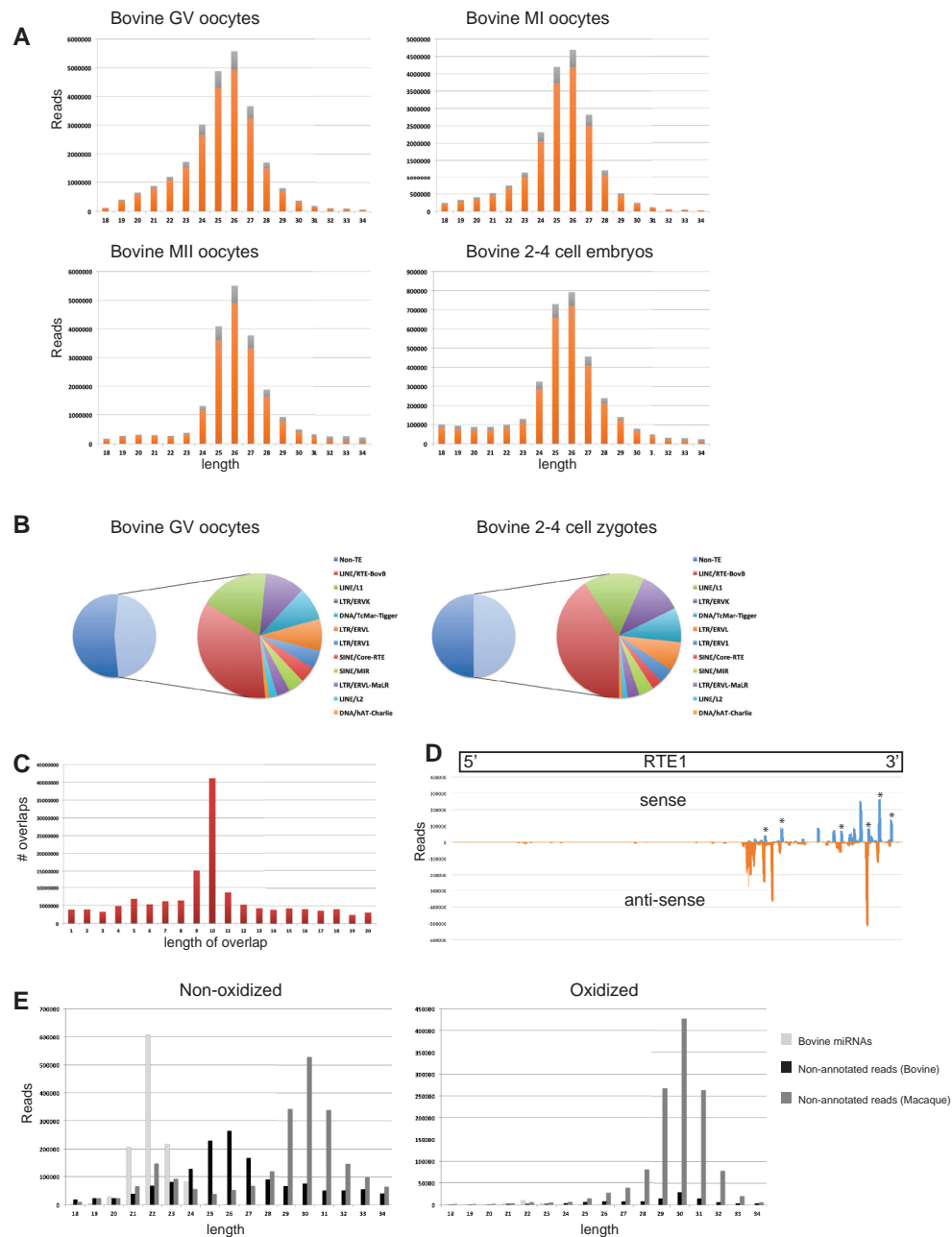


Figure 5. piRNAs from bovine oocytes and embryos. (A) Length profiles of the indicated libraries. Only non-annotated reads are depicted. Orange: reads starting with a U. Grey: reads not starting with a U. (B) Diagram depicting the transposon content of bovine piRNA populations from oocytes and embryos of the indicated developmental stages. (C) Overlaps of 5' ends of reads that are mapped to opposite strands of the same locus. (D) Distribution of sense and anti-sense piRNAs over the RTE1 element. The asterisks indicate pairs of sense and anti-sense piRNAs that overlap 10 nucleotides at their 5' ends. (E) Length distribution of reads obtained from small RNA libraries made from bovine oocyte-cumulus complexes. ▶

1% of the oocyte-derived small RNAs were miRNAs (Figure S1), and the majority of sequences were derived from a piRNA-like population (Figure 5A), ~26 nucleotides in length. These RNAs are strongly enriched for transposon sequences (more than 50%), most notably RTE1, SINE2, L1 and the endogenous retroviruses ERV1-3 (Figure 5B). Overall, a strong ping-pong signal could be detected (Figure 5C). Both sense and antisense piRNAs from, for example, RTE1 are strongly enriched for a U at position 1 (93% and 96% respectively) and an A at position 10 (57% and 62% respectively), consistent with homotypic ping-pong interactions. Finally, the distribution of piRNAs along the RTE1 element shows a strong clustering of piRNAs at the 3' end of the element, with clear signs of ping-pong interactions (Figure 5D). We could define a set of 211 clusters that together span 0.13% of the genome and explain ~83% of the GV-stage piRNAs. These oocyte clusters are very consistent between all oocyte stages and embryos (see below), but very different from the piRNA clusters we defined based on total ovarian or testicular RNA (Figure 1E). Accordingly, less than 10% of the oocyte-derived piRNAs can come from testicular piRNA clusters.

We also matured GV stage oocytes into metaphase I (MI) and metaphase II (MII) stage oocytes and analyzed their small RNA content. Also in these stages abundant piRNAs can be detected, and their characteristics are almost identical to those observed at the GV stage (Figure 5A, 1E). Strikingly, a very similar piRNA pool can still be detected in 2-4 cell stage bovine embryos that were obtained after in vitro fertilization (IVF) (Figure 5A, B). No testis-derived piRNAs were obvious in these embryos, as defined by their length preference of ~30 nucleotides, indicating that the paternal contribution is very limited. Based on size profiles, we could not detect a distinct population of siRNAs in any of the oocyte-derived small RNA libraries, and we could not detect piRNA-like populations in the somatic cumulus cells (Figure S6A, B).

All oocyte and embryo libraries were made without NaIO₄ pre-treatment, indicating that the just-described piRNA-like populations represent by far the most abundant small RNA species in these oocytes. Surprisingly, by sequencing RNA from bovine oocyte-cumulus complexes before and after NaIO₄ treatment, we found that the 26-nucleotide piRNAs are as sensitive to oxidation as miRNAs (Figure 5E). The oxidation efficiency in this experiment was demonstrated by the strong increase in cloning frequency of macaque testis piRNAs. These were mixed into the oocyte-cumulus RNA prior to oxidation and library preparation.

Adenylation is prevalent on piRNAs from follicular bovine oocytes

Non-templated nucleotides at the 3' ends of small RNAs have been described for multiple small RNA classes in multiple species (Ameres et al., 2010; Katoh et al., 2009; Li et al., 2005; van Wolfswinkel et al., 2009). Adenylation and uridylation are the most common forms of these post-transcriptional modifications. Uridylation has been coupled to target-dependent de-

- ▶ Mixed into the RNA was macaque testis RNA. Annotated bovine miRNAs are indicated in light grey, non-annotated bovine-specific RNA fragments in black. Macaque-specific reads are in dark grey. Left panel: non-oxidized. Right panel: oxidized. Only 20% of the macaque reads are displayed in both panels in order to improve visualization of the data.

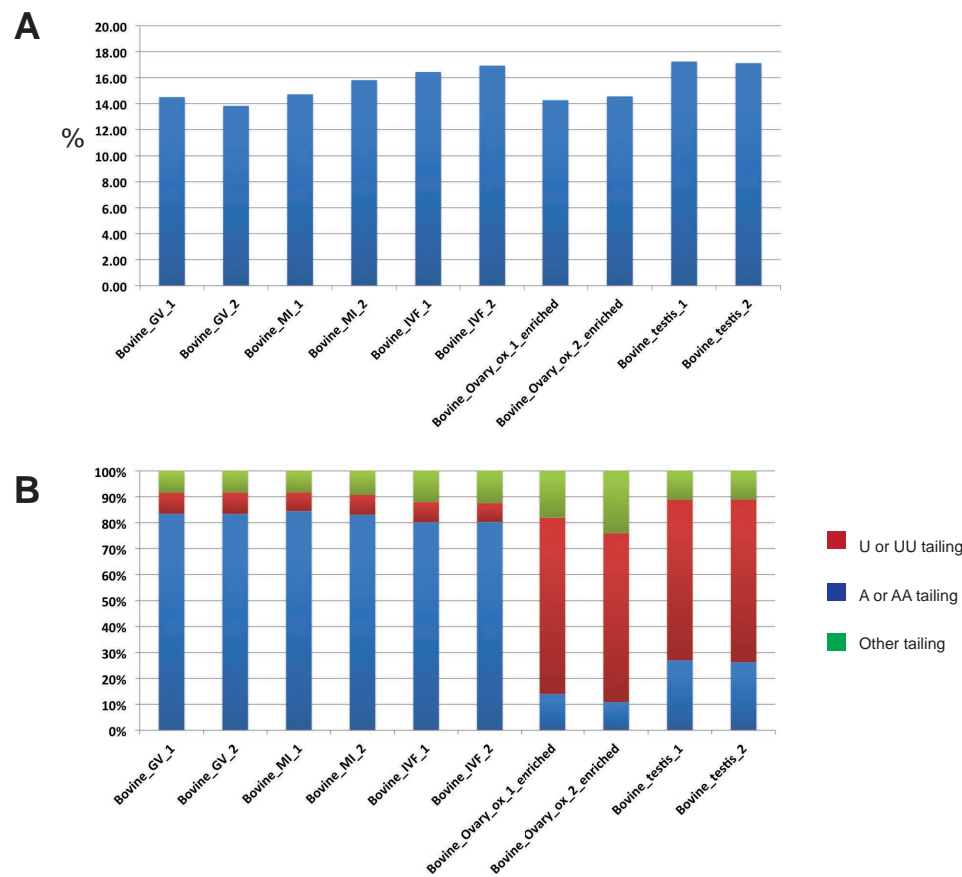


Figure 6. Nontemplated nucleotide analysis. (A) Frequencies of non-templated bases at the 3' end of the reads of the indicated libraries. (B) Separation of the identified non-templated nucleotides into 'A', 'U' or 'other' tails.

stabilization of small RNAs (Ameres et al., 2010), while adenylation has recently been shown to be involved in the clearance of maternal miRNAs (Lee et al., 2014). We therefore checked non-templated nucleotide addition in all bovine samples. In all datasets we see that around 15% of the piRNAs is marked by such extensions of 1-2 nucleotides (Figure 6A). We observed a very interesting difference for non-templated nucleotides on the 26-nucleotide piRNAs from antral follicle oocytes compared to the 30-nucleotide piRNA populations. In all cases, the 30-nucleotide populations mostly displayed uridylation, while the 26-nucleotide piRNAs displayed strong adenylation (Figure 6B). These results are consistent with the proposed role for adenylation in the clearance of maternal miRNAs in embryos. Given the abundance of piRNAs in the 2-4 cell stage embryos, the maternal clearance pathway may not yet be activated in the IVF embryos we have analyzed. In accordance with this, the bovine zygotic genome is activated after the 4-cell stage (Graf et al., 2014).

PIWIL3 is expressed in bovine oocytes

To see which known PIWI pathway-, or RNAi-related proteins are expressed in oocytes, we analyzed the proteome of bovine GV- and MII-stage oocytes using label-free quantitative mass spectrometry. As comparison, we also determined the proteomes of bovine testis and of the oocyte-associated cumulus cells. In the complete dataset, we identified 7031 bovine protein groups (FDR=0.01) and quantified 5359 in at least half of the replicates of any sample. The individual replicates of each sample show excellent reproducibility (Figure 7A), with a correlation between 0.76 in testis and 0.94 in oocytes (Figure S7A). On the overall proteome, oocytes of both maturation stages (GV and MII) are extremely similar (Figure 7B; correlation of 0.87), but clearly distinct from testis and cumulus cells (Figure 7B, S7B). DICER and AGO2 were lowly expressed in our testis samples, but escaped detection in oocytes and cumulus cells. In contrast, many PIWI pathway components, including PLD6, VASA, MAEL, TDRD1 and others, were expressed at significant levels in the oocytes, but not in the cumulus cells (Figure 7C). Strikingly, we detected just one of the four bovine PIWI paralogues in the oocyte samples: PIWIL3. In fact, PIWIL3 was among the most highly expressed proteins in oocytes, and displayed extensive peptide coverage (Figure S7C), while we could not identify it in testis. In contrast, PIWIL1, and to some extent PIWIL2, both known to be expressed in testis, were easily detectable in testis (Figure 7C). Finally, in the cumulus cells no expression of PIWIL3 could be detected (Figure 7C, S7B). We conclude that PIWIL3 is strongly and specifically expressed in bovine oocytes from GV stage onwards. We note that PIWIL3 might also be expressed during earlier stages of oogenesis.

DISCUSSION

PIWI proteins and piRNAs in mammalian ovaries

In the adult ovaries of three mammalian species we have detected populations of small RNAs that bear resemblance to the PIWIL1-bound, pachytene piRNAs found in the testis of these species. Their transposon-targeting competence is relatively limited, and hence a role outside transposon regulation seems likely. In parallel, we describe expression of PIWIL1, and in human and macaque also of PIWIL2, in oocytes. Based on our data we cannot determine the functions of these ovarian piRNAs, but extrapolating from work performed previously on testicular piRNAs (Aravin et al., 2007; Di Giacomo et al., 2013; Grivna et al., 2006; Hirano et al., 2014; Li et al., 2013; Robine et al., 2009; Gou et al., 2014) we may anticipate diverse roles for these piRNAs in the mammalian ovary. One of these roles could be inhibition of transposon activity. In fact, even though in mouse oocytes siRNAs are the dominating small RNA species, repressive effects of MILI on transposon-derived transcripts have been described (Lim et al., 2013; Watanabe et al., 2008). Nevertheless, no obvious effects of mouse Piwi genes on female fertility have been reported. This contrasts starkly with the strong phenotype of zebrafish Piwi mutant oocytes (Houwing et al., 2008). Possibly, siRNAs have taken over parts of the PIWI pathway in mouse oocytes. Indeed, Dicer is essential for oocyte development in mice (Murchison et al., 2007), while Dicer is non-essential in zebrafish oocytes (Giraldez et al., 2005). Extrapolating from

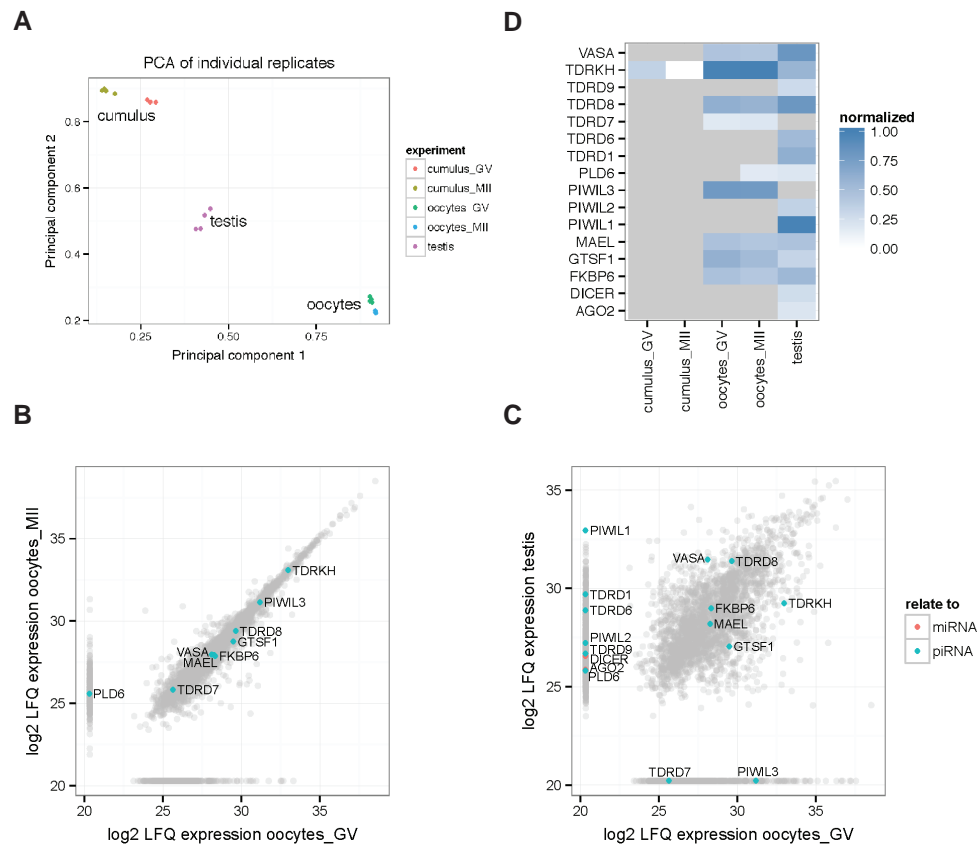


Figure 7. Quantitative mass spectrometry on bovine oocytes and testis. (A) Principle component analysis (PCA) calculated for the LFQ (label-free quantitation) protein expression level measured by mass spectrometry measurements show specific groups for testis, oocytes and cumulus cells. The quadruplicate measurements cluster together according to sample type. (B) Scatterplot comparing the LFQ expression levels of protein groups between GV and MII stage oocytes. Important members of the piRNA pathway are labeled. DICER and AGO1-4 involved in miRNA processing were not detectable in oocytes. (C) Scatterplot comparing the LFQ expression levels of protein groups between GV stage oocytes and testis. Selected piRNA (marked in blue) and miRNA (marked in red) pathway components are labeled. (D) Heatmap showing the normalized relative expression of several piRNA and miRNA related proteins. PIWIL3 expression could only be detected in oocytes, while PIWIL1 and PIWIL2, but not PIWIL3, are expressed in testis. Grey: not detected.

these data, and the data we here report, we would predict that piRNAs are required for proper development of non-Muridae oocytes.

It has been shown that the transcription factor A-MYB drives pachytene piRNA cluster transcription, as well as transcription of core PIWI pathway factors, in mouse testes (Li et al., 2013). Consistent with this finding we observe strong enrichment of A-MYB binding sites at piRNA clusters in all our testis datasets. While A-MYB sites seems to be enriched in ovarian clusters in the various species we analyzed, this enrichment does not reach significance in

most cases (not shown). Interestingly, in the most deeply-sequenced ovarian dataset (macaque oxidized ovary, Figure 2A), A-MYB enrichment is as significant as in testis ($p < 10^{-20}$), suggesting that transcription of the ovarian clusters may also be driven by A-MYB. Possibly, the coverage of the other ovarian datasets is too limited to reach significance.

Finally, we show that the ~26-nucleotide, potentially PIWIL3-bound piRNAs in maturing oocytes (also see below) are not methylated. Hence, since we depend on oxidation to detect piRNAs in ovarian tissue, our set-up would not be able to detect these ~26-nucleotide piRNAs during earlier oogenesis stages. An efficient PIWIL3-specific antibody would help to address this issue.

A secondary piRNA pathway in human fetal ovaries?

In human fetal ovary, starting around week 9, we have uncovered a piRNA population that bears resemblances to the MILI-MIWI2-bound piRNAs from mouse male primordial germ cells (PGCs) (Aravin et al., 2008; Aravin et al., 2007; Kuramochi-Miyagawa et al., 2008). In human fetal ovaries, starting at the second trimester, we could detect PIWIL2, but not PIWIL1 or PIWIL4. Even though we currently lack piRNA sequencing data from immune-precipitates, the size profile, and transposon-coverage of the total piRNA population resembles the piRNA profile from mouse gonocytes closely, suggesting that PIWIL4 may be expressed in these human fetal oocytes. Everything considered, a PIWIL2-PIWIL4-driven piRNA pathway that primarily targets transposons might be active in human fetal oocytes. This should be further tested by more extensive analysis of PIWIL4 expression, and if feasible, immune-precipitation of the Piwi proteins followed by sequencing.

PIWI proteins and piRNAs in mammalian follicular oocytes and early embryos

We find that oocytes of GV and MII stage, and early embryos contain an abundant pool of non-2'O-methylated, ~26-nucleotide-long piRNAs, that is strongly enriched for transposon-derived sequences. Given that PIWIL3 is the dominating Piwi protein in these maturing oocytes, it seems likely that this RNA pool is bound by PIWIL3. While we could not detect PIWIL1 in our mass spectrometric analysis, we did detect a minor fraction of ~30-nucleotide piRNAs in one of our oocyte-data sets, suggesting a relatively low expression of PIWIL1 in these later stage oocytes. However, this particular data set was derived from a sample that also contains macaque testis RNA, and it is possible that these PIWIL1-like piRNAs are derived from non-sequenced parts of the macaque genome. It is presently unclear whether the piRNAs in these oocytes are required for oogenesis, or whether they are made to serve later. As we could not detect PIWIL3 in testis, PIWIL3 might be specialized for oocyte-specific functions or for early embryonic processes. Indeed, we can still detect these piRNAs in 2-4 cell-stage bovine embryos. Given the striking adenylation marks on these piRNAs, and the recent evidence that adenylation plays a role in clearing maternal miRNAs (Lee et al., 2014), it is unlikely that these piRNAs will still be present when the PGCs are specified. Hence, they most likely do not serve to prime the piRNA

3 pathway in the PGCs, as has been described in *Drosophila* (Le Thomas et al., 2014a; Le Thomas et al., 2014b). Possibly, mammalian embryonic piRNAs act during the early embryonic genome-reprogramming steps, much like piRNAs in mouse gonocytes help to prevent transposon activity during genome reprogramming. Alternatively, these piRNAs might execute a continuous regulatory program during both oogenesis and early embryogenesis. Testing such, not mutually exclusive, hypotheses will require the establishment of a model system in which PIWIL3 is naturally expressed and in which it can be manipulated. Given the current power of genome editing techniques such experiments should be feasible.

PIWIL3 loss from Muridae oocytes?

Why have PIWIL3 and its accompanying piRNAs disappeared from mouse oocytes and why do other mammals appear not to have a significant siRNA mediated pathway in their oocytes? These questions can only be speculated about. The fact that the Muridae constellation, loss of PIWIL3 combined with gained expression of DICER in oocytes, is the most restricted one, it seems reasonable to assume that a state of PIWIL3 expression combined with low levels of siRNAs, represents the ancestral state that arose after the emergence of PIWIL3 in a common ancestor of eutherians. It is possible that the acquisition of DICER within the oocyte, driven by the Muridae-specific retro-transposon insertion in the Dicer gene (Flemer et al., 2013), had an adaptive value in maintaining PIWIL3-related functionality during and after the elimination of PIWIL3 function from ancestral populations, thus facilitating the decrease of Piwi-mediated regulation in Muridae oocytes. The fact that only PIWIL3 has been lost from the Muridae, and that PIWIL1, 2 and 4 have been conserved, is in line with our observation that only PIWIL3 specifically functions in oocytes, while the other Piwi paralogs are also expressed in male germ cells. Our findings reinforce the notion that piRNA pathways are extremely flexible on evolutionary time-scales and that we should be very careful to extrapolate from one species to another when it comes to piRNA-related biology. In addition, we should consider an impact of Piwi/piRNAs on human reproductive biology and early embryonic development.

EXPERIMENTAL PROCEDURES

RNA isolation, oxidation and library construction and sequencing

Detailed information about these procedures can be found in the Supplemental Information. Library statistics, and information about which library is used for which figure, are provided in Supplementary Table S1. Sequencing data is accessible at GEO, GSE64942.

Bovine oocytes collection, maturation and fertilization

Whole bovine ovaries and testes were collected from a local slaughterhouse and transported to the laboratory within 2 h after slaughter at room temperature. For oocyte collection, after washing and removal of connective tissue, the ovaries were transferred to 0.9% NaCl supplemented with 1% penicillin/streptomycin (Gibco-BRL) and maintained at 30°C in a water bath. Cumulus

oocyte complexes (COCs) were recovered by aspiration from follicles with a diameter between 2 and 8 mm. Only COCs with an intact cumulus oophorus were used for further experiments.

For the collection of Metaphase I (MI) and MII oocytes, COCS were cultured from 8 and 23 h respectively in M199 medium supplemented with 12.2 mM NaHCO₃ at 39°C, 5% CO₂ in a humidified incubator (van Tol et al., 2008). Cumulus cells were removed by vortexing for 3 min.

Embryos at the 2-4 cell stage were obtained by culturing COCs for 23 h as described after which they were co-incubated with 0.5 x 10⁶ /ml frozen-thawed sperm from a bull of proven fertility. After 20 h, presumed zygotes were denuded by vortexing and were subsequently cultured in synthetic oviductal fluid (SOF) in a humidified incubator at 39°C, 5% CO₂ and 7% O₂. At 10-14 h of culture in SOF, 2-4 cell stage embryos were snap-frozen in liquid nitrogen and stored at -80°C until further use.

Immunohistochemistry of bovine material

Tissues were fixed in 4%PFA in PBS overnight at 4°C. Upon fixation, tissues were embedded in paraffin and sectioned (7µm). Sections were deparaffinised, underwent antigen retrieval in citrate buffer (pH 6.0) and blocked in block buffer (2%sheep serum/2%BSA in PBS) for 1 hour at RT. The samples were incubated in primary antibodies in block buffer overnight at 4°C. Primary antibodies used were rabbit anti-PIWIL1 (1:500, ABIN502092, Antikoerper, Deutschland), rabbit anti-PIWIL2 (1:500, GWB-MU678C, Genway Bio, San Diego) and goat anti-VASA (1:500; AF2030, R&D Systems, Minneapolis, MN). Secondary antibodies used were Alexa Fluor 488 donkey anti-goat IgG (1:500; ab150129) and Alexa Fluor 647 donkey anti-rabbit (1:500; ab150075) (both Abcam). Sections were mounted in Vectashield® with DAPI (Vector laboratories). Image acquisition was performed on a Leica TCS SP5 confocal microscope (Leica, Wetzlar, Germany).

Human ovary collection and immunofluorescence

Cryopreserved adult and fetal ovaries were thawed and washed in Leibovitz medium (Life Technologies, USA) with 2.5% human serum albumin (Sigma, The Netherlands) and 1,5 M; 1,0 M; 0,5 M; 0 M DMSO (Sigma, The Netherlands) respectively at RT for 5 minutes each step. For immunofluorescence, adult and fetal ovaries were fixed in 4% paraformaldehyde overnight at 4°C, followed by dehydration in ethanol and inclusion in paraffin. Sections (5 µm) were deparaffinised, underwent antigen retrieval in citrate buffer (pH 6.0), were blocked with 1% bovine serum albumin (BSA; Sigma-Aldrich) and were incubated 1h at room temperature in primary antibodies diluted in blocking solution, followed by incubation with secondary antibodies. First antibodies used were rabbit anti-PIWIL1 (1:200 fetal; 1:500 adult; ABIN502092, Antikoerper, Deutschland), rabbit anti-PIWIL2 (1:200 fetal; 1:500 adult; GWB-MU678C, Genway Bio, San Diego), rabbit anti-PIWIL4 (1:200; ab87939, Abcam, Cambridge, UK), goat anti-OCT4 (1:100; sc-8628, Santa Cruz Biotechnology Inc., Santa Cruz, CA, USA), goat anti-VASA (1:500; AF2030, R&D Systems, Minneapolis, MN). Secondary antibodies used were Alexa Fluor 488 donkey

3 anti-rabbit IgG (1:500; A-21206) and Alexa Fluor 594 donkey anti-goat (1:500; A-11058) (both Life Technologies, Carlsbad, CA, USA). Nuclei were stained with 4',6-diamidino-2-phenylindole (DAPI; Vector Laboratories, Peterborough, UK) and sections were mounted in Prolong Gold anti-fade reagent (Life technologies). For antibody negative controls, slides were treated with the only omission of primary antibodies. Image acquisition for fetal samples was performed on a Leica DMRA fluorescence microscope (Leica, Wetzlar, Germany) equipped with a CoolSnap HQ2 camera (Photometrics, Tucson, USA). Image acquisition for adult ovary sections was done on a Leica TCS SP5 confocal microscope (Leica, Wetzlar, Germany).

Biological samples

Ovary (Animal-Nr. 9858, 21y old, and 11825, 9y old) and testis (Animal-Nr. 10787, 7y old) tissue samples from *M. fascicularis* were provided by the German Primate Center (Deutsches Primatenzentrum, DPZ), Göttingen. Human adult ovary samples used were from cancer patients that underwent unilateral oophorectomy for fertility preservation and have signed informed consent. The human fetal material used was from elective abortions and donated for research with informed consent. The research on human material was approved by the Medical Ethical Committee of the Leiden University Medical Center (CME P08.087 and CME 05/03 K/YR).

Bioinformatic data processing

Most bioinformatics processing steps were conducted using in-house Perl scripts that are indicated below and available upon request. Raw sequence reads were scanned for 3' adapter sequences (NNNNAGATCGGA). Identical insert sequences with identical 5' and 3' random tags were assumed to represent cloning products of the same molecule and the datasets were collapsed to non-identical cloning products in order to eliminate PCR amplification artifacts (cliplinker_random.pl). Only sequences in the size range of 18-34 nt after adapter clipping were considered for further analysis.

The remaining sequences were mapped to ncRNA datasets from Ensembl database (release 77), miRBase (Kozomara and Griffiths-Jones 2014, release 21), Genomic tRNA database (Chan and Lowe 2009) and Silva rRNA database (Quast et al. 2013, release 119) in order to identify miRNAs or fragments of miRNA precursors, lincRNA, miscRNA, rRNA, snoRNA, snRNA or tRNA using SeqMap (Jiang and Wong, 2008). In addition, low sequence complexity reads ($\geq 75\%$ of the sequence consists of 1-4 nt repeats) were removed from the datasets (kill_simples.pl).

Sequences that were not annotated as (fragments of) known ncRNA, thus representing putative piRNAs, were mapped to the genomes of *Homo sapiens* (assembly: GRCh38, GenBank accession: GCA_000001405.15), *Macaca fascicularis* (assembly: Macaca_fascicularis_5.0, GenBank accession: GCA_000364345.1) and *Bos taurus* (assembly: UMD3.1, GenBank accession: GCA_000003055.3) requiring a perfect match for nucleotides 1-18 but allowing one internal mismatch from nucleotide 19 and up to two non-template 3' nucleotides considering that sRNAs might exhibit modified 3' ends (piRmapper_1.0.pl). The length distribution and

positional nucleotide composition for each length fraction was analyzed for successfully mapped reads only (length_base.pl).

In order to detect ping-pong signatures we analyzed the amount of 5' overlaps of mapped sequences and computed scores on the basis of the product of the read counts of overlapping sequences according to the method applied by Zhang et al., 2011 (ppchecker.pl).

Repetitive sequences in the genomes of *Homo sapiens*, *Macaca fascicularis* and *Bos taurus* were annotated using RepeatMasker (version 4.0.2, RepeatMasker library 20140131) applying the most sensitive settings (cross_match, slow search). Subsequently, the amount, strand polarity, positional nucleotide composition and length distribution of piRNAs related to transposable elements was analyzed by comparing the location of mapped sequences with RepeatMasker annotation (RMvsMAP.pl) taking account of overlapping transposon annotation, piRNAs that span adjacent transposons and multiple genomic hits of piRNAs.

piRNA clusters were predicted using proTRAC (version 2.0.4, command line options: -clstrand 0.5 -pimin 24 -pimax 31 -distr 1-90 -nr -nh -mmr). To obtain a universal set of piRNA clusters for each species we defined piRNA producing loci by conflating proTRAC piRNA cluster annotation for all piRNA containing probes of a species merging piRNA clusters closer than 25kb. piRNA read coverage was calculated for each piRNA producing locus and probe and normalized by the total number of mapped piRNA reads for the probe in question. The results were used to create heatmaps (heatmapper.pl) with colors encoding the relative expression of a piRNA-producing locus across different probes.

We performed chi square tests for an enrichment of A-Myb binding motifs in piRNA clusters comparing the overall genomic abundance of these motifs with their abundance in predicted piRNA clusters. P values represent the right-tailed probability of the chi square distribution (degrees of freedom = 1).

The piRNA read coverage of transposable elements was calculated by mapping piRNA sequence reads to all genomic copies of a certain transposable element. The coverage was then displayed along the respective consensus sequence by analyzing how each genomic copy behaves to the consensus based on the alignments obtained from RepeatMasker runs (trueTEcov.pl).

Mass spectrometry

Samples were boiled in LDS sample buffer (Life Technologies) and separated on a 4-12 percent gradient NOVEX gel (Life Technologies). Each lane was processed independently by in-gel digestions (Butter et al., 2013). Digested peptides were further separated on a 30 cm long reverse-phase capillary (75 μm inner diameter) packed with repositil C18 (1.9 μm , Dr. Maisch). Peptides were eluted within a 4 hour gradient from 5 to 60 percent acetonitrile at 200 nl/min using an Easy LC1000 HPLC system (Thermo) directly mounted to a Q Exactive Plus mass spectrometer (Thermo). The mass spectrometer was operated with a Top10 data-dependent MS/MS acquisition method per full scan. The raw files were processed with MaxQuant (Cox and Mann, 2008) standard settings, except quantitation was solely performed on unique peptides with the ENSEMBL bovine protein database (UMD3.1, 22118 sequences). The data have been

deposited to the ProteomeXchange Consortium (Vizcaino et al., 2014) via the PRIDE partner repository with the dataset identifier PXD001741.

Prior to statistical analysis, peptides mapped to known contaminants, reverse hits and protein groups only identified by side were removed. Only protein groups identified with at least 2, one of them unique, peptides and 2 quantitation events (in minimum one condition) were considered for further data analysis.

We imputed missing values with a normal distribution. We downshifted the mean of the imputed values by 8 (log₂) and applied very small variation (SD=0.00001). For calculating the mean values only with the quantitated values, the missing ones were removed. If there was no quantitative value for the protein group within this condition, the mean was calculated with the imputed values. For PCA analysis the psych package (<http://CRAN.R-project.org/package=psych>) was used with the default settings. Graphics were created with the ggplot2 package (H. Wickham. ggplot2: elegant graphics for data analysis. Springer New York, 2009). The whole analysis was done with R (<http://www.R-project.org/>).

AUTHOR CONTRIBUTIONS

EFR and MM executed and analyzed the experiments. DR performed the bioinformatic analysis. NH, SMCSL and LAJW collected human ovary material and performed IF experiments. CTH optimized small RNA cloning procedures and made libraries. FB performed and analyzed the mass spectrometric analysis. HZ co-designed evolutionary analyses and provided primate material. BAJR provided bovine material and helped design the study. RFK designed the study, interpreted the data and wrote the manuscript. All authors contributed to and approved the final version of the manuscript.

ACKNOWLEDGEMENTS

We thank Mario Dejung for help with proteomics data analysis and Leni van Tol for help with bovine oocyte collection and culture. We thank the sequencing facility at the IMSB, Johannes Gutenberg Universität Mainz for sequencing and Stefan Redl for stimulating discussions. We thank Dr. Franz-Josef Kaup and Dr. Kerstin Mätz-Rensing from the DPZ, Göttingen for providing ovary and testis tissue samples from *M. fascicularis*. This work was supported by a VICI grant (724.011.001) from the Dutch Organization for Scientific Research (RFK), by the Rhineland-Palatinate Forschungsschwerpunkt GeneRED (FB, HZ), by the Iranian Ministry for Science, Research and Technology (42/5/33681) (MM), and by the research funding program MAIFOR of the University Medical Center Mainz (DR).

REFERENCES

1. 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* 328, 1534-1539.
2. 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.
3. 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. *Molecular cell* 31, 785-799.
4. Aravin, A.A., Sachidanandam, R., Girard, A., Fejes-Toth, K., and Hannon, G.J. (2007). Developmentally regulated piRNA clusters implicate MILI in transposon control. *Science* 316, 744-747.
5. 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.
6. Butter, F., Bucerius, F., Michel, M., Cicova, Z., Mann, M., and Janzen, C.J. (2013). Comparative proteomics of two life cycle stages of stable isotope-labeled *Trypanosoma brucei* reveals novel components of the parasite's host adaptation machinery. *Molecular & cellular proteomics* : MCP 12, 172-179.
7. 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. *Developmental cell* 12, 503-514.
8. Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature biotechnology* 26, 1367-1372.
9. Deng, W., and Lin, H. (2002). miwi, a murine homolog of piwi, encodes a cytoplasmic protein essential for spermatogenesis. *Developmental cell* 2, 819-830.
10. Di Giacomo, M., Comazzetto, S., Saini, H., De Fazio, S., Carrieri, C., Morgan, M., Vasiliauskaite, L., Benes, V., Enright, A.J., and O'Carroll, D. (2013). Multiple epigenetic mechanisms and the piRNA pathway enforce LINE1 silencing during adult spermatogenesis. *Molecular cell* 50, 601-608.
11. Ding, X., Guan, H., and Li, H. (2013). Characterization of a piRNA binding protein Miwi in mouse oocytes. *Theriogenology* 79, 610-615 e611.
12. Flemr, M., Malik, R., Franke, V., Nejepinska, J., Sedlacek, R., Vlahovicek, K., and Svoboda, P. (2013). A retrotransposon-driven dicer isoform directs endogenous small interfering RNA production in mouse oocytes. *Cell* 155, 807-816.
13. Ghildiyal, M., and Zamore, P.D. (2009). Small silencing RNAs: an expanding universe. *Nature reviews. Genetics* 10, 94-108.
14. Giraldez, A.J., Cinalli, R.M., Glasner, M.E., Enright, A.J., Thomson, J.M., Baskerville, S., Hammond, S.M., Bartel, D.P., and Schier, A.F. (2005). MicroRNAs regulate brain morphogenesis in zebrafish. *Science* 308, 833-838.
15. 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.
16. Gkoutela, S., Li, Z., Vincent, J.J., Zhang, K.X., Chen, A., Pellegrini, M., and Clark, A.T. (2013). The ontogeny of cKIT+ human primordial germ cells proves to be a resource for human germ line reprogramming, imprint erasure and in vitro differentiation. *Nature cell biology* 15, 113-122.
17. Gou, L.T., Dai, P., Yang, J.H., Xue, Y., Hu, Y.P., Zhou, Y., Kang, J.Y., Wang, X., Li, H., Hua, M.M., et al. (2014). Pachytene piRNAs instruct massive mRNA elimination during late spermiogenesis. *Cell research* 24, 680-700.
18. Graf, A., Krebs, S., Zakhartchenko, V., Schwalb, B., Blum, H., and Wolf, E. (2014). Fine mapping of genome activation in bovine embryos by RNA sequencing.

20. Proceedings of the National Academy of Sciences of the United States of America 111, 4139-4144.
21. Grivna, S.T., Pyhtila, B., and Lin, H. (2006). MIWI associates with translational machinery and PIWI-interacting RNAs (piRNAs) in regulating spermatogenesis. Proceedings of the National Academy of Sciences of the United States of America 103, 13415-13420.
22. 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.
23. Hirano, T., Iwasaki, Y.W., Lin, Z.Y., Imamura, M., Seki, N.M., Sasaki, E., Saito, K., Okano, H., Siomi, M.C., and Siomi, H. (2014). Small RNA profiling and characterization of piRNA clusters in the adult testes of the common marmoset, a model primate. Rna 20, 1223-1237.
24. 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. Current biology : CB 17, 1265-1272. Houwing, S., Berezikov, E., and Ketting, R.F. (2008). Zili is required for germ cell differentiation and meiosis in zebrafish. The EMBO journal 27, 2702-2711.
25. 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.
26. Ipsaro, J.J., Haase, A.D., Knott, S.R., Joshua-Tor, L., and Hannon, G.J. (2012). The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis. Nature 491, 279-283.
27. 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. The EMBO journal 29, 3688-3700. 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 & development 23, 433-438.
28. Kawaoka, S., Izumi, N., Katsuma, S., and Tomari, Y. (2011). 3' end formation of PIWI-interacting RNAs in vitro. Molecular cell 43, 1015-1022.
29. Ketting, R.F. (2011). The many faces of RNAi. Developmental cell 20, 148-161.
30. 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., 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 & development 22, 908-917.
31. 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.
32. Le Thomas, A., Marinov, G.K., and Aravin, A.A. (2014a). A transgenerational process defines piRNA biogenesis in *Drosophila* virilis. Cell reports 8, 1617-1623. Le Thomas, A., Rogers, A.K., Webster, A., Marinov, G.K., Liao, S.E., Perkins, E.M., Hur, J.K., Aravin, A.A., and Toth, K.F. (2013). Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. Genes & development 27, 390-399.
33. Le Thomas, A., Stuwe, E., Li, S., Du, J., Marinov, G., Rozhkov, N., Chen, Y.C., Luo, Y., Sachidanandam, R., Toth, K.F., et al. (2014b). Transgenerationally inherited piRNAs trigger piRNA biogenesis by changing the chromatin of piRNA clusters and inducing precursor processing. Genes & development 28, 1667-1680.
34. Lee, M., Choi, Y., Kim, K., Jin, H., Lim, J., Nguyen, T.A., Yang, J., Jeong, M., Giraldez, A.J., Yang, H., et al. (2014). Adenylation of Maternally Inherited MicroRNAs by Wispy. Molecular cell 56, 696-707.
35. 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. Current biology: CB 15, 1501-1507.
36. Li, X.Z., Roy, C.K., Dong, X., Bolcun-Filas, E., Wang, J., Han, B.W., Xu, J., Moore, M.J., Schimenti, J.C., Weng, Z., et al. (2013). An ancient transcription factor initiates the burst of piRNA production during early meiosis in mouse testes. Molecular cell 50, 67-81.
37. Lim, A.K., Lorthongpanich, C., Chew, T.G., Tan, C.W., Shue, Y.T., Balu, S., Gounko, N., Kuramochi-Miyagawa, S., Matzuk, M.M., Chuma, S., et al. (2013). The nuage mediates retrotransposon silencing in mouse primordial ovarian follicles. Development 140, 3819-3825.
38. Malone, C.D., and Hannon, G.J. (2009). Small RNAs as guardians of the genome. Cell 136, 656-668.
39. Murchison, E.P., Stein, P., Xuan, Z., Pan, H., Zhang, M.Q., Schultz, R.M., and Hannon, G.J. (2007). Critical roles for Dicer in the female germline. Genes & development 21, 682-693.
40. 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. Current biology : CB 19, 2066-2076.
41. Rosenkranz, D., and Zischler, H. (2012). protract-a software for probabilistic piRNA cluster detection, visualization and analysis. BMC bioinformatics 13, 5. Rozhkov, N.V., Hammell, M., and Hannon, G.J. (2013). Multiple roles for Piwi in silencing *Drosophila* transposons. Genes & development 27, 400-412.
42. Saito, K., Sakaguchi, Y., Suzuki, T., Suzuki, T., Siomi, H., and Siomi, M.C. (2007). Pimet, the *Drosophila* homolog of HEN1, mediates 2'-O-methylation of Piwi-interacting RNAs at their 3' ends. Genes & development 21, 1603-1608.
43. Saito, K., and Siomi, M.C. (2010). Small RNA-mediated quiescence of transposable elements in animals. Developmental cell 19, 687-697.
44. Sienski, G., Donertas, D., and Brennecke, J. (2012). Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression. Cell 151, 964-980.
45. 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.
46. 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.
47. van Tol, H.T., van Eerdenburg, F.J., Colenbrander, B., and Roelen, B.A. (2008). Enhancement of Bovine oocyte maturation by leptin is accompanied by an upregulation in mRNA expression of leptin receptor isoforms in cumulus cells. Molecular reproduction and development 75, 578-587.
48. 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.
49. Vizcaino, J.A., Deutsch, E.W., Wang, R., Csordas, A., Reisinger, F., Rios, D., Dianes, J.A., Sun, Z., Farrah, T., Bandeira, N., et al. (2014). ProteomeXchange provides globally coordinated proteomics data submission and dissemination. Nature biotechnology 32, 223-226.
50. Voigt, F., Reuter, M., Kasaruho, A., Schulz, E.C., Pillai, R.S., and Barabas, O. (2012). Crystal structure of the primary piRNA biogenesis factor Zucchini reveals similarity to the bacterial PLD endonuclease Nuc. Rna 18, 2128-2134.
51. Watanabe, T., Totoki, Y., Toyoda, A., Kaneda, M., Kuramochi-Miyagawa, S., Obata, Y., Chiba, H., Kohara, Y., Kono, T., Nakano, T., et al. (2008). Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. Nature 453, 539-543.
52. Weick, E.M., and Miska, E.A. (2014). piRNAs: from biogenesis to function. Development 141, 3458-3471.
53. Xiol, J., Spinelli, P., Laussmann, M.A., Homolka, D., Yang, Z., Cora, E., Coute, Y., Conn, S., Kadlec, J., Sachidanandam, R., et al. (2014). RNA clamping by Vasa assembles a piRNA amplifier complex on transposon transcripts. Cell 157, 1698-1711.

SUPPLEMENTARY FIGURES

3

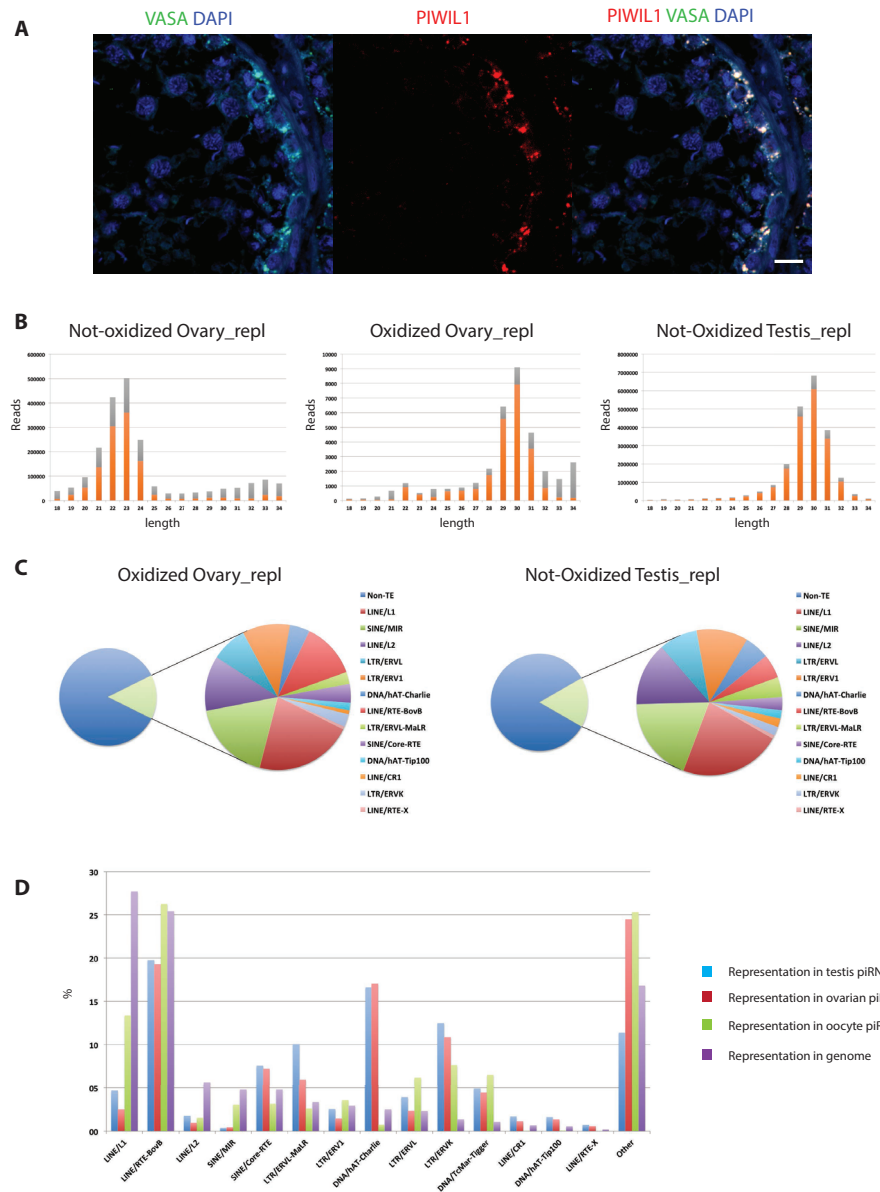


Figure S1. Biological replicate of bovine ovarian and testicular piRNAs. (A) Confocal image of PIWIL1 and VASA staining on bovine testis. Scale bar is 15µm. (B) Length profiles of the indicated libraries. Only non-annotated reads are depicted. Orange: reads starting with a U. Grey: reads not starting with a U. (C) Diagram depicting the transposon content of bovine ovarian and testicular piRNA populations. These represent biological replicates compared to Figure 1. (D) Bar-diagram depicting the representation (in %) of the indicated transposable elements (x-axis) in piRNA pools or in the genome. Representation in piRNA pools reflects % within all transposon-related piRNA species. Representation within the genome reflects the amount of sequence coverage within the total genome.

3

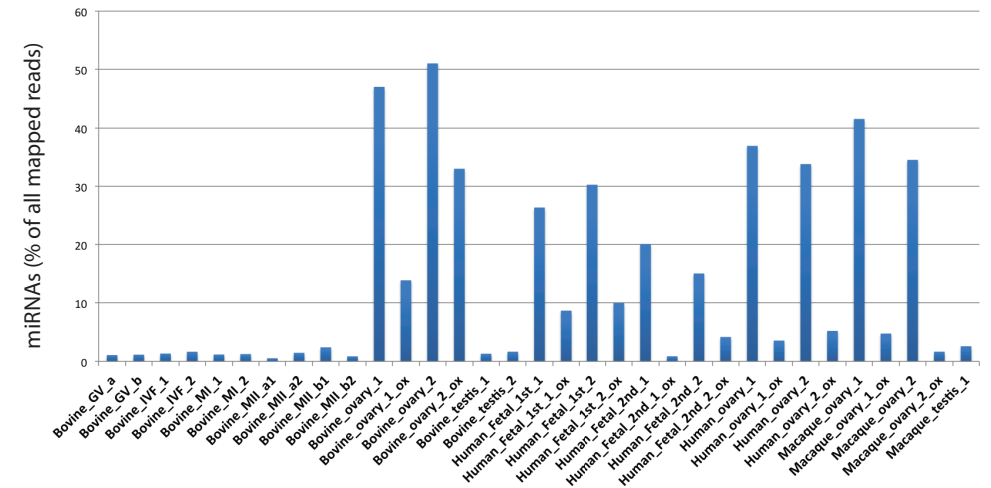


Figure S2. Cloning frequencies of miRNAs in the various indicated libraries. Shown is the percent of miRNAs (mature plus reads matching to precursors) found in the various libraries. Consistently, following oxidation lower miRNA cloning frequencies are observed. The depletion of miRNAs in the 2nd bovine ovary sample is relatively inefficient, possibly due to very low oocyte abundance in the tissue sample.

3

3

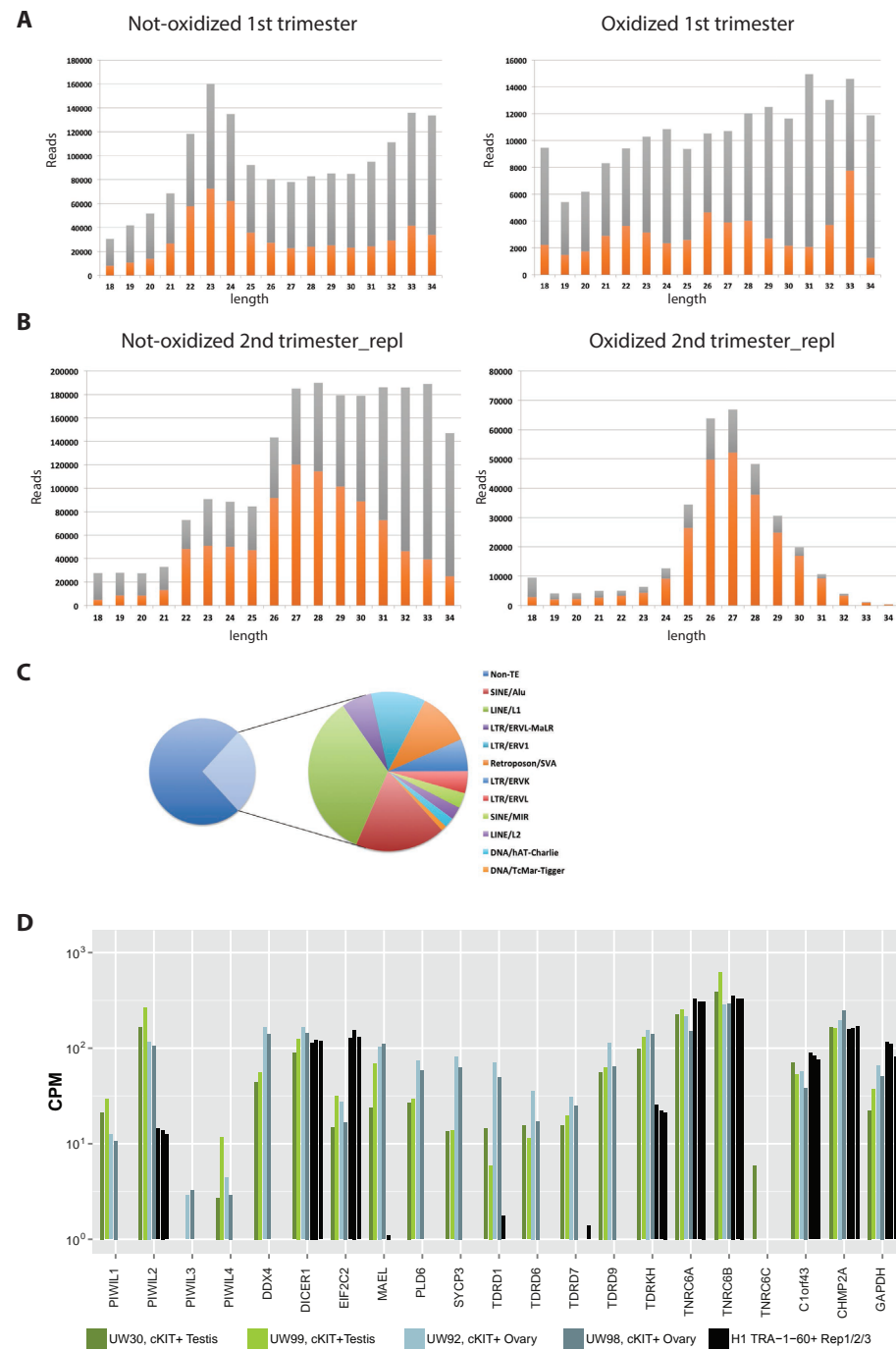


Figure S5. Human fetal ovarian piRNAs. (A) Length profiles of the indicated libraries. Only non-annotated reads are depicted. Orange: reads starting with a U. Grey: reads not starting with a U. (B) Length profiles of the indicated libraries. Only non-annotated reads are depicted. Orange: reads starting with a U. Grey: reads not starting with a U. This represents a biological replicate compared to Figure 4. (C) Diagram depicting

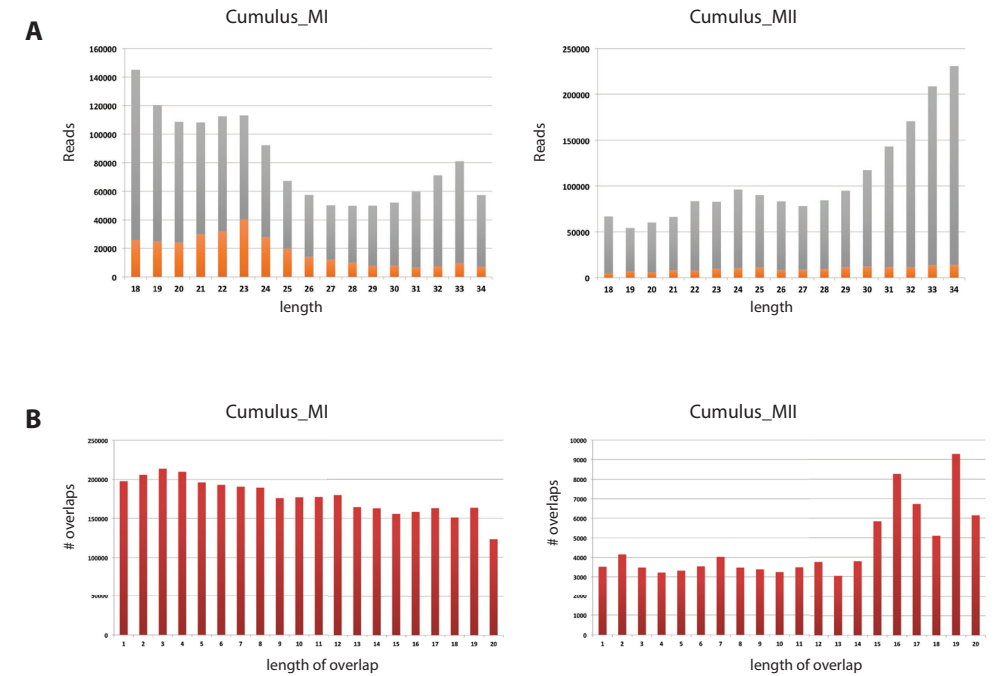
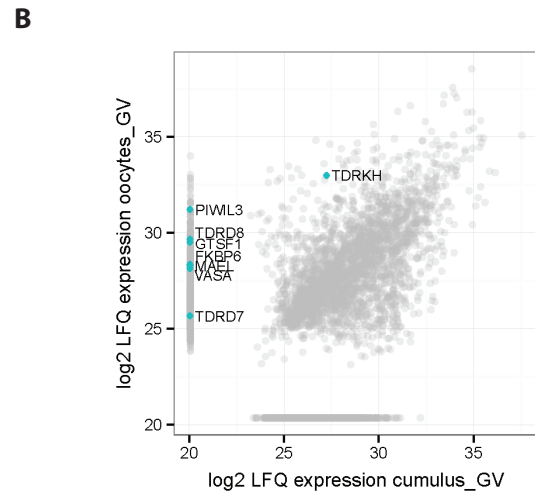
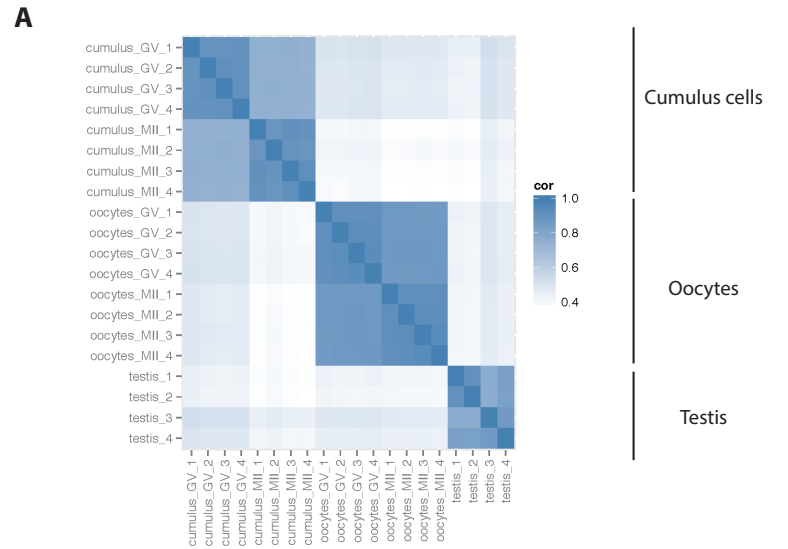


Figure S6. Small RNA profiles of bovine cumulus cells. (A) Length profiles of the indicated libraries. Only non-annotated reads are depicted. Orange: reads starting with a U. Grey: reads not starting with a U. (B) Overlaps of 5' ends of reads that are mapped to opposite strands of the same locus. No ping-pong signal is evident (lack of strong signal at 10 nucleotides overlap).

► the transposon content of human ovarian piRNA populations. This represents a biological replicate compared to Figure 3. (D) Expression of the indicated genes in various fetal cell types. Green: fetal testis. Blue/grey fetal ovary. Dark: hESC. CPM: counts per million.

3

3



C

MTGRARVHARGRRRDTTPPSEPAQPEAARGHPMPVPR TDPPLSFADLPRQIQP
 RLQSLTGVRHVAEAGPPAGAEATSQPSEELKPASESQESLLRKKRDVGRDYRD
 FVVNTRQNLVHVRESTKGTQGSVMMLFSNHVRLKSCSQRHLYKYNVIYTPDIE
 DGRKREALLSELEKLLGNRCIYDGNLPHSLGETKKEVSVKLNKNEPVKITF
 ELSRELQTTSPDCLRYYNILFRKMLEKMDLNQIGRNYYNKKNKTEFNEYRWL
 EIWPYVTSILPYETGLTLCADVSHKLLRMETAYDLISHTREKARGEDAKEYI
 LRKLVGSSVLTKYNNRTYRVDIIWEMSPSSFTFKSDGSEISFVDYYKERYGT
 VVTTLNQPLLITK GKWKKSQRDTPHQPIMLVPELCHLTGLTDDMRKDYRMRD
 LSTHTRMDDPDRRQHKLLTFMDALRKNNTVQKELRDWNLELEEGFLSFSGRTLQ
 DVRIHQGRRMFDNHKADWSKNTREAPLLRAMSLDHWLVIVYTKGNYETALTLLQQ
 NLQRVTPKMGITVRNAKLLEAADTVQSYIR TLEKHASQKTQMVLCCLPTDNKE
 IYDGIKRYLCINYPISQCVLKRITLDRPKTPVTIATKIALQMNCKLGGALWKV
 DIGLQNAFIGIDCFHDIERRRKS VAGFVASIDPDLTRWFSQCIIQESGQELV
 KGLTTCCLYTALKLWSEQNSSPPRSIIIVYRDGVGDGQLQALIDQEVKQMESYLE
 NSYRGQKVR LTFIVVKRINTRFFRQNGKPPFFILKVPFYLKMYDFIVSQS
 VREGTGTVTPHYNV IHD TLYLSPDAVQGLTYRLCLFFFLKGIIRVPAPCHY
 AHKLAYLVGQSLHQQPHELSLSKSLFYL

Figure S7. Quantitative mass spectrometry on bovine germline tissues. (A) Correlation plot of the quantitative mass spectrometry data obtained from bovine oocytes of the indicated stages and bovine testis. This plot accompanies the PCA analysis shown in Figure 7A. (B) Scatterplot comparing protein abundance between GV stage oocytes (stripped of their cumulus cells) and the cumulus cells in isolation. (C) Peptide coverage of bovine PIWIL3 obtained in our mass spectrometric analysis. The two different colors are used to reflect distinct, or partially overlapping identified peptides.

Chapter

4

Tdrd6a regulates the aggregation of
Buc into functional subcellular
compartments that drive germ
cell specification

ABSTRACT

Phase separation represents an important class of subcellular compartmentalization. However, relatively little is known about how the formation or disassembly of such compartments is regulated. In zebrafish, the Balbiani body (Bb) and the germ plasm (Gp) are intimately linked phase-separated structures essential for germ cell specification and home to many germ cell-specific mRNAs and proteins. Throughout development, these structures occur as a single large aggregate (Bb), which disperses throughout oogenesis and upon fertilization accumulates again into relatively large assemblies (Gp). Formation of the Bb requires Bucky ball (Buc), a protein with prion-like properties. We found that the multi-tudor domain-containing protein Tdrd6a interacts with Buc, affecting its mobility and aggregation properties. Importantly, lack of this regulatory interaction leads to significant defects in germ cell development. Our work presents new insights into how prion-like protein aggregations can be regulated and highlights the biological relevance of such regulatory events.

4

Elke F. Roovers^{1,*}, Lucas J.T. Kaaij^{1,*}, Stefan Redl¹, Alfred W. Bronkhorst¹, Kay Wiebrands², António M. de Jesus Domingues¹, Hsin-Yi Huang², Chung-Ting Han^{3,4}, Stephan Riemer⁵, Roland Dosch⁵, Willi Salvenmoser⁶, Dominic Grün^{2,7}, Falk Butter⁸, Alexander van Oudenaarden², René F. Ketting¹

¹ Biology of Non-coding RNA Group, Institute of Molecular Biology, Ackermannweg 4, 55128, Mainz, Germany

² Hubrecht Institute, Royal Netherlands Academy of Arts and Sciences and University Medical Center Utrecht, Uppsalalaan 8, 3584 CT, Utrecht, The Netherlands

³ Genomics Core Facility, Institute of Molecular Biology, Ackermannweg 4, 55128, Mainz, Germany

⁴ CeGaT GmbH, Center for Genomics and Transcriptomics, Paul-Ehrlich-Straße 23, 72076, Tübingen, Germany

⁵ Institute of Developmental Biochemistry, Justus-von-Liebig-Weg 11, 37077, Göttingen, Germany

⁶ Institute of Zoology, Center of Molecular Bioscience, University of Innsbruck, Technikerstraße 25, 6020 Innsbruck, Austria

⁷ Max Planck Institute of Immunology and Epigenetics, Stübeweg 51, 79108, Freiburg, Germany

⁸ Quantitative Proteomics Group, Institute of Molecular Biology, Ackermannweg 4, 55128, Mainz, Germany

*Contributed equally

Adapted from: Roovers, E.F.*, Kaaij, L.J.T.*, Redl, S., Bronkhorst, A.W., Wiebrands, K., de Jesus Domingues, A.M., Huang, H.Y., Han, C.T., Riemer, S., Dosch, R., Salvenmoser, W., Grün, D., Butter, F., van Oudenaarden, A., and Ketting, R.F. (2018). Tdrd6a Regulates the Aggregation of Buc into Functional Subcellular Compartments that Drive Germ Cell Specification. *Dev. Cell.* 46, 285–301.e9.

INTRODUCTION

Phase-separating mechanisms have been acknowledged as important aspects of cell biology. After the initial description of the liquid-like behavior of P-granules, peri-nuclear RNA-rich protein aggregates in the *C. elegans* germline, many other RNA-containing granules have been shown to have similar properties (Brangwynne et al., 2009, 2011; Kroschwald et al., 2015). Important players in the formation of these structures are proteins containing intrinsically disordered regions (IDRs) and/or prion-like domains (PrDs) (Kato et al., 2012; Kroschwald et al., 2015). Such proteins have the propensity to self-aggregate and potentially trigger other proteins to phase separate as well (Prusiner, 1998; Shorter and Lindquist, 2005). In many ways, biologically functional protein assemblies such as P-granules resemble pathogenic protein-aggregation states. It has been suggested that such disease-causing aggregations are an extreme manifestation of an abundantly used mechanism to form membrane-less compartments (Shin and Brangwynne, 2017). This suggests that mechanisms are in place that prevent healthy, functional aggregates to transform into pathological forms.

In many organisms germ cell fate is imposed on cells through the cytoplasmic inheritance of P-granule-like structures, called germ plasm (Gp) (Ikenishi, 1998; Raz, 2003). In zebrafish, Gp originates from an evolutionary conserved electron-dense aggregate in the oocyte, called the Balbiani body (Bb) (Kloc et al., 2004). The mRNAs enriched in the Bb and Gp are often germline-specific and in zebrafish these include *vasa*, *nanos3* and *dazl* (Hashimoto et al., 2004; Kopranner et al., 2001; Yoon et al., 1997). Depletion of single Gp mRNAs can have detrimental effects on primordial germ cell (PGC) numbers, showing that individual Gp components are important for PGC specification and survival (Kopranner et al., 2001; Slaidina and Lehmann, 2017; Tzung et al., 2015; Weidinger et al., 2003).

Bucky ball (Buc) is a protein known to play a key role in the formation of the Bb in zebrafish (Bontems et al., 2009; Marlow and Mullins, 2008). Overexpression of Buc in zygotes revealed that Buc is sufficient to induce ectopic PGCs, suggesting it is also involved in the formation of the Bb-related Gp structure (Bontems et al., 2009). Buc contains a PrD and elegant studies on its homolog in *Xenopus* (Xvelo) have demonstrated that these proteins self-aggregate into membrane-less organelles that display amyloid-like features (Boke et al., 2016).

Core piRNA-pathway components, such as Ziwi in zebrafish and Aub in *Drosophila*, are present in the Gp as well (Harris and Macdonald, 2001; Houwing, 2009). Furthermore, it has been shown in *Drosophila* that piRNA-pathway components inherited via the Gp are essential for transposon silencing in the offspring (Brennecke et al., 2008), and piRNA-mRNA interactions have been proposed to drive mRNA localization to Gp (Barckmann et al., 2015; Vourekas et al., 2016). Many proteins involved in the piRNA-pathway have been identified through genetic and biochemical approaches including multi-Tudor domain-containing proteins (Tdrds) (Siomi et al., 2010). Tdrds play important roles in the formation of nuage, a peri-nuclear protein-RNA aggregate that associates closely with mitochondria. For some Tdrds it has been shown that they bind to symmetrically dimethylated arginine (sDMA) residues on their interaction partners. In zebrafish for instance, the interaction between Tdrd1 and the Piwi protein Zili is mediated via a specific sDMA site in Zili (Huang et al., 2011).

4

One of the Tdrds that has received relatively little attention is Tdrd6, the closest vertebrate homolog to *Drosophila* Tudor (Tud). Tud has been shown to interact with Piwi proteins Aub and Ago3 and plays a role in localization of Aub to Gp and polar granule formation (Kirino et al., 2010; Nishida et al., 2009; Thomson and Lasko, 2004). In mice, TDRD6 plays a role in establishing the chromatoid body, a testis-specific structure that resembles Gp, and the localization of piRNA pathway components to this body (Vasileva et al., 2009). In addition, it is involved in spliceosome assembly in primary spermatocytes (Akpınar et al., 2017). However, a specific molecular function of Tdrd6 or Tud has thus far not been demonstrated.

We show that Tdrd6a is required for coordinated loading of essential Gp components into PGCs through fine-tuning of the aggregating properties and mobility of the Bb organizer Buc. The Tdrd6a-Buc interaction represents one of the few documented cases that demonstrate how the aggregation of a prion-like protein is regulated *in vivo*. We speculate that similar phase separation-regulating mechanisms may act in other cell types as well.

RESULTS

Tdrd6a is gonad-specific and localizes to nuage, the Bb and Gp

The zebrafish genome encodes three Tdrd6 paralogs: *tdrd6a-c*. In this study we focused on *tdrd6a*. Tdrd6a contains seven Tudor domains and is 2117 amino acids in length (Figure S1A). Germline-specific expression of *tdrd6a* was validated by RT-PCR (Figure S1B). Immunohistochemistry (IHC) confirmed that Tdrd6a is expressed in the ovary, where it localizes to nuage (Figure 1A, arrowhead) and to the Bb (Figure 1A, arrow). Tdrd6a is also maternally provided and localizes to the Gp in 4-cell stage embryos (Figure 1B, arrowheads). 24 hours post fertilization (hpf), Tdrd6a is restricted to PGCs, where it again localizes to nuage (Figure 1C, arrowheads). We confirmed the identity of the Tdrd6a-containing structures using established markers for the nuage, Bb and the Gp, using both IHC and localization of transgenic Tdrd6a-mCherry (Figure S1C). These results demonstrate that Tdrd6a is maternally contributed and localizes to three conserved and related structures involved in germline specification and maintenance: the Bb, Gp and nuage.

Identification and characterization of a *tdrd6a* mutant allele

We isolated a *tdrd6a* allele harboring a premature stop codon (Q158X) from an ENU mutagenized library (Wienholds, 2002). Western blot analysis confirmed loss of Tdrd6a in homozygous mutant animals (Figure S1D). *Tdrd6a*^{-/-} oocytes show complete loss of Tdrd6a staining in perinuclear nuage (Figure S1E, arrowhead) and Gp in 4-cell stage embryos (Figure S1F, arrowheads). Some residual staining remains in the Bb in *tdrd6a* mutants (Figure S1E, arrow), however, a strong Tdrd6a-related Bb phenotype (see later) and the presence of a Tdrd6a-mCherry transgene in both nuage and the Bb suggest that this is due to cross reactivity of the antibody in IHC. Homozygous zygotic (Z) and maternal-zygotic (MZ) *tdrd6a* mutants are fertile, indicating that Tdrd6a is not essential for fertility. We conclude that *tdrd6a*^{Q158X} represents a strong loss of function allele.

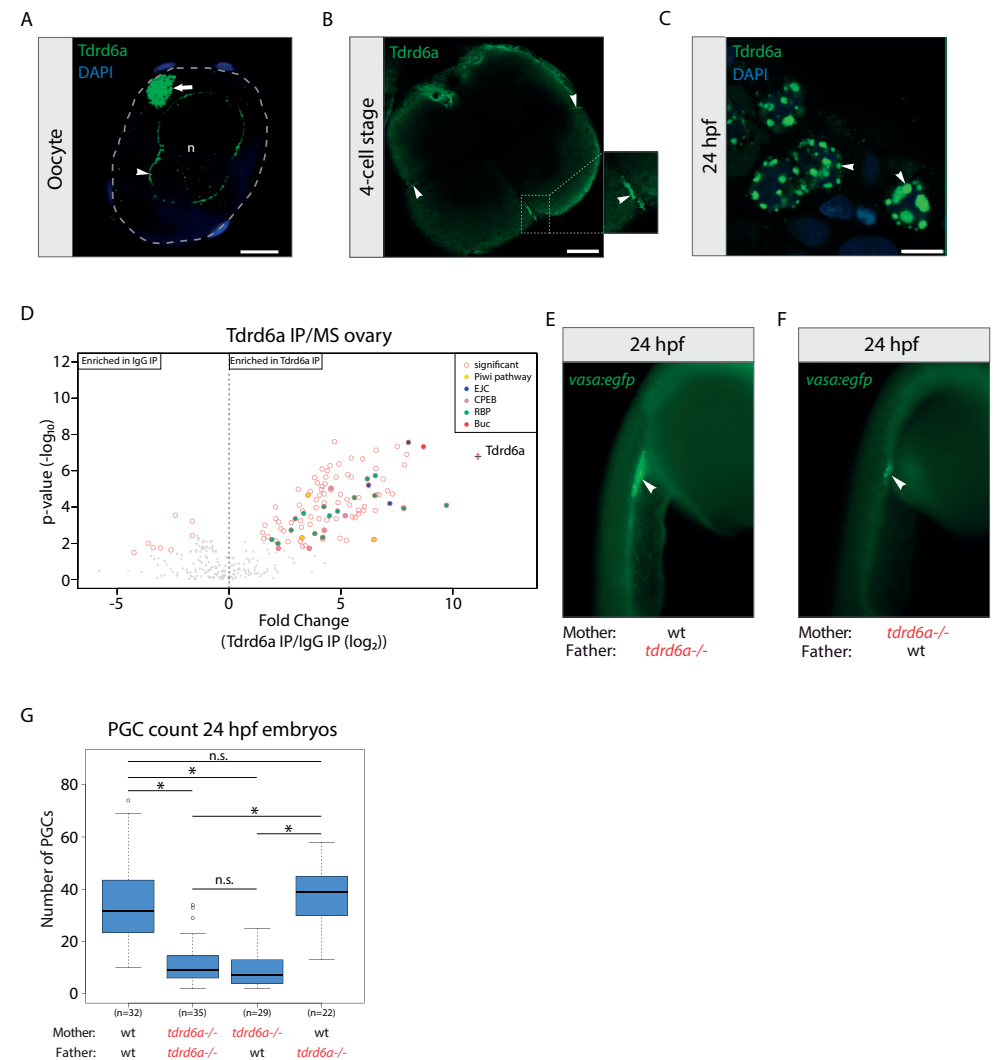


Figure 1. Tdrd6a is germline specific and required for PGC formation. (A) IHC for Tdrd6a in oocytes. Arrowhead and arrow indicate Tdrd6a staining in the nuage and Bb, respectively. Grey dashed line outlines the cell, n = nucleus. Scale bar = 10µm. (B) IHC for Tdrd6a in 4-cell stage embryos. Arrows indicate Tdrd6a localization to the Gp. Scale bar = 100µm. (C) Tdrd6a localizes to perinuclear nuage granules (arrowheads) in PGCs at 24hpf. Scale bar: 7.5 µm. (D) MS of Tdrd6a IPs on ovary, compared to IgG control. (E,F) 24hpf embryos derived from wt (E) or *tdrd6a* mutant mothers (F) in a *vasa:egfp* background. Arrowheads indicate the PGCs. (G) Quantification of PGC numbers in 24hpf embryos from the crosses indicated on the X-axis (* indicates p-value < 0.0001, n.s. = non-significant, calculated by Wilcoxon test).

Tdrd6a does not affect piRNAs

Next, we performed a Tdrd6a immunoprecipitation (IP) on ovary lysates, followed by label-free quantitative mass spectrometry (Figure 1D). Besides Tdrd6a we found strong enrichments

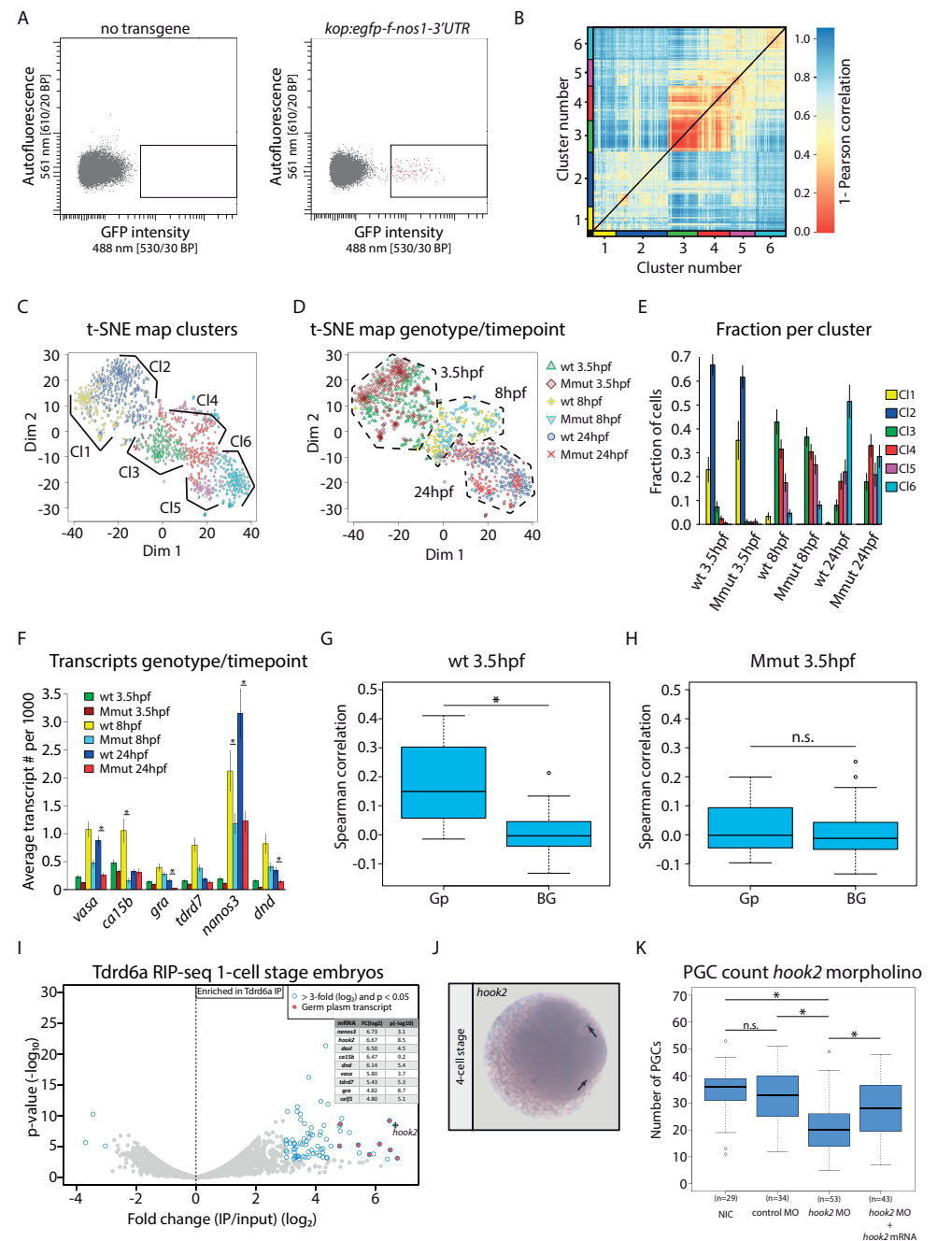


Figure 2. Single cell RNA-seq analysis reveals that maternal Tdrd6a mediates positive correlation of loading of Gp-residing mRNAs into PGCs. (A) Flow cytometry plots of the sort strategy used in this study. Representative FACS plots of embryos 8hpf without or with the *kop:egfp* transgene are shown. Positive events are indicated in red. (B) Heatmap indicating transcriptome distances of ~1100 PGCs computed as 1-Pearson's correlation coefficient. K-medoids clustering identified six clusters, which are color coded on the x and y-axis. (C) Similar as in (B), but now visualized in a t-SNE map. Clusters identified by k-medoids

for several complexes containing RNA-binding proteins (RBPs), including the Exon Junction Complex (EJC) and the CPEB complex. In addition, we identified the Piwi pathway components Ziwi, Zili, and Tdrd7. Finally, we found that Buc was highly enriched.

Given the interaction with Ziwi and Zili, we probed for a role of Tdrd6a in the piRNA-pathway. We first validated the Tdrd6a interaction with Ziwi and Zili (Figure S2A). Despite these interactions, small RNA (smRNA) sequencing of total ovary did not show significant differences between piRNAs of *tdrd6a*^{+/+} and *tdrd6a*^{-/-} animals (Figure S2B-D). We only observe a small but significant reduction in the typical antisense bias for piRNAs mapping to retro-transposons (Figure S2E). When we roughly divided oocytes into early ($\phi < 300 \mu\text{m}$) and later stages ($\phi > 300 \mu\text{m}$), we noticed that this represents a defect in accumulation of antisense piRNAs during early oogenesis only (Figure S2F-S2J). In conclusion, while Tdrd6a associates with Ziwi and Zili, its absence barely affects piRNA populations.

Tdrd6a affects PGC formation

MZ *tdrd6a* mutants have a strong tendency to develop into males. Since the amount of PGCs can have an impact on sex-determination in zebrafish (Tzung et al., 2015), we examined the effect of Tdrd6a on PGC formation. In both wild-type (wt) and MZ *tdrd6a*^{-/-} embryos, PGCs marked by the *vasa:egfp* transgene (Krøvel and Olsen, 2002) are at the genital ridge at 24hpf (Figure 1E and F, arrowhead). However, we observe a significant reduction in PGC number in the offspring from *tdrd6a*^{-/-} females, irrespective of the genotype of the father (Figure 1G).

Tdrd6a affects coordinated loading of Gp mRNAs into PGCs

To learn more about the underlying cause of the PGC defect, we performed single-cell RNA-sequencing (scRNA-seq) on PGCs isolated from embryos spawned by *tdrd6a*^{+/+} (wt) and *tdrd6a*^{-/-} (Mmut) mothers, both crossed with *tdrd6a*^{+/+} males. PGCs were marked using the *kop:egfp-f-nos1-3'UTR* transgene (Blaser et al., 2005) and isolated by fluorescence-activated

clustering are color coded as in (B). (D) t-SNE map highlighting the genotype and developmental timepoint of the individual PGCs as indicated. (E) Barplot displaying the fraction of cells per clusters identified in (B) for the different genotype-developmental time combinations. Error bars were derived from error propagation. (F) Barplot showing the average transcript counts per 1000 transcripts per cell of six Gp transcripts in all six different genotype-developmental timepoint combinations, as indicated. Error bars represent the SEM (* indicates p-value < 0.01, calculated by negative binomial statistics and corrected for multiple testing (*Benjamini-Hochberg*)). (G, H) Boxplots displaying the Gp-Gp and BG-BG correlations in wt and Mmut embryos, respectively (* indicates p-value < 0.001, n.s. = non-significant, calculated by Wilcoxon test). (I) Volcano plot displaying the fold difference between Tdrd6a RIP-seq and input on the x-axis (average of three biological replicates). Y-axis: p-value belonging to the observed differences between Tdrd6a RIP-seq and input. Listed are the values of enriched Gp transcripts. (J) ISH against *hook2* at the 4-cell stage. Arrows indicate Gp. (K) Quantification of PGC numbers observed in embryos in morpholino knock-down (MO KD) injection experiment. NIC = non-injected control, the control MO targets the *fus* transcript, the *hook2* mRNA contained mismatches at the *hook2* MO target site and rescues the KD (* indicates p-value < 0.01, n.s. = non-significant, calculated by Wilcoxon test).

cell sorting (FACS) (Figure 2A). Three time-points were analyzed: 1) when PGCs can be first identified using transgenic GFP expression (3.5hpf), 2) during migration of the PGCs (8hpf) and 3) when the PGCs have reached the genital ridge (24hpf).

Roughly 1100 PGCs were sequenced and analyzed using RaceID2 (Figure S3A) (Grün et al., 2014, 2016; Hashimshony et al., 2012). Representation of the pairwise distances of the single cell transcriptomes in a heatmap revealed two main clusters, which can be further subdivided into clusters 1-2 and clusters 3-6 by k-medoids clustering (Figure 2B). Representation of this data in t-distributed stochastic neighbor embedding (t-SNE) maps (Van Der Maaten and Hinton, 2008) revealed that clusters 1-2 predominantly harbor 3.5hpf old PGCs, whereas clusters 3-6 consist of PGCs from 8hpf and 24hpf (Figure 2C and 2D). Consistent with this, the pluripotency gene *nanog* is selectively expressed in clusters 1-2 (Figure S3B) (Takahashi and Yamanaka, 2006). In contrast, the *rps* gene family, which has been shown to be upregulated after the MZT (Siddiqui et al., 2012), is expressed in cluster 3-6 (Figure S3C). No strong differences between genotypes could be observed for 3.5hpf and 8hpf PGCs (Figure 2E). However, a significant fraction of 24hpf Mmut PGCs was enriched in cluster 4 (Figure 2E and Figure S3D), which is dominated by wt PGCs of 8hpf, suggesting that PGCs lacking maternal Tdrd6a experience developmental delay between 8hpf and 24hpf.

Since individual Gp transcripts can influence PGC numbers (Kopranner et al., 2001; Tzung et al., 2015; Weidinger et al., 2003), we tested if PGCs lacking Tdrd6a generally have lower Gp mRNA levels. Of the 8 known zebrafish Gp transcripts (Hashimoto et al., 2004; Kopranner et al., 2001; Strasser et al., 2008; Wang et al., 2013; Weidinger et al., 2003; Yoon et al., 1997), 6 transcripts passed our filtering criteria (see experimental procedures). While at 8hpf and 24hpf PGCs lacking Tdrd6a indeed tend to have significantly fewer Gp mRNAs than wt, at 3.5hpf no significant difference was found (Figure 2F). In line with this, bulk RNA-seq at the 1-cell stage did not reveal significant effects on mRNA levels (Figure S3E and S3F). Hence, the reduction in PGC number observed upon loss of maternal Tdrd6a most likely is not due to an overall reduction of Gp transcripts provided by the mother.

We then computed all pairwise correlations between the individual Gp mRNAs in wt PGCs at 3.5hpf and compared these to pairwise correlations of non-Gp background (BG) mRNAs (see experimental procedures). This revealed a general positive correlation for Gp mRNAs in wt PGCs (Figure 2G and Figure S3G), indicating that relatively fixed ratios of individual Gp transcripts are loaded into PGCs. Strikingly, in absence of Tdrd6a this positive correlation is completely lost (Figure 2H and Figure S3G). Together, these data show that the stoichiometry of Gp mRNAs in single PGCs is tightly controlled and that this depends on maternally provided Tdrd6a.

Tdrd6a interacts with known Gp mRNAs

Since Tdrd6a is required for correct loading of Gp transcripts into PGCs, we next explored whether Tdrd6a interacts with Gp-residing mRNAs through RNA-IP followed by sequencing (RIP-seq). Strikingly, all known Gp mRNAs were strongly enriched in the Tdrd6a RIP-seq compared to input (Figure 2I). We validated these findings using Tdrd6a RIP-qPCR for the Gp markers *vasa*, *dazl*, and *nanos3*, revealing between fifty- and a hundred-fold enrichment in

the Tdrd6a RIPs (Figure S3H). The mRNA that was most strongly enriched in the RIP-seq was *hook2* (Figure 2I). *Hook2* is an unknown Gp component in zebrafish, but reported to be present in *Xenopus* Gp (Owens et al., 2017). Interestingly, in our scRNA-seq data, *hook2* behaves similar to other Gp markers and also displays the typical Tdrd6a-dependent positive correlation with other Gp transcripts (Figure S3I and S3J). Indeed, *in situ* hybridization (ISH) confirmed presence of *hook2* in zebrafish Gp (Figure 2J). Finally, translation inhibition morpholino (MO) injections revealed that *hook2* affects PGC numbers (Figure 2K), substantiating that *hook2* is a *bona fide* Gp component.

It has been reported that in *Drosophila*, the PIWI protein Aub plays a role in regulating Gp mRNA stability and localization (Barckmann et al., 2015; Vourekas et al., 2016). In analogy, we performed Ziwi RIP-qPCR experiments on 1-cell stage embryos for *vasa*, *dazl* and *nanos3*, using the same experimental conditions used for the Tdrd6a RIP-qPCR experiment. The enrichment values for the tested mRNAs were all below three-fold (Figure S3K), while Western blot confirmed that the IPs were successful (Figure S3L). These enrichment values for Ziwi are in sharp contrast to the values obtained in Tdrd6a RIPs, indicating that Ziwi-mRNP interactions are not very prominent in zebrafish.

Tdrd6a affects Bb organization

We next probed for Bb integrity in the presence and absence of Tdrd6a by doing whole mount fluorescent ISH (FISH) on oocytes against *dazl*. In *tdrd6a* mutant oocytes the Bb often appears to be smaller relative to the entire oocyte, lacking a well-defined edge or even being further distorted. We quantified these defects by classifying the observed structural abnormalities (Figure 3A) and calculating the size ratio between the Bb and the oocyte using various probes (Figure 3B).

We extended these experiments by combining double single-molecule FISH (smFISH) with IHC for Tdrd6a in a Buc-eGFP background (Riemer et al., 2015). In the Bb, Buc-eGFP and Tdrd6a form a continuous structure in which Gp mRNAs are embedded (Figure 3C). SmFISH shows that different Gp transcripts display diverse sub-localization within the Bb. The *dazl* signal is found as a rather compact core in the Bb, whereas *vasa* is found more throughout the entire Bb (Figure 3C). Interestingly, the smFISH signals do not overlap with each other, but rather form transcript-specific networks (Figure 3C, line graph). In *tdrd6a* mutant oocytes, the Buc-eGFP signal is more irregular (Figure 3D). Gp-transcripts still localize to the Bb, indicating that Tdrd6a is not essential for these transcripts to accumulate in the Bb (Figure 3D).

Electron microscopy (EM) revealed that the electron-dense structures in the Bb display a heterogeneous, fibrillary appearance (Figure 3E, S4A, yellow overlays). In contrast, Bbs without Tdrd6a have larger and more homogenous, electron-dense areas than with Tdrd6a (Figure 3F, S3B, yellow overlays). A more widely conserved function of the Bb is mitochondrial selection, which are therefore highly represented in the Bb (Bilinski et al., 2017). We found that mitochondria still accumulate in the Bb in the absence of Tdrd6a (Figure 3E, 3F, S4A and S4B, cyan overlays). In conclusion, Tdrd6a is required for the overall organization of the Bb, even though mRNAs and mitochondria are still present.

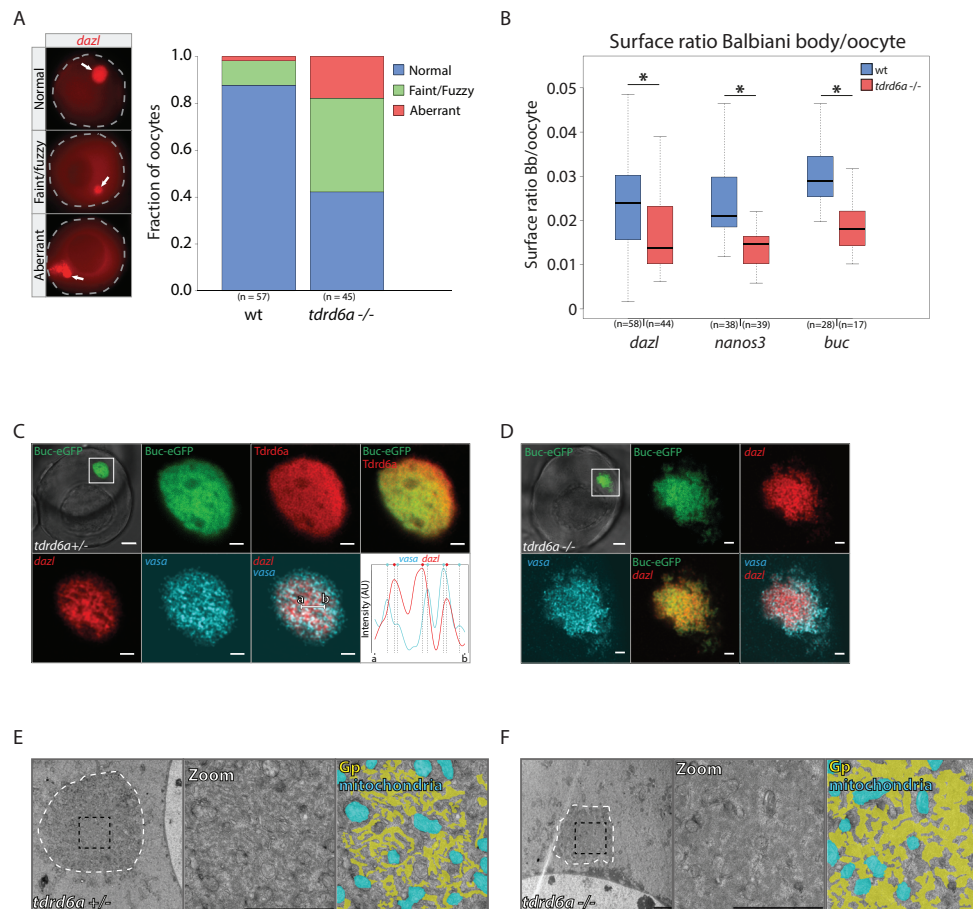


Figure 3. Tdrd6a is required for Bb integrity. (A) Quantification of Bb phenotypes as indicated based on *dazl* FISH on oocytes (examples indicated on the left, arrows indicate Bb). (B) Surface ratio of Bbs in wt versus *tdrd6a* mutant oocytes (* indicates p-value < 0.001, Wilcoxon test). (C, D) Confocal images of Buc-eGFP positive oocytes in *tdrd6a* +/- (C) and *tdrd6a* -/- (D) background. IHC for Tdrd6a and double smFISH was performed and displayed as indicated. *Dazl* and *vasa* signals typically do not overlap, illustrated in the line graph. Intensity for *dazl* (red) and *vasa* (cyan) signals over line a-b (see overlay), with vertical lines indicating fluorescence peaks per smFISH signal (highlighted by colored circle on top) showing transcript peaks are in a separate phase. (E, F) Electron micrographs of Bbs of *tdrd6a* +/- (E) and *tdrd6a* -/- (F) oocytes (white dotted line). The zoom (black square) is shown with (right) and without (middle) overlays that mark the Gp (yellow) and mitochondria (cyan). Scale bars: 10µm (overview A, B), 2µm (zoom) (A, B) and 2µm (E, F).

Tdrd6a is required for merging particles with distinct mRNA content into mature Gp structures in the embryo

In late oogenesis, the Bb disperses into fragments at the vegetal cortex of the oocyte. Upon fertilization, these Buc-containing assemblies accumulate at the cleavage planes to form larger

Gp structures (Riemer et al., 2015). Using smFISH and IHC, we found that in 1-cell stage embryos, Buc and Tdrd6a form isolated particles, decorated with discrete mRNA foci at their periphery (Figure 4A, arrowheads). Buc forms the core of the Gp particles, whereas the Tdrd6a signal is predominantly found at the edge (Figure 4A, arrowheads). As in the Bb, transcript signals do not overlap. At the 2-cell stage, the smaller Buc-Tdrd6a units organize themselves along the cleavage planes and start to cluster together (Figure 4A). The Buc signal often bridges individual granules (Figure 4A, arrows). Furthermore, Tdrd6a appears to be localized around the Buc-assemblies, similar to the mRNA (Figure 4A, arrowheads).

The Gp grows further towards the 4-cell stage into a larger structure, in which the Tdrd6a signal surrounds the Buc signal (Figure 4A). We also observe that in these parts of the Gp transcripts have mostly moved inwards, forming large, intermingled networks (Figure 4A, line graph). Overall, the smFISH signals for different mRNAs are very well mixed within the larger Gp structure, but areas of overall enrichment for one or the other mRNA can still be observed. We note that structures similar to the internal smFISH signal were found using antibody-mediated FISH (Figure S5A), suggesting that the peripheral Tdrd6a signal on Gp does not result from issues related to general antibody penetration into the structure.

In embryos lacking Tdrd6a, mRNAs still associate with Buc particles (Figure 4B), showing that like in the Bb, Tdrd6a is not required for this association. However, without Tdrd6a the Gp structure fails to grow and remains relatively small and highly fragmented (Figure 4C, 4D and S5B). We do observe some apparent fusion of Buc particles, but typically, also in these cases mRNA remains at the periphery (Figure 4B, arrows). These observations lead us to propose that Gp forms through the ongoing accumulation of small granules, containing Buc, Tdrd6a and mRNPs. Tdrd6a contributes to the accumulation of these granules and for the mRNP particles to move into the Buc structure, where they intermingle and form networks with mRNPs of the same kind. We speculate that it is the lack of Gp growth that ultimately results in the above described Gp mRNA defects we see in PGCs lacking Tdrd6a.

Tdrd6a interacts with Buc via symmetrically dimethylated arginines

The IP/MS experiments on ovary extracts identified Buc as a strong interactor of Tdrd6a (Figure 1D). We also found Buc, as well as the close Tdrd6a paralog Tdrd6c, to be among the strongest interactors of Tdrd6a in freshly laid embryos (Figure 5A). We verified the Buc-Tdrd6a interaction on Western blot and show resistance to RNase A treatment (Figure 5B). Tdrds often bind symmetrically dimethylated arginine (sDMA) residues in a binding partner (Siomi et al., 2010). Indeed, analysis of our MS results identified two dimethylated arginine residues within the C-terminus of Buc, residing in a tri-RG (RG(X₀₋₄)RG(X₀₋₄)RG) motif (Figure 5C) (Thandapani et al., 2013). In order to test their relevance for interaction with Tdrd6a, we performed pull-down experiments using biotinylated peptides covering these arginines in either an sDMA- or non-methylated state followed by MS. In the pull-down using the methylated Buc-peptide, Tdrd6a was highly enriched (Figure 5D). Interestingly, another Tdrd6 paralog, Tdrd6c was also among the few enriched proteins. The pull-down with

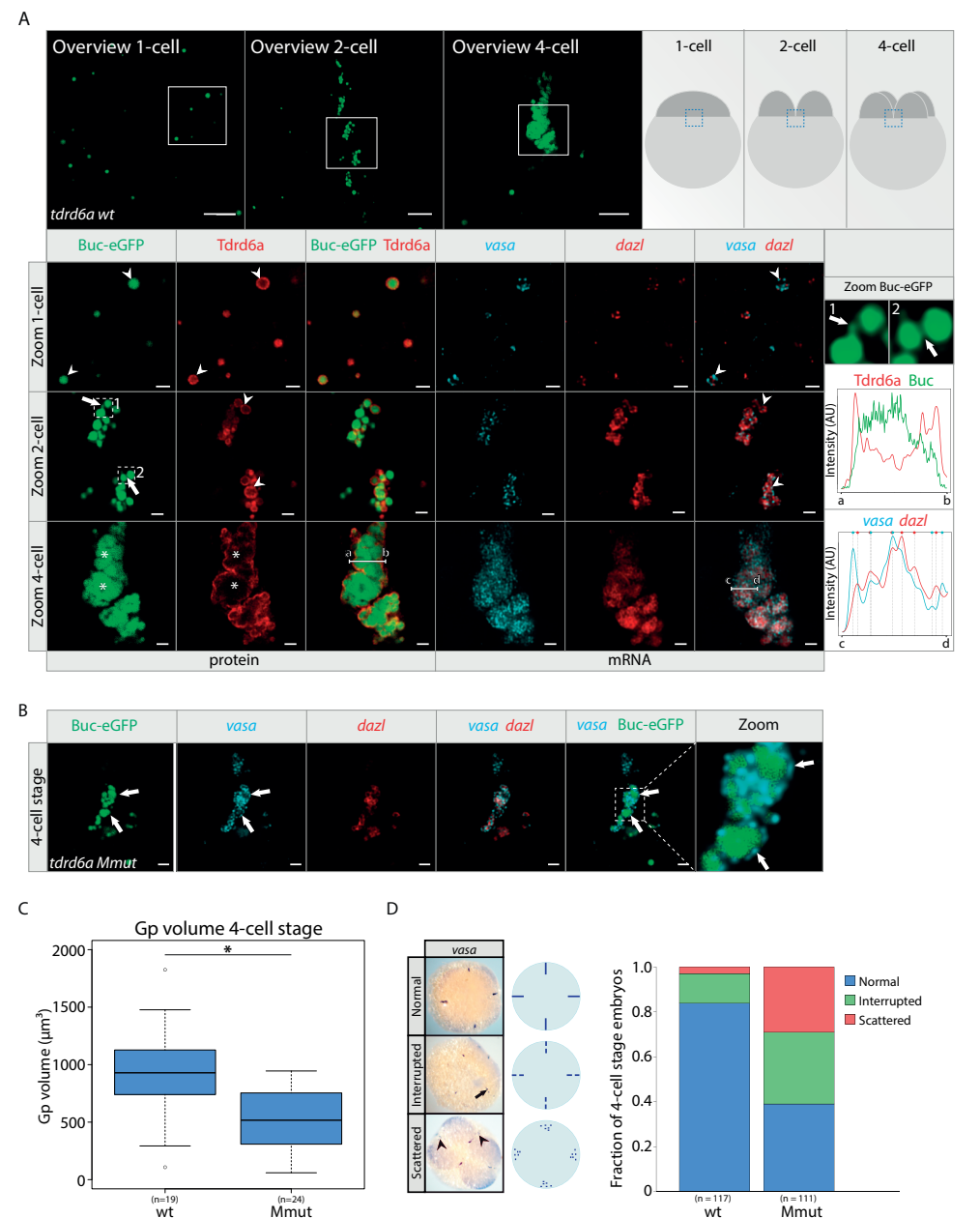


Figure 4. Tdrd6a is required for Gp integrity. (A) Confocal images of Buc-eGFP positive embryos from *tdrd6a +/-* mothers, at 1-, 2- and 4-cell stage, focusing on Gp as schematically indicated. In zoom of the Gp, IHC for Tdrd6a and double smFISH was performed and displayed as indicated. Arrowheads indicate mRNA and Tdrd6a that is peripherally localized on the Gp granule, arrows indicate Buc-eGFP bridges (zoom 1 and 2 on the right). Asterisks mark mature Gp, containing fused Buc-eGFP and mRNA networks inside the structure. Line graphs display intensity for Buc (green) versus Tdrd6a (red) signals over line a-b and *dazl* (red) versus *vasa* (cyan) intensity over line c-d, with vertical lines indicating fluorescence peaks per smFISH signal (highlighted by colored circle on top). (B) Confocal images of Buc-eGFP positive Gp of

the non-methylated peptide showed enrichment for two members of the serine/arginine-rich protein kinase complex Srpk1a and Srpk1b (Figure 5D), confirming that also this pull-down was successful and revealing potential additional post-translational regulation of Buc besides sDMAs. To test the specificity of the Buc peptide for Tdrd6a, we repeated the pull-down using peptides derived from Ziwi, Zili and the analogous C-terminal region of the *Xenopus* Buc homolog X Velo. Clear enrichment of Tdrd6a was found using the methylated Buc peptide, but not using the non-methylated Buc peptide (Figure 5E). We do see some affinity of Tdrd6a for the methylated Zili peptide previously shown to interact with Tdrd1 (Huang et al., 2011). Since Zili is not maternally provided, this affinity could be biologically relevant in ovarian nuage where Zili interacts with Tdrd6a (Figure S2A). Tdrd1 displayed affinity only for the methylated Zili228 peptide, as shown before (Huang et al., 2011), and did not interact with the sDMAs residing in the Buc C-terminus (Figure 5E). Moreover, we found that all three sDMAs on the Buc peptide are required for Tdrd6a interaction (Figure 5F). Interestingly, MS analysis demonstrated that Tdrd6c does bind to both the di- and tri-methylated Buc peptides (Figure 5G). Lastly, Tdrd6a IP/MS from *buc +/-* and *buc -/-* embryos showed that without Buc, Tdrd6c (and also Ziwi) is lost from Tdrd6a IPs (Figure 5H). We conclude that both Tdrd6a and Tdrd6c specifically interact with sDMA-modified Buc and that they may be responsible for recruitment of different protein complexes to Gp.

Tdrd6a affects the aggregation behavior of Buc

It has been shown previously that X Velo is a protein with an N-terminal PrD that has a tendency to self-aggregate. We showed that Buc aggregation, illustrated by detailed Bb and Gp imaging, is highly regulated *in vivo*, and affected by Tdrd6a. We reasoned that Tdrd6a might be involved in spatio-temporal regulation of Buc aggregation. We first tested this in a heterologous cell culture system, using silkworm-derived BmN4 cells. These are of ovarian origin and cultured at 27°C, the same temperature at which zebrafish are kept, thereby mimicking natural conditions for Tdrd6a and Buc. IP/MS experiments on transfected Buc-eGFP revealed the same arginine methylation on Buc as observed in zebrafish (data not shown). Expression of Buc in BmN4 cells results in abundant, cytoplasmic, small granules (Figure 6A). In contrast, Tdrd6a displays a ubiquitous cytoplasmic signal (Figure 6A). Upon co-transfection we observe two possible outcomes: The presence of both Tdrd6a and Buc either results in co-localization in enlarged, cytoplasmic aggregates with a broad variety in size (Figure 6A, middle row and S6A), or in diffuse cytoplasmic localization of both proteins (Figure 6A, bottom row). We then performed consecutive-, rather than co-transfection and quantified protein behavior. If we first transfect

- ▶ a 4-cell stage embryo from a *tdrd6a -/-* mother (Mmut). SmFISH was performed and displayed as indicated. Arrows indicate areas where Buc-eGFP has fused, but mRNA remains peripherally localized. (C) Boxplot representing volumes of the Buc-eGFP signal at the cleavage planes of wt and Mmut 4-cell stage embryos. The largest Gp fragment of each embryo was measured (also see Figure S5B) after 3D reconstruction in Imaris. (* indicates p -value < 0.001, Wilcoxon test). (D) Quantification of Gp phenotypes of 4-cell stage embryos using an ISH against *vasa*. Scale bars: 10 μm (overview A), 2 μm (zoom) (A) and 2 μm (F).

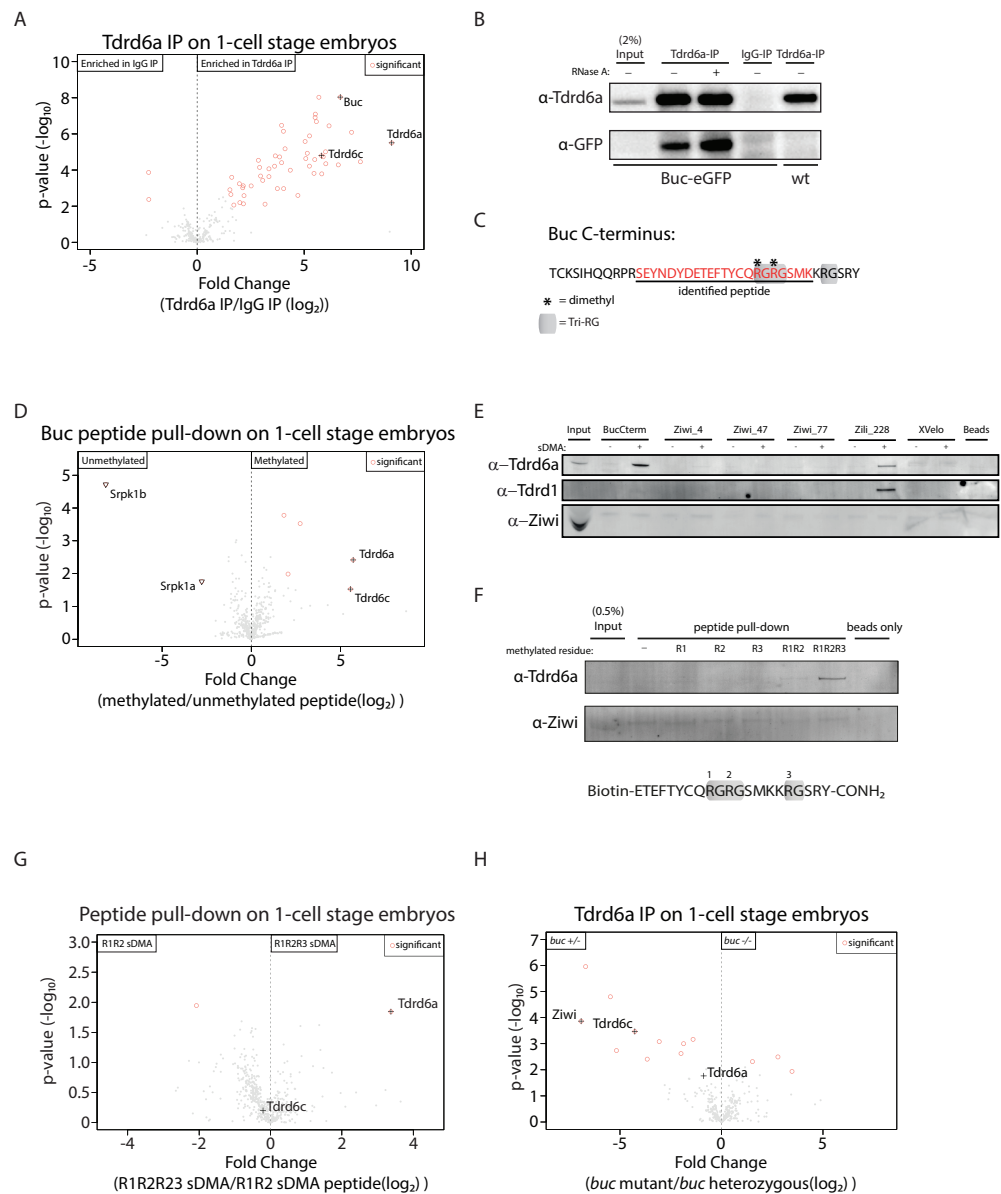


Figure 5. Tdrd6a and Buc interact via sDMAs in the C-terminus of Buc. (A) Volcanoplot of Tdrd6a IP compared to IgG IP on embryo extracts, followed by MS. (B) Confirmation of Tdrd6a co-IP with Buc using the Buc-eGFP transgenic line. (C) C-terminus of Buc with the identified dimethylated peptide underlined. Asterisk indicates residues that were found to be dimethylated by MS. Three RG sites together form a tri-RG motif, indicated in grey. (D) Volcanoplot of peptide pull-down on embryo extracts followed by MS. On the 'Methylated' peptide all 3 RG motifs were symmetrically dimethylated. (E) Peptide pull-down followed by Western for multiple methylated (sDMA) and non-methylated peptides derived from proteins known to contain sDMA modifications and the Buc homolog XVelo on ovary extracts. Listed are all peptides used. (F) Peptide pull-down of Buc C-terminus peptides with different methylation states on embryo extracts. (G) MS of pulldowns of double and triple sDMA modified peptides. (H) MS of Tdrd6a IP in the *buc*^{+/-} background.

Buc, followed by Tdrd6a the next day, we always observed enlarged granules that are positive for both Buc and Tdrd6a (Figure 6B). When the order of transfection is reversed, Buc mostly localizes throughout the cytoplasm (Figure 6C). Only when the Tdrd6a signal is low, Buc seems to be able to form enlarged granules. This mutual effect between Buc and Tdrd6a is specific, since co-transfection of Buc and Tdrd6a with Dcp1, a P-body marker, leaves both proteins unaffected (Figure 6D). These results demonstrate that Tdrd6a can either stimulate the accumulation of Buc into larger granules or prevent its aggregation altogether.

We investigated the properties of these granules in more detail using Fluorescence Recovery After Photobleaching (FRAP). Tdrd6a recovers rapidly upon bleaching of Buc-Tdrd6a double positive granules, reflecting a high mobility in and out of the granule (Figure 6E, n=17). Buc alone only recovers up to ~35% (n=17) of the initial fluorescence intensity. Interestingly, Buc recovery increases to ~55% (n=17) in the presence of Tdrd6a (Figure 6E). Quantification of the FRAP experiments shows that this increase in Buc recovery in the presence of Tdrd6a is significant (Figure S6B). However, we did observe a rather broad distribution in recovery in Tdrd6a-Buc double positive granules (Figure S6B). We hypothesized that this variation in Buc recovery could be due to differences in relative Buc and Tdrd6a concentrations in the granules that were studied. Hence, we normalized the protein amounts in the FRAP experiments by calibrating relative fluorescence using an mCherry-eGFP construct (Figure S6C). This revealed that the more Tdrd6a is present in a granule, the better Buc can recover (Figure 6F). Furthermore, without Tdrd6a Buc-eGFP cannot be detected in the soluble fraction of BmN4 lysates and is predominantly found in the pellet (Figure 6G and 6H). In contrast, in presence of Tdrd6a significant amounts of Buc-eGFP were soluble (Figure 6G). We conclude that Tdrd6a positively stimulates Buc mobility and solubility and that this can contribute to growth of Buc granules.

The tri-RG motif of Buc is required for Bb formation

We then aimed to test the *in vivo* relevance of the FRAP results and the Tdrd6a-Buc interaction data we describe in Figure 5. First, we performed FRAP on the Bb of Buc-eGFP positive oocytes in *tldr6a*^{+/-} and *-/-* backgrounds. These studies showed a remarkable decrease of mobility of Buc within the Bb in the absence of Tdrd6a (Figure 7A). Second, we created a line that expresses a modified version of Buc-eGFP in which we replaced the arginine residues in the C-terminal tri-RG motif by lysines (Buc-RtoK). In the presence of wt Buc, Buc-RtoK can interact with Tdrd6a (Figure 7B), can be incorporated into the Bb and Gp, and is found in PGCs 1dpf (Figure 7C). In contrast, in the absence of wt Buc, Buc-RtoK does not interact with Tdrd6a (Figure 7B) and fails form a Bb in stage Ib oocytes (Figure 7C, S7A). Tdrd6a still localizes to nuage in these oocytes (Figure S7A). Furthermore, in the absence of wt Buc, Buc-RtoK embryos neither have Gp at the 4-cell stage cleavage planes nor form PGCs (Figure 7C). Buc-RtoK alone does form small granules in early stage I oocytes ($\phi < \sim 30\mu\text{m}$), but these are detached from

► compared to *buc*^{-/-} background. Tdrd6c and Ziwi are specifically enriched in the *buc*^{+/-} background, indicating they require the presence of Buc to associate with Tdrd6a.

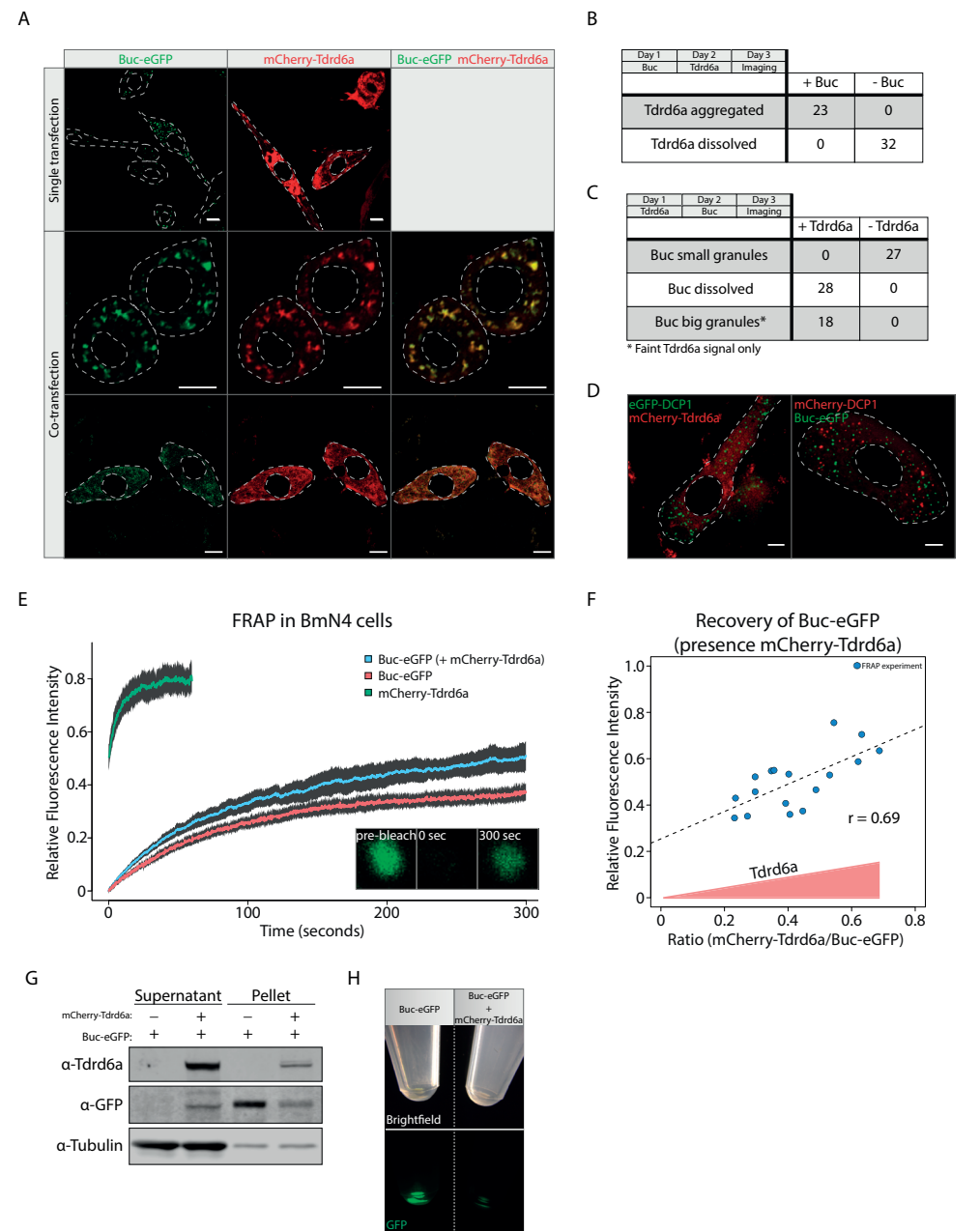


Figure 6. Tdrd6a stimulates Buc mobility in BmN4 cells. (A) Localization of Buc-eGFP and mCherry-Tdrd6a in BmN4 cells in a single transfection (upper panel) and when they are co-transfected (middle and bottom panel). Co-transfected BmN4 cells displaying enlarged Buc-eGFP granules to which mCherry-Tdrd6a co-localizes (middle panel) or dissolved Buc-eGFP (bottom panel). Scale bars: 10 μ m. (B) Quantification of localization of Tdrd6a transfected 1 day after Buc. (C) Quantification of localization of Buc, transfected 1 day after Tdrd6a. (D) Co-transfection of Dcp1-mCherry with Buc-eGFP (left) or Dcp1-eGFP with mCherry-Tdrd6a (right). Scale bars: 5 μ m. (E) FRAP recovery curves of mCherry-Tdrd6a and Buc-eGFP (with or without the presence of mCherry-Tdrd6a as indicated). Fluorescence intensity

the nucleus and never progress to form a Bb (Figure 7D, arrowheads). Despite the absence of a Bb, Buc-RtoK embryos are polarized and are partially viable (Figure 7E, S7B), indicating that Buc-RtoK can partly rescue the *buc* phenotype and that polarization is Bb-independent. However, most embryos do show severe defects in cell division and/or subsequent development (Figure 7E, S7B).

In conclusion, our data show that Tdrd6a and its interaction with arginine-methylated Buc affects the aggregation behavior of Buc-containing structures by stimulating their growth, heterogeneity and mobility (Figure 7F), both in cell culture as well as *in vivo*, and that this is directly relevant for germ cell formation and embryonic development.

DISCUSSION

Proteins like Tdrd6a, with multiple Tudor domains in tandem, are well known to act in germ cells, in particular in small RNA pathways and their organization in peri-nuclear granules. Their precise molecular functions, however, are far from resolved. Other highly abundant components of germ cells are proteins with low-complexity regions and/or prion-like domains, such as Buc. Other examples are MUT-16 and MEG proteins in *C. elegans*, and Xvelo in *Xenopus* which nucleate a variety of subcellular aggregates (Boke et al., 2016; Phillips et al., 2012; Wang et al., 2014). However, insights into how their aggregation behavior is regulated remain scarce. We demonstrate that Tdrd6a regulates the aggregation of Buc. More specifically, it promotes solubility and mobility of Buc, and thereby growth of Buc aggregates into larger structures containing well-determined amounts of germ cell-specifying mRNPs and other Gp components. Various aspects related to our findings will be further discussed here.

Tdrd6a does not affect piRNA generation

Even though the Piwi proteins appear to interact with Tdrd6a in ovary extracts, lack of Tdrd6a does not have an effect on piRNA accumulation. Given the intimate connection between piRNA biogenesis and function, a mechanistic role for Tdrd6a in the piRNA pathway does not seem likely. Our results suggest, however, that Tdrd6c, instead of Tdrd6a, is the more relevant interaction partner for the piRNA pathway. Analysis of *tdrd6c* mutants will be required to clarify this.

Molecular basis behind the PGC phenotype of *tdrd6a* mutants

We observed that Gp arises from the continuous merging of smaller Buc-Tdrd6a granules in embryos. In these granules, Buc is found at the core, while we detected Tdrd6a mainly

- is the calculated fraction of the pre-bleach intensity, and plotted with the 95% confidence interval. (F) FRAP recovery of Buc-eGFP plotted against increasing relative amounts of mCherry-Tdrd6a present in the bleached granule. (G) Western blot for GFP and Tdrd6a on transfected BmN4 cell lysates and corresponding pellets. (H) Pellets of lysates of Buc-EGFP-expressing BmN4 cells in the presence or absence of Tdrd6a as indicated.

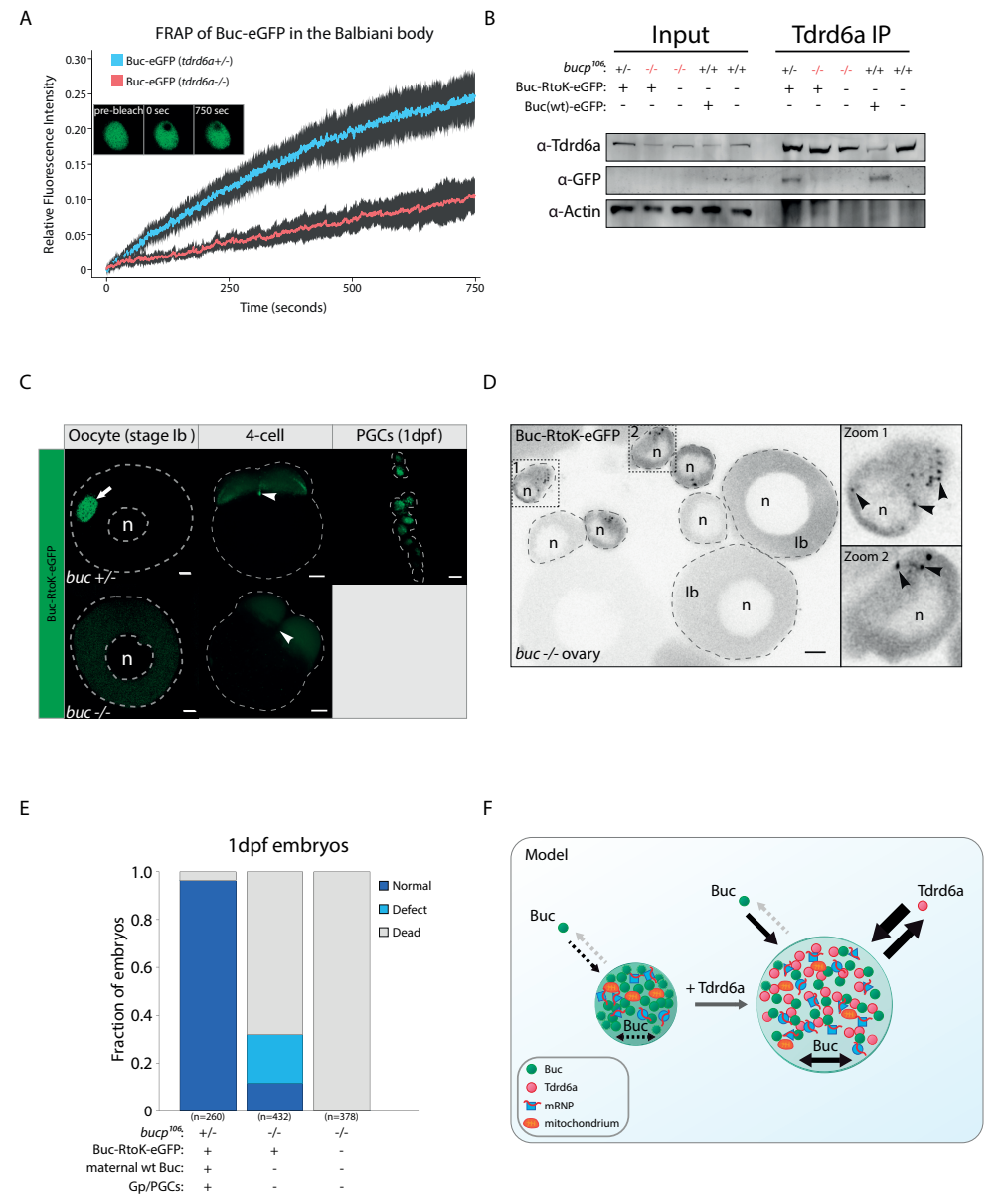


Figure 7. Tdrd6a stimulates Buc-eGFP mobility *in vivo*. (A) FRAP recovery curves of Buc-eGFP in *tdrd6a* heterozygous or mutant Balbiani bodies. Fluorescence intensity is the calculated fraction of the pre-bleach intensity and plotted with a 95% confidence interval. (B) Tdrd6a IPs probed for the indicated proteins by Western blot. *Buc^{p106}* = *buc* loss of function allele. Note that Buc-eGFP is typically very hard to detect in total lysates. (C) Localisation of Buc-RtoK-eGFP in the *buc+/-* and *buc-/-* background. Arrow indicates Bb, arrowheads indicate Gp (*buc+/-*), or where Gp should be (*buc-/-*). Scale bars for oocyte and 1dpf: 10 μ m. Scale bar for 4-cell: 100 μ m. n=nucleus. (D) Overview of *buc-/-* ovary (whole mount) positive for Buc-RtoK-eGFP. Zoom 1 and 2 are examples of stage I oocytes $\phi < \sim 30\mu$ m, containing small Buc-RtoK positive granules (arrowheads). These granules are never detected in stage Ib oocytes, where Buc-RtoK

at the periphery. We note, however, that at present, we cannot be certain that this apparent substructure is real, or whether it represents an experimental artefact due to very high local Tdrd6a concentrations surrounding the Gp structure.

The individual Buc-Tdrd6a granules in 1-2 cell embryos display discrete mRNA foci at their circumference. Interestingly, these foci move more internally and start to form networks when larger assemblies arise. Homotypic assemblies of mRNPs have been described recently in *Drosophila*, where it has been demonstrated that Gp mRNAs initially form homogenous mRNP granules, followed by fusion into heterogeneous mRNP aggregates, in which the quantities of Gp mRNAs positively correlate (Little et al., 2015; Trcek et al., 2015). In zebrafish, we could infer that mRNA quantities in mature Gp positively correlate as well, and that Tdrd6a is required for this.

Why does this correlation between Gp mRNAs depend on Tdrd6a? In absence of Tdrd6a mostly small, incomplete Gp-like structures are found. Given that each Tdrd6a-Buc granule at the 1-2 cell stage only carries a limited number of individual mRNPs, sufficiently large numbers of Buc-Tdrd6a granules need to accumulate to attain the ratios as found on all granules combined. Without Tdrd6a, these numbers may not be reached. Since single Gp mRNAs can have a strong impact on of PGC formation, such unstable ratios may directly relate to the observed PGC specification and/or maintenance defects.

mRNP recruitment and organization in Gp

Single Buc-containing particles contain individual foci of various Gp transcripts at their periphery. We show that this does not require Tdrd6a. How then are these transcripts recruited? Intrinsic properties, such a primary sequence or secondary structure of the mRNPs may play a role (Knaut et al., 2002; Koprunner et al., 2001; Trcek et al., 2015). Furthermore, the fact that we identify the cytoplasmic EJC complex and the CPEB complex in our Tdrd6a interactome may reveal an additional aspect: mRNPs that have not undergone translation could be prone to be incorporated into Gp-related structures. Indeed, these complexes have been demonstrated to play a role in translational control and/or Gp transcript localization in *Drosophila* and *Xenopus* (Hachet and Ephrussi, 2004; Minshall et al., 2007; Nelson et al., 2004).

In more enlarged Gp structures, we observe bigger transcript networks, each consisting of single types of mRNA that spread throughout the Gp. Possibly, intrinsic properties of mRNPs trigger such network formation when local concentrations are sufficiently high. Indeed, intrinsic tendency of transcripts of the same kind to cluster is a phenomenon that has been suggested previously in *Drosophila* (Little et al., 2015; Trcek et al., 2015). These larger homotypic structures may subsequently be further stabilized by their continued interaction with the growing Buc

► is diffusely cytoplasmic. Scale bar: 10 μ m, n = nucleus. (E) Quantification of progeny viability at 1dpf spawned by mothers with background as indicated, crossed with wt males. (F) Model of Buc-containing granules, with or without Tdrd6a. Arrows indicate movement in and out of the structure or mobility within the structure itself.

containing structure, in which they intermingle with other homotypic networks. Our data show that Tdrd6a is required for this higher level organization of Gp mRNPs. Whether this results from its effect on Buc or on mRNPs directly cannot be distinguished at present.

How does Tdrd6a regulate Buc aggregation?

We describe potentially contradicting effects of Tdrd6a on Buc behavior: On the one hand Tdrd6a promotes Buc solubility/mobility and on the other hand Tdrd6a drives the formation of larger Buc aggregates. Based on our cell culture experiments, we speculate that the effect of Tdrd6a on Buc may critically depend on relative and absolute concentrations of both proteins. Possibly, the multi-tudor domain organization of Tdrd6a/c allows it to increase local Buc concentrations, and hence Buc aggregation behavior. But at the same time, the high mobility of Tdrd6a, and possibly also Tdrd6c, may drive constant remodeling of Buc aggregates, and prevent the formation of too rigid, or too much condensed Buc aggregates and allow fusion and/or growth of Buc aggregates. To address these possibilities, *in vitro* systems will need to be established, such that protein-protein interactions and aggregation behavior can be studied in much greater detail.

Regardless of the exact mechanisms, our work reveals that PTMs can play an important role in how aggregations are regulated: loss of Buc arginine methylation, and hence Tdrd6a interaction, severely affects Buc behavior *in vivo*. In fact, the RtoK mutations in Buc result in a much more severe phenotype than that observed in *tdrd6a* mutants. We consider it likely that this is caused by additional loss of Tdrd6c interactions with Buc. Possibly, additional PTMs, and their dynamics are involved in the complex aggregation behavior that Buc displays *in vivo*, but also in other scenarios. For instance, during early embryogenesis in *C. elegans*, phosphorylation and dephosphorylation of MEG-1/3 control P-granule disassembly and assembly, respectively (Wang et al., 2014). Since kinases were identified in the non-modified Buc peptide pulldown it is tempting to speculate that besides arginine dimethylation, phosphorylation may also regulate Buc aggregation dynamics.

The Bb is not required for generating oocyte polarity

We found that Buc-RtoK can rescue the oocyte polarization defect of *buc* mutants, even though a Bb never forms. This shows that the Bb as such is not essential for oocyte polarity establishment, and that Buc may have a Bb-independent role that helps to maintain or establish polarity. Nevertheless, we did observe that Buc-RtoK could not fully rescue the loss of endogenous Buc, because many embryos did not develop properly. This is unlikely to be due to tag-interference, since wt transgenic Buc that also carries a GFP tag at its C-terminus can fully rescue (Riemer et al., 2015). Therefore, Buc, and/or the Bb may play important roles downstream of polarity establishment in the oocyte as well.

Buc-Tdrd6a interaction as a model for regulated protein aggregation

Phase separation of proteins with IDRs has been recognized as a research field of major importance. It represents a pivotal type of compartmentalization, which mediates diverse

cellular processes. There appears to be a wide range of aggregation states, spanning from liquid-like droplets to almost solid aggregations (Brangwynne et al., 2011; Patel et al., 2015; Shin and Brangwynne, 2017). It has been proposed that pathogenic protein aggregates, such as found in Alzheimer's disease or ALS, represent a detrimental state of normally occurring protein aggregation. Hence, knowledge about how aggregation states can be regulated *in vivo* will be directly relevant to the understanding of these types of disease. Since Buc aggregation is very dynamic during zebrafish oogenesis and embryogenesis it represents a powerful model to study the spatio-temporal regulation of protein aggregation, both by trans-acting factors as well as PTMs.

Analogous to previous studies, Buc typically behaves like a 'scaffold': recovering slowly and only partially. Tdrd6a recovery is typical for a granule 'client', displaying rapid, near complete recovery, indicating high mobility in and out of the Buc-aggregate (Woodruff et al., 2017). This may mean that in other scenarios in which Tdrd6a-like proteins have been described to affect aggregations, such as for example the chromatoid body in mammalian spermatocytes or perinuclear nuage, scaffold proteins such as Buc are still to be discovered. Alternatively, well-known proteins may in fact act as such scaffolds. For instance, Piwi proteins typically have rather long and seemingly unstructured N-terminal tails, and also other well-studied proteins, like Vasa, contain disordered regions and can phase separate *in vitro* (Nott et al., 2015). Indeed, these proteins are rich in RG motifs that could be sites of aggregation modulation by Tdrd proteins.

ACKNOWLEDGEMENTS

We thank the members of our laboratory for fruitful discussions. Yasmin El Sherif is thanked for extensive experimental support. We thank the Cuppen group (Hubrecht) for identifying the *tdrd6a*^{Q158X} allele. We would like to thank Jeroen Krijgsveld and the proteomics core facility at EMBL for performing initial proteomics analysis and Eugene Berezikov (Hubrecht/ERIBA) for initial bioinformatics support. We thank the following IMB Core Facilities for their contributions and valuable services: Genomics, Microscopy, Bioinformatics, Flow Cytometry, and the Media Lab. In particular, we thank Mária Hanulová for support in FRAP experiments. We thank Edward Lemke for critical reading of the manuscript. This work was supported by the Rhineland Palatinate Forschungsschwerpunkt GeneRED, a Marie Curie fellowship 623119 (LJTK), the ERC (ERC-StG202819, RFK, and ERC-AdG294325-GeneNoiseControl, AvO) and Vici awards (RFK, AvO) from the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO).

AUTHOR CONTRIBUTIONS

EFR, LJTK and RFK conceived the study and designed experiments. LJTK performed scRNA-seq experiments and data analysis, RIP-seq analysis and some of the IP experiments. EFR performed most other experiments. S Redl performed electron microscopy experiments. AWB performed BmN4 cell culture experiments. AMJD performed smRNA-seq analysis. HH performed initial analyses of the *tdrd6a* mutant and Tdrd6a localization. SR and RD provided the wt Buc-eGFP construct and transgenic line. KW performed scRNA-seq experiments and DG and AvO devised

scRNA-seq data analyses. CTH optimized RIP-seq protocols and FB performed MS analysis. RFK supervised the project. LJTK, EFR and RFK wrote the paper with input from all authors.

MATERIALS AND METHODS

Zebrafish strains

Zebrafish strains were kept under standard conditions. The *tdrd6a*^{Q185X/+} mutant allele zebrafish was derived from ENU mutagenized libraries using target-selected mutagenesis as described before (Wienholds, 2002). Animals carrying *tdrd6a*^{Q185X/+} were out crossed against wt fish (TL), *kop:egfp-f-nos1-3'UTR*, *vasa:egfp* or *buc:buc-egfp-buc3'UTR* transgenic fish (Kopranner et al., 2001; Krovel and Olsen, 2004; Riemer et al., 2015) and subsequently incrossed to obtain *tdrd6a*^{Q185X/Q185X} 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 CG-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'-CAG TGT ACA ATT TCT TTG CAA AGC CCA TT -3'. The lesion induces a truncation after amino acid Q185. This residue precedes the epitope used for immunization. The *buc*^{p106} allele was amplified with *buc*p106seq_F: TCT CCC CAA AGG GAG AAC TCC ATT G and *buc*p106seq_R: GTT TAA CAT TTT AAA CTG CTC AAC ATA CCT CTG and sequenced with the reverse oligo.

Whole mount smFISH and IHC

PFA fixed oocytes/embryos were collected as described above and were incubated overnight with 1:100 anti-Tdrd6a and 1:100 of both smFISH probes (Stellaris™ custom design, Quasar 570 (*vasa*) and Quasar 670 (*dazl*) labelled) stocks (12.5μM in TE buffer) in hybridization buffer (10% dextran sulfate, 10% formamide, 1mg/mL tRNA, 0.02%BSA, 2mM vanadyl-ribonucleoside complex (NEB S1402S) in 2xSSC) at 30°C. Next day, wash 15 minutes in wash buffer (10% formamide, 2xSSC) and incubate in 1:500 anti-rabbit alexa-405 for 30 minutes in wash buffer. Then wash 2 x 15 minutes in wash buffer and mount in ProLong™ Gold Antifade Mountant (ThermoFisher).

Peptide pulldown

Peptides were synthesized by Peptide Specialty Laboratories GmbH. 20μg peptide in 500μL IP buffer (25mM Tris pH 7.5, 150mM NaCl, 1.5mM MgCl₂, 1% Triton-X100, 1mM DTT) was pre-incubated with streptavidin-coupled magnetic beads for 30 minutes at RT, rotating. 20μL resin (ThermoFisher 65001) was used per pulldown. Then, the respective lysate was added to the washed beads and incubated for 1hr at 4°C while rotating. The beads were then washed with wash buffer (25mM Tris pH 7.5, 300mM NaCl, 1.5mM MgCl₂, 1mM DTT) and either used for Western blot analysis or MS.

FRAP

FRAP was performed on a TCS SP5 Leica confocal microscope, equipped with a FRAP-booster, using a 63x oil objective with an NA of 1.4 (BmN4 cells) or a 63x water objective with an NA of 1.2 (Bb). In BmN4 cells, entire granules were bleached in a fixed region of 1.5μm ø and recovery was followed for 1500 frames (0.2s/frame). Bbs were bleached partially in a fixed region of 2.5μm ø and recovery was followed for 1500 frames (0.5s/frame). Regions (pre- and post-bleach) were tracked using TrackMate (Tinevez et al., 2017). 10 pre-bleach frames were recorded and after background subtraction, the average intensity was used as pre-bleach intensity. Post-bleach frames were background subtracted and to make replicates comparable, post-bleach frame #1 of each measurement was set to 0 and corresponding pre-bleach intensity was corrected for this. Normalization of Buc-eGFP and mCherry-Tdrd6a was performed by plotting intensities of mCherry-eGFP using the same microscope settings as for the FRAP. Intensities were plotted after background subtraction and the resulting curve was used to calculate protein ratios using the initial/pre-bleach intensities of each experiment.

Immunohistochemistry

IHC was performed as described before (Houwing et al., 2007). Tdrd6a antibodies were raised in rabbits with the synthetic peptide H₂N-QAVVHEPESEKEKRD-CONH₂. Antisera were subsequently purified against the synthetic peptide (Eurogentec). Embryos and adult gonad tissue were fixed in 4% PFA/PBS at RT for 3 hours. Adult tissue was embedded in paraffin and sectioned prior to the staining. Tdrd6a antibody was used 1:100, rat-anti-Ziwi was used 1:100. Anti-rabbit alexa-647 (Abcam, ab150075) and anti-rat alexa-488 (Abcam ab150153) was used 1:500.

Confocal Imaging

Samples were imaged using a TCS SP5 Leica confocal microscope using a 10x dry objective (NA 0.3), 40x oil (NA 1.3), 63x oil (NA of 1.4) or a 63x water objective (NA 1.2). The following figures were deconvolved using the Huygens software: Figure 3C, 3D, 4A, 4B, 6A, 6D, 7D and Supplemental Figure 1C, 4A and 6A.

Immunoprecipitation

Immunoprecipitations were performed as described before (Huang et al., 2011). Tdrd6a antibody was used at a dilution of 1 to 100. For pulldown of the GFP-tagged constructs, 25μL GFP-Trap (GFP-Trap_A, Chromotek) per lysate was used. For q-RT-PCR and RIP-seq analysis, beads were eluted in TRIzol LS (ThermoFisher) for RNA isolation. For western blot and mass spectrometry analysis, beads were eluted in crack buffer and NuPAGE LDS sample buffer (ThermoFisher) respectively.

Western blot

Samples were heated to 95°C for 5 minutes prior to loading on a 4%-12% NuPage NOVEX gradient gel (ThermoFisher) and blotted on an Immobilon-FL PVDF membrane (Merck) overnight at 15V, RT. The membranes were incubated next day with primary antibodies (Rb- α -Tdrd6a 1:1000, Rt- α -Ziwi 1:1000, Rb- α -Tdrd1 1:500, Rb- α -Zili 1:10,000, Ms- α -GFP (Santa Cruz) 1:1000) and upon washing incubated with 1:10,000 800CW IRDye α -Rb and IRDye 680RD α -Rt (LI-COR) and imaged on an Odyssey CLx imaging system (LI-COR) and quantified with Image Studio.

Library construction and high-throughput sequencing

Total RNA was subjected to 15% TBE-urea gel for size selection of 15– 35 nt. This excised gel fraction was eluted in 0.3 M NaCl for N16 h and precipitated with 100% isopropanol and Glycoblue for 1h at –20 °C. The precipitated RNA pellet was washed once with 75% ethanol and dissolved in nuclease-free water. The purified RNA fraction was confirmed by Bioanalyzer Small RNA assay (Agilent). Library preparation was based on the NEBNext® Small RNA Library Prep Set for Illumina® (New England Biolabs) with minor modifications. To counteract ligation bias and to remove PCR duplicates, small RNA was first ligated to the 3' adapter and then the 5' adapter, both of which contained four random bases at the 5' and 3' end, respectively. Adapters with random bases were chemically synthesized by Bioo Scientific. Adapter-ligated RNA was reverse-transcribed and PCR amplified for 14 cycles using index primers. The PCR amplified cDNA construct was checked on the Bioanalyzer (Agilent) using High Sensitivity DNA assay. We performed a size selection of the small RNA library on LabChip XT instrument (PerkinElmer) using the DNA 300 assay kit. All libraries were pooled to obtain 10 nM, which was denatured to 9 or 10 pmol with 5% PhiX spiked-in and sequenced as single-read for 50 cycles on an Illumina MiSeq or HiSeq 2500 instrument in either rapid or high-output mode.

Bioinformatic analysis

The quality of raw sequenced reads was accessed with FastQC, Illumina adapters were then removed with cutadapt (-O 8 -m 26 -M 38), reads with low-quality calls were filtered out with fastq quality_filter (-q 20 -p 100 -Q 33). Using information from unique molecule identifiers (UMIs) added during library preparation, reads with the same sequence (including UMIs) were collapsed to removed putative PCR duplicates using a custom script. Prior to mapping, UMIs were trimmed (seqtk trimfq) and library quality re-assessed with FastQC. Reads were aligned against the Zebrafish (*Danio rerio*) genome assembly Zv9 with bowtie v0.12.8 (—tryhard —best —strata —chunkmbs 256 -v 1 -M 5).

The locations of transposable elements were downloaded from the UCSC genome browser (repeat masker track, Zv9) and used to select reads mapping to either RNA (SINE, LINE and LTR) or DNA transposons. The strength of the ping-pong cycle was assessed as the 5' overlap of reads in opposite strands (Brennecke et al., 2007) and the Z-scores scores were calculated as $Z\text{-score} = (P10 - M) / SD$, where P10 is the number of read pairs with an offset of 10 bases, M the mean of read pairs with 1-9 and 11-30 bases, and SD the standard deviation.

LC-MS/MS

Mass spectrometry sample preparation

Proteins were heated to 80°C for 10 min prior to loading on a 4%-12% NuPage NOVEX gradient gel (Life Technologies). The proteins were separated using MOPS buffer (ThermoFisher) at 170V for 10 min. In-gel digestion was performed essentially as previously described (Shevchenko et al., 2007). Peptides were desalted on StageTips and stored on them until MS measurement.

Mass spectrometry measurement

Peptides were eluted from StageTips with 80 percent ACN/0.5% formic acid. The mixture was separated using an EASY-nLC1000 with a reversed phase column (25 cm, 75 μ m inner diameter, packed in-house with ReproSil-Pur C18-AQ 1.9 μ m (Dr. Maisch GmbH)) mounted directly at a Q Exactive Plus mass spectrometer (ThermoFisher). A 88 minute gradient of 2% to 40% acetonitrile at a flow of 225 nl/min was combined with a wash-out of 95% acetonitrile in an overall 105 min instrument method. Spray voltage was set to ca. 2.4 kV. The instrument performed a top10 data-dependent acquisition with up to 10 HCD fragmentations per MS full scan (70k resolution, 300-1650 m/z).

MS analysis

The raw files were processed with MaxQuant v.1.5.2.8 and searched with the incorporated Andromeda search engine against a Uniprot/Trembl *Danio rerio* fasta file (58,793 entries) (Cox and Mann, 2008). Carbamidomethylation was set as fixed modification while methionine oxidation, protein N-acetylation, phosphorylation and lysine/arginine dimethylation were considered as variable modifications. The search was performed with an initial mass tolerance of 7ppm mass accuracy for the precursor and 20 ppm for the MS/MS spectra in the HCD fragmentation mode. Standard settings were applied except match between runs and the LFQ quantitations were activated. Search results were filtered at a false discovery rate of 0.01 on protein and peptide level.

Data Analysis

For statistical analysis, protein groups identified by site, known contaminants and reverse hits were excluded. The dataset was further filtered for at least 2 peptide identifications (at least 1 unique and 1 razor) per protein group. Missing LFQ values were imputed using lower values of a beta distribution derived from the measured values. Provided LFQ values were log2 transformed and for the volcano plot, the mean of the LFQ intensity of the Tdrd6a IP subtracted with the mean of the LFQ intensity of IgG IP (x-axis) against the p-value from a Welch t-test between both groups (y-axis) were plotted.

Fluorescence-activated cell sorting (FACS)

Embryos were collected at the stage of interest in 20 ml E3 medium (5 mM NaCl, 0.17 mM KCl, 0.33 mM CaCl₂, 0.33 mM MgSO₄) in a glass beaker. 500 μ l 10mg/ml pronase was added

to remove the chorions. The chorions were washed away with E3 and the embryos were taken up in 1 mL TrypLE Express (1x) (Life Technologies) and dissociated with the help of a syringe. The samples were added to 6 mL E3, 5 μ L 1M EDTA and 0.5 mL FCS was added. Samples were centrifuged for 5 minutes at 1000g. Supernatant was removed and cells were taken up in PBS. Cells were pipetted through a cell strainer and 10 μ L DNase was added. Prior to sorting, DAPI was added. Single GFP-positive cells were sorted using a BD FACSAria III (Becton Dickinson) with an 85 μ m nozzle in 96 well plates containing 150 μ L TRIzol Reagent (ThermoFisher).

4

CEL-Seq library preparation.

Library preparation was performed as described (Grün et al., 2014).

Quantification of transcript abundance.

Paired end read processing obtained by CEL-Seq was essentially done as described with minor modifications (Grün et al., 2015). The transcriptome of all Ensembl genes (version Zv9) was downloaded and all isoforms were merged into single genes. We aligned the reads with BWA using standard settings to the improved Ensembl transcriptome as described (Junker et al., 2014).

RaceID2

RaceID2 was used with default settings. The lower limit of transcript counts per cell we used was 1100 transcripts per cell. Thereafter the transcript counts of all cells passing this threshold were down sampled to 1100 transcripts per cell. The data described in Figure S3I and S3J was obtained from cells down sampled to 1750 transcripts to be able to pick up more lowly expressed genes as well. K-medoids clustering was performed similarly as the k-means clustering as described (Grün et al., 2015), but optimal cluster number was not determined by gap statistics, but by determining the saturation point of the within cluster dispersion (Grün et al., 2016). For t-SNE-map generation a seed of 2500 was used. Standard deviations on the individual Gp correlations were obtained by bootstrapping (n=100).

Detection of differentially expressed genes in scRNA-seq data

To identify differentially expressed genes we applied an approach akin to previously published method (Anders and Huber, 2010). To compare two subsets of cells from the same dataset the down-sampled version computed by RaceID2 was used as input. A p-value for a significant difference in mean expression of a gene between the two subsets was computed as described in (Anders and Huber, 2010) using the RaceID background model (Grün et al., 2015) to estimate the dispersion parameter of the negative binomial expression distribution within the two subgroups. These p-values were corrected for multiple testing by the Benjamini-Hochberg method.

Background correlation model and random cell generation

To compute the expected Spearman correlations values of transcripts with a particular average expression we ranked the genes in our dataset according to their average expression. For every Gp pair for which we computed the pairwise correlation we took the three genes above and below the two Gp genes in question from the ranked list and computed all their pairwise correlations. The pairwise correlation between Gp markers was done with all cells at 3.5hpf except one, which was discarded after manually checking of atypical Gp expression levels. This particular cell expressed *tdrd7* > 200-fold higher than the average cell in the 3.5hpf dataset and was considered a technical artifact.

Generation of random cell was done by randomly ascribing Gp counts derived from cells at the 3.5hpf timepoint and we subsequently computed the pairwise correlation of the Gp mRNAs in these hypothetical cells.

qPCR

From total RNA from input and Tdrd6a IP samples, cDNA was synthesized with M-MLV reverse transcriptase (RNase H point mutant, Promega), using random hexamers (Promega). qPCR was performed using iQ SYBR Green supermix (BioRad) on a CFX384 Real-Time thermal cycler (Bio Rad). P-values were calculated on the delta Ct values.

RIP-seq

RNA was isolated from input samples and Tdrd6a IP experiments on freshly laid embryos by TRIzol extraction according to manufacturer's instructions. 1.5 ng of immunoprecipitated or input material as total RNA was amplified to cDNA with Ovation RNA-seq System V2 (NuGEN) followed by purification with QIAquick PCR Purification Kit (Qiagen). The purified full-length cDNA was sheared to around 200 bp by a focused-ultrasonicator (Covaris). 1 μ g of sheared cDNA was end-repaired and adapter-ligated following manufacture's instruction of TruSeq DNA Sample Prep Kit (Illumina). Size selection of adapter-ligated cDNA between 200-400 bp was done by LabChip XT (PerkinElmer). This size-selected fraction was PCR amplified for 3 cycles before pooling for paired-end sequencing on HiSeq 2500 for 50 bp read length. Around 30 million reads per sample were available for processing. Reads were aligned to Zv9 using TopHat with default settings. DESeq was used to compute significance of observed differences between Tdrd6a RIP and input counts per gene (Anders and Huber, 2010). Only genes with on average more than 50 RPM were used for further analysis.

Morpholino knockdown and PGC quantification

A morpholino was designed antisense to the region containing the start-codon of *hook2* (GeneTools). 1nl was injected into 1-cell stage *pvasa:egfp* positive embryos at a concentration of 0.5mM and co-injected with 1/10th volume of rhodamine dextran. This was also performed using a morpholino targeting the *fus* transcript at the same concentration. Rescue experiments were performed by mixing 200ng/ μ L final concentration in the MO injection mix. *Hook2*

4

mRNA was amplified using an oligo containing an Sp6 promoter and mismatches at the MO binding site. mRNA was synthesized using the Sp6 mMESSAGE MACHINE kit (Invitrogen), followed by poly-A-tailing using the poly(A) tailing kit (Invitrogen). At 24hpf, the larvae were dechorionated and fixed in 4%PFA/PBS so GFP positive PGCs could be counted at the same developmental stage. 1dpf, embryos were dechorionated and fixed in 4%PFA/PBS and observed through Leica stereo microscope for counting PGCs.

4 Whole mount *in situ* hybridization

Embryos were collected and fixed at 4-cell stage in 4% PFA/PBS ON at 4°C. Next, they were washed with PBST and dechorionated. Followed by storage in MeOH at least ON at -20°C. Upon rehydration, embryos were blocked in Hyb+ (50% de-ionized formamide, 5xSSC, 0.1%Tween-20, 5mg/ml yeast RNA, 50µg/ml heparin) for 2 hrs at 70°C. Next, samples were incubated ON at 70°C with DIG-labelled probe in Hyb+. After probe removal, samples were washed at 70°C: 2 x 20 minutes in Hyb- (Hyb+ without yeast RNA and heparin), 2 x 20 minutes in 2xSSCT, 2 x 20 minutes in 0.2xSSCT. Samples were washed twice in TBST at RT and blocked 1 hour in 10% BSA. Next, samples were incubated with 1:2000 anti-DIG-AP Fab fragments (Roche) in 10% BSA ON at 4°C. Next day, they were washed 3 x 20 min. in TBST and 2 x 10 min. in AP-staining buffer (100mM NaCl, 100mM Tris pH9.5, 50mM MgCl₂, 0.1% Tween-20), followed by incubation with NBT (4.5µl/ml AP staining buffer; Roche) and BCIP (3.5µl/ml AP staining buffer; Roche) to stain the embryos.

Whole mount fluorescent *in situ* hybridization

Ovary tissue was put in OR2 medium (82mM NaCl, 2mM KCl, 1mM MgCl₂, 5mM HEPES pH7.5) and filtered through a 300µm mesh to collect only stage I-III oocytes. These oocytes were fixed in 4% PFA/PBS 3 hrs at RT, followed by dehydration in MeOH and storage ON at -20°C. The same procedure was followed as described above for colorimetric ISH, only after blocking in 10% BSA, samples were incubated with 1:1000 sheep-anti-DIG (Roche) in 10% BSA ON at 4°C. After 3 x 20 min. washing in TBST, samples were incubated with 1:500 anti-sheep-Alexa555 (Life Technologies) for 1 hr/RT. Samples were washed 3 x 10 min. in TBST and mounted in 80% glycerol, followed by imaging under DM6000 Leica microscope.

Electron microscopy

Gonads of 5wpf Tdrd6a Mut and WT fish were fixed with half-strength Karnovsky fixative (pH 7.4, Karnovsky 1965) and postfixated with 1% OsO₄ in 0.1M Cacodylate buffer (pH7.4), dehydrated in an Acetone series and embedded in EPON (Karnovsky, 1965). Semi-thin sections (1.5µm) were cut on a Reichert Ultracut 2040 and a Butler diamond knife (Diatome) until the desired area in the gonad was reached. Ultra-thin sections (90nm) were cut on a Reichert Ultracut E and collected on pioloform coated copper slot grids, dried and stained with leadcitrate under oxygen free conditions for 2 minutes. Sections were examined with a Zeiss LIBRA 120.

BmN4 cell culture and transfection

BmN4 cells were cultured at 27°C in IPL-41 (Gibco) medium supplemented with 10%FBS (Gibco) and 0.5% Pen-Strep. For imaging, cells were grown in 8-well µ-slides (ibidi 80826). 300ng construct mix (co-transfection 300ng total or single transfection 150ng plus 150ng empty vector) was incubated with 0.9µl X-tremeGENE™ HP (Roche) filled up to 30µL with insect medium without additives for 30 minutes prior to transfection.

Surface calculation Buc granules in BmN4 cells

The surface of the 2D information of the granules was calculated in Fiji. First, images were binarised using the Otsu method. Next, the surface of the granules was measured using 'Analyze Particles' (size: > 0.1µm²).

Cloning

Tdrd6a-mCherry-polyA construct

The transgene was made using Tol2 kit for multisite Gateway cloning. The *tldr6a* CDS was amplified and cloned into pDonr221 using BP clonase in order to obtain pME_tdrd6a. Next, an LR reaction was performed using p5E_pziwi (ziwi promoter), pME_tdrd6a, p3E_mCherry-polyA and tol2CG2 (pDest) (Kwan et al., 2007; Leu and Draper, 2010). Injected embryos were screened for *cmc2:gfp* to create a line. The expression pattern of this transgene was limited to oocytes up to stage Ib to early stage II (data not shown).

Buc-RtoK-eGFP mutant construct

The Buc construct used to make the Buc-eGFP line (Riemer et al., 2015) was modified with PCR in order to obtain the RtoK mutated version. First, the plasmid was digested with NotI and SalI and the resulting 4064 bp fragment was subcloned in pre-digested pCS2+. Arginine codons were replaced by Lysine codons using RtoKmut_GFPstart_F (phosphorylated): GGCTCA AGATACGGCGGAAGCGGCATGGTGAGCAAGGGCGAGGAG and RtoKmut_R: **CTT** TTT CTT CAT AGA ACC **TTT** GCC **CTT** CTG GCA GTA GGC. This also causes loss of intron 6, however, since the other introns remained unaffected and the transgene was translated and could localize normally in the presence of wt Buc, this does not seem to affect the functionality of the transgene. After the PCR, the original plasmid was digested with DpnI and the PCR fragment was circularized by ligation and sequenced. Next, the mutated fragment was ligated back into the pre-digested *buc-egfp* plasmid. The mutated construct was injected into wt embryos together with *tol2* transposase and adult females were screened for GFP-positive gonads. The F0 females were then outcrossed in order to obtain a stable line, prior to crossing it with the *buc^{p106}* allele.

Cloning of the BmN4 expression constructs

Dcp1 was amplified and digested with BamHI and NotI and ligated into pBEMBL-NHA. Next, mCherry and eGFP were amplified and ligated N-terminally into the BamHI site in order to

create mCherry-Dcp1 and eGFP-DCP1 respectively. Tdrd6a was amplified and digested with NotI and XbaI and ligated into pre-digested pBEMBL-NHA-mCherry. Buc was amplified using Buc F, which was 5' phosphorylated, and Buc_R, containing an overhang for the pBEMBL-NHA-eGFP. Next, the amplicon was mixed with pBEMBL-NHA-eGFP and together with the pBEMBL_R oligo, a PCR was performed on the entire plasmid. Next, the original backbone was digested with DpnI and the linear PCR product was ligated (and thereby circularized) in order to obtain pBEMBL-NHA-Buc-eGFP.

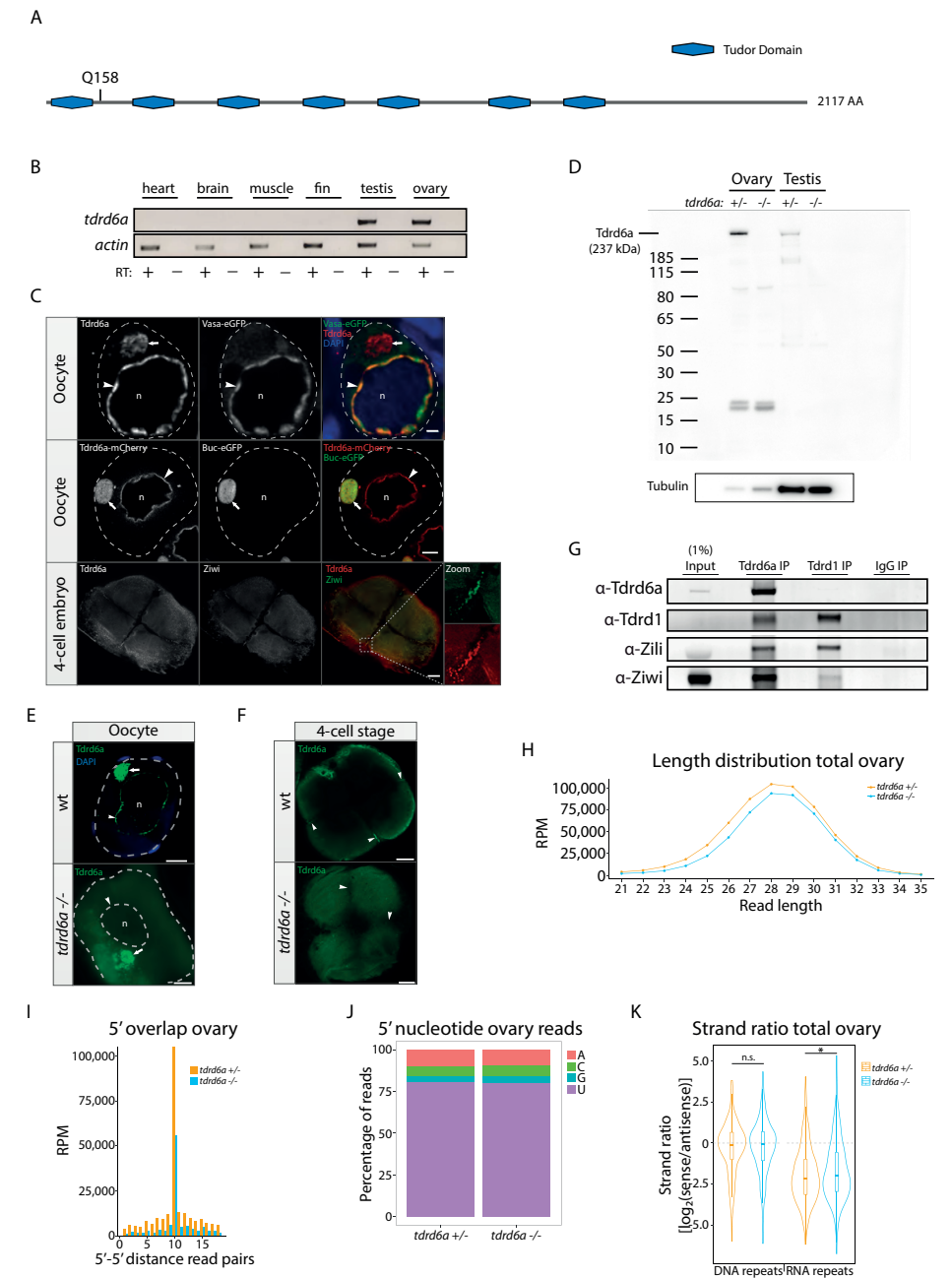
REFERENCES

- Akpınar, M., Lesche, M., Fanourgakis, G., Fu, J., Anasstasiadis, K., Dahl, A., and Jessberger, R. (2017). TDRD6 mediates early steps of spliceosome maturation in primary spermatocytes. *PLoS Genet.* *13*.
- Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biol.* *11*, R106.
- Barckmann, B., Pierson, S., Dufourt, J., Papin, C., Armenise, C., Port, F., Grentzinger, T., Chambeyron, S., Baronian, G., Desvignes, J.P., et al. (2015). Aubergine iCLIP Reveals piRNA-Dependent Decay of mRNAs Involved in Germ Cell Development in the Early Embryo. *Cell Rep.* *12*, 1205–1216.
- Bilinski, S.M., Kloc, M., and Tworzydło, W. (2017). Selection of mitochondria in female germline cells: is Balbiani body implicated in this process? *J. Assist. Reprod. Genet.* 1–8.
- 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.
- Boke, E., Ruer, M., Wühr, M., Coughlin, M., Lemaitre, R., Gygi, S.P., Alberti, S., Drechsel, D., Hyman, A.A., and Mitchison, T.J. (2016). Amyloid-like Self-Assembly of a Cellular Compartment. *Cell* *166*, 637–650.
- 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.
- Brangwynne, C.P., Eckmann, C.R., Courson, D.S., Rybarska, A., Hoegge, C., Gharakhani, J., Julicher, F., and Hyman, A.A. (2009). Germline P Granules Are Liquid Droplets That Localize by Controlled Dissolution/Condensation. *Science* (80-.). *324*, 1729–1732.
- Brangwynne, C.P., Mitchison, T.J., and Hyman, A.A. (2011). Active liquid-like behavior of nucleoli determines their size and shape in *Xenopus laevis* oocytes. *Proc. Natl. Acad. Sci.* *108*, 4334–4339.
- 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* (80-.). *322*, 1387–1392.
- Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* *26*, 1367–1372.
- Grün, D., Kester, L., and van Oudenaarden, A. (2014). Validation of noise models for single-cell transcriptomics. *Nat. Methods* *11*, 637–640.
- Grün, D., Lyubimova, A., Kester, L., Wiebrands, K., Basak, O., Sasaki, N., Clevers, H., and van Oudenaarden, A. (2015). Single-cell messenger RNA sequencing reveals rare intestinal cell types. *Nature* *525*, 251–255.
- Grün, D., Muraro, M.J., Boisset, J.C., Wiebrands, K., Lyubimova, A., Dharmadhikari, G., van den Born, M., van Es, J., Jansen, E., Clevers, H., et al. (2016). De Novo Prediction of Stem Cell Identity using Single-Cell Transcriptome Data. *Cell Stem Cell* *19*, 266–277.
- Hachet, O., and Ephrussi, A. (2004). Splicing of oskar RNA in the nucleus is coupled to its cytoplasmic localization. *Nature* *428*, 959–963.
- 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.
- Hashimshony, T., Wagner, F., Sher, N., and Yanai, I. (2012). CEL-Seq: Single-Cell RNA-Seq by Multiplexed Linear Amplification. *Cell Rep.* *2*, 666–673.
- Houwing, S. (2009). Piwi-piRNA complexes in the zebrafish germline.
- 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.
22. Huang, H.-Y., Houwing, S., Kaaij, L.J.T., Meppelink, A., Redl, S., Gauci, S., Vos, H., Draper, B.W., Moens, C.B., Burgering, B.M., et al. (2011). Tdrd1 acts as a molecular scaffold for Piwi proteins and piRNA targets in zebrafish. *EMBO J.* 30, 3298–3308.
23. Ikenishi, K. (1998). Germ plasm in *Caenorhabditis elegans*, *Drosophila* and *Xenopus*. *Dev. Growth Differ.* 40, 1–10.
24. Junker, J.P., Noël, E.S., Guryev, V., Peterson, K.A., Shah, G., Huisken, J., McMahon, A.P., Berezikov, E., Bakkers, J., and Van Oudenaarden, A. (2014). Genome-wide RNA Tomography in the Zebrafish Embryo. *Cell* 159, 662–675.
25. Karnovsky, M.J. (1965). A Formaldehyde-Glutaraldehyde Fixative of High Osmolality for Use in Electron Microscopy. *Cell Biol. Source J. Cell Biol.* 27, 137A.
26. Kato, M., Han, T.W., Xie, S., Shi, K., Du, X., Wu, L.C., Mirzaei, H., Goldsmith, E.J., Longgood, J., Pei, J., et al. (2012). Cell-free formation of RNA granules: Low complexity sequence domains form dynamic fibers within hydrogels. *Cell* 149, 753–767.
27. Kirino, Y., Vourekas, A., Sayed, N., de Lima Alves, F., Thomson, T., Lasko, P., Rappsilber, J., Jongens, T.A., and Mourelatos, Z. (2010). Arginine methylation of Aubergine mediates Tudor binding and germ plasm localization. *RNA* 16, 70–78.
28. 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.
29. Knaut, H., Steinbeisser, H., Schwarz, H., and Nusslein-Volhard, C. (2002). An evolutionary conserved region in the vasa 3'UTR targets RNA translation to the germ cells in the zebrafish. *Curr Biol* 12, 454–466.
30. Kopranner, 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.
31. Kroschwald, S., Maharana, S., Mateju, D., Malinowska, L., Nüske, E., Poser, I., Richter, D., and Alberti, S. (2015). Promiscuous interactions and protein disaggregases determine the material state of stress-inducible RNP granules. *Elife* 4.
32. 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.
33. Krøvel, 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.
34. Kwan, K.M., Fujimoto, E., Grabher, C., Mangum, B.D., Hardy, M.E., Campbell, D.S., Parant, J.M., Yost, H.J., Kanki, J.P., and Chien, C. Bin (2007). The Tol2kit: A multisite gateway-based construction Kit for Tol2 transposon transgenesis constructs. *Dev. Dyn.* 236, 3088–3099.
35. Leu, D.H., and Draper, B.W. (2010). The zivi promoter drives germline-specific gene expression in zebrafish. *Dev. Dyn.* 239, 2714–2721.
36. Little, S.C., Sinsimer, K.S., Lee, J.J., Wieschaus, E.F., and Gavis, E.R. (2015). Independent and coordinate trafficking of single *Drosophila* germ plasm mRNAs. *Nat. Cell Biol.* 17, 558–568.
37. Van Der Maaten, L.J.P., and Hinton, G.E. (2008). Visualizing high-dimensional data using t-sne. *J. Mach. Learn. Res.* 9, 2579–2605.
38. 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.
39. 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.
40. Nelson, M.R., Leidal, A.M., and Smibert, C.A. (2004). *Drosophila* Cup is an eIF4E-binding protein that functions in Smaug-mediated translational repression. *EMBO J.* 23, 150–159.
41. 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.
42. Nott, T.J., Petsalaki, E., Farber, P., Jervis, D., Fussner, E., Plochowitz, A., Craggs, T.D., Bazett-Jones, D.P., Pawson, T., Forman-Kay, J.D., et al. (2015). Phase Transition of a Disordered Nuage Protein Generates Environmentally Responsive Membraneless Organelles. *Mol. Cell* 57, 936–947.
43. Owens, D.A., Butler, A.M., Agüero, T.H., Newman, K.M., Van Booven, D., and King, M. Lou (2017). High-throughput analysis reveals novel maternal germline RNAs crucial for primordial germ cell preservation and proper migration. *Development* 144, 292–304.
44. Patel, A., Lee, H.O., Jawerth, L., Maharana, S., Janel, M., Hein, M.Y., Stoykov, S., Mahamid, J., Saha, S., Franzmann, T.M., et al. (2015). A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* 162, 1066–1077.
45. Phillips, C.M., Montgomery, T.A., Breen, P.C., and Ruvkun, G. (2012). MUT-16 promotes formation of perinuclear Mutator foci required for RNA silencing in the *C. elegans* germline. *Genes Dev.* 26, 1433–1444.
46. Prusiner, S.B. (1998). Prions. *Proc. Natl. Acad. Sci. U. S. A.* 95, 13363.
47. Raz, E. (2003). Primordial germ-cell development: the zebrafish perspective. *Nat. Rev. Genet.* 4, 690–700.
48. Riemer, S., Bontems, F., Krishnakumar, P., Gömann, J., and Dosch, R. (2015). A functional Bucky ball-GFP transgene visualizes germ plasm in living zebrafish. *Gene Expr. Patterns* 18, 44–52.
49. Shevchenko, A., Tomas, H., Havlíš, J., Olsen, J. V., and Mann, M. (2007). In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* 1, 2856–2860.
50. Shin, Y., and Brangwynne, C.P. (2017). Liquid phase condensation in cell physiology and disease. *Science* (80-.). 357, eaaf4382.
51. Shorter, J., and Lindquist, S. (2005). Prions as adaptive conduits of memory and inheritance. *Nat. Rev. Genet.* 6, 435–450.
52. Siddiqui, N.U., Li, X., Luo, H., Karaiskakis, A., Hou, H., Kislinger, T., Westwood, J.T., Morris, Q., and Lipshitz, H.D. (2012). Genome-wide analysis of the maternal-to-zygotic transition in *Drosophila* primordial germ cells. *Genome Biol.* 13, R11.
53. 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.
54. Slaidina, M., and Lehmann, R. (2017). Quantitative Differences in a Single Maternal Factor Determine Survival Probabilities among *Drosophila* Germ Cells. *Curr. Biol.* 27, 291–297.
55. Strasser, M.J., Mackenzie, N.C., Dumstrei, K., Nakkrasae, L.-I.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.
56. Takahashi, K., and Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* 126, 663–676.
57. Thandapani, P., O'Connor, T.R., Bailey, T.L., and Richard, S. (2013). Defining the RGG/RG Motif. *Mol. Cell* 50, 613–623.
58. 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.
59. Tinevez, J.Y., Perry, N., Schindelin, J., Hoopes, G.M., Reynolds, G.D., Laplantine, E., Bednarek, S.Y., Shorte, S.L., and Eliceiri, K.W. (2017). TrackMate: An open and extensible platform for single-particle tracking. *Methods* 115, 80–90.
60. Trcek, T., Grosch, M., York, A., Shroff, H., Lionnet, T., and Lehmann, R. (2015). *Drosophila* germ granules are structured and contain homotypic mRNA clusters. *Nat. Commun.* 6, 7962.
61. Tzung, K.W., Goto, R., Saju, J.M., Sreenivasan, R., Saito, T., Arai, K., Yamaha, E., Hossain, M.S., Calvert, M.E.K., and Orbán, L. (2015). Early depletion of primordial germ cells in zebrafish promotes testis formation. *Stem Cell Reports* 4, 61–73.
62. 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.

63. Vourekas, A., Alexiou, P., Vrettos, N., Maragkakis, M., and Mourelatos, Z. (2016). Sequence-dependent but not sequence-specific piRNA adhesion traps mRNAs to the germ plasm. *Nature* 531, 390–394.
64. Wang, H., Teng, Y., Xie, Y., Wang, B., Leng, Y., Shu, H., and Deng, F. (2013). Characterization of the carbonic anhydrases 15b expressed in PGCs during early zebrafish development. *Theriogenology* 79, 443–452.
65. Wang, J.T., Smith, J., Chen, B.C., Schmidt, H., Rasoloson, D., Paix, A., Lambrus, B.G., Calidas, D., Betzig, E., and Seydoux, G. (2014). Regulation of RNA granule dynamics by phosphorylation of serine-rich, intrinsically disordered proteins in *C. elegans*. *Elife* 3.
66. 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.
67. Wienholds, E. (2002). Target-Selected Inactivation of the Zebrafish *rag1* Gene. *Science* (80-.). 297, 99–102.
68. Woodruff, J.B., Ferreira Gomes, B., Widlund, P.O., Mahamid, J., Honigsmann, A., and Hyman, A.A. (2017). The Centrosome Is a Selective Condensate that Nucleates Microtubules by Concentrating Tubulin. *Cell* 169, 1066–1077.e10.
69. 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.

SUPPLEMENTARY FIGURES

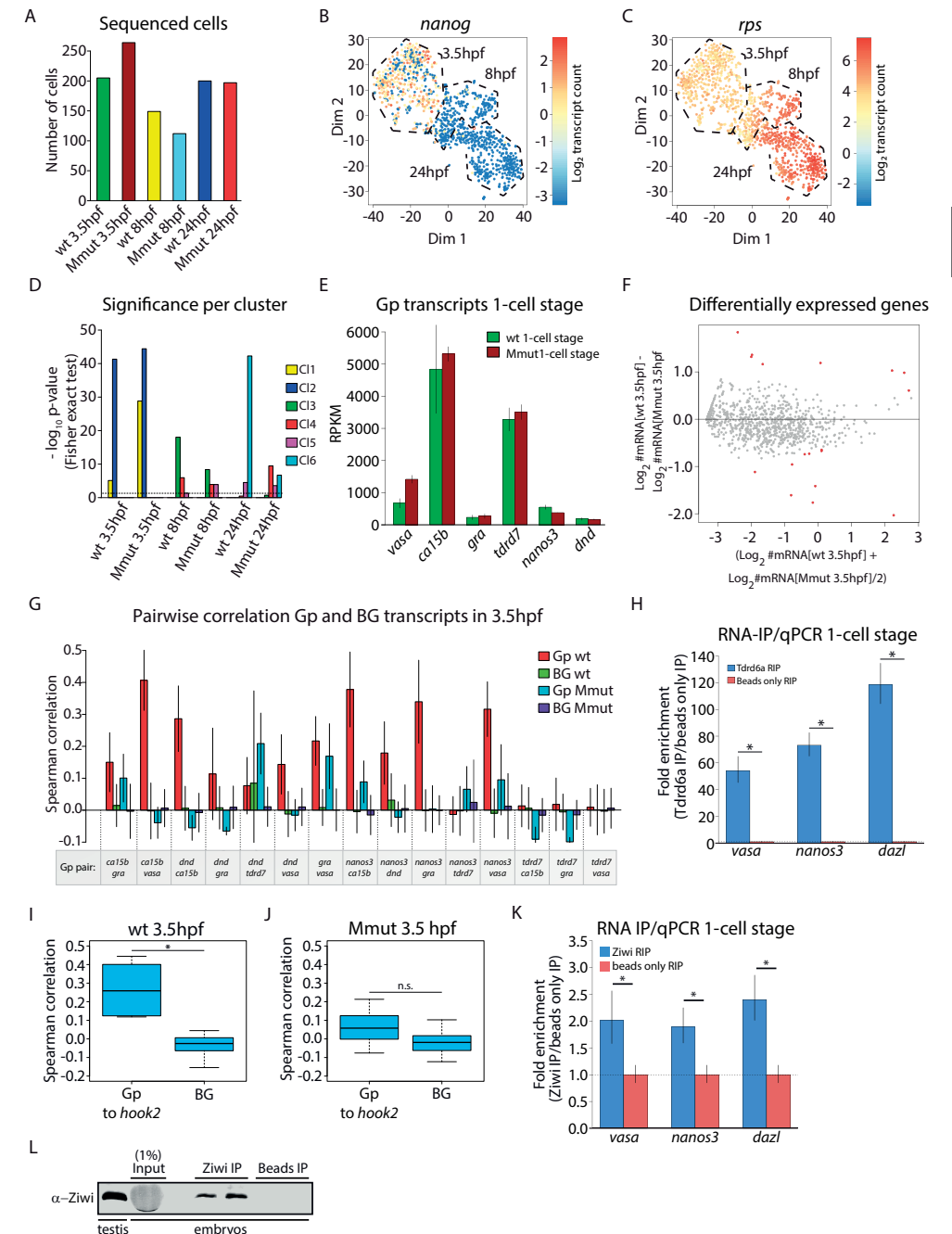


Supplemental Figure 1, Related to Figure 1. (A) Schematic overview of Tdrd6a. (B) RT-PCR of different tissues to confirm germline expression. (C) Tdrd6a co-localizes to Vasa-eGFP in nuage, to Buc-eGFP in the Bb (Riemer et al., 2015) and to Ziwi in the Gp. A Tdrd6a-mCherry-polyA3'UTR was used under a Ziwi promoter, which expresses until ~stage II oocytes. Scale bars (top-middle-bottom): 2µm, 10µm and 100µm

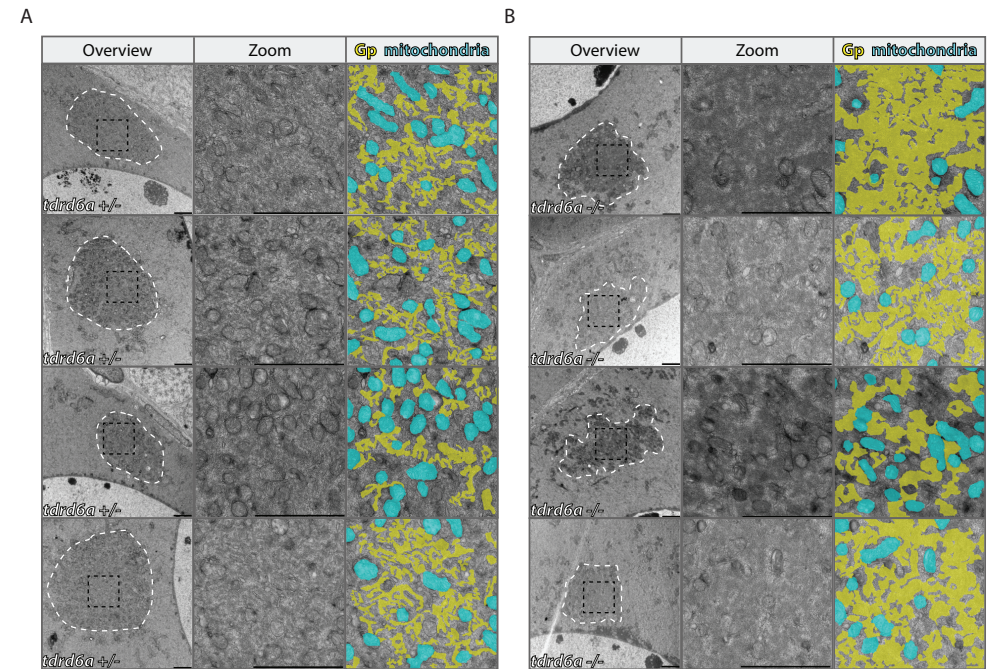
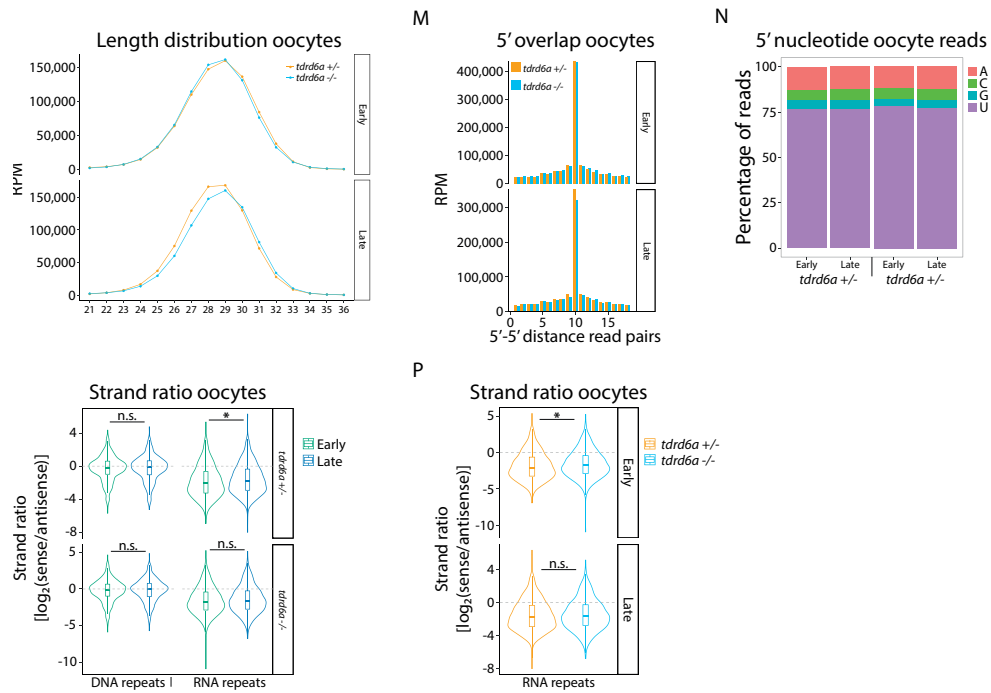
4

4

► (D) Western blot for Tdrd6a in wt and *tdrd6a*^{-/-} ovary and testis. (E) Tdrd6a staining in wt and *tdrd6a* mutant ovary. Tdrd6a is lost from nuage in the *tdrd6a* mutant (arrowhead) but some signal from the Bb remains (arrow). Wt is taken along from Figure 1A for comparison. Scale bars indicate 10µm (F) Tdrd6a staining in wt and a MZ *tdrd6a* mutant 4-cell stage embryo. Arrowheads indicate the cleavage planes. Wt is taken along from Figure 1B for comparison. Scale bars indicate 100µm. (G) Quantitative Western blot for Tdrd6a, Tdrd1 and IgG IPs (based on IRDye detection with LI-COR). Ziwi is 3.3x more present in Tdrd6a IPs compared to Tdrd1 IPs, when normalized to Zili. (H) Length distribution sRNA reads mapping to TEs for *tdrd6a*^{+/-} and *tdrd6a*^{-/-} ovary, as indicated. (I) 5' overlap of piRNAs in *tdrd6a*^{+/-} and *tdrd6a*^{-/-} ovary. Ping-pong Z-scores are 39 and 41 for *tdrd6a*^{+/-} and *tdrd6a*^{-/-} respectively. (J) 5' nucleotide bias displaying the typical 5'U bias in both genotypes. (K) The sense/antisense bias observed in ovary of *tdrd6a*^{+/-} and *tdrd6a*^{-/-} siblings is plotted for all DNA and RNA transposons. *Tdrd6a*^{-/-} ovary shows a reduction in antisense piRNAs (p = 0.03) mapping to RNA elements. (L) Length distribution sRNA reads mapping to TEs for *tdrd6a*^{+/-} and *tdrd6a*^{-/-} early and late oocytes, as indicated. (M) 5' overlap of piRNAs in *tdrd6a*^{+/-} and *tdrd6a*^{-/-} early and late oocytes. Ping-pong Z-scores are 30 and 32 in early oocytes and 32 and 33 in late oocytes, for *tdrd6a*^{+/-} and *tdrd6a^{-/-} respectively. (N) 5' nucleotide bias displaying an unaffected 5'U bias. (O,P) The sense/antisense bias observed in ovary of *tdrd6a*^{+/-} and *tdrd6a*^{-/-} siblings plotted for reads mapping against DNA and RNA transposons, early versus late (O) and *tdrd6a*^{+/-} versus *tdrd6a*^{-/-} (* indicates p < 0.01, Mann-Whitney-Wilcoxon Test) (P) RPM = Reads per million.*

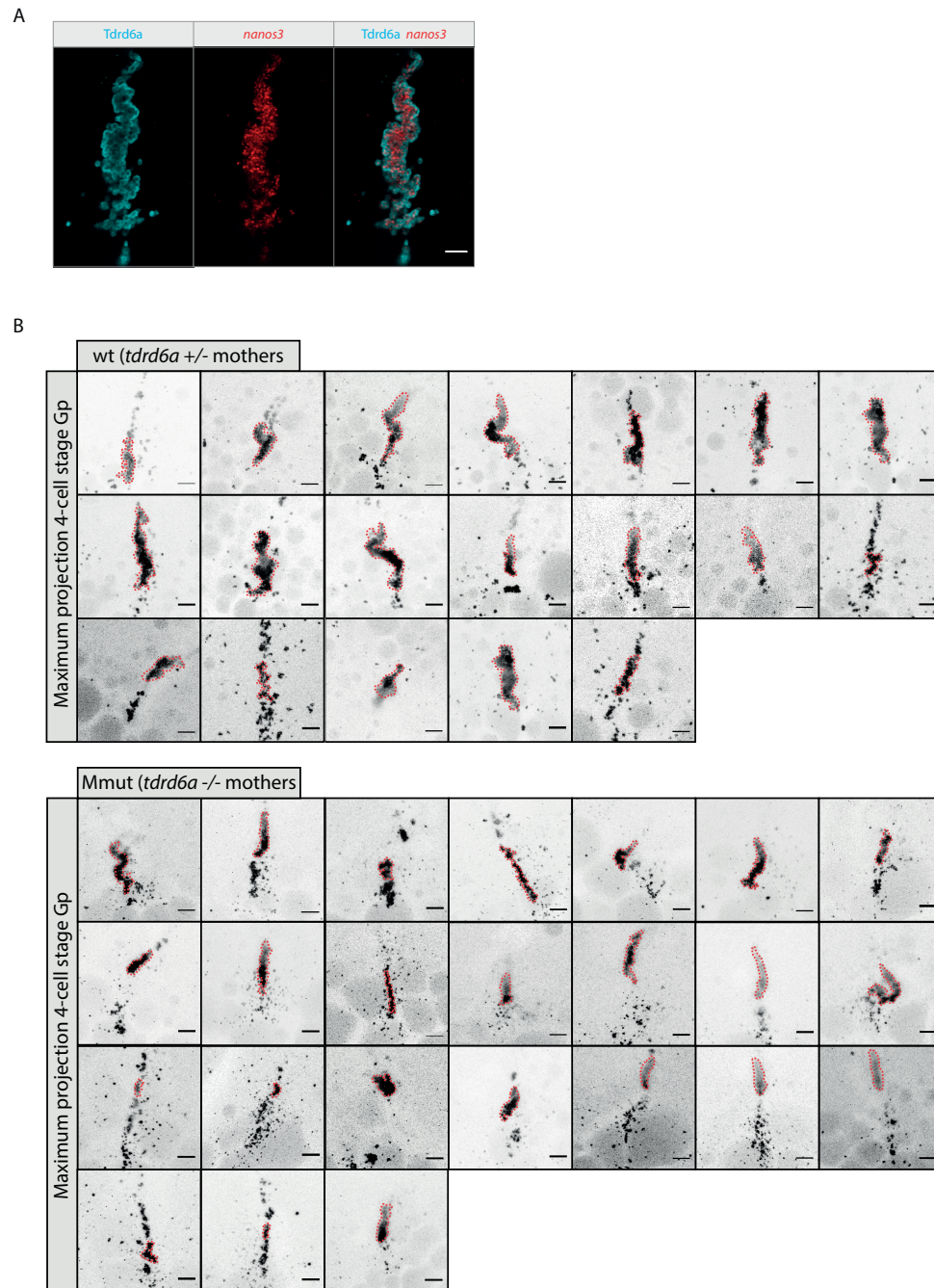


Supplemental Figure 2, Related to Figure 2. (A) Barplot displaying the number of cells per genotype and developmental timepoint used in this study. (B and C) t-SNE maps showing transcript counts of *nanog* and the *rps* gene group, respectively. (D) Barplot displaying the significance of enrichment for the different genotype-developmental time combinations in the six clusters identified in (2B). (E) Barplot displaying the RPKM counts for six Gp mRNAs from bulk RNA-seq of 1-cell stage wt and Mmut *tdrd6a* embryos, as ►

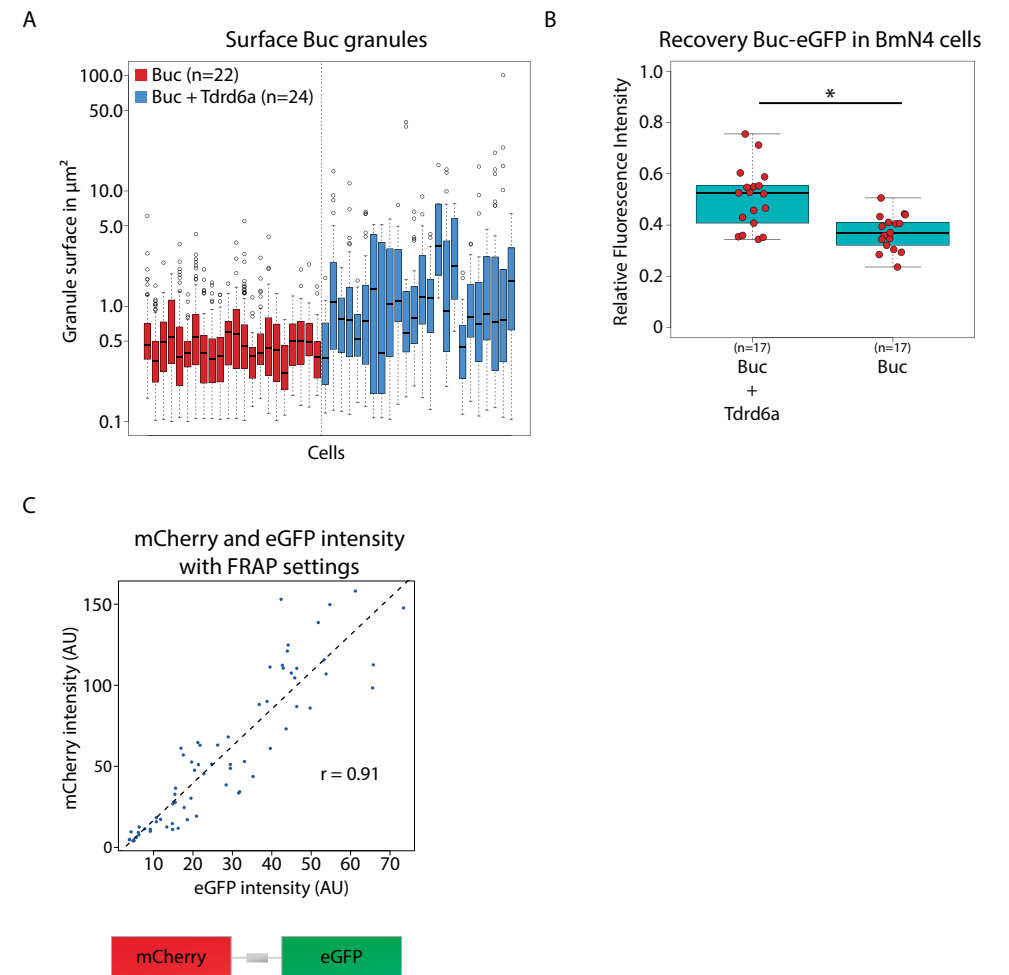


Supplemental Figure 3, Related to Figure 3. (A, B) Additional electron micrographs to further illustrate *tdrd6a* heterozygous (A) and mutant (B) Balbiani bodies, including the Bbs from Figure 3E and F (bottom row). Overlays in the right panel indicate the Gp regions (yellow) and mitochondria (cyan) that can be appreciated in the middle panel.

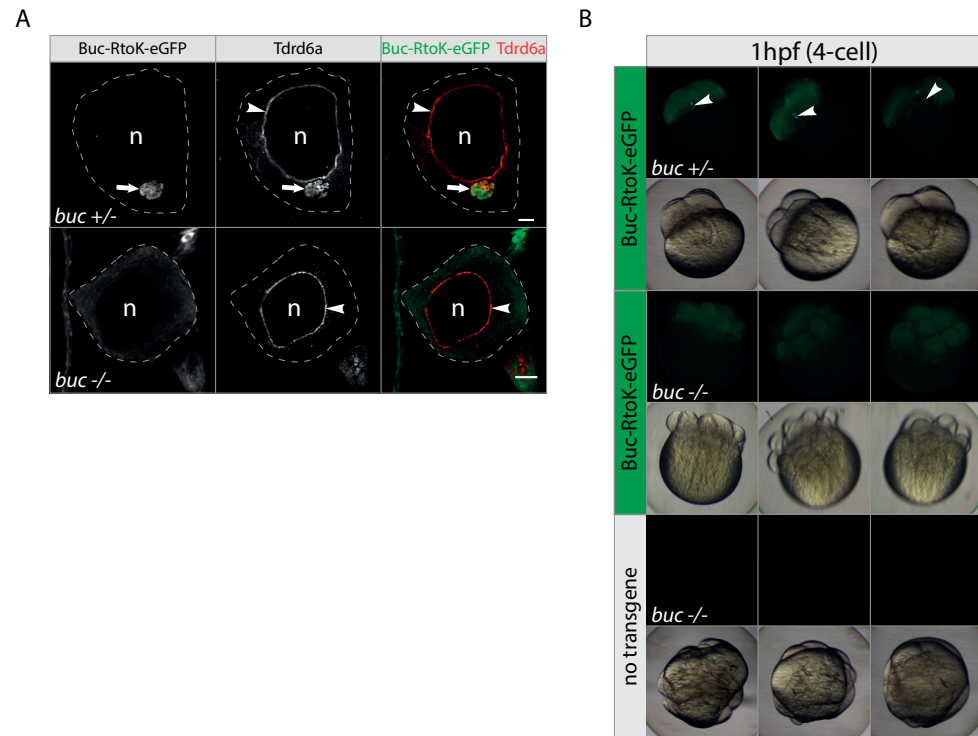
► indicated. Error bars indicate standard deviation obtained from three biological replicates. (F) Scatterplot displaying transcript counts in wt and Mmut PGCs at 3.5hpf based on scRNA-seq. Genes highlighted in red are 2 fold up or down regulated between genotypes with a p-value < 0.01 (p-value is calculated by negative binominal statistics and corrected for multiple testing (*Benjamini–Hochberg*)). (G) Barplot representing the Spearman correlation between Gp-Gp transcript and BG-BG transcripts in 3.5hpf old embryos, as indicated. Standard deviations of the Gp-Gp correlations were obtained by bootstrapping. (H) Tdrd6a RIP-qPCR analysis for *nanos3*, *dazl* and *vasa*. Enrichments were calculated compared to beads only RIP and normalized to β -actin. Error bars represent standard deviation of two biological replicates (* indicates p-value < 0.001, p-value obtained by two-sided Student's t-test). (I, J) Boxplots displaying the *hook2*-Gp and BG-BG correlations in 3.5hpf PGCs from wt and Mmut embryos, as indicated (* = p-value < 0.001, n.s. = non-significant, calculated by Wilcoxon test). (K) Ziwi RIP-qPCR analysis for *nanos3*, *dazl* and *vasa*. Enrichments were calculated compared to beads only RIP and normalized to β -actin. Error bars represent standard deviation of two biological replicates (* = p-value < 0.05, p-value obtained by two-sided Student's t-test). (L) Western blot control for successful IP of (K).



Supplemental Figure 4, Related to Figure 4. (A) FISH against *nanos3* using anti-DIG IHC combined with Tdrd6a IHC visualizing Gp at a 4-cell stage embryo cleavage plane. (B) Maximum projections of Z-stacks of wt and Mmut Gp of 4-cell stage Buc-eGFP signals, used to calculate Gp volumes of Figure 4C. Red dotted line indicates the largest Gp fragment, of which the volume in μm^3 was measured. Scalebars represent $5\mu\text{m}$ (A) and $10\mu\text{m}$ (B).



Supplemental Figure 5, Related to Figure 6. (A) Surface calculation of Buc-containing granules in transfected BmN4 cells in μm^2 . Overall, granules increase in size in the presence of Tdrd6a, with large variation between co-transfected cells. (B) Boxplot of the recovery of separate FRAP experiments of Buc-eGFP as in Figure 6E, based on the average intensity value of the last 20 frames (* indicates p-value < 0.05 , calculated by Wilcoxon test). (C) Fluorescence intensity of an mCherry-eGFP fusion construct as indicated, using the FRAP settings used in Figure 6E.



Supplemental Figure 6, Related to Figure 7. (A) Tdrd6a localization in sections of *buc +/-* and *buc -/-* oocytes in the Buc-RtoK-eGFP background. Arrow: Bb, arrowheads: nuage. Scale bar represents 10µm (B) Examples of 4-cell stage embryos of *buc +/-* and *-/-* mothers, with and without Buc-RtoK-eGFP. Without the presence of wt Buc, Buc-RtoK can rescue the lack of polarity of the *buc* phenotype (*buc -/-*, no transgene), even though most embryos display severe developmental defects (*buc -/-*, Buc-RtoK-eGFP). Arrowheads: Gp.

Chapter

5

Bucky ball stabilizes late vegetal
pathway components and
oocyte polarity in
the absence of a Balbiani body

ABSTRACT

Throughout zebrafish oogenesis, crucial patterning events take place for viability and fertility of the progeny. The earliest embryonic axis, the animal-vegetal (AV) axis, is determined during prophase of the first meiotic division. This coincides with the formation of a large, membrane-less condensate, the Balbiani body (Bb), which is the first step in the formation of germ plasm. The nucleator of this structure, Bucky ball (Buc), is essential for the Bb and for oocyte polarity. Without Buc, the Bb cannot be established and downstream germ plasm factors cannot become localized and stored for their function post fertilization. Additionally, loss of oocyte polarization results in polyspermy and hence embryonic lethality. We previously identified a mutated version of Buc, Buc-RtoK, which cannot interact with Tdrd6 paralogs. Buc-RtoK initially phase-separates into small granules that readily dissolve during the early vegetal pathway and fail to form a Bb, while embryos do remain polarized. We demonstrate that Buc-RtoK re-condenses during the late vegetal pathway and forms structures similar to germinal granules in a polarized fashion. Our results therefore uncouple the Bb from egg polarization.

5

INTRODUCTION

Animal development is intimately related to the specification of axes in order to form the body plan. How these axes are positioned is often determined by the asymmetric localization of transcripts and proteins. The first axis, the animal-vegetal (AV) axis, develops prior to fertilization during oogenesis when maternal determinants become spatially distributed. In zebrafish, this axis emerges as early as the zygotene stage of meiosis I (oocyte $\varnothing = \sim 10\text{-}15\mu\text{m}$), when telomeres of all chromosomes come together at the vegetal side of the nucleus in order to form the so-called ‘chromosomal bouquet’ (Elkouby et al., 2016). The bouquet forms through interactions with the SUN/KASH complex, a complex that spans the inner and outer nuclear membrane and links intra-nuclear telomeres to cytoplasmic microtubules that arise from the centrosome, a mechanism that is conserved among eukaryotes (Scherthan, 2001). When microtubules are inhibited, the bouquet fails to form (Elkouby et al., 2016). The bouquet-centromere orientation is the most upstream symmetry breaking event in zebrafish known to date.

Another structure that forms at this stage is the Balbiani body (Bb), a large phase-separated aggregate containing mRNA, proteins and organelles (Kloc et al., 2004). The Bb components, including its nucleating factor Bucky ball (Buc), accumulate around the centrosome in a nuclear cleft, upon which the Bb assembles (Bontems et al., 2009; Elkouby et al., 2016; Marlow and Mullins, 2008). Buc is a protein with an N-terminal prion-like domain (PLD), a protein stretch that is rich in uncharged polar amino acids, which stimulates liquid-liquid phase separation (LLPS) of proteins in solution into a physically distinct phase. Initially identified as causing agents of neurodegenerative diseases – examples are Alzheimer’s disease, Creutzfeldt-Jakob disease and Parkinson’s disease – it has now been acknowledged that these PLDs serve many biological functions in healthy contexts as well. PLDs are commonly found in proteins that play a role in the assembly or regulation of biological condensates, including the Bb (Boke et al., 2016; Brangwynne et al., 2009, 2011; Franzmann et al., 2018; Molliex et al., 2015; Nott et al., 2015; Wang et al., 2014; Woodruff et al., 2017).

Without Buc, the Bb fails to form, resulting in loss of polarization, even though upstream events like the bouquet and the centrosome still take place (Bontems et al., 2009; Elkouby et al., 2016; Marlow and Mullins, 2008). This means that the bouquet and the vegetal centromere, both transient structures, are not sufficient to maintain oocyte polarity and require downstream events in order to stabilize this. It has been assumed that the Bb is essential for this, however, a recent study demonstrated that a mutant version of Buc (Buc-RtoK) that fails to form a Bb, is still capable of forming polarized oocytes (Roovers et al., 2018). This mutant has a disrupted interaction site between Buc and Tdrd6 paralogs (Tdrd6a, b and/or c) which prevents Buc from Bb establishment, even though it can still phase-separate into small granules around the zygotene stage (Roovers et al., 2018). Still, severe developmental defects occur causing lethality to most embryos. This could reflect general patterning defects that occur without the Bb, since vegetal markers do not localize in *buc* mutants and animal pole markers localize ectopically (Bontems et al., 2009; Marlow and Mullins, 2008; Nojima et al., 2010). Furthermore, it must be noted

5

Elke F. Roovers¹, Hanna Braun², Falk Butter², René F. Ketting¹

¹ Biology of Non-coding RNA Group, Institute of Molecular Biology, Ackermannweg 4, 55128, Mainz, Germany

² Quantitative Proteomics Group, Institute of Molecular Biology, Ackermannweg 4, 55128, Mainz, Germany

that *buc* transgenes require introns and at least part of the 3'UTR in order to completely rescue oocyte polarity and embryo patterning, suggesting additional roles for introns, UTRs and/or splicing (Heim et al., 2014).

Formation of the Bb is a key process in RNA localization in the oocyte. Previous work in *Xenopus* and zebrafish has demonstrated that two pathways act in vegetal localization of RNA, the early and late vegetal pathway, and a third pathway for transport towards the animal pole (Abrams and Mullins, 2009; Forristall et al., 1995; Kloc and Etkin, 1995). The Bb (in *Xenopus* also named 'mitochondrial cloud') is part of the early vegetal pathway or 'METRO' pathway (for **M**essenger **T**ransport **O**rganizer) during which several transcripts are recruited to the Bb, including transcripts that are required for germ cell specification (Forristall et al., 1995; Kosaka et al., 2007). In zebrafish, transcripts like *dazl*, *vasa* but also *buc* itself localize during this stage (Bontems et al., 2009; Hashimoto et al., 2004; Marlow and Mullins, 2008). Furthermore, many proteins including Tdrd6a, Rbpms2a and Rbpms2b localize to the Bb as well (Kaufman et al., 2018; Roovers et al., 2018).

In addition to the Bb-driven localization of components there is a second wave of RNA localization during the so-called late vegetal pathway. Towards the end of oogenesis stage I, the Bb is transported towards the vegetal pole and disperses into germinal granules, which will move more cortically during later stages. Transcripts like *magoh* and *bruno-like* are recruited towards the vegetal pole during the late vegetal pathway, whereas other transcripts, including *buc* and *vg1rbp* move towards the animal pole (Bontems et al., 2009; Kosaka et al., 2007; Zhang et al., 1999). The cytoskeleton plays an important role in this movement. Studies in *Xenopus* have demonstrated that a specific pool of microtubules, oriented to the vegetal pole with their plus ends, and kinesin motors are required for the late vegetal transport (Messitt et al., 2008; Yisraeli et al., 1990). Upon transport, germinal granules also require 'docking' at the cortex, for which actin is required (Escobar-Aguirre et al., 2017; Yisraeli et al., 1990). In zebrafish, the microtubule-actin crosslinking factor 1 (*Macf1*) is required for vegetal dispersion of the Bb during the transition from the early to the late vegetal pathway (Escobar-Aguirre et al., 2017; Gupta et al., 2010).

In *buc* mutants, transcripts of both the early and late pathway fail to localize (Bontems et al., 2009; Dosch et al., 2004; Marlow and Mullins, 2008). This indicates that *buc* is not only responsible for localization of transcripts that require initial Bb localization, but also downstream of this structure. It remains unknown whether this is due to lack of a Bb structure itself, or that these transcripts reflect other functions of Buc and/or its mRNA during downstream localization steps.

Here, we explore Buc-RtoK function in the presence and absence of a wild-type copy of *buc* (*buc*^{p106+/-} and *buc*^{p106-/-} background, hereafter called *buc*^{+/-} and *buc*^{-/-}) and investigate protein and transcript localization throughout oogenesis. We find that even in the absence of a Bb, Buc-RtoK does phase-separate at the onset of the late vegetal pathway, forming structures that behave similar to germinal granules. This restores localization of some, but not all vegetal markers. Furthermore, we identify novel Buc interactors that could play a role during these events, since they, too, are mostly disordered or prion-like. In conclusion, our work disconnects

the formation of the Balbiani body from oocyte polarity, further characterizes RNA localization during the late vegetal pathway and identifies novel proteins that may play a role in this process.

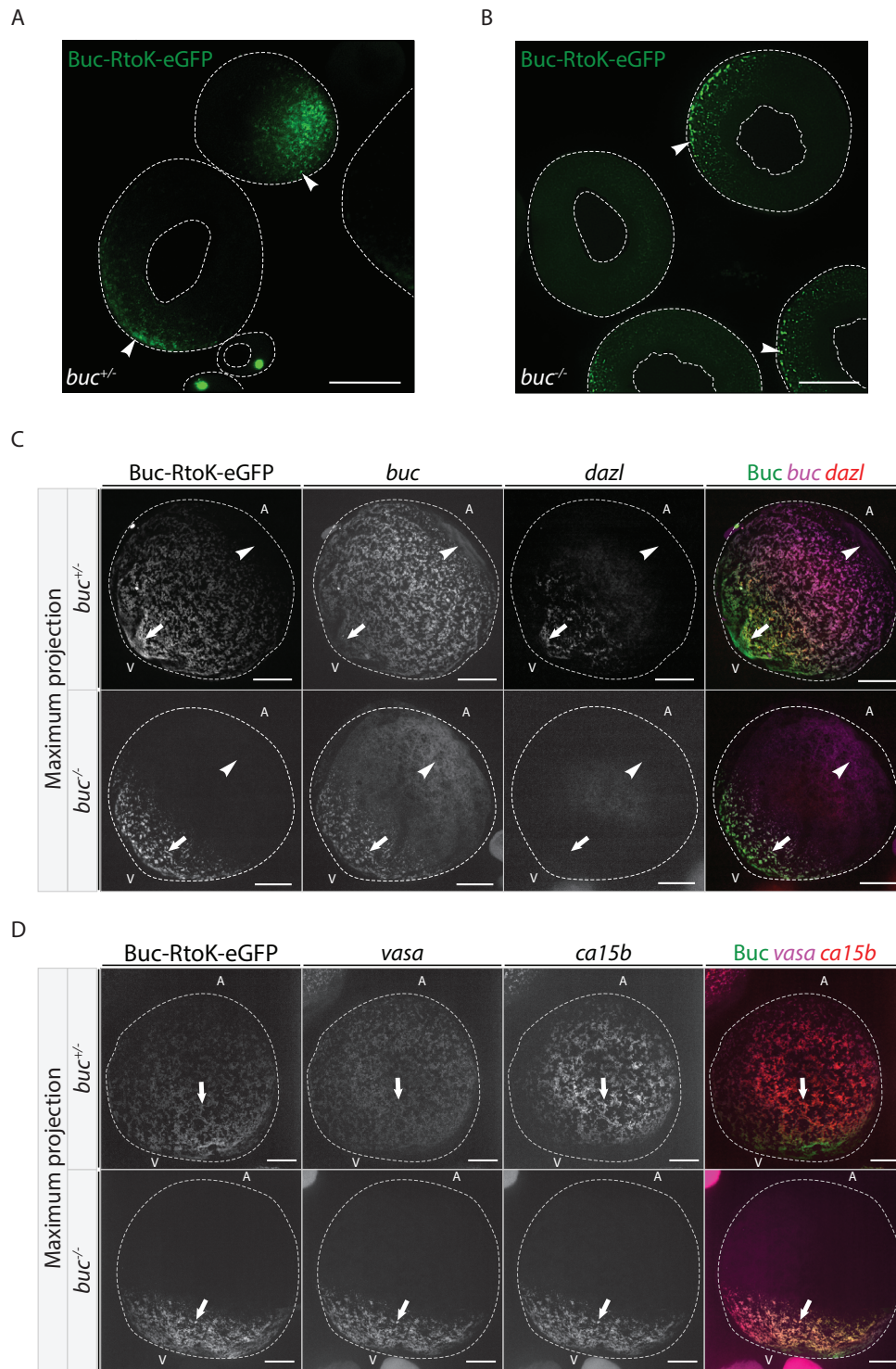
RESULTS

Buc-RtoK localizes to germinal granules during the late vegetal pathway

Without wild-type Buc protein, Buc-RtoK forms small granules before it becomes cytoplasmic as the oocytes grow and mature further (Roovers et al., 2018). Since AV-polarity is maintained, we investigated what happens during the late vegetal pathway and the animal pole transport pathway in such Buc-RtoK embryos. We first examined Buc-RtoK localization itself. In *buc*^{+/-} oocytes, Buc-RtoK followed the typical dispersion pattern of wild-type Buc towards the vegetal pole (Figure 1A, arrowheads). To our surprise, we observed that in *buc*^{-/-} oocytes, Buc-RtoK re-condenses at the onset of the late vegetal pathway (Figure 1B, arrowheads). These structures closely resemble germinal granules in a heterozygous or wild-type context, indicating that the late vegetal pathway still takes place in Buc-RtoK/*buc*^{-/-} oocytes, even though a Bb never forms.

Several early pathway transcripts localize during the late pathway in Buc-RtoK/*buc*^{-/-}

We next looked at the localization behavior of various early vegetal pathway transcripts, which are normally found in the Bb before they are transported vegetally in order to form germinal granules during the late pathway. We performed double smFISH at these stages in order to investigate their localization compared to the granules marked by Buc-RtoK. Stacks were taken sagittally through approximately half the oocytes and revealed that germinal granules form thread-like structures and together form a large network spanning the cortex (maximum projections, Figure 1C, D). In *buc*^{+/-} oocytes, transcripts of both *buc* and *dazl* localize to this network, however, *dazl* mRNA is more concentrated at the vegetal side than *buc* mRNA (Figure 1C and S1A, arrow). *Buc* mRNA is also found at the animal pole as published previously (Figure 1C and S1A, arrowhead) (Bontems et al., 2009). The animal pole localization is different from its vegetal localization since it is not granular but rather cytoplasmic (Figure 1C and S1A, arrowhead). Buc-RtoK protein is found throughout the granular network, but not at the animal pole. In *buc*^{-/-} oocytes, *dazl* localization is lost (Figure 2C and S1A, arrow). *Buc* is still found in the germinal granules (Figure 2C and S1A, arrow) and at the animal pole, however, the animal pole region is much broader compared to *buc*^{+/-} oocytes (Figure 2C and S1A, arrowhead). Two other early pathway components, *vasa* and a novel early pathway marker *ca15b* (Figure S1B), localized to the germinal granules marked by Buc-RtoK in both *buc*^{+/-} and *buc*^{-/-} oocytes (Figure 1D and S1C, arrows). These experiments revealed that in embryos with only Buc-RtoK protein, the animal pole region extends significantly and some, but not all mRNAs that normally localize to the vegetal pole, no longer do so.



Tdrd6a localizes to Buc-RtoK-containing granules during the late pathway

The Tudor domain-containing protein Tdrd6a has been demonstrated to localize to structures containing Buc, including the Bb and the germ plasm (Roovers et al., 2018). Tdrd6a has strong affinity for dimethylated arginine residues at the Buc C-terminus that are mutated in Buc-RtoK (Roovers et al., 2018). We therefore next studied localization of Tdrd6a in the Buc-RtoK/*buc*^{-/-} oocytes during the late vegetal pathway as well, using immunohistochemistry (IHC) against Tdrd6a. Interestingly, even in the absence of a wild-type copy of Buc and therefore an intact interaction site, we find Tdrd6a localizing to the same granules as Buc-RtoK, similar to its localization in the *buc*^{+/+} background (Figure 2). Given that we previously showed that Buc-RtoK does not bind to Tdrd6a anymore, the presence of Tdrd6a in these Buc-RtoK granules likely is driven by interactions with other proteins than Buc. The protein Buc2-like, that we identify below, is a good candidate for this.

IP/MS reveals novel Buc interactors Buc2-like, Alg13, Dazl and Mad211

In order to gain more insight into the Buc-RtoK phenotype, we performed an immunoprecipitation followed by label-free quantitative mass spectrometry (IP/MS) on Buc-RtoK ovaries in a *buc*^{+/+} and *buc*^{-/-} background. We also compared this to IP/MS on the wild-type Buc-eGFP transgene (Riemer et al., 2015). With standard IP buffer as described before (Roovers et al., 2018), we did not find strong enrichments for Buc itself or its potential interactors in such experiments (data not shown). Because of the tendency of Buc to aggregate, like its homolog Xvelo, we next performed IPs in the presence of 500mM L-arginine in order to have more Buc available in solution (Boke et al., 2016; Tsumoto et al., 2004). In these conditions we consistently found strong enrichments for Buc, indicating the IP was efficient and specific (Figure 3A, 3B and S2A). Interestingly, the different genetic backgrounds (*buc*^{+/+} or *buc*^{-/-}) and transgenes (producing either wild-type Buc or Buc-RtoK) did not display strong differences but all revealed robust interactions with four main candidates: Mad211, Alg13, Buc paralog Buc2-like and Dazl (Figure 3A, 3B and S2A). Interestingly, just like *dazl* mRNA, which was lost from Buc-RtoK germinal granules in *buc*^{-/-} embryos, Dazl enrichment lost significance in the *buc*^{-/-} background (Figure 3B).

Mad211 (*Mitotic arrest deficient*) is a spindle checkpoint protein which gets recruited to unattached kinetochores during prometaphase in mitosis, but has also been demonstrated to

◀ **Figure 1. Buc-RtoK(*buc*^{-/-}) localizes to germinal granules during the late pathway.** (A) Spinning disk image of Buc-RtoK-eGFP positive oocytes in the presence of a wild-type copy of *buc*. (B) Confocal image of oocytes with condensed or non-condensed Buc-RtoK in the *buc* mutant background. (C, D) Representative examples of maximum projections of smFISH signals of early pathway components *buc* & *dazl* (C) and *vasa* & *ca15b* (D) in the Buc-RtoK/*buc*^{+/+} or Buc-RtoK/*buc*^{-/-} background as indicated. Most factors still localize to the Buc-RtoK granules, except for *dazl*. Scalebars represent 50µm.

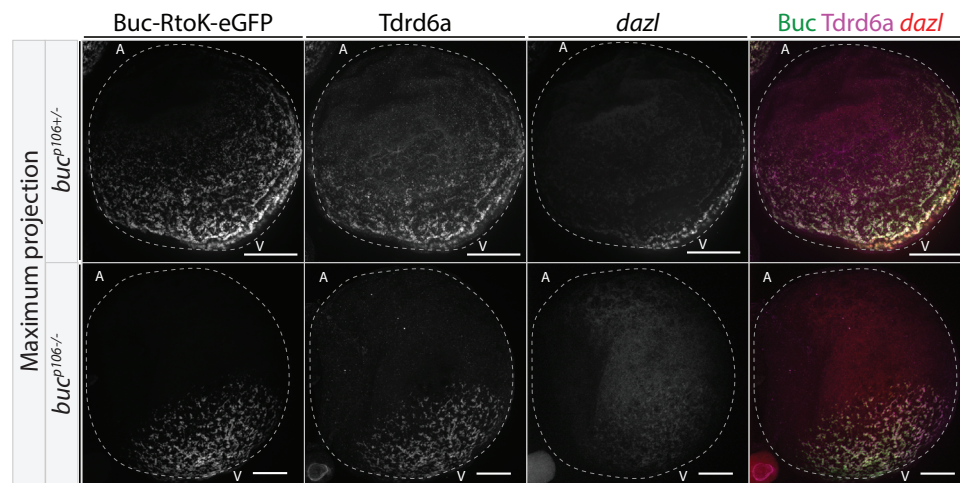


Figure 2. Localization Tdrd6a during the late pathway. Tdrd6a localizes to germinal granules marked by Buc-RtoK in both the *buc*^{+/+} background as well as in the *buc*^{-/-} background. Similar to Figure 1, *dazl* is not detected in the *buc*^{-/-} background. Scalebar represent 50μm.

be involved in meiosis in mouse (Homer et al., 2005; Wassmann et al., 2003). Interestingly, Mad211 has also been detected on spindle poles (Howell et al., 2000). Alg13 (Otu/Z-OTU) is expressed in zebrafish ovary and early embryogenesis and contains an OTU-like domain, a deubiquitinase domain, with an intact catalytic triad (Mo et al., 2005). Furthermore, Alg13 contains a Tudor domain, a domain that is involved in protein-protein interactions via symmetrically dimethylated arginine (sDMA) residues present on their interaction partners (Côté and Richard, 2005; Sprangers et al., 2003; Tripsianes et al., 2011). Dazl is an RNA-binding protein that recognizes motifs in the 3'UTR of germline-specific transcripts, thereby protecting them from miR-430 degradation, comparable to Dnd function (Kedde et al., 2007; Takeda et al., 2009). Additionally, Dazl increases translation specificity of its targets when this is required in PGCs (Takeda et al., 2009). Two transcripts demonstrated to interact with Dazl are *dazl* and *buc* (Heim et al., 2014). Lastly, Dazl contains a DAZ-like motif with which it can interact with Daz-associated proteins (Dazap), Dazap1 and Dazap2 (Tsui et al., 2000). Buc2-like is a Buc paralog that is widespread among teleosts, but does not contain any annotated protein domain (Škugor et al., 2016). Not much is known thus far besides that its homolog in salmon (Buc2a) localizes to the Bb, just as salmon Buc (Škugor et al., 2016).

Novel Buc interactors are disordered proteins of which some have PLDs

Buc is a protein with an N-terminal PLD and is largely disordered, according to predictions by the PLAAC and RONN algorithms (Figure S2B) (Lancaster et al., 2014; Yang et al., 2005). We next examined whether any of the MS candidates has either disordered regions or PLDs

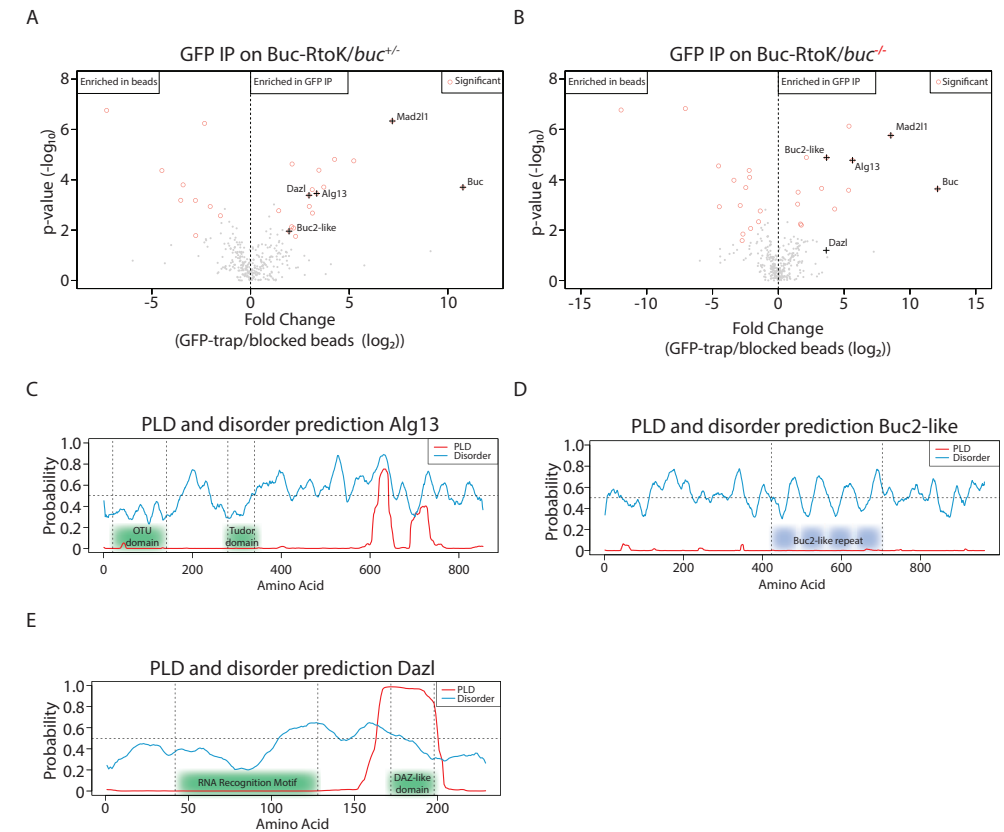


Figure 3. Buc interacts with Alg13, Buc2-like, Dazl and Mad211. (A, B) pulldowns using anti-GFP nanobody-coupled beads, compared to uncoupled, blocked beads in the *buc*^{+/+} (A) and *buc*^{-/-} background (B). (C-D) Protein order prediction (RONN) and PLD prediction (PLAAC) for Alg13 (C), Buc2-like (D) and Dazl (E).

as well. Besides Mad211, all candidates identified in the Buc IPs were largely disordered (Figure 3C-E). Moreover, both Alg13 and Dazl have regions that are predicted to be prion-like (Figure 3C, E). Interestingly, the predicted PLD of Dazl covers the entire conserved DAZ-like domain (Figure 4C). This domain is required for its interaction with DAZ associated proteins Dazap1 and Dazap2, both of which contain PLDs as well (Figure S2C) (Tsui et al., 2000). Buc2-like does not contain a PLD, however, we identified a repeated motif of unknown function in its sequence (Figure 3D and S2D). We conclude that Buc can interact with several proteins that are disordered, prion-like or both, like Buc itself.

Buc interactors co-localize with Buc in BmN4 cells

We have previously demonstrated that Buc forms granules when transfected into silkworm ovary-derived BmN4 cells and that these granules are enlarged or dissolved in the presence of Tdrd6a

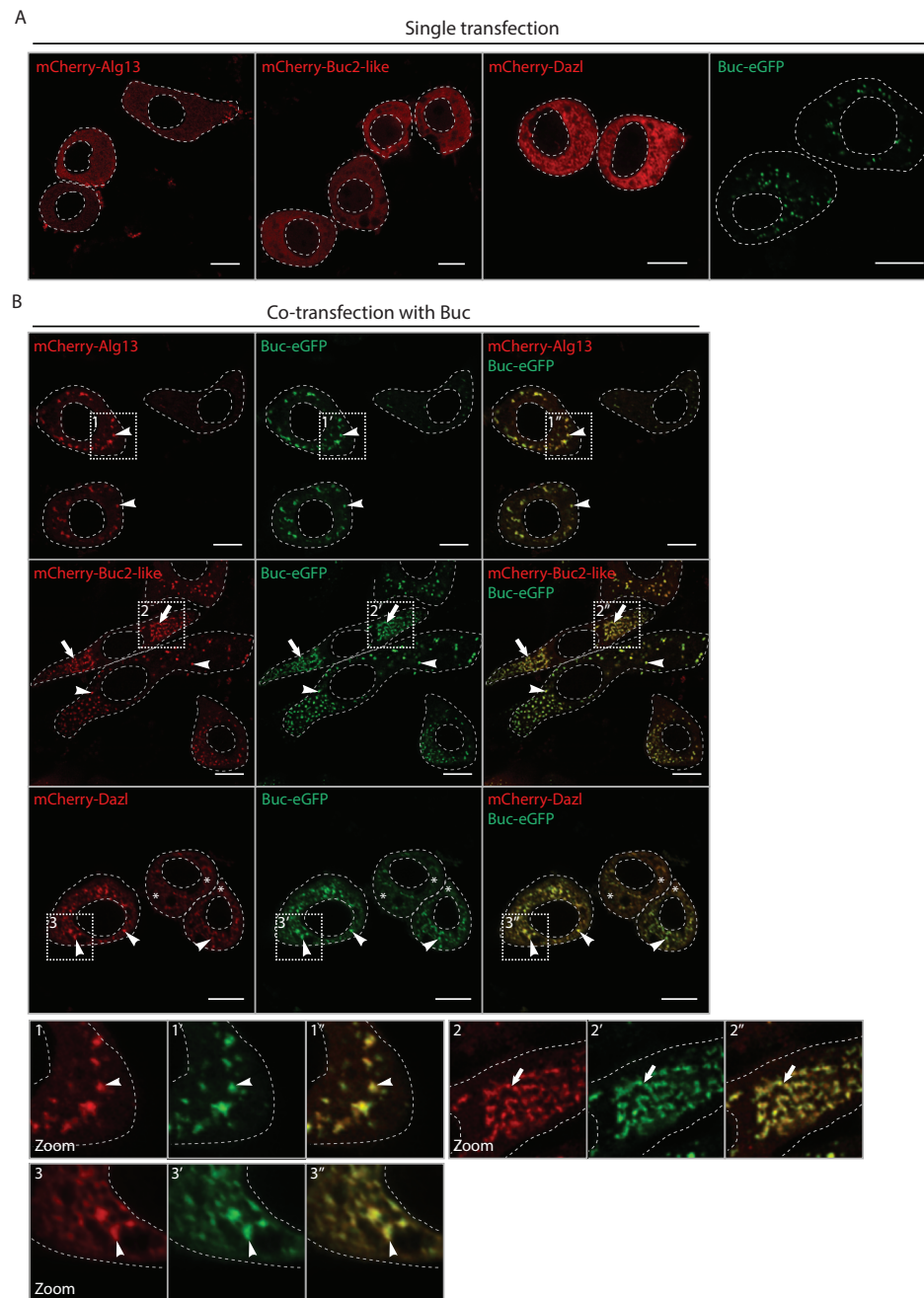


Figure 4. Buc interactors are disordered and PLD-containing proteins that co-localize with Buc. (A) Transfected BmN4 cells with MS candidates all localize throughout the cytoplasm, whereas Buc localizes to small, round granules. (B) Co-transfections with Buc recruits the candidates to Buc-structures (arrowheads). Furthermore, Buc2-like and Buc together form networks in most cells as well (arrows). For Buc2-like, examples of different scenarios (more granular and more threadlike) are displayed. In co-transfected Buc/Dazl cells, Buc is often more cytoplasmic rather than in granules (asterisks). Scalebars represent 10 μm.

(Roovers et al., 2018). Buc protein became more mobile within the granules in the presence of Tdrd6a, which mimicked the *in vivo* situation (Roovers et al., 2018). These results indicated that the environment in BmN4 cells allows us to study the behavior of Buc and other granule components *ex vivo*, and we used them to study the Buc-interacting proteins in more detail. We transfected mCherry-tagged versions of the IP/MS candidates Alg13, Buc2-like and Dazl and also performed co-transfections with Buc. We first confirmed that our expressed fusion constructs are of the predicted size (Figure S3). Next, we checked their subcellular localization. In single transfections Alg13, Buc2-like and Dazl all localize throughout the cytoplasm and no granule formation was observed (Figure 4A). Buc alone localizes to small, round granules as reported previously (Figure 4A) (Roovers et al., 2018). When co-transfected with Buc, all candidates localize to the Buc-containing structures, which become more amorphous in the presence of our candidates, compared to the round granules of Buc in single transfections (Figure 4B, zoom, arrowheads). The presence of Dazl also stimulates Buc granules to become more cytoplasmic (Figure 4B, asterisks). Interestingly, the presence of Buc2-like causes the Buc structures to become thread-like and form networks in the majority of the co-transfected cells, which resemble the structures that are formed during the late vegetal pathway (Figure 4B, zoom, arrow). We did not observe such structures in cells expressing only Buc, or Buc together with other Buc-interacting proteins. We conclude that members of the Buc-complex affect the morphology of Buc granules in cell culture, which can potentially contribute to our understanding of the behavior of zebrafish germ granules.

Tdrd6a interacts with novel Buc interactors in a Buc-dependent manner

We previously identified strong enrichments of Buc in IPs of Tdrd6a (Roovers et al., 2018). We therefore reanalyzed Tdrd6a IPs and found the novel Buc-interacting proteins identified above strongly enriched in Tdrd6a IPs as well, except for Dazl, which was not detected (Figure 5A). Additionally, we find enrichments for Rbpms2a and Rbpms2b, which were not found in the Buc IPs, but known to localize to the Bb and interact with Buc (Figure 5A) (Heim et al., 2014; Kaufman et al., 2018). When we perform a Tdrd6a IP in the *buc* mutant background, we lose all candidates of the Buc-complex, as well as Rbpms2a and Rbpms2b. This suggests that the interaction of Tdrd6a with the Buc-complex, including Rbpms2a and Rbpms2b, requires Buc or the structural presence of the Bb and/or germinal granules in order to be stable (Figure 5B and 5C). Furthermore, Tdrd6a interacted independently of Buc with the Exon Junction Complex (EJC) (Magoh, Y14, eIF4A3, Btz) and with RNA binding proteins Elavl1 and Elavl2 (Figure 5B and 5C).

Finally, we compared the Tdrd6a interacting proteins in *buc*^{-/-} embryos that either do, or do not express Buc-RtoK. This comparison should give insights into the late vegetal pathway granules that are found in the Buc-RtoK expressing embryos, without seeing interactions that take place within the Bb, which lacks in both samples. Not many proteins are enriched in this analysis, but Mad211 and Buc2-like are, suggesting they may play a role in the late vegetal

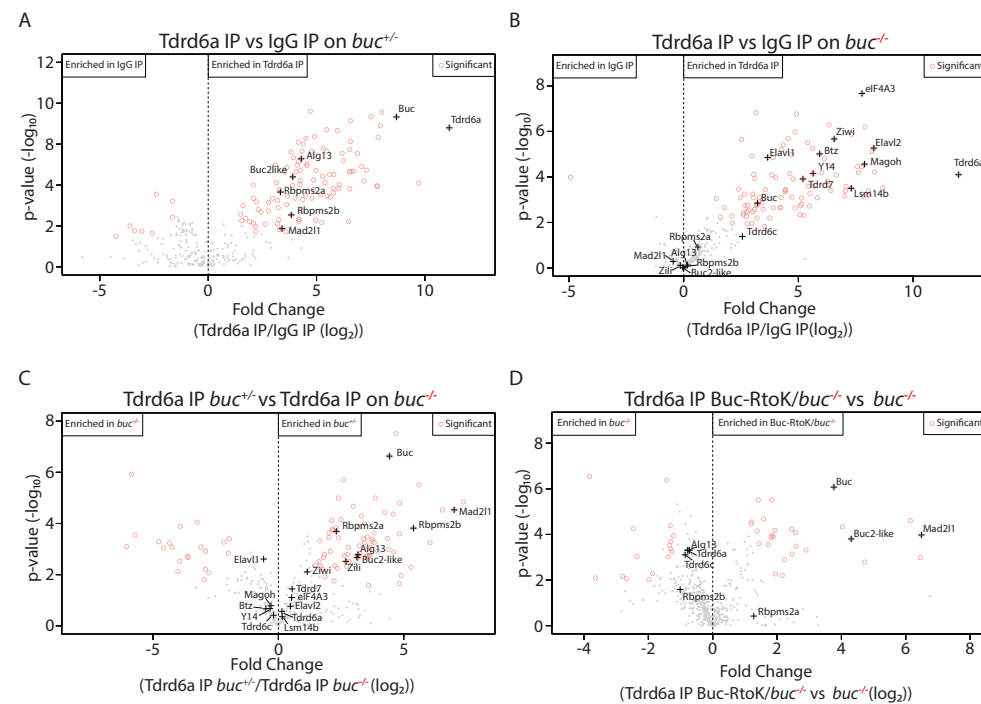


Figure 5. Tdrd6a co-precipitates with the Buc-complex. (A-D) Tdrd6a IPs in different genetic backgrounds as indicated, compared to an IgG control (A, B) or compared to different backgrounds (C, D). Without a wild-type copy present, the interaction with Buc is less strong than Buc2-like, indicating that Buc2-like could recruit Buc-RtoK (D).

pathway (Figure 5D). Two other Buc-interacting proteins, Alg13 and Dazl were not enriched. Surprisingly, we also detected Buc in the Buc-RtoK embryos in the Tdrd6a IPs, even though it should not bind to Tdrd6a directly anymore (Figure 5D). We hypothesize that this interaction between Tdrd6a and Buc-RtoK is indirect via Buc2-like. Our finding that Buc2-like and Buc indeed interact in BmN4 cells supports this idea.

DISCUSSION

During oogenesis, many patterning events take place that are essential for viability of the progeny. Two vegetal localization pathways are known to occur in zebrafish and *Xenopus*, during which factors are recruited to the vegetal side. Additionally, an animal localization pathway is in place for animal pole identity. Vital events that depend on these pathways are the formation of a single micropyle, the accumulation of intact germ plasm and the localization of other patterning factors, like *wnt8* and *syntabulin* (Bontems et al., 2009; Heim et al., 2014; Lu et al., 2011; Marlow and Mullins, 2008; Nojima et al., 2010).

In wild-type oocytes, the chromosomal bouquet and the centrosome are upstream of these events, followed by Bb formation, a structure that depends on Buc. Formation of the Bb has been regarded to be a key event in polarity maintenance. In *buc* null mutants, these pathways are drastically compromised and axis formation fails, resulting in non-viable eggs. We show that the Buc-RtoK transgene in a *buc* mutant background can rescue polarity and that the late vegetal pathway can occur without a Bb. Below, our findings will be further discussed.

Buc-RtoK enables the late vegetal pathway in absence of a Bb

Without a wild-type copy of Buc, Buc-RtoK still phase-separates into small granules early on in oogenesis, but fails to become stabilized into a Bb and becomes cytoplasmic (Roovers et al., 2018). Interestingly, we found that there exists another trigger for Buc-RtoK to phase-separate and localize to the vegetal pole by the end of oogenesis stage I. This indicates that the late vegetal pathway can be rescued by Buc, independent of Tdrd6a-sDMA interactions and a Bb. This is in agreement with the rescued polarity of the progeny of Buc-RtoK mothers, in which axis positioning is maintained (Roovers et al., 2018).

In a wild-type situation, the late vegetal pathway is initiated following transition of the Bb into a vegetally dispersed pattern. Many factors from the early pathway follow this re-localization and therefore overlap with the late vegetal pathway, such as Buc and Tdrd6a, and transcripts like *dazl* and *vasa*. It could therefore be that in order to follow vegetal localization, early pathway factors rely on the Bb structure in order to be transported vegetally, especially since in the absence of Buc in *buc*^{-/-}, and therefore a Bb, not only the early, but also the late pathway does not take place.

We found that different early pathway components displayed dissimilar localization. For instance, *dazl* transcripts localize only to a restricted vegetal region and not to all granules, while *buc* transcripts display both granular and cytoplasmic localization. Furthermore, in the Buc-RtoK/*buc*^{-/-} background, some early pathway components – *buc*, *vasa* and *ca15b* – still localized to germinal granules, whereas *dazl* failed to localize. These data suggest that individual transcripts have different localization cues. This results in different sub-structural localization patterns and partial loss of localization of transcripts in Buc-RtoK/*buc*^{-/-} oocytes.

Certain proteins and transcripts could depend on interactions of Buc with Tdrd6a paralogs during the early pathway. Some factors were recruited by Tdrd6a independently of Buc, like members of the EJC and Elav1 and Elav2. In *Drosophila* and *Xenopus* oocytes, these proteins are required for localization of germ plasm (Arthur et al., 2009; Hachet and Ephrussi, 2001; Micklem et al., 1997) and patterning factors which could therefore potentially be depleted or lost without Tdrd6-Buc interactions during the early vegetal pathway in zebrafish. The absence of *dazl* and possibly other late vegetal pathway components from the germinal granules could well explain the downstream developmental defects of embryos derived from Buc-RtoK/*buc*^{-/-} mothers (Roovers et al., 2018). Regardless of apparent recovery during the late pathway, Buc-RtoK is still insufficient for embryonic germ plasm formation and normal embryo patterning (Roovers

et al., 2018), since particular critical components may not have been localized and/or stabilized properly during oocyte development.

Mad211 as a potential link between Buc and the centrosome

In IP/MS experiments, we found that Mad211 consistently co-precipitated with Buc-RtoK. In wild-type Buc IPs Mad211 came down as well, indicating this is not an artefact caused by the Buc-RtoK phenotype. Mad211 is mostly known for its role as a checkpoint protein, since it can bind unattached kinetochores together with Mad1 and can prevent progression into anaphase, by sequestering Cdc20 (Yu, 2002). Neither Cdc20 nor Mad1 was identified in the IPs, indicating Buc potentially interacts with a fraction of Mad211 that could function distinctly from its role in the mitotic or meiotic checkpoint. It has been reported that Mad2 can localize to γ -Tubulin-containing spindle poles depending on microtubules during prometaphase (Howell et al., 2000). Mad211 could therefore provide a link between Buc and the centrosome. Since the centrosome is upstream of the Bb, it is likely that at least one Bb component can associate with the centrosome in order to recruit other components. Future experiments, involving IHC in early stage I oocytes, and biochemical dissection using BmN4 cells, should provide more insights on such a potential link between the centrosome and the Bb and whether this depends on Mad211.

Buc IP/MS reveals candidates involved in granule morphology

In the other identified Buc-interactors – Alg13, Buc2-like and Dazl – we found disordered regions and predicted PLDs. These could be involved in driving Buc-RtoK into phase separated structures during the late pathway. Phase-separated structures typically contain multiple disordered and PLD-containing proteins in order to fine-tune their aggregation properties. We demonstrated that in BmN4 cells, our candidates all localize to granules in the presence of Buc and affect Buc-granule morphology. It is unclear how their behavior is affected by Buc, but since our candidates do not phase-separate in single transfections, this is somehow triggered by the presence of Buc. It is possible that this is mediated through their disordered regions or PLDs, which should be investigated further by studying granule behavior in deletion mutants of our candidates.

Alg13, another Tudor domain-containing protein that modulates Buc granules

One candidate that could modulate the germ plasm aggregate is Tudor domain-containing protein Alg13. Proteins with Tudor domains have been demonstrated previously to alter granule behavior and composition, like Tdrd7 in structures in the mouse lens and nuage, and Tdrd6a in the Bb of zebrafish (Lachke et al., 2011; Roovers et al., 2018; Tanaka et al., 2011). In BmN4 co-transfections, we found that Alg13 alters the Buc granules, and therefore possibly affects *in vivo* germ plasm structures, containing Buc, as well.

It would be interesting to see whether Alg13 requires its Tudor domain in order to associate with the Buc granules. In our Tdrd6a IP/MS in the Buc-RtoK/*buc*^{-/-} background, Alg13 was not significantly enriched, which could indicate that Alg13 requires an intact tri-RG motif, an interaction site for Tudor domains, at the Buc C-terminus in order to co-precipitate. Furthermore, its deubiquitinase domain could be involved in germ granule dynamics, for instance by preventing degradation of ubiquitylated targets, or by affecting phase-separating properties directly, since the presence or absence of ubiquitin can influence this as well (Dao et al., 2018; Herhaus and Dikic, 2018; Zhang et al., 2017).

The RNA-binding protein Dazl localizes to Buc granules

Dazl is a protein with RNA-binding properties and a PLD. It has been shown to interact with *dazl* transcript in ovaries and regulate *tldr7* and *elavl2* transcripts in embryos (Heim et al., 2014; Takeda et al., 2009; Wiszniak et al., 2011). In addition to temporary storage and protection of mRNA in subcellular granules, RNA itself can function as a structural component of these condensates as well. This way, Dazl could contribute to the early and the late pathway through its own PLD and that of Dazap1 and Dazap2, but also by recruiting and accumulating transcripts in these granules.

In BmN4 cells, Dazl localized to the Buc-containing granules, but also seemed to be able to solubilize these granules. It is possible that this is due to overexpression since Dazl expression was very high. Expression under weaker promoters could therefore provide a better protein balance in order to study its effect on Buc granules more subtly.

Dazl transcript localization and potentially stability was not rescued by Buc-RtoK, indicating that *dazl* relies more on the Bb or Tdrd6 paralogs compared to the other candidates. Furthermore, Dazl protein was also not significantly enriched in the Buc-RtoK/*buc*^{-/-} IP. It is possible that loss of the interaction with Dazl results in loss of *dazl* from the germinal granules and that the IP result reflects its destabilization in general. Since the IP/MS was performed on total ovary lysates, this possibility can be explored in more detail by performing the IP stage-specifically.

A possible function of Buc2-like in the late vegetal pathway

The disordered protein Buc2-like, a paralog of Buc, was co-precipitating with Buc and Tdrd6a. Moreover, in Tdrd6a IPs on Buc-RtoK/*buc*^{-/-} oocytes, in which only the late pathway occurs, we found that Buc2-like strongly associated with Tdrd6a. This could indicate that Buc2-like is an interactor of Tdrd6a in germinal granules specifically. Two RG-motifs are present in Buc2-like, however, it is unknown whether these motifs are involved in direct interaction and recruitment of Tdrd6a during the late vegetal pathway, in the absence of the C-terminal interaction motif of Buc.

Buc2-like could alter Buc granules into a thread-like network in BmN4 cells, which makes Buc2-like an interesting candidate for germ granule organization during the late vegetal pathway. This can be examined further by dissecting the role of different domains of Buc2-like in BmN4

cells, like the Buc2-like repeat motif. Furthermore, it will be interesting to see whether the late pathway is abolished specifically in *buc2-like* mutants, or if its absence alters germinal granule content and/or morphology.

Concluding remarks

Zebrafish germ plasm, which ultimately gets deposited in the embryo, starts to become shaped in the oocyte when Buc-containing granules accumulate around the centrosome during the bouquet stage. These events are followed by the formation of the Bb, which remains as one round structure while the oocyte grows from ~15µm to ~100µm in diameter. Next the Bb moves vegetally and changes from a single condensate into fragmented, thread-like structures, while the oocyte keeps expanding (up to ~700µm in diameter at the moment of the egglay). These substantial changes in oocyte volume and organization require tight control over complex processes such as germ plasm formation. The oocyte therefore would need to regulate spatio-temporal expression of factors that are involved in for instance adjusting the Bb 'phase' into a fragmented germ granule 'phase'. Our results indicate that the novel Buc interactors are promising candidates for future research of Bb and germinal granule composition and regulation. How the late pathway is triggered and which components are essential for germ plasm stabilization such that it sustains until after fertilization remains unknown. Studying mutants like Buc-RtoK, in which various aspects of Buc-aggregation are affected differentially, will be important to further understand how these events are triggered and coordinated.

MATERIALS AND METHODS

Immunohistochemistry and double smFISH

Ovaries were dissected and dissociated in cold PBS. They were fixed in 4%PFA/PBS-0.5%Triton-X100 overnight at 4°C, rotating. The next day, they were washed with PBS-Triton and put in 100% MeOH. They were kept at -80°C for at least 3 hours. Next, the samples were put in PBST followed by overnight incubation with Rb-anti-Tdrd6a and the smFISH probes in smFISH hybridization buffer (10% dextran sulfate, 10% formamide, 1mg/mL tRNA, 0.02%BSA, 2mM vanadyl-ribonucleoside complex (NEB S1402S) in 2xSSC) at 30°C. The next day, the samples were washed once with smFISH wash buffer (10% formamide, 2xSSC) followed by 45 minutes incubation with Gt-anti-Rb-Alexa555. This was followed by two more washes of 30 minutes before the samples were mounted in GLOX buffer and imaged. The oocytes were imaged on ibidi 8-well µ-slides on a Leica Sp5 confocal, using a 63x water objective and for 3D imaging on the Visitron Spinning Disk, using a 40x water objective. The confocal pictures of Figure 1A, 4 and Supplemental Figure 1 were deconvolved using the Huygens software.

Immunoprecipitation

Ovaries were dissected after laying, in order to make the tissues more comparable. They were taken up in IP buffer with L-Arginine (25 mM Tris pH 7.5, 150 mM NaCl, 1.5 mM MgCl₂, 1% Triton-X100, 1mM DTT, protease inhibitor and in case of the Buc IPs: 500mM L-Arginine

monohydrochloride) and mashed with a pestle followed by sonication (3 x 30 seconds). The samples were spun down 10 minutes at 12,000 x g at 4°C and the supernatant was collected. Next, the samples were incubated with Magnetic-Agarose GFP binder beads (in house production, 25µL slurry per IP) and control samples with Magnetic-Agarose control beads (blocked with L-Cysteine, in house production, 25µL slurry per IP) for 1 hour at 4°C while rotating. For Tdrd6a IPs, 1:200 Rb-anti-Tdrd6a was added and samples were rotated for 1.5 hours at 4°C, followed by addition of 25µL pre-washed slurry Dynabeads Protein G (ThermoFisher) per IP. After incubation with the beads, samples were washed 3 times with wash buffer (500mM L-Arginine, 25 mM Tris pH 7.5, 300 mM NaCl, 1.5 mM MgCl₂, 1mM DTT) and afterwards, the beads were eluted in NuPAGE LDS sample buffer and used mass spectrometry analysis.

Mass spectrometry

Immunoprecipitates were resuspended in 1x NuPAGE LDS Sample Buffer (Thermo) supplemented with 0.1 M DTT (Sigma) and heated for 10 minutes at 70 °C. The respective samples were run on a 4%-12% NuPAGE Bis-Tris gradient gel (Thermo) in 1x MOPS buffer (Thermo) at 180 V for 10 min. The samples were processed by in-gel digestion as essentially previously described (Shevchenko et al., 2007) and desalted using C18 StageTips (Rappsilber et al., 2007).

The digested peptides were separated by an EASY-nLC1000 uHPLC system with a C18 column (25 cm long, 75 µm inner diameter, packed in-house with ReproSil-Pur C18-AQ 1.9 µm (Dr. Maisch GmbH)) and sprayed via ESI online into a Q Exactive Plus mass spectrometer (Thermo). A 2h MS run consisting of a gradient from 2% to 40% acetonitrile and a washout of 95% acetonitrile at a flow rate of 200 nl/min was used. The instrument performed a top10 data-dependent acquisition with up to 10 HCD fragmentations per MS full scan (70k resolution, 300-1650 m/z).

Raw files were processed using MaxQuant v.1.5.2.8 (Cox and Mann, 2008) searching against a Uniprot/Trembl *Danio rerio* (www.uniprot.org) fasta file. Carbamidomethylation was set as fixed modification while methionine oxidation, protein N-acetylation, phosphorylation (STY) and lysine/arginine dimethylation were considered as variable modifications. The false discovery rate for peptide and protein identifications was set to 0.01. LFQ quantitation was performed with standard settings and the match between runs option activated.

The resulting proteome data was filtered for at least two peptide identifications per protein group, one unique and one razor. Protein Groups identified by site, known contaminants and reverse hits were disregarded. LFQ values were then log₂ transformed. The mean LFQ intensities of each experimental condition (measured in quadruplicates) as well as the p-values (Welch t-test) comparing these conditions were calculated and used for the determination of enriched proteins.

Cloning

Dazl, *buc2l* and *alg13* were amplified from zebrafish ovary cDNA with the oligos as indicated below and amplicons were digested with NotI and XbaI. Next, they were ligated in pre-digested pBEMBL-NHA-mCherry, in order to obtain N-terminally tagged fusion constructs, which were confirmed by sequencing.

BmN4 cell culture

BmN4 cells were kept at 27°C in IPL-41 (Gibco) medium with added 10% FBS (Gibco) and 0.5% Pen-Strep. Imaging was performed as described previously (Roovers et al., 2018). For western blot, cells were collected and spun down at 500 x g for 5 minutes. The pellet was washed with ice-cold PBS and spun down again at 500 x g for 5 minutes at 4°C. Next, the pellets were taken up in lysis buffer (50 mM Tris pH 8.0, 150 mM NaCl, 5 mM MgCl₂, 0.5% Triton X-100, 10% glycerol, freshly added 1mM DTT and Protease inhibitor) and sonicated. The lysates were spun down at 12,000 x g at 4°C and the supernatant was collected and measured. 20µg protein was loaded on a 4%-12% NuPage NOVEX gradient gel (ThermoFisher) and blotted on an Immobilon-FL PVDF membrane (Merck) overnight at 15V, RT. The membrane was incubated with 1:1000 Ms-α-HA, washed and incubated with 1:10,000 IRDye 680RB Gt-α-Ms (LI-COR). The membrane was imaged using an Odyssey CLx imaging system (LI-COR).

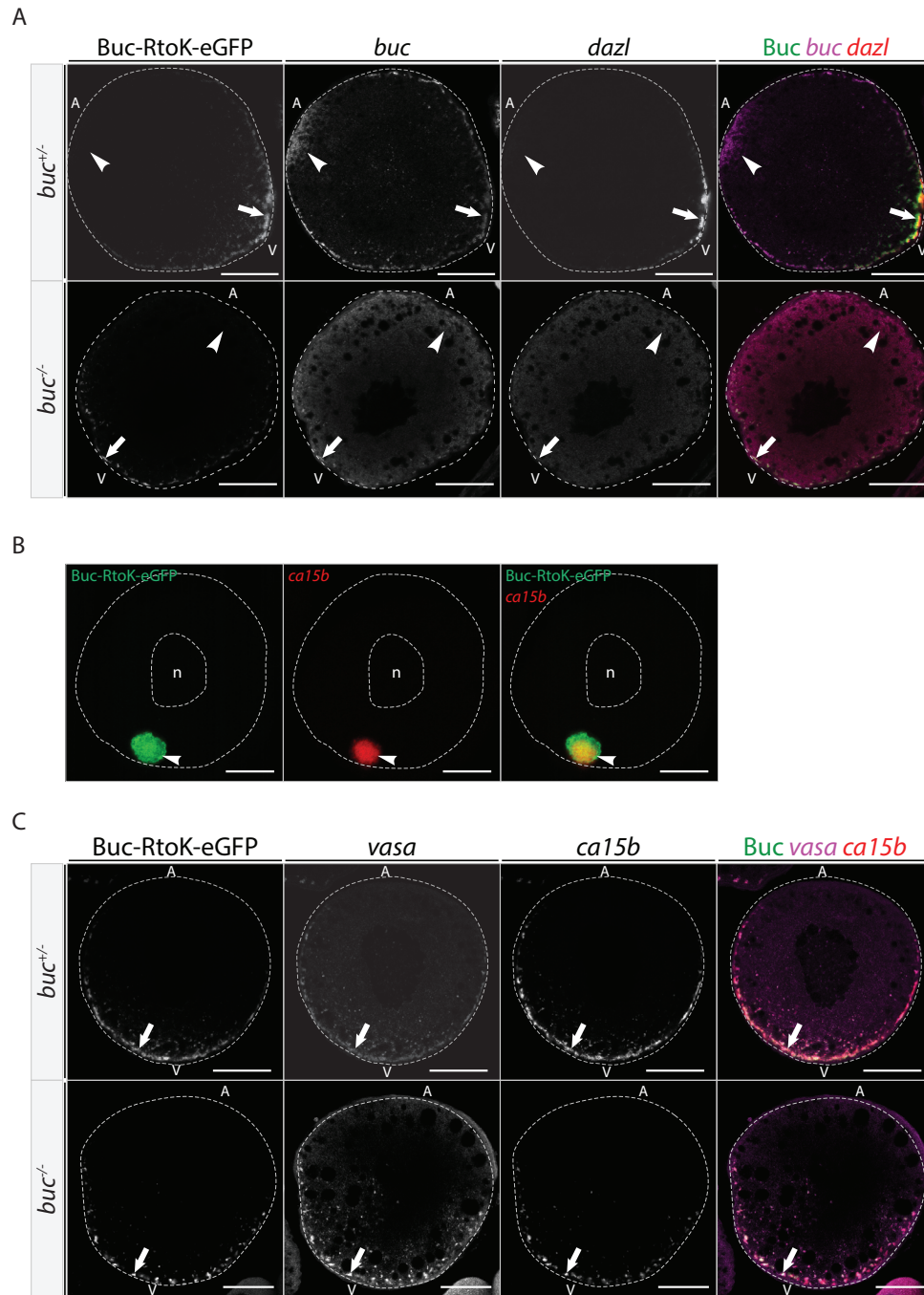
Oligo	sequence
dazl_NotI_F	gggcggccgcC GTT CAG GGG GTT CAG TTA CCC GTG TG
dazl_XbaI_R	gcgtctagaCAT AAG GGT TAG CAA AGT CTG CAC TCC
buc2l_NotI_F	gggcggccgcCGCT GCT ATA TCT ACT CAT GAA CCC ATT G
buc2l_XbaI_R	gcgtctagaTTC ACG GCC AGT GAC ATG TTC CCC
alg13_NotI_F	gggcggccgcCCAA AAA GCC TTG AAG AAG TAT TTT GTA AGC
alg13_XbaI_R	gcgtctagaGAC AAA CTG TGA GTG TGG CAG ATG GT

REFERENCES

- Abrams, E.W., and Mullins, M.C. (2009). Early zebrafish development: It's in the maternal genes. *Curr. Opin. Genet. Dev.*
- Arthur, P.K., Claussen, M., Koch, S., Tarbashevich, K., Jahn, O., and Pieler, T. (2009). Participation of *Xenopus* Elr-type proteins in vegetal mRNA localization during oogenesis. *J. Biol. Chem.*
- Boke, E., Ruer, M., Wühr, M., Coughlin, M., Lemaitre, R., Gygi, S.P., Alberti, S., Drechsel, D., Hyman, A.A., and Mitchison, T.J. (2016). Amyloid-like Self-Assembly of a Cellular Compartment. *Cell* 166, 637–650.
- 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.
- Brangwynne, C.P., Eckmann, C.R., Courson, D.S., Rybarska, A., Hoegge, C., Gharakhani, J., Julicher, F., and Hyman, A.A. (2009). Germline P Granules Are Liquid Droplets That Localize by Controlled Dissolution/Condensation. *Science* (80-.). 324, 1729–1732.
- Brangwynne, C.P., Mitchison, T.J., and Hyman, A.A. (2011). Active liquid-like behavior of nucleoli determines their size and shape in *Xenopus laevis* oocytes. *Proc. Natl. Acad. Sci.* 108, 4334–4339.
- Côté, J., and Richard, S. (2005). Tudor domains bind symmetrical dimethylated arginines. *J. Biol. Chem.*
- Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 26, 1367–1372.
- Dao, T.P., Kolaitis, R.M., Kim, H.J., O'Donovan, K., Martyniak, B., Colicino, E., Hehnl, H., Taylor, J.P., and Castañeda, C.A. (2018). Ubiquitin Modulates Liquid-Liquid Phase Separation of UBQLN2 via Disruption of Multivalent Interactions. *Mol. Cell.*
- 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.*
- Elkouby, Y.M., Jamieson-Lucy, A., and Mullins, M.C. (2016). Oocyte Polarization Is Coupled to the Chromosomal Bouquet, a Conserved Polarized Nuclear Configuration in Meiosis. *PLoS Biol.* 14.
- Escobar-Aguirre, M., Zhang, H., Jamieson-Lucy, A., and Mullins, M.C. (2017). Microtubule-actin crosslinking factor 1 (Macf1) domain function in Balbiani body dissociation and nuclear positioning. *PLoS Genet.* 13.
- Forristall, C., Pondel, M., Chen, L., and King, M.L. (1995). Patterns of localization and cytoskeletal association of two vegetally localized RNAs, Vg1 and Xcat-2. *Development* 121, 201–208.
- Franzmann, T.M., Jahnel, M., Pozniakovsky, A., Mahamid, J., Holehouse, A.S., Nüske, E., Richter, D., Baumeister, W., Grill, S.W., Pappu, R. V., et al. (2018). Phase separation of a yeast prion protein promotes cellular fitness. *Science* (80-.). 359.
- Gupta, T., Marlow, F.L., Ferriola, D., Mackiewicz, K., Dapprich, J., Monos, D., and Mullins, M.C. (2010). Microtubule actin crosslinking factor 1 regulates the balbiani body and animal-vegetal polarity of the zebrafish oocyte. *PLoS Genet.* 6.
- Hachet, O., and Ephrussi, A. (2001). *Drosophila* Y14 shuttles to the posterior of the oocyte and is required for oskar mRNA transport. *Curr. Biol.*
- 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.
- Heim, A.E., Hartung, O., Rothhamel, S., Ferreira, E., Jenny, A., and Marlow, F.L. (2014). Oocyte polarity requires a Bucky ball-dependent feedback amplification loop. *Development* 141, 842–854.
- Herhaus, L., and Dikic, I. (2018). Ubiquitin-induced phase separation of p62/SQSTM1. *Cell Res.*
- Homer, H.A., McDougall, A., Levasseur, M., Yallop, K., Murdoch, A.P., and Herbert, M. (2005). Mad2 prevents aneuploidy and

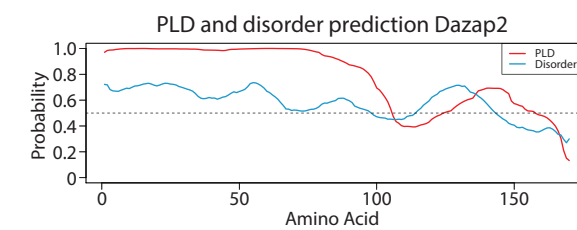
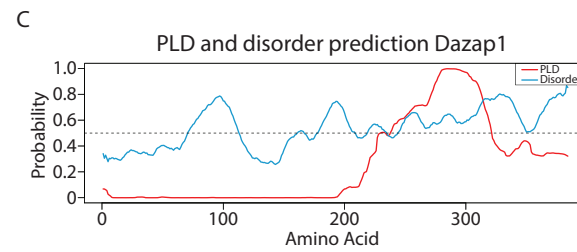
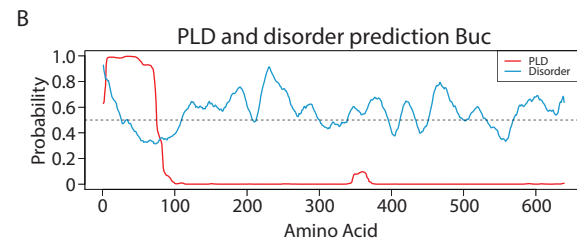
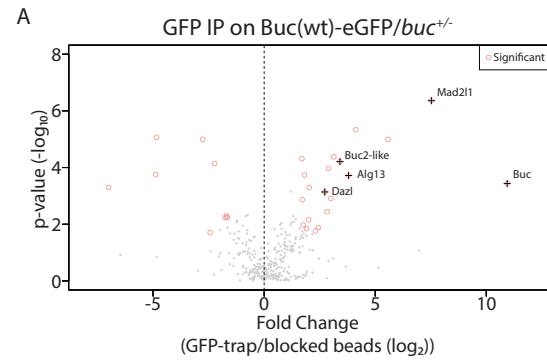
- premature proteolysis of cyclin B and securin during meiosis I in mouse oocytes. *Genes Dev.*
21. Howell, B.J., Hoffman, D.B., Fang, G., Murray, A.W., and Salmon, E.D. (2000). Visualization of Mad2 dynamics at kinetochores, along spindle fibers, and at spindle poles in living cells. *J. Cell Biol.*
 22. Kaufman, O.H., Lee, K.A., Martin, M., Rothhämel, S., and Marlow, F.L. (2018). *rbpms2* functions in Balbiani body architecture and ovary fate. *PLoS Genet.*
 23. 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.
 24. Kloc, M., and Etkin, L.D. (1995). Two distinct pathways for the localization of RNAs at the vegetal cortex in *Xenopus* oocytes. *Development.*
 25. 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.
 26. Kosaka, K., Kawakami, K., Sakamoto, H., and Inoue, K. (2007). Spatiotemporal localization of germ plasm RNAs during zebrafish oogenesis. *Mech Dev* 124, 279–289.
 27. 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.-H., et al. (2011). Mutations in the RNA granule component TDRD7 cause cataract and glaucoma. *Science* 331, 1571–1576.
 28. Lancaster, A.K., Nutter-Upham, A., Lindquist, S., and King, O.D. (2014). PLAAC: A web and command-line application to identify proteins with prion-like amino acid composition. *Bioinformatics.*
 29. Lu, F.-I., Thisse, C., and Thisse, B. (2011). Identification and mechanism of regulation of the zebrafish dorsal determinant. *Proc. Natl. Acad. Sci.*
 30. 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.
 31. Messitt, T.J., Gagnon, J.A., Kreiling, J.A., Pratt, C.A., Yoon, Y.J., and Mowry, K.L. (2008). Multiple Kinesin Motors Coordinate Cytoplasmic RNA Transport on a Subpopulation of Microtubules in *Xenopus* Oocytes. *Dev. Cell.*
 32. Micklem, D.R., Dasgupta, R., Elliott, H., Gergely, F., Davidson, C., Brand, A., González-Reyes, A., and St. Johnston, D. (1997). The mago nashi gene is required for the polarisation of the oocyte and the formation of perpendicular axes in *Drosophila*. *Curr. Biol.*
 33. Mo, S., Song, P., Lv, D., Chen, Y., Zhou, W., Gong, W., and Zhu, Z. (2005). Zebrafish *z-otu*, a novel Otu and Tudor domain-containing gene, is expressed in early stages of oogenesis and embryogenesis. *Biochim. Biophys. Acta - Gene Struct. Expr.*
 34. Molliex, A., Temirov, J., Lee, J., Coughlin, M., Kanagaraj, A.P., Kim, H.J., Mittag, T., and Taylor, J.P. (2015). Phase Separation by Low Complexity Domains Promotes Stress Granule Assembly and Drives Pathological Fibrillization. *Cell.*
 35. Nojima, H., Rothhämel, S., Shimizu, T., Kim, C.-H., Yonemura, S., Marlow, F.L., and Hibi, M. (2010). Syntabulin, a motor protein linker, controls dorsal determination. *Development.*
 36. Nott, T.J., Petsalaki, E., Farber, P., Jarvis, D., Fussner, E., Plochowitz, A., Craggs, T.D., Bazett-Jones, D.P., Pawson, T., Forman-Kay, J.D., et al. (2015). Phase Transition of a Disordered Nuage Protein Generates Environmentally Responsive Membraneless Organelles. *Mol. Cell* 57, 936–947.
 37. Rappsilber, J., Mann, M., and Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* 2, 1896–1906.
 38. Riemer, S., Bontems, F., Krishnakumar, P., Gömann, J., and Dosch, R. (2015). A functional Bucky ball-GFP transgene visualizes germ plasm in living zebrafish. *Gene Expr. Patterns* 18, 44–52.
 39. Roovers, E.F., Kaaij, L.J.T., Redl, S., Bronkhorst, A.W., Wiebrands, K., de Jesus Domingues, A.M., Huang, H.Y., Han, C.T., Riemer, S., Dosch, R., et al. (2018). *Tdrd6a* Regulates the Aggregation of Buc into Functional Subcellular Compartments that Drive Germ Cell Specification. *Dev. Cell.*
 40. Scherthan, H. (2001). A bouquet makes ends meet. *Nat. Rev. Mol. Cell Biol.*
 41. Shevchenko, A., Tomas, H., Havli[sbrev], J., Olsen, J. V, and Mann, M. (2007). In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* 1, 2856–2860.
 42. Škugor, A., Tveiten, H., Johnsen, H., and Andersen, Ø. (2016). Multiplicity of Buc copies in Atlantic salmon contrasts with loss of the germ cell determinant in primates, rodents and axolotl. *BMC Evol. Biol.*
 43. 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.
 44. Takeda, Y., Mishima, Y., Fujiwara, T., Sakamoto, H., and Inoue, K. (2009). DAZL relieves miRNA-mediated repression of germline mRNAs by controlling poly(A) tail length in zebrafish. *PLoS One.*
 45. 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 chromatin bodies during spermatogenesis. *Proc. Natl. Acad. Sci.* 108, 10579–10584.
 46. Tripsianes, K., Madl, T., MacHyna, M., Fessas, D., Englbrecht, C., Fischer, U., Neugebauer, K.M., and Sattler, M. (2011). Structural basis for dimethylarginine recognition by the Tudor domains of human SMN and SPF30 proteins. *Nat. Struct. Mol. Biol.* 18, 1414–1420.
 47. Tsui, S., Dai, T., Roettger, S., Schempp, W., Salido, E.C., and Yen, P.H. (2000). Identification of two novel proteins that interact with germ-cell-specific RNA-binding proteins DAZ and DAZL1. *Genomics.*
 48. Tsumoto, K., Umetsu, M., Kumagai, I., Ejima, D., Philo, J.S., and Arakawa, T. (2004). Role of arginine in protein refolding, solubilization, and purification. *Biotechnol. Prog.*
 49. Wang, J.T., Smith, J., Chen, B.C., Schmidt, H., Rasoloson, D., Paix, A., Lambrus, B.G., Calidas, D., Betzig, E., and Seydoux, G. (2014). Regulation of RNA granule dynamics by phosphorylation of serine-rich, intrinsically disordered proteins in *C. elegans*. *Elife* 3.
 50. Wassmann, K., Niaux, T., and Maro, B. (2003). Metaphase I arrest upon activation of the Mad2-dependent spindle checkpoint in mouse oocytes. *Curr. Biol.*
 51. Wiszniak, S.E., Dredge, B.K., and Jensen, K.B. (2011). HuB (*elavl2*) mRNA is restricted to the germ cells by post-transcriptional mechanisms including stabilisation of the message by DAZL. *PLoS One.*
 52. Woodruff, J.B., Ferreira Gomes, B., Widlund, P.O., Mahamid, J., Honigsmann, A., and Hyman, A.A. (2017). The Centrosome Is a Selective Condensate that Nucleates Microtubules by Concentrating Tubulin. *Cell* 169, 1066–1077.e10.
 53. Yang, Z.R., Thomson, R., McNeil, P., and Esnouf, R.M. (2005). RONN: The bio-basis function neural network technique applied to the detection of natively disordered regions in proteins. *Bioinformatics.*
 54. Yisraeli, J.K., Sokol, S., and Melton, D. a (1990). A two-step model for the localization of maternal mRNA in *Xenopus* oocytes: involvement of microtubules and microfilaments in the translocation and anchoring of Vg1 mRNA. *Development* 108, 289–298.
 55. Yu, H. (2002). Regulation of APC-Cdc20 by the spindle checkpoint. *Curr. Opin. Cell Biol.*
 56. Zhang, Q., Yaniv, K., Oberman, F., Wolke, U., Git, A., Fromer, M., Taylor, W.L., Meyer, D., Standart, N., Raz, E., et al. (1999). Vg1 RBP intracellular distribution and evolutionarily conserved expression at multiple stages during development. *Mech Dev* 88, 101–106.
 57. Zhang, Z., Fan, Y., Xie, F., Zhou, H., Jin, K., Shao, L., Shi, W., Fang, P., Yang, B., Van Dam, H., et al. (2017). Breast cancer metastasis suppressor OTUD1 deubiquitinates SMAD7. *Nat. Commun.*

SUPPLEMENTARY FIGURES



Supplemental Figure 1, Related to Figure 1. (A) Confocal images of double smFISH for *buc* and *dazl* in Buc-RtoK positive oocytes in the presence of a wild-type copy of *buc*. (B) Confocal image of a novel Bb marker *ca15b* in a stage I oocyte. (C) Confocal images of double smFISH for *vasa* and *ca15b* in Buc-RtoK positive oocytes in the presence of a wild-type copy of *buc*. Scalebars represent 50µm (A, C) and 10µm (B).

5



D

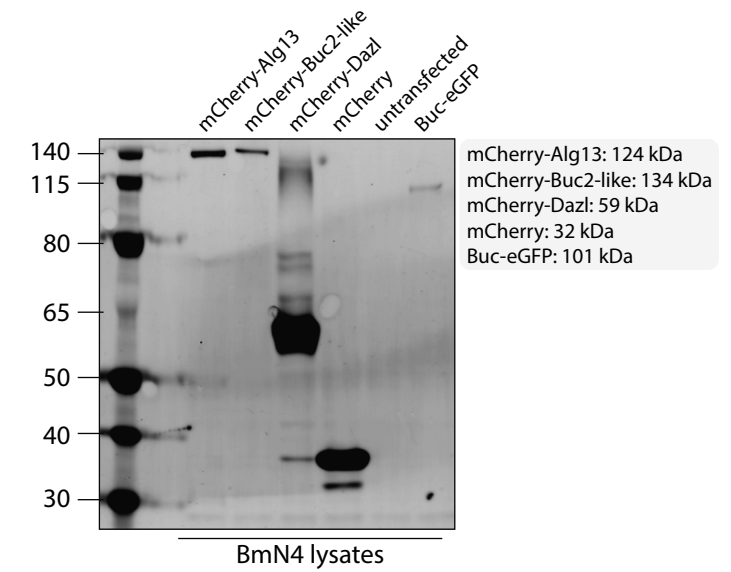
Buc2-like motif

```

motif1  KNLQSGPSAKYCVSPHQLEASIWSIDSLMPYVPSNEWMVENGFTSPQKGVSPQIKSSDGHSSTDGIPLR
motif2  KNLQACGSKYHGVSHQLEASIWSVDSLIPYVPPNELMVDNGFSTPQKSVSPQIKSSDGFSTDGIPLR
motif3  KNLQTCGSKYHRVSHQLEASIWSVDSLMPYVPSNELMVDNGFSTPQKGISPQIKSSDGYSSTDGIPLK
motif4  KNLQTCGSAKYHGVSHQLEASIWSVESLKPYPVPSNGYMDNGYLTSHQALSPPFKSSNVVSEPNQTSS
****.  *:**  *****:** ***** * **:** * **:*** :***:  .:
    
```

Supplemental Figure 2, Related to Figure 3. (A) GFP IP on the wild-type *Buc-eGFP* transgene in a *buc* heterozygous background, compared to blocked beads. (B, C) Protein order prediction (RONN) and PLD prediction (PLAAC) for *Buc* (B) and *Daz*-associated proteins *Dazap1* and *Dazap2* (C). (D) Alignment of the 4 repeated motifs found in *Buc2-like*.

5



Supplemental Figure 3, Related to Figure 4. Western blot of BmN4 lysates. mCherry fusions with Alg13 (124 kDa), *Buc2-like* (134 kDa) and *Dazl* (59 kDa) all migrate at the expected size.

Chapter

Discussion

6

In this thesis, we have explored several aspects that play roles in germ cell development and embryogenesis. We looked at piRNA biogenesis in zebrafish, piRNAs in the mammalian ovary and the role of liquid-liquid phase separation in oogenesis and germ plasm formation. In this chapter I will discuss different aspects of these matters and directions of future research required to solve unanswered questions.

THE PIWI PATHWAY: TO EACH SPECIES THEIR OWN?

The use of model organisms – compare with care

Since the discovery of piRNAs, many in-depth studies have been performed in order to unveil molecular aspects spanning from biogenesis, to mode of action, to inheritance. Although based on only few models, many fundamental differences already became apparent. For instance, murine females do not depend on the PIWI pathway, while males do (Aravin et al., 2008; Carmell et al., 2007). This is in contrast to other species studied, in which the PIWI pathway is highly active in females (Brennecke et al., 2007; Houwing et al., 2007; Kawaoka et al., 2009). Furthermore, silkworm ovary-derived cells have an intact cytoplasmic PIWI pathway, whereas a nuclear counterpart seems to be absent, in contrast to mouse and *Drosophila* (Aravin et al., 2008; Brennecke et al., 2007; Xiol et al., 2014). Also in zebrafish a nuclear branch of the piRNA pathway has not been discovered thus far. Ziwi and Zili have not been detected in the nucleus, due to absence, or concentrations below the detection limit of IHC, indicating that their main site of action of is in the cytoplasm.

In chapter 2, we studied the function of a Tudor-domain containing protein (Tdrd) TdrKH in zebrafish. During piRNA biogenesis, many piRNAs require 3'end trimming to a mature, PIWI-compatible length. The identification of the piRNA trimmer PNLDC1 in mouse and silkworm has been an important step forward in the understanding of this aspect of piRNA biogenesis. However, in *Drosophila*, PARN-like exonucleases, comprising PNLDC1, are absent (Ding et al., 2017; Izumi et al., 2016; Zhang et al., 2017b). Conversely, another trimmer, Nibbler (MUT-7, Exd3), is involved in piRNA trimming in *Drosophila*, but is absent in mouse (Feltzin et al., 2015; Hayashi et al., 2016; Wang et al., 2016). In mouse, PNLDC1 requires TdrKH in order to be recruited to pre-piRNAs, whereas the mechanism of Nibbler recruitment towards pre-piRNAs in *Drosophila* is unknown. *Tdrkh(papi)* mutants do display a mild trimming defect in *Drosophila*, but it is unclear whether this is due to loss of Nibbler recruitment, recruitment of another exonuclease or a different unknown function of TdrKH (Hayashi et al., 2016). PNLDC1 and Nibbler seem to function in different arms of the piRNA pathway: TdrKH and PNLDC1 process Zuc-dependent piRNAs whereas Nibbler acts predominantly on pre-piRNAs generated by PIWI-slicing, such as ping-pong cycle products (Hayashi et al., 2016; Mohn et al., 2015; Nishimura et al., 2018). Zebrafish, silkworm and humans encode both PARN-exonucleases (although zebrafish does not have PNLDC1) and Nibbler. To what extent these enzymes are involved in piRNA trimming relative to each other remains to be investigated, but the presence of both enzymes could be important for the ability to switch from one – Zuc(/PARN?)-dependent – system to the other (Zuc-independent) in case one mechanism is compromised.

In chapter 3 we investigated the presence of a functional PIWI pathway in the female mammalian germline, an issue that had been marginalized since standard models – mouse and rat – have a dispensable piRNA-dependency in the ovary (Aravin et al., 2008; Carmell et al., 2007; Lau et al., 2006). Model organisms other than mammals display ovarian piRNA pathways that are perhaps even more vital than in testis, which is the case in zebrafish but even more so in *Drosophila* (Brennecke et al., 2007; Houwing et al., 2007, 2008; Huang et al., 2011; Kamminga et al., 2010). We found high PIWIL3 and piRNA expression in bovine GV oocytes and IVF embryos. Also in macaque and human ovary, and human fetal samples, we found strong similarities with piRNA pathway features of the male mouse germline.

These findings imply caution when inferring function from one organism to the next. On the other hand, seemingly unrelated findings can in turn share general concepts. Altogether, a message from chapter 2 and 3 is that the use of alternative model organisms can lead to different and sometimes even contradictory conclusions, but still provide important pieces of the puzzle of how piRNAs act in germ cells.

PIWIL3 and maternally inherited piRNAs – a function in mammals

In addition to protecting their own germline using the piRNA pathway, the female gonad of species with an inherited germline also plays a role in protection of the germline of the next generation. Oocytes deposit piRNA pathway factors in the embryo, packaged in the germ plasm in order to end up in the PGCs of the progeny (Houwing et al., 2007; Kirino et al., 2010). This way, an inherited *memory* of piRNA targets is provided in order to assist the progeny to generate a piRNA repertoire of their own. Illustrative of the significance of this inheritance is sterility in *Drosophila* progeny caused by a phenomenon called hybrid dysgenesis. This occurs when two strains are crossed that have a different transposon repertoire, which is the case for wild-caught flies and lab strains, the latter lacking certain transposon insertions that were introduced after their isolation (Bucheton et al., 1984; Kidwell, 1983). Sterility in the progeny only occurs when the novel transposon comes from the father. Mothers lacking these transposons do not accumulate piRNAs against these sequences and consequently, deposited piRNAs in the embryo fail to recognize transposons that they inherit from the father (Brennecke et al., 2008). Also in zebrafish, dominant effects of maternal piRNAs over paternal piRNAs can be traced over generations (Kaaij et al., 2013).

In the alternative mammalian systems we investigated in chapter 3 – macaques, cows and human – we found three different populations of piRNAs in oocytes. The first and second population were reminiscent to the pre-pachytene and pachytene piRNAs found in the male germline of the mouse. These similarities are based on their transposon-mapping features (high and low, respectively), their resistance to NaIO₄ indicating methylation by Hen1, and the presence of upstream A-MYB binding motifs in the case of the pachytene-like population in the adult ovary. The third population accumulated during later stages of oogenesis, together with PIWIL3 protein. Interestingly, this population is inherited and was found in bovine IVF embryos. What they have in common with maternally inherited piRNAs from *Drosophila* and

zebrafish is that they are rich in sequences mapping to transposons. This is in analogy with the transposon memory in animals with an inherited germline. However, in contrast, inherited piRNAs are readily degraded in bovine IVF embryos and almost entirely depleted around the time of zygotic genome activation (~8-cell stage)(Figure 1) (Graf et al., 2014). These piRNAs were not methylated at their 3' ends, reflecting reduced stability and indicating that perhaps they intentionally serve a role *only* during late oogenesis and/or early embryonic development. It therefore seems unlikely that they function in the germ cells of the progeny directly, but to date, the exact role of PIWIL3-piRNA complexes remains unclear.

PIWIL3 absence in Muridae coincides with excess siRNAs produced by an overly active Dicer variant, Dicer^o (Flemer et al., 2013). Since these siRNAs derive mostly from retrotransposon transcripts, it has been hypothesized that in mice, siRNAs fulfil the function piRNAs normally have in female germlines that depend on the piRNA pathway instead (Flemer et al., 2013). Dicer is required for successful meiosis I and fertility in females, analogous to pachytene piRNA function in males, even though the substrates arise from gene-pseudogene pairs and inverted repeats in females and not from A-MYB-driven transcription, and targets are involved in meiotic spindle formation (Murchison et al., 2007; Tam et al., 2008; Watanabe et al., 2008). What the function of the *inherited* siRNAs is in the mouse is still under investigation. Maternally deposited siRNAs into zygotes target transposons and rapidly decrease upon fertilization, analogous to what we found for the piRNAs that correlate with PIWIL3 expression (Cuthbert et al., 2018; García-López et al., 2014; Ohnishi et al., 2010). They have been demonstrated to be able to target and cleave transposon transcripts, suggesting they could protect the genome from damage by transposition events (Ohnishi et al., 2010). Genome-wide reprogramming that occurs during early mammalian development involves DNA demethylation and replacement of protamines of

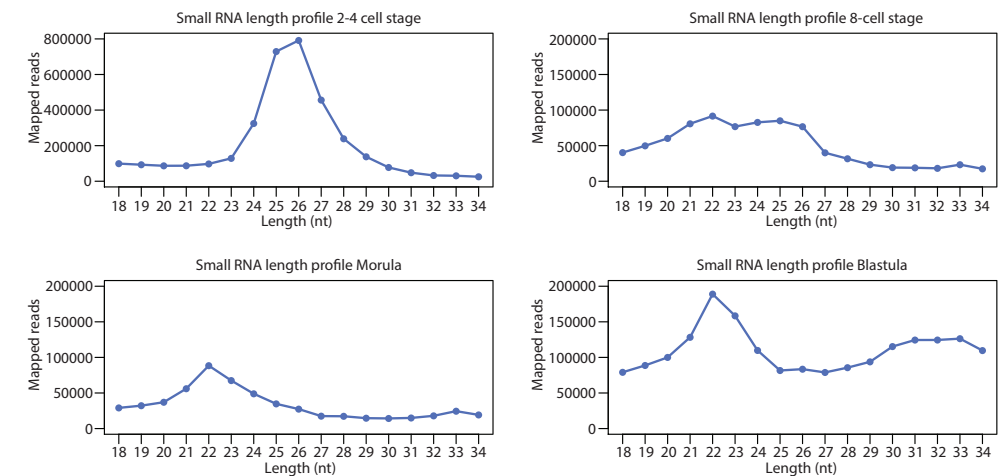


Figure 1. Small RNA length profiles from bovine IVF samples at different stages. After fertilization, 2-4 cell stage embryos display high levels of ~26 nt long piRNAs (chapter 3). These piRNAs have been degraded by the time of zygotic genome activation (8-cell stage).

the paternal chromatin by histones (Reik et al., 2001). Inherited RNA interference machineries, siRNAs in the mouse and piRNAs in PIWIL3-expressing animals, could therefore play a role in suppression of unwanted transposon activity during these crucial timepoints, until they are epigenetically silenced again (Ohnishi et al., 2010).

Interestingly, piRNA reduction in bovine IVF embryos from our own study and the siRNA depletion in mouse around zygotic genome activation coincide with a burst of LTR expression in the zygote (Evsikov and Vries, 2004; Franke et al., 2017). It has been suggested that activation of these elements could provide robust control over gene expression, since random insertions co-regulate genes simultaneously (Gifford et al., 2013; Peaston et al., 2004). Considering this high transposon expression during early development, it has been assumed they adapted a role that is required by the host organism during these early developmental stages. Relieving siRNA or piRNA-mediated silencing around these stages could provide the organism a tool to time transposon expression. In order to explore this possibility, the relationships between transposon expression, siRNA/piRNA sequences and siRNA/piRNA pathway activity have to be investigated in more detail.

Our study has the major limitation that genetic manipulations are not realistic in the organisms studied – bovine, macaque and human – and that multiple attempts to obtain an antibody against PIWIL3 failed for unknown reasons. The emergence of CRISPR technology could allow for studies in models, other than mouse, that comprise a copy of PIWIL3. Obtaining mutant alleles and tagged lines for PIWIL3 could help us understand the role of inherited, transposon-targeting piRNAs in mammals. Another advantage would be that since PIWIL3 and its piRNAs are more restricted to late oogenesis/early embryogenesis compared to Dicer in mouse. *Dicer* phenotypes in mouse occur much earlier in oogenesis because of its role there, and therefore, studying later phenotypes depend on knockdowns or inducible knockouts instead. Potential candidates for PIWIL3 follow-up studies are for instance rabbits, hamsters or guinea pigs. Studying the role of PIWIL3 could therefore provide us insights in the role of maternally inherited piRNAs and potentially regulation of transposon activity during early mammalian development.

PHASE SEPARATION IN GERM CELL BIOLOGY AND BEYOND

The usage of protein condensates: maternal control over PGC-content in the offspring

The phenomenon of inherited germlines has evolved independently several times, although the reason for this remains unclear. What species, that make use of this mechanism, have in common is that they package essential germ cell-specifying factors in phase-separated condensates. A characteristic feature of germ cells is the presence of a wide diversity of phase-separated structures, which have been recognized as being ‘electron-dense’ compared to the surrounding cytoplasm. These germ *granules* are conserved structures, also in species with an induced germline, like mammals, and are essential for their function. Examples are nuage, chromatoid bodies and the Balbiani body (Kloc et al., 2004; Kulkarni and Extavour,

2017; Voronina et al., 2011). They can either be inherited through germ plasm or formed upon germ cell induction. Even though several components are shared, germ granules are not per se equivalent to germ *plasm*: heritable structures that promote cells to become germ cells in the early embryo (Kulkarni and Extavour, 2017; Voronina et al., 2011).

A switch from the inductive mode towards the inherited mode would require the temporary storage, precise localization and delivery, and accurately timed function of maternal factors in the embryo. Perhaps the easiest way to achieve this is to use pre-existing structures with a related function in the cell (such as translational control) and adapt this machinery. Components in germ plasm structures (partially) overlap with components of germ granules, and germ granules are in turn related to P-bodies and stress granules. They share proteins involved in RNA-binding (e.g. Staufen), RNA-processing (e.g. Sm-like or ‘LSm’ proteins), RNA helicases (e.g. Me31B/DDX6) and proteins required for translational control (e.g. eIF4E) (Anderson and Kedersha, 2006; Voronina et al., 2011). It could therefore be that the mere presence of germ granules is an incentive for convergent evolution of germ plasm, since they could relatively easily be adapted such that germ cell-specifying factors for the offspring are incorporated. This way, these factors become translationally and spatially controlled and can be inherited.

In order for this adaptation to happen, germ plasm requires a feature that sets them apart from the other granules present. Though adaptation to the inherited mode has to have occurred independently several times, a common event is the evolution of individual germ plasm *nucleators* or *assemblers* that function analogously in organizing germ plasm. There is accumulating evidence that these nucleators – like the *C. elegans* PGL and MEG proteins, *Drosophila* Oskar and zebrafish/*Xenopus* Buc – have evolved independently, since they are exclusively present in different animal lineages (Figure 2) (Ephrussi et al., 1991; Kawasaki et al., 1998; Kulkarni and Extavour, 2017; Leacock and Reinke, 2008; Marlow and Mullins, 2008; Wang et al., 2014). They typically contain intrinsically disordered regions or prion-like domains, suggesting that this way, they can dictate a specialized ‘germ plasm phase’ to the adapted germ granules in order to stay physically and functionally distinct from other granules present. However, since they are related, they can still incorporate factors from other granules, in case they are required for PGC inheritance too, like PIWI proteins. This enables germ plasm to remain separate from granules that are not directly inherited, like nuage, while factors can still selectively be recruited.

So many granules, so many Tdrds?

Germ cells are rich in phase-separated structures that are distinct in size, localization and ‘electron-densities,’ and consequently composition and function (Figure 3A) (Voronina et al., 2011). The co-existence of these different granules with often (partially) overlapping components, requires inclusion and exclusion mechanisms in order to remain intact and distinct from other types. Examples are PIWI proteins that are selectively recruited to germ plasm for inheritance, like Aubergine, but not Ago3, in *Drosophila* and Zwi, but not Zili, in zebrafish, even though they all also localize together to perinuclear nuage (Brennecke et al., 2007; Harris and Macdonald, 2001; Houwing et al., 2007, 2008). Also in animals that do not have inherited germlines

6

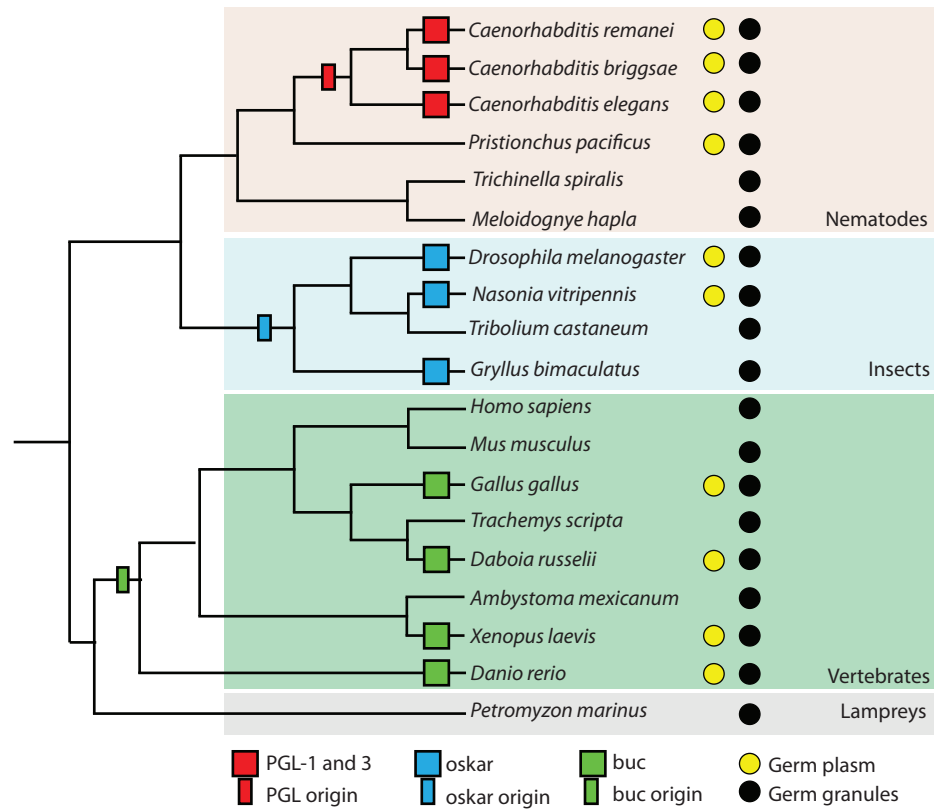


Figure 2. Scenarios for the evolution of germ plasm nucleators, replicated from Kulkarni and Extavour, 2017. Possible convergent evolution of germ plasm nucleators is indicated(Kulkarni and Extavour, 2017).

like mouse, a wide variety of granules co-exists in the male germline, including chromatoid bodies, intermitochondrial cement/pi-bodies and piP-bodies. These granules contain common as well as unique components, however, they still ‘collaborate’ and depend on each other (Aravin et al., 2009).

How are germ cells capable of maintaining these structures? A remarkable high number of Tudor domain-containing proteins (Tdrds) is expressed in germ cells, compared to other cells. These domains have been implicated in many different functions, particularly in the PIWI pathway and germ plasm. Their function is typically associated with the composition and stability of phase-separated structures. For instance, Tdrd6a in zebrafish is required for an intact Balbiani body, mouse Tdrd7 maintains the chromatoid body as a single structure, since in its absence, the chromatoid body disintegrates into contiguous, non-fusing bodies of dissimilar electron-density (Figure 3B, left and middle panel) (Tanaka et al., 2011). Furthermore, zebrafish Tdrd1 is required for maintenance of nuage, which can be visualized with electron microscopy as well (Figure 3B, right panel) (Huang et al., 2011). Without Tdrd1, nuage disappears during early development resulting in infertility (Huang et al., 2011).

6

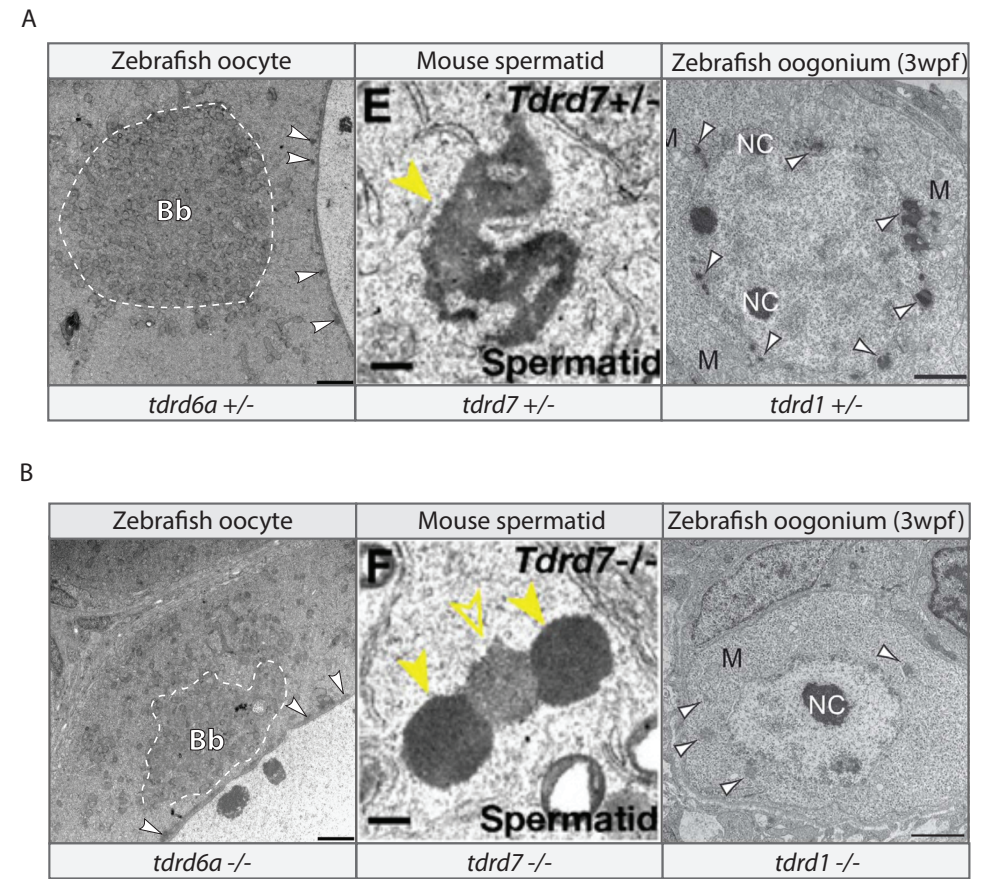


Figure 3. (A) Examples of phase separated structures in germ cells. Left: the Balbiani body (Bb) and peri-nuclear nuage (arrowheads) in zebrafish, middle: the chromatoid body in mouse, right: peri-nuclear nuage in 3 week old zebrafish oogonia. (B) The same structures as (A) but in mutant backgrounds of Tdrd proteins as indicated. The overall structure is affected such as in the Bb in a *tdrd6a* mutant background (left), the chromatoid body separates into two distinct adjacent aggregates with different electron densities without Tdrd7 (middle), the perinuclear nuage in oogonia detaches and becomes less electron-dense without Tdrd1 (right). The pictures were replicated from Tanaka et al., 2011 (Tdrd7) and Huang et al., 2011 (Tdrd1) and chapter 4 of this thesis(Huang et al., 2011; Tanaka et al., 2011).

Another aspect of these granules is that many proteins present in these structures, like PIWI proteins, Vasa and LSM proteins, contain arginine-glycine(RG)-motifs. These RG motifs are potential interaction sites for Tdrds, often in a symmetrically dimethylated state of the arginine (sDMA). Tdrds are capable of recognizing protein stretches sequence-specifically, and the presence – or sometimes rather the absence – of sDMAs provides another dimension of specificity, which causes them to be able to selectively bind with a wide range of interactors (Côté and Richard, 2005; Huang et al., 2011; Liu et al., 2010a, 2010b; Roovers et al., 2018; Zhang et al., 2017a). Furthermore, the presence of multiple Tudor domains in one protein allows

them to interact and recruit certain combinations of proteins, which could ultimately result in different granule types. For instance, Mili-Tdrd1 complexes in mouse reside in pi-bodies, whereas Miwi2-Tdrd9 complexes localize to piP-bodies; two distinct, but adjacent, granular types that require each other for a functional PIWI pathway, implying communication without merging (Aravin et al., 2008, 2009; Shoji et al., 2009). This indicates that they require active inclusion, exclusion and shuttling.

The exact role of Tdrds in granules is not entirely clear, however, they are often essential for their stability and composition (Huang et al., 2011; Sato et al., 2015; Shoji et al., 2009; Tanaka et al., 2011). They are also required for their phase-behavior, which can be demonstrated at the EM level (Figure 3) but also using FRAP (Huang et al., 2011; Tanaka et al., 2011). In chapter 4, we measured Buc mobility, which was reduced in the absence of Tdrd6a (Roovers et al., 2018). Also Spn-E (Tdrd9) in *Drosophila* has reduced mobility in nuage granules in the absence of Qin (Tdrd4) (Andress et al., 2016). The fact that germ cells contain many different granule variants could explain why these cells in particular require such a large collection of Tdrds. Furthermore, Tdrds are also abundant in other cell types – for instance SMN, Tdrd3 and Polycomb-like (PCL) – and play roles in splicing, transcriptional regulation and stress responses, however, the role of Tdrds in granules outside of germ cell biology remains poorly studied (Casanova et al., 2011; Gao et al., 2010; Pellizzoni et al., 1998; Weissbach and Scadden, 2012; Yang et al., 2014). Taken together, Tdrds provide an effective mechanism for composition and properties of granules in germ cells, and potentially many other cell types.

The role of post-translational modifications in phase separation regulation

Over the last years, it has become clear that post-translational modifications (PTMs) can strongly influence granule dynamics and phase behavior. The above-discussed arginine methylation, but also phosphorylation are becoming more acknowledged as important features in the control of granule behavior. Phosphorylation has been demonstrated to be a main switch in assembly or disassembly of *C. elegans* P-granules (Wang et al., 2014). The kinase MBK-2/DYRK3 and the phosphatase PPTR-1 are responsible for these dynamics. When MBK-2 phosphorylates *C. elegans* MEG proteins (P-granule scaffolds), the granules dissolve, which is reversed upon removal by PPTR-1 (Wang et al., 2014). DYRK3, can also ‘dissolve’ several organelles during mitosis in cultured cells, including stress granules (Rai et al., 2018; Wippich et al., 2013). Another example is phosphorylation of serines and threonines in the PLD of FUS, which inhibits phase separation as well (Monahan et al., 2017). The idea is that electrostatic repulsion between phosphorylated versions of PLDs lower the tendency of self-assembly. It has been hypothesized that for this reason, negatively charged amino acids are underrepresented in PLDs (March et al., 2016). Phosphorylation kinetics therefore provide an effective mechanism to reversibly control phase separation.

Phase separation of FUS is not just affected by the phosphorylation state of its N-terminal PLD, but depends on interactions of tyrosines in its PLD with arginines in the RNA-binding

domain through so-called cation- π interactions (Bogaert et al., 2018; Hofweber et al., 2018; Qamar et al., 2018). Cation- π interactions are non-covalent interactions between the electrons from a π -system, in this case the benzene ring of tyrosine, and a cation, in this case the sidechain of arginine. It has been estimated that in proteins, on average 1 in 77 residues takes part in such interactions, and that they therefore contribute widely to protein folding and interaction (Gallivan and Dougherty, 1999). It has been demonstrated that asymmetric dimethylation of arginines in FUS reduces the strength of this interaction, thereby interfering with its phase separation (Hofweber et al., 2018; Qamar et al., 2018). Arginine methylation, however, does not change the charge of the cationic residue, but rather redistributes the charge and has a steric effect, influencing interaction properties (Evich et al., 2016). It was suggested that this ‘bulkiness’ causes disruption of FUS phase separation, either directly or through binding of TNPO1 (Hofweber et al., 2018; Qamar et al., 2018). By now, several studies focusing on intramolecular interactions have demonstrated a role of arginines and arginine methylation in phase separation (Hofweber et al., 2018; Nott et al., 2015; Qamar et al., 2018). Arginine methylation therefore provides another mechanism of control.

Furthermore, other PTMs are increasingly being linked to phase separation as well, including acetylation, ubiquitination, PARylation and SUMOylation (Altmeyer et al., 2015; Banani et al., 2016; Dao et al., 2018; Ferreon et al., 2018; Herhaus and Dikic, 2018; Patel et al., 2015). In chapter 5, we identified a novel Buc interactor, Alg13, which has a catalytically intact Ovarian Tumor(OTU)-deubiquitinase domain, which could be involved in stabilizing its target or more general signaling via ubiquitin modifications. PTMs could be essential in modulation of the germ plasm aggregate, similar to what has been demonstrated for P-granules. Finally, in treatments of disease phenotypes, like the alteration or even reversion of detrimental aggregates, manipulation of PTMs could be of great impact.

A function for arginine dimethylation in the Buc-Tdrd6 interaction

In chapter 4, we identified dimethylated arginine residues at the Buc C-terminus. Buc has 9 tyrosines in its N-terminal PLD which could interact with arginine residues in the rest of the protein, in order to stimulate phase separation. FRAP experiments on Buc without Tdrd6a showed that Buc is less mobile compared to when Tdrd6a is present. By dimethylating its C-terminus, the intramolecular contacts of Buc could be disrupted and make the protein available for intermolecular interactions, for instance with Tdrd6 paralogs. If interactors like Tdrd6a bind and ‘stabilize’ this disruption, Buc remains more liquid-like and is prevented from becoming a homogenous, immobile, gel-like condensate, a hallmark of detrimental phase-separated aggregates (Patel et al., 2015). Possibly, Tdrd6a could couple multiple Buc molecules and together, a larger and more mobile structure could be formed.

Furthermore, Tudor domain-arginine interactions are cation- π contacts too. NMR studies demonstrated that arginine dimethylation make the cation ‘fit’ into the Tudor domain. Through dimethylation, the cation acquires an ideal position relative to the aromatic sidechains in the Tudor domain cage, such that cation- π interactions can stabilize the interaction (Tripsianes

et al., 2011). Considering the affinity of Tudor domains for dimethylated arginines, it can be speculated that arginine methylation could have a dual effect, by disrupting one type of interaction while stimulating another one.

Germ cells as a model for granule dynamics

How phase separation of biological condensates is regulated is still poorly understood. Many *in vitro* assays are performed in which isolated features come forward. This includes mostly intramolecular interactions or interactions with limited, controlled factors like added RNA. The problem with these types of experiments is that biomolecular condensates in an *in vivo* setting are far from homogenic and there are numerous factors that contribute to the outcome of the physical properties and biological function. Among factors that contribute to phase behavior are protein concentration, the presence of RNA or other molecular seeds, the presence of many different disordered proteins (all contributing different properties to the granule), pH(-changes) and post-translational modification status of the entire combination of proteins present in an *in vivo* condensate. Many aspects of phase behavior can be learned from components in isolation, over-expression or stressed settings, and many reports have come out in recent years that demonstrate this, but translating outcomes to an *in vivo* situation is often not yet being done (Langdon et al., 2018; Maharana et al., 2018; Murray et al., 2017; Nott et al., 2015; Patel et al., 2015; Zhang et al., 2015).

Many neurological diseases are caused by cytotoxic effects of detrimental protein aggregates. Our knowledge of how these diseases develop remains scarce. Understanding how assembly and reversibility of such protein aggregates is regulated and how different aspects of these biological condensates could lead to disease would allow us to develop treatments. For this reason, studying phase separation dynamics of germ cell biology can be of importance in order to study these events in healthy, *in vivo* contexts.

Germ plasm is a highly dynamic phase-separated structure that contains many aspects of which the concepts could be translated to a wide variety of condensates. It (1) is subject to assembly and dispersion in a programmed, spatiotemporally controlled manner, (2) contains many proteins that could be prone to PTMs that affect the condensate, like phosphorylation and arginine methylation, (3) contains a mixture of PLD-containing proteins and by timing expression, it can alter physiochemical properties, content and morphology of the condensate, (4) contains a range of transcripts that display sub-structural localization, (5) germ plasm dynamics occur without the requirement of an external stress stimulus, like the exposure to arsenite or X-rays. After a large number of *in vitro* and *ex vivo* studies, more research would have to focus on *in vivo* knowledge, derived from structures such as germ plasm, in order to understand the mechanisms of phase separation in health and disease.

REFERENCES

- Altmeyer, M., Neelsen, K.J., Teloni, F., Pozdnyakova, I., Pellegrino, S., Gröfte, M., Rask, M.B.D., Streicher, W., Jungmichel, S., Nielsen, M.L., et al. (2015). Liquid demixing of intrinsically disordered proteins is seeded by poly(ADP-ribose). *Nat. Commun.*
- Anderson, P., and Kedersha, N. (2006). RNA granules. *J. Cell Biol.*
- Andress, A., Bei, Y., Fonslow, B.R., Giri, R., Wu, Y., Yates, J.R., and Carthew, R.W. (2016). Spindle-E cycling between nuage and cytoplasm is controlled by Qin and PIWI proteins. *J. Cell Biol.*
- 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., 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.
- Banani, S.F., Rice, A.M., Peeples, W.B., Lin, Y., Jain, S., Parker, R., and Rosen, M.K. (2016). Compositional Control of Phase-Separated Cellular Bodies. *Cell.*
- Bogaert, E., Boeynaems, S., Kato, M., Guo, L., Caulfield, T.R., Steyaert, J., Scheveneels, W., Wilmans, N., Haeck, W., Hersmus, N., et al. (2018). Molecular Dissection of FUS Points at Synergistic Effect of Low-Complexity Domains in Toxicity. *Cell Rep.* 24, 529–537.e4.
- 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* (80-.). 322, 1387–1392.
- Bucheton, A., Paro, R., Sang, H.M., Pelisson, A., and Finnegan, D.J. (1984). The molecular basis of I-R hybrid Dysgenesis in *drosophila melanogaster*: Identification, cloning, and properties of the I factor. *Cell.*
- Carmell, M.A., Girard, A., van de Kant, H.J.G., 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.
- Casanova, M., Preissner, T., Cerase, A., Poot, R., Yamada, D., Li, X., Appanah, R., Bezstarosti, K., Demmers, J., Koseki, H., et al. (2011). Polycomblike 2 facilitates the recruitment of PRC2 Polycomb group complexes to the inactive X chromosome and to target loci in embryonic stem cells. *Development.*
- Côté, J., and Richard, S. (2005). Tudor domains bind symmetrical dimethylated arginines. *J. Biol. Chem.*
- Cuthbert, J.M., Russell, S.J., White, K.L., and Benninghoff, A.D. (2018). The maternal-to-zygotic transition in bovine *in vitro*-fertilized embryos is associated with marked changes in small non-coding RNAs†. *Biol. Reprod.*
- Dao, T.P., Kolaitis, R.M., Kim, H.J., O'Donovan, K., Martyniak, B., Colicino, E., Hehnly, H., Taylor, J.P., and Castañeda, C.A. (2018). Ubiquitin Modulates Liquid-Liquid Phase Separation of UBQLN2 via Disruption of Multivalent Interactions. *Mol. Cell.*
- Ding, D., Liu, J., Dong, K., Midic, U., Hess, R.A., Xie, H., Demireva, E.Y., and Chen, C. (2017). PNLDC1 is essential for piRNA 3' end trimming and transposon silencing during spermatogenesis in mice. *Nat. Commun.* 8.
- Ephrussi, A., Dickinson, L.K., and Lehmann, R. (1991). Oskar organizes the germ plasm and directs localization of the posterior determinant nanos. *Cell* 66, 37–50.
- Evich, M., Stroeva, E., Zheng, Y.G., and Germann, M.W. (2016). Effect of methylation on the side-chain pK_a value of arginine. *Protein Sci.*
- Evsikov, A., and Vries, W. De (2004). Systems biology of the 2-cell mouse embryo. ... *Genome Res.*
- Feltzin, V.L., Khaladkar, M., Abe, M., Parisi, M., Hendriks, G.J., Kim, J., and Bonini, N.M.

- (2015). The exonuclease Nibbler regulates age-associated traits and modulates piRNA length in *Drosophila*. *Aging Cell* 14, 443–452.
21. Ferreón, J.C., Jain, A., Choi, K.J., Tsoi, P.S., Mackenzie, K.R., Jung, S.Y., and Ferreón, A.C. (2018). Acetylation disfavors tau phase separation. *Int. J. Mol. Sci.*
22. Flemr, M., Malik, R., Franke, V., Nejepinska, J., Sedlacek, R., Vlahoviček, K., and Svoboda, P. (2013). A retrotransposon-driven Dicer isoform directs endogenous siRNA production in mouse oocytes. *Cell* 155, 807–816.
23. Franke, V., Ganesh, S., Karlic, R., Malik, R., Pasulka, J., Horvat, F., Kuzman, M., Fulka, H., Cernohorska, M., Urbanova, J., et al. (2017). Long terminal repeats power evolution of genes and gene expression programs in mammalian oocytes and zygotes. *Genome Res.*
24. Gallivan, J.P., and Dougherty, D.A. (1999). Cation-pi interactions in structural biology. *Proc. Natl. Acad. Sci.* 96, 9459–9464.
25. Gao, X., Ge, L., Shao, J., Su, C., Zhao, H., Saarikettu, J., Yao, X., Yao, Z., Silvennoinen, O., and Yang, J. (2010). Tudor-SN interacts with and co-localizes with G3BP in stress granules under stress conditions. *FEBS Lett.*
26. García-López, J., Hourcade, J. de D., Alonso, L., Cárdenas, D.B., and Del Mazo, J. (2014). Global characterization and target identification of piRNAs and endo-siRNAs in mouse gametes and zygotes. *Biochim. Biophys. Acta - Gene Regul. Mech.*
27. Gifford, W.D., Pfaff, S.L., and MacFarlan, T.S. (2013). Transposable elements as genetic regulatory substrates in early development. *Trends Cell Biol.*
28. Graf, A., Krebs, S., Zakhartchenko, V., Schwalb, B., Blum, H., and Wolf, E. (2014). Fine mapping of genome activation in bovine embryos by RNA sequencing. *Proc. Natl. Acad. Sci.*
29. 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.
30. Hayashi, R., Schnabl, J., Handler, D., Mohn, F., Ameres, S.L., and Brennecke, J. (2016). Genetic and mechanistic diversity of piRNA 3'-end formation. *Nature.*
31. Herhaus, L., and Dikic, I. (2018). Ubiquitin-induced phase separation of p62/SQSTM1. *Cell Res.*
32. Hofweber, M., Hutten, S., Bourgeois, B., Spreitzer, E., Niedner-Boblenz, A., Schifferer, M., Ruepp, M.D., Simons, M., Niessing, D., Madl, T., et al. (2018). Phase Separation of FUS Is Suppressed by Its Nuclear Import Receptor and Arginine Methylation. *Cell* 173, 706–719.e13.
33. 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.
34. 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.
35. Huang, H.-Y., Houwing, S., Kaaij, L.J.T., Meppelink, A., Redl, S., Gauci, S., Vos, H., Draper, B.W., Moens, C.B., Burgering, B.M., et al. (2011). Tdrd1 acts as a molecular scaffold for Piwi proteins and piRNA targets in zebrafish. *EMBO J.* 30, 3298–3308.
36. Izumi, N., Shoji, K., Sakaguchi, Y., Honda, S., Kirino, Y., Suzuki, T., Katsuma, S., and Tomari, Y. (2016). Identification and Functional Analysis of the Pre-piRNA 3' Trimmer in Silkworms. *Cell.*
37. Kaaij, L.J., Hoogstrate, S.W., Berezikov, E., and Ketting, R.F. (2013). piRNA dynamics in divergent zebrafish strains reveals long-lasting maternal influence on zygotic piRNA profiles. *RNA.*
38. Kamminga, L.M., Luteijn, M.J., Den Broeder, M.J., Redl, S., Kaaij, L.J.T., 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.
39. 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.*
40. Kawasaki, I., Shim, Y.H., Kirchner, J., Kaminker, J., Wood, W.B., and Strome, S. (1998). PGL-1, a predicted RNA-binding component of germ granules, is essential for fertility in *C. elegans*. *Cell.*
41. Kidwell, M.G. (1983). Evolution of hybrid dysgenesis determinants in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci.*
42. Kirino, Y., Vourekas, A., Sayed, N., de Lima Alves, F., Thomson, T., Lasko, P., Rappsilber, J., Jongens, T.A., and Mourelatos, Z. (2010). Arginine methylation of Aubergine mediates Tudor binding and germ plasm localization. *RNA* 16, 70–78.
43. 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.
44. Kulkarni, A., and Extavour, C.G. (2017). Convergent evolution of germ granule nucleators: A hypothesis. *Stem Cell Res.* 24, 188–194.
45. Langdon, E.M., Qiu, Y., Niaki, A.G., McLaughlin, G.A., Weidmann, C.A., Gerbich, T.M., Smith, J.A., Crutchley, J.M., Termini, C.M., Weeks, K.M., et al. (2018). mRNA structure determines specificity of a polyQ-driven phase separation. *Science* (80-.).
46. 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* (80-.).
47. Leacock, S.W., and Reinke, V. (2008). MEG-1 and MEG-2 are embryo-specific P-granule components required for germline development in *Caenorhabditis elegans*. *Genetics.*
48. Liu, H., Wang, J.Y.S., 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.*
49. 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.*
50. Maharana, S., Wang, J., Papadopoulos, D.K., Richter, D., Pozniakovskiy, A., Poser, I., Bickle, M., Rizk, S., Guillén-Boixet, J., Franzmann, T.M., et al. (2018). RNA buffers the phase separation behavior of prion-like RNA binding proteins. *Science* (80-.).
51. March, Z.M., King, O.D., and Shorter, J. (2016). Prion-like domains as epigenetic regulators, scaffolds for subcellular organization, and drivers of neurodegenerative disease. *Brain Res.*
52. 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.
53. Mohn, F., Handler, D., and Brennecke, J. (2015). PiRNA-guided slicing specifies transcripts for Zucchini-dependent, phased piRNA biogenesis. *Science* (80-.). 348, 812–817.
54. Monahan, Z., Ryan, V.H., Janke, A.M., Burke, K.A., Rhoads, S.N., Zerze, G.H., O'Meally, R., Dignon, G.L., Conicella, A.E., Zheng, W., et al. (2017). Phosphorylation of the FUS low-complexity domain disrupts phase separation, aggregation, and toxicity. *EMBO J.*
55. Murchison, E.P., Stein, P., Xuan, Z., Pan, H., Zhang, M.Q., Schultz, R.M., and Hannon, G.J. (2007). Critical roles for Dicer in the female germline. *Genes Dev.*
56. Murray, D.T., Kato, M., Lin, Y., Thurber, K.R., Hung, I., McKnight, S.L., and Tycko, R. (2017). Structure of FUS Protein Fibrils and Its Relevance to Self-Assembly and Phase Separation of Low-Complexity Domains. *Cell.*
57. Nishimura, T., Nagamori, I., Nakatani, T., Izumi, N., Tomari, Y., Kuramochi-Miyagawa, S., and Nakano, T. (2018). PNLDC1, mouse pre-piRNA Trimmer, is required for meiotic and post-meiotic male germ cell development. *EMBO Rep.* e44957.
58. Nott, T.J., Petsalaki, E., Farber, P., Jervis, D., Fussner, E., Plochowitz, A., Craggs, T.D., Bazett-Jones, D.P., Pawson, T., Forman-Kay, J.D., et al. (2015). Phase Transition of a Disordered Nuage Protein Generates Environmentally Responsive Membraneless Organelles. *Mol. Cell* 57, 936–947.
59. Ohnishi, Y., Totoki, Y., Toyoda, A., Watanabe, T., Yamamoto, Y., Tokunaga, K., Sakaki, Y., Sasaki, H., and Hohjoh, H. (2010). Small RNA class transition from siRNA/piRNA

- to miRNA during pre-implantation mouse development. *Nucleic Acids Res.*
60. Patel, A., Lee, H.O., Jawerth, L., Maharana, S., Jahnel, M., Hein, M.Y., Stoykov, S., Mahamid, J., Saha, S., Franzmann, T.M., et al. (2015). A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* 162, 1066–1077.
 61. Peaston, A.E., Evsikov, A. V., Graber, J.H., de Vries, W.N., Holbrook, A.E., Solter, D., and Knowles, B.B. (2004). Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Dev. Cell.*
 62. Pellizzoni, L., Kataoka, N., Charroux, B., and Dreyfuss, G. (1998). A novel function for SMN, the spinal muscular atrophy disease gene product, in pre-mRNA splicing. *Cell.*
 63. Qamar, S., Wang, G.Z., Randle, S.J., Ruggeri, F.S., Varela, J.A., Lin, J.Q., Phillips, E.C., Miyashita, A., Williams, D., Ströhl, F., et al. (2018). FUS Phase Separation Is Modulated by a Molecular Chaperone and Methylation of Arginine Cation- π Interactions. *Cell* 173, 720–734.e15.
 64. Rai, A.K., Chen, J.X., Selbach, M., and Pelkmans, L. (2018). Kinase-controlled phase transition of membraneless organelles in mitosis. *Nature.*
 65. Reik, W., Dean, W., and Walter, J. (2001). Epigenetic reprogramming in mammalian development. *Science* (80-).
 66. Roovers, E.F., Kaaij, L.J.T., Redl, S., Bronkhorst, A.W., Wiebrands, K., de Jesus Domingues, A.M., Huang, H.Y., Han, C.T., Riemer, S., Dosch, R., et al. (2018). Tdrd6a Regulates the Aggregation of Buc into Functional Subcellular Compartments that Drive Germ Cell Specification. *Dev. Cell.*
 67. Sato, K., Iwasaki, Y.W., Shibuya, A., Carninci, P., Tsuchizawa, Y., Ishizu, H., Siomi, M.C., and Siomi, H. (2015). Krimper Enforces an Antisense Bias on piRNA Pools by Binding AGO3 in the Drosophila Germline. *Mol. Cell.*
 68. 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 Is Essential for piRNA-Mediated Retrotransposon Silencing in the Mouse Male Germline. *Dev. Cell* 17, 775–787.
 69. 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.*
 70. 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.* 108, 10579–10584.
 71. Tripsianes, K., Madl, T., MacHyna, M., Fessas, D., Englbrecht, C., Fischer, U., Neugebauer, K.M., and Sattler, M. (2011). Structural basis for dimethylarginine recognition by the Tudor domains of human SMN and SPF30 proteins. *Nat. Struct. Mol. Biol.* 18, 1414–1420.
 72. Voronina, E., Seydoux, G., Sassone-Corsi, P., and Nagamori, I. (2011). RNA granules in germ cells. *Cold Spring Harb. Perspect. Biol.*
 73. Wang, H., Ma, Z., Niu, K., Xiao, Y., Wu, X., Pan, C., Zhao, Y., Wang, K., Zhang, Y., and Liu, N. (2016). Antagonistic roles of Nibbler and Hen1 in modulating piRNA 3' ends in Drosophila. *Development* 143, 530–539.
 74. Wang, J.T., Smith, J., Chen, B.C., Schmidt, H., Rasoloson, D., Paix, A., Lambrus, B.G., Calidas, D., Betzig, E., and Seydoux, G. (2014). Regulation of RNA granule dynamics by phosphorylation of serine-rich, intrinsically disordered proteins in *C. elegans*. *Elife* 3.
 75. Watanabe, T., Totoki, Y., Toyoda, A., Kaneda, M., Kuramochi-Miyagawa, S., Obata, Y., Chiba, H., Kohara, Y., Kono, T., Nakano, T., et al. (2008). Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature.*
 76. Weissbach, R., and Scadden, A.D.J. (2012). Tudor-SN and ADAR1 are components of cytoplasmic stress granules. *RNA.*
 77. Wippich, F., Bodenmiller, B., Trajkovska, M.G., Wanka, S., Aebersold, R., and Pelkmans, L. (2013). Dual specificity kinase DYRK3 couples stress granule condensation/dissolution to mTORC1 signaling. *Cell.*
 78. Xiol, J., Spinelli, P., Laussmann, M.A., Homolka, D., Yang, Z., Cora, E., Couté, Y., Conn, S., Kadlec, J., Sachidanandam, R., et al. (2014). RNA clamping by Vasa assembles a piRNA amplifier complex on transposon transcripts. *Cell.*
 79. Yang, Y., McBride, K.M., Hensley, S., Lu, Y., Chedin, F., and Bedford, M.T. (2014). Arginine Methylation Facilitates the Recruitment of TOP3B to Chromatin to Prevent R Loop Accumulation. *Mol. Cell.*
 80. Zhang, H., Elbaum-Garfinkle, S., Langdon, E.M., Taylor, N., Occhipinti, P., Bridges, A.A., Brangwynne, C.P., and Gladfelter, A.S. (2015). RNA Controls PolyQ Protein Phase Transitions. *Mol. Cell.*
 81. Zhang, H., Liu, K., Izumi, N., Huang, H., Ding, D., Ni, Z., Sidhu, S.S., Chen, C., Tomari, Y., and Min, J. (2017a). Structural basis for arginine methylation-independent recognition of PIWIL1 by TDRD2. *Proc. Natl. Acad. Sci.*
 82. Zhang, Y., Guo, R., Cui, Y., Zhu, Z., Zhang, Y., Wu, H., Zheng, B., Yue, Q., Bai, S., Zeng, W., et al. (2017b). An essential role for PNLDC1 in piRNA 3' end trimming and male fertility in mice. *Cell Res.* 27, 1392–1396.

Addendum

Nederlandse samenvatting
Dankwoord / Acknowledgements
Curriculum Vitae
List of Publications



NEDERLANDSE SAMENVATTING

De bescherming van de geslachtscellen door de PIWI route

De geslachtscellen zijn verantwoordelijk voor het overdragen van het genetische materiaal van generatie op generatie. De genetische informatie die ligt opgeslagen in het DNA is essentieel voor het functioneren van het nageslacht. Om die reden moet het DNA zo intact mogelijk worden gehouden, met name in dit celtype. Naast de aanwezigheid van genen die coderen voor onze lichaamsfunctie bevinden zich in het DNA ook zogenaamde ‘mobiele elementen,’ ook wel ‘transposons’ genoemd. Dit zijn regio's in het DNA die zich kunnen verplaatsen naar een andere plek, of zichzelf kunnen kopiëren, waarna de kopietjes zich elders nestelen. Dit kan uiteraard schadelijke gevolgen hebben voor genen die nodig zijn voor het normaal functioneren van onze cellen. Zo kan een insertie van een transposon in een bepaald gen dit gen bijvoorbeeld inactief maken, of het kan de ‘gen-expressie,’ de manier waarop genen aan of uit staan, verstoren.

Om het genetische materiaal, ondanks deze potentieel gevaarlijke activiteit van transposons, toch zo goed mogelijk intact te houden wordt het DNA in de geslachtscellen beschermd de zogenaamde PIWI route (PIWI of piRNA ‘pathway’). Dit moleculaire beschermingsmechanisme maakt gebruik van het feit dat actieve genen de eiwitten waar ze voor coderen aanmaken. Dit kan alleen als het desbetreffende stuk DNA wordt afgeschreven in RNA: een tijdelijke kopie van een klein stukje DNA. Dit RNA molecuul, een ‘transcript,’ heeft dus de code van een eiwit dat vervolgens gemaakt kan worden. Als transposons actief willen zijn, moeten ze dus de hiervoor benodigde eiwitten aanmaken wat betekent dat ze ook RNA produceren.

De PIWI route maakt hiervan gebruik, aangezien dit betekent dat er RNA kopietjes van transposons die actief zijn (degenen die voor het gevaar kunnen zorgen) zullen worden gemaakt, welke bestreden kunnen worden. De PIWI route heeft een aantal essentiële componenten nodig om deze RNA moleculen te kunnen herkennen en te degraderen, waaronder PIWI eiwitten en PIWI-interacterende RNAs ofwel ‘piRNAs.’ PIWI eiwitten kunnen transcripten klieven en hierdoor inactiveren. Omdat alleen de transcripten die je uit de cel wil verwijderen moeten worden gekleefd is er een andere component nodig om het doelwit *specifiek* te kunnen herkennen. Hiervoor gebruikt het PIWI eiwit korte RNA sequenties: de piRNAs. De piRNAs bevatten fragmenten van transposon sequenties en zodra ze worden gebonden door een PIWI eiwit kan deze transcripten gaan ‘scannen.’ Als ze dan een transcript van een transposon herkennen middels de piRNA sequentie die hiermee ‘matcht,’ kunnen deze transcripten gekleefd worden en zijn dan dus geïnactiveerd.

In dit proefschrift worden twee aspecten van dit mechanisme verder onderzocht. In hoofdstuk 2 beschrijven we een onderdeel van de synthese van piRNAs in zebrafissen. piRNAs moeten worden getrimd zodat ze niet te lang zijn voor een PIWI eiwit. We inactiveren een gen wat hierbij betrokken is, *tdrkh*, en omschrijven wat er vervolgens misgaat om zo te kunnen leren wat er nodig is om de productie goed te laten verlopen. In hoofdstuk 3 kijken we naar de rol van de PIWI route in zoogdieren. Het lijkt erop dat in muizen, die normaalgesproken ‘model’ staan voor allerlei biologische processen in zoogdieren, een uitzondering zijn als het om



de PIWI route gaat. In muizen lijkt de PIWI route namelijk een minimale functie te bekleden in de eicellen, terwijl het in andere dieren, zoals de zebravis en de fruitvlieg, juist essentieel is voor eicellen om te kunnen overleven en dus voor de vruchtbaarheid. We onderzoeken dit aan de hand van weefsels van andere zoogdieren en we vinden bewijzen dat de PIWI route in andere zoogdieren, waaronder de mens, wel degelijk een belangrijke rol kan spelen.

Vloeistof-achtige compartimenten in de cel

Het tweede deel van dit proefschrift richt zich op aggregaten die in geslachtscellen voorkomen. De cel is opgebouwd uit verschillende compartimenten, zoals de celkern, mitochondriën en het Golgi systeem. Deze onderdelen zijn allemaal gescheiden van het celplasma door een membraan van lipiden, waardoor er aan de binnen- en buitenkant verschillende processen plaats kunnen vinden. Naast deze membraan-omsloten compartimenten, zijn er ook compartimenten zonder membraan, die tóch nog fysiek gescheiden zijn van het celplasma. Dit komt doordat ze eiwitten bevatten die de neiging hebben om te aggregeren en hierdoor niet meer 'in oplossing' zijn, maar een soort druppels vormen. Men kan dit vergelijken met druppeltjes olie in water, of bijvoorbeeld de lavalamp: Er is sprake van twee vloeistoffen die niet met elkaar mixen omdat ze andere eigenschappen hebben. In de biologie zijn deze compartimenten vaak niet het resultaat van de eigenschap van slechts één component, maar bestaan meestal uit een uiterst complexe mix van eiwitten met verschillende aggregatie- danwel oplossings-eigenschappen, in combinatie met andere moleculen zoals RNA.

In alle celtypes wordt gebruik gemaakt van deze manier om extra compartimenten te creëren. Doordat ze geen membraan bevatten kunnen cellen ze makkelijker aanmaken en afbreken, bijvoorbeeld als ze snel moeten reageren op een stress situatie. Ze kunnen middels deze aggregaten ook op een kleine plek een hoge concentratie van een bepaald type moleculen accumuleren, waardoor processen sneller en efficiënter kunnen verlopen. Vooral de geslachtscellen worden gekenmerkt door het hoge aantal verschillende typen membraan-loze compartimenten. Zo vindt bijvoorbeeld de PIWI route grotendeels plaats in dit soort aggregaten, die zich om de celkern bevinden. Als deze compartimenten door een mutatie verdwijnen, begeeft de PIWI route het, wat erop duidt dat deze aggregaten essentieel zijn om te kunnen functioneren en wellicht door deze activiteit worden gestabiliseerd.

Een ander kenmerkend compartiment in geslachtscellen is het zogenaamde kiemplasma. Deze structuur wordt in de eicel aangemaakt en meegegeven aan het embryo. Na de bevruchting gaat het embryo zich delen en het kiemplasma wordt meegegeven aan een aantal van deze cellen. Door de informatie die in het kiemplasma zit, worden deze cellen uiteindelijk geïnstrueerd om de geslachtscellen van het embryo te vormen. Op deze manier worden er in de eicel al voorbereidingen getroffen voor de geslachtscellen van het nageslacht. Door allerlei belangrijke componenten die voor geslachtscel-specificatie nodig zijn te verpakken in een aggregaat, kan het lang genoeg bewaard blijven totdat ze nodig zijn, zelfs al is dat pas na de bevruchting. Dit gebeurt bij sommige dieren, zoals fruitvliegen en zebravissen, maar niet bij alle. Andere dieren instrueren cellen pas later tijdens de embryonale ontwikkeling om geslachtscel te worden, zoals bij de muis en de mens.

We kijken in hoofdstuk 4 en 5 naar het kiemplasma van zebravissen. Twee eiwitten die bij dit kiemplasma betrokken zijn, Bucky ball en Tdrd6a, worden in hoofdstuk 4 beschreven. Bucky ball is een eiwit dat aggregerende eigenschappen heeft, doordat het een 'prion-achtig domein' bevat. Prionen zijn eiwitten die aggregerende eigenschappen hebben die zo sterk kunnen zijn dat ze andere eiwitten ertoe aan kunnen zetten om dit ook te doen. Als hierdoor allerlei afzettingen in of om de cel ontstaan, kan dit de celfunctie belemmeren of zelfs leiden tot het afsterven van cellen. Eiwitten met prion-achtige eigenschappen zijn daarom vooral berucht vanwege hun potentie als ziekteverwekker. Vooral neurologische ziektebeelden worden vaak gekenmerkt door over-aggregatie. Denk hierbij aan de ziekte van Alzheimer of Parkinson, of het type ziekten waarbij prionen ontdekt zijn, zoals Bovine Spongiform Encephalopathy (BSE), ook wel bekend als de 'gekke koeziekte'.

Ondanks dat ze dit soort levensbedreigende ziektes kunnen veroorzaken worden deze eiwitten door onszelf aangemaakt. Het is nu bekend dat prion-achtige eiwitten een belangrijke rol spelen bij het vormen van membraan-loze compartimenten, die weer van vitaal belang zijn voor celfunctie, ook al betekent dit dat de aanwezigheid van deze eiwitten mogelijk ziektes kan veroorzaken. Om die reden bestaan er mechanismen die deze aggregaten controleren om over-aggregatie te voorkomen, hoewel er nog vrij weinig over bekend is. In hoofdstuk 4 kijken we daarom naar een eiwit dat de aggregatie eigenschappen van het kiemplasma beïnvloedt, Tdrd6a. Tdrd6a kan met Bucky ball interacteren en is één van de componenten die het kiemplasma gezond en functioneel houdt. In hoofdstuk 4 meten we onder andere dat Bucky ball meer rigide is zonder Tdrd6a, dat de samenstelling van het kiemplasma verandert en dat dit uiteindelijk leidt tot een afname van de geslachtscellen in het nageslacht. Verder kijken we in meer detail naar hoe Tdrd6a met Bucky ball interacteert en vinden een regio in het Bucky ball eiwit waar Tdrd6a veel affiniteit voor heeft. Als we dit stukje vervolgens muteren, zien we dat de eerste kiemplasma structuur die in de eicel gevormd wordt, de zogehete 'Balbiani body,' niet meer kan vormen. Het mechanisme waarop Tdrd6a het kiemplasma beïnvloedt kan ons uiteindelijk iets leren over hoe aggregaten, ook zoals die andere celtypes zoals zenuwcellen, in meer algemene zin regulatie ondergaan.

In hoofdstuk 5 proberen we onze kennis van betrokken eiwitten bij kiemplasma regulatie verder uit te breiden. Ten eerste vinden we dat de gemuteerde versie van Bucky ball, die geen Balbiani body kan vormen, toch later tijdens de eicel ontwikkeling kan aggregeren en een soort kiemplasma ontwikkelt. Het werd tot nu toe verondersteld dat de Balbiani body eerst moet worden gemaakt omdat de latere stappen hieruit voortkomen. In hoofdstuk 5 laten we zien dat dit dus niet per se hoeft, alhoewel het kiemplasma niet helemaal intact blijkt te zijn en daarnaast ook andere defecten heeft die betrekking hebben op de algehele ontwikkeling van het embryo, niet alleen de geslachtscellen. Verder analyseren we eiwitcomplexen van Bucky ball, om zo meerdere betrokken factoren te identificeren. We testen voor een aantal geïdentificeerde kandidaten of ze op dezelfde plek als Bucky ball lokaliseren in celcultuur, wat zo lijkt te zijn. Dit kan erop wijzen dat ze in de daadwerkelijke eicellen ook naar het kiemplasma lokaliseren. Ook kijken we naar de eiwitsequentie en zien dat een aantal kandidaten ook prion-achtig zijn of in ieder geval niet een duidelijke eiwitvouwing hebben in een waterige omgeving (zoals het



celplasma). Ook dit soort eiwitten zijn typisch aanwezig in membraanloze compartimenten. Toekomstig onderzoek moet uitwijzen welke functie onze nieuwe kandidaten vervullen in de eicel ontwikkeling en wat voor effect ze op het kiemplasma hebben.



DANKWOORD / ACKNOWLEDGEMENTS

Na vele jaren en nog een beetje is het dan zover, het proefschrift is dan toch eindelijk werkelijkheid geworden. Ik heb een ontzettend leerzame maar vooral ook erg leuke tijd gehad de afgelopen jaren en daar ben ik een aantal mensen erg dankbaar voor.

De eerste die ik wil bedanken is natuurlijk René. Jij hebt me jaren geleden als totaal onervaren masterstudent een kans gegeven in je lab en ik kwam maar wat graag terug voor een PhD project, ook al zou dat wel een verhuizing betekenen naar Mainz, maar dat had ik er wel voor over. Ik kreeg alle vrijheid om mezelf te ontwikkelen, naar congressen te gaan en om projecten op te zetten. Tegelijkertijd maakte je altijd tijd voor discussie en zorgde je er zo voor dat ik je kantoor altijd weer enthousiaster uitliep! Ik heb me in jouw lab heel erg op mijn plek gevoeld en dat ik het na al die tijd nog steeds moeilijk vind om nu toch ècht eens afscheid te nemen heeft daar wellicht alles mee te maken! Heel erg bedankt voor alle mooie jaren!

Dan Josien, bij wie ik als eerste alle aspecten van het 'lab leven' heb leren kennen. Ik had het geluk dat ik mijn eerste stage bij jou mocht doen. Toen ik begon had ik nog zo mijn twijfels of het überhaupt wel iets voor mij zou zijn, maar al snel werd ik aangestoken door jouw enthousiasme! Het was echt een ontzettend leuke en leerzame tijd, vol gezelligheid, avondjes in het lab (de ochtenden waren voor ons allebei iets minder...), van mijn aller allereerste agarose gelletje tot een enorme Y2H screen en vooral... Heel veel Koffie!

Lieve Maartje L! Ook jij was er al vanaf het begin bij en wat hebben we altijd een lol gehad! Binnen en buiten het lab was jij echt m'n maatje en hebben we heel wat Hubrecht borrels, congressen en avondjes in de kroeg (nauwelijks) overleefd. Ik weet niet of iedereen om ons heen ons geschater en flauwe humor op prijs kon stellen maar dat is nou eenmaal wat je in me losmaakt en ik vind het helemaal geweldig! Nu maar hopen dat je binnenkort ook weer lekker in de buurt komt wonen want ik heb je de afgelopen jaren wel heel erg gemist hoor!

Then my PhD partner in crime Stefan, with whom I started together and also (almost!) finished together. From our first night onwards with some Duvels in 't Oude Pothuys our friendship was born and there were many more nights to follow! It was incredibly helpful to discuss every single aspect of our research with you on a daily basis. Next to that we introduced the Borrel to 'the Germans,' made sure (almost) all the fish came from Utrecht to Mainz and we brewed our first beer together! I had so many laughs with you and I will miss you a lot! I wish you, as a true ungeschoren Tiroler Bergschaf, an amazing time back in Innsbruck together with Vroni and little Felix. I'm so happy that we have met and that you are standing by my side during my defense!

Maartje V. dankzij jou heb ik echt een briljante start van mijn PhD gehad (misschien niet op wetenschappelijk vlak maar zeker op alle andere vlakken). Tsja, je weet al precies waar ik het over heb natuurlijk, ik noem een kaassoufflé, 100% Elsheimer positive en natuurlijk een legendarisch



Kafé dat gek genoeg nog steeds niet op google maps is te vinden. Het werd al snel duidelijk dat we een groot fan van een (1) klein biertje waren, hebben heel wat concerten afgewerkt samen met PJ en kwam graag ieder jaar weer terug voor een festival reünie! Onvergetelijke tijden waren het en ik mis jullie nog altijd! FYI Kafé Els maakt binnenkort een doorstart op de de tweede so cu there!

Lucas, bedankt voor al die jaren samenwerken aan het Tdrd6a project dat er na lang zwoegen toch eindelijk doorheen is gekomen, maar vooral ook de gezelligheid! Het was altijd superfijn om met jou over onze projecten te praten onder het genot van oneindig veel triple espresso's, maar misschien nog wel het allerbelangrijkste: de Roddels!

Suus! We leerden elkaar kennen toen we allebei masterstudent op de 3^e waren en het was zo gezellig dat je toen meteen mee bent gegaan naar Madrid! Hilarische tijden! Later zijn we in Utrecht aan onze PhDs begonnen en kon ik altijd bij je terecht om lief en leed te delen, uiteraard onder het genot van een onmisbare schaal bittergarnituur! Het zit er bijna op en ik ben blij dat je me aan mijn zijde bijstaat tijdens mijn verdediging!

Silja, together with Suzanne we donated our first salary to a greater cause and we further misbehaved on our trips to New York and Berlin! Wishing you all the best in Innsbruck but remember you are always welcome for a 'bessen' in Utrecht! (you can also choose something else ;-)

Sveti, in you I found another crazy catlady! Besides the fact that I learned a lot from you in the lab you were also a fanatic part of the brewing team! Without you I would never have danced the Chacarera and there was of course this one time you 'lost' your glasses... I miss you a lot but I'm sure we meet again in a place where they serve wine for breakfast.

Wallie, ik moet altijd lachen om je droge humor! Ik had nooit gedacht dat ik overstag zou gaan om me aan cell culture te wagen maar door jou is dat toch gebeurd en het heeft het in 2 hoofdstukken tot belangrijke figuren en inzichten geleid! Bedankt voor alle samenwerkingen, al dan niet in ruil voor een mooie Nel, en voor het behoud van wat Nederlandse gezelligheid in het Ketting lab!

Nadine, thanks for translating (and improving...) all our vague proposals into German and for making sure there were fish at all! It must be some bloody good chocolate you keep in your drawer! Wishing you all the best as the new lab grandma!

Patsy I was lucky to have you as a student for a while. I probably gained at least 10 kilos during your internship because of the Best Kuchen Ever that you brought almost every week! Thanks to you I was introduced to the Funzelfahrt concept which was an important step for me in

accepting German culture. I am glad you always followed cake orders while still being very productive in the lab as well, but besides that I really enjoyed your humor and I always had a great laugh with you!

Jan, I know that when you started I was 'almost leaving,' but I'm happy I was wrong so we had more time to get to know each other and have a lot of food together. Besides the 'special sauce' only good memories!

Edo, I'm glad someone did the negative control for injecting a few molar of free nucleotides in fish eggs to prove that it indeed doesn't work. But I am even happier about the positive control that potatoes in lasagna definitely work and it is now one of my favorite dishes!

Maria (Puffeta) and Svenja, I'm glad I'm not the only one who needs to go to Hiro on a regular basis and that you are always there to join me! Maria, thanks for hosting food parties at your place and taking care of us as a true Italian mama. I'm looking forward to the next one!

Then Sheriff, you genotyped so many fish for me you probably reduced my PhD from 10 to 'only' 7 years. Thanks for all the help and kicking my butt when I need to clean something!

No Ketting lab without Portuguese! Miguel, Ricardo, Bruno and António ('I am Jesus!'), I never thought I would hear so much 'Russian' on a daily basis. Miguel, we met when you did your internship and we ended up both doing our PhDs in Germany. I admire your great memory for lots of (scientific) details, which also makes me a bit scared for my movie. António thanks for being a bioinformatician who can explain things to the clueless wetlab people and for collaborating on the TdrKH project. I will miss your cheerfulness but I'm sure we stay connected through our chocolate snob membership.

Ale and Nadia, you only recently joined but you were definitely a great addition to the lab. Nadia, you are just hilarious and you always make me smile! I didn't think it was possible that a person could make so much protein in a lifetime and you managed to do this in only one year. You therefore deserve all the stroopwafels in the world! Ale, keep up the great music, 'the ATPs' is definitely one of the best inventions ever made at the IMB!

I would also like to thank all former labmembers Yi-Yen, Leonie, Suzanne, Marjo, Régis and Hsin-Yi, you have all contributed to a time of my life I can look back on with a smile!

Then I want to thank the other people at the IMB, especially Frotto who makes sure everything runs smoothly and always knows who is doing what & when, the Ketting lab would be completely lost without you!



Also thanks to the Ukuladies with whom I had the best gig(gle)s of my life! Laura (THE BEAUTIFUL PEOPLE), Mary (Ukes not nukes!), Sabrina, Kete, Daniela and Lara. I'll be back for you ladies!

The results in this thesis would not have been possible without collaborations outside of the Ketting lab. Falk and Hanna, thanks for performing and analyzing all the mass spectrometry experiments which made every chapter of this thesis a lot stronger. Tina, you made me feel not so bad about all the very low input samples I gave you and you somehow magically turned them into sequencable libraries that made it to both published chapters of this thesis!

Furthermore, I would like to thank David, Mahdi, Bernard, Susana, Nannan and Lucette for collaboration on the PIWIL3 project. Even though we were on different locations, it went very smoothly and resulted in our Cell Reports paper!

&

In Utrecht had ik mijn allerleukste buuf Ankelien! Vele spontane etentjes thuis en bij de Shiiiiii hebben me er destijds doorheen geholpen! Ik had me geen leukere buurvrouw in Loena kunnen voorstellen dan jij en de volgende valentijnsdag bij de Mr. Shi schuif ik graag weer aan!

Gijs, Kasper, Noëlle en Joren, mijn lievelings Zeeuwen in Utrecht(/Aruba), we kennen elkaar al zo lang en ondanks alle verhuizingen en drukke schema's is het altijd weer ontzettend gezellig als we elkaar weer zien! Gelukkig lijkt het erop dat het binnenkort weer wat vaker kan!

Opa en oma, jullie zijn altijd erg geïnteresseerd en enthousiast en hebben het beste met me voor al zolang ik me kan herinneren. Dat ik hier nu sta is daarom ook zeker aan jullie te danken!

Mijn lieve broertjes en (schoon)zusjes, Eva, Stefan, Bas en Jolijn, ik ben zo blij dat jullie er zijn! Dank jullie voor alle gezelligheid tijdens Sinterklaas, bezoeken aan Mainz, borrels in Breda, Funzelfahrten enzovoorts! Dat er nog veel leuke feestjes en gezelligheid samen in het verschiep mag liggen! Zoals het motto luidt: Immerfest! ;-)

Dan mijn lieve ouders, die ervoor gezorgd hebben dat ik altijd alle kansen heb gekregen die ik me maar kon wensen en er altijd voor me zijn. Pap, door jou was mijn fascinatie voor 'het lab' al vroeg aanwezig (al had ik toen ik klein was waarschijnlijk niet helemaal door wat die leuke konijntjes daar deden...) en het heeft ongetwijfeld een groot aandeel gehad in de keuzes die ik uiteindelijk gemaakt heb. Mam, door de juiste balans tussen enerzijds vrijheid en ruimte en anderzijds goede raad ben ik gekomen waar ik wilde zijn! Ik kan jullie niet genoeg bedanken voor alle liefde en support!

Then finally Angie, I feel like such a lucky person that we've met and you were an important part of the last couple of years of this PhD. You make me very happy and after all this time of travelling between cities I can't wait for our next adventure together in Utrecht!

CURRICULUM VITAE

Elke Roovers was born on the 21st of June, 1986 in Kloetinge, a village in the municipality of Goes, the Netherlands. In 2004 she obtained her secondary school (vwo) diploma from the 'Goese Lyceum.' She then started her studies Biomedical Sciences at the University of Utrecht. During the Master program 'Cancer Genomics & Developmental Biology' she did an internship in the Hubrecht Institute in Utrecht in the lab of René Ketting, under the supervision of Josien van Wolfswinkel. Her second internship was performed at the University of Sydney in the lab of Peter Waterhouse. Upon her graduation in 2011, she started her PhD in the lab of René Ketting at the Hubrecht Institute and later at the Institute of Molecular Biology (IMB) in Mainz. The results of this research are described in this thesis.

&

LIST OF PUBLICATIONS

Roovers, E.F.*, Kaaij, L.J.T.*[†], Redl, S., Bronkhorst, A.W., Wiebrands, K., de Jesus Domingues, A.M., Huang, H.Y., Han, C.T., Riemer, S., Dosch, R., Salvenmoser, W., Grün, D., Butter, F., van Oudenaarden, A., and Ketting, R.F. (2018). Tdrd6a Regulates the Aggregation of Buc into Functional Subcellular Compartments that Drive Germ Cell Specification. *Dev. Cell*. 46, 285–301.e9.

Barton, D.A., **Roovers, E.F.**, Gouil, Q., da Fonseca, G.C., Reis, R.S., Jackson, C., Overall, R.L., Fusaro, A.F., and Waterhouse, P.M. (2017). Live Cell Imaging Reveals the Relocation of dsRNA Binding Proteins Upon Viral Infection. *Mol. Plant-Microbe Interact.* 30, 435–443.

Rosenkranz, D., Han, C.T., **Roovers, E.F.**, Zischler, H., and Ketting, R.F. (2015). Piwi proteins and piRNAs in mammalian oocytes and early embryos: From sample to sequence. *Genomics Data* 5, 309–313.

&

Roovers, E.F.*, Rosenkranz, D.*[†], Mahdipour, M.*[†], Han, C.-T.T., He, N., de Sousa Lopes, S.M.C., van der Westerlaken, L.A.J., Zischler, H., Butter, F., Roelen, B.A.J., and Ketting, R.F. (2015). Piwi proteins and piRNAs in Mammalian Oocytes and early embryos. *Cell Rep.* 10, 2070–2083.

Phillips, C.M., Montgomery, B.E., Breen, P.C., **Roovers, E.F.**, Rim, Y.S., Ohsumi, T.K., Newman, M.A., Van Wolfswinkel, J.C., Ketting, R.F., Ruvkun, G., and Montgomery, T. (2014). MUT-14 and SMUT-1 DEAD box RNA helicases have overlapping roles in germline RNAi and endogenous siRNA formation. *Curr. Biol.* 24, 839–844.

Luteijn, M.J.*[†], Van Bergeijk, P.*[†], Kaaij, L.J.T., Almeida, M.V., **Roovers, E.F.**, Berezikov, E., and Ketting, R.F. (2012). Extremely stable Piwi-induced gene silencing in *Caenorhabditis elegans*. *EMBO J.* 31, 3422–3430.

Kammaing, L.M.*[†], Luteijn, M.J.*[†], den Broeder, M.J., Redl, S., Kaaij, L.J.T., **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.