Association of germ granules and ribosomes in piRNA-mediated RNA regulation during spermatogenesis

Institute of Biomedicine
MDP in Biomedical Sciences, Drug Discovery and Development
Master's thesis

Author:
Samuli Laasanen

2.5.2024
Turku

Master's thesis

Subject: Institute of Biomedicine, MDP in Biomedical Sciences, Drug Discovery and Development
Author: Samuli Laasanen
Title: Association of germ granules and ribosomes in piRNA-mediated RNA regulation during spermatogenesis
Supervisor: Professor Noora Kotaja
Number of pages: xx pages
Date: 2.5.2024

Male infertility due to various defects in spermatogenesis has been increasing rapidly for the past few decades. The successful completion of spermatogenesis requires strict post-transcriptional RNA regulation, and cytoplasmic germ cell-specific ribonucleoprotein granules, germ granules, are known to be important sites for this regulation.

The purpose of this thesis was to examine the mechanisms of germ granule-associated RNA regulation pathways in male mouse germ cells, with a focus on the piRNA pathway. More specifically, the aim was to broaden the understanding of genic piRNA production, and its possible association with ribosomes and nonsense-mediated RNA decay pathway component SMG6.

To reach this aim, the localization of gene-overlapping piRNA clusters in the genome was characterized visually, followed by characterization of the gene and piRNA expressions during the first wave of spermatogenesis using mRNA and small RNA sequencing. Eventually, the interaction of MIWI, SMG6 and ribosomes was studied using polysome fractionation of mouse testicular extracts.

This study revealed that genic piRNA clusters align almost perfectly with the genes they overlap with, suggesting that all genic piRNAs are produced from mRNAs of the genes. Temporally, the expressions of piRNAs and the piRNA-hosting genes correlated only partially. SMG6, MIWI, piRNAs, and piRNA-producing mRNAs were all found associated with polysomes, suggesting it is possible that SMG6 affects the ribosome-associated production of piRNAs from mRNAs. However, the deletion of SMG6 didn’t affect the interaction of the other studied components, and thus, further research is needed to determine the function of SMG6 in piRNA production and to eventually understand the dynamics between germ granules and translation.

Key words: Germ granule, ribosome, piRNA, MIWI, SMG6

Table of contents

1 Introduction 5

1.1 Spermatogenesis 6

1.1.1 Phases and stages of spermatogenesis 6

1.1.2 Special requirements for RNA regulation during spermatogenesis 8

1.2 Germ granules 9

1.3 piRNA-pathway 13

1.3.1 Introduction to the piRNA pathway 13

1.3.2 piRNA production 15

1.3.3 piRNA function 18

1.4 NMD-pathway 19

1.5 Connection between the CB and ribosomes 22

2 Results 25

2.1 Genic piRNA clusters and their overlapping genes 25

2.1.1 Genic piRNA clusters align almost perfectly with protein coding genes 25

2.1.2 Other observations about piRNA-hosting genes 26

2.2 Expressions of piRNA-producing mRNAs and genic piRNAs correlate only partially 27

2.3 Association of SMG6, MIWI, and piRNAs with ribosomes 28

2.3.1 SMG6, MIWI, and piRNAs associate with ribosomes 28

2.3.2 MIWI and piRNAs associate with ribosomes independently of SMG6 29

2.4 piRNA-producing mRNAs are actively translated independently of SMG6 30

3 Discussion 32

3.1 Genic piRNA production 32

3.2 Connection between germ granules and translation 34

3.3 Significance of the study 35

4 Materials and methods 38

4.1 Characterization of piRNA cluster localization in the genome 38

4.2 RNA sequencing 38

4.3 mRNA data analysis 38

4.4 Small RNA data analysis 39

4.5 Mice 39
4.6 Polysome fractionation 39
4.7 Western blotting 40
4.8 RNA extraction 41
4.9 RNA gel electrophoresis 41
4.10 RT-PCR 41
5 Acknowledgements 42
6 Abbreviations 43
References 44
Appendices 52
Appendix 1 Supplementary Figures 52

1 Introduction

Infertility has become a major problem in the modern world: according to WHO, approximately one sixth of all people in reproductive age experience infertility (World Health Organization, 2023). The total fertility rates (the average number of live births per woman) have been decreasing in industrialized countries in the past decades (Priskorn et al., 2016). For example, the total fertility rate in Finland (1.37) has decreased far below 2.1, which is considered the rate required to sustain the population size, and same is true in many other developed countries as well (Perheentupa et al., 2023). This decrease is of course a sum of several factors, such as socioeconomic and personal reasons, but the medical factors can't be overlooked. Indicating unwanted childlessness, the number of children born through assisted reproduction techniques (ART) has been constantly increasing to more than 5 million by today (World Health Organization, 2023). Although ART is a huge help for people experiencing infertility, decreasing fertility can't be solved solely by ART, but we have to understand the causes of the decline to be able to combat it (Perheentupa et al., 2023).

Male factors are estimated to cause 50% of all infertility (Perheentupa et al., 2023). Indicating major contribution of male factors in the increasing infertility, it has been shown that sperm counts have been decreasing all over the world for decades, and that the decline has even accelerated in the 21st century (Figure 1) (Jacobs et al., 2019; Levine et al., 2023). In addition to genetic mutations, lifestyle factors such as obesity, smoking, and alcohol consumption are known to be able to affect male fertility (Barratt et al., 2017; World Health Organization, 2023). However, in most cases the etiology of male infertility is unknown (Schlegel, 2022).

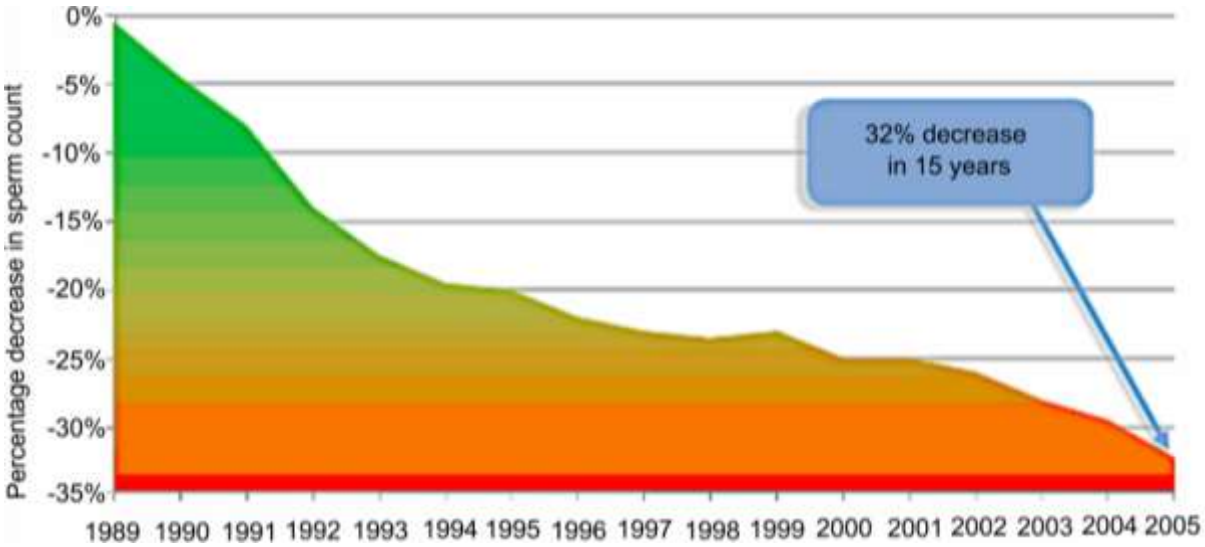


Figure 1. Sperm counts are constantly decreasing. Modified from Jacobs *et al.* (2019).

In addition to determining fertility, processes during sperm production determine the transmission of genetic and epigenetic information to the next generation. Indeed, the father's environmental exposures have been shown to be able to affect the phenotype of his offspring (Pembrey et al., 2005). It has been shown in mice, as well as in other species, that this is possible through a phenomenon called epigenetic inheritance, where the expression of genes in the sperm is altered by the environment, without changes in the gene sequence (Zheng et al., 2021). It is thus also possible that the use of ART affects the health of the offspring, as well as the outcome of the treatment, through epigenetic mechanisms that have not been properly studied.

All in all, better knowledge of the production and function of male germ cells is definitely required to ensure the well-being of the next generation and the continuation of life.

1.1 Spermatogenesis

The process of producing mature male germ cells, called spermatozoa – or sperm, is called spermatogenesis. Spermatogenesis is a complex and highly specialized process that includes both mitosis and meiosis, followed by dramatic morphological changes (Hess et al., 2008). Spermatogenesis is yet a very organized and efficient process: it has been estimated that human males are able to produce as many as 90 million sperm per day on average (Johnson et al., 1980).

1.1.1 Phases and stages of spermatogenesis

During embryonic development, the germ line is separated from somatic cells when primordial germ cells (PGCs) are formed. PGCs initiate the germ cell differentiation by becoming prospermatogonia in the fetal testis, which later give rise to the first postnatal male germ cells called spermatogonial stem cells (SSCs). SSCs serve as the source material for spermatogenesis during the whole male reproductive age. In the first part of spermatogenesis, called spermatocytogenesis, SSCs start a series of mitoses by forming spermatogonia (three types in mice: type A, intermediate, and type B), which eventually become primary spermatocytes through the last mitotic division. Primary spermatocytes are the cells to begin meiosis, which includes to meiotic divisions.

The lengthy prophase of meiosis I can be divided into five phases based on the behavior of chromosomes: leptotene, zygotene, pachytene, diplotene, and diakinesis. Pachytene stage,

when maternal and paternal chromosomes pair and undergo synapsis, is the longest of these phases, lasting more than a week in mice, which makes pachytene spermatocytes a very prominent cell type with several important events. After meiosis I, the newly formed secondary spermatocytes promptly undergo the second meiotic division to produce haploid cells called round spermatids.

In the final phase of spermatogenesis, called spermiogenesis, spermatids undergo dramatic morphological changes to eventually form spermatozoa with their required, dynamic properties. Spermiogenesis can be divided to at least 16 transition steps in mice. Steps 1-8 are round spermatids, steps 9-11 are elongating spermatids in which the morphological transformation and nuclear condensation start, steps 12-14 are elongated/condensed spermatids, and steps 15 and 16 are spermatozoa, from which most of the cytoplasm is eliminated. (Meistrich et al., 2013)

The timepoints for each event during spermatogenesis have been determined, and the total duration of the process is around 70 days in humans and 35 days in mice (Amann, 2008; Oakberg, 1956). The major steps of spermatogenesis are summarized in Figure 2.

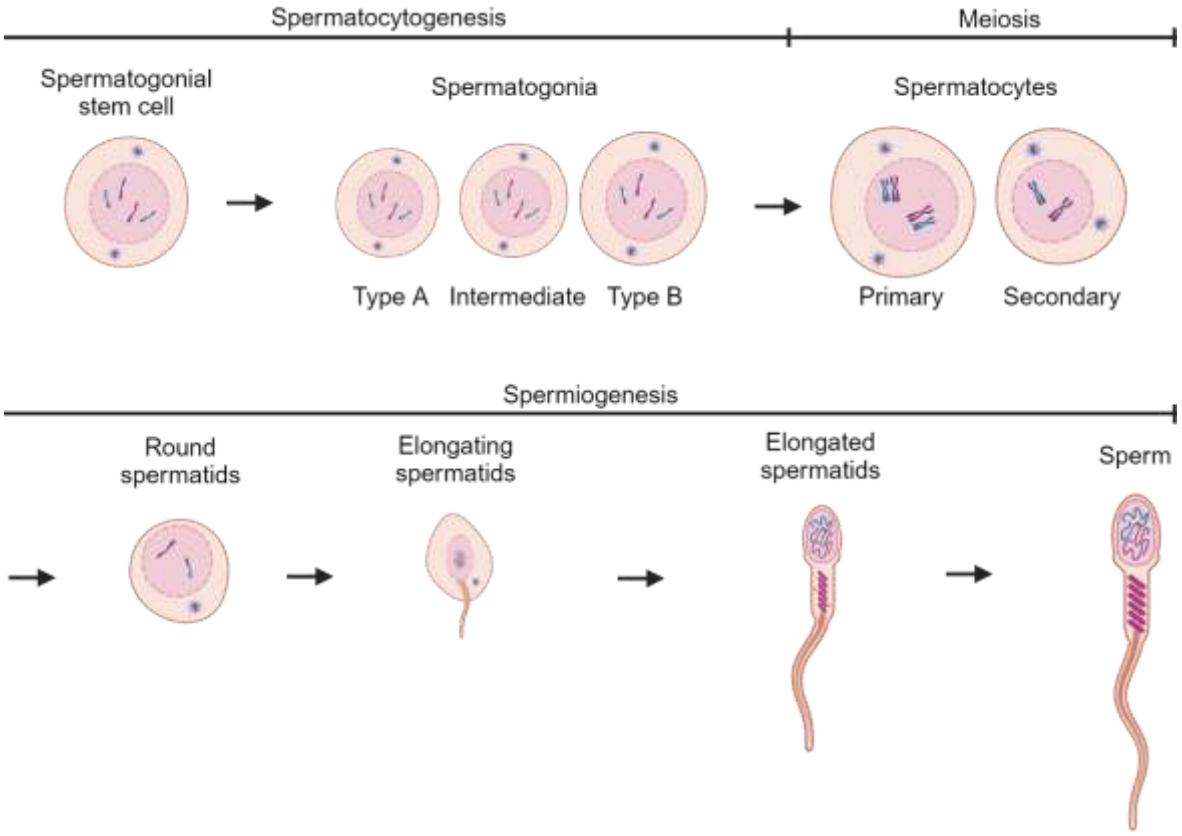


Figure 2. Schematic representation of mouse spermatogenesis. Spermatogonial stem cells divide mitotically to produce spermatogonia, which eventually give rise to spermatocytes that enter meiosis. Meiosis produces haploid spermatids, which evolve into mature sperm. Created in BioRender.com.

The whole process of spermatogenesis takes place in the seminiferous tubules inside the testis. In mice, the structure and function of the seminiferous epithelium are highly organized and synchronized. Spermatogonia rest on the basement membrane, and spermatocytes are located right next to it. As spermatogenesis goes on, cells move closer to the lumen of the tubule, where spermatozoa are eventually released in a process called spermiation. After the release, sperm are passed through the epididymis where they experience further maturation. The complexity of the seminiferous epithelium has been simplified by dividing it into 12 determined stages according to the associated cell types (Figure 3). The cells don't migrate laterally along the tubule, but still the stages appear in successive order and change in cell composition synchronously. (Hess et al., 2008) Organization of the seminiferous epithelium into stages is conserved in humans, but species-specific differences exist (Amann, 2008).

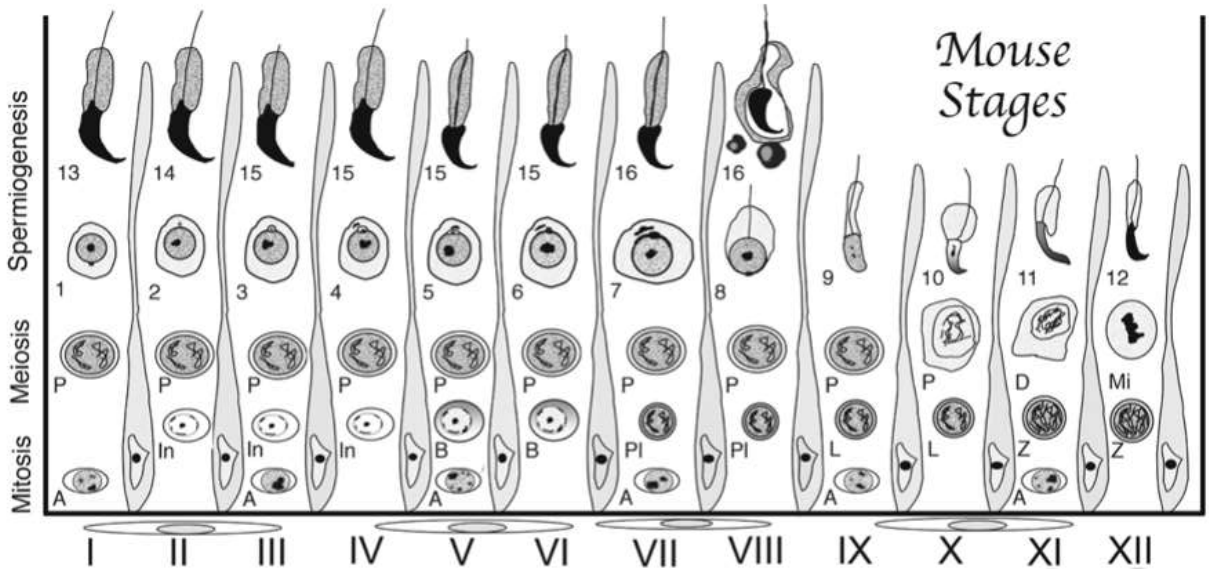


Figure 3. Stages of the mouse seminiferous epithelial cycle. Each column represents the cells of one stage (I-XII), which are separated from each other by drawn Sertoli cells. A, In, and B: spermatogonia; spermatocytes in different phases of meiosis: Pl - preleptotene, L - leptotene, Z - zygotene, P - pachytene, D - diakinesis, Mi - meiotic division; 1-8: round spermatids; 9-16: elongating/elongated spermatids. Modified from Hess and de Franca (2008) and reproduced with permission from Springer Nature.

1.1.2 Special requirements for RNA regulation during spermatogenesis

The high complexity of spermatogenesis comes with a requirement for strict control of gene expression during the process. Each of the major steps is characterized by its unique transcriptomic profile (Laiho et al., 2013). An especially high demand for RNA regulation is created by the remarkably diverse transcriptomes in spermatocytes and round spermatids particularly, including not only mRNAs needed for the differentiation, but also a large

diversity of non-coding RNAs, which have recently been shown to have critical functions in the RNA regulation (Soumillon et al., 2013; X. Wang et al., 2024). This may be a response to the significant challenge that arises during spermiogenesis as the chromatin is condensed and spermatids become transcriptionally inactive (Lehtiniemi et al., 2018). Thus, many mRNAs needed for the morphological changes in the late steps of spermiogenesis are transcribed earlier and translationally repressed for several days (Kleene, 2003). This temporal regulation of translation during spermatogenesis is essential for fertility (Lehtiniemi et al., 2018). The high demand for RNA regulation is reflected by a high number of RNA-binding proteins expressed in male germ cells (Kleene, 2003; Paronetto et al., 2010).

Moreover, a significant amount of RNAs are retained in mature sperm and thus transmitted to the offspring in fertilization. These RNAs, especially small non-coding RNAs, have been shown to be able to carry information about the father's environmental exposures and acquired conditions to the offspring, and thus affect the health of the next generation (Waldron, 2016). This highlights the role of RNA regulation during spermatogenesis not only in determining fertility, but also in determining the health of the future generations through epigenetic inheritance.

1.2 Germ granules

The key site for RNA regulation during spermatogenesis are germ cell-specific ribonucleoprotein (RNP) granules, also known as germ granules. Generally, RNP granules are membraneless organelles, found in several different types of cells, that consist of RNA and RNA-binding proteins. RNP granules form by condensation through protein-protein, RNA-protein, or RNA-RNA interactions. Some RNP granules are constantly present in cells, while others appear only in specific conditions. RNP granules are very diverse and have varying components and functions, but some common functions include RNA processing, translation regulation, RNA localization, and RNA degradation. RNP granules are not stabile organelles but are characterized by constant dynamic of components exiting and entering the granule. (Ripin et al., 2023)

Germ granules are large cytoplasmic RNP granules that are essential for the RNA regulation especially during and after meiosis. Mutations of key germ granule components lead to impaired male fertility, highlighting the important functions of germ granules (Lehtiniemi et al., 2018). Germ granules have some germ cell-specific characteristics, but also share some characteristics with somatic germ granules, such as processing bodies and stress granules

(Voronina et al., 2011). While germ granules have slightly different functions and biological roles in different species, many of the components are evolutionarily conserved (Lehtiniemi et al., 2018). For example, a DEAD-box RNA helicase DDX4 is found in nearly all germ granules and thus considered an important germ granule marker (Lehtiniemi et al., 2018). Other conserved proteins include Tudor domain-containing proteins and PIWI-proteins (Lehtiniemi et al., 2018).

In addition to regulating the processing, localization, stability, and translation of RNAs, germ granules are known to participate in the production and function of small non-coding RNAs that participate also in the regulation of chromatin organization. One important function of germ granules is the temporal regulation of translation, which is critical in the transcriptionally silent late spermatids. mRNAs can be transcribed during meiosis and stored in germ granules for several days before translation. For example, mRNAs for protamines 1 and 2 are transcribed already in pachytene spermatocytes but are translated only in elongating spermatids approximately a week later in mice (Fajardo et al., 1997). This temporal regulation is essential for fertility. (Lehtiniemi et al., 2018)

The best-characterized germ granules in male germ cells are the intermitochondrial cement (IMC) and the chromatoid body (CB) (Figure 4). The IMC originates already in fetal prospermatogonia around clusters of mitochondria and is most prominent in pachytene spermatocytes. After the completion of meiosis, mitochondria are dispersed, and the IMC is no longer detected in haploid cells. (Lehtiniemi et al., 2018) As other RNP granules, the IMC contains many RNA processing proteins, but especially the biogenesis factors of the non-coding piRNAs (Olotu et al., 2023). The exact functions of the IMC, besides piRNA biogenesis, are not well known.

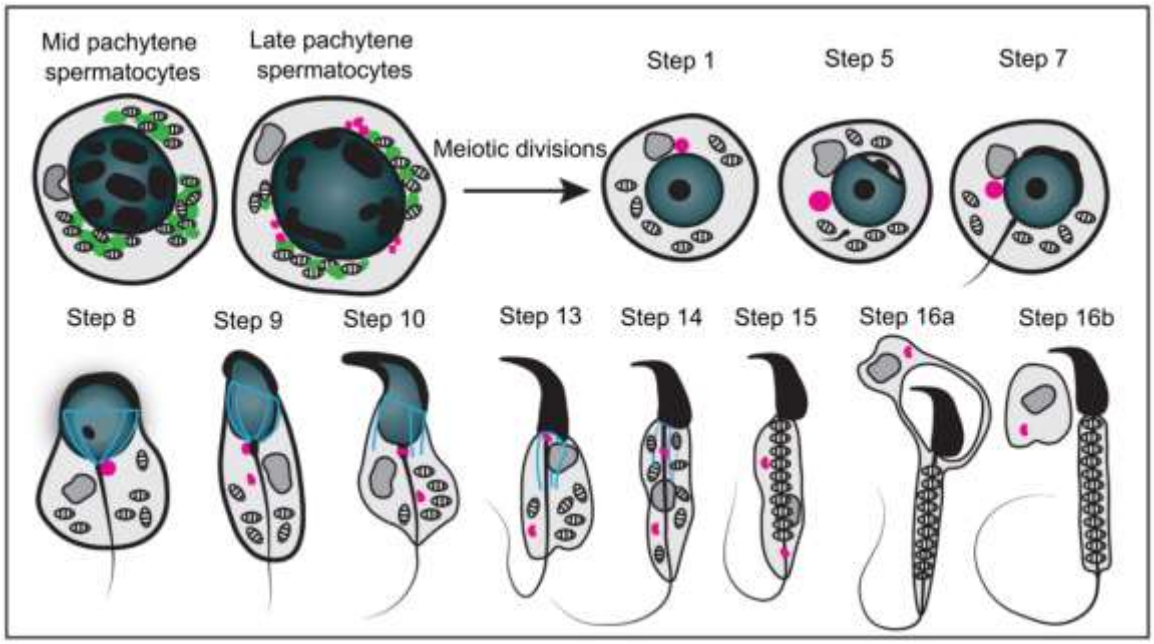


Figure 4. Germ granules during spermatogenesis. A schematic representation of mouse spermatogenesis from pachytene spermatocytes to mature sperm. The IMC (green) is observed as granular material among mitochondria in mid and late pachytene spermatocytes. The precursors of the CB (pink) appear in late spermatocytes and condense into the final form of the CB in early round spermatids, before splitting into two structures in elongating spermatids (steps 9-14) and eventually becoming undetectable in mature sperm (step 16b). The steps of spermatid differentiation (1–16) are indicated. Blue indicates a microtubular structure called manchette. Golgi complex shown in dark grey. Used with permission of Bioscientifica Limited, from Germ granule-mediated RNA regulation in male germ cells, Lehtiniemi and Kotaja, volume 155, issue 2, 2018; permission conveyed through Copyright Clearance Center, Inc.

More is known about the CB, which was discovered already nearly 150 years ago due to its large size of about 1 μm (Brunn, 1876). It emerges in late spermatocytes, co-existing with the IMC for a brief moment, as a pre-CB thought to provide the material for the actual CB that reaches its final form in round spermatids (Lehtiniemi et al., 2018). Unlike the IMC, the CB is not associated with mitochondria. During late spermatogenesis, the CB decreases in size and splits into two parts before becoming undetectable.

The molecular composition of the CB has been extensively described (Figure 5) (Meikar et al., 2014). It contains many RNA regulatory proteins and a wide variety of different RNAs, including mRNAs, non-coding RNAs, as well as non-annotated intergenic transcripts. The majority of the CB consists of RNA-binding proteins and other RNA regulatory proteins, which indicates its functions in post-transcriptional RNA regulation. Approximately 70% of the CB mass is contributable to nine proteins: MILI, MIWI, TDRD7, TDRD6, DDX4, DDX25, D1PAS1, PABP1 and HSPA2. The most represented molecular pathways are the piRNA pathway and the nonsense-mediated mRNA decay (NMD) pathway (discussed in section 1.4).

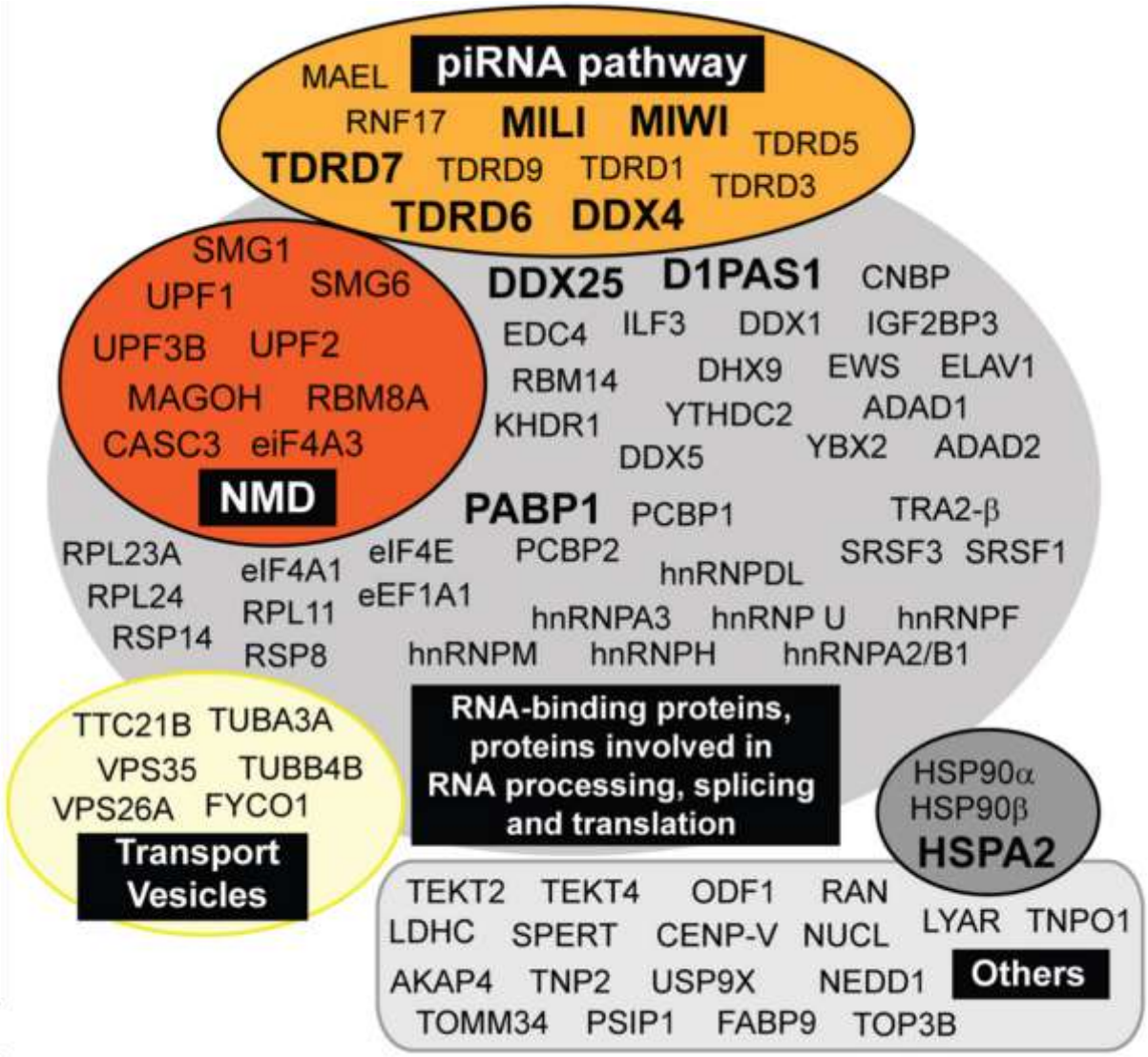


Figure 5. Proteome of the CB. Represented are CB proteins identified by mass spectrometry and immunostaining experiments in several studies. The nine proteins that have been found to account for around 70% of the CB mass in mass spectrometry analysis are in bold. The proteins are grouped in functional pathways, such as piRNA pathway, NMD pathway, RNA-binding/translation-related proteins, proteins involved in vesicle transport, and other uncategorized proteins. Used with permission of Bioscientifica Limited, from Germ granule-mediated RNA regulation in male germ cells, Lehtiniemi and Kotaja, volume 155, issue 2, 2018; permission conveyed through Copyright Clearance Center, Inc.

The CB, like other RNP granules, is not a static structure, but rather a highly dynamic formation that constantly changes in its location, composition, and shape. The CB is located next to the nucleus, where it moves along to nuclear envelope with the help of microtubules (Ventelä et al., 2003). It is suggested that there is a constant transfer of materials between the nucleus and the CB, supported by transcription dependency of the morphology of the CB, and constant flow of RNA to the CB (Meikar et al., 2014; Parvinen et al., 1978; Söderström et al., 1976). In late spermatids, the CB often dissociates from the nucleus (Lehtiniemi et al., 2018). Further highlighting the dynamic nature of the CB, it has been shown to be associated with

the endoplasmic reticulum, the Golgi complex, and cytoplasmic vesicles (Da Ros et al., 2015).

In summary, germ granules have been shown to widely participate in transcriptome regulation throughout spermatogenesis and host various RNAs and RNA-binding proteins, but the precise mechanisms of their function have remained elusive.

1.3 piRNA-pathway

1.3.1 Introduction to the piRNA pathway

Both IMC and CB harbour components of the piwi-interacting RNA (piRNA) pathway (Lehtiniemi et al., 2018). piRNAs are germ cell –specific small non-coding RNAs of 24-35 nucleotides in length that are essential for spermatogenesis (Girard et al., 2006). piRNAs were first identified as silencers of transposable elements (TEs), also called "jumping genes" which cover more than half of the human genome and are involved in various diseases, cellular dysfunction, and genome evolution (Kazazian et al., 2017). piRNAs were later found also to regulate the expression of many protein-coding genes (Gou et al., 2014). piRNAs are the most abundant type of small RNAs in male germ cells, and they account for even 94% of all small RNAs in round spermatids (Gan et al., 2011).

piRNAs act by binding to the PIWI-subfamily proteins of the Argonaute protein family to direct their endonucleolytic activity. The mouse genome encodes three PIWI-proteins: PIWIL1 (MIWI), PIWIL2 (MILI), and PIWIL4 (MIWI2) (X. Wang et al., 2022). MIWI2 is expressed in fetal testis, MILI expression starts before birth and continues at least until the round spermatid stage, and MIWI expression is induced in pachytene spermatocytes and stays strong in round spermatids (Figure 6) (Y. H. Sun et al., 2022). In addition to these three, PIWIL3 is found in human and many other species, mostly in female germ cells, but not in mice (X. Wang et al., 2022). The deletion of any one of the PIWI-proteins leads to male infertility in mice (Aravin et al., 2007; Carmell et al., 2007; Deng et al., 2002). Mutations in the PIWI-proteins as well as in other piRNA-pathway proteins have been identified in cases of human infertility too (Nagirnaja et al., 2022; X. Wang et al., 2022).

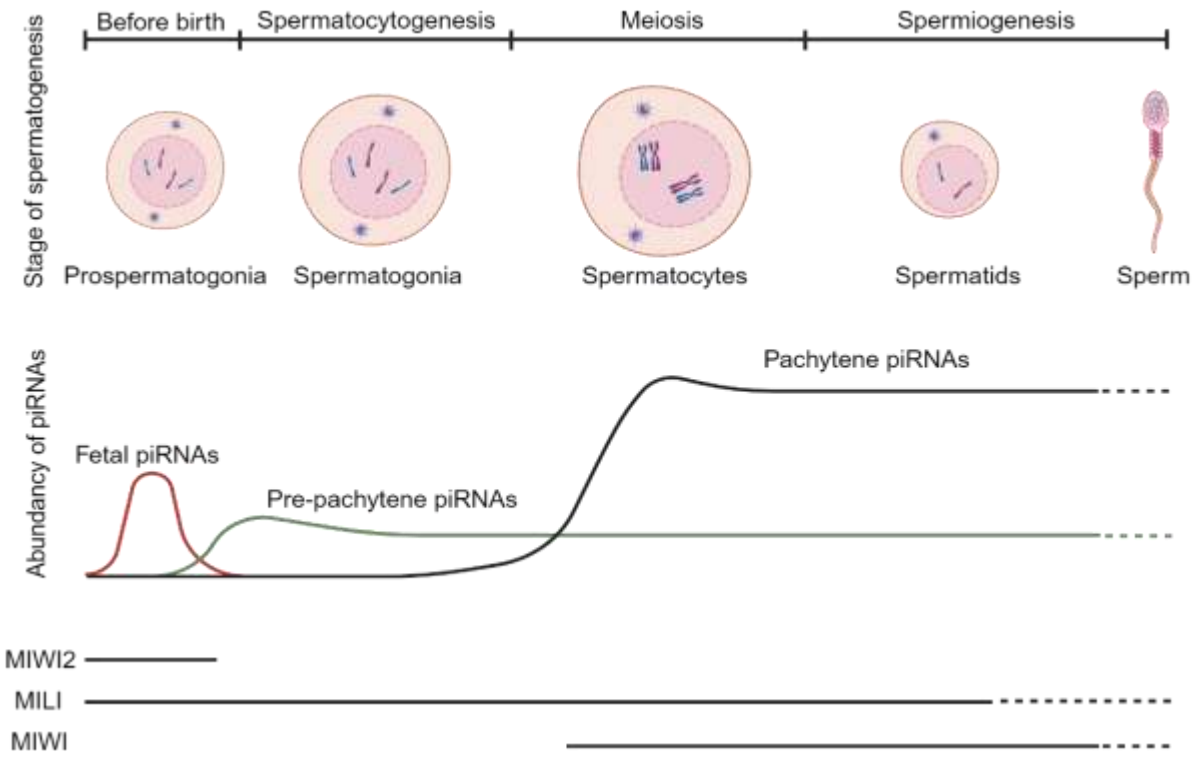


Figure 6. Schematic representation of the expression of PIWI-proteins and piRNAs during spermatogenesis. The upper part represents the progress of spermatogenesis, starting from fetal prospermatogonia and ending in mature sperm. The middle part describes the abundancy of fetal, pre-pachytene, and pachytene piRNAs in different phases of spermatogenesis. The bottom part shows the temporal expression patterns of PIWI-proteins. The piRNA abundancies are based on data from Li *et al.* (2013). Created in BioRender.com.

In addition to being essential for fertility, the piRNA pathway has potential to be involved in sperm-mediated inter-/transgenerational transmission of epigenetic information about the father's acquired conditions, such as obesity. A role for piRNAs in epigenetic inheritance has been identified in *D. melanogaster* (Grentzinger et al., 2012), and changes in male germ cell piRNAs have been reported after various environmental exposures in mice (Gapp et al., 2014; Grandjean et al., 2015; Skinner et al., 2018), suggesting potential role in epigenetic inheritance in mammals as well.

Although considered a germ cell-specific pathway, functions of the piRNA pathway have been reported outside the germ line as well. Some components of the piRNA pathway, including PIWIL1, have been identified as cancer-testis-antigens, which are germ cell-specific proteins that are activated in certain somatic cancers (C. Wang et al., 2016). PIWIL1 has even been shown to promote the progression of some cancers, such as gastric and pancreatic cancers (F. Li et al., 2020; Shi et al., 2020). MIWI and piRNAs have also been identified in the mouse brain (Lee et al., 2011) and differential expression of piRNAs has been suggested in Alzheimer's disease in humans (Qiu et al., 2017). Low expression levels of PIWI-proteins

and piRNAs have been reported in other somatic tissues as well, further indicating their potential involvement in somatic processes (X. Wang et al., 2022). However, there's still a lot of contradicting results and views, and the exact role of the piRNA pathway outside the germ line is highly uncertain.

1.3.2 piRNA production

piRNAs originate from genomic loci called piRNA clusters. A total of 214 piRNA clusters have been identified in mice (X. Z. Li et al., 2013). Individual piRNAs align with several other genomic loci too, but these 214 clusters have been defined based on the piRNA density and abundance. Also, in addition to just aligning the piRNA reads to the genome, the piRNA precursor transcripts of these 214 piRNA clusters have been experimentally identified. Moreover, these 214 clusters reportedly account for 95% of all piRNAs (X. Z. Li et al., 2013).

100 of the 214 piRNA clusters are found between protein-coding genes and are thus called intergenic piRNA clusters. The other 114 clusters are called genic piRNA clusters since they are overlapping with protein-coding genes. It is not known if the genes and piRNA clusters are overlapping totally or only partially, and in the latter case which parts of the gene and piRNA cluster are overlapping. The significance of the overlap is neither currently known.

piRNA clusters are transcribed into long single-stranded piRNA precursors, which are regular mRNA-like RNA polymerase II transcripts (X. Z. Li et al., 2013). Precursor transcripts are then transported from the nucleus to the cytoplasm, where they are further processed. It is suggested that the piRNA processing happens in the IMC, on outer membranes of mitochondria, through three main steps: 5' end formation by endonuclease PLD6 (also known as MitoPLD), loading onto PIWI-proteins, and 3' end formation by exonuclease PNLDC1 and methyltransferase HEN1. These steps create the characteristics of piRNAs, such as preference for 5' uridine and 2' -O-methylation. Several other proteins are involved in the process as well, including mitochondrial Tudor domain-containing proteins TDRKH and TDRD1, and RNA helicase MOV10L1. The described process is the primary pathway of producing piRNAs (Figure 7) (Y. H. Sun et al., 2022). In the secondary pathway of piRNA production, also known as the ping-pong cycle, primary piRNAs recognize their targets, which are then cleaved to produce secondary piRNAs that can again cleave new targets (Ozata et al., n.d.). This cycle has been suggested to be an evolutionary conserved adaptive immune response to allow piRNA-mediated post-transcriptional silencing of TE transcripts, mostly in fetal testes, that continues until the transcripts are diminished (Y. H. Sun et al., 2022).

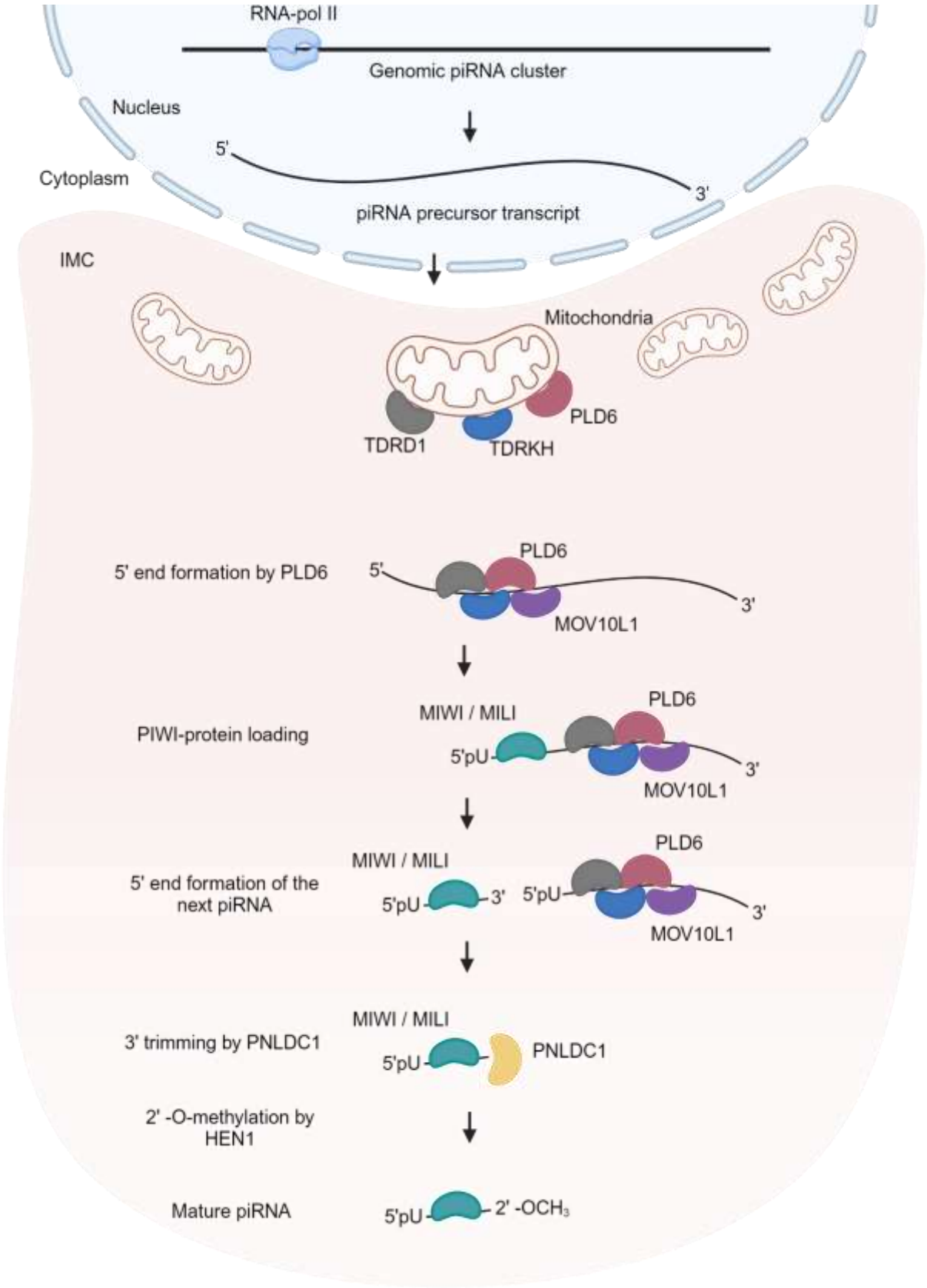


Figure 7. Primary pathway of piRNA biogenesis. Genomic piRNA clusters are transcribed into long precursor RNAs, which are then exported from the nucleus to the IMC, where they are processed by several mitochondria-associated factors into mature piRNAs that are loaded onto PIWI-proteins. Created in BioRender.com based on Sun *et al.* (2018) with a permission under the following license: https://creativecommons.org/licenses/by/4.0/.

In addition to the IMC and its components, both genic and intergenic piRNA production have been shown to be associated with ribosomes in recent studies (Y. H. Sun et al., 2020, 2021). Ribosomes are bound to the piRNA precursor transcripts at sites that are to become piRNAs and are required for efficient piRNA biogenesis, suggesting that they guide the piRNA biogenesis factors (Y. H. Sun et al., 2020). In the case of certain genic piRNA precursors that are protein-coding mRNAs, normal translation and piRNA biogenesis using the same mRNA molecule are shown to be temporally coupled so that some rounds of translation first produce the protein, after which the mRNA is cleaved into piRNAs (Y. H. Sun et al., 2021). This ribosome-guided piRNA production from mRNAs is specific to the 3'UTR's of the mRNAs and leads to the fine-tuning of the expression of proteins that have important roles in the transcriptional and translational regulation during spermatogenesis (Y. H. Sun et al., 2021). However, currently there is no general consensus about the association of germ granules and translation that these proposed mechanisms would require.

piRNAs are usually divided into four groups based on their expression and characteristics: prenatal, pre-pachytene, pachytene, and hybrid piRNAs (Figure 6). Prenatal piRNAs associate with MILI and MIWI2 to regulate TEs in the fetal prospermatogonia (Y. H. Sun et al., 2022). Pre-pachytene piRNA-production starts in the early steps of spermatogenesis, they are mainly genic and also target TEs (DiGiacomo et al., 2013). Pre-pachytene piRNAs are present at least until the round spermatid stage (X. Z. Li et al., 2013). The largest wave of piRNA production happens during the pachytene stage of meiosis I, and these piRNAs are thus called pachytene piRNAs (X. Z. Li et al., 2013). They are mainly intergenic and have been shown to target protein-coding mRNAs (Gou et al., 2014; X. Z. Li et al., 2013). Even more than 95% of piRNAs in postmeiotic cells are pachytene piRNAs (X. Z. Li et al., 2013). The fourth class of piRNAs share the characteristics of pachytene and pre-pachytene piRNAs and are thus called hybrid piRNAs (X. Z. Li et al., 2013).

The piRNAs of the 30 hybrid clusters, which also overlap with protein-coding genes, have been shown to be produced from processed mRNAs of the genes (Y. H. Sun et al., 2021). For the rest of the genic piRNA clusters, it has not been studied if the piRNAs are produced from the mRNAs of the overlapping genes or just some cryptic transcripts that happen to align with them. However, supporting the suggested production of piRNAs from mRNAs, most genic piRNAs map specifically to the 3'UTRs and coding regions of the genes (Gan et al., 2011).

1.3.3 piRNA function

piRNA loading onto MIWI during the late steps of piRNA biogenesis was recently shown to induce the translocation of MIWI from the site of the biogenesis, the IMC, to the other important germ granule, the CB, which is the believed to be the main site of piRNA function (Wei et al., 2024). The PIWI-protein-loaded piRNAs then regulate the expression of TEs and protein-coding genes both transcriptionally and post-transcriptionally depending on the PIWI-protein involved (Grivna et al., 2006).

TE repression is thought to be carried out mainly by prenatal piRNAs, and also by genic piRNAs in the early steps of postnatal spermatogenesis (Y. H. Sun et al., 2021). The cytoplasmic MILI and MIWI participate in the TE repression post-transcriptionally by cleaving target RNAs, guided by the base complementarity between piRNAs and the target RNAs (X. Wang et al., 2022). However, the function of nuclear MIWI2 is required for complete TE silencing (Carmell et al., 2007). MIWI2 can induce transcriptional silencing through DNA methylation and histone modifications, by binding to nascent transcripts via piRNAs and recruiting cofactor proteins (X. Wang et al., 2022).

The regulation of protein-coding genes occurs mainly in postmeiotic cells, where MIWI-bound pachytene piRNAs induce a large elimination of mRNAs (Gou et al., 2014). In addition to direct cleavage of mRNAs, this was also shown to be accomplished by decay of the target RNAs through deadenylation by a deadenylase-enzyme (Gou et al., 2014). Moreover, studies in *D. melanogaster* and *C. elegans* have suggested that piRNAs can regulate protein-coding genes transcriptionally too through DNA methylation and histone modifications (X. Wang et al., 2022).

Interestingly, in addition to TE and gene silencing, piRNAs have been suggested also to stabilize mRNAs and even activate translation of repressed mRNAs in the CB (Dai et al., 2019; Vourekas et al., 2012). Thus, it seems like PIWI-proteins could act as RNA-binding proteins that are able to recruit different cofactors to diversely regulate the transcriptome of male germ cells (X. Wang et al., 2022).

1.4 NMD-pathway

Besides the piRNA pathway, the CB also contain components of another RNA regulatory pathway, nonsense-mediated RNA decay (NMD) pathway (Meikar et al., 2014). The NMD pathway recognizes mRNAs with premature stop codons during translation and degrades them to prevent the production of truncated, non-functional proteins (Kurosaki et al., 2019). In addition, NMD has been shown to target normal, functioning mRNAs too (Kurosaki et al., 2019). NMD is present in all eukaryotic cells, and the main components are evolutionary conserved between species from yeast to humans (Kurosaki et al., 2019). Knockouts of the main components of the NMD pathway are embryonic lethal in mice, and mutations of them have also been associated with human diseases (Kurosaki et al., 2019).

Besides the function in quality control of mRNAs, NMD has been reported to regulate around 10% of cellular transcripts and thus fine-tune protein production. In addition to a premature translation termination codon, the selection of transcripts to be targeted by NMD is based on several identified features, such as an unusually long 3'UTR. NMD usually occurs during or immediately after the export of mRNA from the nucleus and the pioneer round of translation. Translation of the mRNA is required for NMD. (Kurosaki et al., 2019)

The mRNA decay is initiated by the binding of NMD component UPF1, an ATPase and helicase protein, to translating mRNA, phosphorylation of which leads to recruiting of additional proteins that complete the degradation of the mRNA (Figure 8). One option is to recruit SMG5-SMG7-heterodimer, which causes the degradation of the mRNA with the help of other decapping and deadenylation proteins. The other option is to recruit endonuclease SMG6, which can cleave the mRNA directly. (Kurosaki et al., 2019)

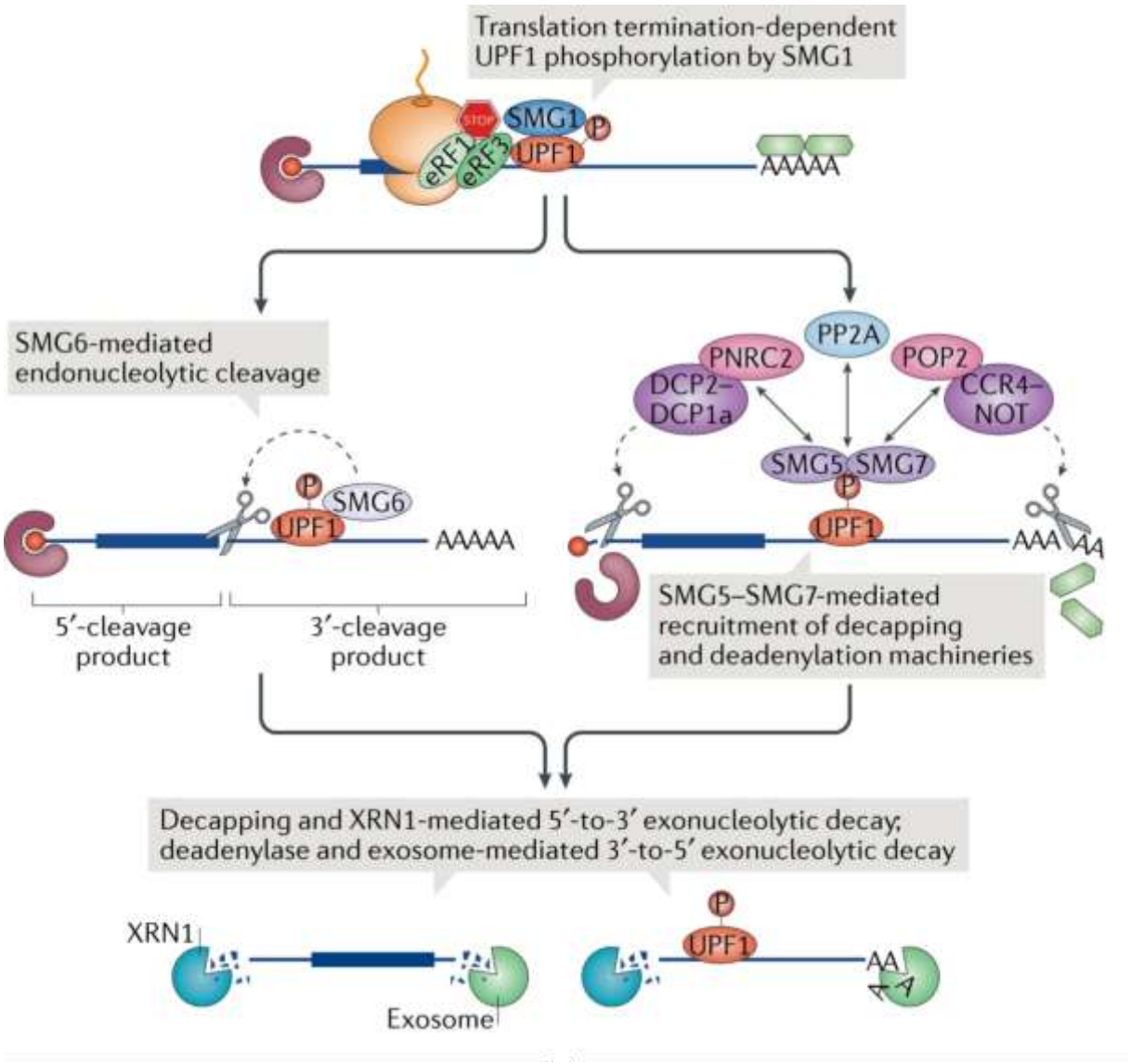


Figure 8. Mechanisms of RNA degradation by the NMD pathway. NMD is initiated by UPF1 binding to the translating target RNA, which is then degraded using either SMG6 or SMG5-SMG7-complex, supported by several other factors. Modified from Kurosaki *et al.* (2019) and reproduced with permission from Springer Nature.

Further validation for the involvement of the CB in the function of NMD comes from studies with mice lacking the protein TDRD6, which is needed for spermiogenesis and the structure and function of the CB (Vasileva et al., 2009). Without TDRD6, UPF1 fails to localize to the CB, and its interaction with UPF2, which is needed for the function of UPF1, is disrupted (Fanourgakis et al., 2016).

The primary endonuclease of the NMD pathway, SMG6, is highly enriched in the CB (Meikar et al., 2014), and thus, further studies on the function of SMG6 in male germ cells have been conducted in our research group (Lehtiniemi et al., 2022). Although SMG6 is expressed in all tissues, as are other NMD components, the expression is especially high in the testis. The

expression is highly concentrated to the CB, but not to the IMC, in mouse and human germ cells. Unlike SMG6, the SMG7-mediated NMD does not localize specifically to the CB and uses different cofactor proteins. (Lehtiniemi et al., 2022)

Our studies with germ cell-specific *Smg6*-knockout (*Smg6*-KO) mice have shown that SMG6 is required for male fertility, particularly the differentiation of round spermatids after meiosis (Lehtiniemi et al., 2022). Loss of SMG6 caused misregulation of meiotic mRNAs in spermatocytes and round spermatids, where also the expression of SMG6 peaks during spermatogenesis. *Smg6*-KO changed also the RNA composition of the CB specifically. Large amount of the misregulated mRNAs was discovered to be NMD target mRNAs with long 3'UTRs. (Lehtiniemi et al., 2022)

A significant amount of these NMD target mRNAs are also targets of MIWI and the piRNA pathway. Additionally, the deletion of SMG6 caused similar spermatogenic arrest in round spermatid stage as the deletion of MIWI. Furthermore, in addition to SMG6 and MIWI co-localizing to the CB, they were shown to interact with each other in mouse male germ cells. Together these findings suggest that the NMD and piRNA pathways might co-operate in the regulation of long 3'UTR-containing mRNAs during the transition from spermatocytes to round spermatids. (Lehtiniemi et al., 2022)

Interestingly, long 3'UTR seems to also be characteristic of piRNA producing mRNAs (Y. H. Sun et al., 2021). These mRNAs have been shown to be highly associated with the NMD-initiating factor UPF1, but in an inactive unphosphorylated form, indicating that the NMD is not degrading these mRNAs (Y. H. Sun et al., 2021). NMD indeed has been shown to be generally repressed in pachytene spermatocytes, when a massive amount of piRNAs are produced (Shum et al., 2016). Further indicating the interaction of NMD and piRNA pathways, pre-pachytene piRNAs are upregulated in *Smg6*-KO round spermatids while the overall production of piRNAs was not altered (Figure 9) (Lehtiniemi et al., 2022). Only 20 of 84 pre-pachytene piRNA clusters were statistically significantly upregulated, but a clear trend towards upregulation in pre-pachytene clusters specifically was observed. 81 of the 84 pre-pachytene piRNA clusters overlap with protein-coding genes (X. Z. Li et al., 2013), suggesting that SMG6 might have a role in the regulation of genic piRNAs.

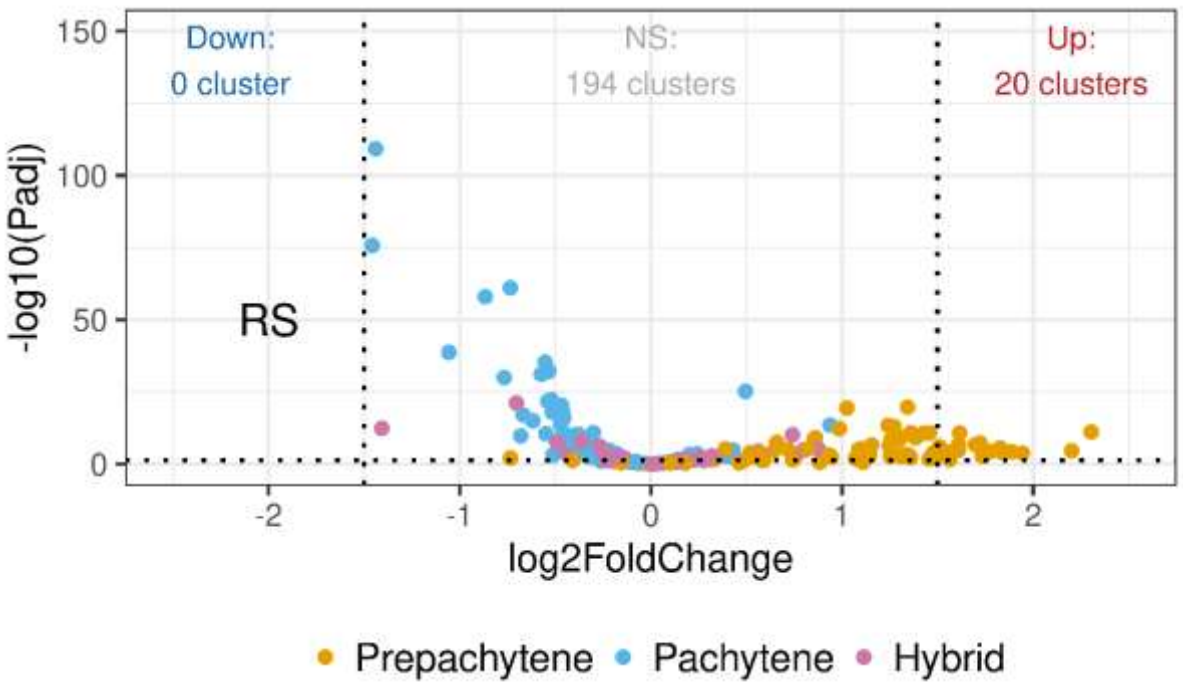


Figure 9. Pre-pachytene piRNAs are upregulated in *Smg6*-KO round spermatids. piRNAs from 20 pre-pachytene piRNA clusters are upregulated above the selected threshold of 1.5 log2FoldChange, but almost all of them are shifted towards upregulation, while pachytene piRNA clusters remain unchanged. Pre-pachytene piRNAs are mostly genic, while pachytene piRNAs are mostly intergenic. Modified from Lehtiniemi *et al.* (2022).

1.5 Connection between the CB and ribosomes

As described above, piRNA and NMD pathways are RNA regulatory pathways that are essential for spermatogenesis and thus for the transmission of genetic and epigenetic information from the father to the offspring. They both are enriched and co-operating in the most prominent germ granule, the CB, which is an important site for RNA regulation in postmeiotic male germ cells. The exact mechanisms of the CB are still largely unknown. Both pathways are also associated with ribosomes and the translational machinery, suggesting some kind of dynamics between the CB and translation. Especially regarding the NMD-component SMG6, which has a very prominent CB-localization, and the NMD has been shown to require ongoing translation (Kurosaki et al., 2019). Somatic RNP granules with similar functions as the germ granules, such as processing bodies and stress granules, have been shown to be associated with ribosomes (Balagopal et al., 2009; Voronina et al., 2011). Supporting the suggestion of an interaction between the CB and translation, the CB has been found to include some components of translational machinery, such as mRNAs, ribosomal RNA and ribosomal proteins (Figure 5) (Meikar et al., 2014). Also, polysome-like structures have been identified near the CB in electron microscopy of rat spermatids (Figueroa et al., 1998). We have also detected similar structures in our own, unpublished electron microscopy

images of mouse CB (Figure 10). Furthermore, inhibition of protein synthesis increases the size of the CB in *Xenopus laevis* (Kalt et al., 1975). Most recently, translation was recently shown to occur at the surface of *D. melanogaster* germ granules (Ramat et al., 2023). Thus, it is strongly suggested that there is an interplay of the CB, ribosomes, and the piRNA and NMD pathways in mice.

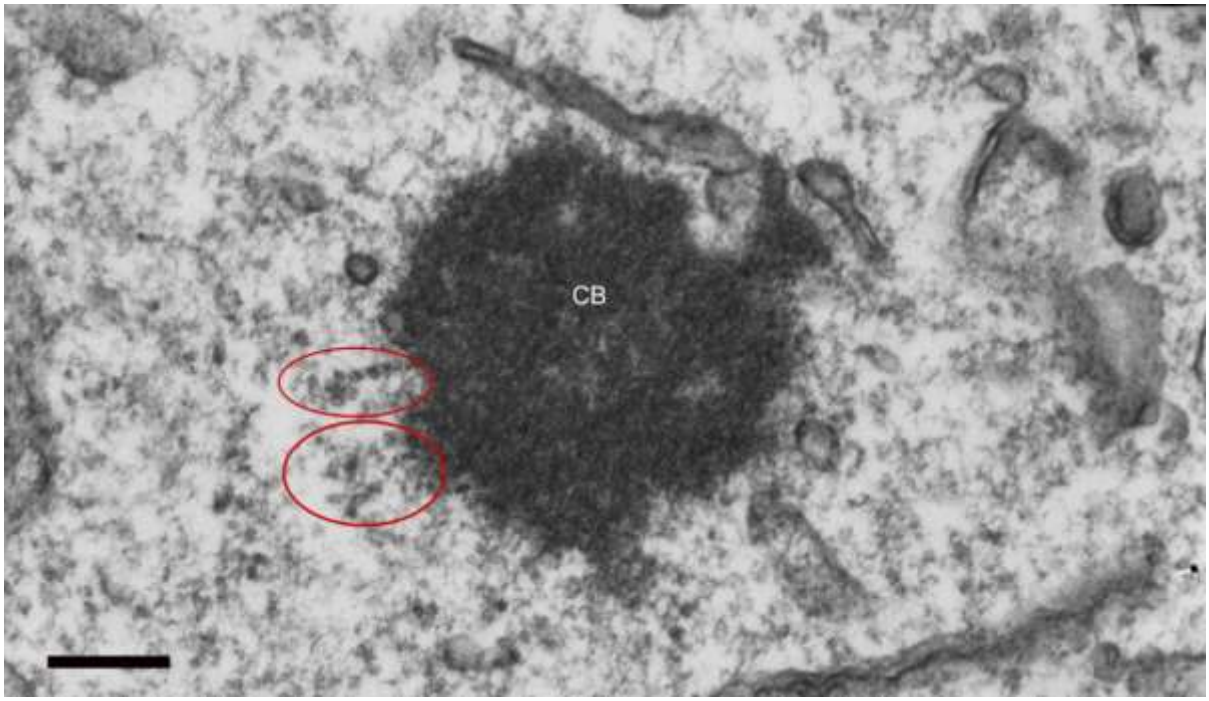


Figure 10. Polysome-like structures near the CB. Electron microscopy image of a CB (dark area in the middle). Polysome-like structures indicated by red circles. Scale bar 200nm.

The purpose of this study was to examine the suggested association of germ granules and ribosomes, especially regarding the production of genic piRNAs. The specific aims of the thesis are:

1. To characterize the localization of genic piRNA clusters in the genome in relation to the genes they are overlapping with.

2. To characterize the expression patterns of the genes overlapping with piRNA clusters in relation to the respective piRNAs during spermatogenesis.

3. To examine if the PIWI-protein MIWI is associated with the NMD pathway component SMG6 at ribosomes.

We hypothesize that every genic piRNA cluster is completely overlapping with the protein-coding gene, which would suggest that all genic piRNAs are produced from mRNA

molecules. We also hypothesize that there is a temporal correlation between the expressions of the gene and the piRNAs mapping to that gene. Lastly, we hypothesize that MIWI, SMG6, and ribosomes are all associated with each other, which would provide an environment where SMG6 could affect the production of genic piRNAs.

2 Results

2.1 Genic piRNA clusters and their overlapping genes

2.1.1 Genic piRNA clusters align almost perfectly with protein coding genes

To clarify the relationship of genic piRNA clusters and the protein-coding genes they overlap with, we visualized their genomic overlap using a genome browser. The extent of the overlap was analyzed by calculating the percentages of the clusters' lengths in nucleotides that was shared with the corresponding gene. If the cluster was completely inside the gene, the percentage of the gene's length overlapping with the piRNA cluster was considered. The results of the analysis are summarized in Table 1.

92 of 114 genic piRNA clusters overlapped with the corresponding gene by more than 95% (Supplementary Figure S1). Given the fact that gene annotations are generally slightly inaccurate, the difference of less than 5% between the piRNA cluster and the gene suggests that the piRNA cluster is the same as the gene.

For the 22 genic piRNA clusters that overlapped less than 95% with their corresponding genes, further evaluation of the nature of the overlap was carried out. For 11 of the 22 clusters, the protein-coding gene was found to encode several transcript isoforms, and the piRNA cluster overlapped with a single transcript by more than 95% (Supplementary Figure S2). For 1 cluster, a 93% matching transcript was found (Supplementary Figure S3). mRNA sequencing data suggested that these transcripts are the most expressed ones (Supplementary Figures S2 and S3). The remaining 10 piRNA clusters overlapped with all the exons of the protein-coding gene, and the only differences were identified outside the coding regions. Seven of these clusters would match the overlapping gene with a longer 3'UTR (Supplementary Figure S4), and one with a shorter 3'UTR (Supplementary Figure S5). In all these cases, mRNA sequencing data supported the suggestion that the UTRs of the protein-coding genes have been inaccurately annotated (Supplementary Figures S4 and S5). The remaining two genic piRNA clusters included sequences from adjacent protein-coding genes, which explains the incomplete overlap (Supplementary Figure S6).

These results revealed an extensive overlap between piRNA clusters and protein-coding genes, which suggests that the piRNA precursor transcripts originating from all 114 genic

piRNA clusters equal the mRNAs of the overlapping protein-coding genes. In other words, it is likely that genic piRNAs are produced from protein-coding mRNAs.

Table 1. piRNA-hosting genes and the nature of their overlap with piRNA clusters.

Gene name	No. of genes in the group	Overlap with piRNA cluster	Nature of the overlap
Bcl2l13, Cmtm4, Luzp1, Marchf8, Cbfa2t2, Ddx19b, Ppp1r12b, Wdfy3, Foxo3, Ing5, Ndst1, Hinfp, Hmbox1, Lpp, Tfcp2l1, Hif1an, Rnf169, Rad54l2, Ppm1f, Dcaf7, Arhgap20, Rc3h1, Hic2, Crkl, Tmem245, Ppp1r15b, Mgll, Uhmk1, Slc43a2, Zfp652, Zbtb37, Klf13, Ctdsp2, Atxn1l, Ccdc117, Abitram, Gtf3c4, Eya3, Il17rd, Igsf9b, Smcr8, Uhrf1bp1, Trim71 ,Phf20, Exog, Fam168b, Gabpb2, Snx30, Gm5878, Gan, Mrs2, Rplp1, Mlec, Ssh1, Ago2, Rcan3, Socs7, Fbxo41, D10Wsu102e, Shank3, Ankrd11, Cramp1l, Bend4, Eif4ebp2, Cbl, Fbxl18, Zfp346, Abl2, Asb1, Nemp1, Mllt6, Dnmt3a, Cdc42ep3, H1f1, Zmat3, Slc25a51, Tef, Wipf2, Mroh4, Nr2c2, Zdhhc23, Noct, Cbx5, Nsd1, Zfp280b, Zyg11b, Tbl2, Tet3, Rab11fip4, Igf2bp1, Mafg, Elk4	92	>95%	Difference explained by inaccurate annotation of gene coordinates
Map3k9, Tktl2, Acvr2b, Hjurp, Naa38, Zbtb49, Exoc8, Fth1, Tacc1, Zfp866, Pou6f1	11	20-90%	Single transcript overlaps >95%
Ipmk	1	45%	Single transcript overlaps 93%
1700006A11Rik, Mlc1, Zyg11a, Kcng3, Zfp382, Klhl11, Strbp	7	60-90%	Would match with longer 3'UTR
Kctd7	1	93%	Would match with shorter 3'UTR
Ip6k1, Fam53b	2	72-87%	Transcript of an adjacent gene completes overlap

2.1.2 Other observations about piRNA-hosting genes

To understand the significance of piRNA-production from mRNAs, we conducted gene ontology (GO) enrichment analysis to study the characteristics of the piRNA-producing genes. The most enriched GO terms for biological processes among the 114 piRNA-hosting genes included several terms related to regulation gene expression, which indicates that the genes have important functions in cells (Figure 11). To further characterize the piRNA-hosting genes, a brief search for tissue expressions of the 114 genes was executed in the NCBI-database. The search revealed that the genes are not testis specific but expressed variably in different tissues, for example in the central nervous system (data not shown). Confirming these observations requires more in-depth analysis.

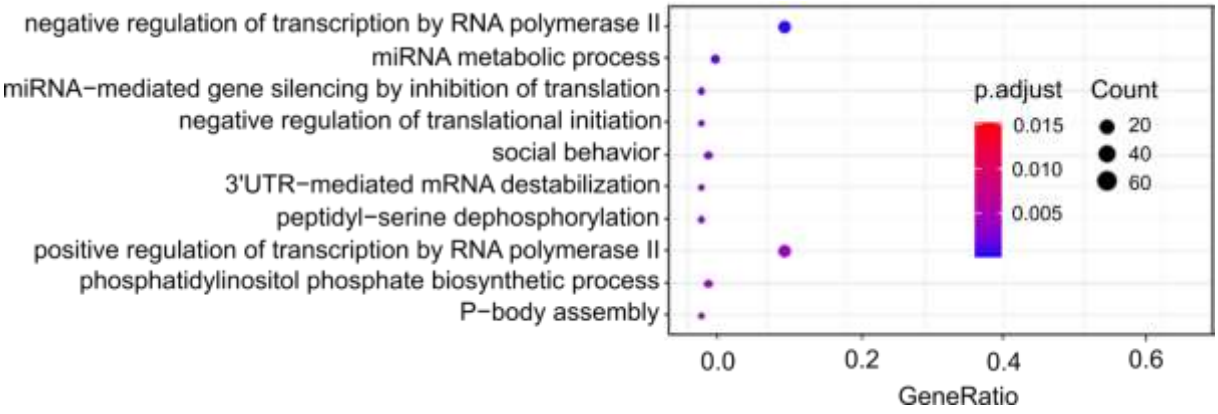


Figure 11. The 10 most enriched GO terms for biological process among the 114 piRNA-hosting genes. The analysis was performed using topGO package, and the results were arranged based on significance.

2.2 Expressions of piRNA-producing mRNAs and genic piRNAs correlate only partially

In order to investigate the correlation of piRNA-producing mRNA expression and piRNA production from those mRNAs, their temporal expression patterns during the first wave of spermatogenesis were analyzed. Both mRNA and small RNA sequencing were performed on total RNA extracted from 1, 2, 3, and 4 weeks old mouse testes, representing the appearance of spermatogonia, pachytene spermatocytes, round spermatids and elongating spermatids, respectively. The analysis revealed that slightly more than half of the mRNAs reach the peak of their expression in the first two weeks, while the rest peak in the weeks 3 and 4 (Figure 11A). Nearly all genic piRNAs, in turn, are most abundantly present in week 1 (Figure 11B). Even many piRNAs produced from mRNAs that are mostly expressed in weeks 3 and 4 were abundantly present already in week 1. Altogether, no clear correlation between piRNA-producing mRNA expression and genic piRNA abundance was observed, and further analyses are required to clarify how the piRNA production from mRNAs is carried out.

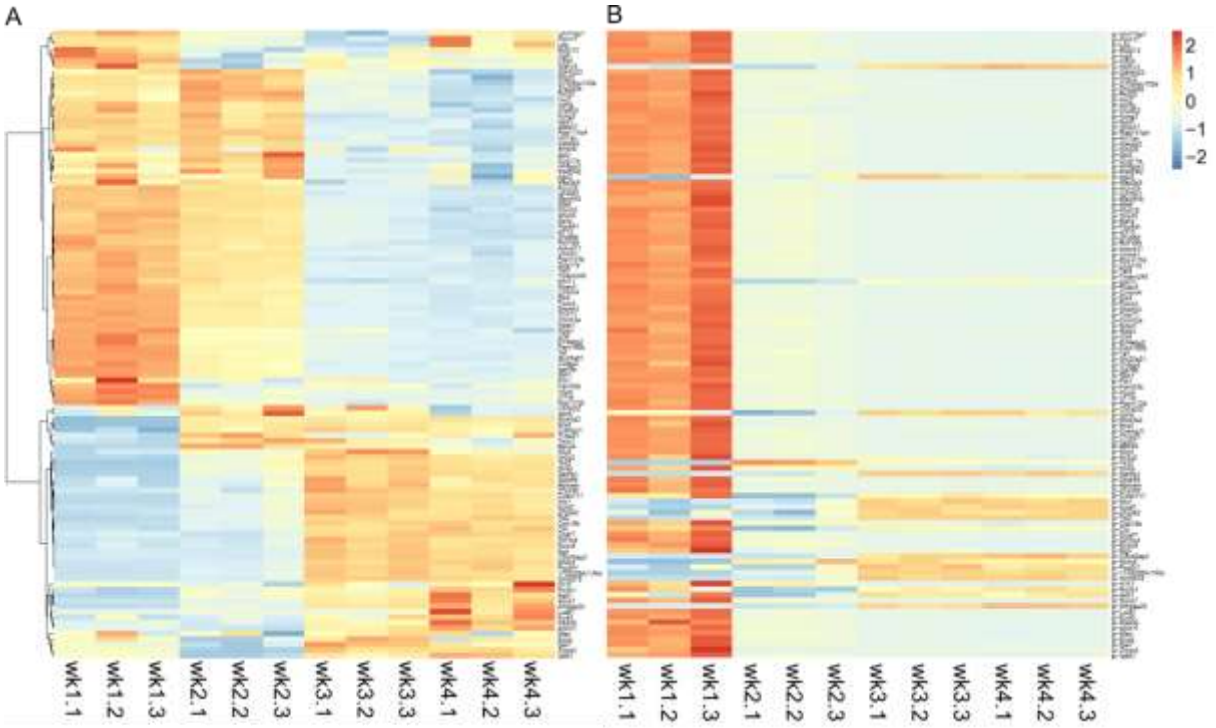


Figure 11. Expression patterns of the 114 piRNA-producing mRNAs and the corresponding piRNAs during the first wave of spermatogenesis. Heatmaps were plotted based on normalized read counts of piRNA-producing mRNAs (A) and piRNAs produced from each mRNA (B) in three replicates of 1, 2, 3, and 4 weeks old mouse testes.

2.3 Association of SMG6, MIWI, and piRNAs with ribosomes

2.3.1 SMG6, MIWI, and piRNAs associate with ribosomes

To start exploring the potential interplay of the CB components SMG6, MIWI and piRNAs with the translation machinery, we first studied if they co-fractionate with polysomes in the testis. This was studied using polysome fractionation of mouse testicular extracts by density gradient ultracentrifugation, followed by protein and RNA analyses. Free RNP, monosome, and polysome fractions were determined based on 260 nm UV absorbance profile of the fractions and by the presence of ribosomal RNA in the fractions (Figure 12).

Western blotting of the fractions using specific antibodies revealed that SMG6 and MIWI co-fractionated with polysomes. Furthermore, gel electrophoresis and staining of RNA extracted from the fractions showed that also piRNAs are present in the polysome fractions. In addition to polysome fractions, MIWI and piRNAs were enriched in the free RNP fractions, and SMG6 in the monosome fractions (Figure 12A). To confirm that the studied components were associated with ribosomes, the testis lysate was treated with EDTA, which disassociates ribosome subunits from each other and from the mRNA. EDTA-treatment before

fractionation shifted SMG6, MIWI, and piRNAs from the polysome fractions to the free RNP fractions, indicating the association with ribosomes (Figure 12B). The dissociation of ribosomes with EDTA-treatment was not complete, but some ribosomes remained intact.

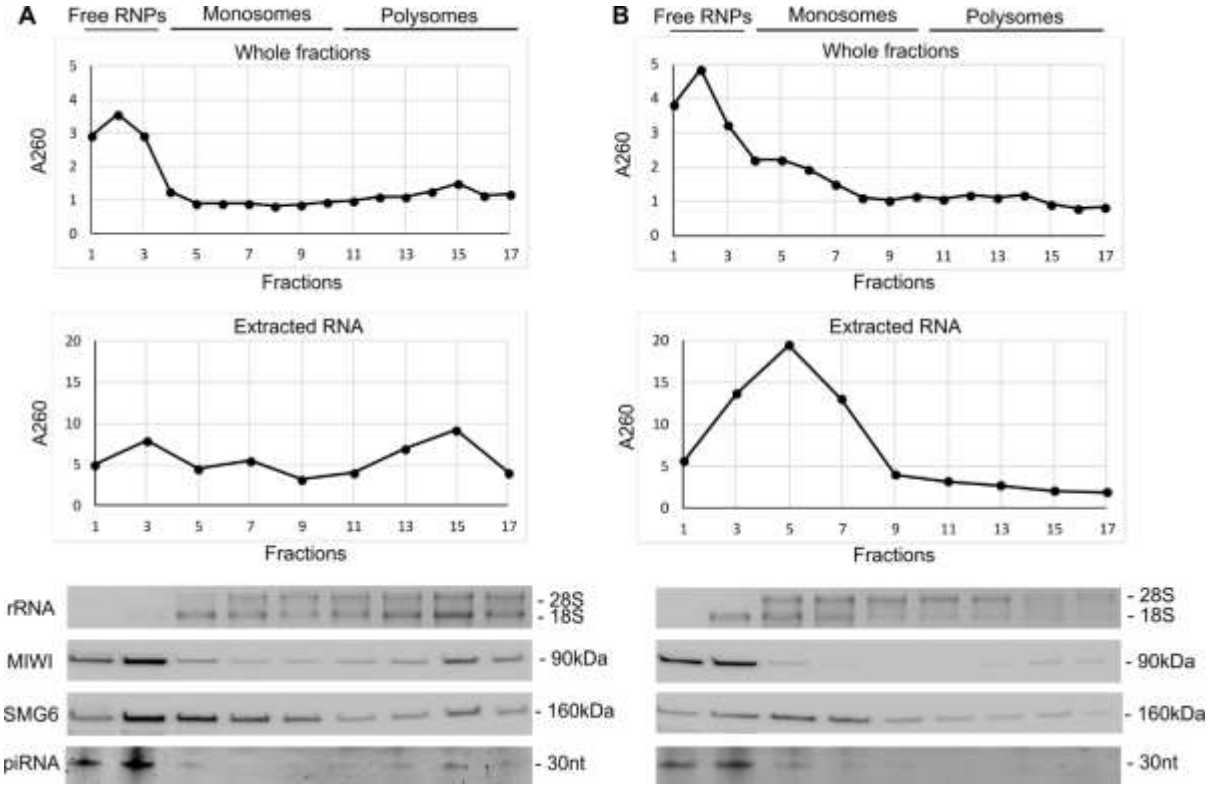


Figure 12. SMG6, MIWI, and piRNAs associate with ribosomes. Untreated (A) and EDTA-treated (B) testis lysates were fractionated using density gradient ultracentrifugation, followed by UV spectrometry, Western blot (WB), and RNA electrophoresis analyses. Fractions for free RNPs, monosomes, and polysomes were determined based on the 260 nm absorption profile of the fractions, as shown in the top of the figures. The second graph from the top shows the 260 nm absorption profile of the RNA extracted from the fractions. The bottom parts of the figure show the presence of ribosomal RNA (agarose gel electrophoresis), MIWI (WB), SMG6 (WB), and piRNAs (urea-PAGE) in the fractions.

2.3.2 MIWI and piRNAs associate with ribosomes independently of SMG6

To study the role of SMG6 in the association of MIWI, piRNAs and polysomes, *Smg6*-knockout (*Smg6*-KO) testes were used. Overall fractionation profile was comparable between *Smg6*-KO and control testes, and no major changes were observed in the fractionation of MIWI and piRNAs in the absence of SMG6 (Figure 13). A small proportion of MIWI was shifted from monosome and polysome fractions to the free RNP fractions, but this could be due to different cell compositions of the WT and *Smg6*-KO testes. Altogether, our results showed that SMG6, MIWI and piRNAs associate with polysomes, and the association of MIWI and piRNAs with polysomes is not dependent on SMG6.

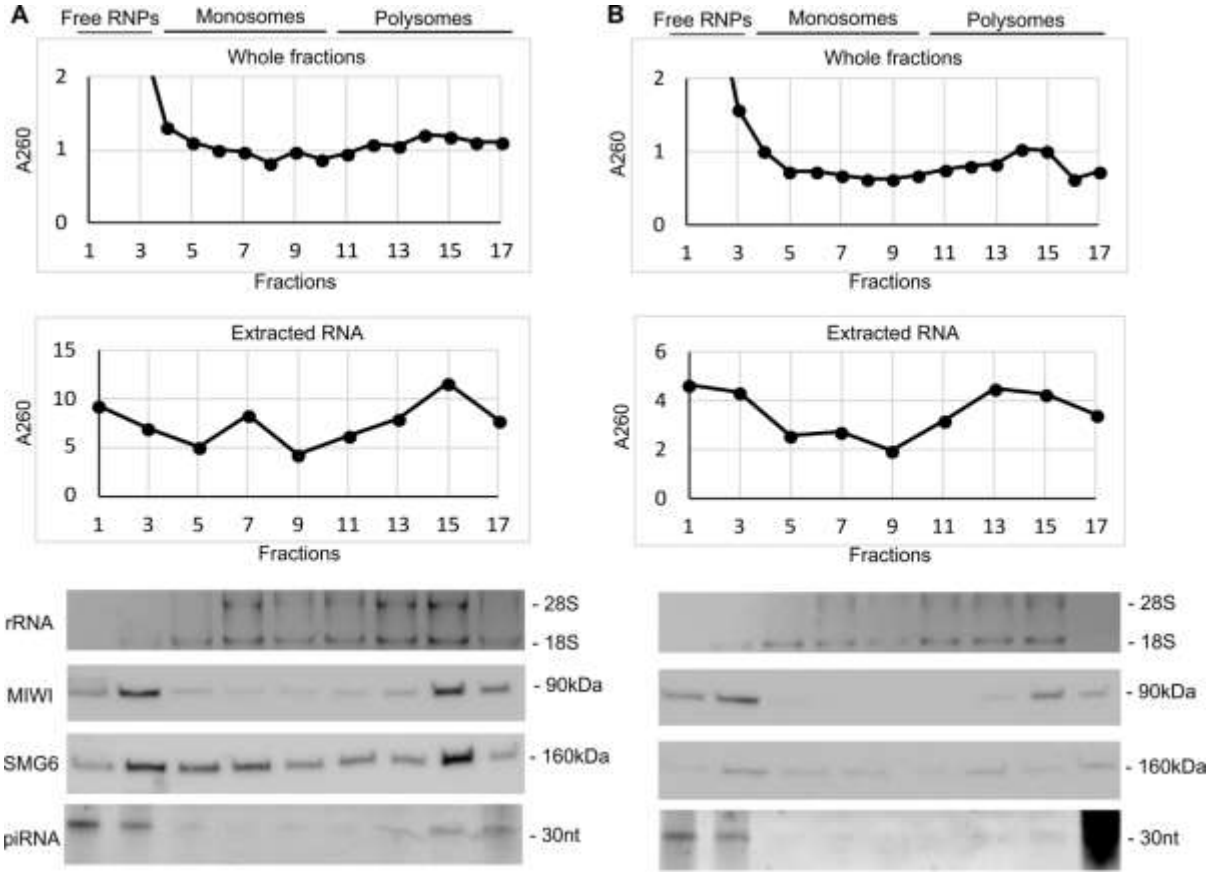


Figure 13. Deletion of SMG6 has no major effects in the association of MIWI and piRNAs with ribosomes. Wild-type (A) and SMG6 knockout (B) testis lysates were fractionated using density gradient ultracentrifugation, followed by UV spectrometry, Western blot, and RNA electrophoresis analyses. Fractions for free RNPs, monosomes, and polysomes were determined based on the 260 nm absorption profile of the fractions, as shown in the top of the figures. The second graph from the top shows the 260 nm absorption profile of the RNA extracted from the fractions. The y-axes of the absorbance graphs were cut at the value of 2 to better visualize the patterns. The bottom parts of the figure show the presence of ribosomal RNA (agarose gel electrophoresis), MIWI (WB), SMG6 (WB), and piRNAs (urea-PAGE) in the fractions. The weak SMG6-signal in *Smg6*-KO testes comes from the somatic cells of the testis, as the KO is germ cell-specific. The RNA from fraction 17 of the *Smg6*-KO testes was degraded during the RNA electrophoreses, but that doesn't affect the interpretation of the results.

2.4 piRNA-producing mRNAs are actively translated independently of SMG6

Our earlier finding that genic piRNAs are produced from protein-coding mRNAs prompted us to investigate if these piRNA-producing mRNAs are actively translated in the testis. RT-PCR was used to study the distribution of 6 piRNA-producing mRNAs in the free RNP, monosome, and polysome fractions to assess the state of their translation. All studied mRNAs were found in the polysome fractions, suggesting active translation. SMG6 was not required for the association of these mRNAs with polysomes as their fractionation was not significantly affected in *Smg6*-KO testes (Figure 14).

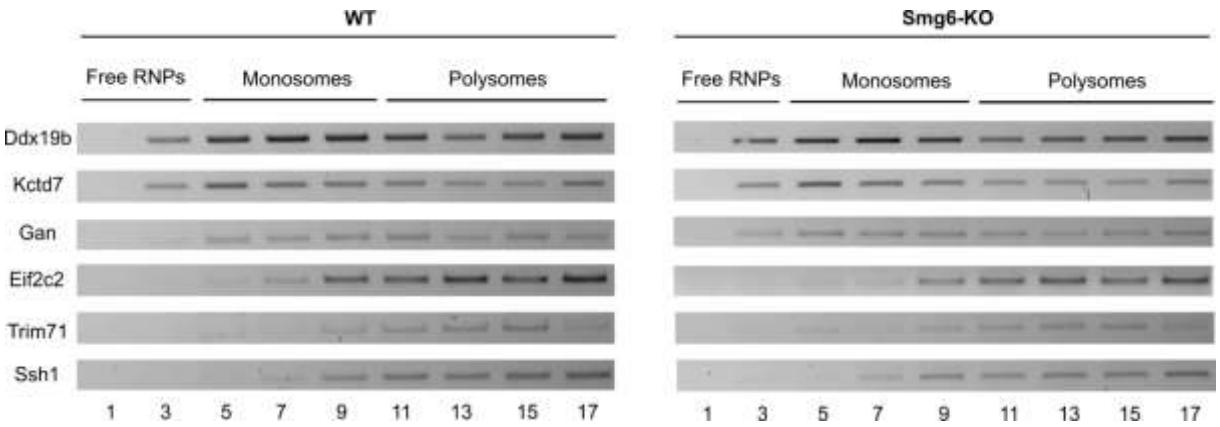


Figure 14. piRNA-producing mRNAs are actively translated in both WT and *Smg6*-KO testes. RT-PCR was used to study the presence of 6 piRNA-producing mRNAs in the testis fractions after polysome fractionation of WT and *Smg6*-KO testis lysates. Fractions for free RNPs, monosomes, and polysomes were determined based on the 260 nm absorption profile of the fractions, as presented in the figure.

3 Discussion

3.1 Genic piRNA production

The existence of piRNAs originating from protein-coding areas of the genome was discovered already soon after the discovery of piRNAs. However, it has remained elusive how the genic piRNAs link to the protein-coding genes the piRNA-clusters overlap with. The results of this study suggested that all 114 annotated genic piRNA clusters (X. Z. Li et al., 2013) are equal to the genes they overlap with. Considering the previous finding that the piRNAs of 30 hybrid clusters, that are also genic, are produced from full-length mRNAs (Y. H. Sun et al., 2021), it is highly probable that all genic piRNAs are produced from mRNAs of protein-coding genes. Interestingly, genic piRNAs mostly map specifically to the 3'UTRs of the genes (Gan et al., 2011). This 3'UTR-targeted production pattern indicates that the genic piRNA production is not random, but a highly coordinated process. The 3'UTR-targeted piRNA production could also explain the discovered appearance of certain mRNA isoforms with truncated 3'UTRs during spermatogenesis (Dong et al., 2002; Pero et al., 2003; Persengiev et al., 1995). It is not known how the piRNA production is directed to the 3'UTRs, but it has been suggested that differences in post-transcriptional processing of 3'UTRs could explain the piRNA production specifically from the those areas (Y. H. Sun et al., 2021).

Based on previous results from our lab, we hypothesized that NMD component SMG6 has a role in the production of piRNAs from mRNAs (Lehtiniemi et al., 2022). The results of the present study support this hypothesis, since SMG6, MIWI, piRNAs, and piRNA-producing mRNAs were all found associated to ribosomes. However, the deletion of SMG6 didn't cause any significant changes in the ribosome-association of the mentioned components. Changes especially in the ribosome-association of piRNA-producing mRNAs, which represents the activity of their translation, would have been expected, since *Smg6*-KO induced the production of piRNAs from those mRNAs in our previous study (Lehtiniemi et al., 2022), and ribosomes are thought to guide the translation-linked piRNA production from mRNAs (Y. H. Sun et al., 2021). Additionally, the piRNA-producing mRNAs were not differentially expressed in *Smg6*-KO testes (Lehtiniemi et al., 2022), which would also have been expected as they are used more in the piRNA production. So, it seems that in the absence of SMG6, the abundance and translational activity of piRNA-producing mRNAs remain unchanged, but the amount of piRNAs produced from them is increased. One explanation for this would be that in *Smg6*-KO testes, the piRNA production compensates for the defective degradation of the

mRNAs by NMD to keep the protein levels constant. Supporting this hypothesis, long 3'UTRs are a characteristic for both NMD-targets and piRNA precursors (Lehtiniemi et al., 2022; Y. H. Sun et al., 2021). It is also possible that only the 3'UTRs of the mRNAs are turned into piRNAs more efficiently in the *Smg6*-KO, and thus the abundance of the mRNAs and their translational activity remain unchanged.

The significance of this 3'UTR piRNA production has also remained unclear. It has been observed that the 3'UTRs of piRNA-producing mRNAs tend to contain TE sequences, and the generated piRNAs seem to target TEs (X. Wang et al., 2022). Even though the piRNA-hosting genes differ between species, the presence of TE sequences in the 3'UTR is suggested to be a conserved feature (Y. H. Sun et al., 2021). Since the production of TE-silencing fetal piRNAs ends perinatally leading to increased expression of TEs in the early spermatogenesis, the production of TE-targeting genic piRNAs is suggested to be crucial for the genome protection throughout spermatogenesis, and could thus be the evolutionary reason for genic piRNA production (Y. H. Sun et al., 2021). Additionally, the production of piRNAs from the 3'UTRs of protein-coding genes has also been found to affect the expression levels of the proteins, including proteins needed for translational and transcriptional regulation during spermatogenesis, such as AGO2 (Y. H. Sun et al., 2021). Sun *et al*. (2021) found that a deficiency in piRNA production increased the level of AGO2, as well as AGO2-mediated miRNA-guided degradation of RNAs that are essential for fertility. Thus, fine-tuning of protein expression could also be the driving factor for genic piRNA production.

Another elusive aspect of genic piRNA production is the selection of mRNAs for piRNA production. The genes are not testis-specific but seem to be highly relevant for the cellular function also in spermatogenesis, being involved in the transcriptional, post-transcriptional, and translational regulation. piRNA-hosting genes in chicken, although being different, share this importance for spermatogenesis (Y. H. Sun et al., 2021). The selection of mRNAs has at least some specificity, since the piRNA production does not correlate with the abundance of the mRNAs (Robine et al., 2009). In *D. melanogaster*, the piRNA clusters are epigenetically marked (Klattenhoff et al., 2009), but similar marks have not been identified in mice. Besides an epigenetic mark, another option would be to post-transcriptionally recognize the mRNAs. piRNA precursors have indeed been discovered to have long first exons and long 3'UTRs that host TE sequences (Y. H. Sun et al., 2021; X. Wang et al., 2022), but no clear evidence of this kind of recognition has been presented yet.

The number of piRNA-hosting genes in humans happens to be similar to the number in mice, which raises more suggestions that the genic piRNA production is somehow targeted and evolutionarily conserved (Özata et al., 2019). However, it has been suggested that even one-third of mouse mRNAs could be processed into piRNAs (Gan et al., 2011). This is based on the alignment of single piRNA sequences with protein-coding genes. The 214 piRNA clusters used in this thesis were determined experimentally and discovered to account for more than 95% of all piRNAs (X. Z. Li et al., 2013). It is of course possible that some piRNAs are produced from several other mRNAs in addition to the 114 discussed in this thesis, but the significance of that is questionable due to the low rate of production.

To further understand the dynamics of piRNA production from mRNAs, we analyzed the correlation of genic piRNA and piRNA-producing mRNA expressions during spermatogenesis. Our analysis suggested that the expressions positively correlate only partially. More than half of the mRNAs were most expressed at the same time as the expression of the respective piRNAs peak, but for the rest of the mRNAs the expressions didn't peak at the same time. A positive correlation has been reported previously by others, using a different list of piRNA-hosting genes (Gan et al., 2011). It would be rational that either the piRNA production decreases the mRNA levels, meaning negative correlation, or that piRNAs are more abundant when also the mRNAs are more abundant, meaning positive correlation. Deeper analysis would be required to determine whether these correlations truly exist and what is their significance. Overall, more in-depth analyses are needed in the future to reveal which mRNAs are turned into piRNAs, when and how it is done, and why.

3.2 Connection between germ granules and translation

This study confirmed that NMD component SMG6 and PIWI-protein MIWI, which are highly enriched in the CB, are also associated with ribosomes. This supports our hypothesis of an association of germ granules and translation, which could be mediated by the NMD and piRNA pathways. As there is a physical interaction of SMG6 and MIWI in male germ cells (Lehtiniemi et al., 2022), we expect them to be interacting at the ribosomes as well, although that wasn't shown in this study. Unexpectedly, the knockout of SMG6 didn't affect the association of MIWI and piRNAs with ribosomes. However, the independency of MIWI localization to the polysomes is consistent with our previous finding that both MIWI and SMG6 localize to the CB independently of each other, even though they are also found

interacting (Lehtiniemi et al., 2022). So even though they are co-operating, their localization inside the cell appears to be independent of each other.

Overall, our results, together with all previous evidence discussed in section 1.5, strongly suggest a close association of germ granules and translation that could be mediated by the piRNA and NMD pathways, but the details and significance of this association are still largely unknown. Ribosomes could be acting as a platform for recognition of RNAs to determine if they are translated, stored for later translation, degraded by the NMD, or turned into piRNAs. MIWI and SMG6 could be co-operating in this recognition of RNAs, as has been previously suggested (Lehtiniemi et al., 2022), along with other unidentified factors. The recognition would probably be happening during the first rounds of translation, as has been suggested for NMD recognition (Ishigaki et al., 2001). The outcome processes could be taking place in the CB, since mRNAs and all the mentioned factors, as well as several RNA binding and modifying proteins, are accumulated there (Meikar et al., 2014). The translation, too, could be happening inside or near the CB, as was shown for *D. melanogaster* germ granules (Ramat et al., 2023). However, further research is required to reveal the processes that take place after the mRNA has been exported from the nucleus to determine the fate of the mRNA. In our lab, the dynamics between germ granules and translation, and the role of NMD and piRNA pathways in this association, will be further studied.

3.3 Significance of the study

This study adds to the pool of valuable information about the principles of piRNA production, and the mechanisms by which the RNA regulatory pathways in male germ cells are functioning together to ensure the production of functioning sperm and healthy offspring. Since both the piRNA and the NMD pathways are required for successful spermatogenesis, the presented results supplement the inadequate understanding of the factors leading to male infertility. Right now, the etiology of male infertility in most cases is not known (Schlegel, 2022). Understanding what is happening during spermatogenesis and how it is regulated is crucial in order to understand, diagnose, treat, and most importantly prevent male infertility and stop the current alarming development. It is not known how, for example, the environment or lifestyle can affect male fertility, or what are the reasons behind unsuccessful in vitro fertilization. Being able to predict the outcomes of ART would mean huge savings in terms of resources as well as human suffering. For example, a variant of the piRNA biogenesis factor PNLDC1 has been found to be associated with decreased fertility and poor

outcomes of ART (Nagirnaja et al., 2021). The components of the piRNA pathway could thus be used as a biomarker of infertility or defective sperm function. Furthermore, as ART enables reproduction even though the sperm might have defects, the long-term consequences of the operation are not known. Epigenetic mechanisms, where sperm RNAs play a big role, might be causing the defects and thus transmitting the consequences to the offspring. Further research on these mechanisms is required to understand and prevent this transmission.

Sperm RNAs, indeed, have been shown to be able to transmit information about father's environmental exposures, such as a high-fat diet, to the offspring and affect their health (Waldron, 2016). It is currently uncertain how the RNA profile of the transcriptionally inactive sperm is formed, but the CB has the potential to be a key contributor to this formation, being associated with several RNA processing machineries. It is possible that the CB accumulates RNAs to deliver them to the mature sperm. Understanding the mechanisms of the sperm RNA profile formation is the basis of creating strategies for prevention of the transmission of harmful information to the offspring and thus for decreasing the rapidly increasing incidence of severe and complex diseases, such as metabolic diseases, the heritability of which can't be solely explained by genetics (Sharma, 2019).

In addition to understanding and preventing the discussed phenomena, the constantly increasing knowledge about RNA regulation during spermatogenesis can be a basis for future drug development. As overpopulation is a major problem in many parts of the world, an urgent need for male contraceptives exists. By inhibiting the germ cell-specific piRNA pathway, sperm production could potentially be halted. However, in spite of active research, no such intervention has successfully been developed yet.

In addition to the germline, the drugs targeting the discussed mechanisms could also be used to treat somatic conditions. The piRNA and NMD pathways have been associated with many somatic diseases, including cancer (Rayford et al., 2021; L. Sun et al., 2023), and some piRNA pathway components, including MIWI, are also known as cancer-germline antigens, which are proteins that are specifically expressed in the germline, but also show some activation in different types of cancer (C. Wang et al., 2016). Because of this restricted expression, they are very potential targets for cancer therapy, with the potential of having very limited adverse effects (Meng et al., 2021). Additionally, the co-operation of MIWI with the NMD pathway has been identified in cancer as well, highlighting the extensive importance of the present study (Shi et al., 2020). Further supporting the connections between

spermatogenesis and cancer, azoospermia has been found to increase risk of cancer (Eisenberg et al., 2013).

On the other hand, the knowledge on RNA regulatory mechanisms and epigenetics could be used to develop RNA therapies that target the epigenome. The potential of such approaches is huge, as the epigenome is involved in a wide variety of conditions, acting as a link between the environment and disease (Cavalli et al., 2019). One such, very intriguing possibility would be to use synthesized piRNAs as epigenome-modifying drugs. This is just one example of various possibilities, and future research will reveal how all the increasing knowledge will be utilized for the well-being of humans.

All in all, the present study is building the crucial understanding about the mechanisms of RNA regulation during spermatogenesis that widely contribute to human health. Regarding infertility, epigenetic inheritance, and somatic diseases like cancer, the thorough understanding of the studied mechanisms could lead to great advances in improving people's quality of life, decreasing the enormous burden on healthcare systems, and creating a healthier tomorrow.

4 Materials and methods

4.1 Characterization of piRNA cluster localization in the genome

The localization of 114 genic piRNA clusters in relation to the protein-coding genes they overlap with was analysed using ReproGenomicsViewer genome browser web tool. Coordinates for piRNA clusters were derived from Li et al. (2013) (X. Z. Li et al., 2013), and for the overlapping genes from Ensembl.

Overlap percentages of the genes and piRNA clusters were calculated by comparing the length of the overlap to the length of the gene or piRNA cluster in question. Percentage of cluster's length overlapping with gene was calculated, unless the cluster was completely inside the gene, in which case the percentage of gene's length overlapping with cluster was considered.

RefSeq and Ensembl reference genomes were used to study different transcript variants. RNA sequencing data from Gan et al. (2013), under accession number GSE35005, was used to visualize the expression of the genomic regions encoding the protein-coding genes and the overlapping piRNA clusters. To present the observations, snapshots were taken from the genome browser web tool.

4.2 RNA sequencing

Three biological replicates of previously extracted total RNA from 1, 2, 3, and 4 weeks old mouse testes were used for RNA sequencing. Both directional mRNA-seq libraries and small RNA-seq libraries were prepared and sequenced with Illumina NovaSeq 6000 platform at Novogene (UK) Company Limited.

4.3 mRNA data analysis

The quality of raw data was evaluated by FastQC (v0.11.9). The adapters and low-quality bases were removed from sequences using trimmomatic (v0.39) (Bolger et al., 2014). The clean data were then mapped to the mouse reference genome (Ensembl: Mus musculus.GRCm38.101) using STAR (v2.7.10a) (Dobin et al., 2013). The mapped reads were assigned and counted using featureCounts (v2.0.3) (Liao et al., 2014). Raw reads were normalized using DESeq2 (v1.40.1) (Love et al., 2014). Heatmap of mRNA expressions was

plotted based on the z-score of normalized counts. GO enrichment analysis for piRNA producing mRNAs was performed by topGO package (v2.52.0) (Alexa et al., 2024).

4.4 Small RNA data analysis

FastQC (v0.11.9) was first used to assess the quality of the raw data. The low-quality reads and reads ranging beyond 23-35 nucleotides were discarded. Cutadapt was used to trim off the adapters (Martin, 2011). The reads were mapped to the mouse genome (GRCm38) using Hisat2 (Kim et al., 2019). Then, reads mapping to tRNA or rRNA locations were discarded and the remaining reads were mapped to the previously annotated 114 piRNA clusters using featureCounts (v2.0.3) 3,8. The reads of piRNA clusters were normalized using DESeq2 (v1.40.1) (Love et al., 2014). Heatmap of piRNA cluster expressions was plotted based on the z-score of normalized counts according to the order of the corresponding mRNAs.

4.5 Mice

Mice were maintained and housed at the central animal facility of the University of Turku, Finland, under controlled pathogen-free conditions, following local laws and regulations (Finnish Act on the Protection of Animals Used for Scientific or Educational Purposes [497/2013], Government Decree on the Protection of Animals Used for Scientific or Educational Purposes [564/2013]). Mice were euthanized by CO_2 inhalation followed by cervical dislocation. The Laboratory Animal Care and Use Committee of the University of Turku approved all the animal experiments.

In all experiments, germ cell-specific *Smg6* conditional knockout mice with a mixed genetic background with C57Bl/6J and SV129, and their wild-type control mice, were used. The knockout was created using Cre-LoxP system with Neurogenin 3 promoter-driven Cre expression to direct the knockout to germ cells, as described previously (Lehtiniemi et al., 2022).

4.6 Polysome fractionation

10-50% continuous sucrose gradient was prepared by carefully overlaying 10%, 20%, 30%, 40%, and 50% sucrose solutions on top of each other, followed by diffusing overnight at +4°C. Sucrose solutions were prepared in a lysis buffer (containing 150 mM potassium acetate, 5 mM magnesium acetate, 2 mM DTT, cOmplete EDTA-free Protease Inhibitor

Cocktail (Roche, 11873580001), 80 U/ml RiboLock RNase Inhibitor (Thermo Scientific, EO0382), and 50 mM HEPES, pH 7.5).

Mouse testes were lysed in the lysis buffer supplemented with 100 µg/ml cycloheximide (Sigma-Aldrich, C4859), 0.5% Triton-X100, and 0.25 M sucrose, with two 5 mm Stainless Steel Beads (Qiagen, 69990), using TissueLyser LT (Qiagen) at 50 oscillations per second for 1 minute, followed by 30 minute incubation on ice. Testis lysates were then centrifuged at 500 g for 15 minutes at +4°C to remove tissue debris. Lysates were loaded on top of the 10-50% sucrose gradient and centrifuged at 29,000 rpm for 3 hours at +4°C. SW 32.1 Ti rotor (Beckman Coulter, 369651) and 17ml Open-Top Thinwall Ultra-Clear Tubes (Beckman Coulter, C14297) were used.

1ml fractions of the gradients were collected. Absorbances at 260 nm were measured using NanoDrop ND-1000 Spectrophotometer (Thermo Scientific) to determine free RNP, monosome, and polysome fractions. 50 µl of a 1ml fraction was collected for Western blotting, supplemented with SDS sample buffer (4% SDS, 20% glycerol, 10% 2-mercaptoethanol, 0.004% bromophenol blue, and 0.125 M Tris HCl), and stored in -20°C. 300 µl was collected for RNA extraction and stored in -80°C. For practical reasons, samples for Western blotting and RNA extraction were collected only from every other fraction (1, 3, 5, *etc.*).

4.7 Western blotting

Samples diluted in SDS sample buffer were incubated for 5 minutes at +95°C. Proteins were separated using 10% SDS-PAGE and transferred to PVDF membranes (Amersham, RPN303F) using wet-blotting system (Bio-Rad). Membranes were blocked with 5% skimmed milk, 0.1% triton X-100 in TBS, and then incubated in primary antibody against MIWI (Cell Signaling Technology, #2079) or SMG6 (Abcam, ab87539) diluted 1:1000 in blocking solution overnight at +4°C. HRP-linked anti-rabbit IgG (NA934) was used as secondary antibody. Proteins were detected using western lightening ECL pro solution (Ne112200IEA, Perkin Elmer, Netherlands) and LAS4000 imaging system (FujiFilm). Images were saved as 8-bit TIFF files and processed using Adobe Photoshop. When appropriate, the immunoblots were stripped Restore™ PLUS Western Blot Stripping Buffer (Thermo Scientific, 46430) and re-probed with a different antibody.

4.8 RNA extraction

RNA was extracted from the testis fractions with Trizol LS Reagent (Thermo Fisher Scientific, 10296010) using manufacturer's protocols. Extracted RNA was treated with DNase I (Sigma-Aldrich, D5307) and analysed using Nanodrop ND-1000 Spectrophotometer (Thermo Scientific).

4.9 RNA gel electrophoresis

Samples for RNA gel electrophoreses were prepared by diluting total RNA extracted from testis fractions to RNA Loading Dye (New England Biolabs, B0363S) followed by incubation for 5 minutes at +70°C. One fifth of the total RNA volume was used for each gel run.

Ribosomal RNA was analyzed using a 1% agarose gel and visualized with Midori Green Advance DNA stain (Nippon genetics, MG04).

piRNAs were analyzed using 15% urea-PAGE followed by staining with SYBR Gold (Invitrogen, S11494).

4.10 RT-PCR

Part of the total RNA from the testis fractions was converted to cDNA using SensiFAST cDNA Synthesis Kit (Meridian Bioscience, BIO-65054). Same volume of each fraction was used, so that the maximum amount of RNA per fraction was approximately 1 µg.

PCR was performed for 6 piRNA-producing mRNAs using 26-30 cycles, depending on primer pair, at 95°C for 30 sec, 55°C for 30 sec, and 72°C for 1 min. The PCR primers used are listed in Table 2.

Table 2. Primers used for RT-PCR.

Gene	Forward primer sequence	Reverse primer sequence
Trim71	TTATGTTCACGCCCCCTGAC	TGGGGCCTGTAGCTCACTAT
Ssh1	CAAGCAACAGCGAGTTGGAG	ATGCTTGTGAGGGTGCTGAA
Gan	AGCCCGTACATCAGGACAAA	CAGTAGTGCAGCGCAAAGTC
Eif2c2	CACGTGTGTTTACGGCAGTG	CCCTCGTCAACGAAAGGGAA
Ddx19b	AGCTCGGAATCCACTTTCCC	AAGCTGTGGTTTCAGGCGAA
Kctd7	CCCCAGGAGTTTCCTGAAGTC	GAACGCAGGAAGTTCAGCAC

5 Acknowledgements

I would like to thank my supervisor, Professor Noora Kotaja for the guidance, support, and feedback during the entire thesis process. I thank all the members of the Kotaja group for the support and help in the experiments, especially Riikka Palimo for conducting some of the experiments, and Lin Ma and Matthieu Bourgery for conducting, and guiding with, the bioinformatic analyses. I also thank all other personnel of the Medisiina C5 floor for support when needed. I thank Veijo Hukkanen for the assistance with ultracentrifugation. Finally, I thank all other people, friends and family, involved in supporting the process.

6 Abbreviations

ART	Assisted reproductive technologies
CB	Chromatoid body
IMC	Intermitochondrial cement
KO	Knockout
NMD	Nonsense-mediated mRNA decay
PGC	Primordial germ cell
piRNA	piwi-interacting RNA
SSC	Spermatogonial stem cell
TE	Transposable element
WB	Western blot

References

Alexa, A., & Maintainer, J. R. (2024). *Package "topGO" Type Package Title Enrichment Analysis for Gene Ontology*. Retrieved from https://git.bioconductor.org/packages/topGO

Amann, R. P. (2008). The Cycle of the Seminiferous Epithelium in Humans: A Need to Revisit? *Journal of Andrology*, *29* (5), 469–487. doi: 10.2164/JANDROL.107.004655

Aravin, A. A., Sachidanandam, R., Girard, A., Fejes-Toth, K., & Hannon, G. J. (2007). Developmentally regulated piRNA clusters implicate MILI in transposon control. *Science*, *316* (5825), 744–747. doi: 10.1126/SCIENCE.1142612/SUPPL_FILE/ARAVIN.SOM.PDF

Balagopal, V., & Parker, R. (2009). Polysomes, P bodies and Stress granules: States and Fates of Eukaryotic mRNAs. *Current Opinion in Cell Biology*, *21* (3), 403. doi: 10.1016/J.CEB.2009.03.005

Barratt, C. L. R., Björndahl, L., De Jonge, C. J., Lamb, D. J., Martini, F. O., McLachlan, R., Oates, R. D., van der Poel, S., John, B. S., Sigman, M., Sokol, R., & Tournaye, H. (2017). The diagnosis of male infertility: An analysis of the evidence to support the development of global WHO guidance-challenges and future research opportunities. *Human Reproduction Update*, *23* (6), 660–680. doi: 10.1093/humupd/dmx021

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30* (15), 2114–2120. doi: 10.1093/bioinformatics/btu170

Brunn, A. v. (1876). Beiträge zur Entwicklungsgeschichte der Samenkörper. *Archiv Für Mikroskopische Anatomie*, *12* (1), 528–535. doi: 10.1007/BF02933904/METRICS

Carmell, M. A., Girard, A., van de Kant, H. J. G., Bourc'his, D., Bestor, T. H., de Rooij, D. G., & Hannon, G. J. (2007). MIWI2 Is Essential for Spermatogenesis and Repression of Transposons in the Mouse Male Germline. *Developmental Cell*, *12* (4), 503–514. doi: 10.1016/J.DEVCEL.2007.03.001

Cavalli, G., & Heard, E. (2019). Advances in epigenetics link genetics to the environment and disease. *Nature 2019 571:7766*, *571* (7766), 489–499. doi: 10.1038/s41586-019-1411-0

Dai, P., Wang, X., Gou, L. T., Li, Z. T., Wen, Z., Chen, Z. G., Hua, M. M., Zhong, A., Wang, L., Su, H., Wan, H., Qian, K., Liao, L., Li, J., Tian, B., Li, D., Fu, X. D., Shi, H. J., Zhou, Y., & Liu, M. F. (2019). A Translation-Activating Function of MIWI/piRNA during Mouse Spermiogenesis. *Cell*, *179* (7), 1566-1581.e16. doi: 10.1016/J.CELL.2019.11.022

Da Ros, M., Hirvonen, N., Olotu, O., Toppari, J., & Kotaja, N. (2015). Retromer vesicles interact with RNA granules in haploid male germ cells. *Molecular and Cellular Endocrinology*, *401*, 73–83. doi: 10.1016/J.MCE.2014.11.026

Deng, W., & Lin, H. (2002). miwi, a Murine Homolog of piwi, Encodes a Cytoplasmic Protein Essential for Spermatogenesis. *Developmental Cell*, *2* (6), 819–830. doi: 10.1016/S1534-5807(02)00165-X

DiGiacomo, M., Comazzetto, S., Saini, H., DeFazio, S., Carrieri, C., Morgan, M., Vasiliauskaite, L., Benes, V., Enright, A. J., & O'Carroll, D. (2013). Multiple epigenetic mechanisms and the piRNA pathway enforce LINE1 silencing during adult spermatogenesis. *Molecular Cell*, *50* (4), 601–608. doi: 10.1016/J.MOLCEL.2013.04.026

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, *29* (1), 15–21. doi: 10.1093/bioinformatics/bts635

Dong, L. Q., Ramos, F. J., Wick, M. J., Lim, M. A., Guo, Z., Strong, R., Richardson, A., & Liu, F. (2002). Cloning and characterization of a testis and brain-specific isoform of mouse 3 0-phosphoinositide-dependent protein kinase-1, mPDK-1b q. *Biochemical and Biophysical Research Communications*, *294* (1), 136–144. Retrieved from www.academicpress.com

Eisenberg, M. L., Betts, P., Herder, D., Lamb, D. J., & Lipshultz, L. I. (2013). Increased risk of cancer among azoospermic men. *Fertility and Sterility*, *100* (3), 681–685. doi: 10.1016/j.fertnstert.2013.05.022

Fajardo, M. A., Haugen, H. S., Clegg, C. H., & Braun, R. E. (1997). Separate Elements in the 3' Untranslated Region of the Mouse Protamine 1 mRNA Regulate Translational Repression and Activation during Murine Spermatogenesis. In DEVELOPMENTAL BIOLOGY (Vol. 191).

Fanourgakis, G., Lesche, M., Akpinar, M., Dahl, A., & Jessberger, R. (2016). Chromatoid Body Protein TDRD6 Supports Long 3' UTR Triggered Nonsense Mediated mRNA Decay. *PLOS Genetics*, *12* (5), e1005857. doi: 10.1371/JOURNAL.PGEN.1005857

Figueroa, J., & Burzio, L. O. (1998). Polysome-like structures in the chromatoid body of rat spermatids. *Cell and Tissue Research*, *291* (3), 575–579. doi: 10.1007/S004410051027/METRICS

Gan, H., Lin, X., Zhang, Z., Zhang, W., Liao, S., Wang, L., & Han, C. (2011). piRNA profiling during specific stages of mouse spermatogenesis. *RNA*, *17* (7), 1191–1203. doi: 10.1261/rna.2648411

Gapp, K., Jawaid, A., Sarkies, P., Bohacek, J., Pelczar, P., Prados, J., Farinelli, L., Miska, E., & Mansuy, I. M. (2014). Implication of sperm RNAs in transgenerational inheritance of the effects of early trauma in mice. *Nature Neuroscience*, *17* (5), 667–669. doi: 10.1038/nn.3695

Girard, A., Sachidanandam, R., Hannon, G. J., & Carmell, M. A. (2006). A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature*, *442* (7099), 199–202. doi: 10.1038/nature04917

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., Zhao, S., Hu, S. Da, Wu, L. G., Shi, H. J., Li, Y., Fu, X. D., Qu, L. H., Wang, E. D., & Liu, M. F. (2014). Pachytene piRNAs instruct massive mRNA elimination during late spermiogenesis. *Cell Research 2014 24:6*, *24* (6), 680–700. doi: 10.1038/cr.2014.41

Grandjean, V., Fourré, S., De Abreu, D. A. F., Derieppe, M. A., Remy, J. J., & Rassoulzadegan, M. (2015). RNA-mediated paternal heredity of diet-induced obesity and metabolic disorders. *Scientific Reports*, *5*. doi: 10.1038/srep18193

Grentzinger, T., Armenise, C., Brun, C., Mugat, B., Serrano, V., Pelisson, A., & Chambeyron, S. (2012). piRNA-mediated transgenerational inheritance of an acquired trait. *Genome Research*, *22* (10), 1877–1888. doi: 10.1101/GR.136614.111

Grivna, S. T., Beyret, E., Wang, Z., & Lin, H. (2006). A novel class of small RNAs in mouse spermatogenic cells. *Genes & Development*, *20* (13), 1709–1714. doi: 10.1101/GAD.1434406

Hess, R. A., & De Franca, L. R. (2008). Spermatogenesis and cycle of the seminiferous epithelium. *Advances in Experimental Medicine and Biology*, *636*, 1–15. doi: 10.1007/978-0-387-09597-4_1/COVER

Ishigaki, Y., Li, X., Serin, G., & Maquat, L. E. (2001). Evidence for a Pioneer Round of mRNA Translation: mRNAs Subject to Nonsense-Mediated Decay in Mammalian Cells Are Bound by CBP80 and CBP20. *Cell*, *106* (5), 607–617.

Jacobs, C., Kerna, N. A., & Tulp, O. L. (2019). *An Overview of the Causes and Consequences of Male Fertility Decline*. *3* (1). doi: 10.31031/PRM.2019.03.000555

Johnson, L., Petty, C. S., & Neaves, W. B. (1980). A Comparative Study of Daily Sperm Production and Testicular Composition in Humans and Rats. *Biology of Reproduction*, *22* (5), 1233–1243. doi: 10.1093/BIOLREPROD/22.5.1233

Kalt, M. R., Pinney, H. E., & Graves, K. (1975). Inhibitor induced alterations of chromatoid bodies in male germ line cells of Xenopus laevis. *Cell and Tissue Research*, *161* (2), 193–210. doi: 10.1007/BF00220368

Kazazian, H. H., & Moran, J. V. (2017). Mobile DNA in Health and Disease. *New England Journal of Medicine*, *377* (4), 361–370. doi: 10.1056/nejmra1510092

Kim, D., Paggi, J. M., Park, C., Bennett, C., & Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology*, *37*, 907–915. doi: 10.1038/s41587-019-0201-4

Klattenhoff, C., Xi, H., Li, C., Lee, S., Xu, J., Khurana, J. S., Zhang, F., Schultz, N., Koppetsch, B. S., Nowosielska, A., Seitz, H., Zamore, P. D., Weng, Z., & Theurkauf, W. E. (2009). The Drosophila HP1 Homolog Rhino Is Required for Transposon Silencing and piRNA Production by Dual-Strand Clusters. *Cell*, *138* (6), 1137–1149. doi: 10.1016/J.CELL.2009.07.014

Kleene, K. C. (2003). Patterns, mechanisms, and functions of translation regulation in mammalian spermatogenic cells. *Cytogenetic and Genome Research*, *103* (3–4), 217–224. doi: 10.1159/000076807

Kurosaki, T., Popp, M. W., & Maquat, L. E. (2019). Quality and quantity control of gene expression by nonsense-mediated mRNA decay. *Nature Reviews Molecular Cell Biology 2019 20:7*, *20* (7), 406–420. doi: 10.1038/s41580-019-0126-2

Laiho, A., Kotaja, N., Gyenesei, A., & Sironen, A. (2013). Transcriptome profiling of the murine testis during the first wave of spermatogenesis. *PloS One*, *8* (4). doi: 10.1371/JOURNAL.PONE.0061558

Lee, E. J., Banerjee, S., Zhou, H., Jammalamadaka, A., Arcila, M., Manjunath, B. S., & Kosik, K. S. (2011). Identification of piRNAs in the central nervous system. *RNA*, *17* (6), 1090–1099. doi: 10.1261/RNA.2565011

Lehtiniemi, T., Bourgery, M., Ma, L., Ahmedani, A., Mäkelä, M., Asteljoki, J., Olotu, O., Laasanen, S., Zhang, F. P., Tan, K., Chousal, J. N., Burow, D., Koskinen, S., Laiho, A., Elo, L. L., Chalmel, F., Wilkinson, M. F., & Kotaja, N. (2022). SMG6 localizes to the chromatoid body and shapes the male germ cell transcriptome to drive spermatogenesis. *Nucleic Acids Research*, *50* (20), 11470–11491. doi: 10.1093/NAR/GKAC900

Lehtiniemi, T., & Kotaja, N. (2018). Germ granule-mediated RNA regulation in male germ cells. *Reproduction*, *155* (2), R77–R91. doi: 10.1530/REP-17-0356

Levine, H., Jørgensen, N., Martino-Andrade, A., Mendiola, J., Weksler-Derri, D., Jolles, M., Pinotti, R., & Swan, S. H. (2023). Temporal trends in sperm count: a systematic review and meta-regression analysis of samples collected globally in the 20th and 21st centuries. *Human Reproduction Update*, *29* (2), 157–176. doi: 10.1093/HUMUPD/DMAC035

Liao, Y., Smyth, G. K., & Shi, W. (2014). FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, *30* (7), 923–930. doi: 10.1093/bioinformatics/btt656

Li, F., Yuan, P., Rao, M., Jin, C. H., Tang, W., Rong, Y. F., Hu, Y. P., Zhang, F., Wei, T., Yin, Q., Liang, T., Wu, L., Li, J., Li, D., Liu, Y., Lou, W., Zhao, S., & Liu, M. F. (2020). piRNA-independent function of PIWIL1 as a co-activator for anaphase promoting complex/cyclosome to drive pancreatic cancer metastasis. *Nature Cell Biology 2020 22:4*, *22* (4), 425–438. doi: 10.1038/s41556-020-0486-z

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., & Zamore, P. D. (2013). An Ancient Transcription Factor Initiates the Burst of piRNA Production during Early Meiosis in Mouse Testes. *Molecular Cell*, *50* (1), 67–81. doi: 10.1016/J.MOLCEL.2013.02.016

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15* (12). doi: 10.1186/s13059-014-0550-8

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, *17* (1), 10–12. doi: 10.14806/EJ.17.1.200

Meikar, O., Vagin, V. V., Chalmel, F., Sõstar, K., Lardenois, A., Hammell, M., Jin, Y., Da Ros, M., Wasik, K. A., Toppari, J., Hannon, G. J., & Kotaja, N. (2014). An atlas of chromatoid body components. *RNA*, *20* (4), 483–495. doi: 10.1261/RNA.043729.113

Meistrich, M. L., & Hess, R. A. (2013). *Assessment of Spermatogenesis Through Staging of Seminiferous Tubules*. 299–307. doi: 10.1007/978-1-62703-038-0_27

Meng, X., Sun, X., Liu, Z., & He, Y. (2021). A novel era of cancer/testis antigen in cancer immunotherapy. In International Immunopharmacology (Vol. 98). Elsevier B.V. doi: 10.1016/j.intimp.2021.107889

Nagirnaja, L., Lopes, A. M., Charng, W. L., Miller, B., Stakaitis, R., Golubickaite, I., Stendahl, A., Luan, T., Friedrich, C., Mahyari, E., Fadial, E., Kasak, L., Vigh-Conrad, K., Oud, M. S., Xavier, M. J., Cheers, S. R., James, E. R., Guo, J., Jenkins, T. G., … Conrad, D. F. (2022). Diverse monogenic subforms of human spermatogenic failure. *Nature Communications 2022 13:1*, *13* (1), 1–18. doi: 10.1038/s41467-022-35661-z

Nagirnaja, L., Mørup, N., Nielsen, J. E., Stakaitis, R., Golubickaite, I., Oud, M. S., Winge, S. B., Carvalho, F., Aston, K. I., Khani, F., van der Heijden, G. W., Marques, C. J., Skakkebaek, N. E., Rajpert–De Meyts, E., Schlegel, P. N., Jørgensen, N., Veltman, J. A., Lopes, A. M., Conrad, D. F., & Almstrup, K. (2021). Variant PNLDC1 , Defective piRNA Processing, and Azoospermia . *New England Journal of Medicine*, *385* (8), 707–719. doi: 10.1056/nejmoa2028973

Oakberg, E. F. (1956). Duration of spermatogenesis in the mouse and timing of stages of the cycle of the seminiferous epithelium. *American Journal of Anatomy*, *99* (3), 507–516. doi: 10.1002/AJA.1000990307

Olotu, O., Dowling, M., Homolka, D., Wojtas, M. N., Tran, P., Lehtiniemi, T., Da Ros, M., Pillai, R. S., & Kotaja, N. (2023). Intermitochondrial cement (IMC) harbors piRNA biogenesis machinery and exonuclease domain-containing proteins EXD1 and EXD2 in mouse spermatocytes. *Andrology*, *11* (4), 710–723. doi: 10.1111/ANDR.13361

Ozata, D. M., Gainetdinov, I., Zoch, A., OCarroll, D., & Zamore, P. D. (n.d.). PIWI-interacting RNAs: small RNAs with big functions. *Nature Reviews Genetics*. doi: 10.1038/s41576-018-0073-3

Özata, D. M., Yu, T., Mou, H., Gainetdinov, I., Colpan, C., Cecchini, K., Kaymaz, Y., Wu, P.-H., Fan, K., Kucukural, A., Weng, Z., & Zamore, P. D. (2019). Evolutionarily conserved pachytene piRNA loci are highly divergent among modern humans. *Nature Ecology & Evolution*, *4*, 156–168. doi: 10.1038/s41559-019-1065-1

Paronetto, M. P., & Sette, C. (2010). Role of RNA-binding proteins in mammalian spermatogenesis. *International Journal of Andrology*, *33* (1), 2–12. doi: 10.1111/J.1365-2605.2009.00959.X

Parvinen, L. -M, Jokelainen, P. T., & Parvinen, M. (1978). Chromatoid body and haploid gene activity: actinomycin D induced morphological alterations. *Hereditas*, *88* (1), 75–80. doi: 10.1111/J.1601-5223.1978.TB01605.X

Pembrey, M. E., Bygren, L. O., Kaati, G., Edvinsson, S., Northstone, K., Sjöström, M., & Golding, J. (2005). Sex-specific, male-line transgenerational responses in humans. *European Journal of Human Genetics 2006 14:2*, *14* (2), 159–166. doi: 10.1038/sj.ejhg.5201538

Perheentupa, A., & Toppari, J. (2023). Male fertility and semen quality are decreasing – Do we have the expertise to deal with this challenge? *Acta Obstetricia et Gynecologica Scandinavica*, *102* (12), 1606–1607. doi: 10.1111/AOGS.14693

Pero, R., Lembo, F., Chieffi, P., Pozzo, G. Del, Fedele, M., Fusco, A., Bruni, C. B., & Chiariotti, L. (2003). Translational Regulation of a Novel Testis-Specific RNF4 Transcript. *MOLECULAR REPRODUCTION AND DEVELOPMENT*, *66*, 1–7. doi: 10.1002/mrd.10322

Persengiev, S. P., Saffert, J. D., & Kilpatrick, D. L. (1995). An alternatively spliced form of the transcription factor Spl containing only a single glutamine-rich transactivation domain. *Biochemistry*, *92*, 9107–9111. Retrieved from https://www.pnas.org

Priskorn, Z. S., & Juul, L. (2016). Male Reproductive Disorders and Fertility Trends: Influences of Environment and Genetic Susceptibility. *Physiol Rev*, *96*, 55–97. doi: 10.1152/physrev.00017.2015.-It

Qiu, W., Guo, X., Lin, X., Yang, Q., Zhang, W., Zhang, Y., Zuo, L., Zhu, Y., Li, C. S. R., Ma, C., & Luo, X. (2017). Transcriptome-wide piRNA profiling in human brains of Alzheimer's disease. *Neurobiology of Aging*, *57*, 170–177. doi: 10.1016/J.NEUROBIOLAGING.2017.05.020

Ramat, A., Haidar, A., Garret, C., & Simonelig, M. (2023). Germ granule higher-order organization coordinates their different functions. *BioRxiv*, 2023.11.24.568558. doi: 10.1101/2023.11.24.568558

Rayford, K. J., Cooley, A., Rumph, J. T., Arun, A., Rachakonda, G., Villalta, F., Lima, M. F., Pratap, S., Misra, S., & Nde, P. N. (2021). Pirnas as modulators of disease pathogenesis. In International Journal of Molecular Sciences (Vol. 22, Issue 5, pp. 1–22). MDPI AG. doi: 10.3390/ijms22052373

Ripin, N., & Parker, R. (2023). Formation, function, and pathology of RNP granules. *Cell*, *186* (22), 4737–4756. doi: 10.1016/J.CELL.2023.09.006

Robine, N., Lau, N. C., Balla, S., Jin, Z., Okamura, K., Kuramochi-Miyagawa, S., Blower, M. D., & Lai, E. C. (2009). A Broadly Conserved Pathway Generates 3′UTR-Directed Primary piRNAs. *Current Biology*, *19* (24), 2066–2076. doi: 10.1016/J.CUB.2009.11.064

Schlegel, P. N. (2022). Human Spermatogenesis: Insights From the Clinical Care of Men With Infertility. *Frontiers in Endocrinology*, *13*. doi: 10.3389/fendo.2022.889959

Sharma, U. (2019). Paternal contributions to offspring health: Role of sperm small rnas in intergenerational transmission of epigenetic information. *Frontiers in Cell and Developmental Biology*, *7* (OCT), 1–15. doi: 10.3389/FCELL.2019.00215/BIBTEX

Shi, S., Yang, Z. Z., Liu, S., Yang, F., & Lin, H. (2020). PIWIL1 promotes gastric cancer via a piRNA-independent mechanism. *Proceedings of the National Academy of Sciences of the United States of America*, *117* (36), 22390–22401. doi: 10.1073/PNAS.2008724117/-/DCSUPPLEMENTAL

Shum, E. Y., Jones, S. H., Shao, A., Chousal, J. N., Krause, M. D., Chan, W. K., Lou, C. H., Espinoza, J. L., Song, H. W., Phan, M. H., Ramaiah, M., Huang, L., McCarrey, J. R., Peterson, K. J., De Rooij, D. G., Cook-Andersen, H., & Wilkinson, M. F. (2016). The Antagonistic Gene Paralogs Upf3a and Upf3b Govern Nonsense-Mediated RNA Decay. *Cell*, *165* (2), 382. doi: 10.1016/J.CELL.2016.02.046

Skinner, M. K., Ben Maamar, M., Sadler-Riggleman, I., Beck, D., Nilsson, E., McBirney, M., Klukovich, R., Xie, Y., Tang, C., & Yan, W. (2018). Alterations in sperm DNA methylation, non-coding RNA and histone retention associate with DDT-induced epigenetic transgenerational inheritance of disease. *Epigenetics & Chromatin*, *11* (1). doi: 10.1186/S13072-018-0178-0

Söderström, K. O., & Parvinen, M. (1976). Transport of material between the nucleus, the chromatoid body and the Golgi complex in the early spermatids of the rat. *Cell and Tissue Research*, *168* (3), 335–342. doi: 10.1007/BF00215311

Soumillon, M., Necsulea, A., Weier, M., Brawand, D., Zhang, X., Gu, H., Barthès, P., Kokkinaki, M., Nef, S., Gnirke, A., Dym, M., deMassy, B., Mikkelsen, T. S., & Kaessmann, H. (2013). Cellular Source and Mechanisms of High Transcriptome Complexity in the Mammalian Testis. *Cell Reports*, *3* (6), 2179–2190. doi: 10.1016/J.CELREP.2013.05.031

Sun, L., Mailliot, J., & Schaffitzel, C. (2023). Nonsense-Mediated mRNA Decay Factor Functions in Human Health and Disease. In Biomedicines (Vol. 11, Issue 3). MDPI. doi: 10.3390/biomedicines11030722

Sun, Y. H., Lee, B., Xin, ·, & Li, Z. (2022). The birth of piRNAs: how mammalian piRNAs are produced, originated, and evolved. *Mammalian Genome*, *33*, 293–311. doi: 10.1007/s00335-021-09927-8

Sun, Y. H., Wang, R. H., Du, K., Zhu, J., Zheng, J., Xie, L. H., Pereira, A. A., Zhang, C., Ricci, E. P., & Li, X. Z. (2021). Coupled protein synthesis and ribosome-guided piRNA processing on mRNAs. *Nature Communications*, *12* (1). doi: 10.1038/s41467-021-26233-8

Sun, Y. H., Zhu, J., Xie, L. H., Li, Z., Meduri, R., Zhu, X., Song, C., Chen, C., Ricci, E. P., Weng, Z., & Li, X. Z. (2020). Ribosomes guide pachytene piRNA formation on long intergenic piRNA precursors. *Nature Cell Biology*, *22* (2), 200–212. doi: 10.1038/s41556-019-0457-4

Vasileva, A., Tiedau, D., Firooznia, A., Müller-Reichert, T., & Jessberger, R. (2009). Tdrd6 Is Required for Spermiogenesis, Chromatoid Body Architecture, and Regulation of miRNA Expression. *Current Biology*, *19* (8), 630–639. doi: 10.1016/J.CUB.2009.02.047

Ventelä, S., Toppari, J., & Parvinen, M. (2003). Intercellular Organelle Traffic through Cytoplasmic Bridges in Early Spermatids of the Rat: Mechanisms of Haploid Gene Product Sharing. *Molecular Biology of the Cell*, *14* (7), 2768. doi: 10.1091/MBC.E02-10-0647

Voronina, E., Seydoux, G., Sassone-Corsi, P., & Nagamori, I. (2011). RNA Granules in Germ Cells. *Cold Spring Harb Perspect Biol.*, *3* (12). doi: 10.1101/cshperspect.a002774

Vourekas, A., Zheng, Q., Alexiou, P., Maragkakis, M., Kirino, Y., Gregory, B. D., & Mourelatos, Z. (2012). Mili and Miwi target RNA repertoire reveals piRNA biogenesis and function of Miwi in spermiogenesis. *Nature Structural & Molecular Biology 2012 19:8*, *19* (8), 773–781. doi: 10.1038/nsmb.2347

Waldron, D. (2016). *Non-coding RNA: Inheritance of diet-induced metabolic changes via tsRNAs*. doi: 10.1126/science.aad7977

Wang, C., Gu, Y., Zhang, K., Xie, K., Zhu, M., Dai, N., Jiang, Y., Guo, X., Liu, M., Dai, J., Wu, L., Jin, G., Ma, H., Jiang, T., Yin, R., Xia, Y., Liu, L., Wang, S., Shen, B., … Hu, Z. (2016). ARTICLE Systematic identification of genes with a cancer-testis expression pattern in 19 cancer types. *Nature Communications*, *7*. doi: 10.1038/ncomms10499

Wang, X., Ramat, A., Simonelig, M., & Liu, M.-F. (2022). Emerging roles and functional mechanisms of PIWI-interacting RNAs. *Nature Reviews Molecular Cell Biology*, *24*, 123–141. doi: 10.1038/s41580-022-00528-0

Wang, X., Zhang, Z.-Y., Zhao, S., & Liu, M.-F. (2024). New insights into small non-coding RNAs during spermatogenesis. *Science Bulletin*. doi: 10.1016/j.scib.2024.02.019

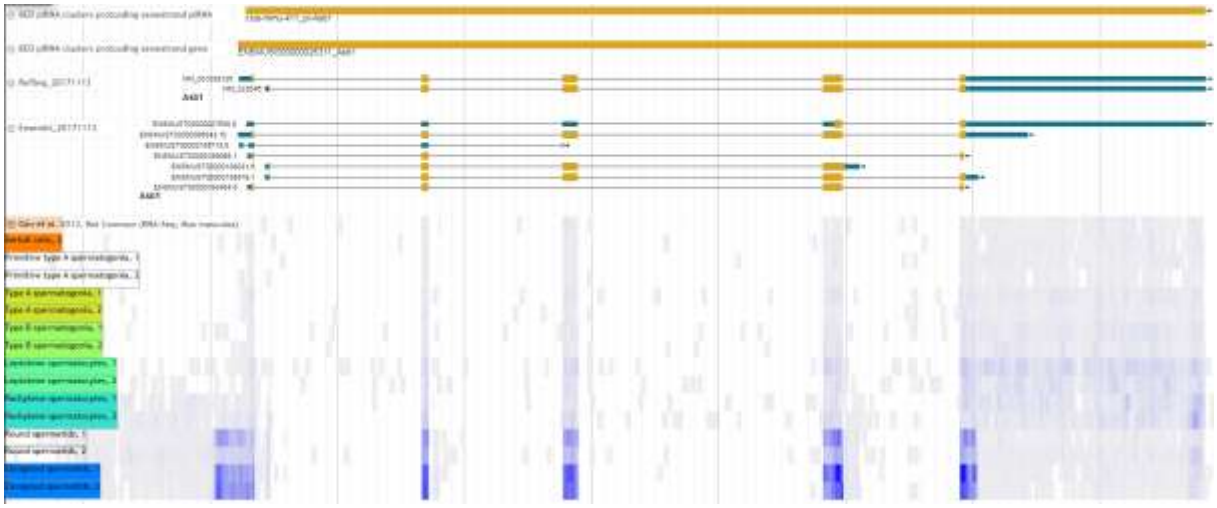
Wei, H., Gao, J., Lin, D.-H., Geng, R., Liao, J., Huang, T.-Y., Shang, G., Jing, J., Fan, Z.-W., Pan, D., Yin, Z.-Q., Li, T., Liu, X., Zhao, S., Chen, C., Li, J., Wang, X., Ding, D., & Liu, M.-F. (2024). piRNA loading triggers MIWI translocation from the intermitochondrial cement to chromatoid body during mouse spermatogenesis. *Nature Communications*, *15*, 2343. doi: 10.1038/s41467-024-46664-3

World Health Organization. (2023, April 3). *Infertility*. Https://Www.Who.Int/News-Room/Fact-Sheets/Detail/Infertility.

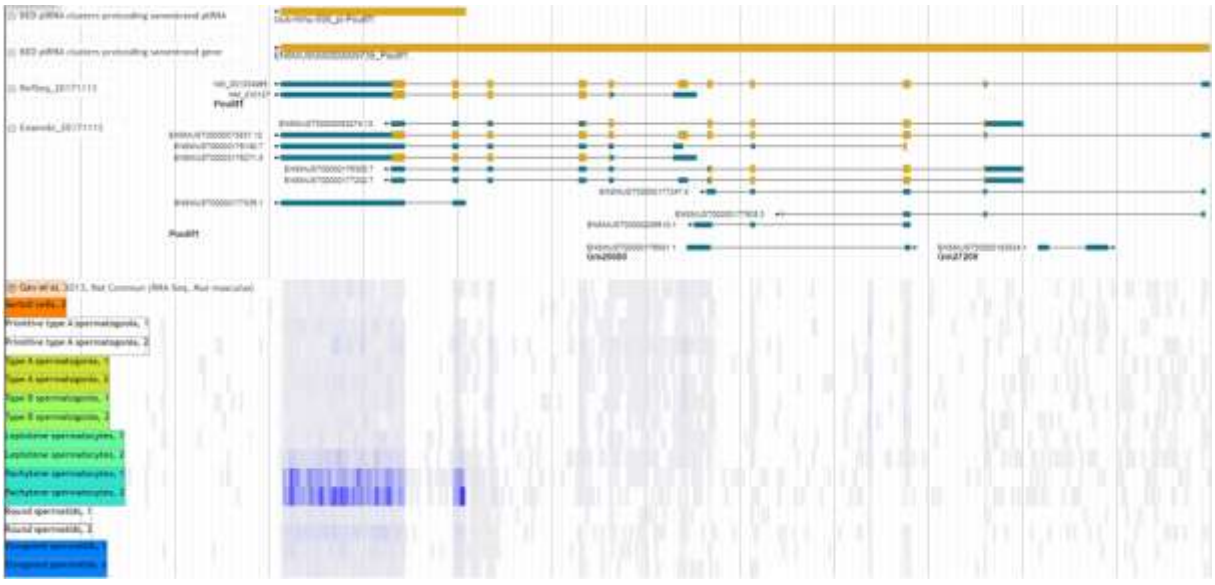
Zheng, X., Li, Z., Wang, G., Wang, H., Zhou, Y., Zhao, X., Cheng, C. Y., Qiao, Y., & Sun, F. (2021). Sperm epigenetic alterations contribute to inter- and transgenerational effects of paternal exposure to long-term psychological stress via evading offspring embryonic reprogramming. *Cell Discovery 2021 7:1*, *7* (1), 1–22. doi: 10.1038/s41421-021-00343-5

Appendices

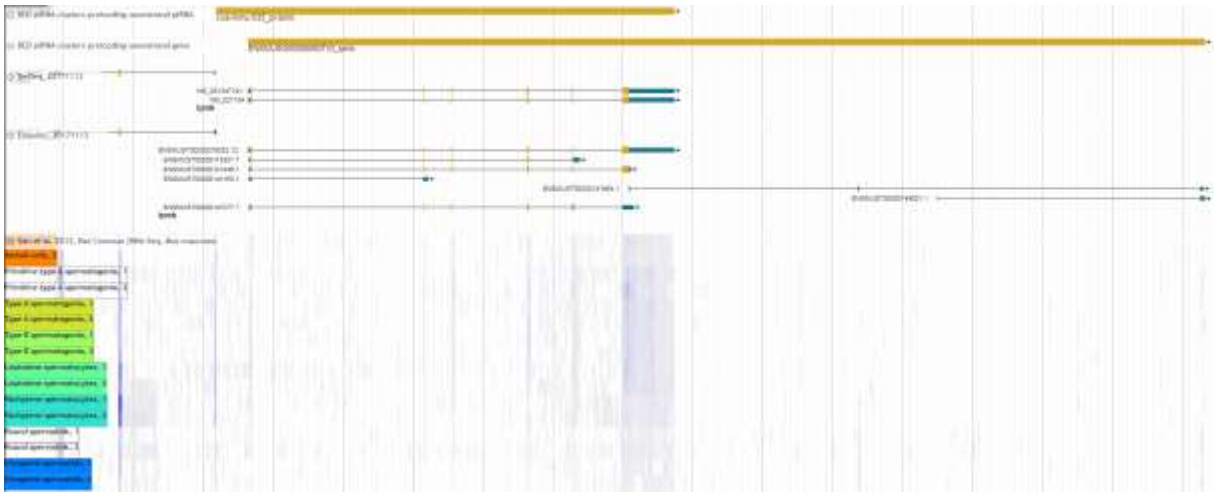
Appendix 1 Supplementary Figures



Supplementary Figure S1. Example of a piRNA cluster overlapping >99% with a protein-coding gene. Genomic coordinates of piRNA cluster pi-Asb1 (top yellow bar) were compared to coordinates of protein-coding gene Asb1 (second yellow bar, from Ensembl), individual transcripts of the gene (middle part, from RefSeq and Ensembl), and RNA sequencing data from different male mouse germ cell types (bottom half, data accession number GSE35005).

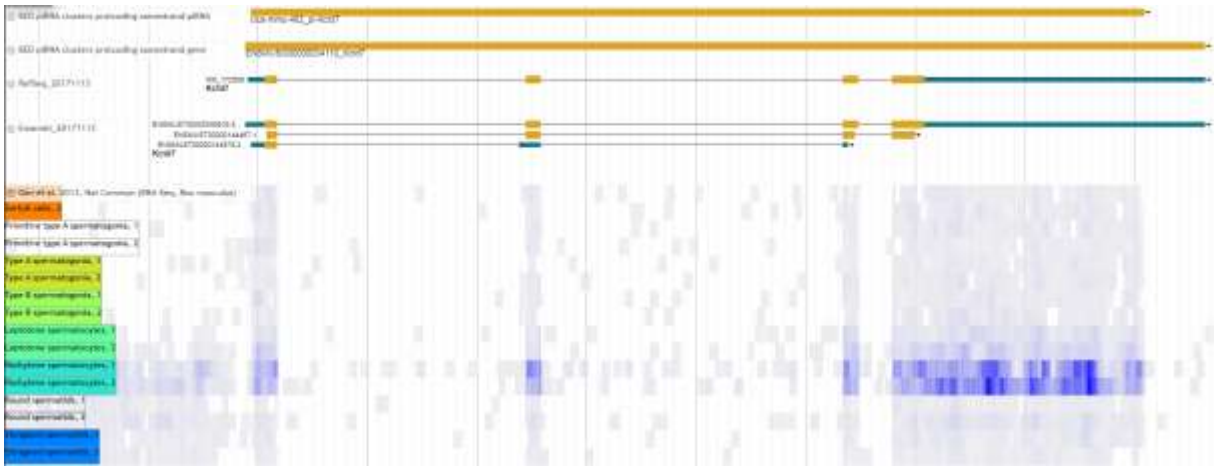


Supplementary Figure S2. Example of a piRNA cluster overlapping >99% with a single transcript of a protein-coding gene. Genomic coordinates of piRNA cluster pi-Pou6f1 (top yellow bar) were compared to coordinates of protein-coding gene Pou6f1 (second yellow bar, from Ensembl), individual transcripts of the gene (middle part, from RefSeq and Ensembl), and RNA sequencing data from different male mouse germ cell types (bottom half, data accession number GSE35005).



Supplementary Figure S3. piRNA cluster pi-Ipmk overlaps 93% with a single transcript of protein-coding gene Ipmk. Genomic coordinates of piRNA cluster pi-Ipmk (top yellow bar) were compared to coordinates of protein-coding gene Ipmk (second yellow bar, from Ensembl), individual transcripts of the gene (middle part, from RefSeq and Ensembl), and RNA sequencing data from different male mouse germ cell types (bottom half, data accession number GSE35005).

Supplementary Figure S4. Example of a piRNA cluster that would match the overlapping protein-coding gene with a longer 3' UTR. Genomic coordinates of piRNA cluster pi-Klhl11 (top yellow bar) were compared to coordinates of protein-coding gene Klhl11 (second yellow bar, from Ensembl), individual transcripts of the gene (middle part, from RefSeq and Ensembl), and RNA sequencing data from different male mouse germ cell types (bottom half, data accession number GSE35005).



Supplementary Figure S5. Example of a piRNA cluster that would match with the overlapping protein-coding gene with a shorter 3'UTR. Genomic coordinates of piRNA cluster pi-Kctd7 (top yellow bar) were compared to coordinates of protein-coding gene Kctd7 (second yellow bar, from Ensembl), individual transcripts of the gene (middle part, from RefSeq and Ensembl), and RNA sequencing data from different male mouse germ cell types (bottom half, data accession number GSE35005).

Supplementary Figure S6. Example of a piRNA cluster that matches with a combination of two protein coding transcripts. Genomic coordinates of piRNA cluster pi-Ip6k1 (top yellow bar) were compared to coordinates of protein-coding genes Ip6k1 and Cdhr4 (second row of yellow bars, from Ensembl), individual transcripts of the genes (middle part, from RefSeq and Ensembl), and RNA sequencing data from different male mouse germ cell types (bottom half, data accession number GSE35005).