TURUN YLIOPISTON JULKAISUJA
ANNALES UNIVERSITATIS TURKUENSIS

SARJA - SER. D OSA - TOM. 1099

MEDICA - ODONTOLOGICA

CHROMATOID BODY MEDIATED RNA REGULATION IN MOUSE MALE GERMLINE

by

Oliver Meikar

TURUN YLIOPISTO
UNIVERSITY OF TURKU
Turku 2013

From the Department of Physiology and Paediatrics, Institute of Biomedicine, University of Turku, and
Turku Doctoral Programme of Biomedical Sciences (TuBS), Turku, Finland

Supervised by

Assistant Professor Noora Kotaja, PhD
Department of Physiology
Institute of Biomedicine
University of Turku
Turku, Finland

and

Professor Jorma Toppari, MD, PhD
Departments of Physiology and Paediatrics
Institute of Biomedicine
University of Turku
Turku, Finland

Reviewed by

Liisa Kauppi, PhD
Biomedicum
University of Helsinki
Helsinki, Finland

and

Mikko Frilander, PhD
Institute of Biotechnology
University of Helsinki
Helsinki, Finland

Dissertation opponent

Professor Miles Wilkinson, PhD
Department of Reproductive Medicine
University of California
San Diego, USA

The originality of this thesis has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.

ISBN 978-951-29-5622-7 (PRINT)
ISBN 978-951-29-5623-4 (PDF)
ISSN 0355-9483
Painosalama Oy – Turku, Finland 2013

to my families and mentors

ABSTRACT

OLIVER MEIKAR
Chromatoid body mediated RNA regulation in mouse male germline

The Department of Physiology and Turku Doctoral Programme of Biomedical Sciences (TuBS) Institute of Biomedicine, Faculty of Medicine, University of Turku, Kiinamyllynkatu 10, 20520 Turku, Finland

Male germ cell differentiation, spermatogenesis is an exceptional developmental process that produces a massive amount of genetically unique spermatozoa. The complexity of this process along with the technical limitations in the germline research has left many aspects of spermatogenesis poorly understood. Post-meiotic haploid round spermatids possess the most complex transcriptomes of the whole body. Correspondingly, efficient and accurate control mechanisms are necessary to deal with the huge diversity of transcribed RNAs in these cells. The high transcriptional activity in round spermatids is accompanied by the presence of an uncommonly large cytoplasmic ribonucleoprotein granule, called the chromatoid body (CB) that is conjectured to participate in the RNA post-transcriptional regulation. However, very little is known about the possible mechanisms of the CB function.

The development of a procedure to isolate CBs from mouse testes was this study's objective. Anti-MVH immunoprecipitation of cross-linked CBs from a fractionated testicular cell lysate was optimized to yield considerable quantities of pure and intact CBs from mice testes. This protocol produced reliable and reproducible data from the subsequent analysis of CB's protein and RNA components. We found that the majority of the CB's proteome consists of RNA-binding proteins that associate functionally with different pathways. We also demonstrated notable localization patterns of one of the CB transient components, SAM68 and showed that its ablation does not change the general composition or structure of the CB. CB-associated RNA analysis revealed a strong accumulation of PIWI-interacting RNAs (piRNAs), mRNAs and long non-coding RNAs (lncRNAs) in the CB. When the CB transcriptome and proteome analysis results were combined, the most pronounced molecular functions in the CB were related to piRNA pathway, RNA post-transcriptional processing and CB structural scaffolding. In addition, we demonstrated that the CB is a target for the main RNA flux from the nucleus throughout all steps of round spermatid development. Moreover, we provided preliminary evidence that those isolated CBs slice target RNAs *in vitro* in an ATP-dependent manner.

Altogether, these results make a strong suggestion that the CB functions involve RNA-related and RNA-mediated mechanisms. All the existing data supports the hypothesis that the CB coordinates the highly complex haploid transcriptome during the preparation of the male gametes for fertilization. Thereby, this study provides a fundamental basis for the future functional analyses of ribonucleoprotein granules and offers also important insights into the mechanisms governing male fertility.

Keywords: RNA regulation, chromatoid body, haploid male germ cells, spermatogenesisis, ribonucleoprotein, MVH, piRNA

TIIVISTELMÄ

OLIVER MEIKAR
Kromatoidikappaleen välittämä RNA-säätely hiiren siittiönkehityksen aikana

Fysiologian oppiaine ja Turun biolääketieteellinen tohtoriohjelma (TuBS), biolääketieteen laitos, lääketieteellinen tiedekunta, Turun yliopisto, Kiinamyllynkatu 10, 20520 Turku

Miesten sukusolujen kehittyminen, spermatogeneesi, on tärkeä ja monella tavoin poikkeuksellinen kehitysprosessi, joka tuottaa suunnattomia määriä geneettisesti yksilöllisiä siittiöitä. Spermatogeneesi on monimutkainen, tarkasti säädelty tapahtumasarja, mikä myös aiheuttaa teknisiä haasteita spermatogeneesin molekyylimekanismien tutkimisessa. Näin ollen prosessin yksityiskohdat ovat vielä pitkälti tuntemattomia. Eräs siittiön haploidien esiasteiden (pyöreät spermatidit) erityisominaisuus on niiden ainutlaatuisen runsas transkriptionaalinen aktiivisuus. Genomin aktiivinen ilmentyminen puolestaan edellyttää tehokkaita ja täsmällisiä RNA:n säätelymekanismeja. Pyöreiden spermatidien solulimassa sijaitsee epätavallisen suuri RNA:ta ja proteiineja sisältävä rakenne, kromatoidikappale (chromatoid body, CB), joka ilmaantuu juuri voimakkaimman transkriptioaallon aikana ja osallistuu RNA-säätelyyn.

Tutkimuksen tavoitteena oli selvittää CB:n toimintaa siittiönkehityksen aikana. Tärkeänä osana tutkimusta kehitimme menetelmän, jonka avulla CB:t voidaan eristää hiiren kiveksestä. Menetelmä on yksinkertainen, nopea ja tehokas, ja sen avulla saadaan eristettyä toistettavasti rakenteeltaan ehjiä CB:ta, joiden puhtaus on hyvä ja määrä riittävä molekyylitason analyysiin. Jatkotutkimukset paljastivat, että suurin osa CB:n sisältämistä proteiineista on erilaisilla RNA-säätelyreiteillä toimivia RNA:ta sitovia proteiineja. Useat CB:n proteiineista pysyvät rakenteessa stabiilisti, mutta näytimme myös, että RNA:ta sitova proteiini SAM68 vierailee CB:ssa vain hyvin hetkellisesti tarkasti määritellyssä kehitysvaiheessa. Poistogeenisen hiirimallin avulla saimme selville, ettei SAM68 proteiinia kuitenkaan tarvita CB:n muodostumiseen. Tulostemme mukaan CB:een kulkeutuu suuri määrä RNA:ta kaikissa pyöreiden spermatidien kehitysvaiheissa. RNA-sekvenointi osoitti, että pienet piRNA (PIWI-interacting RNA) molekyylit ovat rikastuneet CB:ssa. Lisäksi CB sisältää suuren joukon erilaisia lähetti-RNA:ita ja aivan tuntemattomia intergeenisiä ei-koodaavia RNA:ita. Analyysimme mukaan hallitsevin CB:n molekyylireiteistä on piRNA-välitteinen RNA-säätelyreitti, mutta myös lähetti-RNA:n prosessointiin liittyvät tekijät ovat vahvasti edustettuina. CB ei selvästikään ole vain passiivinen RNA:n varastointipaikka, vaan näytimme sen kykenevän ATP:stä riippuvaiseen RNA prosessointiin *in vitro*

Kaikki tutkimuksen tulokset osoittavat, että CB on keskeinen RNA:n säätelykeskus, joka koordinoi sukusolujen erittäin monimuotoista transkriptomia. Tällä toiminnallaan CB osallistuu tärkeänä tekijänä miesten hedelmällisyyden ja sukusolujen geneettisen ja epigeneettisen informaation säätelyyn.

Avainsanat: RNA-säätely, kromatoidikappale, haploidi miesten sukusolu, spermatogeneesi, MVH, piRNA

TABLE OF CONTENTS

ABBREVIATIONS

ADP	adenosine diphosphate
AGO	argonaute protein
ARE	AU-rich element
ATP	adenosine triphosphate
CB	chromatoid body
CBP	cap-binding complex
DEAD	aspartate-glutamate-alanine-aspartate
DIG	digoxigenin
DNA	deoxyribonucleic acid
EDTA	ethylenediaminetetraacetic acid
EJC	exon-junction complex
EM	electron microscopy
emPAI	exponentially modified protein abundance index
EU	5-ethynyl uridine
fig	figure
FPKM	Fragments per kilobase of transcript per million mapped reads
FSH	follicle-stimulating hormon
GnRH	gonadotropin-releasing hormone
GO-term	gene ontology term
GTP	guanosine triphosphate
hnRNP	heterogeneous ribonucleoprotein particle
IP	immunoprecipitation
KDa	kilodaltons
LH	luteinizing hormone
LNA	locked nucleic acid
lncRNA	long non-coding RNA
mM	millimolar
mRNA	messenger RNA (protein-coding RNA)
mRNP	messenger ribonucleoprotein particles
MS	mass scpectrometry
ncRNA	non-coding RNA
ncRNA	non-coding RNA
NMD	nonsense-mediated (mRNA) decay
nt	nucleotide
PABPC	cytoplasmic poly(A)-binding protein
PABPN	nuclear poly(A)-binding protein
PCR	polymerase chain reaction
PFA	paraformaldehyde
piRNA	PIWI-interacting RNA
PTC	premature termination codon
qRT-PCR	quantitative RT-PCR
RNA	ribonucleic acid
RNAi	RNA interference
RNase	ribonuclease
RNP	ribonucleoprotein
RT-PCR	reverse-transcription PCR
siRNA	small interfering RNA
snRNP	small nuclear ribonucleoprotein
SSC	saline-sodium citrate
suppl	supplemental
TE	transposable element
TEI	transgenerational epigenetic inheritance
TREX	transcription-export complex
UTR	untranslated region
UV	ultra-violet
µl	microliter
µm	micrometer
µM	micromolar

LIST OF ORIGINAL PUBLICATIONS

The study is based on the following publications and manuscripts, which are referred to in the text by Roman numerals (I-IV). Unpublished data are also included. The original publications have been reproduced with permission of the copyright holders.

I Meikar O, Da Ros M, Liljenbäck H, Toppari J, Kotaja N. Accumulation of piRNAs in the chromatoid bodies purified by a novel isolation protocol. Experimental Cell Research 316(9):1567-75, 2010

II Meikar O, Kotaja N. Isolation of chromatoid bodies from mouse testis as a rich source of short RNAs. Methods in Molecular Biology. In press.

III Oliver Meikar, Vasily V. Vagin, Frédéric Chalmel, Karin Sõstar, Aurélie Lardenois, Molly Hammell, Ying Jin, Matteo Da Ros, Kaja A. Wasik, Jorma Toppari, Gregory J. Hannon, Noora Kotaja. An Atlas of Chromatoid Body Components. Submitted.

IV Messina V, Meikar O, Paronetto MP, Calabretta S, Geremia R, Kotaja N, Sette C. The RNA binding protein SAM68 transiently localizes in the chromatoid body of male germ cells and influences expression of select microRNAs. PLoS One 7(6):e39729, 2012

1. INTRODUCTION

Humans consist of approximately 10^{14} cells. These cells construct tissues that construct organs, e.g. heart, skin, eye and testis. Organs, in turn, are part of organisms to serve some specific functions. Organisms construct species. The general evolutionary aim of species is survival and improvement. Evolutionary immortality of species is engendered through the procreation of its members carrying their modified genetic individuality on to the offspring. Success in the evolutionary race is achieved by constant adaption in a changing environment. These adaptions are of genetic origin. The cells in an organism that are able to undergo genetic changes and give rise to new organisms are the gametes. Besides the task to drive evolution, the quality control in these germ cells needs to be of highest order since an error at any stage in the development of a gamete can lead to a disastrous outcome in the offspring. This is particularly true for species with limited number of offspring and slow generation times. Therefore, it is difficult to underestimate the importance of different mechanisms in mammalian germ cells that regulate and control the gamete development.

In male mammals, germ cell differentiation, or spermatogenesis occurs in the seminiferous tubules inside the testis. Male germline stem cells form the foundation of spermatogenesis through continuous production of daughter cells that can be assigned to differentiation. The program of spermatogenesis can be divided into three developmental phases. First, spermatogonia proliferate through a series of mitotic divisions. Subsequently they enter the meiotic program as spermatocytes which undergo meiosis to effect the production of haploid cells and to shuffle the genetic material. In the last phase, called spermiogenesis, the haploid spermatids with unique genomes undergo major structural transformation in order to differentiate into mature spermatozoa- small, resistant and motile vehicles of genetic material.

Spermiogenesis includes notable mechanisms that are specific to male gamete differentiation. These unique mechanisms are controlled by strictly regulated, phase-specific expression patterns for both mRNAs and their isoforms as well as for a considerable amount of different non-coding RNAs. Indeed, haploid round spermatids are a cell type with the highest transcriptome complexity in the whole organism. One of the unique visual characteristics of haploid round spermatids is a cytoplasmic large granule, called the chromatoid body (CB). Available data irrefutably demonstrate that the CB has a vital role in spermiogenesis even though information on its possible functions has remained rather limited.

To study a phenomenon in greater detail, it requires isolation and enrichment. To date, all attempts to enrich CBs to an acceptable degree of both purity and integrity have been unsuccessful. Thereby the establishment of a reproducible protocol for CB purification from male germ cells, followed by in depth analyses of isolated CBs, has been the general aim of this current study.

2. REVIEW OF THE LITERATURE

2.1. Regulation of gene expression

A simplified mechanism of eukaryotic gene expression is that a gene in the DNA is transcribed into its RNA copy, transported from the nucleus to the cytoplasm and then translated into a protein. While genes are considered as passive storage components of cells, proteins have a vast array of functions and participate in virtually every process within the cell.

When the Human Genome Project began, it was estimated that humans have more than 150,000 genes to warrant our huge complexity as compared to lower organisms. However, what transpired was that there are only approximately 20,000 genes- roughly the same amount as a roundworm *Caenorhabditis elegans*, yet which has only 10^3 cells, compared to humans with around $\sim 10^{14}$ cells and far more complexity (International Human Genome Sequencing Consortium 2004). How is this possible? Two explanations present both of which are related to the world of RNA. First, mRNAs of protein-coding genes are actively modified before translation, substantially increasing the number of actual proteins (the proteome) to in excess of perhaps a million members. Secondly, it has now transpired that in addition to proteins, non-coding RNAs transcribed from the genome are also involved in the execution of cellular functions and are the major contributors in evolution towards complexity (Mattick, Makunin 2006, Taft et al. 2010a, Soumillon et al. 2013, Kaessmann 2010, Ashe et al. 2012, Pang et al. 2006, Brennicke et al. 1999, Taft et al. 2007, Jablonka, Raz 2009).

Multicellular organisms consist of very different types of cells with specialized assignments. Despite possessing identical DNA, they can still differ dramatically from one another, both morphologically and functionally. This is achieved by differential cell type-specific gene expression patterns. In fact, cells have very complex and multilevel mechanisms to regulate their gene expression, whether to trigger developmental pathways, respond to environmental stimuli, or adapt to new sustenance sources. Virtually any step of gene expression can be modulated from the initiation of transcription to the post-translational modifications of a translated protein (Alberts et al. 2002) (Fig. 1).

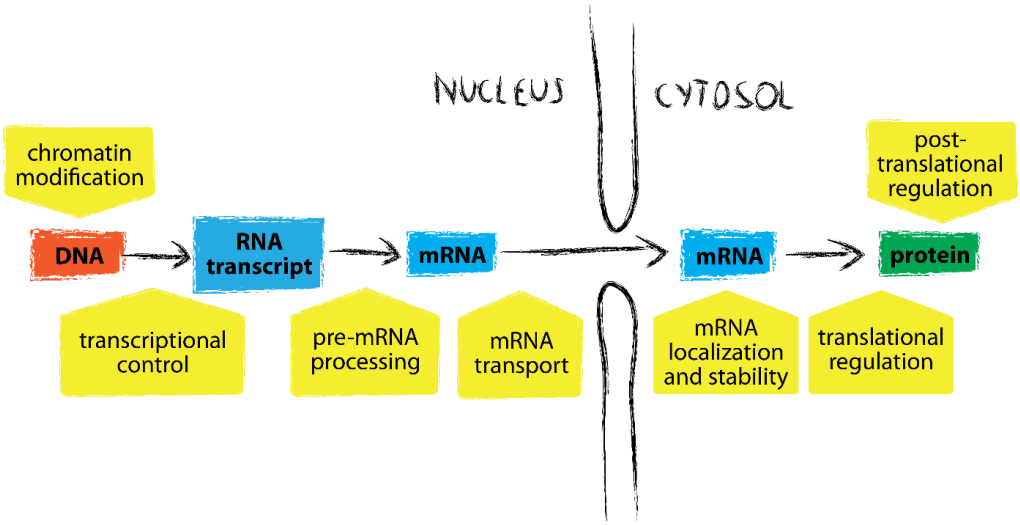

Figure 1. *Steps at which eukaryotic gene expression can be controlled. Various steps at the RNA level provide a flexible means for complex regulation of gene expression.*

2.1.1. Transcriptional and epigenetic regulation

Gene expression at DNA level controls the accessibility of DNA to the transcriptional machinery. This is the most upstream step of gene regulation. Large areas in the genome- or as in case of the Barr body in the somatic cells of females, even the whole chromosome- can be silenced by chromatin compaction (Lyon 1961, Boumil, Lee 2001). DNA methylation is a common biochemical method for gene silencing and is catalyzed by DNA methyltrasferases, which are, in turn, regulated by DNA-binding proteins and non-coding RNAs. DNA methylation results in a reversible but stable silencing of genes (Smith, Meissner 2013). Typically, DNA is methylated on the fifth position of cytosine nucleotides in CpG sequences (in mammals). DNA methylation has a regulatory role in the genome and unsurprisingly, frequent CpG sites (or CpG islands) are concentrated on gene promoter areas (Deaton, Bird 2011). Methylated DNA promotes the de-acetylation of adjacent histones and signals for DNA to be packed more densely, reducing gene expression in that area. There is a large number of different and highly dynamic post-translational histone modifications that affect the expression state of DNA. These include histone acetylation, methylation phosphorylation, ubiquitination and sumoylation. Different combinations of histone modifications form the basis of the histone code which can regulate gene expression, both positively and negatively (Bannister, Kouzarides 2011).

The initiation of transcription is the most important and extensively used regulatory step of gene expression. In the eukaryotic nucleus the responsibility of RNA synthesis is divided between three RNA polymerases (RNAPs). RNAPI and RNAPIII transcribe genes that encode ribosomal RNA, transfer RNA and various small RNAs, while RNAPII transcribes mRNAs. RNAPII machinery includes general transcription factors that help to position the RNA polymerase to the promoter area of the gene. Thereafter other factors bind to form a complete transcription initiation complex (Alberts et al. 2002, Shandilya, Roberts 2012). The

onset of transcription is controlled by DNA and histone modifications, transcription factors, coativators, corepressors, small non-coding RNAs and other *cis* and *trans* elements that regulate the timing and effectiveness of DNA directed RNA synthesis (Alberts et al. 2002, Deaton, Bird 2011, Shandilya, Roberts 2012, Taft et al. 2010b).

2.1.2. Post-transcriptional regulation

Post-transcriptional gene control, with a significant contribution from alternative splicing, significantly accounts for the evolution towards cellular, functional and biological complexity. The RNA transcript of the gene undergoes three main modifications in the nucleus - 5' capping, 3'polyadenylation and splicing. mRNA processing events are integrated and coordinated in space and time (Hocine et al. 2010). The post-transcriptional regulation of mRNAs prior to translation can occur by many means, for example by modulating mRNA's alternative splicing, editing, stability or spatial and temporal expression (Darnell 2013, Alberts et al. 2002). Since this thesis is most related to the post-transcriptional aspect of gene regulation, it is discussed in greater detail here.

2.1.2.1. Pre-mRNA processing

RNA 5'capping is the first modification of eukaryotic pre-mRNAs that already occurs during transcription. Shielding the 5' phosphor group by a 7-methylguanosine "cap" via 5'5' triphosphate link, provides the mRNA with significant resistance to 5'-3' exonucleases plus also distinguishes it from other types of RNA molecules. The 3' end of the pre-mRNA transcript is elongated with a poly(A) tail, based on the polyadenylation and cleavage signal sequences at the end of the gene (Dreyfuss et al. 1993, Han et al. 2010, Gorgoni, Gray 2004, Hocine et al. 2010).

The vast majority of eukaryotic genes is split between coding segments (exons) and intragenic sequences (introns). These split genes are firstly transcribed into long primary transcripts called precursor mRNAs (pre-mRNAs), but during their maturation, introns are removed and exons are combined by two distinct transesterification reactions. This process is called splicing and occurs in the nucleus, mediated by a large macromolecular machine called spliceosome. Each spliceosome is composed of approximately 125 proteins in conjunction with five small nuclear RNAs (snRNA) to form RNA-protein complexes called snRNPs (Lerner et al. 1980, Hocine et al. 2010). Recent evidence supports the concept that pre-mRNA splicing is catalyzed at least in part by RNA (Will, Luhrmann 2011).

SR-proteins (Serine-Arginine rich proteins) mark the splicing status of the RNA transcript by their phosphorylation status. These are splicing activators which in their phosphorylated form, act like anchors to assemble spliceosome and become dephosphorylated upon splicing (Long, Caceres 2009). The heterogenous ribnucleoproteins (hnRNPs) also bind to pre-mRNAs and help to decipher the cellular splicing. SR-proteins and hnRNPs are the main regulators of both constitutive and alternative splicing. During export from the nucleus, SR proteins are released from the mRNA, while some hnRNPs will accompany their mRNAs in cytoplasm (Dreyfuss et al. 1993, Han et al. 2010).

The total number of proteins specified by the genome is substantially greater than the number of genes, which indicates that each gene should encode more than just one protein. Alternative splicing is a mechanism to introduce/hide alternative splice sites in the pre-mRNA which allows different versions of mRNAs to be produced from an individual gene. Of note, it has been estimated that transcripts from around 95% of multiexon genes undergo alternative splicing (Wang et al. 2008, Pan et al. 2008). Typical forms of mammalian alternative splicing are exon skipping and alternative 5' and 3' splice site usage. Alternative splicing is regulated in a tissue-specific manner by positive and negative *trans*-acting factors, in particular the SR and hnRNP family proteins which can promote or mask specific splice sites. Alternative splicing is particularly prevalent in testis (Elliott, Grellscheid 2006).

2.1.2.2. RNA editing

RNA editing is a process, which alters the nucleotide sequences of RNA transcripts after they have been synthesized thereby altering the translational information they carry. RNA editing processes appear to be evolutionarily recent acquisitions that arose independently. There are two general types of RNA editing- by nucleobase modifications (such as cytidine to uridine and adenosine to inosine deaminations) and by insertion/deletion of a nucleotide (Brennicke et al. 1999). This latter is specific for trypanosome mitochondria, being one of a prevalent mechanisms of RNA post-transcriptional regulation there.

A-to-I editing is conserved from humans to sea anemones (Tang et al. 2012). Inosine preferentially base pairs with cytidine and is therefore functionally equivalent to guanosine thereby changing the coding information of mRNA. This conversion is carried out by ADAR enzymes (Adenine Deaminases which act on double-stranded RNA) by hydrolytic deamination. Alu elements are the main targets of A-to-I editing in humans, but such editing is, for example, also crucial for the modulation of glutamate receptor subunit mRNAs (Rueter et al. 1995). Further, A-to-I conversion can potentially cause alterations in splice sites and affect the nuclear retention of transcripts (DeCerbo, Carmichael 2005, Tang et al. 2012). By computational analysis, RNA editing has been conjectured to also be involved in the generation of antibody diversity (Tang et al. 2012, Steele et al. 2006).

Evidence suggests that A-to-I RNA editing antagonizes RNAi-mediated gene silencing efficacy, as the A-to-I edited pri-miRNAs can inhibit Drosha and Dicer cleavages (Nishikura 2010). Controversially, it was recently shown that ADAR1 forms a complex with Dicer to promote microRNA processing and RNA-induced gene silencing (Ota et al. 2013). Interestingly, ADAR1 has been shown to also directly interact with hUPF1 in the cell nucleus, linking the RNA editing with mRNA surveillance pathway too, which demonstrates again an integration of apparently independent pathways of mRNA post-transcriptional modification (Agranat et al. 2008).

TENR/ADAD1 is a testis-restricted nuclear protein present in mid-pachytene spermatocytes until mid-round spermatids and is related to a family of adenosine deaminases. The sequence comparison of TENR with ADARs reveals similar RNA binding and catalytic domains to suggest that TENR may function in the testis as an RNA-editing enzyme; however, a mutation in its catalytic center suggests that it may prevent, rather than promote RNA

editing in spermatogenic cells (Connolly et al. 2005). The disruption of murine *Tenr* results in abnormal sperm morphology and male infertility. Still, the severity of the phenotype is dependent on the genetic background, which indicates a more complex interaction mechanism (Connolly et al. 2005).

2.1.2.3. RNA transport to the cytoplasm and localization

Once mRNA has been transcribed and properly modified, it is ready to be transported to the cytoplasm for translation. This is another vital step in the gene-to-protein pathway which can be regulated. All transport between the nucleus and cytoplasm occurs through nuclear pore complexes. In order for the mRNA to pass through nuclear pores, it has to carry respective adaptor and export factors. The transcription-export complex (TREX) is recruited to the mRNA 5' end in a splicing and cap-dependent manner to render an mRNP export competent. However, if a transcript is not properly processed, it can be recognized by nuclear surveillance machinery, retained in the nucleus and degraded (Hocine et al. 2010). Most mRNA transcripts move through the nucleus by random, but channeled, nuclear diffusion (Mor et al. 2010).

DBP5 is an ATP-dependent RNA DEAD-box helicase which has a central role in mRNA export from the nucleus, mediating the exchange of nuclear mRNA binding proteins by cytoplasmic proteins (Cole, Scarcelli 2006). Non-coding RNAs are also exported from the nucleus as cargos of export proteins, but they are packed into different RNP complexes. A central instigator in the nucleocytoplasmic transport of non-coding RNAs is RAN, a 25 KDa GTPase. In the nucleus, RAN-GTP attaches and modulates the RNA export receptor proteins, karyopherins, which bind their RNA targets. In the cytoplasm, RAN hydrolyses GTP to GDP, undergoes a conformational change and releases its cargo (Katahira, Yoneda 2011).

Once exported from the nucleus, RNA transport coupled with translation is a crucial mechanism to target mRNAs to discrete subcellular locations of a cell or an organism, where their protein products are expressed locally (Martin, Ephrussi 2009, Martin, Zukin 2006, Lecuyer et al. 2007). One of the advantages of this mechanism is its economy, as each localized mRNA can be translated multiple times when compared to the energy expended to individually transport each protein molecule to its target location. Further, localized translation can protect the cell from possibly toxic effects that some proteins might elicit in some cellular compartments (Martin, Ephrussi 2009). In addition, spatially restricted translation allows its quick regulation by local stimuli on-site, instead from an extended command-line through the nucleus. For example, the autonomous translation of synaptic mRNAs in response to local stimuli greatly contributes to a neuron's synaptic plasticity and enhances the computational capacity of the brain (Martin, Zukin 2006).

The correct localization of specific mRNAs in asymmetric cells such as fibroblasts, oligodendrocytes and neurons is vital for their function. A high-resolution fluorescent *in situ* hybridization analysis of early *Drosophila* embryogenesis has revealed that special expression patterns of mRNAs is the major mechanism to organize cellular architecture and function (Lecuyer et al. 2007, Martin, Ephrussi 2009). For spatially and temporally modulated gene expression, mRNAs are assembled into ribonucleoprotein (RNP) granules in order to

transport, protect and inhibit them. The importance of compartmentalized RNA regulation is considered further below.

2.1.2.4. General RNA degradation pathways

RNA as such is unstable in a cell due to a vast number of different ribonucleases (RNases). mRNA turnover plays a key role in the control of gene expression as a response to environmental signals or, for example, the developmental program of the cell. Decay rates of different mRNA molecules can vary greatly. For example, in mammalian cells the mRNA half-lives can vary from minutes to days (Meyer et al. 2004) and the estimated median mRNA half-life in human cells is conjectured to be 10 hours (Yang et al. 2003). There is a variety of RNA degrading enzymes with different specificities, suggesting that distinct subpopulations of transcripts within the cell are specially regulated. The stability of mRNAs is also determined by the *cis* elements in their sequences, which operate as recognition sites for regulatory proteins which can protect the RNA, or stimulate decay (Parker, Song 2004). RNases are also important for the quality control of intact mRNAs and in antiviral defenses (Decker, Parker 2012, Meyer et al. 2004, Alberts et al. 2002).

Degradation of eukaryotic mRNAs is generally initiated with the shortening of their poly(A) tails by the major mRNA deadenylase, the CCR4/POP2/NOT complex (Chen et al. 2002). 3' deadenylation in turn promotes mRNA degradation in a 3'-5' direction by a complex of diverse exonucleases in the cytoplasmic exosome. More often, the shortening of poly(A) tail leads to the removal of 5' cap of mRNA by the DCP1/DCP2 decapping enzyme which then exposes the mRNA transcript to digestion by a 5'-3' exonuclease, XRN1 (Decker, Parker 2012, Muhlrad et al. 1994). MicroRNAs and RNA interference (RNAi) in general make up the whole new dimension of (usually) negative regulation of mRNA stability and discussed further separately below.

2.1.2.5. Nonsense-mediated decay

In the nucleus, during the splicing of pre-mRNA, a set of proteins called the exon junction complex (EJC) is deposited onto the mRNA upstream of each splice site (Le Hir et al. 2000). The core of the EJC can be reconstituted *in vitro* and consist of ATP-binding EIF4A3, CASC3/MLN51/BTZ and MAGOH, RBM8A/Y14 (Ballut et al. 2005, Bono et al. 2006). EJCs remain bound to the mature mRNA until the first round of translation in the cytoplasm when the moving ribosome strips them off. EJCs provide a position-specific memory of the splice sites which are used to ensure that the mRNA transcript does not contain premature termination codons (PTCs). If a stop codon is located before any splice site, the following EJC is not removed from mRNA since the ribosome will have been released before it is reached. This mRNA is then considered to have a truncated open reading frame and is degraded via nonsense-mediated decay (NMD) (Schweingruber et al. 2013, Isken, Maquat 2007).

The NMD in conserved in metazoans and is responsible for the cytoplasmic degradation of aberrant transcripts. The core proteins of NMD are UPF1, UPF2 and UPF3. UPF2 and UPF3 bind to EJC and if any EJC is not removed from the mRNA after the first round of translation

because of a premature stop codon, they come into contact with UPF1 and trigger its phosphorylation by a NMD kinase SMG1. Phosphorylated UPF1 induces mRNA decay by interaction with other NMD factors. The degradation of NMD substrate mRNAs can be initiated by two different decay routes. SMG6 is a functional endonuclease, which also contains EJC binding motifs and results in mRNA cleavage close to a premature termination codon. The resulting 5' and 3' RNA fragments are then rapidly degraded by general exonucleolytic rotes involving XRN1 and the exosome. The SMG5 and SMG7 are not nucleases, but form a stable heterodimer and mediate the deadenylation and decapping followed by general exonucleolytic degradation of the mRNA (Schweingruber et al. 2013, Chang et al. 2007).

The functions of NMD extend beyond the surveillance of mRNAs with premature termination codons, as many non-PTC containing mRNAs are also targeted by NMD. Further, other elements, for example some long 3'UTRs and alternative polyadenylation sites, can trigger NMD (Schweingruber et al. 2013). Moreover, NMD controls alternative splicing and the ablation of NMD results in the upregulation of splice isoforms of 30% of all expressed genes in mouse tissues (Weischenfeldt et al. 2012). All this emphasizes the versatile roles of NMD machinery in the control of gene expression.

Some individual key elements of NMD machinery are shown to possess other vital functions in the cell. UPF1, a 5' to 3' DNA and RNA helicase, physically interacts with replicative DNA polymerase and is essential during DNA replication and repair. Further, the NMD endonuclease SMG6 binds active telomerase ribonucleoprotein complex and is vital to maintain the normal structure of (human) chromosome ends. Furthermore, the NMD kinase SMG1 can phosphorylate p53 upon genotoxic stress and its ablation causes the accumulation of spontaneous DNA damage and increases cellular sensitivity to ionizing radiation (Azzalin, Lingner 2006).

2.1.3 Translational regulation

The mechanisms of translational control of gene expression are, for example, to control the half-life of mRNA, the storage of mRNA or the selective repression of mRNA translation. In eukaryotes, a ribosome recruitment complex EIF4F is formed to the 5' end of an mRNA molecule to be translated. This consists of a cap-binding helicase 4E (EIF4E), the scaffolding protein EIF4G in association with the mRNA unwinding protein, a DEAD-box helicase EIF4A1. In turn, EIF4G binds a 3' poly(A)-binding protein PABP to type an RNA-protein-RNA circular structure. This kind of mRNA pseudo-circular conformation is considered to support efficient translation and ensure that only intact mRNAs are translated (Szostak, Gebauer 2013). Several stimuli, including growth factors, cytokines, and nutrient availability, regulate the formation of EIF4F complex. The general direct regulators are the translational inhibitory EIF4E-binding proteins (4E-BPs) which prevent the interaction between EIF4E and EIF4G; EIF4B, as an activator of the EIF4A1; PDCD4, as an inhibitor of EIF4A1. There are also numerous *trans*-acting regulators of translation for specific mRNAs (Sonenberg, Hinnebusch 2009).

2.2 General properties of RNA

RNA is a remarkable molecule with various properties and crucial functions in nature. Chemically, DNA and RNA are quite similar, with only two minor chemical substitutions- RNA has one extra oxygen atom in its sugar moiety (ribose vs. deoxyribose) and DNA has a methyl group replacement in one of its nucleobases (thymine vs. uracil) (Fig. 2).

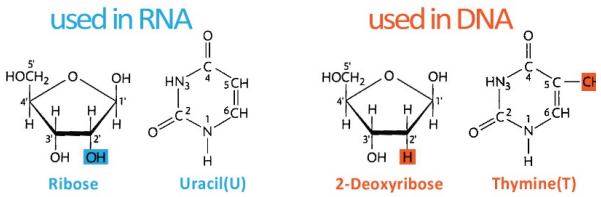

Figure 2. *Differences in the compositions of sugars and bases in RNA and DNA. Otherwise, these molecules are chemically identical. Modified from (Elliott, Ladomery 2011).*

DNA has evolved to use thymine as a base, instead of uracil, as a means to protect the integrity of its genetic information. The reason is that cytosine is an unstable base which can change into uracil by spontaneous deamination. This C-to-U conversion changes the base-pairing properties of the nucleotide and can result in a point mutation during the replication of DNA. As the DNA's analogue of uracil is thymine, its repair machinery can recognize the deaminated cytosines (or uracils) and replace them. Although the deaminated cytosines remain undetected in RNA, this does not pose a problem, since RNA is generally not used for long-term storage of genetic information (Elliott, Ladomery 2011).

While DNA has the single function to store genetic information, the tasks of RNA are diverse also including the recognition, scaffolding, shape-related and catalytic functions (Fig. 3). RNA is believed to be the primordial molecule- the precursor to all current life. This hypothesis of a self-replicating, 3,5 billion year old RNA, is supported by the fact that RNA possesses the properties of both, DNA and protein enzymes- it can store genetic information and catalyze chemical reactions. Moreover, many of the most critical components of cells are composed mostly or entirely of RNA and also many critical cofactors are either nucleotides or substances clearly related to them (ATP, Ac-CoA, NADH) (Alberts et al. 2002).

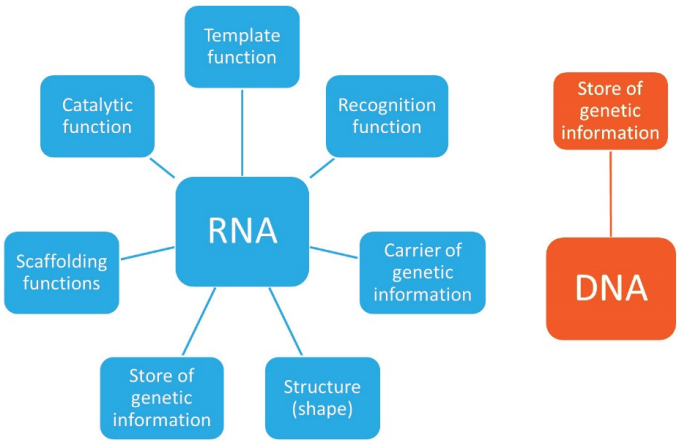

Figure 3. *Functions of RNA and DNA in the cell. Modified from (Elliott, Ladomery 2011).*

The RNA-characteristic 2'-OH group on the ribose sugar of its nucleotides is responsible for changing the base-pair stacking properties of RNAs, by affecting the interphosphate distance and giving rise to different and more compact helical structures. Unlike DNA, ssRNA is synthesized without its complementary partner, so it folds dynamically by itself via stretches of intra- and intermolecular interactions into helixes, loops, pseudoknots, bulges and junctions, which can, in turn, form very compact and highly organized tertiary structures. These shapes can very precisely and specifically recognize and bind to other molecules, as being functionally similar to antibodies (Elliott, Ladomery 2011, Jenison et al. 1994).

RNA is unstable due to the multitude of RNases that are widely expressed. An RNA molecule that can be a carrier of genetic information is, by default, considered dangerous and subjected to degradation. Further, a large part of gene expression control is built on the negative regulation of the half-life of mRNAs. Besides the enzymatically destabilizing environment, RNA is also inherently chemically unstable due to the 2' -OH group in the sugar moiety of its nucleotides. The 2' -OH group can launch a spontaneous nucleophilic attack on its own adjacent phosphodiester bond, which results in cleavage of the RNA chain, leaving the upstream and downstream nucleotides with a 2'3'-cyclic phosphate and 5' -OH, respectively (Fig. 4). Basic solution, metal ions and adjacent nucleotides can greatly enhance this reaction. When accordingly coordinated, the chemically reactive 2' –OH group can become the core of the RNA's catalytic center and launch similar attacks to cleave and to form chemical bonds of its targets. These catalytic RNAs are called ribozymes as based on their functional similarity to enzymes (Zappulla, Cech 2006, Elliott, Ladomery 2011).

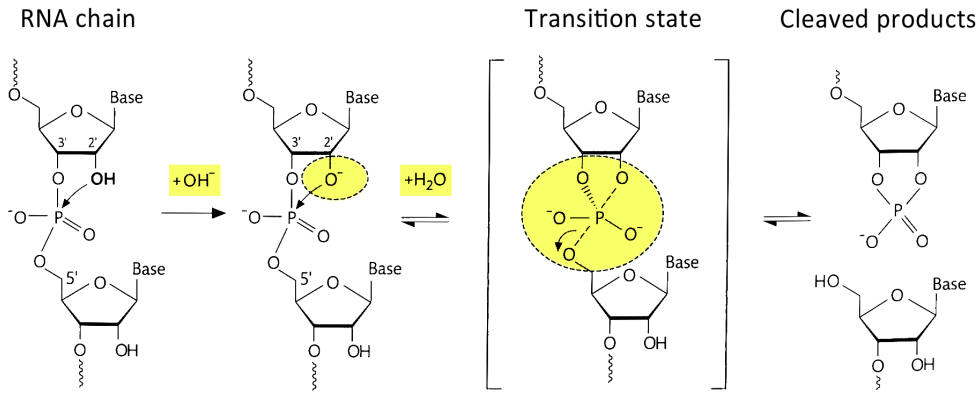

Figure 4. *Spontaneous RNA cleavage, catalyzed by basic solution. The -OH group activates the 2' -OH group of the ribose sugar which becomes the catalytic center. Modified from (Elliott, Ladomery 2011).*

When combining the structural and catalytic potential of RNA molecules, it is not surprising that non-coding RNAs can fold into complex structures capable of unique functions and enzymatic reactions. Some important naturally occurring ribozymes are self-splicing introns, RNAse P and the ribosome. Examples of structurally important but not catalytic RNAs, are the scaffolding RNAs, which bind together larger complexes, riboswatches which change their shape upon binding to their targets and the simple short linear RNA molecules which guide proteins to their complementary DNA or RNA targets.

2.3. The world of non-coding RNAs

In contrast to the relatively modest changes in the protein-coding genes through evolution, the quantity of non-protein-coding DNA has increased dramatically and accounts for >98% of the human genome sequence. It is argued that around 70-90% of the genome is transcribed into RNA. Moreover, this non-protein-coding "junk DNA" is shown to be differentially expressed. Accumulating data links non-coding (nc) RNAs with nearly all fundamental biological pathways which control various levels of gene expression in physiology and development, including chromatin architecture/epigenetic memory, transcription, RNA splicing, editing, translation and turnover (Mattick, Makunin 2006, Taft et al. 2007). NcRNAs are typically defined as RNA transcripts that lack protein-coding capacity and can be roughly divided into small (<200 nt) and long ncRNAs (>200 nt). Further characterization could divide them to functional and regulatory ncRNAs.

2.3.1. Small non-coding RNAs

2.3.1.1. RNA interference

RNA interference (RNAi), or RNA mediated gene silencing, is one of the most important and evolutionarily conserved strategies in higher metazoa for transcriptional and post-transcriptional gene expression control. Generally, small non-coding regulatory RNAs act as negative modifiers of gene expression by routing their target mRNAs to degradation or translational repression (Lee et al. 2004, Krol et al. 2010, Ghildiyal, Zamore 2009). The importance of RNA-mediated gene silencing was recognized in 2006 with the award to Andrew Z. Fire and Craig C. Mello of the Nobel Prize in Physiology or Medicine for their discovery of RNA interference. The amount and complexity of different small RNAs and the versatility of their functions has been increasingly recognized. The best known small regulatory RNAs are microRNAs (miRNAs), small interfering RNAs (siRNAs) and PIWI-interacting RNAs (piRNAs) (Fig. 5).

The miRNA pathway is a well-understood, powerful post-transcriptional silencing mechanism which regulates the expression of virtually all protein-coding genes. miRNA precursors are endogenous long hairpin loop transcripts, which are successively processed in the nucleus and in the cytoplasm by endonucleases into short (usually 21 nt) double-stranded mature miRNAs. The guide strand of the miRNA duplex associates with the RNA-induced silencing complex (RISC) and guides it to the target mRNAs causing their translational inhibition or decay (Fig. 5) (Lee et al. 2004, Krol et al. 2010). Mature siRNAs are similar to mature miRNAs and also use the same generic RISC machinery to silence their targets. Unlike miRNAs, siRNAs are fully complementary with their targets which results in the target RNA degradation rather than translational inhibition. siRNAs are considered to be cells' defense mechanism against viruses and other foreign elements but can also participate in gene silencing (Lau 2010, Song et al. 2011). Other functions, including the epigenetic chromatin modification in the nucleus, have also been suggested (Weinberg et al. 2006).

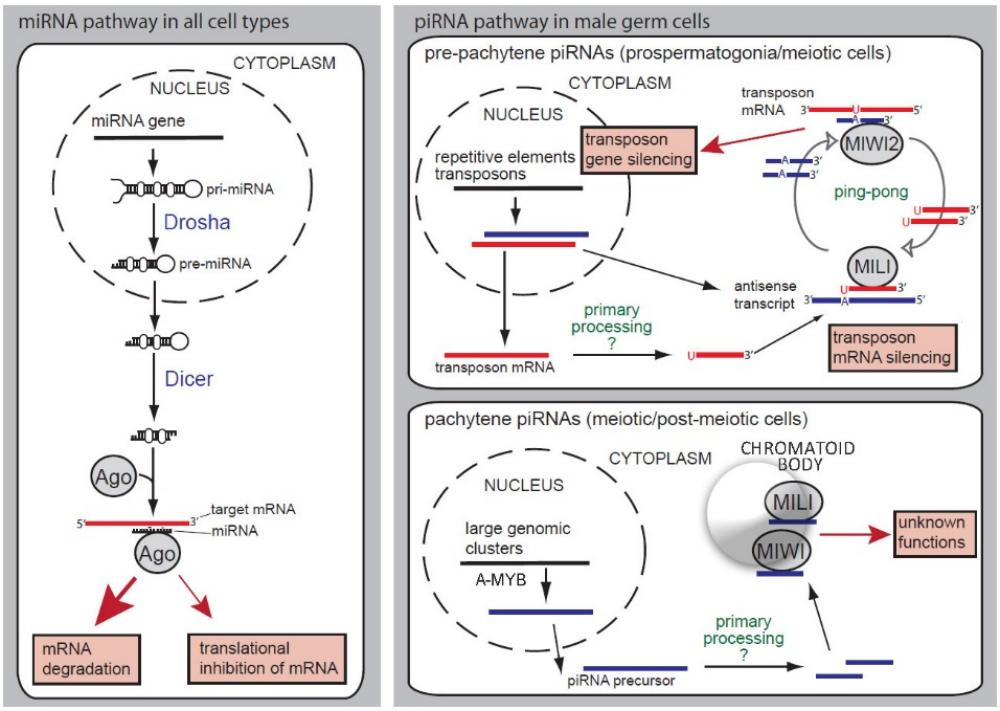

Figure 5. *Biosynthesis and function of miRNAs and piRNAs. miRNA precursors are long imperfect hairpin-loops which are successively cut by DROSHA and DICER endonucleases to produce mature miRNA duplexes which contain guide and passenger strands. The guide strand of miRNA is bound by AGO proteins (discussed in section 2.4.2.4) to form the RNA-induced silencing complex (RISC). RISC uses miRNA to recognize and control the stability and translation efficiency of target mRNAs. Pre-pachytene piRNAs bind to MILI and MIWI2 AGOs. Their primary processing mechanism is not known but once present, they are amplified by the ping-pong mechanism in which piRNAs from the opposite strands promote one another's synthesis. Pre-pachytene piRNAs are shown to silence transposons transcriptionally and post-transcriptionally. Pachytene piRNAs arise later in development and bind to MIWI and MILI AGOs. Their precursors are long primary transcripts, regulated by A-MYB translation factor. Modified from (Meikar et al. 2011).*

The roles of miRNAs, siRNAs and other Dicer-dependant mechanisms in spermatogenesis are demonstrated by various Dicer knockout mouse models (Yadav, Kotaja 2013). Dicer ablation at different stages of spermatogonial development showed that Dicer is dispensable for spermatogonial stem cell renewal and mitotic proliferation, but is subsequently required for the meiotic and haploid phases of spermatogenesis (Korhonen et al. 2011, Romero et al. 2011). Dicer-null mouse embryonic stem cells are viable but defective in differentiation and centromeric silencing (Kanellopoulou et al. 2005). In the assessment of the complexity of secondary effects when the miRNA/siRNA dimension is capped off from the gene regulation network, plus other possible functions of Dicer, it is very difficult to exactly locate all the mechanisms of the Dicer-related actions. It is evident though that Dicer and miRNAs are involved in multiple, fundamental biological processes in a mammalian organism.

Individual miRNAs usually have several target mRNAs and individual mRNAs usually have several different miRNA binding sites. This provides a complex and dynamic association network which gives rise to sophisticated control schemes of specific sets of mRNAs. Clearly, miRNAs would need to be regulated too, and in some situations, for example during the switch of cells' developmental program, it is essential to modulate miRNA expression. An interesting mechanism in the form of circular RNA to control the specific sets of miRNAs during cell development in the brain and testis has been recently described. So-called "miRNA sponges" contain many binding sites for specific miRNAs which provide a theoretical possibility to modulate the active miRNA repertoire in the cells. The circular formation renders these RNAs intrinsically resistant to exonucleases and is conjectured to be a general phenomenon of miRNA sponges (Hansen et al. 2013).

New types of small regulatory RNAs are constantly being discovered and their functions investigated. For example, tiRNAs are yet another new class of metazoan small nuclear RNAs. They are predominantly 18 nt in length, appear to originate from an uncommon biosynthesis pathway, map close to promoters of eukaryotic genes and probably participate in transcriptional regulation and splicing in metazoans (Taft et al. 2010b).

2.3.1.2. piRNAs

Contrary to miRNAs that are conserved and ubiquitous, PIWI-interacting small RNAs (piRNAs) are not conserved and are predominantly expressed in the male germline and are the most abundant type of small non-coding RNAs. piRNAs are mostly known through their role in silencing the transposable elements which become activated during DNA demethylation in the primordial germ cells (Siomi et al. 2011). piRNAs drive their functions in association with the three PIWI proteins (MILI, MIWI and MIWI2 in mice). Each of them bind a specific subset of piRNAs, have a different expression pattern and is demonstrated to be essential for spermatogenesis (Carmell et al. 2007, Kuramochi-Miyagawa et al. 2004, Deng, Lin 2002, Thomson, Lin 2009). PIWIs belong to evolutionally conserved Argonaute protein family from where the effector proteins of miRNA and siRNA driven RISC complexes also originate (Cenik, Zamore 2011).

As with other functional small RNAs, piRNAs have a 5' monophosphate. In addition, piRNAs share a bias for a 5' uridine and have a methylated 2'-oxygen at their 3' end. piRNAs constitute a highly abundant small RNA population of ~30 nt in length in mammalian testis, which comprise hundreds of thousands of unique sequences (Aravin et al. 2007). Primary biogenesis of piRNAs is different to miRNAs and siRNAs. Careful analysis of piRNAs that associate with different PIWI proteins has revealed at least two main distinguishable classes of piRNAs in mammalian testis, which are named by their expression timing as pre-pachytene and pachytene piRNAs. Although similar at the molecular level, these have different tasks and mechanisms of action, associate with different PIWI proteins and derive from different regions of the genome (Aravin et al. 2007) (Fig. 5).

Pre-pachytene piRNAs comprise a strikingly uniform, tiny subclass of piRNAs which originate from repeat sequences related to transposable elements and heterochromatic regions. These repeat-associated small interfering RNAs (also called rasiRNAs) participate in the

silencing of transposable elements both at epigenetic and post-transcriptional level during fetal and neonatal germ cell development (Aravin et al. 2007, Carmell et al. 2007). Deep sequencing and time-specific expression analysis has led to the proposal of a so-called "ping-pong" cycle model for pre-pachytene piRNA biogenesis. This suggests that pre-pachytene piRNAs are produced by a sense-antisense amplification loop between transposons and piRNA transcripts which originate from piRNA clusters that are antisense of transposon sequences. This ping-pong loop slices active transposons while amplifying respective piRNAs, and is mediated by MILI and MIWI2 (Aravin et al. 2008). The primary pre-pachytene piRNAs have a 5' bias to uracil which corresponds to the bias of adenine in the 10th position of their complementary targets. This footprint of 5'U and 10A in the primary and secondary piRNAs, respectively, is characteristic to the piRNA ping-pong cycle, which operates in different organisms as an adaptive immune system by optimizing the piRNA population accordingly to its target elements.

Pachytene piRNA population appears in meiotic spermatocytes and peaks in haploid round spermatids overlapping with the expression patterns of the respective PIWI proteins MILI and MIWI (Aravin et al. 2006, Girard et al. 2006, Grivna et al. 2006). Pachytene piRNAs represent a highly abundant and heterogeneous class of small RNAs and account for the easily detectable characteristic ~30 nt piRNA band in adult total testis RNA (Aravin et al. 2006, 203-207; Girard et al. 2006, 199-202; Grivna et al. 2006, 1709-1714). Compared to their pre-pachytene counterparts, the function of pachytene piRNAs remain largely unknown but is not associated with transposons or repeat sequences. Instead, pachytene piRNAs map into large sparse clusters across the genome, from tens to hundreds of kilobases and mostly derive from a single strand (Aravin et al. 2007, Siomi et al. 2011). The transcription factor A-MYB has been recently found to regulate the expression of pachytene piRNA precursors, which were reported to be long 5' capped and 3' polyadenylated transcripts (Li et al. 2013).

While there is very little conservation of individual piRNA sequences between different mammals, a surprisingly significant conservation of the genomic locations of mammalian piRNA clusters, has been observed (Betel et al. 2007). This conjectures that pachytene piRNAs may act as a population which represents specific regions of chromosomes.

It has been recently suggested that PIWI/piRNAs in the brain of Aplysia (sea slug) epigenetically regulate stable long-term changes in neurons for the persistence of memory (Rajasethupathy et al. 2012). Other recent findings in the *C. elegans* research demonstrate a multi- generational epigenetic inheritance mechanism induced by piRNAs and piRNA machinery. It has been demonstrated that piRNAs can recognize even single- copy foreign sequences, presumably by comparison of these sequences to an epigenetic memory of previously expressed sequences, and thereby initiate a remarkably stable epigenetic memory of their silencing (Shirayama et al. 2012, Ashe et al. 2012). It is fascinating to speculate that this mechanism may serve for considerably more than only defense. For example, it is possible that the piRNA-derived epigenetic memory of the history of gene expression patterns may also be epigenetically transferable to offspring as the parents' adaptive response to the environment. Indeed, mounting evidence supports the idea that RNAs in mammals and other animals can function in epigenetic reprogramming. As a reference, a recent report describes an epigenetic licensing of germline gene expression by maternal RNA in *C. elegans*, where the absence of maternal allele causes the silencing of a

paternally contributed allele in the zygotic germ cells. This would suggest that genes with no history of germline expression, as assessed by comparison against the pool of maternally deposited RNA, are targeted for silencing (Johnson, Spence 2011).

2.3.2. Long non-coding RNAs

Transcriptome analyses have demonstrated that the vast majority of the mammalian genome is transcribed in extremely complicated patterns of interlaced and overlapping transcripts, of which many bear processing signatures of mRNAs, including 5' capping, splicing and poly(A)denylation, but have little or no open reading frames (Carninci et al. 2005). The general anatomy of long non-coding RNA (lncRNA) loci is divergent with intergenic, intronic and bidirectional origin. Notably, most mammalian genes also have antisense non-coding transcripts which appear play a role in the regulation of their expression (Katayama et al. 2005, Taft et al. 2007).

It has been observed already in 1975 that biochemically purified chromatin contains twice as much RNA as DNA, thus raising the question if RNA may influence chromatin structure and gene regulation (Paul, Duerksen 1975). The association of lncRNAs with heterochromatin formation and imprinting was further emphasized when the dose compensation mechanism by X chromosome inactivation in female cells was solved. Xist is a 17 Kbases long (in humans) lncRNA that is expressed from one of the two X chromosomes. It promotes changes in the chromatin structure and consequently causes the epigenetic silencing of the entire chromosome. X chromosomes without Xist expression will not be inactivated, and interestingly, the expression of the Xist gene on another chromosome will cause the inactivation of that chromosome too (Penny et al. 1996, Boumil, Lee 2001). Xist and other long functional ncRNAs are poorly conserved, which demonstrates that the lack of conservation does not necessarily mean a lack of function (Nesterova et al. 2001, Pang et al. 2006).

There are several infrastructural long ncRNAs that have been known for some time and have well-established functions, for example ribosomal RNA and ribonuclease P. However, besides those and the increasing number of other lncRNAs with regulatory functions, there are still tens - if not hundreds of thousands - of uncharacterized lncRNAs. Taking into account the physicochemical properties of an RNA molecule and the already known examples of lncRNA functions, the general features of lncRNA mechanisms of action are as follows:

1. Competitive binding to prevent interactions. RNA can form structures that mimic specific DNA *cis* elements as decoys and associate with DNA-binding proteins, for example transcription factors, to preclude their access to DNA recognition sites (Kino et al. 2010).
2. Scaffolding to assemble larger complexes. RNAs may serve as adaptors that link together proteins into complexes (Tsai et al. 2010, Zappulla, Cech 2006). The formation of lncRNA-protein complexes where the RNA joins several proteins together seems to be a universal concept (Rinn, Chang 2012).

3. Sequence-specific guide of a protein machinery. Besides the a generally known mechanism of small ncRNAs, many individual lncRNAs are also required for the proper localization of specific protein complexes and can serve as guides to target gene - or even whole chromosome silencing. Guide lncRNAs combine two basic molecular functions- binding their specific protein partner and linking it with selective regions of the genome (Boumil, Lee 2001, Huarte et al. 2010).

Moreover, besides the ability to bind to its RNA, DNA and protein targets, and switch its allosteric structure, RNA is also catalytically potent (as described in section 2.2.), which is germane in the consideration of possible functions of lncRNAs.

Nearly half of the human genome is derived from transposed repetitive "selfish" elements and therefore considered as "junk DNA" with no use and function other than self-propagation. However the main mechanism of genome evolution is DNA transposition and duplication that enables to cut, copy, paste, combine and propagate functional cassettes and regulatory regions of genes. Thereby it cannot be ruled out that a large portion of transposon-derived sequences in human genome is actually a critical instigator of our genetic evolution. This suggestions is supported by differential expression patterns of many transposon transcripts, particularly in embryogenesis and in brain development (Taft et al. 2007, Taft et al. 2010a, Mattick, Makunin 2006, Carninci et al. 2005).

With assessment of the regulatory dimension that small and long ncRNAs add to gene expression control, it is not surprising to reveal their association with complex diseases. Further, most of the genetic variation that affects complex human traits appears to stem from regulatory mutations and notably, over 90% of all the loci mapped in genome-wide association studies are non-protein-coding. Intriguingly, many of these non-coding regions are also differentially expressed (Taft et al. 2010a). This refers to the far more important role of ncRNAs- not only in the regulation of human cells' functions but also in the determination of who we are.

2.4. RNA-protein interactions

mRNA is constantly coated with various proteins that can package, organize, protect and prepare RNA for downstream processes. They bind RNA via different types of binding domains and form ribonucleoprotein complexes (RNPs). The repertoire of proteins bound (or not bound) to mRNA can indicate a correctly processed mRNA molecule and reference on its status (Hocine et al. 2010). During mRNA post-transcriptional processing and translation, some RNA-binding proteins are stripped from it and replaced by others. This kind of "mRNA tagging" is used to sense and regulate gene expression. For example, EJC, cap-binding complex and poly(A)-binding proteins, can indicate a new complete and non-translated mRNA molecule, yet the presence of snRNPs would signify incomplete or aberrant splicing (Moore 2005). Further, the cap-binding-complex (CBC) attaches to the 5' cap structure of newly synthesized mRNA to protect and guide mRNA nuclear export. During the first ''pioneering round'' of translation, the CBC proteins CBP20 and CBP80 are replaced by the cap-binding translation initiation factors EIF4E, EIF4A1 and EIF4G (Szostak, Gebauer 2013). Heterogeneous nuclear ribonucleoproteins (hnRNPs) are a multifunctional family of RNA-

binding proteins which are involved in processing pre-mRNAs into mature mRNA. Various hnRNPs can participate in the packing of nascent RNA transcripts, regulation of constitutive and alternative splicing, but also in mRNA export and stability. (Dreyfuss et al. 1993, Han et al. 2010).

Mature mRNAs contain regulatory sequences before the start codon and after the stop codon. These untranslated regions (UTRs) are the repository of regulatory factors- RNA binding proteins (RBP, considered further below), which determine mRNA stability, localization and translational efficiency. Usually, UTRs bind *trans*-acting elements which repress the translation of mRNA (Mignone et al. 2002, Kwon, Hecht 1993). 3' UTR appears to be a more important contributor to the post-transcriptional regulation whose average length in humans is approximately 800 nucleotides, compared to only approximately 200 nucleotides for 5'-UTRs (Mignone et al. 2002, Mignone et al. 2002, Dassi et al. 2013). Longer UTRs permit more regulatory options by *trans*-elements and are therefore in negative correlation with the translational potential of the mRNA (Szostak, Gebauer 2013). Besides regulatory proteins, the 3' UTR is also the major binding target for miRNAs (as considered further below), which can decrease gene expression by inhibiting mRNA translation or directly causing its degradation (Szostak, Gebauer 2013). 5'UTR may contain several *cis* regulatory elements which can directly, or via *trans*-elements, promote or inhibit translation initiation. An example of a negative gene regulation by 5' UTR binding *trans*-elements is the iron regulatory proteins that recognize specific elements on mRNA and bind them upon low iron concentrations (Goss, Theil 2011). Further, translation can be promoted by internal ribosome entry site (IRES) elements- highly structured RNA sequences originally discovered in picornaviruses (Pelletier, Sonenberg 1988)

Recently, a UV-cross-linking based, mRNA interactome-capure study has identified 860 mRNA binding proteins, including unexpectedly many metabolic enzymes. These mRNA-binding proteins may broadly connect intermediary metabolism with RNA biology and post-transcriptional gene regulation (Castello et al. 2012). Thereby, to proceed, it is advisable to consider some RNA-binding proteins and protein families which are important in the context of this thesis.

2.4.1. Specific RNA-binding proteins

2.4.1.1. SAM68

A particularly notable protein in current research is the KH domain-containing, RNA-binding, signal transduction-associated protein 1 (KHDRBS1), also known as Src-associated substrate in Mitosis of 68 kDa (SAM68). SAM68 is one of the RNA binding proteins which recognizes U(U/A)AA direct repeats with a high relative affinity (Chen et al. 1997) It has been assessed to be involved in several signal transduction pathways as an adaptor protein, although one of its major functions is to regulate alternative splicing (Iijima et al. 2011, Huot, Richard 2012). By this mechanism, extensive studies have defined SAM68 as a crucial regulator of neurogenesis, spermatogenesis, osteogenesis and adipogenesis (Huot et al. 2009, Huot, Richard 2012, Iijima et al. 2011).

2.4.1.2. ELAVL1/HUR

An RNA-binding protein, Embryonic Lethal Abnormal Vision (ELAV) L1/Human antigen R (HUR) is a well-studied post-transcriptional RNA regulator. ELAVL1/HUR binds to AU-rich elements (AREs) in 3'UTR of mRNAs. AREs are 50-150 nucleotide copies of AUUUA sequences that bind proteins, which can stabilize or destabilize mRNA (Dassi et al. 2013). Approximately 5-8% of all mRNAs are predicted to contain AREs (Decker, Parker 2012). ELAVL1/HUR contains three ARE-binding RNA recognition motifs (Doller et al. 2008, Chi et al. 2011). AREs signal mRNA degradation and HUR stabilizes them upon binding. Although mainly nuclear, HUR is a nucleo-cytoplasmic shuttling protein (Doller et al. 2008, Fan, Steitz 1998). HUR is essential for male germ cell differentiation and over a hundred of mammalian genes, including Hspa2/Hsp70-2/Hsp72, are under its post-transcriptional control (Chi et al. 2011). HUR has also been demonstrated to dynamically regulate miRNAs, either by synergy or disengagement with them from repressed mRNAs (Meisner, Filipowicz 2011).

2.4.1.3. Mouse VASA homologue

A vital role in RNA processing (and indeed in nearly all aspects of RNA metabolism) is played by RNA DEAD-box helicases, which use ATP energy to unwind and rearrange RNA-RNA and RNA-protein interactions. It is the largest helicase family found in all three domains of life and its members are characterized by a highly conserved Aspartate-Glutamate-Alanine-Aspartate (DEAD) motif (Linder, Jankowsky 2011, Lim et al. 2000). A wide range of biochemical activities has been attributed to DEAD-box helicases. For example EIF4A3/DDX48 is a DEAD-box protein that is part of the core of the EJC, as mentioned above and gonadotropin-regulated testicular RNA helicase (GRTH/DDX25) is DEAD-box helicase which is essential for the completion of spermatogenesis (Tsai-Morris et al. 2010).

Mouse Vasa Homologue, MVH/DDX4 is a central protein in this doctoral thesis. MVH is a DEAD-box helicase that has been generally referred to as a specific germline marker, required for germline determination, maintenance and function (Castrillon et al. 2000). Recent data show that MVH also functions in primordial germ cell specification, stem cell maintenance, regulation of cell cycle and piRNA biogenesis for retrotransposon silencing (Kuramochi-Miyagawa et al. 2010, Yajima, Wessel 2011). The expression of MVH protein is exclusively restricted to the primordial germ cells in the embryos of male and female mice. MVH is also expressed in germ cells during oogenesis and spermatogenesis (Toyooka et al. 2000). MVH knockout results in complete sterility in male mice yet females do not show any obvious reproductive defects (Tanaka et al. 2000). The expression of the VASA-homologue protein is evolutionally conserved in the cytoplasm of spermatocytes and round spermatids of primates, ungulates, marsupials, chiropterans and most likely in all mammals (Toyooka et al. 2000). In round spermatids, MVH localization is concentrated to the male germ granule known as the chromatoid body which is discussed in detail, in section 2.6.

2.4.1.4. Argounaute proteins

Argonaute proteins are the catalytic components in small-RNA-guided gene-silencing processes. These accommodate the small RNA component, such as miRNAs, siRNAs or piRNAs and coordinate the gene-silencing events through a mechanism generally termed as RNA interference (RNAi) (Meister 2013). Argonautes form an evolutionarily conserved family which consists of two subclades, the AGOs and the PIWIs, both of which participate in RNAi. Animals have four ubiquitously expressed AGOs (AGO1-4) and three mainly germ cell-specific PIWIs (MIWI, MILI and MIWI2 in mouse). AGO proteins bind siRNAs and miRNAs while PIWIs bind piRNAs (Fig. 5). In addition to the piRNA pathways, PIWIs have been reported in other, non-piRNA related functions such as stabilizing spermatogenic mRNAs and are associated with cytoplasmic poly(A)-binding proteins (PABPCs) (Vourekas et al. 2012, Kimura et al. 2009).

2.4.2. RNP complexes and RNP granules

RNA and proteins are both important structural and catalytic molecules. As considerd above, the common mechanism of action for lncRNAs would appear to be via the formation of ribonucleic- protein complexes (RNPs). The combination of RNA and protein can create powerful composites with complex structures which have a diverse spectrum of merged functions and unique properties (Zappulla, Cech 2006, Rinn, Chang 2012). RNPs can, in turn, aggregate into larger granules as both a cause and a consequence of altered mRNA translation, decay or editing (Decker, Parker 2012).

RNA granules are key modulators of post-transcriptional and epigenetic gene expression (Anderson, Kedersha 2009). Processing bodies (P-bodies) and stress granules are two main types of cytoplasmic mRNP granules in the somatic cells of eukaryotes. P-bodies are the factories of mRNA decay, containing most enzymes involved in mRNA turnover, for example the decapping enzyme complex Dcp1/Dcp2, 3’ exonuclease Xrn1 and the Ccr4/Pop2/Not deadenylase complex (Kulkarni et al. 2010). P-bodies are conserved in somatic cells of vertebrates, invertebrate, yeast, plants and trypanosomes and they have been shown to play fundamental roles in general mRNA decay, NMD and miRNA pathways (Kulkarni et al. 2010, Decker, Parker 2012). However, P bodies are not required for RNA degradation - instead the active silencing pathways are required for P-body formation, which indicates that P bodies are not the cause, but a consequence, of silencing (Eulalio et al. 2007). The poly(A)-binding protein (PABP) is absent from P-bodies which suggests that the deadenylation of mRNAs - the general initiation of their degradation, precedes their targeting to the P-bodies. Further, translation initiation factors and ribosomal proteins are generally excluded from P-bodies. Notably, P-bodies have also been demonstrated to store and release some transnationally repressed mRNAs (Kulkarni et al. 2010, Decker, Parker 2012).

Stress granules form as a response to cell stress. Translational arrest aggregates elements of translation, inclusive of small ribosomal subunits, translation initiation factors and untranslated mRNAs, into stress granules. Although stress granules and P-bodies are distinct structures that form independently, they seem to be functionally connected in the regulation of cytoplasmic mRNA translation/degradation balance, in response to environmental factures (Anderson, Kedersha 2009, Parker, Sheth 2007, Decker, Parker 2012).

2.5. Spermatogenesis

Spermatogenesis is a complex and continuous process which produces over a thousand spermatozoa every second (in man). A complete wave of spermatogenesis takes around 35 days in mouse and 64 days in man (Clermont 1972, Heller, Clermont 1963) and can be divided into three phases: mitotic, meiotic and post-meiotic (Fig. 6 and 7). In the mitotic phase, proliferating spermatogonia undergo consecutive mitotic divisions to provide a sufficient amount of differentiating cells from a few slowly dividing type A spermatogonial stem cells. During the two meiotic divisions of the second phase of spermatogenesis, the genetic material in spermatocytes is recombined and segregated. In the post-meiotic differentiation phase, haploid spermatids undergo a series of morphological and structural changes to become spermatozoa- vehicles which are responsible for the delivery of genetic information to the egg. These massive reorganizations include the formation of acrosome and flagellum, nuclear reshaping and the compaction of chromatin through replacement of histones with protamines (Hess, Renato de Franca 2008, Clermont 1962, Clermont 1972, Russell et al. 1990).

The production of spermatozoa takes place in seminiferous tubules of the testis (Fig. 6). Seminiferous epithelium consists of male germ cells (spermatogonia, spermatocytes, spermatids) at different developmental stages and the somatic Sertoli cells which provide the necessary microenvironment and mediate spermatogenesis. Sertoli cells extend from the basal lamina to the tubular lumen of the seminiferous epithelium and form an amorphous medium around the germ cells, dividing them between two compartments by the so-called blood-testis barrier (Dym, Fawcett 1970, Russell et al. 1990). The basal compartment houses spermatogonia that reside next to the basal lamina, while the meiotic and post-meiotic male germ cells mature in the adluminal compartment. Mature spermatozoa are released into the lumen in the center of the tubules (Clermont 1972). These seminiferous tubules are in convoluted, occasionally branching, loops that all converge in a duct which directs the released, fully differentiated spermatozoa out of the testis to the epididymis where spermatozoa undergo their final maturation.

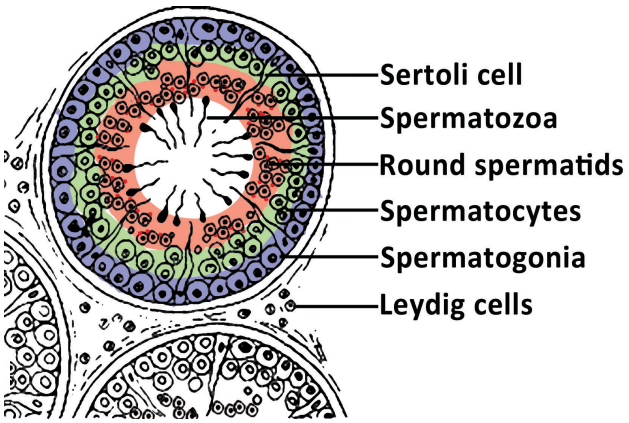

Figure 6. *A schematic illustration showing the cellular architecture of a cross-section of a seminiferous tubule in the mammalian adult testis. Three colors represent the three phases of spermatogenesis: blue- mitotic, green- meiotic, red- post-meiotic. Male germ cells are engulfed by the somatic Sertoli cells. During the development, the germ cells progress inwards, towards the center of the tubule. The mature spermatozoa are released in the tubular lumen.*

The intertubular space of the testis contains a support infrastructure for spermatogenesis, including blood vessels, nerves, myoid cells and Leydig cells. Peritubular myoid cells are contractile and surround the seminiferous tubules to provide them with structural integrity and are involved in the transport of testicular fluid and spermatozoa in tubules. Leydig cells are responsible for producing the male sex hormone, testosterone. Testosterone diffuses locally to the seminiferous tubules where it promotes the development of the germinal cells but also gets distributed throughout the whole organism when transported into blood vessels and is associated with sex drive and secondary sex characteristics (Maekawa et al. 1996).

Spermatogenesis is under the direct control of the hypothalamus-pituitary axis. Gonadotropin-releasing hormone (GnRH) secreted by the hypothalamus induces the production of follicle-stimulating hormone (FSH) and luteinizing hormone (LH) in the pituitary gland, in the brain. LH stimulates Leydig cells to produce testosterone which enhances spermatogenesis. FSH promotes Sertoli cells inside the seminiferous epithelium. Sertoli cells in turn produce inhibin which downregulates FSH synthesis and release. Testosterone is also part of a negative feedback system through inhibition of the secretion of GnRH and FSH and LH (Russell et al. 1990). Several studies have demonstrated that lactate is the central energy metabolite in meiotic and post-meiotic male germ cells. Lactate is converted from glucose by Sertoli cells under the control of the endocrine system, primarily FSH, insulin and insulin growth factor I, and transported to post-meiotic germ cells (Boussouar, Benahmed 2004). At present, the reason to utilize lactate as the primary energy source remains unknown.

Spermatogenesis proceeds sequentially so that each cross-section of a seminiferous tubule represents a specific stage of the seminiferous epithelial cycle and contains the respective cell types of that stage. The seminiferous epithelial cycle is divided into 12 stages in mouse (Fig. 7) (Russell et al. 1990). Each stage of the cycle follows in an orderly sequence along the length of the tubule and the adjacent segments of the tubule communicate in some unknown manner. Notably, the stages of the seminiferous epithelial cycle can be determined by the tubule's transillumination pattern which reflects the absence/presence and special organization of condensed nuclei of elongating spermatids. This allows to study spermatogenesis in greater detail, in a stage specific manner (Parvinen 1982, Toppari, Parvinen 1985, Kotaja et al. 2004).

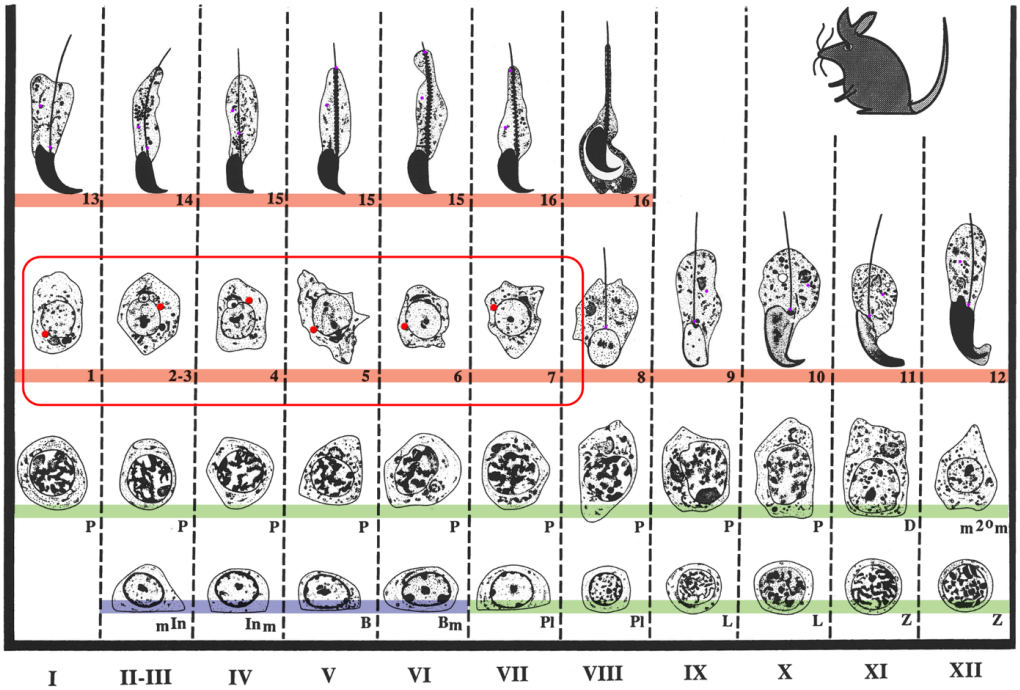

Figure 7. *The cycle of the seminiferous epithelium in mouse, progressing in horizontal columns from left to right, from bottom to top. The vertical columns (dashed lines) depict the specific cell type repertoire in a given stage (Roman numerals) of the seminiferous epithelial cycle. Male germ cell differentiation is divided into three parts- mitotic divisions of spermatogonia (blue line), meiosis in spermatocytes (green line), and the differentiation of haploid spermatids (red line). Arabic numerals refer to different steps of post-meiotic germ cell differentiation: round spermatids (steps 1-8), elongating spermatids (step 9-16). CBs are marked as red dots. The steps at which the MVH-positive CBs in round spermatids become isolated by this study's CB-immunoprecipitation protocol are indicated (red box). In, intermediate spermatogonia; B, type B spermatogonia; Pl, preleptotene spermatocytes; L, leptotene spermatocytes; Z, zygotene spermatocytes; P, pachytene spermatocytes; Di, diplotene spermatocytes; m2°m, meiotic divisions. Type A spermatogonia and Sertoli cells are not included in the figure. Different stages can be visualized by transillumination-assisted microscopy where stages differ from each other by their light absorption characteristics. Modified from (Russell et al. 1990).*

2.5.1. Chromatin remodeling during spermatogenesis

Epigenetic transitions, along with chromatin remodeling, play an important role in the development of the male germline. The first global epigenetic reprogramming occurs during embryonic development in primordial germ cells/prospermatogonia. 80-90% of the somatic DNA methylation pattern is erased and new epigenetic marks are established to replace the somatic program. Such epigenetic resetting also include paternal imprinting- an epigenetic filter which determines the specific expression profile of the male-contributed half of the genome of the developing embryo (Kota, Feil 2010, Ewen, Koopman 2010, Meikar et al.

2012). During the erasure of epigenetic marks in primordial germ cells, the transposons become de-inhibited and need to be silenced in order to protect the genome's integrity. piRNA pathway and piRNAs, as described above, contribute to this crucial alternative mechanism of silencing the activated transposons (Aravin et al. 2007, Carmell et al. 2007).

Meiosis is another occasion of massive chromatin remodeling in male germline. In meiosis, homologous chromosomes become aligned to allow synapsis and recombination eventually, which eventually leads to the formation of genetically unique haploid gametes with shuffled parental genes. Besides the *cis*-acting elements, the regulation of meiotic events is also mediated by, for example, histone modifications and alternative histone variants (Meikar et al. 2012, Kimmins, Sassone-Corsi 2005). In addition, the X and Y chromosomes are subjected to chromatin modifications which lead to their silencing in the so-called XY-body or the sex body (Burgoyne et al. 2009, Meikar et al. 2012). The meiotic sex chromosome inactivation (MSCI) appears to be a general mechanism for transcriptional silencing of any unsynapsed chromosome region during meiosis, to also lead to their substantial post-meiotic repression in spermatids (Turner et al. 2006).

The third global chromatin remodeling in male germ cells occurs during the elongation of post-meiotic spermatids when most of the nucleosomal histones are replaced by sperm-specific protamines. However, during this time, the methylation of DNA does not change detectably any more. The timing of these transitions is strictly regulated and begins with histone modifications and the incorporation of a variety of histone variants. This is followed by an exchange of the histones with specific transition proteins and then subsequently by the protamines. Protamines are small, arginine-rich, highly basic proteins which enable more than 10 fold more compact packing of paternal DNA into the sperm head (Meikar et al. 2012, Kimmins, Sassone-Corsi 2005).

2.5.2. Transcriptional activity and regulation of gene expression

Mammalian testis possesses the highest transcriptome complexity, when compared to other organs (Soumillon et al. 2013). This is due to meiotic spermatocytes and especially post-meiotic round spermatids which have a remarkably diverse RNA content with a pronounced diversity for non-coding transcripts. The appearance of lncRNAs in meiotic and post-meiotic cells is suggested to be at least partly the result of the ''leaky'' transcription of the genome as a consequence of chromatin remodeling (Soumillon et al. 2013). The most abundant and diverse population of all small RNAs- the pachytene piRNAs, is also characteristic to the same cell types (Aravin et al. 2006). Besides ncRNAs, a significantly larger number of autosomal protein-coding genes are transcribed in testis (~18,700) than in brain (~18,000) and liver (~15,500) (Soumillon et al. 2013). Spermatocytes and spermatids are also cells with one of the most complex alternative splicing patterns in the testis and thereby, the whole organism (Elliott, Grellscheid 2006, Soumillon et al. 2013, Laiho et al. 2013).

Further, an extra translational control is required for the terminal stages of spermatid differentiation, when transcription has ceased but *de novo* protein production is required. Thus, the mRNAs for many spermatozoon proteins have to be priorly synthesized in meiotic

cells or round spermatids, and subsequently translationally inhibited and stored for several days (Kleene 2001, Monesi 1964).

The complex transcriptome in male germ cells requires correspondingly complex post-transcriptional control mechanisms. This is reflected in the high number of RNA-binding proteins in spermatogenic cells, many of which are testis-specific (Paronetto, Sette 2010), with other peculiarities, which will be considered further in the next chapter.

2.6. The chromatoid body- the largest RNP granule

The chromatoid body (CB) is a phenomenally large RNP granule which appears in the late pachytene spermatocyte and condenses in the cytoplasm of a post-meiotic round spermatid into a single perinuclear structure (Fig. 8) (Fawcett et al. 1970, Meikar et al. 2011). Due to its size (~1 µm in diameter), the CB is easily visible by a phase contrast microscope and was therefore first observed already in the 19th century (von Brunn 1876). The CB occupies ~0.4% of the cytoplasmic volume, as based on the relative diameters of the CB and the round spermatid cell (10 µm) and the nucleus (5 µm) (Kleene, Cullinane 2011). When prepared for transmission electron microscopy, the CB is visible as a distinct, single finely filamentous, lobulated, dense, non-membranous structure (Kotaja, Sassone-Corsi 2007). The CB forms in the cytoplasm of early round spermatids and remains intact until around step 8 when it will begin to gradually decrease in size. During the course of spermatid elongation, the CB undergoes functional transformation, breaks in to two structures and eventually disappears (Shang et al. 2010, Kotaja, Sassone-Corsi 2007).

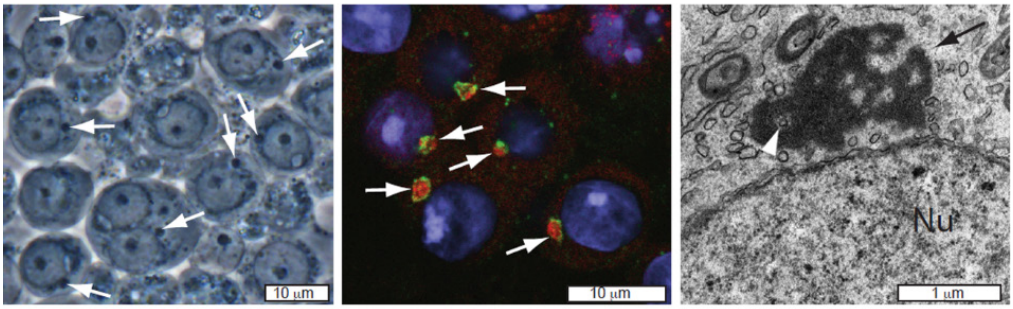

***Figure 8**. The appearance of the CB. Left illustration: Phase contrast microscopy image of CBs (white arrows) in round spermatids prepared by squash preparation (Kotaja et al. 2004). Center illustration: Immunofluorescence microscopy image of paraffin-embedded mouse testis section. CBs (white arrows) are detected by antibodies against two CB-components, MIWI (red) and FYCO1 (green). The nuclei (blue) are counterstained with DAPI. Right illustration: Electron microscopy visualization of the CB (black arrow) in the cytoplasm of a round spermatid in a glutaraldehyde-fixed testis section. The CB lays in close proximity to the nucleus (Nu). One of the surrounding vesicles has been engulfed (white arrowhead). Modified from (Meikar et al. 2011).*

CBs are specific to haploid mammalian cells. Testis sections prepared from several mammals, including the brush-tailed possum (*Trichosurus vulpecula*), horse (*Equus caballus*), pig (*Sus*

domestica), cow (*Bos primigenius*), Japanese monkey (*Macaca fuscata*) and human revealed a specific granular pattern in the perinuclear zone of round spermatids, when stained with anti-MVH antibody, which is considered to be a putative CB, (Toyooka et al. 2000).

Different germ granules are present in germline development. These include pi-bodies and piR-bodies in prospermatogonia, the intermitochondrial cement (IMC) in late spermatocytes and thereafter the CB in haploid spermatids (Aravin et al. 2009, Chuma et al. 2009). These, along with other granules in lower organisms, have different biological functions, but still share many protein components, for example the DEAD-box helicases, Tudor domain-containing proteins and PIWI proteins (Chuma et al. 2009, Meikar et al. 2011). This suggests that germ granules share some similar mechanisms for their actions.

The CB is not a static structure, as its shape, location and composition are dynamically changing. It moves actively and non-randomly in the cytoplasm making frequent contacts with the nuclear envelope and its movement pattern alters during the course of round spermatid development (Parvinen, Parvinen 1979). Electron microscopy studies have demonstrated that euchromatin and nuclear pore complexes are enriched in those areas adjacent to the CB, which appears to receive material directly from the nucleus (Soderstrom, Parvinen 1976, Parvinen, Parvinen 1979). In addition, the CB is constantly surrounded by multivesicular bodies along with different small vesicles, some of which getting occasionally encapsulated in the CB. These vesicles can share similarities to lysosomes and appear to be the part of a larger transport network between the CB, Golgi complex, developing acrosome, mitochondria and endoplasmatic reticulum (Ventela et al. 2003). Haploid round spermatid is functionally diploid, as being connected with the adjacent cell via the intercellular bridges by which material can be shared (Shang et al. 2010, Ventela et al. 2003). Small, CB-derived particles and even larger parts of the CB itself have been demonstrated to travel via these bridges (Ventela et al. 2003).

2.6.1. Known CB components

With immunological detection methods, several proteins have been demonstrated to localize in the CB, most being RNA-binding and associated with RNA processing pathways (Parvinen 2005, Kotaja, Sassone-Corsi 2007, Kimura et al. 2009, Tsai-Morris et al. 2010). One of the major components of the CB is the DEAD-box RNA helicase MVH which is crucial for spermatogenesis and appears to be the general marker for germ granules of divergent animals (Toyooka et al. 2000, Tanaka et al. 2000, Chuma et al. 2009). Although a conserved constituent in both *Drosophila* and mice germ granules, its loss-of-function analyses demonstrate distinct phenotypes in opposite sexes and at different developmental stages of the germline. The same can be observed with Tudor proteins, which are also widely conserved components of germline granules (Chuma et al. 2009).

Tudor proteins are known components of the CB. They can recognize symmetrically dimethylated arginines and function as adaptors to facilitate protein-protein interactions and are required for the assembly of larger ribonucleoprotein complexes (RNPs). Tudor proteins have been implicated in many biological processes such as mRNA splicing, small RNA pathways and transcriptional regulation, but their general well-characterized role is

molecular scaffolding (Pek et al. 2012). In the animal germline, Tudor proteins are required to ensure undisrupted gametogenesis (Chen et al. 2011).

The CB is known to consist of large amounts of RNA, as demonstrated by the accumulation of tritiated uridine and RNase-gold in CB and RNA hybridization experiments with Cy3-labeled oligo(dT) probe (Soderstrom, Parvinen 1976, Meikar et al. 2010, Kotaja et al. 2006, Walt, Armbruster 1984, Figueroa, Burzio 1998). However, ribosomal RNA appears to be excluded from the CB which would indicate that this is not the location for active translation (Kotaja et al. 2006). The characteristics of the later stages of spermatid differentiation, when transcription ceases as the nucleus becomes condensed, suggests that CBs could also operate as a storage of mRNAs. Direct evidence of storage of a translationally repressed mRNA in the CB is, to date, primarily derived from only a single study of the transition protein 2 mRNA in rat testis (Saunders et al. 1992).

2.6.2. Isolation of CBs

To comprehend the functions and mechanisms of the CB, it is essential to know its contents. To achieve this, pure fractions of whole CBs in sufficient amounts must be isolated. CB research has been hindered by the lack of an acceptable purification protocol. The general approach to isolate CBs has been gradient centrifugation. As a powerful method for polysomal fractionation, this lacks the necessary discriminative power to obtain biochemically pure CBs. The last published attempt required 60 testes of 40-day old rats and yielded only 2 mg of protein sample of which 70% were reported to correspond to CBs, while the remainder accounted for other structures, such as dense bodies and satellite-like organelles (Figueroa, Burzio 1998).

Immunoprecipitation (IP) is a relatively new technique which is used to selectively purify an antigen of interest that is present in a complex mixture. IP may offer a possible solution to specifically “fish out” the CB from a cell lysate with a suitable antibody. The antibody itself is attached to a carrier, for example a paramagnetic bead which allows the washing and the separation of the bound CBs from the reminder of the sample mix. As the epitopes for antibodies can be very versatile, the potential of IP is wide. The most important parameter of successful IP is the specificity and sensitivity of the antibody.

3. AIMS OF THE STUDY

Spermatogenesis represents a unique and complex process that is essential for the genetic perpetuation of an individual organism and, ultimately, a species. The mechanisms which account for the control and regulation of this extremely involved process are important not only from a germ cell research point of view, but also to confront the increasing problems with human fertility.

Thereby the aim of this study is to characterize the post-transcriptional regulation of male germ cell differentiation. Its focus is concentrated on the chromatoid body- the conjectured center of RNA regulation in round spermatids.

The specific aims of this study were:

1. Development of a reliable and reproducible protocol for the isolation of whole, pure and intact CB material in sufficient quantities to enable subsequent detailed molecular analyses.

2. Detailed characterization of the CB's composition, namely its proteome and transcriptome content.

3. Based on such acquired data, the assessment of CB functions, mechanisms and roles in post-transcriptional RNA regulation, during haploid male germ cell differentiation.

4. MATERIALS AND METHODS

4.1. Cell and tissue preparations

Mice were maintained in a specific pathogen-free stage at the Central Animal Laboratory of the University of Turku. All the experiments were based on the regulations and guidelines by local laboratory animal authorities and carefully designed to cause minimal harm and distress on animals. For all the persons in charge, and for the animal experiments described, licenses were approved from the Laboratory Animal Care and Use Committee of the University of Turku.

Adult and juvenile mice (C57BL/6) were sacrificed by CO_2 asphyxiation and cervical dislocation. Testes were dissected and decapsulated. The seminiferous tubules were released on a Petri dish containing 1x phosphate buffered saline (PBS) or DMEM (Dulbecco's Modified Eagle's Medium/Nutrient Mixture F-12 Ham [D8437, Sigma]). Transillumination-assisted microdissection method was used to isolate specific segments of the seminiferous tubules (Toppari, Parvinen 1985, Kotaja et al. 2004). Squash preparations of stage-specific pieces of seminiferous tubules were performed as previously described (Kotaja et al. 2004). The preparation of drying-down slides is thoroughly described in (II). Testes from adult mice were used to obtain pachytene enriched populations of spermatocytes, secondary spermatocytes, and round spermatids by centrifugal elutriation as described previously (Barchi et al. 2009).

For the preparation of paraffin sections, mouse testes were fixed in 10% formalin at 4 °C overnight. The fixed samples were dehydrated and embedded into paraffin. 5 µm sections were cut and mounted onto polylysine-coated slides. For the immunochemical analysis, the sections were dewaxed, rehydrated subjected into antigen retrieval by boiling in sodium citrate buffer (10 mM, pH 6.0). For the preparation of frozen sections, the mouse testes were fixed in 4% (w/v) paraformaldehyde (PFA) at 4 °C overnight, followed by the incubation in 1M sucrose at 4 °C overnight. The samples were then embedded into Tissue-Tek matrix (Sakura) at -40 °C. 5 µm sections were cut at -20 °C and mounted onto poly-lysine-coated slides.

4.2. Immunofluorescence analysis

Stage-specific squash preparations of male germ cells, drying-down preparations and prepared tissue sections were post-fixed in 4% PFA for 20 min. This was followed by permeabilization in 0.2% (v/v) Triton X-100 in 1x PBS for 5 minutes and blocking in 10% (w/v) BSA in 1x PBS for 1 hour at room temperature. The immunofluorescent staining was performed in 5% (w/v) BSA in 1x PBS at 4 °C overnight with the specific primary antibodies described in Table 1. AlexaFluor 488/594 conjugated secondary antibodies (Life Technologies) were used in dilution 1:500 for 1 hour at room temperature or at 4 °C overnight. Sections were mounted in a medium with DAPI (Santa Cruz). Photomicrographs were taken with Olympus DP72 digital color camera mounted onto a Leica DMRB

microscope (Leica Microsystems, Wetzlar, Germany) using PL FLUOTAR 40×/0.70 and N PLAN 40x/0.65 PH 2 objectives and cellSens Entry 1.5 (Olympus) digital imaging software. All images were processed using Photoshop (Adobe).

***Table 1**. List of primary antibodies used in immunocytochemical, immunohistochemical stainings, immunoblottings and/or immunoprecipitations.*

ANTIBODY	DILUTION	ANTIGEN NAME	CAT. NO.	MANUFACTURER
ACR	1/500	Acrosin	sc-67151	Santa Cruz
GRTH/DDX25	1/200	Gonadotropin-regulated testicular RNA helicase	sc-51271	SantaCruz
EIF4A3/DDX48	1/200	Eukaryotic initiation factor 4A-III	17504-1-AP	Proteintech,
MIWI/PIWIL1	1/1000	Piwi-like protein 1	G82	Cell Signaling Technol.
MVH/DDX4	1/1000	Mouse VASA homologue	in-house	
MVH/DDX4	1/1000	Mouse VASA homologue	ab13840	Abcam
PABP	1/1000	Poly(A)-binding protein	ab21060	Abcam,
RBM8A/Y14	1/200	RNA binding motif protein 8A	14958-1-AP	Proteintech
SAM68/KHDRBS1	1/1000	Src-Associated substrate in Mitosis of 68 kDa	sc-333	Santa Cruz

4.3. Electron microscopy

Small pieces of testis were cut, fixed in 5% glutaraldehyde and treated with a potassium ferrocyanide-osmium fixative. The samples were embedded in epoxy resin (Merck), sectioned, stained with 5% uranyl acetate and 5% lead citrate, and visualized on a JEOL 1200 EX transmission electron microscope

4.4. Seminiferous tubule cultures

Mouse testes were decapsulated in DMEM medium. Segments of the seminiferous tubules representing various stages of the seminiferous epithelial cycle were incubated on a glass slide in 30-50 µl of medium with 1 mM Ethynyl Uridine (EU) (Life technologies) at 34 °C overnight in highly humidified atmosphere containing 95% air and 5% CO_2. The cultured tubule were subjected to squash preparations and the nascent RNA was visualized by Click-iT® RNA Alexa Fluor® 488 Imaging Kit (Life Technologies). For the chase experiment to follow the translocation of labeled RNAs, EU was removed from the media after 8 hours of incubation and the tubule culture was continued for 12 hours. For transcription inhibition experiments, actinomycin-D (Sigma) was added 1 µg/ml to the culture medium.

4.5. CB isolation

Detailed protocol of the CB isolation and sample preparations is described in (II). In short, testicular cells from adult or juvenile C57BL6 wild-type or knockout mice were liberated by collagenase digestion. The cells were washed, PFA cross-linked and lysed, followed by low-speed centrifugation. The CB-enriched pellet fraction of the lysate used for immunoprecipitation with anti-MVH antibody.

4.6. Western blotting and silver staining

Proteins were separated by 10% SDS–PAGE or 4-20% precast TGX gels (BioRad) and electroblotted to PVDF membrane. After blocking with 5% (w/v) non-fat milk in 1x PBS, the membranes were incubated with primary antibodies (Table1). Horseradish peroxidase-conjugated anti-rabbit IgG, anti-mouse IgG (GE Healthcare) or anti-rabbit light chain (Millipore) secondary antibodies were used and the signals were detected with ECL Plus Western blotting detection reagents (GE Healthcare) and FujiFilm LAS-4000 camera system. silver staining was performed using PageSilver™ silver staining kit (Fermentas) following manufaturer's instructions.

4.7. Mass spectrometry

CB protein samples were separated in the NuPAGE Novex Gel System (Invitrogen) and stained with SimplyBlueSafeStain (Life Technologies). Gel pieces were cut out from the gel lanes avoiding any cross-contamination followed by in-gel digestion with Trypsin (Promega) and the extraction of peptides. The peptides were analyzed with LC-MS/MS using an Agilent 1200 series nanoflow system (Agilent Technologies) connected to a LTQ Orbitrap mass-spectrometer (Thermo Electron, San Jose, CA, USA) equipped with a nanoelectrospray ion source (Proxeon, Odense, Denmark). The analyses were performed in five independent technical and biological replicates. Data were analyzed by Mascot search engine using NCBI MOUSE database. For approximate quantitation Exponentially Modified Protein Abundance Index (emPAI) (Ishihama et al. 2005) was used.

4.8. RNA extraction and gel electrophoresis

RNA was extracted from whole mouse tissues and from elutriated testis cell fractions using Trisure reagent (Bioline) and modified protocol adding extra chloroform and ethanol purification steps to remove traces of phenol. Isolated RNA was analyzed by NanoDrop (Thermo Scientific), Bioanalyzer (Agilent) and RNA PAGE in 15% denaturing urea-polyacrylamide gel, post-stained with SYBR Gold (Invitrogen) and visualized by Fastgene Blue LED Illuminator (Nippon Genetics). For piRNA isolation, the respective ~30 piRNA band was cut and eluted from crushed gel slices in TE buffer for 1 hour and the isolated piRNAs were precipitated with 75% ethanol, 0.4 M NaCl and 20 µg GlucoBlue^{Tm} (Invitrogen) at -80 °C overnight. For CB RNA profile analysis, CB RNAs from adult and juvenile mice were γ[32P] ATP-labeled with exchange reaction using the T4 Polynucleotide Kinase kit (EK0032, Fermentas) and separated in denaturing urea-polyacrylamide gel. The radioactive signals were detected by FujiFilm BAS-5000 phosphoimager system.

4.9. Small RNA library generation and analysis

19-33 nt RNAs were size-selected from CB and control samples. This was followed by small RNA library generation as previously described (Malone et al. 2012). The samples were sequenced on Illumina Genome Analyzer II for 36 cycles in a single end sequencing run. The

3' adaptor sequence was clipped and 15nt or longer reads were selected, collapsed and mapped to mouse genome (2007 NCBI37/mm9) without mismatches. Previously described pachytene cluster coordinates were used for piRNA analysis (Girard et al. 2006).

4.10. Transcriptome libraries and analysis

Transcriptome libraries were prepared using the not-so-random (NSR) priming approach (Armour et al. 2009). The samples were sequenced on Illumina Genome Analyzer II for 50 to 75 cycles in a paired-end sequencing run. After trimming the first 8 nucleotides, the reads were mapped to mouse genome (2007 NCBI37/mm9). UCSC RefSeq July 2011 and piRNA cluster coordinates were used for mapping reads to known transcripts. Cufflinks (Trapnell et al. 2010) was used for the identification of novel non-coding transcripts. Sequences overlapping with exons of known RefSeq transcripts or piRNA clusters were removed from the analysis. Finally all the predicted novel transcripts from all replicates were merged and the redundant calls were removed. Also, the assembled transcripts that were less than 50 nt were removed from the analysis.

4.11. qRT-PCR

Isolated RNA was treated with DNase I, Amplification Grade (Invitrogen). cDNA synthesis and qPCR quantification were performed with DyNAmo cDNA Synthesis Kit (Finnzymes) and DyNAmo Flash SYBR Green qPCR Kit (Finnzymes), respectively. Primer sequences are available in Table 2. The assays were performed in at least three independent technical and biological replicates.

***Table 2**. Primers used for RT-PCR and qRT-PCR analyses. TP2, PRM2 and ODF1 mRNAs are the predicted CB targets, while CD9 and ERM are negative controls. Expression of RPL13A and YWHAZ was used to normalize the expression of noncoding RNA targets (CUFFs).*

GENE/TARGET NAME	FORWARD (5' – 3')	REVERSE (5' – 3')
TP2	GAGCCTTCCCACCACTCAT	TGCACTGGTTACTGGTGTGACT
PRM2	CAGAAGGCGGAGGAGACAC	CTCCTCCTTCGGGATCTTCT
ODF1	TGTGGCCTGTGTGACCTCTA	TTCTATTTGTCGTCCTTCTGAGTCT
CD9	TGCAGTGCTTGCTATTGGAC	GGCGAATATCACCAAGAGGA
ERM	CAAGAGCCCCGAGATTACTG	CTCGGGTACCACGCAAGTAT
CUFF1279	TGTCAATGTCTGAGCATGTTTCC	TCCTGGGGATAAAAGAGACAAAT
CUFF1289	ATGTCCCCTTAGGTCTCAGC	CATGGACACAGCAATGGATGTA
CUFF1617	CCTCTAGGAAACATGCTCAGACA	CTTAGCCTGGGAACACTGCT
CUFF2242	AGGACACAGATGCATCCTAAATGA	CCTTCCCTGGATGTCACTGG
CUFF2476	AACCACCCAGTCTGCTGAAA	GCGTGTAGAATCTCTGCCCA
CUFF3809	CCACCCCAAGCTTGAACATA	TGTTATGAAGCGTGGAGGCA
CUFF3839	CTGTAGTACCCCACAGTTCAGT	AACTTATGGGCCCCTCTCCA
CUFF4273	CTGAGGTCATCTGGGGAGGA	AGCCTTTGACTGGTCTGTGT
RPL13A	ATGGCGGAGGGGCAGGTTCTG	GTACGACCACCACCTTCCGGC
YWHAZ	CGACCACCCATTGTCCCCGC	ACGTCAAACGCTTCTGGCTGC

4.12. Northern blotting

Extracted RNA was separated by 15% denaturing polyacrylamide urea MOPS-NaOH gel, transferred onto nylon membrane and cross-linked with 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (Sigma) (Pall, Hamilton 2008). The membrane was hybridized using EasyHyb (Roche) solution following the protocol suggested by the manufacturer. A previously reported mouse pachytene piRNA, piRNA-030365 (NCBI ID: DQ715868) (Lau et al. 2006) signal was detected with [γ-32P] ATP-labeled LNA probe (Exiqon): 5'-aataAagCtaTctGagCacCtgTgtgatgtt-3' (capital letters stand for LNA and small letters for DNA nucleotides).

4.13. *In situ* hybridization

Paraffin-embedded testis sections were prehybridized in hybridization buffer containing 50% formamide (Sigma), 5x SSC (saline-sodium citrate buffer) (Sigma), 250 µg/ml yeast RNA (Ambion), 1× Denhardt's solution (Fluka) in nuclease-free water. Hybridization was done in the same buffer containing 100 µM Cy5-poly(T) 25mer DNA probe (Exiqon) or 2.5 µM of 5'DIG-labeled LNA probe (Exiqon): 5'-GccAtcActCcaAtaTttGgt-3' (capital letters stand for LNA and small letters for DNA nucleotides) against a previously reported mouse pachytene piRNA, piRNA-038309 (NCBI ID: DQ727400) (Lau et al. 2006) or scrambled probe (Exiqon) at 37 °C overnight. Sections were washed three times in 0.1× SSC and once in 1× SSC. Sections were mounted in medium with DAPI and the photomicrographs were taken as described above (section 4.2.).

4.14. RNA processing assays

Substrate RNA for the CB assay was prepared from several hybridization probes of 500-1000 nt with SP6/T7 Transcription kit (Roche Applied Science) in the presence of [α-32 P] UTP (PerkinElmer) according to the manufacturer's instructions. Before the assay, the crosslinks of the isolated CBs were mildly reversed in RNA assay buffer (20 mM Tris-HCl, 2 mM $MgCl_2$, 75 mM NaCl, pH: 7) at 37 °C overnight. IgG-immunoprecipitates from the CB isolation procedure were used as negative controls. For the RNA slicing assays, different internally radiolabeled substrate RNAs were added to the CB-Dynabead or IgG-Dynabead complexes and incubated at 37 °C for 30 min. When required, ATP (10 mM) and/or Ribolock^{Tm} (1:40) (Fermentas) were added to the assay buffer. The CB-Dynabead complexes were then separated from the processed substrate RNAs with magnet and washed three times with RNA assay buffer. The assay was repeated with a new RNA substrate several times. After the assays, the processed target RNAs and the CB- bound RNAs were purified and separated in 10% denaturing urea-polyacrylamide and 1% agarose gels. The RNA was visualized by SYBRGold (Invitrogen) staining and phosphoimager analysis as described above.

5. RESULTS

5.1. Isolation of CBs (I, II)

A CB isolation procedure was developed which takes advantage of the compact nature and large size of the CB and an excellent antibody against its major component, MVH. This protocol takes 2 days and produces a pure fraction of CBs. The estimated yields of this isolation procedure from only 1–2 adult mice testes are in the order of hundreds of nanograms of CB-associated proteins and RNA, which is efficient for most downstream analyses. The starting material is a mixed population of cells, released from mouse testes by collagenase digestion followed by filtration. No prior fractionation of the cell types for the isolation of round spermatids is required. The released testicular cells are cross-linked with formaldehyde to preserve the molecular interactions inside the CB. Cell lysis with detergents and by sonication is followed by a low-speed centrifugation step which concentrates the CBs and removes the non-CB associated MVH. The CB-enriched pellet fraction of the cell lysate is then subjected to immunoprecipitation (IP) with an anti-MVH antibody. This protocol works comparably well for in-house and commercially available anti-MVH antibody (Abcam, ab13840). As a negative control for the IP, rabbit IgG or antibodies against the non-CB related acrosin, were used. Most of the immunoprecipitation occurs within the first 1-2 hours and reaches the plateau within 8 hours. For convenience, the IP was carried out overnight. After IP, virtually all CBs in the sample became attached to the paramagnetic beads via anti-MVH antibody.

To monitor the success of the CB isolation and validate the purity of the obtained material, the test samples from different steps of the protocol were analyzed by a variety of methods (II, Fig. 1). Western blot and silver staining analyses confirmed the efficiency of the CB isolation (I, Fig. 3A-D). Immunostaining of drying-down slide samples was used to estimate the integrity and ratio of the enriched CBs (II, Fig. 2A; III, Fig. S3A). RNA profile analysis of the purified CBs by polyacrylamide gel electrophoresis revealed a characteristic pattern with the ~30 nt band (III, Fig. S3E). Transmission electron microscopy demonstrated that intact CBs, with similar size and appearance as native CBs in round spermatids, were purified by anti-MVH-bead complexes (III, Fig. S3D). Furthermore, we did not observe any other evident structures attached to the antibody-bead complexes at the electron microscopy level. MS analysis of the protein content from the biological replicates of the isolated CB samples provided the final verification of reproducibility, sensitivity and specificity of the current CB isolation procedure.

Five individual mass-spectrometric analyses and tree independent replicates of RNA deep sequencing analyses of the isolated CBs were used in this study. Overall, more than hundred individual CB isolation procedures were carried out during the course of this study. Current protocol is remarkably consistent in the expected purity and yields of the isolated CB material. Moreover, with minor modifications in the protocol, CBs from juvenile mice and mouse models with defects in late spermatogenesis and altered CBs were also successfully isolated. These mouse models were the full knockouts of SAM68 (Messina et al. 2012), RNF17 and the heterozygous knockout of MIWI.

In summary, a reliable, reproducible, fast, efficient, relatively simple and inexpensive protocol for the isolation of pure fractions of CBs from mice testes, has been developed.

5.2. The RNAs of the CB (I, III)

5.2.1. The CB accumulates large quantities of piRNAs

Total testis RNA has a distinctive small-RNA band around 30 nt which corresponds to piRNAs and can be easily visualized by EtBr/SybrGold staining in polyacrylamide gel, referring to their huge abundance. Following low-speed centrifugation of the testicular cell lysate, it was found that the 30nt RNA band became enriched in the pellet fraction. Moreover, virtually all of the 30 nt RNAs in the pellet fraction co-immunoprecipitated with the CBs (III, Fig. S3E). Since the characteristic 30 nt population relates to piRNAs and the piRNA-binding MIWI and MILI localize in the CB (Aravin et al. 2007), it would appear likely that the 30 nt band which accumulates in CBs, corresponds to piRNAs. To verify this, Northern blot analysis of the RNA samples from the CB isolation procedure with a probe against piRNA-030365, was performed. The specific hybridization signal was present in the fractions of input cells and the isolated CBs, but absent in the control immunoprecipitate (I, Fig. 5C).

In situ hybridization of stage-specific squash preparations of seminiferous tubules further proved the CB localization of piRNAs. In addition to a diffuse cytoplasmic staining, a strong piRNA hybridization signal was found in the cytoplasmic granule of haploid round spermatid which was confirmed to be the CB by phase contrast microcopy. No signal was detected when using a scrambled negative control probe (I, Fig. 5D). This data demonstrates that piRNAs are present in the CB. It should be noted that the analysis with a single piRNA probe may not reflect the behavior of the whole pachytene piRNA population.

Next, several libraries from CB-associated small RNAs, from adult mouse testis were prepared and analyzed by deep sequencing. 78% of the CB small RNAs comprised of piRNAs mapping to the known pachytene piRNA clusters (Fig. 9). Pachytene piRNAs were enriched in the CBs and the miRNAs were not, when compared to round spermatids (III, Fig. 1D). Overall, the CB piRNA content resembles a MIWI-bound piRNA profile (data not shown), which is consistent with the observation that MIWI is one of the most abundant CB components (see chapter 5.3.).

From its size separation profile in gel electrophoresis, CB RNA can be broadly separated into three parts: small RNAs with a predominant ~30 nt band, a medium-sized smear between 50-300 nt and the remainder of the larger RNA molecules (Fig. 9). This represents the sum of CB-RNA profiles from all the developmental steps of round spermatids, since the input material for the CB-isolation was the whole testis. To analyze whether the CB small RNA profile alters with the progression round spermatid development, RNA was extracted from CBs which were isolated from the juvenile mice testes. These testes were beforehand confirmed to contain early round spermatids of step 1-2 (in 22 days old mice) and late round spermatids of step 7-8 (in 26 days old mice). No considerable changes in the characteristic ~30 nt band of the CB-RNA profiles of early and late round spermatids were observed (III,

Fig. S3G), which suggests that piRNAs are evenly represented in the CBs throughout all steps of round spermatid differentiation.

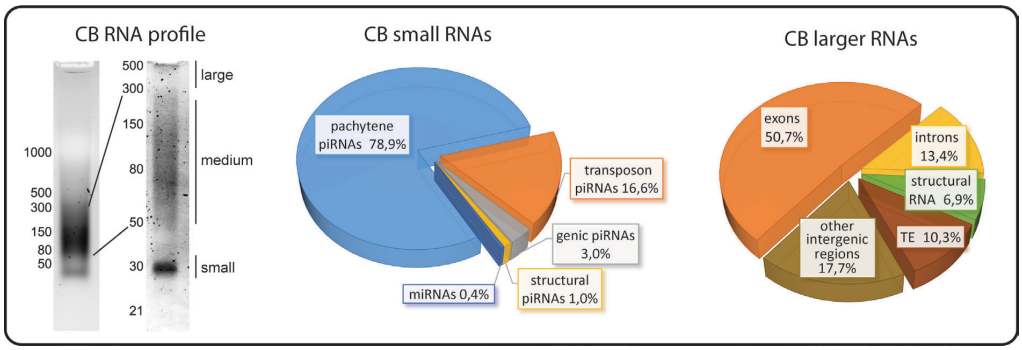

Figure 9. *The CB main RNA components. Left image: The CB RNA profile as seen in the agarose and polyacrylamide gel electrophoresis, visualized with SYBR gold. Center illustration: Small RNAs (<50 nt) in the CB. Right image: Large RNAs (>50 nt) in the CB. TE-transposable elements.*

5.2.2. mRNAs in the CB

The CB was originally hypothesized to be involved in mRNA translational regulation and storage (Kotaja, Sassone-Corsi 2007, Soderstrom, Parvinen 1976, Parvinen 2005). To analyze it, the localization of poly(A) transcripts in the CB throughout all the steps of round spermatid differentiation, was confirmed by stage-specific immunofluorescent analysis with Cy-5 labeled poly(T) probe, of the seminiferous tubule cross sections (III, Fig. S2). Since many non-coding RNAs, including transposons, are polyadenylated, this approach does not tell if and to what extent does this signal account for mRNAs.

Next the RT-PCR analysis of the total RNA, purified from immunoprecipitated CBs was used to qualitatively confirm the CB-location of several mRNAs (I, Fig. 4). One of the CB components, gonadotropin regulated testicular helicase (GRTH), is an RNA-binding protein with several known target mRNAs (Tsai-Morris et al. 2010). For PCR analyses, the primers respective to GRTH-target mRNAs such as transition protein 2 (TP2), protamine 2 (PRM2), and outer dense fiber protein 1 (ODF1) were used. All these haploid-specific mRNAs were present in the input and CB fraction samples of the protocol, but not in the rabbit IgG control immunoprecipitate. As negative controls, transcripts which are expressed in other cell types that lack CBs and MVH were used. The testicular expression of *Ets*-related molecule (ERM) is restricted to Sertoli cells (Chen et al. 2005) and CD9 is a surface marker for spermatogonia (Kanatsu-Shinohara et al. 2004). Neither of these mRNAs was detected in the pellet fraction nor in the anti-MVH immunoprecipitate of the CB-isolation procedure (I, Fig. 4). This suggests that specific mRNAs are indeed present in the CB. However, this method does not discriminate between the full length mRNA transcripts and the degradation products.

To further analyze the RNA content of the CB, several transcriptome libraries from different steps of CB purification procedure and also from round spermatids were prepared for deep sequencing. The NSR protocol that detects both poly(A) and non-poly(A) transcripts, but

excludes ribosomal RNAs (Armour et al. 2009), was used for the generation of the RNA libraries.

The deep sequenced RNA contents of CBs, total testis lysate, round spermatids and Acrosin-immunoprecipitate (negative control) were compared. Reads were mapped to 5 major categories: exons, introns, non-coding RNAs, structural RNAs, tranposable elements (TE) and other intergenic regions (Fig. 9). The most abundant transcripts in the CB originated from exonic regions of genes. Although there were some differences between the ratios of the CB and round spermatid RNA profiles, no overall enrichment for any given category was observed (III, Fig. 3A).

The compiled mRNA dataset consisted of 8058 CB-associated mRNAs with median and mean FPKM (Fragments Per Kilobase of transcript per Million mapped reads) values of 30.3 and 85.3, respectively (III, Table S5). The most abundant CB-associated mRNAs, based on their FPKM values, were transition proteins (TNP1, TNP2) and protamines (PRM1 and PRM2). TNPs and PRMs are involved in chromatin repacking into the sperm head, in elongating spermatids. The transcription of TNP2 and PRM1 is initiated around step 7 spermatids (Mali et al. 1988, Saunders et al. 1992) which suggests that the CB can accumulate large quantities of mRNAs in the later stages of its development.

To investigate whether the CB collects all round spermatid mRNAs, or is rather a place for specialized messages, the sequenced CB mRNA list was compared with the available microarray-obtained transcriptome data of round spermatids (Chalmel et al. 2007). By direct comparison of both spermatid and CB transcriptome datasets, 90.7% (6,686/7,367) of seq-CB mRNAs were found to overlap with the microarray-detected spermatid mRNAs. Concurrently, 45.2% (6,686/14,788) of mRNAs expressed in round spermatids were detected in the CB mRNA list.

Next, a list of more abundant CB mRNAs was created by removal of all the seq-CB mRNAs with FPKM values below the median, resulting 4029 mRNAs and this dataset was named CB-High. The CB-High list was compared with the microarray datasets on mRNAs which are differentially expressed in the transcriptomes of different male germ cells (spermatogonia, pachytene spermatocytes and round spermatids) (Chalmel et al. 2007). Microarray expression data are not directly comparable with RNA deep sequencing data. Nevertheless, for the set of transcripts found in the CB by sequencing, it is possible to display how they are expressed in the different testicular cells based on the Affymetrix GeneChip data. When correlating the internal ratios of the differentially expressed genesets between the trascriptomes of the spermatid and the CB, it was found that the CB-High list is significantly enriched in round spermatid genes which are differentially expressed in meiotic and post-meiotic cells. By comparison, the mRNAs of genes which are differentially expressed in somatic and mitotic cells are not enriched in the CB, as compared to the remainder of the round spermatid mRNA transcriptome (III, Fig. 6, S5). It was also noticed that many genes which are ubiquitously expressed during spermatogenesis appear to be less represented in CBs, in comparison to round spermatids. All this suggests that the CB is more enriched for transcripts which are specifically induced in spermatocytes or round spermatids, rather than for house-keeping-like genes.

5.2.3 Long, non-coding RNAs in the CB

The main classes of long RNAs present in the CB are mRNAs, repeat-derived sequences and long non-coding RNAs (lncRNAs). There is a significant amount of novel uncharacterized transcripts enriched in the CBs when compared the input cell lysate. In the current study, around 4992 novel lncRNA transcripts were mapped which were over 50 nt in length and did not overlap mRNA exons or piRNA clusters (III, Table S4). They represent a diverse population of spliced transcripts with an average of 4,5 exons per transcript (maximally 36) and with individual exon sizes from 18 to 8500 nt. The length of the processed transcripts may reach to approximately 17600 nt, with an average length of 600 nt. These lncRNA transcripts map evenly to all chromosomes except for the Y-chromosome, which appeared to be virtually unrepresented (III, Fig 4A).

The expression of six randomly selected CB novel lncRNA transcripts was analyzed by qRT-PCR. Most of which were predominantly expressed in the testis, with one exception where the highest expression was found in brain (III, Fig 4B). Then, the expression profiles of four novel lncRNAs in the enriched populations of pachytene spermatocytes, round spermatids and elongating spermatids were compared. Two transcripts were found to have a relative higher expression in haploid round spermatids, while the other two were more expressed in meiotic spermatocytes (III, Fig 4C). All the four tested transcripts were downregulated in elongating spermatids. The results suggest that at least these novel CB-associated lncRNAs are of meiotic and post-meiotic origin and function.

5.2.4 Constant accumulation of nascent RNAs in the CB

To assess if newly synthesized RNA is targeted to the CB, specific pools of pieces of seminiferous tubules from a mouse testis were microdissected. These pools represented various stages of the seminiferous epithelial cycle (XII-I, II-V, VII-VIII) and allowed the separate analyses of spermatids and CBs at every developmental step of spermiogenesis (Kotaja et al. 2004). The tubule pieces were cultured in the presence of a nucleotide analog 5-ethynyl uridine (EU) which becomes incorporated into newly synthesized RNA (Jao, Salic 2008).

After 12 hours of incubation, the nuclei of pachytene spermatocytes and round spermatids became intensively marked by EU (III, Fig. S1A). Notably, a clear signal was also found in the CBs of round spermatids at every developmental step analyzed, while the cytoplasm outside the CB was, in general, very weakly labeled by EU (III, Fig. 1A, S1B). This demonstrates that the nascent RNA localizes in the CB and that this route is not restricted to a specific developmental stage of round spermatids. However, a closer comparison revealed a slight reduction of nascent RNA-specific CB staining in the late round spermatids. A chase experiment demonstrated that the EU-labeled RNA signal was retained in the CB for at least 12 hours after the initial EU-incubation (III, Fig. 1B).

5.3. CB proteome comprises of different RNA regulation pathways (I, III, IV)

5.3.1. General analysis of the CB proteome

The full CB proteome was identified by mass spectrometric analysis of the immunopurified CBs. The results of these analyses were combined from five independent experiments to provide a reproducible list of 88 CB-associated proteins. The relative abundance of individual proteins in the CB was calculated from their Exponentially Modified Protein Abundance Index (emPAI) values, divided by the total emPAI value of the CB proteome sample. emPAI offers approximate, label-free, relative quantitation of the proteins in a mixture based on protein coverage by peptide matches in a database search result (Ishihama et al. 2005). The nine most abundant CB proteins were the DEAD box helicases DDX4/MVH, DDX25/GRTH and DDX3L, PIWI proteins PIWIL1/MIWI and PIWIL2/MILI, Tudor domain-containing proteins TDRD6 and TDRD7, poly(A)-binding protein PABP and a heat shock protein HSP72/HSPA2/HSP70-2 (Fig. 10). MVH and MIWI appear to be the two CB core components that together contribute to approximately 40 mole percent to the CB's proteome. The relative weight contributions of the individual CB components can be approximately calculated by multiplying the molecular weights of the proteins with their emPAI values. The relative weight percentages of the individual CB proteins enabled these to be grouped into four general categories, viz. Tudor-domain containing proteins, DEAD-like helicases, piRNA pathway proteins and the remainder, each of which occupy approximately 28, 25, 21 and 26 percent of the CBs mass (Fig. 10).

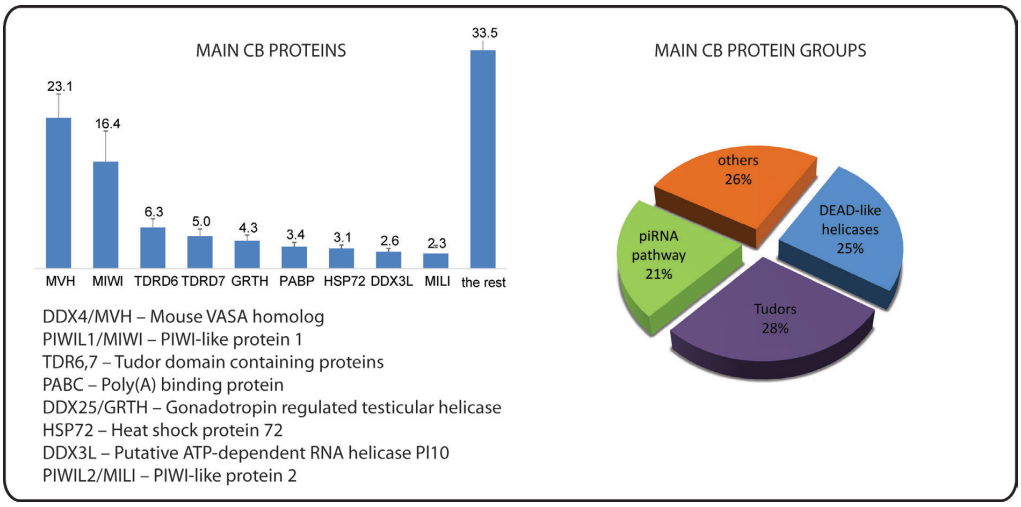

Figure 10. *The CB main protein components. Left illustration: The CB main protein components, calculated by their mole percentages. Longer names of these proteins are provided below. Right illustration: Weight percentages of the CB main protein groups.*

A compiled list of 88 CB proteins which were identified in at least 3 separate MS experiments (III, Table S1) was generated to perform functional annotation and association analyses. The most common domains in the CB proteome are nucleotide and nucleic acid binding domains, DEAD-box helicase domain and the K-homology domain - all of which are related to RNA recognition and binding (III, Fig. 2B). Another well-presented protein domain in the CB

proteome is the Tudor domain which is related to the structural organization of ribonucleoprotein (RNP) particles. Considering the well-characterized role of Tudor proteins as molecular scaffolds (Pek et al. 2012) and the disrupted CB architecture in TDRD6 and TDRD7 knockout animals (Tanaka et al. 2011, Vasileva et al. 2009), it would appear that Tudor-domain proteins serve as structural backbones of the CB. They probably form the CB protein mesh and recruit other proteins by dimethylarginine-binding (Pek et al. 2012) or through other protein-protein interactions.

Some proteins, for example the RNA-binding proteins HUR/ELAV1, have been demonstrated to localize in the CBs only transiently in very early round spermatids (Chi et al. 2011). Most of the known CB proteins seem to remain in the CB during the whole course of round spermatid development until step 7-8. To study the CB proteome composition dynamics in the early and late stage of its development, CBs from 22 and 26 days old juvenile mice were isolated and their protein content analyzed by mass spectrometry. Phase contrast microscopy of the pieces of seminiferous tubules before the CB isolation confirmed that the test samples contained step 1-2 and step 7-8 round spermatids, respectively. The MS and immunostaining results confirmed that the general profile of the most abundant CB proteins remains the same during round spermatid development (data not shown).

Functional annotation clustering of CB proteome resulted in 29 groups of annotations. As expected, the groups with the highest enrichment score can be generalized with the keywords such as: mRNA metabolic processes, post-transcriptional regulation of gene expression, Tudor subgroup, helicases, reproductive cellular processes, chromatoid body and translation regulation.

Further, a protein interaction analysis tool STRING (Franceschini et al. 2013) was used to visualize functional protein association networks by known and predicted protein-protein interactions of the CB proteome. Several pathways related to RNA processing were clearly highlighted, for example the piRNA pathway, pre-mRNA regulation, exon-junction complexes and nonsense mediated decay (Fig. 11).

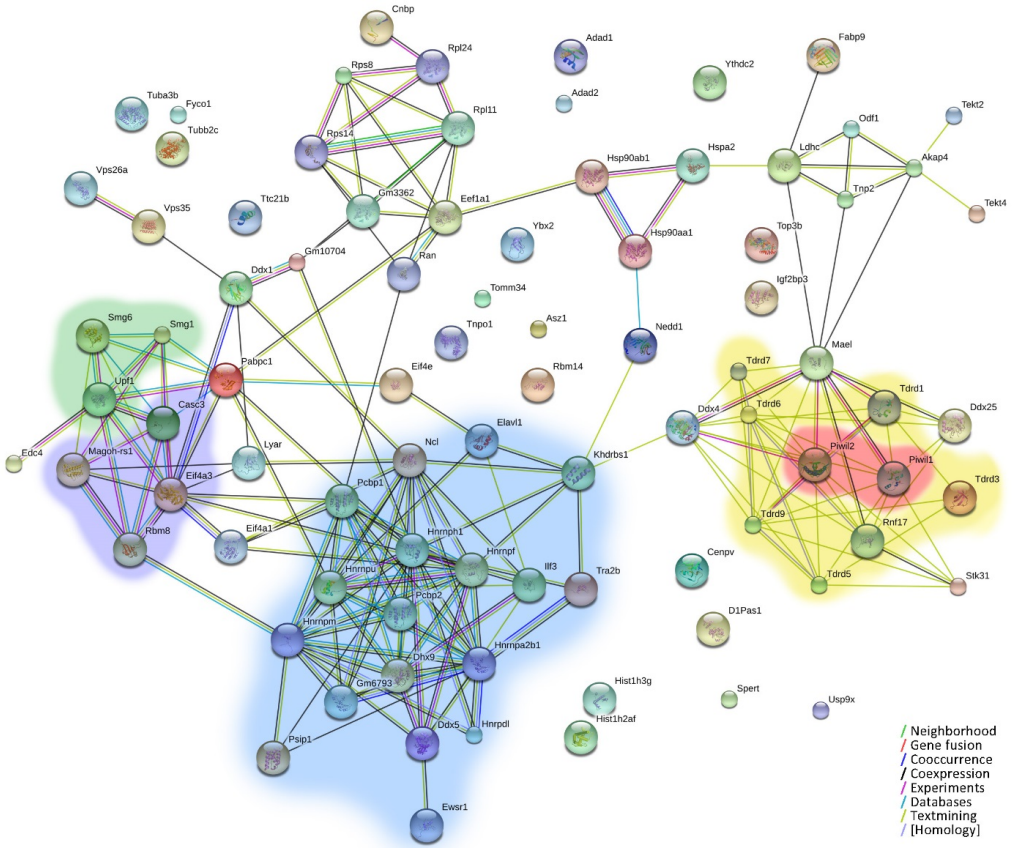

Figure 11. *Functional association network CB proteins from known and predicted protein-protein interactions using STRING software (Franceschini et al. 2013). Different line colors represent the types of evidence for the association. Several pathways and interactions are highlighted: green- NMD; violet- EJC; blue- proteins associated with mRNA maturation; red- piRNA pathway; yellow- Tudor proteins.*

5.3.2. Proteins involved in pre-mRNA processing are targeted to the CB

A large proportion of proteins in the CB are associated with nuclear mRNA processing of which many are heterogeneous nuclear ribonucleoproteins (hnRNPs) and poly(C) or poly(A)-binding proteins, that shuttle between the nucleus and cytoplasm (Dreyfuss et al. 1993, Han et al. 2010, Gorgoni, Gray 2004). There are several CB proteins involved in the regulation of splicing, such as hnRNPA2/B1, DDX5, hnRNPA3, PABP, hnRNPM, TRA2B, PCBP1, hnRNPH, hnRNPF and ELAVL1. However, small nuclear ribonucleoproteins (snRNPs) which are involved in splicing event itself are not localized in the CB. These results would suggest that mRNAs are targeted to the CB after splicing.

Exon-junction-complex (EJC) proteins are clearly present in the CB. EJC is assembled onto the exon borders of the mRNA upon splicing and travels with the mRNP to the cytoplasm. All the EJC core proteins, MAGOH, RBM8A/Y14, EIF4A3 and CASC3 were identified as CB components by MS analysis (III, Fig. 5A). The CB localization of EIF4A3 and RBM8A was further confirmed by

immunofluorescence analysis and/or by immunoblotting of the isolated CB extracts (III, Fig. S4A-B). Moreover, EIF4A3 profiling showed that it was enriched in the CBs of round spermatids during all steps of round spermatid differentiation (III, Fig. S4C).

The pioneer round of translation uses newly synthesized mRNA which is bound by 5' cap-binding complex CBC, nuclear poly(A)-binding protein N1 (PABPN1) and the cytoplasmic poly(A)-binding protein C1 (PABPC1) at the poly(A) tail, and the EJCs deposited close to the splice sites. Translation induces extensive rearrangements of the CBC-bound mRNPs. EJCs are removed by translating ribosomes after the first successful round of translation. Further, other mRNA binding proteins, for example the remaining hnRNPs, are then replaced by the repertoire of translation-associated proteins. By that time, the exchange of PABPN with PABPC is also completed. The CBC is replaced by the cap-binding EIF4E, but importantly, in a translation-independent manner (Rufener, Muhlemann 2013). EIF4E interacts with a scaffolding protein EIF4G, which in turn binds to PABPC1, leading to mRNA pseudo-circularization- a conformation that allows efficient EIF4E-mediated bulk translation (Okada-Katsuhata et al. 2012, Chang et al. 2007).

The presence of nascent mRNA shuttling proteins and the EJC in the CB conjectures that non-translated mRNAs are targeted to the CB. However, there are some notable exceptions. CBC proteins CBP20 and CBP80 were not found in the MS list of the CB proteins (III, Table S1). Instead, EIF4E- a translation-associated replacement of CBC, was located in the CB proteome, however without EIF4G or other core members of the ribosome recruitment complex. PABPN was not detectable in the CB MS list either but instead, several PABPCs were located there. Assessed together, these observations suggest a special transition state of mRNAs in the CB.

5.3.3. Nonsense mediated decay machinery is localized in the CB

Interestingly, key elements of nonsense mediated decay (NMD) machinery were also located in the CB (III, Fig. 5A). NMD is considered both a process of degrading truncated mRNAs and also a means to regulate normal transcripts' expression (Schweingruber et al. 2013). Besides the EJC which can trigger NMD, the CB also contains a set of the core NMD components: UPF1 is the key mediator which is recruited to mRNAs upon translational termination that destines this mRNA to degradation; SMG1 kinase activates UPF1 by phosphorylation; and SMG6 is an endonuclease responsible for the slicing of target mRNA (Chang et al. 2007) (III, Fig. 5B). Notably, the presence of the activated form of UPF1 in the CB was confirmed by using an antibody which recognizes phosphorylated serine at the C-terminus of UPF1. Since the diversity of NMD targets and mechanisms is wide, the presence of these current NMD components provides a potentially powerful mechanism for the regulation of a broad range of transcripts in the CB. This is further emphasized by the fact that besides MIWI- a piRNA-driven AGO slicer, SMG6 is the only known nuclease in the CB.

5.3.4. piRNA machinery occupies the majority of the CB

MIWI is the second-most abundant CB protein, to account for almost one fifth of its total calculated mass. MIWI is the germline-specific AGO engine responsible for piRNA-related

gene silencing (Cenik, Zamore 2011). Notably, also piRNAs that drive MIWI to its RNA targets are highly concentrated in the CB. Moreover, it was found that the whole set of piRNA function related proteins are concentrated in the CB, making it by far the most pronounced molecular pathway in the CB and suggesting that the CB is a platform of pachytene piRNA biogenesis and/or action. According to our analysis, the piRNA-binding PIWI proteins MIWI and MILI and the Tudor-domain containing scaffolding proteins especially TDRD6, TDRD7 are among the most prominent CB components (Fig. 10). Several other proteins that have previously been associated with the piRNA pathways on the basis of their molecular interactions or phenotypic characteristics of the knockout mice models, were also found to be CB components, for example, MAEL and MOV10L (Aravin et al. 2009, Frost et al. 2010, Zheng, Wang 2012).

5.3.5. SAM68 localizes transiently in very early CBs

SAM68 binds a specific set of mRNAs and, in addition to other functions, is likely responsible for the crucial regulation of gene expression during the post-meiotic transition stage (Paronetto et al. 2009). Careful analyses of SAM68 revealed its rare and peculiar stage-specific localization pattern, to demonstrate its transient localization in the CB during the meiotic divisions and in early post-meiotic cells. During spermatogenesis, the second meiotic division and the meiotic-haploid transition state (stages XII-I) is short-lived and occurs within the same cross-section of a seminiferous epithelium. SAM68 is predominantly a nuclear protein, but it localizes in perinuclear granules in secondary spermatocytes and early round spermatids (IV, Fig. 1A-B). Immunofluorescent co-staining with the known CB components MVH and MILI suggest that these granules will form the CB in early round spermatids (IV, Fig. 2A-B). Furthermore, western blot analysis of the enriched secondary spermatid fraction revealed that MVH was co-immunoprecipitated with anti-SAM68 antibody, but not with control IgGs (IV, Fig. 3A). Importantly, this co-immunoprecipitation was not detected in primary spermatocytes or round spermatids (IV, Fig. 3B), which indicates the specific interaction of these proteins in the dividing meiotic germ cells.

5.3.6. SAM68 is not required for CB formation

To further investigate the role of SAM68 in the CB, the effect of SAM68 ablation on the CB structure was analyzed. Transmission electron microscopy analysis of the testes sections of *Sam68* KO mice revealed some clear structural CB differences as compared to wild-type CBs (IV, Fig 4 A-F). However, this phenotype was not persistent as some CBs of the *Sam68* -/- late round spermatids appeared normal-looking. Furthermore, the CBs in early round spermatids (steps I-III) appeared to be overall unaffected.

To test whether SAM68 is responsible for the mislocation of any CB structural components, CBs from the *Sam68* knockout mice testes were isolated and subjected to MS analysis, together with a wild-type positive control. All the main protein components of the CB in the SAM68 KO mice had remained unchanged when compared to the wild-type CB. Furthermore, the internal ratios of the most abundant CB proteins also remained relatively unaltered (IV, Fig. 6E). This infers that SAM68 is not required for CB formation.

All the piRNA pathway components in the SAM68 KO CBs were also detected. Conversely, it was further determined that the ablation of MIWI did not impair the localization of SAM68 in the CB either (IV, Fig. 5B). This suggests that SAM68 is not involved in the piRNA pathway.

5.3.7. Other specific mRNA binding proteins in the CB

The majority of CB-associated proteins are RNA-binding and represent different molecular pathways in addition to pre-mRNA processing (III, Table S1; Fig. 11). The most abundant CB protein, the RNA DEAD-box helicase MVH, was found to account for almost a quarter of all CB proteins by its mole percentage and almost one fifth by its total mass. Other CB-associated helicases are involved in various processes, including splicing (DDX5, EIF4A3), translation (EIF4A1), NMD (UPF1) and post-transcriptional silencing (TDRD9).

TENR/ADAD1 is another notable RNA binding protein which is related to the family of RNA editing enzymes, adenosine deaminases (ADARs). The expression of TENR is highly restricted to spermatocytes and spermatids and mostly considered nuclear (Connolly et al. 2005). ADAD1 is a prominent CB member which suggests its involvement in some specific CB function. Of note, the CB contains also another testis-specific RNA processing protein-TENRL/ADAD2. Both ADADs contain the double-stranded RNA binding motif and the adenosine-deaminase (editase) domain. Based on the suggested functional data on ADAD1, it may likely contribute to the wide-scale mRNA editing or increased translational inhibition, in the CB. Moreover, direct interactions between ADARs with RNAi and mRNA surveillance pathways (Agranat et al. 2008, Ota et al. 2013) would suggest an even wider cross-pathway role for TENR in the CB related mRNA post-transcriptional modification.

5.4. CB *in vitro* RNA processing assay

The presence of a high quantity of RNAs and proteins which represent different RNA metabolic pathways in the CB, raised the question if the CB can process RNAs. To address this issue, CB preparation was optimized to establish a CB *in vitro* processing assay. Through usage of random radiolabelled RNA substrates, the immunopurified CBs were indeed found to process RNA *in vitro* into a smear of smaller products in ATP-enhanced manner, while the negative control samples left the substrate RNA intact (Fig. 12). Heating the CBs at 70 °C, 45 min. and Ribolock™ inhibited this activity. The CBs own general RNA and protein profiles appeared to remain unaffected after assay (data not shown).

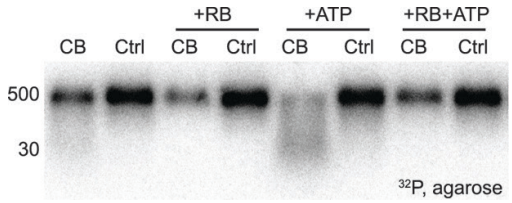

Figure 12. *Isolated CBs can cut RNA substrate in ATP-enhanced manner. CB- CB-Dynabead complexes; Ctrl- negative control IgG-Dynabead complexes; RB- Ribolock.*

6. DISCUSSION

6.1. The isolation of CBs

6.1.1. Why did we want to isolate the CBs?

The outcome of germ cell differentiation determines our survival as a species and an infertile organism is considered a "genetic corpse". Infertility in being constantly rising in humans and results nowadays with increasing number of couples having a problem (Skakkebaek et al. 2006). Yet still the mechanisms underlying spermatogenesis are poorly understood. Male germ cell differentiation is difficult to study yet crucial to understand, not only from a medical, but perhaps even more importantly, from a basic research' objective. The majority of peculiarities involved in male germ cell differentiation are, at least partly, associated with spermatids.

The most visible characteristic of a round spermatid is its large perinuclear granule. Because of its size, the CB was already discovered in the end of 19th century- even before Köhler illumination was developed. It is favorable to assume that this unique and prominent structure has an important role in the development of round spermatids and consequently in spermatogenesis. Interest in the CB as an active regulation center developed after electron microscopy studies revealed its association with the nuclear envelope and a multitude of different vesicles. The description of its active and non-random movements in the cytoplasm and communications with the Golgi complex, ER and even neighboring spermatids (Ventela et al. 2003) clearly suggested the CB's role to somehow function as a regulation center. This conjecture was further supported by localization studies of RNA and specific proteins in the CB. The vital role of the CB for round spermatid differentiation was demonstrated by the complete infertility of several knockout animal models lacking the normal shaped CB structure (Kotaja, Sassone-Corsi 2007). Conclusively, it has been argued that the CB might be involved with the central regulation of post-transcriptional gene expression in round spermatids. Comprehension of the functions of the CB will help to unravel the mechanisms of post-meiotic male germ cell development. Furthermore, CB research can be extrapolated to explain questions on the general formation and movement of RNP granules, the trafficking, sorting, storing and processing mechanisms of RNA, and the interconnections between different molecular pathways.

6.1.2. The strategy of a successful CB isolation

Understandably, the scientific demand for a successful CB isolation protocol has been prevalent since the early studies of the CB. There has been at least one published plus many unpublished attempts of the CB isolation by gradient centrifugation, with limited success (Figueroa, Burzio 1998). Being a powerful method for polysomal fractionation, gradient centrifugation lacks the necessary discriminative power to extract pure CBs. Other methods, such as flow cytometry, laser-capture microscopy and *in situ* MALDI-TOF MS regional analyses, are too unspecific and limited to obtain reliable complex data. Immunoaffinity-

based isolation techniques provides a solution for isolating the CBs. Immunoprecipitation (IP) is quick, reliable, offers a wide range of possibilities and is nowadays extensively used for capturing DNA, RNA and proteins and their associates. The most important parameter of successful immunoprecipitation is the specificity and sensitivity of the used antibody.

The current CB isolation protocol can be divided into two parts- first, the enrichment of CBs from the cell lysate, followed by the IP of the concentrated CBs. The balance between the sensitivity and specificity of the first, CBs enrichment stage can be tuned with the cross-linking, sonication and centrifugation steps. Cross-linking of the testicular cells before their lysis is a crucial step for the successful isolation of CBs, since without cross-linking, the CBs will become fragmented due to the detergents and sonication in the cell lysis. However, without detergents and sonication, membraneous structures are not dissociated and may co-immunopurify with the CBs. We cannot exclude the possibility that some non-CB matter becomes associated to the CBs during cross-linking. It is also possible that some CB components become lost during the sonication and low-speed centrifugation steps. However, overall data suggests that the purified CBs are intact, biochemically pure and suitable for sensitive downstream analyses.

MVH presents a good target for CB IP. It is a highly abundant CB protein which is only expressed in the late meiotic and post-meiotic cells, in testis (Tanaka et al. 2000, Kotaja, Sassone-Corsi 2007). The in-house rabbit polyclonal anti-MVH antibody produces a single band what corresponds to MVH in the western blot analysis of testicular cell lysates. The antibody-antigen interaction is very sensitive and remarkably strong. These are the criteria of a good antibody in the design of an immunoprecipitation (IP) protocol. CBs can also be isolated with a commercial anti-MVH antibody (Abcam, ab13840). We have used the MVH-IP protocol to isolate CBs from different mouse-models with a spermiogenic defect. One of these is the SAM68 knockout mouse-line that is discussed further below.

For the optimization of CB isolation protocol, it is essential to monitor its steps. As in the evaluation of the CB isolation, the different aspects of purity, integrity and quantity arise. Western blot analyses provide the best overview of the efficiency of the CB's enrichment and isolation steps and require only small quantities of samples. Immunostaining of the drying-down slide samples allows the monitoring of the integrity and yields of CBs during the sample's enrichment steps. RNA gel electrophoresis is not informative to monitor cell lysis and CB enrichment steps, as a result of the interfering cellular RNA in samples. Further, this analysis requires ten times more material than the previous two methods. However, the RNA profile of purified CBs is characteristic and should thus be analyzed before any further application with CB-RNA. The optimal method to study the integrity of the isolated CBs and the sample's purity from other particles, is by transmission electron microscopy. In depth analysis of the protein and RNA content from the biological replicates of the isolated CB samples by mass-spectrometry and deep sequencing provide the final piece of information by which to evaluate the sensitivity and specificity of the CB isolation procedure.

6.1.3. Limitations of the CB isolation protocol

Two significant limitations must be considered when studying CBs which are isolated with the current protocol - the isolation of a mix, yet an incomplete mix of the CBs.

Firstly, the input material for the CB isolation is adult testis which contains round spermatids from all developmental stages. Thereby, a mix of CBs from different developmental stages will become isolated. By which those CB-associated proteins and RNAs detected represent the total sum of CB profiles, but not the actual composition and rations of a CB at some specific developmental step. It has to be taken into account that some of the CBs components are there only transiently. For example, SAM68 locates only briefly in the very early CBs, while some mRNAs (encoding for transition proteins and protamines) accumulate only in late CBs.

Secondly, after stages 8 the round spermatid shifts its developmental program and becomes an elongating spermatid. Concurrently the CB begins to diminish, undergo functional transformation and, among other changes, become depleted of its major component, MVH. Since the current CB isolation procedure relies on MVH, CBs from the later stages of elongating spermatids are not isolated. Thereby to study the CB at the final stages of its development, another target, specific to late CBs, must be used for immunoprecipitation. However, the fact that the transformed CBs are not isolated with the current protocol can also be considered an advantage, since the analysis of even more complex mix of CBs might be very challenging.

When comparing the early (step 1-2) and late (step 7-8) CBs, immunoprecipitated from juvenile mice testes, it was concluded that the most abundant CB proteins and their ratios remains the same. Further, no quantitative changes in piRNA population of early and late CBs were detected. This suggests that the general CB composition is stable throughout round spermatid development. It has to be noted that although widely used for analyses, the first wave of spermatogenesis in juvenile mice is different (Yoshida et al. 2006) and may not represent the actual mechanisms in adult testis.

6.1.4. Future perspectives in CB research

The ultimate experiment to characterize the dynamics of the CB composition would require the creation of separate libraries of immunopurified CBs from all developmental stages, by use of microdissected pools of seminiferous tubules as inputs. The comparative transcriptome and proteome analysis of these CB fractions would provide additional information on the dynamics of CB composition. Further, prior to the CB isolation, these tubule pieces could also be incubated separately with ethynyl-uridine and L-azidohomoalanine to mark the newly synthesized RNA and proteins. This would add yet another dimension to CB analysis results, allowing the discrimination between the passive and shuttling components of the CB during its development.

The major challenge of this approach would to downscale the CB isolation protocol, since the input sample levels would be low. Current experience and results from already

performed tests suggest that the CB isolation from pooled sets of seminiferous tubule pieces will yield enough good quality material for the subsequent analyses.

RNA sequencing for *de novo* discoveries clearly has room to evolve. Limitations provided by discriminating library construction strategies, data processing pipelines and the definitions of targets can leave us blind for potentially interesting results. This is especially relevant when studying novel or unknown phenomena. As an example, RNA editing events and circular RNAs are both mechanisms that can accumulate considerable meaningful variability but are currently generally overlooked.

To further advance the proteome analysis of CBs, but also of other multiprotein complexes and ribonucleoprotein granules, cross-linking/mass spectrometry with protein modelling should be considered (Rappsilber 2011). The MS analysis of heteropeptides generated from the cross-linked protein complexes would identify amino acid pairs that are positioned in close proximity to each other. This can provide qualitative and also quantitative data on the structure of proteins and their complexes. The MS analysis of cross-linked protein complexes could certainly contribute to resolving the structural composition of CBs and also yield other fascinating potentials in the proteome research. Similarly, the associations between CB proteins and RNA could be further studied by use of protocols for covalent UV cross-linking of RNA binding proteins to RNA (Castello et al. 2012). These methods could attach CB proteins to their respective RNA targets, making both counterparts available for separate analyses.

6.2. The CB as an RNA processing center

CB is packed with different RNAs and RNA-binding proteins. There is also a general flux of nascent RNA to the CB. The assessment of CB-associated RNA and protein components conjectures that the CB could be a platform where different mechanisms of RNA binding, processing and post-transcriptional gene expression control, converge. There are several known molecular pathways clearly represented in the CB which will be discussed further below.

6.2.1. pre-mRNA-binding proteins tell their story

A large group of CB-associated proteins are related to the processing of pre-mRNAs, especially splicing and alternative splicing, but also to RNA packing and the nuclear export of mRNAs and non-coding RNAs. Numerous hnRNPs and all the core members of the exon-junction complex (EJC, a multiprotein complex that is deposited at exon–exon junctions during splicing) are located in the CB. Significantly, these proteins are generally considered to be the nucleocytoplasmic shuttling components of mature mRNAs. This indicates that at least some of the mature mRNAs localize in the CB prior to their (possible) translation. It is unlikely that the shuttling proteins would accumulate to the CB after being stripped from their mRNAs.

No doubt, the CB exhibits a peculiar repertoire of RNA-binding proteins. These can reflect the status of CB-associated mRNAs. In addition to the evident CB components, such as the EJC, it is also notable to recognize proteins which are expected yet clearly absent from the CBs. For example the CBC components CBP80 and CBP20 were not detected in the CBs by the MS analysis. CBC is the prominent component of spliced mRNPs and indicates, together with EJCs and the hnRNPs, the newly synthesized non-translated mRNAs. Notably, instead of the CBC, the CB accumulates translation-associated cap-binding EIF4E. This may be explained by the reported evidence of translation-independent exchange of CBC with EIF4E at the cap (Rufener, Muhlemann 2013, Sato, Maquat 2009).

The CB is also depleted from the nuclear PABPNs which are characteristic for untranslated mRNAs, and contains instead the PABPCs, which are more related to translation. While the CBC-EIF4E replacement is not dependent on translation (Rufener, Muhlemann 2013), translation does promote the replacement of PABPN1 by PABPC1 (Sato, Maquat 2009). The CB proteome does not contain any members of the ribosome recruitment complex that can bind to EIF4E nor the ribosomes to promote translation. It is possible that EIF4E is transported to the CB by a set of mRNAs which have been translated, but which would still not explain the simultaneous presence of pre-translationally-bound mRNA binding proteins (EJC, hnRNPs) and the complete lack of the prominent CBC or PABPNs, in the CB.

In addition to the ubiquitous PABPC1 in the CB MS proteome list, an equal amount of the testis-specific PABPC2 was also detected. Despite large homology, PABPC2 has been demonstrated to be absent from actively translating polyribosomes, which suggests its function in translational repression (Kimura et al. 2009). Of note, immunostainings of PABPC2 has demonstrated its accumulation in the CB while the PABPC1 has been reported be absent from the CB (Kimura et al. 2009). Further analysis is required to assess this potentially meaningful data.

The localization of pre-mRNA processing proteins in the CB may be partly contributed to the CB-associated long non-coding (lnc) RNAs- as they are spliced and not translated. However, much about lncRNAs still remains obscure and becomes increasingly surprising as their mechanisms and functions are revealed. The possible functions of lncRNAs in round spermatids will be discussed below.

In conclusion, the presence and the specific absence of several mRNA-binding proteins in the CB can refer to some alternative regulatory state of the CB mRNAs.

6.2.2. Alternative routes of mRNAs to the CB- the SAM68 story

CB protein analysis suggests different mechanisms of mRNA transport to the CB. For example, a prominent CB component, gonadotropin regulated testicular RNA helicase (GRTH/Ddx25), transports relevant mRNAs from the nucleus to the CB, presumably for storage to await translation. Downstream, GRTH is also associated with polyribosomes to regulate target gene translation (Tsai-Morris et al. 2010). HUR/ELAVL1 is another CB-associated post-transcriptional mRNA regulator, which can stabilize its target mRNAs by binding to the AU-rich elements (AREs) in their 3UTRs (Doller et al. 2008, Fan, Steitz 1998).

Similarly to HUR, SAM68 is also an ARE-binding nucleocytoplasmic protein with an interesting CB-localization pattern, as nuclear yet occupying the CB and cytoplasm for a very short period during the meiotic-haploid transition (described in 5.3.5). During this period SAM68 seems to reside almost exclusively in the CB. What does this peculiar dynamic translocation pattern indicate? The first task of a newly formed round spermatid is a successful exit from meiosis and the entry to the cell's developmental program towards metamorphosis in becoming a spermatid. SAM68 may mediate this huge transition by the collection and relocation of RNA transcripts that need to be quickly compartmentalized, once the cell's nuclear membrane has re-associated. In addition to implication with alternative splicing, SAM68 is also an ARE-binding protein, so it may participate in the CB-targeted transport of some specific mRNAs. Comparison of the wild-type and SAM68 knockout CB transcriptomes is required to reveal the possible RNAs which are targeted to the CB by SAM68. SAM68 translocation patterns in secondary spermatocytes and round spermatids are precisely timed and transient, which renders it impossible to monitor precisely on fixes tissue sections. A GFP-tagged SAM68 mouse-line and the live visualization of squash prepared stage-specific seminiferous tubules would shed light on to the peculiar localization dynamics of SAM68 and its targets. While the general proteome profile of the CB seems to be stable, the temporary passengers e.g. SAM68 and others, reflect its association with dynamic functions, in round spermatids.

Although the functional role of SAM68 in the CB remains unclear, this study has demonstrated that SAM68 is not required for the assembly and structure of the CB. CBs in the early round spermatids in SAM68 knockout mouse testis appeared normal in electron microscopy. In late round spermatids of the knockout testis, smaller and/or incorrectly assembled CBs were often observed- likely as a result of indirect consequences that co-occur with massive haploid germ cell apoptosis in this mouse-line (Paronetto et al. 2009). In addition, the MS analysis of the CBs from SAM68 KO mice revealed that all the major protein constituents were similarly recruited to the CB, as compared to the wild-type. Currently, it cannot be ruled out that instead of SAM68's involvement in some CB-associated function, the CB itself could actually be required for SAM68 function. For example, the CB could serve as an entry site for nuclear import of SAM68, after the formation of haploid cells and the reestablishment of nuclear compartmentalization.

This study has also demonstrated that the ablation of SAM68 does not affect the localization of piRNA machinery and piRNA population in the CB. All of this further supports the hypothesis, that CB is a docking station for divergent molecular pathways and mechanisms which can function independently from one another.

6.2.3. Nonsense mediated decay components in the CB

NMD is an RNA surveillance pathway that is recruited if the translation ends upstream from the last exon-exon junction or other characteristics such as long 3'UTR. NMD is not yet fully understood but a simplified model suggests that an NMD kinase SMG1 phosphorylates UPF1, which then becomes a binding platform for an endonuclease SMG6 and/or a SMG5-SMG7 complex which leads to the direct or indirect mRNA degradation, respectively (Chang et al. 2007, Okada-Katsuhata et al. 2012) (Fig. 13).

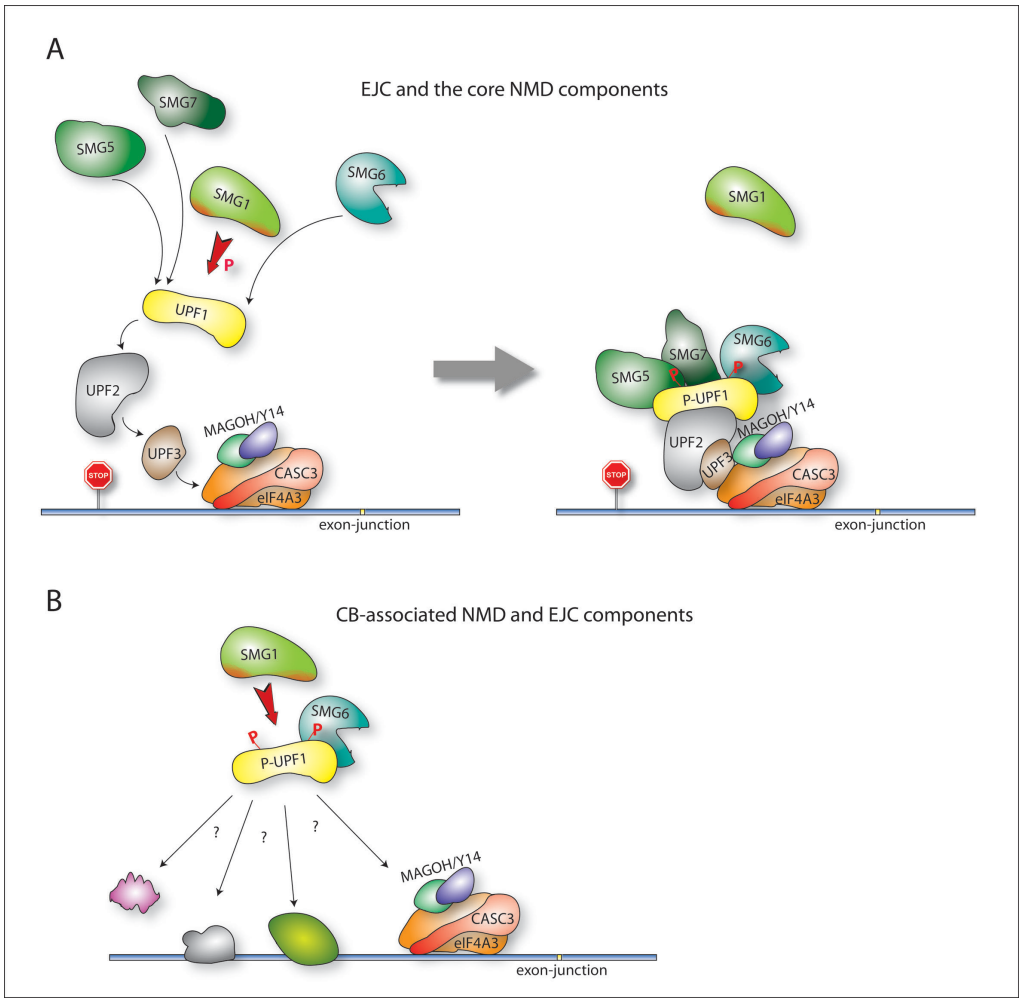

Figure 13. *NMD pathway components. A- Partial model of the classical NMD. Upon encountering a premature stop codon upstream of an exon junction, UPF1 binds to EJC (MAGOH, Y14, CASC3, eIF4A3) via UPF2 and UPF3. SMG1 phosphorylates UPF1 which becomes a binding platform for SMG5, SMG7 and SMG6 leading to mRNA degradation. B- NMD core components in the CB. A putative model of a SMG1-UPF1-SMG6 triad which forms a translation-independent RNA decay pathway.*

We have detected the core EJC proteins (CASC3/MLN51/BTZ, Y14, MAGOH, eIF4A3) in the CB. The repertoire of NMD components in the CB proteins consists of phospho-UPF1, SMG1 and SMG6. It is important to note that we detected the activated form of UPF1 in the CB (III, Fig. 5B). The complete set of the core NMD pathway would also involve the mediating factors UPF2 and UPF3A/B which link the NMD to EJC. However, UPF2 does not appear to be essential, as demonstrated by the presence of an alternative UPF2-independent route to NMD (Chang et al. 2007, Gehring et al. 2005). Notably, another very recent study reports the presence of UPF2- and UPF3- independent branches of NMD, demonstrating that SMG1-mediated UPF1 phosphorylation (and the subsequent dephosphorylation) is the only critical

requirement for both, EJC-enhanced and the alternative, EJC-independent NMDs (Metze et al. 2013).

The CBs lack CBC, but contain EIF4E. NMD was believed to operate only on untranslated mRNAs with a 5' cap-bound CBC. However again, it has been recently demonstrated that NMD can actually target both the CBC- and EIF4E-bound mRNAs, and is not restricted to the pioneering round of "surveillance" translation (Rufener, Muhlemann 2013). Still, the activation of NMD requires ribosomal proofreading which recognizes an EJC positioned downstream from a PTC on mRNA. The general lack of ribosomes in the CB suggests that the recognition of NMD-competent RNA transcripts occurs elsewhere. In addition to the polyribosomes, also a nuclear scanning mechanism for the PTC recognition has been suggested (Buhler et al. 2002).

Considering the above, it is possible that the CB harbors a translation-independent RNA degradation pathway which consists of the following key components- recognizer (some UFP1-binding RBP), the mediator (UPF1), the switcher (SMG1) and the saw (SMG6) (Fig. 13). An example of one possible "recognizer" of that pathway in the CB is a testis-specific putative RNA editing enzyme ADAD1 which occupies about 2% of the whole CB's proteome. It has been demonstrated that the ubiquitous ADAR1, which shares similar RNA binding and catalytic domains with ADAD1, binds directly with UPF1 (Agranat et al. 2008).

Round spermatids have the highest transcriptome complexity of all the cells in the body (Soumillon et al. 2013) and the CB appears to be connected with the general flux of nascent RNAs in round spermatids. In addition to mRNAs, the CBs contain a large amount of long non-coding RNAs with unknown functions. These are suggested to partly originate from leaky transcription in meiosis (Soumillon et al. 2013). Hence, one of the CB's roles in early round spermatids, may be also to collect such potentially harmful transcripts after meiosis for the subsequent degradation. The potential involvement of the SMG1-UPF1-SMG6 triad in some RNA control mechanism other than NMD could provide and answer to the evident regulatory requirement for a broad range of transcripts in round spermatids. This hypothesis is particularly favored since besides MIWI, SMG6 is the only known nuclease in the CB.

6.2.4. piRNA pathway in the CB

piRNA-mediated machinery is the best represented and most pronounced pathway in the CB, at both the protein and RNA level. In addition to the mature pachytene piRNAs, we have also located several long transcripts in the CB which could serve as pre-piRNAs. However, as the majority of the pachytene piRNA population arises already in meiotic cells before the formation of CBs, it is more likely that the CB is the location of piRNA function, rather than a place of primary piRNA biogenesis. This is also supported by the absence of HEN1, a methyltransferase required for piRNA maturation (Kirino, Mourelatos 2007) and PLD6- a nuclease in the primary piRNA biogenesis (Ipsaro et al. 2012), in the CB.

Non-coding RNAs guiding protein complexes to their nucleic acid targets is a wonderful concept of dynamic and adaptive gene regulation, with divergent small non-coding RNAs which drive generic AGO engines. The function of the MIWI-piRNA complexes is still being

studied. In addition to silencing transposons, piRNAs have been directly associated with the creation and maintenance of epigenetic memory (Shirayama et al. 2012, Ashe et al. 2012). piRNA-induced epigenetic inheritance is a fascinating prospect in future studies.

While the functions of pachytene piRNAs remain elusive, their remarkable accumulation in the CB throughout the whole course of round spermatid development suggests that these functions are at least partly commenced in the CB.

6.2.5. A general flux of RNAs in the CB

5-ethynyluridine (EU) is rapidly incorporated into cells, labeling the newly synthesized, bulk RNA transcripts (Jao, Salic 2008). An EU-culture analysis of the seminiferous tubules demonstrated a major flux on newly synthesized RNA to the CB (III, Fig. 1A). EU-RNA signals were detected equally in the nuclei and the CBs of round spermatid, while the cytoplasm was very weakly labeled. Usually, in standard cell culture experiments with EU, a strong nuclear and faint cytoplasmic signal, has been reported (Jao, Salic 2008), which renders the prominent CB-labeling in round spermatids an unprecedented phenomenon. Notably, the CB-targeted flux of nascent RNAs was observed throughout the entire program of round spermatid development.

The hypothesis of an active RNA export from the nucleus to the CB is supported by the concentration of nuclear pore complexes in the nuclear membrane which lays in close proximity to the CB (Parvinen 2005). Furthermore, the presence of a high number of CB-associated pre-mRNA-binding proteins and the EJC also infer that the CB receives nascent RNAs.

Interestingly, most of the EU-labeled RNAs appear to remain in the CB, as revealed from a 12-hour chase experiment (III, Fig. 1B). Unfortunately, a longer experiment was not possible due to imitations of tubule culture viability. However, similar results were obtained, when an RNA polymerase inhibitor actinomycin D was added to the EU-cultured seminiferous tubules (data not shown). This indicates that the CB-targeted RNA is at least partly retained in the CB.

If the CB only accumulated RNAs, it would increase in size, yet the general size of a CB seems to be constant during most stages of round spermatid development. We also did not notice any changes in the general gel electrophoresis profile of early and late CB-RNAs. In addition, *in situ* hybridization experiments of testis sections with poly(T) probe labeled all CBs with equal intensity, implying that the CB transcriptome does not undergo major changes during the development. This indicates, in turn, that the constant import of CB-directed RNA should be coupled with the respective export. The concentration of nascent RNA staining in the CB is considerably and consistently higher in the CB than in the rest of the cytoplasm. However, the dilution factor between the CB and the cytoplasm is over 200-fold, which renders it difficult to measure the overall quantities of CB-bound and “free” cytoplasmic RNAs.

Overall, the acquired data suggest that 1) the CB accumulates large quantities of bulk nascent RNA, 2) the CB receives newly synthesized RNA constantly, 3) of which most is, at

least temporarily, retained in the CB, and 4) there is a parallel import-export flux of RNAs through the CB during the course of round spermatid development.

6.2.6. Can CBs store mRNA?

The previous chapter raises a question if the CB could function as storage for mRNAs. Indeed, whether CB sequesters dormant mRNAs, protects and renders them temporarily inaccessible to the translational apparatus, has been under an emotional debate for decades (Parvinen 2005, Kleene, Cullinane 2011, Soderstrom, Parvinen 1976). There is an obvious necessity to delay translation during spermiogenesis, since many proteins are still required, when transcription has already been inhibited, due to chromatin packing. The precise timing of translation is critical, as indicated by findings that the premature translation of protamine mRNA arrests spermatid differentiation, presumably by disrupting the structure of chromatin (Lee et al. 1995). Undoubtedly, there must be appropriate mechanisms in spermatids to both protect and store mRNAs for delayed translation.

It is generally accepted that CBs contain large quantities of RNA. Gel electrophoresis of isolated CB RNAs demonstrates the presence of different RNA populations in the CB. EU-culture experiments, poly(T) *in situ* hybridization analysis and immunostainings, conjecture the CBs central role in RNA compartmentalization. The majority of CB proteins can bind RNA, which provides another indirect but important aspect on the possible mechanisms of how RNAs localize in the CB. Finally, *in situ* and RT-PCR experiments, combined with deep sequencing analysis further verify that a large proportion of the RNA in the CB accounts for mature mRNAs.

We have indirect and incomplete evidence that mRNAs are stored in the CB. A large quantity of mRNAs in the CB indicates that these may be stored therein. CB-related transcriptome differs from that of round spermatid's, which indicates that the CB may be a target for specially regulated transcripts, rather than a general processing platform of the whole cell's transcriptome. An EU-chase experiment on the cultured seminiferous tubule pieces demonstrated that the newly synthesized RNA is stored in the CB for at least 12 hours. However, if- and to what extent- this EU-labeled RNA corresponds to mRNAs, remains currently undetermined.

Gel electrophoresis of CB RNAs from early and late round spermatids have demonstrated that the general CB RNA profile is consistent. It was further noticed that the CB-associated RNA remains very stable during the CB isolation and the CB *in vitro* assay. This indicates that the RNAs inside the CB are protected.

The specific CB localization of PABPC2- a testis-restricted poly(A) binding protein that is associated with stable, non-translated mRNAs (Kimura et al. 2009)- is a further indication that the CB-associated mRNAs may be stably stored there.

Transition proteins (TNP) and protamines (PRM) are the most abundant CB mRNAs. The mRNA expression of *Tnp2* and *Prm1* begins around step 7 spermatids (Mali et al. 1988, Saunders et al. 1992). It is known that TNP and PRM mRNAs are stored in cytoplasm for 3 to

7 days and translated in steps around 11 and 13, respectively (Meistrich et al. 2003, Kwon, Hecht 1993). The massive abundance of these transcripts in late CBs indicates that they may be translationally inhibited and protected therein. However, the CB localization of MVH disappears in step 7-8 round spermatids, (Toyooka et al. 2000) which means that the transformed CB structures from elongating spermatids cannot be isolated by the current MVH-immunoprecipitation method. Therefore, the investigation of how long are these late transcripts stored in the CB-like structures cannot be undertaken using MVH-IP. To date, there is only a single report of *in situ* hybridization analysis which states that transition protein 2 mRNA is indeed stored in the CB (Saunders et al. 1992). However, this study dates back to over 20 years and requires an updated verification.

Overall, the results obtained demonstrate that large quantities of mRNAs are present in the CB. It is likely that some of these are also stored therein for later translation. At present, there is insufficient reliable data to unequivocally either claim or overrule this. Hence, whether the CB functions as a general reserve to store dormant mRNA stock, still remains to be confirmed.

6.2.7. *In vitro* RNA-processing potential of the CB

We have provided preliminary evidence that CBs can process RNA *in vitro*. Substrate RNA is degraded into smaller fragments with an inconsistent pattern. In addition to the processing ability, the isolated CBs were also able to bind significant amounts of substrate RNA. While the *in vitro* substrate RNA binding appeared to be a general property of the isolated CBs, the processing ability was greatly enhanced by ATP and completely blocked by Ribolock™ RNase inhibitor or by heating at 70 °C for 45 min. Ribolock has been stated to inhibit the activity of RNases A, B, and C by noncompetitive binding. However, these common nucleases do not require ATP for their activity and are stable at higher temperatures (in fact they are boiled for tens of minutes during their isolation to eliminate the contaminating deoxyribonuclease). Further, none of the common nucleases were detected in the CB proteome by MS. This might suggest that the CB-associated *in vitro* degradation of substrate RNA is an enzymatic process, unrelated to the common RNases.

The only known proteins in the CB with reported ribonuclease activity are MIWI and SMG6. MIWI is the AGO engine for piRNA-guided RNA processing. Taking into account the well-represented piRNA pathway in the CB, it is probably functionally active there. However, the MIWI slicer assay does not require ATP and is not inhibited by Ribolock (Reuter et al. 2011). Further, the activity of the CB RNA degradation assay does not depend on manganese (data not shown) while it is required for SMG6 (Glavan et al. 2006). This suggests that MIWI and SMG6 are likely not involved in the *in vitro* RNA slicing activity of the isolated CBs.

Whether CB-associated *in vitro* RNA processing ability is caused by RNase contamination, remains unclear. One possible explanation to the ATP-dependent CB-associated *in vitro* RNA processing would be the involvement of an RNA helicase which utilizes ATP to hydrolyze RNA higher structures to make it accessible to the effector molecules which perform the RNA cleavage. Further experiments would be required to analyze and validate such probabilities.

6.3. The 137 year old question

What is the role of CB? This has been asked since the discovery of the CB by von Brunn (von Brunn 1876). Back in 1960s CB's function was tentatively conjectured to provide basic proteins for the final maturation of the chromatin in the nucleus of late spermatid (Sud 1961). However, currently the CB is considered to be involved in RNA post-transcriptional regulation (Meikar et al. 2011). Despite our detailed analyses on the CB's components, its possible functions still remain largely unknown and evoke more questions and hypothesis than answers (Fig. 14).

The common keyword of most CB-related molecular pathways is the regulation of gene expression, mostly at a post-transcriptional level. Spermatids are cells with a complex developmental program, which require extra mechanisms for gene regulation. These haploid cells are functionally diploid, as connected to their sister cells by intercellular bridges, by which to share the common cytoplasm. CB-associated transport between these sisters cells has been described (Ventela et al. 2003). Therefore, it can be conjectured that spermatids require a localized RNA regulation center to correctly deal with their intra- and intercellular transcriptomes. The CB is present as a single structure in a round spermatid, and shares in this respect a strategic similarity with the nucleus and the Golgi complex. A single central station ensures central control and provides the opportunity for global scanning/sensing/quality-control and grants monopoly over regulation. Considering all this, the CB might be a department for the centralized regulation of the round spermatid's cytoplasmic RNA affairs.

Thereby, in summary, in the case of the storage of dormant mRNAs in the CB - translational inhibition is crucial in spermiogenesis the sequestration of dormant mRNAs inarguably occurs. Remaining questions such as where, to what extent and for how long, are unlikely to be answered conclusively and should not be too overemphasized. mRNAs are translationally inhibited in different smaller and larger mRNP granules, in the cell. For the CB's transcriptome and proteome, it is very likely that at minimum some mRNA storage also occurs there, in addition to other, potentially a lot more fascinating functions. These clearly relate to the observations that round spermatids have one the most complex transcriptomes and that non-coding RNAs are strongly accumulated in the CB.

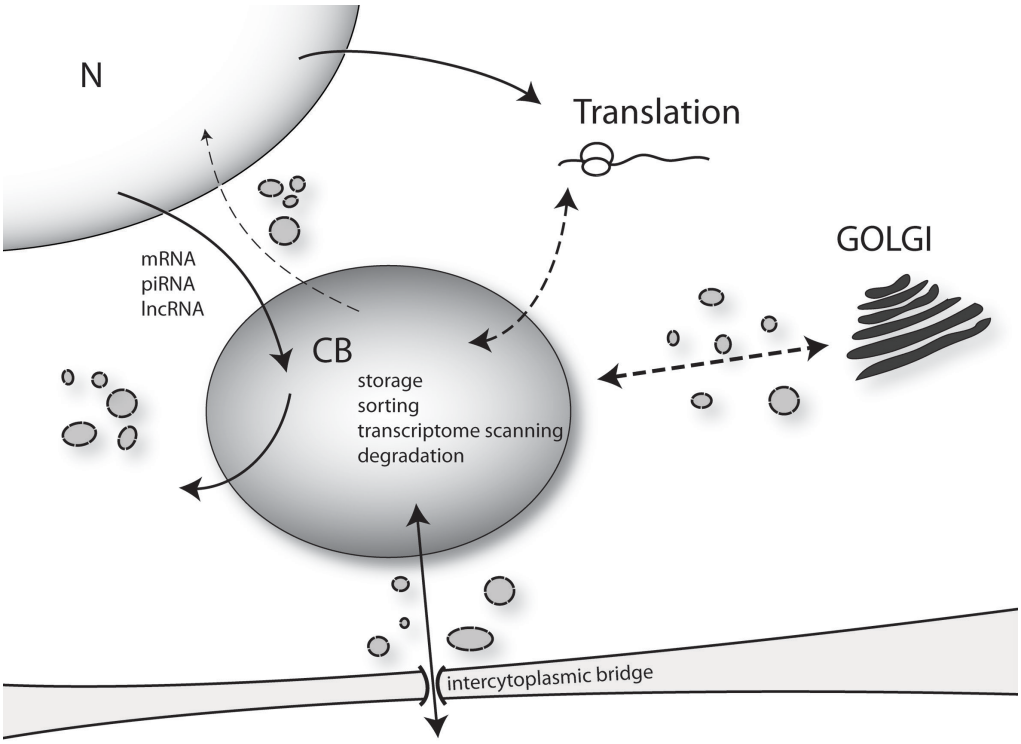

Figure 14. *A hypothetical model of the CB's possible functions. The CB is the department for the centralized regulation of the round spermatid's cytoplasmic RNA affairs. The CB is a large, single structure in the cytoplasm of a round spermatids, in close proximity to the nuclear envelope. In addition to the nucleus, the CB shares actively material also with Golgi, endoplasmatic reticulum and the neighboring haploid cell via vesicles and other particles. A major flux of nascent RNA from the nucleus to the CB is constant and accompanied with a parallel depletion, by the release of RNAs to the cytoplasm or by decay. However some dormant RNAs are stored for longer time. Several dynamic protein components of the CB shuttle actively to share their cargo between the CB and the rest of the cell, while the general RNA and protein profile of the CB remains largely unchanged. CB is the center of piRNA action and RNA post-transcriptional regulation. It consists largely of RNA binding proteins, piRNA pathway components and scaffolding proteins. Besides the abundant piRNAs and mRNAs, also long non-coding RNAs are accumulated in the CB. CB is a central RNA scanning and sorting platform where different molecular pathways are converge to organize the highly complex transcriptome of round spermatids.*

6.4. Testis- the backroom of our evolution?

Germline is a prerequisite for the survival of the species. However, for survival, change and adaption to the environment is essential. Shuffling genetic material and mixing it equally with a partner in offspring is a successful means of evolution. A self-explanatory proof of the evolutionary advantage of this strategy is confirmed by the overwhelming number of

sexually reproducing organisms throughout the world. Every individual with a new modified genome (and epigenome) is exposed to the environment and positively selected, based on the success of the phenotype. Organisms with a slow generation time and relatively low number of offspring should need compensatory mechanisms in order to accelerate their adaptive evolution. In addition to meiosis, there are at least two other hypothetical mechanisms which could be independently responsible for driving our evolution. These are the emergence of new genes (Fig. 15) and the mechanisms of transgenerational epigenetic inheritance (Fig. 16). Available data suggest that these probably RNA-mediated processes occur in the male germline and are possibly connected to the CB.

Male germline stem cells are the only immortal germ cells in the adult human that can give rise to a new organism and are theoretically capable of unlimited divisions. Adult female ovaries have instead a limited number of prophase-arrested primordial oocytes (Von Stetina, Orr-Weaver 2011). Surprisingly, the largest transcriptome complexity in all the cells of a mammalian organism resides in testis - in post-meiotic round spermatids. These haploid cells express the largest amount of different protein-coding genes, predicted lncRNA genes and retrogenes, and are also most affected by alternative splicing events (Soumillon et al. 2013). The current study has demonstrated that the CB appears to be in the center of round spermatid's general RNA flux and provides as a platform for different RNA regulatory pathways. As such, it is tempting to impute some role in the emergence of new genes and the mechanisms of transgenerational epigenetic inheritance to round spermatids and to "the department for the centralized regulation of the round spermatid's cytoplasmic RNA affairs"- the CB.

6.4.1. The birth of new genes and the control of the novel genome

It has been firmly established that new genes are essential contributors to the origin of adaptive evolutionary novelties. Genomic analyses have revealed that the major mechanisms related to the creation of new genes are the duplications and fusions of genetic material (Kaessmann 2010). These can result in new protein-coding and functional non-coding genes from previously non-functional genomic sequences, various types of gene fusions and the formation of new genes, from RNA intermediates. Transposons are functionally capable to cut, copy and paste DNA around the genome. Retrotransposons can do the same, but over the RNA intermediates- RNA however is far more potent molecule for alterations, for example by RNA editase enzymes or alternative splicing. Approximately 42% of the human genome consists of retrotransposons, while DNA transposons account for approximately 2-3% (Lander et al. 2001). By which, it is fascinating to assume that the evolution of our genome is orchestrated by the syncytium of different RNA modifying enzymes and retrotransposons. Transposable elements or so-known "selfish DNA parasites" may in fact provide a vital mechanism for our evolution by creating new combinations of genes. Not surprisingly, testis is the place for very high transposonal activity (Aravin et al. 2008, Aravin et al. 2009, Meikar et al. 2012, Aravin et al. 2007).

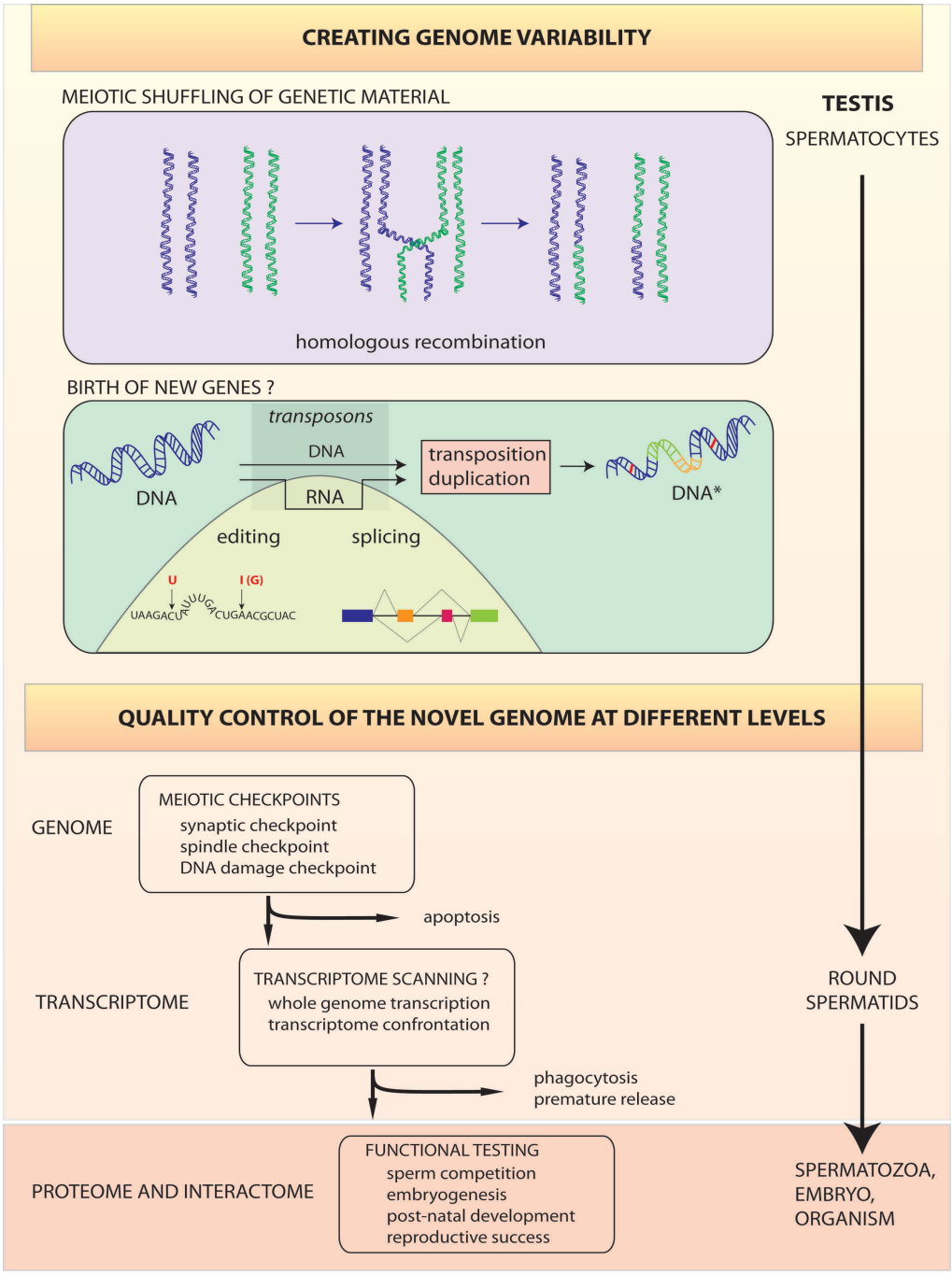

Figure 15. *A model of potential mechanisms of creating genomic variability and the subsequent quality scan, in male germ cells. Meiotic mechanisms (random chromosomal segregation and homologous recombination) and (retro)transposon-derived mechanisms (DNA transposition and duplication; also splicing and editing, in case of a RNA intermediate) create genetic variability. Quality-check of the modified, unique genomes is carried out on genome, transcriptome and, to some extent, proteome level.*

The emergence of new genes also applies to non-coding functional RNA transcripts. While the total number of protein-coding genes in mammals has remained generally the same since the metazoan radiation, the amount of long non-coding RNA in the genome has been increasing, in accordance with the associated complexity of the organisms. These tens, if not hundreds of thousands of lncRNAs are differentially expressed, highly heterogenous and, in general, less conserved than the protein coding genes (Taft et al. 2007). Regulatory RNAs are not restricted to structure-function relationships as the protein-coding genes and are therefore much more flexible to changes in order to respond the demands created by evolutional pressure. Therefore, lncRNAs have been associated with a broad epigenetic regulatory network that provides a parallel dynamic alternative to the coding genes in the evolution of our functional genome (Rinn, Chang 2012, Mercer, Mattick 2013). The extraordinary association between the complexity of an organism and the proportion of its non-coding genome supports the concept that our evolution is largely mediated by non-coding RNAs.

There are two developmental windows in male germ cells when transposons are de-repressed- in fetal prospermatogonia and in early meiotic spermatocytes (Meikar et al. 2011). New genes may likely be introduced to the genome by (retro)transposons in preleptoten spermatocytes at stage VII of the seminiferous epithelial cycle, (Fig. 7). At which time, the first meiotic division coincides with the second activation of transposable elements, which are eventually silenced at the pachytene stage. In comparison to the first transposonal activity in pro-spermatogonia, the spermatocytes do not proliferate, to ensure that the genomic rearrangements in every spermatocyte would be unique to each cell and not multiplied.

The mechanism for the emergence of new genes would surely require a quality control mechanism to somehow exclude the germ cells with too massive genomic, transcriptomic and/or proteomic rearrangements. Complete functional testing of the newly formed genome is only possible in the offspring. However, it is relatively easy to evaluate the general integrity of the genome and its first derivative- the transcriptome, in single cells. The basic test of genomic integrity after transposons silencing could be meiosis. For example, the formation of synaptonemal complexes between homologous chromosomes allows the probing for the presence of pairing abnormalities. The large number of apoptotic, early post-meiotic cells around stage I of the seminiferous epithelial cycle, is likely the result of this control step.

Evaluating the transcriptomes of the unique haploid spermatids by some comparative mechanism would allow the exclusion of cells with major aberrations, since this is generally a sign of non-viability. Thereby, perhaps the exceptionally heterogeneous transcriptome in round spermatids is a result of some form of RNA scanning and confrontation mechanism. The CB in round spermatids matches the criteria necessary for a location where the final confrontation of the spermatid's unique genome, through its transcriptome, could be arranged. The CB is a sole, separate granule in cells, which allows the global transcriptome scanning machinery to be compartmentalized into a single center. CB is the target of the central flux of nascent RNA from the nucleus. The localization patterns of SAM68, but also other RNA binding proteins, may refer to some additional, CB-targeted RNA transport from the cytoplasm. Furthermore, the isolated CBs contain vast amounts of different coding and

non-coding RNAs and a peculiar set of various proteins which are associated with RNA processing. All this infers that the CB could house the scanning mechanism which controls the transcriptional integrity of its round spermatid. If this form of quality-check fails, the spermatid's unique genome is considered potentially unviable and removed by phagocytosis of premature release to lumen of the seminiferous tubule (Fig. 15).

This putative transcriptional quality-scan mechanism of the genome remains yet to be characterized. However, it could functionally explain the permissive chromatin and the biggest known transcriptional complexity in late meiotic spermatocytes and particularly in post-meiotic round spermatids (Soumillon et al. 2013). Moreover, it could also explain the remarkably low number of harmful mutations in the fertilizing spermatids.

6.4.2. Transgenerational epigenetic inheritance

A remarkable phenomenon where environment can epigenetically alter gene expression and that at least some of these epigenetic changes are hereditary, has been reported throughout prokaryote and eukaryote kingdoms (Jablonka, Raz 2009, Bohacek et al. 2013). Although not clearly understood, transgenerational epigenetic inheritance (TEI) infers an impressive mechanism for accelerated adaptive evolution, giving some credit to the generally long abandoned Lamarckism. Furthermore, the role of TEI in the etiology of diseases, such as cancer, diabetes and neurological disorders which have strong heritable components and environmental associations, has garnered increasing interest in the mechanism of epigenetics.

Although the phenomenon of TEI exists, its mechanisms that underlay still need to be revealed. When dividing TEI into steps- first, the environment triggers an epigenetic response in the body; the information of such epigenetic response reaches the developing male germ cells in the testis and becomes included in the (epi)genome of spermatozoa; this epigenetic information is then transferred to the oocyte and ultimately manifest itself in the progeny (Fig. 16). These steps of TEI are considered separately below.

The fact that epigenetic signals play a critical role in the regulation of gene expression and thereby the determination of our phenotypes, has become widely accepted. The epigenetic response modifies chromatin and the transcriptomal outcome, without any effect on the DNA sequence itself. These epigenetic tuning mechanisms include DNA methylation and chromatin remodeling (Daxinger, Whitelaw 2010, Smith, Meissner 2013, Bannister, Kouzarides 2011). Further, small-RNA mediated epigenetic silencing by directing DNA methylation, has been convincingly reported in the germline of different organisms (Aravin et al. 2007, Carmell et al. 2007).

TEI travels through germ cells as these are the only genetic bridge between parents and their offspring. Compared to oogenesis, spermatogenesis is an exposed, active and continuous process which takes approximately 2 months to complete. As such TEI could directly target the developing male germ cells and equip them with the snapshot of the organism's adaptive epigenetic response to the environment. To accomplish the transgenerational transfer, the epigenetic signal would need to reach from its originating tissue to the testis

and manifest itself in the developing germ cells. Of note cell-to-cell and long-distance siRNA movements has been described in plants (Brosnan, Voinnet 2011). Further, recently an extensive list of circulating RNA-based biomarker candidates from mouse blood plasma has also been published (Wang et al. 2013). If and how might the circulating epigenetic signal find its way to the sperm head and in which form it would be presented in the zygote, remain at present completely unknown.

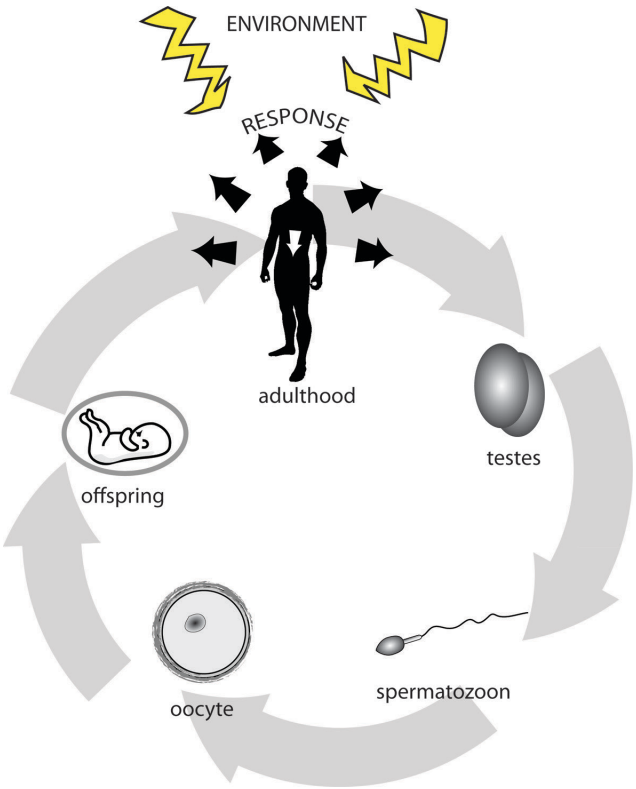

Figure 16. *The circle of transgenerational epigenetic inheritance. Environmentally induced, epigenetic response tunes gene expression (e.g. by small RNAs, DNA methylation, histone modifications). This information is delivered to testis and is incorporated into differentiating male germ cells. Spermatozoa carry the genetic information, together with the current epigenetic metainformation to the egg. Eventually, the inherited epigenetic metadata alters gene expression in the offspring to equip it with prior adaptive responseto the environment.*

Epigenetic changes in germ cells can become transferred on to the next generation. Notably, a multi-generational epigenetic inheritance mechanism, induced by piRNAs, was recently reported in fly (Shirayama et al. 2012, Ashe et al. 2012). Moreover, early-life stress induced behavioral responses were demonstrated to be passed on to the next generation in mice. Furthermore, respective DNA methylation was reported not only in the sperm of the induced fathers, but also in the brain and germline of the offspring too (Franklin et al. 2010).

It is absolutely fascinating to consider an additional mechanism whereby an inherited epigenetic mark in the germ cells could also, with some probability, be prone to become genetically fixed. This would thereby complete the full mechanism of the accelerated and directed evolution of our genome. When considering our long generation time and the small number of progeny, this model appears clearly more favorable than the evolution through only random mutations.

7. SUMMARY AND CONCLUSIONS

Male germ cells represent the immortal line- their unique characteristics and unusual mechanisms make them important, fascinating and challenging subjects to study. As such, keywords and phrases to describe mammalian male germ cell differentiation would include the following- separation from the immune system by the blood-testis barrier; unusual energy metabolism and conserved temperature requirements; accelerated evolutionary differentiation; wide transcriptional profile with transient transcriptional quiescence; extensive alternative splicing and the biggest repertoire of non-coding regulatory RNAs; epigenetic patterning; three rounds of global chromatin remodeling; meiosis in its complexity; direct hormonal control via the hypothalamic-pituitary axis; continuous cyclic production of spermatozoa and huge productivity. The complexity of this system and the limitation of standard tools to studying (very limited *in vitro* culture possibilities, limited transfection efficiency and long generation times) have left large areas in male germ cell research still remaining a mystery.

To overcome these limitations, improvisation with a wide range of different methods and approaches is required. I consider my efforts to fulfill this requirement as the most important and educative personal value of this study and I am very grateful for it. We have provided a working protocol for high-purity CB isolation which, for the first time, has allowed the study of CB components in great detail to further stretch in comprehension of its functions and mechanisms. The in-depth analyses on the CBs proteome, transcriptome and interactome, together with some mechanistic studies, clearly infer to the CB's roles in association of the most complex transcriptome in the body. We conjecture that the CB functions in the post-transcriptional regulation of a round spermatid through various independent or converged pathways, such as the piRNA- and nonsense-mediated-decay-based mechanisms. In addition, the possible contributions of male germ cells and CBs, to the evolutionary race, are considered.

The development of the CB-isolation procedure and an atlas of the CB components provide a fundamental basis for the future studies of CBs and also other RNP granules. To paraphrase Sir Winston Churchill's famous quote, it could be now said "This study is not at its end, it is not even at the beginning of its end. But it is, perhaps, at the end of a beginning of the new era in CB research".

8. ACKNOWLEDGEMENTS

The current study was carried out at the Department of Physiology, Institute of Biomedicine, University of Turku, during years 2006-2013.

My deepest gratitude goes to my supervisors Assistant Professor Noora Kotaja, PhD and Professor Jorma Toppari MD, PhD. Thank you Jorma for your continuous support, appreciation and good influence. This study would be impossible without Noora, who is really the best supervisor I have seen or heard about. Thank you for giving me the opportunity to enjoy science, develop and learn good scientific practice from you.

My special thanks to Professor Emeritus Lauri J. Pelliniemi, MD, PhD for a great scientific discussion in front of Mikro on one summer day in 2006, which resulted in my PhD career in Noora's lab. Ever since, these scientific discussions have been many more, each one of which very fruitful.

I am very grateful to Professor Emeritus Martti Parvinen, MD, PhD - the father of chromatoid body research, for his excellent scientific support and profound professional knowledge, which are so wonderfully combined together with easygoingness and joyful appearance. This is truly encouraging.

I am very thankful to the reviewers of this thesis, Dr. Liisa Kauppi, PhD, and Dr. Mikko Frilander, PhD, for careful and critical review. Your involvement and contribution was very useful, thank you for your time and energy.

I would also like to thank Professor Lea Sistonen, PhD and Dr. Markko Kallio, PhD, for being the members of my thesis committee. Additional thanks to Lea for her direct support in making some scientific infrastructure so easily accessible for me.

All my co-authors are kindly acknowledged for their contribution to this work. I would especially like to thank Dr. Vasily Vagin, PhD for the great time together in Cold Spring Harbor Labs.

One beautiful chapter of my PhD-student life has been the Turku Doctoral Programme Biomedical Sciences (TuBS), where I have had the privilege to belong to. I would like to express my deep gratitude to Professor Olli Lassila, MD, PhD, Mrs. Nina Wildber and Susanna Rosenberg, MSc. for all the scientific events and social activities. TuBS brought science and fun together in to a very motivating environment. I am grateful to all my fellow TuBS family members but I would like to particularly thank the very talented students Johanna Björk, PhD and Claire Hyder, MSc for great time, great work and great results.

Thank you Maire Tohv and Helena Mikk for your support, good company and for being "a piece of Estonia" for me in Turku.

I am indebted to my past and present lab-mates in the Department of Physiology, who have taught, nurtured and tolerated me throughout these years. My sincere apologizes for being

sometimes too hasty, messy and disorganized – thank you from the bottom of my heart for creating a nice working and living environment throughout these years.

I would especially like to thank Adolfo Rivero-Müller, PhD, Karin Sõstar, BSc, Milena Doroszko, MSc, Henriette Undeutsch MSc, Tiina Lehtiniemi BSc, Hanna Korhonen, MSc, Ram Prakash Yadav, MSc and Mari Lehti, MSc for making the immediate lab- and office-life so pleasurable and special. My biggest thanks to Juho Mäkela, PhD, Matteo Da Ros, MSc and Heikki Turunen, PhD for their friendship. All you guys, in your very different way, make me complete. Thank you for being my family.

Many thanks to Allan Nurk, PhD, Maire Peters, PhD, Eerik Jõgi, MSc and Indrek Suitso MSc, for the beginning years of my scientific career. These were great times I will always remember. Big thanks to Rahel, Meelis & co for your friendship throughout many years.

My dear "TK" and "I4" boys: Hendrik, Kalmer, Kristjan, Indrek, Koit, Viljo, Reidar and Tõnis- ma tänan teid, et te olete mul alati olemas olnud, te olete parimad.

There are numerous people that I would like to express my gratitude for the great times and good, memorable emotions. Please, if you recognize that I am referring to you, accept my deep appreciation and consider yourself acknowledged.

My thanks to Dr. Vladimir Benes, PhD for being my friend and mentor. I am very grateful for Katharina Gapp, MSc for showing me that everything is possible. My deepest thanks go to Dr. Maria Kirss, and her excellent kin- I want to express my gratitude for accepting me as part of your family and for memorable time together.

I am grateful to Kersti Türk for being my mentor and always asking me the right questions in the right time. I thank Rene Türk for great company and continuous surprises. I would like to thank the whole Uudam family but, most notably, you Triin- thank you for your love, attention and support during this last very busy year. You are the best.

I am so grateful to my parents Sirje and Toivo for providing me a great childhood and bringing me up with good principles and values. I am exceedingly proud of my twin brother Silver and my sister Margit- I thank you both for being great and reliable companions for me ever since I can remember.

This work has been financially supported by the Academy of Finland, Sigrid Jusélius Foundation, Emil Aaltonen Foundation, the Finnish Cultural Foundation, Centre for Reproductive and Developmental Medicine, Turku University Foundation, and Turku Doctoral Programme of Biomedical Sciences.

Oliver Meikar, Turku, December 2013

9. REFERENCES

Agranat, L., Raitskin, O., Sperling, J., Sperling, R. 2008. The editing enzyme ADAR1 and the mRNA surveillance protein hUpf1 interact in the cell nucleus. Proc.Natl.Acad.Sci.U.S.A. 105, 5028-33.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P. 2002. Molecular biology of the cell. 55, Garland Science, New York.

Anderson, P. and Kedersha, N. 2009. RNA granules: post-transcriptional and epigenetic modulators of gene expression. Nat.Rev.Mol.Cell Biol. 10, 430-6.

Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., Morris, P., Brownstein, M.J., Kuramochi-Miyagawa, S., Nakano, T., Chien, M., Russo, J.J., Ju, J., Sheridan, R., Sander, C., Zavolan, M., Tuschl, T. 2006. A novel class of small RNAs bind to MILI protein in mouse testes. Nature 442, 203-7.

Aravin, A.A., Hannon, G.J., Brennecke, J. 2007. The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. Science 318, 761-4.

Aravin, A.A., Sachidanandam, R., Bourc'his, D., Schaefer, C., Pezic, D., Toth, K.F., Bestor, T., Hannon, G.J. 2008. A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. Mol.Cell 31, 785-99.

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

Armour, C.D., Castle, J.C., Chen, R., Babak, T., Loerch, P., Jackson, S., Shah, J.K., Dey, J., Rohl, C.A., Johnson, J.M., Raymond, C.K. 2009. Digital transcriptome profiling using selective hexamer priming for cDNA synthesis. Nat.Methods 6, 647-9.

Ashe, A., Sapetschnig, A., Weick, E.M., Mitchell, J., Bagijn, M.P., Cording, A.C., Doebley, A.L., Goldstein, L.D., Lehrbach, N.J., Le Pen, J., Pintacuda, G., Sakaguchi, A., Sarkies, P., Ahmed, S., Miska, E.A. 2012. piRNAs can trigger a multigenerational epigenetic memory in the germline of C. elegans. Cell 150, 88-99.

Azzalin, C.M. and Lingner, J. 2006. The double life of UPF1 in RNA and DNA stability pathways. Cell.Cycle 5, 1496-8.

Ballut, L., Marchadier, B., Baguet, A., Tomasetto, C., Seraphin, B., Le Hir, H. 2005. The exon junction core complex is locked onto RNA by inhibition of eIF4AIII ATPase activity. Nat.Struct.Mol.Biol. 12, 861-9.

Bannister, A.J. and Kouzarides, T. 2011. Regulation of chromatin by histone modifications. Cell Res. 21, 381-95.

Barchi, M., Geremia, R., Magliozzi, R., Bianchi, E. 2009. Isolation and analyses of enriched populations of male mouse germ cells by sedimentation velocity: the centrifugal elutriation. Methods Mol.Biol. 558, 299-321.

Betel, D., Sheridan, R., Marks, D.S., Sander, C. 2007. Computational analysis of mouse piRNA sequence and biogenesis. PLoS Comput.Biol. 3, e222.

Bohacek, J., Gapp, K., Saab, B.J., Mansuy, I.M. 2013. Transgenerational epigenetic effects on brain functions. Biol.Psychiatry 73, 313-20.

Bono, F., Ebert, J., Lorentzen, E., Conti, E. 2006. The crystal structure of the exon junction complex reveals how it maintains a stable grip on mRNA. Cell 126, 713-25.

Boumil, R.M. and Lee, J.T. 2001. Forty years of decoding the silence in X-chromosome inactivation. Hum.Mol.Genet. 10, 2225-32.

Boussouar, F. and Benahmed, M. 2004. Lactate and energy metabolism in male germ cells. Trends Endocrinol.Metab. 15, 345-50.

Brennicke, A., Marchfelder, A., Binder, S. 1999. RNA editing. FEMS Microbiol.Rev. 23, 297-316.

Brosnan, C.A. and Voinnet, O. 2011. Cell-to-cell and long-distance siRNA movement in plants: mechanisms and biological implications. Curr.Opin.Plant Biol. 14, 580-7.

Buhler, M., Wilkinson, M.F., Muhlemann, O. 2002. Intranuclear degradation of nonsense codon-containing mRNA. EMBO Rep. 3, 646-51.

Burgoyne, P.S., Mahadevaiah, S.K., Turner, J.M. 2009. The consequences of asynapsis for mammalian meiosis. Nat.Rev.Genet. 10, 207-16.

Carmell, M.A., Girard, A., van de Kant, H.J., 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. Dev.Cell. 12, 503-14.

Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R., Shimokawa, K., Bajic, V.B., Brenner, S.E., Batalov, S., Forrest, A.R., Zavolan, M., Davis, M.J., Wilming, L.G., Aidinis, V., et al. 2005. The transcriptional landscape of the mammalian genome. Science 309, 1559-63.

Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B.M., Strein, C., Davey, N.E., Humphreys, D.T., Preiss, T., Steinmetz, L.M., Krijgsveld, J., Hentze, M.W. 2012. Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell 149, 1393-406.

Castrillon, D.H., Quade, B.J., Wang, T.Y., Quigley, C., Crum, C.P. 2000. The human VASA gene is specifically expressed in the germ cell lineage. Proc.Natl.Acad.Sci.U.S.A. 97, 9585-90.

Cenik, E.S. and Zamore, P.D. 2011. Argonaute proteins. Curr.Biol. 21, R446-9.

Chalmel, F., Rolland, A.D., Niederhauser-Wiederkehr, C., Chung, S.S., Demougin, P., Gattiker, A., Moore, J., Patard, J.J., Wolgemuth, D.J., Jegou, B., Primig, M. 2007. The conserved transcriptome in human and rodent male gametogenesis. Proc.Natl.Acad.Sci.U.S.A. 104, 8346-51.

Chang, Y.F., Imam, J.S., Wilkinson, M.F. 2007. The nonsense-mediated decay RNA surveillance pathway. Annu.Rev.Biochem. 76, 51-74.

Chen, C., Nott, T.J., Jin, J., Pawson, T. 2011. Deciphering arginine methylation: Tudor tells the tale. Nat.Rev.Mol.Cell Biol. 12, 629-42.

Chen, C., Ouyang, W., Grigura, V., Zhou, Q., Carnes, K., Lim, H., Zhao, G.Q., Arber, S., Kurpios, N., Murphy, T.L., Cheng, A.M., Hassell, J.A., Chandrashekar, V., Hofmann, M.C., Hess, R.A., Murphy, K.M. 2005. ERM is required for transcriptional control of the spermatogonial stem cell niche. Nature 436, 1030-4.

Chen, J., Chiang, Y.C., Denis, C.L. 2002. CCR4, a 3'-5' poly(A) RNA and ssDNA exonuclease, is the catalytic component of the cytoplasmic deadenylase. EMBO J. 21, 1414-26.

Chen, T., Damaj, B.B., Herrera, C., Lasko, P., Richard, S. 1997. Self-association of the single-KH-domain family members Sam68, GRP33, GLD-1, and Qk1: role of the KH domain. Mol.Cell.Biol. 17, 5707-18.

Chi, M.N., Auriol, J., Jegou, B., Kontoyiannis, D.L., Turner, J.M., de Rooij, D.G., Morello, D. 2011. The RNA-binding protein ELAVL1/HuR is essential for mouse spermatogenesis, acting both at meiotic and postmeiotic stages. Mol.Biol.Cell 22, 2875-85.

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

Clermont, Y. 1972. Kinetics of spermatogenesis in mammals: seminiferous epithelium cycle and spermatogonial renewal. Physiol.Rev. 52, 198-236.

Clermont, Y. 1962. Quantitative analysis of spermatogenesis of the rat: a revised model for the renewal of spermatogonia. Am.J.Anat. 111, 111-29.

Cole, C.N. and Scarcelli, J.J. 2006. Transport of messenger RNA from the nucleus to the cytoplasm. Curr.Opin.Cell Biol. 18, 299-306.

Connolly, C.M., Dearth, A.T., Braun, R.E. 2005. Disruption of murine Tenr results in teratospermia and male infertility. Dev.Biol. 278, 13-21.

Darnell, J.E.,Jr. 2013. Reflections on the history of pre-mRNA processing and highlights of current knowledge: a unified picture. RNA 19, 443-60.

Dassi, E., Zuccotti, P., Leo, S., Provenzani, A., Assfalg, M., D'Onofrio, M., Riva, P., Quattrone, A. 2013. Hyper conserved elements in vertebrate mRNA 3'-UTRs reveal a translational network of RNA-binding proteins controlled by HuR. Nucleic Acids Res. 41, 3201-16.

Daxinger, L. and Whitelaw, E. 2010. Transgenerational epigenetic inheritance: more questions than answers. Genome Res. 20, 1623-8.

Deaton, A.M. and Bird, A. 2011. CpG islands and the regulation of transcription. Genes Dev. 25, 1010-22.

DeCerbo, J. and Carmichael, G.G. 2005. Retention and repression: fates of hyperedited RNAs in the nucleus. Curr.Opin.Cell Biol. 17, 302-8.

Decker, C.J. and Parker, R. 2012. P-bodies and stress granules: possible roles in the control of translation and mRNA degradation. Cold Spring Harb Perspect.Biol. 4, a012286.

Deng, W. and Lin, H. 2002. Miwi, a Murine Homolog of Piwi, Encodes a Cytoplasmic Protein Essential for Spermatogenesis. Dev.Cell. 2, 819-30.

Doller, A., Pfeilschifter, J., Eberhardt, W. 2008. Signalling pathways regulating nucleo-cytoplasmic shuttling of the mRNA-binding protein HuR. Cell.Signal. 20, 2165-73.

Dreyfuss, G., Matunis, M.J., Pinol-Roma, S., Burd, C.G. 1993. hnRNP proteins and the biogenesis of mRNA. Annu.Rev.Biochem. 62, 289-321.

Dym, M. and Fawcett, D.W. 1970. The blood-testis barrier in the rat and the physiological compartmentation of the seminiferous epithelium. Biol.Reprod. 3, 308-26.

Elliott, D. and Ladomery, M. 2011. Molecular biology of RNA. 11, Qxford University Press, United States.

Elliott, D.J. and Grellscheid, S.N. 2006. Alternative RNA splicing regulation in the testis. Reproduction 132, 811-9.

Eulalio, A., Behm-Ansmant, I., Schweizer, D., Izaurralde, E. 2007. P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. Mol.Cell.Biol. 27, 3970-81.

Ewen, K.A. and Koopman, P. 2010. Mouse germ cell development: from specification to sex determination. Mol.Cell.Endocrinol. 323, 76-93.

Fan, X.C. and Steitz, J.A. 1998. Overexpression of HuR, a nuclear-cytoplasmic shuttling protein, increases the in vivo stability of ARE-containing mRNAs. EMBO J. 17, 3448-60.

Fawcett, D.W., Eddy, E.M., Phillips, D.M. 1970. Observations on the fine structure and relationships of the chromatoid body in mammalian spermatogenesis. Biol.Reprod. 2, 129-53.

Figueroa, J. and Burzio, L.O. 1998. Polysome-like structures in the chromatoid body of rat spermatids. Cell Tissue Res. 291, 575-9.

Franceschini, A., Szklarczyk, D., Frankild, S., Kuhn, M., Simonovic, M., Roth, A., Lin, J., Minguez, P., Bork, P., von Mering, C., Jensen, L.J. 2013. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Res. 41, D808-15.

Franklin, T.B., Russig, H., Weiss, I.C., Graff, J., Linder, N., Michalon, A., Vizi, S., Mansuy, I.M. 2010. Epigenetic transmission of the impact of early stress across generations. Biol.Psychiatry 68, 408-15.

Frost, R.J., Hamra, F.K., Richardson, J.A., Qi, X., Bassel-Duby, R., Olson, E.N. 2010. MOV10L1 is necessary for protection of spermatocytes against retrotransposons by Piwi-interacting RNAs. Proc.Natl.Acad.Sci.U.S.A. 107, 11847-52.

Gehring, N.H., Kunz, J.B., Neu-Yilik, G., Breit, S., Viegas, M.H., Hentze, M.W., Kulozik, A.E. 2005. Exon-junction complex components specify distinct routes of nonsense-mediated mRNA decay with differential cofactor requirements. Mol.Cell 20, 65-75.

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

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

Glavan, F., Behm-Ansmant, I., Izaurralde, E., Conti, E. 2006. Structures of the PIN domains of SMG6 and SMG5 reveal a nuclease within the mRNA surveillance complex. EMBO J. 25, 5117-25.

Gorgoni, B. and Gray, N.K. 2004. The roles of cytoplasmic poly(A)-binding proteins in regulating gene expression: a developmental perspective. Brief Funct.Genomic Proteomic 3, 125-41.

Goss, D.J. and Theil, E.C. 2011. Iron responsive mRNAs: a family of Fe2+ sensitive riboregulators. Acc.Chem.Res. 44, 1320-8.

Grivna, S.T., Beyret, E., Wang, Z., Lin, H. 2006. A novel class of small RNAs in mouse spermatogenic cells. Genes Dev. 20, 1709-14.

Han, S.P., Tang, Y.H., Smith, R. 2010. Functional diversity of the hnRNPs: past, present and perspectives. Biochem.J. 430, 379-92.

Hansen, T.B., Jensen, T.I., Clausen, B.H., Bramsen, J.B., Finsen, B., Damgaard, C.K., Kjems, J. 2013. Natural RNA circles function as efficient microRNA sponges. Nature 495, 384-8.

Heller, C.G. and Clermont, Y. 1963. Spermatogenesis in man: an estimate of its duration. Science 140, 184-6.

Hess, R.A. and Renato de Franca, L. 2008. Spermatogenesis and cycle of the seminiferous epithelium. Adv.Exp.Med.Biol. 636, 1-15.

Hocine, S., Singer, R.H., Grunwald, D. 2010. RNA processing and export. Cold Spring Harb Perspect.Biol. 2, a000752.

Huarte, M., Guttman, M., Feldser, D., Garber, M., Koziol, M.J., Kenzelmann-Broz, D., Khalil, A.M., Zuk, O., Amit, I., Rabani, M., Attardi, L.D., Regev, A., Lander, E.S., Jacks, T., Rinn, J.L. 2010. A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. Cell 142, 409-19.

Huot, M.E., Brown, C.M., Lamarche-Vane, N., Richard, S. 2009. An adaptor role for cytoplasmic Sam68 in modulating Src activity during cell polarization. Mol.Cell.Biol. 29, 1933-43.

Huot, M.E. and Richard, S. 2012. Stay lean without dieting: Lose Sam68. Adipocyte 1, 246-9.

Iijima, T., Wu, K., Witte, H., Hanno-Iijima, Y., Glatter, T., Richard, S., Scheiffele, P. 2011. SAM68 regulates neuronal activity-dependent alternative splicing of neurexin-1. Cell 147, 1601-14.

International Human Genome Sequencing Consortium. 2004. Finishing the euchromatic sequence of the human genome. Nature 431, 931-45.

Ipsaro, J.J., Haase, A.D., Knott, S.R., Joshua-Tor, L., Hannon, G.J. 2012. The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis. Nature 491, 279-83.

Ishihama, Y., Oda, Y., Tabata, T., Sato, T., Nagasu, T., Rappsilber, J., Mann, M. 2005. Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. Mol.Cell.Proteomics 4, 1265-72.

Isken, O. and Maquat, L.E. 2007. Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function. Genes Dev. 21, 1833-56.

Jablonka, E. and Raz, G. 2009. Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. Q.Rev.Biol. 84, 131-76.

Jao, C.Y. and Salic, A. 2008. Exploring RNA transcription and turnover in vivo by using click chemistry. Proc.Natl.Acad.Sci.U.S.A. 105, 15779-84.

Jenison, R.D., Gill, S.C., Pardi, A., Polisky, B. 1994. High-resolution molecular discrimination by RNA. Science 263, 1425-9.

Johnson, C.L. and Spence, A.M. 2011. Epigenetic licensing of germline gene expression by maternal RNA in C. elegans. Science 333, 1311-4.

Kaessmann, H. 2010. Origins, evolution, and phenotypic impact of new genes. Genome Res. 20, 1313-26.

Kanatsu-Shinohara, M., Toyokuni, S., Shinohara, T. 2004. CD9 is a surface marker on mouse and rat male germline stem cells. Biol.Reprod. 70, 70-5.

Kanellopoulou, C., Muljo, S.A., Kung, A.L., Ganesan, S., Drapkin, R., Jenuwein, T., Livingston, D.M., Rajewsky, K. 2005. Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. Genes Dev. 19, 489-501.

Katahira, J. and Yoneda, Y. 2011. Nucleocytoplasmic transport of microRNAs and related small RNAs. Traffic 12, 1468-74.

Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C.C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Frith, M., Ravasi, T., Pang, K.C., Hallinan, J., Mattick, J., Hume, D.A., et al. 2005. Antisense transcription in the mammalian transcriptome. Science 309, 1564-6.

Kimmins, S. and Sassone-Corsi, P. 2005. Chromatin remodelling and epigenetic features of germ cells. Nature 434, 583-9.

Kimura, M., Ishida, K., Kashiwabara, S., Baba, T. 2009. Characterization of two cytoplasmic poly(A)-binding proteins, PABPC1 and PABPC2, in mouse spermatogenic cells. Biol.Reprod. 80, 545-54.

Kino, T., Hurt, D.E., Ichijo, T., Nader, N., Chrousos, G.P. 2010. Noncoding RNA gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. Sci.Signal. 3, ra8.

Kirino, Y. and Mourelatos, Z. 2007. 2'-O-methyl modification in mouse piRNAs and its methylase. Nucleic Acids Symp.Ser.(Oxf) (51), 417-8.

Kleene, K.C. 2001. A possible meiotic function of the peculiar patterns of gene expression in mammalian spermatogenic cells. Mech.Dev. 106, 3-23.

Kleene, K.C. and Cullinane, D.L. 2011. Maybe repressed mRNAs are not stored in the chromatoid body in mammalian spermatids. Reproduction 142, 383-8.

Korhonen, H.M., Meikar, O., Yadav, R.P., Papaioannou, M.D., Romero, Y., Da Ros, M., Herrera, P.L., Toppari, J., Nef, S., Kotaja, N. 2011. Dicer is required for haploid male germ cell differentiation in mice. PLoS One 6, e24821.

Kota, S.K. and Feil, R. 2010. Epigenetic transitions in germ cell development and meiosis. Dev.Cell. 19, 675-86.

Kotaja, N., Bhattacharyya, S.N., Jaskiewicz, L., Kimmins, S., Parvinen, M., Filipowicz, W., Sassone-Corsi, P. 2006. The chromatoid body of male germ cells: similarity with processing bodies and presence of Dicer and microRNA pathway components. Proc.Natl.Acad.Sci.U.S.A. 103, 2647-52.

Kotaja, N., Kimmins, S., Brancorsini, S., Hentsch, D., Vonesch, J.L., Davidson, I., Parvinen, M., Sassone-Corsi, P. 2004. Preparation, isolation and characterization of stage-specific spermatogenic cells for cellular and molecular analysis. Nat.Methods 1, 249-54.

Kotaja, N. and Sassone-Corsi, P. 2007. The chromatoid body: a germ-cell-specific RNA-processing centre. Nat.Rev.Mol.Cell Biol. 8, 85-90.

Krol, J., Loedige, I., Filipowicz, W. 2010. The widespread regulation of microRNA biogenesis, function and decay. Nat.Rev.Genet. 11, 597-610.

Kulkarni, M., Ozgur, S., Stoecklin, G. 2010. On track with P-bodies. Biochem.Soc.Trans. 38, 242-51.

Kuramochi-Miyagawa, S., Kimura, T., Ijiri, T.W., Isobe, T., Asada, N., Fujita, Y., Ikawa, M., Iwai, N., Okabe, M., Deng, W., Lin, H., Matsuda, Y., Nakano, T. 2004. Mili, a mammalian member of piwi family gene, is essential for spermatogenesis. Development 131, 839-49.

Kuramochi-Miyagawa, S., Watanabe, T., Gotoh, K., Takamatsu, K., Chuma, S., Kojima-Kita, K., Shiromoto, Y., Asada, N., Toyoda, A., Fujiyama, A., Totoki, Y., Shibata, T., Kimura, T., Nakatsuji, N., Noce, T., Sasaki, H., Nakano, T. 2010. MVH in piRNA processing and gene silencing of retrotransposons. Genes Dev. 24, 887-92.

Kwon, Y.K. and Hecht, N.B. 1993. Binding of a phosphoprotein to the 3' untranslated region of the mouse protamine 2 mRNA temporally represses its translation. Mol.Cell.Biol. 13, 6547-57.

Laiho, A., Kotaja, N., Gyenesei, A., Sironen, A. 2013. Transcriptome profiling of the murine testis during the first wave of spermatogenesis. PLoS One 8, e61558.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., et al. 2001. Initial sequencing and analysis of the human genome. Nature 409, 860-921.

Lau, N.C. 2010. Small RNAs in the animal gonad: guarding genomes and guiding development. Int.J.Biochem.Cell Biol. 42, 1334-47.

Lau, N.C., Seto, A.G., Kim, J., Kuramochi-Miyagawa, S., Nakano, T., Bartel, D.P., Kingston, R.E. 2006. Characterization of the piRNA complex from rat testes. Science 313, 363-7.

Le Hir, H., Izaurralde, E., Maquat, L.E., Moore, M.J. 2000. The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. EMBO J. 19, 6860-9.

Lecuyer, E., Yoshida, H., Parthasarathy, N., Alm, C., Babak, T., Cerovina, T., Hughes, T.R., Tomancak, P., Krause, H.M. 2007. Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. Cell 131, 174-87.

Lee, K., Haugen, H.S., Clegg, C.H., Braun, R.E. 1995. Premature translation of protamine 1 mRNA causes precocious nuclear condensation and arrests spermatid

differentiation in mice. Proc.Natl.Acad.Sci.U.S.A. 92, 12451-5.

Lee, Y., Kim, M., Han, J., Yeom, K.H., Lee, S., Baek, S.H., Kim, V.N. 2004. MicroRNA genes are transcribed by RNA polymerase II. EMBO J. 23, 4051-60.

Lerner, M.R., Boyle, J.A., Mount, S.M., Wolin, S.L., Steitz, J.A. 1980. Are snRNPs involved in splicing? Nature 283, 220-4.

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. Mol.Cell 50, 67-81.

Lim, J., Thomas, T., Cavicchioli, R. 2000. Low temperature regulated DEAD-box RNA helicase from the Antarctic archaeon, Methanococcoides burtonii. J.Mol.Biol. 297, 553-67.

Linder, P. and Jankowsky, E. 2011. From unwinding to clamping - the DEAD box RNA helicase family. Nat.Rev.Mol.Cell Biol. 12, 505-16.

Long, J.C. and Caceres, J.F. 2009. The SR protein family of splicing factors: master regulators of gene expression. Biochem.J. 417, 15-27.

Lyon, M.F. 1961. Gene action in the X-chromosome of the mouse (Mus musculus L.). Nature 190, 372-3.

Maekawa, M., Kamimura, K., Nagano, T. 1996. Peritubular myoid cells in the testis: their structure and function. Arch.Histol.Cytol. 59, 1-13.

Mali, P., Sandberg, M., Vuorio, E., Yelick, P.C., Hecht, N.B., Parvinen, M. 1988. Localization of protamine 1 mRNA in different stages of the cycle of the rat seminiferous epithelium. J.Cell Biol. 107, 407-12.

Malone, C., Brennecke, J., Czech, B., Aravin, A., Hannon, G.J. 2012. Preparation of small RNA libraries for high-throughput sequencing. Cold Spring Harb Protoc. 2012, 1067-77.

Martin, K.C. and Ephrussi, A. 2009. mRNA localization: gene expression in the spatial dimension. Cell 136, 719-30.

Martin, K.C. and Zukin, R.S. 2006. RNA trafficking and local protein synthesis in dendrites: an overview. J.Neurosci. 26, 7131-4.

Mattick, J.S. and Makunin, I.V. 2006. Non-coding RNA. Hum.Mol.Genet. 15 Spec No 1, R17-29.

Meikar, O., Da Ros, M., Korhonen, H., Kotaja, N. 2011. Chromatoid body and small RNAs in male germ cells. Reproduction 142, 195-209.

Meikar, O., Da Ros, M., Kotaja, N. 2012. Epigenetic regulation of male germ cell differentiation. Subcell.Biochem. 61, 119-38.

Meikar, O., Da Ros, M., Liljenback, H., Toppari, J., Kotaja, N. 2010. Accumulation of piRNAs in the chromatoid bodies purified by a novel isolation protocol. Exp.Cell Res. 316, 1567-75.

Meisner, N.C. and Filipowicz, W. 2011. Properties of the Regulatory RNA-Binding Protein HuR and its Role in Controlling miRNA Repression. Adv.Exp.Med.Biol. 700, 106-23.

Meister, G. 2013. Argonaute proteins: functional insights and emerging roles. Nat.Rev.Genet. 14, 447-59.

Meistrich, M.L., Mohapatra, B., Shirley, C.R., Zhao, M. 2003. Roles of transition nuclear proteins in spermiogenesis. Chromosoma 111, 483-8.

Mercer, T.R. and Mattick, J.S. 2013. Structure and function of long noncoding RNAs in epigenetic regulation. Nat.Struct.Mol.Biol. 20, 300-7.

Messina, V., Meikar, O., Paronetto, M.P., Calabretta, S., Geremia, R., Kotaja, N., Sette, C. 2012. The RNA binding protein SAM68 transiently localizes in the chromatoid body of male germ cells and influences expression of select microRNAs. PLoS One 7, e39729.

Metze, S., Herzog, V.A., Ruepp, M.D., Muhlemann, O. 2013. Comparison of EJC-enhanced and EJC-independent NMD in human cells reveals two partially redundant degradation pathways. RNA 19, 1432-48.

Meyer, S., Temme, C., Wahle, E. 2004. Messenger RNA turnover in eukaryotes: pathways and enzymes. Crit.Rev.Biochem.Mol.Biol. 39, 197-216.

Mignone, F., Gissi, C., Liuni, S., Pesole, G. 2002. Untranslated regions of mRNAs. Genome Biol. 3, REVIEWS0004.

Monesi, V. 1964. Ribonucleic Acid Synthesis during Mitosis and Meiosis in the Mouse Testis. J.Cell Biol. 22, 521-32.

Moore, M.J. 2005. From birth to death: the complex lives of eukaryotic mRNAs. Science 309, 1514-8.

Mor, A., Suliman, S., Ben-Yishay, R., Yunger, S., Brody, Y., Shav-Tal, Y. 2010. Dynamics of single mRNP nucleocytoplasmic transport and export through the nuclear pore in living cells. Nat.Cell Biol. 12, 543-52.

Muhlrad, D., Decker, C.J., Parker, R. 1994. Deadenylation of the unstable mRNA encoded by the yeast MFA2 gene leads to decapping followed by 5'-->3' digestion of the transcript. Genes Dev. 8, 855-66.

Nesterova, T.B., Barton, S.C., Surani, M.A., Brockdorff, N. 2001. Loss of Xist imprinting in diploid parthenogenetic preimplantation embryos. Dev.Biol. 235, 343-50.

Nishikura, K. 2010. Functions and regulation of RNA editing by ADAR deaminases. Annu.Rev.Biochem. 79, 321-49.

Okada-Katsuhata, Y., Yamashita, A., Kutsuzawa, K., Izumi, N., Hirahara, F., Ohno, S. 2012. N- and C-terminal Upf1 phosphorylations create binding platforms for SMG-6 and SMG-5:SMG-7 during NMD. Nucleic Acids Res. 40, 1251-66.

Ota, H., Sakurai, M., Gupta, R., Valente, L., Wulff, B.E., Ariyoshi, K., Iizasa, H., Davuluri, R.V., Nishikura, K. 2013. ADAR1 Forms a Complex with Dicer to Promote MicroRNA Processing and RNA-Induced Gene Silencing. Cell 153, 575-89.

Pall, G.S. and Hamilton, A.J. 2008. Improved northern blot method for enhanced detection of small RNA. Nat.Protoc. 3, 1077-84.

Pan, Q., Shai, O., Lee, L.J., Frey, B.J., Blencowe, B.J. 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat.Genet. 40, 1413-5.

Pang, K.C., Frith, M.C., Mattick, J.S. 2006. Rapid evolution of noncoding RNAs: lack of conservation does not mean lack of function. Trends Genet. 22, 1-5.

Parker, R. and Sheth, U. 2007. P bodies and the control of mRNA translation and degradation. Mol.Cell 25, 635-46.

Parker, R. and Song, H. 2004. The enzymes and control of eukaryotic mRNA turnover. Nat.Struct.Mol.Biol. 11, 121-7.

Paronetto, M.P., Messina, V., Bianchi, E., Barchi, M., Vogel, G., Moretti, C., Palombi, F., Stefanini, M., Geremia, R., Richard, S., Sette, C. 2009. Sam68 regulates translation of target mRNAs in male germ cells, necessary for mouse spermatogenesis. J.Cell Biol. 185, 235-49.

Paronetto, M.P. and Sette, C. 2010. Role of RNA-binding proteins in mammalian spermatogenesis. Int.J.Androl. 33, 2-12.

Parvinen, M. 2005. The chromatoid body in spermatogenesis. Int.J.Androl. 28, 189-201.

Parvinen, M. 1982. Regulation of the seminiferous epithelium. Endocr.Rev. 3, 404-17.

Parvinen, M. and Parvinen, L.M. 1979. Active movements of the chromatoid body. A possible transport mechanism for haploid gene products. J.Cell Biol. 80, 621-8.

Paul, J. and Duerksen, J.D. 1975. Chromatin-associated RNA content of heterochromatin and euchromatin. Mol.Cell.Biochem. 9, 9-16.

Pek, J.W., Anand, A., Kai, T. 2012. Tudor domain proteins in development. Development 139, 2255-66.

Pelletier, J. and Sonenberg, N. 1988. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. Nature 334, 320-5.

Penny, G.D., Kay, G.F., Sheardown, S.A., Rastan, S., Brockdorff, N. 1996. Requirement for Xist in X chromosome inactivation. Nature 379, 131-7.

Rajasethupathy, P., Antonov, I., Sheridan, R., Frey, S., Sander, C., Tuschl, T., Kandel, E.R. 2012. A role for neuronal piRNAs in the epigenetic control of memory-related synaptic plasticity. Cell 149, 693-707.

Rappsilber, J. 2011. The beginning of a beautiful friendship: cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. J.Struct.Biol. 173, 530-40.

Reuter, M., Berninger, P., Chuma, S., Shah, H., Hosokawa, M., Funaya, C., Antony, C., Sachidanandam, R., Pillai, R.S. 2011. Miwi catalysis is required for piRNA amplification-independent LINE1 transposon silencing. Nature 480, 264-7.

Rinn, J.L. and Chang, H.Y. 2012. Genome regulation by long noncoding RNAs. Annu.Rev.Biochem. 81, 145-66.

Romero, Y., Meikar, O., Papaioannou, M.D., Conne, B., Grey, C., Weier, M., Pralong, F., De Massy, B., Kaessmann, H., Vassalli, J.D., Kotaja, N., Nef, S. 2011. Dicer1 depletion in male germ cells leads to infertility due to cumulative meiotic and spermiogenic defects. PLoS One 6, e25241.

Rueter, S.M., Burns, C.M., Coode, S.A., Mookherjee, P., Emeson, R.B. 1995. Glutamate receptor RNA editing in vitro by enzymatic conversion of adenosine to inosine. Science 267, 1491-4.

Rufener, S.C. and Muhlemann, O. 2013. eIF4E-bound mRNPs are substrates for nonsense-mediated mRNA decay in mammalian cells. Nat.Struct.Mol.Biol. 20, 710-7.

Russell, L.D., Ettlin, R.A., Sinha Hikim, A.P., Clegg, E.D. 1990. Histological and Histopathological Evaluation of the Testis. 11, Cache River Press, St. Louis, United States.

Sato, H. and Maquat, L.E. 2009. Remodeling of the pioneer translation initiation complex involves translation and the karyopherin importin beta. Genes Dev. 23, 2537-50.

Saunders, P.T., Millar, M.R., Maguire, S.M., Sharpe, R.M. 1992. Stage-specific expression of rat transition protein 2 mRNA and possible localization to the chromatoid body of step 7 spermatids by in situ hybridization using a nonradioactive riboprobe. Mol.Reprod.Dev. 33, 385-91.

Schweingruber, C., Rufener, S.C., Zund, D., Yamashita, A., Muhlemann, O. 2013. Nonsense-mediated mRNA decay - Mechanisms of substrate mRNA recognition and degradation in mammalian cells. Biochim.Biophys.Acta 1829, 612-23.

Shandilya, J. and Roberts, S.G. 2012. The transcription cycle in eukaryotes: from productive initiation to RNA polymerase II recycling. Biochim.Biophys.Acta 1819, 391-400.

Shang, P., Baarends, W.M., Hoogerbrugge, J., Ooms, M.P., van Cappellen, W.A., de Jong, A.A., Dohle, G.R., van Eenennaam, H., Gossen, J.A., Grootegoed, J.A. 2010. Functional transformation of the chromatoid body in mouse spermatids requires testis-specific serine/threonine kinases. J.Cell.Sci. 123, 331-9.

Shirayama, M., Seth, M., Lee, H.C., Gu, W., Ishidate, T., Conte, D.,Jr, Mello, C.C. 2012. piRNAs initiate an epigenetic memory of nonself RNA in the C. elegans germline. Cell 150, 65-77.

Siomi, M.C., Sato, K., Pezic, D., Aravin, A.A. 2011. PIWI-interacting small RNAs: the vanguard of genome defence. Nat.Rev.Mol.Cell Biol. 12, 246-58.

Skakkebaek, N.E., Jorgensen, N., Main, K.M., Rajpert-De Meyts, E., Leffers, H., Andersson, A.M., Juul, A., Carlsen, E., Mortensen, G.K., Jensen, T.K., Toppari, J. 2006. Is human fecundity declining? Int.J.Androl. 29, 2-11.

Smith, Z.D. and Meissner, A. 2013. DNA methylation: roles in mammalian development. Nat.Rev.Genet. 14, 204-20.

Soderstrom, K.O. and Parvinen, M. 1976. Incorporation of (3H)uridine by the chromatoid body during rat spermatogenesis. J.Cell Biol. 70, 239-46.

Sonenberg, N. and Hinnebusch, A.G. 2009. Regulation of translation initiation in eukaryotes: mechanisms and biological targets. Cell 136, 731-45.

Song, R., Hennig, G.W., Wu, Q., Jose, C., Zheng, H., Yan, W. 2011. Male germ cells express abundant endogenous siRNAs. Proc.Natl.Acad.Sci.U.S.A. 108, 13159-64.

Soumillon, M., Necsulea, A., Weier, M., Brawand, D., Zhang, X., Gu, H., Barthes, P., Kokkinaki, M., Nef, S., Gnirke, A., Dym, M., de Massy, B., Mikkelsen, T.S., Kaessmann, H. 2013. Cellular source and mechanisms of high transcriptome complexity in the Mammalian testis. Cell.Rep. 3, 2179-90.

Steele, E.J., Lindley, R.A., Wen, J., Weiller, G.F. 2006. Computational analyses show A-to-G mutations correlate with nascent mRNA hairpins at somatic hypermutation hotspots. DNA Repair (Amst) 5, 1346-63.

Sud, N.B. 1961. Morphological and Histochemical Studies of the Chromatoid Body and Related Elements in the Spermatogenesis of the Rat. Quarterly Journal of Microscopical Science 102, 495-505.

Szostak, E. and Gebauer, F. 2013. Translational control by 3'-UTR-binding proteins. Brief Funct.Genomics 12, 58-65.

Taft, R.J., Pang, K.C., Mercer, T.R., Dinger, M., Mattick, J.S. 2010a. Non-coding RNAs: regulators of disease. J.Pathol. 220, 126-39.

Taft, R.J., Pheasant, M., Mattick, J.S. 2007. The relationship between non-protein-coding DNA and eukaryotic complexity. Bioessays 29, 288-99.

Taft, R.J., Simons, C., Nahkuri, S., Oey, H., Korbie, D.J., Mercer, T.R., Holst, J., Ritchie, W., Wong, J.J., Rasko, J.E., Rokhsar, D.S., Degnan, B.M., Mattick, J.S. 2010b. Nuclear-localized tiny RNAs are associated with transcription initiation and splice sites in metazoans. Nat.Struct.Mol.Biol. 17, 1030-4.

Tanaka, S.S., Toyooka, Y., Akasu, R., Katoh-Fukui, Y., Nakahara, Y., Suzuki, R., Yokoyama, M., Noce, T. 2000. The mouse homolog of Drosophila Vasa is required for the development of male germ cells. Genes Dev. 14, 841-53.

Tanaka, T., Hosokawa, M., Vagin, V.V., Reuter, M., Hayashi, E., Mochizuki, A.L., Kitamura, K., Yamanaka, H., Kondoh, G., Okawa, K., Kuramochi-Miyagawa, S., Nakano, T., Sachidanandam, R., Hannon, G.J., Pillai, R.S., Nakatsuji, N., Chuma, S. 2011. Tudor domain containing 7 (Tdrd7) is essential for dynamic ribonucleoprotein (RNP) remodeling of chromatoid bodies during spermatogenesis. Proc.Natl.Acad.Sci.U.S.A. 108, 10579-84.

Tang, W., Fei, Y., Page, M. 2012. Biological significance of RNA editing in cells. Mol.Biotechnol. 52, 91-100.

Thomson, T. and Lin, H. 2009. The biogenesis and function of PIWI proteins and piRNAs: progress and prospect. Annu.Rev.Cell Dev.Biol. 25, 355-76.

Toppari, J. and Parvinen, M. 1985. In vitro differentiation of rat seminiferous tubular segments from defined stages of the epithelial cycle morphologic and immunolocalization analysis. J.Androl. 6, 334-43.

Toyooka, Y., Tsunekawa, N., Takahashi, Y., Matsui, Y., Satoh, M., Noce, T. 2000. Expression and intracellular localization of mouse Vasa-homologue protein during germ cell development. Mech.Dev. 93, 139-49.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., Pachter, L. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat.Biotechnol. 28, 511-5.

Tsai, M.C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J.K., Lan, F., Shi, Y., Segal, E., Chang, H.Y. 2010. Long noncoding RNA as modular scaffold of histone modification complexes. Science 329, 689-93.

Tsai-Morris, C.H., Sheng, Y., Gutti, R.K., Tang, P.Z., Dufau, M.L. 2010. Gonadotropin-regulated testicular RNA helicase (GRTH/DDX25): a multifunctional protein essential for spermatogenesis. J.Androl. 31, 45-52.

Turner, J.M., Mahadevaiah, S.K., Ellis, P.J., Mitchell, M.J., Burgoyne, P.S. 2006. Pachytene asynapsis drives meiotic sex chromosome inactivation and leads to substantial postmeiotic repression in spermatids. Dev.Cell. 10, 521-9.

Vasileva, A., Tiedau, D., Firooznia, A., Muller-Reichert, T., Jessberger, R. 2009. Tdrd6 is required for spermiogenesis, chromatoid body architecture, and regulation of miRNA expression. Curr.Biol. 19, 630-9.

Ventela, S., Toppari, J., Parvinen, M. 2003. Intercellular organelle traffic through cytoplasmic bridges in early spermatids of the rat: mechanisms of haploid gene product sharing. Mol.Biol.Cell 14, 2768-80.

von Brunn, A. 1876. Beiträge zur Entwicklungsgeschichte der Samenkörper. Archiv für Mikroskopische Anatomie 12, 528–536.

Von Stetina, J.R. and Orr-Weaver, T.L. 2011. Developmental control of oocyte maturation and egg activation in metazoan models. Cold Spring Harb Perspect.Biol. 3, a005553.

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. Nat.Struct.Mol.Biol. 19, 773-81.

Walt, H. and Armbruster, B.L. 1984. Actin and RNA are components of the chromatoid bodies in spermatids of the rat. Cell Tissue Res. 236, 487-90.

Wang, E.T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P., Burge, C.B. 2008. Alternative isoform regulation in human tissue transcriptomes. Nature 456, 470-6.

Wang, K., Yuan, Y., Li, H., Cho, J.H., Huang, D., Gray, L., Qin, S., Galas, D.J. 2013. The spectrum of circulating RNA: a window into systems toxicology. Toxicol.Sci. 132, 478-92.

Weinberg, M.S., Villeneuve, L.M., Ehsani, A., Amarzguioui, M., Aagaard, L., Chen, Z.X., Riggs, A.D., Rossi, J.J., Morris, K.V. 2006. The antisense strand of small interfering RNAs directs histone methylation and transcriptional gene silencing in human cells. RNA 12, 256-62.

Weischenfeldt, J., Waage, J., Tian, G., Zhao, J., Damgaard, I., Jakobsen, J.S., Kristiansen, K., Krogh, A., Wang, J., Porse, B.T. 2012. Mammalian tissues defective in nonsense-mediated mRNA decay display highly aberrant splicing patterns. Genome Biol. 13, R35,2012-13-5-r35.

Will, C.L. and Luhrmann, R. 2011. Spliceosome structure and function. Cold Spring Harb Perspect.Biol. 3, 1-23.

Yadav, R.P. and Kotaja, N. 2013. Small RNAs in spermatogenesis. Mol.Cell.Endocrinol. in press.

Yajima, M. and Wessel, G.M. 2011. The multiple hats of Vasa: its functions in the germline and in cell cycle progression. Mol.Reprod.Dev. 78, 861-7.

Yang, E., van Nimwegen, E., Zavolan, M., Rajewsky, N., Schroeder, M., Magnasco, M., Darnell, J.E.,Jr. 2003. Decay rates of human mRNAs: correlation with functional characteristics and sequence attributes. Genome Res. 13, 1863-72.

Yoshida, S., Sukeno, M., Nakagawa, T., Ohbo, K., Nagamatsu, G., Suda, T., Nabeshima, Y. 2006. The first round of mouse spermatogenesis is a distinctive program that lacks the self-renewing spermatogonia stage. Development 133, 1495-505.

Zappulla, D.C. and Cech, T.R. 2006. RNA as a flexible scaffold for proteins: yeast telomerase and beyond. Cold Spring Harb.Symp.Quant.Biol. 71, 217-24.

Zheng, K. and Wang, P.J. 2012. Blockade of pachytene piRNA biogenesis reveals a novel requirement for maintaining post-meiotic germline genome integrity. PLoS Genet. 8, e1003038.