ELUCIDATING ROLES FOR AGO PROTEINS IN MALE MOUSE MEIOSIS

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Elizabeth A. Crate

August 2016

ABSTRACT

Meiosis is a specialized cell division during which genomic material of a diploid cell is halved to produce haploid gametes. During meiotic prophase I, autosomes become fully synapsed and undergo homologous recombination. In mammals, males are the heterogametic sex as their cells contain an X-Y chromosome pair. Despite the paring of these two chromosomes at their pseudoautosomal region, the rest of these chromosomes are asynapsed and transcriptionally silenced via Meiotic Sex Chromosome Inactivation (MSCI). Proper MSCI is essential for meiotic progression and gamete production in male mammals. However the mechanism underlying MSCI has not been fully elucidated. Recent work has implicated small non-coding RNA (sncRNA) binding partners, AGO3 and AGO4, in this process. My thesis work: 1. analyzes *Dgcr8* and *Dicer* conditional knockout mice, 2. optimizes a FACS method for prophase I staged spermatocyte isolation, 3. Defines a strategy for the development of *Ago3* and *Ago4* epitope tagged mouse lines.

BIOGRAPHICAL SKETCH

Elizabeth grew up in a military family, traveling the world and moving homes and schools. By the time she was a senior in high school, she had moved ten times. In 2009, she graduated with an honors diploma from AFNorth International School in the Netherlands. She then entered New College of Florida- the Honors College of Florida, planning to pursue a degree in Sociology. During her college summers, she worked in the Smithsonian Museum of Natural History as a botany intern and at Mississippi State University as a NSF-REU fellow. While at New College, she developed a strong interest in Developmental Biology. She successfully defended a thesis focused on the examination of sncRNAs in bull spermatozoa and graduated with a B.A. in Biology. Elizabeth continued to Cornell University, to pursue her interest in Developmental Biology through graduate work. In 2014, she joined the Cohen Lab and began a project examining the role of AGO proteins during male meiosis. During her time at Cornell, she was involved with various organizations including being President of the Graduate and Professional Women's Network, co-founding the Code-4-Kids program at Belle Sherman Elementary School, working as a blogger for Cornell's Life on the Hill Blog, and peer counseling with EARS. She plans to incorporate her Genetics experience from Cornell into a career as a Genetic Counselor, focusing on prenatal genetic diagnosis.

ACKNOWLEDGMENTS

I would first like to thank my mentor, Dr. Paula Cohen, for all of her support and guidance during my time at Cornell University. Thank you for always being available to discuss issues that came up, even when you were busy. Your strength as a woman scientist is something I will always admire and aspire to. I also appreciate all of the mentorship I have received from Dr. Andrew Grimson. It has been a pleasure collaborating with you and your lab. I enjoyed all your clear and thoughtful perspectives during meetings. Thank you also to Dr. Mariana Wolfner; the classes I took with you were my favorite during my time at Cornell and helped me grow as a scientist. Finally, I want to thank all three of you, for being a part of my thesis committee and having patience and understanding as I introspectively searched for my scientific interests. I could not have asked for a better committee.

It has been a life-changing experience being a part of the Cohen lab, starting with the Von Trapp Halloween during my lab rotation to the delicious lab treats at lab meetings. I want to thank all the members of the Cohen Lab for their critique and support. It was great getting to meet a group of such strong personalities. Each one of you, Dr. Stephen Gray, Dr. Miguel Angel Brieno Enriquez, Dr. Kate Grive, Dr. Kim Holloway, Pete Borst, Kadeine Campbell-Peterson, Melissa Toledo, Carolyn Milano, and all the undergraduates, have so many amazing qualities that I will miss!

From the Grimson Lab, I also want to thank Dr. Caterina Schweindenback. It was amazing getting to work so closely with you over the past year. Your encouragement definitely helped me to continue working when things got challenging. Your superb reasoning skills are

something I hope to develop over time. Thank you also to Dr. Stephanie Hilz; it was great working with you during my lab rotation and through all the collaborating projects.

I am also appreciative for my sources of funding: the Eunice Kennedy Shriver National Institute of Child Health and Development P50 Center Grant award and the NIH training grant in Genetics, Genomics, and Development.

Last but not least, I would like to thank my family. John and Mark, I love you both so much. I hope that you both find your passions and work hard to pursue them. Mom, you are the strongest and kindest woman I know. It takes a woman with incredible resilience and adaptability to successfully raise three kids overseas, often alone, and you did it with perfection. There is no one more self-less than you, thank you for listening to me during this whole process. Dad, thank you for always being a voice of reason and clear-minded. And to Arthur and Mochi, you two guys have been my comedic relief and source of love during this process. Thank you!

TABLE OF CONTENTS

CHAPTER 1

Introduction to Male Meiosis and Small RNAs in gametogenesis

Introduction

Meiosis is a hallmark of sexual reproduction. Errors during meiosis can lead to numerous developmental disorders, due to aneuploidy, and infertility [1]. The predominant feature of meiosis is the exchange of genetic material between maternal and paternal chromosomes. The genetic exchange of meiosis is mediated by two crucial processes: synapsis, or the physical pairing of homologous chromosomes through protein:protein interactions, and recombination, the formation of DNA crossovers between homologs [2]. These prophase I events are largely conserved between the sexes in mammals, but there are several regulatory processes that differ between male and female meiosis. Most significantly, cells from heterogametic sex (in this case, males) have a pair of chromosomes, X and Y, that cannot undergo full synapsis along the chromosomes' length due to lack of homology. Therefore, the unpaired regions of the XY chromosomes present a unique situation for the cell.

Meiotic sex chromosome inactivation (MSCI) is a fundamental process by which asynapsed sex chromosomes in the heterogametic sex are transcriptionally silenced [3]. Failure to properly silence the X and Y chromosomes in mice leads to cell death and failed meiosis in the germ cells, highlighting the importance of MSCI for meiotic progression and gametogenesis [4]. While some of the proteins mediating this silencing have been identified [5, 6], the precise mechanism by which the genes on the sex chromosomes are silenced is currently unknown.

Recent work has shown that the small non-coding RNA (sncRNA) binding proteins, AGO4 and AGO3, localize to the sex chromosomes in mouse spermatocytes during prophase I

[7]. Furthermore, loss of AGO4 in mouse spermatocytes leads to defects in MSCI and an influx of RNA Polymerase II into the sex body, the nuclear sub-domain that houses the silenced X and Y [7]. Prior to these findings, it had been thought that mammalian AGO proteins functioned strictly within the cytoplasm to mediate post-transcriptional gene silencing in conjunction with small non-coding RNAs (sncRNAs) [8]. Therefore, the observations that AGO4 is present in the nucleus of spermatocytes raise the interesting possibility that mammalian AGO proteins have a novel nuclear role necessary for proper MSCI during prophase I of meiosis.

Defining events of Prophase I

Meiosis is germline-specific and is well conserved across organisms. Meiosis is characterized by two cellular divisions, meiosis I and meiosis II, that result in viable haploid gametes. Meiosis I begins following premeiotic DNA replication and consists of four stages: prophase I, metaphase I, anaphase I, telophase I [2]. Prophase I is the first and longest stage, during which homologous chromosomes, or homologs, pair, while anaphase I, the final stage, is the point at which the homologs are equally segregated. Prophase I is divided into five substages: leptonema, zygonema, pachynema, diplonema, and diakinesis, as can be seen in Figure 1. These substages are characterized by the state of pairing and synapsis of the homologous chromosomes. In most organisms, any mispairing of homologs or recombination errors will result in apoptosis; so proper prophase I progression is a necessary first step in the larger process of gametogenesis [9].

Homologous chromosome alignment and pairing

During prophase I of meiosis, homologous chromosomes must first align and pair. "Alignment," refers to the process of bringing the homologous chromosomes in closer proximity

along their lengths, while "pairing," occurs when homologs are more intimately associated [10]. Homologous chromosome pairing can generally be classified as double strand break (DSB)-dependent or DSB-independent [11]. The mechanism underlying homolog pairing varies considerably between organisms and is independent of synapsis and recombination [11, 12].

In most organisms, DSB-dependent homolog pairing is initiated by the formation of chromosome breaks. The free 3' DNA ends can then invade potential homologous regions in search of a match. In mammals, a DSB-dependent mechanism appears to be essential for correct homolog pairing, despite the presence of certain DSB-independent mechanisms [10]. Conversely, in *Drosophila melanogaster* and *Caenorhabditis elegans,* DSB-independent mechanisms are sufficient for homolog pairing [13, 14].

Telomere clustering and bouquet formation is one mechanism of DSB-independent homolog recognition that facilitates the homolog search in some organisms, including *Schizosaccharomyces pombe* and mammals [2]. The process of chromosome pairing begins with the alignment of homologous chromosomes in a nearly parallel fashion from the decondensed cloud of chromatin [15]. This alignment is thought to be mediated by associations between the chromosomes and the nuclear envelope during early meiosis I. Through this interaction, the telomeres of the chromosomes cluster to one region of the nuclear envelope and the chromosomes fan out from a single location, in a manner loosely resembling a flower bouquet, giving rise to the term telomere bouquet [16]. Anchoring of the chromosomes to the nuclear envelope is mediated by protein-protein interactions between two proteins, SUN and KASH [16]. SUN-KASH complexes form a nuclear envelope bridge that connects the cytoplasmic microtubules to the intranuclear chromosomes [17]. In fission yeast, the attachment of the telomeres to the spindle pole body is followed by rapid oscillatory chromosome movements,

powered by dynein motors, which aid in homolog pairing [18]. While intranuclear movements have been visualized in mammals, their importance is less well understood [19, 20].

D. melanogaster, C. elegans, and *S. pombe* also utilize non-canonical methods of homolog pairing. In *D. melanogaster* males, there is a lack of recombination and synaptonemal complex formation between homologs during meiosis. Interestingly, it is thought that homolog pairing in *D. melanogaster* males begins somatically, prior to meiosis, and the homologs remain paired during entry into meiosis [13]. *C. elegans* and *D. melanogaster* utilize pairing centers, *cis*-acting sequences that function as centers for chromosome pairing [10]. *S. pombe* utilizes a meiosis specific small non-coding RNA, called meiRNA, and an RNA-binding protein, Mei2 [21]. The noncoding RNA, which originates from the *sme2* locus, is believed to be an important mediator of chromosome pairing at the *sme2* locus [21].

Synapsis is necessary for proper meiotic DSB processing

Synapsis is defined as the physical tethering of the homologous chromosomes through the formation of a proteinaceous structure called the synaptonemal complex (SC) between homologous chromosomes. Synapsis begins at leptonema and completes at pachynema, as can be seen in Figure 1 [22]. For most organisms, double strand breaks are needed prior to SC assembly [23]. This is true in mouse; however in certain organisms, like in *Drosophila melanogaster* females, the SC is able to form in the absence of DSBs. However, *D. melanogaster* male meiosis also lacks homologous recombination, so the process of meiosis is carried out in a different manner in this organism [24].

In mouse, SC assembly begins at leptonema with formation of a lateral element (LE) along each sister chromatid pair. The LE is composed of the synaptonemal complex proteins 3 and 2 (SYCP3 and SYCP2) [25]. It is thought that meiotic cohesion proteins, like Rec8, facilitate

the loading of SYCP3 and SYCP2 and correct lateral element formation through the maintenance of sister chromosome cohesion and by providing a scaffold for the lateral element assembly [26]. In mouse, loss of either *Sycp2* or *Sycp3* through mutagenesis leads to significant impairments in synapsis, DSB repair, and recombination [25, 27]. In both cases, males exhibit sterility due to a block in meiotic progression [25, 28]. Intriguingly, the *Sycp2* mutant mouse, whose SYCP2 lacks the coil-coiled domain, displays a loss of SYCP3 protein localization to the chromosomes, despite the localization of mutant SYCP2 to the chromosomes [25]. These results suggest that the coil-coiled domain of SYCP2 is required for SYCP3 localization to chromosomes.

The initiation of synapsis can vary greatly between organisms, and between the sexes within the same species. In human males, for example, synapsis begins at the telomeres and the eventual chiasmata are also closer to the chromosome ends as compared to females [29].

The physical synapsis of the homologous chromosomes begins with the formation of three distinct structures. Transverse filaments extend from the lateral element of each homolog and meet in the middle in a structure called the central element (CE) [2]. The transverse filament in mammals is composed of SYCP1, with the C terminus being embedded in the lateral element and the N terminus both associating with the CE and other SYCP1 subunits [22]. In male mice, loss of *Sycp1* results in normal lateral element formation but failure to synapse and an absence of central element protein localization to the homologs [22]. Therefore, CE formation requires SYCP1. Two central element proteins, SYCE1 and SYCE2, bind SYCP1 and are presumed to act as a linker between the transverse filament and central element [30]. Two other central element proteins, TEX12 and SYCE3, are also thought to be recruited by SYCP1 [31]. However, TEX12 does not appear to directly interact with SYCP1, but rather associates with SYCP1 through complex formation with SYCE2. This complex formed by TEX12 and SYCE2 is

believed to be more structural in nature. Using a zipper analogy, the TEX12 and SYCE2 complex may act as the slider on a zipper, while the SYCP1-SYCE1-SYCE3 complex would act as the teeth [2]. This suggests that the SYCE2/TEX12 complex is the final complex to load right in the center of the CE and is essential in the completion of synapsis and the extension of synapsis along the chromosome length [31]. Thus, in *Syce2* and *Tex12* mutants, loss of either protein prevents the maturation of the SC central region, resulting in only short CE-like structures made up of SYCP1, SYCE1, and SYCE3 [30, 32].

FK506 binding protein, FKBP6, localizes to the CE, and also appears to be important for proper synapsis; however, FK506's direct role is still unknown. In FKBP6 null mutant males, there is misalignment and abnormal pairing of chromosomes [33]. Using mice with mutations in the central element proteins, the order of central element loading has been determined as SYCE3 loading onto SYCP1, followed by SYCE3 recruiting SYCE1. These three initial CE proteins are essential for synapsis, and a loss of any of them will prevent CE formation. The formation of the SC decreases the distance between the homologous chromosomes through the aid of many proteins, as described above.

Homologous recombination produces crossovers and non-crossovers

One of the ways by which homologous chromosomes remain paired until the first meiotic division is through DNA interactions between the maternal and paternal chromosomes. Homologous recombination is the physical exchange of DNA between the homologous chromosomes. Recombination is initiated by the formation of DSBs in early prophase I [10]. In most organisms, spo11, a topoisomerase-like protein, is responsible for cutting the DNA during leptonema at specific hotspot sites, as shown in Figure 2A [2]. Following the cut, the 5' to 3' end is resected from each strand, leaving a single stranded 3' overhang that can then invade a

homologous DNA molecule. This single strand invasion is mediated by the single strand exchange proteins, RAD51 and DMC1 [34]. Strand invasion into the homolog results in displacement of DNA, creating a structure call the displacement-loop, or D-loop. The invading DNA then extends by DNA synthesis, and the newly synthesized DNA strand is ligated back to the original DNA end from which it initially broke, in a process known as second end capture (Figure 2C) [35]. If the D-loop is unwound prior to the second end capture and recaptures its own DNA arm, then a noncrossover (NCO) will occur through a process called synthesis dependent strand annealing (SDSA) (Figure 2B) [36]. If the D-loop is not unwound, a double Holliday Junction (dHJ) forms, and depending on the way in which the strands are cleaved at each junction, a crossover (CO) or noncrossover (NCO) can occur, with a bias for these DHJ to become COs (Figure 2C) [37]. A dHJ is considered to be resolved as a crossover if there is a reciprocal exchange of the DNA arms between homologs (Figure 2E) [38]. This process of crossing over is mediated by the ZMM proteins, identified in *Saccharomyces cerevisiae* [39].

In mammals, the processing of dHJ to generate a crossover is thought to be partly mediated by MutS homologs 4 and 5, MSH4/5, and MutL homologs MLH1 and MLH3, along with others proteins including RNF212, CNTD1, and HEI10 [2, 37]. These two heterodimer complexes, MutSγ and MutLγ, are thought to act sequentially, narrowing the large number of initial DSBs to the eventual crossover sites [40]. A subset of MutSγ sites will become crossovers, while all MutLγ sites will become crossovers. It has also been shown that some crossover intermediates form crossovers by the endonuclease, MUS81-EME1, through mechanisms that do not require the MutS or MutL heterodimers (Figure 2F) [41]. Crossing over is a reciprocal exchange of DNA between the maternal and paternal homologous chromosomes, which creates physical connections between the homologs, known as chiasmata. The chiasmata physically hold

the homologs together at the metaphase plate and reduce the chances of aneuploidy at the first meiotic division. Every chromosome pair is required to have at least one crossover, known as an obligate crossover [36]. However, COs only make up a small proportion of the total recombination events that occur during meiosis I [35]. Many DSBs are repaired as NCOs. NCOs can be repaired by interhomolog recombination, that does not involve the reciprocal exchange of DNA seen in a crossover, or by sister chromatid recombination [34, 35].

Sexual Dimorphism in Prophase I in mammals

The Y chromosome is sex determining in mammals

In mammals, primary sex determination is dependent on chromosomal content, due to the fact that the Y chromosome encodes a testis-determining factor, called *Sry* (sex determining region of the Y chromosome) [42]. For example, a mammal with two X chromosomes and one Y chromosome would develop a male gonad, and a mammal with two X chromosomes but no Y chromosome would develop a female gonad. When wildtype XX mouse zygotes are injected with *Sry* DNA, some of the embryos develop testes and male accessory organs [43]. These mice do not produce viable sperm however, due to the fact that they are missing other crucial Y-encoded spermatogenesis genes [43].

In mammals, the initial production of sperm or oocytes begins following the differentiation of the gonad. The primordial germ cells, or PGCs, migrate to the genital ridge, the precursor of the male or female gonad, and proliferate and undergo germ cell licensing [44, 45]. In males, the germ cells are mitotically arrested at G1, while in females, the germ cells begin meiosis and arrest at the diplotene stage of prophase I [45]. Meiosis is a sexually dimorphic process [46]. In mammalian females, meiosis is initiated once and semi-synchronously generates a finite pool of

gametes; while in males, meiosis is continuously initiated and germ cells are produced throughout the organism's life [2]. Although the timing of gamete production is considerably different between the sexes, the metabolic signals (retinoic acid and stimulated by retinoic acid 8, Stra8) for meiotic initiation are the same in both sexes [47, 48].

Meiotic silencing in the heterogametic sex is crucial for bypass of the synapsis checkpoint

Meiotic Silencing is a process in which asynapsed chromatin is transcriptionally silenced during prophase I. This process occurs in many organisms, including fungi, birds, flies and mammals [49-52]. There are two types of meiotic silencing in mammals: meiotic silencing of unpaired chromatin (MSUC) and meiotic sex chromosome inactivation (MSCI). MSCI is thought to be a more specific process of MSUC [49]. MSUC occurs when autosome pairs fail to synapse during meiosis and the cell initiates a pathological response to silence the autosomes. In male mammals, the sex chromosomes (XY) only contain a small region of homology called the pseudoautosomal region (PAR). Thus in pachynema, only the PAR will undergo synapsis, while the remainder of the X and Y chromosomes remain asynapsed and thus are transcriptionally silenced by the MSCI mechanism.

MSCI is thought to be needed to prevent the unsynapsed chromosomes from triggering a synapsis checkpoint, to inhibit retrotransposon mobilization and to silence lethal genes from the Y chromosome, including the zinc finger genes *Zfy1* and *Zfy2* [4, 53]. It has also been suggested that MSCI has evolved as a means of suppressing recombination of the non-homologous sex chromosomes [52]. Through the mechanism of MSCI, the unsynapsed X and Y are shielded during pachynema by the formation of a heterochromatic silenced domain in which the X and Y chromosomal DNA reside. This domain, called the sex body can be visualized

using immunofluorescence imaging with antibodies that recognize the modified histones and/or proteins that mediate the silencing, like γH2AX and MDC1 [54].

MSCI is an essential mechanism in primary spermatocytes of male mammals; if it fails, germ cells will undergo apoptosis. Spermatocytes have a mechanism of apoptotic elimination if there are unsilenced X and Y chromosomes and a sex body does not form [55]. This apoptotic elimination is driven by the Y-linked genes *Zfy1* and *Zfy2*, which encode transcription factors [4]. Ectopic expression of *Zfy1/2*, even at low levels, results in apoptosis.

The proposed process of MSCI process involves two groups of proteins: sensors, which sense the asynapsis of the XY and bind at the axial elements, and effectors, which induce silencing at the chromatin loops, Figure 3 [5]. Two Horma domain proteins, HORMAD1 and HORMAD2, as well as breast cancer I protein, BRCA1, act as sensors, with BRCA1 being recruited in a HORMAD1/2 and SYCP3 dependent manner to sites of double strand breaks and asynapsed chromatin during zygotene [5]. ATR, a DNA damage checkpoint related kinase, is then recruited to these sites in a sensor-dependent manner [5]. Following recruitment, ATR recruits more BRCA1 at the axial elements and induces phosphorylation of the HORMADs. This asynapsis signal remains through late zygotene at both the asynapsed autosomes and the sex chromosomes. At pachytene, when all the autosomes have synapsed and only the male sex chromosomes remain unsynapsed, BRCA1 and ATR are no longer found on the autosomes, but the entire length of the asynapsed XY axial elements are coated with ATR and BRCA1, as seen in Figure 3 [3]. ATR then spreads across the XY chromatin loops where it phosphorylates the histone H2AX to form γH2AX [5]. γH2AX and the DNA damage checkpoint 1 protein, MDC1, act as effectors to mediate silencing within the chromatin loops through a positive feedback loop and heterochromatin formation [56]. However, the precise biochemical mechanism

driving transcriptional silencing following the H2AX histone modifications has not been fully elucidated.

Role of small non-coding RNA in gametogenesis

Basics of canonical sncRNA biology

Small non-coding RNAs (sncRNAs) have be implicated in multiple roles throughout meiotic prophase I and may play other undescribed roles. The three major classes of sncRNAs are piwi-interacting RNAs (piRNAs), small interfering RNAs (siRNAs), and microRNAs (miRNAs). These RNAs are defined by their interaction with an Argonaute protein. piRNAs bind the subclass of Argonaute proteins called the PIWI proteins, while miRNAs and siRNAs bind AGO proteins. While piRNAs are known to silence retrotransposons in the germline, miRNAs and siRNAs are traditionally found in the cytoplasm where they are involved in post-transcriptional regulation. However, nuclear roles for miRNAs and siRNAs have been described in some species [57].

The canonical role of microRNAs is to regulate many endogenous target messenger RNAs, mainly through translational repression [58]. miRNAs are single-stranded and typically 19-25 nucleotides in length [59]. Genes encoding miRNAs are found throughout the genome but about half are encoded on the introns of host genes [60]. miRNA processing is complex and differs between species. In the initial steps of mammalian miRNA biogenesis, miRNA genes are transcribed by RNA polymerase II or III into primary miRNA transcripts, or pri-miRNAs, that form hairpin structures. The pri-miRNAs are processed by a microprocessor complex found in the nucleus, containing the RNase III enzyme, Drosha, and DGCR8 [61]. DGCR8 determines the precise cleavage sites by interacting with the pri-miRNA, and the two RNase domains on Drosha

cleave the 3' and 5' arms of the hairpin [61]. The Drosha/DGCR8 processing results in a pre-miRNA which is approximately 70 nucleotides long [62]. The pre-miRNA is then exported into the cytoplasm and cleaved by the endoribonuclease Dicer and its RNA-binding partner, transactivating region binding protein (TRBP), resulting in a double-stranded mature miRNA about 22 nucleotides in length [61]. Following this cleavage, one of the miRNA strands is preferentially loaded into an AGO protein. This complex, along with cofactors forms the RNA induced silencing complex (RISC) that mediates mRNA repression. The RISC complex recognizes target mRNAs via base pairing with the bound miRNA. Specifically, a region of the miRNA termed the seed sequence, typically nucleotides 2-8 at the 5' end of the miRNA, forms a duplex with target mRNAs [63]. miRNAs are thought to play a role in the deadenylation and translational inhibition of mRNAs.

Another class of sncRNAs is short interfering RNAs (siRNAs). Like miRNAs, siRNAs are processed by Dicer and are bound to a RISC. However, unlike miRNAs, siRNAs are derived from double-stranded RNA rather than RNA hairpins, and siRNA biogenesis is Drosha-independent. An important function of siRNAs is to silence transposable elements, and typically, target RNAs must contain exact complementarity to the siRNA [59]. siRNAs play a role in the degradation of targets, either transposons or mRNAs [61].

Small RNAs act to repress target mRNAs through interaction with Argonaute proteins. One subfamily of the Argonaute proteins is the AGO proteins, which includes AGO1-4 in mammals [8]. The four mammalian AGO proteins show high levels of sequence similarity. Of these four, only AGO2 has the ability to cleave target mRNAs [64, 65]. These AGO proteins have a very highly conserved domain, called the Piwi-Argonaute-Zwille (PAZ) domain that binds the 3' end of associated sncRNAs and mediates the mRNA-sncRNA interactions [66].

AGO proteins have only recently been found in the nucleus of mammalian cells [7]. They have been known to play important roles in the nuclei of other organisms, including RNA Polymerase II control, alternative splicing, double strand break repair, and chromatin modifications; some of these known roles can be seen in Figure 4 [8, 62, 67, 68].

Small RNA localization in human/mouse germ cells

Primordial germ cells (PGCs) are the progenitor cells of the oocyte and spermatocyte. Through somatic divisions their numbers are multiplied, and they migrate to the eventual gonad. One of the prime features of PGCs in flies and amphibians is the presence of a germ plasm [42]. The germ plasm contains germline specific proteins and mRNAs, as well as sncRNAs [69]. Within the spermatid, there exists a chromatoid body (CB) that is packed with RNAs including mRNAs, piRNAs, and miRNAs, and it is similar to the germ plasm in other organisms [70].These sncRNAs are thought to play a range of roles, including protection of the genome, maintenance of PGC fate, and involvement in germline differentiation [71]. Although mammals utilize inductive processes rather than germ plasm to specify the differentiation of germ cells, sncRNAs still play critical roles in the germline.

In mammals, small non-coding RNAs are known to play crucial roles in germ cells by controlling the expression of developmental genes and protecting the genome from transposable elements [72]. More specifically, the repressive methylation of retrotransposons has been linked to piRNAs in fetal male germ cells [72]. The role of miRNAs and siRNAs in germ cells has been explored by mutating specific miRNAs or proteins involved in sncRNA biogenesis. A loss of Dicer, which is essential for the biogenesis of both miRNAs and siRNAs, results in fertility defects in males and females mice due to incomplete gametogenesis [73, 74].

Proposed roles for small RNAs in prophase I of mouse spermatocytes

Small non-coding RNAs play various functions during meiosis, and more specifically, during prophase I. For example, in *S. pombe*, a noncoding RNA called meiRNA along with an RNA binding protein Mei2 are essential for entry into meiosis and enhance homolog pairing specifically at the sme2 locus [21]. However, in mammals there is no evidence for the use of pairing sites as a mechanism of homolog pairing. Another possibility is that small RNAs may be active in DSB repair. In both *Arabidopsis* and human somatic cells, there is evidence that small RNAs about 21 nt in length are derived from sites of DSBs and in conjunction with DICER and AGO2, these small RNAs are able to facilitate the repair of DSBs [67]. Could a similar process be active during meiosis, when there is an abundance of double strand breaks?

Small RNAs and their machinery have been shown to participate in transcriptional silencing through an RNA Induced Transcriptional Silencing (RITS) complex or a similar processing complex in several organisms, including *S. pombe* and *C. elegans* [75]. In *S. pombe*, siRNAs, as part of the RITS complex, direct the methylation of histone H3 at lysine 9, H3K9, leading to the formation of heterochromatin [76]. During the progression of RITS processing, an RNA-directed RNA Polymerase Complex (RdRC), which also contains an AGO protein, is shuttled to sites of non-coding RNAs that are transcribed from repeats within the centromere to produce siRNAs [77]. This creates a feedback loop, feeding chromatin-derived siRNAs back to the chromatin to further initiate heterochromatin formation. Although some of the components at play in the RITS complex, such as AGO proteins, have been discovered to be present in the mammalian nucleus, a functional RITS complex in mammals has not yet been elucidated [78].

In mouse, it has been shown that a loss of DICER and DGCR8 in the spermatocyte at prophase I results in disregulation of ATM and sequestration of essential proteins from the sex

body [79]. This sequestration results in a decrease in the chromosomal stability of the sex chromosomes and the presence of chromosomal fusion events [79]. These results bring to question what roles small RNAs and their machinery (AGO proteins, DICER, and DGCR8) may be playing in the spermatocyte nucleus during meiosis and MSCI. While there is no current evidence that small RNAs are derived from sites of meiotic DSBs, there is some controversial evidence that a large number of X-linked miRNAs are able to evade MSCI and are active in the spermatocyte nucleus [80]. In addition, RNA FISH experiments have shown that miRNAs localize to the sex body during prophase I [54, 81, 82]. Might these X-linked miRNAs escape MSCI to play an active role during meiosis? It has been shown in mice that AGO4 localizes to the spermatocyte nucleus during prophase I and that a loss of AGO4 during prophase I leads to premature meiotic initiation and disruption of MSCI [7]. Similarly, it has been shown that a loss of DICER and DGCR8 in the spermatocyte at prophase I results in disregulation of ATM and sequestration of essential proteins from the sex body [79]. This sequestration results in a decrease in the chromosomal stability of the sex chromosomes and the presence of chromosomal fusion events [79]. These results raise the question of what roles small RNAs and their machinery (AGO proteins, DICER, and DGCR8) may be playing in the spermatocyte nucleus during meiosis and MSCI.

Conclusion

The process of meiosis is highly regulated and crucial for the production of viable, haploid gametes. Prophase I is the longest stage of meiosis I, and during this stage two essential processes occur: homologous chromosome synapsis and recombination. During the pachytene stage of meiotic prophase I, all chromosomes synapse at sites of homology, including the X and

Y-chromosomes at the PAR, however the remaining length of sex chromosomes remains asynapsed. At pachytene, asynapsed sex chromosomes are transcriptionally silenced through a mechanism called MSCI. Sex chromosomes that fail to be silence result in failed meiosis, so this is an essential process in the formation of male gametes. My work is concentrated on understanding what elements are controlling this process of MSCI.

The work in this thesis is focused on examining the role of sncRNAs and associated proteins during male meiotic prophase I. The experiments described in this thesis will examine the roles of AGO3 and AGO4 within the spermatocyte nucleus through the characterization of protein and sncRNA binding partners. Chapter 2 describes work examining male meiosis in the absence of two proteins required for miRNA and siRNA biogenesis, DICER and DGCR8. Chapter 3 describes the optimization of a fluorescence activated cell sorting (FACS) method for the isolation of prophase I staged mouse spermatocytes, providing the ability to analyze protein and RNA differences between the substages of prophase I spermatocytes. Finally, Chapter 4 outlines the method of generation of epitope tagged *Ago3* and *Ago4* mouse lines utilizing CRISPR/CAS9 technology. These experiments will provide insight into the function of AGO proteins in MSCI and meiotic progression in male mammals. Ultimately, this work will improve our understanding of male mammalian gametogenesis.

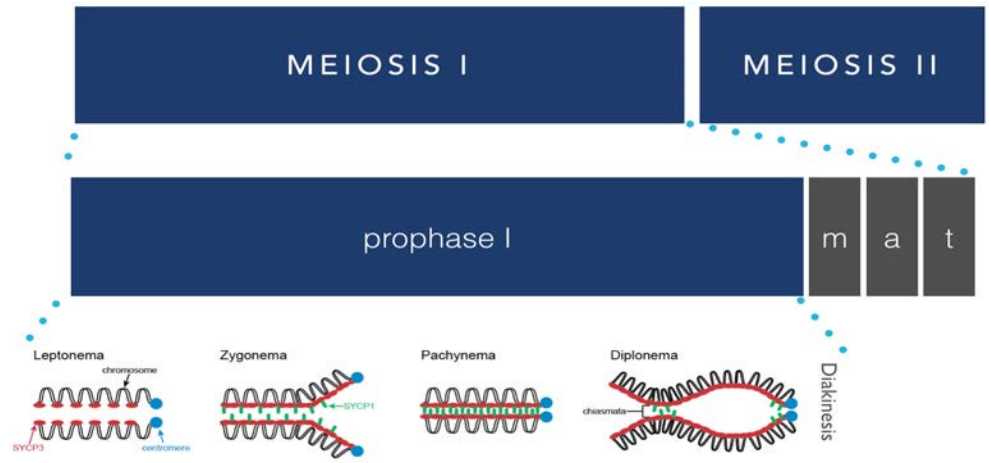


Figure 1: Simplified diagram of male meiosis in mouse. Prophase I can be characterized by five sub-stages. These can be visualized using immunofluorescence against synaptonemal complex proteins. In leptonema, lateral elements assemble along the homologous chromosomes. Then, there is initial assembly of central elements in zygonema. At pachynema, there is full synapsis of homologous chromosomes, except at the sex chromosomes in males, where there is only synapsis of the PAR. At diplonema, there is degradation of synaptonemal complex proteins except at points of chiasmata, or DNA crossovers. And in the final stage, diakinesis, the chromosomes remain tethered at points of chiasmata, while the synaptonemal complex proteins are no longer present on the homologs [1]. (Adapted from figure by Lipkin et al., 2002).

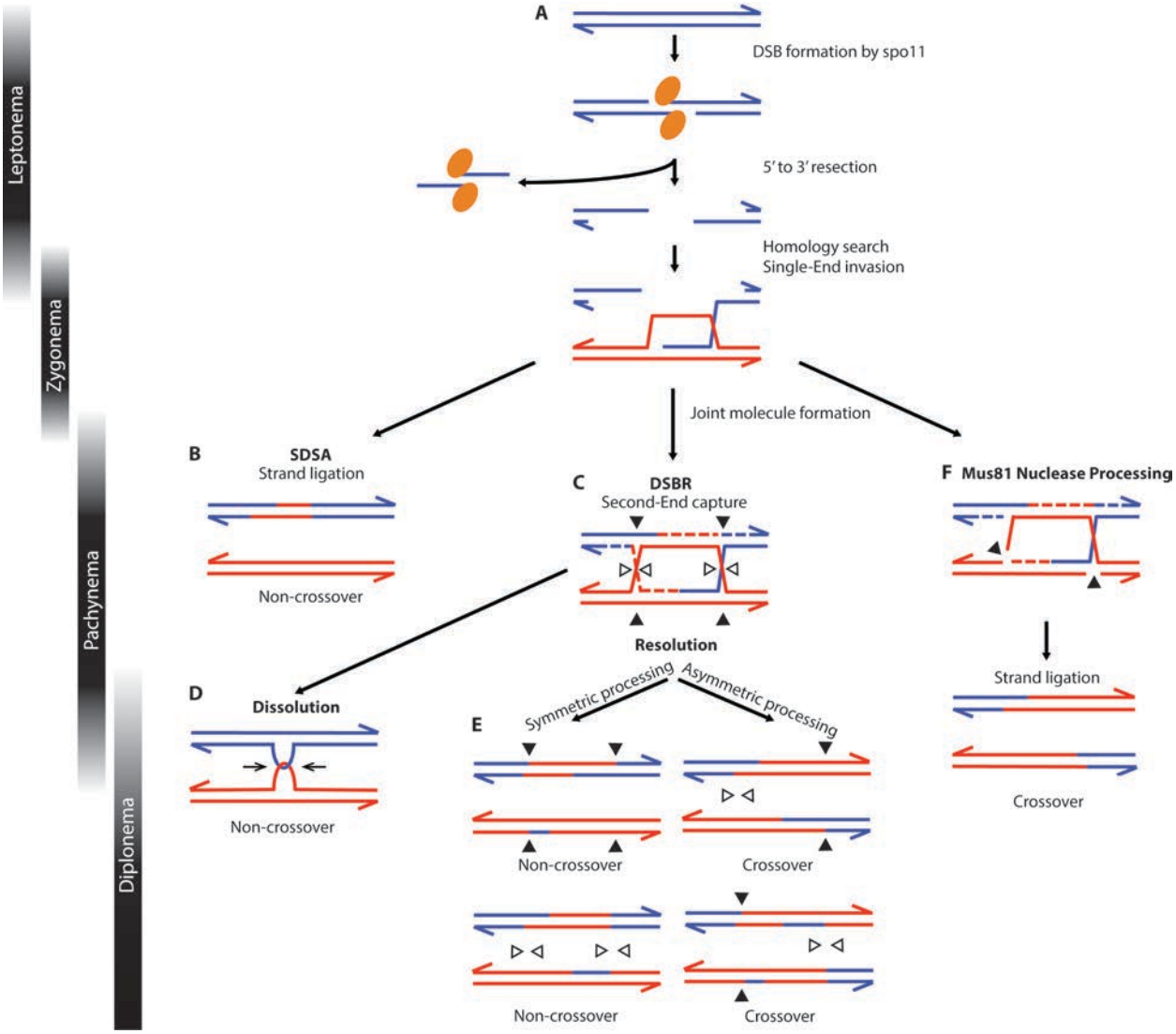


Figure 2: Meiotic Recombination. Meiotic recombination can result in crossovers and non-crossovers. This process begins with the formation of a DSB by spo11, then the 5' to 3' resection of strands, followed by homology search by strand invasion creating a D loop (**A**). Synthesis dependent strand annealing (SDSA) occurs when the D loop in unwound, creating a non-crossover (**B**). Double strand break repair (DSBR) through second-end capture (**C**) can occur through dissolution (**D**), which results in a noncrossover. Or, second-end capture (**C**) can result in resolution (**E**), to create a non-crossover or crossover depending on the orientation of processing of the double Holliday Junction. Another endonuclease, Mus81, can process crossovers (**F**) independent of the MutS or MutL heterodimers required for DSBR. (Adapted from figure by Gray and Cohen, in review).

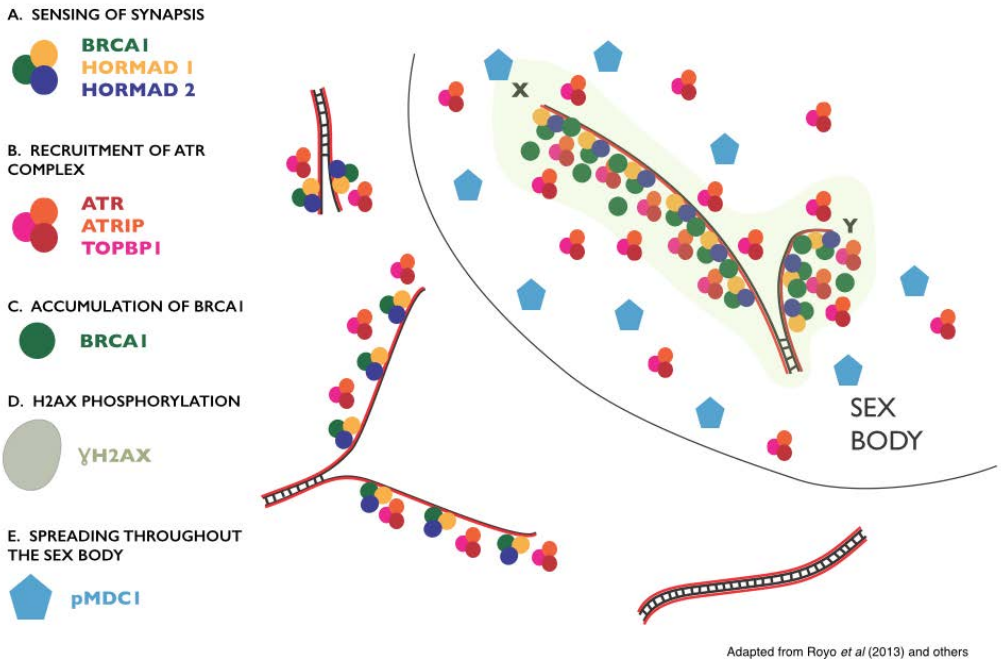


Figure 3: Proposed Mechanism of MSCI. BRCA1 and HORMAD1/2 are present at all sites of double strand breaks at autosomes and sex chromosomes during early prophase I. And recruitment of ATR, ATRIP, and TOBP1 to these sites occurs. At pachytene, these components are lost from the autosomal chromosome cores and BRCA1 accumulates instead at the asynapsed sex chromosomes.

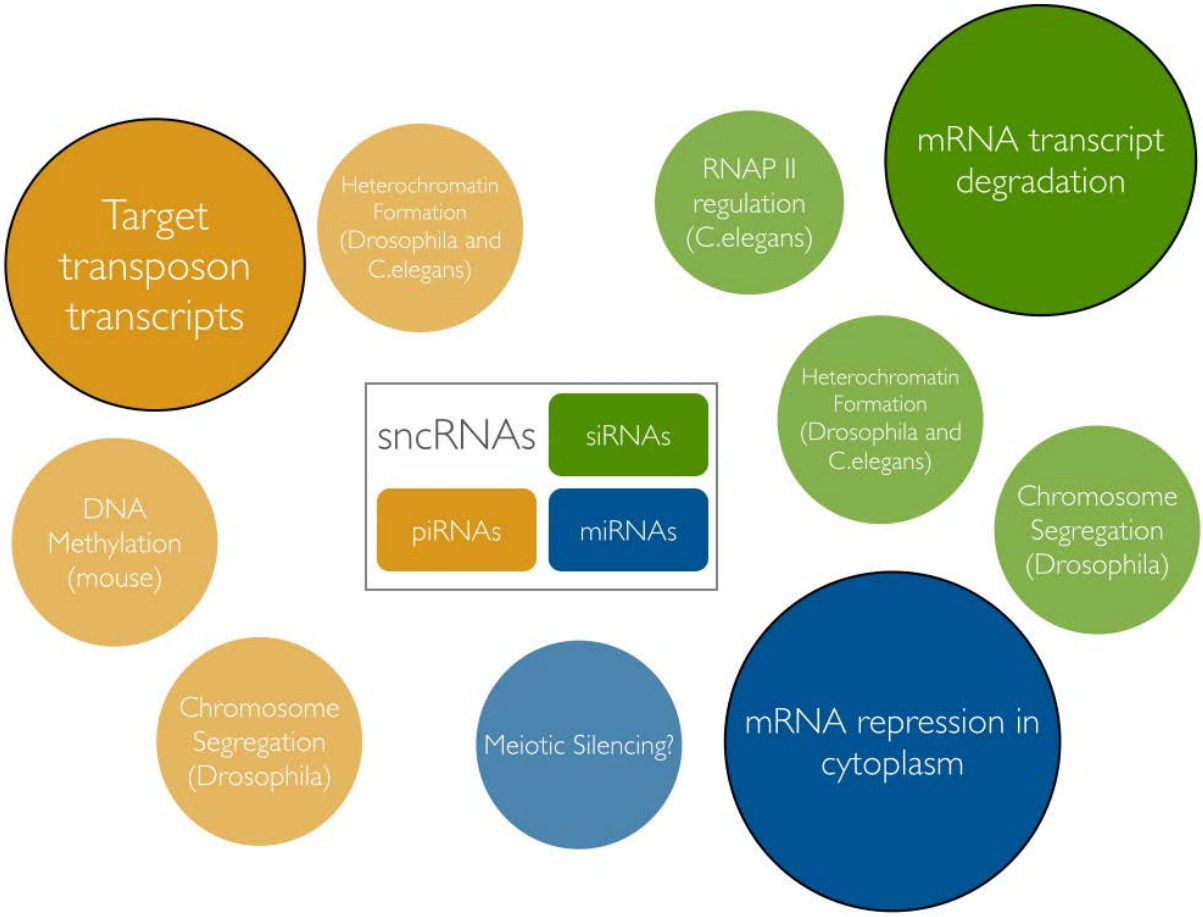


Figure 4: Known roles for sncRNAs in the nucleus. Small non-coding RNAs are known to have various roles in the nucleus of various organisms. This diagram shows some of these known roles and the organisms in which they are seen to play them. The larger, outlined circle shows the most well defined role for each of the sncRNAs.

References

1. Handel, M.A. and J.C. Schimenti, *Genetics of mammalian meiosis: regulation, dynamics and impact on fertility.* Nat Rev Genet, 2010. 11(2): p. 124-36.
2. Plant, T.M. and A.J. Zeleznik, *Knobil and Neill's Physiology of Reproduction Two-Volume Set*. 2014, Academic Press Imprint Elsevier Science & Technology Books: San Diego. p. 1 online resource.
3. Turner, J.M., *Meiotic sex chromosome inactivation.* Development, 2007. 134(10): p. 1823-31.
4. Vernet, N., et al., *Mouse Y-linked Zfy1 and Zfy2 are expressed during the male-specific interphase between meiosis I and meiosis II and promote the 2nd meiotic division.* PLoS Genet, 2014. 10(6): p. e1004444.
5. Royo, H., et al., *ATR acts stage specifically to regulate multiple aspects of mammalian meiotic silencing.* Genes Dev, 2013. 27(13): p. 1484-94.
6. Turner, J.M., et al., *BRCA1, histone H2AX phosphorylation, and male meiotic sex chromosome inactivation.* Curr Biol, 2004. 14(23): p. 2135-42.
7. Modzelewski, A.J., et al., *AGO4 regulates entry into meiosis and influences silencing of sex chromosomes in the male mouse germline.* Dev Cell, 2012. 23(2): p. 251-64.
8. Hutvagner, G. and M.J. Simard, *Argonaute proteins: key players in RNA silencing.* Nat Rev Mol Cell Biol, 2008. 9(1): p. 22-32.
9. Bolcun-Filas, E. and J.C. Schimenti, *Genetics of meiosis and recombination in mice.* Int Rev Cell Mol Biol, 2012. 298: p. 179-227.
10. Gerton, J.L. and R.S. Hawley, *Homologous chromosome interactions in meiosis: diversity amidst conservation.* Nat Rev Genet, 2005. 6(6): p. 477-87.
11. Page, S.L. and R.S. Hawley, *The genetics and molecular biology of the synaptonemal complex.* Annu Rev Cell Dev Biol, 2004. 20: p. 525-58.
12. Loidl, J., H. Scherthan, and D.B. Kaback, *Physical association between nonhomologous chromosomes precedes distributive disjunction in yeast.* Proc Natl Acad Sci U S A, 1994. 91(1): p. 331-4.
13. Vazquez, J., A.S. Belmont, and J.W. Sedat, *The dynamics of homologous chromosome pairing during male Drosophila meiosis.* Curr Biol, 2002. 12(17): p. 1473-83.
14. MacQueen, A.J., et al., *Chromosome sites play dual roles to establish homologous synapsis during meiosis in C. elegans.* Cell, 2005. 123(6): p. 1037-50.
15. Harper, L., I. Golubovskaya, and W.Z. Cande, *A bouquet of chromosomes.* J Cell Sci, 2004. 117(Pt 18): p. 4025-32.
16. Tomita, K. and J.P. Cooper, *The telomere bouquet controls the meiotic spindle.* Cell, 2007. 130(1): p. 113-26.
17. Fridkin, A., et al., *SUN-domain and KASH-domain proteins during development, meiosis and disease.* Cell Mol Life Sci, 2009. 66(9): p. 1518-33.
18. Fennell, A., et al., *Telomeres and centromeres have interchangeable roles in promoting meiotic spindle formation.* J Cell Biol, 2015. 208(4): p. 415-28.
19. Morelli, M.A., et al., *Analysis of meiotic prophase I in live mouse spermatocytes.* Chromosome Res, 2008. 16(5): p. 743-60.
20. Parvinen, M. and K.O. Soderstrom, *Chromosome rotation and formation of synapsis.* Nature, 1976. 260(5551): p. 534-5.

21. Ding, D.Q., et al., *Meiosis-specific noncoding RNA mediates robust pairing of homologous chromosomes in meiosis.* Science, 2012. 336(6082): p. 732-6.
22. Yamashita, A., et al., *RNA-assisted nuclear transport of the meiotic regulator Mei2p in fission yeast.* Cell, 1998. 95(1): p. 115-23.
23. Sato, M., et al., *The fission yeast meiotic regulator Mei2p undergoes nucleocytoplasmic shuttling.* FEBS Lett, 2001. 499(3): p. 251-5.
24. de Vries, F.A., et al., *Mouse Sycp1 functions in synaptonemal complex assembly, meiotic recombination, and XY body formation.* Genes Dev, 2005. 19(11): p. 1376-89.
25. Egel, R. and D.-H. Lankenau, *Recombination and meiosis : crossing-over and disjunction*. Genome dynamics and stability,. 2008, Berlin ; New York: Springer. xiv, 365 p.
26. McKim, K.S. and A. Hayashi-Hagihara, *mei-W68 in Drosophila melanogaster encodes a Spo11 homolog: evidence that the mechanism for initiating meiotic recombination is conserved.* Genes Dev, 1998. 12(18): p. 2932-42.
27. Yang, F., et al., *Mouse SYCP2 is required for synaptonemal complex assembly and chromosomal synapsis during male meiosis.* J Cell Biol, 2006. 173(4): p. 497-507.
28. Eijpe, M., et al., *Meiotic cohesin REC8 marks the axial elements of rat synaptonemal complexes before cohesins SMC1beta and SMC3.* J Cell Biol, 2003. 160(5): p. 657-70.
29. Wang, H. and C. Hoog, *Structural damage to meiotic chromosomes impairs DNA recombination and checkpoint control in mammalian oocytes.* J Cell Biol, 2006. 173(4): p. 485-95.
30. Yuan, L., et al., *The murine SCP3 gene is required for synaptonemal complex assembly, chromosome synapsis, and male fertility.* Mol Cell, 2000. 5(1): p. 73-83.
31. Brown, P.W., et al., *Meiotic synapsis proceeds from a limited number of subtelomeric sites in the human male.* Am J Hum Genet, 2005. 77(4): p. 556-66.
32. Bolcun-Filas, E., et al., *Mutation of the mouse Syce1 gene disrupts synapsis and suggests a link between synaptonemal complex structural components and DNA repair.* PLoS Genet, 2009. 5(2): p. e1000393.
33. Bolcun-Filas, E., et al., *SYCE2 is required for synaptonemal complex assembly, double strand break repair, and homologous recombination.* J Cell Biol, 2007. 176(6): p. 741-7.
34. Hamer, G., et al., *Progression of meiotic recombination requires structural maturation of the central element of the synaptonemal complex.* J Cell Sci, 2008. 121(Pt 15): p. 2445-51.
35. Crackower, M.A., et al., *Essential role of Fkbp6 in male fertility and homologous chromosome pairing in meiosis.* Science, 2003. 300(5623): p. 1291-5.
36. Cole, F., S. Keeney, and M. Jasin, *Comprehensive, fine-scale dissection of homologous recombination outcomes at a hot spot in mouse meiosis.* Mol Cell, 2010. 39(5): p. 700-10.
37. Guillon, H., et al., *Crossover and noncrossover pathways in mouse meiosis.* Mol Cell, 2005. 20(4): p. 563-73.
38. Kohl, K.P. and J. Sekelsky, *Meiotic and mitotic recombination in meiosis.* Genetics, 2013. 194(2): p. 327-34.
39. Oke, A., et al., *Controlling meiotic recombinational repair - specifying the roles of ZMMs, Sgs1 and Mus81/Mms4 in crossover formation.* PLoS Genet, 2014. 10(10): p. e1004690.

40. Lorenz, A. and M.C. Whitby, *Crossover promotion and prevention.* Biochem Soc Trans, 2006. 34(Pt 4): p. 537-41.
41. Lynn, A., R. Soucek, and G.V. Borner, *ZMM proteins during meiosis: crossover artists at work.* Chromosome Res, 2007. 15(5): p. 591-605.
42. Hassold, T., et al., *Cytological studies of meiotic recombination in human males.* Cytogenet Genome Res, 2004. 107(3-4): p. 249-55.
43. Amangyeld, T., et al., *Human MUS81-EME2 can cleave a variety of DNA structures including intact Holliday junction and nicked duplex.* Nucleic Acids Res, 2014. 42(9): p. 5846-62.
44. Gilbert, S.F., *Developmental biology*. Tenth edition. ed. 1 volume (various pagings).
45. Koopman, P., et al., *Male development of chromosomally female mice transgenic for Sry.* Nature, 1991. 351(6322): p. 117-21.
46. Gill, M.E., et al., *Licensing of gametogenesis, dependent on RNA binding protein DAZL, as a gateway to sexual differentiation of fetal germ cells.* Proc Natl Acad Sci U S A, 2011. 108(18): p. 7443-8.
47. Swain, A. and R. Lovell-Badge, *Mammalian sex determination: a molecular drama.* Genes Dev, 1999. 13(7): p. 755-67.
48. Handel, M.A., *Meiosis and Gameteogenesis*. Academic Press. Vol. 37. 1998.
49. Koubova, J., et al., *Retinoic acid activates two pathways required for meiosis in mice.* PLoS Genet, 2014. 10(8): p. e1004541.
50. Kumar, S., et al., *Sex-specific timing of meiotic initiation is regulated by Cyp26b1 independent of retinoic acid signalling.* Nat Commun, 2011. 2: p. 151.
51. Shiu, P.K., et al., *Meiotic silencing by unpaired DNA.* Cell, 2001. 107(7): p. 905-16.
52. Schoenmakers, S., et al., *Meiotic silencing and fragmentation of the male germline restricted chromosome in zebra finch.* Chromosoma, 2010. 119(3): p. 311-24.
53. Federici, F., et al., *Incomplete meiotic sex chromosome inactivation in the domestic dog.* BMC Genomics, 2015. 16: p. 291.
54. Namekawa, S.H. and J.T. Lee, *XY and ZW: is meiotic sex chromosome inactivation the rule in evolution?* PLoS Genet, 2009. 5(5): p. e1000493.
55. Morelli, M.A. and P.E. Cohen, *Not all germ cells are created equal: aspects of sexual dimorphism in mammalian meiosis.* Reproduction, 2005. 130(6): p. 761-81.
56. Campbell, P., J.M. Good, and M.W. Nachman, *Meiotic sex chromosome inactivation is disrupted in sterile hybrid male house mice.* Genetics, 2013. 193(3): p. 819-28.
57. Vernet, N., et al., *The Y-encoded gene zfy2 acts to remove cells with unpaired chromosomes at the first meiotic metaphase in male mice.* Curr Biol, 2011. 21(9): p. 787-93.
58. Song, R., et al., *Many X-linked microRNAs escape meiotic sex chromosome inactivation.* Nat Genet, 2009. 41(4): p. 488-93.
59. Marcon, E., et al., *miRNA and piRNA localization in the male mammalian meiotic nucleus.* Chromosome Res, 2008. 16(2): p. 243-60.
60. Bai, B., H. Liu, and M. Laiho, *Small RNA expression and deep sequencing analyses of the nucleolus reveal the presence of nucleolus-associated microRNAs.* FEBS Open Bio, 2014. 4: p. 441-9.
61. Schimenti, J., *Synapsis or silence.* Nat Genet, 2005. 37(1): p. 11-3.
62. Castel, S.E. and R.A. Martienssen, *RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond.* Nat Rev Genet, 2013. 14(2): p. 100-12.

63. Lau, N.C., *Small RNAs in the animal gonad: guarding genomes and guiding development.* Int J Biochem Cell Biol, 2010. 42(8): p. 1334-47.
64. Stefani, G. and F.J. Slack, *Small non-coding RNAs in animal development.* Nat Rev Mol Cell Biol, 2008. 9(3): p. 219-30.
65. Shomron, N. and C. Levy, *MicroRNA-biogenesis and Pre-mRNA splicing crosstalk.* J Biomed Biotechnol, 2009. 2009: p. 594678.
66. McIver, S.C., et al., *miRNA and mammalian male germ cells.* Hum Reprod Update, 2012. 18(1): p. 44-59.
67. Gregory, R.I., et al., *The Microprocessor complex mediates the genesis of microRNAs.* Nature, 2004. 432(7014): p. 235-40.
68. Brennecke, J., et al., *Principles of microRNA-target recognition.* PLoS Biol, 2005. 3(3): p. e85.
69. Meister, G., *Argonaute proteins: functional insights and emerging roles.* Nat Rev Genet, 2013. 14(7): p. 447-59.
70. Ma, H., J. Zhang, and H. Wu, *Designing Ago2-specific siRNA/shRNA to Avoid Competition with Endogenous miRNAs.* Mol Ther Nucleic Acids, 2014. 3: p. e176.
71. Schraivogel, D. and G. Meister, *Import routes and nuclear functions of Argonaute and other small RNA-silencing proteins.* Trends Biochem Sci, 2014. 39(9): p. 420-31.
72. Wei, W., et al., *A role for small RNAs in DNA double-strand break repair.* Cell, 2012. 149(1): p. 101-12.
73. de Mateo, S. and P. Sassone-Corsi, *Regulation of spermatogenesis by small non-coding RNAs: role of the germ granule.* Semin Cell Dev Biol, 2014. 29: p. 84-92.
74. Findley, S.D., et al., *Maelstrom, a Drosophila spindle-class gene, encodes a protein that colocalizes with Vasa and RDE1/AGO1 homolog, Aubergine, in nuage.* Development, 2003. 130(5): p. 859-71.
75. Banisch, T.U., M. Goudarzi, and E. Raz, *Small RNAs in germ cell development.* Curr Top Dev Biol, 2012. 99: p. 79-113.
76. Kuramochi-Miyagawa, S., et al., *DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes.* Genes Dev, 2008. 22(7): p. 908-17.
77. Hayashi, K., et al., *MicroRNA biogenesis is required for mouse primordial germ cell development and spermatogenesis.* PLoS One, 2008. 3(3): p. e1738.
78. Lei, L., et al., *The regulatory role of Dicer in folliculogenesis in mice.* Mol Cell Endocrinol, 2010. 315(1-2): p. 63-73.
79. Parvinen, M., *The chromatoid body in spermatogenesis.* Int J Androl, 2005. 28(4): p. 189-201.
80. Modzelewski, A.J., et al., *Dgcr8 and Dicer are essential for sex chromosome integrity in male meiosis.* J Cell Sci, 2015.
81. Hammond, S.M., et al., *An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells.* Nature, 2000. 404(6775): p. 293-6.
82. Holoch, D. and D. Moazed, *Small-RNA loading licenses Argonaute for assembly into a transcriptional silencing complex.* Nat Struct Mol Biol, 2015. 22(4): p. 328-35.
83. Motamedi, M.R., et al., *Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs.* Cell, 2004. 119(6): p. 789-802.
84. Maida, Y., et al., *An RNA-dependent RNA polymerase formed by TERT and the RMRP RNA.* Nature, 2009. 461(7261): p. 230-5.

CHAPTER 2[1]

Dgcr8 and *Dicer* are essential for sex chromosome integrity in male meiosis

1. This chapter is adapted from the paper: Modzelewski, A.J., et al., *Dgcr8 and Dicer are essential for sex chromosome integrity during meiosis in males.* J Cell Sci, 2015. 128(12): p. 2314-27. doi: 10.1242/jcs.167148. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed. Elizabeth A. Crate contributed to this paper by performing phenotypic analysis. This included mouse husbandry of the Dicer and Dgcr8 cKO mouse lines, performing and analyzing chromosome spreads and immunofluorescence staining, and constructing figures of the immunofluorescence.

Abstract

Small RNAs play crucial roles in regulating gene expression during mammalian meiosis. To investigate the function of microRNAs and small-interfering RNAs in male meiosis, we generated germ cell-specific conditional deletions of Dgcr8 and Dicer in mice. Analysis of spermatocytes from both conditional knockout lines reveals frequent chromosomal fusions during meiosis, always involving one or both sex chromosomes. RNA sequencing indicates upregulation of Atm in spermatocytes from microRNA-deficient mice, and immunofluorescence imaging demonstrates an increased abundance of activated ATM kinase and mislocalization of phosphoMDC1, an ATM phosphorylation substrate. The Atm 3′UTR contains many potential microRNA target sites; notably, target sites for several miRNAs depleted in both conditional knockout mice are highly effective at promoting repression. RNF8, a telomere-associated protein whose localization is controlled by the MDC1/ATM kinase cascade, normally associates with the sex chromosomes during pachytene, but in both conditional knockouts redistributes to the autosomes. Together, these results suggest that Atm dysregulation in microRNA-deficient germ lines contributes to the redistribution of proteins involved in chromosomal stability from the sex-chromosomes to the autosomes, resulting in sex-chromosome fusions during meiotic prophase I.

Introduction

Small RNA-mediated silencing has emerged as a major mechanism of gene regulation in animals and plants, with clear evidence for both cytoplasmic and nuclear small RNA-based regulatory pathways that act upon RNA and DNA targets [1–4]. Since the discovery of microRNAs (miRNAs) and endogenous small interfering RNAs (siRNAs), comparatively few studies have investigated roles for these small RNA species in mammalian reproduction. In

females, siRNAs (but not miRNAs) are essential during oogenesis for proper maturation, meiotic spindle organization and chromosome alignment [5–7]. During male gametogenesis, a third class of small RNAs, the piwi-interacting RNAs (piRNAs), are expressed in addition to miRNAs and siRNAs. The functions of piRNAs in the male germ line are relatively well-defined: piRNAs deter transposon integration, and also have distinct functions during prophase I [8]. In contrast to piRNAs, the roles and importance of miRNAs and siRNAs in the male germ line are poorly defined. While it is clear that a functional miRNA pathway is necessary to complete spermatogenesis [9–13], the identities of essential miRNAs, and their mRNA targets, remain to be determined. Whether siRNAs are also required in the male germ line is presently unclear and remains controversial, with reports suggesting either no role in spermatogenesis [9] or implicating siRNAs in specific functions in the germ line [14].

The major mode of action for miRNAs, based on a number of studies, is to accelerate the turnover of target mRNAs and inhibit their translation [4,15]. Such activities are consistent with the typical cytoplasmic localization of Argonaute (AGO) proteins, the RNA binding proteins that function in concert with small RNAs and underpin all small RNA regulatory pathways [15,16]. Additionally, in non-mammalian organisms such as nematodes, fungi and plants, Argonautes have been shown to regulate chromatin structure and transcriptional activity [17,18]. Similar nuclear functions have yet to be firmly established in mammals, although some studies have hinted at their existence [19,20]. Importantly, while mammalian AGO proteins have been observed in the nucleus, the functional significance of this localization is currently the subject of some debate [21,22].

We previously described the nuclear localization of AGO4, one of four mammalian Argonaute proteins, in mammalian germ cells during prophase I of meiosis. AGO4 accumulates

within the sex body (SB) of pachytene spermatocytes. The SB is a specialized nuclear sub-domain harboring the X and Y chromosomes, in which the chromatin is transcriptionally silenced: a process known as meiotic sex chromosome inactivation (MSCI). In addition, AGO4 localizes to autosomal regions that fail to pair with their homologous partners; these unpaired regions are also silenced during prophase I as part of the more global meiotic silencing of unpaired chromatin (MSUC). *Ago4*$^{-/-}$ male mice are subfertile, with reduced testis size and lowered epididymal sperm counts. Meiotic prophase I progression is severely impaired in the absence of AGO4, resulting in a high proportion of spermatocytes undergoing apoptosis. Importantly, the formation and function of the SB during pachynema of prophase I requires AGO4, since MSCI is perturbed in *Ago4*$^{-/-}$ male s[23]. Collectively, these data suggest a direct role for AGO4 in transcriptional silencing during mammalian meiosis, specifically in regions of the genome in which chromosomes are unpaired (such as the X and Y chromosomes or asynapsed regions of the autosomes), in a fashion reminiscent of analogous processes in *Caenorhabditis elegans* and *Neurospora crassa* (reviewed by [17]).

Two lines of evidence suggest that AGO4 is not the sole Argonaute contributing to MSCI. First, AGO3 expression is upregulated in the absence of AGO4, and this upregulation is specific to the male germ line and specific to AGO3 alone [23]. Second, like AGO4, AGO3 localizes to the SB during prophase I, suggesting that AGO3 and AGO4 function redundantly in mammalian meiosis. These observations imply that our analysis of *Ago4* mutant animals provides only a glimpse of the repertoire of siRNA and miRNA actions during meiotic prophase I.

In the current study, we created germline-specific conditional deletions of two proteins needed for small RNA biogenesis, DGCR8 and DICER. DGCR8, a component of the

endonuclease microprocessor complex, is essential for the first step in miRNA processing: conversion of the primary miRNA transcript into a precursor miRNA. DICER, a second endonuclease, acts on the precursor miRNA in the cytoplasm to generate small RNAs, and is needed for the synthesis of almost all miRNAs[24,25]. DICER, but not DGCR8, is also required for the processing of siRNAs [26–29]. Phenotypes common to both *Dgcr8* and *Dicer* conditional deletion animals would reveal roles of miRNAs in mammalian meiotic progression, whereas phenotypes specific to *Dicer* conditional deletion animals would likely implicate siRNAs. Our results demonstrate that loss of miRNAs leads to misregulation of the DNA damage repair pathway, including upregulation of *Atm* and improper localization of ATM substrate proteins. We find that the majority of these miRNA-deficient spermatocytes display frequent sex chromosome fusions and fail to progress through meiosis.

Results

Conditional deletion of Dgcr8 or Dicer in the male germ line results in sterility

To investigate and compare roles for miRNAs and siRNAs in the male germ line, we generated germline-specific deletions of *Dgcr8* and *Dicer*, which we analyzed in parallel. Our strategy used floxed alleles of *Dgcr8* and *Dicer*, in combination with a *Ddx4* promoter-driven *Cre* transgene. The *Ddx4-Cre* transgene induces expression of the recombinase in spermatogonia from embryonic (e) day 18, before the initiation of meiosis at around day 8 post-partum (pp), but after the male-specific block of meiotic entry at around e12 [30]. We selected this *Cre* transgene to allow unaltered development of both pre-meiotic germ cells and somatic compartments of the testis, while ensuring that all spermatocytes possess *Dgcr8* and *Dicer* deletions prior to meiotic initiation.

Conditional knock-out mice (cKO) of either *Dgcr8* or *Dicer* are infertile, as no offspring were recovered from matings to wildtype (C57BL/6) females over a 7-day period, despite normal mating behavior. At day 70 pp, both cKO mice exhibited normal body size (Figure S1C), but testis weights were significantly reduced compared to wildtype littermates (Figure 1A and Figure S1A and B). Mean epididymal spermatozoa counts from *Dgcr8* and *Dicer* cKOs were also greatly reduced (Figure 1B), with the degree of reduction significantly more severe in *Dicer* cKO germ lines ($P < 0.05$, Student's T-Test). Histological examination of testes from cKO animals revealed severe morphological abnormalities throughout the seminiferous epithelium, including an increased prevalence of vacuous tubules, together with tubules in which spermatogonia were arrested prior to the onset of spermiogenesis (Figure 1E and F, insets). No mature spermatozoa were observed in *Dicer* cKO males, whereas mature spermatozoa, while rare, were detectable in *Dgcr8* cKO sections (Figure 1E, inset arrow). Consistent with the absence of mature spermatozoa, the frequency of apoptotic spermatocytes, as assessed by TUNEL labeling, was also markedly increased in both *Dgcr8* and *Dicer* cKO testes (Figure 1C, H and I). Overall, these results are consistent with observations from the majority of published studies of *Drosha*, *Dgcr8* and *Dicer* germline cKO mutants [9,11,13,14,31–33]. Our data implicate small RNAs that rely on both *Dgcr8* and *Dicer* processing, such as miRNAs, rather than *Dgcr8*-independent small RNAs, such as siRNAs, as the major non-coding small RNA regulators during early spermatogenesis.

Sex chromosomes fail to synapse and frequently undergo chromosomal fusion in Dgcr8 and Dicer cKO germ lines

To investigate possible roles for small RNAs in spermatocyte nuclei, we examined prophase I chromosome spread preparations visualized with antibodies raised against components of the synaptonemal complex (SC), a meiosis-specific structure that connects

homologous chromosomes during prophase I [34,35]. The status of the SC serves as an indicator of the stages and progression of prophase I. We used antibodies against the chromosome axis protein SYCP3, which localizes along homologous chromosomes prior to synapsis, and SYCP1, a component of the SC central element, which serves to tether homologous chromosomes together from zygotene of prophase I onwards. Since SYCP1 associates only with synapsed chromosomes, it is normally excluded from the unpaired X and Y-chromosomes, except at a short region of homology capable of synapsis, referred to as the pseudoautosomal region (PAR; Figure 2A, white arrow). Additionally, to investigate sex body integrity, we examined localization of H2AX, a histone variant whose phosphorylation (resulting in γH2AX) is essential for MSCI. During pachynema, H2AX localizes exclusively to the sex body of normal spermatocytes and to any autosomal sites of asynapsis.

Examination of *Dgcr8* and *Dicer* cKO spermatocytes in prophase I revealed that SYCP1 and γH2AX were mislocalized in a minority of cells (~10%), in a pattern similar to that of *Ago4*$^{-/-}$ animals [23]. More strikingly, we observed a high frequency of structural abnormalities involving the X and Y-chromosomes of cKO spermatocytes, with very few cells exhibiting normal XY associations and PAR structure (Figure 2). A range of abnormalities were observed, including circularization of the X and Y chromosomes (Figure 2B, E and F; quantitated in Figure 2J), terminal fusion of either the X or Y chromosome to an autosome (Figure 2C and E; quantitated in Figure 2J), and increased frequency of PAR asynapsis. Such defects were never observed in wildtype control littermates (Figure 2A, D and G), nor in *Ago4*$^{-/-}$ animals (A.J.M., unpublished observations). Notably, chromosomal fusions always involved at least one sex chromosome, and were never observed between autosomes.

Despite frequent fusions between sex chromosomes and autosomes in both cKO germlines, staining patterns of γH2AX (Figure 2E and F), the kinase ATR, and pan-ATR/ATM phospho-substrates are largely unaltered (Figure S1), retaining exclusive X and Y chromosome localization and wildtype intensity. In various translocation and asynapsis models, the chromosome that becomes associated with the X or Y is usually engulfed by the silencing machinery and thereby silenced [36]. In both *Dgcr8* and *Dicer* cKO spermatocytes, however, there is no expansion of the γH2AX domain in response to these chromosomal abnormalities, and autosomes fused to sex chromosomes do not convert to a pseudosex-body like state.

In wildtype spermatocytes, RNA polymerase II (RNAPII) localizes throughout the nucleus during pachynema, except where it is actively excluded, such as at centromeres and the sex-body (Figure 2G). In a small fraction of spermatocytes from *Dgcr8* (~9%) and *Dicer* (~5%) cKO animals, RNAPII was found aberrantly localized within the SB domain. However, for many instances of the most extreme XY abnormalities, such as X-to-autosome fusions, both the *Dgcr8* and *Dicer* cKO pachytene spermatocytes still exclude RNAPII from the X chromosome portion of the fusion, while normal RNAPII staining is seen on the fused autosome (Figure 2H and I). This differential localization of RNAPII presumably maintains, at least partially, the silenced status of the X and Y and the active status of the fused autosome. To confirm this, we used RT-qPCR to quantify the transcript levels of *Zfy1* and *Zfy2* [37], key sex chromosome-encoded genes that must be silenced for spermatocytes to exit pachynema. Neither transcript was consistently elevated during pachytene (Figure S1L). This observation is consistent with the ability of spermatocytes from *Dgcr8* cKOs to proceed to diplonema, despite displaying frequent sex chromosome fusions (Figure S1J and K).

Altered distribution of CDK2 and RNF8 in Dgcr8 and Dicer cKO spermatocytes

The frequent chromosomal fusions observed in prophase I spermatocytes of both *Dgcr8* and *Dicer* cKO germ lines implies that one or more miRNAs regulate critical aspects of chromosomal integrity during meiosis. Moreover, the restriction of these fusion events to the sex chromosomes suggests that understanding the molecular basis for this phenomenon might reveal sex chromosome-specific regulatory features unique to meiotic prophase I. Thus, we pursued two complimentary strategies to characterize the chromosomal phenotypes observed in *Dgcr8* and *Dicer* cKO germ lines, and ultimately, to identify the specific small RNAs whose absence results in chromosomal fusions. Firstly, we assessed the chromosomal localization of proteins that might be involved in the fusion events, comparing their localization patterns in wildtype and cKO germ cells. We tested proteins involved in various aspects of meiosis, chromatin organization, telomere stability, and non-homologous end joining. Secondly, we performed RNA-seq on purified meiotic cells from wildtype and cKO germ lines and examined transcriptome profiles for signatures that might explain the underlying phenotype.

MDC1 (mediator of DNA damage checkpoint-1) is essential for spreading of DNA damage response factors after recognition of asynapsed and damaged chromatin. In wildtype pachytene spermatocytes, MDC1 localization is restricted to the sex body, similar to that of γH2AX (Figure 3A). Almost all *Dgcr8* and *Dicer* cKO spermatocytes displayed normal MDC1 accumulation at pachynema, even in the presence of XY chromosomal abnormalities (Figure 3B and C).

Cyclin-dependent kinase 2 (CDK2) plays multiple roles in spermatocytes. Loss of *Cdk2* results in infertility, a dramatic increase in synapsis between non-homologous chromosomes, the appearance of fusions and ring chromosomes, and eventual arrest in pachynema due to improper telomere maintenance [38]. In wildtype pachytene spermatocytes, CDK2 localizes both to

telomeres and sites of double strand break (DSB) repair [39] but is barely detectable on the unpaired XY chromosome core regions (Figure 3D). In *Dgcr8* and *Dicer* cKO spermatocytes, however, the unpaired or circularized X and Y chromosome cores often (~80% of spermatocytes in the *Dgcr8* cKO) display intense CDK2 localization across their entire length, not restricted to the telomeres (Figure 3E and F).

Trimethylation of Lysine 9 on Histone 3 (H3K9me3) is an epigenetic mark associated with repressed heterochromatin. H3K9me3 is enriched at autosomal centromeres and the majority of the sex-body in wildtype pachytene spermatocytes (Figure 3G; [40,41]. In a minor subset of cKO spermatocytes, we observed mislocalization of H3K9me3 throughout the meiotic nuclei; however the majority of spermatocytes display unaltered localization of H3K9me3 despite exhibiting severe XY chromosomal abnormalities at pachynema (Figure 3H and I). The ubiquitin E3 ligase RNF8 ubiquitinates histones in response to DNA damage; RNF8 localization is controlled, in part, by interactions with phosphorylated MDC1 (pMDC1; [42]. Defects in RNF8 have been associated with chromosomal instability in cell culture: mouse embryonic fibroblasts lacking RNF8 show an increase in chromosomal aberrations, including ring chromosome formation [43]. Normally, RNF8 localizes to the sex-body in pachynema (Figure 3J). However, in spermatocytes from *Dgcr8* and *Dicer* cKO males, RNF8 displays reduced SB localization, instead redistributing along autosomes to flare-shaped domains that extend perpendicularly from the chromosome axes into the chromatin (Figure 3K and L). In *Dgcr8* cKOs, we observed that RNF8 mislocalizes in 88% of spermatocytes with chromosomal abnormalities. These observations suggest that proper RNF8 localization is dependent, either directly or indirectly, on normal miRNA biogenesis. It is worth noting that in contrast to the rare

aberrant localization of MDC1 and H3K9me3, the unusual localization of CDK2 and RNF8 in cKO germ lines was observed in the majority of abnormal pachytene spermatocytes.

Atm is upregulated in Dgcr8 and Dicer cKO spermatocytes

To characterize disregulation of mRNA expression levels in cKO spermatocytes, we profiled the transcriptome using RNA-seq at pachynema and prior to pachynema (a mixed population of leptotene and zygotene spermatocytes). As parallel controls, we also profiled spermatocytes derived from wildtype littermates of both *Dgcr8* and *Dicer* breedings. Many transcripts exhibit significant differences in expression (Figure 4A and B) between wildtype and cKO spermatocytes, with many more genes displaying such differences in *Dicer* cKO samples as compared to *Dgcr8* cKO samples (L/Z: $P < 0.0001$; Pach: $P < 0.0001$; X^2 tests). Our initial analyses of these data focused on miRNA targets, together with gene ontology analysis (Figure S3). As expected, the mRNAs upregulated in cKO samples were enriched ($P < 0.0001$; X^2 test) in predicted miRNA target sites corresponding to miRNAs expressed in the germ line. However, a large proportion (26-31%) of genes upregulated in the cKO mutants do not contain predicted sites; such genes are likely downstream of regulatory events governed by direct targets, as are genes downregulated in the cKO mutants. The complexity of altered mRNA expression profiles is likely due to the multitude of disregulated pathways resulting from loss of miRNAs. Candidate genes known to be involved in regulation of pathways important to telomere maintenance and chromosomal integrity were unchanged. However, given the alterations in RFN8 localization and importance of DSB repair in telomere integrity, we noted that *Atm* is significantly upregulated in miRNA-deficient spermatocytes.

The kinase ATM plays multiple roles in DNA damage signaling and genome integrity, primarily acting through phosphorylation of MDC1. In response to phosphorylation, MDC1 recruits downstream targets including RNF8. The results of RNA-seq on *Dgcr8* and *Dicer* cKO

spermatocytes indicate that *Atm* transcript levels are increased, relative to wild type, from leptonema through pachynema (Figure 5A). We confirmed that *Atm* levels are upregulated in leptotene/zygotene samples from both *Dgcr8* and *Dicer* cKOs using qRT-PCR (Figure 5B). Furthermore, immunofluorescence imaging of phosphoATM, the activated form of ATM [44], demonstrates a subtle but discernable increase in protein levels during leptonema and zygonema (Figure 6 D-E and G-H). Moreover, transcriptionally-regulated targets downstream of pATM are also upregulated. Thus, though other genes are also upregulated in *Dgcr8* and *Dicer* cKO spermatocytes, we focused on *Atm* due to its known roles in promoting genomic stability and regulating RNF8 localization.

MicroRNA-mediated regulation of Atm in the male germ line

The simplest explanation for elevated levels of *Atm* is that *Atm* is a target of germline miRNA(s), which are absent in both cKOs. To explore this possibility, we examined the *Atm* 3′UTR for predicted miRNA target sites, discovering many potential sites (Figure 5C). To reduce the number of sites to those with co-expressed miRNAs, we sequenced small RNAs from the same samples used previously for RNA-seq; this approach greatly reduced the number of sites worthy of scrutiny. Although predictions of target site efficacy are of limited accuracy [45] we nevertheless focused on sites predicted by TargetScan [46–49] to be most effective in mediating repression. Strikingly, the miRNA predicted to mediate the strongest repression of *Atm* is miR-18, a microRNA known to exhibit meiosis-preferential expression [50]. The target sites of two other miRNAs, miR-183 and miR-16, cluster in the same region of the *Atm* 3′UTR as the miR-18 target sites; moreover, the miR-183 and miR-16 target sites are predicted as the second and third strongest sites, respectively, within *Atm* (Figure 5C). Furthermore, miR-18, miR-183 and miR-16 are amongst the most dramatically reduced miRNAs in *Dgcr8* and *Dicer* cKO

spermatocytes, compared to wildtype (Figure S2J, M, and N). Notably, none of the other miRNAs predicted to target *Atm* show measurable depletion in the cKOs.

To investigate the efficacy of miR-18 target sites in the *Atm* 3′UTR, we designed a luciferase reporter construct containing 431 nucleotides of the endogenous *Atm* 3′UTR sequence encompassing the two predicted miR-18 target sites. We also generated otherwise identical reporters in which either one or both sites were disrupted. We measured reporter activity in a cell line that does not express miR-18 by co-transfecting the reporter plasmids with either an siRNA corresponding to miR-18 or a control siRNA corresponding to a miRNA that does not target *Atm*. Such reporter assays are typically performed with transfected siRNAs at a concentration of 20-100 nM, which results in a level of target repression comparable to that achieved by endogenous miRNAs expressed at high levels [47,51–53]. Importantly, miR-18 is not one of the most abundant miRNAs in the murine male germ line (Figure S2J, compared to K and L). To measure the response of the *Atm* transcript to low levels of miR-18, we performed reporter assays using a range of concentrations of the miRNA mimetic, from 25 nM to 40 pM. We found that reporter constructs containing only a single miR-18 site were less effective at lower concentrations, while those with both sites intact elicit full repression over more than a 100-fold reduction in concentration of the miR-18 mimetic: repression at 0.2 nM was not significantly different from that measured at 25 nM (Figure 5D). It is worth noting that the miR-18 target sites are located close to each other, an arrangement previously observed to mediate synergistic enhancements to repression by miRNAs [47,54]. To determine whether repression mediated by miR-18 sites in *Atm* is synergistic, we compared the observed repression mediated by both sites to a value extrapolated from that mediated by each individual site (Figure 5D, red and orange-striped bars, respectively). Using this approach, we did not observe evidence of

synergism at high concentrations (25 nM) of the miR-18 mimetic; at lower concentrations (5, 1 and 0.2 nM), however, we observed increasingly strong evidence for synergistic interactions between the sites (Figure 5D). We repeated our reporter assay for the single miR-183 site, as well as for the three miR-16 sites. Like the miR-18 target sites, the single site for miR-183 (Figure 5F) and combined miR-16 sites (Figure 5H) were able to mediate ~2 fold repression of the reporter construct at high-to-moderate concentrations (25 nM-1 nM) of miR-183 and miR-16 mimetic. Notably, the level of repression (~2-fold) corresponds well to the change in levels of the endogenous *Atm* transcript in cKO germ lines. Taken together, these results indicate that miR-18, miR-183, and miR-16 target sites in *Atm* are functional and effective at eliciting downregulation in response to very low levels of miRNA.

Aberrant localization of phospho-MDC1 to autosomes in Dgcr8 and Dicer cKOs

Our results implicate increased levels of ATM in driving RNF8 from the sex-body to the autosomes in miRNA-deficient spermatocytes. ATM does not directly recruit RNF8; rather, recruitment is mediated by ATM-catalyzed phosphorylation of MDC1 (pMDC1). Although MDC1 localization appears normal in the majority of *Dgcr8* and *Dicer* cKO spermatocytes at pachynema (Figure 3B and C), we reasoned that any alteration in RNF8 localization at pachynema would more likely derive from altered pMDC1 distribution prior to this stage of meiosis, especially given the ~2-fold increase in *Atm* expression seen in leptonema/zygonema. In wildtype pachytene spermatocytes, pMDC1 staining accumulates in the sex-body (Figure 7C), and localization of pMDC1 appears unaffected in *Dicer* cKO pachytene spermatocytes (Figure 7F). At zygonema, however, pMDC1 is normally undetectable on the autosomes in wildtype spermatocytes (Figure 7A and B, leptonema and zygonema, respectively), but is dramatically upregulated across the autosomes in *Dicer* cKO spermatocytes (Figure 7D and E). Localization of the core MDC1 protein, regardless of its phosphorylation status, appears similar to the wild

type in both *Dgcr8* and *Dicer* cKOs (Figure S4), suggesting that the majority of the MDC1 pool remains unphosphorylated. These results, indicating elevated ATM kinase activity in *Dicer* cKO spermatocytes, confirm our observations of increased levels of the *Atm* transcript in *Dgcr8* and *Dicer* cKO germ lines. Together, our data suggest that loss of miRNA-mediated control of *Atm* initiates a series of events in cKO spermatocytes, titrating first pMDC1 and then RNF8 away from the sex-body, and ultimately resulting in sex chromosomes that are deficient in their normal complement of DNA-damage surveillance/repair proteins.

Discussion

The roles and significance of miRNAs and siRNAs during meiotic progression in mammals are poorly understood. Here, we used mouse models to demonstrate that the loss of miRNAs in spermatocytes results in drastic alterations in sex chromosome morphology, typified by an increased rate of chromosome circularization and end-to-end fusions reminiscent of telomere fusions. These telomere-related events are associated with misregulation of the DNA damage repair pathway, including increased amounts of the *Atm* transcript, a greater abundance of ATM protein at leptonema and zygonema, and mislocalization of ATM substrates. We also identified alterations in many small RNAs, including miR-18, miR-183, and miR-16, among whose targets is the mRNA encoding ATM. Our results indicate that miRNAs play a critical role in regulating DNA damage repair machinery, and ultimately chromosome stability, during mammalian spermatogenesis.

Previous studies investigating conditional knock-outs of *Dgcr8*, *Drosha*, and *Dicer* in the male germ line have noted many of the same gross morphological defects we observed,

including infertility (or subfertility), decreased sperm count, and disrupted sperm morphology [9,11,13,14,31–33]. A variety of essential roles for small RNAs in spermatogenesis have been proposed, including silencing of expression from the X and Y [14,31], regulating SINE levels [33], repressing centromeric repeat transcripts [32], and functioning in DNA repair at DSBs breaks [55]. Our work provides an additional explanation for the requirement of small RNAs in the male germline: miRNAs play an essential role in spermatocyte development by directly regulating levels of the ATM kinase, resulting in relocalization of several DNA damage repair proteins, and leading to chromosomal fusions. Although this study does not directly address whether the chromosomal fusions in *Dgcr8* and *Dicer* cKOs underlie the failure of such animals to complete spermatogenesis, such gross chromosomal abnormalities almost always lead to meiotic failure and failed chromosome segregation [56]. Importantly, our model is based on the established biology of miRNAs (post-transcriptional regulation of mRNAs) and does not invoke novel functions for miRNAs.

Many miRNAs are predicted to target the *Atm* 3′UTR, but only three are also expressed in spermatocytes and show depletion in *Dgcr8* and *Dicer* cKOs: miR-18, miR-183, and miR-16. We demonstrate that each of these miRNAs can effectively regulate the *Atm* 3′UTR, even at low concentrations; disruption of this regulation results in a ~2 fold upregulation of reporter activity, roughly the same change we observe for *Atm* in miRNA-deficient spermatocytes. Therefore, miR-18, miR-183 and miR-16 are the strongest candidates for miRNA-regulation of *Atm* expression in mammalian spermatogenesis.

Despite the observed upregulation of *Atm* in both leptotene/zygotene and pachtyene spermatocytes, the activated form of ATM's substrate, phosphoMDC1, is only mislocalized during leptonema and zygonema. By pachynema, phosphoMDC1 is restricted to the sex

chromosomes in both cKO and wildtype mice. This result is congruent with previous studies proposing two waves of phosphorylation during mammalian meiosis. The first wave is catalyzed by ATM during leptotene, in response to double stranded breaks [57]. ATR kinase, a relative of ATM, is thought to initiate a second wave of phosphorylation during zygotene and pachytene.

During zygotene, ATR localizes to sites of asynapsis on all chromosomes, and during pachytene it accumulates on the unpaired sex chromosomes, where it functions along with MDC1 and γH2Ax to silence the X and Y [58–60]. Therefore, normal localization of phosophoMDC1 in pachynema is likely due to ATR's ability to rescue MDC1 misregulation at this stage of meiosis. As MDC1 is essential for meiotic silencing, its normal localization on the sex body by pachynema explains why its earlier mislocalization does not result in a detectable defect on MSCI. We do see, however, that RNF8, which is recruited to sites of DSBs by pMDC1, persists on the autosomes during pachynema in *Dgcr8* and *Dicer* spermatocytes and appears depleted from the sex chromosomes. Taken together, these results suggest that while the activity of ATR at the sex body during pachytene is able to correct the mislocalization of some proteins that function downstream of ATM, the mislocalization of other downstream proteins persists throughout pachynema.

Though the *Dgcr8/Dicer* cKO spermatocytes exhibit misregulation of several proteins previously implicated in chromosomal fusions, the precise contribution of these proteins to sex-chromosome abnormalities is unclear. In particular, we observe misregulation of CDK2 and RNF8 in the *Dgcr8/Dicer* cKOs. *Cdk2*$^{-/-}$mice display frequent chromosomal fusions during meiosis, but importantly, these fusions never involve the X and Y [61]. Furthermore, these defects are presumably due to an absence of CDK2 on chromosomes, and we see that CDK2 is present on the sex chromosomes in our cKOs. Mouse embryonic fibroblasts lacking RNF8

exhibit an increase in chromosome fusion events [43], but the role RNF8 plays in chromosome stability during meiosis is less clear. Mice lacking RNF8 are able to progress through meiosis with no observable chromosomal fusion, although post-meiotic defects result in abnormal spermatids and infertility [62]. RNF8 has been implicated in playing opposing roles in fusion events at telomere ends, which can fuse through either classical (C-NHEJ) or alternative (A-NHEJ) non-homologous end joining [63]. RNF8 is required for efficient telomere fusion through the ATM-dependent C-NHEJ pathway in mouse models that promote such fusions[64,65]. However, RNF8 also protects telomere ends from undergoing fusion through the ATR-dependent A-NHEJ pathway by stabilizing the shelterin protein TPP1[43]. TRF1, a member of the shelterin complex typically disrupted in C-NHEJ, appears intact in our cKO pachytene spermatocytes (Figure S4). Therefore, RNF8 redistribution from the sex body onto the autosomes in *Dgcr8* and *Dicer* cKOs might directly contribute to fusions involving the sex chromosomes by promoting an A-NHEJ-like mechanism (Figure 8). Alternatively, we cannot exclude the possibility that mislocalization of RNF8 may be a downstream effect of other events occurring at the sex chromosomes that themselves lead to such fusion events.

An estimated 7% of the male population will encounter fertility problems, and in approximately half of these cases, the causes will be unknown [66]. Patients with Ataxia Telangiectasia, caused by *Atm* mutations, are often infertile, as are mice deficient in *Atm* [67–69]. However, whether *Atm* overexpression can lead to infertility is unknown; indeed, there is a notable absence of studies reporting the consequences of ATM overexpression. This study suggests that over expression of *Atm* in spermatocytes, whether mediated by miRNAs or by other mechanisms, would likely cause fertility defects in the male germ line. Here, we demonstrated that miRNA-deficient spermatocytes display frequent chromosomal fusion events involving the

sex chromosomes and upregulated *Atm* expression. Additionally, we observed mislocalization of other members of the DSB repair machinery downstream of ATM in *Dgcr8* and *Dicer* cKOs. Finally, we have identified candidate miRNA regulators of ATM expression in the male meiotic germ line. Our results underscore the significance of specific miRNAs in ensuring the fidelity of gametogenesis, and point to miR-18, miR-183 and miR-16 as miRNAs that likely play an important role in male fertility.

Materials and Methods

Mouse breeding strategies

Female mice carrying homozygous floxed alleles for either *Dgcr8*$^{Fl/Fl}$ (C57BL/6 strain background, Dr. Rui Yi, Univeristy of Colorado, Boulder) or *Dicer*$^{Fl/Fl}$ (C57BL/6;129S7-Dicer1^{tm1Smr}/J, Jax Stock: 012284) were crossed to male mice carrying the *Ddx4*-cre transgene (FVB-Tg(Ddx4-cre)1Dcas/J, Jax Stock: 006954) to generate the desired *Dgcr8*$^{Fl/+}$*cre*+ and *Dicer*$^{Fl/+}$ cre+ breeder males, which were then crossed to *Dgcr8*$^{Fl/+}$*cre*- and *Dicer*$^{Fl/+}$ cre- breeder females, respectively. These crosses generated the experimental cohorts, which included male mice displaying the *Dgcr8*$^{Fl/\Delta}$ *cre*+ and *Dicer*$^{Fl/\Delta}$ cre+ genotypes, which were homozygous knockouts in the targeted germline cell types (referred to as conditional knockouts, cKO, *Dgcr8*$^{\Delta/\Delta}$ *cre*+ and *Dicer*$^{\Delta/\Delta}$ cre+*)*, together with wildtype *(+/+ cre+, +/+ cre-, or Fl/+ cre-)* littermates. Genotypes were confirmed using DNA isolated from tail snips and PCR assays specific to the wildtype, floxed, and deleted alleles. All mice were fed *ad libitum* with standard laboratory rodent chow, and were maintained under controlled conditions of light and temperature, according to the regulations outlined and approved by the Cornell Institutional Care and Use Committee.

Testes weights, sperm counts, histology and TUNEL staining

Whole testes were removed from wildtype and cKO littermates, weighed and epididymal sperm counts assessed[70]. For histological analysis, testes were fixed in Bouin's fixative (H&E staining) or 10% formalin (TUNEL and all other staining) overnight at 4°C. Paraffin embedded tissues were sectioned at 5 μm and processed for H&E staining or immunohistochemical analyses using standard methods. TUNEL staining was performed using the Apoptag TUNEL staining kit (Chemicon, Temecula CA, USA).

Chromosome spreading and immunofluorescence staining

Prophase I chromosome spreads, antibodies and antibody staining were as previously described[23,71–74], except for antibodies recognizing ATR (GeneTex GTX70133), ATR/ATM substrate (Cell Signaling #5851), RNA polymerase II (Millipore 05-623, Covance MMS-129R), MDC1 and RNF8 (Raimundo Freire, Tenerife, Spain), CDK2 (Abcam ab7954), H3K9me3 (Millipore 07-442), TRF1 (Abcam ab10579), ATM pS1981 (Rockland Antibodies 200-301-400), and pMDC1 (Abcam ab35967). Alexafluor secondary antibodies were used (Molecular Probes Eugene OR, USA) for immunofluorescence staining at 37°C for one hour. Slides were washed and mounted with Prolong Gold antifade (Molecular Probes).

Image acquisition

All slides were visualized using a Zeiss Imager Z1 microscope under 20X, 40X or 63X magnifying objectives, at room temperature. Images were processed using AxioVision (version 4.7, Zeiss).

Statistical analysis

Statistical analyses (X^2 test, unpaired student's t-test and Wilcoxon rank sum test) were performed using GraphPad Prism version 4.00 for Macintosh (GraphPad Software, San Diego California USA), and using online web utilities (http://vassarstats.net and http://www.fon.hum.uva.nl/Service/CGI-Inline/HTML/Statistics.html).

Isolation of mouse spermatogenic cells

Testes from adult *Dgcr8* and *Dicer* cKOs, together with wildtype littermates (day 70-80 pp, a minimum of 2 mice per genotype) were removed, weighed and decapsulated prior to enrichment of specific spermatogenic cell types using the STA-PUT method based on separation by cell diameter/density at unit gravity[75]. Purity of resulting fractions was determined by microscopy based on cell diameter and morphology. Pachytene cells were approximately 90% pure, with residual possible contamination from spermatocytes of slightly earlier or later developmental stages. Leptotene and zygotene cells were isolated in combination, and were >90% pure. Neither pachytene nor leptotene/zygotene fractions visually exhibited evidence of contaminating somatic cells. Nuclear extracts were prepared using the *NE-PER* Nuclear and Cytoplasmic Extraction *Kit* RNA (Thermo Scientific). RNA was extracted from STA-PUT purified cells using TRIzol (Life Technologies), and used as the source material for mRNA and small RNA sequencing and qRT-PCR.

mRNA transcript sequencing and analysis

Non-stranded RNA-seq libraries were prepared (TRUseq, Illumina) and sequenced (Illumina HiSeq 2500). The resulting sequences were mapped to the genome (mm9) using

BWA (v0.7.8; [76]. The number of reads mapping to each transcript was quantified using HTSeq (v0.6.1; [77], and differential transcript expression between samples was assessed using edgeR (v3.6.8; [78]; R version: 3.1.0). Targeting analysis was performed using custom python scripts and the TargetScan Mouse database (v6.2; [46–49].

Small RNA sequencing and analysis

Small RNA sequencing libraries were prepared (TruSeq Small RNA-seq, Illumina) and sequenced on (Illumina HiSeq 2500). High-quality reads were aligned to the genome (mm9) using Bowtie (v0.12.7; [79]. Mouse miRNA sequences were obtained from miRBase (v20; [80,81]. piRNA-like reads were predicted based on their length and proximity to other piRNA-like reads (custom script). Coordinates for the degenerate PRDM9 and random motif sequences were determined by finding matches to their corresponding regular expressions in the genome.

Accession Numbers

Deep sequencing files are available from NCBI GEO (GSE63166).

Quantitative PCR

Complementary DNA was synthesized with Transcriptor reverse transcriptase (Roche Applied Science) according to the manufacturer's instructions using 1ug total RNA. qPCR reactions were run in triplicate on a LightCycler480 (Roche Applied Science). A melt curve for each reaction confirmed amplicon identity, and a standard curve was used to calculate transcript abundance, assaying GAPDH (TGAAGCAGGCATCTGAGGG and

CGAAGGTGGAAGAGTGGGAG); ATM (TCAGGCTGTATCTCAAGCCAT and AAGGGCTGCTAAGATGTGACT).

Luciferase assays

For the miR-18 assay, a 431nt fragment of the *Atm* 3′UTR (chr9:53245275-53245705) containing the two miR-18 target sites was amplified from mouse DNA using primers TGAGTGAGACGGGCTGTTACC and TCCTGGACTGCCTACTGATTCC, and reamplified using primers (cloning sites underlined) ATATGAGCTCTGCCTGAGGACAGAAGACATTG and ATATTCTAGATTCAGGAAACAGCATAACTGAAAAAC, digested and cloned into the luciferase reporter pmirGLO (Promega), to generate the miR-18 wildtype *Atm* reporter. For the miR-183 and miR-16 assay, a slightly larger fragment of the *Atm* 3′UTR, which spans 739 nt (chr9:53245260-53245998) was used so as to include all three miR-16 sites, in addition to the two miR-18 and one miR-183 sites. This fragment was amplified from mouse DNA using primers (cloning sites underlined) ATATGAGCTCTTCAGATTTCTTCAGTGGCTTTGATAAATCTATGTC and ATATTCTAGAAAAACAAACTGATAATTCAGGAAACAGCATAACTG, and cloned into pmirGLO (as described above), to generate the miR-183 and miR-16 wildtype *Atm* reporter. Reporter constructs in which one or more sites were mutated (QuikChange, Agilent) were generated using the wildtype reporter as a template. Sites were mutated as follows: the miR-18 site (GCACCUUA) was mutated to GCAggaUA for both miR-18 sites; the miR-183 site (GUGCCAUA) was mutated to GUGaCgUA, the miR-16 sites (GCUGCU) were mutated to either GCcGaU, GaUGgU, GCcGaU (corresponding to the order of the sites within the reporter construct). Reporter constructs were transfected into A549 cells along with varying levels of

siRNA duplexes corresponding to miR-18 (using RNA oligonucleotides: UAAGGUGCAUCUAGUGCAGAU and CUGCACUAGAUGCACCUUAAU), miR-183 (UAUGGCACUGGUAGAAUUCACU and UGAAUUCUACCAGUGCCAGAUA), miR-16 (UAGCAGCACGUAAAUAUUGGCG and CCAAUAUUUACGUGCUGUUAUU), or, as a control, miR-124[82]. Renilla and firefly luciferase activity (Dual-luciferase assay, Promega) were measured (Veritas luminometer, Turner Biosystems) 24 hours post-transfection. To control for transfection efficiency, firefly luciferase values were normalized to those of Renilla luciferase. These values were then normalized to those of the construct in which all target sites for that particular miRNA were disrupted. Fold repression was calculated by taking the inverse of the fold change. Predicted fold repression for the wildtype reporter was calculated using measurements for each single mutant reporter [47].

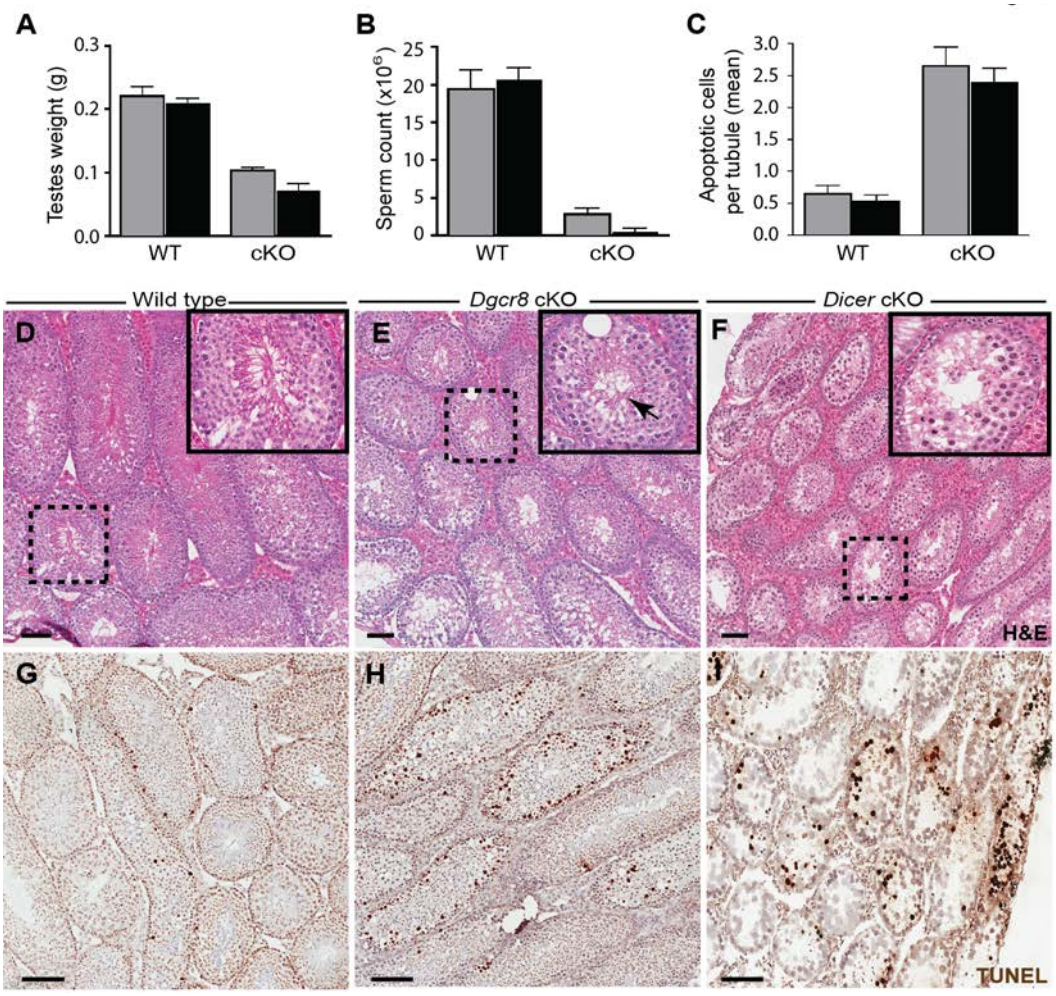

A
Testes weight (g)
0.3
0.2
0.1
0.0
WT
cKO
B
Sperm count (x10^6)
25
20
15
10
5
0
WT
cKO
C
Apoptotic cells per tubule (mean)
3.0
2.5
2.0
1.5
1.0
0.5
0.0
WT
cKO
Wild type
Dgcr8 cKO
Dicer cKO
D
E
F
H&E
G
H
TUNEL

Figure 1. Phenotypic analysis of *Dgcr8* and *Dicer* cKO males.
A Testis weights from *Dgcr8* cKO mice (grey bars; mean = 0.10 g, n = 4) and *Dicer* cKO mice (black bars; mean = 0.07 g, n = 4) are reduced compared to wildtype (WT) littermates (mean = 0.22 g, n = 4, $P < 0.0001$, t test, and, mean = 0.21 g, n = 5, $P < 0.0001$ for WT littermates of *Dgcr8* and *Dicer* cKO litters, respectively).

B Caudal epididymal spermatozoa counts, were significantly lower in *Dgcr8* cKO animals ($P = 0.0007$) than in WT littermates (grey bars). Epididymal spermatozoa were rarely detected in *Dicer* cKO animals (2 of 4 mice; $P < 0.0001$) when compared to WT littermates (black bars).

C Quantification of TUNEL staining from testes sections. TUNEL positive cells were counted from multiple tubules (number specified below) from three mice for each condition to determine the mean number of apoptotic cells per tubule. In both *Dgcr8* and *Dicer* (grey and black bars, respectively) cKO testes, there were significant increases in apoptotic cell counts compared to WT littermates (*Dgcr8*: WT n=152, cKO n=262; 4.07 fold increase, $P < 0.0001$; *Dicer*: WT n=154 cKO n=299; 4.43 fold increase $P < 0.0001$). Error bars indicate SEM (Panels A, B and C).

D, E and F Testes sections from WT (D), *Dgcr8* cKO (E) and *Dicer* cKO (F) mice stained with H&E. Scale bars are 100 μM. Vacuous tubules were observed in both sections of *Dgcr8* (~50% of tubules, E, Inset) and *Dicer* (~50% of tubules, F, inset) cKOs, but never in sections from wildtype littermates (D). Black arrows indicate tubules producing fully elongated spermatozoa in *Dgcr8* cKO animals (E).

G, H and I TUNEL staining of testis sections from wildtype (G), *Dgcr8* cKO (H) and *Dicer* cKO (I) mice, with brown precipitate indicating cells undergoing apoptosis.

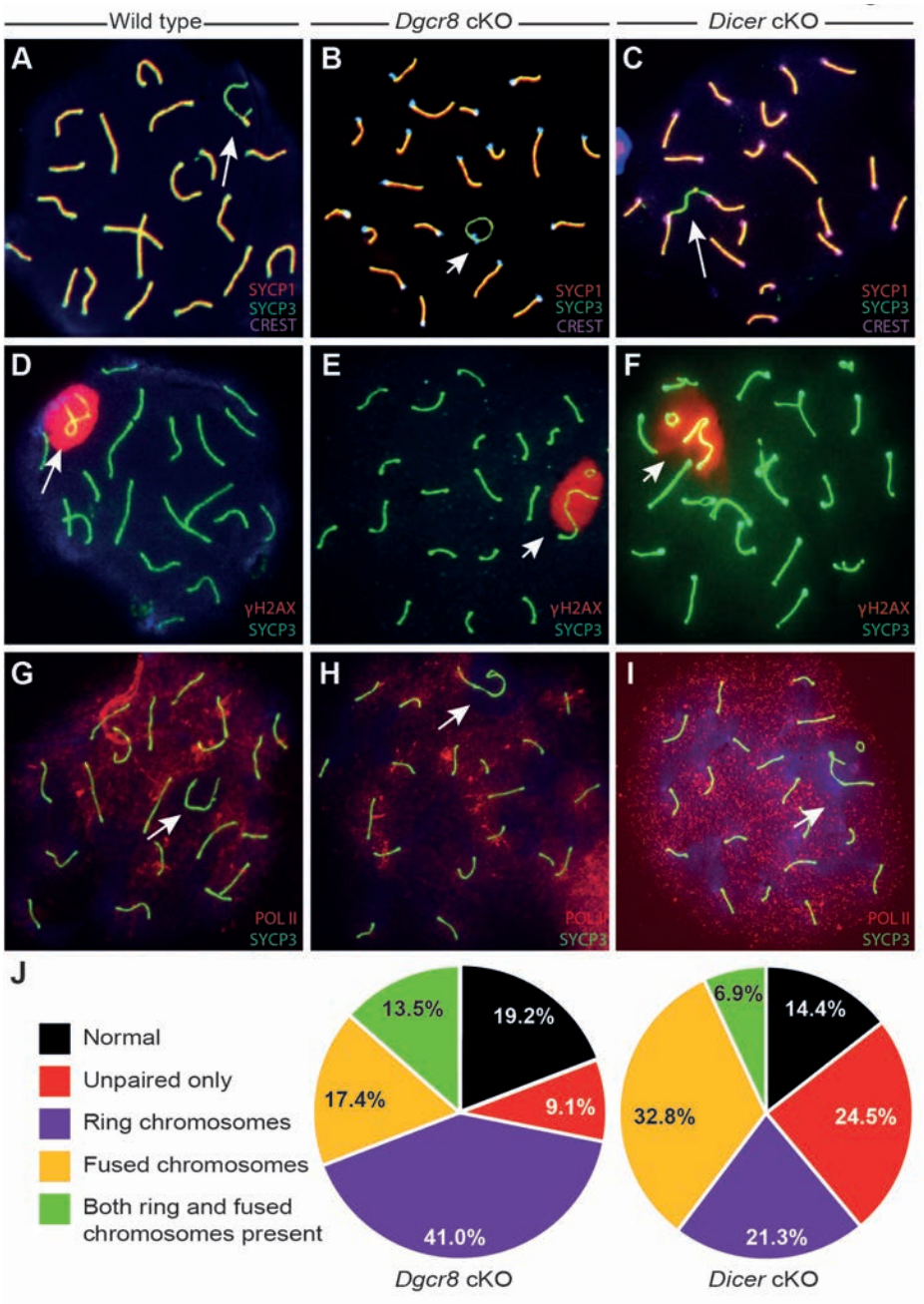

Wild type
Dgcr8 cKO
Dicer cKO
A
B
C
SYCP1
SYCP3
CREST
D
E
F
γH2AX
SYCP3
G
H
I
POL II
SYCP3
J
Normal
Unpaired only
Ring chromosomes
Fused chromosomes
Both ring and fused chromosomes present
19.2%
9.1%
41.0%
17.4%
13.5%
Dgcr8 cKO
14.4%
24.5%
21.3%
32.8%
6.9%
Dicer cKO

Figure 2. Loss of DGCR8 or DICER results in chromosomal fusion abnormalities.

A-I Pachytene staged spermatocytes from wildtype (panels A, D and G) *Dgcr8* cKO (B, E and H) or *Dicer* cKO (C, F and I), each stained with anti-SYCP3 antibody (green); XY chromosomes are indicated by white arrows. Immunostaining was also performed with antibodies specific to SYCP1 and CREST, a human autoimmune serum used to visualize centromeres (A, B and C, in red and purple, respectively), γH2AX (D, E and F, in red), and RNAP II (G, H and I, in red).

J *Dgcr8* cKO and *Dicer* cKO spermatocytes exhibit frequent chromosomal abnormalities. Chromosome spread preparations of pachytene staged spermatocytes were assessed for types and frequencies of observed abnormalities, none of which were ever observed in wildtype pachytene spermatocytes. These were distributed into five indicated categories based on observed frequencies of X and Y chromosomal structures.

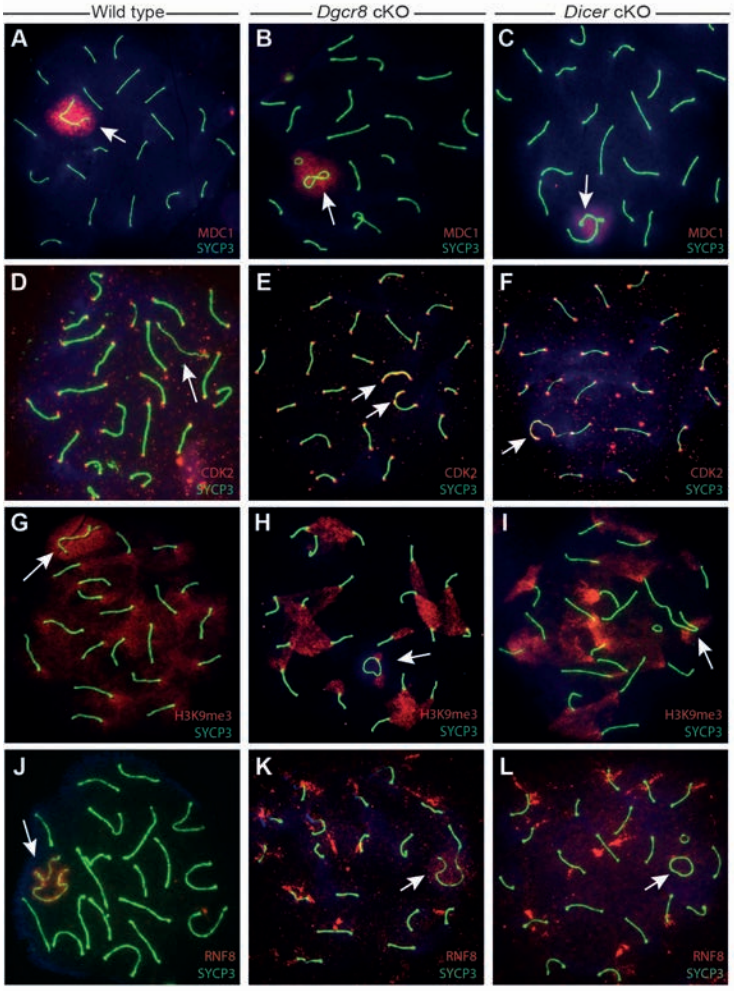

Wild type
Dgcr8 cKO
Dicer cKO
A
B
C
MDC1
SYCP3
D
E
F
CDK2
SYCP3
G
H
I
H3K9me3
SYCP3
J
K
L
RNF8
SYCP3

Figure 3. Loss of DGCR8 or DICER disrupts localization of proteins associated with telomere maintenance.

A-L Pachytene staged spermatocytes from wildtype (A, D, G and J) *Dgcr8* cKO (B, E, H and K) or *Dicer* cKO (C, F, I and L), each stained with anti-SYCP3 antibody (green); XY chromosomes are indicated by white arrows. Immunostaining was also performed with antibodies specific to MDC1 (A, B and C, in red), CDK2 (D, E and F, in red), H3K9me3 (G, H and I, in red) and RNF8 (J, K and L, in red).

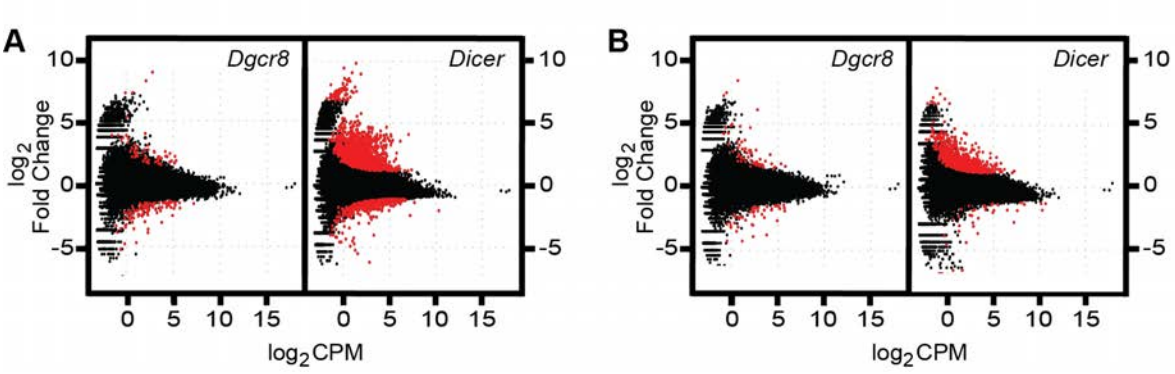

A
Dgcr8
Dicer
log2 Fold Change
log2CPM
B
Dgcr8
Dicer
log2 Fold Change
log2CPM

Figure 4. Differential expression of genes in *Dgcr8* and *Dicer* cKO leptotene/zygotene and pachytene-staged spermatocytes.

A-B The transcriptome was sequenced from purified leptotene/zygotene (L/Z) (A) and pachytene (Pach) (B) cells isolated from either *Dgcr8* or *Dicer* cKO mice, as well as from wildtype littermates as a control. Cells were derived from a minimum of 2 mice per genotype. Each individual cKO sample was compared to the corresponding wildtype littermate using the transcriptome sequencing analysis package edgeR to determine differential expression. Here, each transcript quantified is plotted by its average $\log_2$ CPM (Counts Per Million, average of wildtype and cKO values) and the $\log_2$ of the fold change (cKO CPM/pooled wildtype CPM). In red are genes which are differentially expressed at a 5% false discovery rate (FDR). More genes were upregulated than downregulated in all samples (1.5x-fold-change threshold), and more were differentially expressed in *Dicer* than in *Dgcr8* cKO samples (L/Z: $P < 0.0001$; Pach: $P < 0.0001$; X^2 tests).

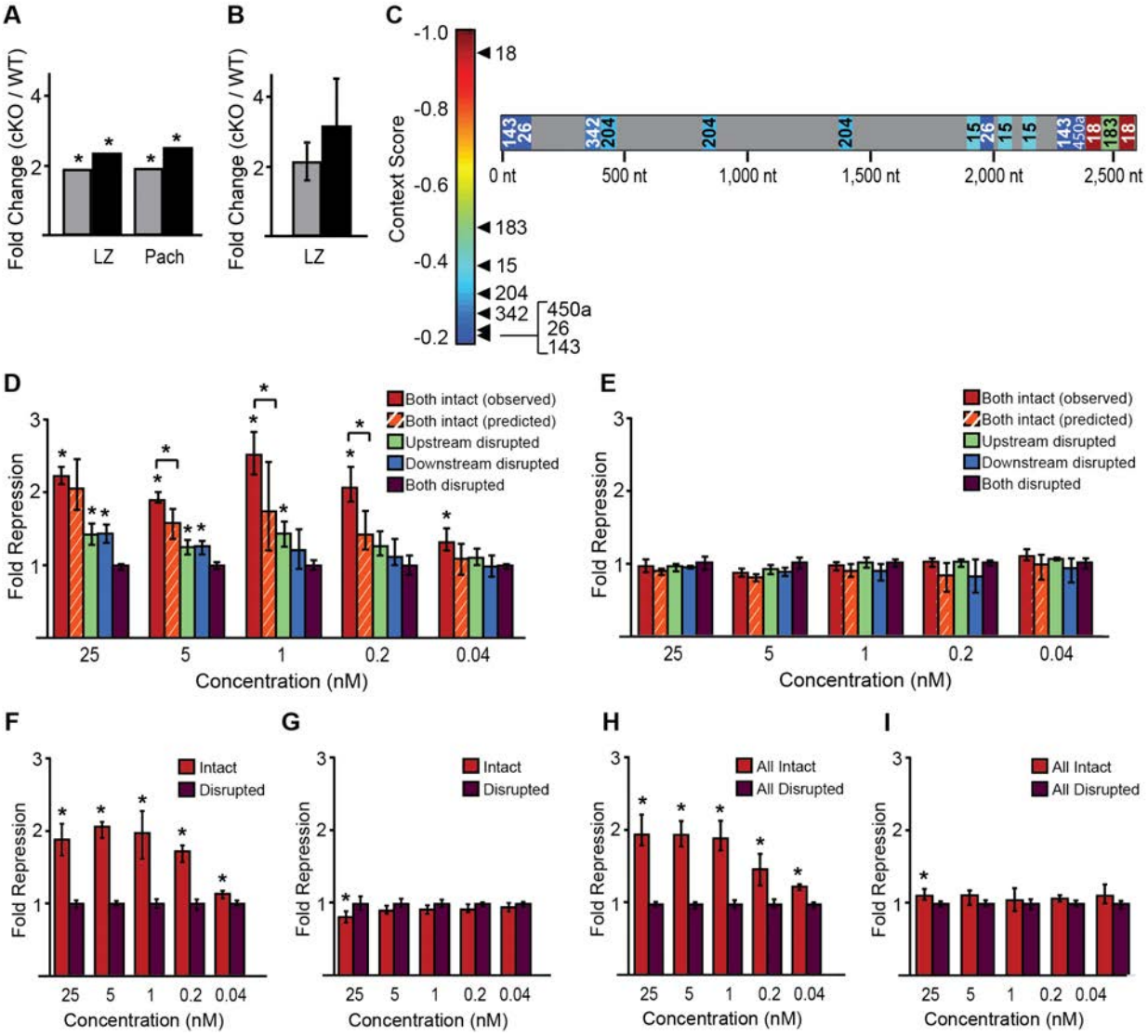

A
Fold Change (cKO / WT)
LZ
Pach
B
Fold Change (cKO / WT)
LZ
C
Context Score
-1.0
-0.8
-0.6
-0.4
-0.2
18
183
15
204
342
450a
26
143
0 nt
500 nt
1,000 nt
1,500 nt
2,000 nt
2,500 nt
D
Fold Repression
Concentration (nM)
Both intact (observed)
Both intact (predicted)
Upstream disrupted
Downstream disrupted
Both disrupted
E
Fold Repression
Concentration (nM)
Both intact (observed)
Both intact (predicted)
Upstream disrupted
Downstream disrupted
Both disrupted
F
Fold Repression
Concentration (nM)
Intact
Disrupted
G
Fold Repression
Concentration (nM)
Intact
Disrupted
H
Fold Repression
Concentration (nM)
All Intact
All Disrupted
I
Fold Repression
Concentration (nM)
All Intact
All Disrupted

Figure 5. *Atm* expression is upregulated in *Dgcr8* and *Dicer* cKOs and a target of germline-expressed miR-18, as well as miR-183 and miR-16.

A Quantification of *Atm* by RNA-seq of purified leptotene and zygotene (L/Z) and pachytene spermatocytes (Pach) from *Dgcr8* (grey bars) and *Dicer* (black bars) cKO animals. The y-axis indicates fold change in read counts of each sample compared to wildtype samples.

B RT-qPCR quantification of *Atm* from *Dgcr8* (grey bars) and *Dicer* (black bars) L/Z spermatocytes plotted as fold change relative to wildtype transcript abundance.

C Schematic of *Atm* 3′UTR (right) illustrating position of all miRNA target sites corresponding to miRNAs expressed in the male germ line. Sites are color-coded by context score[46–49], which predicts site efficacy; increasingly negative scores indicate increasing predicted repression. For miRNAs with multiple target sites, the cumulative context score is plotted; only sites with a context score of -0.2 or below included.

D-E Reporter assays of miR-18 mediated regulation of *Atm*. Luciferase reporter constructs containing a portion of the *Atm* 3′UTR (red bars) were transfected into A549 cells together with different concentrations (X-axis) of a miR-18 mimetic siRNA duplex (D) or miR-124, as a control (E). Fold regulation mediated by miR-18 was calculated by generating reporter constructs containing mutations disrupting both predicted target sites (purple bars) and normalizing luciferase activity to this construct. Similarly, to determine the efficacy of each individual site, reporter constructs containing mutations disrupting each individual site were also assayed (green and blue bars, representing disruption of the upstream and downstream site, respectively). Asterisks indicate reporter constructs mediating statistically significant (i.e. $P < 0.01$) repression ($P < 7\text{x}10^{-4}$ for wildtype reporter at all concentrations, as well as reporters containing a single site at 25 and 5 nM; $P < 3\text{x}10^{-3}$ for the upstream site disrupted at 1nM;

Bonferroni-corrected Wilcoxon-rank sum tests). Orange-striped bars indicate the expected fold repression assuming that the two sites contribute independently to total repression (expected repression = repression of upstream site x repression of downstream site). Concentrations of miR-18 at which target sites contribute synergistically to repression are indicated with an asterisk and bracket above red and orange-striped bars; synergism was inferred when the observed repression of the *Atm* 3′UTR (red bar) significantly ($P < 0.01$) exceeded that expected based on measurements of each site individually (orange-striped bar) (5nM: $P = 0.0022$, 1nM: $P = 0.0029$, 0.2nM: $P = 0.0006$; Bonferroni-corrected Wilcoxon-rank sum test).

F-I Reporter assays of miR-183 and miR-16-mediated regulation of *Atm.* Performed and analyzed as for miR-18. For miR-183, we compared the wildtype *Atm* construct ('Intact') to one in which the single miR-183 site was disrupted ('Disrupted') in the presence of a miR-183 mimetic (F) or a control miR-124 mimetic (G). For miR-16, we compare the wildtype *Atm* construct ('All Intact') to one in which all three miR-16 sites are mutated ('All Disrupted') in the presence of a miR-16 mimetic (H) or a control miR-124 mimetic (I).

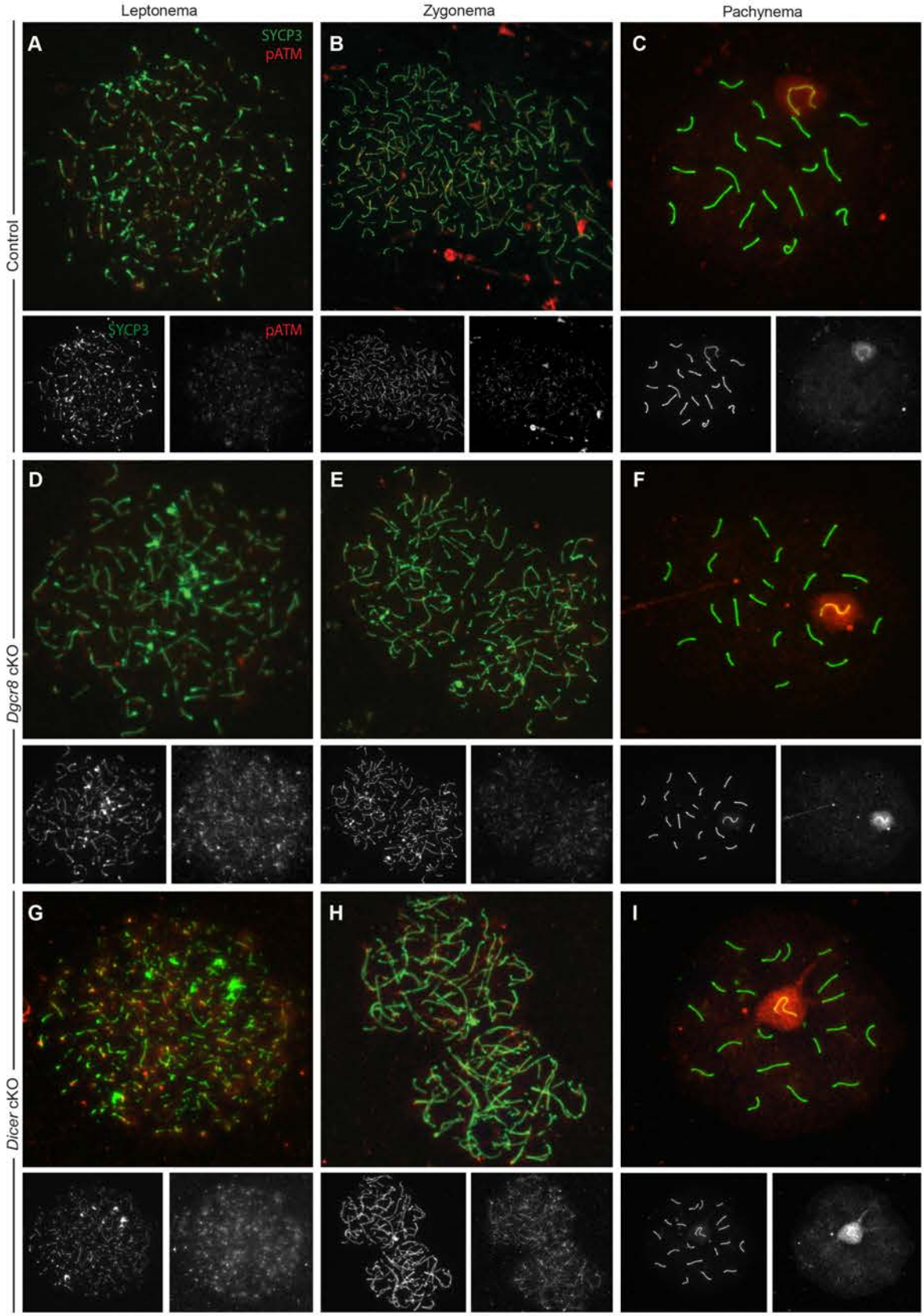

Leptonema
Zygonema
Pachynema
Control
Dgcr8 cKO
Dicer cKO
A
B
C
D
E
F
G
H
I
SYCP3
pATM
SYCP3
pATM

Figure 6. Phospho-ATM protein localization is upregulated in prophase I chromosome spreads from *Dicer* and *Dgcr8* cKO males.

A-I Spermatocytes from wildtype (A-C), *Dgcr8* cKO (D-F), and *Dicer* cKO (G-I) were stained with anti-SYCP3 (green) and anti-ATMpS1981 (red) antibodies. The red and green channels for each image are shown below the respective merged image. Phospho-ATM levels are higher in the *Dicer* cKO (D-E) and *Dgcr8* cKO (G-H) as compared to wildtype (A-B) in leptotene and zygotene staged spermatocytes.

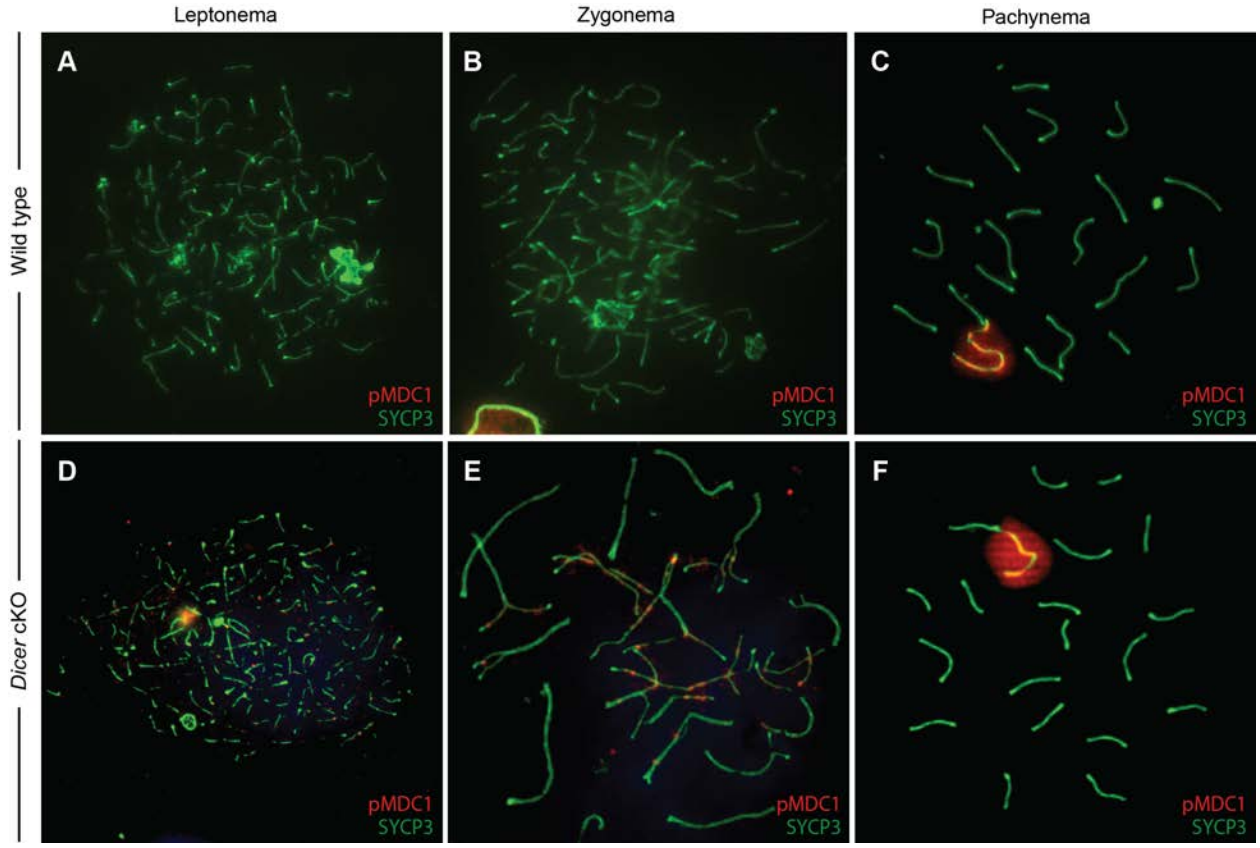

Leptonema
Zygonema
Pachynema
Wild type
Dicer cKO
A
B
C
D
E
F
pMDC1
SYCP3

Figure 7. Phosphorylated MDC1 (pMDC1) mislocalization in *Dicer* cKO prophase I chromosome spreads.

A-F Spermatocytes from wildtype (A-C) and *Dicer* cKO (D-F) mice stained with anti-SYCP3 (green) and anti-phospho-MDC1 (red) antibodies. pMDC1 was mislocalized in 80% of leptotene spermatocytes.

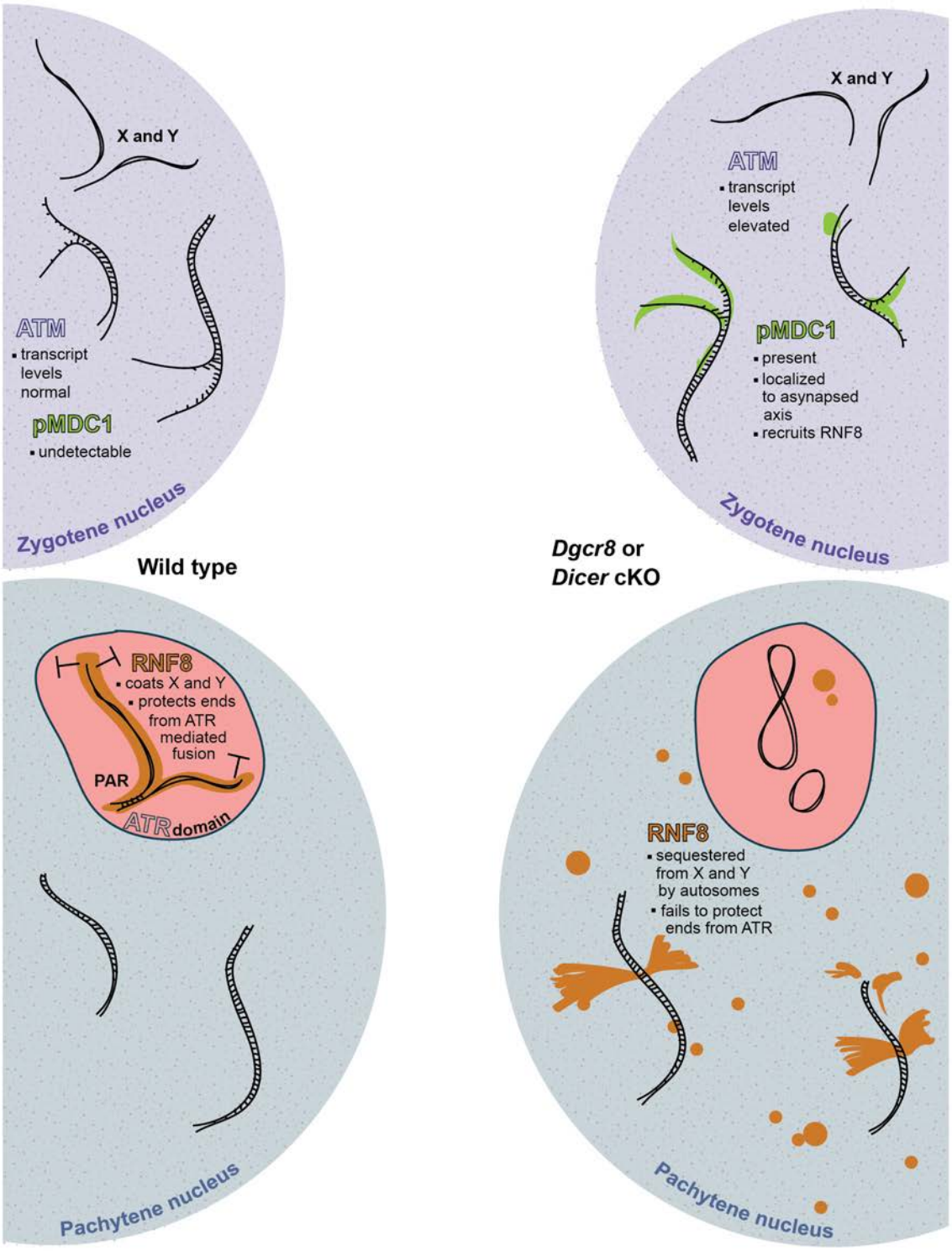

X and Y
ATM
▪ transcript levels normal
pMDC1
▪ undetectable
Zygotene nucleus
Wild type
X and Y
ATM
▪ transcript levels elevated
pMDC1
▪ present
▪ localized to asynapsed axis
▪ recruits RNF8
Zygotene nucleus
Dgcr8 or Dicer cKO
RNF8
▪ coats X and Y
▪ protects ends from ATR mediated fusion
PAR
ATR domain
Pachytene nucleus
RNF8
▪ sequestered from X and Y by autosomes
▪ fails to protect ends from ATR
Pachytene nucleus

Figure 8. Model for the role of miRNA-mediated regulation in controlling sex chromosome stability. In wildtype males at the zygotene stage (left, upper), pMDC1 is undetectable on both the autosomes and sex chromosomes. In *Dgcr8* and *Dicer* cKO (right, upper), pMDC1 is evident and localizes to asynapsed cores. Additionally, the transcript encoding ATM, which phosphorylates MDC1, is upregulated at this stage in cKOs. By the subsequent pachytene stage, RNF8 (whose localization is driven by pMDC1) is enriched within the sex-body in wildtype males (left, lower), coating the X and Y chromosomes, where it functions to protect telomeres from triggering ATR. In males deficient for miRNAs (right, lower), *Atm* elevation and the resulting pMDC1 mislocalization at zygotene recruits RNF8, and likely other members of the DNA damage repair pathway, to the autosomes. Sequestration of RNF8, and potentially other factors, away from the sex chromosomes deprotects telomeres, leading to ATR-mediated chromosomal fusions via non-homologous end-joining.

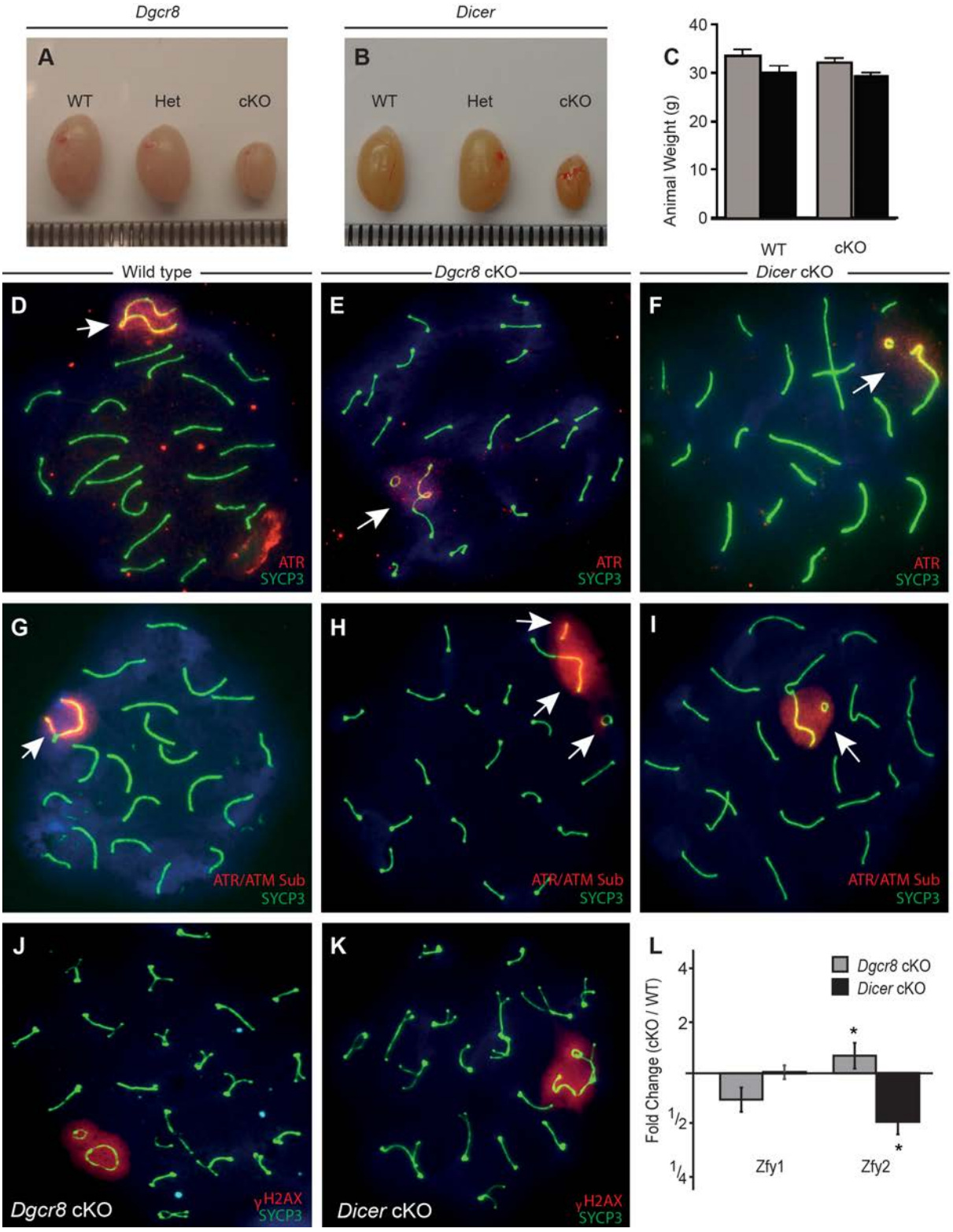

Dgcr8
Dicer
A
WT
Het
cKO
B
WT
Het
cKO
C
Animal Weight (g)
40
30
20
10
0
WT
cKO
Wild type
Dgcr8 cKO
Dicer cKO
D
ATR
SYCP3
E
ATR
SYCP3
F
ATR
SYCP3
G
ATR/ATM Sub
SYCP3
H
ATR/ATM Sub
SYCP3
I
ATR/ATM Sub
SYCP3
J
Dgcr8 cKO
γH2AX
SYCP3
K
Dicer cKO
γH2AX
SYCP3
L
Dgcr8 cKO
Dicer cKO
Fold Change (cKO / WT)
4
2
1/2
1/4
*
*
Zfy1
Zfy2

Figure S1. Phenotypic analysis and assessment of meiotic silencing in *Dgcr8* and *Dicer* cKO males.

A-B Photographic representation of whole testes removed from *Dgcr8* (A) and *Dicer* (B) litters (wild type (WT): Fl/+, cre-, heterozygote (Het): Δ/+ cre and cKO: Fl/Δ cre+).

C Total WT and cKO animal weights are not significantly different from one another. Error bars, SEM.

D-I Impact of loss of DGCR8 or DICER on localization of the kinase ATR as well as subsequent targets. Pachytene-staged spermatocytes with XY chromosomes (white arrows) from wildtype control (D,G), *Dgcr8* cKO (E,H) and *Dicer* cKO (F,I) mice, stained with anti-SYCP3 (green), anti-ATR (D-F, in red, source: GeneTex GTX70133) and anti-ATR/ATM Substrate (G-I, in red, source: Cell Signaling #5851).

J-K Defective spermatocytes able to progress beyond the pachytene checkpoint. *Dgcr8* cKO diplotene-staged spermatocyte in which both X and Y-chromosomes have circularized with themselves with clear γH2AX localization restricted to the typical sex body boundaries (J). *Dicer* cKO diplotene-staged spermatocyte with circularization of Y and fusion of X to an autosome in which silencing mark γH2AX has spread to entire structure (K).

L Sex chromosome abnormalities do not lead to a global failure to silence key sex chromosome genes *Zfy1* and *Zfy2.* The relative abundance of the *Zfy1* and *Zfy2* transcripts were quantified in purified pachytene spermatocytes from sibling mice of wildtype and *Dgcr8* or *Dicer* cKO genotypes. qPCR was performed with TaqMan probes (Applied Biosystems, Carlsbad CA, USA; Zfy1: Mm00494343_g1 Lot- 1061650, Zfy2: Mm00494350_m1 Lot-965381, ActB: 4352341E - 1102015) and run on a Lightcycler 480 (Roche). Error bars indicate the 95% confidence interval. Significantly different fold change values ($P < 0.01$) are marked with an asterisk.

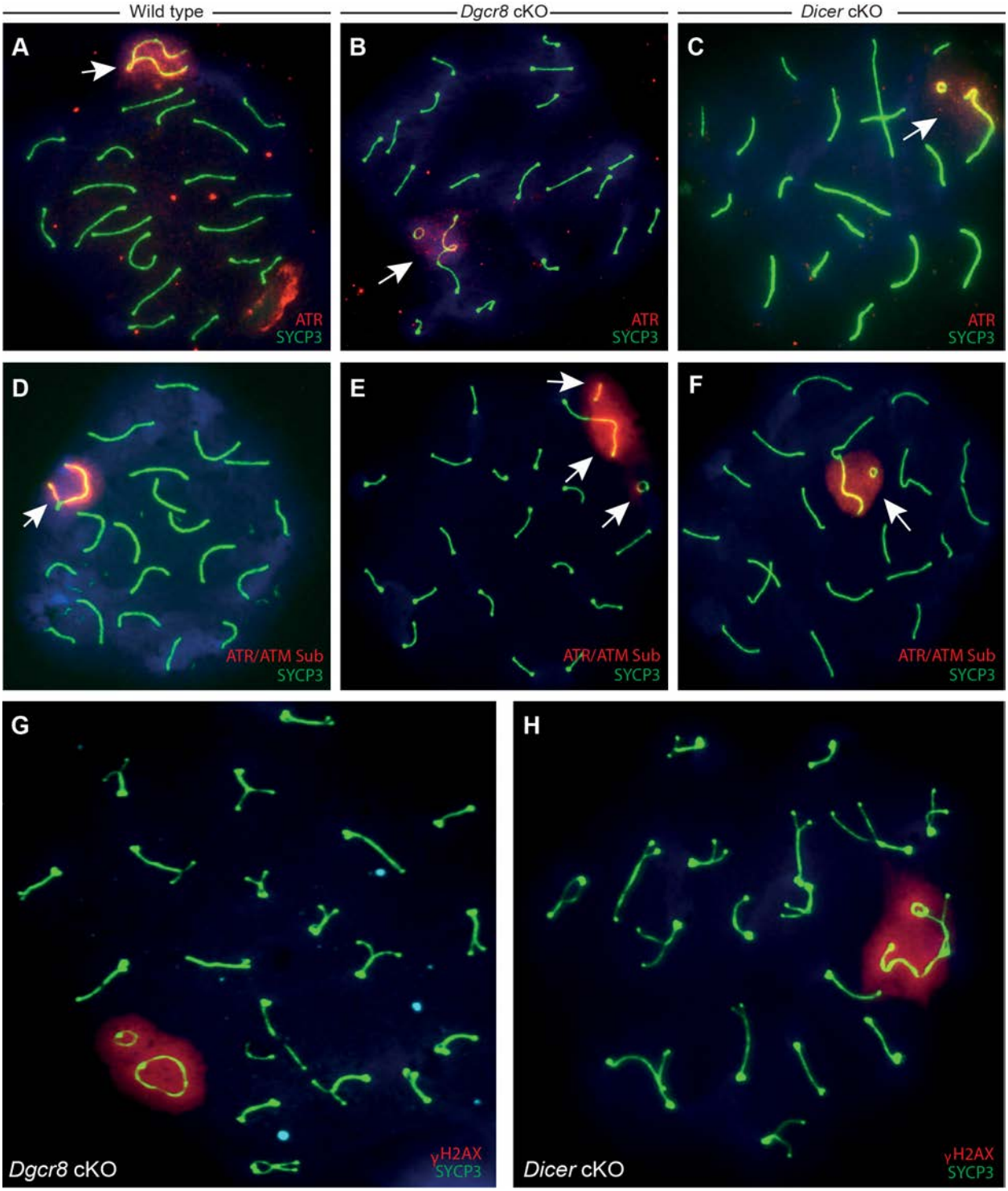

Wild type
Dgcr8 cKO
Dicer cKO
A
ATR
SYCP3
B
ATR
SYCP3
C
ATR
SYCP3
D
ATR/ATM Sub
SYCP3
E
ATR/ATM Sub
SYCP3
F
ATR/ATM Sub
SYCP3
G
Dgcr8 cKO
γH2AX
SYCP3
H
Dicer cKO
γH2AX
SYCP3

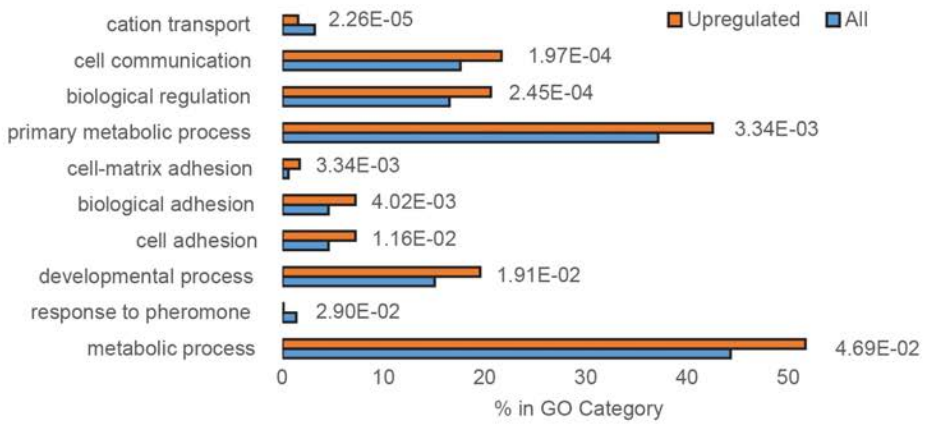
Figure S2. Assessment of gene knockdown efficacy and spermatocyte purification from *Dgcr8* and *Dicer* cKO males.

A-F Immunohistochemical localization of DGCR8 (A-C) protein on testis sections from wildtype (WT; A) and cKO (B) mice, together with a negative control IgG-stained WT section (C). Immunohistochemical localization of DICER (D-F) protein on testis sections from WT (D) and cKO (E) mice, together with a negative control IgG-stained WT slide (F). Rabbit anti-DGCR8 antibody was obtained from Proteintech (10996-1) and rabbit anti-DICER antibody was obtained from Novus Biological (NBP1-71691). DICER localizes to the cytoplasm of spermatogonia and spermatocytes in WT testes (D, black arrows), while DGCR8 localization is strongest within the nucleus of primary spermatocytes (A, black arrowheads). Importantly, while DGCR8 signal is completely absent in cKO spermatocytes (B, white arrowheads), DICER signal persists in the cytoplasm of cKO males, albeit at reduced intensity (E, white arrows). Positive signal corresponds to strong brown staining, background signal intensity can be judged from the IgG controls.

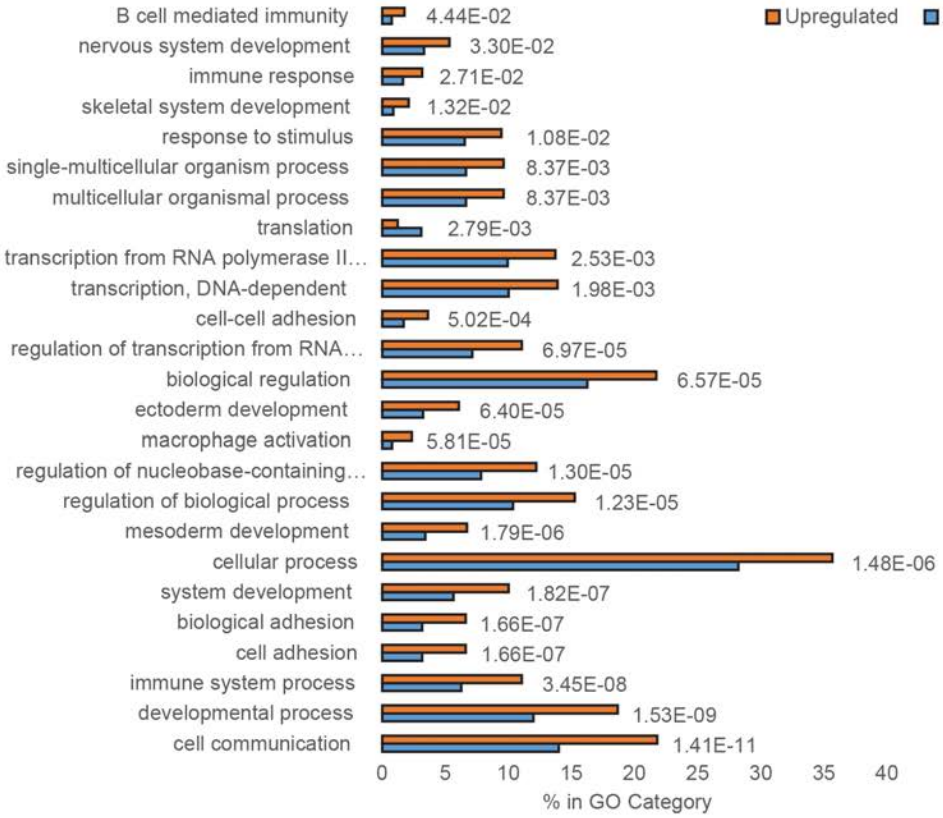
G-H The relative abundance of the WT and knockout transcript isoforms of *Dgcr8* (G) and *Dicer* (H) were quantified in purified cells from two different stages of spermatogenesis, leptotene/zygotene (L/Z) and pachytene (Pach), from sibling mice of each genotype (WT: wild type; HET: heterozygote; and cKO). The knockout isoform of *Dgcr8* was not detected in WT *Dgcr8* samples; conversely the WT isoform of *Dgcr8* was detected at only very low levels in *Dgcr8* cKO samples. While the knockout isoform of *Dicer* was apparently detected at very low levels in WT *Dicer* samples, this is likely an artifact due to the nature of the qPCR assay; however the WT isoform of Dicer constituted about 50% of the *Dicer* transcript detected in both *Dicer* cKO samples. qPCR assays were performed as described for ATM in the Materials and

Methods section, and used the following primers: Dgcr8_wtExon3 (TGGAGAGACAAGTGTACAGCC and AGGCAATGGCTCTGTAGGTG); Dgcr8_ΔExon3 (TTTCTCCTATGAGGTCGTGGC and GATCCATCCATCAGGCAATGG); Dicer_wtExon22/23 (TGGCTTCCTCCTGGTTATGTG and GTTTGCCATTAGCCAGCAAGC); Dicer_ΔExon22/23 (CTGTTTTGCACGTACCCTGATG and TTGGGGACTTCGATATCCTCTTC); GAPDH (TGAAGCAGGCATCTGAGGG and CGAAGGTGGAAGAGTGGGAG).

A



B



C

Pathway	Fold Enrichment	p-value
Apoptosis signaling pathway	4.05	5.33E-07
Integrin signalling pathway	3.19	6.88E-06
Gonadotropin releasing hormone receptor pathway	2.29	1.50E-03
Toll receptor signaling pathway	3.61	1.15E-02
EGF receptor signaling pathway	2.48	4.49E-02

D

Pathway	Fold Enrichment	p-value
Integrin signalling pathway	2.52	8.30E-05
Gonadotropin releasing hormone receptor pathway	2.24	8.16E-04
Angiogenesis	2.46	2.05E-03

Figure S3. Gene ontology and pathway analysis of genes upregulated in both *Dgcr8* and *Dicer* cKOs.

A-D RNA-Seq was used to quantify transcript levels in purified leptotene/zygotene (L/Z) and pachytene (Pach) spermatocytes from *Dgcr8* or *Dicer* WT or cKO mice. We identified those genes which were upregulated 1.5x or more in both the *Dgcr8* and *Dicer* cKO, as compared to the wildtype. We then used PANTHER (*Thomas et al., 2003, Genome Res., 13: 2129-2141*) to identify biological processes (A, B) and pathways (C, D) in which these upregulated genes are overrepresented as compared to all expressed genes at leptotene/zygotene (A, C) and pachytene (B, D).

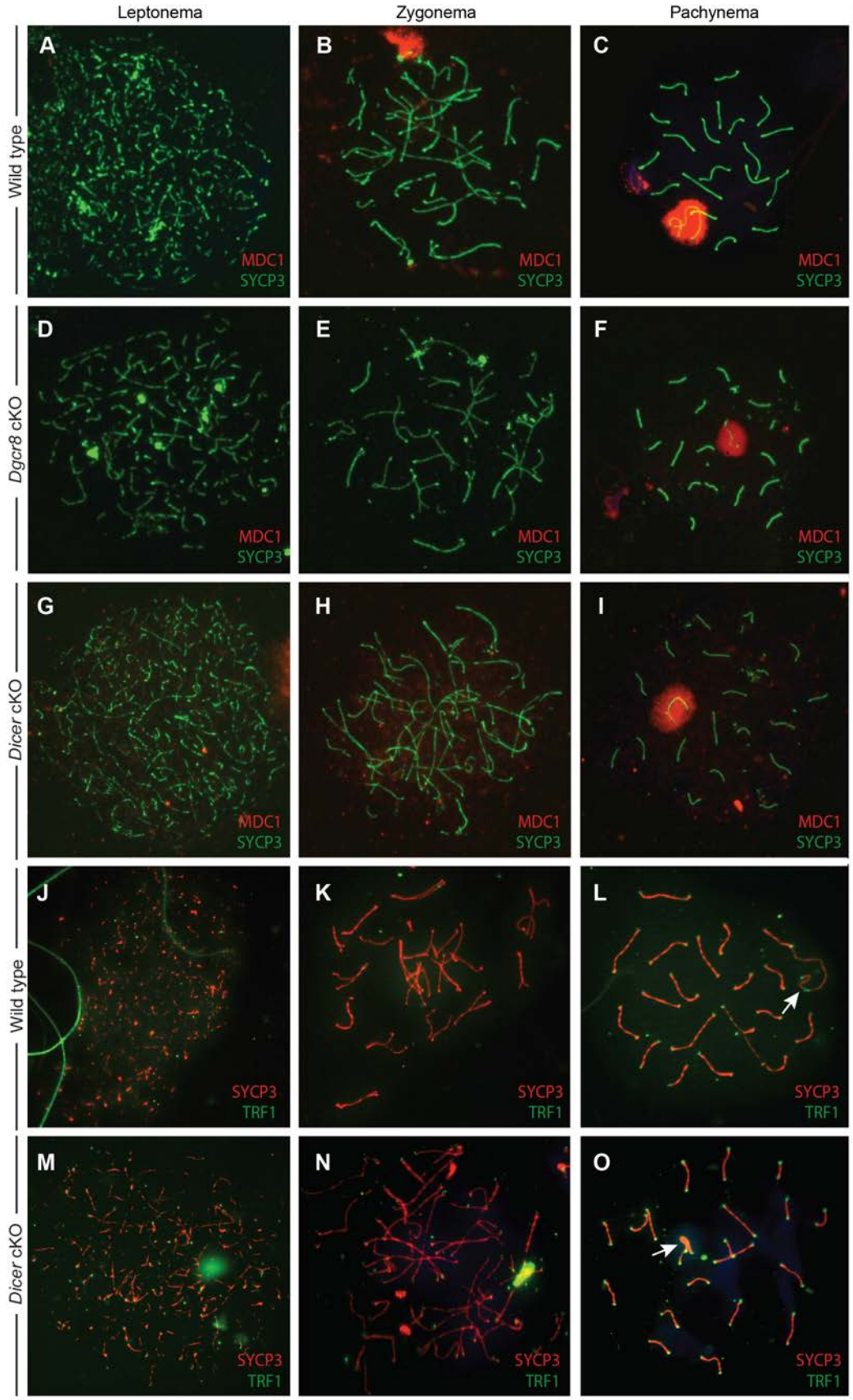

Leptonema
Zygonema
Pachynema
Wild type
Dgcr8 cKO
Dicer cKO
Wild type
Dicer cKO
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
MDC1
SYCP3
SYCP3
TRF1

Figure S4. MDC1 and TRF localization in prophase I chromosome spreads from *Dicer* and *Dgcr8* cKO male mice.

A-I Spermatocytes from wildtype (A-C), *Dgcr8* cKO (D-F), and *Dicer* cKO (G-I) mice were stained with anti-SYCP3 (green) and anti-MDC1 (red; Raimundo Freire, Tenerife, Spain) antibodies. Localization of MDC1 protein does not appear to be significantly changed between WT and the *Dicer* cKO or *Dgcr8* cKO throughout prophase I.

J-O Spermatocyte nuclei from the leptotene, zygotene, and pachytene stages of prophase I from wildtype (A-C) and *Dicer* cKO (D-F) mice stained with anti-SYCP3 (red) and anti-TRF1 (green, Abcam ab10579) antibodies. Localization of TRF1 protein does not appear to change between wild type and the *Dicer* cKO throughout prophase I.

References

1. Fire, Xu, Montgomery, Kostas, Driver, Mello: Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* 1998, 391:806–11.

2. Matzke M, Birchler J: RNAi-mediated pathways in the nucleus. *Nature reviews. Genetics* 2005, 6:24–35.

3. Wassenegger, Heimes, Riedel, Sänger: RNA-directed de novo methylation of genomic sequences in plants. *Cell* 1994, 76:567–76.

4. Ketting R: The many faces of RNAi. *Developmental cell* 2011, 20:148–61.

5. Suh N, Baehner L, Moltzahn F, Melton C, Shenoy A, Chen J, Blelloch R: MicroRNA function is globally suppressed in mouse oocytes and early embryos. *Curr. Biol.* 2010, 20:271–7.

6. García-López J, del Mazo J: Expression dynamics of microRNA biogenesis during preimplantation mouse development. *Biochimica et biophysica acta* 2012, 1819:847–54.

7. Ma J, Flemr M, Stein P, Berninger P, Malik R, Zavolan M, Svoboda P, Schultz R: MicroRNA Activity Is Suppressed in Mouse Oocytes. *Current Biology* 2010, 20:265–270.

8. Yadav R, Kotaja N: Small RNAs in spermatogenesis. *Molecular and Cellular Endocrinology* 2014, 382:498–508.

9. Hayashi K, de Lopes S, Kaneda M, Tang F, Hajkova P, Lao K, O'Carroll D, Das P, Tarakhovsky A, Miska E, et al.: MicroRNA biogenesis is required for mouse primordial germ cell development and spermatogenesis. *PloS one* 2008, 3:e1738.

10. He Z, Kokkinaki M, Pant D, Gallicano, Dym M: Small RNA molecules in the regulation of spermatogenesis. *Reproduction (Cambridge, England)* 2009, 137:901–11.

11. Maatouk DM, Loveland KL, McManus MT, Moore K, Harfe BD: Dicer1 is required for differentiation of the mouse male germline. *Biology of reproduction* 2008, 79:696–703.

12. Papaioannou M, Nef S: microRNAs in the testis: building up male fertility. *Journal of andrology* 2010, 31:26–33.

13. Zimmermann C, Romero Y, Warnefors M, Bilican A, Borel C, Smith LB, Kotaja N, Kaessmann H, Nef S: Germ Cell-Specific Targeting of DICER or DGCR8 Reveals a Novel Role for Endo-siRNAs in the Progression of Mammalian Spermatogenesis and Male Fertility. *PloS one* 2014, 9:e107023.

14. Wu Q, Song R, Ortogero N, Zheng H, Evanoff R, Small CL, Griswold MD, Namekawa SH, Royo H, Turner JM, et al.: The RNase III enzyme DROSHA is essential for microRNA production and spermatogenesis. *The Journal of biological chemistry* 2012, 287:25173–90.

15. Meister G: Argonaute proteins: functional insights and emerging roles. *Nature reviews. Genetics* 2013, 14:447–59.

16. Cenik E, Zamore P: Argonaute proteins. *Current biology : CB* 2011, 21:R446–9.

17. Maine E: Meiotic silencing in Caenorhabditis elegans. *International review of cell and molecular biology* 2010, 282:91–134.

18. Shiu PK, Raju NB, Zickler D, Metzenberg RL: Meiotic silencing by unpaired DNA. *Cell*

2002, 107:905–16.

19. Meister G, Landthaler M, Patkaniowska A, Dorsett Y, Teng G, Tuschl T: Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Molecular cell* 2004, 15:185–97.

20. Robb, Brown K, Khurana J, Rana T: Specific and potent RNAi in the nucleus of human cells. *Nature structural & molecular biology* 2005, 12:133–7.

21. Tan G, Garchow B, Liu X, Yeung J, Morris J, Cuellar T, McManus M, Kiriakidou M: Expanded RNA-binding activities of mammalian Argonaute 2. *Nucleic acids research* 2009, 37:7533–45.

22. Chu Y, Yue X, Younger S, Janowski B, Corey D: Involvement of argonaute proteins in gene silencing and activation by RNAs complementary to a non-coding transcript at the progesterone receptor promoter. *Nucleic acids research* 2010, 38:7736–48.

23. Modzelewski AJ, Holmes RJ, Hilz S, Grimson A, Cohen PE: AGO4 regulates entry into meiosis and influences silencing of sex chromosomes in the male mouse germline. *Developmental cell* 2012, 23:251–64.

24. Murchison E, Hannon G: miRNAs on the move: miRNA biogenesis and the RNAi machinery. *Current opinion in cell biology* 2004, 16:223–9.

25. Kim NV, Han J, Siomi MC: Biogenesis of small RNAs in animals. *Nature reviews Molecular cell biology* 2009, 10:126–139.

26. Han J, Lee Y, Yeom K-H, Kim Y-K, Jin H, Kim: The Drosha-DGCR8 complex in primary microRNA processing. *Genes & development* 2004, 18:3016–27.

27. Bartel D: MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 2004, 116:281–97.

28. Gregory R, Yan K-P, Amuthan G, Chendrimada T, Doratotaj B, Cooch N, Shiekhattar R: The Microprocessor complex mediates the genesis of microRNAs. *Nature* 2004, 432:235–40.

29. Denli A, Tops B, Plasterk R, Ketting R, Hannon G: Processing of primary microRNAs by the Microprocessor complex. *Nature* 2004, 432:231–5.

30. Gallardo T, Shirley L, John G, Castrillon D: Generation of a germ cell-specific mouse transgenic Cre line, Vasa-Cre. *Genesis (New York, N.Y. : 2000)* 2007, 45:413–7.

31. Greenlee AR, Shiao M-SS, Snyder E, Buaas FW, Gu T, Stearns TM, Sharma M, Murchison EP, Puente GC, Braun RE: Deregulated sex chromosome gene expression with male germ cell-specific loss of Dicer1. *PloS one* 2012, 7:e46359.

32. Korhonen HM, Meikar O, Yadav RP, Papaioannou MD, Romero Y, Da Ros M, Herrera PL, Toppari J, Nef S, Kotaja N: Dicer is required for haploid male germ cell differentiation in mice. *PloS one* 2011, 6:e24821.

33. Romero Y, Meikar O, Papaioannou M, Conne B, Grey C, Weier M, Pralong F, Massy B, Kaessmann H, Vassalli J-D, et al.: Dicer1 depletion in male germ cells leads to infertility due to cumulative meiotic and spermiogenic defects. *PloS one* 2011, 6:e25241.

34. Meuwissen, Offenberg, Dietrich, Riesewijk, van Iersel, Heyting: A coiled-coil related protein specific for synapsed regions of meiotic prophase chromosomes. *The EMBO journal* 1992, 11:5091–100.

35. Yuan, Liu, Zhao, Brundell, Daneholt, Höög: The murine SCP3 gene is required for synaptonemal complex assembly, chromosome synapsis, and male fertility. *Molecular cell* 2000, 5:73–83.

36. Heard E, Turner J: Function of the sex chromosomes in mammalian fertility. *Cold Spring Harbor perspectives in biology* 2011, 3:a002675.

37. Royo H, Polikiewicz G, Mahadevaiah SK, Prosser H, Mitchell M, Bradley A, de Rooij DG, Burgoyne PS, Turner JM: Evidence that meiotic sex chromosome inactivation is essential for male fertility. *Current biology : CB* 2011, 20:2117–23.

38. Parra M, Viera A, Gómez R, Page J, Carmena M, Earnshaw W, Rufas J, Suja J: Dynamic relocalization of the chromosomal passenger complex proteins inner centromere protein (INCENP) and aurora-B kinase during male mouse meiosis. *Journal of cell science* 2003, 116:961–74.

39. Ashley, Walpita, de Rooij: Localization of two mammalian cyclin dependent kinases during mammalian meiosis. *Journal of cell science* 2001, 114:685–93.

40. Kim S, Namekawa S, Niswander L, Ward J, Lee J, Bardwell V, Zarkower D: A mammal-specific Doublesex homolog associates with male sex chromatin and is required for male meiosis. *PLoS genetics* 2007, 3:e62.

41. Tachibana M, Nozaki M, Takeda N, Shinkai Y: Functional dynamics of H3K9 methylation during meiotic prophase progression. *The EMBO journal* 2007, 26:3346–59.

42. Mailand N, Bekker-Jensen S, Faustrup H, Melander F, Bartek J, Lukas C, Lukas J: RNF8 ubiquitylates histones at DNA double-strand breaks and promotes assembly of repair proteins. *Cell* 2007, 131:887–900.

43. Rai R, Li J-M, Zheng H, Lok G, Deng Y, Huen M, Chen J, Jin J, Chang S: The E3 ubiquitin ligase Rnf8 stabilizes Tpp1 to promote telomere end protection. *Nature structural & molecular biology* 2011, 18:1400–7.

44. Bakkenist C, Kastan M: DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation. *Nature* 2003, 421:499–506.

45. Thomas M, Lieberman J, Lal A: Desperately seeking microRNA targets. *Nature Structural & Molecular Biology* 2010, 17:1169–1174.

46. Lewis B, Burge C, Bartel D: Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* 2005, 120:15–20.

47. Grimson A, Farh K, Johnston W, Garrett-Engele P, Lim L, Bartel D: MicroRNA Targeting Specificity in Mammals: Determinants beyond Seed Pairing. *Molecular Cell* 2007, 27:91–105.

48. Friedman R, Farh K, Burge C, Bartel D: Most mammalian mRNAs are conserved targets of microRNAs. *Genome Research* 2009, 19:92–105.

49. Garcia DM, Baek D, Shin C, Bell GW, Grimson A, Bartel DP: Weak seed-pairing stability and high target-site abundance decrease the proficiency of lsy-6 and other microRNAs. *Nat. Struct. Mol. Biol.* 2011, 18:1139–46.

50. Björk J, Sandqvist A, Elsing A, Kotaja N, Sistonen L: miR-18, a member of Oncomir-1, targets heat shock transcription factor 2 in spermatogenesis. *Development (Cambridge, England)*

2010, 137:3177–84.

51. Farh KK, Grimson A, Jan C, Lewis BP, Johnston WK, Lim LP, Burge CB, Bartel DP: The widespread impact of mammalian MicroRNAs on mRNA repression and evolution. *Science (New York, N.Y.)* 2006, 310:1817–21.

52. Krützfeldt J, Rajewsky N, Braich R, Rajeev K, Tuschl T, Manoharan M, Stoffel M: Silencing of microRNAs in vivo with "antagomirs". *Nature* 2005, 438:685–9.

53. Rodriguez A, Vigorito E, Clare S, Warren M, Couttet P, Soond D, van Dongen S, Grocock R, Das P, Miska E, et al.: Requirement of bic/microRNA-155 for normal immune function. *Science (New York, N.Y.)* 2007, 316:608–11.

54. Saetrom P, Heale B, Snøve O, Aagaard L, Alluin J, Rossi J: Distance constraints between microRNA target sites dictate efficacy and cooperativity. *Nucleic acids research* 2006, 35:2333–42.

55. Di Fagagna F: A direct role for small non-coding RNAs in DNA damage response. *Trends in cell biology* 2014, 24:171–8.

56. Bolcun-Filas E, Schimenti J: Genetics of meiosis and recombination in mice. *International review of cell and molecular biology* 2011, 298:179–227.

57. Bellani M, Romanienko P, Cairatti D, Camerini-Otero: SPO11 is required for sex-body formation, and Spo11 heterozygosity rescues the prophase arrest of Atm-/- spermatocytes. *Journal of cell science* 2005, 118:3233–45.

58. Subramanian V, Hochwagen A: The Meiotic Checkpoint Network: Step-by-Step through Meiotic Prophase. *Cold Spring Harbor perspectives in biology* 2013, 6.

59. Royo H, Prosser H, Ruzankina Y, Mahadevaiah S, Cloutier J, Baumann M, Fukuda T, Höög C, Tóth A, de Rooij D, et al.: ATR acts stage specifically to regulate multiple aspects of mammalian meiotic silencing. *Genes & development* 2013, 27:1484–94.

60. Ichijima Y, Ichijima M, Lou Z, Nussenzweig A, Camerini-Otero, Chen J, Andreassen P, Namekawa S: MDC1 directs chromosome-wide silencing of the sex chromosomes in male germ cells. *Genes & development* 2011, 25:959–71.

61. Viera A, Rufas J, Martínez I, Barbero J, Ortega S, Suja J: CDK2 is required for proper homologous pairing, recombination and sex-body formation during male mouse meiosis. *Journal of cell science* 2009, 122:2149–59.

62. Lu L-Y, Wu J, Ye L, Gavrilina G, Saunders T, Yu X: RNF8-dependent histone modifications regulate nucleosome removal during spermatogenesis. *Developmental cell* 2010, 18:371–84.

63. Rai R, Zheng H, He H, Luo Y, Multani A, Carpenter P, Chang S: The function of classical and alternative non-homologous end-joining pathways in the fusion of dysfunctional telomeres. *The EMBO journal* 2010, 29:2598–610.

64. Peuscher M, Jacobs J: DNA-damage response and repair activities at uncapped telomeres depend on RNF8. *Nature cell biology* 2011, 13:1139–45.

65. Jacobs J: Fusing telomeres with RNF8. *Nucleus (Austin, Tex.)* 2012, 3:143–9.

66. Krausz C: Male infertility: pathogenesis and clinical diagnosis. *Best practice & research. Clinical endocrinology & metabolism* 2011, 25:271–85.

67. Boder: Ataxia-telangiectasia: some historic, clinical and pathologic observations. *Birth defects original article series* 1974, 11:255–70.

68. Barlow, Hirotsune, Paylor, Liyanage, Eckhaus, Collins, Shiloh, Crawley, Ried, Tagle, et al.: Atm-deficient mice: a paradigm of ataxia telangiectasia. *Cell* 1996, 86:159–71.

69. Barlow, Liyanage, Moens, Tarsounas, Nagashima, Brown, Rottinghaus, Jackson, Tagle, Ried, et al.: Atm deficiency results in severe meiotic disruption as early as leptonema of prophase I. *Development (Cambridge, England)* 1998, 125:4007–17.

70. Edelmann, Cohen, Kane, Lau, Morrow, Bennett, Umar, Kunkel, Cattoretti, Chaganti, et al.: Meiotic pachytene arrest in MLH1-deficient mice. *Cell* 1996, 85:1125–34.

71. Holloway, Booth J, Edelmann W, McGowan C, Cohen P: MUS81 generates a subset of MLH1-MLH3-independent crossovers in mammalian meiosis. *PLoS genetics* 2008, 4:e1000186.

72. Holloway, Mohan S, Balmus G, Sun X, Modzelewski A, Borst P, Freire R, Weiss R, Cohen P: Mammalian BTBD12 (SLX4) protects against genomic instability during mammalian spermatogenesis. *PLoS genetics* 2011, 7:e1002094.

73. Kolas N, Svetlanov A, Lenzi M, Macaluso F, Lipkin S, Liskay, Greally J, Edelmann W, Cohen P: Localization of MMR proteins on meiotic chromosomes in mice indicates distinct functions during prophase I. *The Journal of cell biology* 2005, 171:447–58.

74. Lipkin S, Moens P, Wang V, Lenzi M, Shanmugarajah D, Gilgeous A, Thomas J, Cheng J, Touchman J, Green E, et al.: Meiotic arrest and aneuploidy in MLH3-deficient mice. *Nature genetics* 2002, 31:385–90.

75. Bellvé AR: Purification, culture, and fractionation of spermatogenic cells. *Methods in enzymology* 1993, 225:84–113.

76. Li H, Durbin R: Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)* 2009, 25:1754–60.

77. Anders S, Pyl P, Huber W: HTSeq-a Python framework to work with high-throughput sequencing data. *Bioinformatics (Oxford, England)* 2014, doi:10.1093/bioinformatics/btu638.

78. Robinson M, McCarthy D, Smyth G: edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics (Oxford, England)* 2010, 26:139–40.

79. Langmead B, Trapnell C, Pop M, Salzberg S: Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 2009, 10:R25.

80. Griffiths-Jones S, Grocock R, van Dongen S, Bateman A, Enright A: miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic acids research* 2006, 34:D140–4.

81. Kozomara A, Griffiths-Jones S: miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic acids research* 2014, 42:D68–73.

82. Lim L, Lau N, Garrett-Engele P, Grimson A, Schelter J, Castle J, Bartel D, Linsley P, Johnson J: Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* 2005, 433:769–73.

CHAPTER 3

Fluorescence Activated Cell Sorting of Spermatogonial Cells

Introduction

The focus of this thesis is to examine the role that small non-coding RNAs and their binding partners, AGO proteins, play during male meiosis. Meiosis is a specialized cell division process used to create the viable haploid gametes needed in sexual reproduction [1]. One of the key features of meiosis is the genetic exchange of the maternal and paternal chromosomes, which occurs during meiosis I. More specifically, the exchange of genetic material is mediated through protein interactions, called synapsis, and physical tethering of the DNA, known as recombination, which occur during the first stage of meiosis I, called prophase I [2]. In males, the X and Y chromosome mostly lack homology, and are unable to fully synapse. This presents a special situation for the cell in which the asynapsed X and Y chromosomes have to bypass the mandatory synapsis checkpoint. This bypass is accomplished through a process called Meiotic Sex Chromosome Inactivation (MSCI), in which the sex chromosomes are silenced in the heterogametic sex [3]. While some of the proteins involved in this silencing process have been discovered [4, 5], the exact mechanism mediating silencing is still not known. Recent work has suggested that AGO proteins and sncRNAs may be involved in MSCI [6, 7].

Cell sorting is necessary for the examination of meiotic processes during mouse spermatogenesis

The purpose of the research being presented is to better understand the role sncRNAs and AGO proteins are playing during meiotic prophase I and MSCI. To this end, mouse lines

containing epitope tagged *Ago3* and *Ago4* will be used to co-immunoprecipitate proteins that can then be analyzed by mass spectrometry. Additionally, using these epitope tagged *Ago3* and *Ago4* mouse lines, sncRNAs bound to the AGOs can be analyzed by small RNA sequencing. To examine the roles that the individual AGO proteins are playing in the control of MSCI, it is necessary to examine individual substages of prophase I spermatocytes.

To study the biological mechanisms controlling MSCI in prophase I spermatocytes, these cells have to be isolated in pure form, and concentrated. The meiotic prophase I spermatocytes differentiate within the seminiferous tubule of the testis where there is a range of cell types present including non-spermatogenic populations, including Leydig, Sertoli, and epithelial cells, as well as spermatogenic cells, ranging from spermatogonia to spermatozoa [8]. Fortunately, the spermatogenic cell populations differ enough in size that they can be separated, making a cell sorting method from whole testis tissue possible [9]. Examination of spermatocytes in the earlier and later substages of prophase I, prior to and following pachytene, will provide insight into the mechanism controlling MSCI and the role AGO proteins are playing in this process.

STA-PUT: a spermatogonial cell sorting method

Until recently, the accepted method for the separation of spermatogonial cell populations from the largely heterogeneous testes was STA-PUT [10]. This technique allows separation of isolated cells from the mouse testes based on size and density. STA-PUT is a cell separation technique that uses velocity sedimentation through a Bovine Serum Albumin (BSA) concentration gradient [10].

To isolate spermatocytes from mice, testes are dissected from an adult mouse and de-tunicated. Then cells are isolated from the testes through tissue dissociation with collagenase and

digestion with trypsin. DNase I digestion is also performed to decrease cell clumping by removing DNA released from dead cells. Treatment with those reagents leaves a suspension of both somatic and spermatogenic cells that are then separated by cell sedimentation by size and density. A detailed STA-PUT protocol is described in Bryant et al., 2013 [9].

While, STA-PUT has been a standard technique for spermatogonial cell sorting, there are several challenges with this method. The STA-PUT procedure requires specialized glassware, along with a special apparatus to hold the glassware, and it must be performed in a cold room. Another disadvantage of the STA-PUT protocol is that it requires extensive hands-on collection of the cell fractions, leading to a very time intensive procedure, typically taking a full day. Due to the limitations of special glassware and time, if more then one sample is needed for analysis, wildtype and mutant for example, these would have to be processed on different days. Performing cell isolation of experimental controls on different days can lead to inconsistencies between samples. There is also a relatively high amount of cell population variability in fractions between sort days using the STA-PUT procedure, which makes it difficult to compare different experimental conditions that are collected on different days. The STA-PUT protocol also requires a large number of mouse testes in order to successfully sediment and separate cell types. Due to the shortcomings of the STA-PUT technique, other methods for separating spermatogonial cells have been tried including flow cytometry.

Flow Cytometry is a robust tool to analyze cell populations

Flow Cytometry is a laboratory technique that has been substantially growing in use through the past decade. The ease of flow cytometry methods, range of uses, and improvements made to the instruments, in both cost and ease of use, are all reasons for this growth in use [11].

The process of flow cytometry is to simultaneously measure various elements of an individual particle as is moves through a liquid stream. The size and internal complexity of the particle can be determined through light scattering from different angles. Some of the characteristics that can be measured through these means include RNA/DNA content, cell size, cytoplasmic complexity, and membrane or intracellular proteins [11]. One of the many uses of flow cytometry is the sorting of single cells. Within the flow cytometer, suspended cells are drawn up in an isotonic sheath fluid that allows for laminar flow [12]. Cells then move individually past an interrogation point at which a beam of light, typically a laser, intersects the cell [13]. The light that is emitted from the cells is collected and isolated to a single wavelength signal. This signal is then computed and analyzed, which allows for the distinction of cell populations [13]. Flow cytometry is a fast and powerful tool in the analysis of complex cell populations.

A recent flow cytometry method for spermatogonial cell sorting is fluorescence activated cell sorting (FACS). FACS provides a more refined method of cell sorting of cells in the mouse testes utilizing a flow cytometer to visualize, purify, and isolate meiotic cells from the heterogeneous testes population by DNA staining. The DNA is stained with Hoechst 33342 and cell size and cell density are examined [14]. FACS provides several advantages over STA-PUT. The FACS procedure can be scaled down for collection of a smaller number of cells from distinct populations, which is not possible with STA-PUT. Also, multiple experimental samples can be processed and collected on the same day. Finally, there is very little variability between the fluorescence data readout, or sorting profile, across days, which improves reproducibility of the technique for replicates.

While flow cytometry methods for purification of mouse meiotic cells have previously been described in the literature [14-16], the protocols described have several shortcomings. The

protocols described by Bastos et al., 2005 and Cole et al., 2014 did not separate the spermatogonial cells into the specific prophase I substages, which would be necessary for examination of the role that AGO proteins are playing in the MSCI process [14, 15]. Another protocol described by Getun et al., 2011 did claim separation of spermatocytes into prophase I substages, but was found to not be reproducible by the Cohen or Grimson labs due to cell death when using this protocol. Flow cytometry provides a more reproducible method of cell sorting, and isolated, staged prophase I spermatocytes are needed for analysis of the mechanism controlling MSCI. Thus, this work describes an improved FACS protocol for separation of mouse spermatocytes.

Improvements made to published FACS protocols

I developed an improved FACS sorting protocol to provide a means of separating and concentrating prophase I substage spermatocytes. Using this procedure, cells also have a higher viability both before and after FACS, compared to procedures previously described in the literature. The procedure provides an improved means for separating spermatogonial cells than previously described FACS or STA-PUT protocols.

The FACS procedure (Figure 1) begins with isolation of testes from an adult male mouse. The testes are de-tunicated and incubated in 6 ml of 1X GBSS (Gey's Balanced Salt Solution) with 0.4 mg/ml collagenase for fifteen minutes at 33°C in a shaking waterbath at 150 rpm to dissociate the tubules. The tubules are then washed twice with 1X GBSS. The tubules are then treated with 5ml 1X GBSS, 50 ul of 50 mg/ml trypsin in 1mM HCl and 5 ul of 1mg/ml DNase I to isolate the cells from the tubules. The tubules are then agitated for 15 minutes at at 33°C in a shaking waterbath at 150 rpm. Following incubation, the tubules are manually dissociated by

pipetting up and down for 3 minutes with a cut Pasteur pipette. The trypsin activity is then deactivated by the addition of 400 ul of fetal bovine serum (FBS). The remaining chunks of tubules are then separated by straining the cells through a 70uM cell strainer. Following cell straining, the cells are pelleted at 300 x g for 10 min, and resuspended in 5 ml 1X GBSS + 250 ul fetal bovine serum (FBS) + 5 ul DNase I (1 mg/ml) + 2 ul Hoechst 33342 (5 ug/ml final). Cells are then agitated at 150 rpm for 45 min at 33°C. The final step is the addition of 1 ul propidium iodide (1mg/mL), which is used for detection of dead cells and an additional step to remain any cell clumps by straining the cell suspension through a 40uM cell strainer. The stained cells are then loaded into the flow cytometer and examined based on the Hoechst blue emission and Hoechst red emission forward and side scatter using a UV laser and 488 nm blue laser [16]. Following a quick run of the cells in the flow cytometer, the diagram displaying the cell population is displayed and regions of the curve are selected for collection, also known as cell gating. These cell gates are collected using a 100 μm nozzle. Following collection of the gated cells, the cells are centrifuged at 300 x g for 10 minutes at room temperature to pellet and then resuspended in 20 ul GBSS. To examine cell recovery, 1 ul of cells are used to count on hemocytometer (1 ul cells + 8 ul 1X GBSS + 1 ul 0.4% trypan blue. The cell purity is also examined by performing chromosome spreads: 2.5 ul of cells is added 2.5 ul of 50 mM sucrose solution and incubated 20 minutes at room temperature. Then, 5 ul of the sucrose and cell mixture is transferred to a 30 ul bubble of 1% Paraformaldehyde (PFA) on a slide with wells. The cells are fixed overnight in a humidification chamber, at room temperature, before they undergo immunofluorescence staining. The cells are then fluorescently stained with primary antibodies raised against the synaptonemal complex protein, SYCP3, and the phosphorylated

form of the histone H2AX, γH2AX, as seen in Figure 2. The stained cells are then scored based on synaptonemal complex formation and localization of γH2AX.

This FACS protocol has resulted in a spermatocyte isolation procedure enhanced in efficiency, purity and reproducibility. The reproducibility of this FACS procedure has been significantly improved with the utilization of a purchased, pre-made aqueous Hoechst 33342 dye. The use of this dye has led to less variability, relative to previous attempts at this procedure, in the cell curve displayed by the flow cytometer between sorting procedures on different days, which has improved the reproducibility of cell gating.

Using this FACS procedure, prophase I staged spermatocytes can be successfully collected, as shown in the representative FACS sorting curve, Figure 2. The cell purity of the distinct gated region is still in the improvement stage, however. For each of the FACS-sorted cell populations seen in Figure 2, the cells were collected and examined by light microscopy to examine recovery and homogeneity of the cells. The percent purity was calculated from the individual cell populations on the curve and the resulting percent cell purities were shown in Figure 3. As can be seen in Figure 3, the major contaminating factor currently is cells of varying prophase I substages and not somatic cells. I find that leptotene and zygotene, as well as, pre-leptotene and leptotene staged cells are often collected within the same gated region. Reduction in contamination of prophase I substages can likely be improved be with more testing of the FACS procedure and cell gating.

To compare STA-PUT to FACS cell purities, I scored cells derived from each of the two cell separation techniques that were fluorescently stained with primary antibodies raised against SYCP3 and γH2AX, and DAPI. I compared cells from the determined pachytene region of the FACS curve (p1 in Figure 2) to pachytene fractions from a STA-PUT collection. Any cell that

was DAPI stained without the presence of SYCP3 and γH2AX was scored as a non-prophase I staged cell; this included spermatozoa, spermatids, and somatic cells. When comparing STA-PUT pachytene cell purity to FACS pachytene cell purity, there was 43% non-prophase I staged cell contamination in the STA-PUT samples, while the contamination of non-prophase I staged cells in the FACS pachytene samples was 3%. To examine overall pachytene cell purity of the STA-PUT to FACS, the purity of pachytene cells out of all cells counted, including other prophase I staged cells, was determined. Using these criteria, the overall pachytene staged cell purity using STA-PUT was 28%, while using FACS it was 76%.

One of the advancements with FACS as compared to STA-PUT is the ability to scale down and collect cells in a smaller time frame. Within a two-hour sort, we can collect approximately 20,000-90,000 cells for each of the binned cell populations that are intended for collection of a single prophase I staged cell population. However, the cell recovery following cell sorting is variable. The cell recovery is calculated by estimating the total number of viable cells from each gated region, by counting on a hemocytometer a trypan blue stained cell sample, which stains dead cells. The number of live cells are then divided by the total number of cells (events) collected by the flow cytometer. I find that when more than 50,000 events are collected, the average recovery is 58%, but when fewer than 50,000 events are collected, the average recovery is only 22%. Currently, it seems that cell recovery is reliant on the number of events collected on the flow cytometer and lower numbers of cell events leads to a lower cell recovery following cell pelleting. The reason for this difference in cell recovery is unknown and is an element of the protocol currently being optimized.

These results display the improvements made to the FACS method and are encouraging. There are still several aspects of the protocol still being optimized, including increasing cell

purity of specific prophase I cell populations and improving recovery of cells following sorting. Another possible improvement to the protocol would be the use of antibodies for further cell purification. However, there are not currently any cell surface proteins that are known that are possible for distinguishing cells between the prophase I substages. This FACS technique provides a means of collecting viable, concentrated meiotic prophase I substaged spermatocytes.

Future Directions

The ability to use FACS for isolation and recovery of concentrated prophase I spermatocytes, will be a useful technique for downstream applications like RNA-Seq, immunoprecipitations, Western Blots, Pro-Seq, and other experiments [6, 17, 18]. One of the goals of my project was to examine the role sncRNAs and their associated AGO proteins are playing in the MSCI mechanism. For this work, it is necessary to examine the role AGOs are playing before, after, and during MSCI. Examination of the AGO proteins role at those distinct points requires isolation of spermatocytes at the distinct substages of prophase I. To examine the roles that distinct AGO proteins are playing, mouse lines containing epitope tagged *Ago3* and *Ago4* are currently being generated in the Cohen Lab. The goal is to use these mouse lines in the FACS procedure, isolate pure prophase I staged cells, and then examine the proteins and sncRNAs interacting with AGO3 and AGO4 at leptotene, zygotene, pachytene, and diplotene.

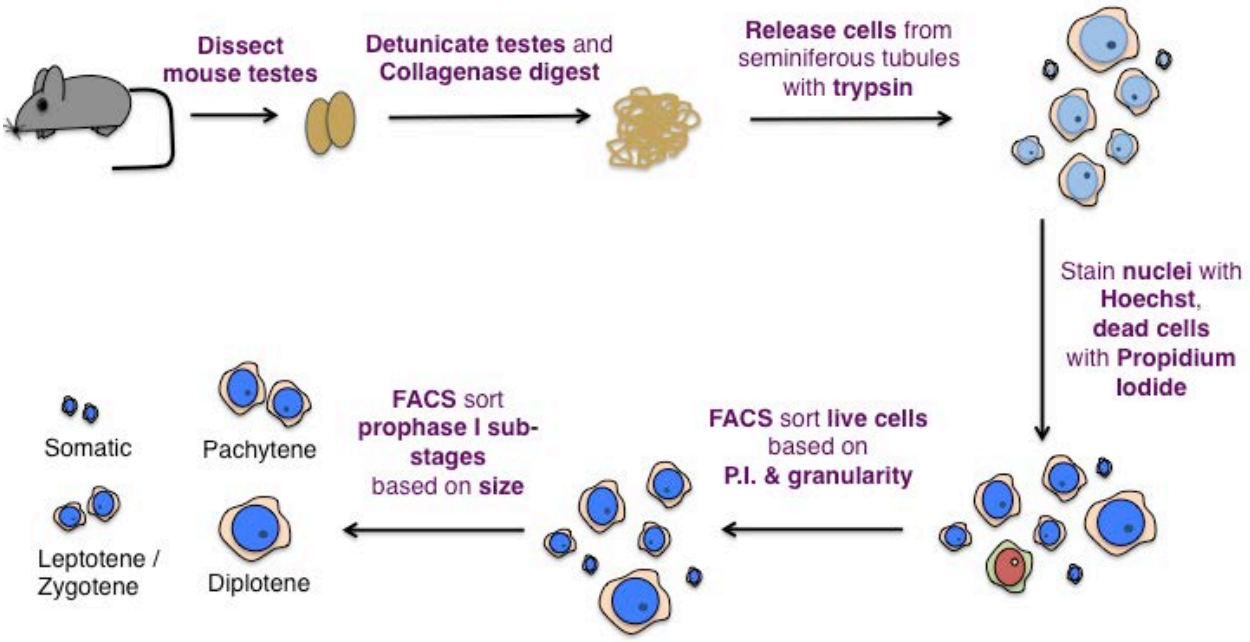


Figure 1. FACS Procedure Flow Chart. This flow chart outlines the FACS procedure in its improved form, beginning with the dissection of mouse testes to the sorting of the prophase I substages. (Adapted from figure by Caterina Schweidenback)

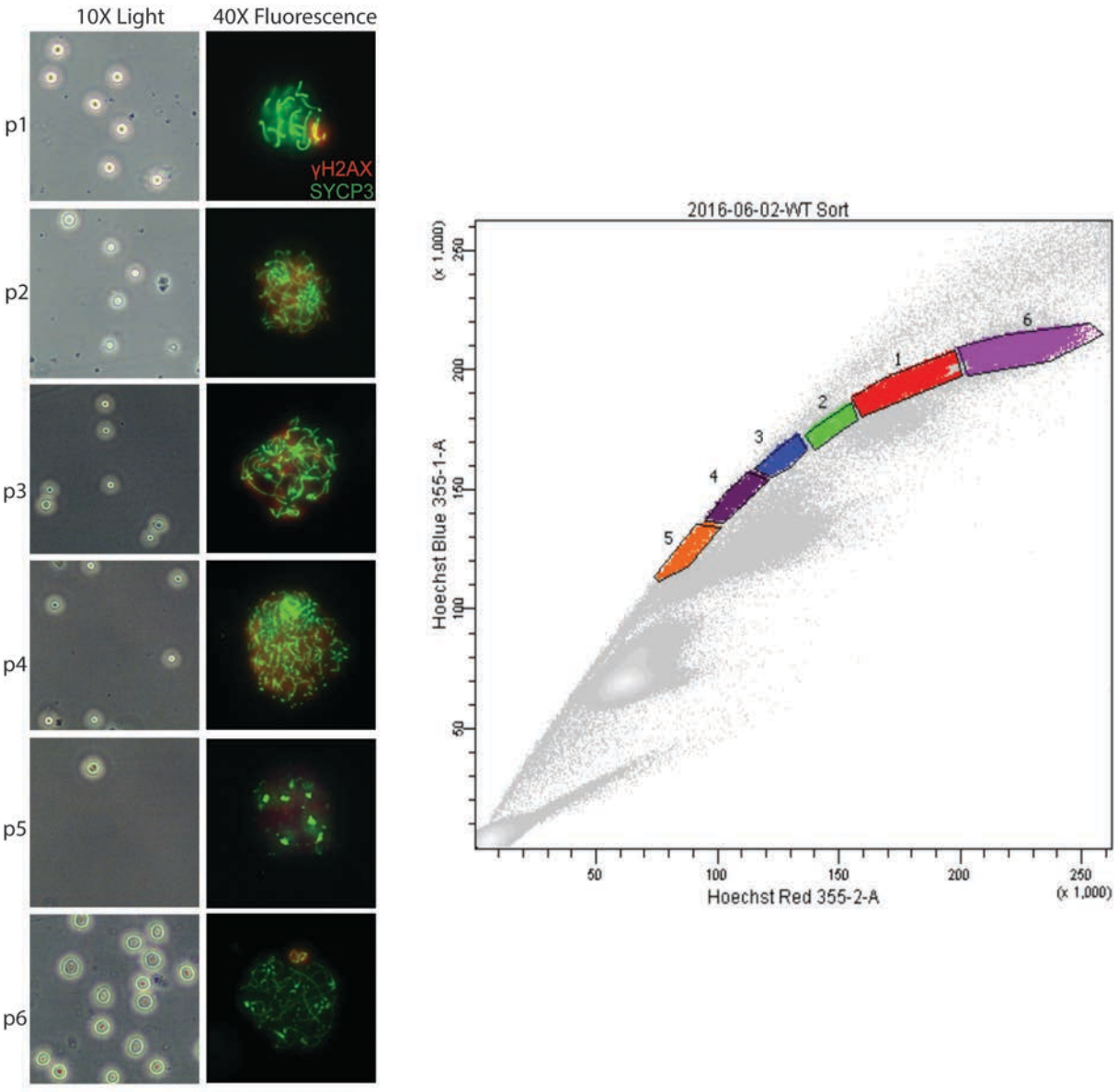


Figure 2. FACS Curve Gating and Cell Morphology. This figure displays a representative FACS sorting profile. Spermatocytes from C57 WT mice were stained with Hoechst and sorting according the gating shown (population 1-6). The populations were then evaluated by staining with anti-SYCP3 (green) and anti- γH2AX (red) antibodies, as seen in the 40X Fluorescence column. Both the 10X Light and 40X Fluorescence cell images, on the left, correspond to the colored gated cell populations (p1-p6) shown on the flow cytometry curve, on the right. The gray parts of the FACS curve correspond to cell populations determined to not be prophase I meiotic cells, including somatic cells, and later staged meiotic cells.

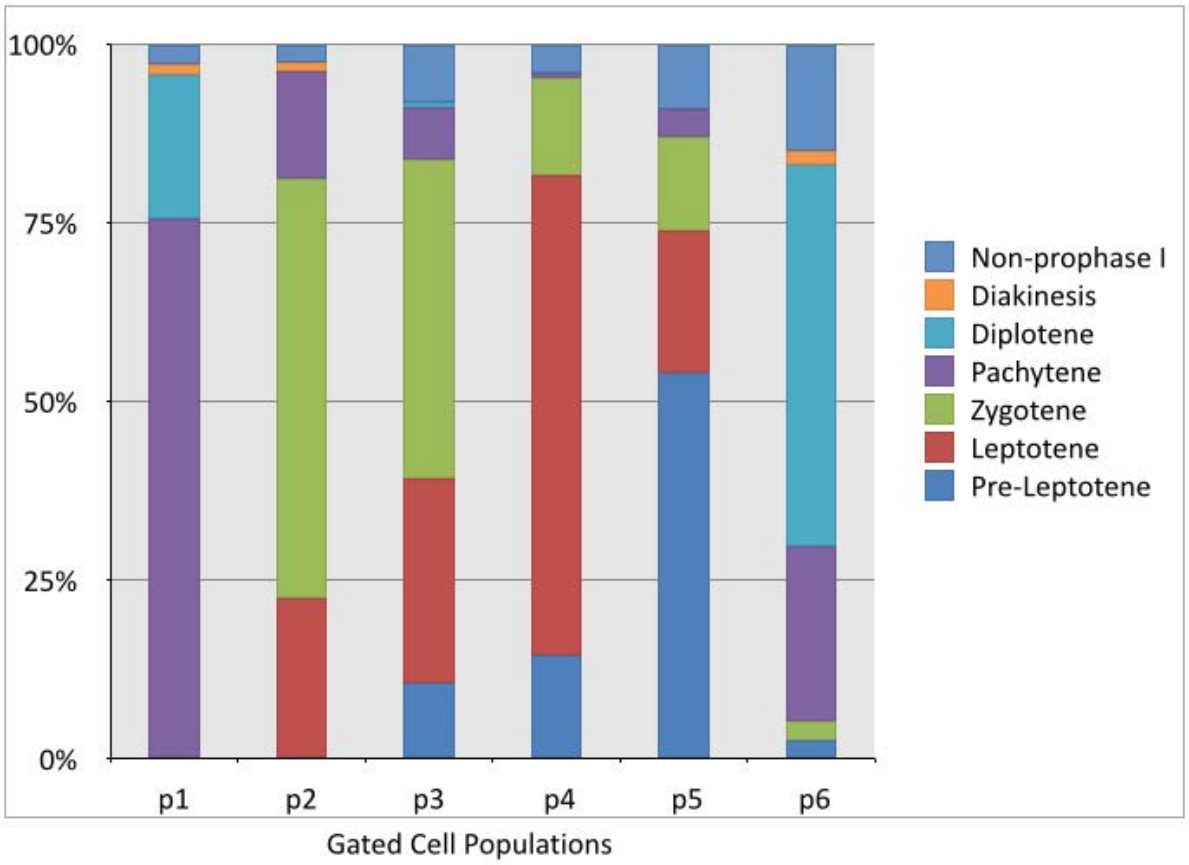


Figure 3. FACS Curve Gating and Cell Purity. Spermatocytes from C57 WT mice stained with anti-SYCP3 and anti- γH2AX antibodies were analyzed and scored according to prophase I stage. The graph shows the determined prophase I purity of each gated cell population from a representative FACS sort done on 06-02-16. The colors denote the different cell types (pre-leptotene, leptotene, zygotene, pachytene, diplotene, diakinesis, and non-prophase I cells) found in each of the populations (p1-6) collected from the FACS curve seen in Figure 2.

References

1. Handel, M.A. and J.C. Schimenti, *Genetics of mammalian meiosis: regulation, dynamics and impact on fertility.* Nat Rev Genet, 2010. **11**(2): p. 124-36.
2. Plant, T.M. and A.J. Zeleznik, *Knobil and Neill's Physiology of Reproduction Two-Volume Set*. 2014, Academic Press Imprint Elsevier Science & Technology Books: San Diego. p. 1 online resource.
3. Turner, J.M., *Meiotic sex chromosome inactivation.* Development, 2007. **134**(10): p. 1823-31.
4. Royo, H., et al., *ATR acts stage specifically to regulate multiple aspects of mammalian meiotic silencing.* Genes Dev, 2013. **27**(13): p. 1484-94.
5. Turner, J.M., et al., *BRCA1, histone H2AX phosphorylation, and male meiotic sex chromosome inactivation.* Curr Biol, 2004. **14**(23): p. 2135-42.
6. Modzelewski, A.J., et al., *AGO4 regulates entry into meiosis and influences silencing of sex chromosomes in the male mouse germline.* Dev Cell, 2012. **23**(2): p. 251-64.
7. Modzelewski, A.J., et al., *Dgcr8 and Dicer are essential for sex chromosome integrity during meiosis in males.* J Cell Sci, 2015. **128**(12): p. 2314-27.
8. Handel, M.A., *Meiosis and Gameteogenesis*. Academic Press. Vol. 37. 1998.
9. Bryant, J.M., et al., *Separation of spermatogenic cell types using STA-PUT velocity sedimentation.* J Vis Exp, 2013(80).
10. Bellve, A.R., et al., *Spermatogenic cells of the prepuberal mouse. Isolation and morphological characterization.* J Cell Biol, 1977. **74**(1): p. 68-85.
11. Brown, M. and C. Wittwer, *Flow cytometry: principles and clinical applications in hematology.* Clin Chem, 2000. **46**(8 Pt 2): p. 1221-9.
12. McCoy, J.P., Jr. and D.F. Keren, *Current practices in clinical flow cytometry. A practice survey by the American Society of Clinical Pathologists.* Am J Clin Pathol, 1999. **111**(2): p. 161-8.
13. Lovett, E.J., 3rd, et al., *Application of flow cytometry to diagnostic pathology.* Lab Invest, 1984. **50**(2): p. 115-40.
14. Cole, F., et al., *Mouse tetrad analysis provides insights into recombination mechanisms and hotspot evolutionary dynamics.* Nat Genet, 2014. **46**(10): p. 1072-80.
15. Bastos, H., et al., *Flow cytometric characterization of viable meiotic and postmeiotic cells by Hoechst 33342 in mouse spermatogenesis.* Cytometry A, 2005. **65**(1): p. 40-9.
16. Getun, I.V., B. Torres, and P.R. Bois, *Flow cytometry purification of mouse meiotic cells.* J Vis Exp, 2011(50).
17. Kwak, H., et al., *Precise maps of RNA polymerase reveal how promoters direct initiation and pausing.* Science, 2013. **339**(6122): p. 950-3.
18. Florez-Zapata, N.M., et al., *Transcriptomic landscape of prophase I sunflower male meiocytes.* Front Plant Sci, 2014. **5**: p. 277.

CHAPTER 4

Generation of epitope tagged *Ago3* and *Ago4* mouse lines

Introduction to epitope tagged Ago3 and Ago4 mouse lines

The focus of this thesis work is to analyze the role that small non-coding RNAs and their binding partners play during male meiosis. As there is little known about what role AGO proteins are playing during mammalian male meiosis. It is important to examine the role that they play during Prophase I through the examination of AGO protein binding partners in isolated prophase I staged spermatocytes.

Previous work in the Cohen lab has shown a role for AGO4 in mammalian meiosis in males. *Ago4*$^{-/-}$ mice show a meiotic phenotype of a disrupted sex body, an influx of RNA Polymerase II to the sex body, and subfertility [1]. Examination of mRNA levels of the other mammalian AGO proteins in the *Ago4*$^{-/-}$ mice showed increased AGO3 levels. Increased AGO3 levels in the absence of AGO4, as well as subfertility, suggested there might be a partially redundant role for AGO3 and AGO4 proteins within the nucleus during male meiosis.

To understand what roles AGO3 and AGO4 are playing during male meiosis and MSCI, the proteins and RNAs binding to the AGO proteins will need to be examined during subsequent stages of meiotic prophase I. To this end, transgenic mice containing epitope tagged AGO3 and AGO4 are being generated using CRISPR/CAS9 genome editing. Each mouse line will be tagged with two epitopes on the N-terminus of the *Ago3* and *Ago4*. The dual tagging method is intended to alleviate lack of specificity of AGO antibodies and provide the ability to tandem purify immunoprecipitates with the second epitope tag [1]. Previous research has shown that N-

terminally tagged AGO proteins are functionally active [2]. The epitope tagged AGO3 line will contain Myc and FLAG tags and the epitope tagged AGO4 line will contain FLAG and HA tags.

Generation of epitope tagged Ago3 and Ago4 mouse lines

The *Ago3-Myc/Flag* and *Ago4-Flag/HA tagged* mice will be generated by CRISPR/CAS9 driven homologous recombination. CRISPR/CAS9-mediated homologous recombination is a tool, co-opted from bacteria, for use in a range of organisms, including mice [3-5]. The CRISPR/CAS9 technique allows for changes to be made in the genomes of living cells. In mice, this technique can be used for developing mouse lines with permanent gene mutations that can be used for downstream experiments. In this case, two AGO proteins' genes are being epitope tagged. To epitope tag a desired gene in the genome with CRISPR/CAS9, three crucial components are required: a single guide RNA (sgRNA), a CAS9 nuclease, and single-stranded oligodeoxynucleotide (ssODN). The targeting process involves CAS9 enzyme guided by the sgRNA to a site in the genome where it induces a double strand break. The ssODN that was designed to include homology with the genome on both ends and the sequence of the desired epitope tags is then used as a homology repair template.

There are several aspects of the CRISPR/ CAS9 technique that have to be considered in the strategy design. CAS9 can be delivered as a protein or mRNA to the embryo; however, injection of mRNA is currently thought to increase efficiency of editing due to the fact that it can be injected directly into the cytoplasm, while the protein needs to be injected directly into the pronucleus, which can be more damaging to the embryos [4]. Also, in choosing the target locus, a protospacer adjacent motif (PAM) has to be present near the intended site of ssODN

incorporation. The PAM is an NGG sequence that directs CAS9 cleavage and is 3-4 nucleotides upstream from the sgRNA recognition site [6].

Generation of the *Ago3-Myc/Flag tagged* line began with designing the sgRNA. To prevent retargeting of the edited allele by CAS9, the sgRNA and its associated PAM site were specifically chosen due to proximity to the *Ago3* start site, with the sequence CCGATTTCCATTCATGGAGC. The sgRNA was predicted and chosen based on the guide quality score using the Zhang Lab MIT CRISPR Design tool [7]. Since the Myc/FLAG epitope tags were being incorporated directly following the ATG start site, after homologous recombination of the ssODN, the sgRNA could no longer target the region. This is due to the fact that following the incorporation of the tag, the sgRNA no longer was homologous with that site in the genome. The sgRNA was then generated using a multi-step process. First, using a primer containing a T7 promoter + sgRNA sequence (without the PAM site) + 20 nt overhang sequence, and a primer with homology to the overhang region sequence, a PCR product containing the double-stranded T7 and sgRNA was produced. This PCR was performed with 16 PCR reactions with 100 uls in each using Phusion High-Fidelity DNA Polymerase. The PCR products were then purified using a single affinity based Qiagen PCR cleanup column. All of the PCR product was put over one column, to concentrate the product and get the highest yield of PCR product following elution. The purified PCR product was then *in vitro* transcribed (IVT) in two reactions using the Ambion MEGAshortscript kit. Purification was performed as described in the kit protocol using the upper limit of 8 ul of template DNA for the reaction and allowing the reaction to proceed overnight. The IVT reactions were DNase treated the following morning. Then, both IVT reactions were put over an Ambion MegaClear column following the included kit protocol and eluted with elution buffer. The resulting sgRNA was of high purity with a concentration of

4150.9 ng/ul. The ssODN containing Myc and FLAG epitope tags flanked by genomic sequence from either side of the ATG start site was synthesized by Integrated DNA Technologies (IDT). The sgRNA and the ssODN were then injected along with CAS9 mRNA into mouse embryos by the Cornell Stem Cell & Transgenic Core Facility. This CRISPR/Cas9 strategy resulted in the birth of twenty pups, seven of which show potential editing and are in being further examined.

Generation of the *Ago4-FLAG/HA tagged* mice proved more challenging. The translation start site, as well as a span of around 400 bp in either direction of the start site, has a very high GC content of around 75%. This led to challenges in synthesizing the sgRNA and ssODN. The sgRNA was successfully synthesized using the improved methods described above for the generation of the *Ago3-Myc/FLAG tagged* sgRNA. Typically, for CRISPR/CAS9 injections performed by the Cornell Stem Cell & Transgenic Core Facility, ssODNs are synthesized by IDT, but this was not possible for the *Ago4-FLAG/HA tagged* ssODN due to the GC richness. To combat this challenge, several possible strategies are being employed simultaneously. The first and simplest strategy is to make the ssODN by amplifying it from several smaller oligos that overlap at their 3' ends. This strategy successfully yielded dsDNA corresponding to the desired repair template; however, it was not possible to separate the two strands of the oligonucleotide. Therefore, the dsDNA was heat-denatured prior to being injected into mouse embryos. I am still awaiting the results of these injections. If this proves to be unsuccessful, another strategy being considered is the generation of a longer oligonucleotide. This would be an oligonucleotide spanning well beyond the GC rich region, using homology arms flanking the region 1-4 kb in length [8]. The resulting homology template would be heat denatured to create more single stranded DNA prior to injection. This could potentially negate the difficulties with GC richness at the *Ago4* start site, but this method is less commonly used and may result in reduced mutation

efficiency as compared with the use of ssODN [9]. A final strategy being considered is to perform recombineering using a bacterial artificial chromosome (BAC) containing AGO4 to make a transgenic mouse containing the FLAG and HA epitope tags at the N terminus of AGO4. This method is more time consuming than CRISPR/CAS9, but has been utilized in the past for gene tagging. The ultimate goal is to successfully use one of the described methods above for generation of the *Ago4-Flag/HA tagged* mouse line.

References

1. Modzelewski, A.J., et al., *AGO4 regulates entry into meiosis and influences silencing of sex chromosomes in the male mouse germline.* Dev Cell, 2012. **23**(2): p. 251-64.
2. Meister, G., et al., *Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs.* Mol Cell, 2004. **15**(2): p. 185-97.
3. Jinek, M., et al., *A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity.* Science, 2012. **337**(6096): p. 816-21.
4. Singh, P., J.C. Schimenti, and E. Bolcun-Filas, *A mouse geneticist's practical guide to CRISPR applications.* Genetics, 2015. **199**(1): p. 1-15.
5. Lackner, D.H., et al., *A generic strategy for CRISPR-Cas9-mediated gene tagging.* Nat Commun, 2015. **6**: p. 10237.
6. Ran, F.A., et al., *Genome engineering using the CRISPR-Cas9 system.* Nat Protoc, 2013. **8**(11): p. 2281-308.
7. Hu, S., et al., *Genome engineering of Agrobacterium tumefaciens using the lambda Red recombination system.* Appl Microbiol Biotechnol, 2014. **98**(5): p. 2165-72.
8. Zhou, D., et al., *Rapid tagging of endogenous mouse genes by recombineering and ES cell complementation of tetraploid blastocysts.* Nucleic Acids Res, 2004. **32**(16): p. e128.
9. Chu, V.T., et al., *Efficient generation of Rosa26 knock-in mice using CRISPR/Cas9 in C57BL/6 zygotes.* BMC Biotechnol, 2016. **16**: p. 4.

CHAPTER 5

Conclusions and Future Directions

My graduate work has examined a role for small non-coding RNAs and proteins associated with small non-coding RNA biogenesis, during male meiosis. In this thesis, I have described how a loss of miRNAs and siRNAs leads to abnormal chromosomal fusions of the sex chromosomes during male meiosis and affects the proteins involved in MSCI signaling. I have also discussed improvements I have made to a spermatocyte isolation procedure using flow cytometry, which will be critical for the examination of specific substages of meiotic prophase I spermatocytes. Finally, I define a strategy for generation of a mouse line containing epitope tagged *Ago3* and *Ago4*. An *Ago3-Myc/FLAG* tagged mouse line has already been generated and is currently in the process of being validated. The work I have completed has paved the way for the examination of the roles of AGO3 and AGO4 in the spermatocyte nucleus during male meiosis.

The future directions of this project will be the validation of AGO3 and AGO4 epitope tagged mouse lines for the incorporation of the tag, examination of any potential off-target effects, and confirmation that the AGO proteins remain functional. Following validation, the ultimate goal will be to examine the proteins and RNAs interacting with AGO3 and AGO4 in isolated prophase I sub-staged spermatocytes. This will aid in understanding the role that AGO proteins are playing in MSCI and male meiosis in general.

Validation of Ago3-Myc/FLAG and Ago4-FLAG/HA tagged mouse lines

The *Ago3-Myc/FLAG* and *Ago4-FLAG/HA* tagged mice will need to be analyzed to: i) validate incorporation of the epitope tags, ii) eliminate mosaicism and any off target-effects as a result of CAS9 activity at loci other than the intended locus, and iii) test the functionality of epitope-tagged proteins.

To validate the incorporation of the epitope tags, the targeted region will need to be amplified and sequenced. To test the functionality of the *Ago3-Myc/FLAG* and *Ago4-FLAG/HA* tagged mouse lines, they will need to be assayed to determine that the *Myc/FLAG* and *FLAG/HA* epitope tags do not disrupt the expression, role, and localization of AGO3 and AGO4, respectively. Expression of the appropriately modified open reading frame will need to be confirmed by alignment of the endogenous *Ago3* and *Ago4* sequence to the cDNA sequencing results from the PCR product using the computational tool MegAlign. To confirm AGO3 and AGO4 protein expression in the *Ago3-Myc/FLAG* and *Ago4-FLAG/HA tagged* mouse lines, western blotting will need to be performed on whole testes lysate for AGO3 or AGO4, as well as the corresponding epitope tags. The negative control will be whole testes lysate from the *Ago3-1-4*$^{-/-}$ mouse line. The *Ago3-1-4*$^{-/-}$ mouse should not express AGO3, AGO4, or any of the epitope tags.

Founder male and female mice that have incorporated the tags, as determined by sequencing and cDNA analysis, will be crossed to the wildtype mouse strain C57bl/6J to eliminate any mosaicism and off-target effects in the founders. Since the sgRNA and CAS9 mRNA are injected into an actively replicating zygote, there is a possibility that genome editing can occur after DNA replication, resulting in some cells having the tag incorporated and others lacking incorporation. To address the potential problem of mosaicism, the founder mice will need to be backcrossed. Although a bioinformatics tool called COSMID will be employed to

help choose sgRNAs with few predicted off target sites, multiple backcrosses to C57bl/6J will eliminate any off-targets that do occur during the genome editing process [1]. Backcrossing of both the *Ago3-Myc/FLAG* and *Ago4-FLAG/HA* mouse lines to the C57bl/6J line will generate $Ago3^{Myc\text{-}FLAG/+}$ and $Ago4^{FLAG\text{-}HA/+}$ heterozygotes. These heterozygotes will then be bred together to generate $Ago3^{Myc\text{-}FLAG/Myc\text{-}FLAG}$ and $Ago4^{FLAG\text{-}HA/FLAG\text{-}HA}$ homozygote experimental mice.

$Ago3^{Myc\text{-}FLAG/Myc\text{-}FLAG}$ and $Ago4^{FLAG\text{-}HA/FLAG\text{-}HA}$ homozygotes will need to be validated for functional AGO3 or AGO4 protein, by assaying complementation of the *Ago4-FLAG/HA* line with the $Ago4^{-/-}$ line. If the incorporation of the tag results in a loss of function of *Ago4*, then crossing $Ago4^{-/-}$ and $Ago4^{Flag/HA/Flag/HA}$ mice would not rescue *Ago4* function, whereas if the epitope tag does not disrupt protein function, the null allele would be rescued. Since $Ago4^{-/-}$ males display decreased testes weights, reduced sperm counts, and altered sex body morphology, a rescue of this phenotype can be examined by quantifying these phenotypes [2]. Validation that CRISPR does not disrupt the protein function will first be done with the *Ago4-FLAG/HA* mouse line, since there is an $Ago4^{-/-}$ line currently available for complementation analysis. To confirm that there are no changes in the relative protein expression of *Ago3* or *Ago4* in wildtype versus the *Ago3-Myc/FLAG* or *Ago4-FLAG/HA* tagged lines, quantitative real-time PCR of AGO3 and AGO4 mRNA will be performed from spermatocyte cell samples.

Changes in the localization patterns of the tagged AGO proteins will need to be assayed by immunofluorescence on spermatocyte chromosome spreads utilizing antibodies specific to the epitopes. It is expected that the epitope-tagged AGO3 and AGO4 proteins will localize to the sex body as do the wildtype proteins [2]. Once the *Ago4-FLAG/HA* line is validated then it would be expected that the *Ago3-Myc/FLAG* tagged line would also exhibit a similar incorporation. An

Ago3$^{-/-}$ mouse line will be generated using CRISPR/CAS9 genome editing that can be used for a complementation assay with the *Ago3-Myc/Flag* tagged line.

Examination of proteins interacting with AGO3 and AGO4 within the spermatocyte nucleus.

AGO3 and AGO4 localize to the spermatocyte nuclei during meiotic prophase I, but there is currently no known mechanism for the localization of AGO proteins in the mammalian nucleus. To examine whether there is a novel nucleus-specific role for AGO proteins, the proteins that specifically interact with AGO3 and AGO4 in the spermatocyte nucleus will need to be isolated [3]. Specifically, AGO3 or AGO4 along with associated proteins will be co-immunoprecipitated from the nuclear fractions of spermatocytes using the *Ago3-Myc/FLAG* and *Ago4-FLAG/HA* tagged lines. Proteins that interact with AGO3 and AGO4 will be sent for mass spectrometry analysis.

The fluorescence-activated cell sorting (FACS) method, described in chapter 2, will be used to isolate spermatocytes from each of the sub stages of prophase I: leptonema, zygonema, pachynema, and diplonema. Following isolation of these cells, nuclear–cytoplasmic fractionation will be performed using the NE-PER Nuclear and Cytoplasmic Extraction Reagents kit (Thermo Fisher Scientific) according to the manufacturer's protocol. Prior to the co-immunoprecipitation assay, the specificity of the epitope antibody will be validated by first performing a Western Blot. As a positive control to verify that the antibodies function on the Western Blot, epitope tag peptide that corresponds to the epitope on the tagged mouse will be used. For example, a HA-peptide would be used with the *Ago4-FLAG/HA* tagged line. The negative control to test that the antibody doesn't recognize other proteins besides the epitope-tagged AGO, will be samples from

a wildtype littermate that does not have an incorporated epitope tag. Once the antibodies have been validated, the co-immunoprecipitation assays can be performed.

Since the *Ago3-Myc/FLAG* and *Ago4-FLAG/HA* tagged mouse lines each will have two incorporated epitope tags, tandem co-immunoprecipitation can be performed on the sample to increase the purity. Alternatively, if antibodies against one epitope tag fail to pull down sufficient protein material, antibodies against the second tag will be used instead. For the co-immunoprecipitation assay, immunoprecipitation will be performed on experimental male and a wildtype littermate prophase I staged cells. The immunoprecipitate from both nuclear and cytoplasmic fractions, from both the wildtype and experimental mice, will then be run on a sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) Gel. Bands that appear unique in the experimental sample and do not appear in the wildtype sample will then need to be sent for Mass Spectrometry analysis [4, 5].

For, the mass spectrometry assay, nuclear and cytoplasmic proteins extracted from the SDS-PAGE gel will be analyzed. Although the focus is on nuclear-specific proteins, the cytoplasmic proteins will be used as a means of comparison. To increase the accuracy of the proteins identified in this assay, tandem mass spectrometry analysis performed by the Cornell University Proteomics and Mass Spectrometry Facility will be used. Tandem mass spectrometry will increase the detection of specific compounds, by determining their characteristic fragmentation patterns through two stages of mass analysis [6].

In analyzing the mass spectrometry results, false positives will be eliminated through the use of a bioinformatics tool, Ingenuity Pathway Analysis [7]. This program will be used to identify proteins not known not to be expressed in germ cells; these proteins will then be eliminated from further analysis. To gauge the fidelity of the nuclear and cytoplasmic

fractionation, nuclear- and cytoplasmic-specific proteins will be assayed. This will be done by using The Nuclear Protein database, which is a searchable database of proteins described in the vertebrate nucleus, since it is currently not known what proteins AGOs might be interacting with in the nucleus [8]. Proteins of interest that have verified involvement in chromatin dynamics, and/or those proteins that have been shown to have roles similar to those seen in the RNA induced transcriptional silencing (RITS) processes will be prioritized. For example, chromodomain proteins that associate with histone variants that are active at the sex body and play roles in heterochromatin formation will be surveyed in the data.

Once specific proteins have been identified, using available commercial antibodies, their functional status will need to be validated by western blot of the AGO3 or AGO4 co-immunoprecipitate. Immunofluorescence on spermatocyte chromosome spreads will also be performed to validate nuclear localization of these proteins. If mouse mutants are available for the genes encoding these candidate proteins, they will be examined for prophase I progression defects. If mouse mutant lines are not available, mutants can be generated rapidly using CRISPR/Cas9 technology [9]. While the traditional ES cell mouse mutant generation often took close to a year to acquire experimental mice, CRISPR/Cas9 editing is a much more efficient process taking only a few weeks to get experimental mice. These experiments will aid in identifying a nuclear-specific AGO-containing complex.

Examination of small non-coding RNAs associating with AGO3 and AGO4 within the spermatocyte nucleus.

To better understand the function of AGO proteins in spermatocyte nuclei, the small RNAs that associate with AGO3 and AGO4 in the nucleus will be compared with those that associate with AGO3 and AGO4 in the cytoplasm at various stages of meiotic prophase. To this

end, the *Ago3-Myc/FLAG* and *Ago4-FLAG/HA* lines will be used in RNA immunoprecipitation and subsequent RNA sequencing experiments that will be carried out in collaboration with the Grimson lab. As described in chapter 3, a FACS method will be used to isolate spermatocytes from each of the sub stages of prophase I: leptonema, zygonema, pachynema, and diplonema. These cells will then undergo nuclear and cytoplasmic fractionation, as described above. Epitope-tagged AGO3 and AGO4 will be immunoprecipitated from the nuclear and cytoplasmic fractions as described above..

Co-precipitated RNA will be isolated from the nuclear and cytoplasmic precipitates of prophase I-staged cells using Trizol [10]. Following Trizol extraction, the RNA will be size-selected using a polyacrylamide gel and then will need to undergo small RNA sequencing performed by the RNA Sequencing Core at Cornell University. As a control for the RNA sequencing assay, small RNAs from the nuclear and cytoplasmic fractions of wildtype mouse spermatocytes at each substage of prophase I will be sequenced. These RNA libraries will serve as a baseline control for the small non-coding RNAs present in each substage. As a negative control, a *Dicer* conditional knockout mouse [11], which lacks active miRNAs and siRNAs, will be crossed to the epitope tagged lines. In this case, it would be expected that there should be no small non-coding RNAs that precipitate with AGO3 or AGO4 proteins.

Following RNA sequencing, the small non-coding RNA results will be filtered to select for siRNAs and miRNA by isolating those transcripts that are between 19-23 nucleotides. These siRNA and miRNA sequences will then be aligned to the mouse genome using the Bowtie short read alignment program. The siRNAs and miRNAs isolated from the spermatocytes nuclear fractions will then be compared to those isolated from the cytoplasmic fractions. In this examination, the Grimson lab will be looking for differentially-expressed small RNAs between

the nuclear and cytoplasmic fractions of each stage, while using the cellular fractions from wildtype mice for comparison. RNAs that appear to be present in the nucleus can be validated for this localization using RNA fluorescent in situ hybridization (RNA FISH).

The sequenced miRNAs and siRNAs will subsequently be examined for differences between the substages of meiotic prophase I. These will be differences in the RNAs association with AGO3 or AGO4 between the different substages. Since AGO3 and AGO4 localize specifically to the sex body during pachynema, we will look for siRNAs or miRNAs that associate with AGO3 or AGO4 and that exhibit increased transcript levels specifically during pachynema. Northern blotting can then be performed to validate an increase by probing for the specific miRNA or siRNA of interest. Finally, the small RNAs isolated from AGO3 versus AGO4 precipitates will then be compared for similarities or differences to explore whether the non-slicing AGO proteins have small RNA binding preferences.

Conclusions

The studies utilizing the epitope tagged AGO3 and AGO4 mouse lines will further our goal of understanding of the potential nuclear role for AGO proteins in male mammalian germ cells. The work outlined in this thesis begins with describing a likely role for sncRNAs in maintaining the integrity of the sex body. In chapter 2, it is described how in *Dicer* and *Dgcr8* cKO mice, where there is a loss of miRNAs and siRNAs, there is abnormal chromosomal fusions at the sex chromosomes and mislocalization of crucial proteins involved in MSCI initiation. Additionally, in chapter 3, I outline improvements made to a FACS protocol, which can be utilized for isolation of meiotic prophase I sub-staged spermatocytes. Finally, in chapter 4, I describe a strategy for the generation of mouse lines containing epitope tagged *Ago3* and *Ago4*

genes that can be used for examination the role AGO proteins are playing during male meiosis and MSCI. It is well accepted that meiotic sex chromosome inactivation (MSCI) is critical to male fertility and there is evidence that sncRNAs and proteins associated with the biogenesis of sncRNAs (AGO, DICER and DGCR8) are likely be involved in the mechanism of sex body silencing. Yet, there is no known role for AGO proteins and small RNAs in the nucleus during mammalian male meiosis. The work outlined here in this thesis sets up experimental means of examining the role that small non-coding RNAs and their associate AGO proteins may be playing in the process of MSCI. Better understanding the role AGO proteins play during MSCI will impact the current understanding of sncRNAs and AGO proteins, as well as male infertility and meiosis.

References

1. Cradick, T.J., et al., *COSMID: A Web-based Tool for Identifying and Validating CRISPR/Cas Off-target Sites.* Mol Ther Nucleic Acids, 2014. **3**: p. e214.
2. Modzelewski, A.J., et al., *AGO4 regulates entry into meiosis and influences silencing of sex chromosomes in the male mouse germline.* Dev Cell, 2012. **23**(2): p. 251-64.
3. Gonzalez-Gonzalez, E., P.P. Lopez-Casas, and J. del Mazo, *The expression patterns of genes involved in the RNAi pathways are tissue-dependent and differ in the germ and somatic cells of mouse testis.* Biochim Biophys Acta, 2008. **1779**(5): p. 306-11.
4. ten Have, S., et al., *Mass spectrometry-based immuno-precipitation proteomics - the user's guide.* Proteomics, 2011. **11**(6): p. 1153-9.
5. Trinkle-Mulcahy, L., et al., *Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes.* J Cell Biol, 2008. **183**(2): p. 223-39.
6. Nesvizhskii, A.I., *Protein identification by tandem mass spectrometry and sequence database searching.* Methods Mol Biol, 2007. **367**: p. 87-119.
7. Kramer, A., et al., *Causal analysis approaches in Ingenuity Pathway Analysis.* Bioinformatics, 2014. **30**(4): p. 523-30.
8. Dellaire, G., R. Farrall, and W.A. Bickmore, *The Nuclear Protein Database (NPD): sub-nuclear localisation and functional annotation of the nuclear proteome.* Nucleic Acids Res, 2003. **31**(1): p. 328-30.
9. Singh, P., J.C. Schimenti, and E. Bolcun-Filas, *A mouse geneticist's practical guide to CRISPR applications.* Genetics, 2015. **199**(1): p. 1-15.
10. Rio, D.C., et al., *Purification of RNA using TRIzol (TRI reagent).* Cold Spring Harb Protoc, 2010. **2010**(6): p. pdb prot5439.
11. Modzelewski, A.J., et al., *Dgcr8 and Dicer are essential for sex chromosome integrity in male meiosis.* J Cell Sci, 2015.