University of Nevada, Reno

Epigenetic regulation of spermatogenesis and transgenerational epigenetic inheritance

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Cell and Molecular Biology

By
Rachel A. Klukovich

Dr. Wei Yan, Dissertation Advisor

May, 2018
THE GRADUATE SCHOOL

We recommend that the dissertation prepared under our supervision by

RACHEL A. KLUKOVICH

Entitled

Epigenetic regulation of spermatogenesis and transgenerational epigenetic inheritance

be accepted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Wei Yan, Advisor

Qi Chen, Committee Member

William Courchesne, Committee Member

Seungil Ro, Committee Member

David Zeh, Graduate School Representative

David W. Zeh, Ph. D., Dean, Graduate School

May, 2018
Abstract

Epigenetics is defined as genomic modifications that alter gene expression without changing the nucleic acid sequence. These modifications come in hundreds of forms and affect both DNA and RNA. Recently, the field of epigenetics has taken to studying two newfound gene expression regulators: the highly abundant and universal N6-methyladenosine modification, and noncoding RNAs. N6-methyladenosine (m6A) has been implicated as a reversible, RNA modification that can cause infertility and other defects when not properly regulated. Noncoding RNAs have also been recently implicated as critical mediators of gene expression control with their ability to alter transcript stability and remodel chromatin. Other epigenetic factors, including noncoding RNAs, can become differentially expressed and transgenerational via epigenetic transgenerational inheritance with the ill effects lasting for generations. Abnormally regulated or persistent epigenetic modifications have been implicated in causing a variety of undesirable phenotypes such as infertility. To better understand the epigenetic causes of male infertility, this dissertation investigates the phenotype and mechanism of improper m6A regulation due to lack of an m6A-specific demethylase, ALKBH5, using a knockout mouse model (Chapter II). The reproductive health of male mice was monitored throughout development, and m6A RNA-immunoprecipitation and sequencing was performed at four different stages of spermatogenesis to follow transcript fate. The knockout mouse model had a delay in meiotic progression starting at P14.
that led to high levels of spermatogenic apoptosis in the adult testis. In addition, there was a correlation between increased splicing events, high m6A levels, and transcript degradation in the knockout. It was concluded that m6A acts in a biphasic manner: in the nucleus, it guides splicing events of pre-mRNAs while in the cytoplasm, it marks transcripts’ 3’UTR for degradation. Then, this dissertation investigates environmentally induced epigenetic causes of prostate disease using a rat model (Chapter III). Pregnant rats were transiently exposed to vinclozolin, and the F3 generation’s prostate stroma and epithelium were examined for epigenetic transgenerational inheritance of differentially expressed epigenetic factors, including noncoding RNAs and differential DNA methylation regions (DMRs). Both noncoding RNAs and the DMRs were found to be differentially expressed with many modifications altered intergenically. DMRs overlapped with the affected genes in both cell types indicating that they may be the primary mediators of transgenerational epigenetic inheritance. The small, noncoding RNAs were found to overlap significantly with the mRNAs in the prostate stroma. Because the modifications were distinct in both the prostate stroma and epithelium, this indicates that ancestral exposure can have different effects on different cell types within the same organ. Together, this dissertation gives a comprehensive overview of the epigenetic regulation of spermatogenesis and transgenerational epigenetic inheritance.
Dedication

I dedicate my dissertation to my parents, Nadia and Richard Klukovich, to my twin brother Eric Klukovich, and to my sister-in-law Sarah Klukovich. Their support, advice, and encouragement have been invaluable to me throughout this journey.
Acknowledgements

I would like to thank Dr. Wei Yan for the opportunity to perform research and to earn my Ph.D. in his laboratory. His support and guidance were essential for me to become the scientist I am today.

I would like to thank my advisory committee for their critical feedback and encouragement throughout this experience. Their insight and advice helped guide me to achieve my goal of earning my Ph.D.

Thank you to the members of Yan lab and Skinner lab for their contributions to this research and their ability to ensure that there was never a dull moment during the long lab days.
Table of Contents

ABSTRACT ............................................................................................................................................ 1
DEDICATION...................................................................................................................................... III
ACKNOWLEDGEMENTS .................................................................................................................. IV
TABLE OF CONTENTS .................................................................................................................... V
LIST OF TABLES ............................................................................................................................... VIII
LIST OF FIGURES ............................................................................................................................. IX
CHAPTER 1: INTRODUCTION TO REGULATORY EPIGENETICS AND INFERTILITY PHENOTYPES ............ 1

ABNORMAL m6A REGULATION CAUSES MALE AND FEMALE INFERTILITY BY CONTROLLING RNA FATE ........ 1
Abstract ................................................................................................................................................ 1
Summary Sentence ............................................................................................................................... 2
Key words: ........................................................................................................................................ 2
Introduction ....................................................................................................................................... 2
New methods of detection .................................................................................................................. 4
Patterns of m6A methylation ............................................................................................................... 11
m6A associated proteins .................................................................................................................... 13
Molecular Roles ................................................................................................................................. 18
Effects on fertility ............................................................................................................................... 30
Maternal effects ................................................................................................................................. 38
Conclusion ......................................................................................................................................... 43

INTRODUCTION TO EPIGENETIC TRANSGENERATIONAL INHERITANCE ............................................... 49

Epigenetic transgenerational inheritance of noncoding RNAs ................................................................ 49
CHAPTER 2: ALKBH5-DEPENDENT M6A DEMETHYLATION CONTROLS SPlicing AND STABILITY OF LONG 3'UTR MRNAS IN MALE GERM CELLS

Abstract

Significance Statement

Results

Discussion

Acknowledgements

Author Contributions

Competing Financial Interests

Materials and Methods

Figures

References

CHAPTER 3: ENVIRONMENTAL TOXICANT INDUCED EPIGENETIC TRANSGENERATIONAL INHERITANCE OF PROSTATE PATHOLOGY AND STROMAL-EPITHELIAL CELL EPIGENOME AND TRANSCRIPTOME ALTERATIONS: ANCESTRAL ORIGINS OF PROSTATE DISEASE

Abstract

Introduction

Results

Prostate Pathology Analysis

DNA Methylation Analysis

noncoding RNA Analysis

Discussion

Methods
List of Tables

Table 1. Comprehensive list of journal articles correlating m6A-associated proteins with fertility phenotypes. ................................................................. 43
List of Figures

Chapter 1

Figure 1-1. RNA-immunoprecipitation and sequencing workflow. .......................... 8

Figure 1-2. m^6^A writer, eraser, and reader proteins influence the fate of the transcript. ............................................................................................................................................. 12

Figure 1-3. The effects of m^6^A on temporal gene regulation. ................................. 42

Figure 1-4. Gestating female's exposure to endocrine-disruptors causes epigenetic transgenerational inheritance in male pups. ............................................. 56

Chapter 2

Fig. 2-1. *Alkbh5* is essential for meiotic and haploid phases of spermatogenesis. ................................................................................................................................................ 91

Fig. 2-2. m6A marks the longer 3'UTR transcripts that are destined to be eliminated during spermiogenesis (from round to elongating/elongated spermatids). .................................................................................................................................. 93

Fig. 2-3. Proper m6A erasure is required for the production of longer 3'UTR mRNAs in pachytene spermatocytes and spermatids. .............................................................................. 95

Fig. 2-4. Fate of the aberrantly spliced short transcripts in *Alkbh5* KO (KO) round spermatids. ..................................................................................................................... 97

Chapter 3

Figure 3-1. Prostate pathology frequency. ................................................................. 133
Figure 3-2. DMR identification.................................................................134
Figure 3-3. DMR chromosomal locations...............................................135
Figure 3-4. DMR genomic features..........................................................136
Figure 3-5. DMR gene associations..........................................................137
Figure 3-6. DMR association gene pathways..........................................138
Figure 3-7. Differential expression of noncoding RNAs .......................139
Figure 3-8. Chromosomal location of differentially expressed large RNAs. .....140
Figure 3-9. Chromosomal locations for differentially expressed small, noncoding
RNAs ........................................................................................................141
Figure 3-10. Overlaps of the DMRs with the differentially expressed noncoding
RNAs ........................................................................................................142
Chapter 1: Introduction to regulatory epigenetics and infertility phenotypes

Abnormal m<sup>6</sup>A regulation causes male and female infertility by controlling RNA fate

Rachel Klukovich<sup>1</sup> and Wei Yan<sup>1,*</sup>

This manuscript is in preparation for submission for publication.

<sup>1</sup>Department of Physiology and Cell Biology, University of Nevada, Reno School of Medicine, Reno, NV 89557

*Correspondence: Department of Physiology and Cell Biology, University of Nevada, Reno School of Medicine, Reno, NV 89557. E-mail: wyan@med.unr.edu

Abstract

Infertility is a common problem around the world, affecting approximately 15% of couples. Throughout the years, it has been the general belief that the majority of these problems stem from the female. It has been recently acknowledged that around 50% of infertility problems arise from the male, often due to low sperm
counts, poor motility, and morphological defects in the spermatozoa. Recent evidence suggests that these defects are the result of abnormal transcript expression, which is often the result of epigenetic modifications. A recently discovered RNA epigenetic modification is m$^6$A, or N$^6$-methyadenosine. This modification is abundant, yet scientists have yet to reach a consensus on what the function is of this epigenetic mark. This is, in part, due to difficulties in precisely detecting this modification. In recent months, numerous high-impact articles have been published showing how the different “writers”, “readers”, and “erasers” of m$^6$A all impact male and female infertility. Here, we review the many prospective methods of detection, the proposed functions, and the link between deregulated m$^6$A content and male and female infertility.

**Summary Sentence**

m$^6$A is a critical regulator of gene expression during gametogenesis, and deregulation of the writers, readers, or erasers can lead to complete infertility or subfertility in both males and females.

**Key words:** Epigenetics; RNA methylation; noncoding RNA; gene regulation; gametogenesis; fertility; female infertility; male infertility

**Introduction**

Epigenetic modifications such as methylation and acetylation affect the expression of nucleic acid transcripts in an occasionally heritable manner and are
consequently able to pass through generations through the germ line. These heritable modifications have the ability to regulate the transcripts at the post-transcriptional level without altering the transcript sequence. This mainly occurs during development and differentiation, so will have lifelong phenotypic effects on the adult organism.

While there are a wide variety of epigenetic modifications, the most frequently occurring and universal mRNA modification in eukaryotes is N\(^6\)-methyladenosine, or m\(^6\)A. Additionally, it is conserved across eukaryotic species and is also found in viruses, bacteria, and archaea. Many different classes of RNA have been discovered to contain this modification including: mRNA, rRNA, tRNA, and small nuclear RNA (1–4). This modification does not affect mRNA message coding ability. It is also believed to be widespread throughout transcripts as it has been located in long internal exons of at least 400 nt (5), and is possibly a part of the 5' cap structure (6). Furthermore, it has been estimated to have 49% conservation between human and mouse orthologues indicating non-random distributions on transcripts (5). This non-random distribution and widespread incorporation in transcripts are strong indications of an important function for these residues. However, the role for particular m\(^6\)A residues, and the function of m\(^6\)A in gene regulation and transcript fate, is currently unknown.

The prevalence of this modification in mammalian transcripts is 3 m\(^6\)A residues per mRNA transcript, or approximately 0.1-0.4% of adenosines in RNA (7). The majority of these residues are believed to be located near stop codons, the 3’ UTR, the 5’ UTR, and in exons (8–10) at the consensus sequence
Pu(G>A)m$^6$AC(U>A>C), also written as RAC/RRACU, where R is a purine.

However, this consensus sequence is more prevalent than actual m$^6$A methylation sites and does not always indicate the presence of an m$^6$A residue: last exons have 70% of the total m$^6$A but 50% of the motifs (5, 11). This gives rise to the need for better methods for determining modification sites than solely the consensus sequence.

Although this modification is so widespread, it is difficult to detect because the modification itself does not affect the base pairing ability and is not able to be chemically modified with current methods (i.e. bisulfite treatment for 5mC). Due to these difficulties, previous methods for m$^6$A detection involved using radiolabeled methionine to detect methylation on adenosine. These radiolabeled nucleotides were then mapped using HPLC or thin-layer chromatography (7). Although these methods are able to correctly detect m$^6$A residues, it does not allow for genome-wide mapping and functional predictions.

**New methods of detection**

**Immunoprecipitation**

Two different groups developed similar immunoprecipitation techniques to overcome these challenges. Dominissini et al. (2012) developed a new method that uses RNA immunoprecipitation to detect m$^6$A residues on the transcripts. The RNA immunoprecipitation and sequencing workflow can be seen in Figure 1-1. RNA was fragmented into 100 nt long segments and was immunoprecipitated using an anti-m$^6$A antibody. The precipitated fragments were then produced into
libraries for next-generation sequencing (NGS). The RNA-seq data is then aligned to the genome. If the anti-m\textsuperscript{6}A antibody successfully enriched for m\textsuperscript{6}A modifications, then the RNA-seq data will overlap in the same locations, creating a peak. This peak indicates the probable location of the m\textsuperscript{6}A residue on the transcript.

It was found that alternatively spliced genes were enriched with m\textsuperscript{6}A, while single isoform coding genes were less methylated indicating m\textsuperscript{6}A has a role in splicing. Further support for this is that proteins that bind to m\textsuperscript{6}A also have roles in alternative splicing and targeting transcripts for nuclear degradation. Additionally, m\textsuperscript{6}A seems to affect expression levels with 33.5% of genes with higher expression were found to have at least one peak in the stop codon region, but not at the actual stop codon. Further, genes with moderate expression had the highest methylation levels over genes with very high or low expression, except in the TSS where high gene expression is correlated with high methylation (5).

It is well known that RNA has abundant secondary structures in its native state. It is interesting to note that N\textsuperscript{6}-substitutions on adenosine are destabilizing when located in RNA duplexes, likely caused by steric hindrance. However, when located at the 3'-terminal nucleotide in an unpaired state, the RNA duplex becomes stabilized (12). Liu et al. (2013) mapped their predicted m\textsuperscript{6}A residues and predicted the secondary structure of the transcripts. They predicted that the m\textsuperscript{6}A residues are located in the stem of a hairpin loop. This may allow interactions with other protein machinery due to the presence of m\textsuperscript{6}A slightly
decreasing the stability of RNA secondary structures (5, 12, 13). It is important to note that the presence of these secondary structures can mask the m\(^6\)A binding site, making immunoprecipitation difficult.

Another group developed a similar technique that also uses two subsequent rounds of RNA-immunoprecipitation and next-generation sequencing (14). They then mapped the m\(^6\)A-enriched transcripts to the DAVID bioinformatics database to predict functions for m\(^6\)A. It was found that transcripts enriched with m\(^6\)A had potential roles in transcriptional regulation and intracellular signaling cascades. They also noted that the enriched transcripts were mapped to genes involved with neurological diseases.

Interestingly, Meyer et al. (2012) used bioinformatics to detect the association between m\(^6\)A residues and miRNA binding sites. Of the 3’ UTR transcripts containing m\(^6\)A residues, two-thirds had a predicted miRNA binding site. It was also noted that the m\(^6\)A residues come sequentially before the miRNA binding site in 62% of studied cases, with the m\(^6\)A residue being located near the stop codon. In contrast, the miRNA binding sites are more abundant at the far end of the 3’ UTR. Further, the miRNAs with the highest expression had target transcripts containing more m\(^6\)A peaks per target gene than did the least expressed miRNAs.

miRNAs are known to function as posttranscriptional gene regulators. They are first produced as primary transcripts, which then go through multiple rounds of modifications before becoming the final, mature, miRNA product. This maturation is highly regulated and modifications on the transcript itself may play
a role in determining the miRNA product. Berulava et al. (2015) hypothesized that the m\textsuperscript{6}A residues play a role in this regulation, and that disruption of the m\textsuperscript{6}A modifications will cause dysregulation of miRNA biogenesis. Indeed, they found that knockdown of FTO, a known m\textsuperscript{6}A demethylase caused such dysregulation, both significantly increasing and decreasing many miRNA transcripts from the wild-type. Further, it was noted that the consensus sequence motif was only present in 20.9% of all miRNAs, but was significantly enriched in the immunoprecipitated miRNA at 32.6% (1).

There are advantages to immunoprecipitation that make this method favorable. The specificity of the antibody to the m\textsuperscript{6}A site allows for accurate detection. This method can also be used genome-wide in all classes of RNAs. However, in order to accurately predict functionality, the sequenced libraries need to be long enough to prevent the loss of specificity when aligned to the genome to prevent mismatches. On the other hand, increasing the length of the libraries does have the disadvantage of making the location of the specific m\textsuperscript{6}A residue at the single-nucleotide level ambiguous.
Figure 1-1. **RNA-immunoprecipitation and sequencing workflow.** The RNA content is fragmented and immunoprecipitated with an anti-m$^6$A antibody. Libraries are constructed for next-generation sequencing (NGS) by adding adaptors (orange) and index primers (navy blue). The RNA-seq data is then aligned to the genome and the relative frequency will create a peak indicating where the m$^6$A site is located.
**Single-nucleotide resolution**

A variation of these immunoprecipitation techniques is m^6^A-CLIP/immunoprecipitation (cross-linking immunoprecipitation), which is used for single nucleotide resolution using very short transcripts (20-80 nt). While this method has the advantages of high specificity and low background with the negative control of a nonspecific rabbit IgG, there are several disadvantages. The cross-linking uses UV light which induced insertions, deletions, or substitutions in some transcripts at +1 position (3′ to m^6^A site) (15) or truncations (16). Further, proteinase is required for elution to remove the crosslinked antibody from the transcript and leaves an additional amino acid residue or peptide attached to the transcript. Finally, this method has a high false discovery rate of 17% (11).

Liu et al. (2013) developed a novel method to assess the presence of the m^6^A residues called SCARLET (Site-specific Cleavage And Radioactive-labeling followed by Ligation-assisted Extraction and TLC). The method uses 2′-OMe/2′-H complementary oligonucleotides to direct RNaseH to cleave the RNA in a site-specific manner 5′ to the possible m^6^A modification. The cut site is radiolabeled, ligated to DNA oligonucleotides, and the RNA is subsequently treated with RNases to digest the RNA. The ligation serves as a shield against the RNases, thus allowing the RNases to digest only non-candidate RNAs. The sample is then digested into single-nucleotide fragments and run through thin-layer chromatography to determine if the m^6^A site is in fact present on the candidate RNA of interest.
The authors show that this method is sensitive enough to detect m^6A modifications at the femtomolar level (13). Furthermore, the use of thin-layer chromatography lends precision to this method. However, there are some disadvantages. This method is not practical at a genome-wide level. Additionally, the use of RNase H would make it difficult, if not impossible, to detect m6A modifications on small RNAs, given that many modifications are present on this RNA class. Such digestion would cleave the short RNA transcripts into fragments that are too short for ligation.

Linder et al. (2015) claim that they have developed a method to locate m^6A at the single nucleotide level by miCLIP (m6A individual-nucleotide-resolution cross-linking and immunoprecipitation). They state that the problem with previously developed methods is the inability to locate specific m^6A residues on the transcript, but rather locate 100-200 nt regions where m^6A modifications may be present. Their method involves UV light to cross-link the anti-m^6A antibodies to the RNA (17). However, cross-links of this kind are known to induce transcriptional mutations, especially when reverse transcription is performed. The most commonly occurring mutations were substitutions and truncations. Many of these mutations were close to the m^6A residue, however many were unpredictable. This makes it nearly impossible to do a thorough and accurate bioinformatics analysis to locate the m^6A residues. Although the authors show that when predictable, these mutations can be used to detect m^6A, the inconsistent nature of mutations makes it impractical.
Patterns of m6A methylation

The abundance of these residues is most common at 1 or 2 residues per transcript: approximately 46% of transcripts have a single m6A residue detected while 37% have two residues. Only 11% have three residues and transcripts with four and more residues only make up 5% of all transcripts. Meyer et al. (2012) noted that 90% of the genes containing more than one m6A residue had these residues adjacent to each other, suggesting a clustering of m6A sites. Approximately 95% of the m6A residues were located in intergenic regions. Of these, the coding regions had 51% while the untranslated regions, specifically the 3’ UTR, had 42%. The highest enrichment in the 3’ UTR is located in close proximity to the stop codon with 25% of total m6A residues being mapped to the first quarter of the 3’ UTR.

Ke et al. (2015) noted that m6A was enriched in beginning of the last exon, not around the stop codon, even if last exon is noncoding. m6A density peaks at 150-400 nt after the start of the final exon, independent of stop codon, which may indicate a connection with mRNA stability and polyA site choice. It was observed that longer final exons have more m6A residues than short final exons. This provides evidence that m6A may be inhibitory for polyadenylation: m6A may be needed in long last exons to prevent proximal polyadenylation while short exons need early polyadenylation. Transcripts with distal polyA site usage had more m6A residues before APA site while proximal polyA site had lower m6A density. They observed that knockdown of m6A causes alternative polyA site usage,
however this is not true for all mRNAs: only some transcripts in brain and liver used different sites, indicating that m^6^A is not just for APA site choice.

**Figure 1-2. m^6^A writer, eraser, and reader proteins influence the fate of the transcript.** The methyltransferase complex adds a methyl group to the sixth position on the adenosine while the demethylases FTO and ALKBH5 remove them. The m^6^A residues are then read by the reader proteins YTHDF or YTHDC to determine the outcome the m^6^A modification will have on transcript fate.
**m^6A associated proteins**

**Methyltransferase Complex**

Adenosine residues are methylated on the sixth position of the purine ring by the methyltransferase complex, which is made up of three core proteins (Figure 1-2). METTL3 (methyltransferase-like 3) encodes a 70 kDa enzyme which is a part of an 800 kDa complex with m6A methylase activity (Bokar et al., 1997). METTL3 and METTL14 form a heterodimer that methylates mRNAs (Liu et al. 2014; Wang et al. 2014). WTAP (Wilms' tumor 1-associating protein) is a mammalian splicing factor apart of the core methyltransferase complex (19, 20). It has been demonstrated that METTL3 knockout is lethal and prevents differentiation in embryonic stem cells (21, 22). METTL3 knockdown has also been shown to cause apoptosis (5), further emphasizing the importance of the ability of transcripts to incorporate m\(^6\)A residues.

**YTH Reader Proteins**

The protein domain family YTH is known to bind single-stranded RNA. A member of this family, YTHDF2, binds selectively to m\(^6\)A residues. Wang et al. (2014) used PAR-CLIP (photoactivatable ribonucleoside crosslinking and immunoprecipitation, m\(^6\)A-seq, and RIP-seq to correspond the m\(^6\)A binding targets to the YTHDF2 protein. It was found that 59% of the PAR-CLIP peaks overlap with m\(^6\)A peaks and that this binding occurs at the stop codon 42% of the
time. This binding also occurs at the coding sequence and UTR 36% and 14% of the time, respectively, indicating that this binding may regulate RNA fate (23).

To assess the role of YTHDF2 in RNA fate, Wang et al. (2014) treated cells with siRNA to knockdown YTHDF2. It was noted that the knockdown had an increase in input target mRNA reads in the knockdown versus the control. But, the ribosome protected (translated) reads were similar to the control, indicating that knocking-down YTHDF2 decreased translation levels. Further, mRNA lifetime was increased by 30% in cells that had YTHDF2 knocked-down, indicating that YTHDF2 has a role in regulating mRNA degradation.

To confirm this, the authors used fluorescent microscopy to determine the localization of YTHDF2 in the cell. It was found that it co-localizes with three markers in the processing bodies that degrade mRNA: DCP1a, GW182, and DDX6. This provides further proof that the m^6A residues function as a marker for RNA fate and that the YTHDF2 reader proteins are able to interpret these markers for mRNA degradation in the P bodies. Therefore, m^6A may have a role in gene expression as a negative regulator.

**FTO**

The DNA demethylase AlkB has nine mammalian homologues: ALKBH1-8 and FTO. FTO is the fat mass and obesity-associated protein apart of the AlkB family Fe^{II}/α-ketoglutarate-dependent dioxygenases. FTO was found to have the ability to oxidatively demethylate m^6A and convert it to adenosine for ssDNA, dsDNA, RNA, and stem loop RNA structures in vitro (Jia et al., 2011). This
process was complete for both ssDNA and for ssRNA, but the stem loop structure only had 40% demethylation under the same conditions while the dsDNA and dsRNA needed longer treatments to achieve partial demethylation (40% and 24%, respectively). Active site mutations of FTO confirm that the demethylation of m$^6$A is both Fe$^{ll}$ and α-ketoglutarate dependent.

Further evidence that m$^6$A is the substrate for FTO oxidative demethylation is that knockdown of FTO increased m$^6$A by 23-42% depending on the cell type, and overexpression of FTO decreased the m$^6$A levels by 18% (24). Immunofluorescence experiments showed that FTO co-localizes in the nucleus with many different splicing factors such as SART1 and SC35, implicating FTO as a possible regulator of m$^6$A levels to coordinate splicing activity.

The 5’ cap on mRNA confers stability to the transcript and is made of many methylated bases including N$^7$-methylguanosine and a methylated ribose at the 2’-hydroxyl position. Approximately 30% of caps have an additional methylation modification on the first nucleotide after the methylated guanine ($A_m$ or 2’-o-methyladenosine); $A_m$ is additionally methylated at the N$^6$ position, generating m$^6$A$_m$ (N6,2’-o-dimethyladenosine), which is the second most abundant mRNA modification after m$^6$A. This additional modification at the 5’ cap further improves stability to the mRNA over other modifications when beginning the transcript sequence, increasing both the transcript levels and the half-life by 2.5 hours. m$^6$A$_m$ also confers resistance to decapping enzymes and microRNA-mediated mRNA degradation.
There are many similarities between m\(^6\)A\(_m\) and m\(^6\)A due to the common methylation at the N\(^6\) position on adenosine; this similarity has caused some ambiguity between the two modifications. First, many antibodies bind to both m\(^6\)A and m\(^6\)A\(_m\) causing many false positives for the modification of interest due to misannotated peaks. In addition, it has been widely reported that FTO is one of the two known demethylases specific for m\(^6\)A (Jia et al., 2011). Now, Mauer et al. (2017) report that FTO is specific for m\(^6\)A\(_m\) and found that m\(^6\)A\(_m\) is readily converted to A\(_m\) by FTO (25). Although it was reported that FTO can demethylate m\(^6\)A, the catalytic efficiency of FTO for m\(^6\)A\(_m\) is 100x higher than for m\(^6\)A, indicating that the specific target for FTO is m\(^6\)A\(_m\), not m\(^6\)A, as previously believed.

However, FTO will demethylate m\(^6\)A when present at high levels, although inefficiently. This does not mean that m\(^6\)A is not a target of FTO, as Mauer et al. (2017) report that FTO-mediated demethylation of m\(^6\)A\(_m\) is dependent on the m\(^7\)G, indicating that recognition of the RNA site is dependent on the structure of the surrounding nucleotides. The surrounding methylation pattern near an m\(^6\)A modification may determine the specificity for FTO to demethylate it. It was also noted that ALKBH5 knockdown increased the levels of m\(^6\)A, but had no effect on the levels of m\(^6\)A\(_m\), indicating that this demethylase is specific for m\(^6\)A. Location and combination of different modifications ultimately determine the effects on the mRNA. For example, both m\(^6\)A and m\(^6\)A\(_m\) on the 5'UTR increase translation efficiency, while in other regions, m\(^6\)A is a degradation marker and m\(^6\)A\(_m\) is a stability marker.
ALKBH5

As mentioned in the previous section, Jia et al. (2011) discovered that FTO, an obesity-associated protein that is a part of the AlkB family of FeII/α-ketoglutarate dependent dioxygenases also functions as RNA demethylase that removes m⁶A from transcripts. It has been previously shown that ALKBH5 is a mRNA-binding protein (26), so the other AlkB protein homologues were examined by HPLC for demethlyase activity. It was observed that one of the homologues, ALKBH5, also has complete demethylation activity of m⁶A (27).

The authors found that m⁶A content was increased by 9% in cells that had ALKBH5 knocked down and was likewise decreased by 29% when ALKBH5 was overexpressed. This suggests that ALKBH5 plays a regulatory role in m⁶A methylation status through its demethylation activity. Further indications that ALKBH5 functions as a demethylase are from the observations that it colocalizes in the nucleus with mRNA-associated proteins: SC35, SM, ASF/SF2. Additionally, RNase treatment reduced the fluorescence of ALKBH5, indicating interactions with RNA. Immunofluorescence experiments also showed that levels of cytoplasmic RNA were increased in the ALKBH5 knock-down cells, likely due to increased RNA export. (For functional roles, see spermatogenesis section).
Molecular Roles

Pluripotent Stem Cell Differentiation

Geula et al. (2015) used small interfering RNA (siRNA) to screen epigenetic regulators in primed epiblast stem cells (EpiSCs) using GFP as a reporter and found that knockout of Mettl3 reduced primed cell viability. Deletion of Mettl3 almost completely depleted m\(^6\)A residues and resulted in embryonic stem cells maintaining their naïve pluripotent state. Differentiation was induced by growth on differentiation media and the knockout cells were unable to cavitate while the knockout embryoid bodies were unable to repress pluripotent gene expression or to form embryo chimeras.

When both wild-type and knockout embryonic stem cells were injected into mice to develop tumors, only the wild-type cells were able to generate mature teratomas. More than 75% of the knockout cells expressed pluripotent markers but only low levels of differentiation markers. Further, when fibroblast growth factor 2-activin was used, the knockout colonies also maintained their naïve pluripotent state. These observations indicate that m\(^6\)A is vital for embryonic stem cells’ ability to prime and differentiate, consequently leading to a perpetual naïve pluripotency state when devoid of m\(^6\)A modifications.

Approximately 80% of the naïve pluripotency-promoting genes had m\(^6\)A modifications. The global transcriptional profiles were also highly correlated to m\(^6\)A levels where the expression level changes between the wild-type and knockout cells were positively correlated to the number of m\(^6\)A modifications.
Depletion of m^6A induced by Mettl3 knockout resulted in increasing levels of mRNA transcripts that had been highly m^6A methylated.

Ythdf proteins control the degradation of mRNA transcripts through m^6A modifications (23), directly affecting the mRNA half-lives. Geula et al., (2015) measured the half-lives of m^6A methylated and unmethylated transcripts using actinomycin D to inhibit transcription, and found that in wild-type cells, the methylated transcripts had significantly shorter half-lives than the unmethylated transcripts. This trend was reversed in the Mettl3 knockout cells with a significant increase in the half-lives of the methylated transcripts. m^6A modifications are therefore a direct mark for transcript degradation. They concluded that the hyperpluripotency observed in the Mettl3 knockout was due to increased stability of the methylated pluripotent transcripts. Further, the translation efficiency was found to be slightly increased in the knockout cells, implicating m^6A to be negatively correlated with translation, specifically in GC-rich transcripts.

Embryonic immunostaining of Mettl3 knockout mice showed that beginning at E6.5, there was a diffuse expression pattern of the pluripotent genes throughout the epiblast that was not observed in the wild-type epiblast. Depleted m^6A profiles in mRNA transcripts caused an inability to prime the lineage for differentiation of embryonic stem cells leading to early embryonic lethality. They concluded that m^6A acts as a regulator of molecular switches and is responsible for the regulation of pluripotency genes necessary for differentiation.
Haematopoietic stem cell differentiation

Zhang et al., 2017 wanted to determine the effects of Mettl3 on haemoatopoietic stem progenitor cells (HSPC) during embryogenesis (Zhang et al., 2017). These cells give rise to all blood cell types in the body and emerge in zebrafish at 28 hpf. The authors used a mettl3 mopholino (MO) to decrease METTL3 protein levels, and consequently, m^6A levels, which corresponded to a decrease of approximately 4,500 genes, leading to the hypothesis that these were targets of METTL3. Many of these genes were indicated to be involved in embryonic developments, suggesting that METTL3-regulated m^6A modifications are required for this process. There was a noticeable decrease in mettl3 morphants for both HSPC and differentiation markers. Further, this led to a decrease in the T lymphocyte production in the morphants and a failure to progress successfully through endothelial-to-haematopoietic (EHT) indicating that METTL3-regulated m^6A modifications are required for the HSPCs to differentiate.

RNA-seq of the METTL3 targets revealed that there were 680 upregulated genes, 30 of which were involved in haemo-vascular development. Notch1a, expression of which represses HPSC differentiation, had greatly increased mRNA abundance and decreased levels of m^6A in the mettl3 MO. HPSC production was rescued by inhibiting notch1a, indicating that HSPC generation is inhibited in the mettl3 MO by regulating notch1a expression.

Finally, the authors looked at the role of m^6A reader YTHDF2 in HSPC generation by mutating the m^6A recognition site and found lower expression of
HSPC markers as well as less HSPC. RNA-seq showed that there were ~2,900 common gene targets between METTL3 and YTHDF2, and notch1a was upregulated. Notch1a also had internal m⁶A modifications that were recognized by YTHDF2 and the ythdf2 morphants had increased notch1a expression, indicating that YTHDF2-recognized m⁶A modifications may lead to the degradation of notch1a. Taken together, these data indicate that m⁶A is required for the balance of mRNA fate for proper stem cell differentiation (Zhang et al., 2017).

T-cell differentiation

m⁶A has been implicated in many differentiation pathways, causing Li et al., (2017) to hypothesize that it may also have a role in the differentiation of T cells, so generated conditional knockout METTL3-null, CD4+ T cells(29). They used an animal model that causes colitis in mice who receive naïve T-cells with IL-7, which was demonstrated with wild-type T-cells. The Mettl3 KO recipient mice had a significant delay of disease onset and KO T-cells were not found in the recipients’ colon. Unlike the differentiated wild-type T-cells, the KO T-cells were found in the lymph nodes in their naïve state and remained unproliferated and naïve 1 month after transfer. Increased levels of apoptosis were not observed, indicating that cell viability was not affected by METTL3 deficiency.

Rescue experiments introducing WT Mettl3 into KO T-cells caused differentiation, further proving that proper m⁶A modifications are needed for differentiation. The m⁶A methyltransferase complex has several different proteins
including METTL3 and METTL14, causing the authors to knockout METTL14 and phenocopy the effects seen with METTL3-deficiency.

Maintenance of the naïve T-cell population is controlled by a combination of phosphorylation and interleukin stimulation, specifically differentiation of such T-cells is controlled by increased signaling of IL-7. To determine if this process is disrupted in by IL-7 stimulation in METTL3-deficient cells, the authors determined the phosphorylation levels of JAK1 and STAT5 and found a marked decrease in the KO. The phosphorylation of other signaling pathways, ERK and AKT increased in the knockout T-cells, indicating that the viability of the KO T-cells is maintained by ERK and AKT signaling pathways and that STAT5 is required for proliferation. RNA-seq showed that the most upregulated genes were the suppressor of cytokine signaling (SOCS) genes that were also found to have high levels of m^6^A in their mRNA 3'UTRs. In addition, the KO cells showed increased levels of mRNA due to decreased decay, but no difference was found in splicing, indicating that loss of m^6^A causes increased levels of the SOCS genes which lead to suppressed IL-7 and STAT5 signaling(29). The induction of different signaling pathways leads to the expression of different groups of genes that are regulated by the presence of m^6^A.

Transcriptional Repression

*XIST* (X-inactive specific transcript) is responsible for silencing gene transcription and is a long, non-coding RNA (lncRNA). Patil et al. (2016) studied *XIST* methylation in human cells and showed that RBM15 and RBM15B, two proteins apart of the m^6^A methylation complex, are responsible for initiating site-
specific methylation within $XIST$. The $m^6A$ reader protein, YTHDC1 is recruited to the $m^6A$ site and causes gene repression.

They used iCLIP (individual-nucleotide resolution UV crosslinking and immunoprecipitation) in HEK293T cells to test the functional redundancy of RBM15 and RBM15B. Male mouse embryonic stem cells that had doxycycline-dependent $Xist$ expression were used and expression was quantified using X-linked gene expression ($Gpc4$ and $Atrx$) quantified by FISH (30). They used short interfering RNAs to knockdown either of the $Rbm15$ or $Rbm15b$ genes, or a combination of both. It was discovered that only knockdown of both of the genes prevented $XIST$-mediated gene silencing, and therefore it was concluded that functional redundancy was observed in RBM15 and RBM15B for $XIST$ gene silencing.

Next, Patil et al. (2016) showed that $m^6A$ is required for transcriptional silencing controlled by $XIST$. They used siRNA to knockdown the $m^6A$ methyltransferase $METTL3$ and found that although $XIST$ was successfully induced, $Gpc4$ and $Atrx$ expression were not repressed. Because METTL3 is the known $m^6A$ methyltransferase and its knockdown represses $XIST$ transcriptional silencing (30), it is therefore reasonable to conclude that $m^6A$ is required for this process.

YTH proteins are the readers of $m^6A$ residues (5). There are different families of YTH proteins that recognize $m^6A$. There is the YTHDF family that has DF1-3 and the YTHDC family of YTHDC1 and YTHDC2. All of these proteins are cytoplasmic except for YTHDC1, which is located in the nucleus (23, 31).
of iCLIP experiments showed that YTHDC1 binds specifically to the m\(^6\)A residues in \textit{XIST}. Furthermore, 3D-SIM (3D structured illumination super-resolution microscopy) was used to show that YTHDC1 was enriched in the \textit{XIST} nuclear sub-compartment.

Finally, knockdown of \textit{YTHDC1} inhibited the \textit{XIST}-mediated gene silencing while the recruitment of YTHDC1 to \textit{XIST} by m\(^6\)A promotes gene silencing. This was proved by tethering YTHDC1 to \textit{XIST} by adding three BoxB hairpins to the 3’ end of the transcript. Silencing of \textit{Gpc4} was observed when \textit{XIST-(BoxB)\textsubscript{3}} expression was induced. Conversely, knockdown of both \textit{Rbm15} and \textit{Rbm15b or Mettl3} reversed this silencing (30). From these experiments, it can be concluded that the close proximity of YTHDC1 to \textit{XIST} without the methylation proteins is adequate to induce gene silencing (30). This indicates that the function of m\(^6\)A is to recruit reader proteins that will ultimately cause a phenotype.

It has also been noted by Linder et al. (2015) that the RBM15 and RBM15B proteins are located next to the m\(^6\)A-containing DRACH consensus sequences. It is therefore plausible that the binding of these proteins is necessary for the proper methylation and consequential recruitment of reader proteins to therefore cause a phenotype.

\textit{Drosophila} sex determination

Due to the previous implications of m\(^6\)A in stem cell differentiation, Lence et al., (2016) looked at the m\(^6\)A levels in \textit{Drosophila melanogaster}. m\(^6\)A levels
were at their highest in early embryogenesis and decreased rapidly until pupal stage. The only body parts that maintain high m⁶A levels in adulthood are the heads and ovaries. *Drosophila* has its own orthologue of METTL3, Ime4, depletion of which decreases m⁶A levels by 70% and causes a held-out wing phenotype. Depletion of *dMettl14* affects locomotion; deletion of both increases phenotype severity. Rescue experiments of neuronal, but not mesodermal, expression of GAL4 fixed the abnormal phenotype implicating m⁶A in neuronal functions. Exploration of the neuromuscular junction (NMJ) in the Ime4 KO showed that the synapses grew quickly with 1.5-fold increase in the number of boutons. RNA-seq on adult heads showed that there was a change in splicing of genes affecting locomotion in the KO. Because both Ime4 and *dMettl14* KO individually show the same phenotype, it is likely that these genes work together and depend on each other for proper function.

The sex determination gene, *Sxl*, was one of the top genes that was aberrantly spliced upon Ime4 KO. Male flies have a premature stop codon in the transcripts due to an extra internal exon. Both Ime4 and *dMettl14* KO females included this extra male exon causing a decrease in female isoforms. This indicates that the methyltransferase complex is required to control the splicing of specific genes, specifically the sex determination gene, *Sxl*, through m⁶A modifications. Notably, YT521-B-depletion also mimics the phenotype of the Ime4 KO.

The authors then wanted to identify the mechanisms of YT521-B involved splicing so used stable isotope labeling with amino acids in cell culture (SILAC)
to find interacting proteins. To determine if these proteins were involved in m⁶A-mediated splicing, these proteins were depleted from cells and three proteins were found to have effects on spicing: Hrb27c, QKr58E-1, and Nito. However, Nito was the only protein to consistently cause splicing defects similar to YT512-B when depleted and interacts independently of RNA. Nito is required for YT521-B to bind to mRNA causing the authors to conclude that Nito (vertebrate homologue RBM15) is a part of the m⁶A methyltransferase complex (32).

Temporal Gene Regulation in Development

He et al. (2017) performed analysis of m⁶A profiles in newborn, suckling, and adult pigs using MeRIP-seq further demonstrating that the m⁶A profile changes throughout developmental stages (33). In all stages, over half of the identified peaks (~5,800) were common in all three stages, and almost 75% of all identified peaks corresponded with intragenic regions. Further, m⁶A modifications were present on approximately 1/3 of expressed genes in all developmental stages. Of the m⁶A modified genes, almost 75% had one or two m⁶A peaks. In addition to being highly enriched at the start and stop codons (~10% and ~27%, respectively) as was previously reported (8–10), m⁶A peaks are also enriched at the coding sequence at a prevalence of ~41%. When plotted by gene segment location of m⁶A modification, it was found that gene expression was moderate for most m⁶A containing genes in all segments analyzed. However, m⁶A peak enrichment was negatively correlated with mRNA abundance, indicating that m⁶A modifications may target transcripts for degradation.
Because of this strong, negative correlation, functional gene analysis was used to determine the function of the genes that were consistently methylated. It was noted that many of the genes were involved in RNA metabolism, transcriptional regulation, and cellular biosynthesis. Analyzing the genes that are differentially expressed and differently methylated in the developmental stages further compounded this, with the newborn pig having the highest differential methylation, demonstrating that m^6A tightly regulates gene expression, especially in development (33).

DNA Damage Response

DNA damage response occurs by the cell first making various DNA modifications at the damaged DNA site to recruit repair factors. When probing for various modifications at damaged DNA sites induced by ultraviolet irradiation (UVA, UVC), Xiang et al., (2017) found that m^6A modifications were present at the damaged site, and increased in a dose-dependent manner as the amount of UVC irradiation increased, implicating that m^6A may be involved in the DNA damage response (34). DNA damage induced by gamma-irradiation or chemicals did not have this effect, indicating that this response is specific to UV damage. Treatment by RNase A eliminated the m^6A signal at the damaged sites, and m^6A modifications peaked in a dose-dependent manner in response to UV irradiation, specifically in the poly-A RNA population suggesting that it is the poly-A tailed RNA population that gets m^6A methylated in response to UV-induced DNA damage.
Two out of the three components of the \( \text{m}^6\text{A} \) methyltransferases complex (METTL3 and METTL14) localized to the DNA damage repair site. Notably, the third component, WTAP, did not. This did not affect the cell’s ability to produce \( \text{m}^6\text{A} \) modifications. The authors wanted to determine if the RNA was being demethylated based off of the quick loss of \( \text{m}^6\text{A} \) on the damaged sites after UV-irradiation. FTO was found to be present at the damage sites, and knockout of FTO shows very strong increases of the \( \text{m}^6\text{A} \) levels at the damage sites post UV irradiation, which persists for longer periods of time than the wild-type. Alkbh5 knockdown cells do not show this increase in \( \text{m}^6\text{A} \) levels, indicating that Alkbh5 is not responsible for modulating \( \text{m}^6\text{A} \) levels at DNA damage sites, but that FTO is, indicating non-redundant function for these two \( \text{m}^6\text{A} \) demethylases.

After UVA irradiation, it is worth noting that the methyltransferases METTL3 preceded the demethylase FTO to localize at the damage site by a little over one minute, allowing time for the methylase to introduce the \( \text{m}^6\text{A} \) modifications before the demethylase could modulate them. The introduction of cellular stresses (i.e. UV damage) led to a slight increase in the \( \text{m}^6\text{A} \) peaks in the 5’ UTR. This \( \text{m}^6\text{A} \) response was found to be required for not only the repair process as the METTL3 KO cells were unable to remove the cyclobutane pyrimidine dimers quickly, but also for transcription to occur after the DNA had been damaged. Further, cells with METTL3 inactivated were found to have a significantly reduced survival rate at less than half that of the WT cells, indicating that the \( \text{m}^6\text{A} \) modifications are required to be introduced to transcripts in response to DNA damaged caused by UV irradiation.
Although these m^6A modifications are required to be introduced to transcripts for repair, they are not required to recruit the components of either the nucleotide excision repair or double stranded break repair proteins, indicating that the m^6A modifications are required for the transcripts they are modifying to change the transcription patterns, but not for recruitment of all proteins to be used in other pathways. DNA polymerase κ was the only polymerase tested that was required to have both METTL3 and METTL14 for its recruitment. No other known m^6A binding proteins (reader proteins YTHDC1, YTHDF1-2) localized to the sites of UV damage or were required to recruit DNA pol κ. DNA pol κ recruitment dependent on METTL3 in the S/G2 phases, remains at the damage site long after METTL3 has reduced methylation. This suggests that METTL3 is required for initial recruitment of DNA pol κ to the damage site, but that other factors are responsible for downstream effects. Additionally, m^6A is required for global post-transcriptional regulation, and that this regulation is responsible for regulating the expression of other genes that are required for the DNA damage response pathway.
Effects on fertility

Spermatogenesis

Clinical infertility

Landfors et al. (2016) wanted to assess how different variants of FTO and ALKBH5 affected human semen quality and used immunofluorescence to show that FTO was located in human testes, most notably in sertoli cells. Further, ALKBH5 was observed to colocalize in the nucleus with SC35, a nuclear speckle protein. The strongest ALKBH5 signal was observed in the nucleus, however, there was also diffuse cytoplasmic expression that could be observed. To determine any sequence variations of these two genes, sequencing and genome mapping were performed on semen samples of 77 men being tested for infertility. Both the ALKBH5 and FTO genes had 21 and 12 sequence variants, respectively. Most notably, two of these variants in the FTO gene had missense mutations that altered the amino acid sequence and the corresponding protein product. It is not currently known what specific effects these mutations have on the fertility of the patients, however it was observed that the morphology of the sperm was dramatically altered in the patients with the FTO mutation p.Cys326Ser, where a cysteine residue was replaced by a serine in the linker region connecting two FTO domains (35).

It is well known that infertile men often have morphological defects in their sperm. Zheng et al. (2013) showed the morphological defects in the tails of the spermatozoa in Alkbh5 knockout mice, where individual sperm had abnormal
morphology including two tails, bent tails, or missing heads (27). As mentioned in the “m\(^6\)A relevant genes” section, these mice also had very low fertility rates and testis that were approximately 50% the mass of the wild-type. The sperm counts and motility from these mice were significantly reduced while the knockout testis had increased levels of apoptosis. Although these phenotypic effects were observed, it remains to be elucidated what the underlying mechanism is for this abnormal phenotype and at what stage of development these abnormalities occur.

Another group studied asthenozoospermia, or decreased motility of sperm, and compared the m\(^6\)A levels in asthenozoospermia patients with normal control patients. Sperm RNA m\(^6\)A levels were measured using LC-ESI-MS/MS in MRM mode (multiple reaction monitoring mode). It was found that asthenozoospermia patients had increased m\(^6\)A content than the controls, indicating that an abnormally high m\(^6\)A content is a risk factor for altered sperm motility (36). The authors then used RT-PCR to measure gene expression levels of m\(^6\)A regulatory genes such as: \(FTO\), \(ALKBH5\), \(YTHDF2\), \(METTL14\), \(METTL3\), and \(WTAP\). Interestingly, it was observed that the asthenozoospermia patients had significantly increased expression levels of \(METTL3\) and \(METTL14\), indicating that abnormally high levels of the methyltransferases account for the increased levels of m\(^6\)A in sperm. Finally, the authors used Pearson correlation analysis and determined that there is a negative correlation between m\(^6\)A content and sperm progressive motility. This correlation is reversed and positive for sperm immotility (36).
Together, these studies indicate that if FTO and ALKBH5 are not properly expressed, then the m$^6$A residue levels are abnormal, affecting RNA fate by processing proteins. As the fate of the transcripts may be degradation or instability, this leads to the transcripts being improperly expressed and morphological defects occur, leading to germ cell depletion and infertility.

**ALKBH5 is required for spermatogenesis**

In the “m6A associated proteins” section, ALKBH5 was implicated as an m$^6$A demethylase. Interestingly, Alkbh5 was detected in many organs, however had the highest level of expression in the testes (27). Alkbh5 knock-out mice survived into adulthood but had significantly smaller testes at almost 50% the mass of the wild-type mice. It was observed that knock-out mice crossed with heterozygotes had very low rates of breeding. Histological analysis revealed that testicle tubular structure was altered and the spermatozoa from the cauda epididymis were reduced in numbers, motility, and were morphologically altered. Further, knock-out mice’s seminiferous tubuli had altered germ cell numbers, with less pachytene and round spermatocytes, with an increase in the number of sertoli cells, indicating an inability of these cells to complete spermiogenesis (27). Finally, apoptosis of pachytene and metaphase-stage spermatocytes was observed, and this may account for the inability to complete spermiogenesis and the low rates of breeding.

Zheng et al. (2013) found that there was a significantly increased number of m$^6$A residues in the knock-out mice, further indicating that ALKBH5 has RNA
demethylase function specific for m⁶A. Gene ontology analysis revealed that the differentially expressed genes were involved in many cellular pathways, most notably, RNA metabolism. This indicates that m⁶A modifications may be a key regulator in these pathways and may be required for not only proper cellular function, but also for the signal transduction necessary to enter these pathways.

**ALKBH4**

A known homologue of ALKBH5 is ALKBH4. Nilsen et al. (2014) studied the histology in mice that have Alkbh4 knocked-out (Alkbh4Δ/Δ). They found that when treated with tamoxifen, an estrogen homologue used to activate enzymes in the presence of an inhibitory estrogen-binding domain, the testes were reduced in size while the seminiferous tubules had a smaller diameter. Germ cell depletion was also observed, and it was concluded that this was due to increased apoptosis in the knock-out, specifically in the pachytene spermatocytes. Germ cells further depleted in subsequent meiotic divisions and led to a reduced number of spermatids, indicating that the loss of pachytene cells affects later stages of spermiogenesis. Staining revealed that the possible cause for this phenomenon is the disruption of the synatonemal complex (37).

Previously, it was discovered that in the nuclei, ALKBH4 had an association with the nucleolus (38). Nilsen et al. (2014) confirmed nucleolar localization, specifically in sertoli cells, and further noted that ALKBH4 localized to euchromatin, close to heterochromatin. They discovered that ALKBH4 disappears in late pachytene.
Taken together, these results indicate that ALKBH4 is required for proper meiotic prophase, the disruption of which leads to apoptosis and a depletion of germ cells, possibly leading to infertility.

**Spermatogenesis initiation**

*Mettl3* is a part of the methyltransferases complex needed to add m⁶A modifications to RNA. Xu et al. (2017) generated germ-cell specific Mettl3-deficient mice to study the role of Mettl3 in spermatogenesis and found that these mice exhibited complete infertility and had 80% smaller testis compared to the control. In addition, no spermatids were found to be present at 8 weeks and the number of germ cells was reduced as early as P8. Differentiation was also significantly reduced in the knockout, and the meiosis initiation marker, STRA8, had reduced expression in both P10 and P12, indicating that Mettl3 is required for proper meiosis to occur. Further, pachytene spermatocytes were not found in the knockout mice, and spermatocytes unable to progress to meiosis undergo apoptosis, indicating that proper m⁶A methylation is required for meiosis (39).

The transcriptomes of the control and conditional knockout testis at P6 and P12 were compared using RNA-seq. The *Mettl3<sup>cKo</sup>* testis at P6 were found to have 157 down-regulated genes that were significantly associated with spermatogenesis regulation including differentiation, meiosis, and stem cell maintenance; this number increased to 699 at P12. The alternative splicing events between the knockout and the control were analyzed using the rMATS software and found that exon skipping was enriched. The notable increase in
alternatively spliced genes in the P6 knockout was the DNA replication and cell division; at P12, the notable increase changed to spermatogenesis. In the control testes, m^6A individual-nucleotide-resolution cross-linking and sequencing (miCLIP-seq) was used to determine m^6A locations. It was found that there was an increase in m^6A residues in exon skipped spliced genes and that genes containing m^6A had low inclusion levels in alternative exon usage in knockout testes. In the knockout, nine genes were shown to have a lower number of exon inclusion genes than the control, causing these genes to be shorter, indicating that Mettl3 is required for alternative splicing. Taken together, these data indicate that m^6A is an essential modification that is required for spermatogenesis to proceed properly, starting from as early as differentiation, all the way through meiosis (39).

**Meiotic prophase I progression**

YTHDC2 is a member of the YTH protein family. Hsu et al 2017 were the first to report that YTHDC2 binds specifically to m^6A, and has a much higher affinity at 3.2-fold to m^6A than unmodified A at the consensus sequence (40). Western blot analysis showed that the YTHDC2 was most highly expressed in the testis out of all of the organs tested. The authors then used RNA from adult mouse testes to perform CLIP-seq, which showed the binding site of YTHDC2 to have a pattern similar to the m^6A modifications: close to the stop codon. RIP-seq was performed to locate the gene targets and indicated that almost all of the gene targets of YTHDC2 contain m^6A modifications.
When knocked out, Ythdc2 caused both male and female infertility showing a loss of germ cells and consequentially smaller testis and ovaries at approximately 30% weight of the wild-type. Male mice had no spermatozoa whatsoever in the testis or epididymis. When mRNA expression levels were examined in the testes, the expression level of Ythdc2 peaked around P9. Because meiosis starts at P8, the authors hypothesized that Ythdc2 is involved in meiotic prophase I. This was confirmed by the observation that KO spermatocytes and oocytes were unable to progress past the zygotene stage and there were high levels of apoptosis in the KO testis at P9.5-10.5 (40). Ythdc2 KO mice had normal sertoli cells and spermatogonia, indicating that YTHDC2 is required for spermatocyte progression through meiotic prophase I, but not beyond.

One of the major spermatogenesis-associated RIP-seq targets of Ythdc2 was analyzed for both mRNA expression and protein abundance in both the WT and KO mice at P8.5. In the KO mice, there was an increase in mRNA abundance and a decrease in protein levels, indicating a strong 37% decrease in translation efficiency. This suggests that YTHDC2 has a role in both the abundance of mRNA and their translation which was further confirmed by the use of a luciferase reporter assay. Translation was increased by 30% and mRNA expression was decreased by 15% in the absence of YTHDC2, resulting in an overall increase in translation efficiency when YTHDC2 is present (40). To determine the effects of Ythdc2 binding to its targets, m⁶A-seq was performed on P8.5 KO mouse testes. Ythdc2 target genes containing m⁶A modifications had a
slightly increased expression level in the KO testes, indicating that YTHDC2 has a role in RNA stability of its target genes.

YTHDC2 likely binds to its targets primarily by recognizing m\(^6\)A: the authors knocked down METTL3, a key component of the methyltransferase complex, and found that translation efficiency was compromised with an increase in mRNA abundance. Polysome profiling followed by Western blot revealed that YTHDC2 was in the 40-80s ribosomal fractions that are involved in translation initiation. Ythdc2-bound protein immunoprecipitation followed by mass spectrometry analysis showed that the top target proteins were involved in both translation and nonsense-mediated decay, indicating that YTHDC2 increases translation efficiency by interacting with translation complex and mediating nonsense-mediated decay, a process that occurs concurrently with translation (40). Meiotic prophase I requires rapid clearance of non-necessary transcripts to allow for increased translation efficiency explaining the requirements of YTHDC2 for proper progression throughout prophase.

YTHDC2 was shown to be an m\(^6\)A reader protein, where fluorescence anisotropy was used to determine that YTHDC2 has a stronger affinity toward m\(^6\)A over unmethylated transcripts (41). Western blot analysis showed that YTHDC2 had highest expression in the testis and little expression in the adult ovaries. Expression in ovaries was observed at E14.5. YTHDC2 KO males and females were both infertile and female ovaries showed atrophy with small follicles that did not progress beyond primary follicles, indicating that YTHDC2 is needed to undergo meiosis.
Testis also showed atrophy and are depleted of all cell types after sertoli cells and spermatogonia, indicating that they were not able to undergo meiosis and were subject to apoptosis. It was observed that 14-19% of P12 KO tubules and 12% adult KO tubules had abnormal metaphase due to premature exit of prophase I, causing apoptosis. RNA-seq from P12 KO and WT testis shows that there was a downregulation of genes involved in meiosis, indicating that YTHDC2 is needed for proper expression of meiotic genes.

M6A-IP-seq showed that at KO P8, there was an upregulation of highly methylated genes. Biochemical analysis showed that YTHDC2 has ATPase activity and RNA helicase activity (3'->5') that interacts with the exoribonuclease XRN1, which is also involved in transcript degradation. Future studies will be needed to determine how the interaction of these binding partners with other proteins affect the fate of the transcripts.

**Maternal effects**

**Oogenesis**

Ivanova et al. (2017) wanted to determine the effects of m6A reader YTHDF2. The authors generated an epitope-tagged, conditional YTHDF2 with GFP-His6-FLAG-HA which was shown to be expressed in a variety of tissues with the highest expression in the testis. Immunofluorescence was used to show that YTHDF2 is a cytoplasmic protein that was present in all stages of spermatogenesis, with the highest expression in the pachytene spermatocytes.
Additionally, YTHDF2 is expressed in the oocyte itself and the surrounding granulosa cells at all stages of folliculogenesis and during oocyte maturation. Ythdf2-null mice were viable but showed half the number of expected KO mice. Fertility tests showed that the KO males were fertile with normal histology, but that the females were completely infertile, but still successfully ovulated.

Because YTHDF2 was present in both somatic and germline cells, the authors used conditional genetics to delete YTHDF2 in oocytes, and not the somatic granulosa cells. Maternal inheritance of Ythdf2 is required for viable pups as revealed by crossing wild-type males and conditional knockout females resulting in no live births. However, the presence of the corpora lutea was noted indicating that ovulation had taken place. Superovulation of the conditional KO females showed that the numbers of MII oocytes were normal and that they were competent to be fertilized, indicating that YTHDF2 is not necessary for this process. Harvesting and examining the zygotes revealed that they were not able to progress past the 2-cell stage, and many were abnormal 2-cell zygotes showing uneven cytokinesis with some cells containing multiple nuclei and others enucleated, showing that YTHDF2 is required for early zygotic development.

Conditional knockout oocytes at the MII stage, the stage where the maternal transcript is first utilized, were shown to have a deregulated transcriptome with a 2-fold change or greater with 201 upregulated genes and 68 downregulated genes. At the germinal vesicle oocyte stage, where there is a mature maternal transcriptome, the transcriptome was normal, indicating that YTHDF2 is required to maintain gene quantities during oocyte maturation. This
trend to upregulate genes when YTHDF2 is knocked out is likely due to the fact that YTHDF2 mediates RNA degradation. The authors wanted to determine if the deregulated genes were direct targets of YTHDF2, so looked for enrichment of the consensus sequence preferred by YTHDF2 GACU/A and found that there was an enrichment of this sequence in the upregulated genes around the stop codon only, which also indicates the presence of an m$^6$A residue (42).

In summary, YTHDF2 is not required for meiosis to be completed successfully nor does it regulate the expression of meiosis-specific genes. It does however, regulate the maternal transcripts so that RNA is not degraded throughout oocyte growth, but is degraded throughout meiotic maturation, overall ensuring oocyte quality and competence.

**Maternal-to-zygotic Transition**

Maternally inherited RNAs are responsible for controlling early embryonic gene expression and are cleared during the maternal-to-zygotic transition (MZT), which takes place when the embryo controls gene expression. This clearance has been shown to take place by both maternal factors and zygotic gene products. Zhao et al. (2017) now show that one maternal factor for this clearance is the use of m$^6$A methylation, which affects the transcript stability. They first showed that during early embryogenesis, the m$^6$A reader, ythdf2, is expressed. When knocked out, the adult F1 generation did not have any noticeable defects. However, when mated, 70% of the F2 generation did not develop past one-cell. Further, maternally ythdf2 knockout embryos (male wild-type) showed a
developmental delay and a block in late G2/M phase of the cell cycle, both of which were able to be partially rescued with ythdf2 mRNA injections, indicating that this delay was derived from the maternal deletion of ythdf2 (43).

When m⁶A-seq was used on embryonic mRNAs, it was discovered that over 36% of maternal mRNAs contained m⁶A modifications. After 4 hours post fertilization (h.p.f.), there were two noticeable changes in the zygote RNA pool: first, the RNAs shifted from maternally derived to zygotically derived and second, the m⁶A population shifted from the maternal RNAs to the zygotic RNAs. It was also observed that the methylated maternal transcripts had higher levels of decay at 4 hpf compared to the non-methylated maternal transcripts, and the methylated transcripts were found to have a role in cell cycle regulation. In maternal ythdf2 knockout embryos, the maternal transcripts were found to be unregulated while the zygotic transcripts were lower in abundance, indicating that maternally derived ythdf2 is required for proper maternal transcript clearance and zygotic control of gene regulation in the MZT.

Bioinformatics analysis revealed that the gene targets of ythdf2 were regulation of the cell cycle progression and phosphorus metabolism, as well as gamete generation. The authors then used reporter genes injected into embryos to determine if m⁶A methylation can induce degradation with yphdf2 and found that during MZT, clearance of the m⁶A reporter gene in the wild-type embryos was faster than the unmethylated reporter gene. In contrast, the maternal ythdf2 knockout embryos had higher levels of the m⁶A reporter gene.
These results indicate that maternally derived Ythdf2 is required for maternal mRNA clearance during MZT and further supports the previously reported hypothesis that m⁶A targets mRNAs for degradation. Because several noticeable mRNA clearance defects were observed 4 h.p.f. in the maternal ythdf2 knockout embryos, recognition of m⁶A modifications on transcripts is required for the careful temporal gene regulation that occurs during development with the reversible m⁶A modification marking particular transcripts for degradation.

Figure 1-3. The effects of m⁶A on temporal gene regulation. (A) Correctly regulated m⁶A residues are marks for transcripts’ fate, marking the gene to be alternatively spliced or marking it for degradation. (B) Altered m⁶A levels, due to the deregulation of the components of the methyltransferase complex (WTAP, METTL3, or METTL14), the m⁶A demethylases (ALKBH5 or FTO), or the m⁶A reader proteins (YTHDC or YTHDF) leads to an unwanted phenotype. One of the common phenotypes is male infertility due to germ cell apoptosis and abnormal spermatozoa.
**Table 1.** Comprehensive list of journal articles correlating m\(^6\)A-associated proteins with fertility phenotypes.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Year</th>
<th>Journal</th>
<th>Gene</th>
<th>Gene Type</th>
<th>Male or Female</th>
<th>Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zheng et al.</td>
<td>2013</td>
<td>Molecular Cell</td>
<td>ALKBH5</td>
<td>eraser</td>
<td>Male</td>
<td>High levels of germ cell apoptosis</td>
</tr>
<tr>
<td>Tang et al.</td>
<td>2017</td>
<td>PNAS</td>
<td>ALKBH5</td>
<td>eraser</td>
<td>Male</td>
<td>Splicing and degradation deregulates spermatogenesis</td>
</tr>
<tr>
<td>Landfors et al.</td>
<td>2016</td>
<td>Fertility and Sterility</td>
<td>ALKBH5,  FTO</td>
<td>eraser</td>
<td>Male</td>
<td>FTO variants had altered sperm morphology</td>
</tr>
<tr>
<td>Yang et al.</td>
<td>2016</td>
<td>Scientific Reports</td>
<td>METTL3,  METTL14</td>
<td>writer</td>
<td>Male</td>
<td>Asthenozoospermia</td>
</tr>
<tr>
<td>Xu et al.</td>
<td>2017</td>
<td>Cell research</td>
<td>METTL3</td>
<td>writer</td>
<td>Male</td>
<td>Pachytene spermatocytes cannot complete meiosis</td>
</tr>
<tr>
<td>Xia et al.</td>
<td>2018</td>
<td>Genetics</td>
<td>METTL3</td>
<td>writer</td>
<td>Male and Female</td>
<td>Subfertility and gametes cannot mature</td>
</tr>
<tr>
<td>Hsu et al.</td>
<td>2017</td>
<td>Molecular Cell</td>
<td>YTHDC2</td>
<td>reader</td>
<td>Male and Female</td>
<td>High levels of germ cell apoptosis</td>
</tr>
<tr>
<td>Wojtas et al.</td>
<td>2017</td>
<td>Molecular Cell</td>
<td>YTHDC2</td>
<td>reader</td>
<td>Male</td>
<td>Spermatocytes unable to complete meiosis</td>
</tr>
<tr>
<td>Ivanova et al.</td>
<td>2017</td>
<td>Molecular Cell</td>
<td>YTHDF2</td>
<td>reader</td>
<td>Female</td>
<td>Ovulated oocytes could not progress past 2-cell stage</td>
</tr>
<tr>
<td>Zhao et al.</td>
<td>2017</td>
<td>Nature</td>
<td>YTHDF2</td>
<td>reader</td>
<td>Female</td>
<td>Unable to clear maternal transcripts during MZT</td>
</tr>
</tbody>
</table>

**Conclusion**

It has been generally accepted that m\(^6\)A is vital for mRNA regulation, however its exact role has remained unclear until recently. Its common location in the 3’ UTR, a portion of the mRNA famous for its roles in RNA regulation by controlling localization and transcript stability, exposes the potential of m\(^6\)A
having a regulatory role. The discovery of the plausible link between miRNA binding sites and m^6^A residues by Meyer et al., (2012) is further evidence of a possible regulatory role. This is especially true considering miRNAs have been implicated in causing both RNA degradation and repression of translation (44). The most promising evidence of a regulatory role was shown with the initial discovery of a possible role in alternative splicing, with the alternatively spliced genes being enriched with m^6^A modifications while the single isoform genes were less methylated (5). This hypothesis was further confirmed with the recent publication by Tang et al., 2017 with the finding that the m^6^A locations correlated with the splicing locations. m^6^A modifications that were unable to be properly erased due to ALKBH5 depletion caused an increase in splicing and subsequent degradation leading to massive germ cell apoptosis and infertility (45).

Additionally, another group implicated m^6^A with the use of poly A sites, where those genes that used proximal poly A sites had lower m^6^A levels, while those that used distal sites had higher levels (11), indicating a role in inhibiting polyadenylation site usage signaling to other proteins the fate of a particular transcript.

A major potential regulatory role includes RNA degradation itself (Figure 1-3a), where an m^6^A mark may target a transcript for degradation. A list of the potential m^6^A regulatory roles and their effects on infertility are provided in Table 1. First, it was discovered that when METTL3 is knocked out, embryonic stem cells maintain a state of naïve pluripotency, and were unable to cavitate even when differentiation was induced. Further, the m^6^A methylated transcripts had
shorter half-lives, which increased upon deletion of METTL3. These knockout cells also had an increase in translation indicating that increased levels of m\(^6\)A decreased translation efficiency (22). Further evidence that m\(^6\)A may be a mark for degradation is that in the maternal-to-zygotic transition, ythdf2 was expressed in early embryogenesis to allow for maternal clearance by recognizing the modification and marking the maternal transcripts for degradation. YTHDF2 was shown to colocalize with mRNA degradation sites while YTHDF2 knockdown cells experienced an increase in the mRNA lifespan by 30% (43), indicating that m\(^6\)A marks decrease transcript stability mediated primarily by YTHDF2, if not other unknown protein players.

The modifications are recognized by the reader proteins YTHDC1 and YTHDF2, which ultimately determine the fate of the RNA. m\(^6\)A marks are also needed to regulate global transcript expression through recruitment of reader proteins for the degradation of transcripts to alter global gene expression. For example, METTL3 knockdown cells inhibited gene silencing by Xist (30), while m\(^6\)A was required to recruit the reader protein, YTHDC1, to induce the gene silencing. Another example of altered gene expression is shown throughout development, specifically in newborn piglets. Evidence for temporal gene regulation by m\(^6\)A was shown when the highest differential methylation patterns were found in the newborn pig, over the suckling and adult stages (33), indicating that m\(^6\)A is required early in postnatal development for proper maturation. Recently, m\(^6\)A has been implicated in cap-independent translation that gets induced when cap-dependent translation is inhibited (46). Because cap-
dependent translation is first in effect after the nucleus and the 5’ mRNA cap are formed, this indicates that this alternative method of translation is important in early development, although it is not as efficient as cap-dependent translation.

As m$^6$A marks the transcripts for degradation, and differential methylation patterns were found in newborn piglets when mRNA expression pattern is also highly differential, it is tempting to speculate that m$^6$A marks are required for temporal gene regulation, as a way to degrade transcripts that are no longer needed. Further supporting this hypothesis is that Ythdf2 knockout embryos experienced a developmental delay resulting from a block in the G$_2$/M phases of the cell cycle (43) highlighting that dysregulation of this process affects gene expression during development.

Because of this vital role in RNA regulation, altered m$^6$A levels lead to incorrectly expressed transcripts, which then lead to an altered phenotype (Figure 1-3b). Many of the enzymes that are responsible for m$^6$A regulation are most highly expressed in the testis, such as ALKBH5 (27), METTL3 (39), and YTHDC2 (40, 41). Maternal inheritance of YTHDF2 is vital for transcript turnover in the maternal to zygotic transition (43) and oocyte competence for female fertility (42). Others, such as FTO and the methyltransferases METTL3 and METTL14, have been shown to be expressed in the testes and sperm, respectively (35, 36). Therefore, abnormal expression of these methylases and demethylases lead to altered m$^6$A levels, and consequently change the fate of the transcripts leading to an unwanted phenotype, specifically male infertility. Due to incorrect splicing and high levels of transcript degradation, these altered
m^6A levels lead to increased apoptosis in the developing germ cells, likely due to inability for the germ cells to pass the meiotic checkpoint. Consequently, there is an inability to complete spermiogenesis, and those few cells that do complete spermatogenesis have altered flagellum morphology and motility due to incorrect transcript expression. Not being able to undergo meiotic maturation or the maternal to zygotic transition in the absence of YTHDF2, and the inability of spermatocytes to correctly undergo meiosis in the absence of ALKBH5, METTL3, YTHDC2 suggest that m^6A modifications are crucial for rapid transcript turnover during transitions in development or maturation. When absent, there are very few germ cells that become fully mature, and those germ cells that do survive into maturation are otherwise altered, in motility, morphology, or both. All of these phenotypes are common characteristics of infertility (47).

Taken together, we propose that m^6A modifications are required for highly dynamic developmental processes such as embryogenesis, oogenesis, and spermiogenesis to allow for proper transcript turnover through splicing and degradation. This modification is ideal for this purpose as it is reversible, able to be added to transcripts early in development, and marks transcripts rapidly. However, it is now required to understand the fate of the transcripts that do not have proper methylation levels due to lack of m^6A writer proteins. This has proven difficult to study due to the lethality of Mettl3 knockout models in early development. Further, it is also crucial to investigate how m^6A affects all aspects of translation, specifically in the ribonucleoprotein (RNP) and polysome. For instance, are m^6A modifications also involved in temporal localization of
transcripts to particular organelles for spatiotemporal regulation? If m\textsuperscript{6}A-translation is truly cap-independent and internal ribosomal entry site (IRES) independent as was suggested by Coots et al., 2017, then what is the complex required for initiation of m\textsuperscript{6}A-dependent translation? Until these questions are answered, we will not fully understand the depth and complexity of m\textsuperscript{6}A regulation and its effects on infertility phenotypes and the possible ways to counteract their effects.
Introduction to epigenetic transgenerational inheritance

Epigenetic transgenerational inheritance of noncoding RNAs

Noncoding RNAs

Acquired traits are those traits that have been gained during the lifetime of the organism, due to external exposures and events, such as mental stress, alteration of diet, or exposure to toxic substances (48). New evidence suggests that these acquired traits can be passed down paternally to the offspring through the father’s sperm. This phenomenon, deemed paternal epigenetic inheritance, is now being widely studied at the RNA level. It has been well established that there are two waves of epigenetic reprogramming that occurs during embryonic development: the first occurs during blastomere implantation and erases the parental epigenetic marks, while the second occurs in rats at E8.5-E14.5 when the primordial germ cells first appear (49). These reprogramming events have always been supposed to eliminate any acquired traits from the paternal side, making these traits unable to pass to the offspring. The continued transmittance from generation to generation (discussed below), makes it likely that there are some acquired traits that are successfully able to escape these reprogramming events.
There are many different kinds of small, noncoding RNAs in the sperm head and large and small noncoding RNAs have been estimated to exceed 20,000 in a single spermatozoa (50, 51). Some of the small, noncoding RNAs in a spermatozoa include PIWI-interacting RNAs (piRNAs) that are primarily present in the spermatogonium through the round spermatids and are still present, but to a much lesser extent in elongating and maturing spermatozoa. This class of small RNAs is believed to regulate germline transcript stability. Others, such as miRNA and siRNA, are believed to regulate spermatogenesis at the level of translation with miRNAs thought to work by binding to the 3’UTR and affecting RNA metabolism. In addition, tsRNAs (derived from the 5’ end of tRNAs) are also present throughout spermatogenesis but are at very low levels during the spermatogonia stage through the round stage (52, 53). They increase significantly during elongation and through epididymal maturation. In fact, tsRNAs are the most abundant of the small, noncoding RNAs in the paternal germline and have complementary sequences to transposable element gene promoters. Other known RNAs in the sperm include: mitochondrial genome-encoded small RNAs (mitosRNAs) and long noncoding RNAs (IncRNAs) that possibly alter the chromatin structure. Extracellular vesicles are another source of small RNAs that can fuse with sperm cells to allow regulatory effects. A new complex has been discovered that allow RNAs to be mobile by being taken up by somatic cells and possibly germ cells: G-quadruplex conformation (guanine-rich nucleic acid that forms tetrads) (53). There are three mechanisms that can be used to alter a phenotype: RNA editing, chromatin modification, and RNA
methyltransferases. They have been shown to transmit the paternal phenotype such as obesity, glucose intolerance, and stress dysregulation to the offspring.

Long, noncoding RNAs are those RNAs that were once referred to as “junk” in the genome because they do not code for proteins, and yet are widely transcribed. They have been hypothesized to regulate transcripts both in cis or trans, and to regulate proteins (49) or RNA itself (54). It is believed that these regulatory elements are not sequence-specific, but rather localize in the nucleus at low levels to act on enhancers or promoters and can regulate local expression in three ways: the RNA can recruit regulatory factors, the process of producing the lncRNA in a sequence-independent manner gives the RNA regulatory functions, or its regulatory abilities depend on the DNA elements at the promoter. In addition to cis-acting transcripts, lncRNAs can act in trans by regulating chromatin states, directing nuclear organization, or by interacting directly with other RNA and proteins (49). Perhaps the most well-known lncRNA is Xist (X-inactive specific transcript), a cis-acting lncRNA that silences one of the two female X chromosomes during early embryonic development. Xist has been shown to be regulated by m^6^A by both the methyltransferase, METTL3, and the reader YTHDC1 (30). When the methyltransferase was knocked down, gene expression was not repressed. Similarly, knockdown of YTHDC1 prevented gene silencing, indicating that long, noncoding RNAs are subject to tight epigenetic regulation.
Transgenerational Inheritance

Transgenerational inheritance refers to germline transmission of epigenetic modifications causing a phenotype across multiple generations after an initial, ancestral (F0) exposure to an external agent. These external agents can vary from environmental toxins, irradiation, and chemotherapy and can be transmitted either through a genetic or epigenetic mechanism such as DNA methylation, noncoding RNA changes, or histone modifications. A pregnant female's (F0) exposure also exposes the F1 fetus and the F1 fetus's germline (F2), with the first non-directly exposed generation is the F3 generation, making any inherited epigenetic changes transgenerational (Figure 1-4). A non-pregnant female's or male's exposure (F0) also exposes their germline (F1), making F2 the transgenerational generation.

The first demonstrated instance of transgenerational effects was in 2002 with the use of ionizing radiation in F0 male mice (55). It was discovered that the original irradiation caused genetic mutations that did not decrease in subsequent generations when the irradiation affected the premeiotic spermatogonia, but no effects were observed if the postmeiotic spermatogonia were exposed. The F1-F3 male and female offspring had similarly high mutation rates as the original exposed male, indicating that the genetic instability seen must stem from an epigenetic cause: a genetic cause would have subsequent generations with lower levels of mutation due to Mendelian segregation. Since then, many exposures have been studied for epigenetic transgenerational inheritance. To
date, a wide variety of environmental toxicants have been documented to cause epigenetic transgenerational inheritance, ranging from the herbicide atrazine in causing hyperactive behavior (56) to dioxin causing prostate and polycystic ovarian disease (57) and to the pesticide DDT (dichlorodiphenyltrichloroethane) in causing heritable sperm epimutations (58). By far, the agricultural fungicide vinclozolin has been implicated as the toxicant with the widest variety of disease susceptibility including: testis disease (59), prostate disease, kidney disease, tumor development (60), and behavioral changes (61).

Anway et al. (2005) were the first to report epigenetic transgenerational inheritance after initial exposure to vinclozolin, a known endocrine disruptor and agricultural fungicide (62). Androgen and estrogen receptors are present on sertoli cells and can be affected by androgen substances during gonadal sex determination (E12-E15). Abnormal androgen substances can cause an unintended germ line reprogramming that can lead to epigenetic inheritance.

Anway et al. (2005) determined the transgenerational effects of vinclozolin exposure on gestating females by injecting pregnant female rats with vinclozolin during E8 and E15. F1 rats from different litters were mated for 4 generations and adult male testis and caudal epididymal sperm from each generation were collected. In all four generations levels of vinclozolin spermatogenic cell apoptosis were double that of the control, in 90% of males, while sperm counts and motility were reduced by 20% and 30%, respectively. Methoxychlor is a pesticide with anti-androgenic properties and was shown to have the same apoptotic effect as vinclozolin, but only on the F1 and F2 generations, indicating
that like vinclozolin, it can cause transgenerational defects. To determine if these transgenerational defects were transmitted through the male germline, F2 vinclozolin males were outcrossed with wild-type females, and it was discovered that the male offspring had the same phenotypes as their fathers, indicating that the male germ line is responsible for transmitting transgenerational phenotypes caused by endocrine disruptors. It should be noted that a reverse vinclozolin outcross with F2 vinclozolin females and wild-type males showed no significant difference.

Vinclozolin lineage rats from F1 to F4 were examined for signs of infertility, and it was found that after P90, 8% of males developed complete infertility with small testis and abnormal spermatogenesis compared to a 0% incidence for the control rats. The >90% frequency of the apoptotic spermatogenic cells and decreased sperm numbers and motility in four generations suggests that this is an epigenetic, rather than genetic mechanism, due to the lack of phenotypic dilution in subsequent generations. A series of methylation sensitive restriction digests and bisulfite sequencing showed that approximately 25% of the vinclozolin lineage sperm samples had both hyper- and hypomethylation levels compared to the control. Together, these results indicate that high levels of exposure to endocrine disruptors during gonadal sex determination (E8-E15) can cause an epigenetically induced phenotype that passes through at least four generations through the male germ line.

Due to the previous discovery that the endocrine-disruptor vinclozolin could cause transgenerational infertility effects on male rats, Guerrero-Bosagna
et al., 2010 decided to examine whether the promoter regions were altered in the sperm from rats from either the vinclozolin or control lineages (63). Using immunoprecipitation and MeDIP-Chip (promoter tiling microarray chip hybridizations) they discovered that 52 differential methylation regions in 48 sperm promoters had altered methylation levels either increasing or decreasing in the F3 vinclozolin lineage sperm. In addition, a consensus sequence motif was found in 59% of imprinted gene promoters, compared to 16.8% of random promoters. Promoters were also shown to have an increase in long terminal repeats (LTRs) in 13% of cases, specifically in the endogenous retrovirus-like (ERV) class I and class II (IAP) elements, indicating that transgenerational inheritance of promoters with altered DNA methylation also have an enrichment for ERV class I and II elements.
Figure 1-4. Gestating female’s exposure to endocrine-disruptors causes epigenetic transgenerational inheritance in male pups. When exposed during gonadal sex determination, the male pup’s sperm is also affected this leading to transgenerational inheritance. Adapted from Nilsson and Skinner 2015 (64).
Mode of Paternal Epigenetic Transmission

Vinclozolin has been shown to cause epigenetic transgenerational inheritance in the unexposed F3 and F4 generations with initial in utero exposure (F0 generation) (62). This inheritance was shown to be caused partially by DNA methylation, but not by genetic inheritance or mutations. Vinclozolin is a known endocrine disruptor that has been previously shown to cause kidney abnormalities and high levels of spermatogenic cell death when exposed during gonadal sex determination (62). Because the vinclozolin-derived phenotypes are primarily caused by sperm, Schuster et al. (2016) examined the F3 generation sperm that had been derived from either the vinclozolin or control lineage. They specifically looked at the small, noncoding RNAs including: miRNAs, piRNAs, tsRNAs, mitochondrial genome-encoded small RNAs. A differential expression analysis software (DESeq2) was used to determine that the vinclozolin lineage sperm had 222 sncRNAs with altered expression levels from the control with 8% difference in the miRNA class, 16% for piRNA, and 14% for rRNA-derived sncRNAs. mitosRNAs showed the most drastic difference with 79% upregulation. The tRNA reads were aligned based off of 5’ or 3’ tRNA halves, and it was found that the 5’halves and tRNA fragments from the 5’end (tRF-5s) were upregulated while those from the 3’ halves were downregulated in the vinclozolin sperm.

Gene ontology analysis was used to determine the function of specific genes that matched to multiple sncRNAs. Specifically for the miRNA class, it was noted that only miRNA targets that were redundantly downregulated had significant functional results being involved in apoptosis, transcription regulation,
axon guidance, and kinase activity. Genes that were redundantly targeted in the 5’ UTR by 5’ halves were also enriched in apoptosis and neuronal development, suggesting a redundancy in the miRNAs and 5’ halves derived from vinclozolin-induced expression alteration.

Next, the authors correlated the DNA methylation regions (DMR) with the differentially expressed sncRNAs. It was noted that of the differentially expressed sncRNA, 76% of the miRNAs and 88% of the 5’ halves were located within 5 Mb of the DMR. The 5’ halves derived from the F3 vinclozolin lineage sperm showed a significant increase in the actual matches versus the predicted target, suggesting that this class of sncRNAs has the ability to post-transcriptionally regulate genes in close proximity to DMR. The transgenerational DMRs in primordial germ cells (PGCs) isolated from E13 testis also showed a strong correlation of the 5’ half targets with genes proximal to the DMRs, suggesting that there is a connection between transgenerationally inherited DMRs in PGCs and sperm with altered sncRNA expression. Taken together, these results indicate that ancestral exposure to vinclozolin causes an upregulation in sperm 5’ halves that are the carriers of epigenetic transgenerational inheritance responsible for the associated phenotype (65).
Conclusion

Despite numerous studies correlating a single environmental exposure during the critical developmental time of pregnancy to a disease phenotype later in life, it is still unknown how the exposed germline of the F1 offspring would affect all somatic tissues of the F2 generation and beyond sired by the F2 males. Both long, noncoding RNA and mRNA have been shown to be dynamically regulated throughout spermatogenesis with spermatogonia having the highest expression of both lncRNA and mRNA, and pachytene spermatocytes having the least (66), making the timing of the exposure important in determining the extent of the epigenetic alterations. As the exposed germline of the F1 offspring will be a genetic source for all F2 tissues, which seem to escape reprogramming during embryonic development, all F2 tissues will have the affected epigenome of the father’s sperm. The epigenome ultimately controls the organ’s fate and therefore may be responsible for many of the diseases that occur during adulthood (67). It is therefore crucial to determine how different transmitted epigenetic modifications from ancestral exposures will affect the somatic tissues of future generations.
References


Chapter 2: ALKBH5-dependent m6A demethylation controls splicing and stability of long 3’UTR mRNAs in male germ cells

Chong Tang\textsuperscript{a,1}, Rachel Klukovich\textsuperscript{a,1}, Hongying Peng\textsuperscript{a}, Zhuqing Wang\textsuperscript{a}, Tian Yu\textsuperscript{a}, Ying Zhang\textsuperscript{a}, Huili Zheng\textsuperscript{a}, Arne Klungland\textsuperscript{b} and Wei Yan\textsuperscript{a,c,2}

\textsuperscript{a}Department of Physiology and Cell Biology, University of Nevada, Reno School of Medicine, 1664 North Virginia Street, MS575, Reno, NV 89557, USA; \textsuperscript{b}Department of Microbiology, Oslo University Hospital, Rikshospitalet, 0027 Oslo, Norway and Department of Molecular Medicine, Institute of Basic Medical Sciences, University of Oslo, 0317 Oslo, Norway; \textsuperscript{c}Department of Biology, University of Nevada, Reno, 1664 North Virginia Street, MS575, Reno, NV 89557, USA

\textsuperscript{1}C.T. and R.K. contributed equally to this work.

This chapter can be found in: \textit{PNAS} (2017), 115 (2) E325-E333.

\textsuperscript{2}To whom correspondence should be addressed:

Wei Yan M.D., Ph.D.
\textit{University of Nevada Reno Foundation Professor}
Department of Physiology and Cell Biology
University of Nevada School of Medicine
Center for Molecular Medicine, Room 207B
1664 North Virginia Street, MS/0575
Reno, NV 89557
Tel: 775 784 7765
Fax: 775 784 4362
Email: wyan@med.unr.edu
Abstract

N6-methyladenosine (m6A) represents one of the most common RNA modifications in eukaryotes. Specific m6A writer, eraser and reader proteins have been identified. As an m6A eraser, ALKBH5 specifically removes m6A from target mRNAs and inactivation of Alkbh5 leads to male infertility in mice. However, the underlying molecular mechanism remains unknown. Here, we report that ALKBH5-mediated m6A erasure in the nuclei of spermatocytes and round spermatids is essential for correct splicing and the production of longer 3'UTR mRNAs, and failure to do so leads to aberrant splicing and production of shorter transcripts with elevated levels of m6A that are rapidly degraded. Our study identified reversible m6A modification as a critical mechanism of post-transcriptional control of mRNA fate in late meiotic and haploid spermatogenic cells.

Keywords: RNA methylation | alternative splicing | mRNA stability | 3'UTR shortening | fertility
**Significance Statement**

N6-methyladenosine (m6A) represents one of the most common RNA modifications. Biochemical analyses have identified ALKBH5 as an eraser of m6A. The present study represents the first molecular characterization of the Alkbh5 knockout mouse model. Our data associate m6A erasure with mRNA length control. Specifically, proper m6A demethylation is required for correct splicing and selective degradation of longer 3'UTR transcripts, which are abundant in mitotic and meiotic male germ cells, but these longer 3'UTR transcripts become rapidly degraded in the haploid male germ cells. Aberrant m6A levels in spermatogenic cells are incompatible with normal spermatogenesis and male fertility.
Similar to DNA and histones, mRNAs and large noncoding RNAs can also be chemically modified (1). Over 100 various chemical modifications have been discovered in RNAs (2, 3), and N\(^6\)-methyladenosine (m6A) represents the most abundant (~3-5 m6A sites per mRNA) and also the most studied in eukaryotes (4, 5). Several active components of the m6A methyltransferase complex, including METTL3, METTL4 and WTAP, have been the regarded writer proteins of m6A (6-8). Interestingly, the m6A modification is reversible in mammalian cells, and it can be removed by two m6A demethylases, ALKBH5 and FTO (9, 10). A number of RNA-binding proteins recognize m6A on RNAs and interactions between m6A sites of mRNAs and these m6A readers exert a wide variety of effects on the mRNA fate (8, 11). The functions of these m6A writers, erasers and readers were initially established mainly in cultured cells in vitro (12). Recently, a flurry of reports has been published in which the physiological roles of these m6A regulators are demonstrated through genetic ablation in animals of various species (12-21). Inactivation of Ythdf2, an m6A reader protein, in either zebrafish(16) or mice (20) causes female infertility by affecting maternal transcript turnover and consequently disruptions of maternal-to-zygotic transition during early embryonic development. Global Mettl3 inactivation in mice has revealed critical roles of m6A in embryonic stem cell differentiation (21) and spermatogonial stem cell differentiation (13). Mutant mice lacking Ythdc2, one of the m6A reader proteins, also display a meiotic arrest phenotype in both males and females (19). FTO KO mice are fertile but display an obesity phenotype (18). As another known m6A eraser protein, ALKBH5 has also been shown to be
involved in glioblastoma (22), and to play an essential for spermatogenesis and male fertility (23) although the underlying mechanism remains unknown.

Of great interest, these studies point to an essential role of m6A in gametogenesis and fertility control, which is in agreement with earlier studies using model organisms (24). This is not surprising given that germ cells, both male and female, all display multi-layered, sophisticated post-transcriptional regulation of gene expression (25, 26). In females, a large number of maternal transcripts are synthesized in developing oocytes toward the end of folliculogenesis, and these maternal transcripts function before or shortly after zygotic genome activation and are ultimately eliminated by the 2-cell stage of early embryonic development (27-29). The maternal transcripts are pre-synthesized in developing oocytes, but they are subjected to translational suppression so that they remain stabilized for an extended period of time until being translated shortly after fertilization (27-29). A similar phenomenon of extended delay in translation also occurs during haploid male germ cell development (also termed spermiogenesis) in the testis (25). At the beginning of spermatid elongation (step 9 in mice), nuclear condensation starts and histones are rapidly replaced by protamines (30, 31). Consequently, the transcriptional machinery is completely shut down. To provide proteins for the final seven steps (steps9-16) of sperm assembly, mRNAs have to be pre-made in late pachytene spermatocytes and round spermatids prior to the onset of nuclear condensation in elongating spermatids. These pre-synthesized mRNAs have been shown to be compartmentalized into ribonuclear protein particles (RNPs) and physically
segregated from the translational machinery in the cytoplasm (25, 30, 31). RNPs exist as nuage (also called inter-mitochondrial cements) in spermatocytes and chromatoid bodies in round spermatids (32, 33). When proteins are needed for final assembly of spermatozoa, those RNP-localized mRNAs are released and loaded onto polysomes for translation. Recent reports have revealed that RNP enrichment of mRNAs is a dynamic process, through which the overall length of 3’UTRs in RNPs become increasingly shorter compared to polysome-enriched mRNAs from late pachytene spermatocytes to round and elongating spermatids (34). In other words, longer 3’UTR transcripts, which are mainly transcribed during mitotic and meiotic phases of spermatogenesis, are no longer needed once spermatogenesis enters the haploid phase; thus, these longer 3’UTR transcripts gradually come out of RNPs for the final round of translation and/or degradation. In contrast, the shorter 3’UTR transcripts, which are mostly those required for late spermiogenesis, enter RNPs for stabilization and translational suppression, and become increasingly enriched from late pachytene spermatocytes to round spermatids (34). The continuous shuffling of longer 3’UTR mRNAs out of RNPs and of shorter 3’UTR mRNAs into RNPs results in that the overall 3’UTR length of the entire mRNA transcriptome in elongating spermatids becomes shorter and shorter. While this global transcriptomic shortening phenomenon has been noticed for a long time (35-37), the underlying mechanism was only unveiled recently. It turns out that UPF1-3, the three mRNA quality control proteins known to function in the nonsense mRNA decay (NMD) pathway, function to selectively degrade longer 3’UTR transcripts, leading
to relative enrichment of shorter 3'UTR mRNAs (35, 38, 39). It is believed that shorter 3'UTRs bind fewer regulatory factors, e.g., RNA-binding proteins and small RNAs, thus enhancing translational efficiency although the stability of these shorter transcripts may be compromised. Clearly, the post-transcriptional regulation of gene expression during late spermiogenesis features two events: 1) compartmentalization of mRNAs into RNPs for enhanced stability and suppressed translation, and 2) global shortening of 3'UTRs to enhance translational efficacy and quick turnover through selective degradation of longer 3'UTR transcripts.

Given that m6A is predominantly enriched in 3'UTRs close to the stop codon and numerous in vitro and in vivo studies have suggested a role of m6A in the control of mRNA dynamics (7, 8, 11), it is highly likely disruptions of spermatogenesis in mice lacking writer (Mettl3) (13), eraser (Alkbh5) (23) and reader (Ythdc2) (19) result from dysregulated mRNA fate control. As an m6A eraser, ALBKH5 has been shown to play an essential role in spermatogenesis (23). Male mice lacking Alkbh5 are completely infertile due to severe germ cell depletion and oligoasthenoteratozoospermia (OAT) (23). However, the underlying mechanism of spermatogenic disruptions in Alkbh5 KO males remains largely unknown. More importantly, data on the effects of ALKBH5 ablation on the mRNA m6A profiles in spermatogenic cells are lacking. To reveal the function of ALKBH5-dependent m6A during spermatogenesis, we purified three types of spermatogenic cell types from wild-type control and Alkbh5 KO testes and analyzed the changes in mRNA transcriptome and m6A profiles using RNA-
Seq and m6A RNA immunoprecipitation sequencing (m6A-RIP-Seq), respectively. Our data suggest that proper m6A erasure is required for correct splicing of longer 3'UTR transcripts in the nucleus, and that m6A enrichment in 3'UTRs of mRNAs correlates with enhanced degradation in the cytoplasm. ALKBH5-dependent m6A is required for meiotic and haploid phases of spermatogenesis by controlling both splicing and stability of mRNAs.

**Results**

**ALKBH5 is required for the late meiotic and haploid phases of spermatogenesis.** The testis expresses the highest levels of *Alkbh5* mRNAs in mice, and global inactivation of *Alkbh5* (herein called *Alkbh5* KO mice) leads to neither discernable development defects nor adult diseases except for male infertility, suggesting an essential role of *Alkbh5* in spermatogenesis and male fertility (23). To study the causes of male infertility in *Alkbh5* KO mice, we first examined testicular development at both gross and histological levels (Fig. 2-1). Adult *Alkbh5* KO testes were ~half of the size of WT controls (Fig. 2-1A), and the testis weight was similar at postnatal day 14 (P14) between WT and KO mice, but the KO testes became significantly smaller compared to the WT controls at P21 and thereafter (Fig. 2-1B). The most advanced spermatogenic cells are mid-pachytene spermatocytes and round spermatids in P14 and P21 testes, respectively. Therefore, the decreased testis size between P14-P21 and thereafter likely results from depletion of late pachytene spermatocytes and spermatids in the KO testes. Indeed, histological examination revealed that there
was a delay in meiotic progress in KO testes, as compared to WT controls at P14 and P21 (SI Appendix, Fig. S1). For example, mid-pachytene spermatocytes were present in most of the seminiferous tubules in WT testes at P14, but mostly absent in KO testes. At P21, round spermatids were detected in WT testes, but absent in the KO ones (SI Appendix, Fig. S1). The adult KO testes contained numerous vacuoles of variable sizes and the seminiferous epithelium was disorganized with much fewer spermatocytes and spermatids, suggesting active germ cell depletion (Fig. 2-1C; SI Appendix, Fig. S1). Supporting this notion, numerous depleted germ cells, mainly spermatids, were observed in the cauda epididymis (Fig. 2-1D). TUNEL analyses detected a much greater number of apoptotic germ cells in KO testes starting from P14 and the enhanced germ cell apoptosis persisted into adulthood, as compared to WT controls (Fig. 2-1E, F; SI Appendix, Fig. S2). Occasionally, some spermatozoa, although very few, were detected in the KO cauda epididymis; these KO sperm were all deformed showing a wide variety of structural abnormalities (Fig. 2-1G, H) without motility (Fig. 2-1I). Given that the male KO mice are otherwise completely healthy, \textit{Alkbh5} appears to be indispensable only for spermatogenesis and male fertility. An essential role of \textit{Alkbh5} in the testis is consistent with the fact that the testis is the organ with the highest expression levels of \textit{Alkbh5} (23).

To determine the stage-specific role of ALKBH5, we examined the testicular localization of ALKBH5 using immunofluorescence. Specificity of the anti-ALKBH5 antibody used was validated by Western blots, in which specific bands were detected in WT, but absent in the KO testes (Fig. 2-1J). ALKBH5
was localized in the nuclei of almost all testicular cell types except elongating and elongated spermatids (steps 9-16) (Fig. 2-1K, L; SI Appendix, Fig. S3). Among spermatogenic cells, ALKBH5 appeared to be the most abundant in spermatocytes; medium levels were detected in spermatogonia and lower levels in round spermatids (Fig. 2-1K, L; SI Appendix, Fig. S3). Consistent with an earlier report (23), ALKBH5 was partially co-localized with SC35, a marker for nuclear speckles (Fig. 2-1K, L; SI Appendix, Fig. S3). The localization pattern of ALKBH5 is consistent with the testicular phenotype observed, i.e., depletion of pachytene spermatocytes and round spermatids, in which the ALKBH5 is abundantly expressed. The phenotype of Alkbh5 KO male mice suggests an essential role of ALKBH5 in the meiotic and haploid phases of spermatogenesis.

**m6A tends to mark the 3’UTRs of longer mRNAs that are destined to be degraded during spermiogenesis.** The biggest change in the overall transcriptome between the late meiotic and haploid phases of spermatogenesis lies in the global shortening of mRNAs (34, 35). The global transcriptomic shortening allows for efficient translation and quick mRNA/protein turnover during late spermiogenesis (34, 35), and was once believed to be achieved through increased production of transcripts with shorter 3’UTRs via alternative polyadenylation (36, 37). However, recent data suggest that enhanced degradation of longer mRNAs (usually with longer 3’UTRs) by the UPF1-3-mediated, non-canonical NMD pathway contributes significantly to the global transcriptomic shortening process (35, 38-41). We purified pachytene
spermatocytes, round and elongating spermatids from the wild-type (WT) adult mouse testes and performed RNA-Seq. As expected, a global shortening trend of all transcripts was observed from round and elongating spermatids (Fig. 2-2A). As reported in previous studies, larger transcripts, which are mainly those with longer 3'UTRs expressed in spermatocytes and round spermatids, are downregulated through degradation in elongating/elongated spermatids; consequently, the shorter mRNAs, which are mostly needed for sperm assembly in the final several steps of spermiogenesis, get increasingly enriched in elongating/elongated spermatids during late spermiogenesis (34, 35). In this study, we found that ~60% of the longer 3'UTR (>1,500nt) transcripts enriched in round spermatids were significantly downregulated in the elongating spermatids, and levels of ~70% of the shorter 3'UTR (<500nt) transcripts were steadily increased. Given that m6A has been implicated in the control of mRNA stability (7, 8, 11), we explored whether m6A is involved in the degradation of those longer 3'UTR transcripts during spermiogenesis.

To map m6A sites in mRNAs, we established and optimized a m6A RIP-Seq protocol based on previous reports (42, 43). We also adopted the latest method reported for bioinformatic analyses of m6A RIP-Seq data (16). We validated our method by performing m6A-RIP-Seq using HEK293 cells and our method detected ~70% of the m6A sites reported previously using high-resolution m6A CLIP-Seq (42, 43) (SI Appendix, Fig. S4). To explore the relationship between 3'UTR length and m6A levels and sites, we selected the top 30 transcripts with the longest 3'UTRs (>3,000nt) that were enriched in round
spermatids, but drastically downregulated in elongating spermatids, and compared them with 200 randomly chosen, shorter 3'UTR (<500nt) transcripts showing increasing abundance from round to elongating spermatids (Fig. 2-2B). Interestingly, we found that the longer 3'UTR (>3,000nt) transcripts contained much higher levels of m6A in their 3'UTRs close to the stop codon, as compared to the shorter 3'UTR (<500nt) mRNAs (Fig. 2-2B). These data suggest that m6A predominantly marks the longer 3'UTR transcripts that are downregulated in elongating spermatids during spermatogenesis. It is noteworthy that m6A levels in those long 3'UTR transcripts appeared to be higher in elongating spermatids than in pachytene spermatocytes and round spermatids (Fig. 2-2B, peak 1 in lower left panel). Indeed, two representative genes (Uhmk1 and Traf3ip) showed much higher m6A levels in elongating spermatids than in round spermatids (Fig. 2-2C, D). Given that m6A is mainly present in the 3'UTRs of longer transcripts (Fig. 2-2B), the increased m6A levels may serve as a signal for decreased stability in longer 3'UTR transcripts when round spermatids develop into elongating and elongated spermatids. Consistent with published data (8, 42, 43), we also found the conserved m6A motifs [G/A/U][G>A]m6AC[U>A>C] (Fig. 2-2E). Together, these data suggest that m6A tends to mark the transcripts with longer 3'UTRs that are destined to be degraded when spermiogenesis progresses from round to elongating spermatids.

**m6A controls correct splicing of long 3'UTR transcripts in spermatocytes and round spermatids.** We performed RNA-Seq and m6A-RIP-Seq using
pachytene spermatocytes, round, and elongating spermatids purified from Alkbh5 KO and WT adult testes. A significant decrease in transcript length was observed in all three KO spermatogenic cell types, as compared to the WT controls (Fig. 2-3A). We chose the same 30 longer 3'UTR (>3,000nt) transcripts, which are mainly synthesized in round spermatids and degraded in elongating spermatids under physiological conditions (Fig. 2-2B), and analyzed their expression levels. Interestingly, almost all 30 longer 3'UTR mRNAs analyzed were significantly downregulated in the KO cells (Fig. 2-3B). Meanwhile, we noticed that the shorter isoforms of these 30 long transcripts were all upregulated (Fig. 2-3C). Gene ontology (GO) term enrichment analyses revealed that these longer 3'UTR mRNAs were almost all involved in the regulation of spermatogenesis (SI Appendix, Fig. S5). Where did these upregulated shorter 3'UTR transcripts come from? To answer this question, we plotted the shortening ratios (defined as average length of the shorter isoform/length of the longest transcript) against the total splicing events (combined counts of intron skipping/retention, exon skipping/inclusion and alternative 5'/3'UTR splicing, etc.), and found an interesting trend: the shorter the transcripts became, the more splicing events were detected (Fig. 2-3D), suggesting that the upregulated shorter transcripts in the three Alkbh5 KO spermatogenic cell types are most likely derived from enhanced splicing events of those longer transcripts normally expressed in WT spermatogenic cells. We then analyzed the total splicing events in all of the transcripts detected in the three spermatogenic cell types purified from KO and WT testes. Exon skipping/inclusion (ESI) appeared to be
upregulated and intron skipping/retention (ISR) was drastically downregulated in
Alkbh5 KO compared to WT controls (Fig. 2-3E), implicating enhanced splicing
events in the absence of ALKBH5 in these three spermatogenic cell types.

To further reveal the relationship between m6A levels and exon
skipping/inclusion (ESI) events, we compared the m6A density between the
transcripts that not only were upregulated in Alkbh5 KO round and elongating
spermatids, but also contained >3 ESI events with those enriched in WT round
and elongating spermatids without ESI events (Fig. 2-3F). Significantly higher
m6A levels were observed in coding sequences (CDS) in Alkbh5 KO round and
elongating spermatids, as compared to WT cells (Fig. 2-3F), further supporting a
role of m6A in the control of splicing. More intriguingly, ~41% of the m6A sites
overlapped with ESI/ISR sites (Fig. 2-3G), and m6A appeared to be enriched
near the sites where the aberrant splicing events took place (Fig. 2-3H). Indeed,
two transcripts (Unc50 and Traf3ip1) showed that the dominant m6A retention
sites were close to the sites with exon skipping events in Alkbh5 KO round
spermatids (Fig. 2-3I). These data suggest that m6A sites may guide the splicing
machinery during pre-mRNA processing in the nucleus, as proposed in previous
studies (6, 10, 14). This notion is also supported by the nuclear localization of
ALKBH5 in pachytene spermatocytes and round spermatids during
spermatogenesis (Fig. 2-1K, L; SI Appendix, Fig. S3). qPCR analyses confirmed
that these the longer 3’UTR transcripts of these two genes were indeed
dramatically downregulated in Alkbh5 KO round and elongating spermatids
compared to WT controls (SI Appendix, Fig. S6). Together, our data strongly
suggest that proper erasure of m6A from the pre-mRNAs by ALKBH5 is critical for the production of longer 3’UTR transcripts in pachytene spermatocytes and round spermatids in WT testes, and failure to do so leads to aberrant splicing and consequently the production of aberrant shorter transcript isoforms in the *Alkbh5* KO testes.

**Fate of aberrantly spliced transcripts due to m6A erasure failure.** From round to elongating spermatids, the overall transcript, 3’UTR, and 5’UTR lengths all further decreased in wild-type mice (Fig. 2-2A). However, an opposite trend was observed in *Alkbh5* KO spermatogenic cells (Fig. 2-4A), i.e., the overall mRNA length significantly increased, instead of decreased (Fig. 2-2A), from round to elongating spermatids. Given that the KO spermatogenic cells lack longer 3’UTR transcripts due to m6A-induced aberrant splicing (Fig. 2-3), we examined the fate of ~600 shorter 3’UTR (<500nt) transcripts derived from aberrant splicing, and found that ~67% of the shorter 3’UTR transcripts that were upregulated in KO round spermatids became downregulated in KO elongating spermatids (Fig. 2-4B). GO term enrichment analyses identified that these dysregulated, shorter transcripts in the KO spermatids were mostly derived from genes involved in RNA splicing, cilium development and spermatogenesis (*SI Appendix*, Fig. S7). It is interesting to see that genes encoding splicing factors (e.g., *Sfswap*, *U2af2*, *Srsf1*, *Khdrbs3* and *Snrnp70*) were aberrantly spliced in the absence of ALKBH5 (*SI Appendix*, Table S1) because dysregulation of these splicing factors inevitably would lead to more splicing errors in their target genes,
thus amplifying the initial adverse effects and generating a vicious cycle of aberrant splicing. Aberrantly spliced transcripts due to m6A erasure failure in pachytene spermatocytes and round spermatids quickly degraded in elongating spermatids (Fig. 2-4B). Many of the implicated genes, e.g., Foxj1 and Dnaaf3, are involved in ciliogenesis (SI Appendix, Table S2), and are known important for sperm flagellar development (44, 45). Among 53 dysregulated spermatogenesis-related genes (SI Appendix, Table S3), many play an essential role in spermiogenesis. For example, Crem encodes a transcriptional factor serving as a master regulator of the spermatid differentiation program (46). Prm2 encodes protamine 2, which works together with other protamine proteins, as well as transition proteins, to properly pack the chromatin to achieve high degree of nuclear condensation in spermatozoa (47). The OAT phenotype observed in our Alkbh5 KO testes likely represent combined effects of all dysregulated genes on spermiogenesis.

Large-scale degradation of those short transcripts may result from the relative increase in the overall length of the transcripts from round to elongating spermatids in the KO testes (Fig. 2-4A). By mapping the m6A sites in those rapidly degraded short transcripts in KO round spermatids, we observed much higher m6A levels at their 3’UTRs (Fig. 2-4C), and this m6A pattern is very different from that of short 3’UTR transcripts in WT round and elongating spermatids (Fig. 2-2B). The elevated m6A levels in the short 3’UTR transcripts in the KO testes likely represent m6A carried over from their precursors, i.e., those longer 3’UTR mRNAs in WT spermatogenic cells, but aberrantly spliced in the
Alkbh5 KO cells. This notion is consistent with the finding that the degradation of longer 3’UTR transcripts in WT spermatogenic cells contain higher m6A levels at the 3’UTRs compared to shorter 3’UTR mRNAs (Fig. 2-2B). Together, these data suggest that the KO-unique shorter transcripts, which represent those spliced from the WT longer mRNAs and marked with elevated levels of m6A, are not stable and become quickly degraded when round spermatids develop into elongating spermatids in Alkbh5 KO testes.

Discussion

Although m6A was first reported >40 years ago, a comprehensive mapping of m6A in total mRNAs was only accomplished recently using anti-m6A antibody-based m6A RIP-Seq or m6A CLIP-Seq (13, 16, 18, 19, 42, 43). Over 12,000 m6A sites have been identified in mRNAs encoded by >7,000 genes in mammalian cells. In general, m6A is often enriched in 3’UTRs close to the stop codon, where factors regulating alternative polyadenylation and splicing, subcytoplasmic compartmentalization and stability (e.g., RNA-binding proteins and small RNAs) tend to bind (42, 43). This suggests that m6A most likely plays a role in post-transcriptional regulation. Indeed, by manipulating genes encoding m6A writers, erasers and readers in cultured cells in vitro and in the whole animals in vivo (e.g., Drosophila, zebrafish and mice), recent studies have demonstrated that m6A in mRNAs appear to be involved in the control of alternative splicing, translational efficiency and stability (6-8). In the present study, we studied the molecular consequences of m6A erasure failure in
spermatogenic cells in murine testes. As an m6A erasure, ALKBH5 has been shown to be nuclear (22, 23). In the testis, ALKBH5 is predominantly nuclear and partially co-localized with SC35, a DNA speckle marker, suggesting that ALKBH5 most likely acts on pre-mRNAs during splicing events in the nucleus. Indeed, our data do support a role of m6A in the control of correct splicing. First, elevated m6A levels are associated with altered splicing events because transcripts enriched in Alkbh5 KO pachytene spermatocytes and round spermatids display increased splicing events (Fig. 2-3E, F). Second, m6A sites are often proximal to the sites of splicing events (Fig. 2-3G, H). Third, genes encoding splicing factors appear to be preferentially affected by m6A retention due to Alkbh5 inactivation (SI Appendix, Table S1). Dysregulated splicing factors, in turn, can further cause aberrant splicing in their target genes. Taken together, ALKBH5-dependent m6A erasure appears to be able to protect longer transcripts from aberrant splicing in the nuclei of pachytene spermatocytes and round spermatids.

The most abundant expression of ALKBH5 in spermatocytes suggests a critical role in meiotic progression. Indeed, a delay in spermatocyte development occurs in the KO testes at P14; dysregulation of many genes known to be critical for meiotic progression, e.g., Syce1, Syce2, and Marf1 (SI Appendix, Table S3) due to failure in m6A demethylation in Alkbh5 KO spermatocytes could contribute to the meiotic defects observed. Levels of ALKBH5 are much reduced in round spermatids and there is no expression in elongating/elongated spermatids. This expression pattern is consistent with the role of m6A in the control of enhanced
degradation of longer transcripts (with higher levels of m6A) and enrichment of shorter mRNAs (containing lower levels of m6A) when spermiogenesis progresses from round to elongating/elongated spermatids.

In KO elongating spermatids, the overall length of mRNAs appears to be increased from round to elongating spermatids, which is in contrast to the continuous decrease in transcript length from WT round to elongating spermatids. Failure of m6A erasure in the Alkbh5 KO pachytene spermatocytes and round spermatids causes retention of more m6A in pre-mRNAs for longer transcripts than in those for shorter 3’UTR mRNAs; consequently, the longer 3’UTR transcripts are affected more severely due to enhanced splicing events. These aberrantly spliced, KO spermatogenic cell-unique short transcripts are likely much shorter than normal shorter transcripts in WT cells, thus leading to a relative increase in length from round to elongating spermatids in the Alkbh5 KO testes. These KO spermatogenic cell-unique shorter mRNAs contain more m6A compared to normal shorter isoforms and thus, are subjected to quick degradation in KO elongating spermatids.

Based on the dynamic changes in m6A and the transcriptomic profiles in the Alkbh5 KO spermatogenic cells, we propose here a biphasic action of m6A, including 1) nuclear action by affecting splicing of the longer transcripts, and 2) cytoplasmic activity by marking transcripts with m6A for degradation (Fig. 2-4D). In WT pachytene spermatocytes and round spermatids, m6A is effectively erased by ALKBH5 to ensure correct splicing of longer 3’UTR transcripts. Despite m6A erasure, the longer 3’UTR transcripts tend to contain much greater levels of m6A.
compared to shorter mRNAs, and the former become less stable in elongating spermatids. The degradation of longer transcripts leads to enrichment of shorter transcripts during late spermiogenesis. In KO spermatogenic cells, failure to reduce m6A levels in longer pre-mRNAs results in enhanced splicing and the production of shorter transcripts. These aberrant short transcripts are different from those normally expressed shorter transcripts in that these are derived from longer mRNAs normally expressed in WT spermatogenic cells, thus retaining much greater levels of m6A, which may serve as the degradation signals in elongating spermatids. Therefore, ALKBH5-dependent m6A plays a critical role in the production of longer mRNAs and their subsequent degradation. The fact that spermatogonia are normal in Alkbh5 KO testes suggests that ALKBH5-dependent m6A has no significant effects on the mitotic phase of spermatogenesis. In contrast, Ythdc2 KO mice display spermatogonial differentiation defects, suggesting different m6A pathways are in operation in mitotic (spermatogonia) vs. meiotic (spermatocytes) spermatogenic cells. Interestingly, inactivation of the writer METTL3 in mice causes a much earlier meiotic arrest (in zygotene spermatocytes) compared to Alkbh5 KO mice (depletion of pachytene spermatocytes and spermatids) (13, 23). The phenotypic differences likely reflect the cell- and stage-specific roles of m6A during spermatogenesis.

In summary, we have demonstrated that ALKBH5-dependent m6A mainly controls mRNA fate in spermiogenesis. Appropriate erasure of m6A from pre-mRNAs is required for correct splicing and production of longer 3'UTR transcripts
in spermatocytes and round spermatids, and elevated levels of m6A in 3'UTRs tend to induce degradation in elongating spermatids.
Acknowledgements

This work was supported by grants from the NIH (HD071736 and HD085506 to WY) and the Templeton Foundation (PID: 50183 to WY). RNA-Seq and bioinformatics were conducted in the Single Cell Genomics Core of University of Nevada, Reno School of Medicine, which was supported, in part, by a COBRE grant from the NIH (1P30GM110767).

Author Contributions

W.Y. conceived and designed the research. C.T., R.K., H. P., T.Y., Z. W., Y.Z. and H.Z. performed the experiments. C.T. performed the bioinformatics analyses. A.K. provided the Alkbh5 KO mice and also participated in data analyses. W.Y., C.T. and R.K. wrote the manuscript. All reviewed the manuscript.

Competing Financial Interests

The authors declare no competing financial interest.
Materials and Methods

Animal use, histology, PAS staining and TUNEL analyses. The male *Alkbh5* KO mice used in this study were described previously (23). Other details are provided in SI Appendix, Materials and Methods.

Gross morphology and sperm analysis. Testis and the whole epididymis were dissected from morphological and sperm analyses. Details are provided in SI Appendix, Materials and Methods.

Immunofluorescence staining and confocal microscopy. Testicular cryosections were used for the analyses. Details are provided in SI Appendix, Materials and Methods.

Western blots. Total protein lysates were extracted from whole testis for the analyses. Details are provided in SI Appendix, Materials and Methods.

Purification of spermatogenic cells. Pachytene spermatocytes, round and elongating/elongated spermatids were purified from adult mouse testes using the STA-PUT method (34). Details are provided in SI Appendix, Materials and Methods.
**RNA extraction.** RNA was extracted from HEK293 cells using the mirVana miRNA Isolation Kit (ThermoFisher, Cat#AM1560), according to the manufacturer's instructions. Extracted RNA quantification was done using the Qubit RNA High Sensitivity Assay Kit (Invitrogen No. Q32855) measured on the Qubit 2.0 Fluorometer (Invitrogen).

**m6A RNA immunoprecipitation.** Rabbit anti-m\(^6\)A antibody (Abcam, Cat#ab151230) or normal rabbit IgG (Invitrogen, Cat#10500C) (6μg each) were used for the analyses. All immunoprecipitations were performed in triplicate. The representative results of the procedures are shown in *SI Appendix Fig. S8*. Other details are provided in *SI Appendix, Materials and Methods*.

**RNA library construction.** Immunoprecipitated RNA (1ng) and non-immunoprecipitated RNA (300ng) were constructed into next-generation sequencing libraries (Illumina) using the KAPA Stranded RNA-Seq Library Preparation Kit (KK8400) according to the manufacturer’s instructions. Details are provided in *SI Appendix, Materials and Methods*.

**RNA-Seq data analyses.** Quality control of RNA-Seq data are shown in *SI Appendix Fig. S9*. Trimmomatic was used to remove adaptor sequences, and low-quality reads from the sequencing data (48). To identify all the transcripts, we used Tophat2 and Cufflinks to assemble the sequencing reads based on the UCSC MM9 mouse genome (49). The differential expression analysis was
performed by Cuffdiff (49). The global statistics and quality controls are presented in Supplemental Figure S1. The UTR and alternative splicing analyses were performed using the SpliceR pipeline (50). The sequences without identified coding frames were extracted and subjected to coding potential calculating (51). The sequencing depth and alignment ratio were listed in SI Appendix Fig. S10. We performed the data mining using our in-house R script.

**Data normalization.** FPKMs counts are scaled in Cuffdiff analyses via the median of the geometric means of fragment counts across all libraries, as described in (52). The principle was identical to the one used by DESeq (53).

**m6a RIP-Seq data analyses.** Trimmomatic was used to remove adaptor sequences and low-quality reads from the sequencing data (48). Other details are provided in SI Appendix, Materials and Methods.
Fig. 2-1. Alkbh5 is essential for meiotic and haploid phases of spermatogenesis. (A) Alkbh5−/− (Alkbh5 KO) testes are much smaller than wild-type controls. (B) Significant decrease in testicular weight started at postnatal day 14 (P14) and persisted thereafter into adulthood. (C) Paraffin-embedded, periodic acid-Schiff (PAS)-stained testis sections showing robust spermatogenesis in wild-type mice, but disrupted spermatogenesis in Alkbh5 KO testes, characterized by drastically reduced number of meiotic (spermatocytes) and haploid (spermatids) male germ cells in the seminiferous epithelium and the presence of numerous vacuoles, a hallmark of active germ cell depletion. Scale bars=50µm. (D) Paraffin-embedded, HE-stained epididymal sections showing the presence of fully developed spermatozoa in wild-type caput and cauda epididymides, whereas the Alkbh5 KO caput and cauda epididymides contain no mature spermatozoa, but degenerated germ cells resembling round, elongating, or elongated spermatids, which were most likely those depleted from the seminiferous epithelium (inset). Scale bars=50µm. (E) Representative images showing TUNEL staining of apoptotic germ cells in 6-week-old wild-type and
Alkbh5 KO testes. (F) Quantitative analyses of apoptotic (TUNEL+) germ cells in developing wild-type and Alkbh5 KO testes. The number of TUNEL+ cells were counted per 100 cross sections of seminiferous tubules. Data are presented as means ± SEM (n=3). *p<0.01, **p<0.001, student’s t-test, 2-tailed, homoscedasticity assumed. (G) HE-stained spermatozoa collected from wild-type and Alkbh5 KO epididymides (very rarely seen). All panels were at the same magnification and the scale bar=10µm. (H) Sperm counts in wild-type and Alkbh5 KO mice. Data are presented as means ± SEM (n=10). (I) Sperm motility in wild-type and Alkbh5 KO mice. Data are presented as means ± SEM (n=10). (J) Representative Western blots showing the detection of ALKBH5 protein in wild-type testes, and the absence of ALKBH5 in Alkbh5 KO testes. β-ACTIN was used as a loading control. (K) Immunofluorescent localization of ALKBH5 (green) and SC35 (maker for nuclear speckles) in stage VI seminiferous tubules of wild-type and Alkbh5 KO testes. ALKBH5 is localized to the nuclei of Sertoli cells (SCs), spermatogonia (Sg), spermatocytes with higher levels in pachytene (P) spermatocytes, and step 6 round spermatids (Sd6). Step 15 spermatids (sd15) are devoid of ALKBH5. ALKBH5 is only partially co-localized with SC35. Scale bars = 20µm. (L) Schematic illustration showing ALKBH5 localization in the murine seminiferous epithelia. Sc: Sertoli cells; A4: type A4 spermatogonia; In: intermediate spermatogonia; B: type B spermatogonia; PL: preleptotene spermatocytes; L: leptotene spermatocytes; Z: zygotene spermatocytes; P: pachytene spermatocytes; Di: diplotene spermatocytes; M: dividing spermatocytes; 1-16: steps 1-16 spermatids.
Fig. 2-2. m6A marks the longer 3’UTR transcripts that are destined to be eliminated during spermiogenesis (from round to elongating/elongated spermatids). (A) Density plots showing that the total transcript, 5’UTR and 3’UTR lengths all decreased when round spermatids developed into elongating spermatids although such a trend is less obvious during late meiotic and early haploid phases of spermatogenesis (i.e., from pachytene spermatocytes to round spermatids). Total transcript, 5’UTR and 3’UTR lengths were determined based on RNA-Seq data using SpliceR. (B) m6A sites and levels in longer vs. shorter 3’UTR transcripts in pachytene spermatocytes, round and elongating spermatids. Two types of transcripts were analyzed: i) 30 longer 3’UTR (>3,000nt) mRNAs.
that were mainly synthesized in round spermatids, but were drastically downregulated when round spermatids developed into elongating spermatids. ii) 200 shorter 3'UTR (<500nt) mRNAs, whose levels continuously increased from round to elongating spermatids. These transcripts are most likely those required for final several steps of sperm assembly and are subjected to delayed translation in pachytene spermatocytes and round spermatids. Density of m6A reads detected in m6A RIP-Seq datasets were plotted against the total mRNA length. Longer 3'UTR mRNAs contain much higher levels of m6A, which is mainly enriched in 3'UTRs proximal to the stop codon. In contrast, levels of m6A in shorter 3'UTR transcripts are much lower and no significant enrichment was noticed. Note that m6A levels of the longer 3'UTR mRNAs were noticeably higher in elongating spermatids than in pachytene spermatocytes and round spermatids (peak 1). (C, D) Two example genes (Uhmk1 and Traf3ip1) showed higher m6A levels in the 3'UTRs close to the stop codon in elongating spermatids than in round spermatids. (E) Common motifs detected surrounding the m6A sites.
Fig. 2-3. Proper m6A erasure is required for the production of longer 3’UTR mRNAs in pachytene spermatocytes and spermatids. (A) Density plots showing that the total transcript length was significantly decreased in Alkbh5 KO (KO) pachytene spermatocytes, round and elongating spermatids compared to corresponding wild-type (WT) spermatogenic cells. (B) Heat maps showing the 30 longer 3’UTR (>3,000nt) transcripts predominantly expressed in round spermatids in wild-type testes were significantly downregulated in all three spermatogenic cell types in Alkbh5 KO testes. (C) Heat maps showing that the shorter 3’UTR isoforms of these 30 longer transcripts were all up-regulated in three spermatogenic cell types in Alkbh5 KO testes compared to wild-type (WT) controls. (D) Dot plot showing the relationship between shortened transcripts and splicing events in the three Alkbh5 KO spermatogenic cell types (pachytene
spermatocytes, round and elongating spermatids) analyzed. Shortening ratios (defined as the length of shorter isoform/the length of the longest transcript) were plotted against the number of total splicing events detected based on the RNA-Seq data. The shorter the transcript isoforms become, the more splicing events they tend to have, suggesting that those shorter transcript isoforms were derived from enhanced splicing of those longer transcripts in the three Alkbh5 KO spermatogenic cell types. (E) Histograms showing splicing events, including exon skipping/inclusion and intron skipping/retention, in the three spermatogenic cell types in WT and Alkbh5 KO testes. (F) Comparison of m6A density in transcripts enriched in Alkbh5 KO round (upper panel) and elongating (lower panel) spermatids with >3 exon skipping/inclusion (ESI) events and those enriched in wild-type (WT) round and elongating spermatids without ESI events. Note that frames indicate elevated m6A levels in the CDS region of the transcripts. (G) Venn diagram showing ~41% of the m6A sites overlap with the sites with splicing events (+/- 200nt distance), including exon skipping/inclusion (ESI) and intron skipping/retention (ISR). (H) Density plots showing correlations between splicing and m6A sites in wild-type (WT) and Alkbh5 KO (KO) round spermatid-enriched transcripts. The transcripts enriched in KO round spermatids appear to contain more splicing sites proximal to the m6A sites, as compared to those enriched in WT round spermatids, suggesting enhanced splicing events due to m6A accumulation in the KO cells. (I) Examples of two mRNAs (Unc50 and Traf3ip1) upregulated in Alkbh5 KO round spermatids showing exon skipping events (circled) near the m6A accumulation sites.
Fig. 2-4. Fate of the aberrantly spliced short transcripts in Alkbh5 KO (KO) round spermatids. (A) Density plot showing the average length of total transcripts increased from round to elongating spermatids in Alkbh5 KO testes; this expression pattern is opposite to the shortening trend in wild-type controls, as shown in Fig. 2-2A (upper right panel). (B) Heat map showing ~2/3 of the shorter 3’UTR (<500nt) transcripts upregulated in Alkbh5 KO round spermatids were quickly down-regulated when round spermatids develop into elongating spermatids. Given that the upregulated shorter 3’UTR transcripts in Alkbh5 KO round spermatids were mostly those aberrantly spliced from the longer transcripts in the wild-type cells, this result suggests that these KO cell-specific shorter 3’UTR transcripts are not stable in the Alkbh5 KO cells. (C) Elevated levels of m6A in the 3’UTRs close the stop codon in Alkbh5 KO cell-unique shorter 3’UTR transcripts. Note that this m6A pattern is typical to the longer, but not the shorter 3’UTR transcripts in WT round and elongating spermatids, as shown in Fig. 2-2B, right lower panel. (D) Schematic presentation showing the physiological functions of ALKBH5-dependent m6A erasure in late meiotic (pachytene spermatocytes) and haploid (round and elongating spermatids) phases of spermatogenesis in the wild-type testes, and the molecular consequences of m6A erasure failure due to Alkbh5 inactivation.
References


Chapter 3: Environmental Toxicant Induced Epigenetic Transgenerational Inheritance of Prostate Pathology and Stromal-Epithelial Cell Epigenome and Transcriptome Alterations: Ancestral Origins of Prostate Disease

Rachel Klukovich\(^2\)*, Eric Nilsson\(^1\)*, Ingrid Sadler-Riggleman\(^1\), Hayden McSwiggin\(^2\), Yeming Xie\(^2\), Daniel Beck\(^1\), Wei Yan\(^2\)+, Michael K. Skinner\(^1\+)

\(^1\)Center for Reproductive Biology
School of Biological Sciences
Washington State University
Pullman, WA, 99164-4236, USA

\(^2\)Department of Physiology and Cell Biology
University of Nevada, Reno School of Medicine
Reno, NV, 89557, USA

(*) Co-first authors
(+) Co-senior authors

This manuscript is in preparation for submission for publication.

Correspondence: Michael K. Skinner
Center for Reproductive Biology
School of Biological Sciences
Washington State University
Pullman, WA 99164-4236
Phone: 509-335-1524
Fax: 509-335-2176
Email: skinner@wsu.edu
Abstract

Prostate diseases include prostate cancer, which is the second most common male neoplasia, and benign prostatic hyperplasia (BPH), which affects approximately 50% of men. The incidence of prostate disease is increasing, and some of this increase may be attributable to ancestral exposure to environmental toxicants and epigenetic transgenerational inheritance mechanisms. The goal of the current study was to determine the effects that exposure of gestating female rats to vinclozolin has on the epigenetic transgenerational inheritance of prostate disease, and to characterize by what molecular epigenetic mechanisms this has occurred. Gestating female rats (F0 generation) were exposed to vinclozolin by daily injections during E8-E14 of gestation. F1 generation offspring were bred to produce the F2 generation, which were bred to produce the transgenerational F3 generation. F3 generation vinclozolin lineage males at 12 months of age were evaluated for prostate disease, compared to the control lineage. Prostate epithelial and stromal cells were isolated from F3 generation 20-day old rats, prior to the onset of disease, and used to obtain DNA and RNA for analysis. Results indicate that there were changes in gene expression, noncoding RNA expression, and DNA methylation in both cell types. Our results suggest that ancestral exposure to vinclozolin at a critical period of gestation induces epigenetic transgenerational inheritance of prostate stromal and epithelial cell changes in both the epigenome and transcriptome that ultimately lead to prostate disease and may serve as a source of the increased incidence of prostate disease observed in recent years.
**Introduction**

Prostate disease is very common in older men in North America with 50% of men between the ages of 50 and 60 having evidence of pathologic benign prostatic hyperplasia (BPH) (1). The incidence of prostate cancer has been increasing worldwide in the past decades with prostate cancer now being the second most common neoplasia in men (2-4). While some of this increase can be attributed to an aging population other factors such as toxicant exposures and epigenetic transgenerational inheritance of disease susceptibility appear to be of importance.

Epigenetics refers to molecular factors and processes around the DNA that regulate genome activity independently of DNA sequence, and are often mitotically stable (5). Epigenetic factors include DNA methylation, histone modifications, expression of non-coding RNA (ncRNA), RNA methylation, and adjustments to chromatin structure (6). Epigenetic transgenerational inheritance is defined as the germline transmission of epigenetic information and phenotypic change across generations in the absence of any continued direct environmental exposure or genetic manipulation (5). Epigenetic changes can be induced by environmental factors such as nutrition or toxicant exposure and are an important mechanism by which organisms change their gene expression in response to their environment. While epigenetic changes must be inherited via germ cells (i.e. sperm or eggs), it is the epigenetic changes that these germ cells induce in the early embryo and embryonic stem cells that then promote altered epigenome and transcriptomes in
all derived somatic cells of the individual that lead to disease susceptibility in tissues and organs. Therefore, disease development in organs such as the prostate gland can be due to ancestral exposures and epigenetic inheritance (7).

The prostate’s epithelium is responsible for contributing secretions to semen. Prostatic epithelial cells contain a large endoplasmic reticulum and golgi apparatus, as well as many secretory granules. There are multiple tubuloalveolar glands in a prostate that are lined by prostatic epithelium. These glands are separated from each other by prostatic stroma. The stroma of the prostate is considered to be the interstitial tissue and is made up of smooth muscle cells, blood vessels, fibroblasts, and nerves. These muscle cells are believed to work in unison with the epithelial cells to expel secretions to the semen (8). In the current study the cell type isolated as prostatic stroma is primarily fibroblasts.

The prostate develops from the urogenital sinus (UGS) which branches to form the prostate in response to androgens, with the prostatic buds appearing in rats at E18-E19 and the majority of prostate branching occurring postnatally (reviewed in (9)). Rodent prostates have three prostatic lobes consisting of the anterior prostate, the dorsolateral prostate, and the ventral prostate, which has the most extensive branching. In the current studies prostatic epithelial cells are isolated from the ventral prostate. After prostate growth is complete in the adult organism, the epithelium of the prostate has low levels of proliferation and cell death, maintaining a constant prostate size in the presence of androgens. The epithelial-stromal ratio is also believed to be critical in determination of final prostate size (reviewed in (9)).
Smooth muscle cells in the prostate are known to have many androgen receptors and are believed to regulate epithelial cells through androgen signaling (10). Further, high levels of testosterone have been shown to induce proliferation of prostate stromal cells (11). A constant source of androgens is required to maintain a healthy prostate and is essential throughout development. Testosterone and its metabolites produced by fetal testis are required for branching during prostate development, while an absence of androgens or their receptors causes the UGS to form as a vagina. In addition, estrogenic compounds can also have a wide variety of effects on the developing prostate, with low levels resulting in an enlarged prostate and high levels stunting prostate growth (reviewed in (12)).

Initial studies of the ability of environmental toxicants to promote the epigenetic transgenerational inheritance of prostate disease showed that ancestral exposure to the agricultural fungicide vinclozolin increased rates of prostatic epithelial atrophy, cystic hyperplasia and prostatitis in the unexposed transgenerational F3 and F4 generations (13). These effects were accompanied by transgenerational changes in mRNA expression in F3 generation ventral prostate epithelial cells, as determined by microarray analysis (14). Associated epigenetic changes in these cells were not investigated at that time.

In the current study, transgenerational changes to the epigenome of prostate epithelial and stromal cells were characterized in F3 generation rats after ancestral vinclozolin exposure, compared to controls. Stromal-epithelial cell interactions are critical for normal prostate development and function and
abnormal interactions can lead to prostate diseases, including cancer (15-19). The transgenerational epigenetic changes investigated in the current study include changes to DNA methylation, which have previously been associated with ancestral toxicant exposures in both germ cells (20, 21) and somatic cells (22, 23). Additionally, changes in expression of mRNAs and non-coding RNAs (ncRNAs) are determined for both prostatic epithelium and stroma.

Noncoding RNAs are any type of RNA other than messenger RNA and are derived from the noncoding portions of the genome. It is currently believed that long, noncoding RNAs (IncRNAs) are responsible for maintaining epigenetic memory through a variety of mechanisms including: regulating DNA methylation, chromatin remodeling, regulating histone modifications, or affecting transcription and translation by altering transcript stability (reviewed in (24)). A recent study has shown that the sperm of transgenerational males that were ancestrally exposed to DDT have differentially expressed IncRNAs (25). Small, noncoding RNAs have also been shown to have a role in epigenetic transgenerational inheritance. In C. elegans an increased sncRNA population was induced upon starvation and persisted until the F3 generation, resulting in longer lifespans (26). There are many different kinds of small, noncoding RNAs that are found in the spermatozoa, and both the large (>200 nucleotide (nt)) and the small (<200 nt) noncoding RNAs have been found to have differential expression throughout spermatogenesis (reviewed in (27)). This suggests that the expression of noncoding RNAs is crucial for proper regulation of gene expression and that environmental disruption of this expression can negatively impact development and the health of future offspring.
Elucidation of epigenetic and gene expression changes that occur in the prostate after ancestral exposure to an environmental toxicant provides insight into the molecular etiology of the epigenetic transgenerational inheritance of prostate disease and improves our understanding of the risk factors that must be considered when investigating the increasing incidence of prostate disease in the human population.
Results

Prostate Pathology Analysis

For this study, pregnant F0 generation rats were treated with vinclozolin or control vehicle from days 8-14 of gestation, as described in Methods. The fetal exposed offspring, F1 generation rats, were bred to produce the F2 generation, and similarly F2 generation animals were bred to produce the transgenerational F3 generation. No sibling or cousin crosses were used to avoid inbreeding artifacts. Only the F0 generation rats received the experimental treatments. Ventral prostatic tissue was harvested from F3 generation males at 18-21 days of age. Prostate epithelial and stromal cells were isolated and analyzed so as to characterize DNA methylation, mRNA gene expression and ncRNA expression as described in Methods. Additional F3 generation vinclozolin lineage and control lineage rats were aged to one year and their prostates subjected to histopathological evaluation to detect signs of prostate disease.

Prostate disease was defined as the presence of prostatic epithelial atrophy, epithelial hyperplasia, and/or as the presence of vacuoles in glands at rates two standard deviations above those found in controls (see Methods). There was a significant increase in prostate disease in transgenerational F3 generation vinclozolin lineage rats at one year of age compared to F3 generation controls (28) (Figure 3-1A). A recent study also demonstrated a significant prostate disease increase in F3 generation DDT lineage males (29). Interestingly, previous studies have shown no transgenerational prostate disease was detected following plastic
derived compound exposures (bisphenol A (BPA) and phthalates (DBT & DEHP))
(30), dioxin (TCCD) (31), pesticides permethrin and DEET (32), jet fuel
hydrocarbons (33), or methoxychlor (34) exposures, Figure 3-1B. Therefore,
observations suggest an exposure specificity in the induction of transgenerational
inheritance of prostate disease. There was also no increase in prostate disease in
F1 or F2 generation vinclozolin lineage rats compared to controls (data not shown).
This indicates that there was an epigenetic transgenerational increase in
susceptibility to prostate disease in rats ancestrally exposed to vinclozolin.

**DNA Methylation Analysis**

Differences in sites of DNA methylation (*i.e.* Differential DNA Methylation
Regions, DMRs) between F3 generation control and vinclozolin lineage rats were
characterized for both prostatic epithelial cells and stroma cells using methylated
DNA immunoprecipitation (MeDIP) followed by next generation sequencing for a
MeDIP-Seq procedure and bioinformatics techniques as described in Methods. A
number of p-value thresholds are assessed. In prostate epithelial cells there were
304 DMRs at a p-value of p<1x10^{-6}, of which 42 DMRs comprised multiple
sequential genomic windows (Figure 3-2A). A list of these DMRs are presented in
Supplemental Table S1. In prostate stromal cells there were 1249 DMRs at a p-
value of p<1x10^{-6}, of which 307 DMRs were comprised of multiple sequential
genomic windows (Figure 3-2B). A list of these DMRs are presented in
Supplemental Table S2. At p<1x10^{-6} there were 50 DMRs in common between
stromal and epithelial cells (Figure 3-2C) and the list of these DMRs are presented
in Supplemental Table S3. Chromosomal locations of the DMRs were examined. The DMRs were present on all chromosomes except the small Y chromosome and on mitochondrial DNA (Figure 3-3A and 3-3B). The red arrowheads identify the DMRs and black boxes clusters of DMRs.

Examination of the characteristics of the genomic sites where DMRs reside shows that most DMRs are present in areas having an average of 1 or 2 CpG sites per 100 base pairs (Figure 3-4A and 3-4C). A CpG is a cytosine adjacent to a guanine on the DNA, and it is primarily these cytosine bases that are methylated. This indicates that most of the DMRs identified occur in areas of low CpG density termed CpG deserts (35). Most DMRs for both prostate epithelial cells and stroma cells were shown to be less than one kilobase (kb) in length (Figure 3-4B and 3-4D).

Some DMRs occurred in the vicinity (within 10 kb) of known genes. These DMR associated genes were categorized and evaluated for potential function. For both prostate epithelial cells and stroma cells the DMR associated genes were most often related to signaling, metabolism, transcription and receptor functions (Figure 3-5A and 3-5B). The list of DMR associated genes was also compared to well-characterized physiological pathways in the KEGG database (http://www.kegg.jp/kegg/kegg2.html). Those pathways having the most DMR associated genes are presented in Figure 3-6. Five of the ten pathways for both the epithelial cells and stromal cells were in common and indicated.
noncoding RNA Analysis

The differentially expressed large (mRNA, lncRNA) and small RNAs (sncRNA) were determined using RNA-seq. Different p-value thresholds were used for both prostate epithelial and stromal (Figure 3-7A and 3-7B, respectively) cell lines. A significance of \( p < 0.001 \) was chosen for subsequent analysis. In both epithelial and stromal cells, the mRNA had the highest differential expression, closely followed by the long, noncoding RNA. The small, noncoding RNAs from the epithelial prostate cells were differentially expressed at almost twice that of the stromal prostate cells (165 vs 76, respectively). Differentially expressed sncRNAs were broken down into categories by type (Figure 3-7C), with both epithelial and stromal prostate cells having piRNA as the most affected category of small noncoding RNAs by far. Noticeably, the tRNAs or other small, noncoding RNAs were not differentially expressed in the epithelial cells. Therefore, the different classes of sncRNAs are affected differently in the same tissue in different cell types, indicating that even within the same organ, ancestral exposures can have vastly different effects on different cell types.

Next, the chromosomal locations of differentially expressed large RNAs were analyzed. The altered long, noncoding RNAs from both the epithelial (Figure 3-8A) and stromal (Figure 3-8B) cell lines were equally dispersed and present on all chromosomes except for the Y chromosome. This was also the case for the mRNAs from the epithelium and stroma (Figure 3-8C and 3-8D, respectively). Interestingly, the small, noncoding RNAs had different chromosomal locations than the large RNA. While both the epithelial and stromal sncRNAs were absent from 6
of the 22 chromosomes (Figure 3-9A and 3-9B, respectively), only the stromal cell line had a differentially expressed sncRNA on the mitochondrial (MT) chromosome. Since differential RNA expression on the MT chromosome only occurs in 1 of the 6 conditions (lncRNA, mRNA, sncRNA for each stromal and epithelial), this indicates that differential RNA expression in transgenerational vinclozolin lineage is not random.

Finally, all differentially expressed epigenetic modifications from both the epithelial prostate and the stromal prostate were compared to see if the modifications overlapped (i.e. lncRNA, mRNA, sncRNA, and DMRs). In the epithelial prostate (Figure 3-10A), DMRs overlapped individually with either the lncRNA, or the mRNA, but not both. In addition, very little overlap was observed between the lncRNA and the mRNA, and no overlap was observed with any of the sncRNAs. In contrast, in the stromal prostate, there was a significant overlap between the mRNA and the sncRNA and no overlap between the lncRNA and the mRNA (Figure 3-10B). The DMRs did have a sizable overlap with each the mRNA and the lncRNA, but not both. In both the stromal and epithelial prostate, the DMRs had overlap for each the mRNA and lncRNA, but not both. This indicates that the DMRs are critical regulators of epigenetic transgenerational inheritance regardless of cell type in the same organ, but that the small, noncoding RNAs are critical regulators of gene control in only some cell types of the same organ.
Discussion

The results of these studies indicate that ancestral exposure to the toxicant vinclozolin induces an epigenetic transgenerational increase in susceptibility to prostate disease in F3 generation rats. These results are in agreement with previous studies, which also found a transgenerational increase in susceptibility to prostate disease after exposure of F0 generation pregnant rats to vinclozolin (13, 14). Interestingly, exposure of pregnant rats to some other environmental toxicants, such as jet fuel, the plastics compounds bisphenol A (BPA) and phthalates, the insecticides permethrin and methoxychlor, and the industrial pollutant dioxin, did not promote a transgenerational increase in prostate disease (30-34). A recent study also demonstrated DDT can promote transgenerational prostate disease (29). This suggests that the ability to promote prostate disease transgenerationally may be specific to certain compounds or environmental exposures.

Changes in DNA methylation were observed in F3 generation vinclozolin lineage prostate epithelial and stromal cells, compared to the control lineage. The sites of these DMRs were in genomic regions of relatively low CpG density ‘CpG deserts’. This finding is consistent with previous work in which transgenerational DMRs in sperm were most often found in regions of low CpG density after ancestral toxicant exposure (30-34).

Examination of the noncoding RNAs showed that the two prostate cell types, epithelial and stromal, had very different classes of differentially expressed
genes. While the epithelial prostate had a dramatically different number of differentially expressed sncRNAs compared to the stroma (165 vs 76), none of the differentially expressed epithelial sncRNAs overlapped with the other differentially expressed epithelial modifications. This is in direct contrast to the stromal sncRNAs, where 12.5 sncRNAs were shown to overlap with mRNAs. That is almost 16% of the total differentially expressed sncRNAs, one of which was also found to be located on the MT chromosome. As sncRNAs are known to affect gene expression, it is likely that the mechanism of epigenetic transgenerational inheritance in the stromal prostate is primarily through sncRNAs. Unlike the stroma, the epithelial prostate did not have any sncRNA overlap with either the lncRNA or the mRNA. In contrast, DMRs were found to overlap with both the mRNAs and lncRNAs. In addition, 1.5 mRNAs overlapped with lncRNAs, indicating that the mechanism of epigenetic transgenerational inheritance in the prostate epithelium is primarily through control of gene expression by methylation and lncRNAs. Since the majority of the differential epigenetic modifications do not overlap with each other, this indicates that the majority of these modifications are intergenic. In the future, it will be necessary to determine the exact gene targets of these epigenetic modifications to determine the mechanism by which the prostate diseases occur due to ancestral exposure to vinclozolin.

It is interesting to note that these epimutations are present in prostatic epithelium and stroma even at 18-21 days of age, which is long before any visible signs of prostate disease are detectable. This indicates that the underlying factors
that can contribute to an adult-onset disease like prostate disease can be present early in life.

In summary, these studies show that exposure to the environmental toxicant vinclozolin can promote the epigenetic transgenerational inheritance of susceptibility to prostate disease. Prostate epithelial and stromal cells from young vinclozolin lineage animals had epigenetic changes in DNA methylation and ncRNA expression, as well as in mRNA gene expression. These changes likely contribute to the dysregulation of the prostate gland that occurs in later life. Future studies should investigate if similar mechanisms are at work in human males who have adult-onset BPH or prostate cancer.
Methods

Animal studies and breeding

Female and male rats of an outbred strain Hsd:Sprague Dawley®™SD®™ (Harlan) at about 70 to 100 days of age were fed ad lib with a standard rat diet and ad lib tap water for drinking. To obtain time-pregnant females, the female rats in proestrus were pair-mated with male rats. The sperm-positive (day 0) rats were monitored for diestrus and changes in body weight. If pregnant, then on days 8 through 14 of gestation (36), the females were administered daily intraperitoneal injections of vinclozolin (25 mg/kg BW/day, Chem Services, Westchester PA, USA) or dimethyl sulfoxide (vehicle) as previously described (20). Treatment groups were designated ‘vinclozolin’ and ‘control’ lineages. The gestating female rats treated were considered to be the F0 generation. The offspring of the F0 generation rats were the F1 generation. Non-littermate females and males aged 70-90 days from F1 generation control or vinclozolin lineages were bred to obtain F2 generation offspring. The F2 generation rats were bred to obtain F3 generation offspring. Only the pregnant F0 generation rats were treated directly with vinclozolin. The control and vinclozolin lineages were housed in the same rooms with lighting, food and water as previously described (5, 13, 20). All experimental protocols for the procedures with rats were pre-approved by the Washington State University Animal Care and Use Committee (IACUC approval # 02568).
Tissue harvest and histology processing

Rats at 12 months of age were euthanized by CO₂ inhalation and cervical dislocation for tissue harvest. Ventral prostates were removed and fixed in Bouin’s solution (Sigma) followed by 70% ethanol, then processed for paraffin embedding by standard procedures for histopathological examination. Tissue sections (5 µm) were cut and were stained with H & E stain and examined for histopathologies.

Histopathology examination and disease classification

Prostate histopathology criteria included the presence of vacuoles in the glandular epithelium, atrophic epithelial layer of ducts and hyperplasia of prostatic duct epithelium as previously described (14, 37). A cut-off was established to declare a tissue ‘diseased’ based on the mean number of histopathological abnormalities plus two standard deviations from the mean of control tissues by each of the three individual observers blinded to the treatment groups. This number was used to classify rats into those with and without prostate disease in each lineage. A rat tissue section was finally declared ‘diseased’ only when at least two of the three observers marked the same tissue section ‘diseased’. Results were expressed as the proportion of affected animals and were analyzed using Fisher’s exact test.

Prostatic epithelial and stromal cell collection

Ventral prostate epithelial and stromal cells were isolated as previously described (38). Briefly, ventral prostates were removed from 19-21 day old rats
and cleaned of fat, then digested in 50 ml Hank’s Buffered Salt Solution (HBSS) with 0.5 mg/ml collagenase type II (Sigma C1764) and 66μg/ml DNAse (Sigma DN25) with agitation at 37° for 4 hours. After gravity settling for 10 min. the supernatant containing the stromal cells was removed. The supernatant was centrifuged at 30xg for 4 min. to pellet contaminating epithelial cells, and then the supernatant centrifuged again at 190xg for 6 min to pellet the stromal cells. This wash is repeated 1-2 more times. To clean the epithelial cells from the original gravity settled pellet, the pellet is resuspended in HBSS, centrifuged at 30xg for 4 min., and the supernatant discarded. This wash is repeated 1-2 more times.

**DNA Isolation**

The cell pellet was resuspended in 820 μL DNA extraction buffer and 80 μl proteinase K (20 mg/ml) added. The sample was incubated at 55°C for 2 hours under constant rotation. Then 300 μl of protein precipitation solution (Promega A795A) were added, the sample mixed thoroughly and incubated for 15 min on ice. The sample was centrifuged at 17,000xg for 30 minutes at 4°C. One ml of the supernatant was transferred to a 2 ml tube and 2 μl of glycoblue and 1 ml of cold 100 % isopropanol were added. The sample was mixed well by inverting the tube several times then left in -20°C freezer for at least one hour. After precipitation, the sample was centrifuged at 17,000xg for 20 min at 4°C. The supernatant was taken off and discarded without disturbing the (blue) pellet. The pellet was washed with 70% cold ethanol and incubated at -20°C for 20 minutes. Samples were
centrifuged for 10 min at 4°C at 17,000xg and the supernatant discarded. Pellet was air-dried at RT (about 5 minutes), then resuspended in 100 µl of nuclease free water.

**Methylated DNA Immunoprecipitation (MeDIP)**

The genomic DNA was fragmented using the Covaris M220 the following way: genomic DNA was diluted to 130 µl with TE buffer (10mM Tris HCl, pH7.5; 1mM EDTA) into the appropriate Covaris tube. Covaris was set to 300 bp program and the program was run for each tube in the experiment. 10 µl of each sonicated DNA was run on 1.5% agarose gel to verify fragment size. The remaining DNA was diluted with TE buffer to 400 µl, heat-denatured for 10min at 95°C, then immediately cooled on ice for 10 min. Then 100µl of 5X IP buffer and 5µg of antibody (monoclonal mouse anti 5-methyl cytidine; Diagenode #C15200006) were added, and the DNA-antibody mixture was incubated overnight with rotation at 4°C.

The following day 50µl of pre-washed anti-mouse IgG magnetic beads (Dynabeads M-280 Sheep anti-Mouse IgG; Life Technologies 11201D) were added to the DNA-antibody mixture, then incubated for 2h on a rotator at 4°C. Then the DNA-antibody-bead mixture was placed into a magnetic rack for 1-2 minutes and the supernatant discarded, then washed with 1xIP buffer 3 times. The washed sample is then resuspended in 250µl digestion buffer (5mM Tris PH8, 10.mM EDT4, 0.5% SDS) with 3.5µl Proteinase K (20mg/ml)) and incubated for 2-3 hours on a rotator at 55°. 250µl of buffered Phenol-Chloroform-Isoamylalcohol
solution were added to the sample and the tube vortexed for 30 sec then centrifuged at 17,000xg for 5min at room temperature. The aqueous supernatant was carefully removed and transferred to a fresh microfuge tube. Then 250μl chloroform were added to the supernatant from the previous step, vortexed for 30sec and centrifuged at 17,000xg for 5min at room temperature. The aqueous supernatant was removed and transferred to a fresh microfuge tube. To the supernatant 2μl of Glycoblue (20mg/ml) (Invitrogen AM9516), 20μl of 5M NaCl and 500μl ethanol were added and mixed well, then precipitated at -20˚C for >1 hour. The DNA precipitate was centrifuged at 17,000xg for 20min at 4˚C and the supernatant removed. The pellet was washed with 500μl cold 70% ethanol and incubated at -20˚C for 15 min. then centrifuged again at 17,000xg for 5min at 4˚C and the supernatant discarded. The pellet was air-dried at RT (about 5 minutes) then resuspended in 20μl H₂O or TE. DNA concentration was measured using a Qubit (Life Technologies) with ssDNA kit (Molecular Probes Q10212).

**MeDIP-Seq Analysis**

The MeDIP DNA was used to create libraries for next generation sequencing (NGS) using the NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB #E7530S) (San Diego, CA) starting at step 1.4 of the manufacturer’s protocol to generate double stranded DNA. After this step the manufacturer’s protocol was followed. Each sample received a separate index primer. NGS was performed at WSU Spokane Genomics Core using the Illumina HiSeq 2500 with a PE50
application, with a read size of approximately 50 bp and approximately 30 million reads per pool. 6 libraries were run in one lane.

**mRNA and ncRNA isolation**

Total RNA (mRNA, IncRNA, rRNA, tRNA, and sncRNA) was extracted from purified prostate cells (epithelial or stromal) using either the Trizol reagent (Thermo Fisher) or mirVana miRNA isolation kit (Life Technologies) following the manufacturer's protocol with some modifications. Control lineage prostate cells were stored as a cell pellet at -80°C until extraction with the mirVana kit. Cell pellets were manually homogenized and heated to 65°C for 10 minutes after lysis buffer was added. The manufacturer's protocol was then resumed. Vinclozolin lineage prostate cells were suspended in 1.2 mL of Trizol and stored at -80°C until use. The manufacturer’s protocol was followed with the exception of increasing the amount of isopropanol added to 1 mL at the RNA precipitation step to recover the small RNA. RNA from both lineages was eluted in 50 μL of water with the addition of 0.5 μL murine RNase inhibitor (NEB).

Quality control analysis for both lineages was performed by running the RNA on an RNA 6000 Pico chip on the Agilent 2100 Bioanalyzer (Agilent). The Qubit RNA HS Assay Kit (Thermo Fisher) was used to determine RNA concentration.
mRNA and ncRNA sequencing

Large mRNA and noncoding RNA libraries were constructed from total RNA using the KAPA RNA HyperPrep Kit with Ribonuclease (KAPA) according to the manufacturer’s protocol, with some modifications. Barcodes and adaptors were from NEBNext Muliplex Oligos for Illumina. Prior to PCR amplification, libraries were incubated at 37°C for 15 minutes with the USER enzyme (NEB). PCR cycle number was determined using qPCR with the KAPA RealTime Library Amplification kit before final amplification. Size selection (200-700 bp) was performed using KAPA Pure beads (KAPA). Quality control was performed using Agilent DNA High Sensitivity chips (Agilent) and concentration was determined using Qubit dsDNA high sensitivity assay (Thermo Fisher). Libraries were pooled (10 samples per pool by equal RNA content) and loaded onto an Illumina HiSeq 4000 sequencer on a paired-end 100 bp flow cell. Bioinformatics analysis was used to separate mRNA libraries from ncRNA libraries and to determine differential expression (see ncRNA bioinformatics section).

Small RNA libraries were constructed using the NEBNext Multiplex Small RNA Library Prep Set for Illumina and were barcoded with the NEBNext Multiplex Oligos for Illumina. After amplification, purification and size selection was performed using the KAPA Pure beads at 1.3x and 3.7x ratios following the manufacturer's instructions. Final size selection (115-160 bp) was performed using the Pippin Prep 3% gel with marker P (Sage Science). Quality control was performed using Agilent DNA High Sensitivity chips (Agilent) and concentration
was determined using Qubit dsDNA high sensitivity assay (Thermo Fisher). Libraries were pooled and concentrated using 2.2x KAPA Pure beads and were loaded onto an Illumina HiSeq 4000 sequencer and sequenced with a single-end 50 bp flow cell. A customized primer was used to sequence the sRNA libraries: 5’-ACA CGT TCA GAG TTC TAC AGT CCG A-3’. Bioinformatics analysis was used to determine differential expression (see ncRNA bioinformatics section).

**DMR Statistics and Bioinformatics**

The basic read quality was verified using summaries produced by the FastQC program. The new data was cleaned and filtered to remove adapters and low-quality bases using Trimmomatic (39). The reads for each MeDIP sample were mapped to the Rnor 6.0 rat genome using Bowtie2 (40) with default parameter options. The mapped read files were then converted to sorted BAM files using SAMtools (41). To identify DMRs, the reference genome was broken into 100 bp windows. The MEDIPS R package (42) was used to calculate differential coverage between control and exposure sample groups. The edgeR p-value (43) was used to determine the relative difference between the two groups for each genomic window. Windows with an edgeR p-value less than an arbitrarily selected threshold were considered DMRs. The DMR edges were extended until no genomic window with an edgeR p-value less than 0.1 remained within 1000 bp of the DMR. CpG density and other information was then calculated for the DMR based on the reference genome.
DMRs were annotated using the biomaRt R package (44) to access the Ensembl database (45). The genes that overlapped with DMR were then input into the KEGG pathway search (46, 47) to identify associated pathways. The DMR associated genes were manually then sorted into functional groups by consulting information provided by the DAVID (48), Panther (49), and Uniprot databases incorporated into an internal curated database (www.skinner.wsu.edu under genomic data). All molecular data has been deposited into the public database at NCBI (GEO # _____) and R code computational tools available at GitHub (https://github.com/skinnerlab/MeDIP-seq) and www.skinner.wsu.edu.

**ncRNA statistics and bioinformatics**

The small ncRNA data were annotated as follows: Low-quality reads and reads shorter than 15nt were discarded by Trimmomatics (v0.33). The remaining reads were matched to known rat sncRNA, consisting of mature miRNA (miR-Base, release 21), precursor miRNA (miRBase, release 21), tRNA (Genomic tRNA Database, rn5), piRNA (piRBase), rRNA (Ensembl, release 76) and mitochondrial RNA (Ensembl, release 76) using AASRA pipeline with default parameters. Read counts generated by AASRA were statistically normalized by DESeq2.

The long ncRNA data were annotated as follows: Trimmomatics (v0.33) was used to remove adaptor sequences and the low-quality reads from the RNA sequencing data of the large RNA libraries. To identify all the transcripts, we used HiSAT2 (v2.1.0) and StringTie (v1.3.4d) to assemble the sequencing reads based
on the Ensembl_Rnor_6.0. The differential expression analyses were performed by Cuffdiff. The coding and the non-coding genes were primarily annotated through rat CDS data ensembl_Rnor_6.0. The non-annotated genes were extracted through our in-house script and then analyzed by CPAT, indicating the true non-coding RNAs.
References


32. Manikkam M, Tracey R, Guerrero-Bosagna C, Skinner M. Pesticide and Insect Repellent Mixture (Permethrin and DEET) Induces Epigenetic


38. Chaudhary J, Schmidt M, Sadler-Riggleman I. Negative acting HLH proteins Id 1, Id 2, Id 3, and Id 4 are expressed in prostate epithelial cells. Prostate. 2005;64(3):253-64.


Figure 3-1. Prostate pathology frequency. (A) Transgenerational prostate disease in F3 generation control, vinclozolin and DDT lineage males at 1 yr of age. The (*) indicates statistical significance of $p<0.05$. (B) Transgenerational prostate disease frequency from previous studies in control, plastics, dioxin, pesticides and jet fuel lineage males at 1 yr of age. No statistical difference from control disease frequency was observed.
Figure 3-2. DMR identification. In each table the number of DMRs found using different p-value cutoff thresholds is presented. The all window column shows all DMRs. The multiple window column shows the number of DMRs containing at least two significant windows. At the base of each table is presented the number of DMR with each specific number of significant windows at a p<1e-06. (A) Prostate epithelial cell F3 generation DMRs p<1e-06. (B) Prostate stromal cell F3 generation DMRs p<1e-06. (C) Venn diagram showing the number of DMR in common between prostate stroma and epithelium.
Figure 3-3. DMR chromosomal locations. The DMR locations on the individual chromosomes for all DMRs are at a p-value threshold of p<1e-06. (A) Prostate epithelial cell. (B) Prostate stromal cell. Red arrowheads indicate positions of DMR and black boxes indicate clusters of DMR.
Figure 3-4. DMR genomic features. (A & C) The number of DMRs at different CpG densities for all DMRs at a p-value threshold of p<1e-06. (B & D) The DMR lengths for all DMRs are at a p-value threshold of p<1e-06. (A & B) Prostate epithelial cell. (C & D) Prostate stromal cell.
Figure 3-5. **DMR gene associations** for (A) Prostate epithelial cell and (B) Prostate stromal cell. Genes in 10 kb proximity to DMR were categorized as to function and the number of DMR associated genes in each category is presented. DMR are at a p-value of ≤1e-6.
Figure 3-6. DMR association gene pathways. (A) Prostate epithelial cell. (B) Prostate stromal cell. DMR associated genes were surveyed for their presence in known physiological pathways (KEGG pathways). The number of DMR associated genes present in each pathway is indicated. Bold indicates common pathways between the cell types.
Figure 3-7. Differential expression of noncoding RNAs between the control and vinclozolin lineage in prostate epithelial (a) and stromal (b) cells. (c) Categories of differentially expressed small, noncoding RNA at P< 0.001.
Figure 3-8. Chromosomal location of differentially expressed large RNAs. Long, noncoding RNAs from the epithelial (a) and stromal (b). mRNAs from the epithelial (c) and stromal (d). Red arrows indicated individual large RNA, while black boxes indicate clusters. P<1 x 10^{-3}. 
Figure 3-9. Chromosomal locations for differentially expressed small, noncoding RNAs from the epithelial (a) and stromal (b) cell lines. Red arrows indicate individual sncRNAs, while black boxes indicate clusters. P<0.001. There are 9 differentially expressed sncRNA with unknown locations on the epithelium and 4 for the stroma.
Figure 3-10. Overlaps of the DMRs with the differentially expressed noncoding RNAs in the prostate epithelium (a) and the prostate stroma (b). p<0.001.
Chapter 4: Conclusion

m6A modifications have always been believed to have a regulatory role, but the exact function of this modification has been under heated debate. When the first techniques for m6A RNA immunoprecipitation were published back in 2012 (1, 2), it revolutionized the functionality studies that could be done. The role of m6A was still under debate with one group showing that alternatively spliced genes were enriched in m6A modifications (2) and another stating that its enrichment in the 3’UTR could potentially serve as an miRNA binding site (1). Later publications indicated that the location of the m6A modification could affect its function: if located in the 5’UTR, translation would be increased (3), while locations on other parts of the transcript led to degradation (4).

Regardless of the function, it became clear that global developmental processes were being affected if the writers, readers, or erasers of m6A were not properly expressed. First, lethality was the result of writer METTL3 depletion (4), highlighting the importance of the m6A modification early in development. Later, the maternal-to-zygotic transition was shown to be impaired if reader, YTHDF2, was knocked out (5). Another reader, YTHDC2, was shown to promote cell cycle progression from prophase to meiosis (6). Immune cell differentiation was also shown to be affected in METTL3 conditional knockout CD4 T-cells, where they maintained a naïve state (7), indicating that m6A is needed continuously to allow for differentiation to adapt to a changing environment.
The first indication that m6A was involved in infertility was in 2013, when it was shown that knockout of the eraser ALKBH5 caused male mice to exhibit an infertility phenotype (8). The authors conclusively demonstrated that lack of ALKBH5 caused an increase in apoptotic spermatocytes and drastically reduced sperm count. Their research was furthered in 2017 with the use of m6A RNA immunoprecipitation to demonstrate the biphasic action of m6A: in the nucleus, m6A marks transcript sites for splicing pre-mRNAs, while in the cytoplasm, it marks transcripts for degradation (Chapter II).

Although the function of ALKBH5 has been elucidated in spermatogenesis, the question remains as to whether there are other truly m6A-specific erasers. FTO was previously believed to be an m6A eraser, but has recently been questioned as to its actual target (9). Any other erasers found could have a similar effect as ALKBH5 knockout or may have an additive effect when both erasers are depleted. Further, the effects of ALKBH5 depletion on female mice remains to be investigated. It is also critical to translate this work to clinical research to allow human patients to benefit.

In addition to improperly regulated epigenetic modifications, epigenetic transgenerational inheritance is also a major contributor to abnormal reproductive health, although its exact contribution is still unknown. It was first discovered that a single exposure to an environmental toxicant could induce the transgenerational phenotype of apoptotic spermatogenic cells and male infertility in 2005 (10). Since then, many studies have investigated the phenomenon of epigenetic transgenerational inheritance stemming from the male. Due to
vinclozolin’s antiandrogenic properties and DDT’s estrogenic properties, these two compounds are the most widely studied for transgenerational effects on reproductive health (11, 12).

The prostate is a male reproductive organ whose growth, development, and maintenance relies upon androgen receptors (13, 14). Because of this reliance on androgens for development, the effects of vinclozolin on the developing prostate were investigated (Chapter III). It was discovered that the transgenerational increased susceptibility to prostate disease was caused by changes in both differential DNA methylation regions (DMRs) and changes in long, noncoding RNA, both large and small. The importance of different cell types in the same tissue having different responses to an ancestral exposure was also highlighted by the fact that the epigenetic modifications were differentially expressed in the prostate stromal and epithelial cells. This was an extension of a previous study demonstrating that ancestral exposure vinclozolin was correlated to prostatic lesions in 45-55% of F1-F4 generation males (15). Interestingly, abnormal histology (atrophy) was only observed in the ventral lobe in the distal, intermediate, and proximal regions. This study showed the clinical relevancy of ancestral vinclozolin exposure on prostate disease because the incidence of prostate lesions in the F1-F4 rats was similar to the 50% incidence seen in men over 50 (16).

These epigenetic transgenerational inheritance studies highlight the importance of limiting exposures during pregnancy to toxic substances, as a single exposure during a critical developmental period can have negative impacts
that last for generations. In the future, it will be necessary to investigate epigenetic transgenerational effects from other periods during embryonic development, rather than limiting studies to only gonadal sex determination (E8-E14). It is also crucial to determine if epigenetic transgenerational inheritance can pass to offspring from the female. It is difficult however, to perform epigenetics experiments such as those done on sperm, due to the limited quantities of oocytes available for analysis. The number of transgenerational F3 females needed to collect a sufficient number of oocytes makes it extremely time consuming and expensive to conduct these studies.

Together, the reversible m6A modification and any heritable alterations of epigenetic states due to ancestral exposure are both crucial aspects of disease susceptibility, both early in development and adult-onset. In addition, both of these epigenetic aspects have important roles in human health and disease.
References


# Appendix

## Reagents used in Chapter II

### Key Resources Table

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Antibodies</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rabbit anti-ALKBH5</td>
<td>Sigma</td>
<td>HPA007196</td>
</tr>
<tr>
<td>Mouse anti-GM130</td>
<td>BD Biosciences</td>
<td>610822</td>
</tr>
<tr>
<td>Mouse anti-sc35</td>
<td>Abcam</td>
<td>ab11826</td>
</tr>
<tr>
<td>Goat anti-rabbit Alexa Fluor 488</td>
<td>Thermo Fisher</td>
<td>A11034</td>
</tr>
<tr>
<td>Goat anti-mouse Alexa Flour 594</td>
<td>Abcam</td>
<td>Ab150116</td>
</tr>
<tr>
<td>Mouse anti-β-actin</td>
<td>Abcam</td>
<td>Ab8226</td>
</tr>
<tr>
<td>Goat anti-rabbit HRP</td>
<td>Abcam</td>
<td>Ab205718</td>
</tr>
<tr>
<td>Goat anti-mouse HRP</td>
<td>Southern Biotech</td>
<td>1030-05</td>
</tr>
<tr>
<td>Rabbit anti-m6A</td>
<td>Abcam</td>
<td>Ab151230</td>
</tr>
<tr>
<td>Normal rabbit IgG</td>
<td>Invitrogen</td>
<td>10500C</td>
</tr>
<tr>
<td><strong>Biological Samples</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HEK293</td>
<td>ATCC</td>
<td>CRL-1573</td>
</tr>
<tr>
<td><strong>Chemicals, Peptides, and Recombinant Proteins</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAPI</td>
<td>Abcam</td>
<td>Ab104139</td>
</tr>
<tr>
<td>4-15% Mini-PROTEAN TGX Stain-free gel</td>
<td>Bio-Rad</td>
<td>456-8084</td>
</tr>
<tr>
<td>StrepTactin-HRP</td>
<td>Bio-Rad</td>
<td>161-0376</td>
</tr>
<tr>
<td>ERKB</td>
<td>Sigma</td>
<td>K-4002</td>
</tr>
<tr>
<td>Protein G Dynabeads</td>
<td>Invitrogen</td>
<td>100040</td>
</tr>
<tr>
<td>SpeedBeads</td>
<td>Sigma-Aldrich</td>
<td>65152105050250</td>
</tr>
<tr>
<td><strong>Critical Commercial Assays</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pierce BCA Protein Assay Kit</td>
<td>Thermo Fisher</td>
<td>23225</td>
</tr>
<tr>
<td>TUNEL kit</td>
<td>Trevigen</td>
<td>4810-30-CK</td>
</tr>
<tr>
<td>Advanta Western Bright ECL Kit</td>
<td>Advanta</td>
<td>K-12045-D20</td>
</tr>
<tr>
<td>mirVana miRNA isolation kit</td>
<td>Life Technologies</td>
<td>AM1560</td>
</tr>
<tr>
<td>KAPA Stranded RNA-Seq Library Preparation Kit</td>
<td>KAPA</td>
<td>KK8400</td>
</tr>
<tr>
<td>KAPA real-time library amplification kit</td>
<td>KAPA</td>
<td>KK2702</td>
</tr>
<tr>
<td><strong>Deposited Data</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sequence Read Achieve database</td>
<td>NCBI</td>
<td>PRJNA420607</td>
</tr>
<tr>
<td><strong>Experimental Models: Organisms/Strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alkbh5− C57BL/6J</td>
<td>Zheng et al. (2013) Molecular Cell</td>
<td>N/A</td>
</tr>
<tr>
<td>Wild-type C57BL/6J</td>
<td>Yan lab’s animal facility</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Oligonucleotides</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEB Multiplex Oligos for Illumina</td>
<td>NEB</td>
<td>E7335, E7500</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>-----</td>
<td>-------------</td>
</tr>
<tr>
<td>Gapdh F</td>
<td>IDT</td>
<td>ACAGTCATGCCATC ACTGCC</td>
</tr>
<tr>
<td>Gapdh R</td>
<td>IDT</td>
<td>GCCGTGCTTCACCACC TTCTT</td>
</tr>
<tr>
<td>Unc50 (447 &amp; 452 com) F</td>
<td>IDT</td>
<td>GGTGGTGGCTTTCTGG ATCTCAACGTTA</td>
</tr>
<tr>
<td>Unc50 (447 &amp; 452 com) R</td>
<td>IDT</td>
<td>CACTGTATCCAGGA AGGTACATAG</td>
</tr>
<tr>
<td>Unc50 (447) short iso R</td>
<td>IDT</td>
<td>CATGAGAGGTGCGA ATGGATAGAGCA</td>
</tr>
<tr>
<td>Unc50 (452) long iso R</td>
<td>IDT</td>
<td>CATATGAGTGAAACC ATAAACACCTCAGGCA</td>
</tr>
<tr>
<td>Traf3ip1 (1506 &amp; 1509 com) F</td>
<td>IDT</td>
<td>CAGGAGCAGAGTAT CACAGACAGTGC</td>
</tr>
<tr>
<td>Traf3ip1 (1506 &amp; 1509 com) R</td>
<td>IDT</td>
<td>AGCAGGCATTTTCTC AGCAGACT</td>
</tr>
<tr>
<td>Traf3ip1 (1506) long iso R</td>
<td>IDT</td>
<td>CTGGAGGCTCTTAAG TAGTAAACACTGTA GC</td>
</tr>
<tr>
<td>Traf3ip1 (1509) short iso R</td>
<td>IDT</td>
<td>ACTGTCTCTCTCCCACA CCGAGAGAC</td>
</tr>
<tr>
<td>Alkbh5 WT F</td>
<td>IDT</td>
<td>IDTCGATCCGTGTT AAATCTG</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------</td>
<td>-------</td>
<td>--------------------------------------</td>
</tr>
<tr>
<td>Alkbh5 WT R</td>
<td>IDT</td>
<td>TAAGTAAGTGCCCT GAATGG</td>
</tr>
<tr>
<td>Alkbh5 KO F</td>
<td>IDT</td>
<td>AATCTGACGGAAT ATCAAAAGACTGGA AAAGG</td>
</tr>
<tr>
<td>Alkbh5 KO R</td>
<td>IDT</td>
<td>AAGGAGACCACAT TCATAGAACTCGA ACTCC</td>
</tr>
</tbody>
</table>

**Software and Algorithms**

<table>
<thead>
<tr>
<th>Software</th>
<th>Bioinformatics</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trimmomatic</td>
<td>Bioinformatics</td>
<td>Bolger et al. 2014</td>
</tr>
<tr>
<td>Tophat</td>
<td></td>
<td>V2.0.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N/A</td>
</tr>
<tr>
<td>Cufflinks</td>
<td>Nat Protoc</td>
<td>Trapnell et al. 2012</td>
</tr>
<tr>
<td>Cuffdiff</td>
<td>Nat Protoc</td>
<td>Trapnell et al. 2012</td>
</tr>
<tr>
<td>SpliceR</td>
<td>BMC Bioinformatics</td>
<td>Vitting-Seerup et al 2014</td>
</tr>
</tbody>
</table>

**Other**

| Other                  |       |                                      |
|                       |       |                                      |
| Superfrost Plus Slides | Fisher | 22-037-246                           |
| Nitrocellulose membrane 0.45 μm | Amersham | 10600003                           |
Reagents used in Chapter III

Key Resources Table

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antibodies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mouse anti-5-methyl cytidine</td>
<td>Diagenode</td>
<td>C15200006</td>
</tr>
<tr>
<td>Chemicals, Peptides, and Recombinant Proteins</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vinclozolin</td>
<td>Chem Service, Inc.</td>
<td>N-13745-250MG</td>
</tr>
<tr>
<td>Collagenase type II</td>
<td>Sigma</td>
<td>C1764</td>
</tr>
<tr>
<td>DNase</td>
<td>Sigma</td>
<td>DN25</td>
</tr>
<tr>
<td>Protein precipitation solution</td>
<td>Promega</td>
<td>A795A</td>
</tr>
<tr>
<td>Dynabeads M-280</td>
<td>Life Technologies</td>
<td>11201D</td>
</tr>
<tr>
<td>Glycobluce</td>
<td>Invitrogen</td>
<td>AM9516</td>
</tr>
<tr>
<td>Trizol Reagent</td>
<td>Thermo Fisher</td>
<td>15596018</td>
</tr>
<tr>
<td>KAPA Pure Beads</td>
<td>KAPA</td>
<td>KK8002</td>
</tr>
<tr>
<td>Critical Commercial Assays</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEBNext Ultra RNA Library Prep Kit for Illumina</td>
<td>NEB</td>
<td>E7530S</td>
</tr>
<tr>
<td>NEB Multiplex Small RNA Library Prep Set for Illumina</td>
<td>NEB</td>
<td>E7330S</td>
</tr>
<tr>
<td>mirVana miRNA isolation kit</td>
<td>Life Technologies</td>
<td>AM1560</td>
</tr>
<tr>
<td>KAPA RNA HyperPrep Kit with RiboErase</td>
<td>KAPA</td>
<td>KK8560</td>
</tr>
<tr>
<td>KAPA real-time library amplification kit</td>
<td>KAPA</td>
<td>KK2702</td>
</tr>
<tr>
<td>Pippin Prep 3% gel with marker P</td>
<td>Sage Science</td>
<td>CDP3010</td>
</tr>
<tr>
<td>Experimental Models: Organisms/Strains</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hsd:Sprague Dawley SD</td>
<td>Harlan</td>
<td>N/A</td>
</tr>
<tr>
<td>Oligonucleotides</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEB Multiplex Oligos for Illumina</td>
<td>NEB</td>
<td>E7335, E7500</td>
</tr>
<tr>
<td>sRNA sequencing primer</td>
<td>IDT</td>
<td>ACA CGT TCA GAG TTC TAC AGT CCG A</td>
</tr>
<tr>
<td>Software and Algorithms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trimmomatic v0.33</td>
<td>Bioinformatics</td>
<td>Bolger et al. 2014</td>
</tr>
<tr>
<td>miRBase</td>
<td>miRBase</td>
<td>Release 21</td>
</tr>
<tr>
<td>Genomic tRNA database</td>
<td>Gtrnadb.ucsc.edu</td>
<td>Rn5</td>
</tr>
<tr>
<td>piRBase</td>
<td>regulatoryrna.org</td>
<td>N/A</td>
</tr>
<tr>
<td>Ensemblli</td>
<td>Ensembl genome browser</td>
<td>Release 76</td>
</tr>
<tr>
<td>HISAT2</td>
<td>John Hopkins University</td>
<td>V2.1.0</td>
</tr>
<tr>
<td>StringTie</td>
<td>John Hopkins University</td>
<td>V1.3.4d</td>
</tr>
<tr>
<td>MeDIPS R package</td>
<td>Bioinformatics</td>
<td>Leinhard et al. 2014</td>
</tr>
<tr>
<td>------------------</td>
<td>---------------</td>
<td>----------------------</td>
</tr>
</tbody>
</table>