2016

Characterization of Small Non-Coding RNAs in the Seminal Plasma of Beef Bulls with Predicted High and Low Fertility

Stephanie Perkins
South Dakota State University

CHARACTERIZATION OF SMALL NON-CODING RNAS IN THE SEMINAL PLASMA OF BEEF BULLS WITH PREDICTED HIGH AND LOW FERTILITY

BY

STEPHANIE PERKINS

A thesis submitted in partial fulfillment of the requirements for the

Master of Science

Major in Animal Science

South Dakota State University

2016

CHARACTERIZATION OF SMALL NON-CODING RNAS IN THE SEMINAL PLASMA OF BEEF BULLS WITH PREDICTED HIGH AND LOW FERTILITY

This thesis is approved as a creditable and independent investigation by a candidate for the Master of Science in Animal Science degree and is acceptable for meeting the thesis requirements for this degree. Acceptance of this does not imply that the conclusions reached by the candidate are necessarily the conclusions of the major department.

Michael Gonda, Ph.D.
Thesis Advisor Date

Joseph Cassady, Ph.D.
Head, Department of Animal Science Date

Dean, Graduate School Date

ACKNOWLEDGEMENTS

First, and foremost, thank you to Dr. Michael Gonda for giving me the opportunity to participate in some incredible experiences. I am so grateful for the time and mentoring that you have provided to me during my time at SDSU. A special thanks to Dr. Tara McDaneld, for hosting me in your lab at the Meat Animal Research Center. You allowed me the chance to learn from you and work with some fantastically knowledgeable (and fun!) people in the process. To Dr. George Perry, thank you for answering my endless questions and coaching me in the laboratory. And a huge thank you to Dr. Joe Cassady, because without you none of this would have happened.

Over the past three years, the support of my fellow graduate students has been invaluable. There are so many people who have helped me along the way and I appreciate each and every one of you.

I would not have survived this process without the support of my family and friends. Mum, Dad, and Kieran, thanks for always loving and believing in me, no matter how far away from home I might go.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

CHARACTERIZATION OF SMALL NONCODING RNAS IN THE SEMINAL PLASMA OF BEEF BULLS WITH PREDICTED HIGH AND LOW FERTILITY

STEPHANIE PERKINS

2016

MicroRNAs (miRNAs) are a family of small non-coding RNAs (sRNA) that play a key role in regulating gene expression by binding to their complementary mRNA and altering translation. It has been reported that this down-regulation of specific areas of the genome plays a role in male fertility. Piwi-interacting RNAs (piRNAs) are another class of germline-specific non-coding RNAs that form silencing complexes to maintain germline genome integrity. In bovine seminal plasma, piRNAs have not previously been characterized. Thus, the objective of this study was to investigate a potential relationship between expression of these sRNAs and fertility in beef bulls and to characterize piRNA expression in bull seminal plasma. Semen samples were collected from 35 beef bulls and analyzed for *Arachis hypogaea*/peanut agglutinin (PNA) and *Lens culinaris* agglutinin (LCA) scores. PNA reveals acrosomal damage or malformation and LCA indicates an altered sperm surface. LCA and PNA are both correlated with ubiquitin, a negative protein biomarker found on sperm cells, which indicates poor overall fertility. Divergent cohorts consisting of 3 bulls with high PNA/LCA scores (H) and 3 bulls with low PNA/LCA scores (L) were chosen. The sRNAs were extracted from the seminal plasma of these bulls. Sequence libraries were prepared using the Illumina TruSeq Small RNA preparation kit and sequenced on an Illumina MiSeq. Using the NCBI database of known human, porcine, mouse, and bovine miRNAs and known human piRNAs, expressed

sequences were identified and mapped to the Bovine genome (Bta_4.6.1). A weighted FDR-corrected t-type test statistic was used to identify differentially expressed sRNAs between the H and L groups. Of the 617 sRNAs that were discovered, 9 miRNAs that were differentially expressed (FDR $P<0.05$) and 49 piRNAs that were differentially expressed (FDR $P<0.05$) were identified. All of the identified piRNAs in seminal plasma have previously been reported in human male germlines. Differentially expressed miR-181a-2 and miR-181b-2 were up-regulated in the L group and have been previously characterized in embryos, small intestine, and as a cluster in the thymus. Also, miR-181a-1 was found to be significantly related to male infertility diagnoses in humans. Additionally, other significant miRNAs identified in this study have been previously identified in embryonic, brain, and mammary tissues. Real-time RT-PCR was performed on 33 bulls with PNA/LCA scores using candidate piRNAs selected based on sequencing results. The association between the PNA and LCA markers and piRNA expression was analyzed using a linear model including age as a covariate. It was found that PNA was significantly related to pi-30961 and pi-32679 ($P<0.05$) and LCA exhibited a trend towards significance with pi-32679 ($P=0.08$). These identified candidate miRNAs and piRNAs support the hypothesis that miRNAs and piRNAs correlate with differences in fertility of beef bulls.

Chapter 1

LITERATURE REVIEW

Introduction

The fertility of beef bulls is a key component to successfully maintaining a cattle herd. In beef cattle herds, bulls are often responsible for successfully impregnating multiple females. It has been found that 20 to 40% of bulls in an unselected population exhibit reduced fertility (Kastelic and Thundathil, 2008). This suboptimal performance of the bull can lead to a greater impact on calf numbers then a single female's infertility. This impact makes being able to identify overall bull fertility important for producers.

The class of small non-coding RNAs (sRNA), known as microRNAs (miRNA) are known to play a key role in regulating gene expression by binding to their complementary mRNA and altering translation. MicroRNAs could play a role in male fertility (Yadav and Kotaja, 2014; Hu *et al*., 2014; Houwing *et al*., 2007). Another class of sRNAs, piwi-interfering RNAs (piRNA) are germline-specific non-coding RNAs that form silencing complexes to maintain germline genome integrity. These piRNAs have not previously been characterized in bovine seminal plasma. The use of next-generation sequencing (NGS) techniques have made researching sRNAs and their potential link to various biological functions much more feasible.

Research into protein biomarkers has discovered numerous proteins that are correlated with semen quality, including ubiquitin and various lectin proteins. Higher ubiquitin levels are positively correlated with abnormal sperm morphology (Odhiambo *et al*., 2011). Increased expression of the lectin proteins *Arachis hypogaea*/peanut agglutinin (PNA) and *Lens culinaris* agglutinin (LCA) are also positively correlated with increased ubiquitin and lower fertility levels in beef bulls (Sutovsky *et al*., 2015).

Identifying connections between sRNA actions and fertility in beef bulls will allow for greater understanding of the underlying biological mechanisms involved in bull fertility. This information could eventually allow for selection of bulls based on predicted fertility measures identified from their genetic makeup. Furthermore, sRNA expression may be manipulated to impact semen quality. This literature review will cover the key components of male fertility, specifically in beef bulls as well as measures that can be used to predict fertility. The other topic covered in detail will be different classes of sRNAs and their link to fertility.

Male Fertility

Reproductive performance in cattle is an economically important trait. Within cow-calf operations, the overall fertility level of the herd affects calving intervals and number of calves produced. An increase in calving interval or a decrease in calf numbers can lead to serious economic losses for a producer. If a cow fails to conceive when she is first bred, it either goes undetected or she is re-bred. If the cow fails to produce a calf, producers who chose to retain the cow will pay for her feed and care without a return on that investment. It is highly unlikely that a non-pregnant cow will be able to generate enough profit in the future to offset a single missed calf (Mathews and Short, 2001). Bulls specifically can have a major impact upon fertility, as individual bulls will service multiple females during a breeding season. The current method of measuring a bull's ability to successfully breed cows is the breeding soundness exam (BSE). However, fertility of bulls can be affected by a plethora of factors at any point in their life and a BSE only provides a snapshot in time rather than an estimate of the bull's fertility throughout its lifetime.

Breeding soundness exams were designed to test the likelihood of a bulls ability to successfully establish pregnancy in more than 25 females over a 65-70 day breeding season (Kastelic *et al.*, 2008). Thus, these exams cannot be performed until a bull reaches puberty. A BSE examines physical traits of the bull as well as characteristics of their ejaculate. These characteristics include scrotal circumference, sperm motility, sperm morphology, and physical defects.

To pass the BSE, bulls under 15 months of age must have a scrotal circumference of at least 30cm. For bulls older than 15 months, scrotal circumference should be at least 32cm. Scrotal circumference is one of the few traits measured in regards to bull fertility that has genetic predictions available. Larger scrotal circumferences have been genetically correlated with decreased days to puberty in males and females (Brinks *et al.*, 1978). Bulls are also examined for physical abnormalities, as abnormalities of the scrotum, testes, or accessory sex glands can cause reduced fertility in a bull. These defects indicate that a bull has been injured, is sick, or that the reproductive tract failed to properly develop (Ellis, 2007). Sperm motility is another factor included in a BSE. To test for motility, a drop of undiluted semen is placed on a warm slide and examined under a light microscope. Sperm should be progressively motile, or moving forward across the field of view (Amann and Hamerstedt, 1993) and should exhibit over 30% motility to pass a BSE (Kastelic and Thundathil, 2008). Extreme temperatures and environmental contamination of the semen sample can decrease motility.

Sperm morphology is examined under a light microscope using oil immersion. This characteristic may have the largest impact on pregnancy rates; a 10% increase in semen quality based on morphology resulted in a 5% increase in pregnancy rates in the U.S. (Wiltbank *et al.*, 1986). Abnormalities can be classified as primary or secondary abnormalities. Primary defects include double, elongated, misshapen and detached heads, crater defect, and ruffled acrosome, and are often considered developmental abnormalities. Secondary defects include coiled or bent tail, doubled midpiece or tail of the spermatozoan, and the majority of these abnormalities (i.e. coiled and bent tails) are considered to occur due to semen handling. Other abnormalities involve proximal or

distal droplets. Bulls must exhibit at least 70% morphologically normal sperm to pass a BSE (Saacke, 2008).

Spermatogenesis results in the generation of millions of spermatozoa being produced each minute. However, the molecular events occurring to produce these spermatozoa must be absolutely accurate in order to allow for the transmission of genetic information to potential offspring (Yadav and Kotaja, 2014). Spermatogenesis occurs in the seminiferous tubules of the testes. In order for spermatozoa to be produced, there are three main endocrine requirements: (1) gonadotropin-releasing hormone (GnRH) secretion from the hypothalamus, (2) follicle stimulating hormone (FSH) and luteinizing hormone (LH) release from the pituitary gland, and (3) testosterone and estrogen secretion. In females, the hypothalamus contains a tonic and surge center that controls GnRH release. The tonic center releases basal levels of GnRH at all times, while the surge center releases a preovulatory surge of GnRH once during the estrous cycle (Senger, 2003). The male lacks this surge center. Instead, hormone pulses occur every few hours (Figure 1.1) and each hormone causes a different reaction throughout the male body. A pulse of GnRH causes the release of LH and FSH from the anterior pituitary gland. The actions of these hormones eventually leads to negative feedback on GnRH, stopping the release of LH and FSH and creating the pulses seen in Figure 1.1. The LH binds to the Leydig cells in the testes, which produce progesterone to be converted into testosterone. In the seminiferous tubules, testosterone is essential for the continuation of spermatogenesis. Part of the testosterone will enter circulation while the rest will be transported to the Sertoli cells. The FSH binds to the Sertoli cells, located in the seminiferous tubules, which helps to convert testosterone into dihydrotestosterone (DHT)

and estrogen. Without FSH, the Sertoli cells do not function. Sertoli cells also produce inhibin, which travels to the anterior pituitary and suppresses the release of FSH. Estradiol and DHT exert negative feedback effects on the hypothalamic release of GnRH (Sharpe, 1994). All of these processes remaining in balance within their self-regulating system are necessary for the maintenance of spermatogenesis.

Spermatogenesis is split into 3 phases: the proliferation phase, the meiotic phase, and the differentiation phase. The proliferation phase involves primitive spermatogonia undergoing mitotic divisions, which is regulated by FSH actions upon Sertoli cells. These mitotic divisions generate large numbers of spermatogonia. Most of the spermatogonia enter the next phase of spermatogenesis, while a few spermatogonia revert back to primitive spermatogonia for later mitosis. This renewal is what allows for the continuity of spermatogenesis. The meiotic phase involves primary and secondary spermatocytes. During this step, DNA replication and crossing over of the chromosomes occurs, producing genetically diverse haploid spermatids. The differentiation phase (spermiogenesis) does not involve any more division of cells. This phase is when the round spermatids transform into the oblong spermatozoon with a head, midpiece, and tail. All major morphological changes like acrosome and flagellum formation and chromatin condensation occur during spermiogenesis (Yadav and Kotaja, 2014).

The differentiation stage is also divided into four stages: the Golgi phase, the cap phase, the acrosomal phase, and the maturation phase. During the Golgi phase, the acrosomic vesicle, containing the acrosomic granule, forms. Centrioles also migrate towards the nucleus during the Golgi phase. Eventually, the centrioles will develop into the sperm tail, which is imperative for movement after ejaculation. During the cap phase,

the acrosomic vesicle spreads over the nucleus and the primitive flagellum, or tail, forms. The acrosomal phase involves continued spreading of the acrosome and elongation of the cytoplasm and nucleus (Leblond and Clermont, 1952). The acrosome plays a key role in fertilization because it contains proteolytic enzymes. When the sperm binds to the embryo, the acrosome releases its contents, which allows the sperm to penetrate the zona pellucida of the embryo. Finally, the maturation phase results in a spermatozoon. The mitochondria assemble themselves along the midpiece and the postnuclear cap made from microtubules has been completed (Senger, 2003). Mitochondrial assembly along the midpiece functions as the 'powerplant' of the sperm, propelling the flagellum and giving the sperm its ability to swim.

Mitotic divisions of the proliferation phase result in 256 spermatids from each starting cell. Thus, if any defects are present in the DNA of the initial spermatids, the defect is replicated over and over again in each spermatid. Once the spermatids have undergone spermiogenesis, the DNA is tightly packaged within the head of the sperm. Due to limited space, it is not possible for DNA modifications to be made, meaning that abnormalities at this point will not be corrected. The process of spermatogenesis is continuous and at any given time, each step involved in generating a sperm cell is occurring simultaneously in different areas of the seminiferous tubules (Leblond & Clermont, 1952).

After spermatogenesis, spermatozoa pass into the head and body of the epididymis and finally into the tail of the epididymis. Sperm are stored in the epididymal tail until the sperm travel through the ductus deferens and into the pelvic urethra, where the sperm cells are combined with seminal plasma and are ready for ejaculation.

In a bull, spermatogenesis takes 61 days (Senger, 2003). Should any insult to the testes occur, 61 days would elapse before the bull can ejaculate normal spermatozoa. These insults can involve a physical injury to the testes or elevated temperatures. The testes are pendulous in many species because heat (body temperature included) has a negative effect on quality of sperm. Understanding the timeframe of the spermatogenic process is particularly important when preparing bulls for the breeding season. A BSE may be able to identify fertility problems but it will take at least 2 months after resolution of the problem for the bull to be ready for breeding.

Although a helpful tool for producers, the BSE has limitations. The BSE can detect changes in sperm quality, but it does not diagnose the problem causing these changes. Even though a BSE is performed, sperm motility and morphology measures are subject to the variation between performing veterinarians and are only representative of a bull's potential at that single point in time. A BSE does not test for the presence of disease or libido level, and will only pass or fail a bull. Being able to identify bulls with inherently superior fertility earlier in their lifetime would be extremely helpful to producers' breeding programs.

Measures for Predicted Fertility

A number of proteins have been associated with sperm quality. One such protein is ubiquitin. Ubiquitin is a 76 amino acid protein that weighs 8.5kDa and targets substrates for proteolysis. Ubiquitination is a mechanism for protein recycling that involves the attachment of one or more ubiquitin molecules that serve as a marker for degradation (Sutovsky *et al*, 2001a). Research has reported that this protein serves as a marker for defective sperm in semen of bulls, rhesus monkeys, humans, mice, and boars (Sutovsky *et al*., 2001b). Ubiquitin is expressed by epithelial cells of the epididymis, which is the location at which ubiquitination of sperm cells occurs. Ubiquitin concentration per sperm cell has been highly correlated with lifelong, average percentage of primary and total sperm cell abnormalities. Beef bulls vary in their median ubiquitin fluorescence concentrations and the proportion of sperm with abnormal morphology is positively correlated with ubiquitin fluorescence (r=0.63; P<0.001; Sutovsky *et al*., 2002).

These fluorescence concentrations are detected using the Sperm-Ubiquitin Tag Immunoassay. This assay works by using specific antibodies with fluorescent dye that attach to the ubiquitin proteins. The intensities of the ubiquitin fluorescence are detected using flow cytometry. Thus this assay allows the amount of ubiquitin to be quantified and serve as a measurement for the quality of sperm cells. This process has been used in male humans and farm animals such as bulls, boars, and rams (Sutovsky, 2009).

Other proteins have been identified which correlate with ubiquitin concentrations and sperm cell quality. Lectins specifically can be used as a marker for acrosomal integrity. Lectins will bind rapidly to specific sugar residues, which are exposed when

the acrosome is damaged (Cheng *et* al, 1996). Two of these lectins are *Arachis hypogaea*/peanut agglutinin (PNA) and *Lens culinaris* agglutinin (LCA).

The precedent for using lectins to assess acrosomal integrity began in the late 1980s and early 1990s with research in humans, mice, boars, stallions, and bulls (Cheng *et al*, 1996). In cattle, 12 lectins that would bind to bull sperm were examined in order to identify one that preferentially attached at the acrosomal region. Of these 12 lectins, PNA exhibited almost no detectable binding to sperm surface receptors, but exhibited intense binding to areas of the acrosome (Cross and Watson, 1994). The lectin PNA has a high affinity for binding to disaccharides with a terminal galactose, which means PNA binds to the exposed outer acrosomal membrane when it is exposed which occurs with acrosomal damage or during sperm capacitation (Kennedy *et al*., 2014). Ubiquitin is also positively correlated with PNA levels (Odhiambo *et al*., 2011). The lectin LCA has an affinity for D-glucose and D-mannose residues, so LCA binds to the acrosomal surface of normal spermatozoa and the entire surface of defective spermatozoa (Graham, 2001). Both lectin markers have been validated in multiple species. Their correlation with ubiquitin and sperm morphology makes them useful as predictors for sperm quality and therefore predictors of fertility based on an ejaculate.

When comparing PNA and LCA, PNA is the preferred measurement due to its limited tendency to bind to other areas of the spermatozoon (Odhiambo *et al*., 2011). The largest limitation of these assays when looking at fertility is that lectins only indicate sperm morphology. Admittedly, sperm morphology and acrosome integrity is imperative to overall fertility. However, other defects could cause serious decreases in fertility that are not recognized by lectin binding.

Small RNAs

The central dogma of molecular biology states that the transfer of genetic information proceeds from deoxyribonucleic acid (DNA) to ribonucleic acid (RNA) to protein (Crick, 1958). Historically, approximately 1.5% of the genome contained protein-coding sequences, while 98% was thought to be non-coding "junk" (Collins, *et al.*, 1998). Further research into this "junk" DNA has proved that the "junk" DNA (non-protein coding sequences) often has a function in cells. Approximately 80% of noncoding sequences play a role in gene expression as short- and long-non-coding RNAs (sRNA & lncRNA; Lander *et al.*, 2001).

Multiple categories of each type of non-coding RNA exist. The two main categories are long non-coding RNAs, such as transfer RNA or ribosomal RNA, and short non-coding RNAs (sRNA), which include small interfering RNAs (siRNAs), repeat-associated small interfering RNAs (rasiRNAs), small nuclear RNAs (snRNAs), microRNAs (miRNAs), and piwi-interacting RNAs (piRNAs). Each of these short non-coding RNAs has a different function: siRNAs are involved with gene regulation and viral defense, rasiRNAs orient heterochromatin in centromere activity, snRNAs function with the spliceosome complex, miRNAs are involved with gene regulation, and piRNAs help regulate transposon activity (Gavazzo *et al.*, 2013; Dogini *et al*., 2014).

The importance of miRNAs was first identified in nematodes. Researchers working with *C. elegans* observed that the sRNA lin-4 bound to and prevented translation of lin-14 (Lee *et al.*, 1993; Wightman *et al.*, 1993). Although initially considered a highly uncommon occurrence, a second miRNA, let-7, was discovered in *C. elegans* (Reinhart *et*

al., 2000). Let-7 appears to control late temporal transitions during development of species like *Drosophila*. Let-7 was also found to be present in a wide range of animal species including vertebrates, arthropods, and molluscs (Pasquinelli *et al*., 2000). Further research into *C. elegans* discovered 55 more miRNAs, several of which had potential cohorts in the human and *Drosophila* genome (Lau *et al*., 2001). As the number of identified miRNAs grew, a collaborative effort worked to establish a centralized naming system and database called miRBase (Griffiths-Jones *et al*., 2006). The latest release for miRBase (Release 21, http://www.mirbase.org) contains over 28,000 entries.

A miRNA is a small sequence ranging from 18 to 26 nucleotides long that can alter genome expression by down-regulating the translation of mRNA (Hutvagner, 2005). These small sequences act by binding to their complementary mRNA and down-regulating the translation of that area by blocking the translational activities. Approximately 60% of mammalian genes undergo regulation by miRNAs (Friedman *et al.*, 2009). The miRNA genes span the entire genome, including introns, exons, and intergenic regions (Rodriguez et. al., 2004).

In mammals, the biogenesis of miRNAs begins in the nucleus (Figure 1.2). The miRNA genes are transcribed by RNA polymerase-II to form the primary miRNA. Primary miRNAs have an imperfectly base-paired stem and hairpin loop structure that is approximately 70 nucleotides in length (Lee *et al*., 2002; Ambros *et al.*, 2003). The stem and loop structure contains both the mature miRNA and the hairpin turn; the stem and hairpin loop will be cleaved from the rest of the primary miRNA by Drosha, creating pre-miRNAs. These pre-miRNAs, which are still stem and loop structures, are transported into the cytoplasm by exportin-5 (Yi *et al.*, 2003). Pre-miRNAs are a complex of the

mature miRNA and the miRNA*. Dicer cleaves the hairpin loop, leaving only the stem (miRNA:miRNA* duplex). The RNA-induced silencing complex (RISC) identifies the mature miRNA and then targets specific binding sites on the 3' messenger RNA (mRNA) transcripts (Lai, 2005). The miRNA* is usually degraded, although recent research has found that some miRNA* act to alter translation themselves (Lytle *et al.*, 2007). Since determining that miRNA* can both regulate translation, miRNA and miRNA* are now referred to as 3' miRNA and 5' miRNA. These miRNAs may target only a specific, single mRNA, or target multiple mRNAs for translational down-regulation (Lim *et al*., 2005).

In miRNAs, the 'seed' region is the nucleotides from base 2 to base 8 or 9 (Lewis *et al.*, 2005). The seed area is particularly important for target recognition although the entire miRNA sequence is important for down-regulation of transcripts (Grimson *et al.*, 2007). A single mRNA can be regulated by numerous miRNAs (Yanaihara *et al.*, 2006). In mammals, around a third of miRNAs show tissue specificity and other miRNAs appear to vary in expression level across tissues (Malone and Hannon, 2009).

The piRNAs were discovered while screening for factors involved in germline cell maintenance in *Drosophila melanogaster* (Lin and Spradling, 1997). After their initial discovery in *D. melanogaster*, piRNAs were also found in other organisms (Cox *et al*., 1998), including mouse and rat germ cells (Aravin *et al*., 2006, Girard *et al*., 2006; Grivna *et al*., 2006; Lau *et al*., 2006; Watanabe *et al*., 2006). The piRNAs are slightly larger than miRNAs, (26 to 30nt in length) and are chiefly found in germ cell lines (Ghildiyal and Zamore, 2009). These piRNAs primarily function to regulate transposon activity and chromatin states through argonaute (AGO) and Piwi proteins (Malone and Hannon, 2009). Mutations disrupting piRNA pathways in fish and mice have been found

to cause germline-specific cell death and sterility (Aravin *et al*., 2007). Little is known about the biogenesis of piRNAs, apart from the fact that piRNAs are generated from RNaseIII cleavage of single-stranded DNA (Brennecke *et al*., 2007). Biogenesis of piRNAs differs among species (Carmell *et al*., 2007). Biogenesis of piRNAs does not appear to require a double-stranded RNA precursor or the activity of Dicer, leading to the suggestion that precursors of piRNAs are long segments of single-stranded DNA that are preferentially cleaved at uridine residues (Houwing *et al*., 2007). In mice, three Piwi proteins are required for male fertility: MILI, MIWI, and MIWI2, each of which is expressed at different stages of development (Carmell *et al*., 2007). Expression of MILI begins in the embryonic stages and persists through adulthood, coinciding with primordial germ cell migration into the developing gonad. The MIWI protein is expressed in adult testes, coinciding with the onset of meiosis in spermatogenesis. Expression of MIWI2 begins after MILI and ends three days postpartum, correlating with cell cycle arrest (Aravin *et al*., 2008).

The function of piRNAs primarily involves the suppression of transposons. Transposon activity is repressed by piRNAs through a transcriptional mechanism called the ping-pong cycle. Primary piRNAs recognize their complementary targets, at which point Piwi proteins are recruited to slice the RNA strand on either side of the transcript, producing a secondary piRNA. The secondary piRNAs then bind to the targets. Due to the high occurrence of Piwi proteins to contain catalytic domains, these proteins are very capable of target slicing (Brennecke *et al*., 2007).

In both mammals and *Drosophila*, piRNAs are transcribed from genomic clusters ranging from 50 to 100 kilobases (kb). The piRNAs can also be found within the 3'

untranslated region (UTR) of protein coding genes. Regardless of their derivation, the majority of piRNA clusters are specifically active in germ cells (Brennecke *et al*., 2007). It would appear that piRNAs play a larger role in epigenetic regulation rather than post-transcriptional regulation, as in the case of miRNAs (Kim *et al*., 2006). Epigenetic regulation describes the occurrence of external or environmental factors causing the modification of gene expression through changes in DNA methylation patterns, histone modification, or small RNAs. This revelation has led to the suggestion that the Piwi-piRNA pathway helps maintain biological robustness (Gangaraju *et al*., 2011). This process, also known as canalization, describes a population's ability to continue exhibiting the same phenotype, despite environmental or mutational affects that might drive a change in phenotype.

Previous Research in miRNA and piRNA

Extensive research has revealed the role that miRNAs play in mammals. A study involving Dicer mutations in zebrafish showed abnormal morphogenesis during gastrulation, brain formation, somitogenesis, and heart development (Giraldez *et al*., 2005). These results suggest that miRNAs play an essential role in overall animal development. A series of studies performed in knockout mouse models helps to illustrate the potential importance that miRNA and piRNA pathways play in murine development and spermatogenesis. Dicer plays a role in the function of Sertoli cells, which help to produce sperm cells, as well as the miRNA generation pathway. One study looked at the removal of Dicer in primordial germ cells (PGCs) of mice after embryonic day 10. This removal resulted in proliferation defects and issues with post-natal spermatogenesis (Maatouk *et al*., 2008). Due to the damage to PGCs, this model was not appropriate for

investigating the exact role of Dicer in spermatogenesis. When Dicer was deleted in mice after birth, meiotic progression of spermatocytes was delayed and haploid differentiation was severely affected (Greenlee *et al*., 2012; Wu *et al*., 2012). Chang *et al*. (2012) used a Dicer knockout model that was activated later in the spermatogenic process after birth, resulting in post-meiotic defects like abnormal head morphology. The takeaway message from all of these studies is that Dicer is important for maintenance of spermatogenesis. The earlier in time the Dicer knockout occurred, the more defects were seen. Dicer plays a key role in the processing of miRNAs, suggesting that some of the effects seen in the knockout models could be due to a lack of miRNA regulation. In a study characterizing the piRNAs of the mouse, there were differences in the predicted piRNAs for elongating spermatids and round spermatids. These results also support the hypothesis that piRNAs suppress specific genes to ensure germ cell development (Yuan *et al*., 2016).

Research involving miRNAs in livestock species usually investigate miRNAs in relation to traits with economic importance. Studies have characterized miRNAs in cattle embryos, brain, liver, and muscles, to mention a few studies (Coutinho *et al*., 2007; Jin *et al*., 2009). Comprehensively, these studies show that miRNAs may be expressed in all tissues or be tissue-specific. Some appear to play key roles in altering translation via high expression abundance while others appear to be lowly expressed. Expression of miRNAs has been reported to change at different points of development. In regards to fertility, much of the research has focused on the females, investigating oocytes, ovarian tissues, and embryonic tissue (Miles *et al*, 2012; Huang *et al*., 2011; Coutinho *et al*., 2006). The miRNAs in the tissue of the testes across species have been characterized,

but to the best of our knowledge no research investigating the relationship between expression of miRNAs in seminal plasma and fertility levels in bulls has been published.

Populations of miRNAs also differ based on developmental stage of the animal. Lian and others (2012) completed a study that involved the sequencing of small RNAs from swine testes. Samples of testicular tissue were collected from sexually immature piglets 30 days after birth and mature boars 180 days after birth. The RNA was extracted and samples from each group (mature and immature) were pooled for sequencing. The sRNA sequences were generated using Solexa deep sequencing and mapped to the swine genome. Despite having a similar number of clean reads, the mature pigs had 1.5-fold higher numbers of unique sequencing reads than the immature pigs, suggesting a difference in the composition of miRNAs present. Differences in the average length of small RNAs between the two groups were also observed. The immature pool reads were primarily 21 to 23nt in length, indicating that most of the sRNAs expressed were miRNAs. In the mature pool, the majority of the read lengths ranged from 26 to 30nt, indicating higher levels of piRNA expression. A total of 122 known miRNAs were differentially expressed between mature and immature pig testes (Lian *et al*., 2012). The results from this study support the idea that miRNA pathways are activated or inactivated at certain stages of mammalian development in the gonadal tissue. This study also reveals differences in sRNA populations at different stages in development, suggesting miRNA drives development and piRNAs play a larger role after sexual maturity is reached. These results are supported by a study characterizing sRNAs in human spermatozoa. The researchers found that, of the sRNAs extracted, 17% were piRNAs and 7% were miRNAs (Krawetz *et al*., 2011).

A similar study was completed on Holstein bulls and cows slaughtered at two different time points (Huang *et al*., 2011). Two cows and two bulls were slaughtered 3 days after birth and then two cows and two bulls were slaughtered at 2.5 years of age. Tissue samples from the testis and ovaries were collected to compare miRNA expression pre- and post-sexual maturity. All samples of the same gender were pooled together prior to sequencing. Solexa deep sequencing was performed and samples were compared to miRBase and mapped to the bovine genome. A total of 122 and 136 novel miRNAs were discovered in the testes and ovarian tissue, respectively. Six of these novel miRNAs are considered to be cattle-specific. In the testes, eight known miRNAs accounted for 63.5% of the total sequencing reads. In the ovarian tissue, seven miRNAs were dominantly expressed, each with more than 100,000 reads. Around 30.5% of the known bovine miRNAs showed >2-fold expression differences between the different tissue types (Huang *et* al., 2011).

Although less work has been completed investigating sRNA expression and fertility, at least one bovine study and several studies in humans have been published. Sperm cell miRNAs from high and low fertility bulls based on an industry progeny testing program were isolated to examine the relationship between miRNA expression and male fertility (Govindaraju *et al*, 2012). Four Holstein bulls were selected and ranked based on conception rates from their breeding records adjusted for environmental and herd factors. All of the bulls had acceptable progeny rates, however they were the most divergent of the 998 bulls considered. The miRNAs were extracted from the sperm cells of each bull. The samples were centrifuged to separate sperm cells and seminal plasma, which was discarded. The sperm cell pellet left behind was the source of RNA

used for sequencing. Only seven differentially expressed miRNAs that matched with miRNAs previously identified in humans were discovered between high and low fertility groups. The differentially expressed miRNAs appear to be conserved across species, but have not yet been annotated so the functions remain unknown. These results supported their hypothesis that miRNAs played a potential role in the regulation of genes, which affect aspects of spermatogenesis (Govindaraju *et al*., 2012).

Differences in seminal miRNA expression in men with conditions that cause male infertility have previously been discovered (Wu *et al*., 2012; Wu *et al*., 2013; Wang *et al*., 2011). Human male infertility research has the convenience of using specific infertility diagnoses within their research. Males who exhibited azoospermia (no sperm cells in the semen) and asthenozoospermia (highly decreased motility of sperm cells) were sampled and compared with fertile controls (Wang *et al*., 2011). Samples were pooled for sequencing and RT-PCR was used to verify the results. The sequencing analysis found 19 differentially expressed miRNAs between the patient groups and controls. RT-PCR results indicated seven miRNAs that were significantly decreased in azoospermia and increase in asthenozoospermia patients. These results indicate that differences in miRNA expression could contribute to different types of infertility. Other studies also examined the relationship between miRNAs and infertility diagnoses. An initial study investigating a few specific miRNAs in conjunction with azoospermia and oligozoospermia (low concentrations of sperm cells) led to an overall characterization of miRNAs in the seminal plasma of patients with these infertility diagnoses (Wu *et al*., 2011; Wu *et al*., 2012).

Within a semen sample, sRNAs may originate from the testes, epididymis, seminal vesicles, prostate, or bulbourethral glands. The objective of this next study was to identify whether sRNAs originated from the testes and epididymis or other male reproductive tissues (Hu *et al*., 2013). The major goal of this study was to identify potential biomarkers for male infertility. For a biomarker to be reliable, it should be specific to the tissue or organ of interest, or secreted specifically from the tissue of interest. For male infertility, this meant identifying sRNAs that are specifically secreted into the seminal plasma by the testes or epididymis. Vasectomized semen samples were compared with normal semen samples. The miRNA profiles between healthy donors and vasectomized men (no sperm in seminal fluid) were compared, providing an interesting perspective on miRNA profiles not seen in the livestock studies. The healthy donors' seminal plasma contained secretions from the testis, epididymis, seminal vesicles, prostate, and bulbourethral glands. This study found 84 miRNAs that appeared to be predominantly derived from the testes and epididymis. This study only identified piRNAs in seminal plasma from the healthy donors. The miRNA results were validated using qRT-PCR and additional samples, and confirmed that 61 of these miRNA were present in samples including testis and epididymal secretions at considerably higher levels (>4-fold) than in the vasectomized samples (Hu *et al*., 2013). This study also found that there were certain miRNAs that were expressed at similar levels in both treatment groups, implicating the seminal vesicles and prostate and potential secretory sources of miRNAs.

Conclusion

The studies described previously make an argument for investigating the miRNA and piRNA profiles in the seminal plasma of beef bulls. Although miRNAs have been studied in the spermatozoa of beef bulls, miRNAs in the seminal plasma have not yet been investigated. The discovery of miRNA expression differences between fertility levels in the seminal plasma of human males suggests it could be worth exploring the sRNA expression in seminal plasma. Studies have also shown that piRNAs play a role in fertility in mice and human males, which suggests that they might also be a factor in bull fertility.

Literature Cited

Amann, R.P. and R.H. Hammerstedt. 1993. *In vitro* evaluation of sperm quality: an opinion. J Androl 14(6): 397-406.

Ambros, V., B. Bartel, D.P. Bartel, C.B. Burge J.C. Carrington, X. Chen, G. Dreyfuss, S.R. Eddy, S. Griffiths-Jones, M. Marshall, M. Matzke, G. Ruvkun, and T. Tuschl. 2003. A uniform system for microRNA annotation. RNA 9: 277-279.

Aravin, A.A., R. Sachidanandam, D. Bourc'his, C. Schaefer, D. Pezic, K. Toth, T. Bestor, and T.G. Hannon. 2008. A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. Mol Cell 31: 785-799.

Aravin, A.A., G.J. Hannon, and J. Brennecke. 2007. The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. Science 314: 761-764.

Aravin, A.A., D. Gaidatzis, S. Pfeffer, M. Lagos-Quintana, P. Landgraf, N. Iovino, *et al*. 2006. A novel class of small RNAs bind to MILI protein in mouse testes. Nature 442: 203-207.

Brennecke, J., A.A. Aravin, A. Stark, M. Dus, M. Kellis, R. Sachidanandam, and G.J. Hannon. 2007. Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. Cell 128: 1089-1103.

Brinks, J.S., M.J. McInerney, and J.P. Chenoweth. 1978. Relationship of age at puberty in heifers to reproductive trains in young bulls. Proc West Sect Am Soc Anim Sci 29: 28-30.

Chang, Y.F., J.S. Lee-Chang, J.S. Imam, K.C. Buddavarapu, S.S. Subaran, A.P. Sinha-Hikim, M. Gorospe, and M.K. Rao. 2012. Interaction between microRNAs and actin-associated protein Arpc5 regulates translational suppression during male germ cell differentiation. Proc Natl Acad Sci 109:5750-5755.

Cheng, F.P., A. Fazeli, W.F. Voorhout, A. Marks, M.M. Bevers, and B. Colenbrander. 1996. Use of the peanut agglutinin to assess the acrosomal status and the zona pellucida-induced acrosome reaction in stallion sperm. J Androl 17(6): 674-682.

Collins, F.S., A. Patrinos, E. Jordan, A. Chakravarti, R. Gesteland, and L. Walters. 1998. New Goals for the U.S. Human Genome Project: 1998-2003. Science 282: 682-689.

Cox, D.N., A. Chao, J. Baker, L. Chang, D. Qiao, and H. Lin. 1998. A novel class of evolutionarily conserved genes defied by piwi are essential for stem cell self-renewal. Genes Dev 12: 3715-3727.

Coutinho, L.L., L.K. Matukumalli, T.S. Sonstegard, C.P. Van Tassell, L.C. Basbarre, A.V. Capuco, and T.P. Smith. 2007. Discovery and profiling of bovine microRNAs from immune-related embryonic tissues. Physiol Genomics 29: 35-43.

Crick, F.H. 1958. On protein synthesis. Symp Soc Exp Biol 12: 138-163.

Cross, N.L. and S.K. Watson. 1994. Assessing acrosomal status of bovine sperm using fluoresceinated lectins. Therio 41(1): 89-98.

Dogini, D.B., V.D.B. Pascoal, S.H. Avansini, A.S. Veiri, T.C. Pereira, and I. Lopes-Cendes. 2014. The new world of RNAs. Genet and Mol Bio 37(1): 285-293.

Ellis, R.W. 2007. Bull breeding soundness exams and beyond: changes in reproductive soundness. Proceedings, Applied Reproductive Strategies in Beef Cattle, Billings, MT.

Fatima, A. and D.G. Morris. 2013. MicroRNAs in domestic livestock. Physiol Genomics 45: 685-696.

Friedman, R.C., K.K. Fahr, C.B Burge, and D.P Bartel. 2009. Most mammalian mRNAs are conserved targets of miRNAs. Genome Res 19: 92-105.

Gangaraju, V.K., H. Yin, M.M. Weiner, J. Wang, X.A. Huang, and H. Lon. 2011. *Drosophila* Piwi functions in Hsp90-mediated suppression of phenotypic variation. Nat Genet 43: 153-158.

Gavazzo, P., M. Vassalli, D. Costa, and A. Pagano. 2013. Novel ncRNAs transcribe by Pol III and elucidation of their functional relevance by biophysical approaches. Front Cell Neurosci 7: 203-215.

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

Giraldez, A.J., R.M. Cinalli, M.E. Glasner, A.J. Enright, J.M. Thomson, S. Baskerville, S.M. Hammond, D.P. Bartel, and A.F. Schier. 2005. MicroRNAs regulate brain morphogenesis in zebrafish. Science 308: 83-838.

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

Govindaraju, A., A. Uzun, L. Robertson, M.O. Atli, A. Kaya, E. Topper, E.A. Crate, J. Padbury, A. Perkins, and E. Memili. 2012. Dynamics of microRNAs in bull spermatozoa. Repro Bio and Endo 10: 82-92.

Graham, J.K. 2001. Assessment of serm quality: a flow cytometric approach. Anim Reprod Sci 68: 239-247.

Greenlee, A.R., M.S. Shiao, E. Snyder, F.W. Buaas, T. Gu, T.M. Stearns, M. Sharma, E.P. Murchison, G.C. Puente, and R.E. Braun. 2012. Deregulated sex chromosome gene expression with male germ cell-specific loss of Dicer1. PLoS One 7: e46359.

Griffiths-Jone, S. R. Grocock, S. van Dongen, A. Bateman and A. Enright. 2006. MiRBase: microRNA sequences, targets and gene nomenclature. Nucl Acids Res 34: D140-D144.

Grimson, A., K.K. Fahr, W.K. Johnston, P. Garrett-Engele, L.P. Lim, and D.P. Bartel. 2007. MicroRNA targeting specificity in mammals: determinants beyond seed pairing. Mol Cell 27: 91-105.

Grivna, S.T., E. Beyret, Z. Wang, and H. Lin. 2006. A novel class of small RNAs in mouse spermatogenic cells. Genes Dev 20: 1709-1714.

Gu, Z., S. Eleswarapu, and H. Jiang. 2007. Identification and characterization of microRNAs from the bovine adipose tissue and mammary gland FEBS Lett 581: 981-988.

Hershko, A. and A. Ciechanover. 1998. The ubiquitin system. Annu Rev Biochem 67: 425-479.

Houwing, S., L.M. Kamminga, E. Berezikov, D. Cronembold, A. Girard, H. van den Elst, D.V. Filippov, H. Blaser, E. Raz, C.B. Moens, R.H. Plasterk, G.J. Hannon, B.W. Draper, and R.F. Ketting. 2007. A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. Cell 129: 69-82.

Hu, L., C. Wu, C. Guo, H. Li, and C. Xiong. 2014. Identification of microRNAs predominately derived from testis and epididymis in human seminal plasma. Clin Biochem 47(10-11): 967-972.

Huang, J., Z. Ju, Q. Li, Q. Hou, C Wang, J. Li, R. Li, L. Wang, T. Sun, S. Hang, Y. Gao, M. Hou, and J. Zhong. 2011. Solexa sequencing of novel and differentially expressed microRNAs in testicular and ovarian tissues in Holstein cattle. Int J Biol Sci 7(7): 1016-1026.

Hutvagner, G. 2005. Small RNA asymmetry in RNAi: Function in RISC assembly and gene regulation. FEBS Lett. 579: 5850-5857.

Hutvagner, G., J. McLachlan, A.E. Pasquinelli, E. Balint, T. Tuschl, and P.D. Zamore. 2001. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. Science 293: 834-838.

Jin, W., J.R. Grant, P. Stothard, S.S. Moore, and L.L. Guan. 2009. Characterization of bovine miRNAs by sequencing and bioinformatics analysis. BMC Mol Biol 10: 90.

Kastelic, J.P. and J.C. Thundathil. 2008. Breeding soundness and semen analysis for predicting bull fertility. Reprod Domest Anim 43(suppl. 2): 368-373.

Kennedy, C.E., K.B. Krieger, M. Sutovsky, W. Xu, P. Vargovic, B.A. Didion, M.R. Ellersieck, M.E. Hennessy, J. Verstegen, R. Oko, and P. Sutovsky. 2014. Protein expression pattern of PAWP in bull spermatozoa is associated with sperm quality and fertility following artificial insemination. Mol Reprod Dev 81: 436-449.

Khazaie, Y. and M.H.N. Esfahani. 2014. MicroRNA and Male Infertility: A Potential for Diagnosis. Int J Fertil Steril 8(2): 113-118.

Kim, V.N. 2006. Small RNAs just got bigger: Piwi-interacting RNAs (piRNAs) in mammalian testes. Genes Dev 20: 1993-1997.

Kowalczykiewicz, D., A. Swiercz, L. Handschuh, K. Lesniak, M. Figlerowicz, and J. Wrezesinski. 2014. Characterization of *Sus scrofa* small non-coding RNAs present in both female and male gonads. PLoS One 9(11): e113249.

Krawetz, S.A., A. Kruger, C. Lalancette, R. Tagett, E. Anton, S. Draghici, and M.P. Diamond. 2011. A survey of small RNAs in human sperm. Hum Reprod 26(12): 3401-3412.

Lai, E.C. 2005. Whys and wherefores of miRNA-mediated regulation. Curr Biol 15: R458-R460.

Lander, E.S., L.M. Linton, B. Birren, C. Nusbaum, M.C. Zody, J. Baldwin *et al.* 2001. Initial sequencing and analysis of the human genome. Nature 409: 860-920.

Lau, N.C., A.G. Seto, J. Kim, S. Kuramochi-Miyagawa, T. Nakanoo, D.P. Bartel, and R.E. Kingston. 2006. Characterization of the piRNA complex from rat testes. Science 313: 363-367.

Lau, N.C., L.P. Lim, E.G. Weinstein, and D.P. Bartel. 2001. An abundant class of tiny RNA with probable regulatory roles in *Caenorhabditis elegans*. Science 294: 858-862.

Lee, R.C. and V. Ambros. 2001. An extensive class of small RNA in *Caenorhabditis elegans*. Science 294: 862-864.

Lee, R.C., R.L. Feinbaum, and V, Ambros. 1993. The *C. elegans* heterochronic gene *lin-4* encode small RNAs with antisense complimentary to *lin-14*. Cell 75: 843-854.

Leblond, C.P. and Y. Clermont. 1952. Definition of the stages of the cycle of the seminiferous epithelium in the rat. Biol Testes 55: 548-573.

Lewis, B.P., C.B. Burge, and D.P. Bartel. 2005. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 120(1): 15-20.

Lian, C, B. Sun, S. Niu, R. Yang, B. Lui, C. Lu, J. Menga, Z. Qiu, L. Zhang, and Z. Zhao. 2012. A comparative profile of the microRNA transcriptome in immature and mature porcine testes using Solexa deep sequencing. FEBS J 12: 1742-4658.

Lim, L.P., N.C. Lau, P. Garrett-Engele, A. Grimson, J.M. Schelter, J. Castle, D.P. Brtel, P.S. Linsley, and J.M. Johnson. 2005. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature 433(7027): 769-773.

Lin, H. and A.C. Spradling. 1997. A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the *Drosophila* ovary. Develop 124: 2463-2476.

Lytle, J.R., T.A. Yario, and J.A. Steitz. 2007. Target mRNAs are repressed as efficiently by microRNA-binding sites in the 5' UTR as in the 3' UTR. Proc Natl Acad Sci 104(23): 9667-9672.

Maatouk D.M., K.L. Loveland, M.T. McManus, K. Moore, and B.D. Harfe. 2008. Dicer1 is required for differentiation of the mouse male germline. Biol Reprod 79: 696-703.

Malone, C. and G. Hannon. 2009. Small RNAs as guardians of the genome. Cell 136: 656-668.

Mathews, K.H. and S.D. Short. 2001. The beef cow replacement decision. J Agribus 19(2): 191-211.

McDaneld, T.G. 2009. MicroRNA: Mechanism of gene regulation and application to livestock. J Anim Sci 57(E. Suppl.): E21-E28.

Miles, J.R., T.G. McDaneld, R.T. Wiedmann, R.A. Cushman, S.E. Echternkamp, J.L. Vallet, and T.P.L. Smith. 2012. MicroRNA expression profile in bovine

cumulus-oocyte complexes: possible role of let-7 and miR-106a in the development of bovine oocytes. Anim Reprod Sci 130: 16-26.

Odhiambo, J.F., M. Sutovsky, J.M. DeJarnettte, C. Marshall, and P. Sutvosky. 2011. Adaptation of ubiquitin-PNA based sperm quality assay for semen evaluation by a conventional flow cytometer and a dedicated platform for flow cytometric semen analysis. Theriogenology 76(6): 1168-1176.

Pasquinelli, A.E., B.J. Reinhart, F. Slack, M.Q. Martindale, M.I. Kuroda, B. Maller, D.C. Hayward, E.E. Ball, B. Degnan, P. Muller, J. Spring, A. Srinivasan, M. Fishman, J. Finnerty, J. Corbo, M. Levine, P Leahy, E. Davidson, and G. Ruvkun. 2000. Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. Nature 408: 86-89.

Reinhart, B.J., F.J. Slack, M. Basson, A.E. Pasquinelli, J.C. Bettinger, A.E. Rougvie, H.R. Horvitz, and G. Ruvkun. 2000. The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. Nature 403: 901-906.

Rodriguez, A., S. Griffiths-Jones, J.L. Ashurst, and A. Bradley. 2004. Identification of mammalian microRNA host genes and transcription units. Genome Res 14: 1902-1910.

Saacke, R.G. 2008. Sperm morphology: its relevance to compensable and uncompensable traits in semen. Theriogenology 70: 473-478.

Senger, P.L. 2003. Pathways to pregnancy and parturition. 2nd ed. Current Conceptions, Pullman WA. p. 216-223.

Sharpe, R.M. 1994. Regulation of spermatogenesis. In: E. Knobil and J.D. Neill, editors, The physiology of reproduction. Raven Press, New York, NY. P. 1363-1434.

Sutovsky, P. 2009. Proteomic analysis of mammalian gametes and sperm-oocyte interactions. Soc Reprod Fertil Suppl 66: 103-116.

Sutovksy, P., M. Aarabi, A. Miranda-Viquete, and R. Oko. 2015. Negative biomarker-based male fertility evaluation: sperm phenotypes associated with molecular-level anomalies. Asian J Androl 17: 554-560.

Sutovsky, P., E. Neuber, and G. Schatten. 2002. Ubiquitin-dependent sperm quality control mechanism recognizes spermatozoa with DNA defects as revealed by dual ubiquitin-TUNEL assay. Mol Reprod Dev 61(3): 406-413.

Sutovsky, P., R. Moreno, J. Ramalho-Santos, T. Dominko, W.E. Thompson, and G. Schatten. 2001a. A putative, ubiquitin-dependent mechanism for the recognition and elimination of defective spermatozoa in the mammalian epididymis. J Cell Sci 114(9): 1665-1675.

Sutovksy, P., Y. Terada, and G. Schatten. 2001b. Ubiquitin-based sperm assay for the diagnosis of male factor infertility. Human Reprod 16(2): 250-258.

Wang, C., C. Yang, X. Chen, B. Yao, C. Yang, C. Zhu, L. Li, J. Wang, X. Li, Y. Shao, Y. Liu, J. Ji, J. Zhang, K. Zen, C.Y. Zhang, and C. Zhang. 2011. Altered profile of seminal plasma microRNAs in the molecular diagnosis of male infertility. Clin Chem 57(12): 1722-1731.

Watanabe, T., A. Takeda, T. Tsukiyama, K. mise, T. Okuno, H. Sasaki, N. Minami, and H. Imai. 2006. Identification and characterization of two novel classes of small RNAs in the mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes. Genes Dev 20: 1732-1743.

Wightman, B., I. Ha, and G. Ruvkun. 1993. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in *C. elegans*. Cell 75: 855-862.

Wiltbank, J.N. and N.R. Parish. 1986. Pregnancy rate in cows and heifers bred to bulls selected for semen quality. Theriogenology 25(6): 779-783.

Wu, Q., R. Song, N. Ortogero, H. Zheng, R. Evanoff, C.L. Small, M.D. Grisworld, S.H. Namekawa, H.Royo, J.M. Turner, and W. Yan. 2012. The Rnase III enzyme DROSHA is essential for microRNA production and spermatogenesis. J Biol Chem 287: 25173-25190.

Wu, W., Y. Qin, J. Dong, J. Dai, C. Lu, X. Guo, Y. Zhao, W. Zhang, B. Hang, J. Sha, H. Shen, Y. Xia, Z. Hu, and X. Wang. 2013. Genome-wide microRNA expression profiling in idiopathic non-obstructive azoospermia: significant up-regulation of miR-141, miR-429, and miR-7-1-3p. Hum Reprod 28(7): 1827-1836.

Wu, W., Z. Hu, Y. Qin, J. Dong, J. Dai, C. Lu, W Zhang, H. Shen, Y. Xia, and X. Wang. 2012. Seminal plasma microRNAs: potential biomarkers for spermatogenesis status. Mol Hum Reprod 18(10): 489-497.

Yadav, R.P. and N. Kotaja. 2014. Small RNAs in spermatogenesis. Mol and Cell Endo 382: 498-508.

Yanaihara, N., N. Caplen, E. Bowman, M. Seike, K. Kumanoto, M. Yi, R.M. Stephen, A. Okamoto, J. Yokota, T. Tanaka G.A. Calin, C.G. Liu, C.M. Croce, and C.C. Harris. 2006. Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. Cancer Cell 9: 189-198.

Yi, R., Y. Qin, I.G. Macara, and B.R. Cullen. 2003. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes Dev 17: 3011-3016.

Yuan, J., P. Zhang, Y. Cui, J. Wang, G. Skogerbo, D. Huang, R. Chen, and S. He. 2016. Computational identification of piRNA targets on mouse mRNAs. Bioinfor 32(7).

Figure 1.1 Relationship between GnRH, LH, and FSH in the Male (Senger, 2003)

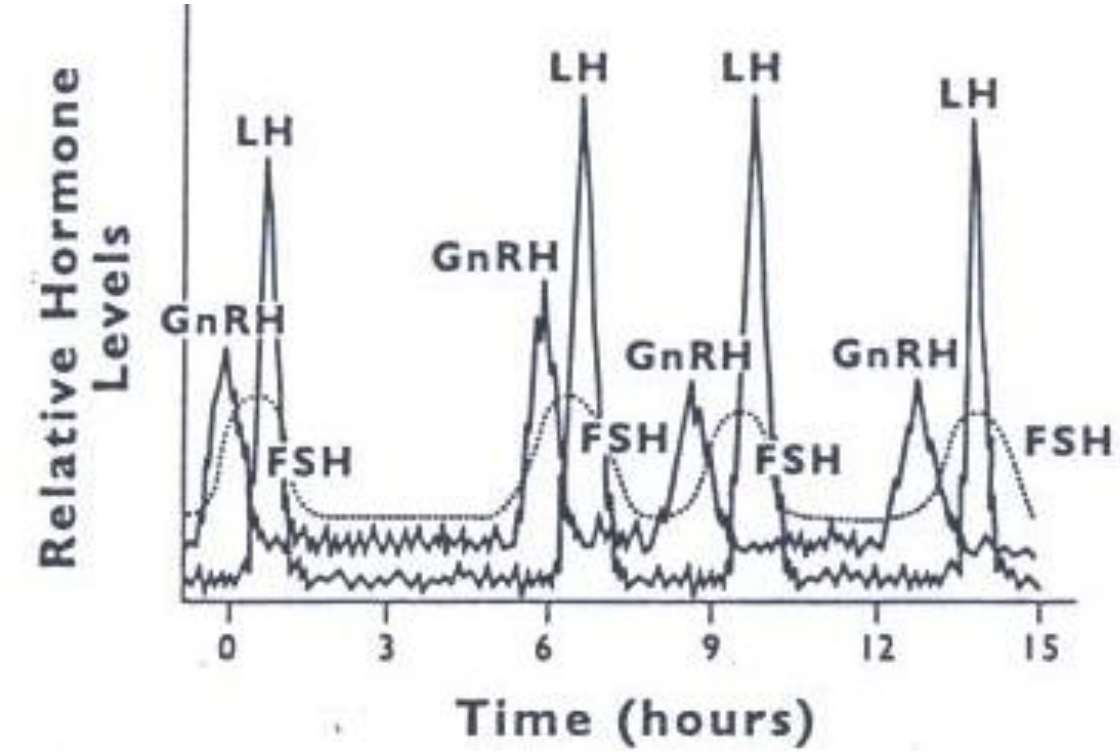
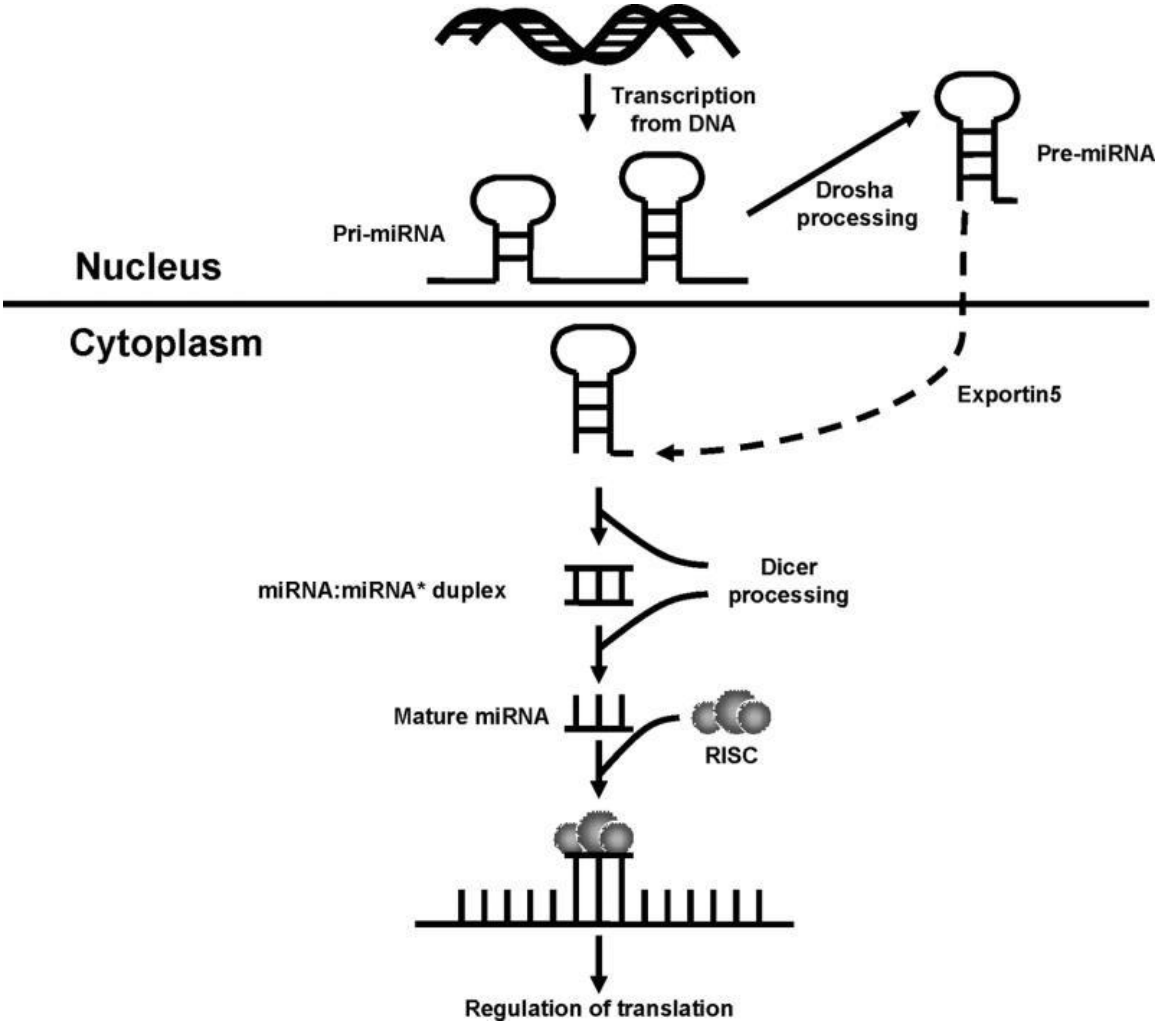


Figure 1.2 MiRNA Biogenesis in the Animal Cell (McDaneld, 2009).



Chapter 2

Characterization of miRNAs and piRNAs in the Seminal Plasma of Beef Bulls with Predicted High and Low Fertility

Introduction

Reproductive performance in cattle has a significant economic impact on the beef industry (Ramsey *et al*., 2005). A failure to breed in a single bull is much more economically costly to a producer than the same failure in a single female. A single bull will service multiple females in a breeding season and therefore has a higher impact on the number of successful pregnancies within a herd. Although most bulls undergo breeding soundness exams (BSE) prior to the breeding season, the information gathered provides a single snapshot in time of their semen quality. Being able to consistently identify bulls with superior fertility earlier in their lifetime would be helpful to producers' breeding programs.

One of the key processes contributing to bull fertility is spermatogenesis, the process during which the male produces sperm samples for ejaculation. Spermatogenesis requires a precise, specific series of molecular events to prevent genetic abnormalities in the offspring. Part of this process involves the development of the acrosome. The acrosome contains proteolytic enzymes which degrade the zona pellucida of an oocyte, allowing sperm to penetrate the oocyte and complete fertilization. Ubiquitin is a protein that targets substrates for proteolysis. Ubiquitin concentrations have been correlated with overall fertility in bulls (Sutovsky *et al*., 2001). Lectins, such as *Arachis hypogaea*/peanut agglutinin (PNA) and *Lens culinaris* agglutinin (LCA), are also correlated with ubiquitin and can be used to estimate fertility of a bull (Sutovsky *et al*., 2015). However, these assays are not conducive to large-scale implementation. If a biomarker were discovered, a quick PCR test could be developed, making implementation into large-scale production more feasible. Finding an easily identifiable

marker in semen samples indicating overall fertility would be beneficial to beef producers.

Small non-coding RNAs, specifically microRNAs (miRNAs) and piwi-interacting RNAs (piRNAs), have been reported to play a role in fertility in multiple species (Yadav and Kotaja, 2014; Houwing *et al*., 2014). The miRNAs post-transcriptionally regulate gene expression, while piRNAs are germline-specific sRNAs that help to regulate transposon activity. However, the relationship of miRNA and piRNA expression with markers of semen quality in beef bulls is unclear. Further, piRNAs have not been characterized in seminal plasma of beef bulls. Our objective was to characterize and identify differentially expressed sRNAs in the seminal plasma of bulls with different predicted fertility levels of semen samples, as determined by expression of lectins PNA and LCA. Discovered sRNAs have the potential to be used as a biomarker for overall fertility in beef bulls.

Materials and Methods

Sample Collection

Semen samples were collected from 35 bulls 1.2 ± 0.6 years old via electro-ejaculation. Bulls belonged to three different herds: the SDSU Research and Teaching Unit, Brookings, SD (n = 24), a herd in Scotland, SD (n = 9), and a herd in Olivet, SD (n = 2). Breed composition varied from purebred Angus to Angus-influenced. Semen samples were snap frozen using liquid nitrogen. Samples were centrifuged at 1200 x g for 15 minutes at 4 °C to separate seminal plasma and sperm cells. Seminal plasma for each sample was removed and stored in a new microcentrifuge tube. Phosphate buffered saline (PBS; 1.5mL) was added and tubes were centrifuged a second time for 15 minutes at 1,200 x g at 4 °C to ensure complete separation of sperm cells and seminal plasma. Supernatant was removed and stored in a new microcentrifuge tube. Both sets of tubes were stored at -80 °C prior to use. The sperm cell pellet was mixed with another 1.5 mL of PBS with 4% paraformaldehyde. The mixture of PBS and paraformaldehyde ensured sperm cells were fixed for later analysis. These cells were subsequently shipped to the University of Missouri for lectin analyses.

Lectin Analyses

Sperm samples were analyzed by Dr. Peter Sutovsky's laboratory at the University of Missouri (UM) for fertility markers PNA and LCA by flow cytometry (Odhiambo *et al*., 2011). The cut-off point used to discard debris and non-cellular junk was determined based on the forward scatter (FSC) vs side scatter (SCC) plots and the FSC log vs SCC log plots. These plots are based on size distribution of cells, allowing

lectins to be identified and quantified. Samples identified by the laboratory at UM as potentially bad samples due to poor staining or excess cellular debris were discarded. Samples were then ranked based on their PNA values. The three samples with the highest PNA scores and the three lowest PNA scores were identified. The same process was followed with the LCA values for each sample and the six samples identified using PNA matched those found using LCA. Based on these results, three semen samples that had high predicted fertility and three semen samples with low predicted fertility were selected.

RNA Isolation

Total RNA, including sRNA, was extracted from all seminal plasma samples using the Qiagen (Qiagen, Redwood City, CA) miRNeasy Serum/Plasma kit following manufacturer's instructions (Appendix A). Quantity and quality of RNA was assessed using a Nanodrop1000 (Thermo Scientific, Waltham, MA). The RNA quality of samples being sequenced was assessed using an Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA).

Sequencing Libraries

The six samples selected based on divergent PNA and LCA binding were prepared for sequencing at the USDA Meat Animal Research Center (MARC) in Clay Center, NE. Samples were prepared using the Illumina TruSeq Small RNA Library PrepKit v2 (Illumina, San Diego, CA) according to manufacturer's recommendations. The first step of the library preparation involved ligating adaptors onto the 5' phosphate and 3' hydroxyl group of the miRNAs. The adaptors are specifically designed to bind to

the 3' hydroxyl group and 5' phosphate groups that result from the enzymatic cleavage by processing enzymes, including Dicer, that process miRNAs in the bull. The single-stranded cDNA sequence was synthesized through RT-PCR using primers that can bind to the adaptors. This cDNA strand was amplified via PCR with a common primer and a primer containing an index sequence. Each of the six selected samples was prepared with a different index sequence. These indices were used to identify the sequence products of each sample after sequencing. After index sequences were added, cDNA constructs were purified and selected based on size (15-40 bp) using polyacrylamide gel electrophoresis. These final samples were checked for quality using an Agilent Bioanalyzer 2100. Samples were sequenced on an Illumina MiSeq using the MiSeq Reagent Kit v2 – 50 cycles at the USDA MARC facility. Paired-end reads were sequenced on a single-lane flow cell.

Bioinformatics Approach

After being sequenced, samples were uploaded to BaseSpace (http://basespace.illumina.com), Illumina's web-based platform for analysis and sharing of sequencing results. Using BaseSpace, the adaptors were trimmed from the sequence reads and the paired-end reads were separated. A number of pathways could be followed in order to analyze the sequencing data and identify differentially expressed miRNAs and piRNAs. The process followed for this study involved comparing the sequences to already discovered sRNAs and testing for differential expression between predicted semen fertility levels (Figure 2.1).

Paired-end reads were transferred into CLC Genomics Workbench (CLC Bio, Qiagen, Waltham, MA 02451). Adaptors were trimmed from sequences and bases with

quality scores less than 20 were discarded. Reads were compared to known miRNAs and piRNAs. MiRBase (Release 21, http://www.mirbase.org) is the database for all known miRNAs and is linked with CLC Genomics workbench. Known bovine, porcine, murine, and human miRNAs were included. No central location for all known piRNAs exists. Therefore, all known piRNAs that could be identified through the NCBI database were uploaded to CLC Genomics workbench for comparison. Expression levels of known sRNAs discovered within the sequence data were compared between high and low fertility marker groups. The CLC Genomics software utilized a Baggerley's test, or a weighted T-test, to test for statistical significance between groups. The Baggerley's test also compares the proportions of counts in one sample group versus another and the groups are weighted based on the size of their read counts (Baggerly *et al*., 2003). Fold changes, weighted proportions, Bonferroni-adjusted P-values, and false discovery rate (FDR)-corrected P-values were also calculated between high and low fertility marker groups.

After identifying the differentially expressed ($P < 0.05$) miRNAs, the potential targets of these miRNAs were identified using Exiqon's miRSearch 3.0 program (Exiqon Inc, Woburn, MA). Exiqon does not include bovine miRNAs, so the sequences were compared with the human miRNAs. miRSearch provides potential target genes, disease information, and tissues where the miRNA has been previously discovered.

Real-Time RT-PCR

Candidate piRNAs were selected based on sequencing results of the initial six seminal plasma samples. From the list of piRNAs that were differentially expressed

between high and low fertility samples, candidates were chosen based on having fold change >5 between high and low samples and a high number of read counts. Low numbers of reads do not indicate a reliable difference in sRNA expression.

Complementary DNA (cDNA) was reverse-transcribed using the miScript II RT kit (Qiagen) following manufacturer's instructions (Appendix B). These reactions included 50ng of RNA in a 20 uL reaction. A pooled RNA sample was prepared by mixing equal amounts of each RNA sample. This RNA pool was used to generate cDNA to be used for generation of standard curves and the no reverse-transcriptase control.

Real time-polymerase chain reaction was performed using the Qiagen SYBR Green master mix kit and universal primers (Appendix C). Sequences for the candidate piRNAs were acquired using the National Center for Biotechnology Information (NCBI, www.ncbi.nlm.nih.gov) database and miScript forward primers were designed using the custom design option offered by Qiagen. The universal primer is a reverse primer designed by Qiagen to work in conjunction with their miScript forward primer assays. In place of a housekeeping gene, a spiked-in synthetic miRNA (*C. elegans* miR-39-1) was added to each sample during the RNA extraction at a concentration of $5.6x10^8$ copies/uL. Since research in miRNA has not identified any single candidate miRNA that is consistently present in tissues that could be used to normalize PCR results, the spike-in control was used instead. The 25 uL reactions contained 1 uL of cDNA, 2.5 uL of the universal primer, 2.5 uL of the pi-RNA specific primer assay, 12.5 uL of SYBR green master mix, and 6.5uL of RNase-free water. The standard curve was generated using a serial dilution of the pooled RNA sample. This dilution resulted in an exponential difference in copies of the spike-in control: $1x10^6$ copies/uL, $1x10^5$ copies/uL, $1x10^4$

copies/uL, and $1x10^3$ copies/uL. The cDNA was amplified on an MxPro3005P real-time thermocycler (Agilent Technologies, Santa Clara, CA). The thermocycler was programmed for an initial incubation of 15 minutes at 95 °C and then 42 cycles of 15 seconds at 94 °C, 30 seconds at 55 °C, and 30 seconds at 72 °C. A dissociation curve was generated at the end of this process. A standard curve, RNA-free control, and no reverse-transcriptase control was included on each plate. Based on the known copy numbers of spike-in control for the standard curve, relative concentrations of each sample were calculated by the MxPro software.

Statistical Analysis

The association between the PNA and LCA markers and piRNA expression was analyzed using a linear model in SAS with PROC REG with the following model:

$$y = b_{Lec}Lec + b_{Ctrl}Ctrl + b_{Age}Age + e$$

In this equation, y represents the piRNA concentration, x_1 represents PNA or LCA, x_2 represents the spike-in control concentrations, x_3 represents age of the bull, e represents residual effects, and b_1. b_2, and b_3 represent parameter estimates for each independent variable. The relative concentrations of piRNA were dependent variables, LCA, PNA, the spike-in control concentrations and age of bull were independent variables. The LNA and PCA values were evaluated independently.

Results

Average PNA and LCA values were 42.7 (±19.9%) and 57.7 (±21.5), respectively (Table 2.1), with the PNA and LCA values ranging between 16 - 93.5% and 18.3 – 92%, respectively. Eight PNA and nine LCA samples were marked as potentially poor samples by the laboratory due to incomplete staining or excessive debris. These samples were excluded from consideration during the selection of samples for sequencing. However, all lectin results were included for real-time RT-PCR results.

Five of the samples sequenced had between 5.5 million reads and 6.5 million reads. However, sample 5123 (high PNA/LCA group) had only 2.8 million reads. A negligible number of reads were removed based on the quality standards for all libraries (less than 1%). For the initial comparison of sequence data to miRBase 21 and known piRNAs in the NCBI database, 617 known sRNAs matched at least one sequencing read. After differential expression analysis between high PNA/LCA and low PNA/LCA groups, 58 sRNAs were differentially expressed (FDR $P<0.05$; Table 2.2). Of these significant sRNAs, nine were miRNAs and 49 were piRNAs. All 49 of these piRNAs were previously identified in human male germlines but their functions remain unknown (Girard *et al*., 2006). Of these piRNAs, seven piRNAs were down-regulated in the high PNA/LCA group while the other 42 piRNAs were up-regulated. All nine of the miRNAs were up-regulated in the high PNA/LCA groups. Overall, there was a trend of up-regulated sRNAs in the bulls with low PNA/LCA scores.

These nine miRNAs were analyzed for putative mRNA targets using Exiqon. The number of putative targets for each miRNA ranged from less than 100 to nearly 1,000

(Table 2.5). Of note, miR-181-a-2 was predicted to bind to all four of the "*deleted in azoospermia*" (Daz) mRNAs. Azoospermia is the lack of sperm cells in a semen sample. The miR-3184 sequence was predicted to bind to Tex-26 mRNA, a testes-specific protein with unknown function. Most importantly, mir-181-a-1 was previously identified in a human male fertility study in seminal plasma (Wang *et al*., 2011). The other miRNA putative targets were less specific. The miR-17 had been identified in human blood serum and was associated with numerous cancers, including prostate cancer. Other significant miRNAs found in blood serum were miR-21, miR30a, and miR-181b-2. Each had a number of putative targets and were associated with cancers. Another association for miR-21 was heart failure and miR-181b-2 has been characterized in conjunction with inflammatory responses. Found in both lymphocytes and blood serum, mir-30d is also associated with a number of cancers. Conversely, mir-1291 had 971 putative targets and many disease associations, suggesting this is a ubiquitous miRNA.

Seven piRNAs were evaluated for an association with semen fertility markers in 31 of the samples. This experiment was designed to investigate the potential relationship of significantly expressed sRNAs across our sample population. The piRNAs were chosen because of their association with fertility markers in the RNA-seq experiment. These piRNAs were pi-30961, pi-32374, pi-32679, pi-36037, pi-37213, pi-44984, and pi-57498. Three primer sets (pi-32374, pi-44984, and pi-57498) did not amplify the appropriate piRNA and thus were not analyzed further. The PNA fertility marker was significantly associated with pi-30961 and pi-32679 expression ($P<0.05$; Table 2.3). The LCA fertility marker approached statistical significance with pi-32679 expression ($P=0.08$; Table 2.4). The direction of association from the regression model matches the

predicted direction from the sequencing results for all but one of the associations tested. The associations between PNA and pi-36037 were not found to be statistically significant, nor were the associations between LCA and pi-30961 and pi-36037.

Discussion

The units used to measure lectin is a percentage of staining, indicated as 'PNA + %PI' and "LCA + %PI'. Propidium Iodide (PI) is a DNA stain that is used to verify that the fluorescent lectins were bound to sperm cells and not cellular debris. Both PNA and LCA are positively correlated with ubiquitin (r = 0.66; r = 0.65; Kennedy *et al*., 2014). The PNA fertility marker was positively correlated with damaged sperm cell acrosomes (r=0.47; P<0.001; Odhiambo *et al*., 2011). The use of LCA as a fertility marker has been investigated but is not as well-supported as PNA (Sutovsky *et al*., 2001). The lectin PNA binds to the acrosome only when there is damage, while LCA binds to normal spermatozoa at low levels and damaged spermatozoa at high levels. Thus, interpreting PNA results is more straightforward than LCA results (Kennedy *et al*., 2014). During processing of semen samples, the formaldehyde used as a fixative to ship the sperm cells was not ultrapure grade, which led to difficulty during staining. More cellular debris was found than expected and a number of the samples exhibited poorer fluorescence than is considered ideal.

Consistent with a previous study examining human male fertility, miR-181a-1 was found to be up-regulated in the low PNA/LCA group. In the previous study, miR-181a-1 was found to be up-regulated in patients that exhibited asthenozoospermia (lack of sperm motility) and down-regulated in patients exhibiting azoospermia (lack of sperm cells altogether) when compared to the control group (Wang *et al*., 2011). Exiqon software predicted that miR181a-2 targets all four DAZ mRNAs. Deleted in azoospermia (DAZ) genes are germ cell proteins, which are important fertility factors in many animals (Fu *et al*., 2015). The loss of DAZ family proteins can cause an extreme

reduction in the production of sperm (Ferlin *et al*., 2007). Testis-expressed sequence 26 protein (TEX26) is a non-functional protein that was characterized in the transcriptome of human males during spermatogenesis (Zhu *et al*., 2016). The TEX26 mRNA was found to be a target for miR-3184.

All of the significantly expressed piRNAs were previously characterized in the human male germline (Girard *et al*., 2006). However, the goal of that study was to characterize piRNAs present without investigating their functions within germ cell lines. Discovering these piRNAs in bovine seminal fluid reiterates the possibility that some of these piRNAs might play a role in male fertility. Further research into the precursors and targets of these piRNAs might shed more light on how these piRNAs are interacting with the genome to affect male fertility.

Two samples, one from the high PNA/LCA group and one from the low group, showed decreased numbers of reads compared to the other four libraries. Sample 5123 had about half of the number of input reads compared to the other libraries. This difference could be attributed to an error during the library preparation or poor RNA quality. Sample 327 showed consistently lower reads compared to the other libraries in the low PNA/LCA groups for the candidate piRNAs that we used for real-time RT-PCR. However, since the statistical analysis weighted the reads based on size before calculating differential expression, this should not have affected the results.

Sequencing the sRNAs of more bulls could provide more information as to whether these differences in reads were specific to the bulls we sampled or are consistent between many bulls. Due to the cost of sequencing, only six samples were chosen for the

sequencing step of the experiment. A larger sample size would help determine whether the results that were shown here were specific to these chosen bulls or were representative of the population. Sequencing more sRNAs from the seminal plasma samples of bulls would also allow for the verification of the 58 significant differentially expressed sRNAs found in this experiment.

Although the piRNAs that were selected for real-time RT-PCR were chosen based on their relative fold change from RNA-seq, overall the piRNAs exhibited low read counts for each sample. The low expression levels of these piRNAs could have contributed to our ability to demonstrate a significant association between piRNA expression and fertility biomarkers. All of the bulls sampled came from three separate herds. The majority of the bulls were 1 year olds at the time of sampling, and those bulls that were not came from two different herds. Thus, age and herd were confounded variables and so only age was included in the statistical model. However, herd may also have had an effect on the fertility of the bulls tested. Finally, sRNAs may also not have a linear relationship with these bull fertility markers. For example, only low levels of these sRNAs may be associated with lower fertility, while moderate to high levels of sRNAs have no effect. The trend of up-regulated sRNAs in the group of bulls with low PNA scores indicates the possibility that the sRNAs are blocking transcription of mRNAs that have a negative effect on fertility.

Another factor to consider is that most of the bulls tested passed their BSE. Although PNA and LCA values ranged from 16 – 93.5% and 18.3-92%, respectively, 33 of the bulls sampled passed a BSE. The three bull included in the low predicted fertility group all passed the BSE. It is possible that more widely applicable candidate sRNAs

could be found if more bulls that did not pass the BSE were included. In human fertility studies that investigated relationships between miRNAs and male fertility, the infertility phenotype used was more specific than the lectin analyses used in this study (Wu *et al*., 2012; Wu *et al*., 2013; Wang *et al*., 2011). These studies evaluated sperm based on quantity of sperm cells and sperm motility, rather than just acrosome status. Both PNA and LCA are both measures of acrosome integrity. Acrosome integrity is an important factor in the fertilization of an embryo, but it is not the only factor that affects sperm quality.

Conclusions

Based on predicted fertility measures using PNA and LCA, 59 differentially expressed sRNAs were discovered in bovine seminal plasma. Several of these sRNAs had previously been associated with fertility in other species. Specifically, miR-181-a-1 was previously found to have a relationship with both sperm quantity and motility in human males. For the first time, piRNAs in bovine seminal fluid have been sequenced and 49 of these piRNAs were differentially expressed between the high and low PNA/LCA groups. Association of two of these piRNAs with PNA fertility markers was confirmed by real-time RT-PCR. Overall, the study supported the hypothesis that miRNAs and piRNAs had a relationship with fertility in beef bull seminal plasma. Further investigation into the sRNAs that were found to be differentially expressed could lead to the establishment of a biomarker for fertility levels in beef bulls.

Literature Cited

Baggerly, K., L. Deng, J. Morris, and C. Aldaz. 2003. Differential expression in SAGE: accounting for normal between-library variation. Bioinformatics 19(12): 1477-1483.

Carmell, M.C., A. Girard, H.J.G. van de Kant, D. Bourc'his, T.H. Bestor, D.G. de Rooij, and G.J. Hannon. 2007. MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. Devel Cell 12(4): 503-514.

Ferlin, A., B. Arredi, E. Speltra, C. Cazzadore, R. Selice, A. Garolla, A. Lenzi, and C. Foresta. 2007. Molecular and clinical characterization of Y chromosome microdeletions in infertile men: a 10-year experience in Italy. J Clin Endocrinol Metab 92(3): 762-770.

Fu, X.F., S.F. Cheng, L.Q. Wang, S. Yin, M. De Felici, and W. Shen. 2015. DAZ family proteins, key players for germ cell development. Int J Biol Sci 11(10): 1226-1235.

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

Houwing, S., L.M. Kamminga, E. Berezikov, D. Cronembold, A. Girard, H. van den Elst, D.V. Filippov, H. Blaser, E. Raz, C.B. Moens, R.H. Plasterk, G.J. Hannon, B.W. Draper, and R.F. Ketting. 2007. A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. Cell 129: 69-82.

Hu, L., C. Wu, C. Guo, H. Li, and C. Xiong. 2014. Identification of microRNAs predominately derived from testis and epididymis in human seminal plasma. Clin Biochem 47(10-11): 967-972.

Kennedy, C.E., K.B. Krieger, M. Sutovsky, W. Xu, P. Vargovic, B.A. Didion, M.R. Ellersieck, M.E. Hennessy, J. Verstegen, R. Oko, and P. Sutovsky. 2014. Protein expression pattern of PAWP in bull spermatozoa is associated with sperm quality and fertility following artificial insemination. Mol Reprod Dev 81: 436-449.

Odhiambo, J.F., M. Sutovsky, J.M. DeJarnettte, C. Marshall, and P. Sutvosky. 2011. Adaptation of ubiquitin-PNA based sperm quality assay for semen evaluation by a conventional flow cytometer and a dedicated platform for flow cytometric semen analysis. Theriogenology 76(6): 1168-1176.

Ramsey, R., D. Doye, C. Ward, J. McGrann, L. Falconer, and S. Bevers. 2005. Factors affecting beef cow-herd costs, production, and profits. J Ag Appl Econ 37(1): 91-99.

Sutovksy, P., M. Aarabi, A. Miranda-Viquete, and R. Oko. 2015. Negative biomarker-based male fertility evaluation: sperm phenotypes associated with molecular-level anomalies. Asian J Androl 17: 554-560.

Sutovsky, P., R. Moreno, J. Ramalho-Santos, T. Dominko, W.E. Thompson, and G. Schatten. 2001. A putative, ubiquitin-dependent mechanism for the recognition and elimination of defective spermatozoa in the mammalian epididymis. J Cell Sci 114(9): 1665-1675.

Wang, C., C. Yang, X. Chen, B. Yao, C. Yang, C. Zhu, L. Li, J. Wang, X. Li, Y. Shao, Y. Liu, J. Ji, J. Zhang, K. Zen, C.Y. Zhang, and C. Zhang. 2011. Altered profile of seminal plasma microRNAs in the molecular diagnosis of male infertility. Clin Chem 57(12): 1722-1731.

Wu, W., Y. Qin, J. Dong, J. Dai, C. Lu, X. Guo, Y. Zhao, W. Zhang, B. Hang, J. Sha, H. Shen, Y. Xia, Z. Hu, and X. Wang. 2013. Genome-wide microRNA expression profiling in idiopathic non-obstructive azoospermia: significant up-regulation of miR-141, miR-429, and miR-7-1-3p. Hum Reprod 28(7): 1827-1836.

Wu, W., Z. Hu, Y. Qin, J. Dong, J. Dai, C. Lu, W Zhang, H. Shen, Y. Xia, and X. Wang. 2012. Seminal plasma microRNAs: potential biomarkers for spermatogenesis status. Mol Hum Reprod 18(10): 489-497.

Yadav, R.P. and N. Kotaja. 2014. Small RNAs in spermatogenesis. Mol and Cell Endo 382: 498-508.

Zhu, Z., C. Li, S. Yang, R. Tian, J. Wang, Q. Yuan, H. Dong, Z. He, S. Wang, and Z. Li. 2016. Dynamics of the transcriptome during human spermatogenesis: predicting the potential key genes regulating male gametes. Sci Rep 6: 19069.

Figure 2.1 Bioinformatics Flowchart for Analyzing Sequencing Data

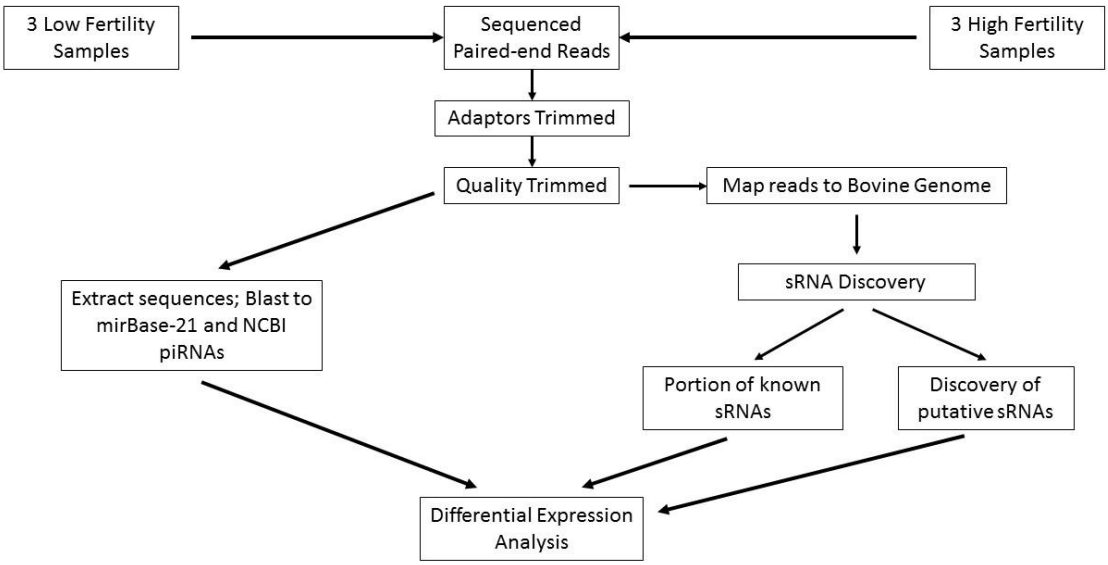


Table 2.1 PNA and LCA Percentages for Sperm Samples

Sample ID	PNA+%PI	LCA+%PI
304	42.5	63.2
310	20.9*	60.9*
314	29.2	33.6
327	21.9	34.1
329	32.8	86.7
333	39.8	52.3
334	40.2	45.2
339	24.9*	84*
344	47.6	67.8
353	40.2	48.4
355	16	18.3
357	23	32
359	27*	77*
366	54.3*	62.7*
367	36.4*	86.9*
378	35	40.2
382	41.1	54.2
389	29	19.6
392	40*	72.5*
3110	25.4	35.3
3101	35.8	40.4
502	67.9	64
742	93.5	87.6
832	70.2	67.1
866	84.1	92
2852	63.2	71.9
2111	65.6	58.9
5123	79.2	83.5
Lot 4	16.6*	28.4*
Lot 5	67.2	89
953	33.2	42.4
345	54.7*	62.8*
355b	18.1	29.9
1633	36.9	42.3*
3108	40.9	84.4
Averages	42.7	57.7

‘*’ indicates potentially bad samples

Table 2.2 Differentially Expressed Small RNAs between High and Low PNA/LCA Samples Sequencing Results

			High PNA Read Numbers				Low PNA Read Numbers			
ID	Fold Δ	FDR-adj p-value	742	866	5123	Avg	357	355b	327	Avg
piR-31355	4.37	0.0	1658	1324	13	998.33	10153	2782	164	4366.33
piR-44984	5.49	0.0	141	174	0	105.00	1235	469	25	576.33
piR-33423	1.04	0.0	70	139	11	73.33	37	33	158	76.00
piR-35468	4.81	0.0	14	73	1	29.33	258	3	162	141.00
piR-36038	4.16	0.0	1561	1168	10	913.00	8948	2315	141	3801.33
piR-36039	4.37	0.0	1653	1324	11	996.00	10132	2768	167	4355.67
piR-36040	4.37	0.0	1665	1324	10	999.67	10156	2769	169	4364.67
piR-36041	4.37	0.0	1659	1325	12	998.67	10154	2782	165	4367.00
piR-36339	1.03	0.0	186	357	18	187.00	93	96	391	193.33
piR-57516	1.23	0.0	5	15	2	7.33	14	8	5	9.00
piR-43773	-1.05	0.0	50	102	5	52.33	39	18	92	49.67
piR-36245	-3.04	0.0	52	19	2	24.33	15	4	5	8.00
piR-35407	-2.95	0.0	1027	1408	56	830.33	605	174	65	281.33
piR-56450	-1.24	2.32E-74	175	323	7	168.33	263	61	84	136.00
piR-33160	-2.82	1.10E-45	14	33	1	16.00	10	1	6	5.67
piR-45884	1.70	7.14E-17	6	16	1	7.67	28	6	5	13.00
piR-31924	6.50	4.44E-12	4	10	0	4.67	43	46	2	30.33
piR-32679	3.34	8.88E-11	13	19	0	10.67	53	52	2	35.67
piR-35467	5.01	1.42E-10	11	74	0	28.33	258	3	165	142.00
piR-35466	4.82	5.17E-10	12	76	0	29.33	258	2	164	141.33
piR-31925	5.65	6.23E-10	4	13	0	5.67	43	50	3	32.00
piR-35469	4.55	4.52E-09	13	80	0	31.00	255	3	165	141.00
piR-49145	2.72	9.27E-08	85	237	1	107.67	406	175	299	293.33
mir-181a-2	2.16	1.88E-07	31	42	0	24.33	72	43	43	52.67

mir-181a-1	2.18	2.08E-07	30	41	0	23.67	72	42	41	51.67
piR-33387	3.47	4.57E-07	9	6	0	5.00	31	19	2	17.33
piR-30961	32.00	1.04E-06	1	0	0	0.33	20	12	0	10.67
piR-36256	1.60	2.34E-05	21	26	0	15.67	18	6	51	25.00
piR-35229	2.00	8.49E-05	12	14	0	8.67	23	29	0	17.33
piR-37213	23.00	9.88E-05	0	1	0	0.33	13	4	6	7.67
piR-61645	5.67	2.33E-04	1	5	0	2.00	16	17	1	11.33
piR-35952	21.00	2.63E-04	1	0	0	0.33	14	7	0	7.00
piR-50725	N/A	3.91E-04	0	0	0	0.00	12	5	1	6.00
piR-35176	N/A	3.91E-04	0	0	0	0.00	12	6	0	6.00
piR-61647	2.91	4.97E-04	4	7	0	3.67	17	14	1	10.67
piR-30229	1.02	8.85E-04	59	56	0	38.33	58	31	28	39.00
piR-57498	6.00	8.85E-04	1	3	0	1.33	11	12	1	8.00
piR-61646	4.13	9.11E-04	2	6	0	2.67	18	15	0	11.00
piR-61648	2.38	1.92E-03	5	8	0	4.33	16	15	0	10.33
piR-35175	N/A	2.69E-03	0	0	0	0.00	11	3	0	4.67
mir-3184	1.27	5.55E-03	56	49	1	35.33	95	26	14	45.00
mir-21	2.55	5.79E-03	31	229	0	86.67	347	69	247	221.00
piR-61644	3.50	7.11E-03	2	6	0	2.67	13	15	0	9.33
piR-36243	1.25	7.12E-03	20	24	0	14.67	16	7	32	18.33
mir-1291	N/A	7.12E-03	0	0	0	0.00	3	8	1	4.00
piR-32374	14.00	7.20E-03	0	1	0	0.33	11	3	0	4.67
piR-36242	1.28	8.74E-03	19	21	0	13.33	16	5	30	17.00
piR-36706	N/A	1.11E-02	0	0	0	0.00	7	4	0	3.67
mir-17	N/A/	1.11E-02	0	0	0	0.00	8	0	3	3.67
mir-30a	1.53	1.77E-02	135	302	2	146.33	391	84	196	223.67
piR-36255	1.19	1.90E-02	20	22	0	14.00	13	5	32	16.67
piR-41435	1.92	1.97E-02	7	6	0	4.33	9	2	14	8.33

piR-33487	2.00	3.36E-02	8	7	1	5.33	21	8	3	10.67
mir-30d	1.76	3.72E-02	207	575	0	260.67	921	140	312	457.67
mir-181b-2	1.77	4.38E-02	5	8	0	4.33	16	3	4	7.67
piR-41209	-2.55	4.54E-02	19	30	2	17.00	14	3	3	6.67
piR-36037	40.00	4.79E-02	0	2	0	0.67	68	10	2	26.67
piR-36249	-3.12	4.99E-02	292	215	1	169.33	124	13	26	54.33

Table 2.3 Linear Model of RT-PCR Expression, Spike-in Control, Age of bull, and PNA

piRNA	Variables	Effect Size	SE	*P* value	R^2
pi-30961	**PNA**	0.98	0.37	**0.014**	0.272
	Spike-in Ctl	-6.28	11.75	0.598	
	Age of Bull	-14.00	10.30	0.186	
pi-32679	**PNA**	0.05	0.02	**0.009**	0.319
	Spike-in Ctl	-0.63	0.55	0.269	
	Age of Bull	-0.34	0.49	0.489	
pi-36037	**PNA**	0.02	0.01	0.171	0.084
	Spike-in Ctl	-0.17	0.35	0.630	
	Age of Bull	-0.40	0.31	0.210	

Table 2.4 Linear Model of RT-PCR Expression, Spike-in Control, Age of bull, and LCA

piRNA	Variables	Effect Size	SE	*P* value	R^2
pi-30961	**LCA**	8.50	5.88	0.160	0.146
	Spike-in Ctl	10.40	10.08	0.312	
	Age of Bull	-5.68	10.77	0.602	
pi-32679	**LCA**	0.50	0.28	**0.082**	0.208
	Spike-in Ctl	0.20	0.47	0.676	
	Age of Bull	-2.38E-17	0.51	1.000	
pi-36037	**LCA**	0.09	0.17	0.596	0.025
	Spike-in Ctl	0.11	0.29	0.704	
	Age of Bull	-0.21	0.31	0.506	

Table 2.5 Significant miRNA Targets from Exiqon

miRNA	Targets	Tissues IDed in	Applicable associations
miR-17	356	Blood serum	Associated with prostate cancer
miR-21	230	Blood serum	Associated with heart failure
miR-30a	79	Blood serum	Associated with ovarian cancer
miR-30d	82	Lymphocytes; Blood serum	Associated with various cancers
miR-181a-1	171	Blood serum; Seminal Plasma	Associated with human male infertility
miR-181a-2	370	Plasma	Associated with DAZ1-4 proteins
miR-181b-2	349	Blood serum	Associated with cancers and inflammation
miR-1291	971	Ubiquitous	Many associations
miR-3184	156	Blood serum	Associated with Tex26

Appendix A: QIAGEN MiRNA Extraction Protocol

Seminal plasma samples were thawed. 750uL of QIAzol Lysis reagent was added to 150uL of seminal plasma. This incubated at room temperature for 5 minutes. 3.5uL of the miRNeasy Serum/Plasma Spike-In Control was added at the concentration of 1.6 x 10^8 copies/uL. 150uL of chloroform was added to each sample, vortexed for 15 seconds, and then incubated for 3 minutes at room temperature. Samples were centrifuged for 15 minutes at 12,000xg at 4°C. 500uL of the upper aqueous phase was transferred to a new collection tube. 750uL of 100% ethanol was added to each sample, mixed, and then 600uL of the solution was pipetted onto the provided MinElute spin column with collection vial. Tubes were centrifuged at 12,000xg for 15 seconds at room temperature. The flow through was discarded and this step was repeated until all of the ethanol-sample mixture was filtered. 700uL of RWT Buffer was added to the spin columns, which were then centrifuged for 15 seconds at 12,000xg. This flow-through was also discarded. 500uL of RPE Buffer was added to the spin column and then centrifuged for 15 seconds at room temperature. Finally, 500uL of 80% ethanol was added to the spin column. The tubes were centrifuged for 2 minutes at 12,000xg and the collection tube was discarded. Columns were spun dry for 5 minutes at 12,000xg. A new 1.5mL collection tube was attached to the spin column and 14uL of RNAse-free water was added. Column was spun for 1 minute at 12,000xg. RNA was stored at -80°C.

Appendix B: QIAGEN RT II miRNA cDNA synthesis protocol

cDNA reactions were designed to use 50ng of RNA. For each reaction, 2ul of 5X HiFlex Buffer was used. This specific buffer was chosen based on the variety of small RNAs involved in the downstream applications. 1uL of 10X nucleics mix, 1uL of reverse transcriptase mix, 50ng of cDNA, and variable amounts of water to bring the final volume of the reaction up to 20uL are also added. Once all the reagents are combined, the tubes are incubated for 60 minutes at 37°C, 5 minutes at 95°C, and then stored directly at -20°C.

This procedure was performed for each individual RNA sample, pooled RNA samples with spike-in control for generating a standard curve, and a no reverse-transcriptase control samples.

Appendix C: QIAGEN SYBR Green PCR Protocol

The cDNA samples generated for the standard curve were diluted to a concentration of $1x10^6$ copies/uL, $1x10^5$ copies/uL, $1x10^4$ copies/uL, and $1x10^3$ copies/uL. The reaction mix for the standard curve samples involves 2uL of cDNA/spike-in control, 12.5uL of SYBR green master mix, 2.5uL of Universal primer, 2.5uL of the specific miScript Primer assay, and 5.5uL of water for a total volume of 25uL. For all other cDNA samples, the reaction mix involves 1uL of cDNA/spike-in control, 12.5uL of SYBR green master mix, 2.5uL of Universal primer, 2.5uL of the specific miScript Primer assay, and 6.5uL of water for a total volume of 25uL. cDNA was pipetted onto 96-well plates, master mix containing the SYBR green, two primers, and water was added. PCR plates were briefly centrifuged and loaded onto the thermocycler. The plate was incubated for 15 minutes at 95°C and then went through 42 cycles of 15 seconds at 94°C, 30 seconds at 55°C, and 30 seconds at 70°C. The samples then underwent a gradual increase in temperature to generate a dissociation curve. Plates were then stored at -20°C.