PATTERNS OF DEGRADATION IN *PRM1* AND *SEMG1* TRANSCRIPTS IN SEMEN DURING 6 MONTHS OF AGING

By

ALLISON J SHERIER

Bachelor of Science in Animal Science

Oklahoma State University

Stillwater, OK

2014

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
July, 2016

PATTERNS OF DEGRADATION IN *PRM1* AND *SEMG1* TRANSCRIPTS IN SEMEN DURING 6 MONTHS OF AGING

Thesis Approved:

Dr. Robert Allen
Thesis Adviser

Dr. Gerwald Koehler

Dr. Ron Thrasher

Name: ALLISON J SHERIER

Date of Degree: July, 2016

Title of Study: PATTERNS OF DEGRADATION IN *PRM1* AND *SEMG1* TRANSCRIPTS IN SEMEN DURING 6 MONTHS OF AGING

Major Field: FORENSIC SCIENCES

Abstract: The purpose of this study was to study the degradation rates of mRNA molecules in semen stains as a possible approach to estimating the age of semen samples found at crime scenes. Specific primers directing amplicons of different lengths from transcripts from the sperm specific PRM1 gene and the seminal fluid specific SEMG1 gene were designed and synthesized. Using two fluid-specific primers allowed for the investigation of the difference in degradation rates between sperm and seminal fluid. No previous study had a large enough sample size to investigate if there is a difference in degradation rate among individuals. In this study, ten young adult males provided semen samples which were dried in 50µl aliquots and stored under controlled conditions. Aliquots were extracted using TRizol every 2 weeks for 6 months to obtain RNA for analysis. The degradation of 18S rRNA and GAPDH transcripts was analyzed as possible stable control transcripts along with PRM1 and SEMG1. 18S started showing increased degradation at 14 weeks while GAPDH seemed to having an increase in degradation over 6 months of aging. Degradation in control genes prevented meaningful statistical analysis from being completed. Looking at raw CT values there was great variation of mRNA concentration between individuals and time-points, but overall PRM1 and SEMG1 showed little to no indication of degradation from 0 to 6 months of aging. It is possible PRM1 and SEMG1 could still be useful in determining the age of semen samples if the samples were over 6 month of age. Additional research with a longer time course needs to be conducted to determine at what time PRM1 and SEMG1 start degrading quickly enough to be used to determine that age of semen samples.

TABLE OF CONTENTS

Chapter Page

I. INTRODUCTION 1

Hypothesis.....3
Research Purpose.....3
Methodology Overview 4

II. REVIEW OF LITERATURE.....6

Properties of RNA vs DNA 7
Instability of RNA.....7
Degradation of RNA 9
Forensic Relevance of RNA 11
Using RNA to Identify Semen.....12
Summary 13

III. METHODOLOGY 14

Sample Collection.....15
RNA Isolation 15
cDNA Production.....16
Primer Design 17
Control Genes 18
Primer Optimization.....19
Primer Efficiencies.....21
Statistical Analysis.....22
Summary.....23

Chapter Page

IV. RESULTS 24

Selection of Primers 24
C_T Analysis 26
Delta Delta C_T Analysis 34
Control Genes 38
Summary 39

V. CONCLUSION 41

Processing Semen for RNA 43
Initial RNA Concentrations 44
Factors Affecting Degradation Rate 45
Changes in Methodology 45
Applications of Research 47
Future Investigation 47
Summary 48

REFERENCES 50

APPENDICES 53

APPENDIX A Approval by the Oklahoma State University Institutional Review Board 54

APPENDIX B Nanodrop Values of RNA (ng/µl) for Each Donor at Each Time-Point 56

APPENDIX C Tables showing the variation in ΔC_T in *PRM1* (96 bp and 200 bp) and *SEMG1* (99 bp and 195 bp) minus the average of *18S* (187 bp, labeled ENDO in the table) 58

LIST OF TABLES

Table Page

1. Selected Nucleotide Sequences for Primers and Probes for *PRM1* and *SEMG1* 18

2. Information about Selected Control Genes.....19

3. Matrix for Optimization of Primers and Probes 20

4. Optimized Primer and Probe Concentrations for Real-Time Polymerase Chain Reaction.....20

5. Efficiencies of Control and Designed Targets in Singleplex.....21

LIST OF FIGURES

Figure Page

1. Efficiency of Primers and Probes Graphed on a Log Scale.....26

2. C_T Values for Time-point 0 in *PRM1* (96 bp & 200 bp) and *SEMG1* (195 bp) Based on GAPDH Being Normalized for Each Donor.....28

3. A-J. Line graphs of *PRM1*, *SEMG1*, *18S*, and *GAPDH* C_T Values for All Time-points for Each Donor.....28

4. Individual Donor's (1-5) 18S (187 bp) C_T Values Over 24 Weeks 35

5. Individual Donor's (6-10) 18S (187 bp) C_T Values Over 24 Weeks 35

6. A-C. Changes in Abundance of *PRM1* and *SEMG1* mRNA Transcripts During 24 Weeks of Storage at Room Temperature.....37

7. Average Raw C_T Values for *18S* (61 bp), *18S* (187 bp), and *GAPDH*.....39

CHAPTER I

INTRODUCTION

Today, the public believes they know everything about crime scene investigation. This belief is based, in part, on television series such as CSI, NCIS, Bones, and other popular crime scene investigation type shows. What the general public does not understand is many crime scenes, especially murders, are a mess. There is potential deoxyribonucleic acid (DNA) everywhere. A crime scene investigations unit could collect hundreds of samples for analysis from one crime scene alone.

Since 1901, identification of biological fluids has played a significant role in forensic science. Over the past 115 years the method for identifying unknown biological samples has greatly improved and current research aims to analyze the decomposition of messenger ribonucleic acid (mRNA) to estimate the age of biological samples present at crime scenes. This new research could distinguish between biological samples left at a crime scene when the crime was committed versus those that may have been present previously. Such information would change the way we process DNA evidence in criminal cases.

Forensic laboratories currently use presumptive and confirmatory tests to identify unknown biological evidence, often reducing the number of samples that must be processed from a crime scene. Techniques used to process semen in particular include illumination of dried stains

with alternate light sources, microscopy, and antibodies-based techniques.[1]

Semen identification has presented problems to the forensic field because of the unique nature of seminal fluid and sperm cells (spermatozoa). While prostatic fluid, approximately 30% of ejaculate, has high concentrations of acid phosphatase and prostate-specific antigens that can be used for semen identification, these chemicals are also found in other materials.[1] Acid phosphatase can be found in bone, blood, and several major organs. Additionally, prostate-specific antigens can be found in urine, fecal material, sweat, and milk.[1] Acid phosphatase and prostate-specific antigen tests can lead to false positives during presumptive testing for semen.[1]

Confirmatory testing for semen is more accurate, but only the laborious searching for spermatozoa microscopically is 100% accurate for the confirmation of a semen stain. In this method, spermatozoa in dried semen stains are liberated from the matrix through rehydrating the stain and then subjecting it to vigorous mixing with a vortex mixer.[1] The extract is then applied to a microscope slide which may be subsequently stained in a variety of ways before microscopic examination is performed in an attempt to confirm the presence of spermatozoa. Multiple staining methods are available to help visualize spermatozoa more easily. [1] This confirmatory method assumes that all semen stains will have spermatozoa present which is not always the case. Some men have very low sperm count or have had a vasectomy and no spermatozoa are present in their ejaculate. Studying mRNA in semen could develop a new more accurate confirmatory test that may also provide the age of the samples being tested.

If RNA analysis could provide useful information to forensic scientists and investigators, RNA analysis would become more commonplace. Species- and tissue-specific RNA markers will allow for another confirmatory test for human semen; and, with the selection of fluid-specific markers, could also yield information about whether a suspect has been vasectomized.[2] Currently

however, RNA analysis is not used in forensic laboratories because of the past view of the instability of RNA and because of time constraints for analyzing evidence.[3]

Hypothesis

The underlying hypothesis for this study is that the mRNA transcripts for two genes expressed in spermatozoa and seminal fluid encoding protamine 1 (*PRM1*) and semenogelin 1 (*SEMG1*) respectively, will allow for the confirmatory identification of an unknown body fluid stain and the transcripts for these genes will degrade in a predictable manner as stains age. The rate of disappearance of these transcripts will ultimately allow for estimates of the age of the semen stains recovered from a crime scene. [4-6] Ultimately, analysis of these two human semen-specific multiplexed markers may allow for the identification and aging of a semen sample in one process.

Research Purpose

The purpose of this study was to examine the degradation of messenger RNAs (mRNA) produced by two tissue-specific genes that are only found in human semen. Messenger RNA degradation begins when biological fluids leave the human body. Although mRNA degradation does occur, Anderson et al have shown that *18S* ribosomal RNA (*rRNA*) is relatively stable in aged bloodstains.[4,5] Other mRNAs encoding housekeeping genes like glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) or beta-actin (*ACTB*) have also been used as transcripts whose abundances are claimed to remain stable.[7-9] Thus, abundance levels of mRNA encoding *PRM1* and *SEMG1* can be normalized against the rather stable abundance of *18S rRNA* or another control gene transcript to produce estimates of degradation rates for these tissue specific transcripts. The rationale for this study is that the degradation rates for the *PRM1* and *SEMG1* transcripts may be informative as to the age of the semen stain.

No systematic study of a large number of semen samples, stored under controlled conditions, has been done to characterize the age dependent changes in semen transcript abundance. Nakanishi et al looked at semen samples ranging from 33 – 56 years old that were stored at room temperature before extraction, but this was a limited number of samples, and each sample was from a different individual.[7] Additionally, several research projects have looked at a small number of samples over a short period, but these studies do not allow a conclusion to be drawn about variation of mRNA expression between donors because of the small sample size.[8,10,11] Moreover, the variability of transcript abundance in semen samples collected from a number of donors has not been studied and little to nothing is known about possible variation in degradation rates of mRNA molecules from one individual to the next. This project examined transcript abundance for 2 fluid-specific markers: *SEMG1*, a gene expressed in seminal fluid, and *PRM1*, a gene expressed in spermatozoa. Fresh semen samples were collected from 10 volunteer donors and spotted onto sample collection cards. At bi-weekly intervals over a period of 6 months, RNA was extracted from semen stains and converted to complimentary DNA (cDNA) using reverse transcriptase so that *SEMG1* and *PRM1* transcripts could be quantified using real-time PCR (qPCR).

Methodology Overview

Human semen-specific markers were selected and multiplexed based upon prior research performed in this laboratory.[2] Protamine 1 (*PRM1*), specific to spermatozoa, and semenogelin 1 (*SEMG1*), specific to seminal fluid were selected as tissue specific markers.[7-9,12,13] The molecular strategy devised to assess degradation of the *PRM1* and *SEMG1* transcripts involved designing polymerase chain reaction (PCR) primers that would direct the amplification of a short amplicon (i.e. ~100 basepairs) and a separate set of primers that would direct amplification of a longer amplicon (i.e. ~150 basepairs) from a separate region of the transcripts. Sequences within both

the short and long amplicon were used to design Taqman probes for amplicon detection so that the abundance of each amplicon produced from each gene marker could be quantified using real-time PCR (qPCR). The abundance of transcripts could therefore be estimated from two qPCR targets of differing length. This approach has been suggested in work performed by others. [8,9,12]

Ejaculates were self-collected by 10 young adult males with approval from Oklahoma State University Center for Health Science's Internal Review Board. Samples were collected and stored for a maximum of 24 hours at 4°C before 50µl aliquots were spotted on standard blood collection cards (705 Classic Collection Cards, Fitzco Inc., Spring Park, MN) in the laboratory. RNA was extracted from semen stains using standard extraction methods from each donor every 2 weeks for a total of 6 months of aging.

Complimentary DNA was produced from the total RNA extract using a protocol supplied with the High Capacity cDNA Reverse Transcription Kit (Life Technologies, Carlsbad, CA) producing about 60µl of cDNA. Real time quantitative PCR reactions were then used to quantify the relative abundance of *PRM1* and *SEMG1* transcripts normalized to transcript abundance for *18S rRNA* (61 bp and 187 bp) and the enzyme glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*). Data was analyzed for delta C_T values (comparing abundance of *PRM1* or *SEMG1* transcript to the abundance of *18S* (187 bp) *rRNA* and *GAPDH*), or delta-delta C_T comparing each age time point against the time zero value. [8,9,12]

CHAPTER II

REVIEW OF THE LITERATURE

Messenger ribonucleic acid (mRNA) can estimate the age of biological samples.[4,14-17] Investigating the transcriptome through the analysis of RNA provides researchers with fluid- and tissue-specific information about forensically relevant biological samples.[18] While deoxyribonucleic acid (DNA) analysis is common in forensic laboratories, RNA analysis is not typical. However, RNA analysis in forensic laboratories has been successful in the identification of unknown biological samples, such as blood, semen, saliva, menstrual blood, and vaginal secretions based upon tissue specific transcripts present in these body fluids.[4,5,17] Thus, research focusing on the types and abundance of RNA sequences in biological evidence will undoubtedly provide new tools that can help investigators solve crime.

With the availability of RNA analysis methodologies, forensic scientists can determine the identity of an unknown stain, and potentially at what time the stain was left at the crime scene. Collection of semen is common in rape cases. Semen recovered from the victim's body, or from clothing, bedding, the floor, etc. is common in sexual assaults. In cases where sexual activity is occurring normally, such as a bedroom, it can be difficult to determine which semen samples belong to a consensual act or from the assault. Knowing when a stain was deposited at a crime scene is important information for the investigation, as the question often arises in court about when a sample was deposited at a crime scene. Being able to answer such questions would greatly assist in the investigation of the crime and probably strengthen the case against a guilty defendant.

Properties of RNA vs DNA

DNA and RNA have some similarities and many differences that allow them to perform specific jobs that a cell needs to stay alive. Both DNA and RNA are macromolecules made up of nucleotides.[19] DNA contains four nucleotides: deoxyadenine, deoxythymine, deoxycytosine, and deoxyguanosine.[19] In RNA, the nucleotides contain an oxygen atom at the position within the ribose sugar that is deoxygenated in DNA and thymine is replaced by uracil.[19] When phosphate bonds link three nucleotides together, a codon forms, which specifies the location of a particular amino acid in the polypeptide sequence of the gene product.[19] Amino acids are the building blocks of proteins that make up the body and allow the body to function.[19] DNA is a double-stranded helix while RNA is usually single-stranded.[19] DNA's helical structure is more stable than RNA's single-stranded structure, due, in part to the double stranded character of DNA and also because the oxygenated position within the ribose sugar in ribonucleotides is especially vulnerable to chemical attack.[19]

RNA has many forms, the most common being transfer RNA (tRNA), ribosomal RNA (*rRNA*), and messenger RNA (mRNA).[19] These three RNAs are involved in protein synthesis. Transfer RNA participates in protein synthesis as the transport vehicle bringing amino acids into the protein synthetic complex.[19] Ribosomal RNA provides the structural scaffold for ribosomes that engage with mRNA during translation to produce a protein.[19] The balance between the transcription of mRNA and its degradation helps to regulate gene expression in cells.[19]

Instability of RNA

As discussed earlier, RNA is less stable than DNA due to their differences in macromolecular structure and chemical composition.[19] While RNA can hybridize with itself in some regions to become double-stranded, RNA is generally single-stranded.[20] This single-

stranded nature and the chemical structure of RNA make it vulnerable to chemical attack in the cell, resulting in the hydrolysis of the RNA chain accompanying the decomposition process.[20] As stated above, the ribose sugar of RNA contains a hydroxyl group that is replaced by hydrogen in DNA.[20] This hydroxyl group participates in chemical reactions, that ultimately cleave the phosphodiester bonds that link the ribonucleotides in the RNA chain.[19] DNA has historically been utilized more than RNA by forensic scientists, for different reasons. DNA is the repository of genetic markers, a handful of which are targets of DNA typing activities that aim to identify the source of biological evidence associated with crime (the question of who). In contrast, RNA studies focus more on questions of what and when as they relate to evidence. Because of its inherent stability, DNA is of little value in attempting to determine the age of evidentiary stains. Likewise, DNA harbors all of the genes of the individual, making studies on tissue specific gene expression impossible.

Multiple factors can affect the rate of RNA degradation. For example, the steady state of abundance of RNA in the cell will contribute to the half-life of a transcript.[18] In addition, the ability of an RNA molecule to fold onto itself and make secondary and tertiary structures also helps reduce the rate of degradation by hydrolysis.[21] Forensic scientists researching the differences between *rRNA* and mRNA folding discovered how RNA folding changes degradation rates.[21] Understanding the degradation of *rRNA* and mRNA has allowed scientists to suggest that RNA degradation can provide clues to the post-mortem interval of deceased individuals.[21,22]

A complication associated with using RNA degradation to estimate time is the environment in which a stain is deposited.[21,23] Exposure of a body or body fluid stain to rain, high humidity, extremes temperature, and exposure to sunlight or other forms of radiation will increase the rate of degradation of RNA. Thus, a thorough understanding of RNA degradation and the factors that affect it is critical to be able to use RNA degradation as a measure of time.[21,23]

Degradation of RNA

In living cells, mRNA quantities are controlled to manage the process of gene expression.[24] Messenger RNA is subject to three types of chemical modification: RNA turnover, RNA processing, and quality control.[20] Turnover of mRNA in a viable cell is essentially controlled degradation by enzymes and other chemical processes, and the half-life of any given mRNA in the cell is determined by the cell's need.[24] The half-life of mRNA can vary from minutes to days depending upon the needs of the cell.[24] The half-life of mRNA will never be longer than the time required for a cell to double because additional mRNA would hinder the regulation of genes. [24,25] During the maturation of mRNA in the cell, processing takes place, thereby removing intronic sequences present in the initial transcript in preparation for translation.[20] Quality control results in removal of incorrectly produced mRNA or non-coding RNAs that cannot be used in translation.[20]

If a biological sample is dried, normal regulation of RNA abundance is disrupted and RNA degradation becomes determined by physical and chemical factors.[16,26] Biological samples collected as evidence at a crime scene are generally dried for transportation and storage. In contrast, it has been proposed that RNases may be responsible for RNA degradation in tissues in deceased individuals.[21] Numerous studies show that mRNA can be extracted from forensically relevant samples that have aged several years.[4,5,11,16,26]

The most common method for detection of degradation is quantitative real-time polymerase chain reaction (qPCR).[4,5,15,27] In this procedure, complementary DNA (cDNA) is synthesized from mRNA extracted from the sample of interest. Then cDNA is used as a template for qPCR assays, which use primers to bind to the target of interest in the cDNA preparation. A fluorescent oligonucleotide probe, hybridized to the target sequence, provides real time fluorescent quantitation of the accumulating amplified segments of cDNA produced during

qPCR. This methodology, known as the Taqman assay, is widely used to quantify mRNA transcripts in biological samples. Accumulating fluorescence in a qPCR reaction is proportional to the relative abundance of mRNA that was present in the sample initially. To control for possible technical and biological variability from one qPCR reaction to the next, a constitutively expressed housekeeping gene or a ribosomal RNA species are co-amplified in qPCR reactions with the mRNA target of interest resulting in fluorescence from an internal RNA control in each reaction whose abundance level is expected to remain constant.[16] Housekeeping genes, like beta actin and glyceraldehye 3-phosphate dehydrogenase, have been reported to exhibit abundance levels that are more stable than mRNA transcripts whose expression can be induced within the cell.[15,27] The same rationale employed in analyzing changes in gene expression in the living cell can be applied to the study of changes in mRNA abundance in non-living material due to RNA degradation. Some studies have shown that total RNA abundance can decrease due to degradation at different rates depending on the tissue type.[15,18,27] Researching RNA transcripts provides insight into protein and gene regulation, as levels of mRNA are a reflection of gene expression in a particular cell.[21] Additionally, different tissues contain different populations of mRNA transcripts based on cell function.[21] The difference in mRNA between tissue types has led to research of tissue-specific mRNA markers for forensically relevant biological samples.[4,6,21,26,28]

Several studies have shown that mRNA degrades in a linear fashion and that some mRNAs degrade slower than others.[4,15,25] Several studies indicate that samples collected from deceased individuals remain stable from a few hours after death up to several days.[14,15,27] While dried stains perhaps degrade differently than post-mortem samples, research on deposited samples nonetheless provides important information about how environmental conditions may affect mRNA degradation.

Forensic Relevance of RNA

For RNA analysis to be worthwhile in forensic laboratories, DNA and RNA must be extracted from the same sample. Often, only one type of biological sample can be collected from a crime scene. Precedence is given to DNA, which provides the profile that can link a suspect to the crime scene. Extraction of RNA from the same sample has the potential to provide additional information to the investigators about the nature of the sample and the time of the crime. For extraction of DNA, many crime labs use Promega DNA IQ™ which also allows for RNA to be collected after DNA is extracted.[29] There are several other methods for co-isolation of DNA and RNA, but thought must be given to what method will be easily incorporated into current protocols used by forensic laboratories.

Messenger RNA can now be used to identify forensic samples such as blood, saliva, semen, sweat, vaginal fluids, and menstrual blood. [4,6,9,10,12,13,21,26,28,30] Studies have also investigated the statistical effects of sex, age, ethnicity, lifestyle, eating habits, and diseases on mRNA levels.[4,5,22,27] While no statistically significant effect was found for sex, age, and ethnicity,[4,5,22,27] lifestyle, diets, and disease appear to affect on mRNA levels.[22,27]

The key to analysis of forensically relevant samples is to target species- and tissue-specific targets.[4] The use of defined PCR oligonucleotide primers confers specificity to a molecular analysis that will allow for the identification of a biological sample, which will in turn allow for the study of mRNA degradation. If investigators had the ability to determine the specific time a biological sample was left at a crime scene, they could determine the times the crime was committed, whether the sample was deposited during the crime, or if the sample is even relevant to the case.

Using RNA to Identify Semen

When identifying semen through the detection of tissue-specific mRNA, a gene associated with cells in seminal fluid and a second gene specific for spermatozoa are used. Targeting a sperm-cell specific mRNA can provide insight to whether or not an assailant may be vasectomized or have a low sperm count.[2] In vasectomized individuals, there will be no expression of sperm-cell specific genes.[2] The most commonly investigated sperm-cell specific genes are protamine 1 (*PRM1*) and protamine 2 (*PRM2*). [7,9-13,31,32] Numerous genes have also been researched for seminal fluid, and a common gene being analyzed is semenogelin 1 (*SEMG1*).[7,8,10,33] Setzer et al performed a study that exposed semen to multiple environmental conditions and was able to detect *PRM2* up to 365 days and *PRM1* up to 547 days after deposit in select samples.

In 2014, Nakanishi et al completed a study that showed that *18S* ribosomal RNA and *PRM2* mRNA remained detectable (perhaps as mRNA fragments) in semen samples collected from crime scenes ranging from 33 to 56 years old.[7] *SEMG1* was not detected in any of the aged semen samples, and *PRM2* was not detected in two samples that were 41 and 44 years old.[7] Differences in the storage environment could explain why *PRM2* was found in some older samples but not in younger samples. There is speculation that *PRM2* can be detected in older samples because of the rigidity of the sperm cell membrane compared to that of the epithelial cells sloughed from the seminal vesicles, ductwork, and prostate gland.[7] Many studies have investigated tissue-specific mRNA transcripts for use to identify biological samples in a forensic setting.

Summary

Research on RNA has expanded in recent years disproving the original thought that RNA was not stable and could not be extracted from old biological samples. The new information on mRNA markers has developed a new field of study that is useful to forensic scientists. Analysis of fluid-specific markers can be incorporated into already existing forensic laboratory protocols to assist with estimating the time a evidentiary sample was deposited at a crime scene and also to help identify unknown biological fluids. While mRNA analysis will take more time with current technology to perform than DNA analysis, it can provide useful insight to known biological samples that could help determine relevant samples to a particular case.

CHAPTER III

METHODOLOGY

In recent years, the use of molecular methodologies to study RNA in forensic samples has demonstrated the value of such techniques in assisting with the identification of unknown body fluid stains.[4,5,7,9,11,12] RNA degradation patterns can also be informative for estimating the age of biological stains.[4,5,7,11] In this project, semen samples were self-collected by 10 young adult males in support of a study on the degradation of semen/sperm specific mRNA transcripts as a tool to estimate the age of semen stains. Semen, less than 24 hours after collection was spotted in 50µl aliquots and allowed to air dry on standard sample collection cards and total RNA was extracted from each dried stain every 2 weeks for 6 months. A protocol allowing for the simultaneous recovery of deoxyribonucleic acid (DNA), ribonucleic acid (RNA), and protein was used to demonstrate that a DNA profile could also be obtained using this extraction protocol.[4] Using a process that allows for DNA and RNA extraction is important for the incorporation into standardized forensic laboratory procedures.

Procedures utilized in this study included sample collection, spotting on body fluid collection cards, RNA isolation using Trizol reagent (Sigma-Aldrich, St. Louis, MO), complementary DNA (cDNA) production, primer design, primer optimization, RNA quantification, and statistical analysis. Sample were collected and processed in accordance with Oklahoma State University Center for Health Sciences' Institutional Review Board approval (Appendix A).

Sample Collection

Samples were self-collected from 10 male undergraduate and graduate students with at least a general knowledge of molecular biology. All participants were between the ages of 18 and 28. No additional information was collected about the individuals who donated. Each sample had a 10-digit identification code, consisting of a unique identification number from 001 - 010, a letter to indicate semen (E), and the date of collection (mmddyy) in accordance with previous studies on human biological fluids performed in the School of Forensic Sciences.[18] Fifty microliters aliquots of each semen sample were spotted onto labeled body fluid collection cards (705 Classic sample cards, non-FTA treated, Fitzco Corp., Spring Park, MN). Stain cards were stored in the dark in the laboratory at room temperature and RNA was extracted from each 50 µl stain on weeks 0, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, and 24.

RNA Isolation

RNA isolation was performed by a combination of TRI Reagent® (Sigma-Aldrich, St. Louis, MO), chloroform, and the RNA Clean and Concentrator™ (Zymo Research, Irvine, CA) following the manufacturer's protocol. TRI Reagent® allows for co-isolation of DNA, RNA, and protein. The semen stains (approximately 1 cm^2) were cut from the stain cards at the appropriate assay time-point and put into a 1.5 mL conical tube with 1 mL of TRI Reagent®. Samples were vortexed at low speed for 30 minutes at room temperature. Chloroform (200 µl) was added to each tube which was then shaken vigorously for 15 seconds. After the chloroform and TRI Reagent® were thoroughly mixed, tubes were left standing in a tube rack at room temperature for 3 minutes. Samples were then centrifuged at 12,000g for 15 minutes at 4°C. Centrifuging separated the mixture into three phases, thereby allowing for the collection of RNA by the pipetting of the upper aqueous layer into a new collection tube.

The RNA Clean & Concentrator™ Kit was used to remove any debris or DNA from the collected RNA according to the manufacturer's instructions. The RNA was eluted from the RNA Clean & ConcentratorTM Kit binding tube with the addition 20µl of RNase-free water followed by centrifugation for 30 seconds. This elution step was repeated with an additional 15µl aliquot of RNase-free water. By eluting bound RNA in 35µl of RNase-free water, sufficient volume of RNA extract was produced to allow for crude RNA quantitation using a Nanodrop ND-1000 microspectrophotometer (Thermo Scientific, Wilmington, DE), and to make 60µl of cDNA.

The resulting RNA was then treated with DNase using TURBO DNA-*free*™ (Life Technologies, Grand Island, NY) following manufacturer instructions to remove any remaining DNA from the sample. TURBO DNase (1.5µl) and 10X TURBO DNase buffer (1µl) were added to the RNA and incubated at 37°C for 30 minutes. Then DNase inactivation reagent (1.5µl) was added and mixed by pipet every minute for 5 minutes at room temperature. In the next step, centrifuging the sample at 10,000g for 1.5 minutes caused the added DNase inhibitor to pellet to the bottom of the tube. With a pipet, the purified RNA was recovered and placed in a new conical tube. A Nanodrop ND-1000 (Thermo Scientific, Waltham, MA) reading was recorded for RNA extractions at each storage time-point. The Nanodrop readings ranged from 6.8 ng/µl – 100.3 ng/µl for RNA samples.

cDNA Production

Complimentary DNA (cDNA) was produced from the RNA extracts using a High Capacity cDNA Reverse Transcription Kit (Life Technologies, Carlsbad, CA) along with Ribonuclease Inhibitor (Affymetrix®) according to the manufacturer's protocol. This process, known as reverse transcription, produced the cDNA template that was subsequently used as qPCR template to quantify the relative amounts of mRNA transcripts. All samples were stored at

-20°C until quantification was completed. It is assumed that the efficiency of the cDNA synthesis reactions were equal for each RNA extract from each donor at each time-point.

Primer Design

Protamine 1 (*PRM1*) and semenogelin 1 (*SEMG1*) mRNA transcripts were selected to be quantified for this project based on published research on semen degradation profiles.[12,18,32,34] Two sets of primers and 2 probes were selected for each sequence. (See Table 1) One set of primers chosen to direct the amplification of a PCR product of approximately 100 base-pairs (bp) while the second primer set was chosen to produce an amplicon of be approximately 200 base-pairs. The primers did not overlap with one another and had similar annealing temperatures. The National Center for Biotechnology Information's (NCBI) website was used as a source of the nucleotide sequence of *PRM1* and *SEMG1* genes and was used for primer and probe design. Primer-BLAST, also available on the NCBI website, was used to design the primers and probes. Primer lengths were selected to allow for the quantification of one long and one short segment of each fluid-specific marker and all primers and probes exhibited comparable melting-point values. All *PRM1* and *SEMG1* primers/probes were multiplexed separately.

Table 1. Selected Nucleotide Sequences for Primers and Probes for *PRM1* and *SEMG1*

	Length	Probe Fluorescence		Sequence
PRM1	200 bp	FAM	Forward	5'CACCATGGCCAGGTACAGAT3'
			Reverse	5'TCCTTCGTCTGCGACTTCTT3'
			Probe	5'GGAGCAGATATTACCGCCAG3’
	96 bp	NED	Forward	5'ACCAAACTCCTGCCTGAGAA3'
			Reverse	5'CCTTAGCAGGCTCCTGATTTT3'
			Probe	5’CTTGCCACATCTTGAAAATGCCACC5’
SEMG1	195 bp	FAM	Forward	5'CTCAAATCCAGGCACCAAAT3'
			Reverse	5'GTGCCAAATGACGATCACTG3'
			Probe	5’AACATGGATCTCATGGGGGATTGGA3’
	99 bp	NED	Forward	5'TTTCCCTGCTCCTCATCTTG3'
			Reverse	5'CGTGTGGAAATTGGGAAAAT3'
			Probe	5’AGCAAGCAGCTGTGATGGGACAAAA3’

Control Genes

Three control genes were selected as possible candidates to normalize *PRM1* and *SEMG1* quantitation for delta delta C_T analysis. Ribosomal RNA *18S* (61 bp and 187 bp) and glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) were selected based on past studies on RNA degradation.[7,9,11] Using two different *18S* targets allowed for degradation analysis on the control genes to maximize our ability to detect degradation among the control transcripts if it occurred.

Table 2. Information about Selected Control Genes

Gene ID	Length	Fluorescence	Catalog Number
18S	187 bp	VIC	4319413E
18S	61 bp	VIC	NM_002046
GAPDH	122 bp	VIC	4326317E

Control genes were bought from Applied Biosystems and ThermoFisher Scientific.

Primer Optimization

Multiplexing is a procedure that allows multiple primer and probe sets to be mixed in a single reaction and reduces the number of quantitative polymerase chain reactions (qPCR) necessary for quantification by real-time qPCR. Optimization of each primer and probe set must occur in singleplex to access the efficiency of the PCR reaction before multiplexing can occur. Optimization of primer and probe concentrations were also performed in singleplex according to a predetermined matrix. Table 2 shows the matrix demonstrating the concentration of primers and probes used to determine optimal concentrations for qPCR. Each reaction was analyzed for delta R_n (ΔR) and threshold cycle (C_T) values to optimize amplification conditions. A high ΔR value and a low C_T are most ideal, but these multiplexed values must be compared to the singleplex reaction values to make sure no inhibition of qPCR is taking place. C_T is the cycle number at which the fluorescence passes the baseline threshold, which is automatically set by the ABI 7500 to remove any background noise. The R_n is the ratio of fluorescence emission of the reporter dye over the reference dye, ΔR is calculated by subtracting the baseline from R_n.

Table 3. Matrix for Optimization of Primers and Probes

Probe (nM)	Primer Concentrations (nM)					
450	900	450	300	250	150	75
350	900	450	300	250	150	75
250	900	450	300	250	150	75
150	900	450	300	250	150	75

An ABI 7500 thermal cycler (Applied Biosystems, Foster City, CA) was used for qPCR for primer optimization and quantification of samples. The amplification cycling parameters were 1 minute at 95°C, with the repetition of 40 cycles of amplification for 15 seconds at 95°C and extension for 34 seconds at 60°C, with a final hold at 4°C until the plate was removed. Optimized primer/probe concentrations can be viewed in Table 4.

Table 4. Optimized Primer and Probe Concentrations for Real-Time Polymerase Chain Reaction

		Concentration (nM)
***PRM1* (200 bp)**	Forward	250
	Reverse	250
	Probe	450
***PRM1* (96 bp)**	Forward	250
	Reverse	250
	Probe	450
***SEMG1* (195 bp)**	Forward	450
	Reverse	450
	Probe	450
***SEMG1* (99 bp)**	Forward	900
	Reverse	900
	Probe	450

Primer Efficiencies

To be able to normalize qPCR results for a target transcript versus the control transcript, the efficiency of the amplification process during qPCR must be determined for each set of primers and probes in singleplex. A comparable efficiency for each qPCR target (i.e. not varying by more than 3%) indicates the acceptability of normalizing the abundance of a specific transcript in different samples (i.e. *PRM1* and *SEMG1*) to one or more control genes whose abundance does not change. It is necessary to also evaluate the efficiency of amplicon detection using the Taqman probes used to detect the amplicons produced during qPCR. The efficiency of qPCR reaction process was determined by analyzing a dilution series with 5X 5-fold dilutions of a cDNA preparation tested in triplicate.

To determine the efficiency of the qPCR method, the C_T values for both target and control transcripts are plotted on the y-axis against the *log* of the template amount on the x-axis. The slope of the resulting standard curve is determined and efficiency is calculated using the formula $10^{(-1/slope)}$-1. (Table 5) Efficiency should be 100% +/- 10% for singleplex and multiplexed primers and probes. When multiplexed, the difference in efficiencies of the various targets should differ by <0.3. *PRM1* long and short could be multiplexed but control genes had to be run in singleplex, the same was true for *SEMG1*.

Table 5. Efficiencies of Control and Designed Targets in Singleplex

	Slope	R2	Efficiency
***18S* (61 bp)**	-3.4408	0.9947	0.9527
***18S* (187 bp)**	-3.5989	0.9990	0.8961
GAPDH	-3.0502	0.9892	1.1274
***PRM1* (96 bp)**	-3.5545	0.9977	0.9113
***PRM1* (200 bp)**	-3.4787	0.9988	0.9385
***SEMG1* (99 bp)**	-4.3378	0.9983	0.7003
***SEMG1* (195 bp)**	-3.4598	0.9972	0.9455

Statistical Analysis

Threshold value (C_T) is the PCR amplification cycle number at which the fluorescence abundance associated with PCR amplicons passes an arbitrary threshold value.[8] The threshold value prevents any background noise unrelated to the amplified segments from being evaluated. The cut off value for target genes was C_T=35 which for the purposes of this study was the limit of detection of transcripts present in the cDNA preparation and presumably indicated that a mRNA transcript was not present in a sample. The cutoff C_T value of 35 was determined from the standard curves produced during the development of the qPCR method used here and thus minimized skewing statistical analysis of the results. [8] Additionally, *18S* (187 bp) was used to normalize the total amounts of RNA to account for the different amounts of RNA present in each sample by Delta delta C_T ($\Delta\Delta C_T$) analysis.[8]

The equations for $\Delta\Delta C_T$ statistical analysis:

$$\Delta C_T = C_{T\,(\text{target})} - C_{T\,(18S)}$$

$$\Delta\Delta C_T = \Delta C_T - \Delta C_{T\,(\text{time-point 0})}$$

$$\text{Ration} = 2^{(-\Delta\Delta C_T)}$$

C_T is determined by the ABI 7500 by taking into account background fluorescence, automatic baseline value, and at what point fluorescence begins to accumulate exponentially in the qPCR reaction with each round of PCR. In a degradation study it would be expected for the C_T value to increase the older the sample indicating the decomposition of mRNA transcripts in the semen stain. When the C_T of the control gene, in this study *18S* (187 bp), is subtracted from

the C_T of the *PRM1* or *SEMG1* transcripts, the normalized amount of the target mRNA can be determined. The normalized value of the target (known as ΔC_T) should be increasing as the semen stains age if mRNA degradation is taking place. It is further possible to subtract the ΔC_T values for each time-point during the storage period from that determined for the time-point zero samples (producing a value known as $\Delta\Delta C_T$) to compare the ΔC_T of all time-points to the initial ΔC_T of the sample. Subtracting ΔC_T from each of the time-points from the value determined for time 0 calculates the increase or decrease in transcript abundance as a function of sample aging. If the amount of the target sequence is decreasing then we would expect a higher $\Delta\Delta C_T$ value in each consecutive time-point. The final step of $\Delta\Delta C_T$ analysis is calculating $2^{(-\Delta\Delta C_T)}$, this calculation converts $\Delta\Delta C_T$ values back into a number of transcript molecules present in a sample and provides the most accurate interpretation of C_T value because C_T value is calculated in log form. If the degradation of the target sequence is occurring, we would expect a $2^{(-\Delta\Delta C_T)}$ value to decrease, with a starting value of 1 at time-point 0.

Summary

Analyzing RNA degradation can be done in an efficient manner. Multiplexing primers for fluid- and tissue-specific markers can help streamline qPCR and provide more accurate degradation information. Finding the correct control gene for RNA degradation studies is important to the overall success of the study. RNA degradation information can be of use to forensic scientists and investigators to determine the relationship of evidence to a crime.

CHAPTER IV

RESULTS

The goal of this project was to study the degradation patterns of two human tissue-specific messenger RNA (mRNA) transcripts isolated from semen samples from volunteer donors. Ten donated semen samples were used as a source of mRNA for protamine 1 (*PRM1*) and semenogelin 1 (*SEMG1*). Dried stains prepared from donated samples were subjected to mRNA extraction every 2 weeks for 6 months to determine the abundance levels of the *PRM1* and *SEMG1* mRNA transcripts. It was hypothesized that one or both of these transcripts would degrade during storage at room temperature in a predictable way. It was further hypothesized that the degradation kinetics of these transcripts could be used as a measure of time, which would be useful to estimate when a semen stain was deposited at a crime scene.

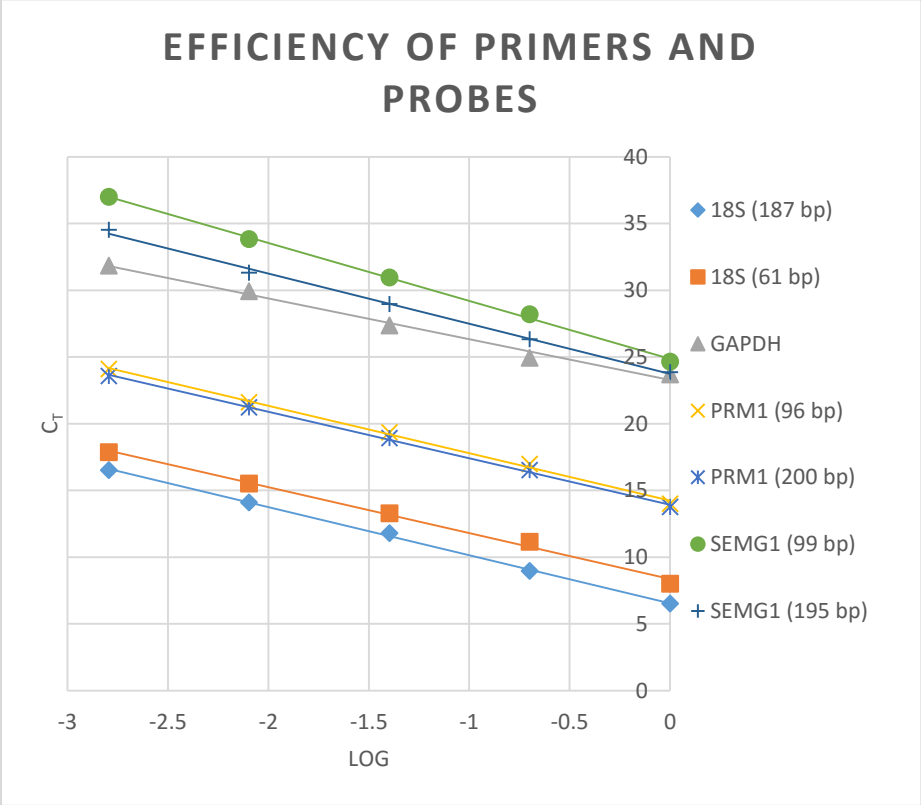
Selection of Primers

PRM1 and *SEMG1* were selected for this project based on a past study performed at Oklahoma State University Center for Health Sciences that showed both markers degraded in a linear fashion over a period of 6 months when analyzed using RNA sequencing.[18] Two sets of primers and Taqman qPCR probes were designed for each gene using the National Center for Biotechnology Information's website (www.ncbi.nlm.nih.gov/). One set of primers and probe were designed to amplify and detect a large PCR amplicon and the other set targeted the amplification and detection of a small PCR amplicon that did not overlap with the larger product. The rationale of this approach was that the large PCR amplicon might be a more sensitive

indicator of mRNA degradation since a larger transcript size would be necessary to successfully amplify the larger target during qPCR analysis.

The control genes *18S* ribosomal RNA (*rRNA*) sizes 61 bp and 187 bp, along with glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) were selected based on past literature which suggests that they are relatively stable and abundant in semen.[7,9,11] Primers were optimized following the protocol Taqman® Multiplex PCR Optimization developed by Applied Biosystems (Foster City, CA). Once primers were optimized, the amplification efficiency was determined for each set of primers and probe. (Figure 1) All primers and probes had a R^2 of 0.98 or higher, suggesting that all dilution curves were accurate.

Figure 1. Efficiency of Primers and Probes Graphed on a Log Scale



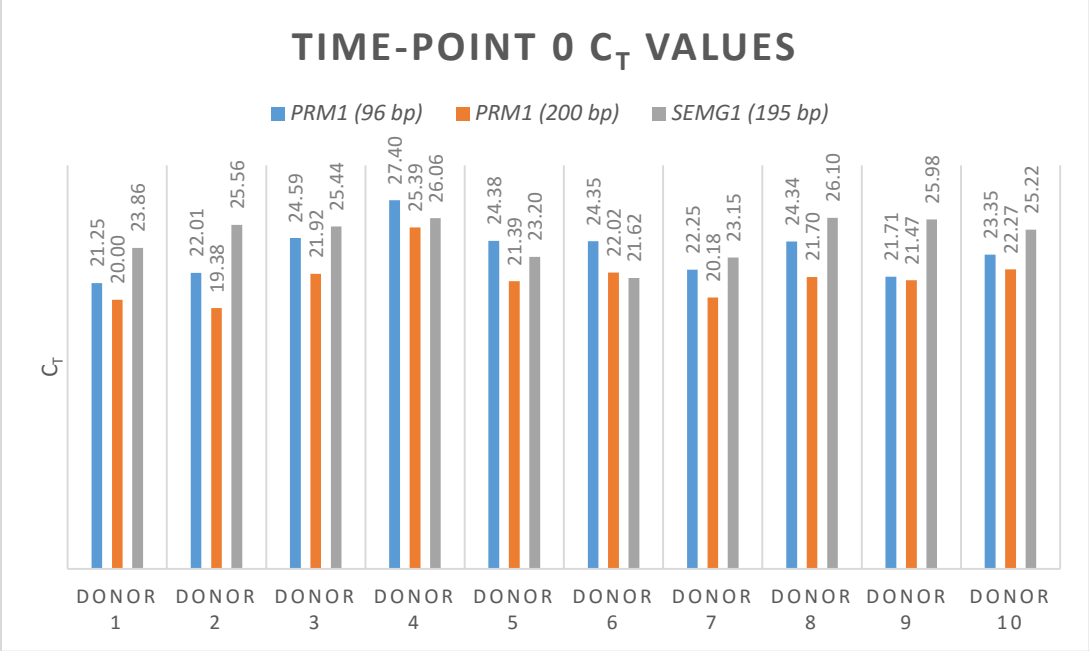
The log graph shows the standard curves for all targets and control genes. The slope of the standard curves was used to calculate the efficiency of each set of primers and probes to determine if control genes could be used as normalizers for $\Delta\Delta C_T$ analysis.

C_T Analysis

The time-point 0 C_T values varied from about 19 to 30 C_T for *PRM1* (200 bp) and *SEMG1* (195 bp). Compared to *18S* (187 bp) that ranged from 10 to 13 C_T and *GAPDH* that ranged from 23 to 27 C_T. The time-point 0 C_T values provide insight into the variation in mRNA abundance for all the different gene markers in a small population sampling. Shown in Figure 2 are the abundance levels of the transcripts for *PRM1* (96 bp, 200 bp) and *SEMG1* (195 bp) which have been normalized to the abundance of *GAPDH* transcript present. Each donor's time-point 0 *GAPDH* C_T was divided by the *GAPDH* C_T of donor 1. The difference in C_T value for each

donor relative to donor 1 was then multiplied by the C_T values for the specific gene transcripts. This approach normalized each donor to the other donors based upon the expression of a control gene whose abundance is expected to be comparable among the donors. This provides an idea of individual variation in *PRM1* and *SEMG1* abundance when RNA values are equivalent.

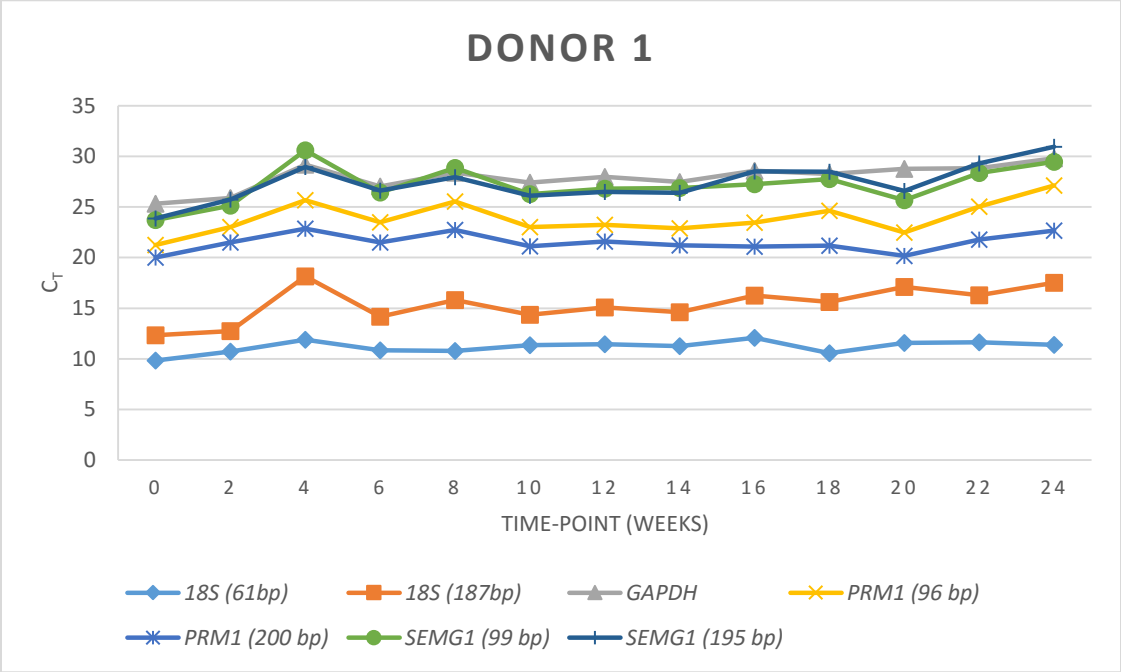
Figure 2. C_T Values for Time-point 0 in *PRM1* (96 bp & 200 bp) and *SEMG1* (195 bp) Based on *GAPDH* Being Normalized for Each Donor

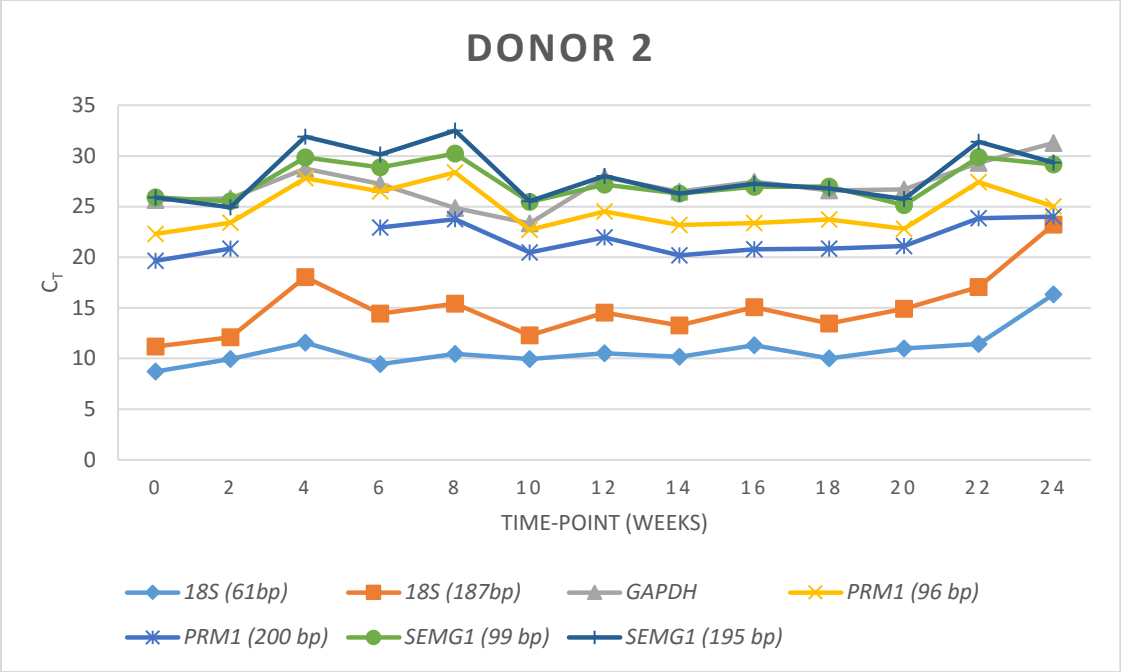


Figures 3A – 3J show C_T values for all targets and control genes for each individual. Large variations in C_T values were seen among individuals and time-points. For example, donors 9 and 10 had time-points producing C_T values greater than 35, meaning there was evidently little or no *PRM1* and *SEMG1* present. However, these individuals exhibited an abundance of control gene transcript that was comparable to the other donors. Thus, there appears to be variation in the expression of *PRM1* and *SEMG1* genes among some males which may reflect differences in some aspect of spermatozoa function and ultimately fertility for the donor. It should be noted that all donors had at least 1 time-point where there was a sharp increase or decrease of targets and

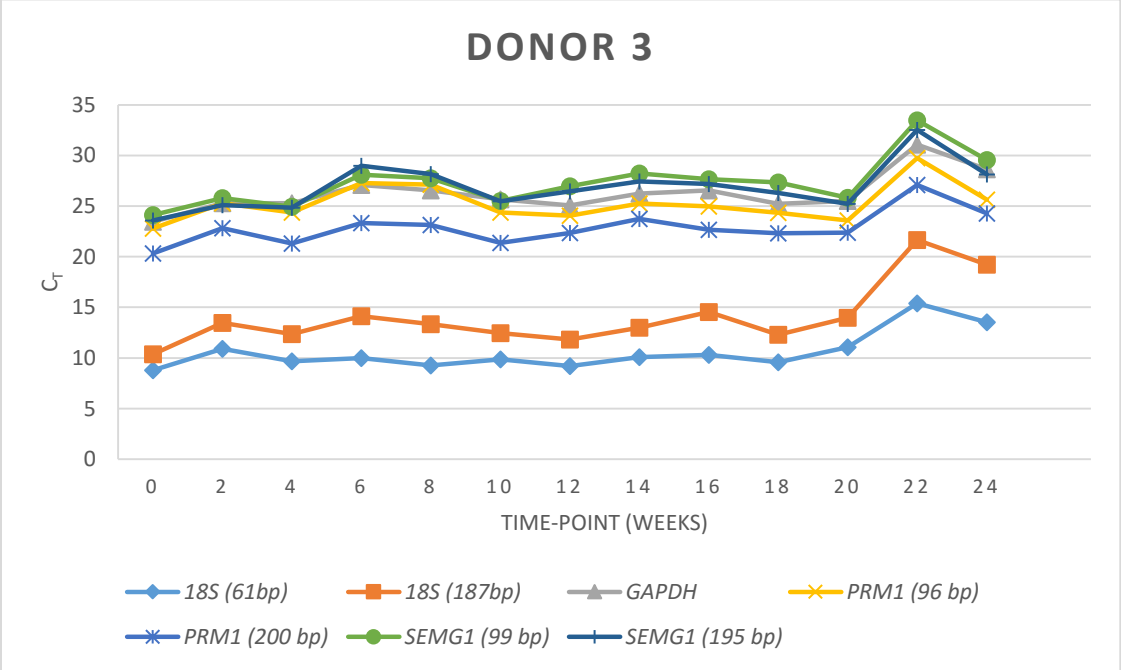
control genes, this is most likely due to inherent technical variability in the qPCR methodology, or low RNA recovery from the dried semen stains.

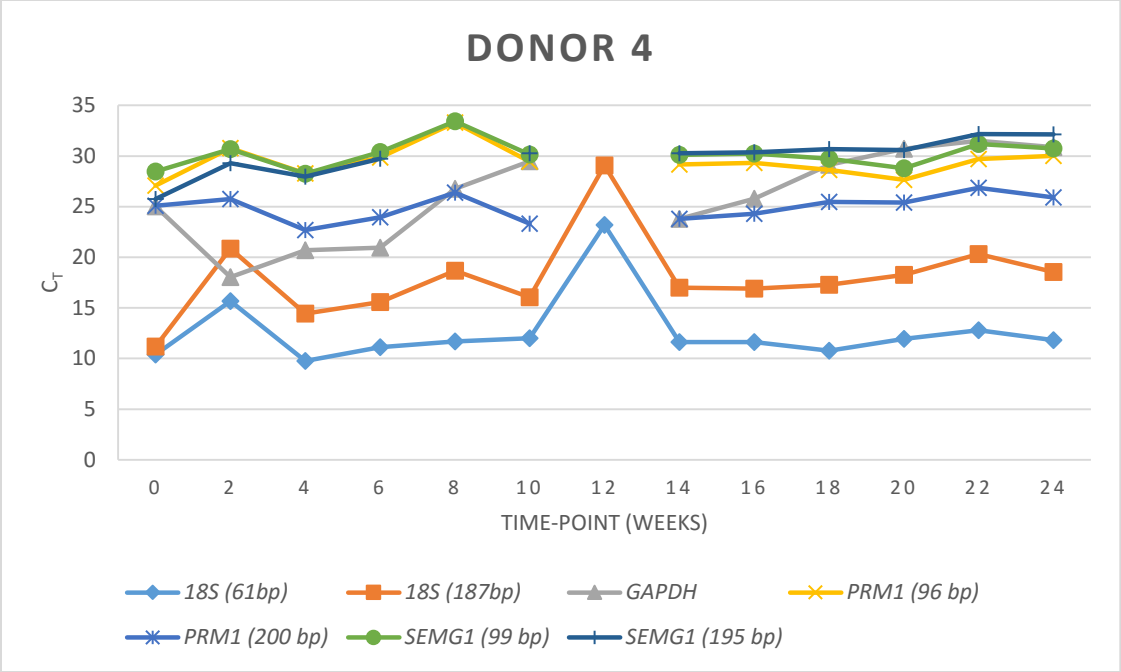
Figure 3A – 3J. Line graphs of *PRM1*, *SEMG1*, *18S*, and *GAPDH* C_T Values for All Time-points for Each Donor.



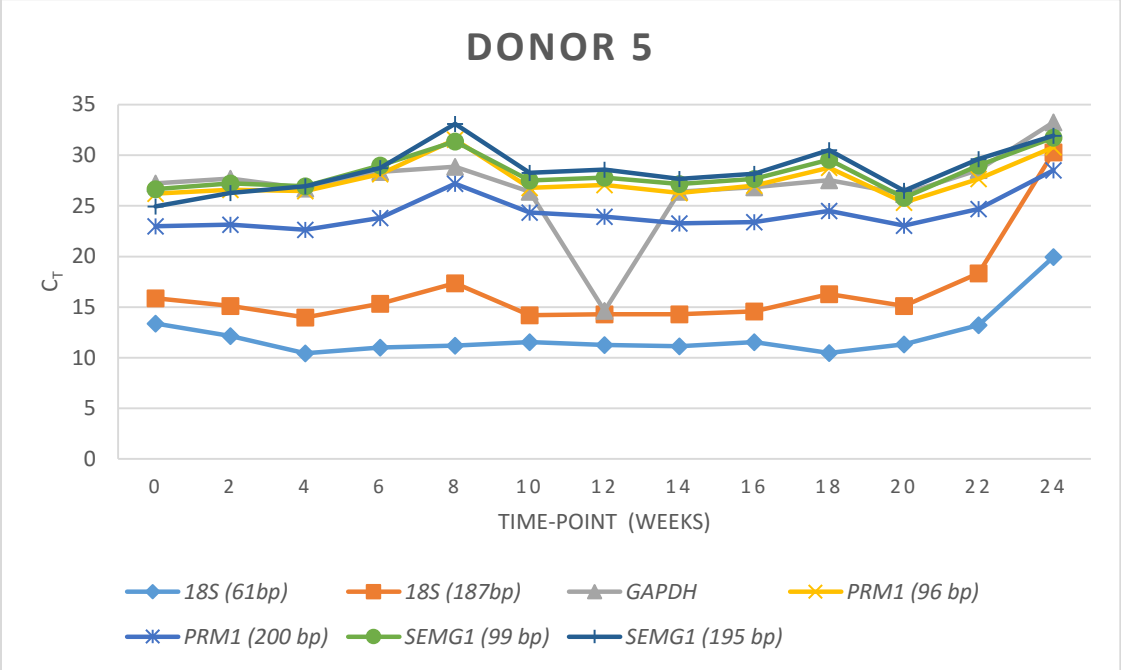


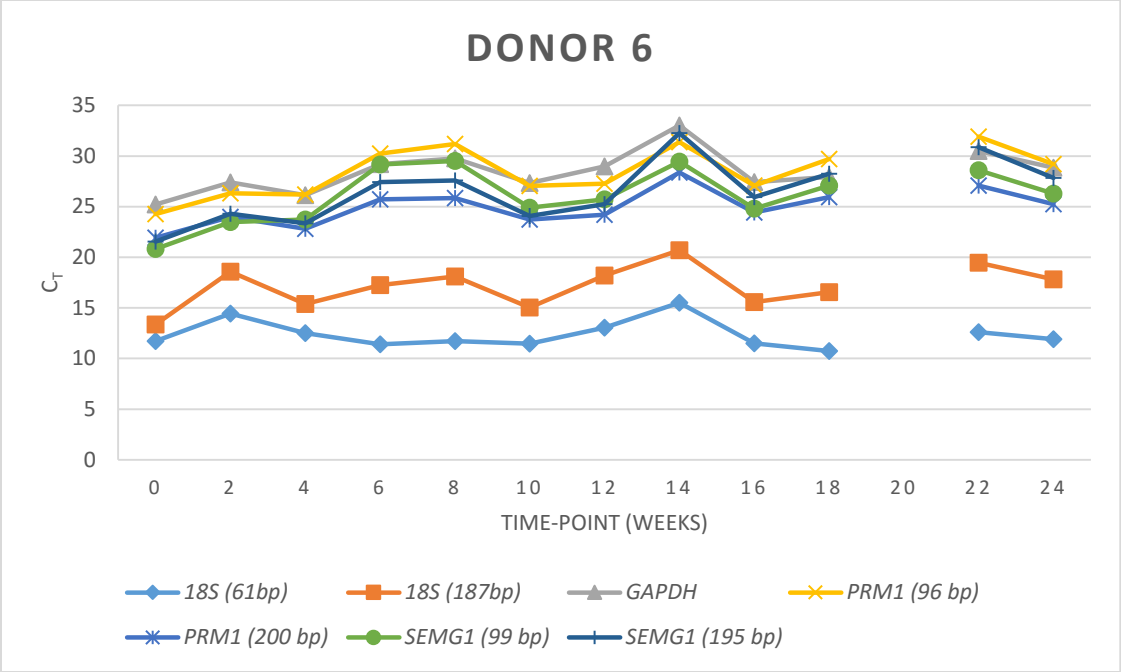
* any missing data points had no CT value



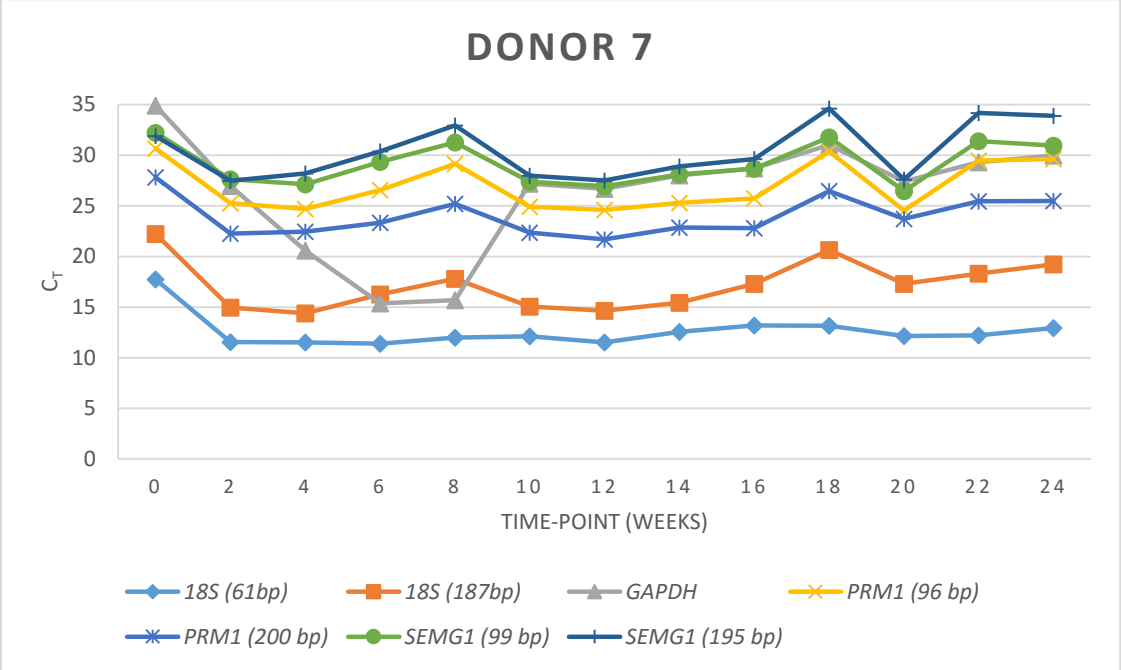


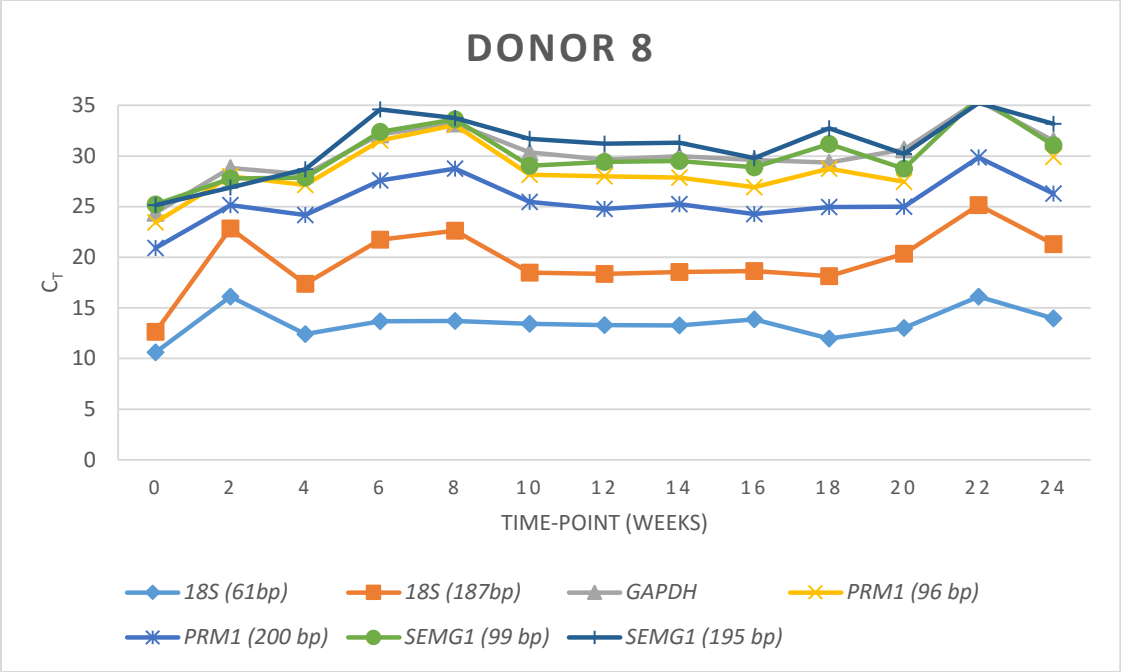
* any missing data points had no CT value

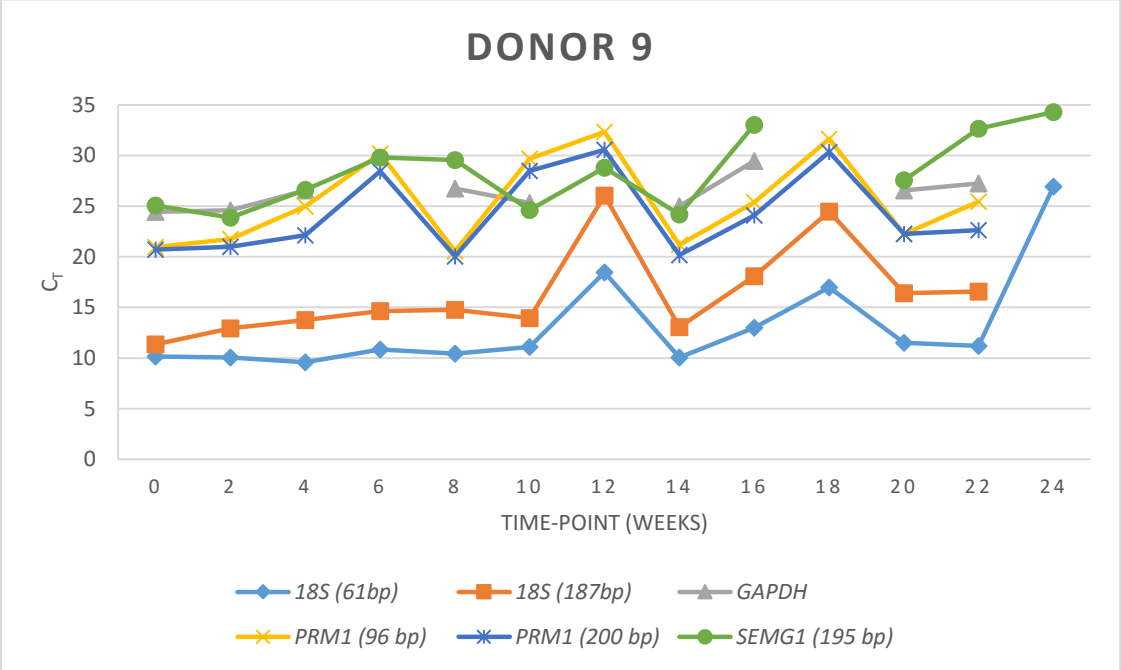




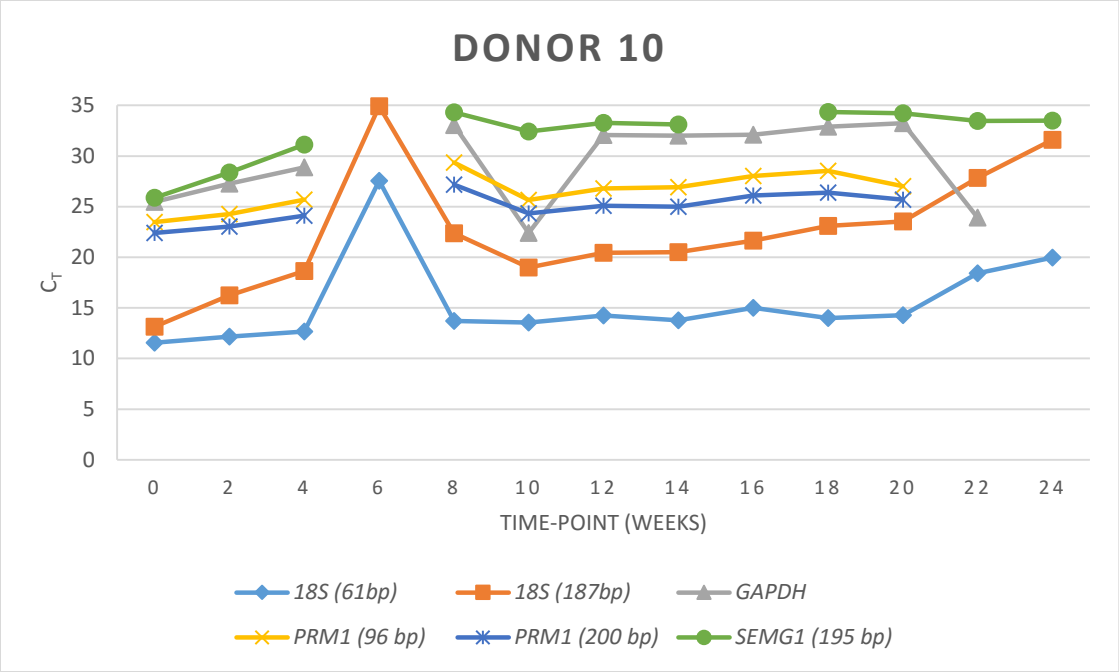
* any missing data points had no CT value







* any missing data points had no CT value



* any missing data points had no CT value

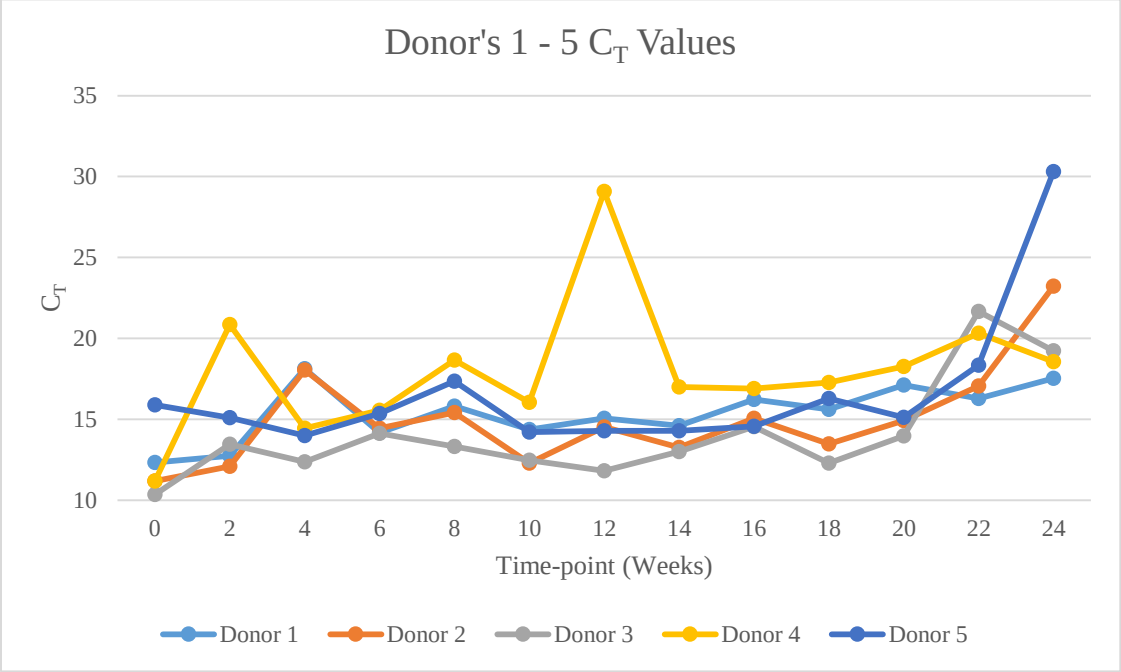
Variability in the results produced from this study was significant and prevented meaningful statistical analysis from being performed. Several factors may have contributed to the variation observed among donors and time-points. One source of variation could arise from variation in the nature of the biological sample provided by the different donors in terms of sperm count, age and overall health of the donor, diet, or frequency of ejaculation. There are a lot of characteristics that could affect the overall quantity and quality of RNA recovered from semen stains of each individual. Variability could also be contributed from technical variation in setting up qPCR reactions or variation introduced during cDNA synthesis. Finally, variation could also be introduced into the results from the basic nature of semen in which sperm often clump together with protein present making uniform pipetting difficult during preparation of the semen stains.[35,36] The nature of semen made it difficult for every dried stain to be exactly 50µl, it has been shown that overall cell concentration can affect quantity and quality of RNA when extracted from a

sample.[37] The decreased quality of RNA, due to having to few cells or to many cells, would reduce the efficiency of cDNA synthesis in turn affecting the accuracy of qPCR quantification.

Delta Delta C_T Analysis

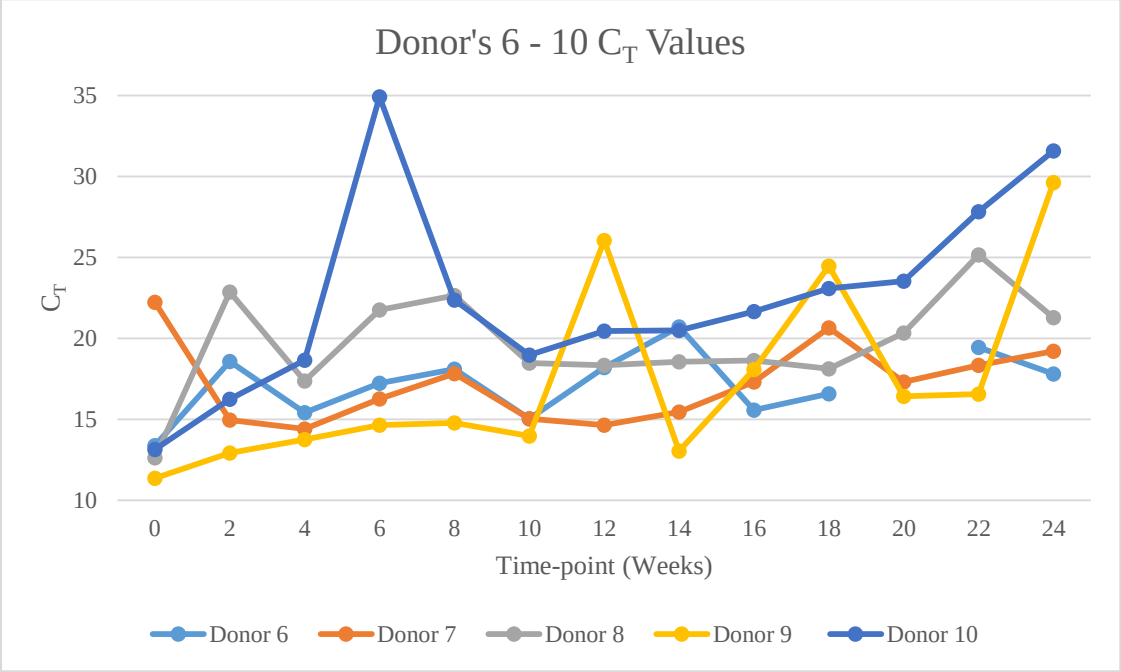
The threshold for C_T was set automatically by the ABI 7500 to remove any background noise from analysis. Only C_T values with a normal amplification curve of fluorescence accumulation and C_T values between 10-35 were analyzed. C_T values meeting these criteria were analyzed using the delta C_T (ΔC_T) method to normalize *PRM1* and *SEMG1* transcript abundance to *18S* (187 bp) *rRNA*. Figure 4 and 5 show variation among semen donors in *18S rRNA* (187 bp) transcript abundance over the 24 week time-course. This variation could due to the difference in overall RNA content in each individual's donation or the amount of RNA extracted from the dried stain.

Figure 4. Individual Donor's (1-5) *18S* (187 bp) C_T Values Over 24 Weeks



A lot of variation was seen in 18S, this is most likely due to variation in RNA extraction from each time-point. Donor 4's large increase at time-point 12 could be due to a low input of RNA in cDNA synthesis. No other extreme changes was linked to low or high cDNA input in qPCR.

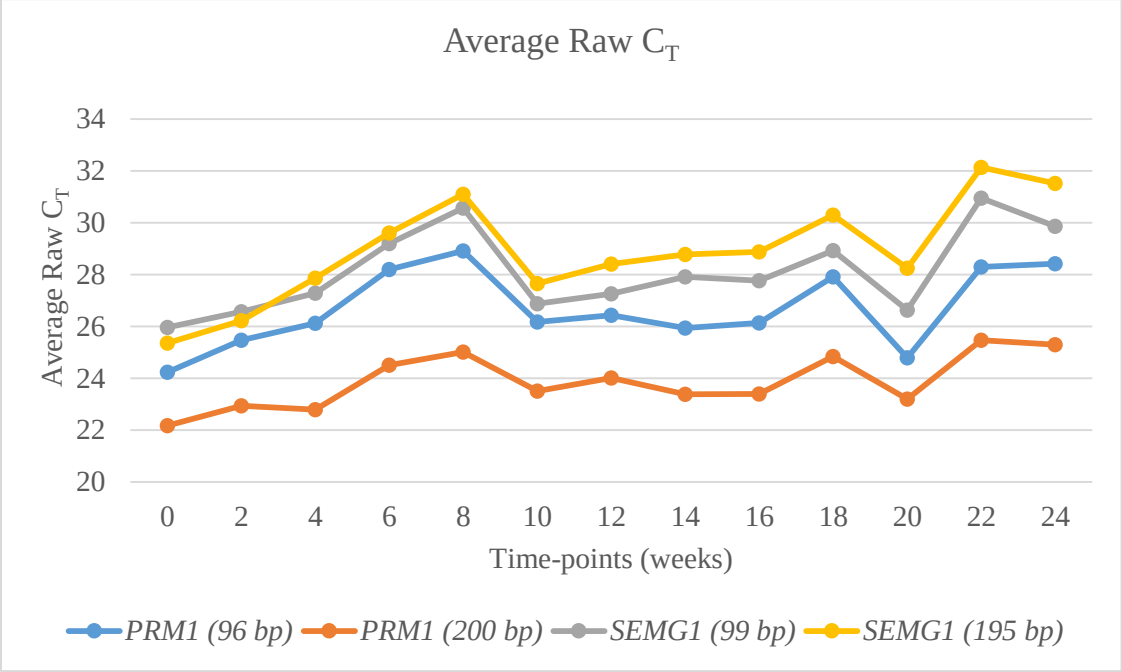
Figure 5. Individual Donor's (6-10) *18S* (187 bp) C_T Values Over 24 Weeks



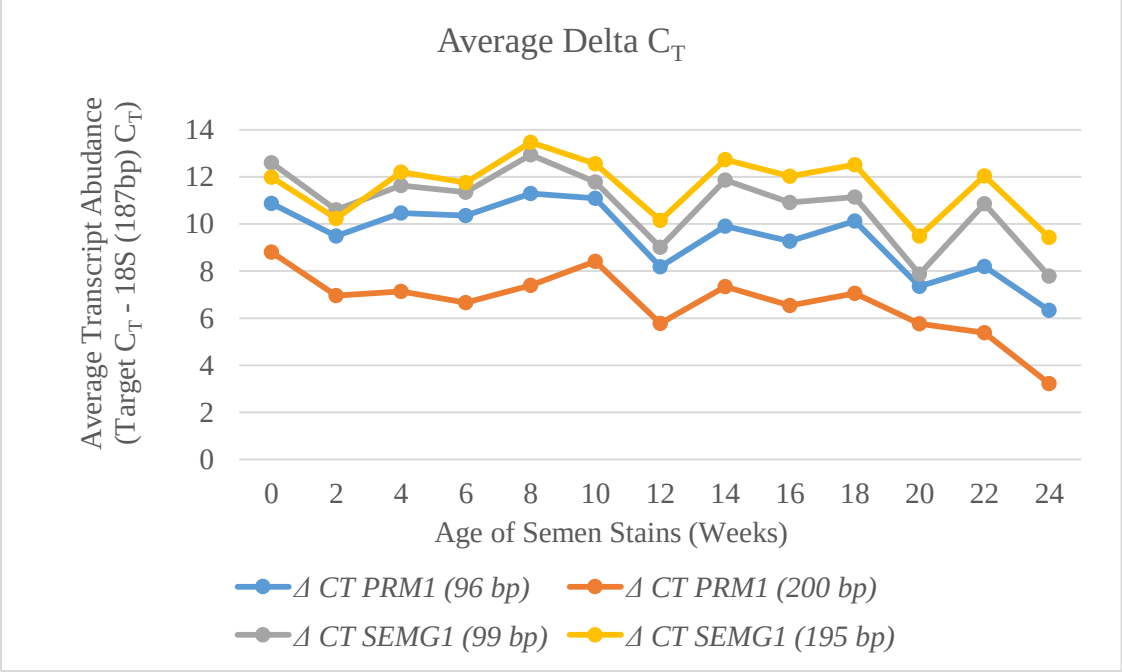
Donor 10's large increase at time-point 6 could be due to low cDNA input in qPCR. Donor 9's time-point 12 large increase is most likely due to inhibition caused by a large input of cDNA in qPCR. Donor 6 had no values for any targets at time-point 20, a low RNA recovery could have caused no detectable targets.

While some samples did appear to have a difference in abundance between 24 weeks and 0 weeks, the abundance differences were not significant. Additionally, no significant trends were detected when examining the abundance data either by C_T, ΔC_T, or $2^{(-\Delta\Delta C_T)}$. (Figures 6A – 6C)

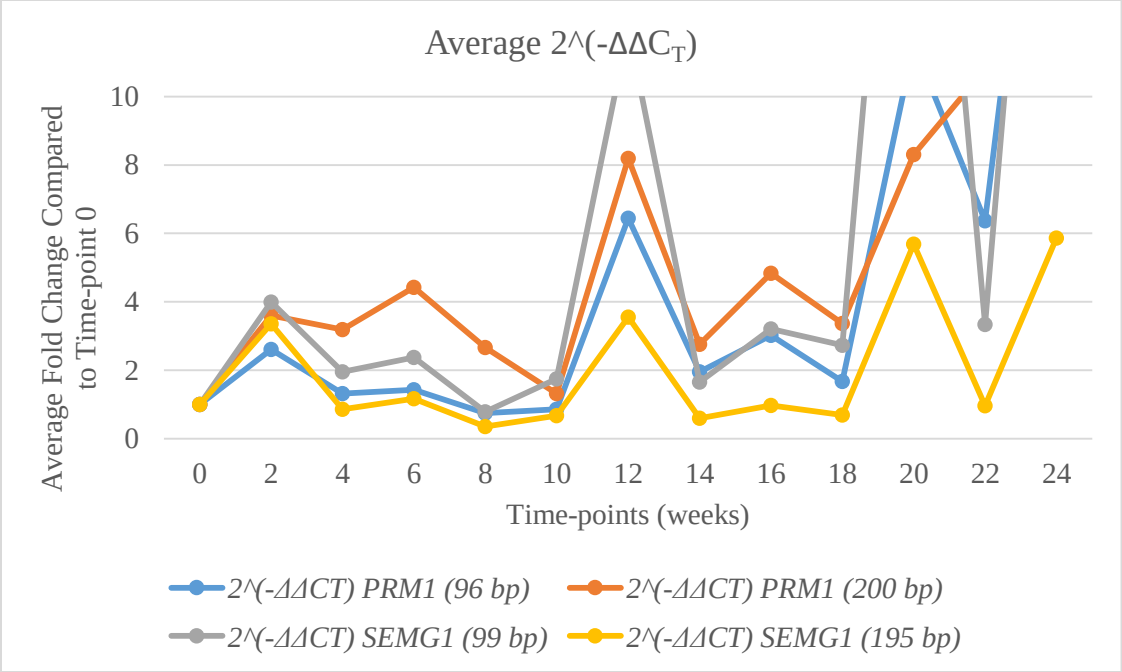
Figure 6A-6C. Changes in Abundance of *PRM1* and *SEMG1* mRNA Transcripts During 24 Weeks of Storage at Room Temperature



There was great variation between donors, but the graph of averaged CT values provides an idea of overall trends among all 10 donors.



The standard deviation and standard error for average ΔC_T values can be seen in Appendix B. Being unable to eliminate outliers form collected data could be influencing the average ΔC_T graph and providing an inaccurate visualization of the data as a whole.



There was great variation among donors in addition to the possible degradation of 18S could be causing the large fold changes that are indicated in the $2^{(-\Delta\Delta CT)}$ analysis. With a large variation among donors and the difference in primer efficiencies this data is most likely an inaccurate representation of the possible change in abundance of designed target sequences.

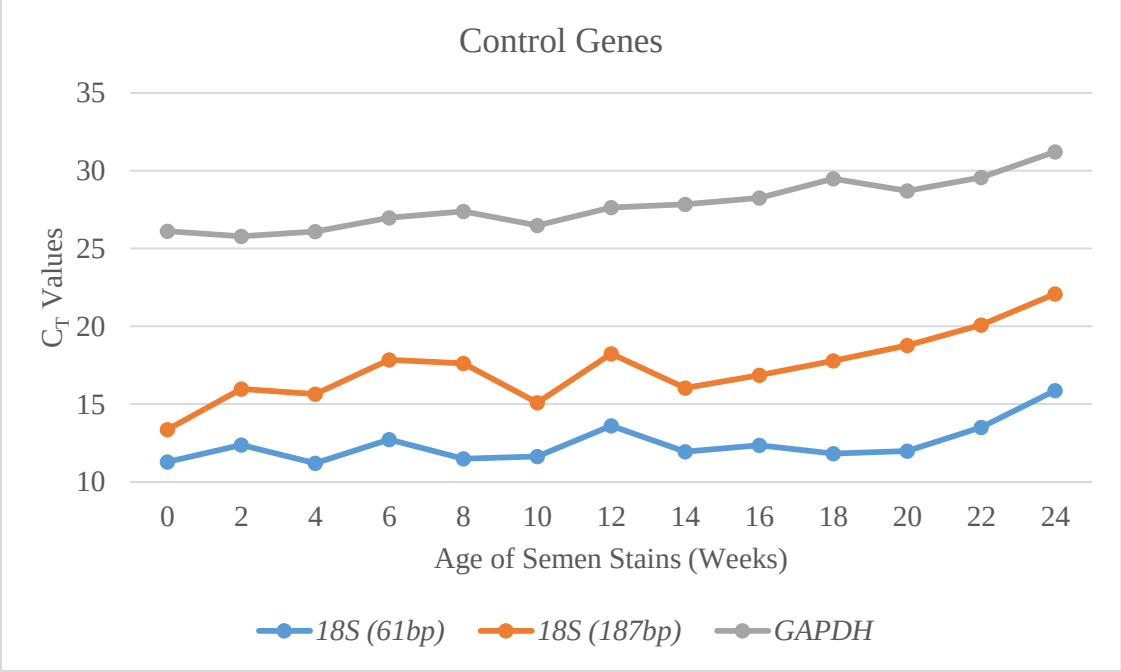
Control Genes

18S and *GAPDH* are common control genes used for studies in which differences in the abundance of specific transcripts are to be measured relative to a stably expressed internal control.[4-9] Control gene abundance can also be used as an interaction control for mRNA degradation studies. Selecting a short (61 bp) and long (187 bp) *18S rRNA* amplicons allowed for the comparison of *SEMG1* and *PRM1* abundance levels to the internal control *18S rRNA* of the same relative sizes as those from the tissue-specific transcripts.

While completing the statistical analysis, a trend of increasing raw C_T values were noticed in the selected control genes. The C_T value for the *GAPDH* appears to increase slightly over the 24 week time-course and its starting abundance is less than that observed for *18S rRNA*. *18S* (187 bp) *rRNA* appears relatively stable up to 14 weeks and then its C_T appears to increase at

a faster rate suggesting enhanced degradation. The C_T observed for the 61 bp *18S rRNA* transcript began to increase at about 20 weeks and seems to have an increase that is comparable to *18S* (187 bp). (Figure 7) While additional data from more than 24 weeks is needed, this increase in C_T suggests that degradation may be taking place in the control genes.

Figure 7. Average of 10 Donors C_T Values for *18S* (61 bp), *18S* (187 bp), and *GAPDH*



The line graph shows the C_T values of all 10 donors averaged together. An increase of C_T values was observed over the 24 week time course indicating possible degradation in housekeeping genes.

Summary

With the great variability among donors and the possible degradation of control genes it is difficult to determine if subtle degradation of the transcripts from *PRM1* and *SEMG1* genes is taking place. Because of the high degree of variation associated with results obtained in this

study, tests for significance could not be performed. However, overall our results suggest that the mRNAs for *PRM1* and *SEMG1* do not degrade appreciably in semen stains aged for 6 months at room temperature in a dark environment. A continued study with a longer time course would be able to provide more insight into the rate of degradation of the target and control genes used in this study.

CHAPTER V

DISCUSSION

In this study primers and probes were designed for use in a qPCR assay that will detect and quantify messenger RNAs (mRNAs) which are species- and semen-specific for the gene transcripts of protamine 1 (*PRM1*) and semenogelin 1 (*SEMG1*). Protocols developed specifically for biological fluid identification and age detection were used in this study, as well as the genes and target transcripts previously described in degradation studies.[4,5,7,11,18]

Setzer et al study showed *PRM1* remained stable up to 180 days in different environmental conditions.[11] Our data supports Setzer's conclusion that *PRM1* is a relatively stable mRNA transcript and has potential for being used to determine the age of samples over a year old.[11] Nakanishi et al preformed a study demonstrating the presence of *PRM2* in samples as old as 56 years.[7] While *SEMG1* was not detected in samples ranging from 33 – 56 years old,[7] the presence of *SEMG1* in fresh samples, with minimal degradation taking place in 6 months may support the use of *SEMG1* as a good marker to determine the age of semen samples older than 6 months. Targets used for determining the age of a semen sample need to degrade at a noticeable rate over time. While no degradation was seen in *PRM1* and *SEMG1* in 6 months, Nakanishi's study shows that *SEMG1* degrades to an undetectable level around or before 33 years and decreased amounts of *PRM2* are present in samples up to 56 years old compared to fresh samples.

Research done by Weinbrecht showed that *PRM1* and *SEMG1* transcripts degrade over 6 months when using next generation sequencing to look at the semen transcriptome as a whole.[18] In sequencing the transcriptome, the Ion Torrent PGM™ used RNA fragments that were approximately 200 bp in length.[18] The data collected by Weinbrecht demonstrated that there was a gradual decrease in the total amount of *PRM1* and *SEMG1* mRNA present in samples aged to 6 months.[18] Compared to the data collected in this study, Weinbrecht showed a higher degradation rate with a linear slope over 6 months.[18] The difference in the results of these two studies could be due to the quantification methods. Degradation is believed to happen randomly, but evenly, throughout a sequence and it is expected to see a greater amount of degradation in a larger segment of RNA. The cDNA synthesized from RNA extracts in the Weinbrecht study was sheared sonically to an average size of ~200 basepairs. However, review of the distribution of cDNA fragments sequenced in each run demonstrated that the average size of sequenced fragments was in the 180 basepair range. So, the slight increase in degradation of SEMG1 and PRM1 observed using the Ion Torrent PGMTM technology versus qPCR could be due to the size differences of the transcript fragments analyzed.

There is a lack of published literature in which a population of semen donors has been subjected to transcript abundance analysis for any gene transcript. Thus it is difficult to know if the variation in *SEMG1* and *PRM1* transcript abundance we observed among our 10 random donors is due to individual variation in gene expression or the handling of semen samples before RNA extraction was performed. The great variability among transcript abundance between time-points from the same donor, is most likely due to technical difficulties in performing the qPCR reaction setup and/or variation in the recovery of comparable amounts of RNA from each dried semen stain. The high degree of variation prevented meaningful statistical analysis from being performed to assess whether there is an enhanced rate of degradation in specific transcripts versus control genes and whether any degradation observed is steady of the 6 month time course of

storage or if there is an enhancement in the rate of degradation beyond a certain time point. Human spermatozoa are complex and much is still not known about the transcriptome in semen and spermatozoa, and the purpose of mRNA (like *PRM1*) in the spermatozoa head.[37,38] Additional research needs to be done on semen to better understand how RNA is being used for production of spermatozoa/seminal fluid, and on RNA's role in fertilization. It is difficult without this knowledge to fully understand the results that are being seen in degradation studies.

Processing Semen for RNA

Semen presents unique problems when it comes to RNA isolation.[35,36,39] Seminal fluid has a high content of protein and polysaccharides which can hinder RNA isolation and promote clumping of spermatozoa in an ejaculate.[36,37,39] The high content of polysaccharides results in a high A260/A280 ratio when evaluating RNA concentration using spectrophotometry, and this prevents the A260/A280 ratio from being used to determine if residue from the reagents used during extraction are present the sample.[37] While the quality of RNA can still be predicted with Nanodrop analysis, other methods of determining RNA quantity and quality should be used if possible. The Qubit flurometer (Invitrogen), Spectraophotometer 3000 (Quawell), or a bioanalyzer (Agilent) could provide more accurate information about the quantity and quality of extracted semen RNA.[37,39,40]

While the quality of RNA has not been correlated in previous studies with the number of spermatozoa in the sample, a positive correlation between RNA quality and cell number has been observed for other cell types.[37] Spermatozoa clump together and are difficult to pipette, so the clumping of spermatozoa may have caused an uneven number of spermatozoa and epithelial cells, to be pipetted during preparation of the individual semen stains used for time-point analysis in this study. The variation in the amount of cells in each sample could impact the quantity and quality of RNA extracted from the sample adding variation between time-points that is not due to

aging. Although many methods have been used for the extraction of RNA from semen,[7,12,37,39,40] high quality RNA cannot be collected from semen using the same RNA extraction methods used with other biological fluids in degradation studies. Semen presents a unique set of problems because of its chemical make-up and these factors must be considered when designing a degradation study. Geordiadis et al proposed an extraction method using TRIzol heated to 37°C and a reduced extraction time to avoid additional degradation while the semen sample was being handled.[39] This would reduce any additional degradation from occurring that was not due to regular aging by environmental conditions.

Initial RNA Concentrations

Semen presents problems with pipetting and homogeneity that other biological fluids do not. Sperm cells stick together and can become difficult to make uniform stains. This difficulty impacts the total amount of sample pipetted on each blood card. With sperm and seminal fluid having unique properties and different concentrations of DNA and RNA present, each 50 µl sample pipetted onto the stain card may have varied greatly in the amount of RNA extracted at each time-point. The A260 absorbance of RNA in extracted ranged from ~7 ng/ul of extract to ~100 ng/ul of extract, a difference of greater than 10 fold. Moreover, within the set of semen stains for a particular individual, this variation in cells deposited could contribute to the great variation in data produced for one time point relative to the next. Even with the variation encountered, there did appear to be a trend among donors, because if a donor had a high amount of RNA present at time-point 0 then the donors Nanodrop values generally stayed close to the initial RNA value at all extraction time-points. (Appendix B) This suggests that there is individual variation in how much RNA is present based on the donor, this could be due to the total number of cells present in the ejaculate or it could be individual variation that occurs between males.

Factors Affecting Degradation Rate

Several factors can affect degradation rates in studies where samples are kept under stable environmental conditions with minimal exposure to light. In this study, volunteers collected semen samples at their convenience, and delivered them to the laboratory within 24 hours of the first RNA extraction (time-point 0 sample). Volunteers were asked to not use lubricant during collection, but there is no way to verify that this instruction or other instructions in the IRB consent paperwork were followed.

Additionally, all volunteers were asked to store semen samples in the refrigerator until they were delivered to the laboratory. Refrigerators can have a range of temperatures that are acceptable to keep perishable items cold, but the difference of a few degrees could have an impact on the RNA contained within the sample. Once picked up, samples were placed in a cooler with ice packs and transported to the laboratory, which introduced an additional temperature change. Another variable in temperature change was that a few volunteers transported their samples to the laboratory and placed them in the refrigerator at their convenience on the day of the first RNA extraction. With so many variables affecting temperature change before spotting of samples took place, it is likely that RNA degradation had already begun taking place at different rates in each sample. The variation in the environment may have impacted the samples in such a way to have an overall impact on the degradation pattern of the sample for the entire study.

Changes in Methodology

After completing this project, it is evident that there are some flaws in the methodology. Self-collection of semen samples should have been performed on site at Oklahoma State University – Center for Health Sciences (OSU-CHS) allowing for more control over the many variables, such as storage temperature and time until extraction for time-point 0. While this

would have put more stress on the volunteer participants, it would have eliminated some uncertainty about storage and about how much time had passed since self-collection. Collection on site at OSU-CHS could also have eliminated unknowns, such as the use of lubrication, length of time until extraction, and changes in storage temperatures. Additional instructions should have been included asking donors not to ejaculate 72 hours before donating to the study, this would have reduced another variable.

Also an estimate of sperm concentration should have been completed before spotting the samples on blood cards. Knowing an estimate of sperm present would have provided additional information about RNA concentration and about whether the amount of sperm present in a sample had any effect on degradation rates. Sperm has a stronger membrane than an epithelial cell and the RNA in the sperm head is better preserved than cells in the seminal fluid. This additional protection from external elements will affect the degradation rate.

Finally, the possibility of an alternative statistical method should be investigated in a future study. Delta delta C_T ($\Delta\Delta C_T$) presents problems in degradation studies. These problems together with the reduced amount of RNA present in some samples due to non-uniform stains could cause decreased synthesis of cDNA due to increased time of finding the target sequence. Other statistical methods have been used in qPCR studies,[4,10,23,41] but the majority of biological fluid degradation studies are now using $\Delta\Delta C_T$ analysis. Currently, $\Delta\Delta C_T$ has flaws that makes analyzing degradation information difficulty and could be causing false negative results. Looking at raw C_T values provides little information about degradation as well, because many factors can affect C_T values cycle to cycle making results difficult to compare. Additional study and analysis of statistical procedures will help reduce inaccurate degradation information.

Applications of Research

Using a multiplexed primer kit with species- and tissue-specific markers can provide another confirmatory test to assist forensic scientists in identifying unknown biological samples. With most forensic laboratories using automated extraction, incorporation of RNA for qPCR analysis could be easily integrated into laboratory protocols. Having a quick and easy confirmatory test could also speed up the processing of unknown samples.

Degradation rates of mRNA could provide forensic scientists and investigators with crucial information about biological samples at crime scenes. More research needs to be done to determine the most effective way to determine the amount of degradation occurring over time in biological fluids. Results from additional research would help establish how the degradation rate correlates to the sample's time of deposit. Additionally, studies looking at how environment affects RNA degradation will help produce a protocol that can be applied in a real world environment.

Future Investigation

Additional research needs to be done focusing on control genes, *18S* and *GAPDH*, and how they degrade over time. While we know *18S* is a rather stable ribosomal RNA (*rRNA*), this study suggests that *18S* is degrading at a notable rate over 6 months. Being able to determine how much degradation is taking place in control genes will enable a more accurate normalization formula for the target sequences. Knowing more about the degradation patterns of *rRNA* and mRNA will help improve primer and probe development for real-time PCR (qPCR) protocols. Furthermore, looking at where primers are located in the selected target sequence could provide information about the pattern of degradation in *rRNA* and mRNA. Investigating whether mRNA degrades more from the 5' end or the 3' end, or if degradation is more likely to happen in a certain part of the sequence would also help with the accuracy of age prediction.

More research should also be done with a larger sample size, and with reduced variability in collection and extraction. Having control over the variability of collection has the potential to add valuable information to semen degradation studies. Adding variables like collecting a sample when a spermicide, lubricant, or condom is known to have been used will also add valuable information to semen degradation studies. If mRNA analysis is adopted into a crime lab setting, semen degradation analysis will most commonly be used in sexual assault cases. In sexual assault cases, it is rare for a semen sample to be collected without lubricant or vaginal secretions. These will likely have an effect on the degradation rate of collected semen samples.

Summary

Complications arose during this study that provided valuable information for further mRNA degradation studies. The stability of control genes like *18S* and *GAPDH* are not as stable as previously thought. Using *18S* and *GAPDH* may still be useful in mRNA degradation studies, but using them as normalizer in $\Delta\Delta C_T$ analysis provides inaccurate statistics. The physical nature of semen presents its own issue in degradation studies because of the two types of cells present and the nature of spermatozoa. Methodology improvements can help reduce the variability between each studied sample and provide new information about how spermatozoa and seminal fluid mRNA degrade differently.

Analysis of mRNA has great potential in forensic laboratories. The more information that can be provided to forensic scientists and investigators the easier it will be to solve a criminal case. This study designed two species- and tissue-specific primers that can be used to identify spermatozoa and seminal fluid. With additional study these same primers may be able to be used to determine the age of semen stains.

REFERENCES

1. Li R. *Forensic biology.* Boca Raton: CRC Press/Taylor & Francis; 2008.

2. Habib S. *Identification of sperm-specific mRNA markers through RNA sequencing*: Forensic Science, Oklahoma State University; 2014.

3. Darnell JE. *RNA : life's indispensable molecule.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 2011.

4. Anderson S, Howard B, Hobbs GR, Bishop CP. A method for determining the age of a bloodstain. *Forensic Sci Int.* 2005;148(1):37-45.

5. Anderson SE, Hobbs GR, Bishop CP. Multivariate analysis for estimating the age of a bloodstain. *J Forensic Sci.* 2011;56(1):186-193.

6. Nussbaumer C, Gharehbaghi-Schnell E, Korschineck I. Messenger RNA profiling: a novel method for body fluid identification by real-time PCR. *Forensic Sci Int.* 2006;157(2-3):181-186.

7. Nakanishi H, Hara M, Takahashi S, Takada A, Saito K. Evaluation of forensic examination of extremely aged seminal stains. *Leg Med.* 2014;16(5):303-307.

8. Sakurada K, Ikegaya H, Fukushima H, Akutsu T, Watanabe K, Yoshino M. Evaluation of mRNA-based approach for identification of saliva and semen. *Leg Med.* 2009;11(3):125-128.

9. Juusola J, Ballantyne J. mRNA profiling for body fluid identification by multiplex quantitative RT-PCR. *J Forensic Sci.* 2007;52(6):1252-1262.

10. Park SM, Park SY, Kim JH, et al. Genome-wide mRNA profiling and multiplex quantitative RT-PCR for forensic body fluid identification. *Forensic Sci Int Genet.* 2013;7(1):143-150.

11. Setzer M, Juusola J, Ballantyne J. Recovery and stability of RNA in vaginal swabs and blood, semen, and saliva stains. *J Forensic Sci.* 2008;53(2):296-305.

12. Juusola J, Ballantyne J. Multiplex mRNA profiling for the identification of body fluids. *Forensic Sci Int.* 2005;152(1):1-12.

13. Richard ML, Harper KA, Craig RL, Onorato AJ, Robertson JM, Donfack J. Evaluation of mRNA marker specificity for the identification of five human body fluids by capillary electrophoresis. *Forensic Sci Int Genet.* 2012;6(4):452-460.

14. Marchuk L, Sciore P, Reno C, Frank CB, Hart DA. Postmortem stability of total RNA isolated from rabbit ligament, tendon and cartilage. *Biochim Biophys Acta.* 1998;1379(2):171-177.

15. Inoue H, Kimura A, Tuji T. Degradation profile of mRNA in a dead rat body: basic semi-quantification study. *Forensic Sci Int.* 2002;130(2-3):127-132.

16. Bauer M, Polzin S, Patzelt D. Quantification of RNA degradation by semi-quantitative duplex and competitive RT-PCR: a possible indicator of the age of bloodstains? *Forensic Sci Int.* 2003;138(1-3):94-103.

17. Bauer M, Gramlich I, Polzin S, Patzelt D. Quantification of mRNA degradation as possible indicator of postmortem interval--a pilot study. *Leg Med (Tokyo).* 2003;5(4):220-227.

18. Weinbrecht K. *RNA-seq of biological fluids for the evaluation of mRNA degradation in relation to sample age*: Forensic Science, Oklahoma State University; 2014.

19. Russell PJ. iGenetics : a Mendelian approach. 2006; xx. 842 p. ill. (chiefly col.) 829 cm. + 841 CD-ROM (844 843/844 in.). Available at.

20. Meister G. *RNA biology : an introduction.* Weinheim: Wiley-VCH; 2011.

21. Fordyce SL, Kampmann ML, van Doorn NL, Gilbert MT. Long-term RNA persistence in postmortem contexts. *Investig Genet.* 2013;4(1):7.

22. Koppelkamm A, Vennemann B, Lutz-Bonengel S, Fracasso T, Vennemann M. RNA integrity in post-mortem samples: influencing parameters and implications on RT-qPCR assays. *Int J Legal Med.* 2011;125(4):573-580.

23. Vennemann M, Koppelkamm A. Postmortem mRNA profiling II: practical considerations. *Forensic Sci Int.* 2010;203(1-3):76-82.

24. Belasco JG, Brawerman G. *Control of messenger RNA stability.* San Diego, CA: Academic Press; 1993.

25. Vennemann M, Koppelkamm A. mRNA profiling in forensic genetics I: possibilities and limitations. *Forensic Sci Int.* 2010;203(1-3):71-75.

26. Zubakov D, Kokshoorn M, Kloosterman A, Kayser M. New markers for old stains: stable mRNA markers for blood and saliva identification from up to 16-year-old stains. *Int J Legal Med.* 2009;123(1):71-74.

27. Kimura A, Ishida Y, Hayashi T, Nosaka M, Kondo T. Estimating time of death based on the biological clock. *Int J Legal Med.* 2011;125(3):385-391.

28. Hanson EK, Lubenow H, Ballantyne J. Identification of forensically relevant body fluids using a panel of differentially expressed microRNAs. *Anal Biochem.* 2009;387(2):303-314.

29. Bowden A, Fleming R, Harbison S. A method for DNA and RNA co-extraction for use on forensic samples using the Promega DNA IQ system. *Forensic Sci Int Genet.* 2011;5(1):64-68.

30. Zubakov D, Boersma AW, Choi Y, van Kuijk PF, Wiemer EA, Kayser M. MicroRNA markers for forensic body fluid identification obtained from microarray screening and quantitative RT-PCR confirmation. *Int J Legal Med.* 2010;124(3):217-226.

31. Steger K, Pauls K, Klonisch T, Franke FE, Bergmann M. Expression of protamine-1 and -2 mRNA during human spermiogenesis. *Mol Hum Reprod.* 2000;6(3):219-225.

32. Bauer M, Patzelt D. Protamine mRNA as molecular marker for spermatozoa in semen stains. *Int J Legal Med.* 2003;117(3):175-179.

33. Li H, Wu C, Gu X, Xiong C. A novel application of cell-free seminal mRNA: non-invasive identification of the presence of germ cells or complete obstruction in men with azoospermia. *Hum Reprod.* 2012;27(4):991-997.

34. Roeder AD, Haas C. mRNA profiling using a minimum of five mRNA markers per body fluid and a novel scoring method for body fluid identification. *Int J Legal Med.* 2013;127(4):707-721.

35. Huang S, Li H, Ding X, Xiong C. Presence and characterization of cell-free seminal RNA in healthy individuals: implications for noninvasive disease diagnosis and gene expression studies of the male reproductive system. *Clin Chem.* 2009;55(11):1967-1976.

36. Wang L, Lv J, Guo C, Li H, Xiong C. Recovery of cell-free mRNA and microRNA from human semen based on their physical nature. *Biotechnol Appl Biochem.* 2014;61(3):342-348.

37. Barragán M, Martínez A, Llonch S, Pujol A, Vernaeve V, Vassena R. Effect of ribonucleic acid (RNA) isolation methods on putative reference genes messenger RNA abundance in human spermatozoa. *Andrology.* 2015;3(4):797-804.

38. Martins RP, Krawetz SA. RNA in human sperm. *Asian Journal of Andrology.* 2005;7(2):115-120.

39. Georgiadis AP, Kishore A, Zorrilla M, et al. High Quality RNA in Semen and Sperm: Isolation, Analysis and Potential Application in Clinical Testing. *The Journal of Urology.* 2015;193(1):352-359.

40. Xu Y, Xie J, Cao Y, et al. Development of highly sensitive and specific mRNA multiplex system (XCYR1) for forensic human body fluids and tissues identification. *PLoS One.* 2014;9(7):e100123.

41. Grabmuller M, Madea B, Courts C. Comparative evaluation of different extraction and quantification methods for forensic RNA analysis. *Forensic Sci Int Genet.* 2015;16:195-202.

APPENDICES

APPENDIX A

Approval by the Oklahoma State University

Institutional Review Board

College of
Osteopathic Medicine

CENTER FOR HEALTH SCIENCES

Office of Research and Sponsored Programs
1111 West 17th Street
Tulsa, Oklahoma 74107-1898
(918) 561-1400
Fax (918) 561-1416

Institutional Review Board
FWA # 00005037

Memo

To: Allison Sherier
Forensics

Cc: Robert Allen, PhD
Forensics

From: Richard Wansley, PhD
Chairman, Institutional Review Board

Date: August 26, 2015

Re: **Expedited Approval of Amendment**
IRB Protocol # 2013008

Titled: Optimization of RNA-Seq to Determine Degradation Profiles of Transcriptome and Identity of Forensically Relevant Samples

Board members of the OSU-CHS Institutional Review Board (IRB) reviewed and approved the below listed items for IRB Protocol # 2013008. It has been determined the items meet expedited criteria under federal guidelines, 45 CFR 46.

- Research Protocol (Aim #2), version 8/28/15
- Additional investigator: Allison Sherier
- Participant Consent Form, version 08/19/2015

As principal investigator it is your responsibility to assure:

- The research is conducted in accordance with the IRB-approved protocol, including, when applicable, the approved recruitment and consent procedures;
- When informed consent is required, informed consent is obtained prior to the initiation of any study-related procedures;
- When written informed consent is required, informed consent is obtained and documented using the current IRB-stamped approved research consent form;

APPENDIX B

Nanodrop Values of RNA (ng/µl) for Each Donor at Each Time-Point

		Donor Number									
		1	2	3	4	5	6	7	8	9	10
Time-points (Weeks)	**0**	21.28	33.23	22.00	17.21	15.42	26.19	31.91	18.84	61.96	22.32
	2	18.20	41.40	24.60	14.50	15.50	35.10	23.00	15.50	51.30	27.50
	4	17.20	27.80	36.90	16.60	33.50	24.30	26.00	10.70	76.30	12.20
	6	20.70	35.30	35.90	15.50	8.50	22.40	17.00	10.80	66.20	19.80
	8	24.10	37.50	33.50	12.20	18.20	21.70	22.40	13.40	54.10	14.20
	10	18.80	35.80	30.60	12.10	16.60	17.80	29.90	9.80	77.80	15.00
	12	18.00	39.20	42.40	6.80	13.80	20.40	24.90	12.20	100.30	12.30
	14	13.60	41.00	32.10	13.60	12.90	21.30	27.50	38.30	53.10	10.90
	16	23.70	38.60	31.50	17.40	14.00	16.90	20.80	13.60	66.90	6.90
	18	17.90	36.20	34.20	11.20	13.70	18.70	12.00	21.90	80.70	8.00
	20	20.70	46.00	30.30	11.30	10.30	12.80	26.00	7.40	61.70	20.60
	22	18.60	21.20	19.70	11.20	18.10	14.00	36.30	11.00	89.40	21.21
	24	25.90	25.70	15.80	27.70	11.20	11.00	19.10	7.70	22.50	17.10

APPENDIX C

Tables showing the variation in ΔCT in *PRM1* (96 bp and 200 bp) and *SEMG1* (99 bp and 195 bp) minus the average of *18S* (187 bp, labeled ENDO in the table).

PRM1 S - ENDO

Age (weeks)	N	Mean	S.D.	S.E.	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
0	10	7.31	1.79	0.57	6.75	7.88	5.42	11.53
2	10	7.42	2.13	0.67	6.75	8.10	5.39	12.60
4	10	8.48	2.11	0.67	7.81	9.14	5.61	13.32
6	9	9.01	2.18	0.73	8.29	9.74	6.13	14.00
8	10	10.08	3.52	1.11	8.97	11.20	3.24	14.29
10	10	8.44	2.11	0.67	7.77	9.11	5.31	12.89
12	9	6.60	2.70	0.90	5.70	7.50	4.54	13.67
14	10	7.33	2.17	0.68	6.65	8.02	4.83	11.70
16	10	6.98	2.24	0.71	6.27	7.69	4.48	11.23
18	10	8.21	2.09	0.66	7.55	8.87	5.21	11.29
20	9	5.51	1.66	0.55	4.96	6.06	3.32	7.78
22	8	7.24	1.57	0.55	6.69	7.80	6.11	11.08
24	9	5.59	3.12	1.10	4.49	6.70	1.41	9.67

PRM1 L -ENDO

Age (weeks)	N	Mean	S.D.	S.E.	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
0	10	5.25	1.77	0.56	4.69	5.81	2.85	9.55
2	10	4.90	1.33	0.42	4.48	5.32	2.57	7.55
4	9	5.14	1.39	0.46	4.68	5.61	3.14	7.73
6	9	5.32	1.59	0.53	4.79	5.85	4.16	9.01
8	10	6.18	2.11	0.67	5.51	6.85	2.74	10.02
10	10	5.76	2.33	0.74	5.03	6.50	3.40	11.69
12	9	4.19	2.39	0.80	3.39	4.98	2.84	10.54
14	10	4.77	1.41	0.45	4.33	5.22	2.92	7.28
16	10	4.24	1.54	0.49	3.76	4.73	2.13	6.26
18	10	5.14	1.56	0.49	4.65	5.63	3.04	7.51
20	9	3.92	1.56	0.52	3.40	4.44	1.03	5.54
22	10	4.42	0.95	0.32	4.10	4.74	2.85	6.23
24	8	2.48	2.02	0.71	1.76	3.19	0.40	5.72

SEMG1 S - ENDO

Age (weeks)	N	Mean	S.D.	S.E.	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
0	9	9.04	2.61	0.87	8.17	9.91	4.06	12.94
2	9	8.53	2.72	0.91	7.62	9.44	3.36	12.48
4	9	9.64	2.11	0.70	8.93	10.34	5.72	13.28
6	9	10.01	2.23	0.74	9.26	10.75	8.76	14.99
8	9	11.73	2.08	0.69	11.04	12.43	9.65	16.10
10	9	9.14	1.20	0.40	8.73	9.54	6.95	10.97
12	7	7.43	2.71	1.02	6.40	8.45	5.65	14.40
14	9	9.30	1.78	0.59	8.71	9.89	6.37	12.65
16	8	8.62	1.67	0.59	8.03	9.21	6.63	12.12
18	8	9.23	1.03	0.36	8.86	9.59	8.65	11.66
20	7	6.82	0.81	0.30	6.51	7.12	6.53	8.97
22	8	9.90	1.16	0.41	9.49	10.31	7.75	11.43
24	8	6.81	2.39	0.85	5.97	7.66	3.91	10.33

SEMG1 S - ENDO

Age (weeks)	N	Mean	S.D.	S.E.	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
0	10	8.43	1.93	0.61	7.82	9.04	4.76	10.72
2	10	8.18	2.28	0.72	7.46	8.90	4.15	11.09
4	10	10.21	2.27	0.72	9.49	10.93	5.34	12.99
6	9	10.42	2.51	0.84	9.58	11.26	8.16	16.05
8	10	12.26	3.08	1.03	11.23	13.29	7.72	17.78
10	10	9.91	2.17	0.68	9.23	10.60	6.15	14.11
12	9	8.57	3.79	1.26	7.31	9.83	2.04	15.17
14	10	10.17	1.35	0.43	9.74	10.60	8.13	12.79
16	9	9.73	1.58	0.53	9.20	10.25	7.75	12.85
18	9	10.60	1.23	0.41	10.19	11.01	9.83	13.02
20	9	8.43	0.98	0.33	8.11	8.76	7.45	10.56
22	10	11.08	1.83	0.58	10.51	11.66	9.60	14.30
24	10	8.46	2.99	0.95	7.51	9.40	4.09	13.18

VITA

Allison J Sherier

Candidate for the Degree of

Master of Science

Thesis: PATTERNS OF DEGRADATION IN *PRM1* AND *SEMG1* TRANSCRIPTS IN SEMEN DURING 6 MONTHS OF AGING

Major Field: FORENSIC SCIENCES

Biographical:

Education:

Completed the requirements for the Master of Science in Forensic Science at Oklahoma State University Center for Health Sciences, Tulsa, Oklahoma in July, 2016.

Completed the requirements for the Bachelor of Science in Animal Science – Biotechnology at Oklahoma State University, Stillwater, Oklahoma in 2014.

Professional Memberships: American Academy of Forensic Scientists