A Survey of RNA Methylation Within Exon 7 of the Bovine *PRKAB1* gene and its Potential Role in Residual Feed Intake Within Livestock Species

Sean Quigley

Honors College Thesis

Abstract

Nucleic acids are able to undergo covalent modification in a variety of ways resulting in functional regulation at the epigenetic level. Understanding these epigenetic changes may provide insight into gene regulation, which can be applicable to any field within the life sciences that deal with DNA. Methylation of the carbon at the fifth position in the nucleic acid cytosine (5mC) within DNA is one of the most well documented epigenetic modifications. In eukaryotic organisms, 5mC is associated with down-regulating the gene in which the methylated cytosine is associated with. Although cytosine methylation is well documented at the genetic level, little is known about it at a post-transcriptional level. Cytosine methylation within RNA has been observed in mRNA, tRNA, and rRNA; however, demonstrable rationality for this epitranscriptomic modification has yet to be elucidated. There has been documentation of stable and long-lived RNA molecules containing methylated cytosines, possibly insinuating that a role in RNA stability. The dearth of knowledge regarding post-transcriptional modification of RNA can be explained by the fact 5mC is not easily detectable through most sequencing methods and that 5mC does not affect base pair binding, precluding the use of any probing assays. In this study, a survey of mRNA methylation within exon 7 of the bovine *PRKAB1* gene is conducted using skeletal muscle of six individual heifer cows in an effort to better understand the biochemical mechanisms driving differences in observed residual feed intake (RFI) among commercial cattle. Skeletal muscle tissue is used since feed efficiency affects muscle growth within cattle. Here we show

methylation of cytosines at the post-transcriptional level is detectable through bisulfite treatment of RNA, followed by conversion to cDNA, PCR amplification and finally restriction digest, showing a potential association of differential methylation within animals that have high and low RFIs. These results give preliminary insight into the bovine AMPK epitranscriptome, while also providing more points of data for post-transcriptional 5mC which could possibly help elucidate patterns in a relatively new sub-field of translational regulation. Finally, this study is the first ever survey of epitranscriptomic markers and their documentation conducted within livestock species.

Introduction

The character of a cell is defined by its constituent proteins, which are the result of specific patterns of gene expression. Crucial determinants of gene expression patterns are DNA-binding transcription factors that choose genes for transcriptional activation or repression by recognizing the sequence of DNA bases in their promoter regions (Bird, A. 2002). The proteins expressed in a cell dictate the function of that cell. In other words, protein expression is what differentiates one cell type from another. Ontological studies of cell types have documented over 680 different types of known cells (Bard et al. 2005). It is known that as a pluripotent cell undergoes differentiation into a specific cell type, the cell is imprinted with developmental 'history'. The genomic sequence of a differentiated cell is thought to be identical in most cases to that of the zygote from which it is descended (mammalian B and T cells being an obvious exception). This means that the marks of developmental history are unlikely to be caused by widespread somatic mutation (Bird, A. 2002). Although transcription factors are a prominent component of gene expression, there is much more regulation that governs protein expression and ultimately cell type differentiation. Processes less irrevocable than mutation fall under the umbrella term “epigenetic” mechanisms. The prefix ‘epi’ meaning above, implies that this regulation of gene expression is occuring not by perturbations within the genetic code itself but rather by alterations of the expression of the genetic code. A current definition of epigenetics is: “The study of mitotically and/or meiotically heritable changes in gene function that cannot be explained by changes in DNA sequence” (Russo et al. 1996). Epigenetic changes within organismal genomes is often in response to exogenous pressure placed on the cell from its surrounding environment as well as intrinsic signals from the cell itself (Jaenisch et al. 2003). The field of epigenetics has entered the scientific spotlight over the past 20 years as more studies have documented both somatic and germline heritable epigenetic changes, which could have important applications in fields ranging from medicine to agriculture (Bird, A. 2002).

Epigenetic changes within a cell can occur at both the nucleic acid and protein levels through a variety of mechanistic modifications. One of the most prevalent and well understood epigenetic modifications is the addition of a methyl group (-CH3) onto the fifth position of the nucleic acid cytosine, with the apt name of 5-methylcytosine (5mC) (Murphy et al 2013). Although cytosine methylation is well documented at the genetic level (DNA), little is known about it at a post-transcriptional level (RNA). Cytosine methylation within RNA has been observed in mRNA, tRNA, and rRNA however demonstrable rationality for this epitranscriptomic modification has yet to be elucidated. The research of Jaenisch, R et al. suggest that RNA methylation may be related to RNA stability, given that it has been observed in rRNA and tRNA, which are both stable and long-lived molecules. It should be noted that there has yet to be conclusive documented evidence explaining the rational for the impetus of RNA methylation. Jaenisch, R et al. goes on to explain that the dearth of knowledge regarding post-transcriptional modification of RNA can be explained by the fact 5mC is not easily detectable through most sequencing methods and that 5mC does not affect base pair binding, precluding the use of any probing assays.

Given that there has been very little documented on 5mC as a post-transcriptional modification, our study is currently the only one to survey it within livestock species. Here we survey differential RNA methylation within the AMPK gene family from skeletal muscle samples of heifer cows derived from a composite of breeds primarily consisting of Charolais. The cattle used in this study are classified based on a measurement of feed efficiency known as Residual Feed Intake (RFI), a trait often defined as the difference between an animal's observed and expected feed intake in relation to the animal's body weight and growth rate during a specified feeding period (Seabury et al. 2017). The expected feed intake of animals is based on body weight and gain, where animals with higher body weights and higher levels of gain are assigned a higher value for expected feed intake, relative to an animal that is lighter or growing slower. For example, if an analyst at a breeding association decided to analyze the feed efficiency through an analysis of the RFIs of two hypothetical steers, steer X and steer Y, both weighing approximately the same at the beginning of a two week long feed testing period and each steer having approximately the same average daily gain (the rate of weight gain per day over a specified period of time), they would each have approximately the same expected feed intake values (20 pounds per day). At the end of the two week testing period, if steer X consumes less food than steer Y, the RFI of steer X will be a smaller value than the RFI of steer Y, since the observed feed intake value for steer X (18 pounds per day) will be smaller than that of steer Y (23 pounds per day). Knowing the expected and observed feed intakes for each steer, the analyst can subtract the expected from the observed, generating RFI values of -2 and 3 for steer X and steer Y, respectively. Steer X would therefore have more optimal feed efficiency with respect to steer Y. A negative value for RFI is considered desirable in a livestock animal because it indicates the animal consumed less feed than was predicted based on its weight and gain.

Measurements of feed efficiency are of great importance to cattle breeders as it allows them to identify the animals that are most efficient at converting feed into profit. Feed constitutes a majority of the cost when it comes to raising beef cattle, making the feed efficiency of an animal is economically beneficial to the farmer. Traditional measures of feed efficiency, such as the ratio of feed consumed to observed body weight gain (i.e., feed conversion ratio), are likely to be very successful in selecting for increased growth rates in beef cattle (i.e., increased producer income) but selection for enhanced weaning or yearling weights may also lead to increased costs associated with maintaining larger mature cows (i.e., increased nutrient requirements and feed costs; increased calving difficulty due to larger birth weights, etc) (Archer et al. 1999; Strohbehn et al. 1993; Deutscher et al. 1995). If traditional measures of feed efficiency, such as an animal's feed conversion ratio, were to be used in a selective breeding program, farmers would be indirectly selecting for an animal's increased genetic merit in gain, not necessarily feed efficiency. This is because the cows selected through this measurement, albeit physically larger, would require more feed intake, not necessarily making them more feed efficient, resulting in little economic gain for the breeders.

The issues associated with using traditional feed efficiency measurements to help guide selective breeding has led to many farmers turning to RFI as an alternative way to measure and estimate feed efficiency (Seabury, C.M et al. 2017). The recent gain in popularity of RFI as a feed intake measurement is due to its main advantage over other feed intake measurements; that being its phenotypic independence from daily gain as well as the traits used to calculate RFI, thereby making RFI a preferred measure for dissecting the underlying biology related to feed efficiency and for enabling genomic selection (Saatchi et al. 2014; Kennedy et al. 1993; Do et al. 2014). The biological forces driving RFI differences within phenotypically indistinguishable cattle has yet to be elucidated and is of great importance within livestock agriculture because of the potential economic impact it could have for farmers. In an attempt to better understand RFI as a complex trait, previous researchers have conducted Genome Wide Association Studies (GWAS) on cattle with differing RFI but failed to find conclusive evidence of either single nucleotide polymorphisms (SNPs) or quantitative trait loci (QTL) to serve as potential genetic markers driving observed differences in RFI, concluding that higher density analyses are needed (Seabury, C.M et al. 2017).

One of the central regulators of cellular and organismal metabolism in eukaryotes is the AMP-activated protein kinase (AMPK), which is activated when intracellular ATP levels lower. AMPK plays critical roles in regulating growth and reprogramming metabolism, and recently has been connected to cellular processes including autophagy and cell polarity (Mihaylova 2012). AMPK was chosen for analysis due to its large role in cellular processes as well as its role as an epigenetic regulator. A paper by Gongol et al. 2018 states AMPK phosphorylation of histones, DNA methyltransferases and histone post-translational modifiers establish AMPK as a key player in epigenetic regulation. It would make sense that given this protein’s role in the cell that it would be regulated at multiple levels, potentially even at the post-translational level. Given that RFI is related to an animal’s ability to metabolize food, it would be logical to assume that the activity of the AMPK gene may differ among animals with higher and lower RFIs.

Given that so little is understood about 5mC in RNA, we believe that this could be a potential area of focus for the elucidation of the biological forces driving differences in RFI within domestic cattle populations. We therefore propose that differences in RFI may be driven through post-transcriptional modifications of mRNA, specifically differential 5mC within the mRNA transcript. This is proposed because of documented speculation of the impact methylation has on mRNA stability. Along with the proposed stability of transcripts, differential methylation of mRNA may also be theoretically associated with the expression level of the protein it encodes.

In this study, we hypothesize that differential epigenetic changes within the mRNA transcript for exon 7 of the bovine *PRKAB1* gene, a gene encoding a subunit of the protein AMPK, may be associated with differing RFI values in domestic cattle. This hypothesis is tested through an RNA Combined Bisulfite Restriction Analysis (COBRA), a process which involves the conversion of unmethylated cytosines of the RNA into uracils, followed by cDNA generation and subsequent PCR amplification, ending with a digestion of the PCR product, which will reveal any sites of potential differential methylation. This technique is routinely performed with DNA; however, we have modified it for the use of RNA and present results of the modified technique here. Within the AMPK gene family, focus will be on the mRNA transcript of exon 7 of the *PRKAB1* gene, which encodes part of the beta-subunit of the protein. This portion of the beta-subunit encodes for a regulatory subunit of the AMPK protein as well as the αγ-binding domain, which is integral to the structural composition of the AMPK protein. The analyses performed in this study serve as a novel exploration into a relatively new and not well documented sector of epigenetics and potentially protein regulation. This is the first ever study done analyzing post-transcriptional epigenetic modification within commercial livestock. Our results serve as a starting point for future studies of post-transcriptional modification and the potential role it has in economically desirable traits within livestock species.

Methods and Materials

Tissue Sample Collection and Phenotypes

Skeletal muscle tissues were harvested from phenotyped cattle at time of slaughter by collaborators (University of Illinois at Urbana-Champaign) using established protocols that have produced high quality RNA yields in previous studies (Saatchi et al. 2014). Immediately after harvest, samples were flash frozen in liquid nitrogen, ensuring the preservation and integrity of cellular RNA. Frozen samples were then shipped overnight to UVM on dry ice, received at Terrill Hall and immediately stored at -80°C until needed for future use.

These cattle represent a composite breed which is primarily derived from the Charolais breed. Residual feed intake phenotypes were measured by including a partial linear regression on average daily gain (ADG) and mid-test metabolic weight (MMWT) as previously described (Seabury et al. 2017; Saatchi et al. 2014).

RNA Extraction and Bisulfite Conversion

RNA was extracted from the skeletal muscle of six total samples, three samples having low RFIs and the other three having high RFIs, using a QIAGEN® RNeasy® Kit

(Qiagen, Hilden, Germany). Prior to extraction, muscle samples were stored at -80°C to ensure native RNA integrity. Five milligrams per tissue were used per extraction and RNA was isolated using a QIAGEN® TissueLyser II system according to protocol (Qiagen, Hilden, Germany). When not in use, RNA was kept at -80°C at all times to ameliorate any potential of RNA degradation. The concentration and purity of each extracted RNA sample was obtained using a NanoDrop spectrophotometer (Thermo Fisher Scientific, Wilmington, DE) and the quality of the samples were checked further by running a gel electrophoresis of 250ng of extracted RNA on a 1% agarose gel at 80 volts for 1 hour using ethidium bromide gel dye and bromophenol blue loading dye (**Figure 1**).

In preparation of differential methylation analysis, unmethylated cytosines within the extracted RNA samples are converted into uracils with a bisulfite treatment. This was done using a Zymo EZ RNA Methylation kit (Zymo Research, Irvine, CA, USA). RNA conversion reagent was added to each sample for a total reaction volume of 150μl containing 750ng per reaction. Conversions were carried out at 70°C for 5 minutes followed by 54°C for 45 minutes. Each sample was incubated at room temperature for 30 minutes to promote desulfination. Finally, bisulfite treated RNA was eluted in a final volume of 20μl and stored at -80°C.

cDNA generation

In preparation of PCR amplification, cDNA was synthesized from the previously bisulfite treated RNA according to Promega's GoScript™ Reverse Transcription System protocol (Promega Corporation, Madison, WI, USA). 5μl of bisulfite treated RNA from each sample was used for each reverse transcription reaction, creating a total reaction volume of 20μl. Samples were incubated at 25°C for 5 minutes followed by a second incubation at 47°C to promote first strand synthesis and finally an incubation at 75°C for 15 minutes to inactivate the reverse transcriptase.

Primer Design

Existing primer design programs developed for standard PCR cannot handle primer design for bisulfite-conversion-based PCRs due to changes in DNA sequence context caused by bisulfite treatment and many special constraints both on the primers and the region to be amplified for such experiments (Li et al. 2002). To circumvent this issue, MethPrimer (http://www.urogene.org/methprimer/) was used to design primers for methylation specific PCR. The primer sequences are as follows:

Forward Primer: TTGATTTAGGAAGAAAGGGGTT

Reverse Primer: CCCTCCAAATAAAAACTTCAAA

MethPrimer's CpG island prediction algorithm was programed to find islands greater than 100bp in size with GC percent greater than 50. These primers were determined to have a Tm of 49°C and after amplification, create a PCR product of 296bp in length.

PCR

Polymerase Chain Reaction was performed using Zymo ZymoTaq PreMix (Zymo Research, Irvine, CA, USA) according to protocol with synthesized 5μL cDNA product and 0.5μM methylation specific primers. The conditions for PCR amplification are as follows: 95°C for 10 minutes to allow full denaturation, followed by 40 cycles of 95°C for 30 seconds to promote denaturation of the sample cDNA, 49°C for 30 seconds to allow primer annealing, and 72°C for 30 seconds to allow for elongation; ending with incubation at 72°C for 7 minutes before being held at 4°C indefinitely. PCR product of this reaction was then used as the template for an additional round of PCR, with the same parameters, due to the low amplicon concentration from the first reaction. The product of the second PCR reaction is then visualized on a 1.5% agarose gel with ethidium bromide.

Digest

A restriction enzyme digest was performed on the secondary PCR product to determine the presence of methylated cytosines. This was done using the restriction enzyme HpyCH4IV, which was obtained from New England BioLabs (New England BioLabs, Ipswich, MA, USA). HpyCH4IV was chosen because its cutsite, A*CGT, appears twice within the sequence of exon 7 of the *PRKAB1* gene (**Figure 2**). After bisulfite conversion of the mRNA, cDNA generation, and PCR amplification, HpyCH4IV will only cut at sites that contained cytosine methylation, as this would be the only way to ensure the sequence integrity of its cutsite. Each digestion consisted of 10µl of the PCR amplicon, 2µl of HpyCH4IV, 2.5µl of NEB's cut smart buffer, Q.S.'d to a total reaction volume of 25µl. The digestion reaction mix was incubated for 4 hours at 37°C to promote enzymatic cleavage of DNA followed by a 20 minutes at 65°C to thermally inactivate the enzyme.

The product of the restriction digest was visualized by gel electrophoresis, allowing the products to run alongside a 100bp ladder (Thermo Fisher Scientific, Wilmington, DE, USA) for 3 hours at 80 volts on a 3% agarose gel using SYBR SAFE gel dye and bromophenol blue loading dye.

Analysis of Methylation

Samples were then analyzed for differential methylation based on banding patterns observed. If samples exhibited the expected enzyme digest banding pattern, the samples are then classified as fully methylated. If the PCR product was present and no enzyme digest banding pattern was observed, the sample was classified as unmethylated. Finally, samples that exhibited the digested banding pattern and the undigested PCR product were classified as partially methylated.

Further analysis was performed by using Bio-Rad's Image Lab software (Bio-Rad Laboratories, Hercules, California, USA) to generate band percent values for each sample digest. Band percent values are calculated by using the band's volume, which is the sum of all of the intensities within the band boundaries (boundaries can be either defined by the software or the user). The value of a band's volume is compared to that of all the other band volume values in a given lane and then band percent can be calculated.

Statistical Analysis

Statistical analyses of RFI values and band percentages among samples were performed using XLMiner Analysis ToolPak and HeatMapper. A two sample unpaired *t* test was employed to compare the mean values of high and low RFI samples. A colorimetrically scaled heatmap of band percentage data and hierarchical clustering using a Euclidean distance measurement paired with average linkage of samples was performed using HeatMapper (http://www2.heatmapper.ca/expression/). Furthermore, two sample unpaired *t* tests were used to analyze band percentage for each produced band among high and low RFI samples. A significance of $p<0.05$ was used *t* test analyses using the google sheets Add-on XLMiner Analysis ToolPak (Frontline Systems, Incline Village, Nevada, USA.).

.

Results

Statistical significance of the mean RFI values between the low and high groups was first conducted, establishing that the groupings of tissue samples were appropriate for downstream analysis and hypothesis formation (**Table 1**). A two sample unpaired two-tailed *t* test was employed to analyze the difference in means. By conventional criteria, the RFI groups were considered to have statistical significance between their mean values, having a two-tailed p-value of less than 0.0001 and a standard error of difference of 0.326.

Digestion of PCR amplified product was observed at various levels across all samples tested, revealing and documenting methylation of ribonucleic cytosine within exon 7 of the bovine *PRKAB1* gene (**Figure 3**). All samples contained six detectable bands in their lanes, except for sample UI19, which was not observed to have the 162bp band, indicating that there may be differential methylation within exon 7. No complete digestion of the PCR product was observed among samples, resulting in all analyzed samples being classified as partially methylated. Upon gel visualization, it is apparent that band intensities from the digested PCR products varied among samples. Bio-Rad's Image Lab software was used to generate percentage values for the bands of each digested sample. Band values for each sample were then entered into a matrix and a heatmap was generated and colorimetrically scaled vertically based on the band percentage values for each digest product (**Figure 4**). Analysis of band percentage for the PCR product was used to determine the percent unmethylation for each sample.

Percentage of unmethylation among samples ranged between 14.7% in the low RFI sample UI19 and 79.7% in the low RFI sample UI23. Percent values of digest product bands range from 1.2% in UI21 for the 64bp product to 49.2% in UI19 for the 134bp band (**Figure 4**). Hierarchical clustering of band percentages within each sample did not result in separate clusters for high and low RFI groups.

As band intensity is directly proportional to percent methylation, band intensities were then analyzed using a tailed two sample *t* test to determine if statistical significance exists between the mean differential methylation values of band size between high and low RFI groups. This analysis was performed for each PCR band and only the 162bp was found to have statistically different methylation between high and low RFI groups with a p-value of 0.02411. This finding indicates the presence of differential RNA methylation within exon 7 of the bovine *PRKAB1* gene between groups of animals with extreme measures of RFI.

Discussion

The results of our study contain an initial recording of RNA methylation within livestock species. This epitranscriptomic modification was identified using a novel experimental method for RNA methylation detection and validation (**Figure 5**), which when optimized, can be used as an alternative to the current more intensive detection methods such as the ones described in (Helm et al. 2017), obviating the need to use such methods.

It should first be addressed that one of the high RFI samples, UI16, failed during the bisulfite conversion step. PCR product was not visible via gel electrophoresis during pre-COBRA product analysis. Subsequent gel electrophoresis of 4μL UI16 bisulfite treated RNA product was performed to determine if there was an RNA product after the bisulfite treatment step and UI19 bisulfite treated RNA was used as a control for visualization. Upon gel image analysis, nothing was able to be visualized in the UI16 lane, suggesting that the bisulfite treated product of UI16 had been lost. Analysis of the remaining five samples continued.

Although RNA methylation was observed in all samples, differential RNA methylation between RFI groups was only made apparent after statistical analysis of mean band percentage values among the high and low groups; supporting our hypothesis that differential RNA methylation of AMPK mRNA would be present between groups of animals with extreme measures of RFI. The overall banding pattern remained the same, differing only in intensity, across four of the five samples analyzed. Low RFI sample UI19 showed no clear band for the predicted digestion product of 162bp. **Figure 2** shows a map of the PCR amplified region of *PRKAB1* exon 7 with *HpyCH4IV* cut sites and the six potential band sizes that we expected to see upon COBRA gel visualization. The lack of a 162bp band in UI19 would indicate that there is nearly full methylation of cytosine at the first enzyme cut site within the amplicon. This is further supported through the band percent values of UI19's 232bp, 134bp, 98bp and 64bp bands. In order for the 162bp band to not appear, the PCR product must be cut at both the first and second cutsite, rendering full digestion, or cut at only the first, yielding a 64bp band

and a 232bp band. U19 has the second lowest band percentage, 1.9%, for the 232 fragment and the highest band percentage values for the 134bp, 98bp and 64bp bands, indicating that it has the highest level of 5-methylcytosine within RNA of all samples.

Methylation of cytosine within mRNA is not a frequent occurrence within eukaryotic organisms and its functionality has yet to be fully elucidated (Schaefer et al. 2009). As mentioned previously, the AMPK protein plays many roles within the cell including actively regulating growth and metabolism; making the need to regulate its function of high biological importance for proper cellular function (Gongol et al. 2018). Because of this, we suspect AMPK and its regulation to be an important biological factor partially responsible for the range of RFI farmers observe in their livestock. RFI is an economically important phenotype that can be selected for through livestock breeding programs; however, it is not known what is biochemically responsible for the observed differences in RFI between animals that would normally be phenotypically similar. Our study suggests that differential RNA methylation of the exon 7 of the bovine *PRKAB1* gene could be associated with observed RFI values. The *PRKAB1* gene encodes the beta-subunit of the AMPK protein which helps facilitate the protein's tertiary structural integrity as well as serving as a site for regulation. Our study serves as a starting point, upon which future data can be generated to one elucidate AMPK's true role in the phenotypic selection of RFI in livestock species.

This study uses a new methodology, which is based on the pre-existing protocol of DNA methylation detection, to analyze 5-methylcytosine within RNA (Xiong and Laird, 1997) . Given that the protocol used here is still in its academic infancy, there are certain points of weakness that should be acknowledged. First, this study did not have a positive control, which would have ideally been used during the bisulfite conversion, gel electrophoresis and image analysis steps. Currently, SssI treatment is used as the universal standard for positive controls in DNA methylation studies. Through SssI treatment, a sample is artificially methylated at every cytosine that is part of a CpG dinucleotide in the treated sequence. This artificially methylated sample can be run parallel with samples of interest to monitor bisulfite conversion and restriction digest. The efficacy of SssI treatment of RNA has not been documented and we did not attempt to treat our own samples with SssI for this study; however, this will be attempted in future studies conducted by our lab. Ideally, there will one day be a commercially available option for a reagent that induces the artificial methylation of cytosine within RNA to serve as a positive control for future studies. Furthermore, data interpretation through images generated by COBRA are error prone. Although Bio-Rad's Image Lab software could not pick up on a 162bp band in UI19's lane, that does not mean there is an absence of digestion product at this position. Hypothetically, this band could have been at such a low concentration that it was not able to reveal itself on the gel image. This problem will be alleviated through bisulfite pyrosequencing, a process which will be elaborated on later in this discussion. Finally, the sample size of this experiment is relatively small compared to conventional scientific studies; however, smaller sample sizes are common in epigenetic studies as seen in (Sevane et al. 2019; Ziller et al. 2014). The failure of sample UI16 and subsequent shrinkage of sample size only weakens the integrity of our results. Statistical analysis with such small experimental group sizes leads to an increase in type II statistical errors, which may make interpreting results more difficult. Ideally, increasing the sample size of our future studies will reduce the potential of type II errors as well as increase the amount of data we generate which may help elucidate the questions that drive our research.

Future directions of this project include the incorporation of bisulfite pyrosequencing into our analysis of methylation. Bisulfite pyrosequencing is a sequencing-by-synthesis method used to quantitatively determine the methylation of individual CG cytosines from PCR amplicons of a region up to 115 bases in length (Bassil et al. 2013). Using this sequencing method in our analysis, we can acquire an accurate quantitative representation of an amplicon's methylation profile, giving the actual percent methylation for each individual cytosine, obviating the need for restriction digest and band percentage analysis. Pyrosequencing involves the creation of specific primers which is more time consuming than that of normal PCR or methylation specific primers and it is also more expensive than traditional COBRA studies. Being that 5-methylcytosine is rarely observed in mRNA, COBRAs will continually be a cheap and accurate method to validate or detect post-transcriptional RNA before investing in pyrosequencing. In conclusion, based on the results of this study, we plan on moving forward with pyrosequencing to gain further insight into the patterns of ribonucleic methylation of cytosine within exon 7 of the bovine *PRKAB1* gene and the potential role these patterns have within the RFI of livestock species.

Acknowledgements

I would like to first and foremost acknowledge my research mentor, Dr. Stephanie McKay, who, through her guidance and support, has truly turned me into both the scientist and person I am today. Joining Stephanie's lab my freshman year has been undoubtedly the best choice I have made as an undergraduate at the University of Vermont. Her passion for science is truly inspiring, and it has been an honor to be a part of her lab. That being said, I would like to thank all of the members, both past and present, of the McKay lab for the support throughout my academic career, specifically Suraj Bhattarai, whos help and guidance was monumental to the work in this thesis. I would consider myself very lucky to have been part of such an amazing lab.

Outside of Dr. McKay's lab, I cannot begin to express how thankful I am for my parents and the constant love and support they have provided me throughout my time at UVM. Finally, I would like to acknowledge the music of MF DOOM and Parquet Courts for providing impetus in times of apprehension and doubt.

REFERENCES

Archer JA, Richardson EC, Herd RM, Arthur PF. (1999) Potential for selection to improve efficiency of feed use in beef cattle: A review. Austr J Agric Res.;50:147–62.

Babicki, S., Arndt, D., Marcu, A., Liang, Y., Grant, J.R., Maciejewski, A., and Wishart, D.S. (2016). Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Research* 44, W147-W153.

Bard, J., Rhee, S.Y., and Ashburner, M. (2005). An ontology for cell types. *Genome Biology* 6, R21.

Bassil, C.F., Huang, Z., and Murphy, S.K. (2013). "Bisulfite Pyrosequencing," in *Ovarian Cancer: Methods and Protocols,* eds. A. Malek & O. Tchernitsa. (Totowa, NJ: Humana Press), 95-107.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes & Development* 16, 6-21.

Burmeister, D.W., Smith, E.H., Cristel, R.T., Mckay, S.D., Shi, H., Arthur, G.L., Davis, J.W., and Taylor, K.H. (2017). The expression of RUNDC3B is associated with promoter methylation in lymphoid malignancies. *Hematological Oncology* 35, 25-33.

Deutscher GH. Reducing calving difficulty by heifer and sire selection and management. Proc. of the Range Beef Cow Symp. XIV, December 5-7, Gering: University of Nebraska-Lincoln; 1995.

Do DN, Ostersen T, Strathe AB, Mark T, Jensen J, Kadarmideen HN. Genome-wide association and systems genetic analyses of residual feed intake, daily feed consumption, backfat and weight gain in pigs. BMC Genet. 2014;15:27.

Gongol, B., Sari, I., Bryant, T., Rosete, G., and Marin, T. (2018). AMPK: An Epigenetic Landscape Modulator. *International journal of molecular sciences* 19, 3238.

Helm, M., and Motorin, Y. (2017). Detecting RNA modifications in the epitranscriptome: predict and validate. *Nature Reviews Genetics* 18, 275-291.

Jaenisch, R., and Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nature Genetics* 33, 245-254.

Kennedy BW, van der Werf JH, Meuwissen TH. Genetic and statistical properties of residual feed intake. J Anim Sci. 1993;71:3239–50.

Li, L.-C., and Dahiya, R. (2002). MethPrimer: designing primers for methylation PCRs. *Bioinformatics* 18, 1427-1431.

Mihaylova, M.M., and Shaw, R.J. (2011). The AMPK signalling pathway coordinates cell growth, autophagy and metabolism. *Nature cell biology* 13, 1016-1023.

Murphy, T.M., Mullins, N., Ryan, M., Foster, T., Kelly, C., Mcclelland, R., O'grady, J., Corcoran, E., Brady, J., Reilly, M., Jeffers, A., Brown, K., Maher, A., Bannan, N., Casement, A., Lynch, D., Bolger, S., Buckley, A., Quinlivan, L., Daly, L., Kelleher, C., and Malone, K.M. (2013). Genetic variation in DNMT3B and increased global DNA methylation is associated with suicide attempts in psychiatric patients. *Genes, Brain and Behavior* 12, 125-132.

Richards, E.J. (2006). Inherited epigenetic variation — revisiting soft inheritance. *Nature Reviews Genetics* 7, 395-401.

Russo, V.E., Martienssen, R.A., and Riggs, A.D. (1996). *Epigenetic mechanisms of gene regulation.* Cold Spring Harbor Laboratory Press.

Saatchi, M., Beever, J.E., Decker, J.E., Faulkner, D.B., Freetly, H.C., Hansen, S.L., Yampara-Iquise, H., Johnson, K.A., Kachman, S.D., Kerley, M.S., Kim, J., Loy, D.D., Marques, E., Neibergs, H.L., Pollak, E.J., Schnabel, R.D., Seabury, C.M., Shike, D.W., Snelling, W.M., Spangler, M.L., Weaber, R.L., Garrick, D.J., and Taylor, J.F. (2014). QTLs associated with dry matter intake, metabolic mid-test weight, growth and feed efficiency have little overlap across 4 beef cattle studies. *BMC Genomics* 15, 1004.

Schaefer, M., Pollex, T., Hanna, K., and Lyko, F. (2009). RNA cytosine methylation analysis by bisulfite sequencing. *Nucleic acids research* 37, e12-e12.

Seabury, C.M., Oldeschulte, D.L., Saatchi, M., Beever, J.E., Decker, J.E., Halley, Y.A., Bhattarai, E.K., Molaei, M., Freetly, H.C., Hansen, S.L., Yampara-Iquise, H., Johnson, K.A., Kerley, M.S., Kim, J., Loy, D.D., Marques, E., Neibergs, H.L., Schnabel, R.D., Shike, D.W., Spangler, M.L., Weaber, R.L., Garrick, D.J., and Taylor, J.F. (2017). Genome-wide association study for feed efficiency and growth traits in U.S. beef cattle. *BMC Genomics* 18, 386.

Sevane, N., Martínez, R., and Bruford, M.W. (2019). Genome-wide differential DNA methylation in tropically adapted Creole cattle and their Iberian ancestors. *Animal Genetics* 50, 15-26.

Strohbehn D, Bryce S, Maxwell D, Anderson L, Wilson D. Influence of Angus sire birth weight EPD on performance of crossbred first-calf heifers: A progress report. Iowa State Beef Res. Rep. 1993;54-60.

Xiong, Z., and Laird, P.W. (1997). COBRA: a sensitive and quantitative DNA methylation assay. *Nucleic acids research* 25, 2532-2534.

Ziller, M.J., Hansen, K.D., Meissner, A., and Aryee, M.J. (2014). Coverage recommendations for methylation analysis by whole-genome bisulfite sequencing. *Nature Methods* 12, 230.

Tables/Figures

This table (Table 1) shows the sample name and its associated RFI value. RFI values that are positive indicate an undesirable feed efficiency, while RFI values that are negative indicate a desirable one. These numbers were used for an unpaired two sample *t* test to determine the statistical significance means between low and high RFI groups. The difference in means for the two groups had a two-tailed p-value of less than 0.0001, indicating that by conventional criteria, they are statistically different. This data was provided by collaborators at the University of Illinois at Urbana-Champaign.

Sample Name	**RFI**
UI15	**2.75505**
UI16	**2.59651**
UI19	**-3.01356**
UI21	**3.19062**
UI23	**-2.21583**
UI24	**-2.17566**

Table 1: Sample Name and its associated RFI value.

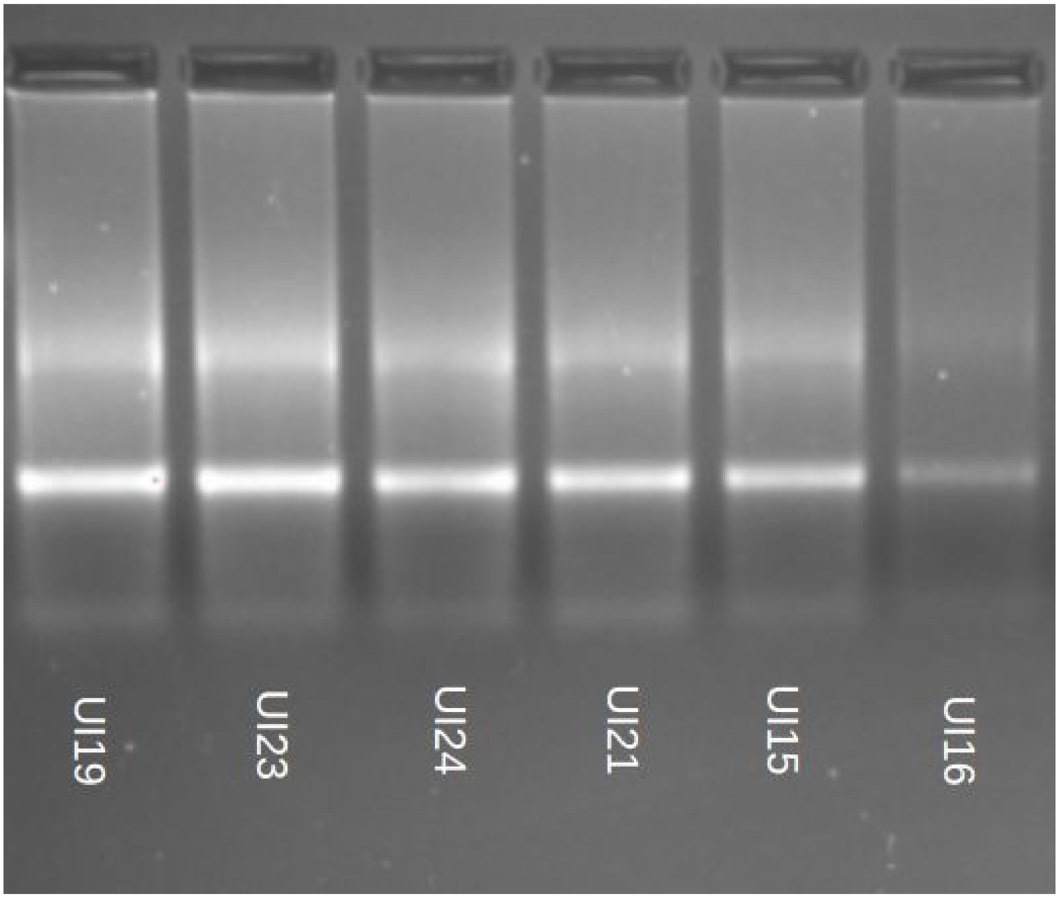

Figure 1: Visualization of RNA Extracts

This figure (Figure 1) displays a gel image that was used as a second form of verification alongside analysis by NanoDrop spectrophotometer (Thermo Fisher Scientific, Wilmington, DE) of successful RNA extraction. This gel was made with agarose and 1X TAE to have a final concentration of 1% agarose. Electrophoresis was conducted at 80 volts for 1 hour using ethidium bromide gel dye and bromophenol blue loading dye. 250ng of extracted RNA was loaded into each well.

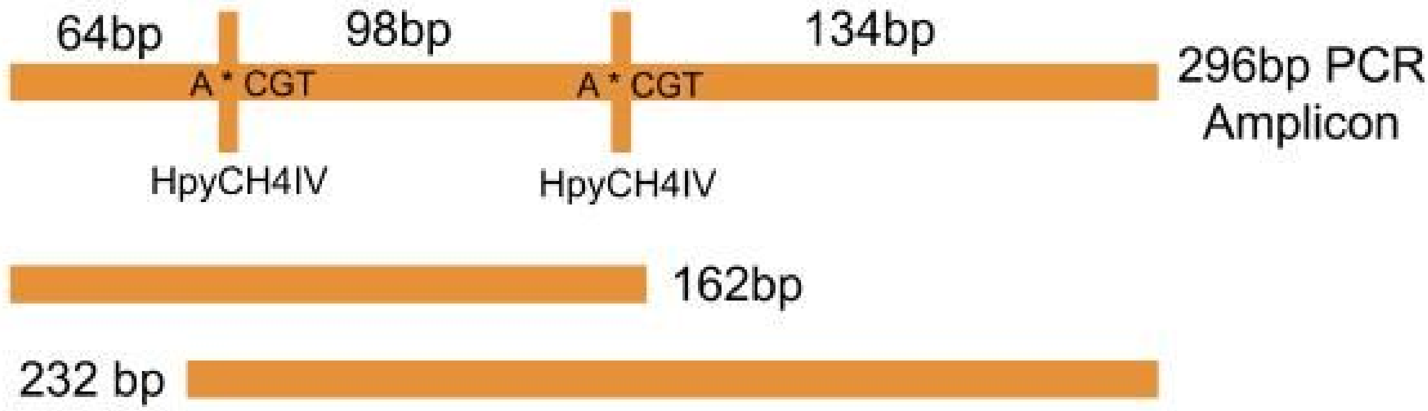

Figure 2: Map of *PRKAB1* Exon 7 Amplicon With *HpyCH4IV* Cut Sites

This figure (Figure 2) displays the PCR amplicon of *PRKAB1* exon 7 with the sizes of predicted digestion products. There are two potential *HpyCH4IV* cut sites within the generated *PRKAB1* exon 7 PCR product. The six potential band sizes are as follows: 296bp (no methylation), 232bp, 162bp, 134bp, 98bp and 64bp. The presence of only 134bp, 98bp, and 64bp bands would mean complete digestion of the PCR product occurred, indicating that all RNA cytosines at these two sites were methylated. After bisulfite treatment of RNA, cut site sequence integrity will only be retained if the cytosine within the RNA (which codes for the guanine in the PCR amplified cDNA) the site is methylated.

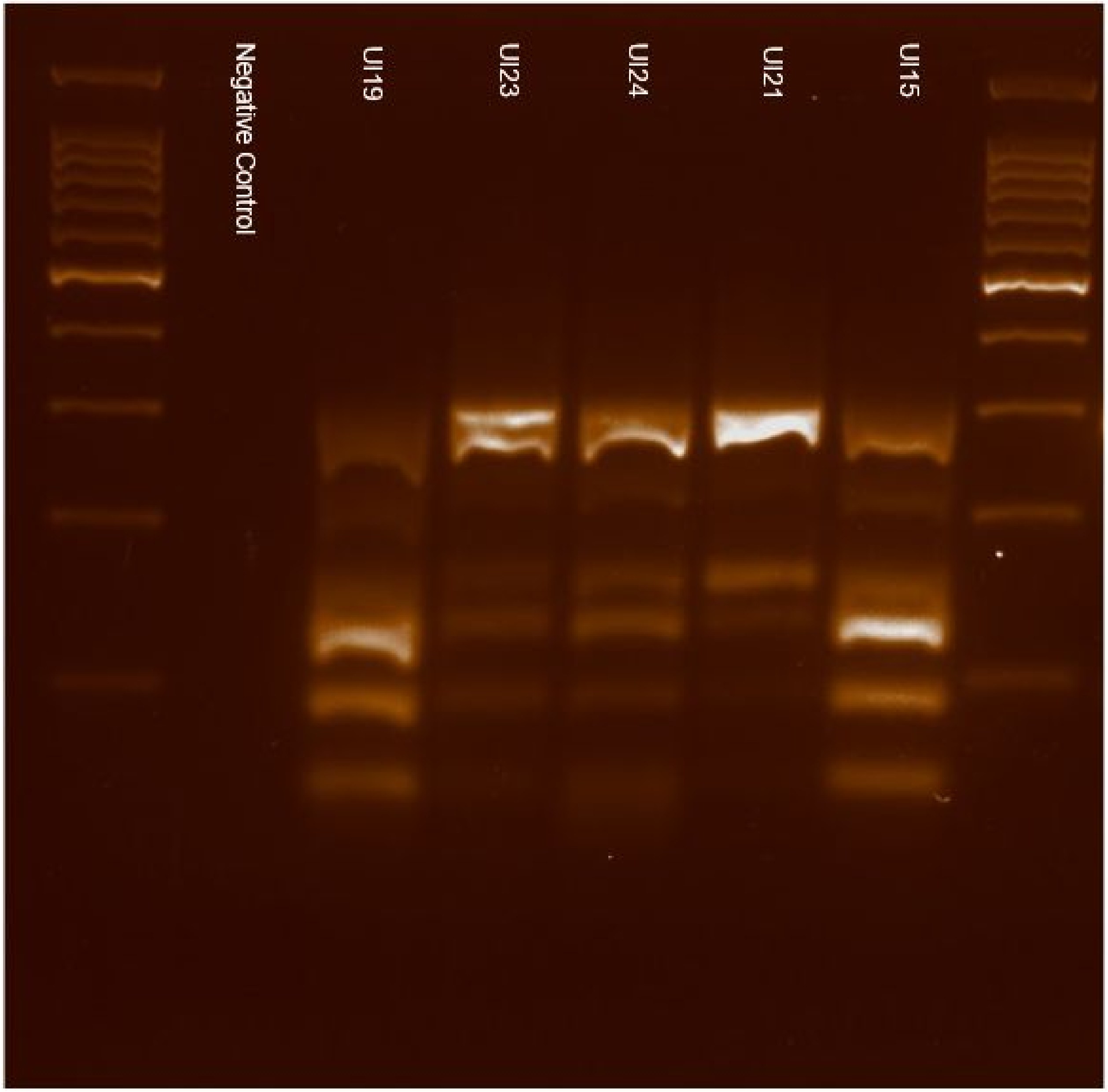

Figure 3: Gel Image displaying Combined Bisulfite Restriction Analysis (COBRA) of *PRKAB1* Exon 7 Amplicon With *HpyCH4IV*.

This figure (Figure 3) displays a 3% agarose gel which was run for 3 hours at 80 volts using SYBR SAFE gel dye and bromophenol blue loading dye to visualize the restriction digestion of sample *PRKAB1* Exon 7 Amplicon. Flanking each side of the gel is a 100bp SYBR SAFE DNA ladder for band size reference.

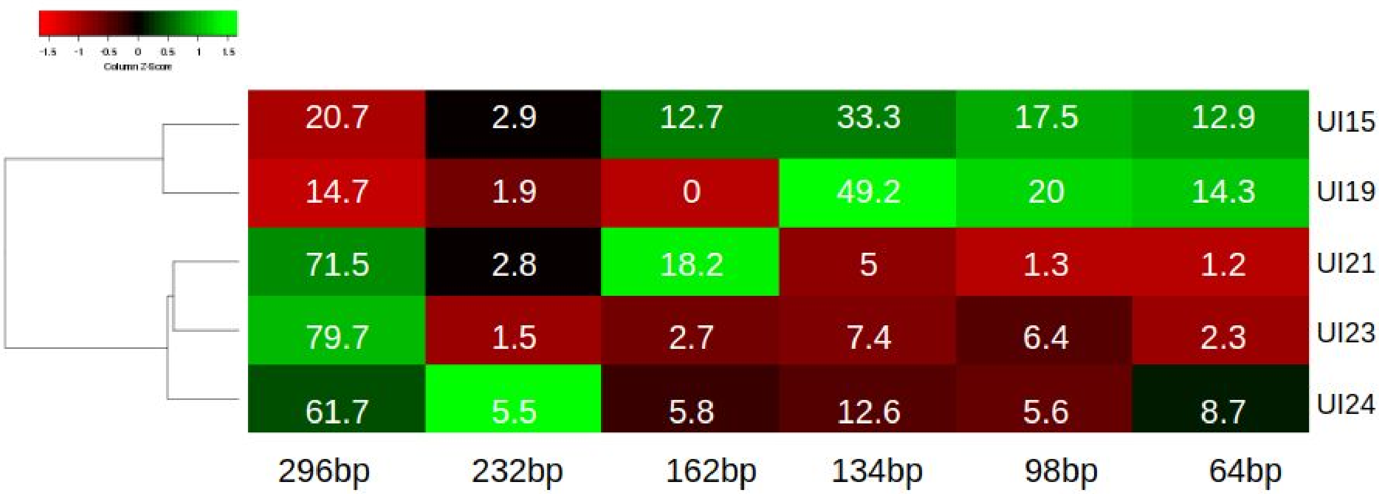

Figure 4: Hierarchically Clustered Heatmap of Digest Band Percentages and Samples.

This figure (Figure 4) displays a heatmap of band percentage data, which was quantified from the digest gel using Bio-Rad's Lab Image software. This heatmap is coloriometically scaled by column based on the band percentage of observed within each sample for the respective band length in each column. Hierarchical clustering was performed using average linkage between sample banding percent patterns using Euclidean distance measurement.

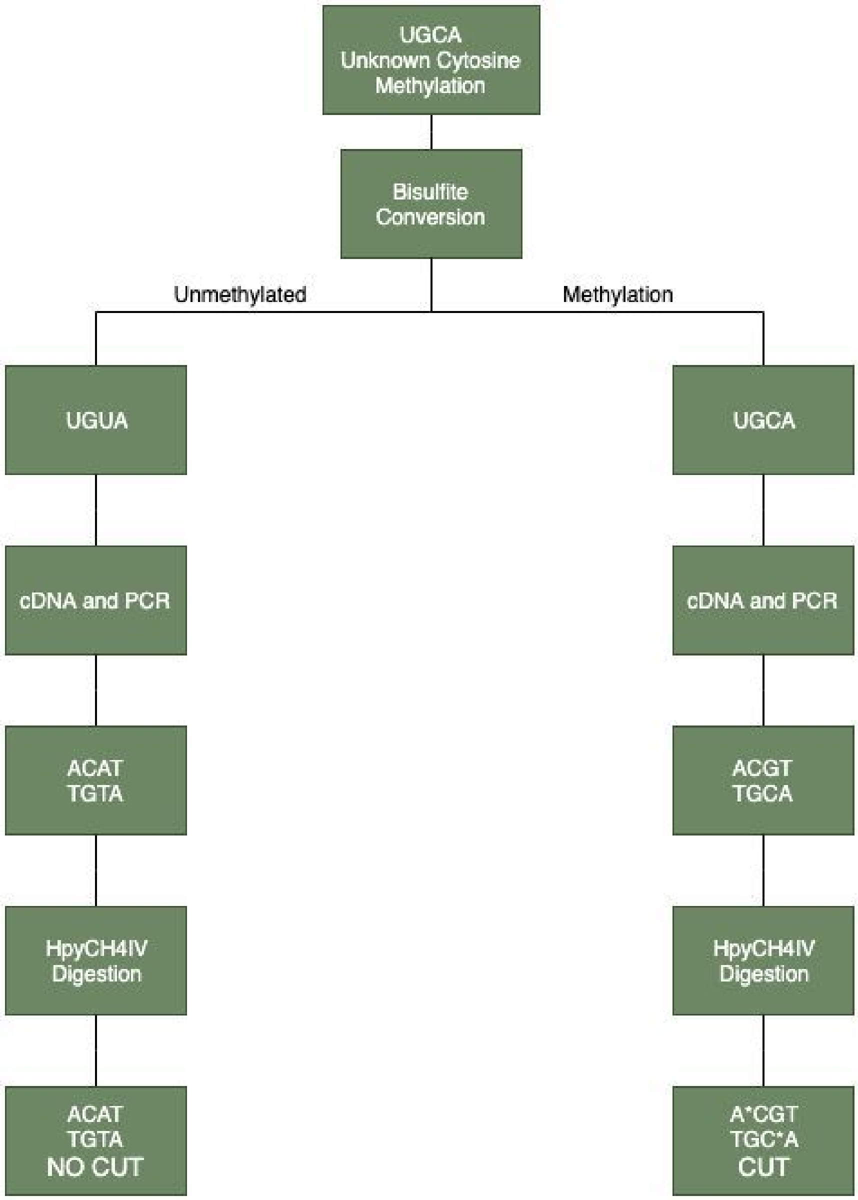
UGCA
Unknown Cytosine
Methylation
Bisulfite
Conversion
Unmethylated
Methylation
UGUA
UGCA
cDNA and PCR
cDNA and PCR
ACAT
TGTA
ACGT
TGCA
HpyCH4IV
Digestion
HpyCH4IV
Digestion
ACAT
TGTA
NO CUT
A*CGT
TGC*A
CUT

Figure 5: Flow Chart of RNA COBRA Method

This figure (Figure 5) shows a visual representation of the RNA COBRA method, a novel method for RNA methylation detection and analysis, used in this study to analyze post-transcriptional 5-methylcytosine. The bottom two squares depict the *HpyCH4IV* cut when methylation is or is not in present in RNA and how it affects the enzyme's ability to cleave PCR amplified cDNA, leading to the difference in band percentages.