New Zealand Society of Animal Production online archive

This paper is from the New Zealand Society for Animal Production online archive. NZSAP holds a regular annual conference in June or July each year for the presentation of technical and applied topics in animal production. NZSAP plays an important role as a forum fostering research in all areas of animal production including production systems, nutrition, meat science, animal welfare, wool science, animal breeding and genetics.

An invitation is extended to all those involved in the field of animal production to apply for membership of the New Zealand Society of Animal Production at our website www.nzsap.org.nz

View All Proceedings

Next Conference

Join NZSAP

The New Zealand Society of Animal Production in publishing the conference proceedings is engaged in disseminating information, not rendering professional advice or services. The views expressed herein do not necessarily represent the views of the New Zealand Society of Animal Production and the New Zealand Society of Animal Production expressly disclaims any form of liability with respect to anything done or omitted to be done in reliance upon the contents of these proceedings.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

You are free to:

Share— copy and redistribute the material in any medium or format

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

NonCommercial — You may not use the material for commercial purposes.

NoDerivatives — If you remix, transform, or build upon the material, you may not distribute the modified material.

http://creativecommons.org.nz/licences/licences-explained/

BRIEF COMMUNICATION: Genome wide DNA methylation patterns and transcription analysis in sheep muscle

C Couldrey[a]*, R Brauning[b], PH Maclean[a], HV Henderson[a] and JC McEwan[b]

[a]AgResearch Ruakura, Private Bag 3123, Hamilton 3240, New Zealand; [b]AgResearch Invermay, Private Bag 50034, Mosgiel 9053, New Zealand
**Corresponding author. Email: christine.couldrey@agresearch.co.nz*

Keywords: epigenetics; DNA methylation; sheep; Carwell; muscular hypertrophy; epigenomic pipeline

Introduction

The Carwell phenotype is a desirable inherited muscular hypertrophy of the *Longissimus dorsi* muscle in Poll Dorset sheep that does not have any detrimental effect on eating quality, fat deposition or detectable meat tenderness. The Carwell haplotype has been localized to a 350kb region on the distal end of OAR18 near the Callipyge locus (Nicoll et al. 1998) containing three known genes. However, in spite of considerable effort in sequencing and resequencing, the causative mutation has not been identified. Although unsuccessful in identifying the causative mutation, these studies have revealed that the Carwell phenotype displays an imprinted inheritance pattern, with increased muscling associated with inheritance from the sire but not the dam.

The wealth of data generated describing the Carwell region, without the identification of a causal mutation together with the genomic imprinting known to be present at this genomic location, suggests that the Carwell phenotype may not be genetic in origin. The Carwell locus therefore offers a unique opportunity for use as a model system in which to develop data pipelines for analysing whole genome bisulfite sequencing data in order to find epigenetic differences that correlate with economically important productive phenotypes in farmed animals, across the whole genome.

DNA methylation is the most widely studied epigenetic mechanism and is known to play a central role in regulating many aspects of growth and development (Smith & Meissner 2013). The development of high-throughput sequencing has proven to be a vital tool in the analysis of DNA methylation, and reduced representation bisulfite sequencing (RRBS) has been used effectively in understanding DNA methylation landscapes in humans and rats (Meissner et al. 2008; Smith et al. 2009; Hartung et al. 2012). However, to date, almost nothing is known about epigenetics, including DNA methylation in farmed animals. Here we report the use of RRBS to examine the methylation patterns across the genome in sheep muscle. The aim was to determine if epigenetic patterns identified in other organisms are also observed in sheep and to compare DNA methylation levels at a single nucleotide resolution in the Carwell region between sheep homozygous for the Carwell haplotype and genetically similar non- Carwell sheep from the same flock. The latter group are referred to a Control.

Materials and methods

A sample of *Longissimus dorsi* muscle from three Control and three Carwell eight-month-old lambs were collected at slaughter in a commercial abattoir. High quality DNA and RNA were extracted (Sambrook et al. 1989). Illumina sequencing libraries for quantifying DNA methylation levels across the genome were generated using a previously published RRBS protocol (Smith et al. 2009; Couldrey et al. 2012).

Quality control of data was undertaken using FastQC software (Babraham Bioinformatics, Cambridge, UK). Sequences were mapped to sheep genome assembly OARv3.1(International Sheep Genomics Consortium, Brisbane, Australia) using Bismark software (Babraham Bioinformatics, Cambridge, UK). Sequencing read counts and levels of methylation at individual CpG sites were calculated and visualised using Seqmonk software (Babraham Bioinformatics, Cambridge, UK). Genome wide DNA methylation analysis was performed using a 1000bp sliding window with a step size of 50bp. Average DNA methylation levels were extracted from Seqmonk data files using python scripts developed in-house.

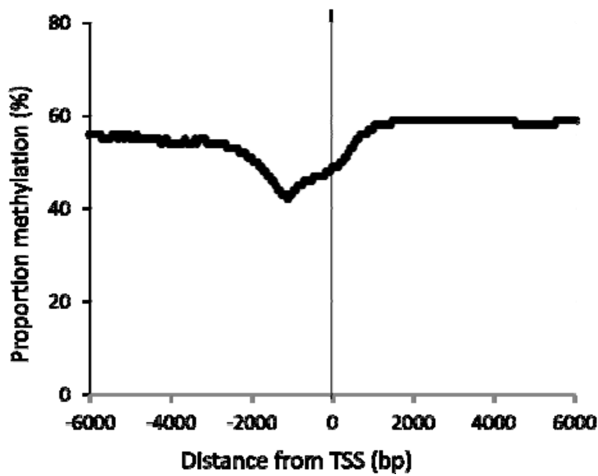
Next generation sequencing of mRNA and microRNA was performed by Macrogen (Seoul, Korea). Data were downloaded and regions of low quality sequence and Illumina primers and adapters remaining from the sequencing process were removed from the reads prior to mapping to the sheep genome. The genes were divided into three groups based on their expression level as determined by transcriptome sequencing, namely repressed genes not expressed at all (60% of all genes), low expression genes (20% of genes), and highly expressed genes (20% of genes).

Results and discussion

Genome wide methylation

Sequencing 1% of the genome using RRBS to target CpG sites resulted in precise and accurate analysis of >10% of all CpG sites. Data indicated that 60% of CpG sites were methylated. While 60% CpG site methylation is lower than traditionally reported in somatic tissues (80%) using techniques able to only determine total methyl content (Ehrlich et al. 1982), this discrepancy is likely to be due to the requirement

Figure 1 Average proportion methylation in *Longissimus dorsi* muscle tissue from three control animals calculated using a 1000 base pair (bp) window with 50bp step size surrounding the transcription start sites (TSS) of all genes annotated in the sheep genome.



of unique mapping to the sheep genome when using next generation sequencing methods to interrogate DNA methylation. Sequences that do not map uniquely are unable to be used in calculating methylation as there is no certainty as to where in the genome they originate from. Previous work (Couldrey & Lee 2010; Couldrey et al. 2011) has indicated that repetitive sequences show high levels of methylation in somatic tissues and likely cause the differences in DNA methylation levels reported here with those reported using historical techniques that did not have the resolution to be informative.

Methylation around transcription start sites

DNA methylation levels around transcriptional start sites (TSS) are known to be involved in regulation of gene expression (Deaton & Bird 2011). On average the DNA sequence directly upstream of the TSS showed less methylation than surrounding sequence (Figure 1). This finding corresponds well with TSS methylation reports in other species (Hartung et al. 2012). However, simply knowing average methylation across all genes does not provide significant insight into the involvement of gene regulation. Whole genome transcriptome sequencing (mRNA and microRNA sequencing) was used to determine gene expression levels in the exact same tissues that were used for DNA methylation profiling.

Methylation around transcription start sites relative to gene expression and DNA sequence surrounding transcription start sites

Based on the grouping of genes according to their expression level, the average DNA methylation around the TSS indicated that genes that were highly expressed, on average, appeared to have lower levels of DNA methylation immediately upstream of the TSS. This finding fits well with the notion that less methylated DNA is likely in a more relaxed state and therefore the gene more open to transcription. However, the difference in methylation levels was not as great as expected suggesting that other factors are also involved. One such factor may be the DNA sequence around the TSS. Due to the non-binomial distribution of these data, direct statistical comparison of averages is difficult.

DNA sequence 6kb upstream and downstream of the TSS was analysed for CpG content. CpG content was highest at the TSS (Figure 2a). This high CpG content corresponds to the presence of CpG dense “CpG islands” know to reside at the start of many genes (Bird 1986). Further analysis of CpG content for each individual gene around 500bp upstream and 500bp downstream of the TSS revealed that the CpG content at the TSS varied from 0% to 10.4%. The frequency distribution (Figure 2b) of TSS CpG content peaks at 1.2% and does not show the bimodal distribution seen in other species (Hartung et al. 2012). It is not clear from these data, whether this difference represents a true species difference, or rather reflects

Figure 2 (a) Average CpG content at each base pair (bp) around the transcription start sites (TSS) of all genes annotated in the sheep genome. (b) Percentage CpG content in1000bp region around TSS in all genes annotated in the sheep genome.

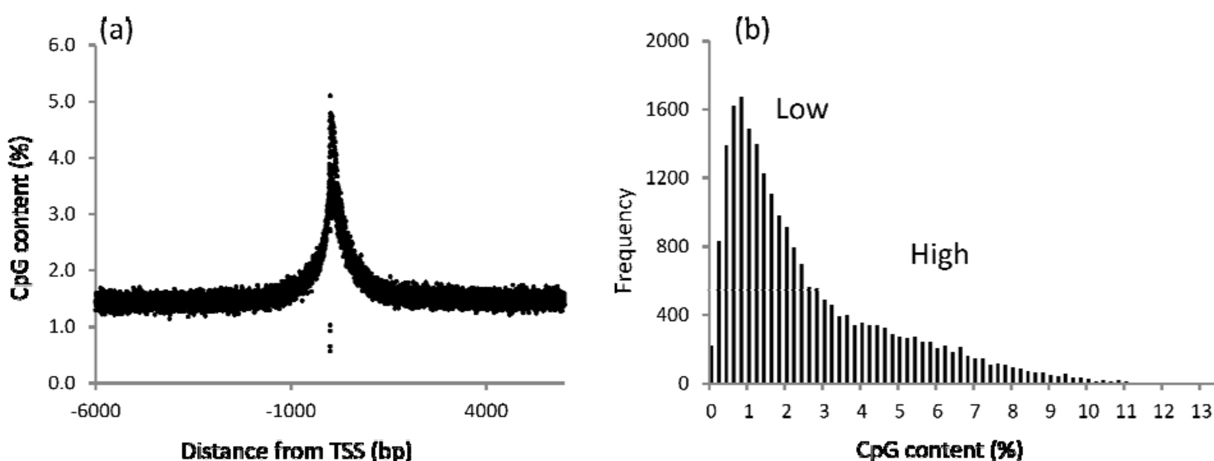
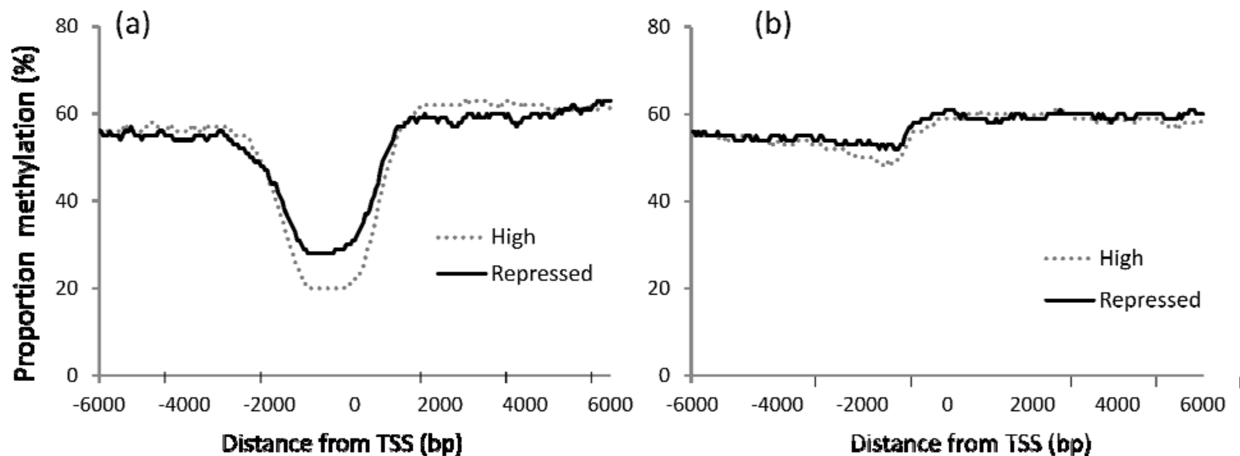


Figure 3 Average proportion methylation in *Longissimus dorsi* muscle tissue from three control animals calculated using 1000 base pair (bp) windows with 50bp step size surrounding the transcription start sites (TSS) of genes measured to be highly expressed or repressed in the same tissue, (a) genes with high CpG content around TSS, (b) genes with low CpG content around TSS



the state of completion of the reference genome as it is known that the current assembly of the sheep genome (OARv3.1) is lacking in sequence around the start of genes (BP Dalrymple, Personal communication), a region that typically contains higher than average levels of CpG dinucleotides. As no bimodal distribution was seen in examining frequency of TSS CpG content, genes were divided into "high" or "low" CpG content based on the distribution seen in other species with the split made at 3.5% (Figure 2b).

Average DNA methylation levels around the TSS were then assessed based on CpG content (high or low) and expression level (high or repressed). This analysis illustrates that genes with low TSS CpG content show little change in DNA methylation regardless of expression level (Figure 3a) while genes with high TSS CpG content have DNA methylation levels around the TSS that correlate with gene expression (Figure 3b).

DNA methylation at single nucleotide resolution

No differences in DNA methylation level were identified between Carwell homozygote and Control sheep muscle when looked at from a genome wide angle. However, RRBS is sensitive enough to analyse individual CpG sites. Focusing on the region of the genome known to give rise to the Carwell phenotype (Nicoll et al. 1998), DNA methylation levels between Carwell and Control animals were examined at the single nucleotide resolution. Four CpG sites were found to show 100% methylation in Control and 0% methylation in Carwell animals. These differences were further examined and were determined to be due to previously unidentified single nucleotide polymorphisms (SNPs) that destroyed the CpG dinucleotide sequences at these locations. A further five CpG have been identified as potentially showing greater DNA methylation in Control compared with Carwell samples and are currently being verified on a larger number of samples including not only Carwell homozygotes and Controls, but also heterozygotes using Sequenom technology previously described (Couldrey & Lee 2010).

Taken together, this work has identified both similarities and differences in genome wide CpG and DNA methylation patterns between sheep and other animals. This work also highlights the complexity of epigenetic regulation of genes, it appears unlikely that regulation of gene expression will be entirely controlled by a particular epigenetic process; rather, it is likely that a combination of epigenetic processes, together with the underlying genomic sequence will interact synergistically to ensure appropriate growth and development.

Acknowledgements

The authors would like to thank Ovita Ltd for sheep muscle samples and Jeremy Bracegirdle for RRBS library construction. Work was supported by an AgResearch Core Funding grant.

References

Bird AP 1986. CpG-rich islands and the function of DNA methylation. Nature 321: 209–213.

Couldrey C, Lee RS 2010. DNA methylation patterns in tissues from mid-gestation bovine fetuses produced by somatic cell nuclear transfer show subtle abnormalities in nuclear reprogramming. BMC Developmental Biology 10: 27.

Couldrey C, Wells DN, Lee RS 2011. DNA methylation patterns are appropriately established in the sperm of bulls generated by somatic cell nuclear transfer. Cellular Reprograming 13: 171–177.

Couldrey C, Brauning R, Retzel EF, McEwan JC 2012. Development of epigenomic pipelines for use in agricultural animals. Proceedings of New Zealand Society of Animal Production 72: 163–165.

Deaton AM, Bird A 2011. CpG islands and the regulation of transcription. Genes and Development 25: 1010–1022.

Ehrlich M, Gama-Sosa MA, Huang LH, Midgett RM, Kuo KC, McCune RA, Gehrke C 1982. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. Nucleic Acids Research 10: 2709–2721.

Hartung T, Zhang L, Kanwar R, Khrebtukova I, Reinhardt M, Wang C, Therneau TM, Banck MS, Schroth GP, Beutler AS 2012. Diametrically opposite methylome-transcriptome relationships in high- and low-CpG promoter genes in postmitotic neural rat tissue. Epigenetics 7: 421–428.

Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein BE, Nusbaum C, Jaffe DB and others 2008. Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature 454: 766–770.

Nicoll GB, Burkin HR, Broad TE, Jopson NB, Greer GJ, Bain WE, Wright CS, Dodds KG, Fennessy PF, McEwan JC 1998. Genetic linkage of microsatellite markers to the Carwell locus for rib-eye muscling in sheep. Proceedings 6th World Congress Genetics Applied Livestock Production, Armidale, Australia: 529–532.

Sambrook J, Fritsch EF, Maniatus T 1989. Molecular cloning a laboratory manual. 2 ed. Cold Spring Harbor, New York, Cold Spring Harbor Laboratory Press.

Smith ZD, Meissner A 2013. DNA methylation: roles in mammalian development. Nature Reviews 14: 204–220.

Smith ZD, Gu H, Bock C, Gnirke A, Meissner A 2009. High-throughput bisulfite sequencing in mammalian genomes. Methods 48: 226–232.