Methylation patterns in ejaculated sperm from Nellore-Angus crossbred bulls

M. R. Prause, G. Wang, & C. A. Gill

Texas A&M University, Department of Animal Science, College Station, TX 77843, USA
clare-gill@tamu.edu (Corresponding Author)

Summary

Methylation of the carbon 5 position of cytosine is a common epigenetic mark in vertebrates, and has been associated with male infertility. Methylation can be affected by environmental changes in utero and our long-term goal is to establish whether maternal nutrition affects the bovine sperm methylome and variation in male fertility. As a first step, our objective here was to establish the general pattern of methylation in ejaculated bovine spermatozoa. We performed whole genome bisulfite sequencing for five F1 Nellore-Angus crossbred bulls and sequences were aligned to a bisulfite-converted version of the UMD3.1 bovine reference assembly. The average level of CpG methylation was 88.2%. Gene bodies were heavily methylated, whereas promoter and CpG islands tended to be unmethylated. For promoters with unmethylated CpG islands, there was enrichment for pathways associated with transcription.

Keywords: bovine, methylation, sperm, whole genome bisulfite sequencing

Introduction

DNA methylation describes the covalent addition of a methyl group to cytosine, and is one of the most widely studied epigenetic mechanisms, which is fundamental for cellular processes and regulation (Portela and Esteller, 2010; Urdinguio et al., 2015; Lim et al., 2016). Cytosine methylation is the addition of a methyl group to the fifth position of the pyrimidine ring. In eukaryotes, DNA methylation occurs mainly in CpG dinucleotide sequences accounting for 60 to 90% of methylated cytosines (Bird, 2002; Cisneros, 2004; Portela & Esteller, 2010).

Several studies have demonstrated that CpG methylation in mammalian sperm is very high. Frommer et al. (1992) found CpG dinucleotides in non-repetitive sequences in human sperm were hypermethylated, whereas CpG islands (CGI) were not. Approximately 70% of the DMR in humans are associated with CGI shores (Portela & Esteller, 2010). According to Popp et al. (2010), 80-90% of CpG in mature mouse spermatozoa are methylated, which is the highest level of methylation observed for specific cell types. Although exons are highly methylated, methylation of promoters is considerably lower (35-40%). Promoters associated with genes
important in early development tend to be hypomethylated in the sperm genome (Hammoud et al., 2011). The objective of this study was to establish the general pattern of methylation in ejaculated bovine spermatozoa.

Materials and methods

Semen samples were collected from five Nellore-Angus bulls using standard electroejaculation methods and flash frozen without extender in liquid nitrogen. Sperm DNA was extracted by standard proteinase K digestion followed by extraction with phenol-chloroform.

Library preparation, bisulfite conversion, and whole genome bisulfite sequencing were performed by Novogene Ltd. (Hong Kong). The bisulfite treated library was size selected and amplified by PCR prior to cluster generation and paired-end 150 bp sequencing with indexing on an Illumina HiSeq 2500.

FastQC (Bock, 2012) was used to trim reads containing adapters, trim reads with > 10% N, and trim reads with low quality. Bisulfite reads were converted to C-to-T and G-to-A and aligned to C-to-T and G-to-A converted versions of the UMD3.1.1 bovine reference genome using Bismark v0.16.1 (Krueger & Andrews, 2011). No mismatches were allowed between the converted read and reference sequence and the minimum alignment score was set as a function \( f(x) = 0 + (-0.6) \times \text{L} \) where L is the 150 bp read length. PCR duplicates were filtered using Bismark deduplicate and methylation status was summarized using Bismark methylation extractor.

Finally, methylation was classified relative to annotated genome features: CpG islands (CGI), CGI shores, promoters and gene bodies. Annotations for CGI were downloaded from UCSC. CpG island shores were defined as sequences 1000 bp upstream and downstream of each CGI (Doerks et al., 2002; Wu et al., 2015). Annotations for RefSeq genes were from Bos taurus annotation release 104. Promoters were defined as sequences 2200 bp upstream to 500 bp downstream of the first nucleotide of a RefSeq gene (Jin et al., 2014). Gene bodies were defined as beginning 500bp into a RefSeq gene and ending at the last nucleotide of the annotation. Gene ontology (GO) and KEGG pathway enrichment analyses were performed using DAVID v6.8beta (Huang et al., 2009a,b). Within category Benjamini-Hochberg correction was applied to control the false discovery rate.

Results and discussion

After quality control filtering, clean sequence data ranged from 48.06 Gb to 69.4 Gb and mapping rates for the bisulfite-converted sequences ranged from 69.1% to 74.1%.
Based on the average methylation in 1 Mb non-overlapping windows across the 5 samples, global methylation in mature sperm was 88.2%. CpG sites in gene bodies were strongly methylated (median 94.76%). Extensive methylation of gene bodies is common in bovine somatic tissues (Huang et al., 2009b), but has not been previously described for mature sperm. The CpG sites in CGI-shores also were heavily methylated (median 65.9%), whereas in promoters (median 24.0%) and CGI (median 12.5%) the CpG sites tended to be unmethylated, similar to previous studies (Gardiner-Garden & Frommer, 1987; Irizarry et al., 2009; Portela & Esteller, 2010; Suzuki et al., 2013; Su et al., 2014). However, there was much more variability in the level of methylation of CGI, CGI shores, and promoters than in gene bodies. In particular, 25% of CGI had <2% methylation and 25% of the CGI had >98.8% methylation. Methylation of CGI in promoters is associated with stable silencing of gene expression (Bird, 2002; Illingworth & Bird, 2009). Conversely, active promoters tend to be unmethylated (Cisneros, 2004; Saxonov et al., 2006; Aran et al., 2011). Enrichment analysis using the set of genes near unmethylated CGI showed strong enrichment in pathways associated with transcription (Table 1).

Table 1. Enrichment analysis of unmethylated CpG islands near genes.

<table>
<thead>
<tr>
<th>Term</th>
<th>Count</th>
<th>%</th>
<th>Enrichment P-value</th>
<th>Benjamini P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Membrane-enclosed lumen</td>
<td>202</td>
<td>5.3</td>
<td>1.7e-21</td>
<td>9.1e-19</td>
</tr>
<tr>
<td>Intracellular organelle lumen</td>
<td>195</td>
<td>5.1</td>
<td>3.9e-21</td>
<td>1.1e-18</td>
</tr>
<tr>
<td>Organelle lumen</td>
<td>195</td>
<td>5.1</td>
<td>4.9e-21</td>
<td>8.9e-19</td>
</tr>
<tr>
<td>RNA processing</td>
<td>104</td>
<td>2.7</td>
<td>2.8e-14</td>
<td>8.7e-11</td>
</tr>
<tr>
<td>RNA binding</td>
<td>116</td>
<td>3.1</td>
<td>1.2e-11</td>
<td>1.4e-8</td>
</tr>
<tr>
<td>Transcription</td>
<td>148</td>
<td>3.9</td>
<td>1.8e-9</td>
<td>1.8e-6</td>
</tr>
<tr>
<td>DNA replication</td>
<td>37</td>
<td>1.0</td>
<td>8.5e-7</td>
<td>2.6e-4</td>
</tr>
<tr>
<td>ATP binding</td>
<td>243</td>
<td>6.4</td>
<td>5.1e-6</td>
<td>8.0e-4</td>
</tr>
<tr>
<td>tNRA metabolic process</td>
<td>40</td>
<td>1.1</td>
<td>9.5e-9</td>
<td>7.3e-6</td>
</tr>
<tr>
<td>tRNA aminoacylation</td>
<td>21</td>
<td>0.6</td>
<td>6.0e-6</td>
<td>1.2e-3</td>
</tr>
</tbody>
</table>

List of References


