Genome-wide ChIP-seq Mapping & Analysis of Butyrate-induced Histone Acetylation in Bovine Cells

Cong-jun Li, Ph. D.
Bovine Functional Genomics Laboratory
ARS, USDA

PAG XX, San Diego, CA, 2012
The Sequence of the Bovine Genome Reveals New Insights into the history and Biology of Cattle
Epigenetics: DNA Isn’t Everything
Histone winds up long DNA into a compact state
Molecular basis of epigenetics

Epigenetic regulations:

1) **Histone modifications:**
   - acetylation
   - methylation
   - phosphorylation
   - ubiquitination

2) **DNA methylation**

3) **Genomic imprinting**

4) **miRNA**

5) **Prion proteins**

6) **etc. etc. etc...**
Ruminant species have evolved to metabolize VFA (acetate, propionate, and butyrate) to fulfill up to 70% of their nutrient energy requirements.

VFA have important implication for ruminant physiology.

Butyrate links nutrition to histone acetylation and epigenetic regulation.
Global gene expression profiles indicated:

1) Only 2~3% of genes are regulated by butyrate
2) There are more genes down regulated than up-regulated by butyrate
3) Other regulatory pathway may be involved too

Funct Integr Genomics. 2007 7(3):193-205.
Workflow of ChIP-CHIP/ChIP-seq
**Overall Processing Procedure**

1. **Illumina GAII**
   - 45bp single-end reads

2. **Genome Alignment**

3. **Peak Calling**
   - Enriched regions
   - **Bowtie**
     - Single-end ChIP-Seq
     - Reference genome used: bosTau4
   - **MACS**
     - Y. Zhang et al., *Genome Biology* (2008) 9: R137

4. **Post-processing**

5. **Motif Discovery**
   - **MEME**

6. **Visualization**
   - **PeakAnalyze**
     - Peak annotation, Nearest Downstream Genes analysis, Transcription Start Site analysis
   - **Integrative Genomics Viewer**
### Bowtie Alignment Results

<table>
<thead>
<tr>
<th>Sample</th>
<th>Total reads</th>
<th>Mapped</th>
</tr>
</thead>
<tbody>
<tr>
<td>H3</td>
<td>29,780,699</td>
<td>27,078,053 (90.92%)</td>
</tr>
<tr>
<td>H3K9</td>
<td>23,326,139</td>
<td>21,368,435 (91.61%)</td>
</tr>
<tr>
<td>H3K27</td>
<td>30,901,336</td>
<td>28,318,383 (91.64%)</td>
</tr>
</tbody>
</table>

### MACS Peak Detection Results

<table>
<thead>
<tr>
<th>Sample</th>
<th># peaks detected</th>
</tr>
</thead>
<tbody>
<tr>
<td>H3</td>
<td>16,927</td>
</tr>
<tr>
<td>H3K9</td>
<td>9,160</td>
</tr>
<tr>
<td>H3K27</td>
<td>50,754</td>
</tr>
</tbody>
</table>
Nearest Downstream Gene (NDG) Analysis

H3

H3K9

H3K27

Funct & Integr Genomics, 2012, In press
Transcription Start Site (TSS) Analysis

Funct & Integr Genomics, 2012, In press
UCSC browser view of Genomic data of representative chromosome regions (peak similarity)
UCSC browser view of Genomic data of representative chromosome regions (peak similarity)
UCSC browser view of Genomic data of representative chromosome regions (peak dissimilarity)
Correlation of gene perturbed by butyrate & ChIP-seq detected binding peaks

ChIP-seq peaks located up-stream of the genes

ChIP-seq peaks overlapping the gene structure

Funct & Integr Genomics, 2012, In press
Genome browser view of genomic data of ORC1L gene region in chromosome 3

0 2.5 5.0 10 (BT, mM)

- MCM3
- ORC1
- ORC2
Dispersed epigenetic 'domains' close to TSS of IGF2
Motif discovery: Histone H3, acetyl-H3K9 and acetyl-H3K27 have different binding preferences.
A high degree of conservation in histone binding: consensus sequencing of motifs

Degenerate consensus sequence
X = A or C
Z = A, C, G, T

<table>
<thead>
<tr>
<th></th>
<th>H3</th>
<th>H3K9</th>
<th>H3K27</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T T C A</td>
<td>T C T C</td>
<td>G A G T</td>
</tr>
<tr>
<td></td>
<td>C T C A</td>
<td>T C T C</td>
<td>C A G G</td>
</tr>
<tr>
<td></td>
<td>T T C C</td>
<td>C A T C</td>
<td>G A G A</td>
</tr>
<tr>
<td></td>
<td>C G G C</td>
<td>G T G A</td>
<td>T G T G</td>
</tr>
<tr>
<td></td>
<td>C T C A</td>
<td>T C T C</td>
<td>G A G A</td>
</tr>
<tr>
<td></td>
<td>T G C C</td>
<td>T C T C</td>
<td>G A G A</td>
</tr>
<tr>
<td></td>
<td>C G G C</td>
<td>G T G A</td>
<td>T G T G</td>
</tr>
</tbody>
</table>

Degenerate consensus sequence:
T C X T C T C G A G Z

Funct & Integr Genomics, 2012, In press
H3

H3K9

H3K27

Funct & Integr Genomics, 2012, In press
A consensus sequence with CACGTG elements (G-BOX)

H3: GATCAGTGGCT
H3K9: AACACGTTTT
H3K27: TCACTGGCTG

T C A C G T G G C T

A: 0 0 3 0 0 0 0 0 0 0
C: 1 3 0 3 0 0 0 0 2 0
G: 0 0 0 0 3 0 3 3 0 0
T: 2 0 0 0 0 3 0 0 1 3

Funct & Integr Genomics, 2012, In press
### Degenerate consensus sequence:

Y = A,C,T

<table>
<thead>
<tr>
<th></th>
<th>T</th>
<th>C</th>
<th>A</th>
<th>C</th>
<th>G</th>
<th>T</th>
<th>G</th>
<th>G</th>
<th>Y</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>H3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>C</td>
<td>4</td>
<td>12</td>
<td>1</td>
<td>11</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>G</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>11</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>T</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>10</td>
</tr>
</tbody>
</table>

**Funct & Integr Genomics, 2012, In press**
Summary:

1. Butyrate-induced acetylation in H3K9 and H3H27 change the sequence-based binding preference of H3.
2. A high degree of conservation in the histone binding evidently presented.
3. Epigenetics is intimately related to robustness as both lie between genotypes and phenotypes.
Acknowledgments:

ChIP-seq:

Dr. Robert W. Li
Dr. Congjun Li
Bovine Functional Genomics Lab
ARS, USDA

Bioinformatics:

Dr. Yuan Gao
Dr. Joo Heon Shin
Lieber Institute for Brain Development
Johns Hopkins University
Think epigenetics...

Thanks!