Satellite Education Workshop (SW4)

DNA Methylomics

Ting Wang
Department of Genetics
Center for Genome Sciences and Systems Biology
Washington University

ABRF2012
March 17, 2012
Part II

Analysis of DNA Methyonomics Data
Outline

• Case 1: Comparison of popular methods

• Case 2: Function of Intragenic DNA methylation

• Integrative methods (combining MeDIP and MRE)
  – Case 3: Insights into monoallelic gene regulation and imprinting
  – Case 4: Predicting DNA methylation at single CpG resolution
  – (Case 5: Cancer methylomes)

• DNA methylomes/Epigenomes resources
Case 1: Comparison of DNA methylation mapping technologies

Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications

R Alan Harris, Ting Wang, Cristian Coarfa, Raman P Nagarajan, Chibo Hong, Sara L Downey, Brett E Johnson, Shaun D Fouse, Allen Delaney, Yongjun Zhao, Adam Olshen, Tracy Ballinger, Xin Zhou, Kevin J Forsberg, Junchen Gu, Lorigail Echpare, Henriette O'Geen, Ryan Lister, Mattia Pelizzola, Yuanxin Xi, Charles B Epstein, Bradley E Bernstein, R David Hawkins, Ping Ren, Wen-Yu Chung, Hongcang Gu, Christoph Bock, Andreas Gnírke, Michael Q Zhang, David Haussler, Joseph R Ecker, Wei Li, Peggy J Farnham, Robert A Waterland, Alexander Meissner, Marco A Marra, Martin Hirst, Aleksandar Milosavljevic, & Joseph F Costello

Quantitative comparison of genome-wide DNA methylation mapping technologies

Sequencing DNA methylomes

- Coverage
- Resolution
- Transposons
- Assign epigenetic state to genetic allele
- Mutation detection
- Copy number profile
# Comparison of Methylome coverage

<table>
<thead>
<tr>
<th>Method</th>
<th>CpG coverage</th>
<th>CpG island coverage</th>
<th>Resolution (bp)</th>
<th>Illumina Lanes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genome total</td>
<td>28 M</td>
<td>28 K</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>BS shot gun</td>
<td>26 M</td>
<td>27 K</td>
<td>1</td>
<td>207 (2009)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10 (now)</td>
</tr>
<tr>
<td>RRBS</td>
<td>0.2-1M</td>
<td>15 K</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>MRE-seq and MeDIP-seq</td>
<td>25 M</td>
<td>27 K</td>
<td>1 and 200</td>
<td>8 (2009)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1 (now)</td>
</tr>
<tr>
<td>Golden-Gate</td>
<td>1,500</td>
<td>800</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td>Infinium</td>
<td>27,500</td>
<td>12,000</td>
<td>1</td>
<td>NA</td>
</tr>
</tbody>
</table>
Genome-wide and CGI Coverage

% CpG covered
Genome-wide

% CpG covered
In CpG islands

Read coverage threshold for CpGs

Graph showing MethyC-seq and RRBS data with comparison of % CpG covered genome-wide and in CpG islands.
Comparison of MethylC-seq and RRBS

At 0.25 difference, MethylC-seq and RRBS are 82% concordant

At 0.25 difference, MethylC-seq and RRBS are 82% concordant

Difference in methylation level called by methylC vs RRBS (minimum 5-read coverage)
Comparison of MeDIP-seq and MBD-seq

<table>
<thead>
<tr>
<th>Minimum Read Depth</th>
<th>1000bp Windows</th>
<th>200bp Windows</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number of Windows</td>
<td>% Genome-wide CpGs</td>
</tr>
<tr>
<td>2</td>
<td>1,189,545</td>
<td>61.82%</td>
</tr>
<tr>
<td>5</td>
<td>446,096</td>
<td>32.65%</td>
</tr>
<tr>
<td>10</td>
<td>162,661</td>
<td>15.07%</td>
</tr>
</tbody>
</table>

![Graph showing comparison of MeDIP-seq and MBD-seq](image)
Comparison of MethylC, RRBS, MeDIP, MeDIP, MBD
Methylome Methods Comparison

- Shotgun bisulfite
  - Base resolution
  - Absolute quantitation
  - Higher cost/sample
  - mC not distinguished from hmC

- Enrichment
  - 150bp resolution
  - Relative quantitation
  - Much lower cost/sample

- Integrative
  - 1-150bp resolution
  - Detection of intermediate and allelic methylation states
Conserved role of intragenic DNA methylation in regulating alternative promoters

Alika K. Maunakea1*†, Raman P. Nagarajan1*, Mikhail Bilenky2, Tracy J. Ballinger3, Cletus D’Souza2, Shaun D. Fouse1, Brett E. Johnson1, Chibo Hong1, Cydney Nielsen2, Yongjun Zhao2, Gustavo Turecki4, Allen Delaney2, Richard Varhol2, Nina Thiessen2, Ksenya Shchors5†, Vivi M. Heine6, David H. Rowitch6, Xiaoyun Xing7, Chris Fiore7, Maximiliaan Schillebeeckx7, Steven J. M. Jones2, David Haussler3,8, Marco A. Marra2, Martin Hirst2, Ting Wang3,7 & Joseph F. Costello1
What is the function of gene body methylation?

Zhang et al 2006, MeDIP-chip, Arabidopsis

Cokus et al 2008, BS-seq, Arabidopsis

Lister et al 2008, MethylC-seq, Arabidopsis

Ball et al 2009, bisulfite padlock probes, human cell line

Lister et al 2009, MethylC-seq, human ES cells

Maunakea et al 2010, MeDIP+MRE, human brain
Genome-wide CpG site coverage

28 million CpGs

- MeDIP only (22.65M)
- MRE only (1.04M)
- Both (0.91M)
- None (3.3M)

- Sample: human brain (frontal cortex)
- MeDIP-seq: 100 million reads
- MRE-seq: 30 million reads
Intragenic CpG island methylation is common; 5’ methylation is rare.
Intragenic CpG island methylation is common; 5’ methylation is rare

Mouse data from: Meissner, Mikkelsen et al, 2008
**Methylation status of orthologous CpG islands is evolutionarily conserved**

<table>
<thead>
<tr>
<th></th>
<th>Promoter CGIs (n=1245)</th>
<th>Intragenic CGIs (n=502)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unmethylated in human</td>
<td>Methylated in human</td>
</tr>
<tr>
<td>Unmethylated in mouse</td>
<td>1234</td>
<td>3</td>
</tr>
<tr>
<td>Methylated in mouse</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>P-value</td>
<td>&lt; 0.0001</td>
<td>&lt; 0.0001</td>
</tr>
</tbody>
</table>
Overlap of H3K4me3, CAGE and Intragenic CpG island

CAGE data from Carninci et al 2005, 2008
5’ CpG island: unmethylated

CpGi1

<table>
<thead>
<tr>
<th></th>
<th>Mouse</th>
<th>Human</th>
<th>Percent Methylation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brain</td>
<td>8</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Keratinocytes</td>
<td>9</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>PBL</td>
<td>8</td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sample Type</th>
<th>n</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>Percent Methylation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brain</td>
<td>8</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>4.2%</td>
</tr>
<tr>
<td>Keratinocytes</td>
<td>9</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>1.7%</td>
</tr>
<tr>
<td>PBL</td>
<td>8</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>2.7%</td>
</tr>
<tr>
<td>Brain</td>
<td>9</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>1.6%</td>
</tr>
<tr>
<td>Keratinocytes</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>1.5%</td>
</tr>
<tr>
<td>PBL</td>
<td>10</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>1.5%</td>
</tr>
</tbody>
</table>
Intragenic CpG island: tissue specific methylation
Intragenic CpG island initiates transcripts
Methylation level anti-correlates with transcript level
Initial observations from brain epigenomes

- Tissue-specific methylation is common in intragenic CpG islands, but rare in 5’ islands

- Genome sites of tissue-specific intragenic methylation overlap significantly with markers of TSS, and promoter-enriched histone modifications
  - Suggesting a major role for intragenic methylation in regulating cell context-specific alternative promoters in gene bodies
Case 3: Combining MeDIP and MRE to investigate intermediate methylation

- Integrative method identifies intermediate methylation states
- Genetic variation detected by epigenotyping
- Initial catalogue of novel imprinted gene candidates
Allevle specific methylation

- Imprinting
- Parent-of-origin specific expression
- Normal growth and brain development
- 100 genes known, but many DMRs unknown (Chaofani, 2011, others)

- X-chromosome inactivation
  - In Females, promoters vs gene bodies

- Monoallelic gene regulation
  - Widespread on somatic chromosomes
    (Hellman and Chess, etc)
Integrative Method

• Methyl-sensitive restriction enzyme – sequencing (MRE-seq)
  each read is a single unmethylated CpG site

• Methyl DNA immunoprecipitation - sequencing (MeDIP-seq)
  higher read density at methylated regions

MRE-seq

MeDIP-seq
Allele Specific Methylation (ASM)

Unmethylated CpGs (MRE-Seq)

Methylated CpGs (MeDIP-Seq)

Reference Allele

Variant Allele

Harris et al, Nature Biotechnology 2010
The genome is divided into methylated and unmethylated domains
Complementarity of MeDIP-seq and MRE-seq

![Graph showing the complementarity of MeDIP-read density and MRE-read density.](image)
The “iMethylome” algorithm
Allele-specific methylation at imprinted genes

Chr 7:
SNURF/SNURP
MeDIP-seq
MRE-seq

Chr 7:
MEST
MeDIP-seq
MRE-seq

Chr 18:
SMAD4
MeDIP-seq
MRE-seq

imprinted
not
imprinted
Enriching for Intermediate Methylation

Bisulfite Score Distributions: Genome-wide vs. IM Regions

- Genome-wide
- IM Regions
- Known DMRs

Density

Bisulfite Score

Pie chart:
- MRE 8%
- IM 1%
- Undetermined 31%
- MeDIP 60%
Intermediate methylation levels at imprinted genes

- **SNRPN**
  - Unmethylated
  - Methylated
  - H3K4me3

- **H19**
  - Unmethylated
  - Methylated
  - H3K4me3

- **PEG10**
  - Unmethylated
  - Methylated
  - H3K4me3

- **GNAS**
  - Unmethylated
  - Methylated
  - H3K4me3
Intermediate methylation levels in \textit{POTE\textsubscript{B}}

<table>
<thead>
<tr>
<th>Location</th>
<th>Medip Allele</th>
<th>Count</th>
<th>MRE Allele</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>chr15:19346666-19350003</td>
<td>G</td>
<td>9</td>
<td>A</td>
<td>30</td>
</tr>
</tbody>
</table>
Validation of monoallelic DNA methylation in *POTEB*
Putative Imprinted Gene

```
Scale
chr19
CpG Islands
H3K4me3
MRE-seq
MeDIP-seq
Bisulfite region
ZNF331
```

```
Bisulfite region 1
```

```
Bisulfite region 2
```
Conservation of IM

Mouse IM
3850

Human IM
1614
(converted to mouse)

Candidate Imprinted Mouse Genes
308

Total = 6080

Human IM
154

Mouse IM
38

Candidate Imprinted Mouse Genes
378

Total = 6080

Total = 2772

Total = 1032
IMs are regulatory elements

Distance to DNaseI HS

- 0kb: 66%
- >0-0.5kb: 17%
- >0.5-2kb: 10%
- >2-10kb: 6%
- >10kb: 2%
Case 4: Predicting single CpG methylation level with Conditional Random Field (methylICRF)

1. A novel statistical framework for integrative analysis of MeDIP and MRE data
DNA Methylation is locally correlated
A conditional random field model

\[HMM \Rightarrow P(5mC)P(MeDIP,MRE,\ldots | 5mC)\]
\[= P(5mC, MeDIP, MRE, \ldots)\]
\[= P(5mC | MeDIP, MRE, \ldots)P(MeDIP, MRE, \ldots)\]

\[CRF \Rightarrow P(5mC | MeDIP, MRE, \ldots)\]
Predicting single CpG methylation level with Conditional Random Field (methylCRF)
Predicting single CpG methylation level with Conditional Random Field (methylCRF)
Sequence-based DNA Methylome/Epigenome Resources

  - Data visualization hub for Roadmap Epigenomics Project
  - [http://epigenomegateway.wustl.edu/](http://epigenomegateway.wustl.edu/)
    - Next generation genome browser
- [http://www.roadmapepigenomics.org/](http://www.roadmapepigenomics.org/)
  - Roadmap Consortium
- [http://www.genboree.org/epigenomeatlas/index.rhtml](http://www.genboree.org/epigenomeatlas/index.rhtml)
  - Data coordination center at Baylor
  - NCBI, where you can download data
Roadmap Epigenomics Visualization Hub (VizHub)

Click Here to Access VizHub
Powered by UCSC Genome Browser

Click Here to Access UCSC Browser's Remote Data Hub Page

Click Here to Access New WashU-Epigenome Browser (Beta Version)

Roadmap Consortium
NCBI Epigenomics Gateway
Epigenome Data and Analysis Coordination Center

http://VizHub.wustl.edu
Integrative/summary tracks

• New (well, a year old) UCSC- technology
• Overlay and summarize many data tracks
  – Different experiments for the same sample;
  – Same experiments for different samples;

MeDIP peaks, or methylated regions
MRE peaks, or unmethylated regions
Combined MeDIP/MRE tracks of 34 samples
Integrative/summary tracks

- Can choose to integrate on samples or on epigenetic marks
- “Rainbow tracks”

MeDIP peaks, or methylated regions

MRE peaks, or unmethylated regions
Summary View of Epigenome Atlas 1

Methylation

- EDACC medip Summary
- EDACC Mre Summary
- EDACC BisulfiteSeq Summary
- EDACC H2AK9ac Summary
- EDACC H2BK120ac Summary
- EDACC H2BK12ac Summary
- EDACC H2BK15ac Summary
- EDACC H2BK20ac Summary
- EDACC H3K14ac Summary
- EDACC H3K18ac Summary
- EDACC H3K23ac Summary
- EDACC H3K27ac Summary
- EDACC H3K27me3 Summary
- EDACC H3K36me3 Summary
- EDACC H3K4ac Summary
- EDACC H3K4me1 Summary
- EDACC H3K4me3 Summary
- EDACC H3K56ac Summary
- EDACC H3K79me2 Summary
- EDACC H3K9ac Summary
- EDACC H3K9me3 Summary
- EDACC H4K20me1 Summary
- EDACC H4K5ac Summary
- EDACC H4K8ac Summary
- EDACC H4K9ac Summary
- EDACC Input Summary

RNA

- EDACC mRNASeq Summary
- EDACC smRNASeq Summary
Roadmap Epigenomics Visualization Hub

Click Here to Access VizHub Powered by UCSC Genome Browser

Click Here to Access UCSC Browser's Remote Data Hub Page

Click Here to Access New WashU-Epigenome Browser (Beta Version)

http://genome.ucsc.edu/
Go to **http://genome.ucsc.edu/**, choose genomes

---

**About the Human Feb. 2009 (GRCh37/hg19) assembly (sequences)**

The February 2009 human reference sequence (GRCh37) was produced by the [Genome Reference Consortium](http://www.genome.org/about/grc.htm). For more information, please refer to the [User's Guide](http://genome.ucsc.edu/hg18.html).

**Sample position queries**

A genome position can be specified by the accession number of a sequenced genomic clone, an mRNA or EST marker, a chromosomal coordinate range, or keywords from the GenBank description of an mRNA. The following list shows examples of valid position queries for the human genome. See the [User's Guide](http://genome.ucsc.edu/hg18.html) for more information.

**Request:**

- **chr7**
  - Displays all of chromosome 7
- **chrUn_g0000212**
  - Displays all of the unplaced contig g0000212
- **chr3:1-1000000**
  - Displays first million bases of chr 3, counting from p-arm telomere
- **chr3:1000000+2000**
  - Displays a region of chr3 that spans 2000 bases, starting with position 1000000
- **RH18061;RH80175**
  - Displays region between genome landmarks, such as the STS markers RH18061 and RH80175. This syntax may also be used for other range queries, such as between uniquely determined ESTs, mRNAs, refSeqs, etc.
- **D16S3046**
  - Displays region around STS marker D16S3046 from the Genethon/Marshfield maps. Includes 100,000 bases on each side as well.
- **AA205474**
  - Displays region of EST with GenBank accession AA205474 in BRCA1 cancer gene on chr 17
- **AC008101**
  - Displays region of clone with GenBank accession AC008101
Track data hubs are collections of tracks from outside UCSC, which can be imported into the Genome Browser. To import a public hub, check the box in the list below. After import, the hub will show up as a group of tracks with its own blue bar and label underneath the main browser graphic, and in the configure page.

**genome:** Human  assembly: Feb. 2009 (GRCh37/hg19)

### Public Hubs

<table>
<thead>
<tr>
<th>Display</th>
<th>Hub Name</th>
<th>Description</th>
<th>URL</th>
</tr>
</thead>
</table>

Contact genome@soe.ucsc.edu to add a public hub.
Scale chr2:

1729000001
1729500001
1730000001

25 kb

H3K27Ac Mark (Often Found Near Active Regulatory Elements) on 7 cell lines from ENCODE

Digital DNase Hypersensitivity Clusters from ENCODE

Transcription Factor ChIP-seq from ENCODE

UCSD H1 Bisulfite Seq Library combined EA Release 2

UCSF-UBC H1 MeDIP-Seq Library HS1376 EA Release 2 Hotspot_Score=0.3687 Pcnt=62

UCSF-UBC H1 MeDIP-Seq Library HS1303 EA Release 2 Hotspot_Score=0.3021 Pcnt=37

UCSF-UBC H1 MRE-Seq Library HS1153 EA Release 2

UCSF-UBC H1 MRE-Seq Library HS1052 EA Release 2

BI H1 Histone H3K27me3 Donor Solexa-12523 EA Release 2 Hotspot_Score=0.3386 Pcnt=87

BI H1 Histone H3K27me3 Donor Solexa-8039 EA Release 2 Hotspot_Score=0.2783 Pcnt=71

BI H1 Histone H3K4me3 Donor Solexa-12522 EA Release 2 Hotspot_Score=0.2734 Pcnt=23

BI H1 Histone H3K4me3 Donor Solexa-8038 EA Release 2 Hotspot_Score=0.4388 Pcnt=61

BI H1 Histone Input Donor Solexa-10531 EA Release 2 Hotspot_Score=0.0894 Pcnt=73

BI H1 Histone Input Donor Solexa-12532 EA Release 2 Hotspot_Score=0.0809 Pcnt=81

MAP1D

DLX1

DLX2

Layered H3K27Ac

UCSD H1 BisulfiteSeq combined

UCSF-UBC H1 MeDIP HS1376

UCSF-UBC H1 MeDIP HS1303

UCSF-UBC H1 MRE HS1153

UCSF-UBC H1 MRE HS1052

BI H1 Histone H3K27me3 Solexa-12523

BI H1 Histone H3K27me3 Solexa-8039

BI H1 Histone H3K4me3 Solexa-12522

BI H1 Histone H3K4me3 Solexa-8038

BI H1 Histone Input Solexa-10531

BI H1 Histone Input Solexa-12532

UCSC Genes Based on RefSeq, UniProt, GenBank, CCDS and Comparative Genomics

H3K27Ac Mark (Often Found Near Active Regulatory Elements) on 7 cell lines from ENCODE

Digital DNase Hypersensitivity Clusters from ENCODE

Transcription Factor ChIP-seq from ENCODE

UCSD H1 Bisulfite Seq Library combined EA Release 2

UCSF-UBC H1 MeDIP-Seq Library HS1376 EA Release 2 Hotspot_Score=0.3687 Pcnt=62

UCSF-UBC H1 MeDIP-Seq Library HS1303 EA Release 2 Hotspot_Score=0.3021 Pcnt=37

UCSF-UBC H1 MRE-Seq Library HS1153 EA Release 2

UCSF-UBC H1 MRE-Seq Library HS1052 EA Release 2

BI H1 Histone H3K27me3 Donor Solexa-12523 EA Release 2 Hotspot_Score=0.3386 Pcnt=87

BI H1 Histone H3K27me3 Donor Solexa-8039 EA Release 2 Hotspot_Score=0.2783 Pcnt=71

BI H1 Histone H3K4me3 Donor Solexa-12522 EA Release 2 Hotspot_Score=0.2734 Pcnt=23

BI H1 Histone H3K4me3 Donor Solexa-8038 EA Release 2 Hotspot_Score=0.4388 Pcnt=61

BI H1 Histone Input Donor Solexa-10531 EA Release 2 Hotspot_Score=0.0894 Pcnt=73

BI H1 Histone Input Donor Solexa-12532 EA Release 2 Hotspot_Score=0.0809 Pcnt=81

UCSC local data

Roadmap Remote Hub data

ENCODE

BS

MeDIP

MRE

K27me3

K4me3

Input
Roadmap Epigenomics Visualization Hub (VizHub)

Click Here to Access VizHub
Powered by UCSC Genome Browser

Click Here to Access UCSC Browser’s Remote Data Hub Page

Click Here to Access New WashU-Epigenome Browser (Beta Version)

http://epigenomegateway.wustl.edu/
(Nature Methods, in press)
Welcome to the gateway to the Human Epigenome Browser. This tool provides researchers with a cutting-edge resource for visualizing and interacting with whole-genome datasets. The browser currently hosts Human Epigenome Atlas data produced by the Roadmap Epigenomics project, but its use of advanced, multi-resolution data formats and its user-friendly interface make it possible for investigators to upload and visualize their own data as custom tracks.

We invite and encourage you to explore the multitude of functions provided by the Human Epigenome Browser including zooming and scrolling, hypothesis testing and correlation analysis on groups of data, and data juxtaposition in which users can concentrate visualization on a set of genes or genomic features of interest.

The Human Epigenome Browser is developed and maintained by the Epigenome Informatics Group at Washington University in St. Louis. If you have questions or comments related to the tools or data on this website, please feel free to contact us on our public mailing list.
Summary

- **Motivation**
  - Larger datasets (whole genome, sequencing based)
  - More datasets (hundreds, thousands)
  - Rich metadata (clinical parameters, phenotypes)
  - Multi-dimensional datasets

- **New way of browsing the genome**
  - Hundreds of tracks in one view
  - Google-map style pan and zoom, drag and drop
  - Display epigenomic data alongside with their metadata
  - Heatmap, wiggle map, going from whole genome to single base
  - Viewing data on specific genomic features, genesets or pathways
  - Statistical analysis (comparing two or more samples, or groups of samples)
  - Support custom tracks
  - Support sessions
  - ENCODE data integrated

- **Help page, mock data and video tutorial available**
  - http://epigenomegateway.wustl.edu/
Acknowledgement

UCSF
Joseph Costello
Ivan Smirnov
Raman Nagarajan
Chibo Hong
Brett Johnson
Shaun Fouse
Rob Bell
Michael McManus
Thea Tlsty

UC Santa Cruz
David Haussler
Jim Kent
Tracy Ballinger
Donna Karolchik
Bob Kuhn
Ann Zweig
Brian Raney
Kate Rosenbloom

GSC Canada
Marco Marra
Martin Hirst
Steven Jones
Cydney Nielsen
Yongjun Zhao
Gordon Robertson

USC
Peggy Farnham

Washington University
Xin Zhou
Mingchao Xie
Brett Maricque
Eric Martin
Brian Koebbe

EDACC
Aleks Milosavljevic
Alan Harris
Christian Coarfa
Rob Waterland

Other REMCs
John
Stamatoyannopoulos
Richard Sandstrom
Alex Meissner
Brad Bernstein
Bing Ren
Joe Ecker
Noam Shoresh
Bob Thurman

NCBI
Greg Schuler
Tanya Barrett
Case 5: Cancer Methylomes

- Comparing endometrial cancers
Methylation level across gene structure

- Increased DNA methylation over genic regions;
- More methylation in promoters in type I than in type II;
MLH1 Promoter

Scale
chr3: EPM2AIP1

CpG Islands
CpG sites

Endo Normal Medip Density
Endo Normal MRE CpG
Endo 1099 Medip Density
Endo 1099 MRE CpG
Endo 1113 Medip Density
Endo 1113 MRE CpG
Endo 1647 Medip Density
Endo 1647 MRE CpG
Endo 2242 Medip Density
Endo 2242 MRE CpG
Endo 2249 Medip Density
Endo 2249 MRE CpG
Endo 2263 Medip Density
Endo 2263 MRE CpG

Normal

Type II

Type I