Systematic mapping of chromatin state landscapes during mouse development

Dave Gorkin, PhD
Associate Director of Epigenomics
UCSD Center for Epigenomics
dgorkin@ucsd.edu
Mapping the mouse epigenome through development

San Diego Production Group:
Bing Ren
Joe Ecker
Len Pennacchio
Axel Visel
Wei Wang
Mapping the mouse epigenome through development

**Developmental Stages (N=8)**

<table>
<thead>
<tr>
<th>Tissue</th>
<th>E10.5</th>
<th>E11.5</th>
<th>E12.5</th>
<th>E13.5</th>
<th>E14.5</th>
<th>E15.5</th>
<th>E16.5</th>
<th>P0</th>
</tr>
</thead>
<tbody>
<tr>
<td>forebrain (fb)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>midbrain (mb)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hindbrain (hb)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>heart (ht)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>liver (lv)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>intestine (in)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kidney (kd)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lung (lu)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>stomach (st)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>neural tube (nt)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>limb (lm)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>craniofacial</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>prominence (cf)</td>
<td>Not sampled</td>
<td>Not sampled</td>
<td>Not sampled</td>
<td>Not sampled</td>
<td>Not sampled</td>
<td>Not sampled</td>
<td>Not sampled</td>
<td>Not sampled</td>
</tr>
</tbody>
</table>

- 72 distinct tissue • stages
- 2 biological replicates each

**ChIP-seq**
- 8 histone mods per tissue, stage
- 1,128 ChIP-seq assays total

**Available or in progress:**
- Methyl-seq (Ecker)
- ATAC-seq (Ren)
- RNA-seq (Wold)
- scRNA, scATAC, HiC, PLAC-seq (Ren, select tissue-stages)
Several dimensions captured in this data series

- **Tissues (N=12)**
- **Genomic loci**
- **Developmental Stages (N=8)**
- **Histone Mods (N=8)**

**Mark:** ALL shown
**Tissue:** Forebrain
**Stage:** e15.5

**H3K27ac**
- **H3K4me3**
- **H3K4me2**
- **H3K4me1**
- **H3K27ac**
- **H3K9ac**
- **H3K36me3**
- **H3K27me3**
- **H3K9me3**

**H3K27ac/me3**
- **fb**
- **mb**
- **hb**
- **ht**
- **iv**
- **in**
- **kd**
- **lu**
- **st**
- **nt**
- **im**
- **cf**

**Forebrain**

**ALL shown**
Histone mods across tissues & developmental time

Increasing distance in developmental time

Color key
- H3K27ac
- H3K9ac
- H3K4me3
- H3K4me2
- H3K4me1
- H3K27me3
- H3K36me3

Developmental progression
- forebrain
- midbrain
- hindbrain
- neural tube
- heart
- liver
- intestine
- stomach
- lung
- kidney
- limb
- cranioface

Developmental stages separating datasets

Peak strength: replicate 1, Heart, H3K27ac

Pearson Correlation Coefficient

Correlation of peak strength between datasets
From histone mods to chromatin states

→ integrate histone mods into a unified set of chromatin state annotations (chromHMM, Ernst & Kellis, 2012)
Polycomb repression of disease-related TFs

Top 20 enriched GO terms:
- sequence-specific DNA binding TF activity
- nucleic acid binding TF activity
- sequence-specific DNA binding
- DNA binding
- regulatory region DNA binding
- transcription regulatory region DNA binding
- nucleic acid binding
- heterocyclic compound binding
- organic cyclic compound binding
- transcription regulatory region DNA binding
- transcription factor binding
- sequence-specific DNA binding TF activity
- protein binding
- double-stranded DNA binding
- RNA pol2 distal DNA binding TF activity
- RNA pol2 regulatory DNA binding
- RNA pol2 regulatory region DNA binding
- chromatin binding
- DNA binding, binding
Polycomb repression of disease-related TFs

Top 20 enriched GO terms

- sequence-specific DNA binding TF activity
- nucleic acid binding TF activity
- sequence-specific DNA binding
- DNA binding
- regulatory region DNA binding
- transcription regulatory region DNA binding
- nucleic acid binding
- heterocyclic compound binding
- organic cyclic compound binding
- transcription regulatory...DNA binding
- transcription...factor binding
- sequence-specific DNA binding...TF activity
- protein binding
- double-stranded DNA binding
- RNA pol2 distal...DNA binding TF activity
- RNA pol2 regulatory DNA binding
- RNA pol2 regulatory region DNA binding
- chromatin binding
- DNA binding, binding

OMIM = Online Mendelian Inheritance in Man

# tissue-stages repressed

% of genes in set

- all protein coding
- OMIM TFs
- all TFs
- OMIM non-TFs
Polycomb repression of disease-related TFs

Top 20 enriched GO terms
- sequence-specific DNA binding TF activity
- nucleic acid binding TF activity
- sequence-specific DNA binding
- DNA binding
- regulatory region DNA binding
- transcription regulatory region DNA binding
- nucleic acid binding
- heterocyclic compound binding
- organic cyclic compound binding
- transcription regulatory...DNA binding
- transcription factor binding
- sequence-specific DNA binding...TF activity
- protein binding
- double-stranded DNA binding
- RNA pol2 distal...DNA binding TF activity
- RNA pol2 regulatory...DNA binding
- RNA pol2 regulatory region DNA binding
- chromatin binding
- DNA binding, binding

OMIM = Online Mendelian Inheritance in Man

# tissue-stages repressed

% of genes in set

GO enrichment p-value (GREAT binomial test)
Enhancer dynamics reveal regulators of tissue development

- Identified “Dynamic” enhancers in each tissue

“Dynamic” enhancers = significant change in activity (measured by H3K27ac signal) between any 2 adjacent stages within a tissue

(e.g. forebrain, N=10,575)

- MEF2C: Mental retardation, stereotypic movements, epilepsy, and/or cerebral malformations
Predicting enhancer target genes

Topologically Associating Domains (TADs)
Dixon et al., 2012

predicted Ascl1 enhancers

non-target gene used for comparison
nearest non-target gene to enhancer
nearest gene >=10kb from enhancer

fraction of enhancer-gene pairs supported by 3D interactions

distance between enhancer and non-target gene

predicted target gene

* Andrey et al., Gen Res. 2017
Insights into human disease genetics

Dias-Logan syndrome (MIM#617101)
Intellectual developmental disorder with hereditary persistence of fetal hemoglobin

Enhancers linked to:
- OMIM gene
- Other gene
- Other gene, same TAD

GWAS SNPs per bp coverage (x $10^{-5}$)

$p = 7.29e-4$

$p = 4.3e-8$

Bcl11a

Orthologous regions in human

Liver

Forebrain

Orthologous
regions in human

Attention deficit hyperactivity disorder
$p = 7e-6$

Educational attainment
years of education
$p = 4e-12$

Educational attainment
(college completion)
$p = 1e-6$

Fetal hemoglobin level
$p = 7e-35$

Fetal hemoglobin level
$p = 3e-130$

Hemoglobin A2 levels in sickle cell anemia
$p = 3e-16$

Sickle cell anemia
hemolysis
$p = 9e-7$

Hemoglobin E disease
$p = 1e-10$

Hemoglobin
$p = 3e-10$

Fetal hemoglobin level
$p = 4e-53$

Fetal hemoglobin levels in sickle cell anemia
$p = 5e-31$

Fetal hemoglobin
level
$p = 3e-10$

Fetal hemoglobin
level
$p = 4e-53$

Fetal hemoglobin
distribution
$p = 5e-31$
In summary...

- **Large resource data series** on mouse development. Resources include
  - Histone mod ChIP-seq
  - Chromatin State maps
  - Dynamic enhancers with enriched motifs
  - Enhancer target gene predictions
  - ~150 in vivo enhancer validations

- Widespread role for polycomb in repressing TFs outside of normal expression domain
  - Particularly TFs that can lead to disease when mutated

- Dynamic enhancers provide insight into tissue development and master regulators

- Regulatory variants at Mendelian disease genes contribute to common disease

**Data availability**
- ENCODE DCC ([https://www.encodeproject.org/](https://www.encodeproject.org/))
- Biorxiv: 166652 ([https://doi.org/10.1101/166652](https://doi.org/10.1101/166652))
Acknowledgements

**Lawrence Berkeley National Laboratory**
- Dr. Len Pennacchio
- Dr. Axel Visel
- Dr. Diane Dickel
- **Dr. Iros Barozzi***

**Salk Institute**
- Dr. Joe Ecker’s group
- Yupeng He

**ENCODE DCC**
- Dr. J. Michael Cherry
- Dr. Jean Davidson
- Dr. J Seth Stratton
- Dr. Cricket Sloan

**ENCODE DAC**
- Dr. Zhiping Weng
- Dr. Anshul Kundaje

*Co-leads

---

**UCSD**
- Dr. Bing Ren’s group
- **Dr. Yanxiao Zhang***
- Ah Young Lee
- Yuan Zhao
- Sora Chee
- Dr. Sebastian Preissl
- Zhen Ye
- Samantha Kuan

- Dr. Wei Wang’s group
- Dr. Bo Ding
- Dr. Andre Wildberg
- Menchi Wang

**Ren Lab, UCSD**

**UCSD Center for Epigenomics**