Players and Models of Transcription Regulation in 3D genome

Yijun Ruan (yijun.ruan@jax.org)
The Jackson Laboratory for Genomic Medicine
Department of Genetics and Genome Sciences, UConn Health Center

at
ENCODE users meeting, Stanford University
June 10, 2016
Players and Models of Transcription Regulation in 3D genome

The players: DNA, Protein, RNA

Protein-mediated chromatin interactions

RNA-mediated chromatin interactions

3D Nucleome


Rinn and Guttman Science 2014
ChIA-PET for 3D genome mapping
(multiplex datasets)

Inclusive:
• Protein bindings,
• Enriched chromatin interaction,
• Non-enriched contacts (Hi-C like data)

High specificity and resolution:
• Functional element specific,
• Haplotype resolved,
• Single nucleotide resolution

In situ Hi-C vs. CTCF ChIA-PET (GM12878)

Total contact reads = 4.9 Billion

Rao et al 2014 Cell

Total PET reads: 1 Miseq = 6.4 Million; 1 Hiseq = 30.8 Million

Tang et al 2015 Cell
Connecting loops to form **CTCF Contact Domains (CCD)**

- Double knot

Chr:4:109556994-113054287 (3.5Mb)

- CTCF loops
- CCD
- CTCF motifs
- CTCF peaks
- Hi-C TAD
- *in situ* Hi-C loops
- ChromHMM
- RNA-Seq
- Gene

**Freq.**

1000

**Den.**

600

**CCD**

2,267

Boundaries

<table>
<thead>
<tr>
<th>3,852</th>
<th>159</th>
<th>166</th>
</tr>
</thead>
</table>

**Inward motifs**

<table>
<thead>
<tr>
<th>4,177 (97%)</th>
</tr>
</thead>
</table>

**Outward motifs**

<table>
<thead>
<tr>
<th>6</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>63</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>56</th>
</tr>
</thead>
</table>

**Size distribution of loop and domain**

- Indiv. CTCF loop
- *in situ* Hi-C loop
- CCD
- Hi-C TAD

**CCD = TAD**
RNAPII associated chromatin interactions

Extra-genic enhancer-promoter interaction (EP)
Single gene complex

chr2:225000000-228000000

Inter-genic promoter-promoter interaction (PP)
Multi-gene complex

chr2:220000000-220180000

Li et al 2012 Cell
Przemysław Szala, Dariusz Plewczynski

Paweł Trzaskoma, Grzegorz M. Wilczynski,
Common & cell-type specific chromatin looping structures

CTCF

GM12878

R = 0.61

RNAPII

GM12878

R = 0.42

Density

MCF7

HeLa

K562

±2.5K

±2.5K

±2.5K

Common & cell-type specific chromatin looping structures
Common & cell type-specific CCD structure
Genetic variations alter chromatin domains

chr6:31426075-31930740 (504 kb)

Biased SNP

CTCF binding motif
Haplotype-resolved interactions linked to genetics

SNP rs12936231

CTCF ChIA-PET

All loop

All peak

M loop

M peak

P loop

P peak

Biased SNP

Gene

IKZF3

ZPBP2

GSDMB

GSDMA

ORMDL3

M: AGTTACTTACATTAGCCCCCAGATGGAGTGAACCATCAAGTA

P: AGTTACTTACATTAGCCCCCAGATGCAGTGAACCATCAAGTA

High-risk SNPs for asthma and autoimmune disease alter domain-wide transcription of certain genes (Verlaan 2009)

SNP-based validation of CTCF binding and looping

Individual 1

Heterozygous

Phased Loop Peak

Individual 2

Homozygous

Individual 3

Homozygous

D' with low LOD

D' with high LOD

CEU

chr17:37911048-38179492

0.0

0.5

1.0

0

1.0

0.5

D' with high LOD

High-risk SNPs for asthma and autoimmune disease alter domain-wide transcription of certain genes (Verlaan 2009)
3D Genome Structure → Genome Function

3D chromatin architecture

- Topological domains
- CTCF-mediated contact domain

CTCF
RNAPII

Open → Active
Closed → Inactive

Haploype chromatin interaction

Individual 1
Heterozygous functional

Individual 2
Homozygous functional

Individual 3
Homozygous dysfunctional

Genetic variations → Traits/diseases
Players and Models of Transcription Regulation in 3D genome

The players: DNA, Protein, RNA

Protein-mediated chromatin interactions

RNA-mediated chromatin interactions

3D Nucleome


Rinn and Guttman Science 2014
**Genome-wide approach for RNA-chromatin interactions**

**RNA Interaction with Chromatin by Paired End Tag Sequencing, RICH-PET**

A. RNA-chromatin contacts
   - Crosslinking fragmentation
   - DNA linker ligation
   - RNA linker anneal
   - RT to make 1st cDNA
   - DNA linker ligation
   - Proximity ligation
   - 2nd strand cDNA
   - DNA tags
   - PET template
   - Amplification
   - Sequencing
   - PET mapping

**RNA tags**
- chr3R:7047557-7050982
- RICH-PET RNA tag
- RNA-Seq
- snoRNAs
- protein-coding

**DNA tags**
- XIST target distribution in human
- roX2 target distribution in Drosophila

**Meizhen Zheng, Oscar Luo**
**Most ncRNAs target active open chromatin loci**

- DNA peak loci
- RICh-PET DNA tag
- DNA peak loci
- DHS
- ChromHMM
- RNAPII binding peak loci
- RNA-Seq

- H3K27ac
- H3K27me3

- chr2L:16200000-18200000
- chr2L:17300000-175500000

- DHS site
  - n = 5573

- RNAPII peak
  - n = 5770

- RICh-PET DNA peak
  - n = 6599

- Shared
- Not Shared

- 3-way

- Supported
- Unsupported

- Enriched in promoter & enhancer regions
RNAPII ChIP / RICH-PET for detecting transcriptional RNA-chromatin interactions

Pol2-ab

- RNA tag
- DNA tag

Intra-chromosome
Inter-chromosome

Pol2-ab

RNA-Seq (RPKM)
RNAPII RICH-PET (RPKM)

coding
non-coding

chr2R:14474361-14615784

ChromHMM
DNase HS

RICH-PET DNA tag
Peak call

Rep1
Rep2
Rep3

RNAPII ChIP-Seq
RNA-Seq
H3K27ac
H3K27me3

100
100
60
10

6000
400
30
15
100
500
ncRNA contacts enriched at TAD boundary regions

chrX:11600000-13000000 (1.4 Mb)

- RNA-Seq
- Dnase HS
- RNAPII peak
- RNAPII loop
- RNAPII Rich-PET DNA peaks
  - roX2
  - 7SK
  - Hromega
- H3K27ac
- H3K27me3

5kb resolution
Combinatory binding by multiplex protein and RNA factors

chr3R:26198479-26310882

RNA-seq
DHS
RNAPII RICH-PET
DNA tag
7SK
Hsr omega

RNA
factors

Protein
factors
(modENCODE)

H3K27ac
H3K27me3

3819 ncRNA target sites
n=1207
n=1451
n=1023

TF ChIP-Seq (modENCODE)
Modulatory
Constitutive
Others

H3K4me3
H3K4me1
RNA-Seq
snRNA-7SK
Hsr omega
To test the role of ncRNAs, ...
Acknowledgements

Lab members
Meizhen Zheng
Wang Ping
Zhonghui Tang
Oscar Luo
Xingwang Li
JianHua Cao
Danjuan Wang
Emaly Piecuch
Jufen Zhu
Guoliang Li

Collaborators
JAX
CZ Zhang
Xiaoan Ruan
Zhengqing Ouyang
Duygu Ucar
Michael Stitzel
Jeff Chuang
Paul Michalski
Krishna Karuturi
Karolina Palucka
Jacques Banchereau
Ken Paigen
Charles Lee
Ed Liu

UConn
Brent Graveley
Marc Lalande

JGI
Chia-Lin Wei

NIH
Rafael Casellas
John O’Shea
Warrant Leonard

University of Warsaw, Poland
Dariusz Plewczynski
Przemek Szałaj
Grzegorz M. Wilczynski

MNHN, Paris
Laurent Sachs

Roux Family
The Jackson Laboratory
National Institutes of Health
National Cancer Institute
4D Nucleome
The NIH Common Fund