WORKSHOP

Transcriptional circuitry and the regulatory conformation of the genome

Ofir Hakim
Faculty of Life Sciences

Chromosome conformation capture (3C)
Most GR Binding Sites Are Distant From Regulated Genes

chr10:79,970,001-80,260,000
Distribution of GR Binding Sites

8,373 GR Binding Sites

Position of GR relative to TSS

John et al., Nature Genet., 2011
The genome is not randomly organized

**Chromosomes territories**

ES cells: 
- Albumin
- β-actin
- Chromosome 5

Macrophages:

Cell-type specificity

Functional compartmentalization

Active chromatin
DNasel

Expression foci

RNA polymerase II

Gene position is
correlated with gene activity

Beta globin

Eraf

Early

Late

Hutchison and Weintraub,
Cell (1985)

Osborne et al.

Schermelleh et al.
Chromosome Res. (2001)
- Increase resolution
- Increase throughput
- Reduce bias
1. Crosslinking With Formaldehyde

FA% - 0% - 2% we commonly use 1%
Time - commonly 10 minutes
Temperature - commonly 37°C
2. lysis?

In vivo formaldehyde cross-linking: it is time for black box analysis

Gavrilov et al., Brief Funct Genomics, in press
3. Endonuclease digestion

Search the literature for validated enzymes capable digesting in 3C conditions (SDS, Triton…)

6 b.p. cutter
4 b.p. butter ➔ We will discuss resolution issues
4. Ligation

Cross link → Endonuclease digestion → Ligation

Quality control

From Splinter et al., Methods 2012
4. Ligation

Low DTT in ligation buffer is important for the following DNA precipitation

Multiple ligases can work in low DTT concentrations suitable for 3C

Schwartz et al., Biotechniques in press
5. Reverse cross-links

Cross link → Endonuclease digestion → Ligation → Reverse cross-links

This is the 3C DNA library
Chromosome Conformation Capture (3C)

- **Chromosome Conformation Capture (3C)**

**Enhancer**

- **Principle**: Contacts between two defined regions
- **Coverage**: Commonly <1Mb
- **Detection**: Locus-specific PCR
- **Resolution**: High
- **Limitations**: Low throughput and coverage
- **Examples**: Determine interaction between a known promoter and enhancer

Dekker et al., Nature Reviews Genetics 14, 390–403 (2013)
Carbon Copy 3C (5C)

<table>
<thead>
<tr>
<th>Principle</th>
<th>All against all</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coverage</td>
<td>Commonly &lt;1Mb</td>
</tr>
<tr>
<td>Detection</td>
<td>Multiplex PCR, HT-sequencing</td>
</tr>
<tr>
<td>Resolution</td>
<td>High</td>
</tr>
<tr>
<td>Limitations</td>
<td>Limited coverage</td>
</tr>
<tr>
<td>Examples</td>
<td>Determine comprehensively higher-order chromosome structure in a defined region</td>
</tr>
</tbody>
</table>

Multiplex PCR

Limb-specific contacts

Circular 3C (4C)

3C DNA

Second digestion

Ligation

PCR

Sequencing
Circular 3C (4C)

<table>
<thead>
<tr>
<th>Principle</th>
<th>All contacts with a point of interest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coverage</td>
<td>Genome-wide</td>
</tr>
<tr>
<td>Detection</td>
<td>PCR, HT-sequencing</td>
</tr>
<tr>
<td>Resolution</td>
<td>High</td>
</tr>
<tr>
<td>Limitations</td>
<td>Limited to one view point</td>
</tr>
<tr>
<td>Examples</td>
<td>All genes and genomic elements associated with a known LCR</td>
</tr>
</tbody>
</table>
Lipocalin 2 (Lcn2) as a bait for

Hierarchical Contact Landscape

View point

Avital Sarusi
Lipocalin 2 (Lcn2) as a bait for resolution - proximal range 10kb-1000kb.

High contact frequency

6b.p. cutter

4b.p. cutter

Bait
Resolution - far cis (Mb) and trans

Low contact frequency

6b.p. cutter

100k.b.  100k.b.

4b.p. cutter

100k.b.  100k.b.

Bait
Mknk2 Long-Range Contacts
High-Resolution

Contacts 1h after induction

Rachel Deitch, Dana Raz, Moran Tal

Analysis method- van de Werken et al., Nat. Methods 2012
ChIA-PET

- **Principle**: All contacts associated with a given protein
- **Coverage**: Genome wide *
- **Detection**: paired end HT-sequencing
- **Resolution**: High
- **Limitations**: Rely on one factor, disregarding other contacts
- **Examples**: Map chromatin interaction network of a known transcription factor

Keifer-Kwon et al., 2013
Capture-C

<table>
<thead>
<tr>
<th>Principle</th>
<th>Focused all against all</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coverage</td>
<td>Up to genome-wide</td>
</tr>
<tr>
<td>Detection</td>
<td>HT-sequencing</td>
</tr>
<tr>
<td>Resolution</td>
<td>High</td>
</tr>
<tr>
<td>Limitations</td>
<td>Limited coverage</td>
</tr>
<tr>
<td>Examples</td>
<td>Determine comprehensively higher-order chromosome structure in multiple defined regions</td>
</tr>
</tbody>
</table>

Sub-sampling of Hi-C

Hughes et al., 2014
Circular 3C (4C)
What does a C signal mean?

200-500 nuclei

DNA FISH

R² = 0.89

Interaction frequency (%) by DNA FISH

d = <0.5 micron

Correlation with 4C signal

3D distance threshold (microns)
Hi-C

<table>
<thead>
<tr>
<th>Principle</th>
<th>All against all</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coverage</td>
<td>Genome-wide *</td>
</tr>
<tr>
<td>Detection</td>
<td>Paired end HT-sequencing</td>
</tr>
<tr>
<td>Resolution</td>
<td>Low *</td>
</tr>
<tr>
<td>Limitations</td>
<td></td>
</tr>
<tr>
<td>Examples</td>
<td>All intra- and inter- chromosomal associations</td>
</tr>
</tbody>
</table>

Dekker et al., Nature Reviews Genetics 14, 390–403 (2013)
Resolution is critical

~ 20 million read pairs ~ 1 billion read pairs

Liebermann-Aiden et al., 2009  Dixon et al., 2012
Topological associated domain (TAD)

Negabase-sized local chromatin interaction domains

Dixon et al., 2012
Domains within domains

Khalor et al., 2011
Chromosomes are organized in territories

Nature Reviews Genetics 6, 429
Liebermann-Aiden et al., 2009
Preferential contacts within and between chromosomes

There are two compartments: A and B

Are there sub-compartment structures?

Liebermann-Aiden et al., 2009
Genomic associations:
Gene density, Gene activity

Compartment A is gene rich

Compartment A
Genes Spearman’s $\rho = 0.431$
Expression Spearman’s $\rho = 0.476$
Accessible chromatin, Spearman’s $\rho = 0.651$
H3K36 trimethylation, Spearman’s $\rho = 0.601$ (active)
H3K27 trimethylation, Spearman’s $\rho = 0.282$ (repressive)
A is more closely associated with open, accessible, actively transcribed chromatin.

Liebermann-Aiden et al., 2009
Shopland, et al 2006
Contacts between TADs are cell-type specific

**Ifng 4C Th1**

**Ifng 4C Th2**

**Ifng 4C Th1**

**Ifng 4C Th2**

**Ifng 4C Th1**

**Ifng 4C Th2**

**Ifng 4C Th1**

**Ifng 4C Th2**

**Furin**

**Il2rb**

**ANKRD4**

**Traf3ip2**

**Socs2**

% of genes

Response to bacterium

Immune response

Ifng contacts

Global
Contacts between TADs are cell-type specific

High enrichment for adipogenic genes

Avital Sarusi