## **Analyzing 3-D genome organization**

#### Martin Aryee







http://www.aidenlab.org

Massachusetts General Hospital Harvard T.H. Chan School of Public Health Broad Institute



Short region of DNA double helix



"Beads on a string" form of chromatin



30-nm chromatin fibre of packed nucleosomes



Section of chromosome in an extended form



Condensed section of chromosome





Entire mitotic chromosome

Felsenfeld and Groudine. Nature, 2003

#### Chromosomes exhibit non-random spatial organization



Dekker and Misteli, Cold Spring Harb Perspect Biol (2015)

#### Gene repositioning is associated with gene activation

#### Silenced gene



Active, expressed gene

#### Green = anti-lamin B Red = IgH locus

Kosak et al., Science (2002)

## Enhancer - promoter looping



© 2010 Nature Education

http://www.nature.com/scitable/topicpage/gene-expression-14121669

# Disruption of genome topology in cancer



"Structural" vs. "Functional" loops

Flavahan and Drier et al., Nature (2015) Image from Grimmer & Costello, Nature (2015)

## Genome topology assays

a 3C: converting chromatin interactions into ligation products



| 3C                       | 4C                        | 5C                         |
|--------------------------|---------------------------|----------------------------|
| One-by-one<br>All-by-all | One-by-all                | Many-by-many               |
|                          | 4)                        | <del>f</del>               |
| PCR or sequencing        | Inverse PCR<br>sequencing | Multiplexed LMA sequencing |

~200M-500M Reads

Nature Reviews | Genetics

• Key metric for both Hi-C and ChIA-PET: Fraction of paired-end read ("PETs") supporting long-range interactions



https://www.bioinformatics.babraham.ac.uk/projects/hicup/scripts\_description/

- Poor Hi-C library: <20% long-range (>20kb) pairs
- Good Hi-C library: ~40% long-range (>20kb) pairs
- A shallow sequencing run (~5M reads) is sufficient to assess library quality

## QC: Hi-C

#### HiC data: From read pairs to an interaction matrix



Chromosome A

Adapted from Lun et al., BMC Bioinformatics, 2015, Volume 16, Number 1, Page 1

| Resolution<br>(bin size) | Number of<br>matrix cells |  |  |  |  |
|--------------------------|---------------------------|--|--|--|--|
| 1 Mbase                  | ~10M                      |  |  |  |  |
| 10 kbase                 | ~100B                     |  |  |  |  |

#### HiC data: From read pairs to an interaction matrix



From Ulianov, S.V., Gavrilov, A.A., Razin, S.V., 2015. Nuclear Compartments, Genome Folding, and Enhancer-Promoter Communication. In: Jeon, K.W. (Ed.), International Review of Cell and Molecular Biology, pp. 183-244. ISBN: 9780128022825 Copyright © 2015 Elsevier Inc. All rights reserved. Academic Press



Dixon, J. R., Gorkin, D. U. & Ren, B. Chromatin Domains: The Unit of Chromosome Organization. Molecular Cell 62, 668–680 (2016).





## 3D genome organization at multiple scales

Α

| Chromosor                           | 00,000 -                    | 1<br>Scale (kp) |
|-------------------------------------|-----------------------------|-----------------|
|                                     | 10,000 -                    |                 |
| Topologica<br>Associatir<br>Domains | 1000 -                      |                 |
| Gene loop                           | 100 <b>-</b><br>10 <b>-</b> |                 |
|                                     | 1 -                         |                 |
| Nucleosom                           | •                           |                 |



Dowen, J. M. et al. Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. - PubMed - NCBI. Cell 159, 374–387 (2014).



## HiChIP / ChIA-PET data: From read pairs to an interaction matrix



#### **Processing steps:**

- 1. Identify loop anchors (i.e. ChIP peaks)
- 2. Anchor pairs joined by PETs are putative loops
- 3. Count PETs (paired-end reads) per putative loop
- 4. Determine loop significance (i.e. is the number of PETs higher than the background?)



PET = Paired End Tag



## How long are chromatin loops?



#### **Distance between PETs**

#### Hi-C All pairwise interactions



Chr 1

Adapted from: Dixon, J. R., Gorkin, D. U. & Ren, B. Chromatin Domains: The Unit of Chromosome Organization. Molecular Cell 62, 668–680 (2016).

#### **ChIA-PET** Pairwise interactions between loop anchor loci





#### HiC

#### ChIA-PET: CTCF loops

ChIA-PET: POL2 loops

http://dnalandscaper.aryeelab.org

## Hi-C and ChIA-PET are complementary





#### Differential topology analysis: POL2 ChIA-PET

K562







- RP11-661B11.3





Legend

enhancer - promoter

CTCF - CTCF

#### http://dnalandscaper.aryeelab.org

## Pol2 ChIA-PET loops



Principal Component 2



#### **PCA of ChIA-PET Counts Matrix**

PC1

**ENCODE** Project Data





#### Sample 2

#### Need to define a common anchor set:







#### Anchors



Home » Bioconductor 3.4 » Software Packages » diffloop

diffloop

http://bioconductor.org/packages/diffloop/

#### Differential loops: diffloop

Loops

#### PET counts

#### PET = paired end tag



### Genome-wide analysis of differential looping

- Genome-wide identification of POL2 ChIA-PET loops in MCF7 and K562
- At an FDR of 1%, there are > 2,600 differential loops
- Nearly 2,000 were classified as enhancer-promoter loops (another 500 p-p)
- Target genes in MCF7 enriched for Estrogen Response Pathways

## POL2 loop strength correlates with expression



Gene expression

POL2 ChIA-PET

#### Epigenetic correlates of enhancer-promoter looping: DNA Methylation



POL2 ChIA-PET

#### Epigenetic correlates of enhancer-promoter looping: H3K27Ac



#### <sup>7</sup>ac log fold change (K562/MCF7)

POL2 ChIA-PET

•

#### HiChIP: efficient and sensitive analysis of proteindirected genome architecture

Maxwell R Mumbach, Adam J Rubin, Ryan A Flynn, Chao Dai, Paul A Khavari, William J Greenleaf & Howard Y Chang

Affiliations | Contributions | Corresponding author

Nature Methods 13, 919–922 (2016) | doi:10.1038/nmeth.3999 Received 02 May 2016 | Accepted 10 August 2016 | Published online 19 September 2016



< 🔒

## HiChIP

- Protocol allows for < 1 million cells</li>
- Higher read efficiency than ChIA-PET
- Shorter protocol

# aryeelab / hichipper Code Issues 1 Pull requests 0 Project Branch: master - hichipper / README.md caleblareau added varying parameter comparison contributors 2 contributors 2 2 contr

#### hichipper

This package is maintained by Caleb Lareau in the install version is provided through PyPi.



HiChI

|           |             |             | O Unwa   | atch <del>-</del> | 3      | ★ Sta   | ar O    | ş    | Fork    | 0   |
|-----------|-------------|-------------|----------|-------------------|--------|---------|---------|------|---------|-----|
| ects O    | 💷 Wiki      | Pulse       | III Grap | ohs               | 🗘 Set  | tings   |         |      |         |     |
|           |             |             |          |                   |        |         | Find f  | ile  | Сору р  | ath |
|           |             |             |          |                   |        | 1       | 18fb7fe | on D | ec 4, 2 | 016 |
|           |             |             |          |                   |        |         |         |      | , -     |     |
|           |             |             |          |                   |        |         |         |      |         |     |
|           |             |             |          | Raw               | Blam   | e Hi    | story   | Ţ    |         | Ī   |
|           |             |             |          |                   |        |         |         |      |         |     |
| Aryee Lab | . Source co | ode is made | freely a | ivailabl          | e here | e and a | a pacl  | kage | d       |     |
|           | Pe          | ak          |          |                   |        |         |         |      |         |     |
|           |             |             |          |                   |        |         |         |      |         |     |
| ))))      |             |             |          | X                 | X      | 0(      | XĮ      | (    |         |     |
|           | HiChIP      | Anchor      |          |                   |        |         |         |      |         |     |
|           |             | 000         | 00       | 0                 |        |         |         | X    |         |     |
| P ANCNO   | Ľ           |             |          |                   |        |         |         |      |         |     |

## Acknowledgements

## Caleb Lareau Brac Zack McCaw Sarah Jose Malagon Lopez Esm

Brad Bernstein

#### Rafael Irizarry

Sarah Johnstone

#### Alejandro Reyes

Esmat Hegazi