Protein determinants of chromosome domains

Laura Braverman

MIT PRIMES Presentation
May 22, 2016
Background
Hi-C

Loops and domains

Chromosome contact maps

High

Low
Loops and domains

2009  2012  2015

A/B Compartments

Inactive

Active

Loops and domains

DNA - Binding Proteins

Functional aspect

Rao et al/Cell 2014
Binding factors at boundaries

- Architectural Proteins
Binding factors at boundaries

- Architectural Proteins
- Transcription Factors
Chromatin Immunoprecipitation Seq (ChIP-seq)
Chromatin Immunoprecipitation Seq (ChIP-seq)

Cell 1

Cell 2

Cell 3

Position 1

Position 2
Methods and Results
Protein Enrichment

Peak enrichment vs percentage states occupied

Rao et. al. 2014
Percentage of loop or domain boundaries with protein
Percentage of loop or domain boundaries with protein

Example:
- 10 total boundaries
- 4 with green protein
- $\frac{4}{10} = 40\%$ green protein
Enrichment of proteins at loop and domain boundaries

Protein Bonding Enrichment At Loop Bases

CTCF

Outside Loop

Inside Loop

Distance from boundary

Outside

Inside

Boundary
Architectural and transcription proteins are enriched at loop boundaries

- **All domains**
  - Average fold enrichment vs. Percent of domains bound by protein (%)
  - Proteins: RNA Polymerase II, ZNF143, SMC3, RAD21, YY1, CTCF

- **All loops**
  - Average fold enrichment vs. Percent of domains bound by protein (%)
  - Proteins: RNA Polymerase II, ZNF143, SMC3, RAD21, YY1, CTCF
Architectural proteins enriched at loop boundaries

CTCF
- Known to be related to 3D genome structure and loop formation

Cohesin
- RAD21 and SMC3 subunits
- Forms chromatin loops in Interphase
Considering boundaries with a particular protein

- Only interested in boundaries with a particular protein (green)
Considering boundaries with a particular protein

- Only interested in boundaries with a particular protein (green)

- 2 boundaries with both CTCF and the blue protein/ 4 boundaries with CTCF
Architectural proteins CTCF, SMC3, RAD21 are co-enriched at loop boundaries
RNA Polymerase II is more common on boundaries without CTCF than with CTCF
ZNF143 is more co-enriched with CTCF than YY1
YY1 and RNA Polymerase II are correlated

With YY1

Without YY1
YY1 bound loci are bound by CTCF with a similar frequency but less enrichment.
RNA Polymerase II bound loci are bound by CTCF with a similar frequency but less enrichment.
Testable hypotheses: CTCF could be replaced with RNA Polymerase II after loops are created.

Cell 1
Cell 2
Cell 3

Position 1
Position 1 later
Conclusion

- Hypothesized classifications:
  - ZNF143 = Architectural protein
  - YY1 = Transcription / RNA Polymerase II associated boundary factor
- Enrichment of RNA Polymerase II is anticorrelated with CTCF
- Testable hypothesis: RNA Polymerase II replaces CTCF at boundaries?
Acknowledgements

Mirny lab:

Prof. Leonid Mirny
Nezar Abdennur
Boryana Doyle
Dr. Geoff Fudenberg
Dr. Maxim Imakaev
Anton Goloborodko
The Espresso Machine

Dr. Slava Gerovitch and
Dr. Pavel Etingof for
dealing with our poor
adherence to deadlines

And my parents for
supporting me through
this journey