

Exploring polymer models to explain the appearance of TADs

Ofir Shukron

Ecole Normale Supérieure

July 22, 2014

1 The experimental setting

2 The data

3 Analysis of the data

- TAD D and E
- Peaks of the encounter data

4 Theoretical model

5 Simulations

- Simulations with a simple rouse chain
- Loops corresponding to the peaks of the encounter data
- Dynamic loops model
- Dynamic loops model with beads' affinity
- Enlarging the encounter distance
- 3C experiment with stiff connectors

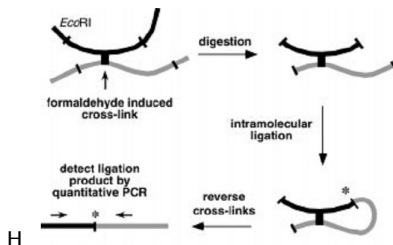
6 Future Perspectives

Chromosome Conformation Capture (3C) and its relatives

A method to simultaneously record millions of looping events occurring within the genome. The 4C/5C/HiC are all derived from the 3C method and uses a 3C segment library for subsequent steps.

The general 3C steps are:

- 1 intact nuclei are extracted from millions of cells
- 2 Formaldehyde induces protein-DNA and protein-protein cross-links
- 3 restriction enzymes digest the cross-linked DNA
- 4 DNA is diluted and ligated
- 5 cross-links are reversed
- 6 PCR to amplify ligation junctions
- 7 histogram of segment encounter is produced



The data

- 5C experiments were conducted by Nora et. al 2012.
- A subset of the data, showing X-chromosome self interactions in mouse embryonic stem cells, will be analyzed here.
- The focus is on regions harboring the Xist enhancer and Tisx promoter, related to X inactivation mechanism.
- A total length of 920,432 bp.
- Two replicates of the experiment were made.
- we have the encounter matrix of different segments.

Topologically Associating Domains (TADs)

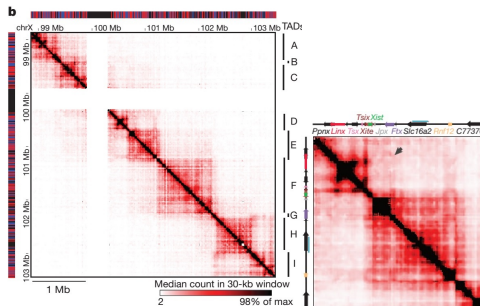
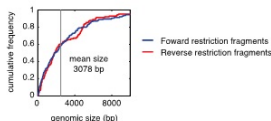


Figure: The 4.5 Mb region analyzed by Nora et al, and TAD D and E regions. Displayed median count in a 30kb window every 6kb

From restriction segments to beads

- To coarse-grain the data, Luca et. al. chose a bead-length of 3000 bp, corresponding to the mean restriction segment length of EcoRIII enzyme.



- the genomic section was evenly partitioned by 3000 bp beads. Each segment receives a start and end index according to the beads it covers.

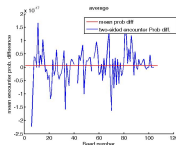
- for example,

bp range	start ind	end ind
500-3500	1	2
4000-4500	2	2
5000-12001	2	4

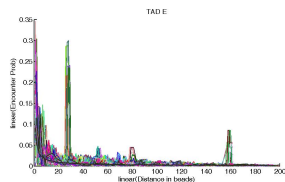
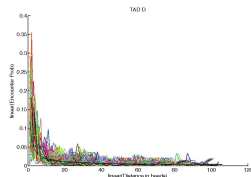
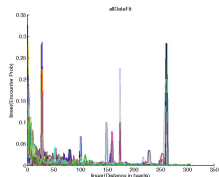
Analysis of the data

TAD D and E

- A total length of 920,432 bp - resulting in 307 beads
- We calculate the 'one-sided' encounter probability vs distance (bead units) for each bead
- the mean encounter probability difference, shows that the data is left-right symmetric



- the encounter data showed



- TAD E has several strong specific interactions.
- TAD D has almost no specific interactions.

Peaks of the encounter data

- About half of the peaks in the encounter data result from specific interactions **between TADs**
- The other half comes from specific internal interactions of **TAD E**.
- To get an impression, a manual marking of the peaks shows

Bead numbers	Encountered beads	TAD
23-26	280-290	$D \leftrightarrow E$
49-53	148-155	$D \leftrightarrow E$
56-59	80-90	$D \leftrightarrow D$
115-117	165-170	$E \leftrightarrow E$
161-162	187-190	$E \leftrightarrow E$
182-184	260-264	$E \leftrightarrow E$
185-186	253-255	$E \leftrightarrow E$
234-236	184-189	$E \leftrightarrow E$
234-236	4-11	$E \leftrightarrow D$
243	88	$E \leftrightarrow D$
264	89-90	$E \leftrightarrow D$
274-277	113-120	$E \leftrightarrow E$

Theoretical models

The Rouse model

We start with the classical and most simple model, the Rouse chain.

- A Rouse chain describes polymer dynamics as a stochastic motion of a collection of microscopic "beads" connected by harmonic springs
- the 3D motion of bead n in the chain of N beads

$$\frac{dR_n}{dt} = -\frac{3D}{b^2}(2R_n(t) - R_{n+1}(t) - R_{n-1}(t)) + f_n(t)$$

- R_n - the position of bead n
 b - the standard deviation of the distance between adjacent beads
 D - the diffusion constant
 f_n - white Gaussian noise
- From the theory, $Pr(\|R_n - R_m\| < \epsilon) \sim |n - m|^{-1.5}$

Simulation with simple rouse chain

- we first check whether a simple model can produce the TADs.
- we examine the results of simulating chain of 64 beads

Loops corresponding to the peaks of the encounter data

Dynamic Loop Model

- some beads in the same TAD have affinity toward one another
- affine beads located within a distance less than ϵ (the encounter distance) are connected
- the rate of disconnection between beads is k_{off}

Dynamic loops model with beads affinity

Enlarging the encounter distance

3C experiment with stiff connectors

Next, we simulate 64 bead chains with stiff connectors.
Stiff connectors are Rouse spring that stay fixed.

Future perspective

- A model with variable encounter distance
-