

3DAROC16

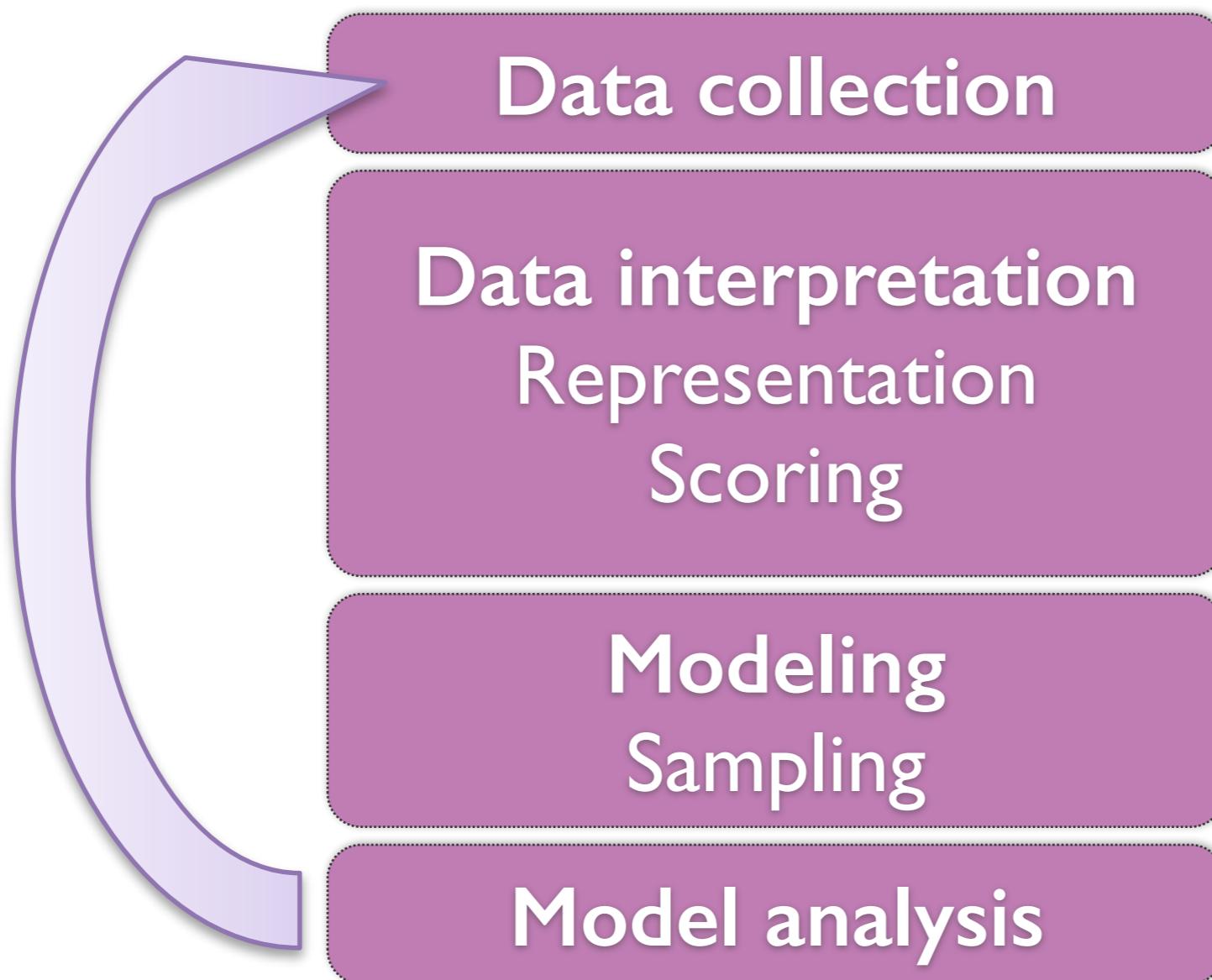
Summary day #2

**Marco Di Stefano, François Serra &
Marc A. Marti-Renom**
Structural Genomics Group (CNAG-CRG)

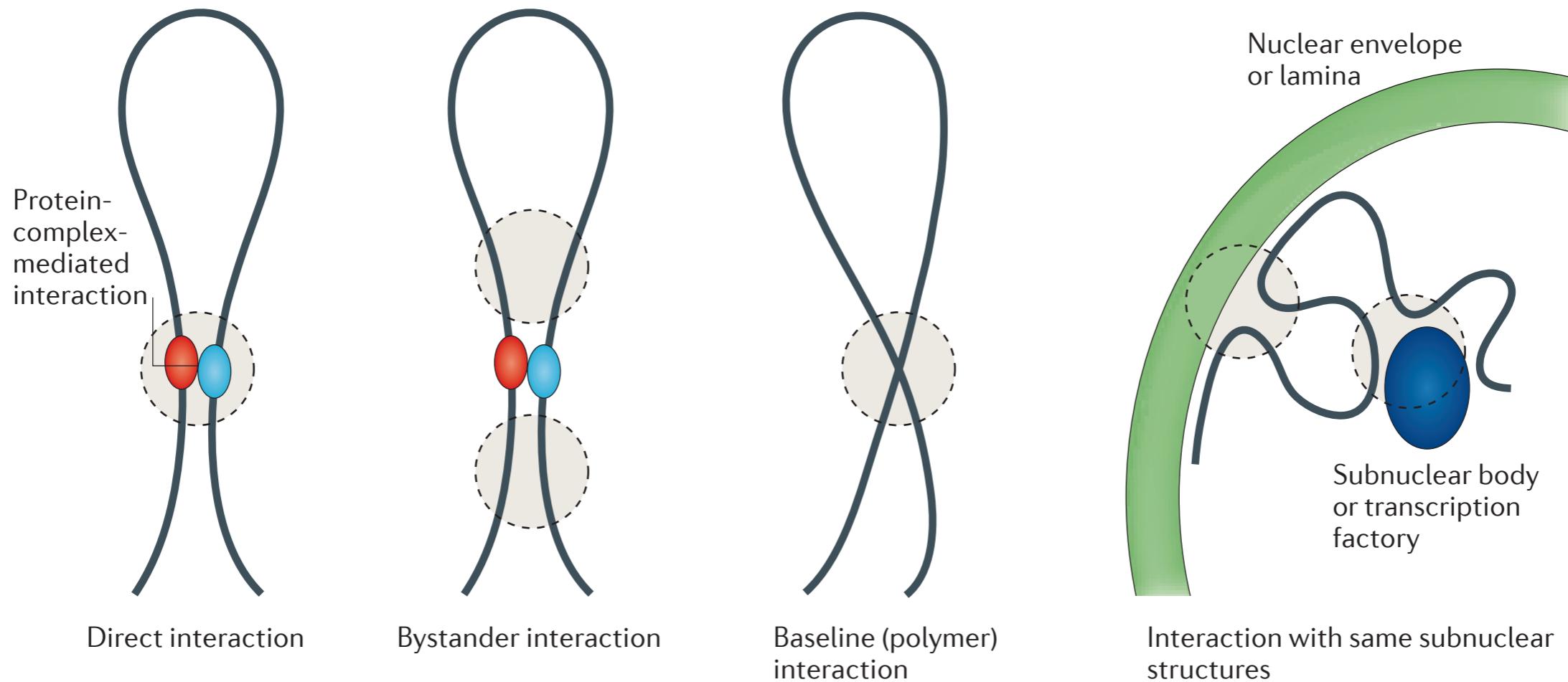


Day 1

Integrative modeling

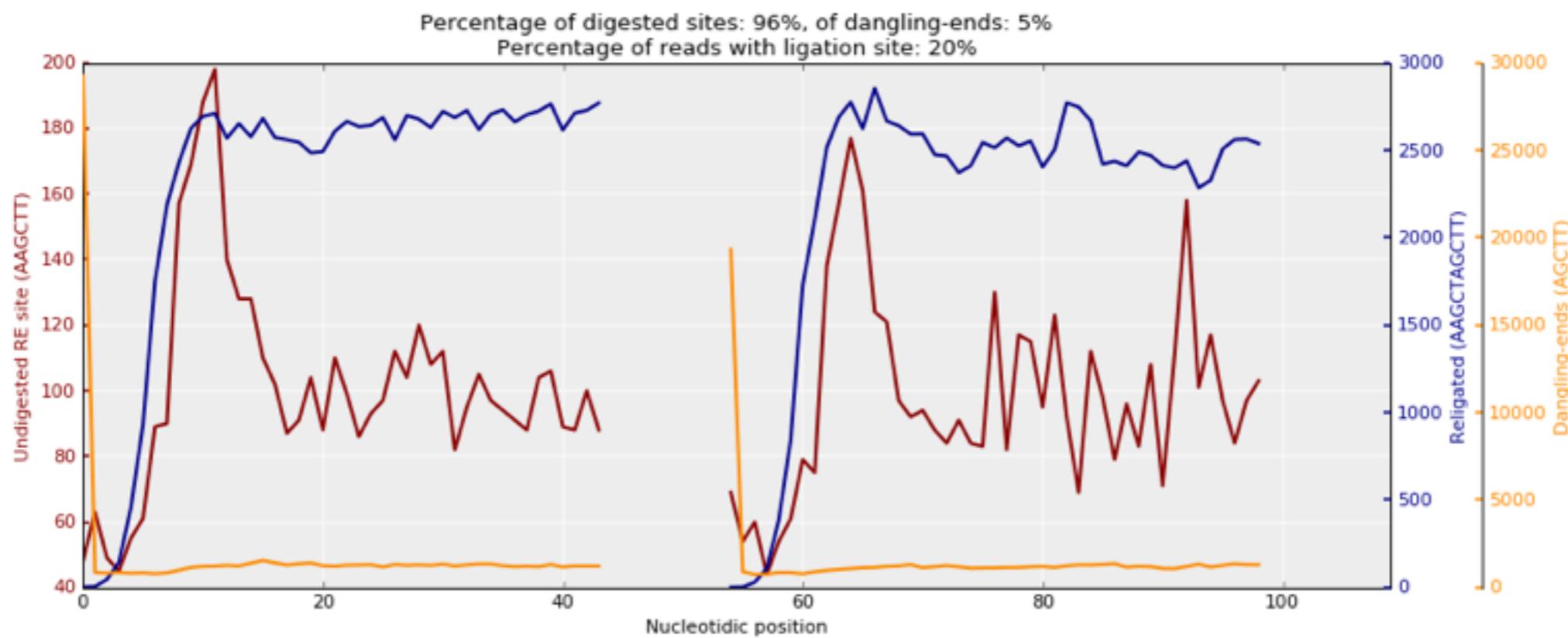
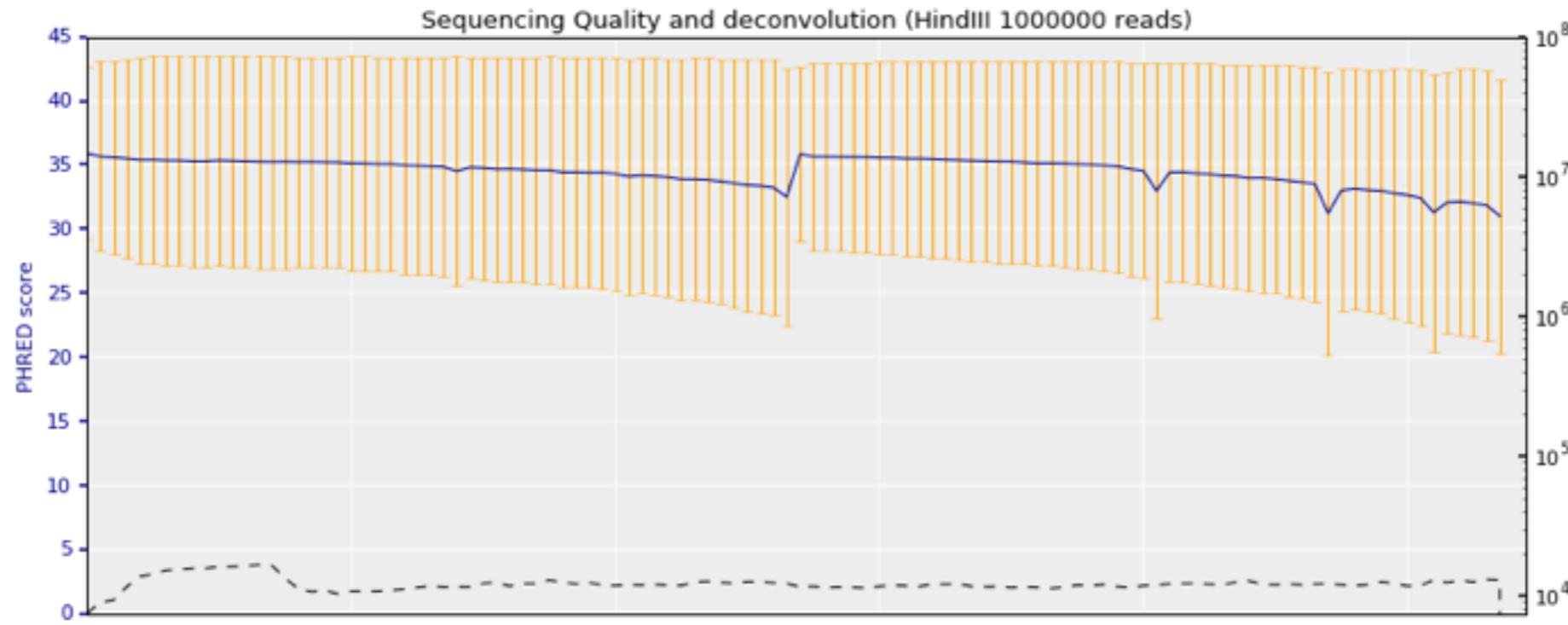


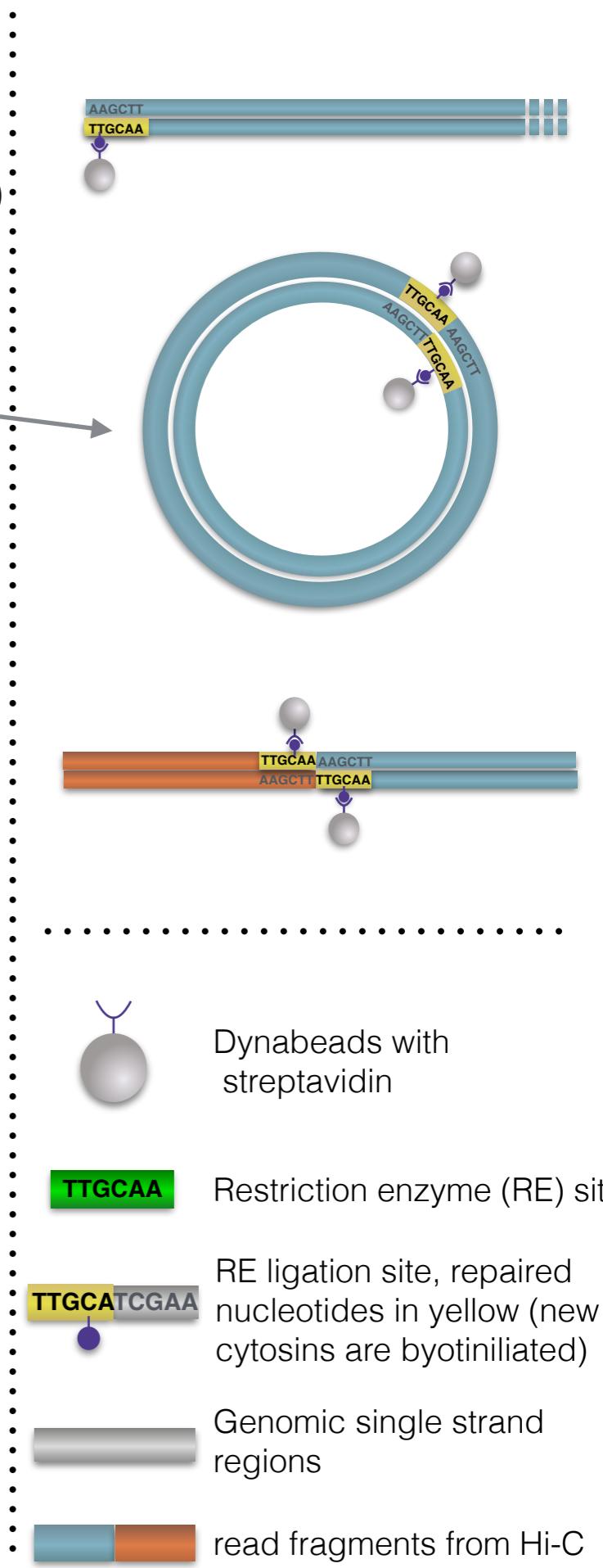
Interpreting chromatin interaction data



Adapted from Dekker et al., (2013) Nat Rev Genetics

Quality plots of the reads





Question on Day1

Raise your hand and score from **1** (very negative) to **5** (very positive):

1 - Would you feel comfortable to explain the general idea of integrative modelling for structural determination to your colleagues in the lab?

Question on Day1

2 - Suppose that someone gives you his (precious & secret) Hi-C fastq files to analyse. Would you feel confident in assessing their quality?

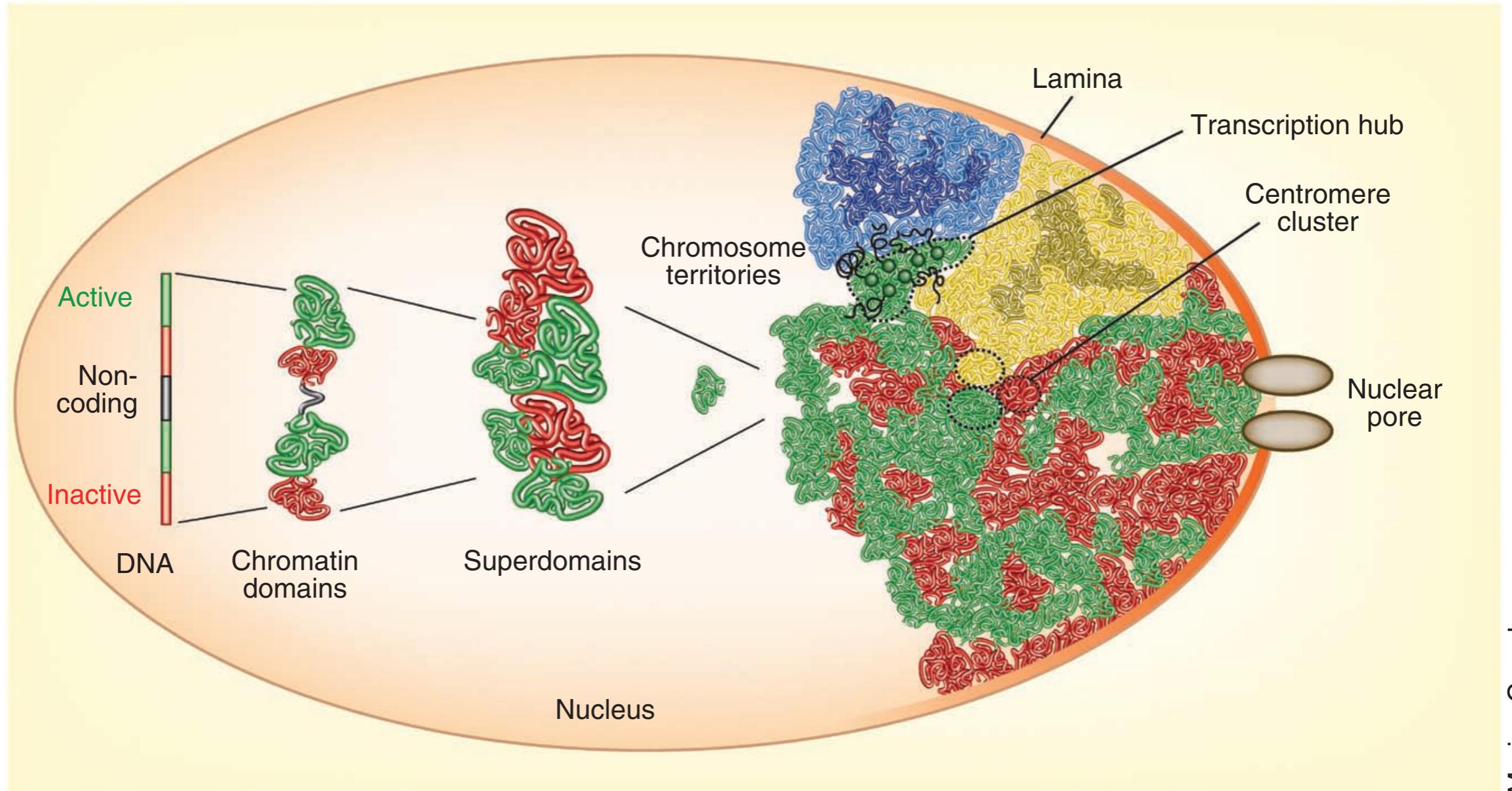
Question on Day1

3 - Would you be confident in the mapping you could produce?

Day 2

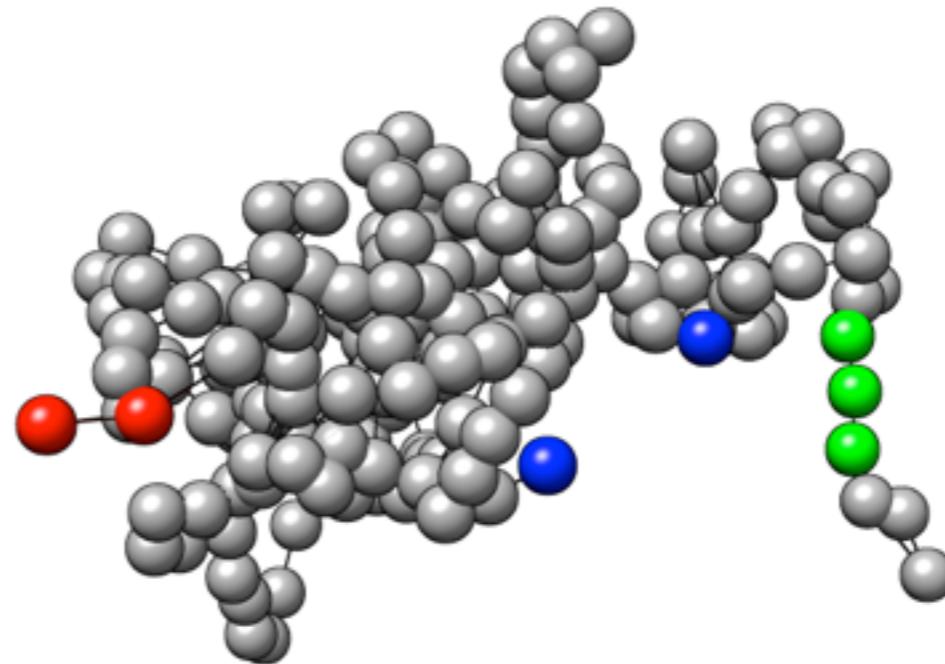
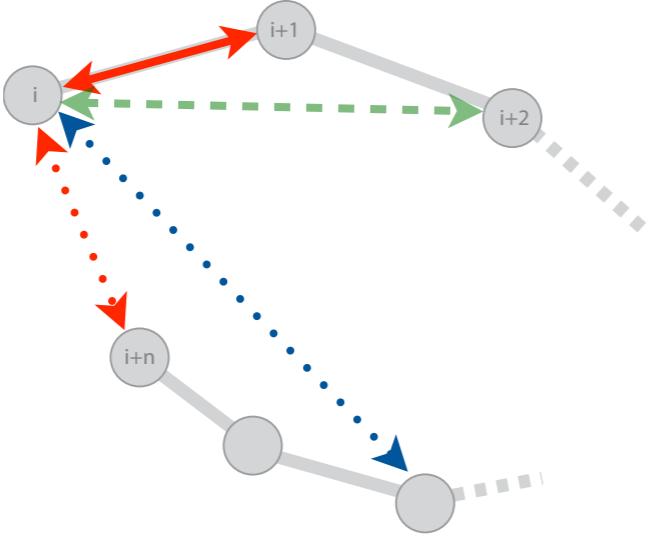
Complex genome organization

Cavalli, G. & Misteli, T. Functional implications of genome topology. *Nat Struct Mol Biol* 20, 290–299 (2013).



Model representation and scoring

Constituent parts of the molecule



$$d < d_0$$



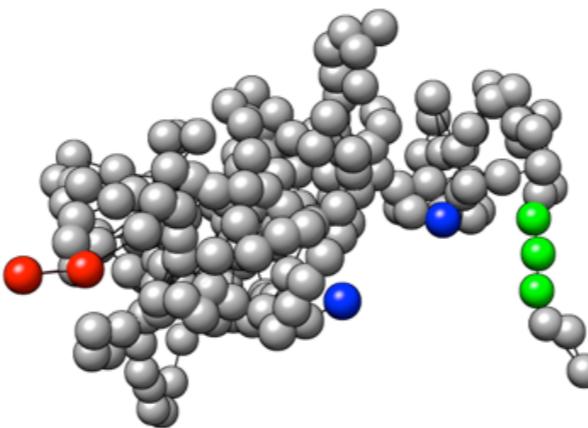
$$d = d_0$$



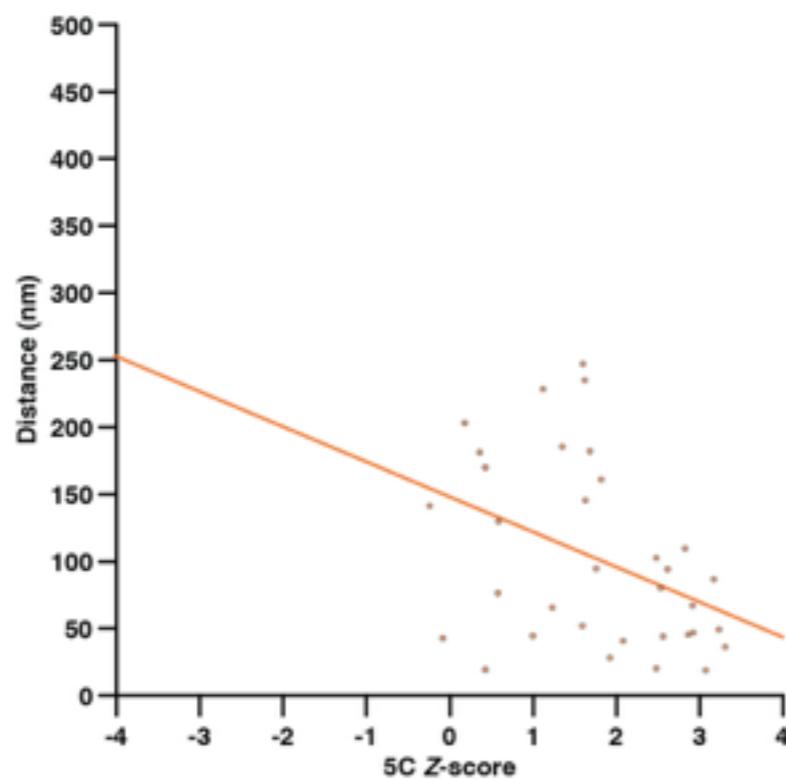
$$d > d_0$$



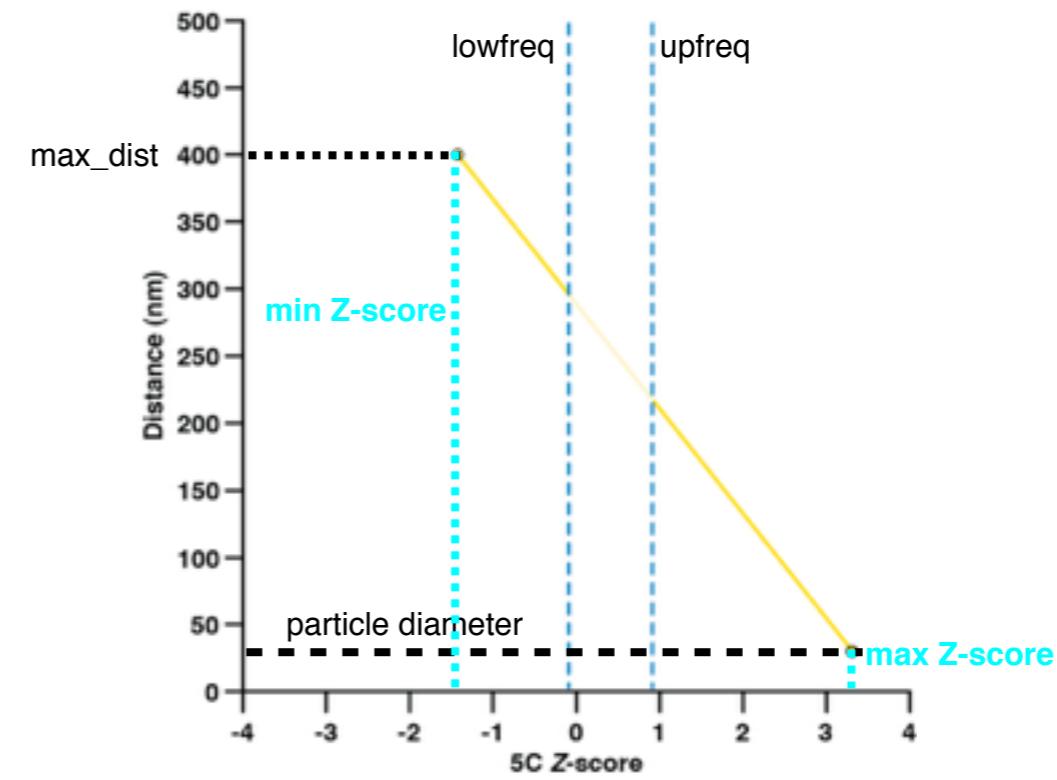
From 5C data to spatial distances



Neighbor fragments



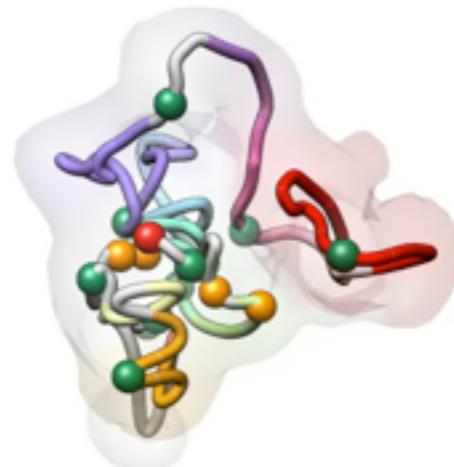
Non-Neighbor fragments



FISH validation

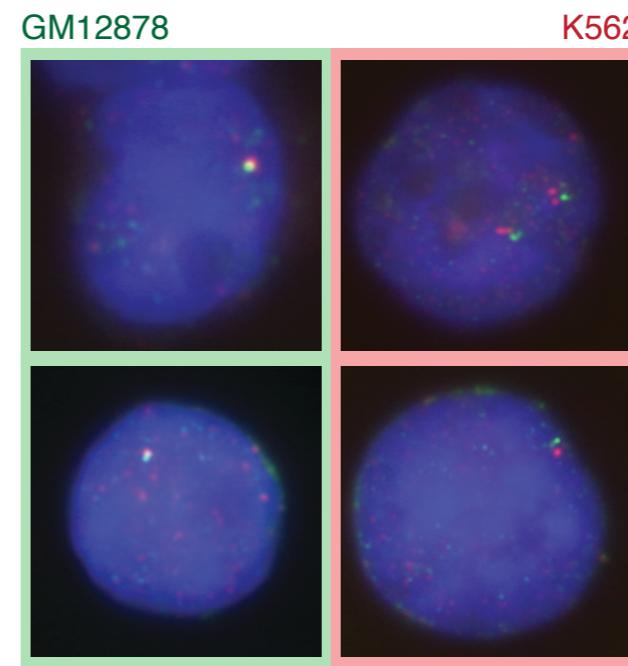
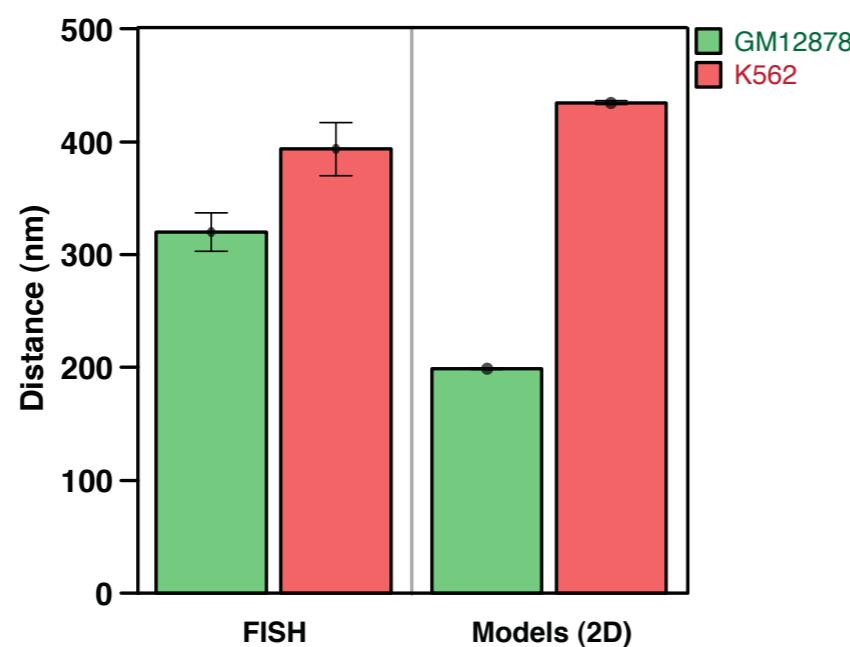
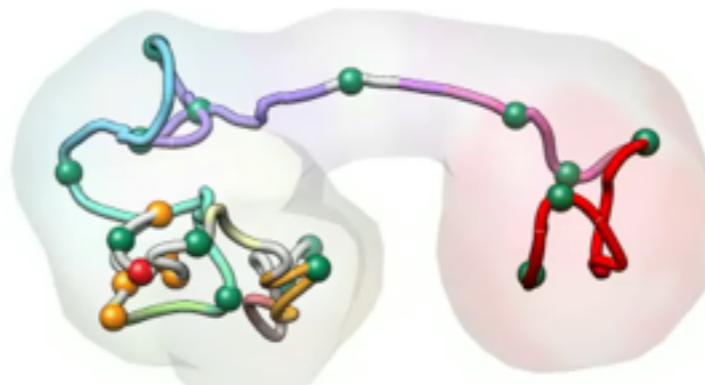
GM12878

Cluster #1
2780 model



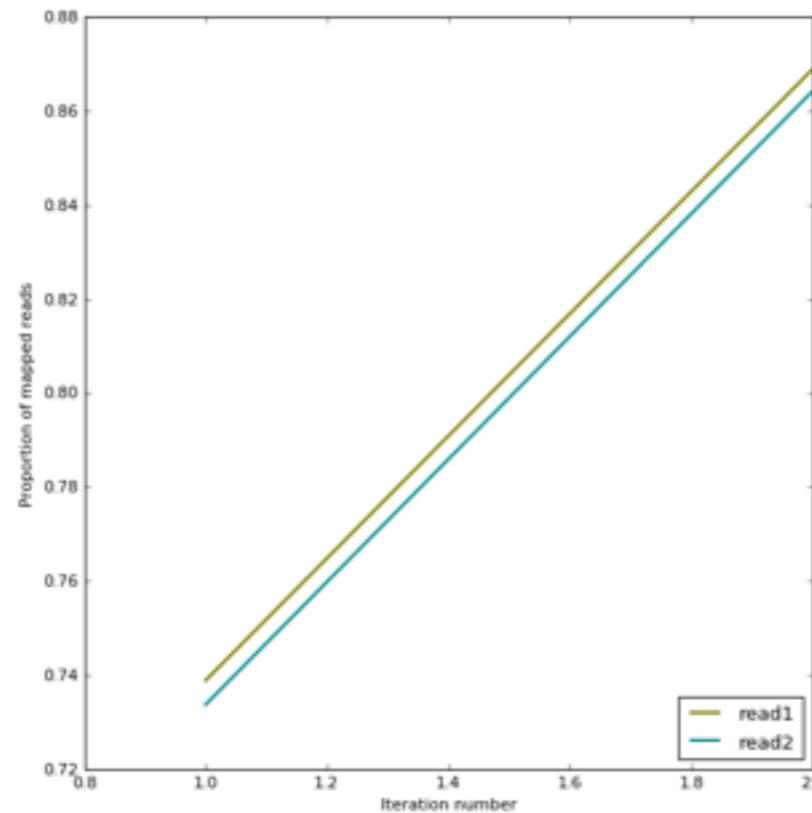
K562

Cluster #2
314 model

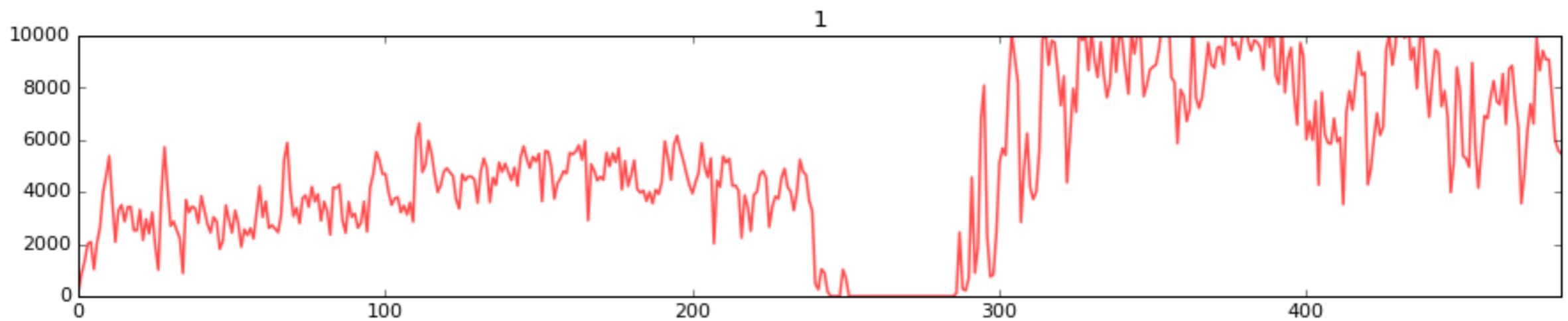
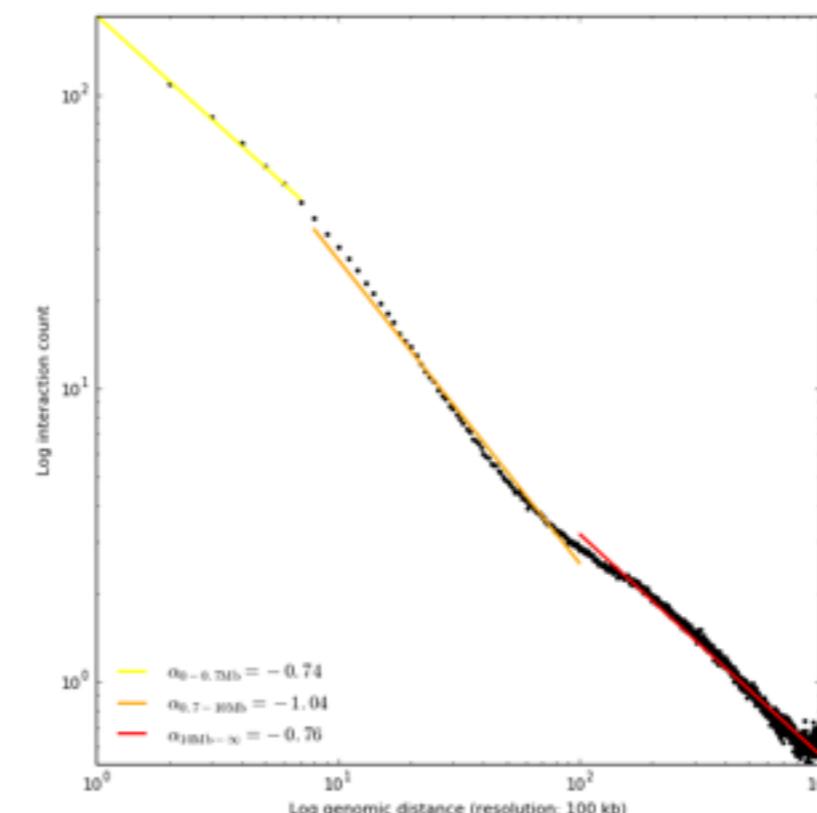


Data quality check

Mapped reads vs full/frag



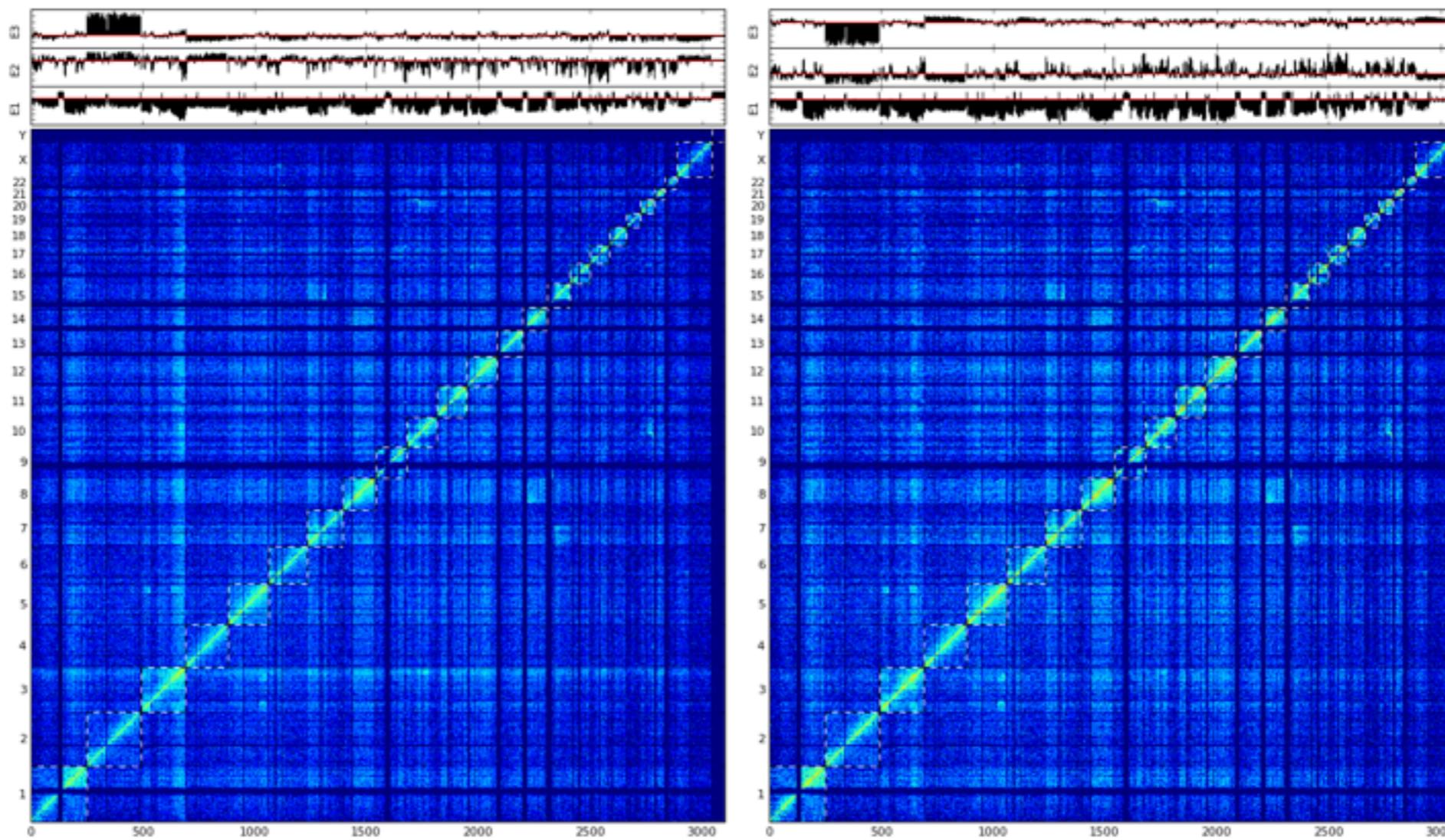
Interactions vs genomic distance



Hi-C map generation and filtering

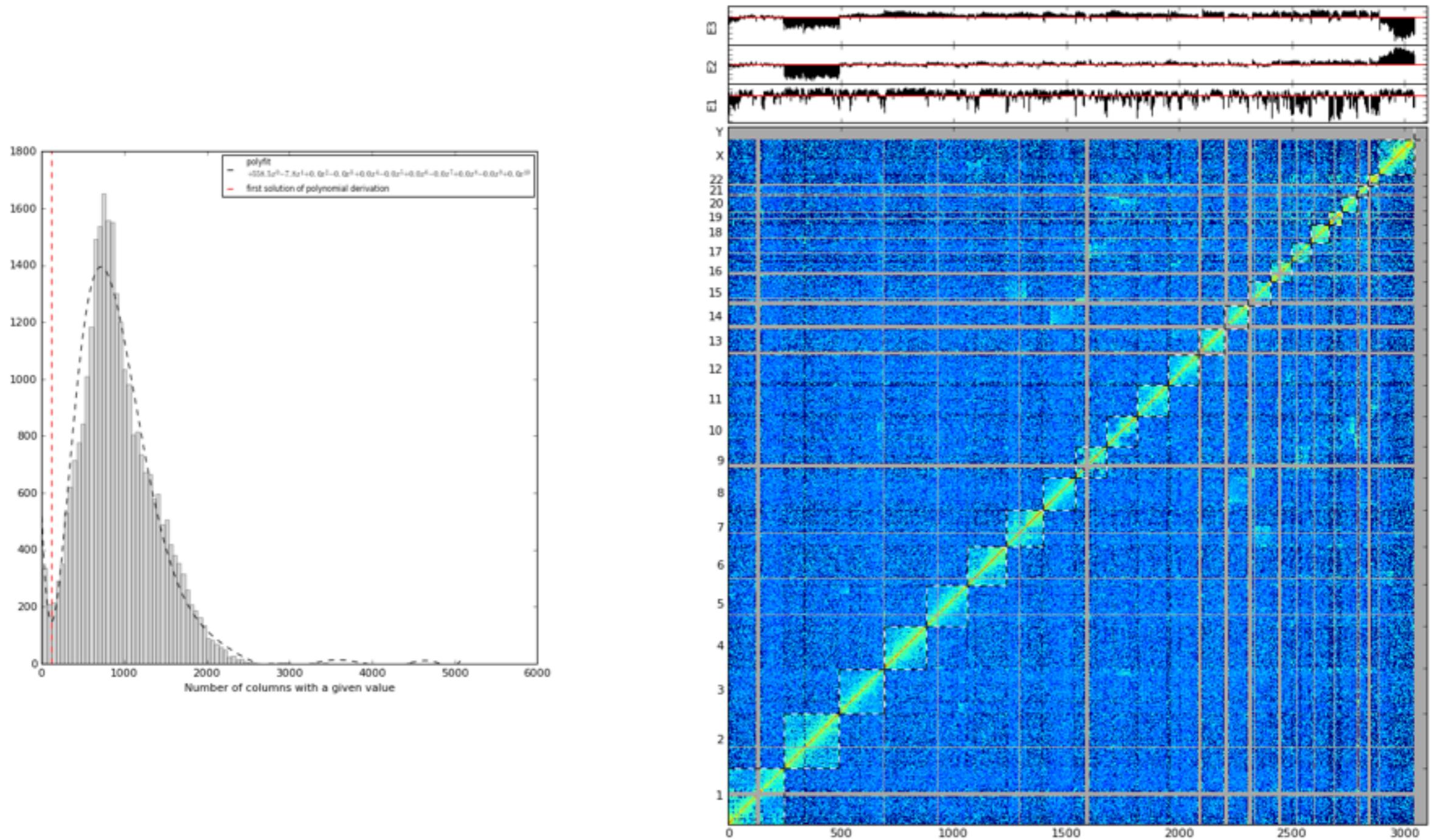
Filtered reads (and percentage of total):

Mapped both	:	21258763 (100.00%)
1- self-circle	:	400576 (1.88%)
2- dangling-end	:	1174450 (5.52%)
3- error	:	26229 (0.12%)
4- extra dangling-end	:	2971724 (13.98%)
5- too close from RES	:	2959353 (13.92%)
6- too short	:	227026 (1.07%)
7- too large	:	2322 (0.01%)
8- over-represented	:	693664 (3.26%)
9- duplicated	:	2123831 (9.99%)
10- random breaks	:	2342747 (11.02%)



Hi-C map filtering & normalising

Giving a uniform “visibility” to all the rows and columns



Questions on Day2

1 - Would it be ok to you to give a class for master students about the genome organisation in the nucleus?

Questions on Day2

2 - Would you feel comfortable to explain the concept of restraint in 3D modelling to your colleagues in the lab?

Questions on Day2

3 - How do you feel to the idea that you will do by your own the filtering and the ICE normalization?
(...sending an email to François is not contemplated!!!)