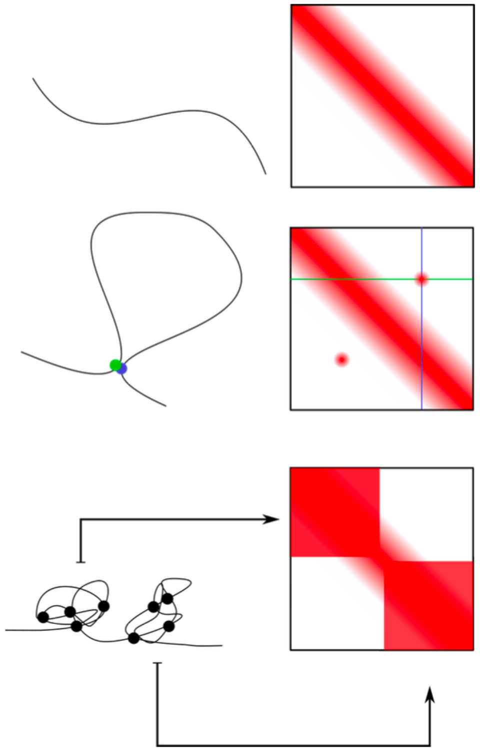
Interpreting Hi-C maps

Epigenomics Data Analysis
Jacques Serizay
Physalia 2025

Hi-C experimental workflow

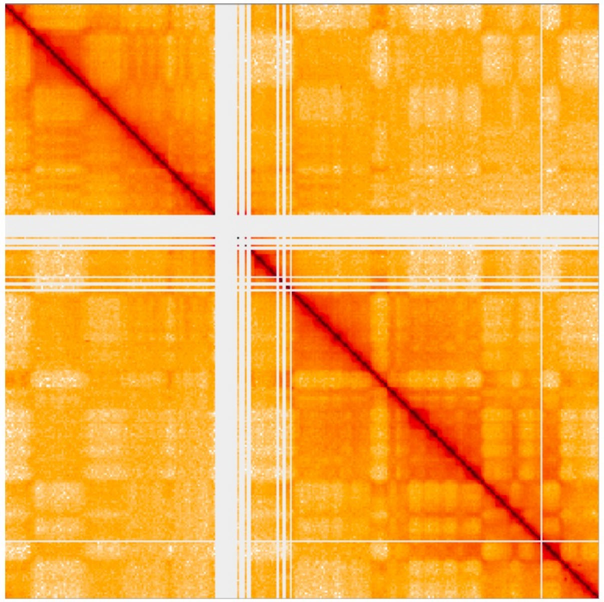


- Diagonal gradient due to polymer behavior
- Various patterns correspond to 3D structures

Chromatin compartments

Active and inactive chromatin is usually classified into A/B compartments

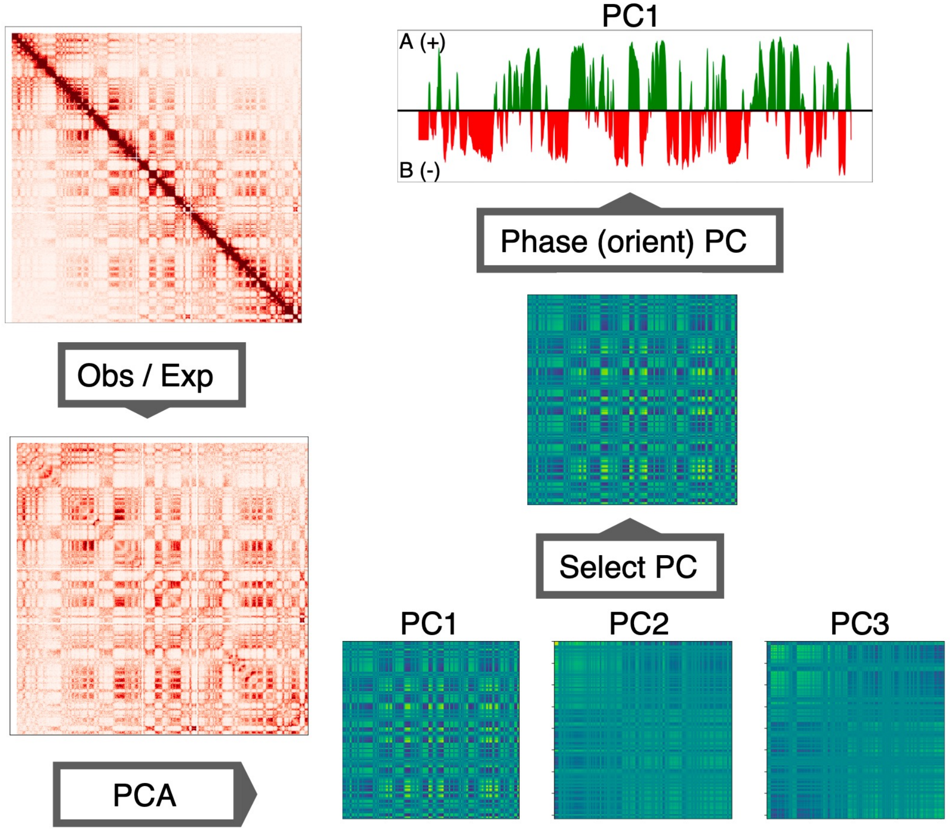
In Hi-C those compartments appear as a plaid-like pattern.



Dataset 4DNESYTUBW2E from Oksuz, Yang et al., bioRxiv 2020.

Chromatin compartments

- Most common method to identify compartments: PCA on the Hi-C matrix
- Eigenvectors explaining the most variance will contain compartments.
- Must be validated with an external correlated signal (e.g. GC%)



Obs / Exp

PCA

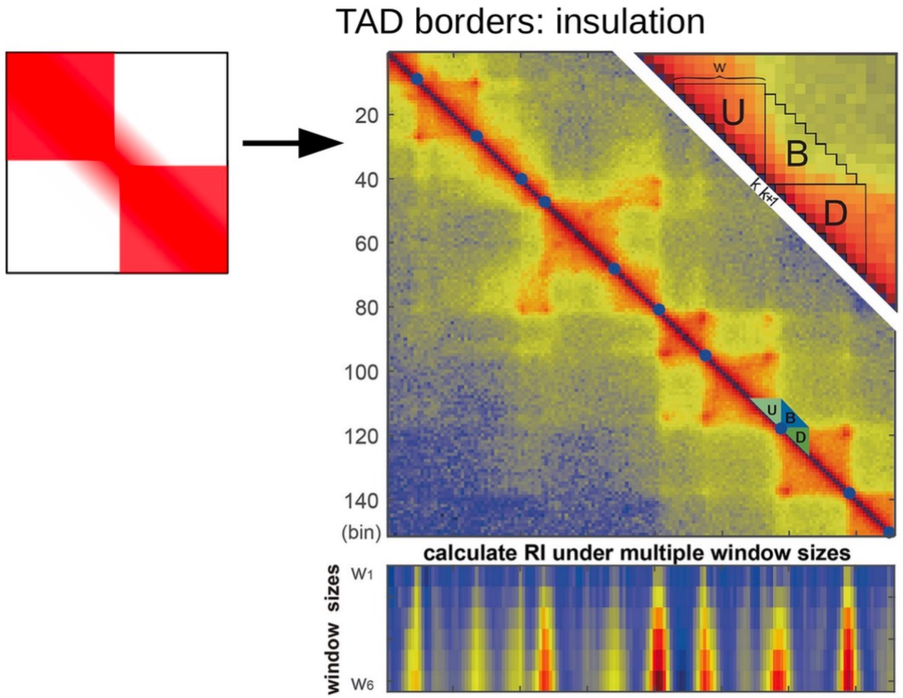
Select PC

Phase (orient) PC

Chromatin domain insulation

Insulation: Contact depletion between domains

Insulation can be quantified with a numeric score.

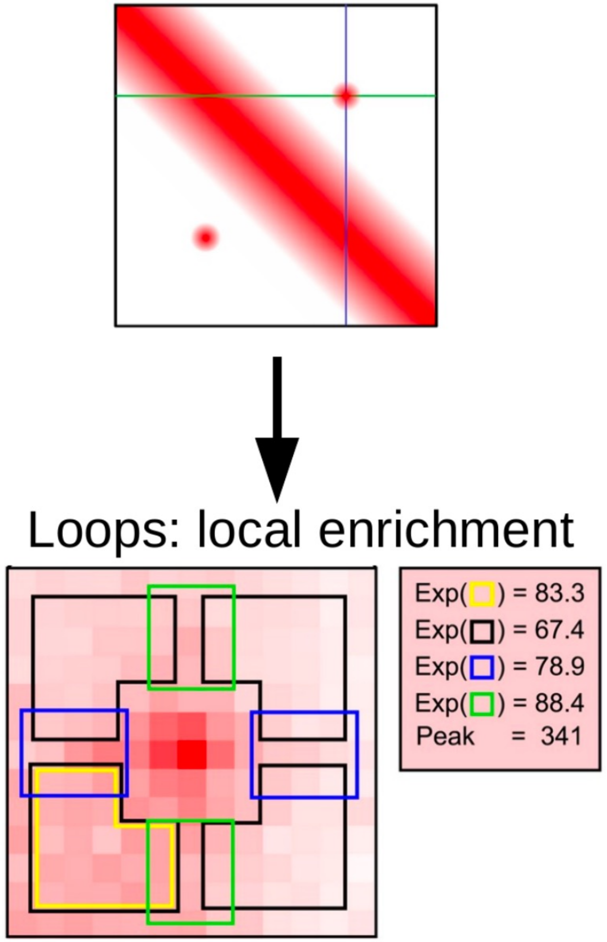


Chen et al., 2018

Chromatin features

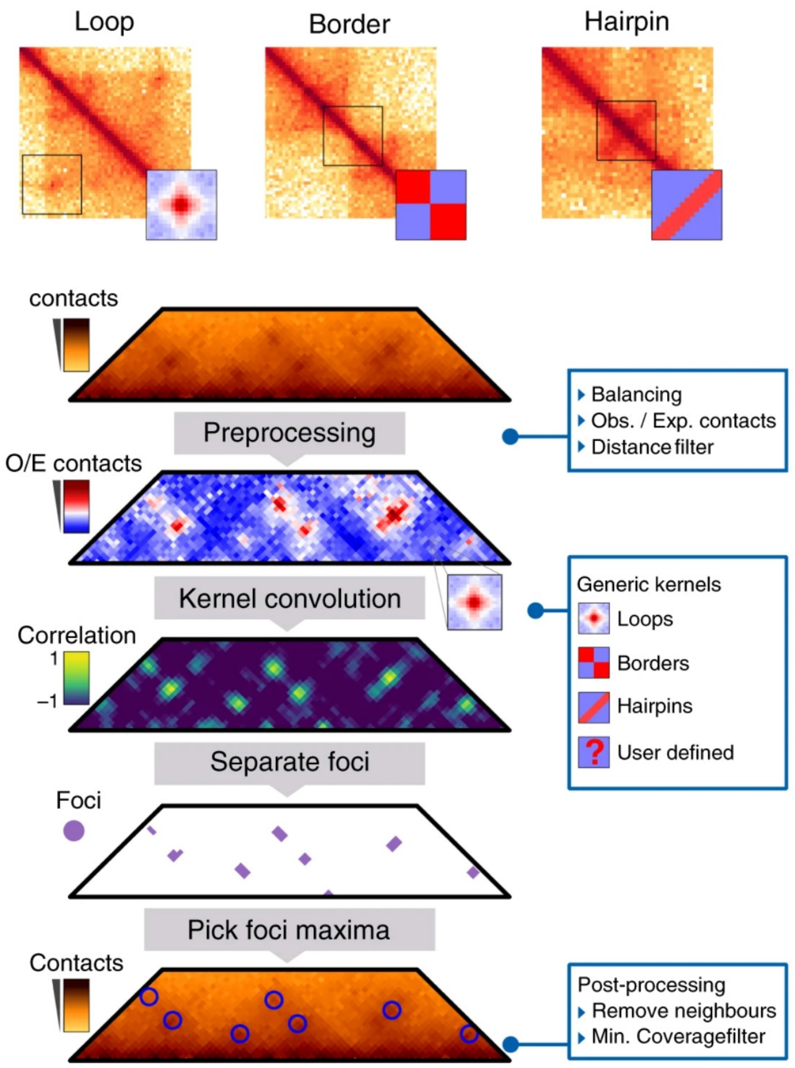
Often we want to automatically find *where* changes are happening in the genome, such as:

- Chromatin loops
- Domains (CID, TADs, …)
- General contact intensity changes



Adapted from Rao et al., 2014"

Chromosight



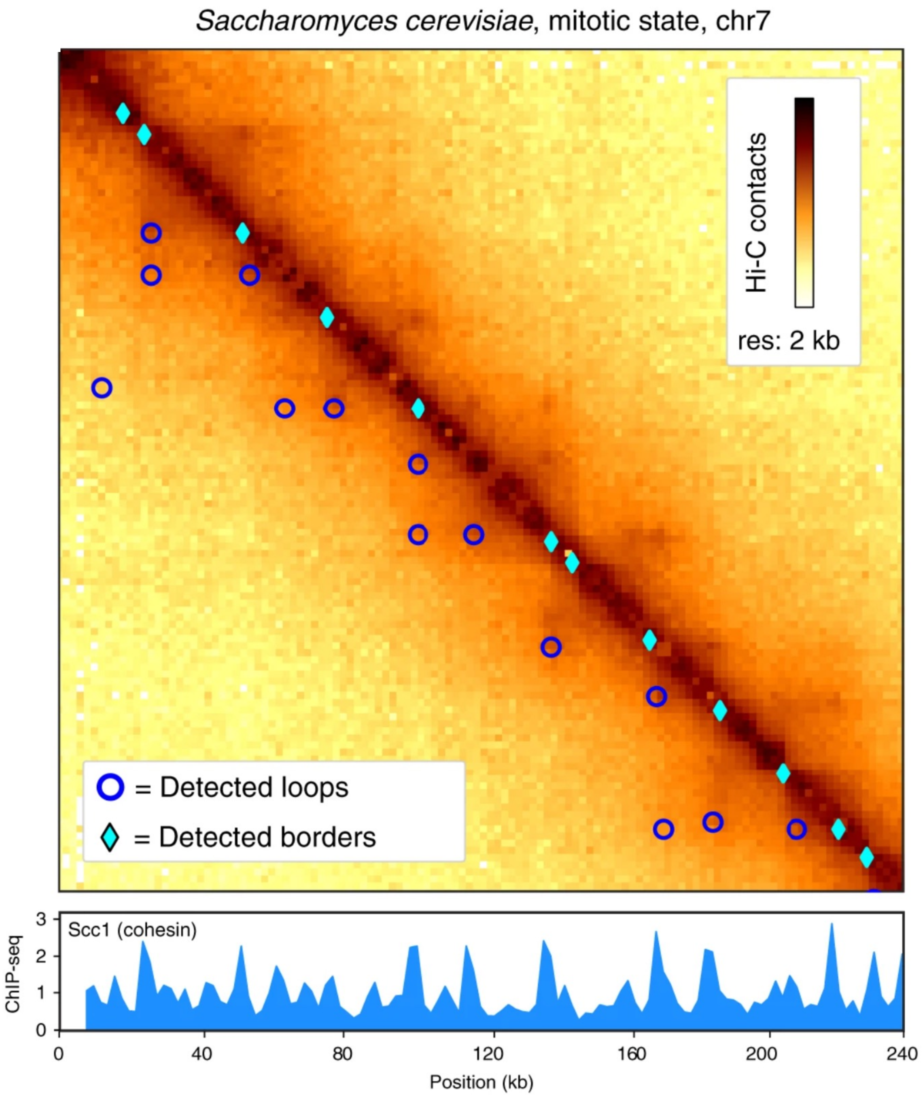
- Balancing
- Obs. / Exp. contacts
- Distance filter

Generic kernels
- Loops
- Borders
- Hairpins
- User defined

Post-processing
- Remove neighbours
- Min. Coveragefilter

https://github.com/koszullab/chromosight

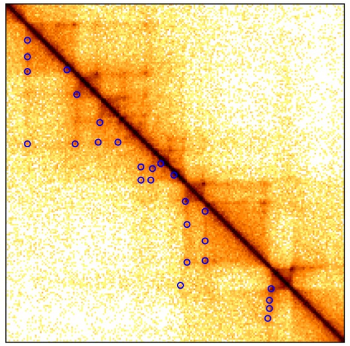
Chromosight



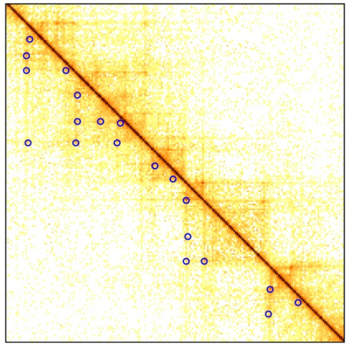
https://github.com/koszullab/chromosight

Chromosight

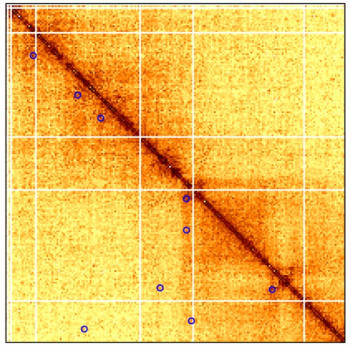
Hi-C, Nc ~250 million
--pattern loops_small



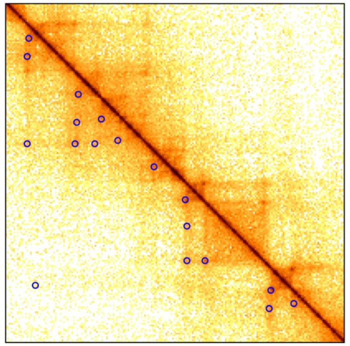
In situ ChIA-PET CTCF, Nc ~ 136 million
--perc-zero 100



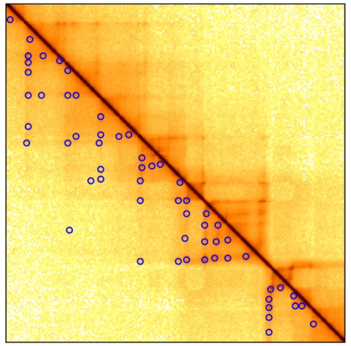
DNA SPRITE, Nc ~ 166 million
default



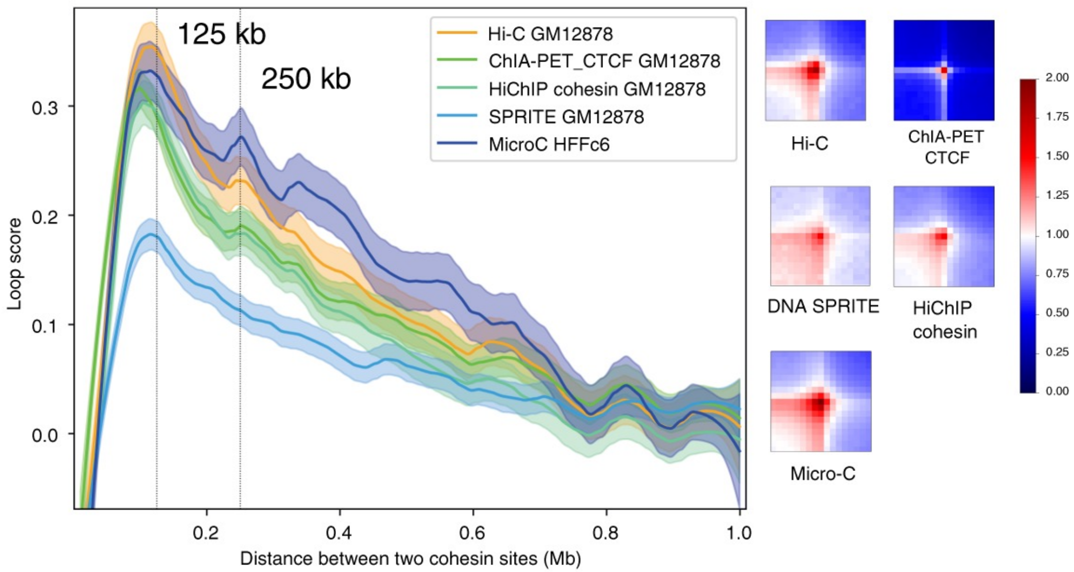
HiChIP cohesin, Nc ~ 132 million
--perc-zero 100



Micro-C, Nc ~ 3,2 billion
--pattern loops_small



Quantification on cohesin network



https://github.com/koszullab/chromosight