Detecting and comparing genomic compartments

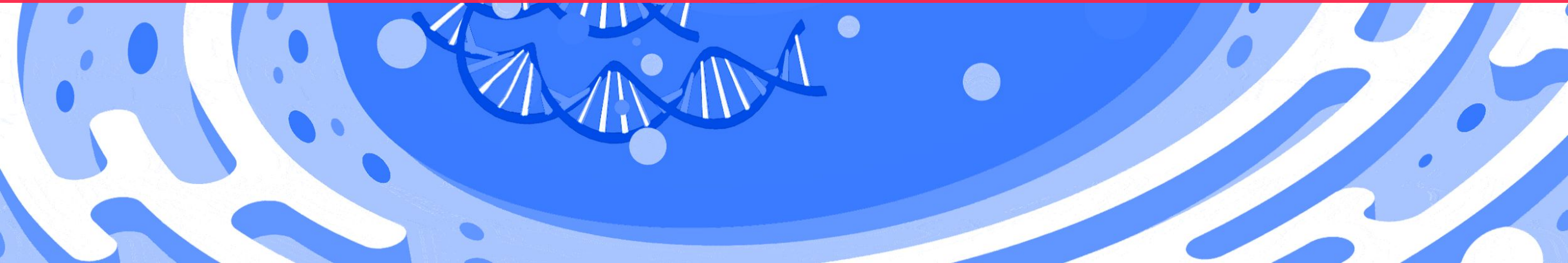
Cyril Kurylo, Sylvain Foissac, Matthias Zytnicki

INRA
SCIENCE & IMPACT

GenPhySE

MIA
TOULOUSE

UNIVERSITÉ
TOULOUSE III
PAUL SABATIER



Genomic structures

Chromosome territories
Segregation of untangled chromosomes

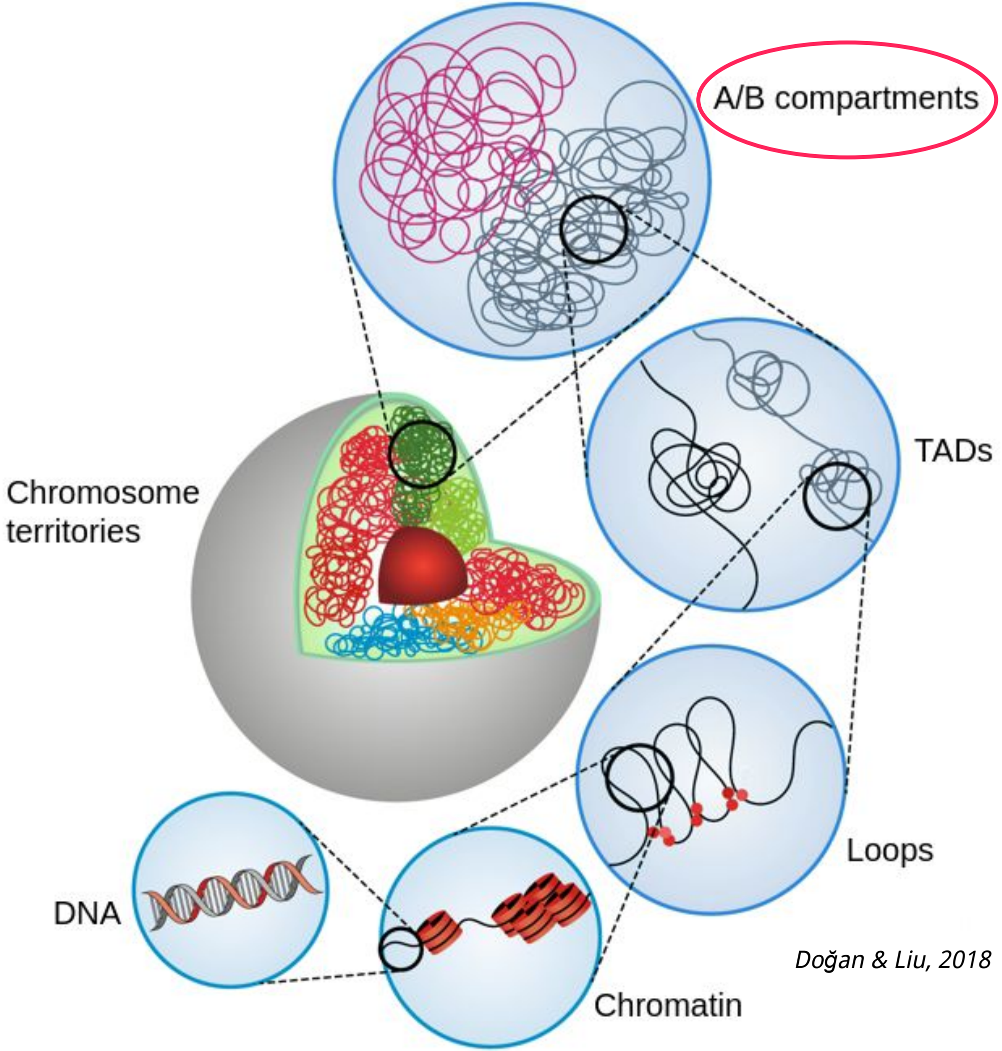
A/B compartments
Impact on gene expression

Topologically Associating Domains
Co-regulation domains

Loops
Interaction of regulatory elements

Chromatin
Compaction of DNA
Accessibility to transcription

DNA
Genetic information



Doğan & Liu, 2018

Analysing compartmentalization

Fetal development in the pig in relation to genetic merit for piglet survival[1]

J. I. Leenhouwers*, E. F. Knol†, P. N. de Groot*, H. Vos*, and T. van der Lende*[2]

*Animal Breeding and Genetics Group, Wageningen Institute of Animal Sciences, Wageningen University, Wageningen, The Netherlands and †IPG, Institute for Pig Genetics B.V., Beuningen, The Netherlands

ARTICLE

OPEN

doi:10.1038/nature14222

Chromatin architecture reorganization during stem cell differentiation

Jesse R. Dixon[1,2]*, Inkyung Jung[1]*, Siddarth Selvaraj[1,3]*, Yin Shen[1], Jessica E. Antosiewicz-Bourget[4], Ah Young Lee[1], Zhen Ye[1], Audrey Kim[1], Nisha Rajagopal[1], Wei Xie[5], Yarui Diao[1], Jing Liang[6], Huimin Zhao[6], Victor V. Lobanenkov[7], Joseph R. Ecker[8], James A. Thomson[4,9,10] & Bing Ren[1,11]

Published online: December 23, 2015

Article

TRANSPARENT PROCESS

OPEN ACCESS

molecular systems biology

Hierarchical folding and reorganization of chromosomes are linked to transcriptional changes in cellular differentiation

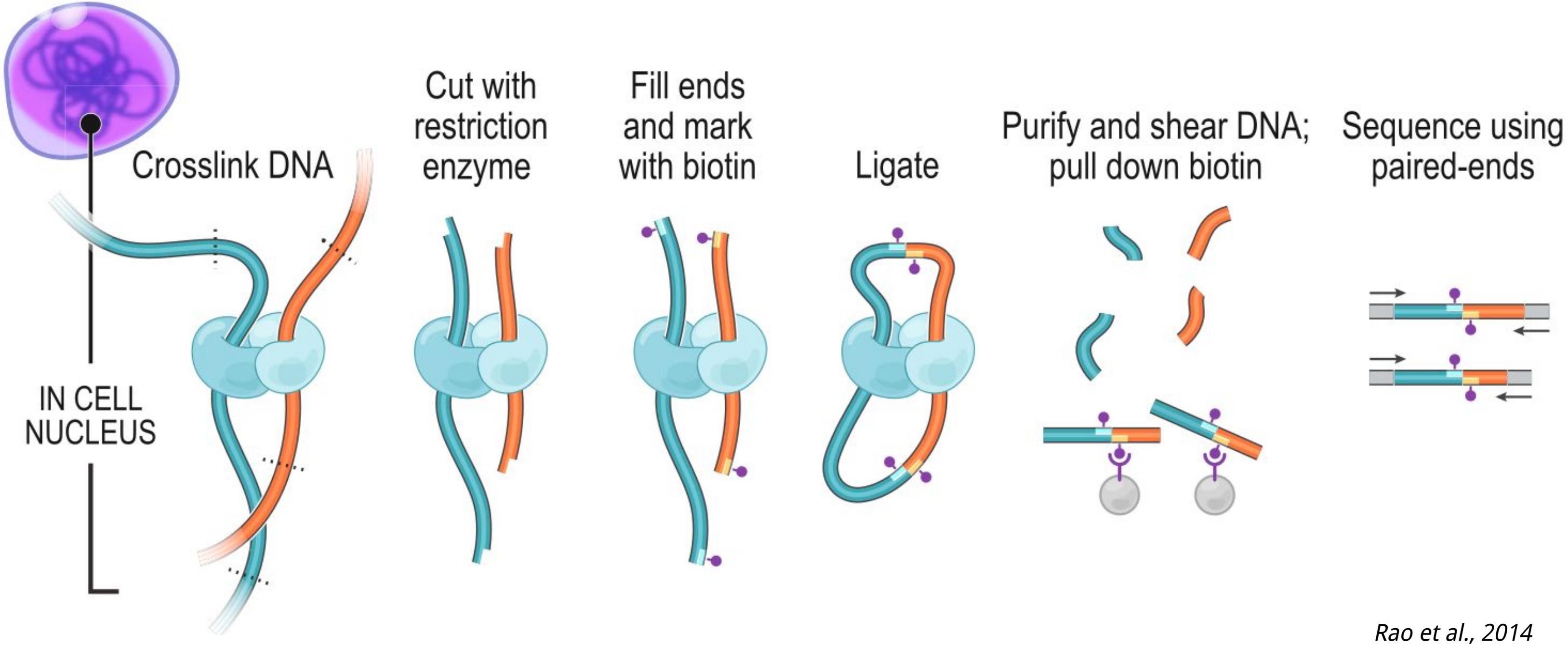
James Fraser[1,†], Carmelo Ferrai[2,3,†], Andrea M Chiariello[4,†], Markus Schueler[2,†], Tiago Rito[2,†], Giovanni Laudanno[4,†], Mariano Barbieri[2], Benjamin L Moore[5], Dorothee CA Kraemer[2], Stuart Aitken[5], Sheila Q Xie[3,‡], Kelly J Morris[2,3], Masayoshi Itoh[6,7], Hideya Kawaji[6,7], Ines Jaeger[8,#], Yoshihide Hayashizaki[6], Piero Carninci[7], Alistair RR Forrest[7,¶], The FANTOM Consortium, Colin A Semple[5,*], Josée Dostie[1,**], Ana Pombo[2,3,***] & Mario Nicodemi[4,****]

Motivations

Compartment changes
linked to transcriptional changes

Large White neonatal mortality
due to poor fetal development

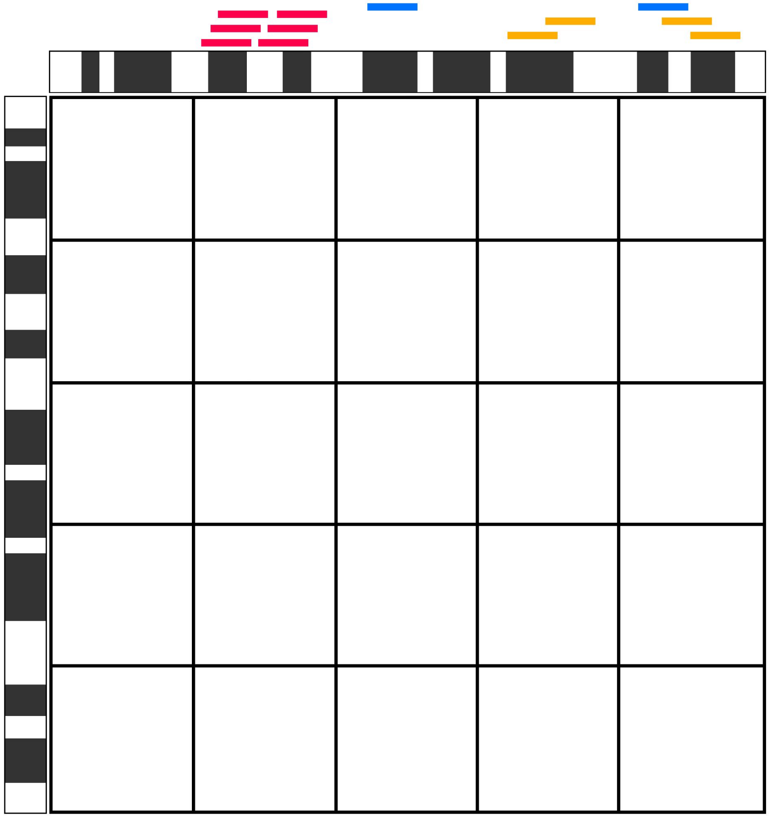
Using Hi-C to expose compartments



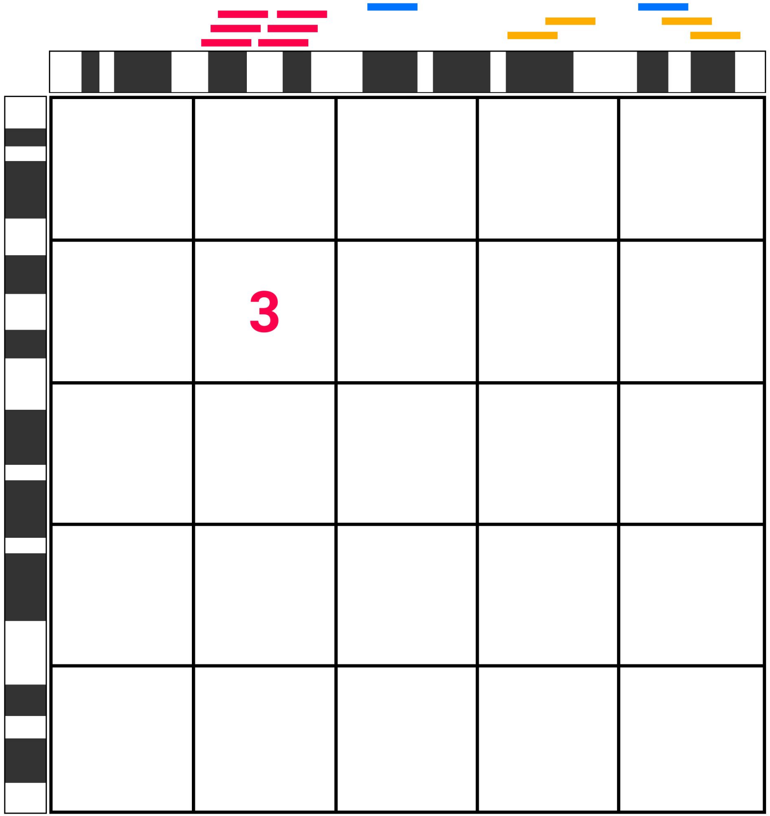
Rao et al., 2014

Using Hi-C to expose compartments

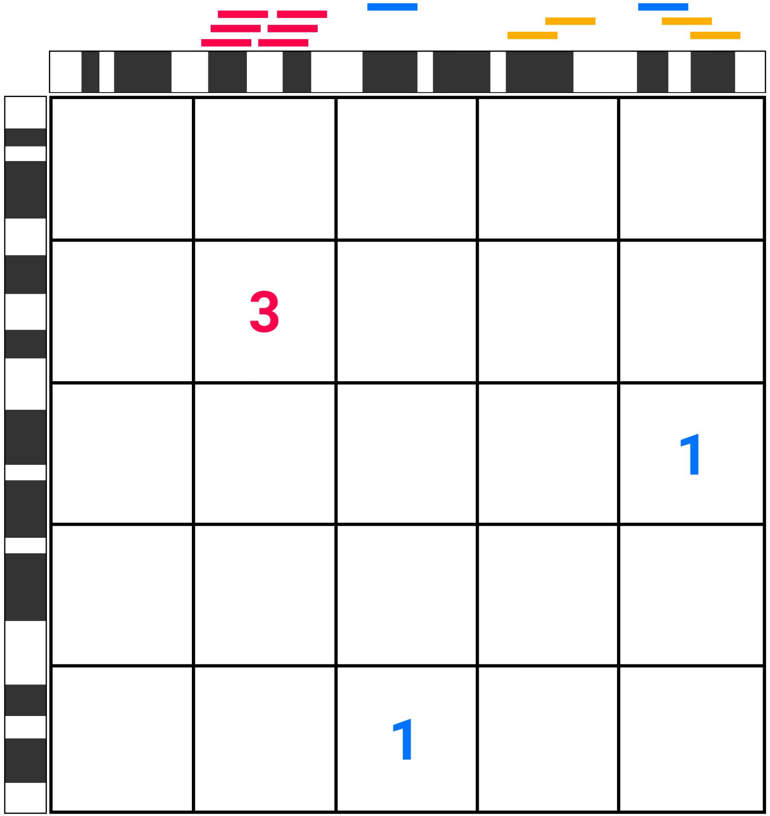
Using Hi-C to expose compartments



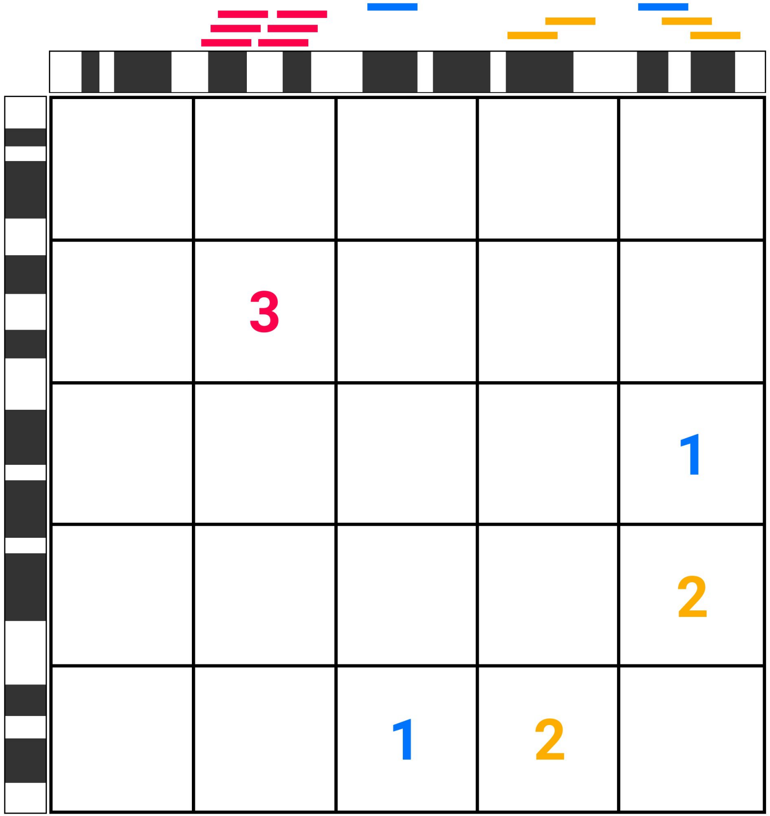
Using Hi-C to expose compartments



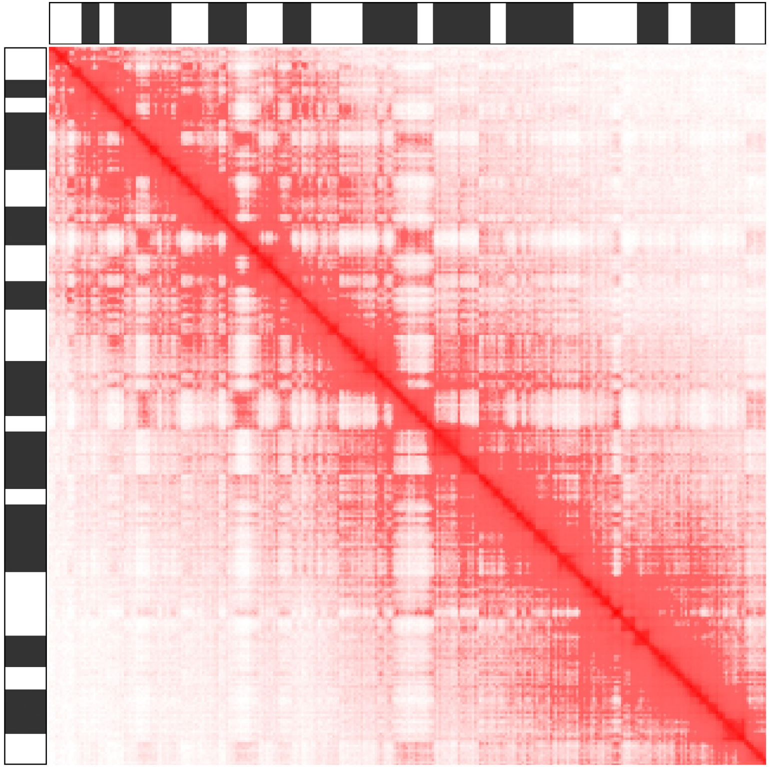
Using Hi-C to expose compartments



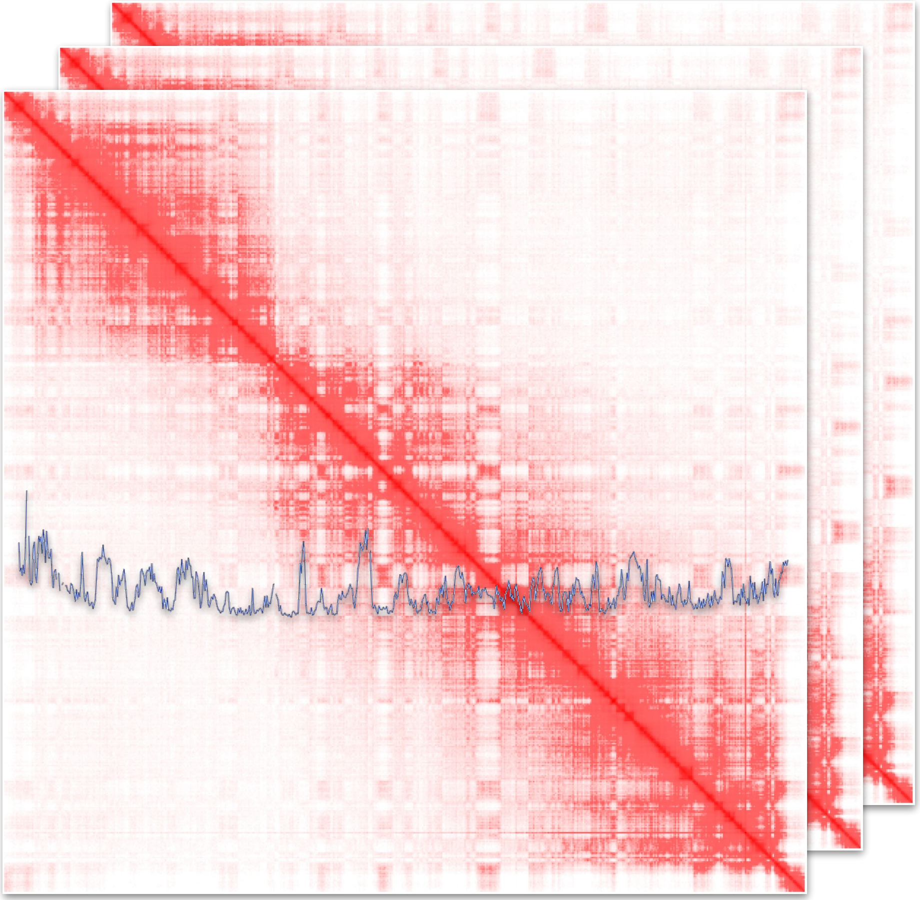
Using Hi-C to expose compartments



Using Hi-C to expose compartments



Analysing compartmentalization



Ambitions

- Computationally detect compartments ?
- Using replicates ?
- Providing a confidence measure ?
- Statistical comparison across conditions ?

Data

2 conditions — 90 and 110 days of development

3 Hi-C replicates per condition

Hi-C DOC: Detection Of Compartments with replicates

available at github.com/mzytnicki/HiCDOC

Cyclic Loess

Knight-Ruiz

Distance Loess Regression

Matrix normalization

Constrained k-means

Compartment detection

Concordance Measure

P-value

Comparison

Correctly normalizing Hi-C matrices

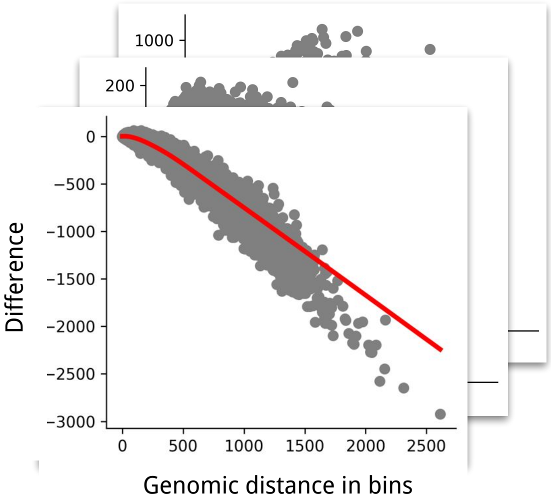
Technical biases

Sequencing depth

Restriction enzyme

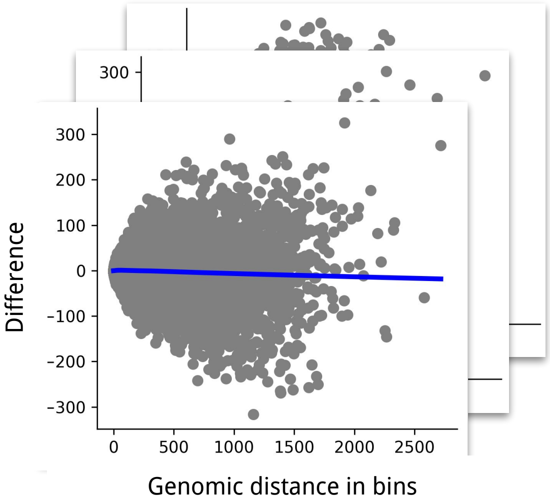
Cross-linking conditions

Experiment quality



Cyclic Loess

multiHiCcompare



Cyclic Loess

Knight-Ruiz

Distance Loess

Matrix normalization

Compartment detection

Comparison

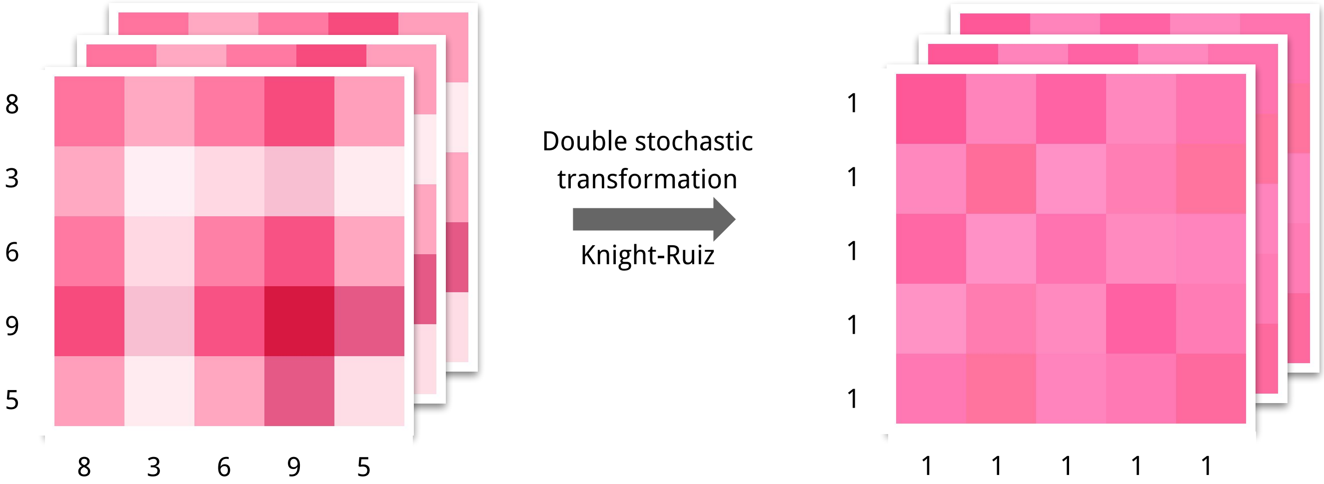
Correctly normalizing Hi-C matrices

Biological biases

GC content

Restriction site distribution

Repeated sequences



Double stochastic transformation

Knight-Ruiz

Cyclic Loess

Knight-Ruiz

Distance Loess

Matrix normalization

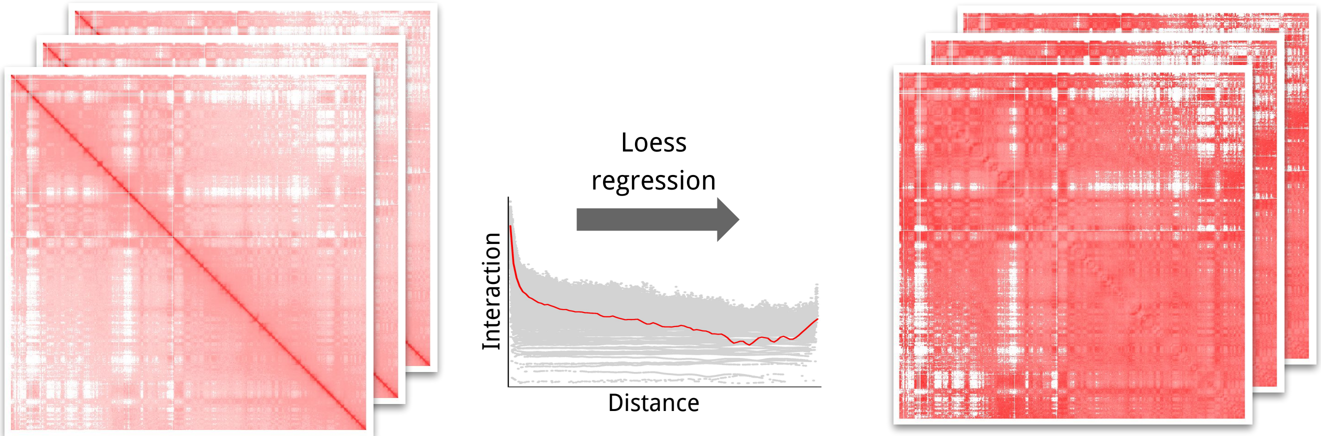
Compartment detection

Comparison

Correctly normalizing Hi-C matrices

Distance effect

Proximity between regions



Cyclic Loess

Knight-Ruiz

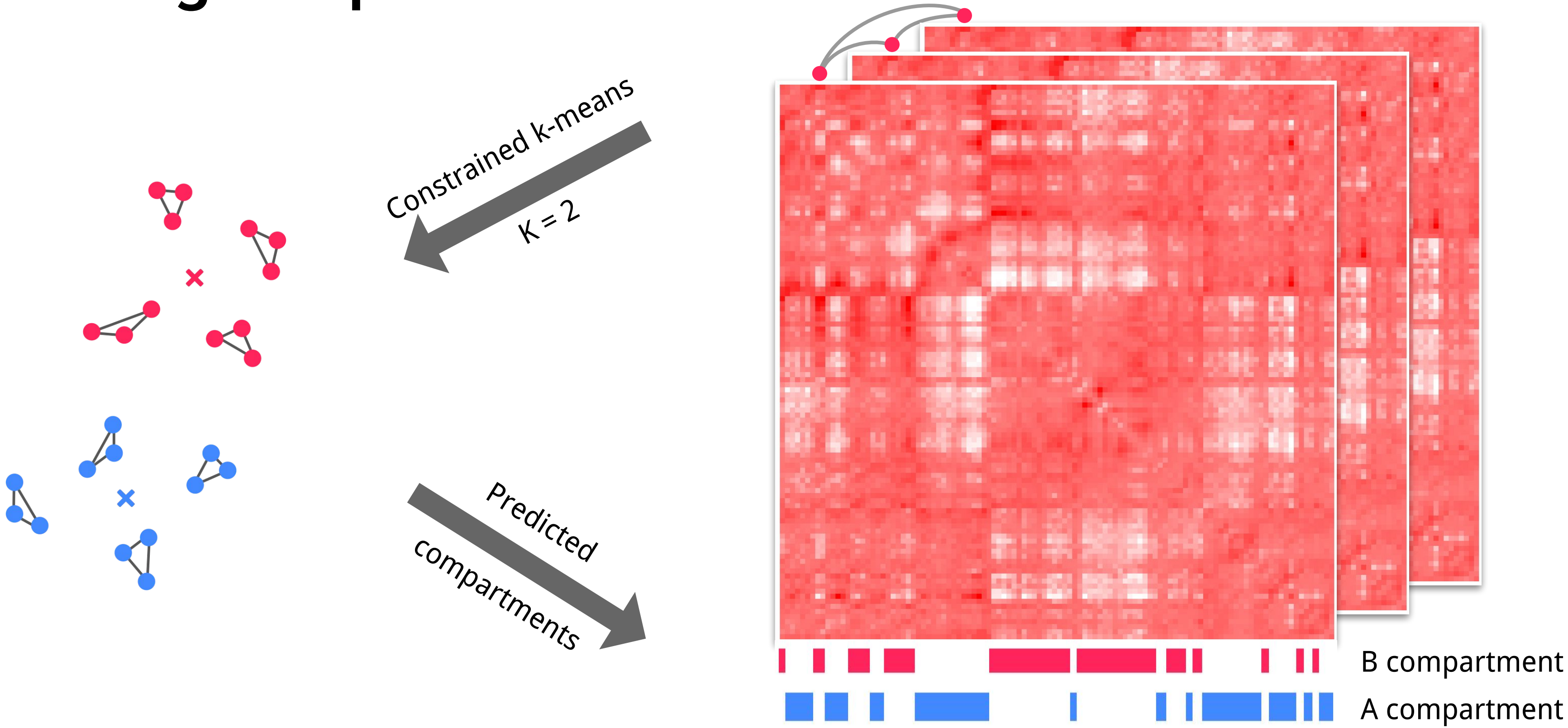
Distance Loess

Matrix normalization

Compartment detection

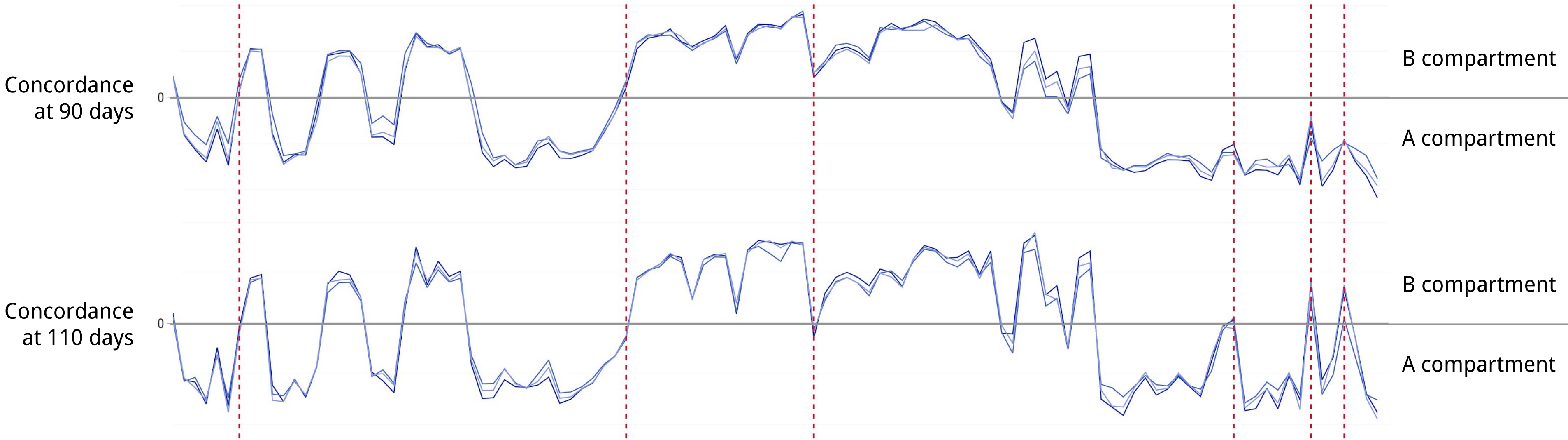
Comparison

Detecting compartments



Matrix normalization | Compartment detection | Comparison

Comparing compartmentalization between conditions



$$2 \times \frac{\log \frac{\text{distance(interactions vector, centroid 1)}}{\text{distance(interactions vector, centroid 2)}} - \log \frac{\text{distance(centroid 1, centroid 1)}}{\text{distance(centroid 1, centroid 2)}}}{\log \frac{\text{distance(centroid 2, centroid 1)}}{\text{distance(centroid 2, centroid 2)}} - \log \frac{\text{distance(centroid 1, centroid 1)}}{\text{distance(centroid 1, centroid 2)}}} - 1$$

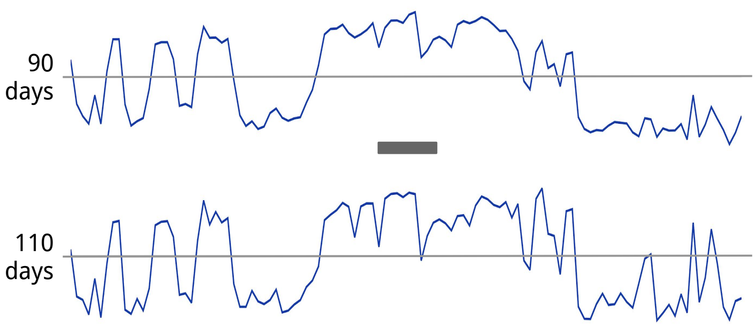
Matrix normalization | Compartment detection | **Comparison**

Comparing compartmentalization between conditions

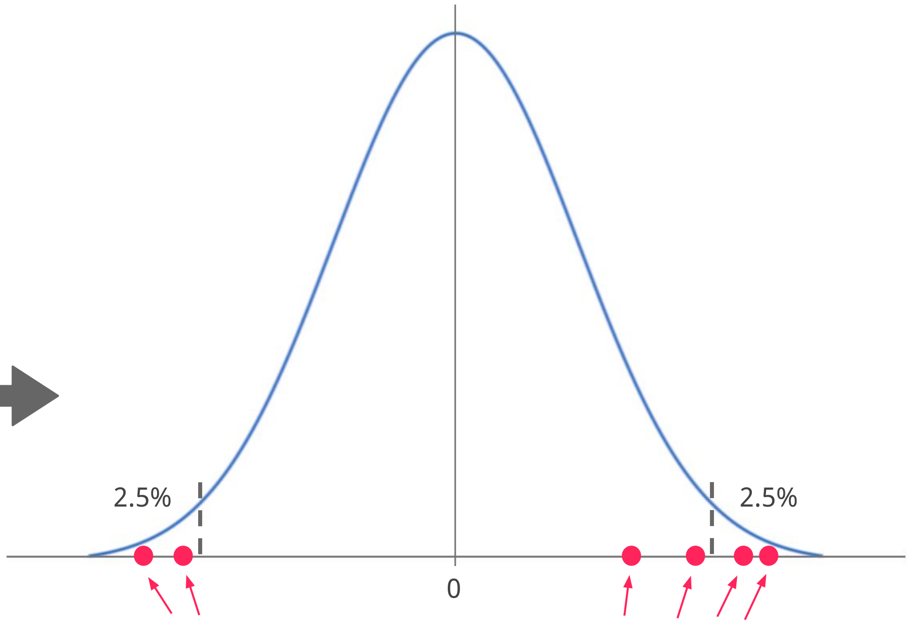
P-value

Probability of observing
a difference between concordances
as extreme or more extreme
when the compartment doesn't change

Differences between concordances



Distribution of the differences
when the compartment doesn't change

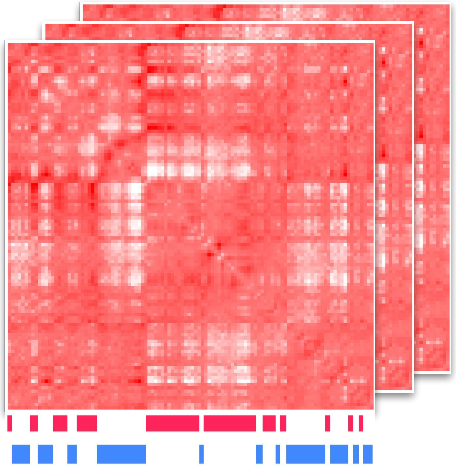


Median differences for predicted compartment changes
(with constrained k-means)

Conclusion and perspectives

Ambitions achieved

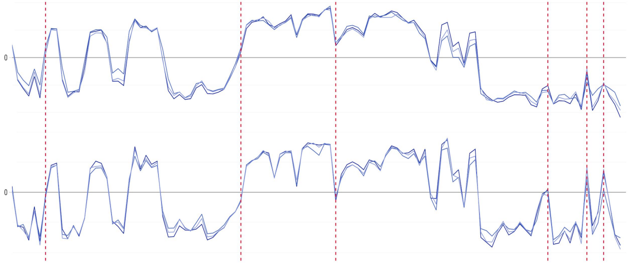
- Computationally detect compartments ✓
- Using replicates ✓
- Providing a quantitative measure ✓
- Statistical comparison across conditions ✓



Preliminary Results

Predicted compartment changes

Ongoing statistical analysis



Perspectives

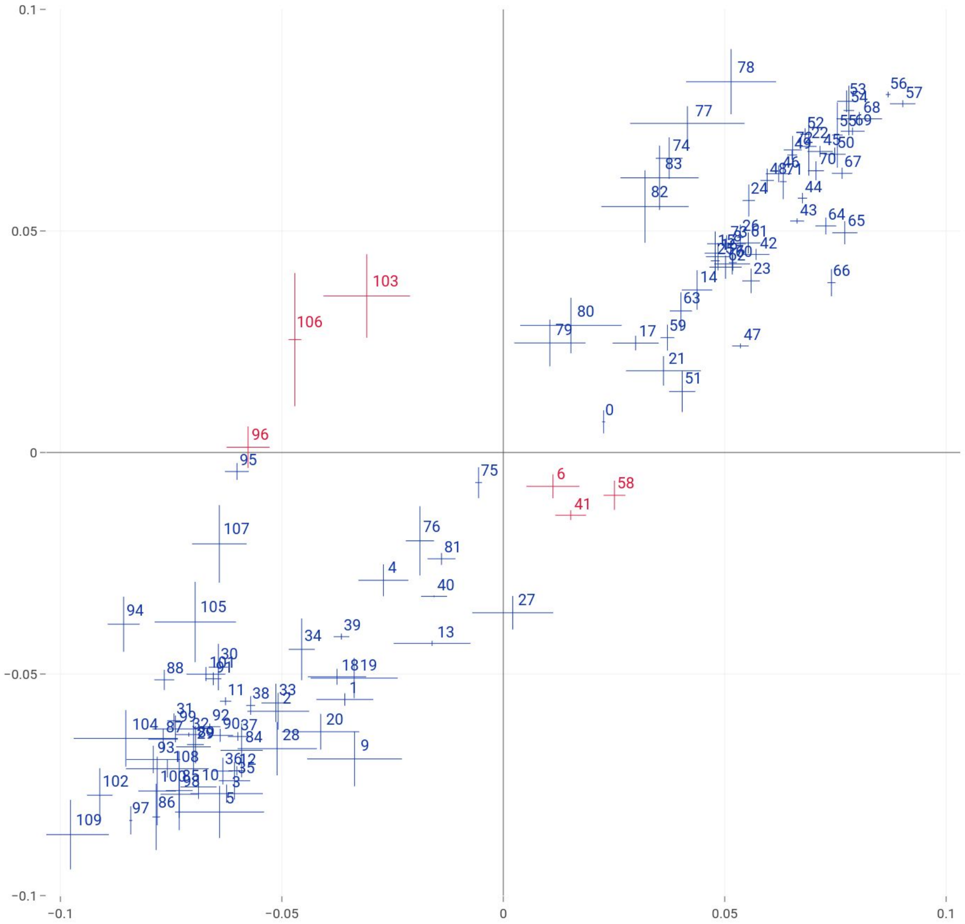
Analyse genes
in switching regions

Publish method and results
for our data

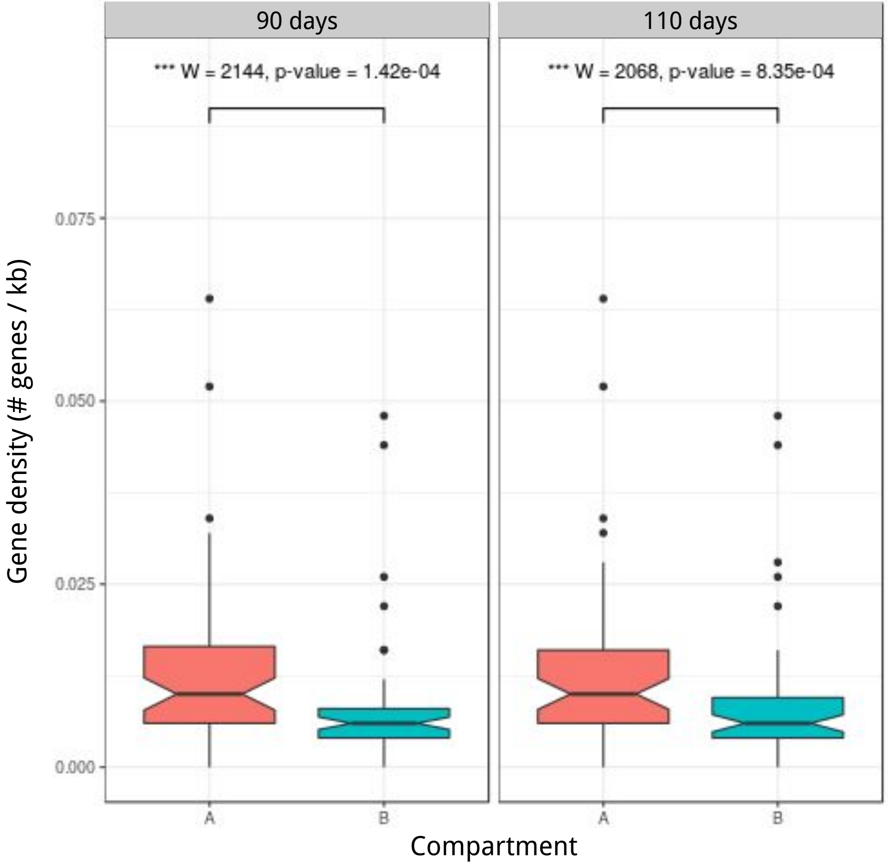
Thank You

github.com/mzytnicki/HiCDOC

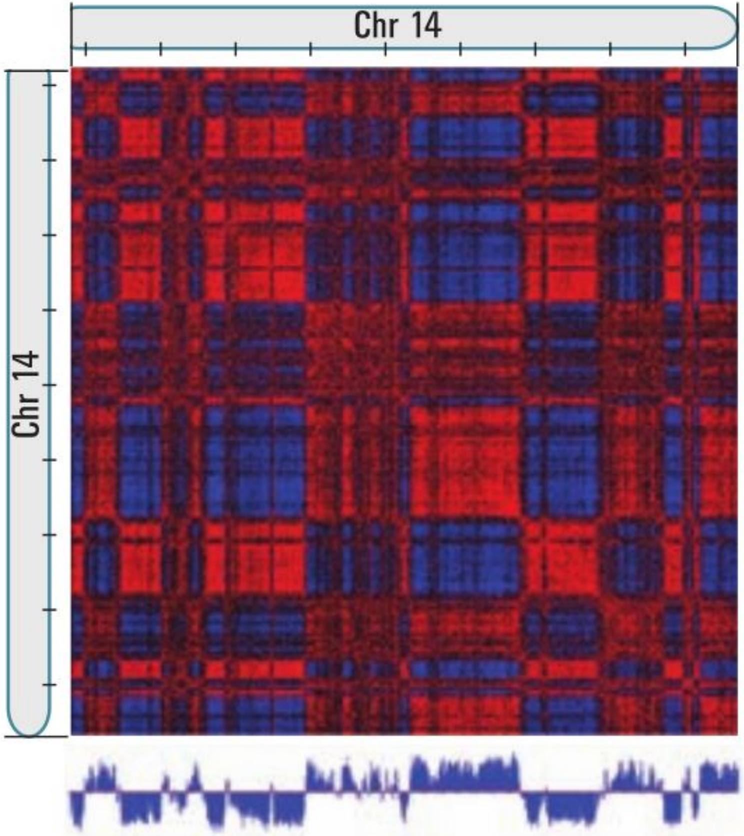
Concordance comparison



Gene density



PCA detection



Lieberman-Aiden et al., 2009