Chromatin 3D organization principles revealed by network theory: gene regulation, replication and beyond

INSA, Toulouse, 4th December 2019

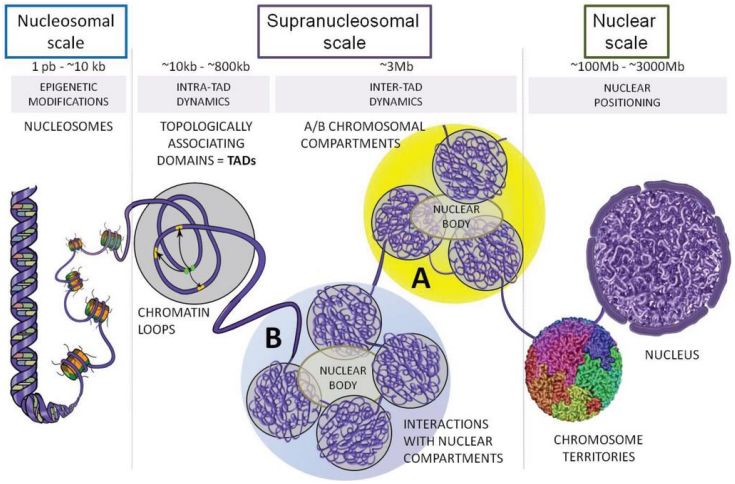
Vera Pancaldi
Cancer Research Center of Toulouse
Barcelona Supercomputing Center

Overview of talk

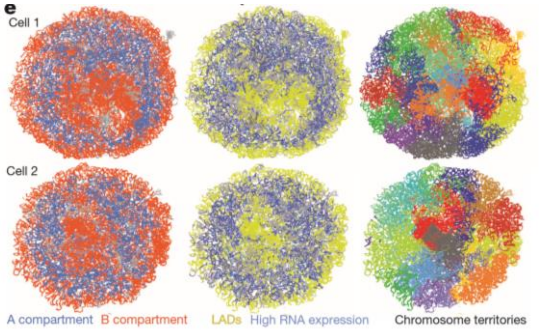
Networks

Chromatin networks
Chromatin Assortativity
Tools
Replication in 3D
Perspectives

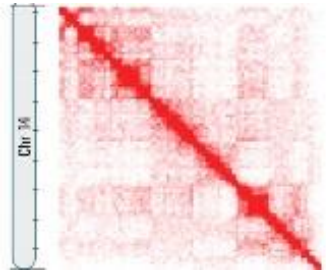
Chromatin 3D structure



Single cell !

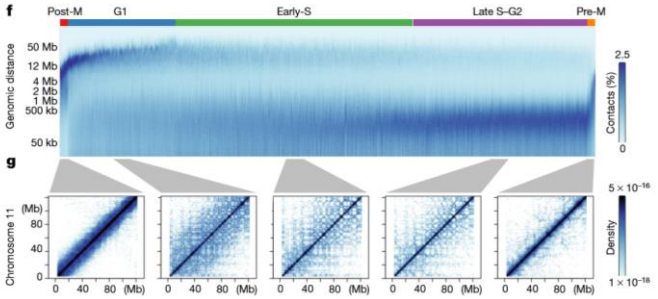


Single cell during the cell cycle



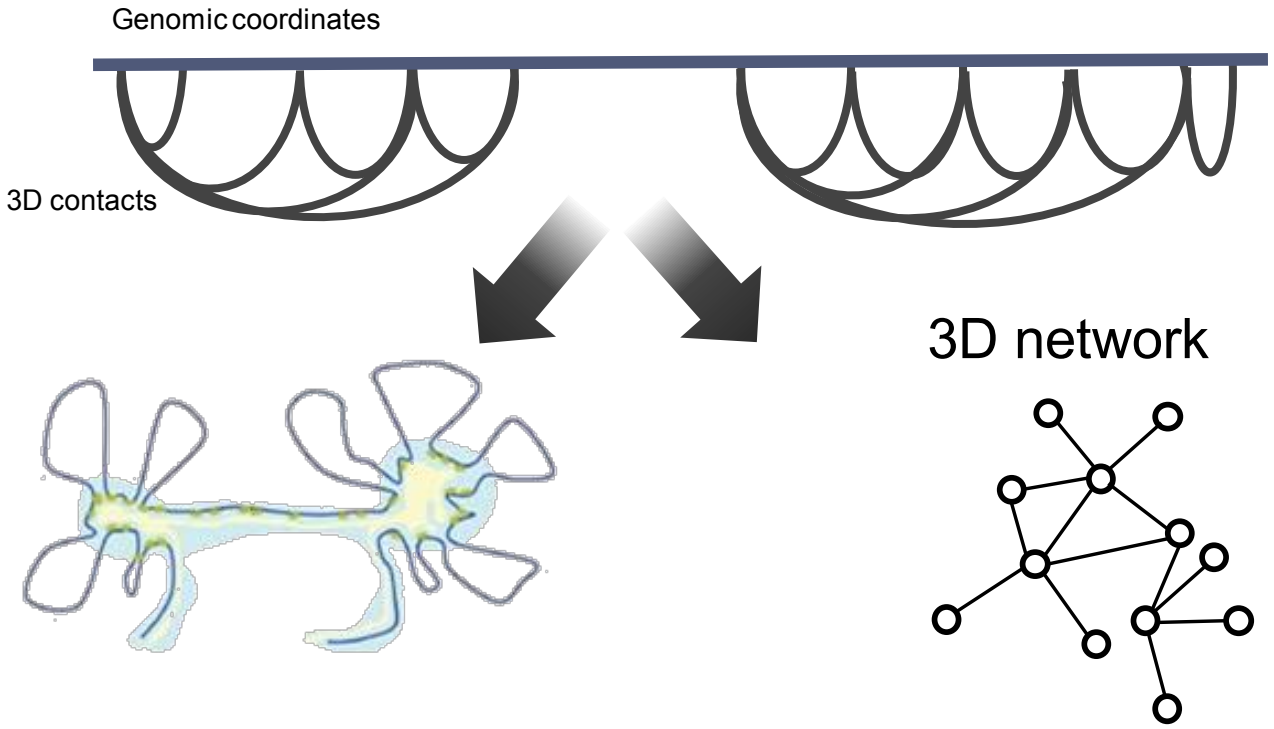
TADs:
More contacts within
Less outside

Compartments:
Acitve/Repressed



Ea et al. 2015 **Contribution of Topological Domains and Loop Formation to 3D Chromatin Organization**.
Stevens et al. 2017 **3D structures of individual mammalian genomes studied by single-cell Hi-C**
Nagano et al. 2017 **Cell-cycle dynamics of chromosomal organization at single-cell resolution**

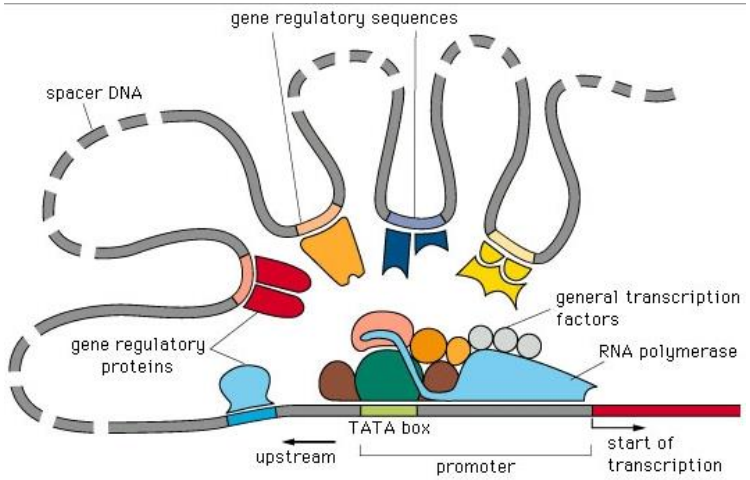
Chromatin networks

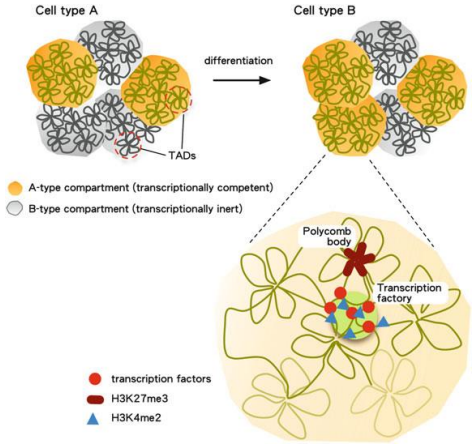


Principal players in gene regulation

Polycomb – gene repression
RNAPII – gene transcription (RNAPIIS2p needed for elongation)

Genes can be co-transcribed (Promoter-Promoter contact, PP)
Regulatory regions bind the gene promoters to activate genes (PO)



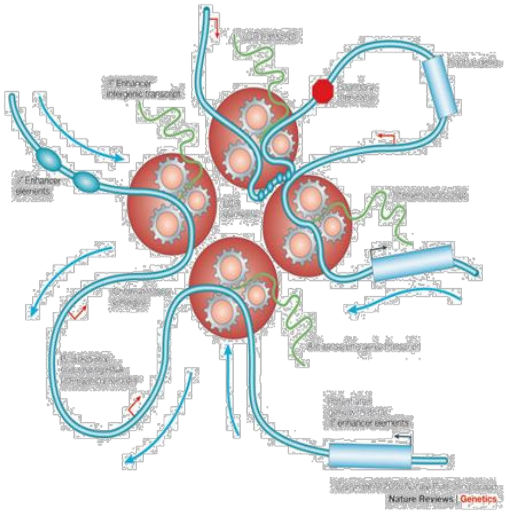


What about genes? PCHiC!

Problem so far: Genome wide interaction networks are dominated by interactions far from genes. Need very high coverage to pick promoters and see their interactions.

Solution: Promoter-Capture HiC (PCHiC)
Add promoter capture step
To ensure only interactions involving at least one promoter are kept.
(No pull-downs, genome-wide)

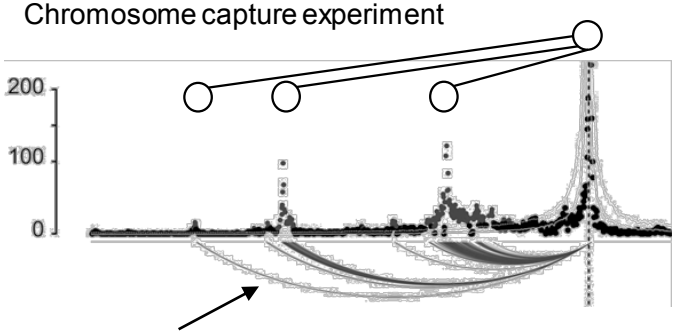
Can look for transcription factories:
Regions where functionally related transcripts are transcribed



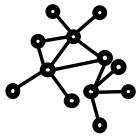
Chakalova et al. 2015, **Replication and transcription: Shaping the landscape of the genome**; Schoenfelder, S. *et al*. 2015, **The pluripotent regulatory circuitry connecting promoters to their long-range interacting elements**.

The 3D genome as a network

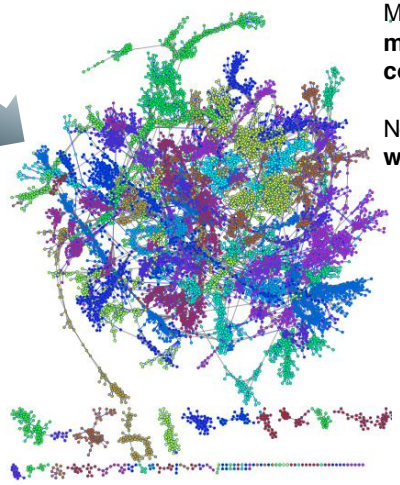
Chromosome capture experiment



3D network



Genomic fragment



Some other network approaches to chromatin:

Botta et al. 2010, **Intra- and inter-chromosomal interactions correlate with CTCF binding genome wide**

KS Sandhu et al. 2012, **Large-scale functional organization of long-range chromatin interaction networks**

Boulos et al. 2017 **Multi-scale structural community organisation of the human genome**

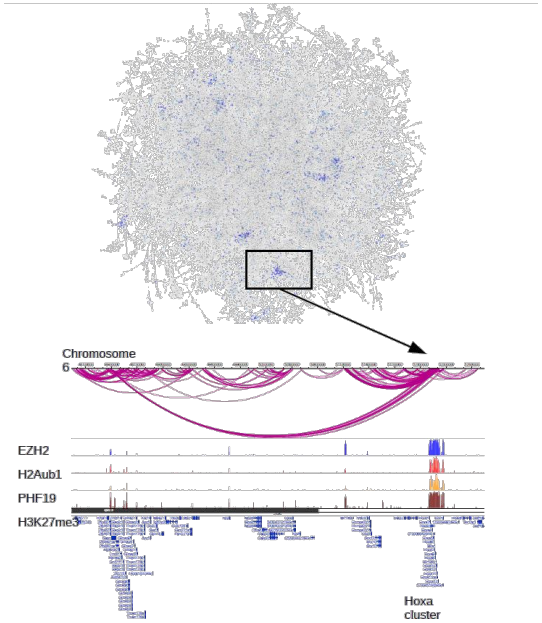
Mourad et al. 2017 **Uncovering direct and indirect molecular determinants of chromatin loops using a computational integrative approach**

Norton et al. 2018 **Detecting hierarchical genome folding with network modularity**

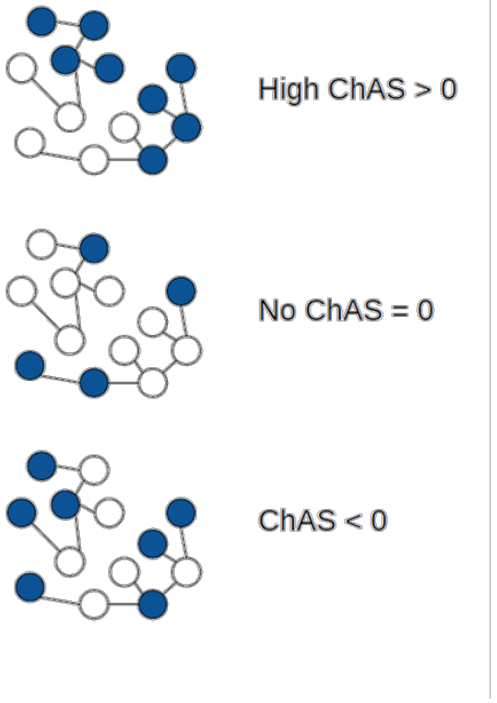
Pancaldi et al. 2016 **Integrating epigenomic data and 3D genomic structure with a new measure of chromatin assortativity**.

Chromatin Assortativity of epigenetic marks

Project ChIP-seq datasets on 3D chromatin interaction network PCHiC networks in mouse Embryonic Stems Cells mESCs)
(Collaboration with Peter Fraser, Babraham Institute)



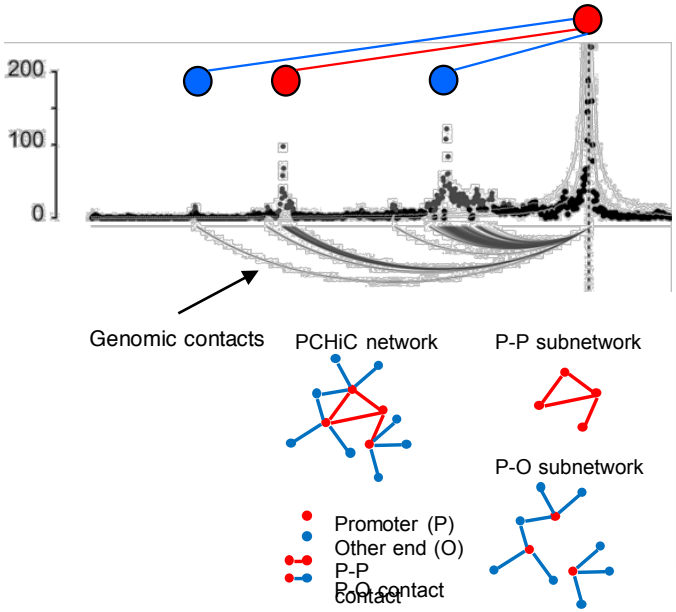
Do regions with specific marks cluster?
Inspired by social networks (the twitter story)
Define Chromatin Assortativity (ChAs)

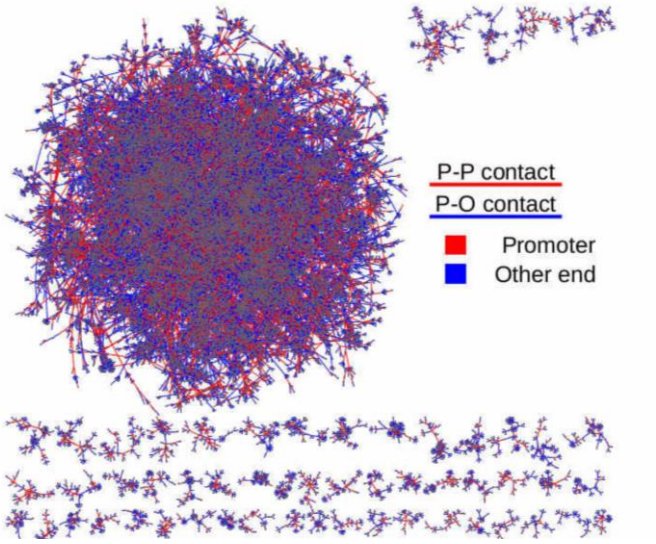


Pancaldi et al. 2016 **Integrating epigenomic data and 3D genomic structure with a new measure of chromatin assortativity**.

Promoter Capture HiC networks in mESC

Nodes are chromatin fragments (5kb median size)
Connections (edges) are 3D contacts
Significant contacts are detected using CHiCAGO





Schoenfelder et al. 2015 **The pluripotent regulatory circuitry connecting promoters to their long-range interacting elements.**
Cairns et al. 2016 **CHiCAGO: robust detection of DNA looping interactions in Capture Hi-C data.**

Comparing ChAs in P-P and P-O subnetworks

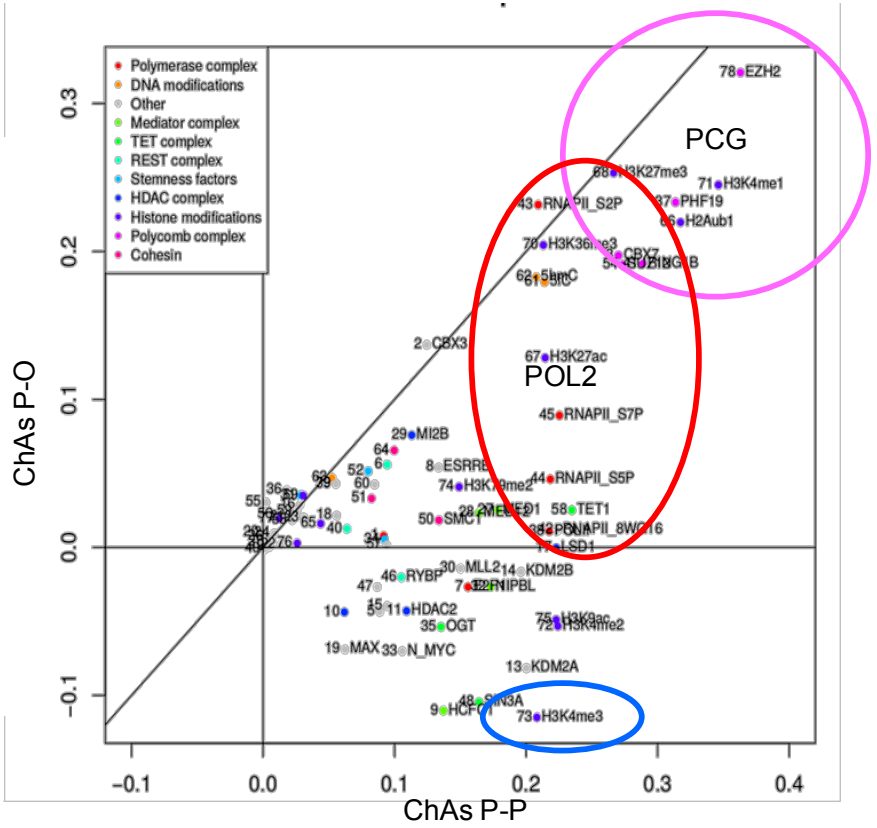
Identify features that have different ChAs in P-P and P-O contacts in mESC

PCG - on diagonal
Similar ChAs>0 in P-P and P-O
Equal importance

RNAPII - spread
Variable ChAs in P-O,
ChAs>0 in P-P

H3K4me3
ChAs >0 in P-P
ChAs<0 in P-O
(only present in promoters)

Fragments that have this mark are more likely to interact
→**Preferential contacts of active gene promoters**

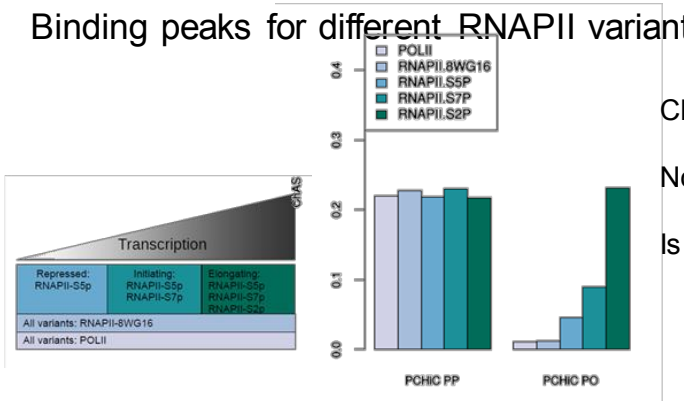


Pancaldi et al. 2016 **Integrating epigenomic data and 3D genomic structure with a new measure of chromatin assortativity**.

Assortativity of RNAPII

5 Different RNAPII features

Binding peaks for different RNAPII variants



ChAs of RNAPII in P-O variable

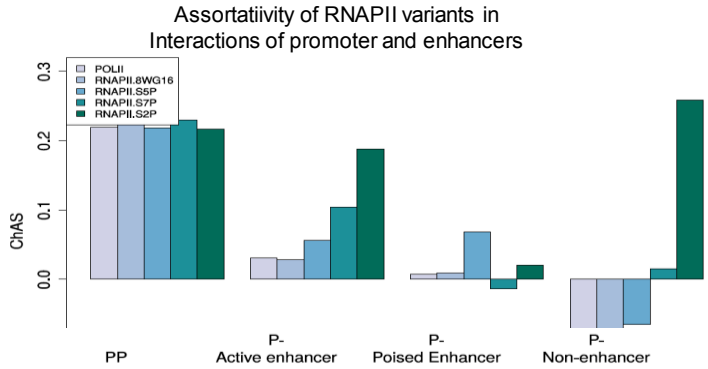
Non-elongating RNAPII has low ChAs in P-0

Is expression enhancing mediated by RNAPII S2p?

Active enhancer
H3K4me1+H3K27ac

Poised Enhancer
H3K4me1+
H3K27me3

Non-enhancer
No H3k4me1

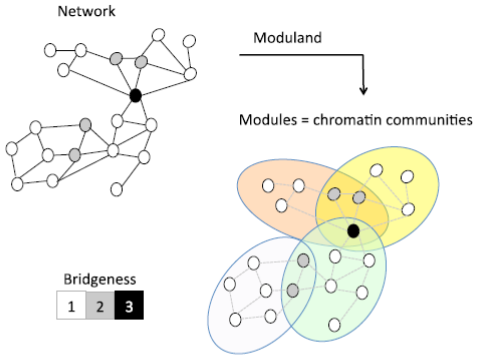


Pancaldi et al. 2016 **Integrating epigenomic data and 3D genomic structure with a new measure of chromatin assortativity**.

Chromatin network analysis

Apply Moduland (Cytoscape plugin) to identify overlapping chromatin communities, measure bridgeness

	Bridgeness	Betweenness centrality	Degree	CC	Party/date
PCG	Low	High	High	Very low	**PARTY**
RNAPII general	High	Low	Low	Low	**DATE**
RNAPII S2p	Low	Very low	Very low	Medium	**NOT HUB**



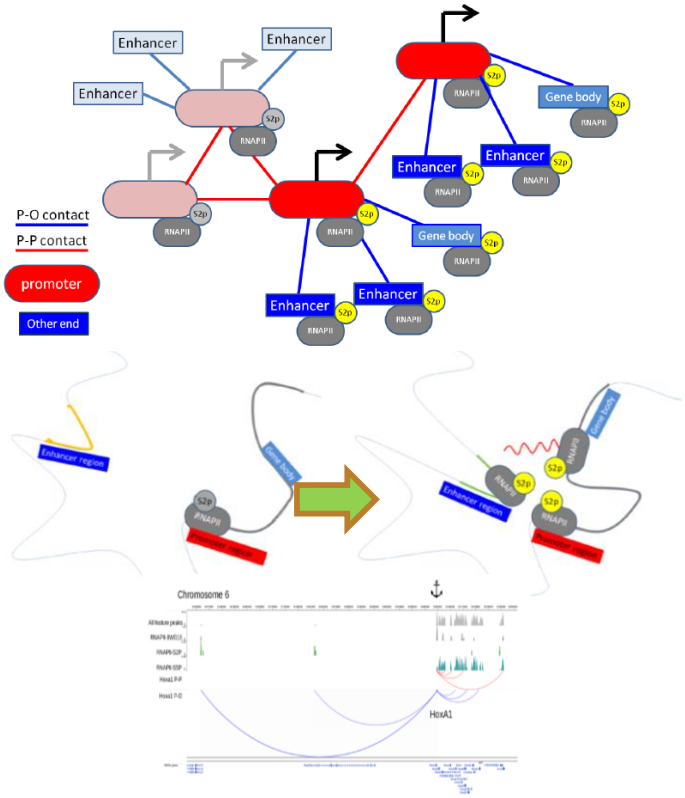
LCC= Large connected component

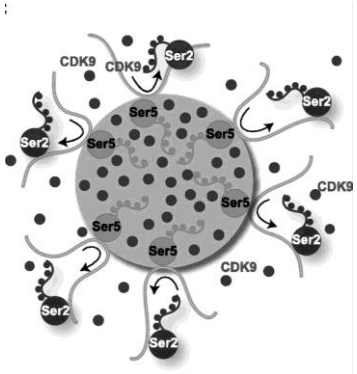
Interpretation:
PCG is a stable hub (across cells, in time?)
RNAPII general dynamic (reflecting transcription regulation?)
RNAPII S2p peripheral

Pancaldi et al. 2016 **Integrating epigenomic data and 3D genomic structure with a new measure of chromatin assortativity**.

The model

Whereas RNAPII S5P accumulates in transcription factories, RNAPII S2p stays peripheral





A model of transcription; gene promoters are loaded with RNAPII-Ser5P (Ser5 light gray) in factories. Elongating RNAPI S2p (Ser2, dark gray) moves to the adjacent nuclear space when it becomes phosphorylated at Ser2 by CDK9

A. Ghamari et al. In vivo live imaging of RNA polymerase II transcription factories in primary cells Genes Dev. 2013;27:767-777

Ghavi-Helm et al. Enhancer loops appear stable during development and are associated with paused polymerase. Nature. 2014;512:96–100.

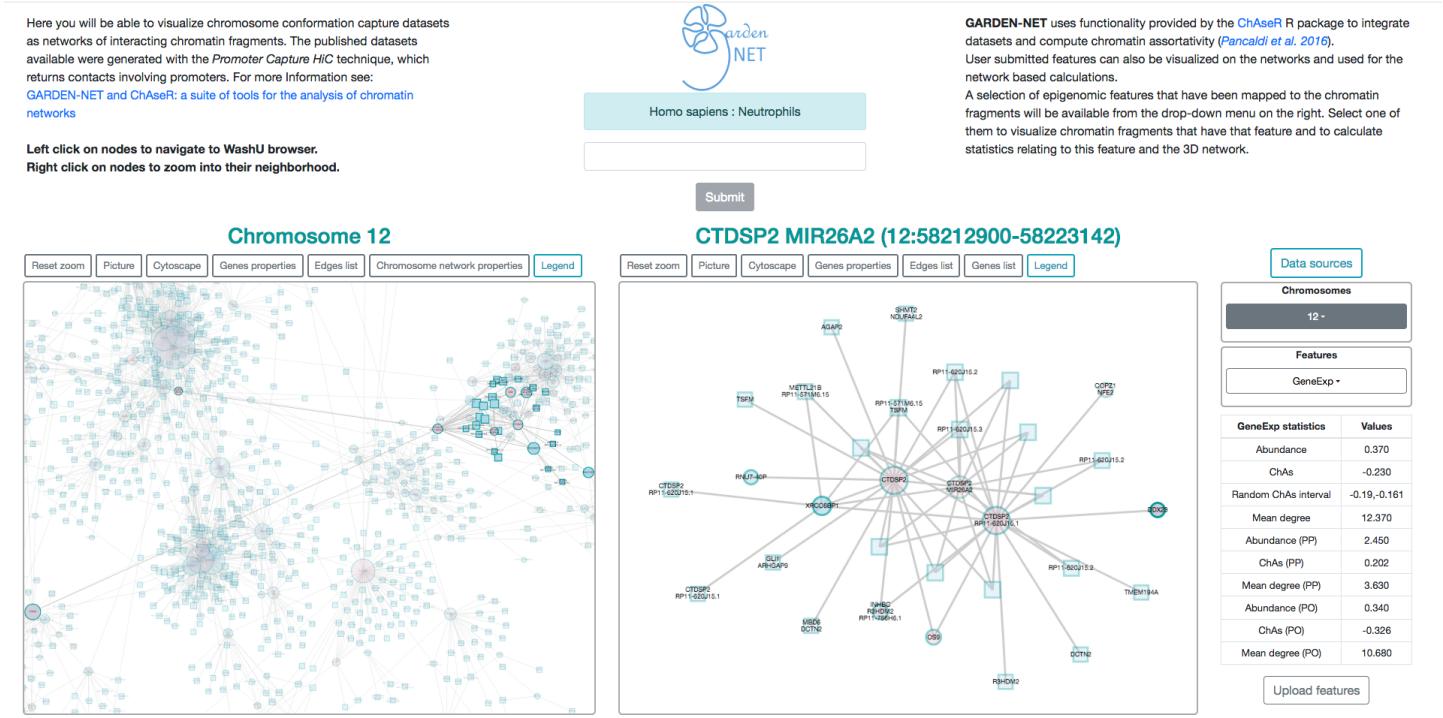
Pancaldi et al. 2016 **Integrating epigenomic data and 3D genomic structure with a new measure of chromatin assortativity**.

Other applications

GARDEN-NET

Genome ARchitecture Data, Epigenome and Nucleome - Network Exploration Tool

https://pancaldi.bsc.es/garden-net

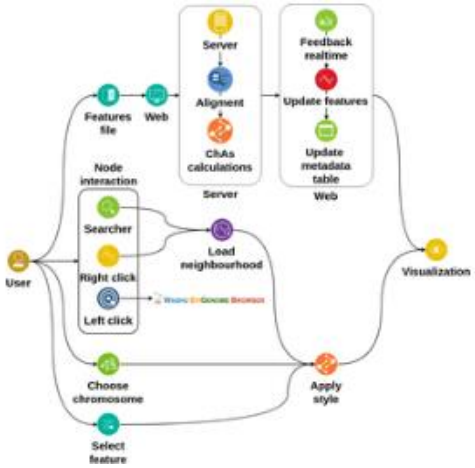


Madrid*, Raineri* and Pancaldi, 2019 GARDEN-NET and ChAseR: a suite of tools for the analysis of chromatin networks BioRxiv 717298, (submitted)

GARDEN-NET

Genome ARchitecture Data, Epigenome and Nucleome - Network Exploration Tool

Interactive and processing in real time



Technologies

Web:

- Yarn (package management)
- Webpack (bundler manager)
- React (Router, Reactrap) with Redux
- Cytoscape.js (network visualization)
- Bootstrap (style)

Server:

- Nginx (HTTP server)
- Python with Flask (REST API)
- R with (ChAseR, Igraph, GenomicRanges, rjson, Tidyverse)
- Bash with pipes, sed & gnu parallel (pipeline integration)
- Celery with redis (task queue system)
- Docker & Docker Compose

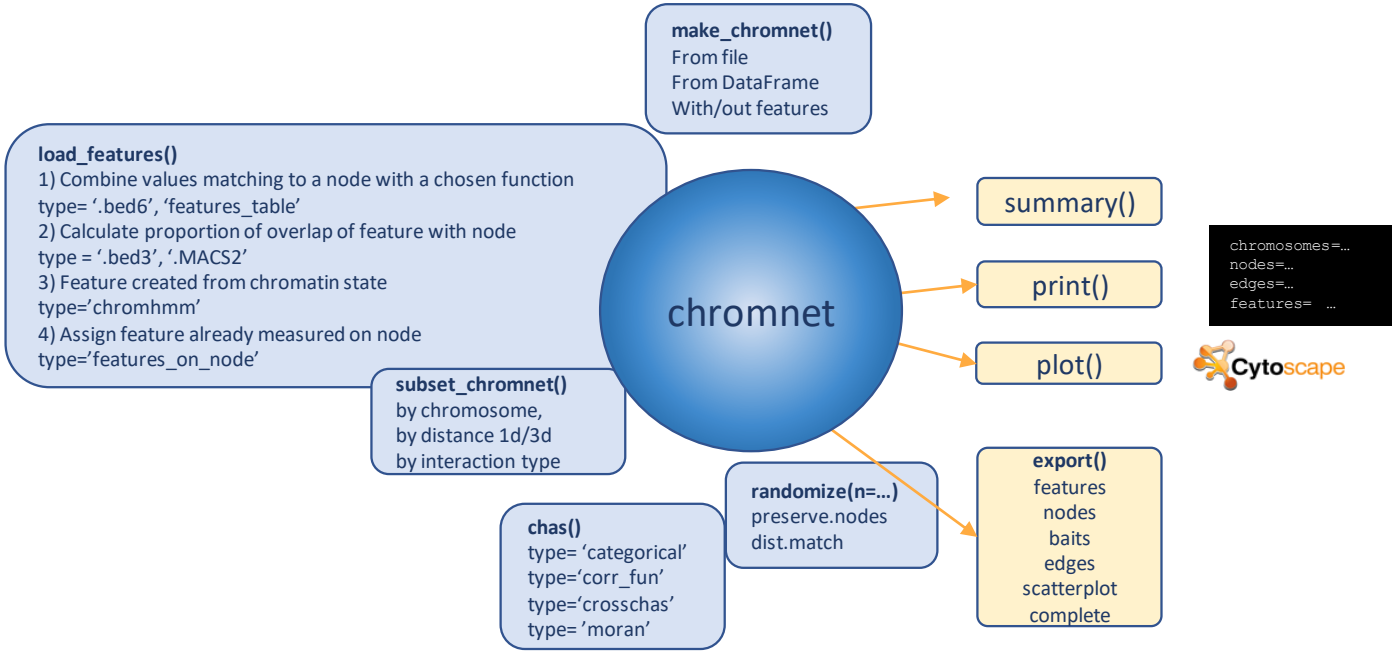
Chromatin Contact networks
Promoter Capture HiC data for haematopoietic cells in human. *Javierre et al. 2016*
Promoter Capture HiC data for mouse embryonic stem cells *Schoenfelder et al. 2015*
Features
Mouse embryonic stem cells histone modifications and 78 ChIP-Seq datasets. *Juan et al. 2016*
GeneExp from *Finotello et al. 2019*
GeneExpEPIVAR for Monocytes, Neutrophils and Tcells from *Chen et al. 2016*
Human Histone modification data: EPIVAR from *Chen et al. 2016*
Human Replication Timing data (GM12878). *Pope et al. 2014*
PCHiC data processed with CHiCAGO. *Cairns et al. 2016*

Technical details at **https://github.com/VeraPancaldiLab/GARDEN-NET**

Madrid*, Raineri* and Pancaldi, 2019 GARDEN-NET and ChAseR: a suite of tools for the analysis of chromatin networks BioRxiv 717298, (submitted)

ChAseR: an R package

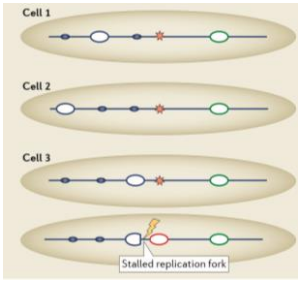
https://bitbucket.org/eraineri/chaser/

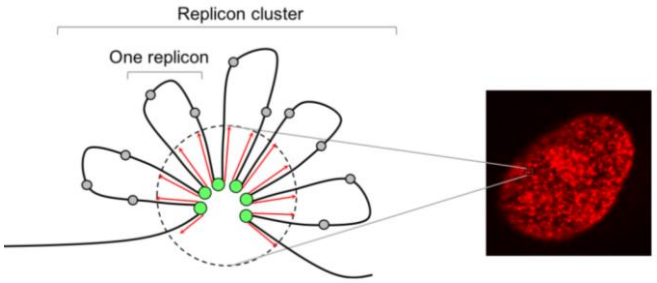


Madrid*, Raineri* and Pancaldi, 2019 GARDEN-NET and ChAseR: a suite of tools for the analysis of chromatin networks BioRxiv 717298, (in revision)

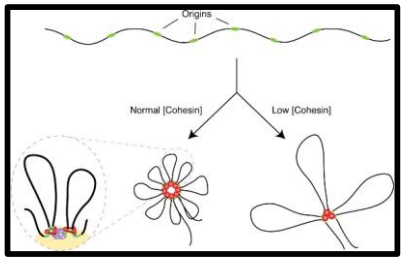
Mammalian DNA replication in 3D

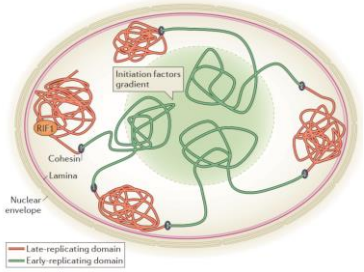
Stochastic firing model: constitutive, flexible, dormant origins





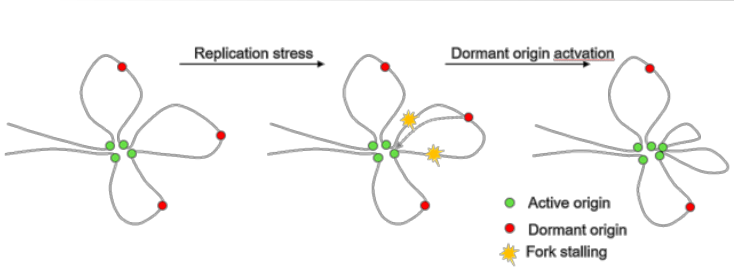
Cohesin might mediate replicon loops, which assemble in Replication Domains (coinciding with TADs)

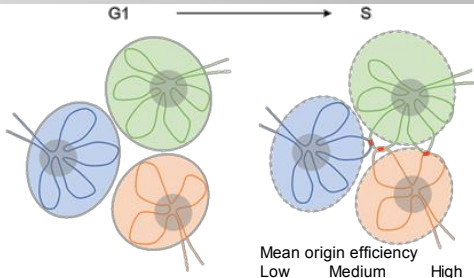


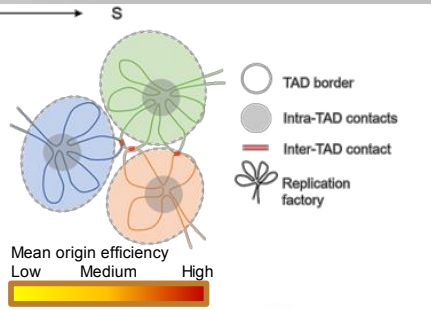


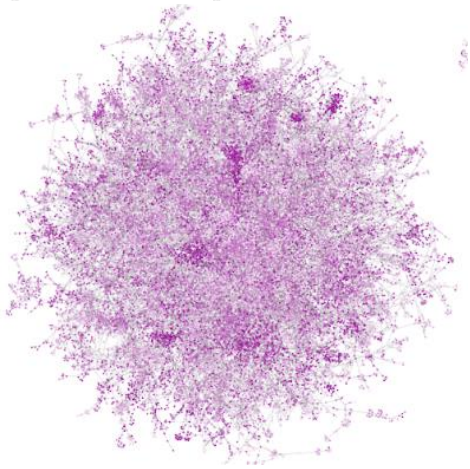
Fragkos et al. **DNA replication origin activation in space and time**. Nat. Rev. Mol. Cell Biol. 2015
Guillou, E. et al. **Cohesin organizes chromatin loops at DNA replication factories**. Genes Dev. 2010

A global view of replication in 3D in mouse



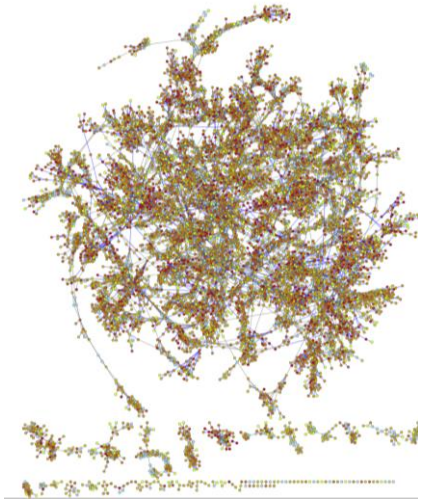






Replication Timing in 3D

?
variability



Origin efficiency in 3D

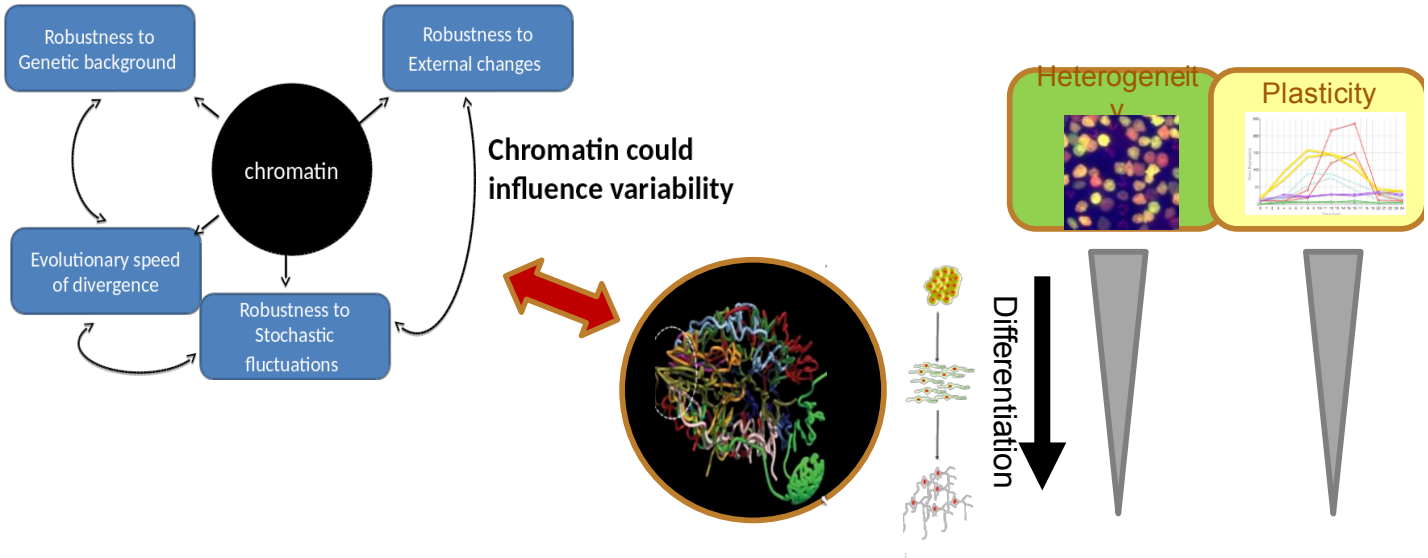
Jodkowska, Pancaldi et al unpublished

Perspectives

Chromatin, heterogeneity, plasticity, stemness…

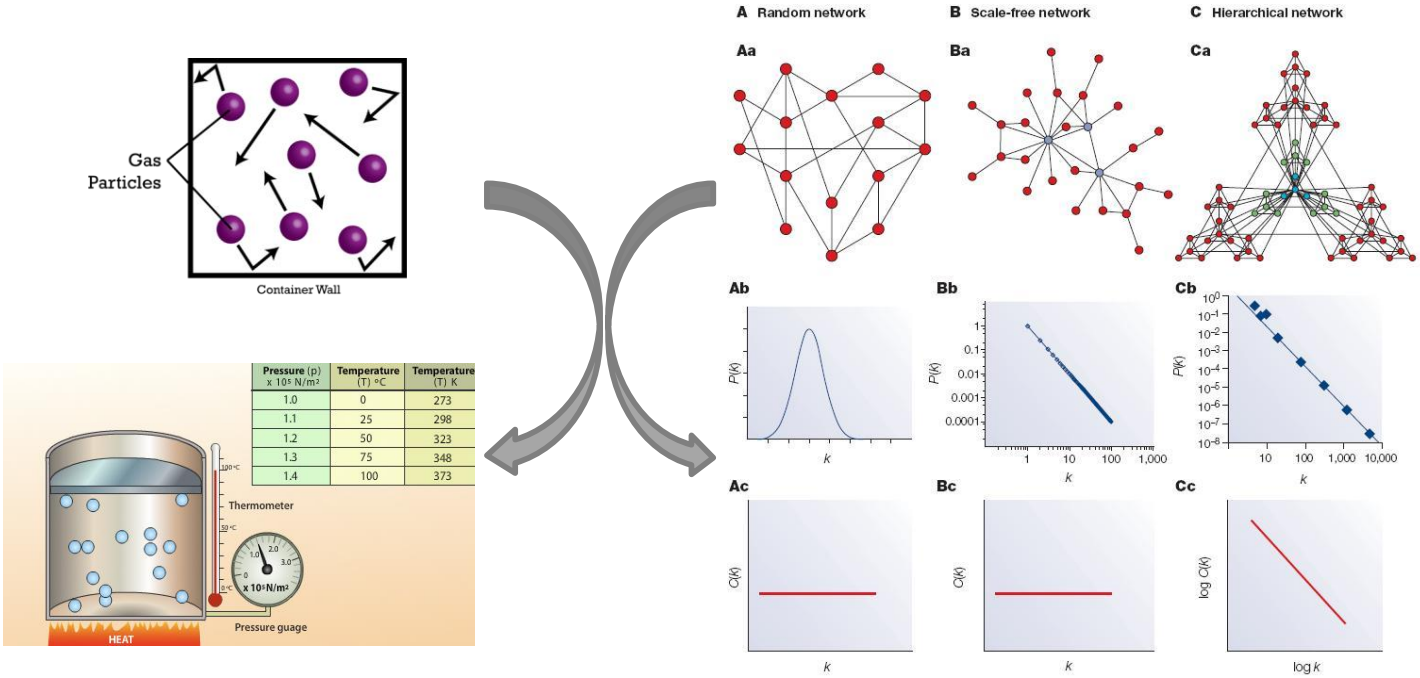
Chromatin state (methylation/histone modifications etc…) can affect

- Plasticity (rapidly regulated stress genes)
- Single cell heterogeneity (noisy promoters)
- Inter-individual differences
- Evolutionary divergence speed



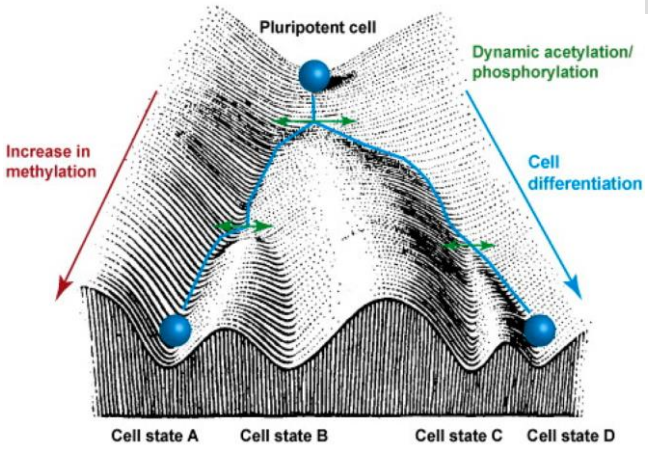
A systems approach

Thermodynamics: from particles' positions and velocities to pressure and temperature



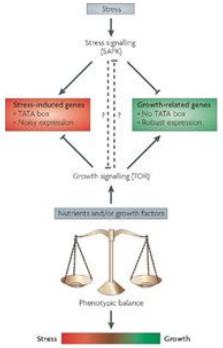
Network theory: from nodes and edges to degree distribution, clustering coefficient, …
Interdisciplinary approach: borrow concepts from studies on other networks

Cellular differentiation and response

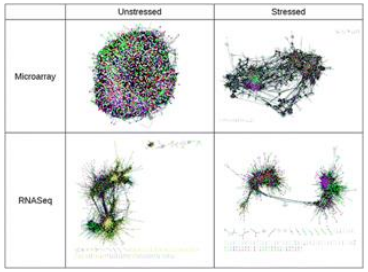


Isogenic heterogeneity
Large fluctuations, motion across landscape
Plasticity – in response to external signals

Reduced heterogenetity/plasticity
Well-defined phenotype



Expression Correlation network



System level Stress Response
Network disaggregates into modules
Increase in heterogeneity

Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation López-Maury L et al. *Nat. Rev. Gen.* 2008.
Pancaldi V, Schubert F and Bahler J. **Meta-analysis of genome regulation and expression variability across hundreds of environmental and genetic perturbations in fission yeast**. *Mol. BioSyst.*, 2010.
Stress induces remodelling of yeast interaction and co-expression networks Lehtinen S, Marsellach FX, Codlin S, Schmidt A, Clément-Ziza M, Beyer A, Bähler J, Orengo C, Pancaldi V. *Mol. BioSyst.*, 2013.

MERCI!

Barcelona Supercomputing Center
Maria Rigau
Alfonso Valencia

CNAG Barcelona
Emanuele Raineri

Newcastle University
Daniel Rico

IMDEA, Madrid
Enrique Carrillo

Universidad Pompeu Fabra
David Juan

Babraham Institute
Biola-Maria Javierre
Mikhail Spivakov
Peter Fraser

CNIO, Madrid
Karolina Jodkowska
Juan Méndez
Osvaldo Graña
Enrique Carrillo
Miriam Rubio

Centro de Biología Molecular "Severo Ochoa"
CSIC-UAM
Jose' Miguel Fernandez
Ricardo Almeida
Sara Rodríguez-Acebes
María Gómez

Team 21 CRCT:

Epigenomics and network modelling of heterogeneity in immuno-oncology


Miguel Madrid Mencia


Ting Xie


Nina Verstraete

Bích Ngọc Cao Trần

Flavien Raynal
(Bystricky lab)

Pierre Fabre

FONDATION TOULOUSE CANCER SANTÉ

Inserm

BSC