3DGenomics

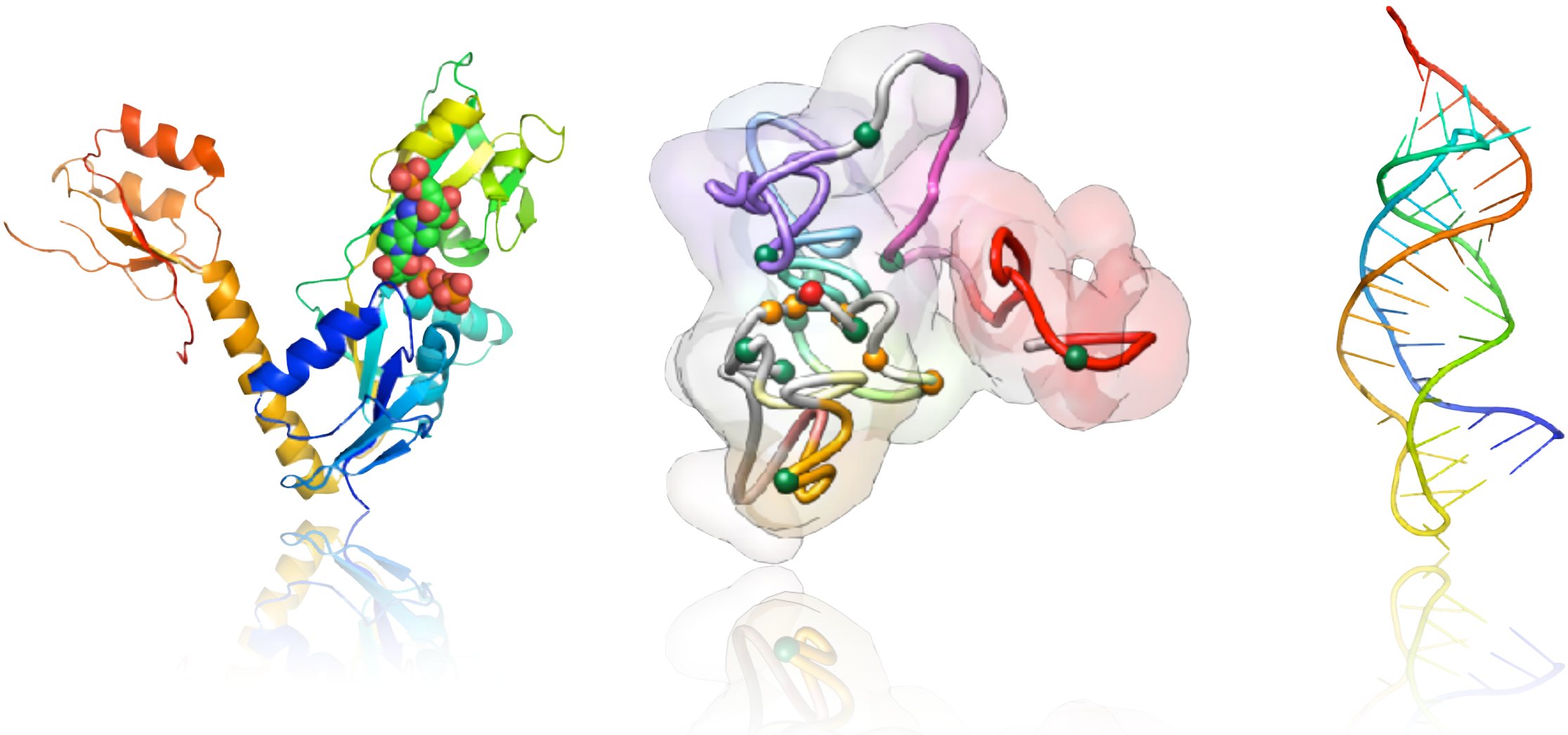
Marc A. Marti-Renom
Genome Biology Group (CNAG)
Structural Genomics Group (CRG)

ICREA

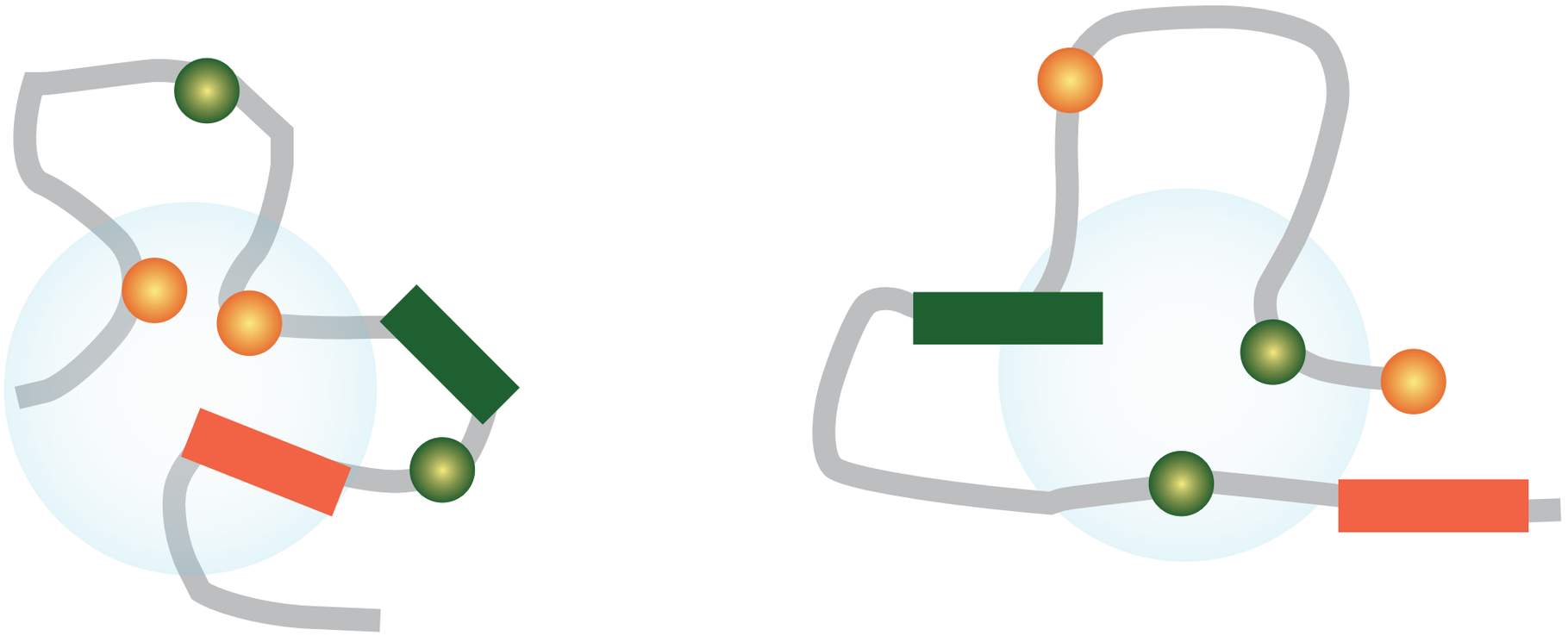
cnag CRG Centre for Genomic Regulation

Structural Genomics Group

http://www.marciuslab.org



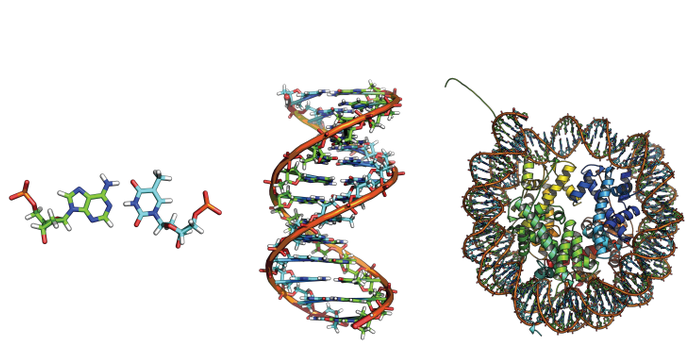
Complex genome organization

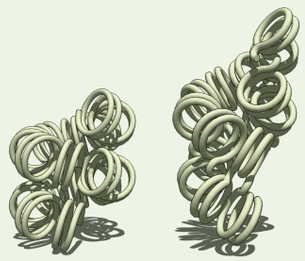


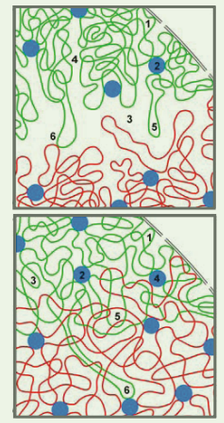
Resolution Gap

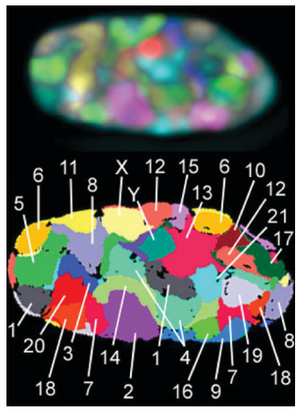
Marti-Renom, M. A. & Mirny, L. A. PLoS Comput Biol 7, e1002125 (2011)

Knowledge









DNA length

10^{0} 10^{3} 10^{6} 10^{9} nt

Volume

10^{-9} 10^{-6} 10^{-3} 10^{0} 10^{3} μm^{3}

Time

10^{-10} 10^{-8} 10^{-6} 10^{-4} 10^{-2} 10^{0} 10^{2} 10^{3} s

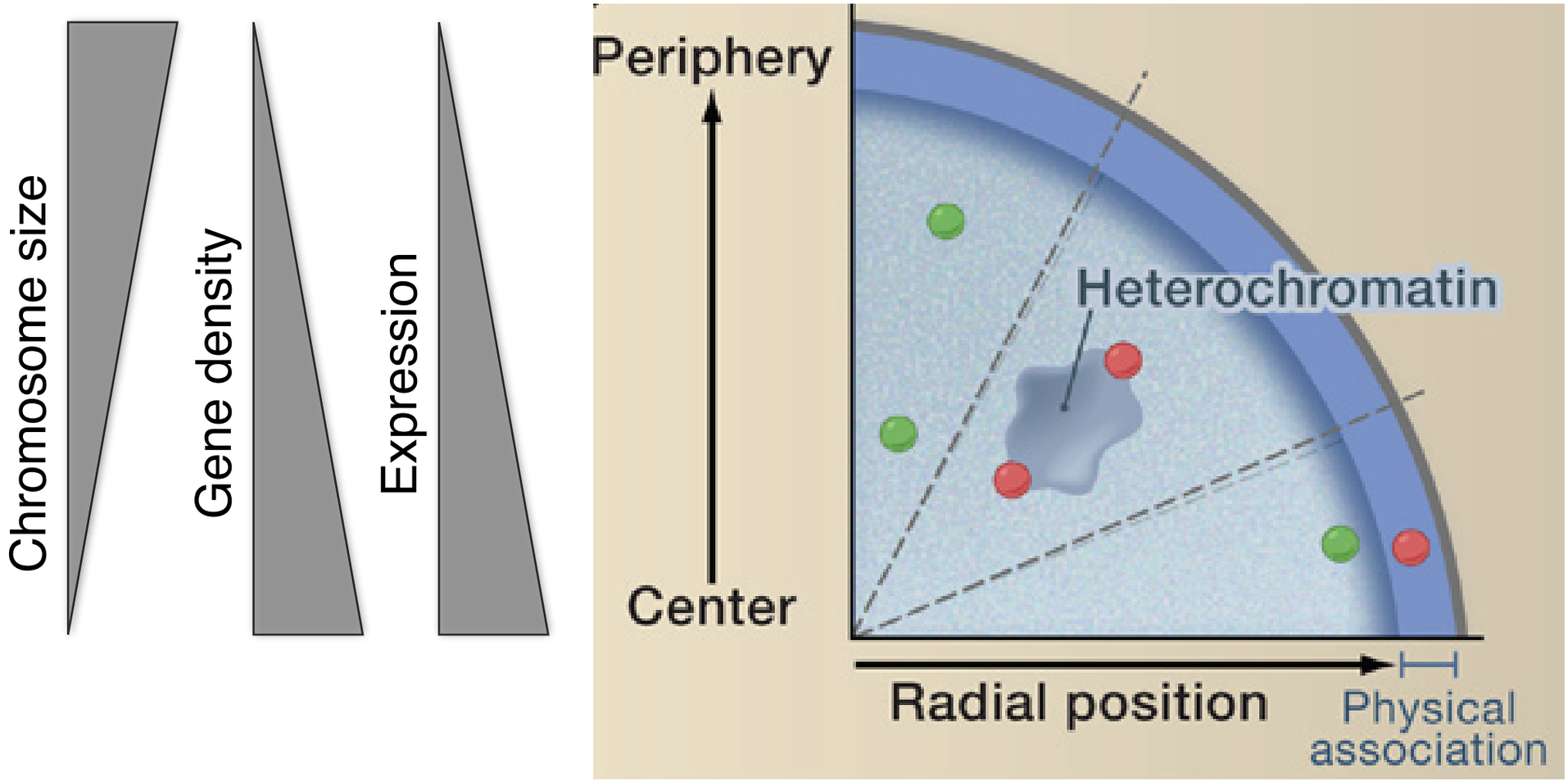
Resolution

10^{-3} 10^{-2} 10^{-1} μ

cnag CRG Centre for Genomic Regulation

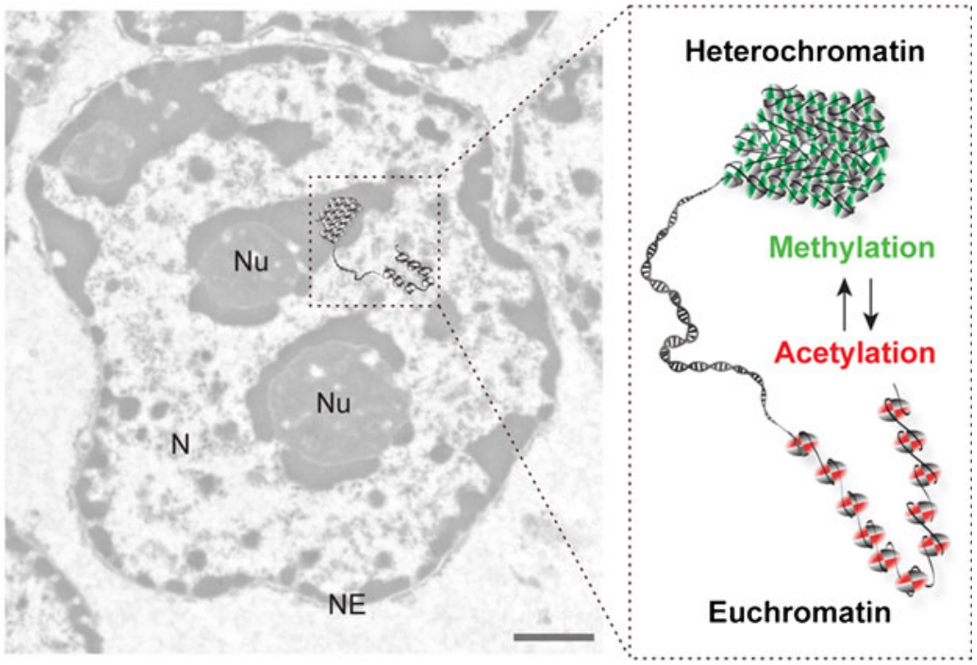
Level I: Radial genome organization

Takizawa, T., Meaburn, K. J. & Misteli, T. The meaning of gene positioning. Cell 135, 9–13 (2008).

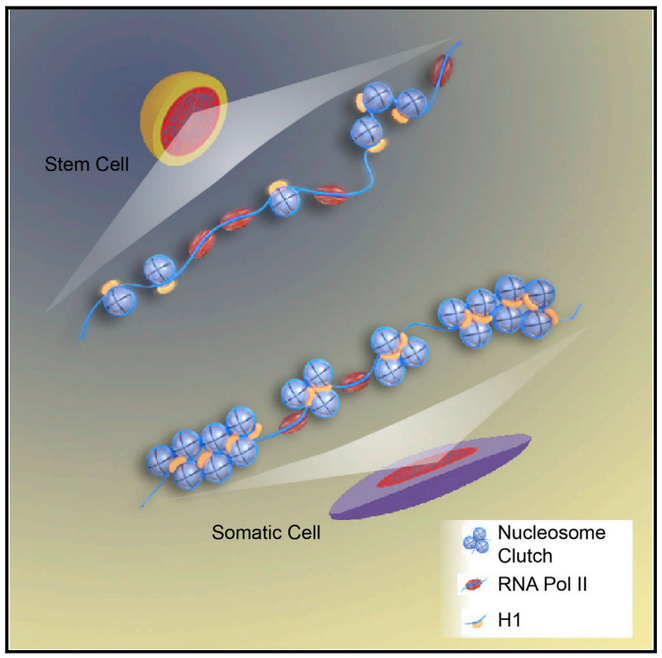


Level II: Euchromatin vs heterochromatin

Electron microscopy



Nanoscopy

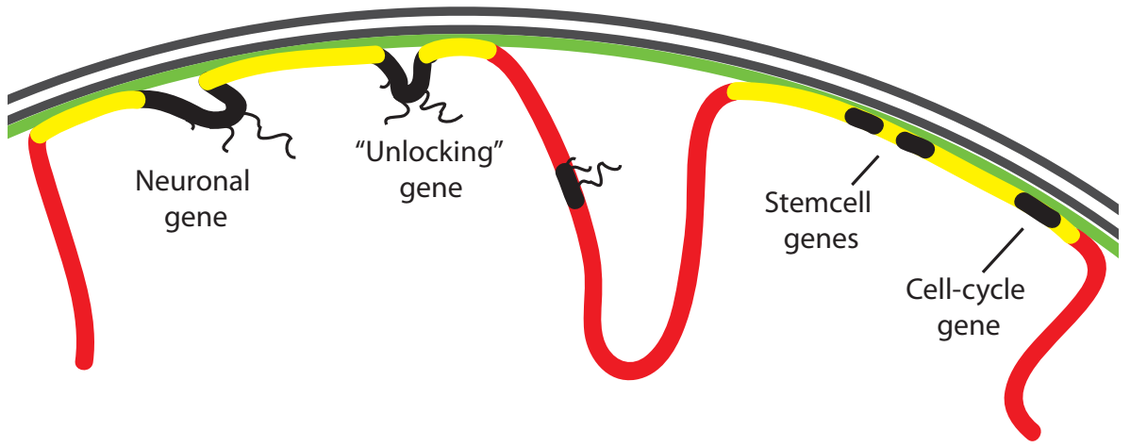


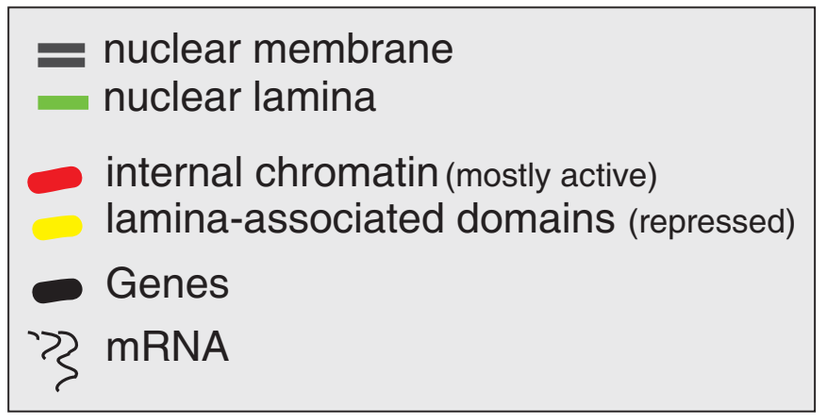
Euchromatin:
chromatin that is located away from the nuclear lamina, is generally less densely packed, and contains actively transcribed genes

Heterochromatin:
chromatin that is near the nuclear lamina, tightly condensed, and transcriptionally silent

Adapted from Cell, 160(6), 1145-1158. 2015

Level III: Lamina-genome interactions



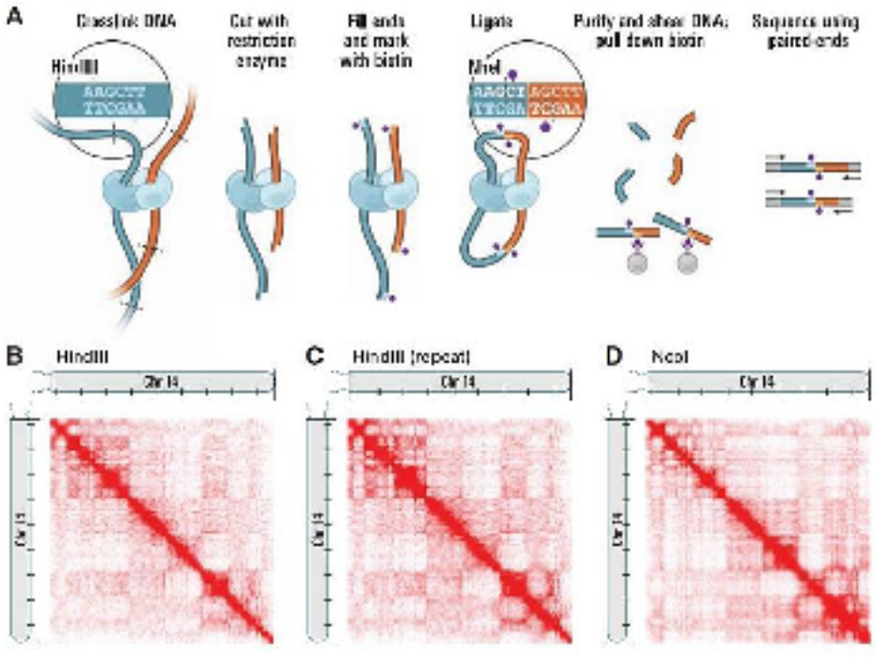


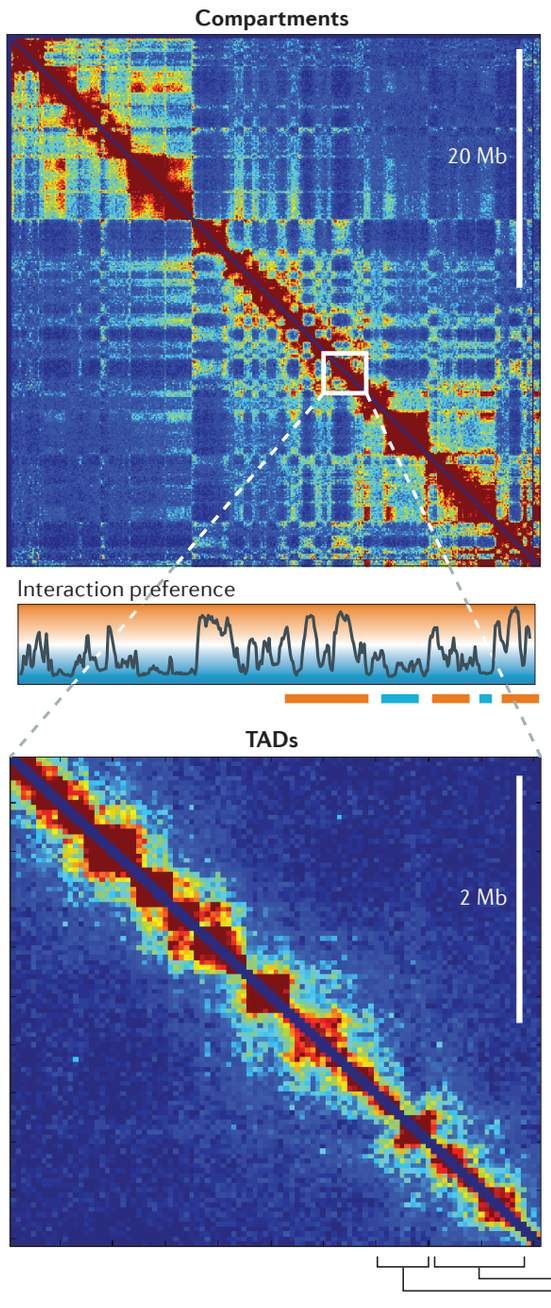
Most genes in Lamina Associated Domains are transcriptionally silent, suggesting that **lamina-genome interactions** are widely involved in the control of **gene expression**

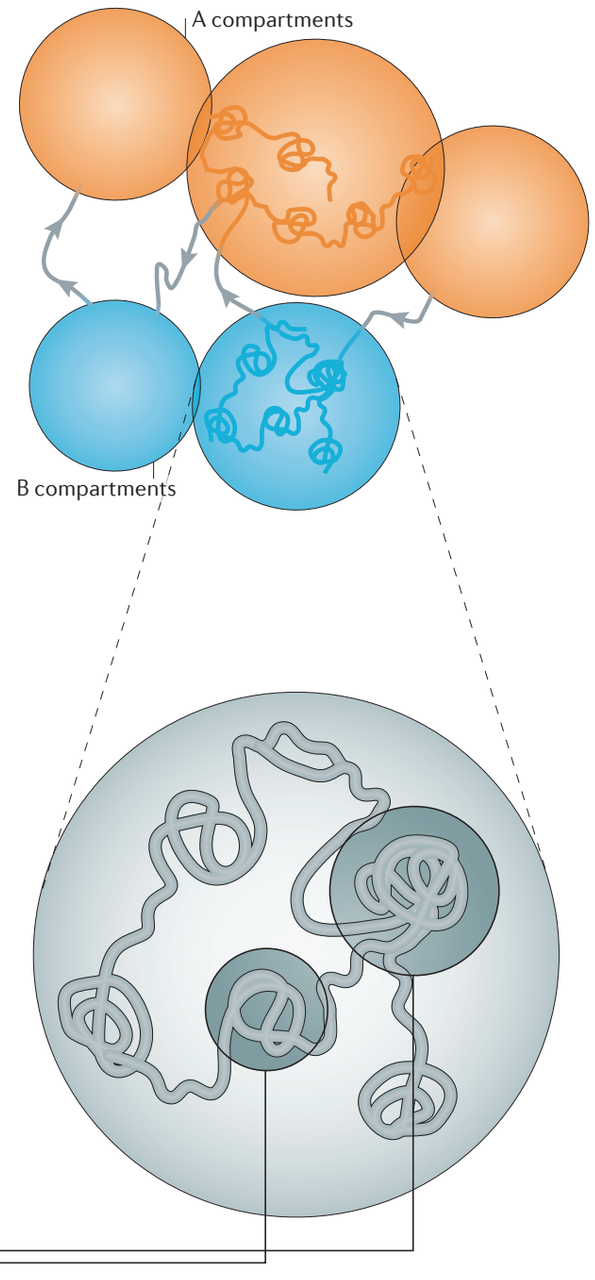
Adapted from Molecular Cell 38, 603-613, 2010

Level IV: Higher-order organization

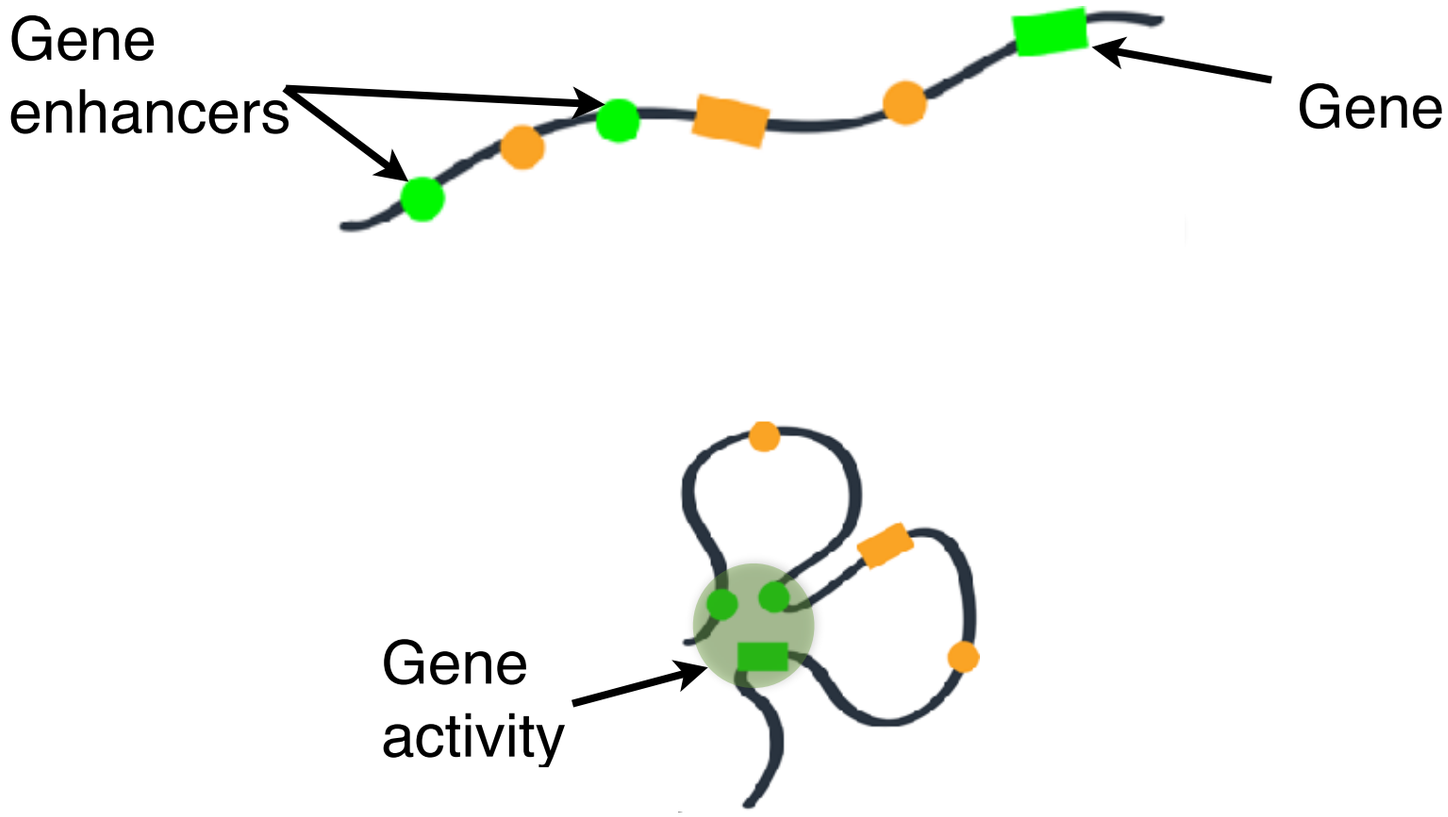
Dekker, J., Marti-Renom, M. A. & Mirny, L. A. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. Nat Rev Genet 14, 390–403 (2013).







Level V: Chromatin loops



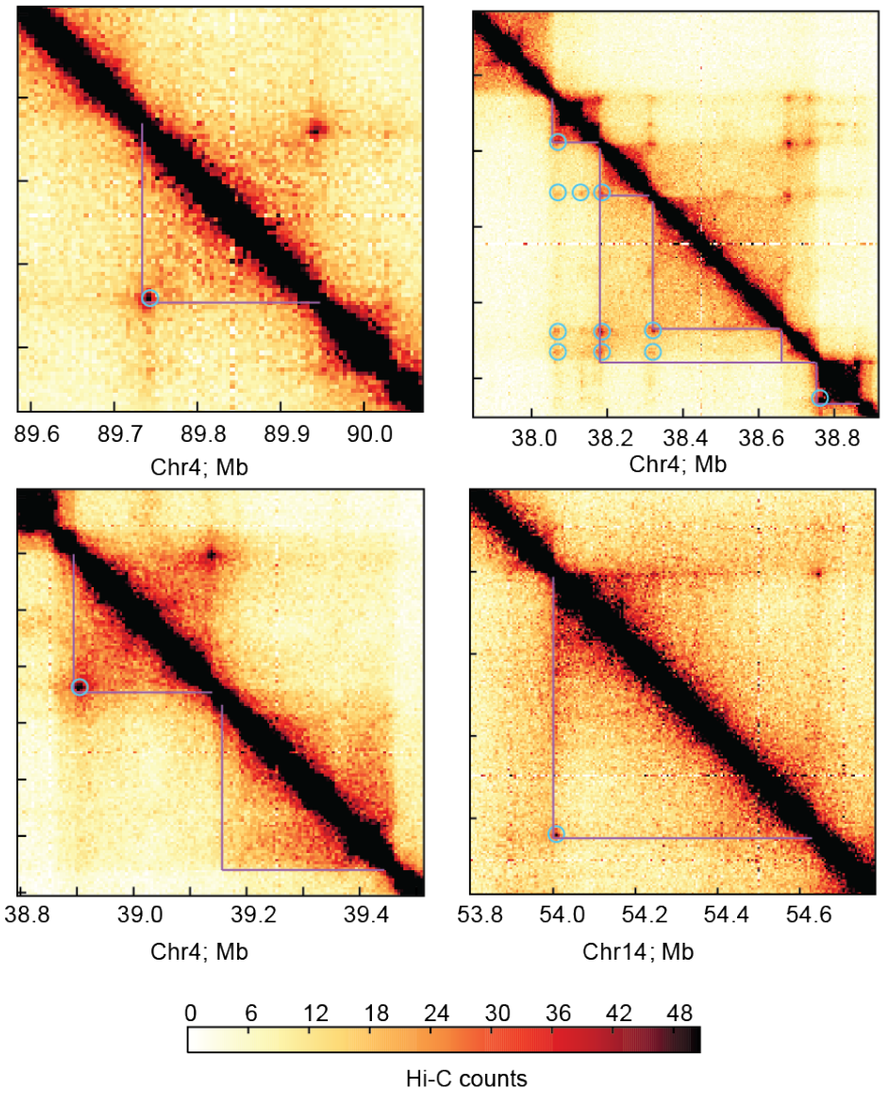
Loops bring distal genomic regions in close proximity to one another

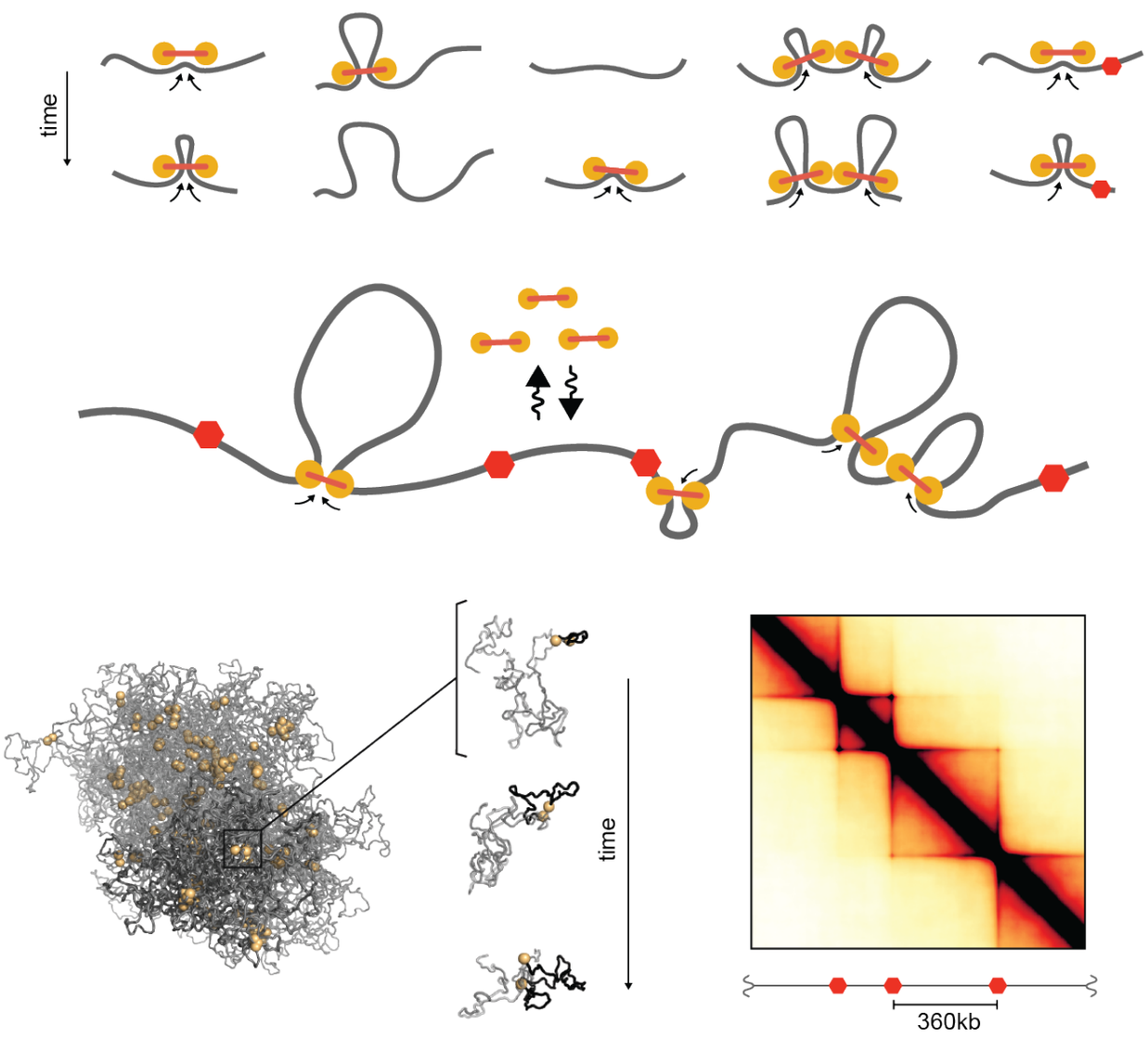
This in turn can have profound effects on gene transcription

Enhancers can be thousands of kilobases away from their target genes in any direction (or even on a separate chromosome)

Level V: Loop-extrusion as a driving force

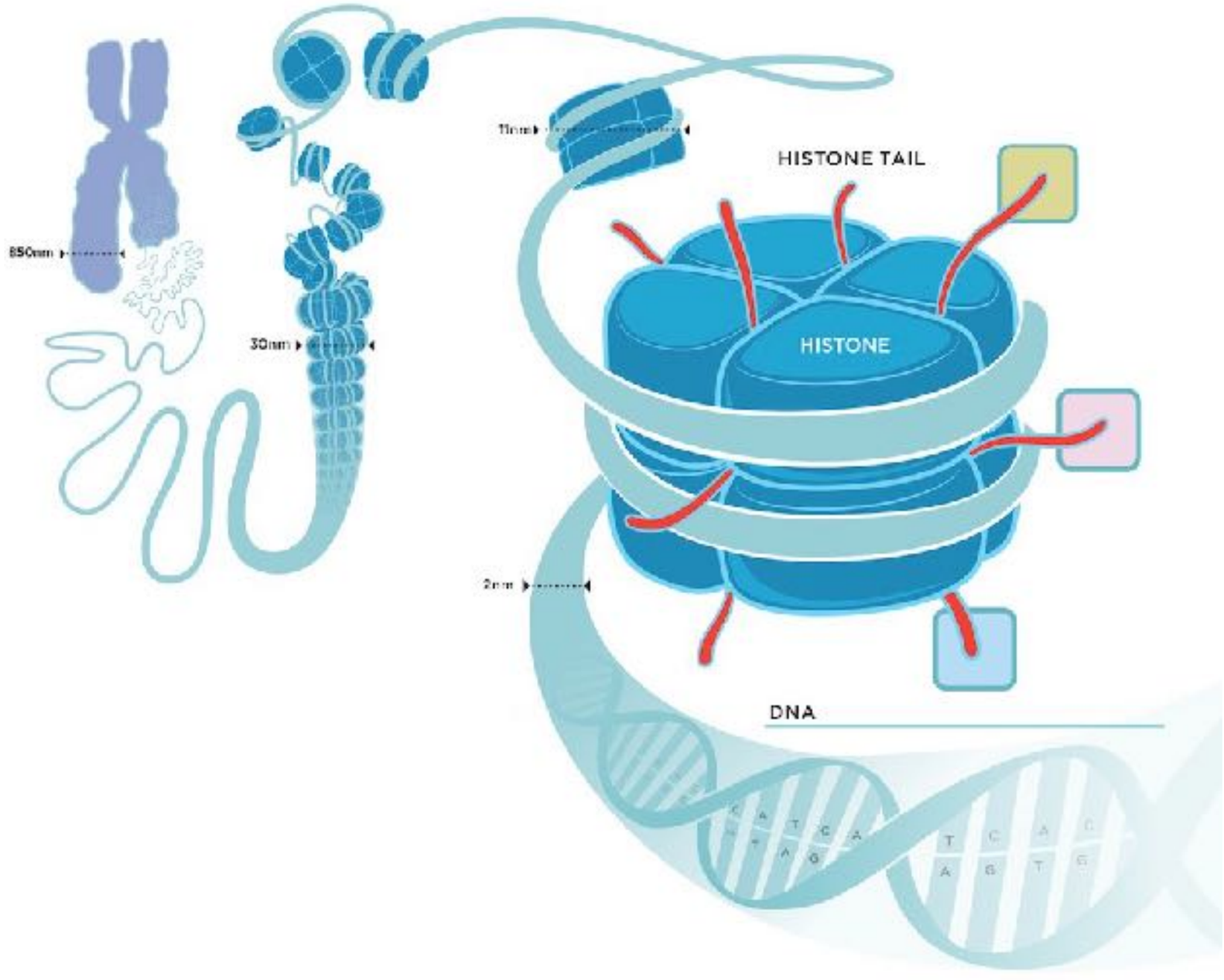
Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., & Mirny, L. A. (2015).
Formation of Chromosomal Domains by Loop Extrusion. bioRxiv.





Level VI: Nucleosome

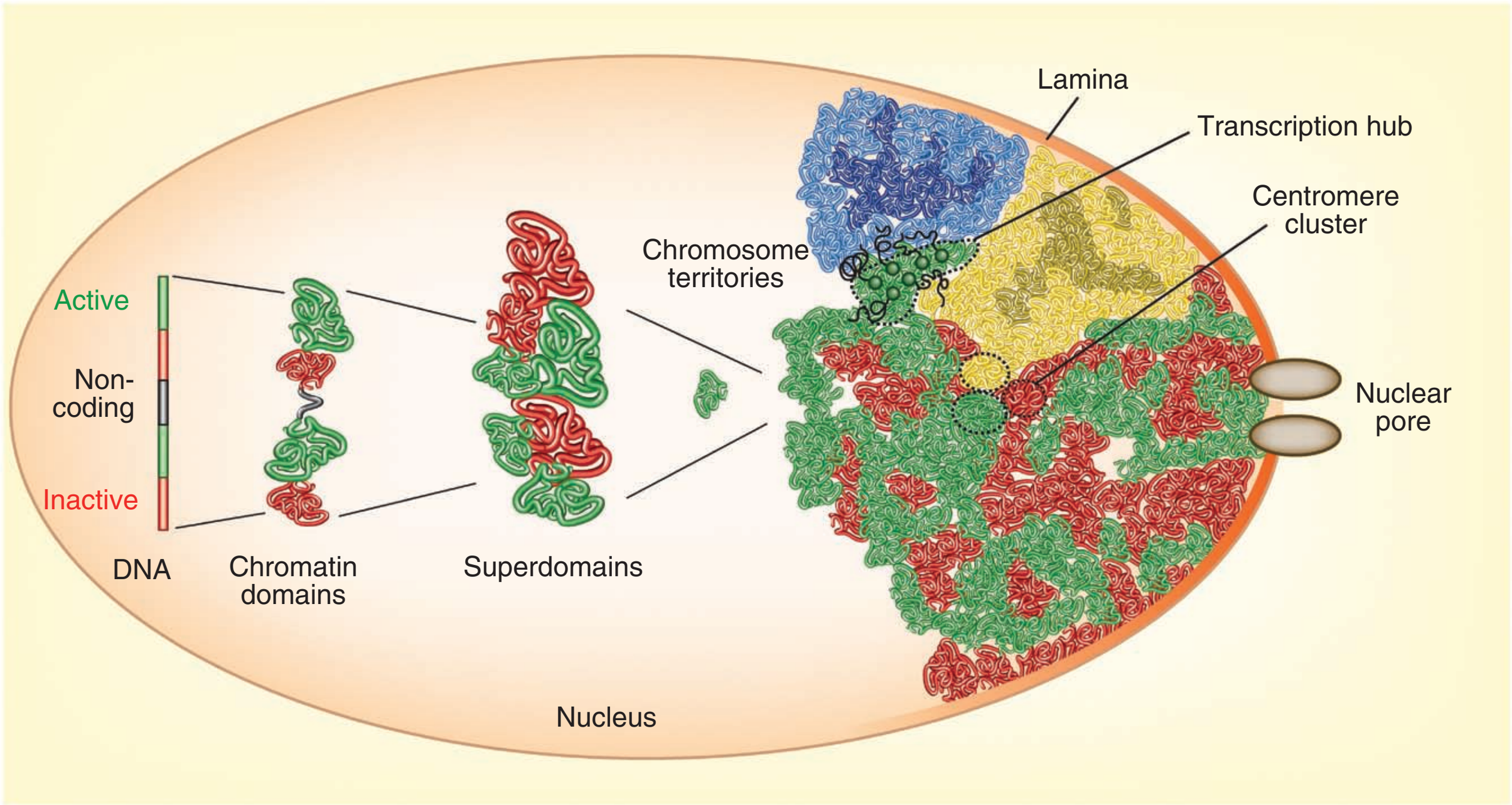
Chromosome **Chromatin fibre** **Nucleosome**



Adapted from Richard E. Ballermann, 2012

Complex genome organization

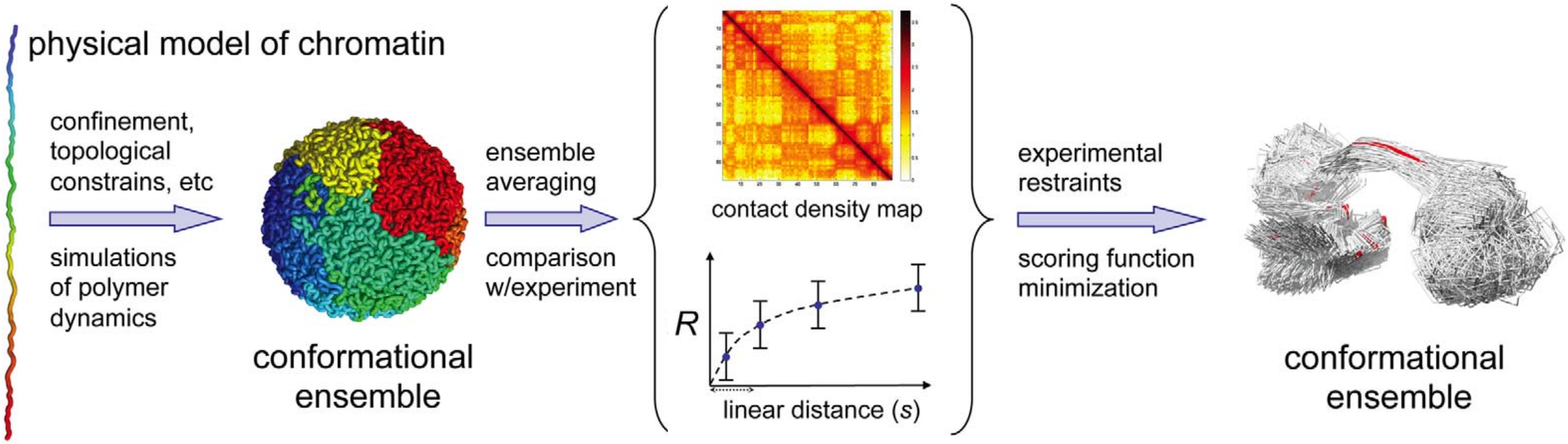
Cavalli, G. & Misteli, T. Functional implications of genome topology. Nat Struct Mol Biol 20, 290–299 (2013).



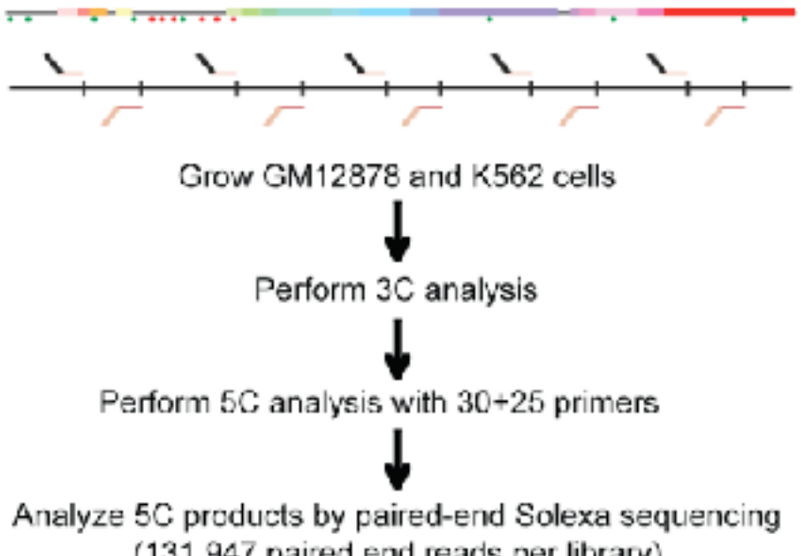
Marina Corral

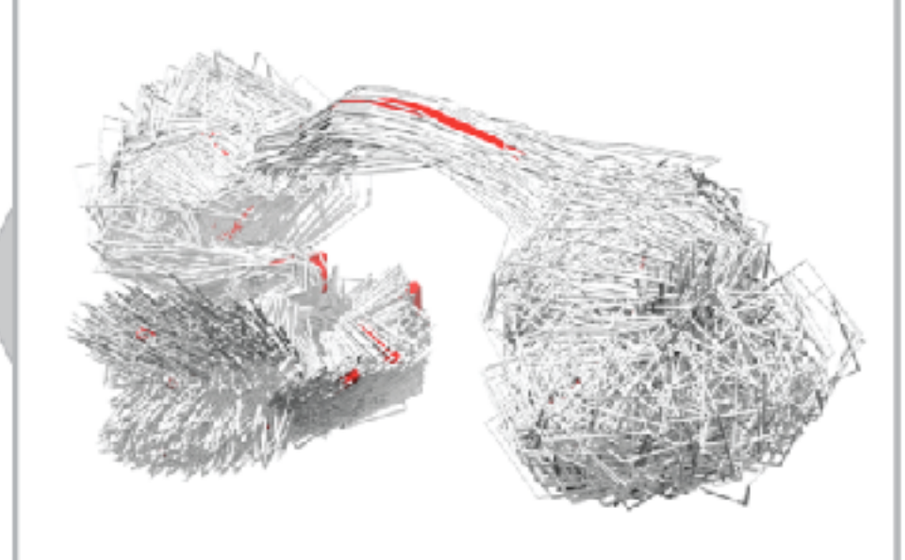
Modeling Genomes

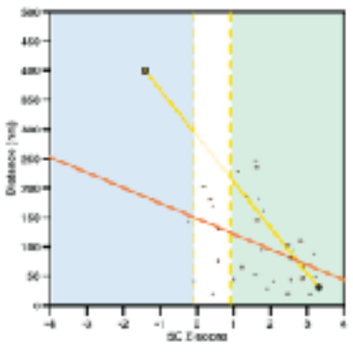
Marti-Renom, M. A. & Mirny, L. A. PLoS Comput Biol 7, e1002125 (2011)

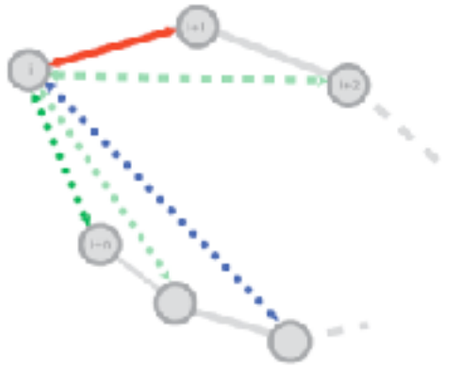


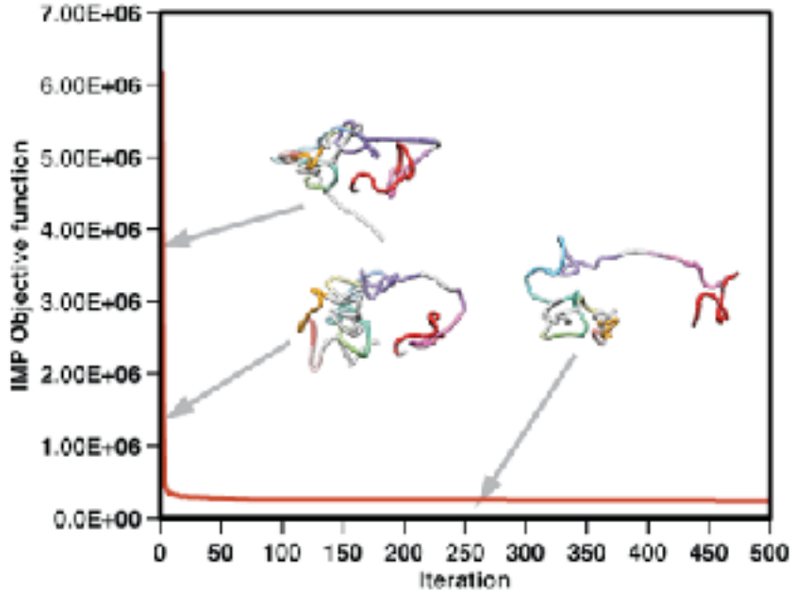
Experiments



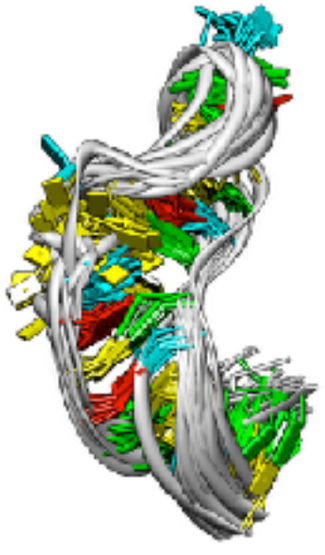


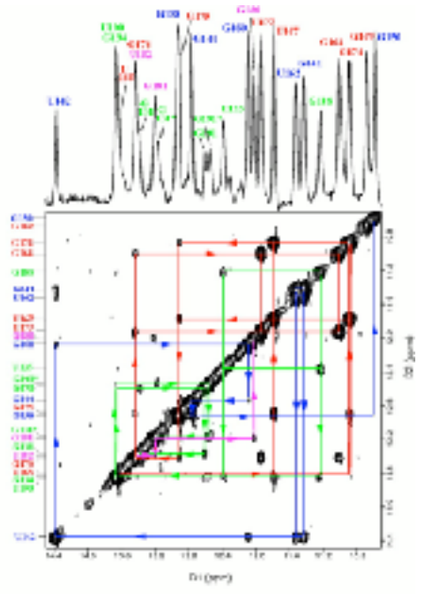






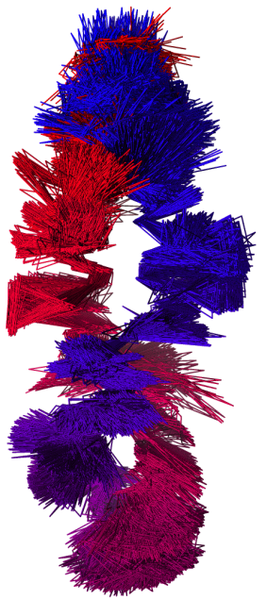
Computation

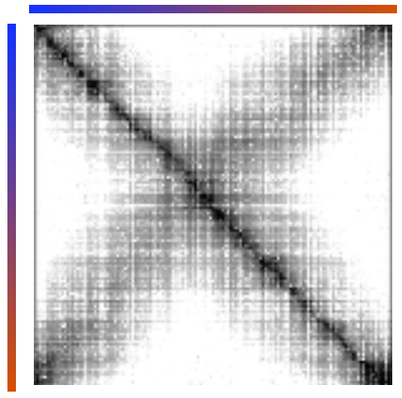




Biomolecular structure determination

2D-NOESY data

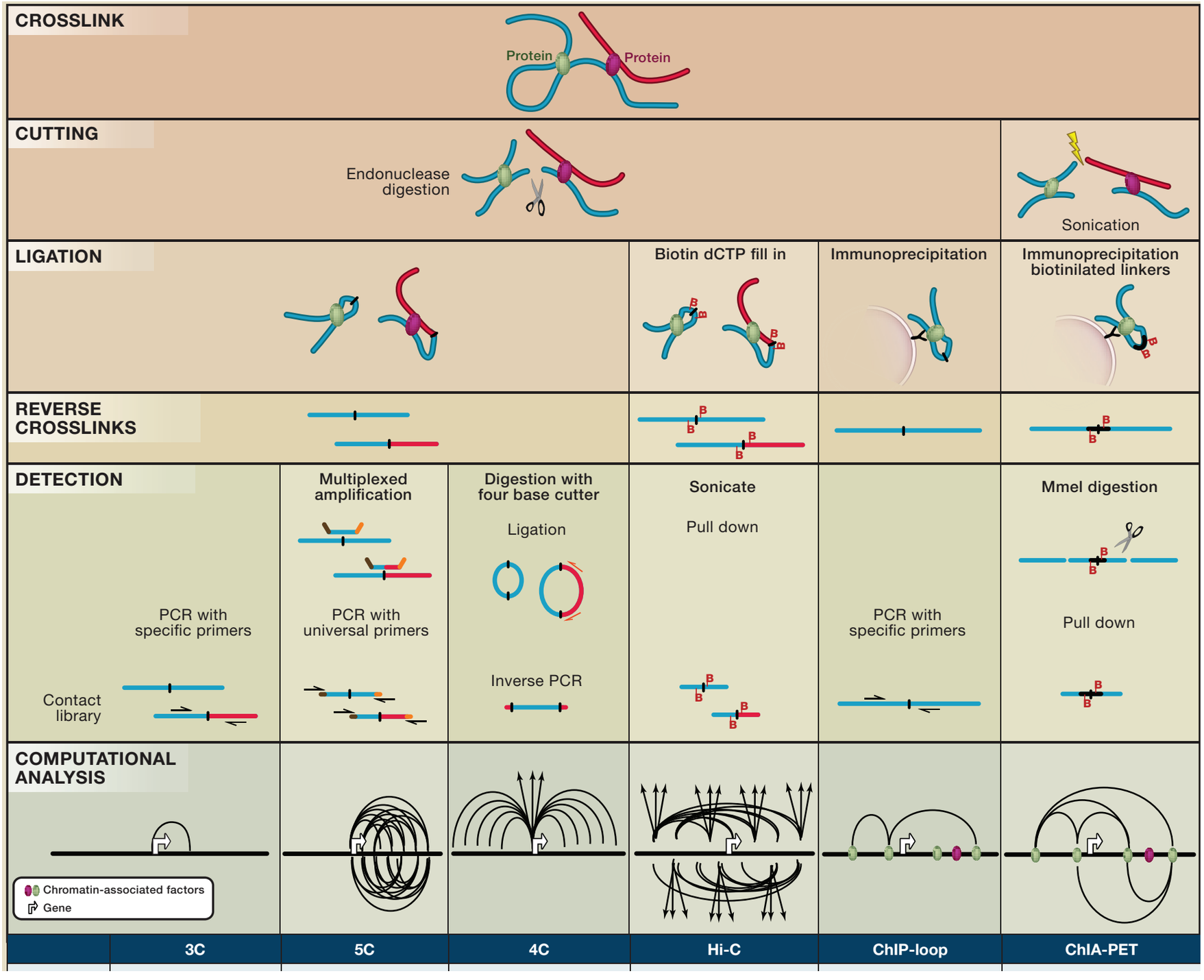




Chromosome structure determination

5C data

Chromosome Conformation Capture



Hakim, O., & Misteli, T. (2012). SnapShot: Chromosome Confirmation Capture. Cell, 148(5), 1068–1068.e2.

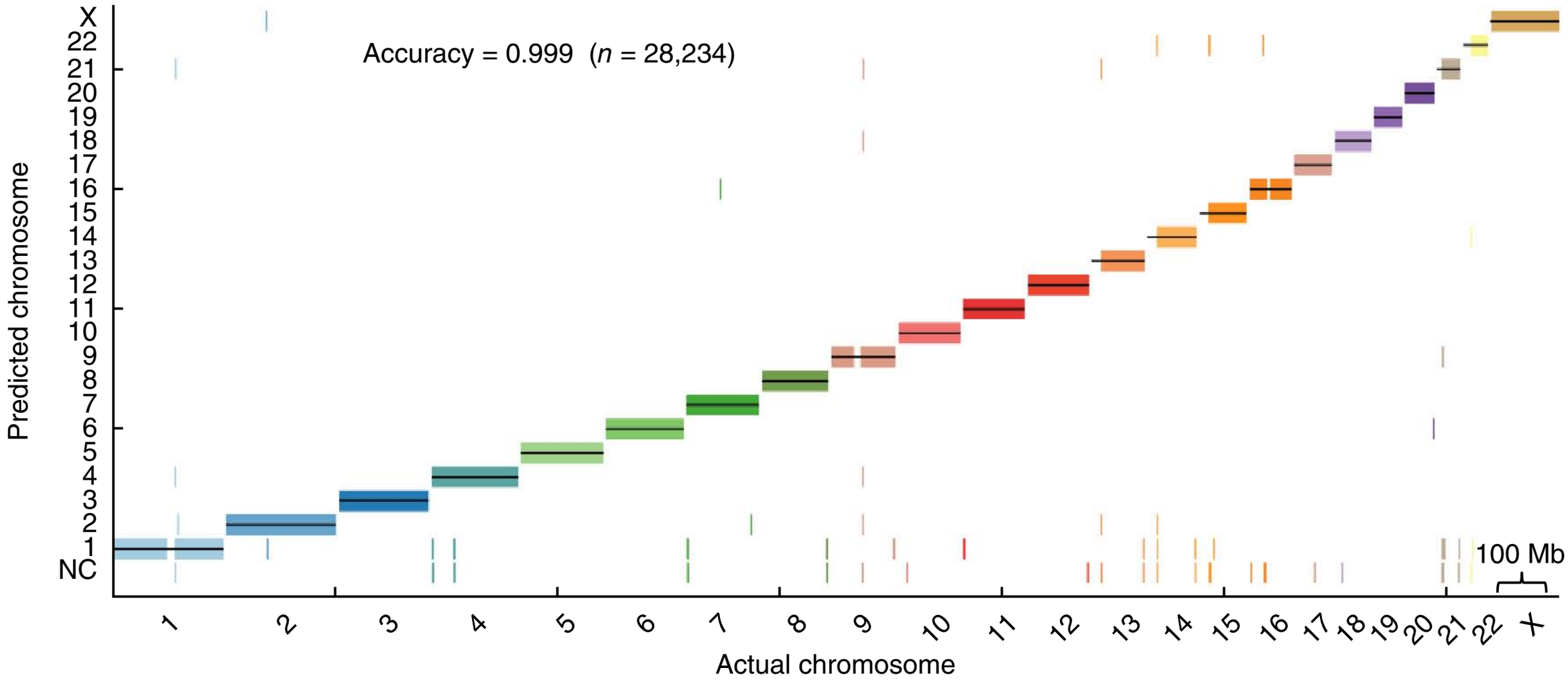
Chromosome Conformation Capture

	3C	5C	4C	Hi-C	ChIP-loop	ChIA-PET
Principle	Contacts between two defined regions[3,17]	All against all[4,18]	All contacts with a point of interest[14]	All against all[10]	Contacts between two defined regions associated with a given protein[8]	All contacts associated with a given protein[6]
Coverage	Commonly < 1Mb	Commonly < 1Mb	Genome-wide	Genome-wide	Commonly < 1Mb	Genome-wide
Detection	Locus-specific PCR	HT-sequencing	HT-sequencing	HT-sequencing	Locus-specific qPCR	HT-sequencing
Limitations	Low throughput and coverage	Limited coverage	Limited to one viewpoint		Rely on one chromatin-associated factor, disregarding other contacts	
Examples	Determine interaction between a known promoter and enhancer	Determine comprehensively higher-order chromosome structure in a defined region	All genes and genomic elements associated with a known LCR	All intra- and interchromosomal associations	Determine the role of specific transcription factors in the interaction between a known promoter and enhancer	Map chromatin interaction network of a known transcription factor
Derivatives	PCR with TaqMan probes[7] or melting curve analysis[1]		Circular chromosome conformation capture[20], open-ended chromosome conformation capture[19], inverse 3C[12], associated chromosome trap (ACT)[11], affinity enrichment of bait-ligated junctions[2]	Yeast[5,15], tethered conformation capture[9]		ChIA-PET combined 3C-ChIP-cloning (6C);[16] enhanced 4C (e4C)[13]

Hakim, O., & Misteli, T. (2012). SnapShot: Chromosome Confirmation Capture. Cell, 148(5), 1068–1068.e2.

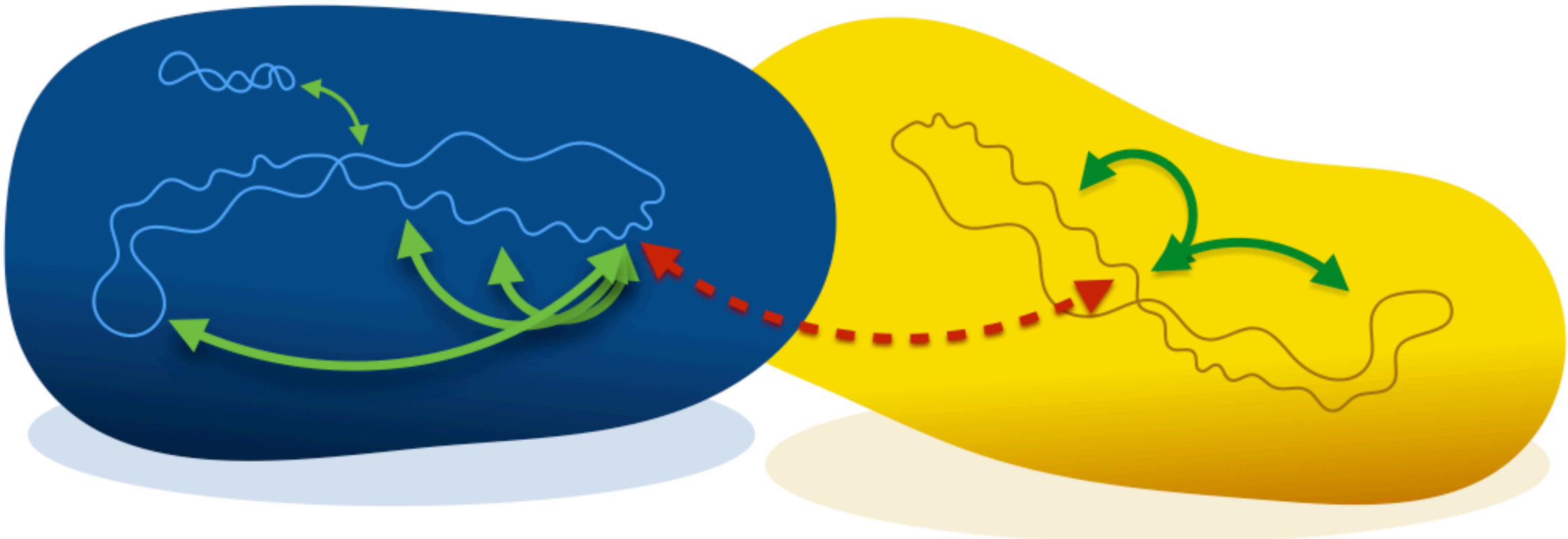
... and one more thing

Chromosome Conformation Capture for de-novo assembly



Kaplan, N., & Dekker, J. (2013). High-throughput genome scaffolding from in vivo DNA interaction frequency. Nature Biotechnology, 31(12), 1143–1147.

Chromosome Conformation Capture for meta genomics

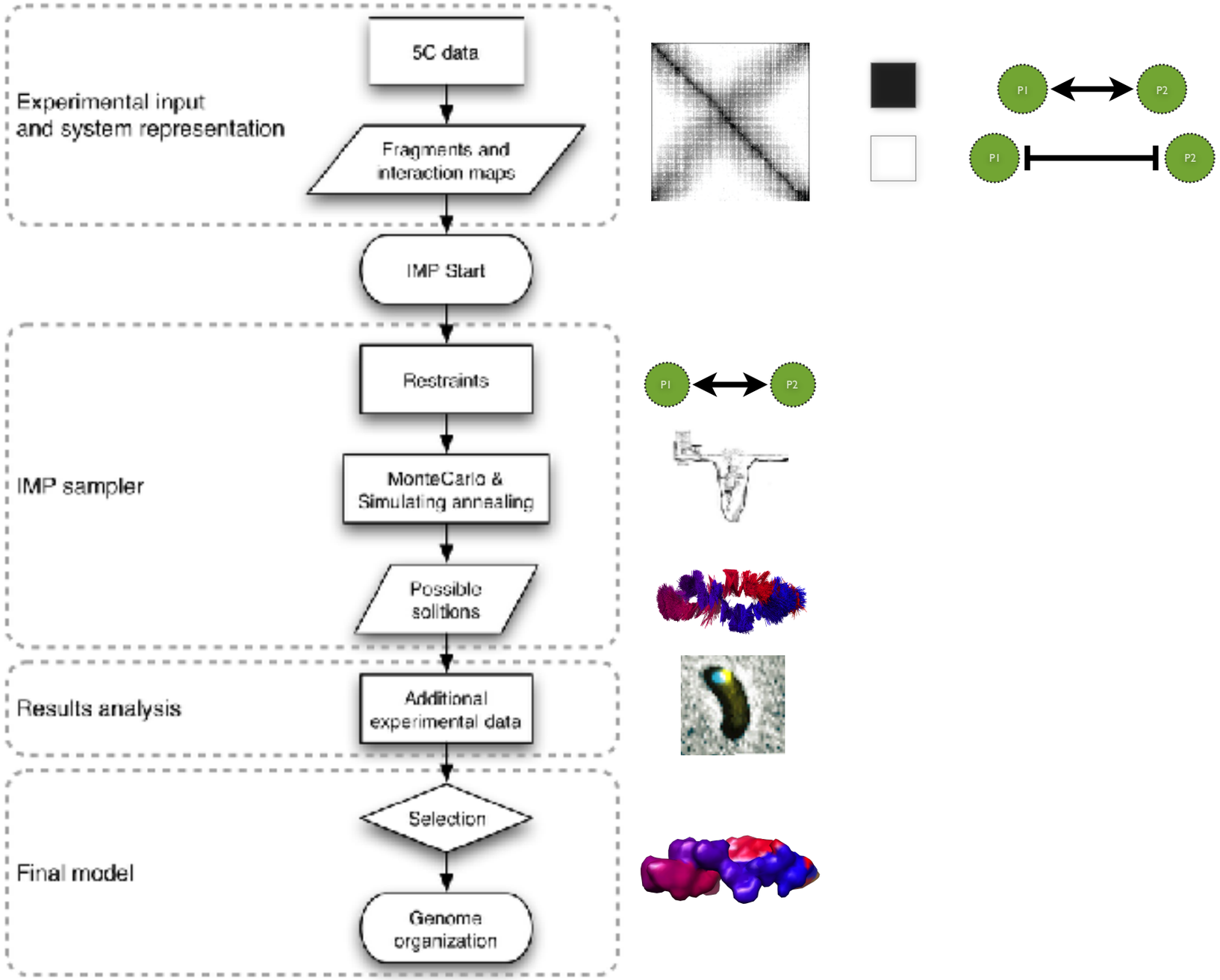


Beitel, C. W., Froenicke, L., Lang, J. M., Korf, I. F., Michelmore, R. W., Eisen, J. A., & Darling, A. E. (2014). Strain- and plasmid-level deconvolution of a synthetic metagenome by sequencing proximity ligation products. doi:10.7287/peerj.preprints.260v1

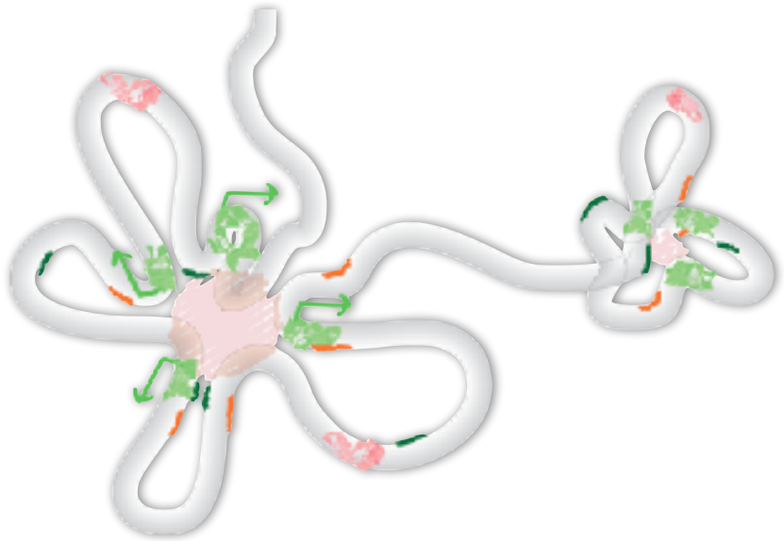
cnag CRG Centre for Genomic Regulation

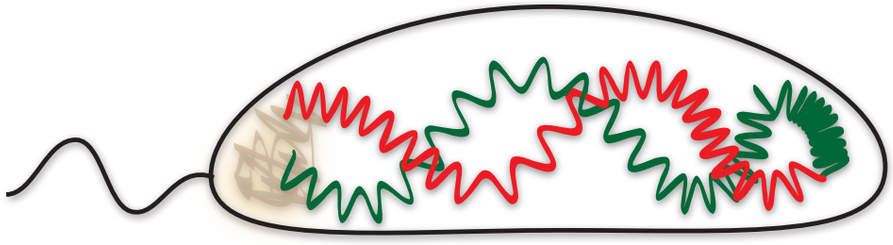
Modeling 3D Genomes

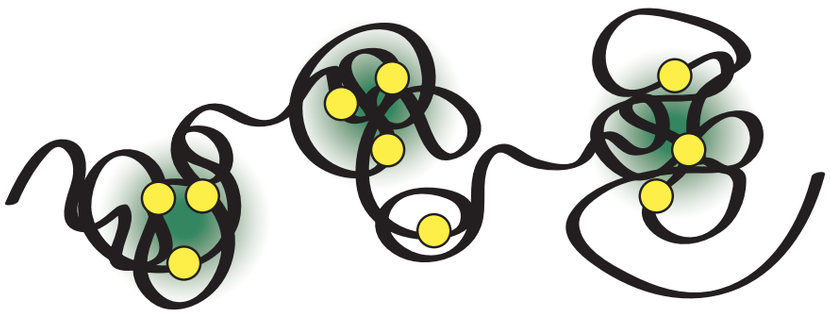
Baù, D. & Marti-Renom, M. A. Methods 58, 300–306 (2012).



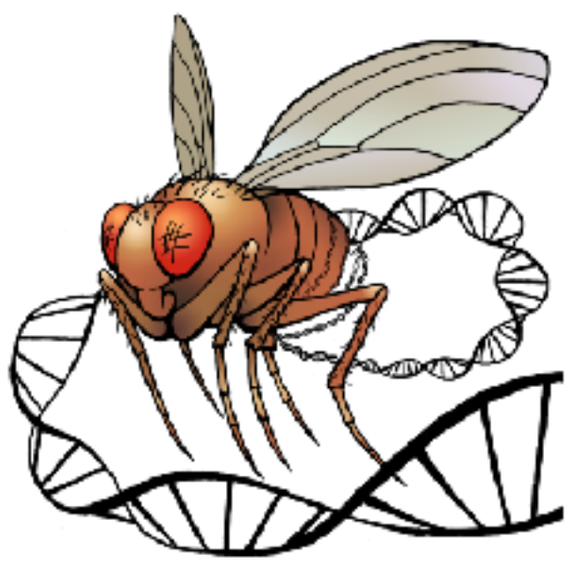
Examples...





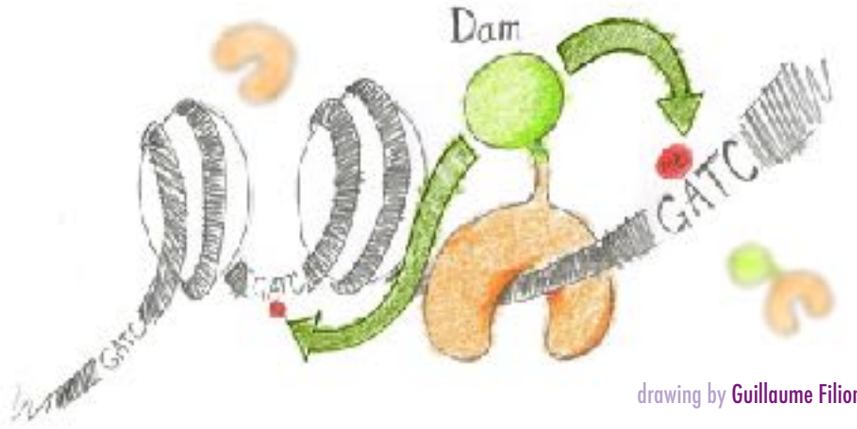


Structuring the COLORs of chromatin

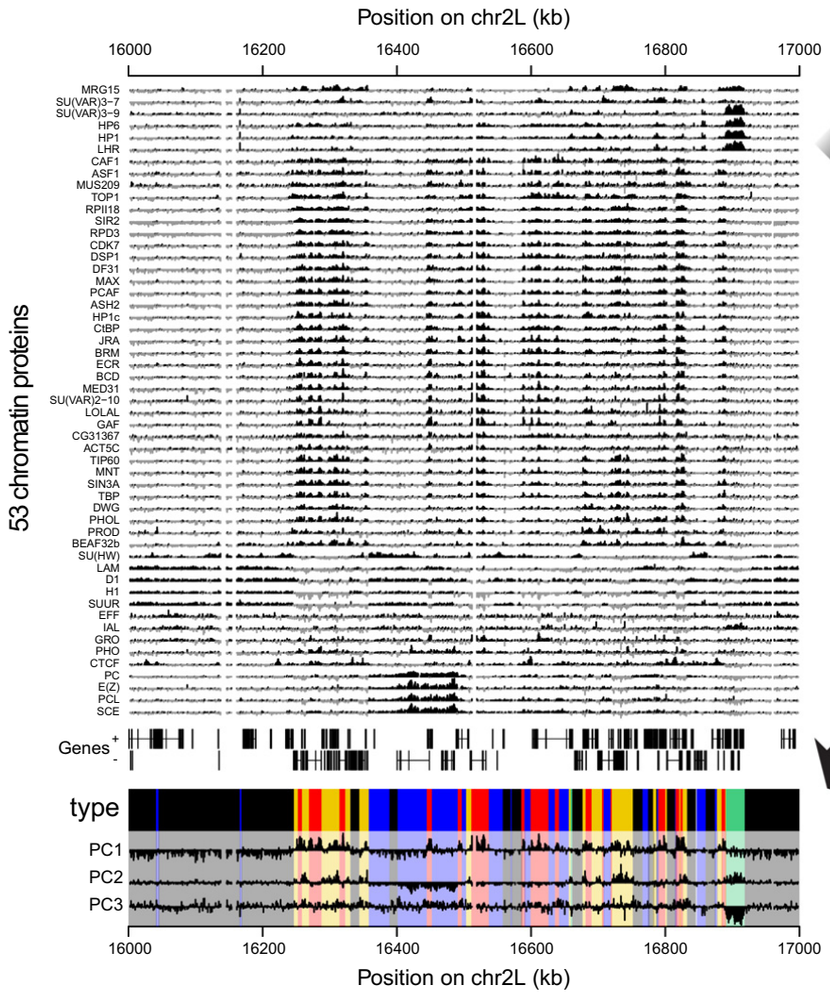


Fly Chromatin COLORs

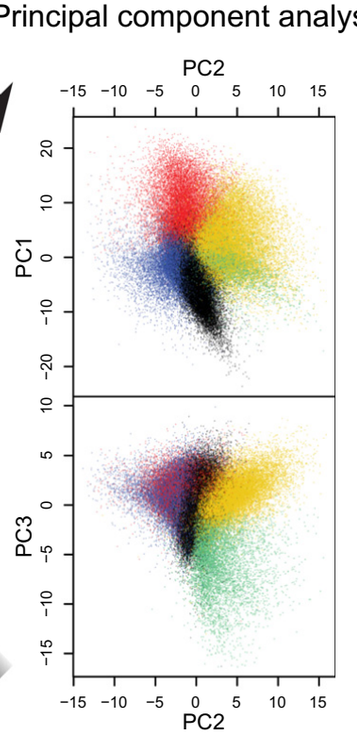
Filion et al. (2010). Cell, 143(2), 212–224.



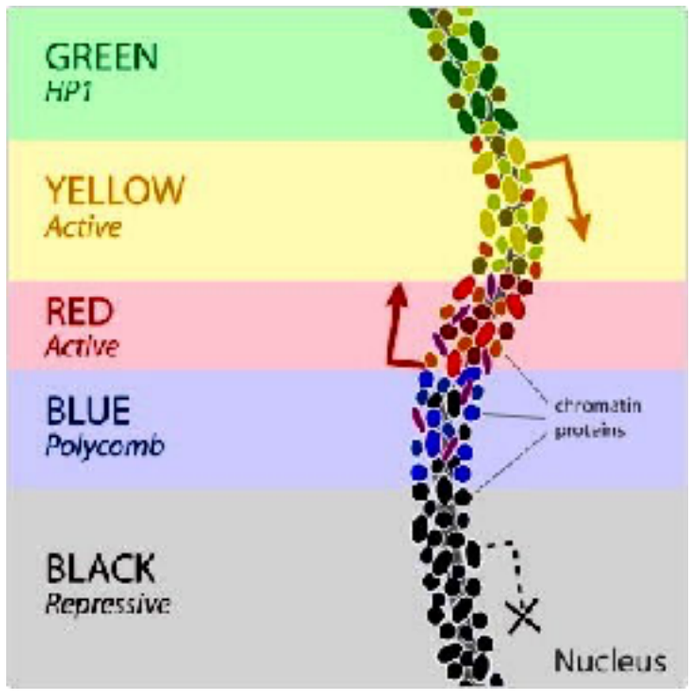
drawing by Guillaume Filion



Principal component analysis

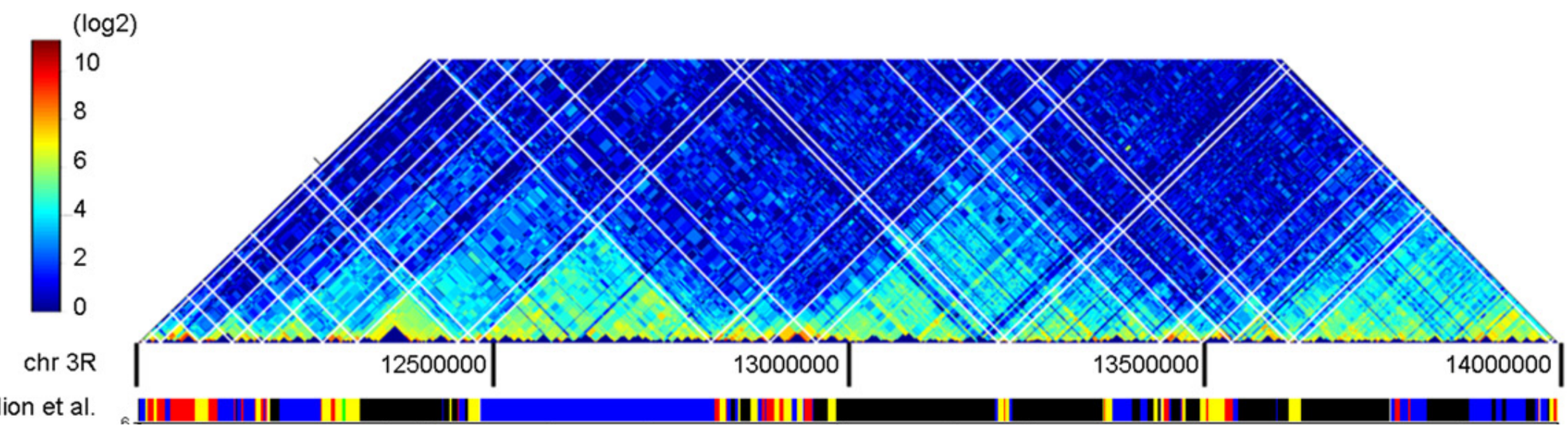


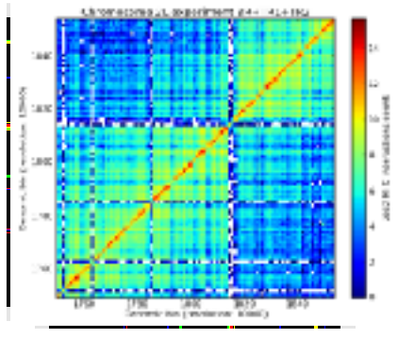
Hidden Markov model

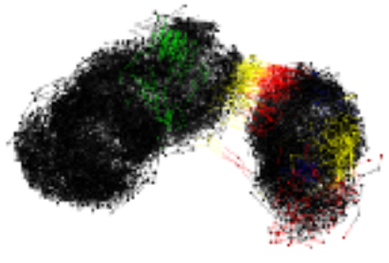


Fly Chromatin COLORs

Hou et al. (2012). Molecular Cell, 48(3), 471–484.



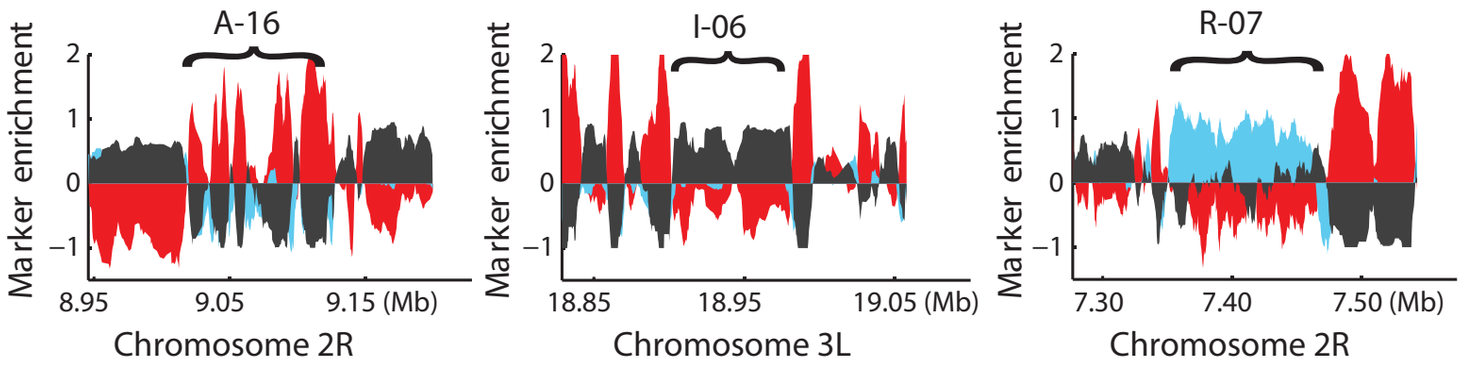


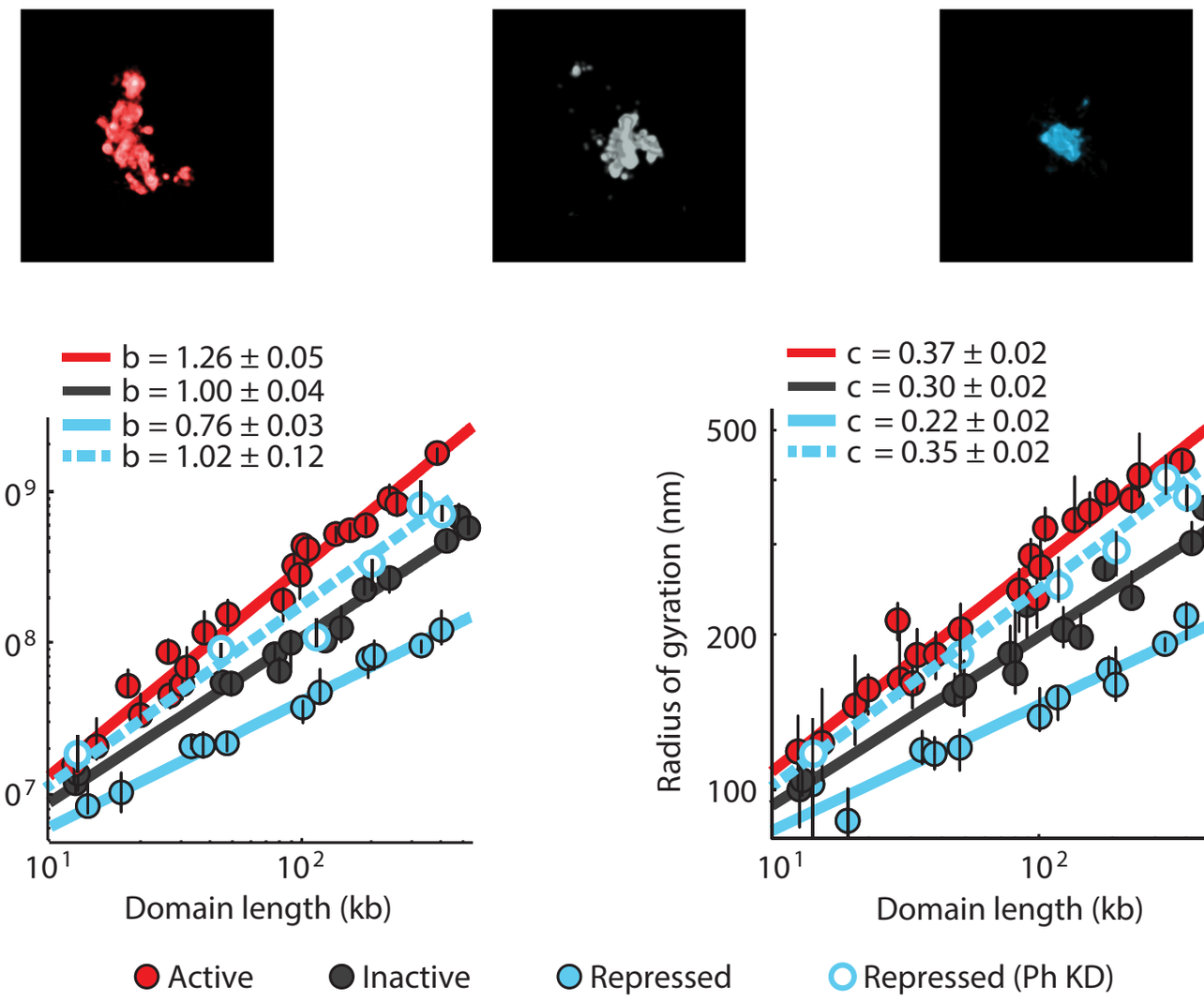


50 ~1Mb regions
10 for each color

Model accuracy

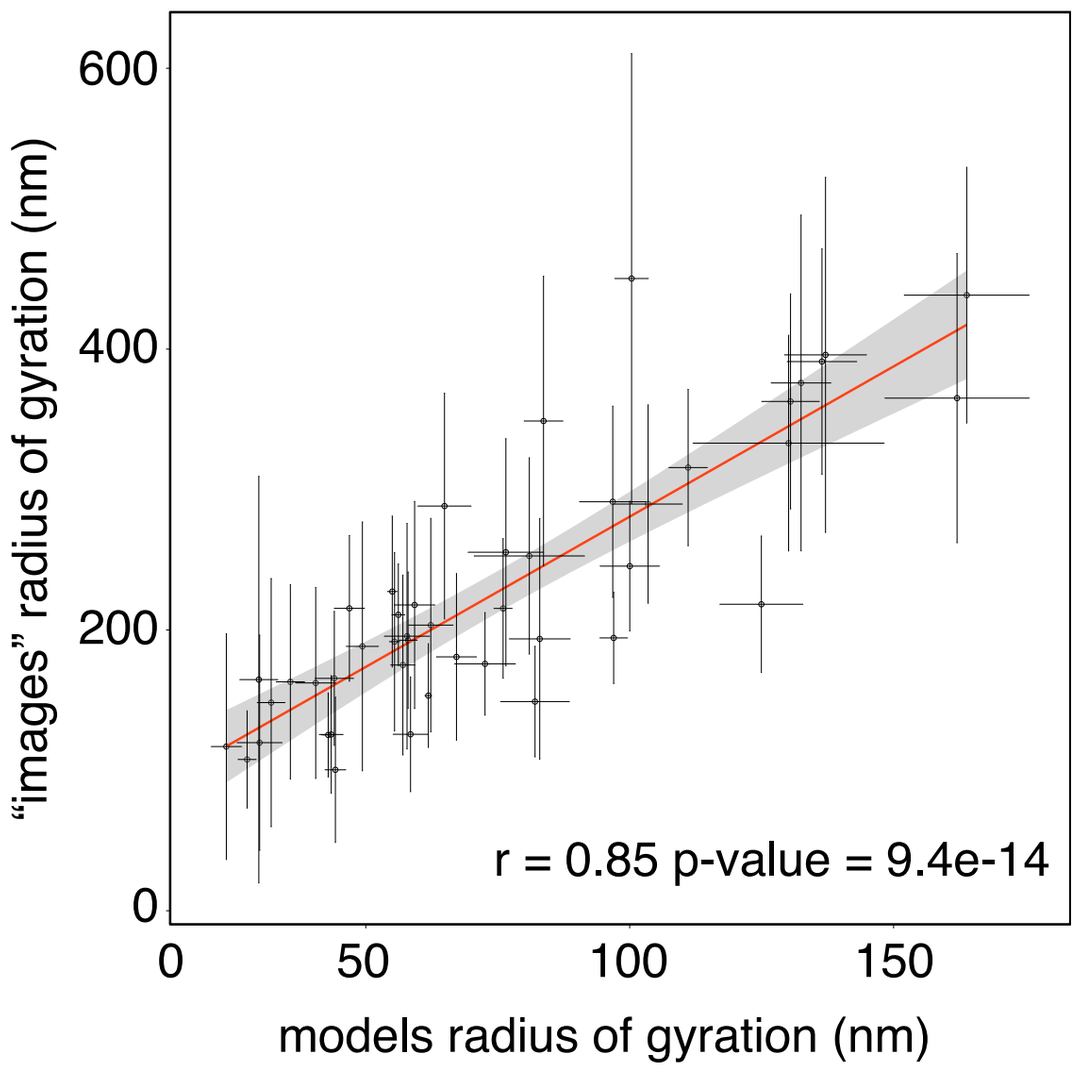
Boettiger, A. N., et al. (2016). Nature, 1–15.

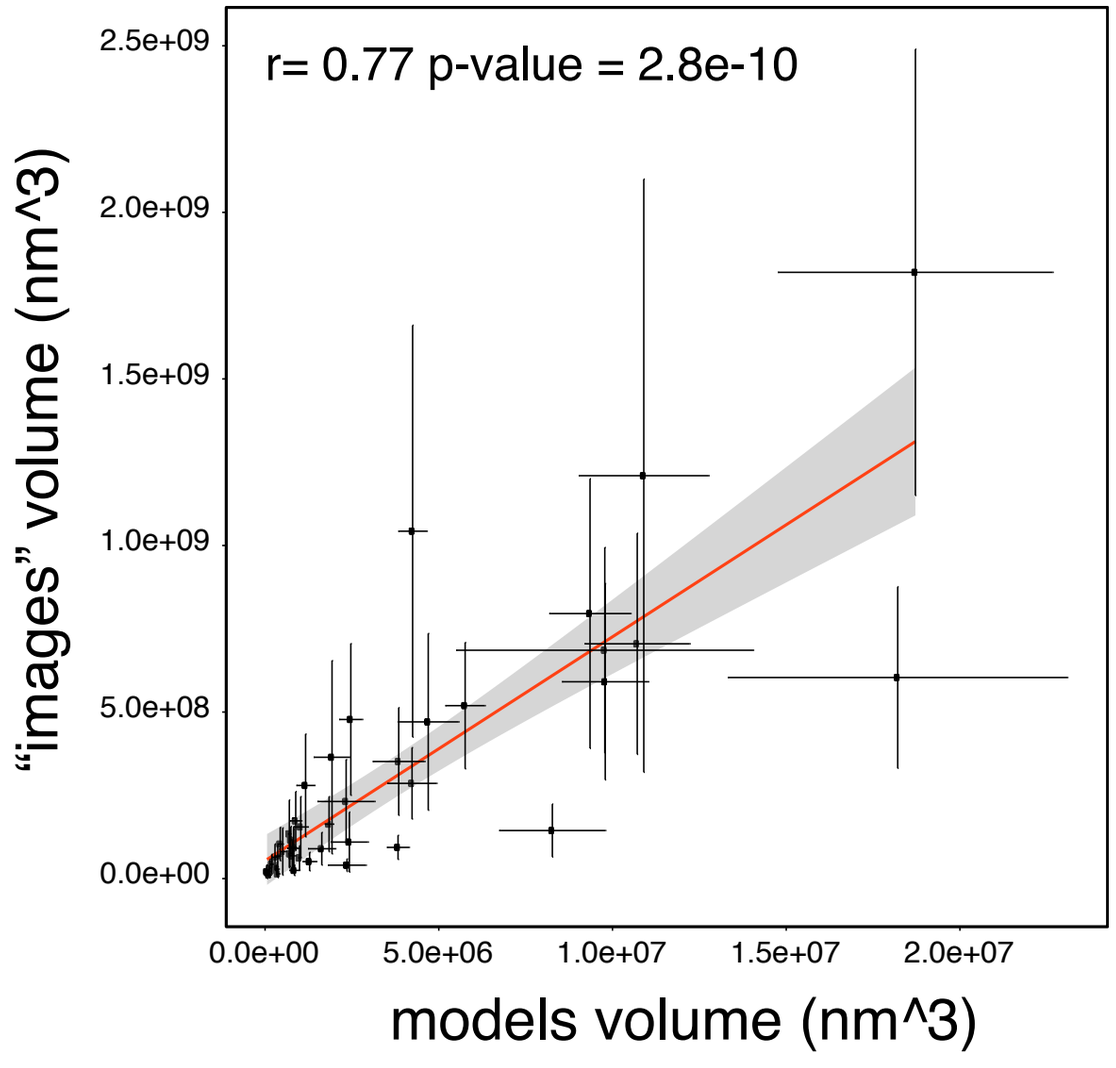




Model accuracy

Boettiger, A. N., et al. (2016). Nature, 1–15.



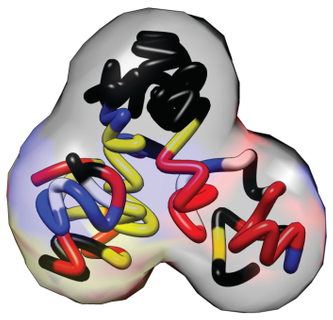


Structural properties

50 1Mb regions. 10 enriched for each color.

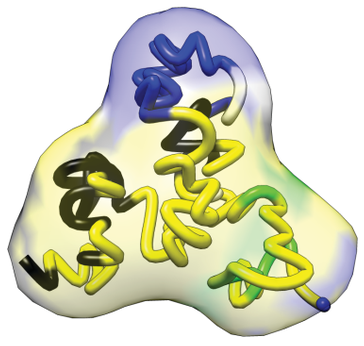
RED dense region
3R:18920000-19920000

22% 17% 0% 11% 45% 6%



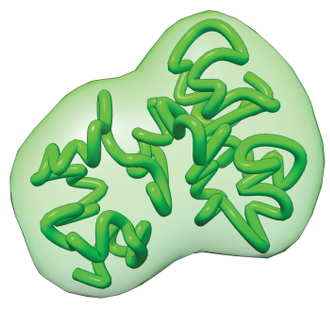
YELLOW dense region
X:15590000-16600000

0% 48% 4% 20% 26% 3%



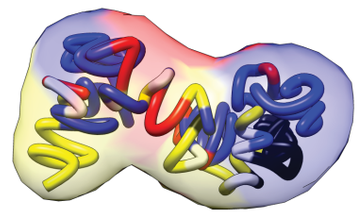
GREEN dense region
2R:510000-1530000

0% 0% 100% 0% 0% 0%



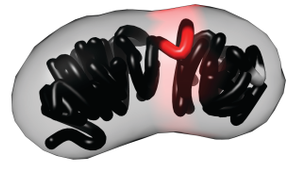
BLUE dense region
3L:210000-1230000

11% 17% 0% 52% 13% 0%

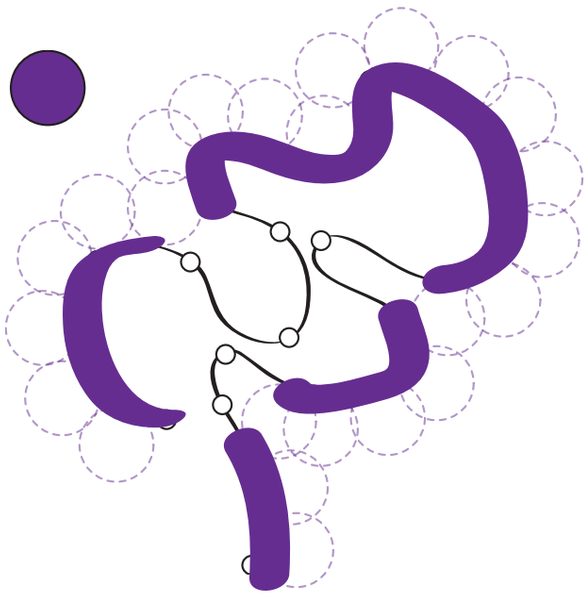


BLACK dense region
2L:17500000-18530000

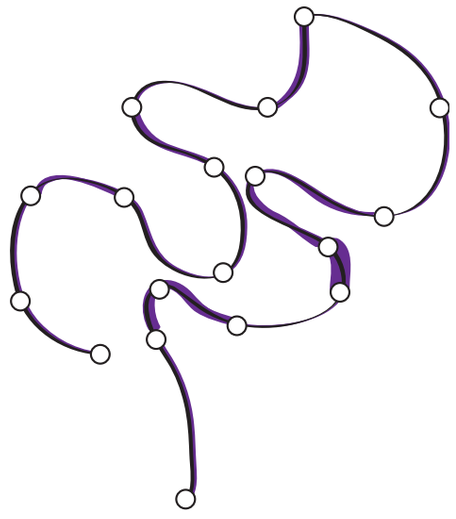
1% 0% 0% 0% 98% 1%



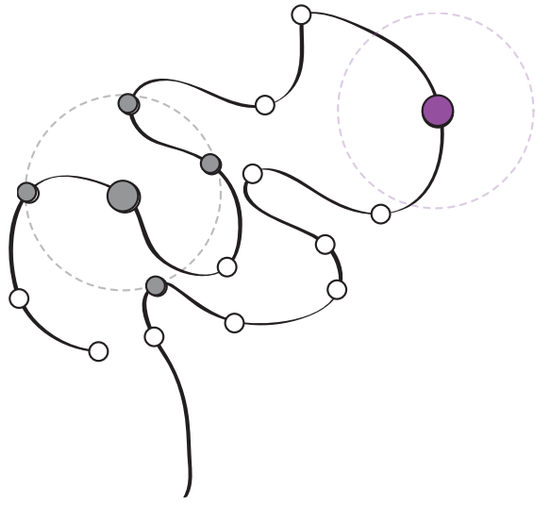
Accessibility (%)



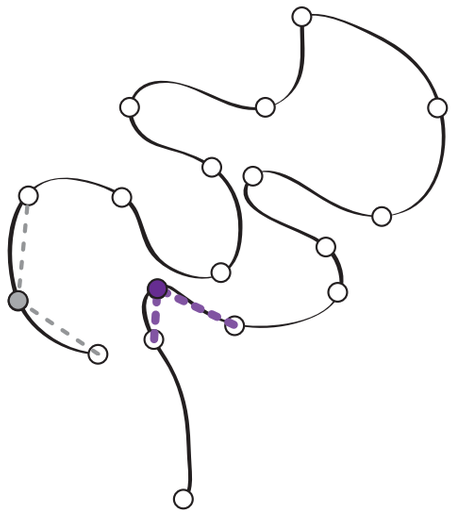
Density (bp/nm)



Interactions



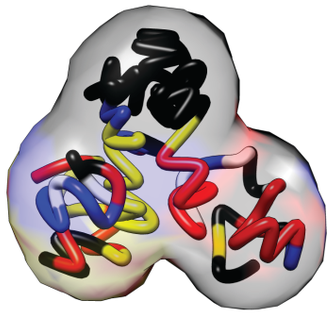
Angle



Structural COLORs

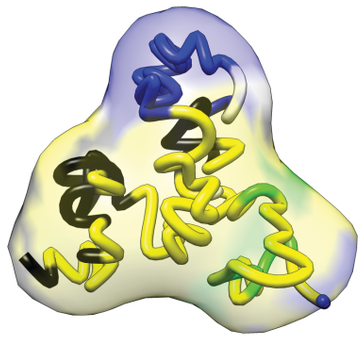
RED dense region
3R:18920000-19920000

22% 17% 0% 11% 45% 6%



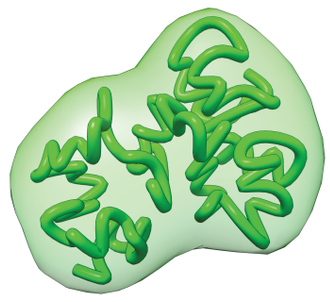
YELLOW dense region
X:15590000-16600000

0% 48% 4% 20% 26% 3%



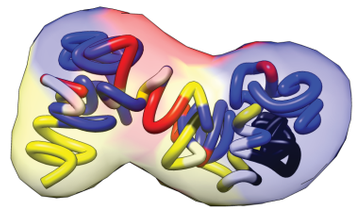
GREEN dense region
2R:510000-1530000

0% 0% 100% 0% 0% 0%



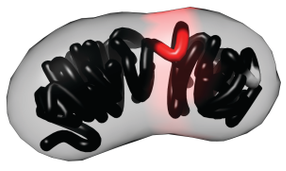
BLUE dense region
3L:210000-1230000

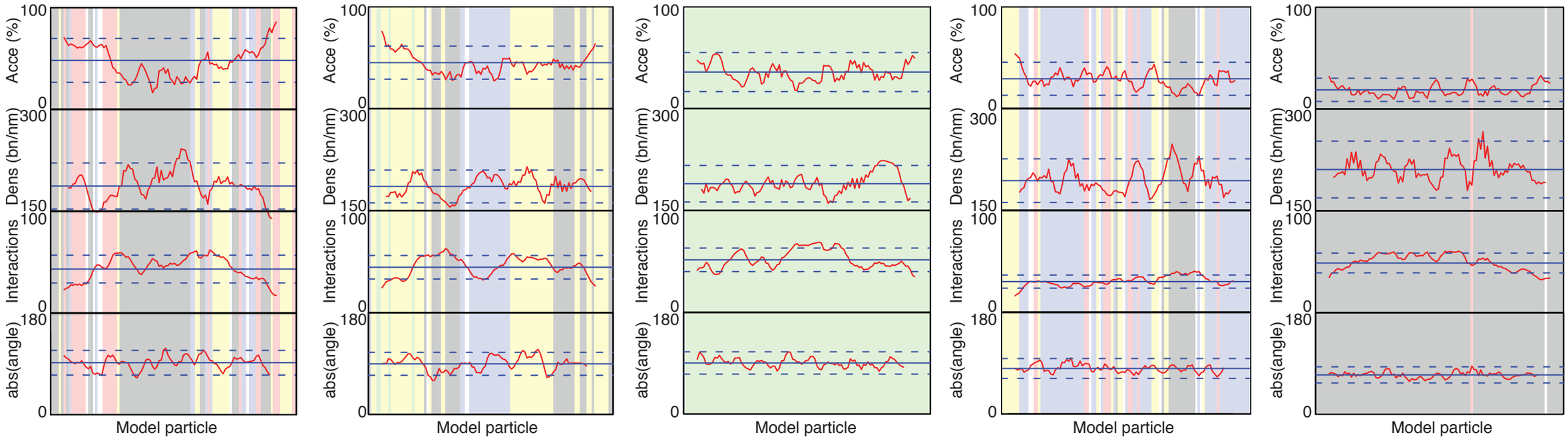
11% 17% 0% 52% 13% 0%



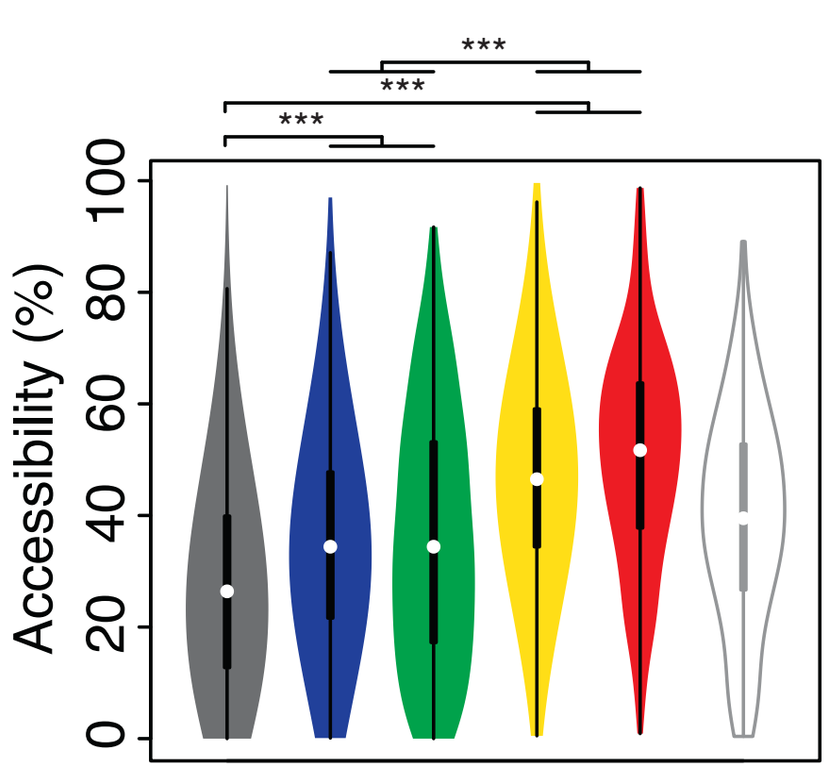
BLACK dense region
2L:17500000-18530000

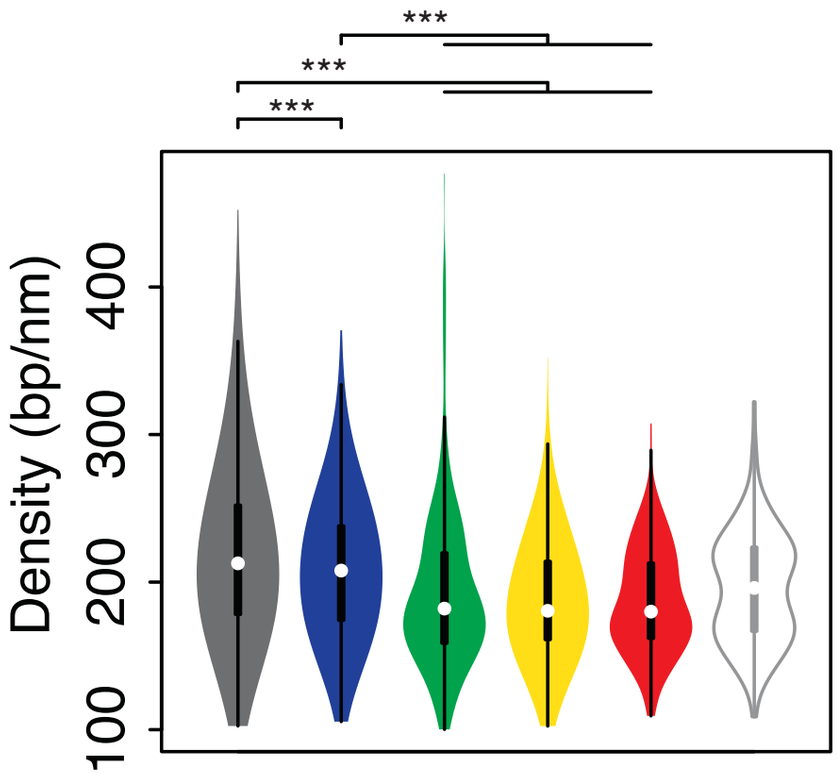
1% 0% 0% 0% 98% 1%

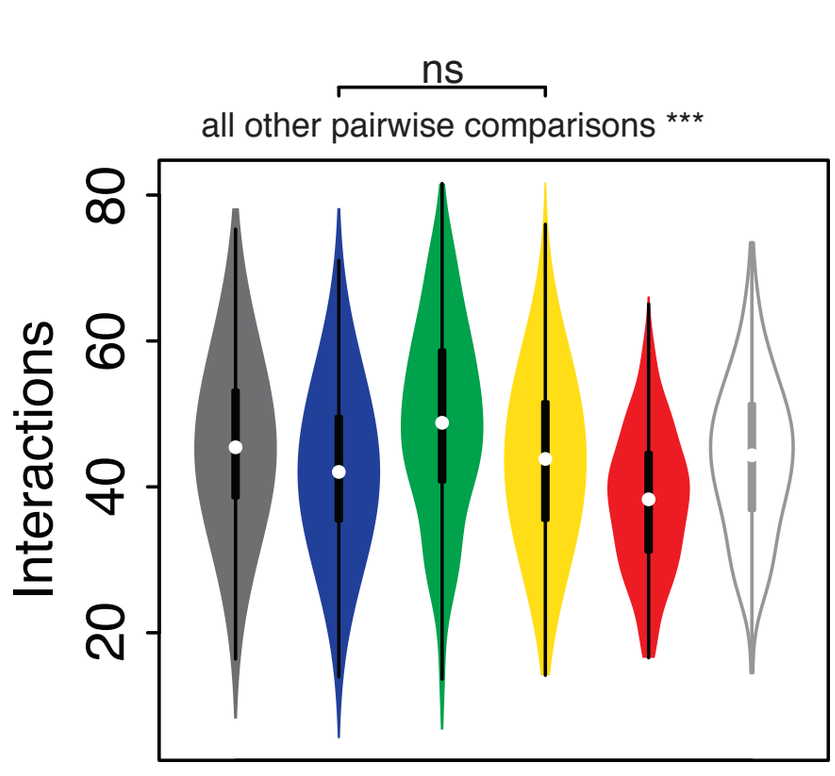


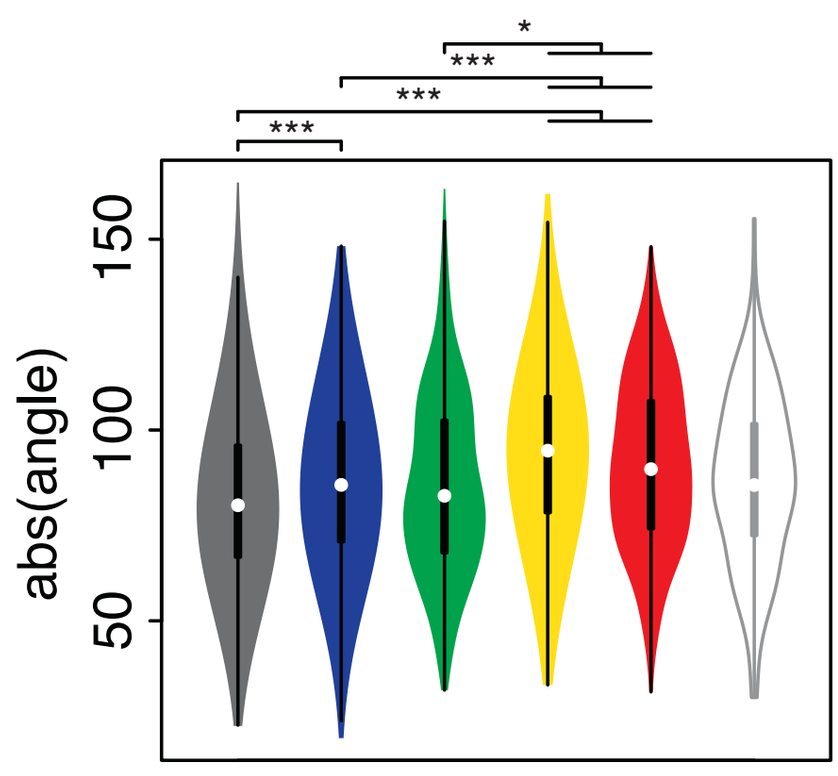


Structural COLORs







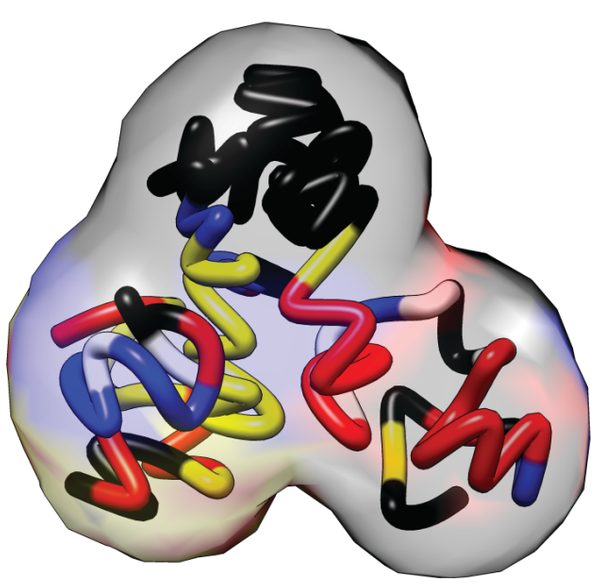


Color prediction by Self Organizing Maps

RED dense region

3R:18920000-19920000

22% 17% 0% 11% 45% 6%



	GainRatio	PCA
interactions all 40 nm	0.25	-0.40
interactions cluster1 40 nm	0.24	-0.40
accessibility radius 20 nm superradius 75 nm	0.18	0.35
stderr of intr2	0.19	-0.34
number of models in cluster 1	0.35	-0.00
stderr of intr1	0.16	-0.34
number of clusters	0.26	-0.02
stderr of dens3	0.13	-0.33
stderr of dens1	0.11	-0.27
unsigned angle with -3 and +3 smoothed over 5 bins	0.10	0.13
density 3 particles (center of mass) cluster 1	0.07	-0.11
stderr of dens2	0.09	-0.31
density 3 particles cluster 1	0.05	0.00
density 3 particles all clusters	0.06	0.00
consistency all 50 nm	0.03	-0.02
unsigned angle with -3 and +3	0.03	0.09
consistency cluster1 50 nm	0.03	0.05
stderr of acces	0.02	0.04

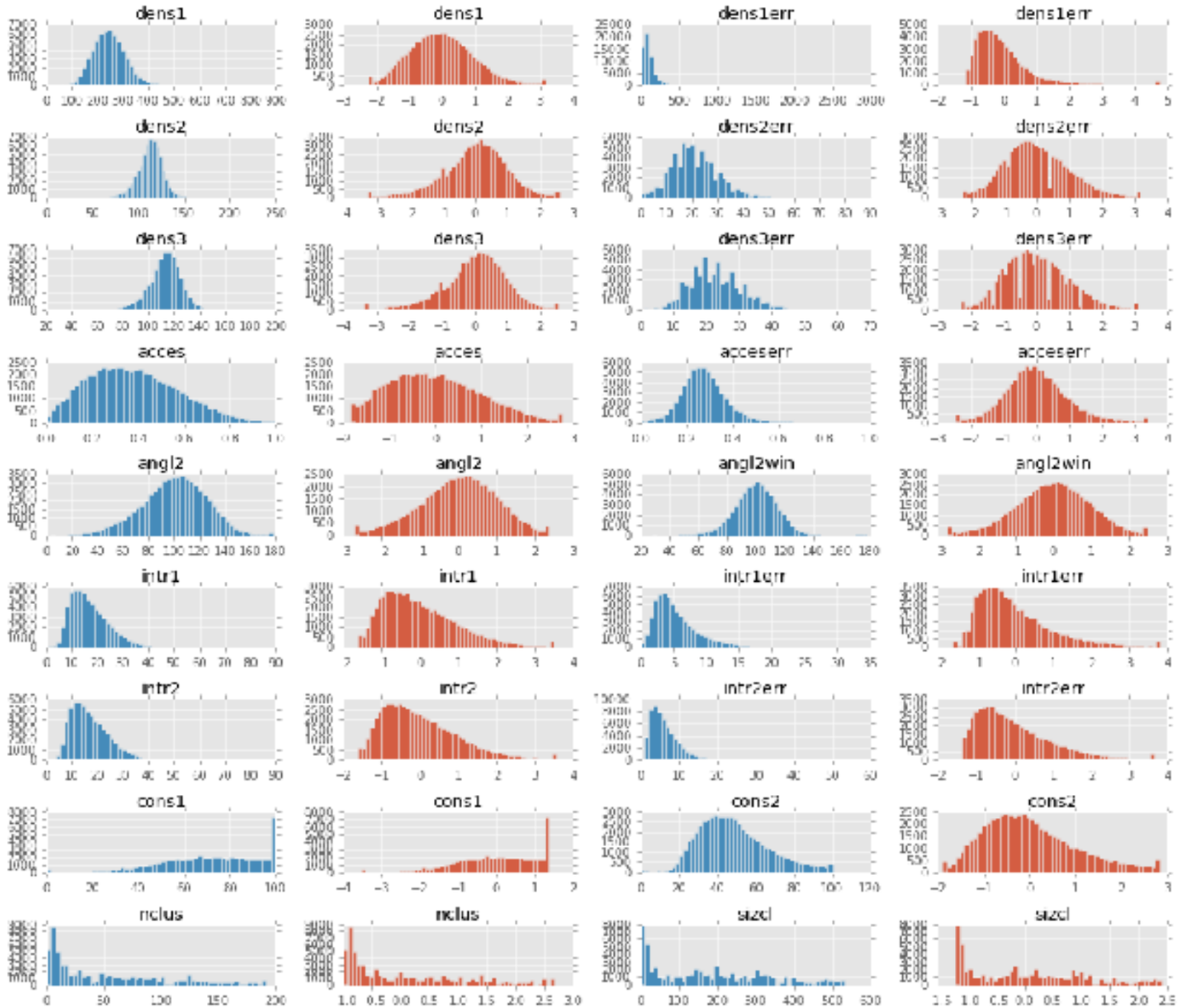
Selected metrics per particle

- extract the **18 metrics for each particle** (if particle is present in several models, an average is calculated)

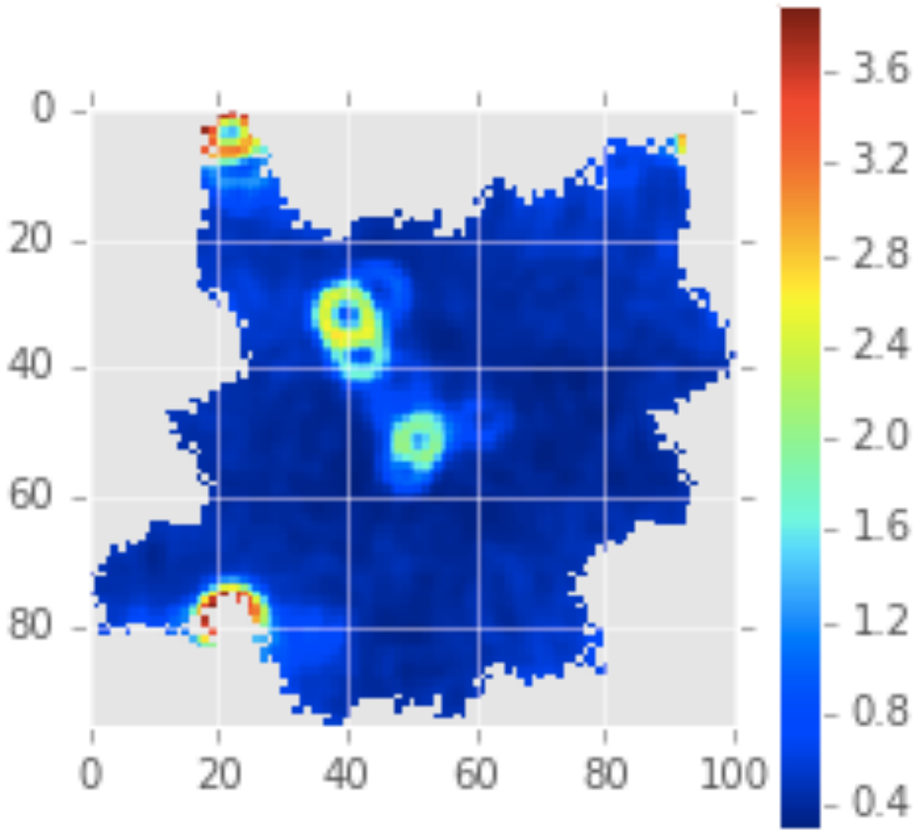
- the **metrics are normalized** (mean=0, std=1), and **outliers removed** (percentile 0.5 and 99.5)

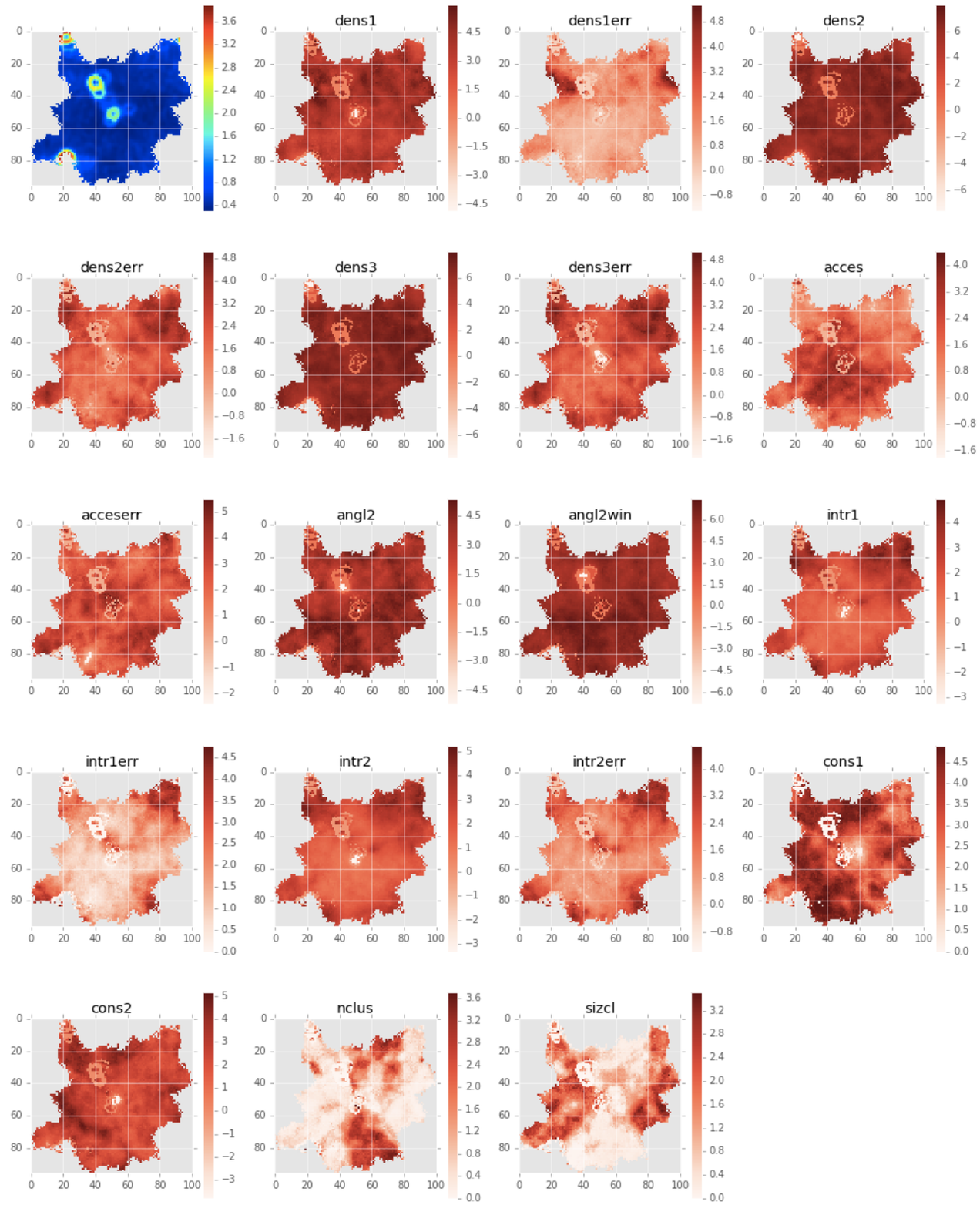
- each of the particles are going to be arranged in the SOM according to their relative **euclidian distance**

- SOM is run with 100.000 iterations

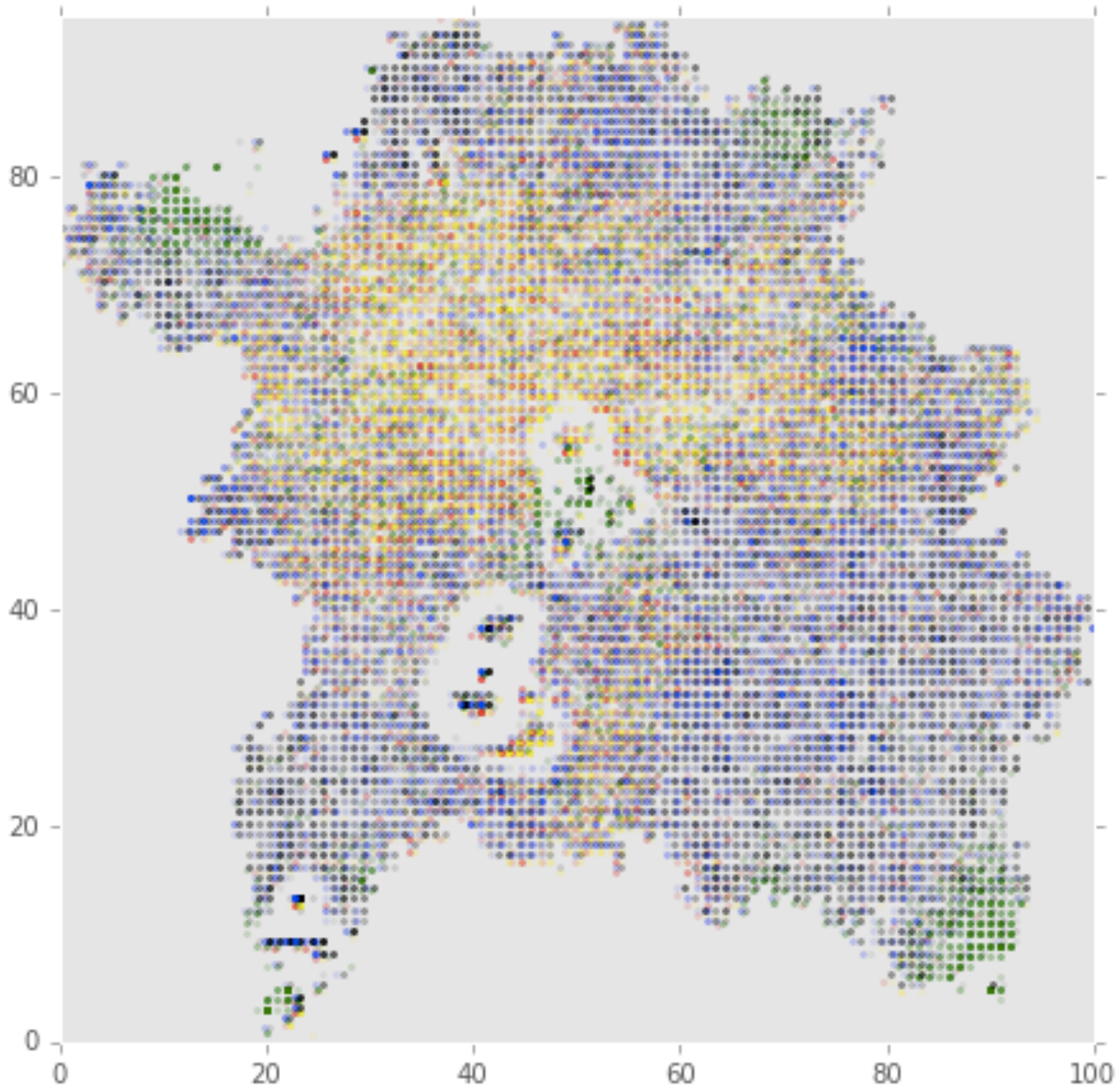


Self Organizing Maps (SOM)

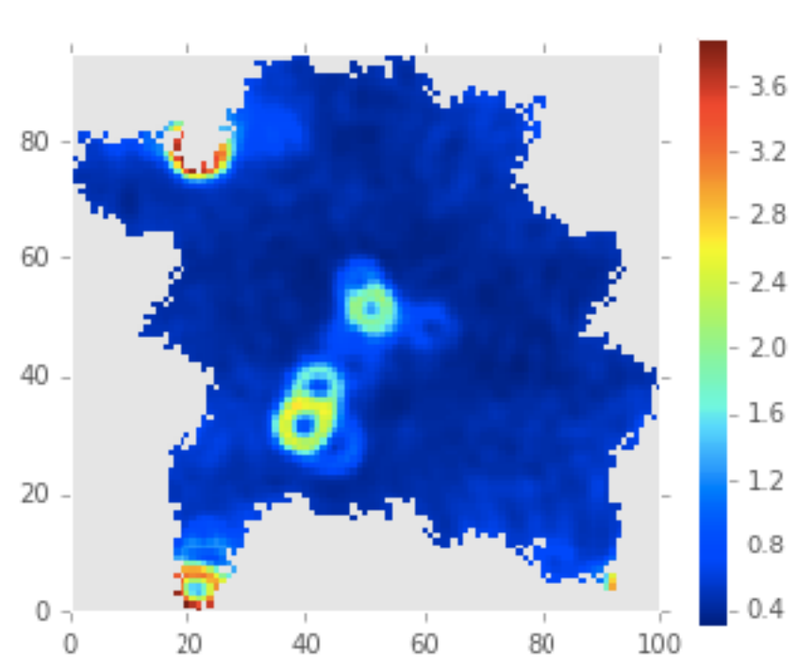


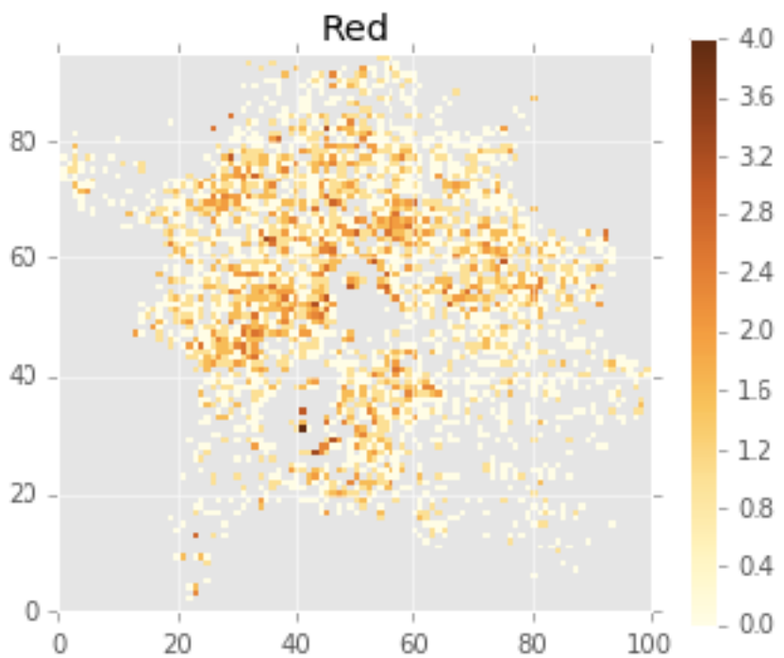


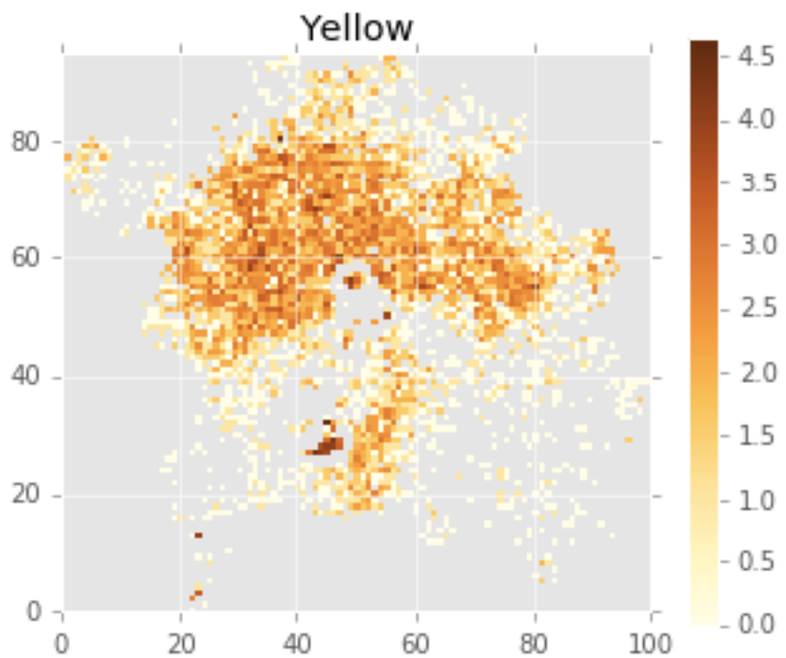
SOM Models

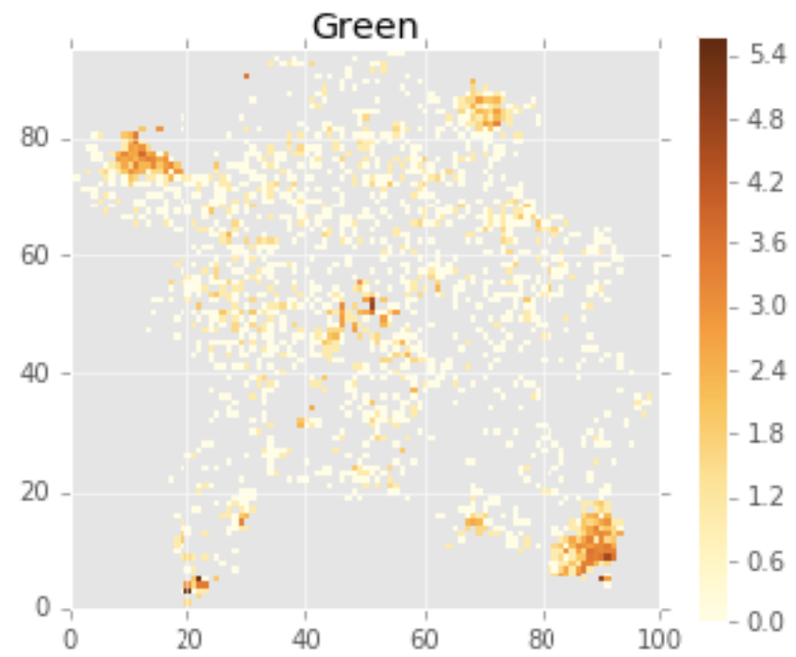


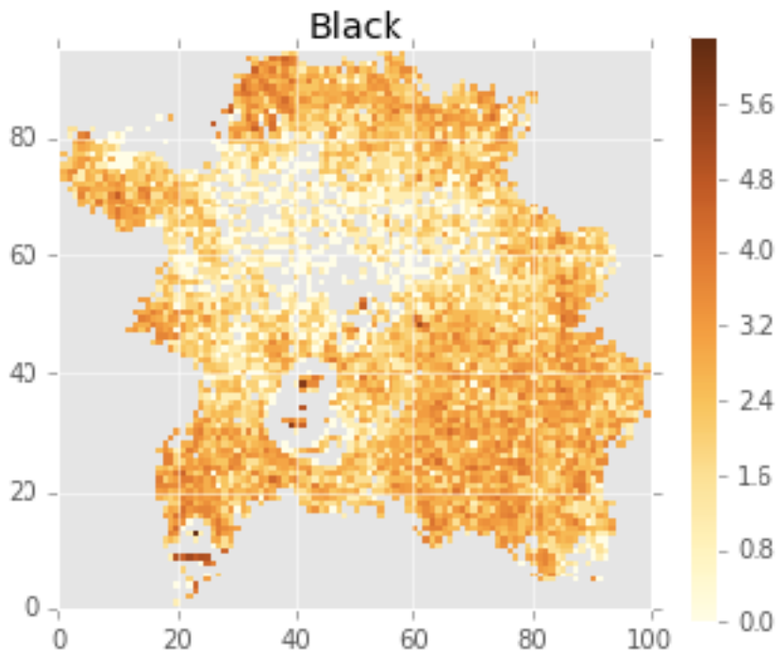
SOM Models

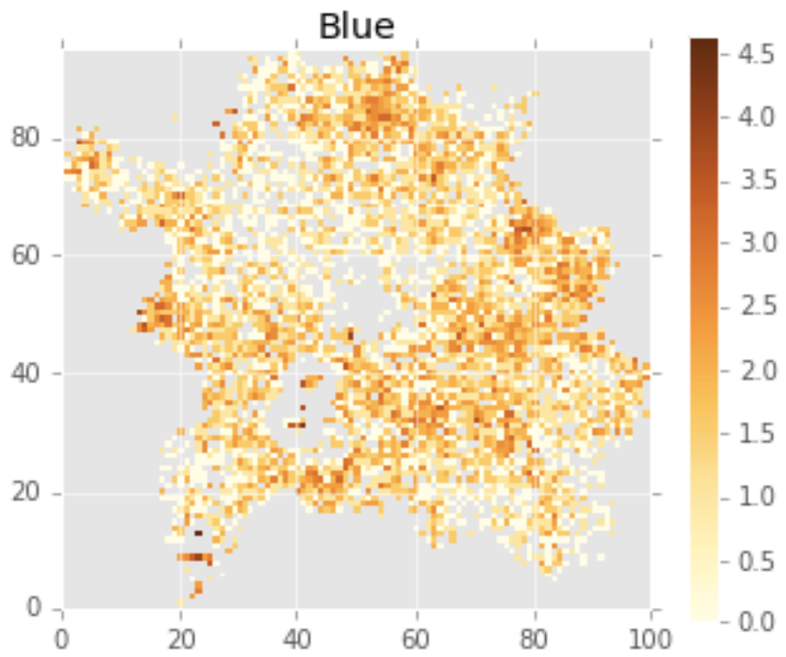




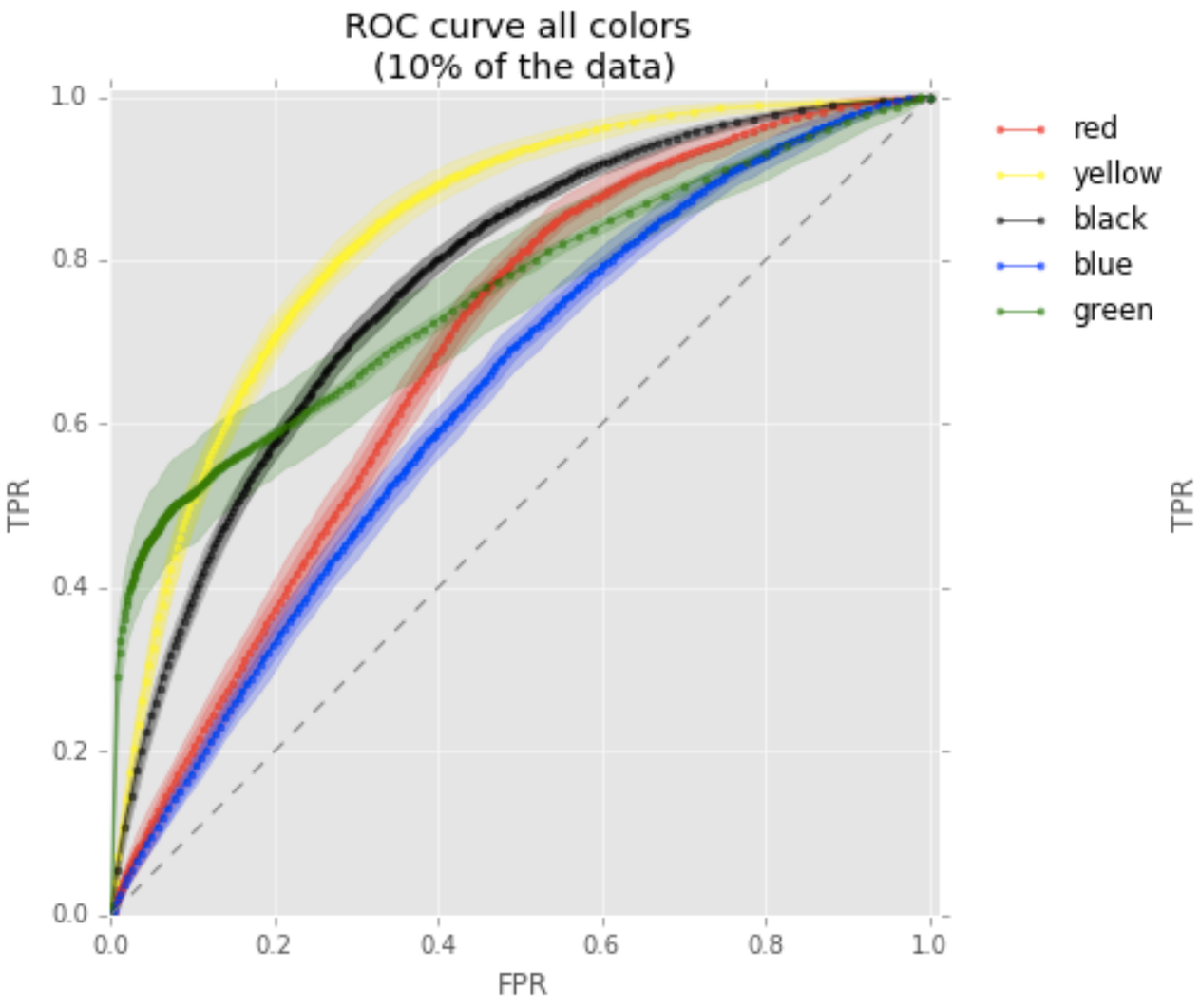


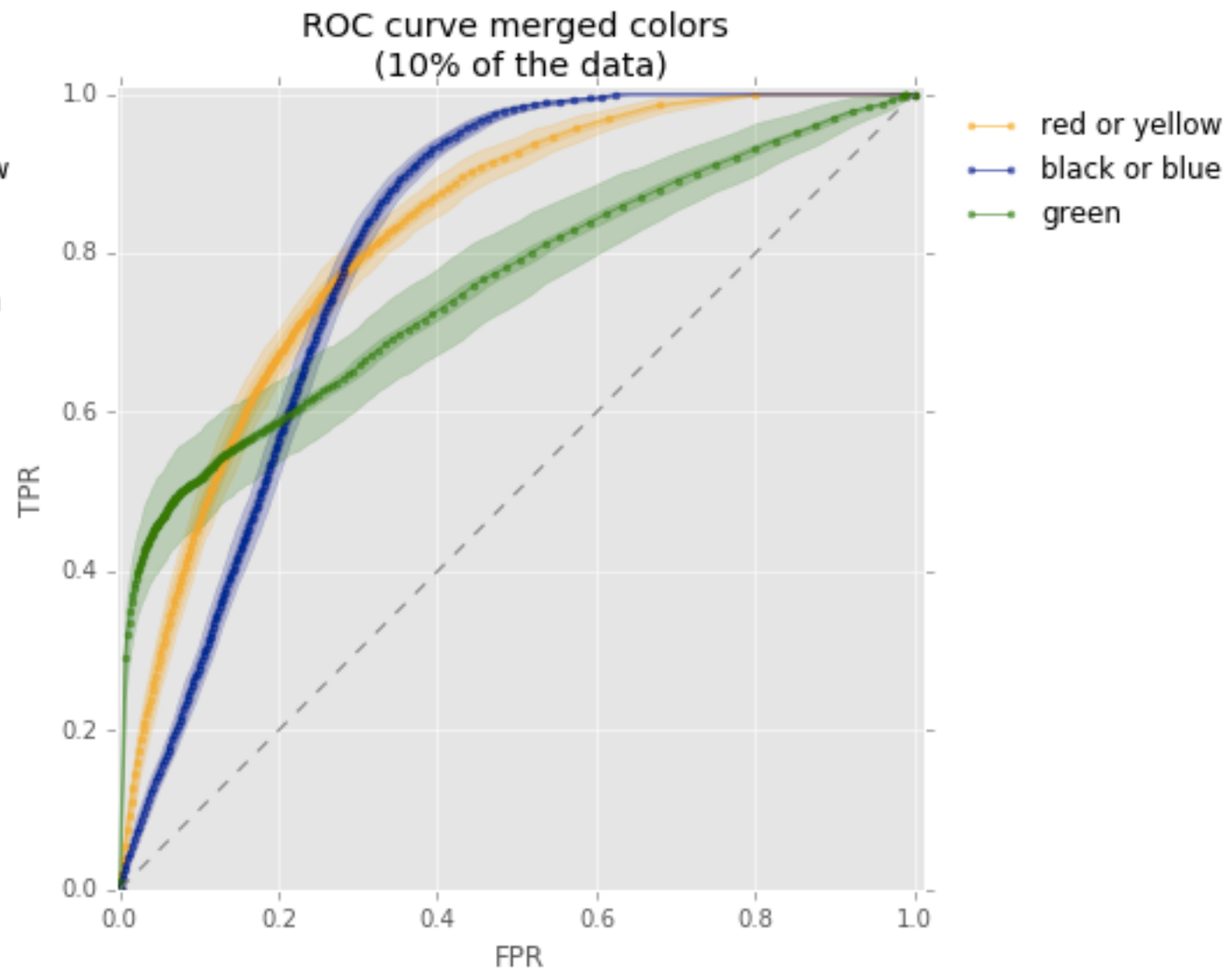






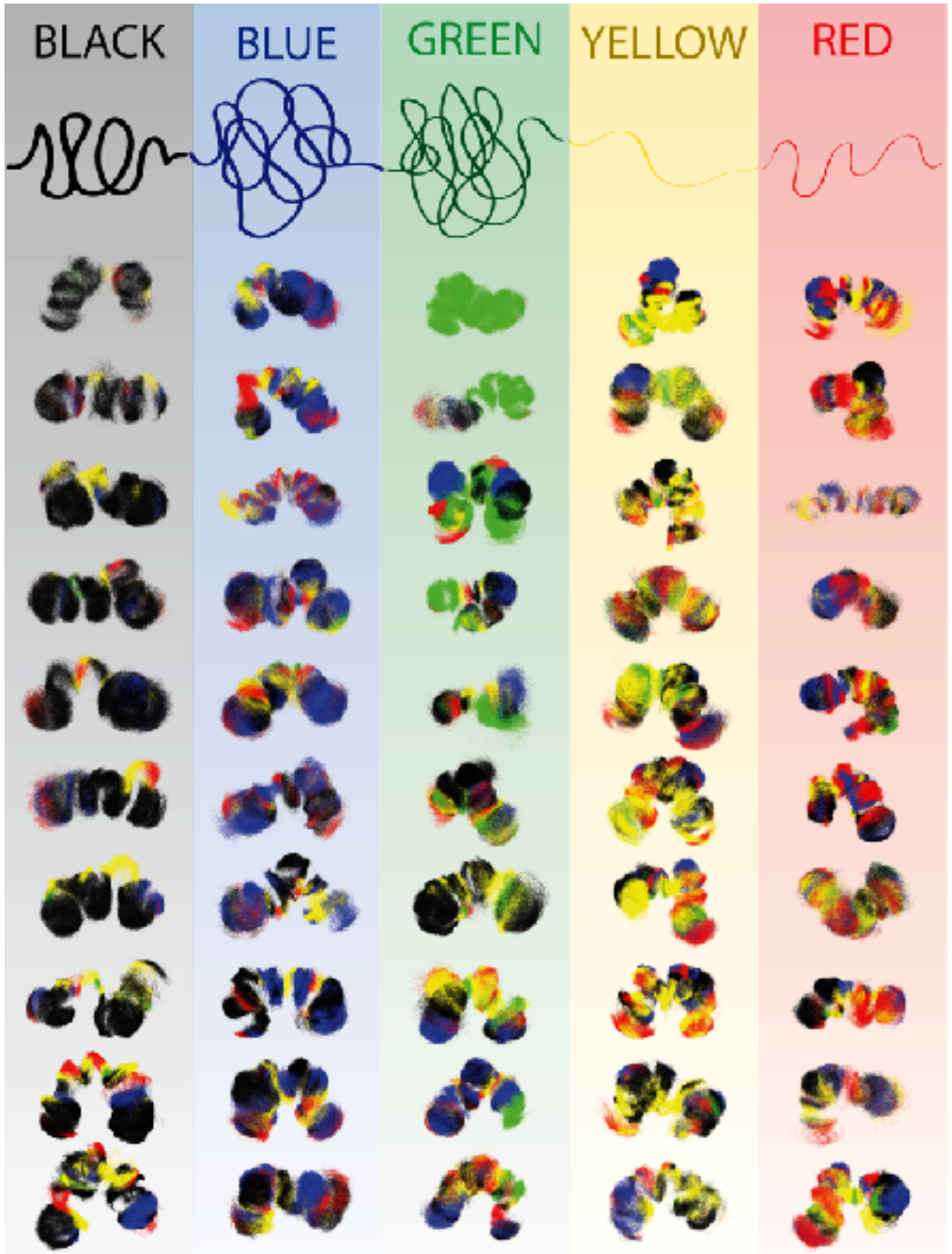
Can we predict the color?

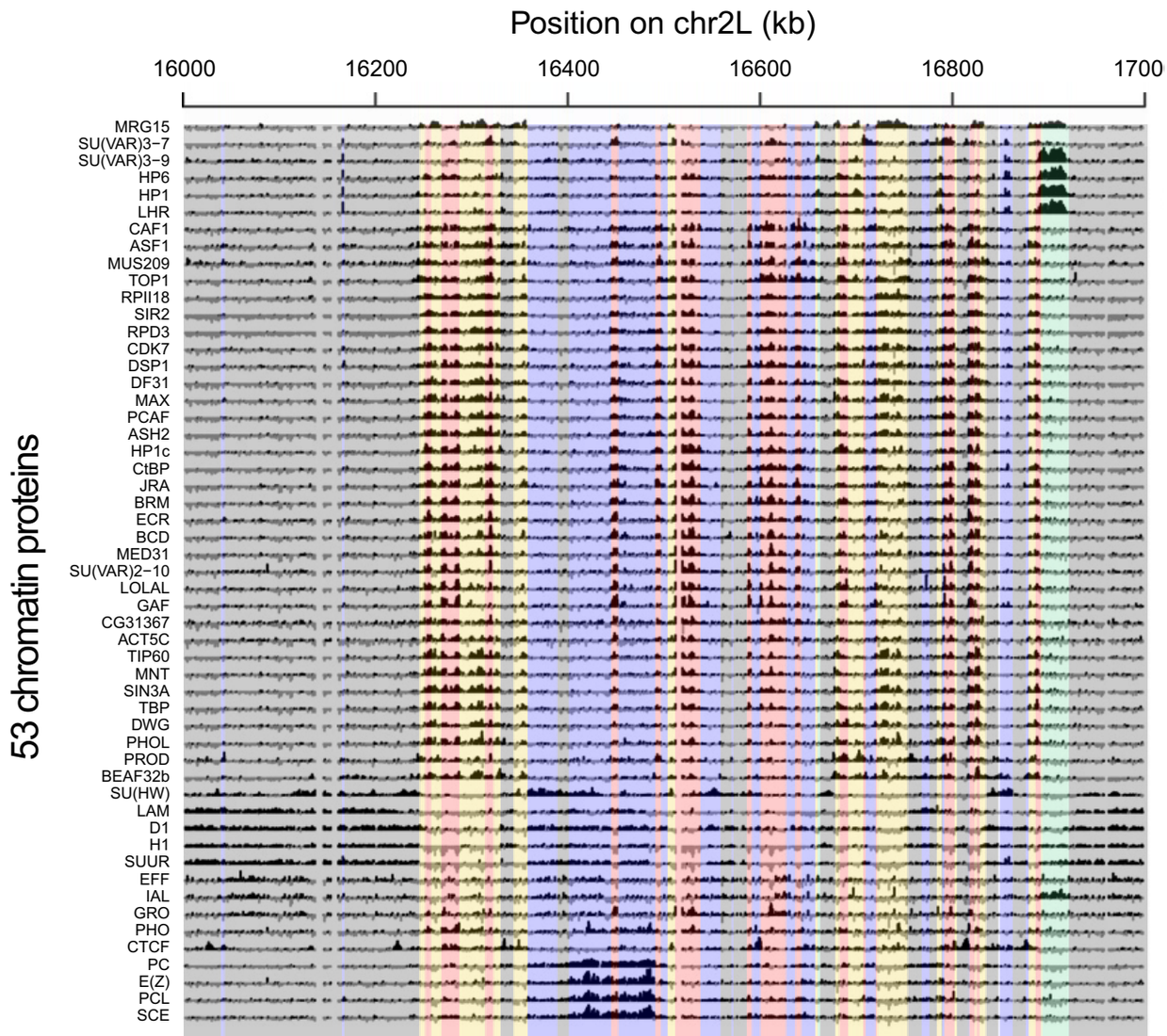


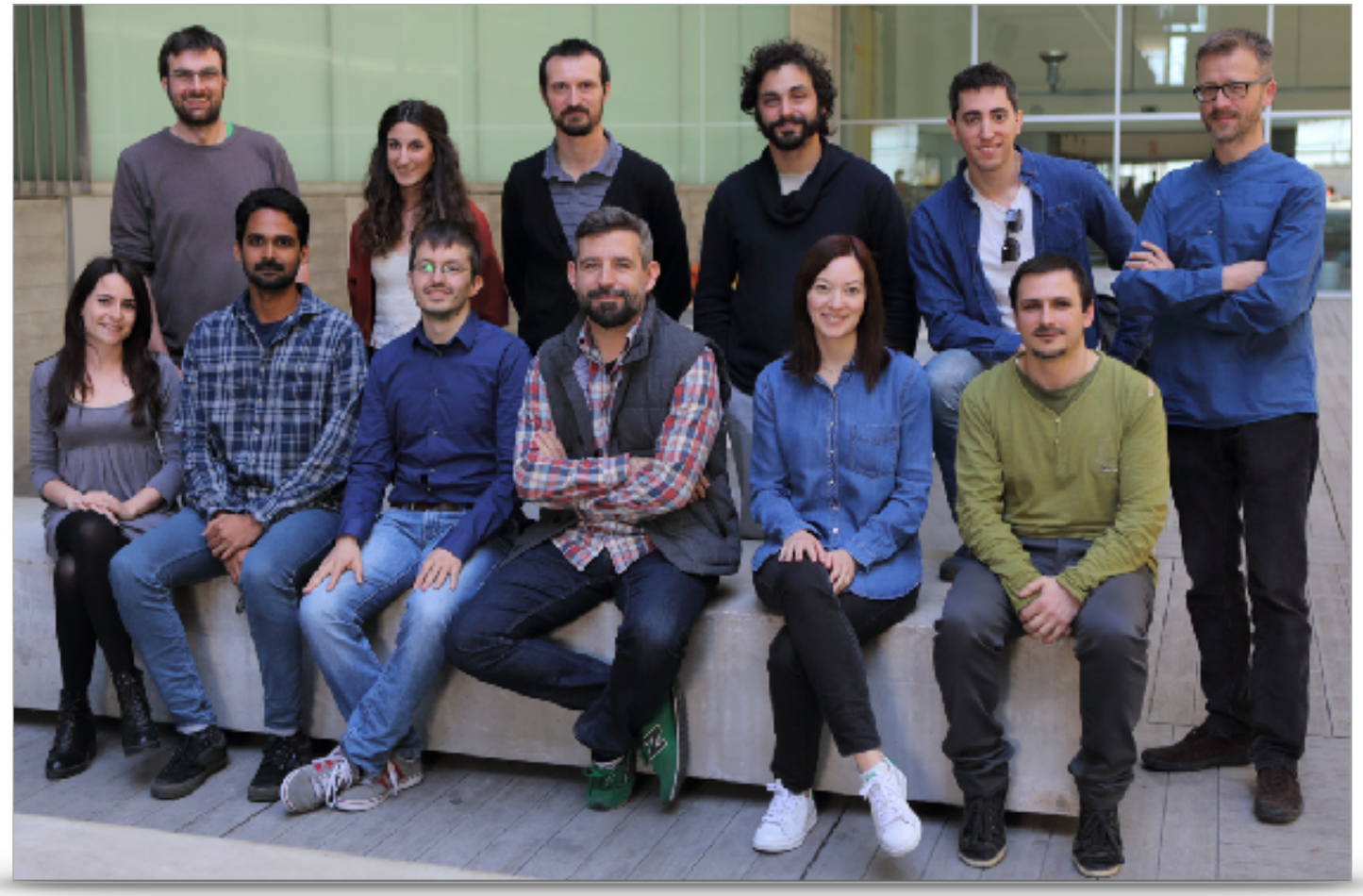



```
black :0.77
 blue :0.64
green :0.80
  red :0.69
yellow :0.83
```

Structural COLORs







Davide Baù
Gireesh K. Bogu
Yasmina Cuartero
François le Dily
David Dufour
Irene Farabella
Silvia Galan
Francesca di Giovanni
Mike Goodstadt
Francisco Martínez-Jiménez
François Serra
Paula Soler
Yannick Spill
Marco di Stefano
Marie Trussart

http://marciuslab.org
http://3DGenomes.org
http://cnag.crg.eu