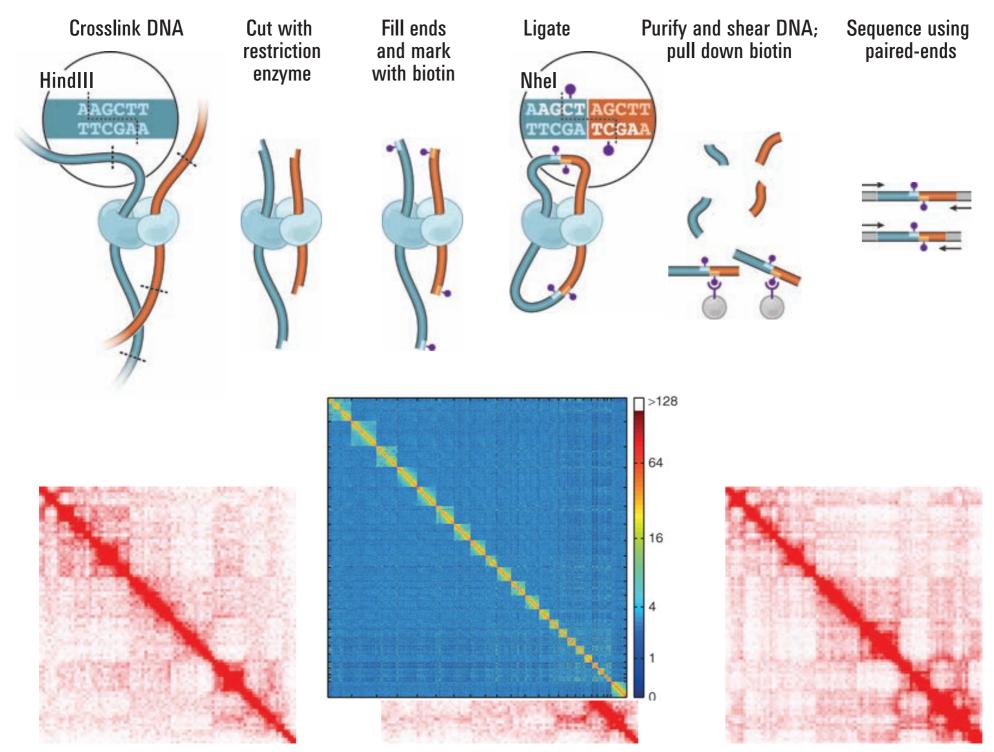
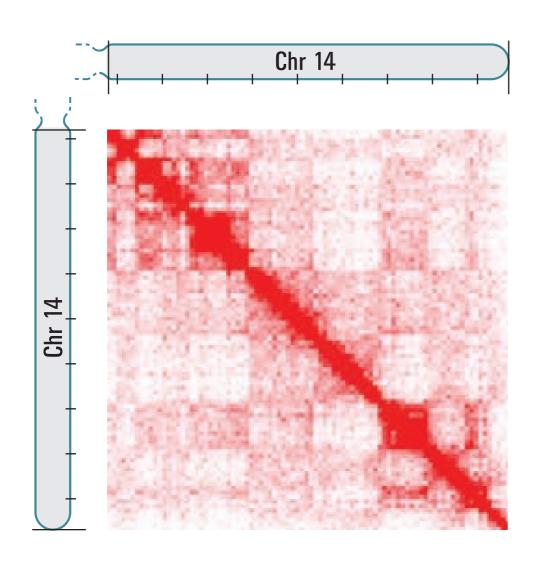
Summary of yesterday lecture

Level IV: Higher-order organization

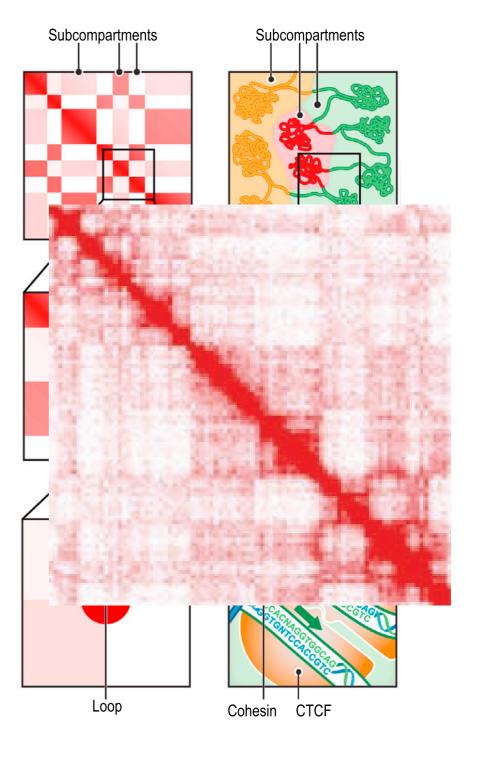


Dekker, J., Rippe, K., Dekker, M., & Kleckner, N. (2002). Science, 295(5558), 1306–1311. Lieberman-Aiden, E., et al. (2009). Science, 326(5950), 289–293.

Hierarchical genome organisation



Le

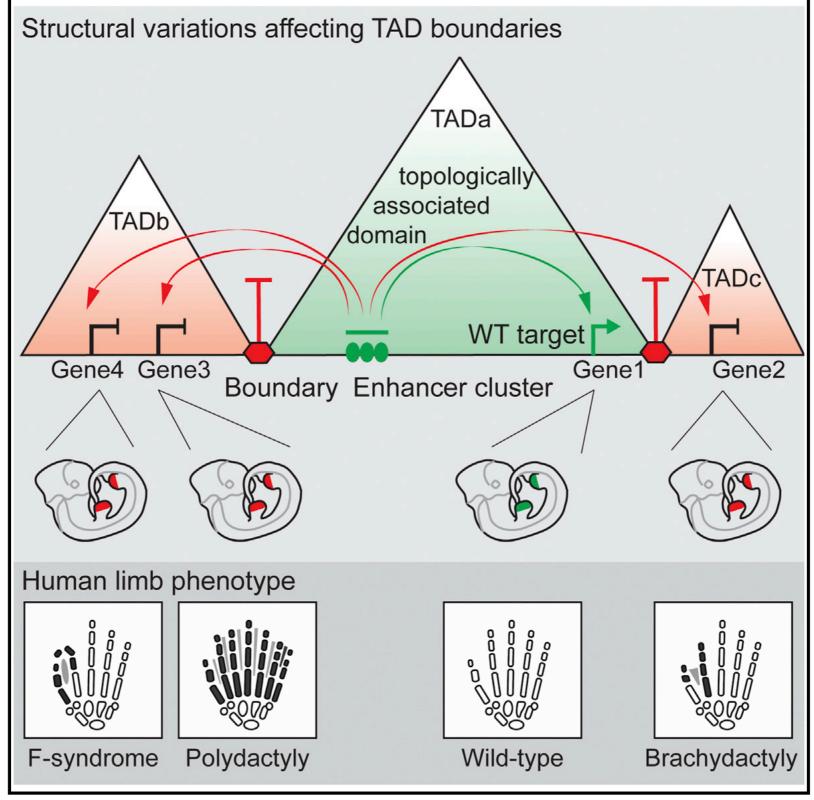


Canization

Lieberman-Aiden, E., et al. (2009). Science, 326(5950), 289–293. Rao, S. S. P., et al. (2014). Cell, 1–29.

Level IV: Higher-order organization

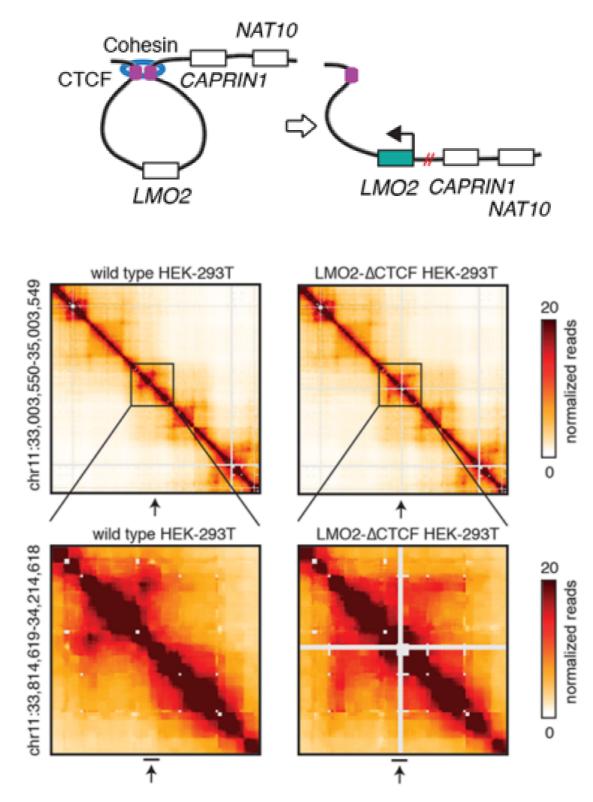
TADs are functional units



Lupiáñez, et al. (2015). Cell, 1–15.

Level IV: Higher-order organization

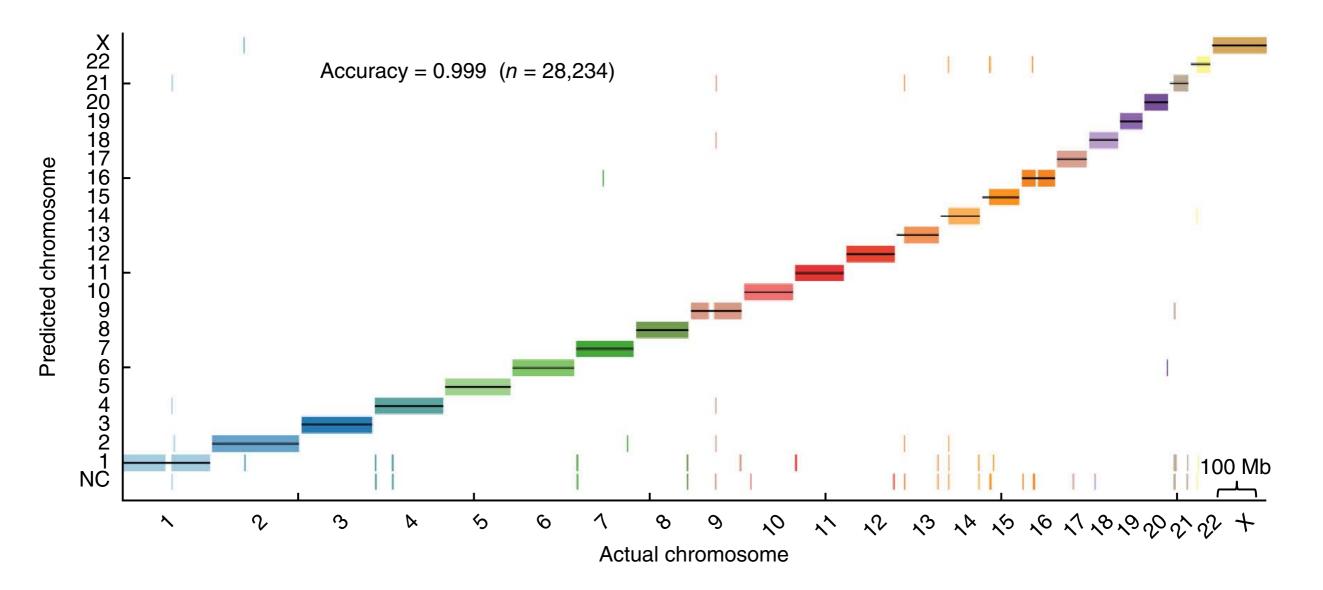
TADs are functional units



Hnisz, D., et al. (2016). Science, on line

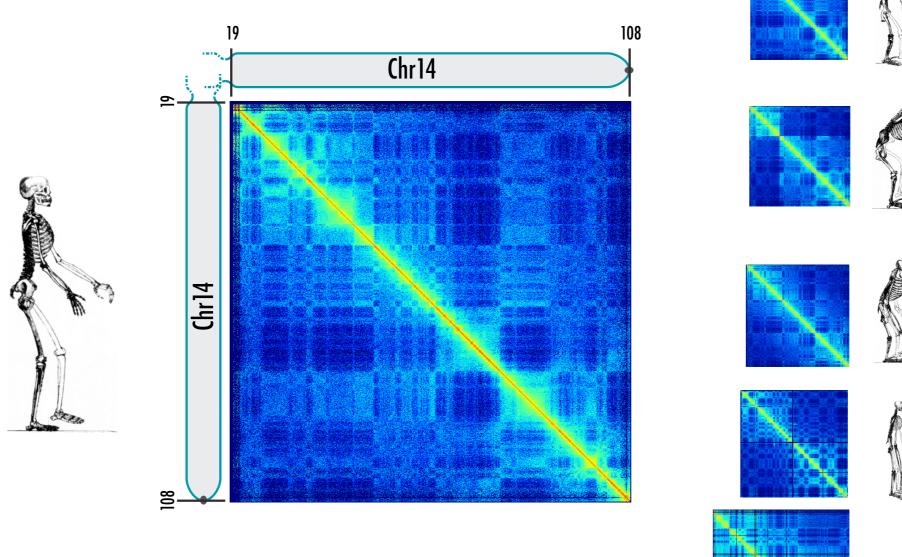
3C Detour... desirable side effects

srmation Capture _____assembly



Kaplan, N., & Dekker, J. (2013). High-throughput genome scaffolding from in vivo DNA interaction frequency. Nature

Great apes lymphoblast maps Chromosome 14

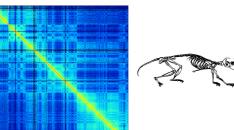


Chimpanzee

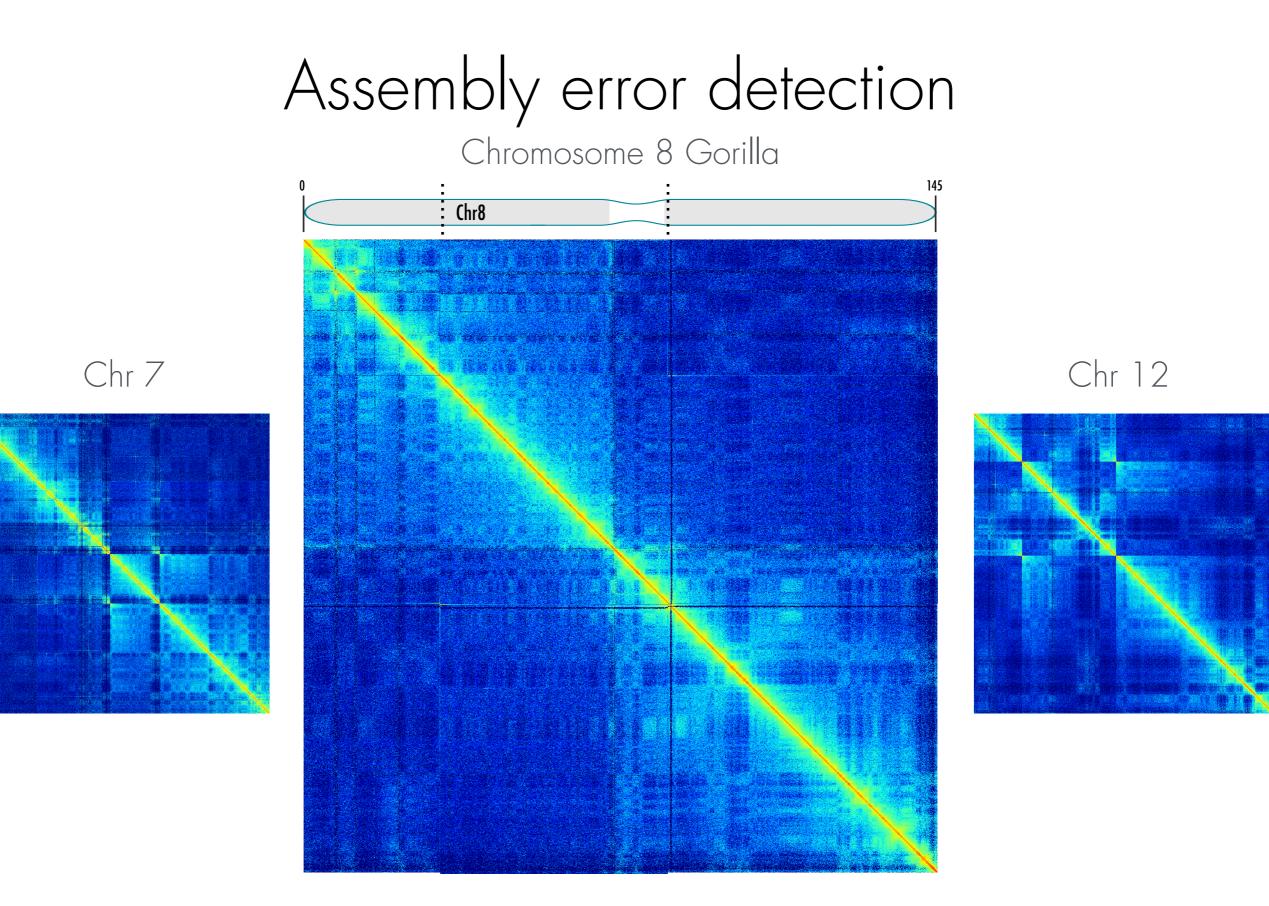
Gorilla

Orangutan

Gibbon

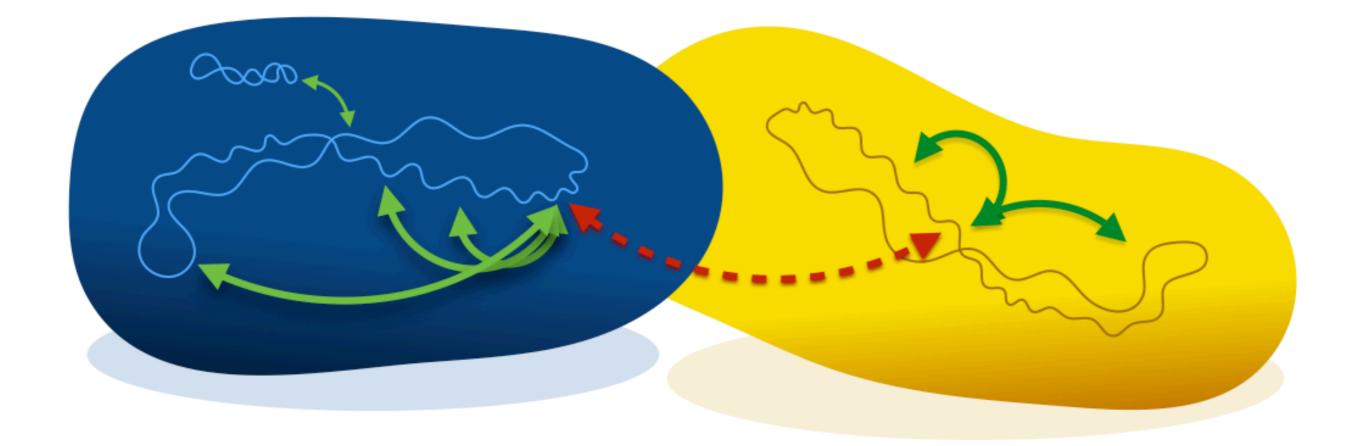


Mouse



GGO8 has an inversion of the region corresponding to HSA8:30.0-86.9Mb Aylwyn Scally (Department of Genetics, University of Cambridge)

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

Structure determination of genomes and genomic domains by satisfaction of spatial restraints

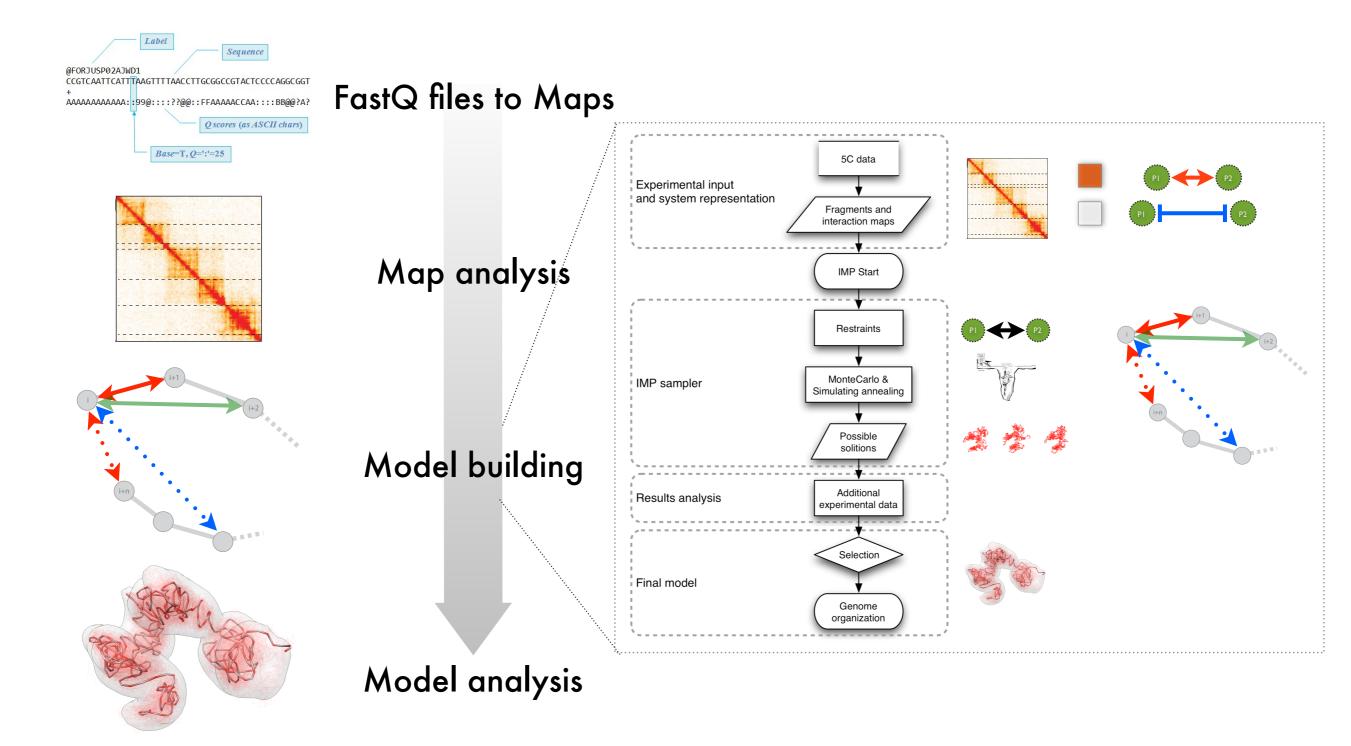
Marc A. Marti-Renom Structural Genomics Group (ICREA, CNAG-CRG)

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

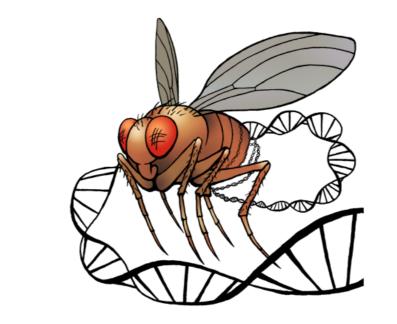


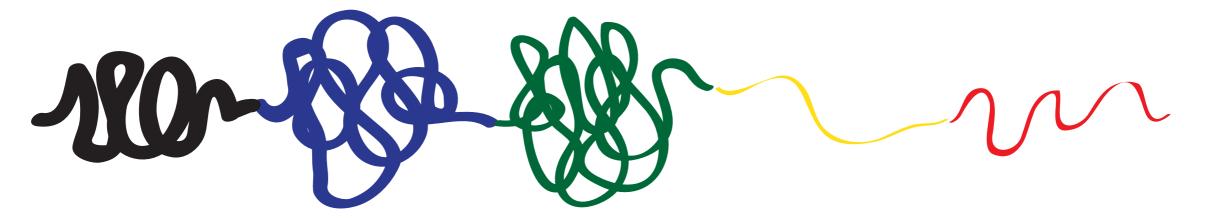


http://3DGenomes.org



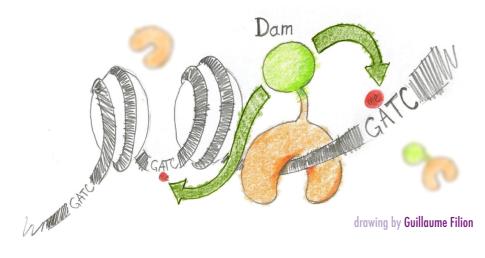
Structuring the **COLORs** of chromatin

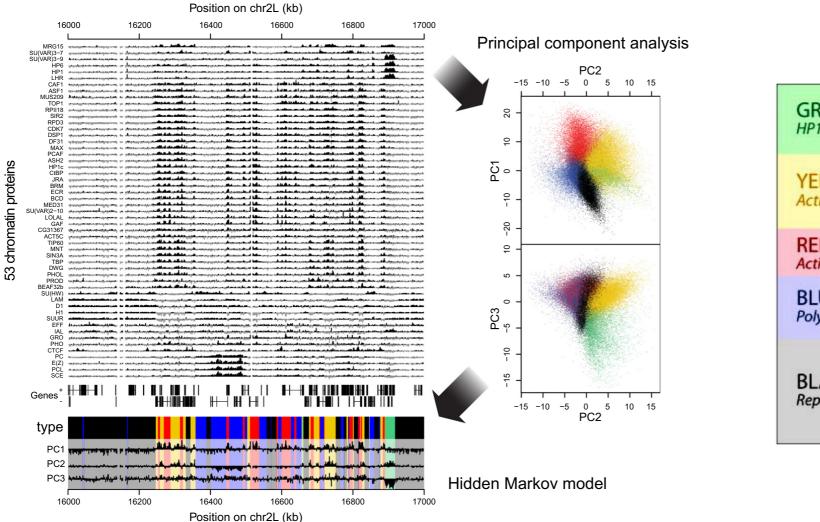


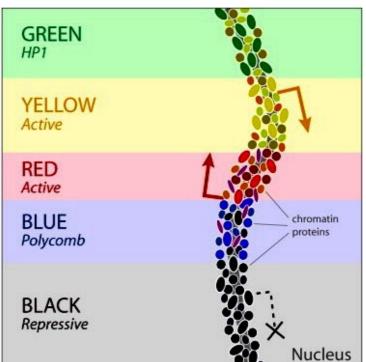


Fly Chromatin **COLORs**

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

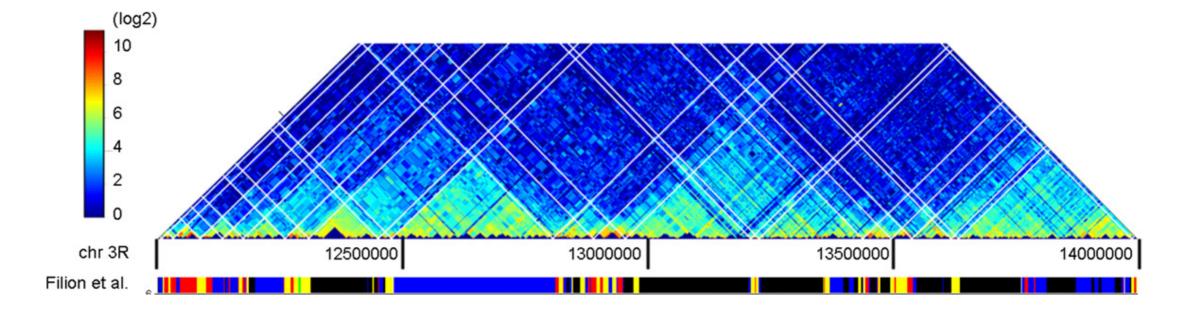


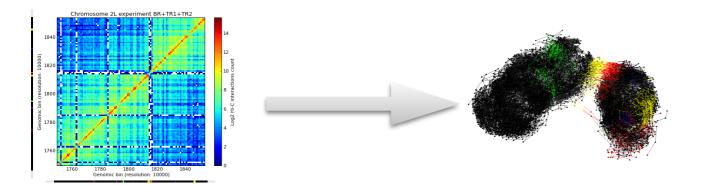




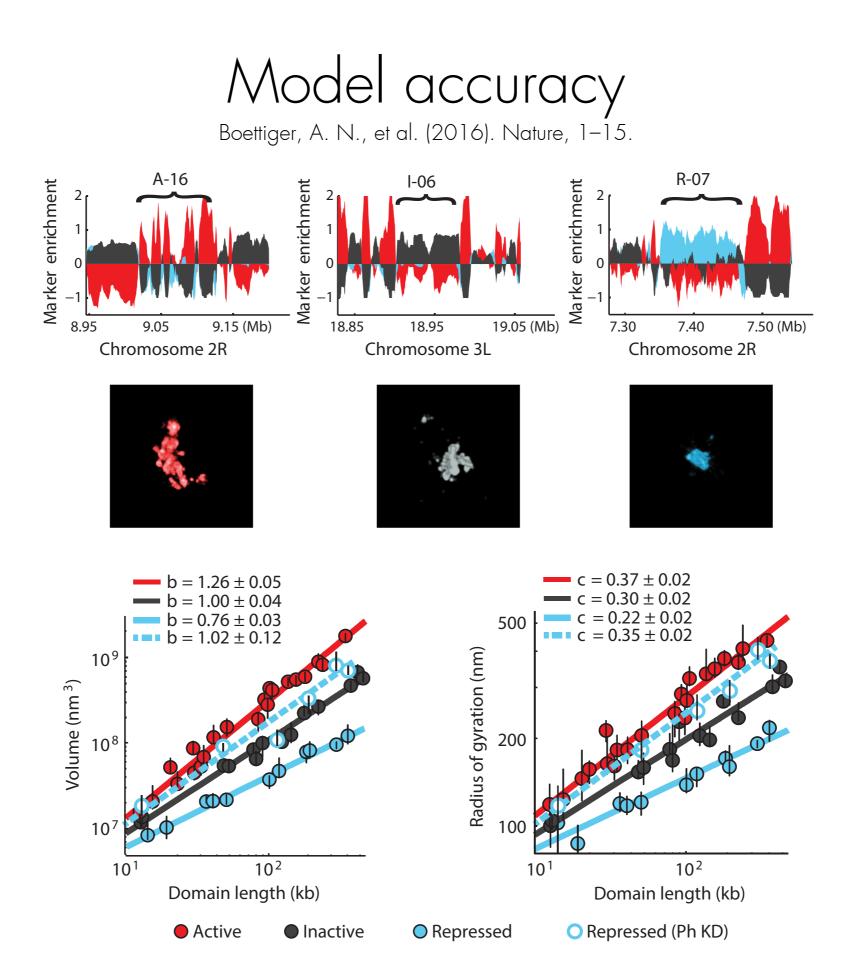
Fly Chromatin **COLORs**

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

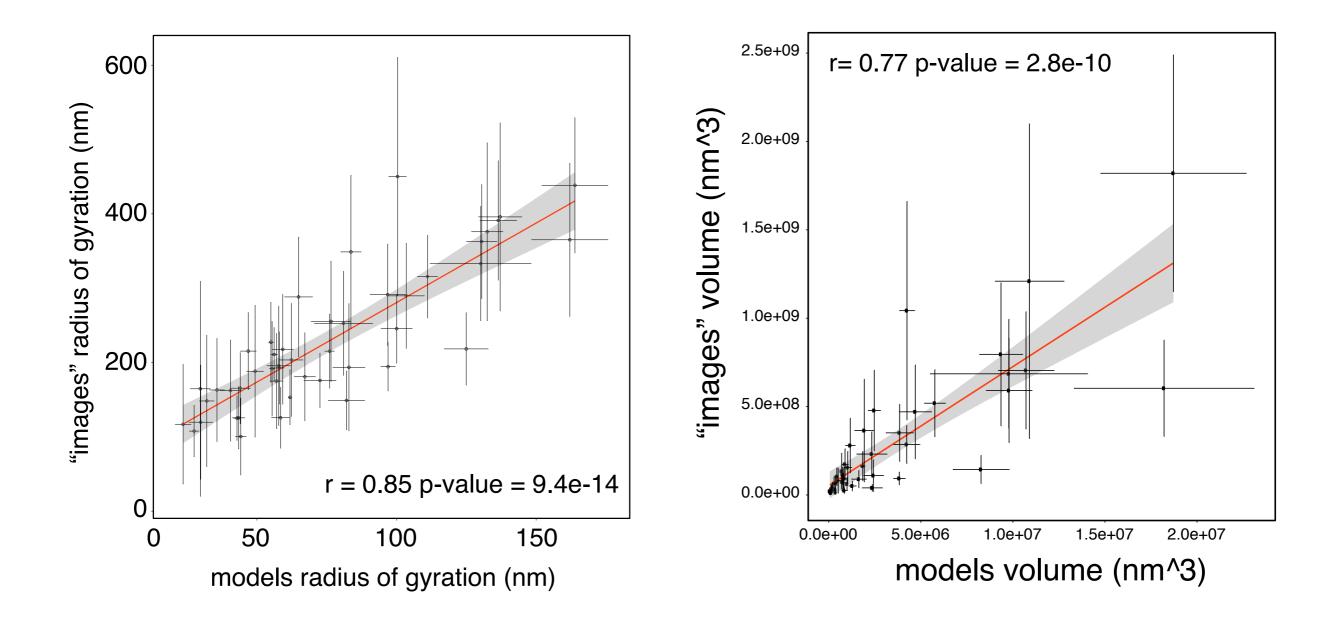




~200 regions of ~5Mb each 2Kb resolution

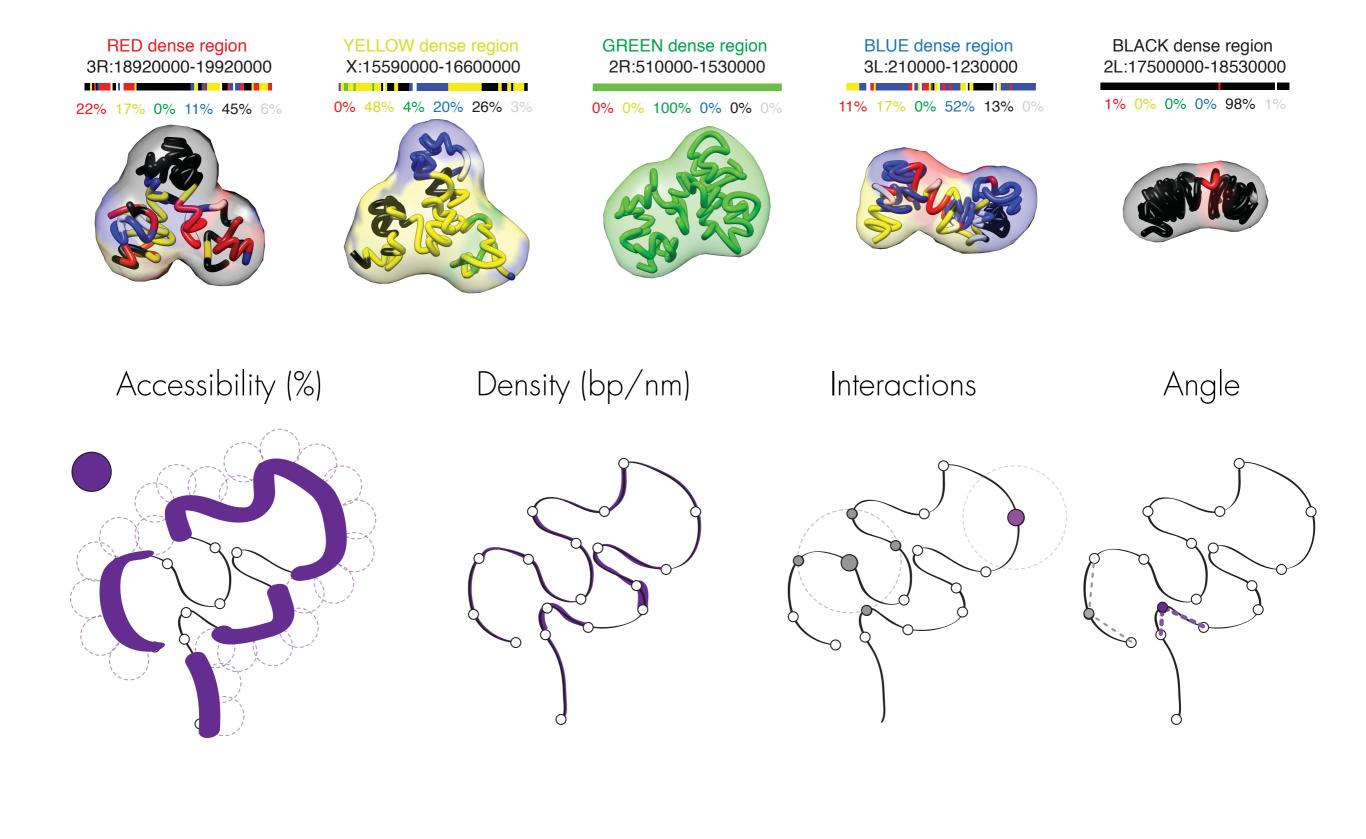


Model accuracy Boettiger, A. N., et al. (2016). Nature, 1–15.

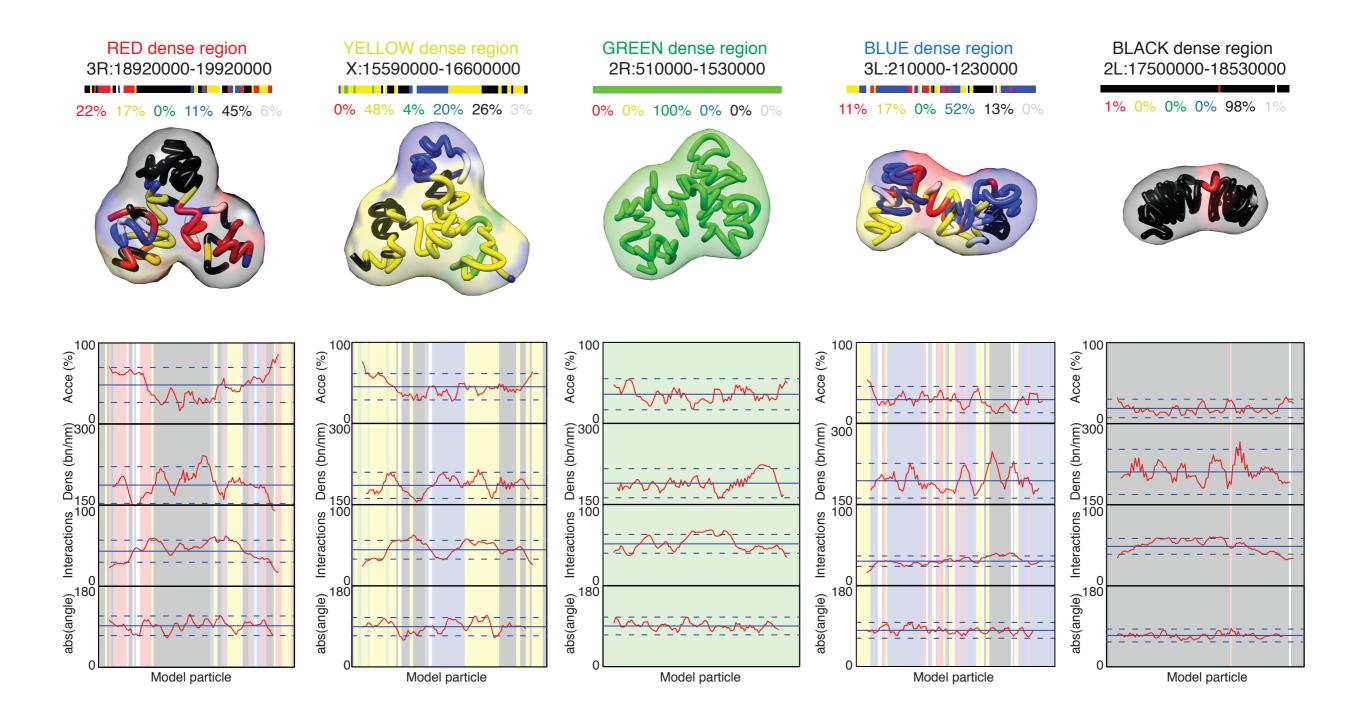


Structural properties

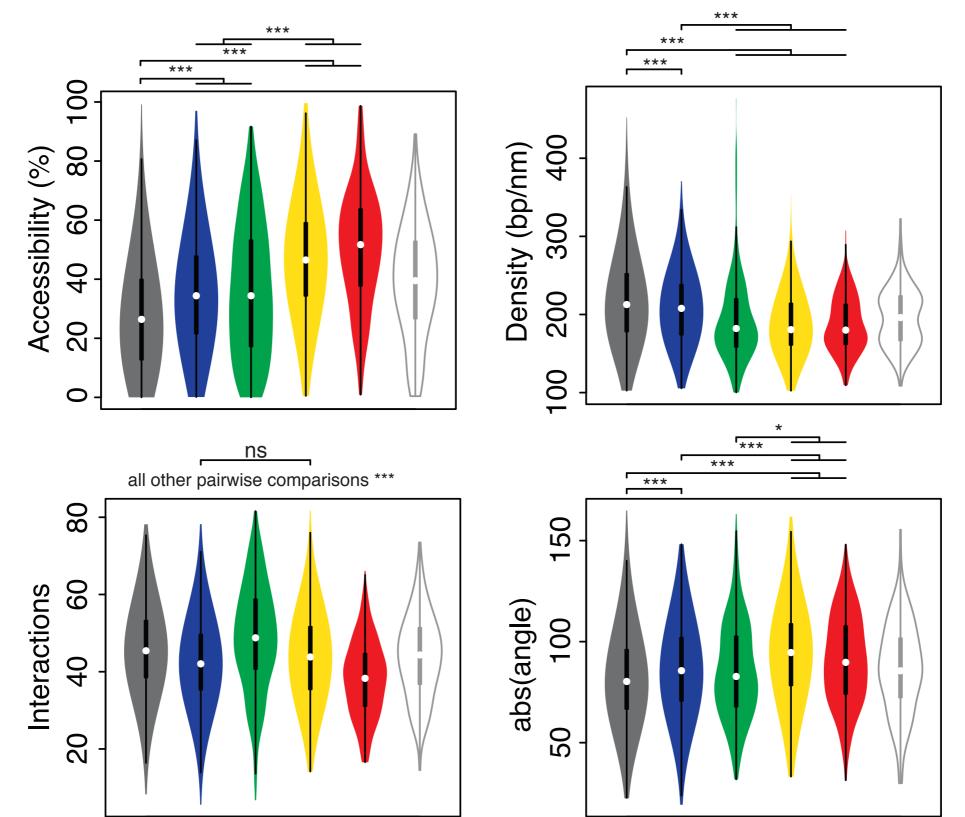
50 1Mb regions. 10 enriched for each color.



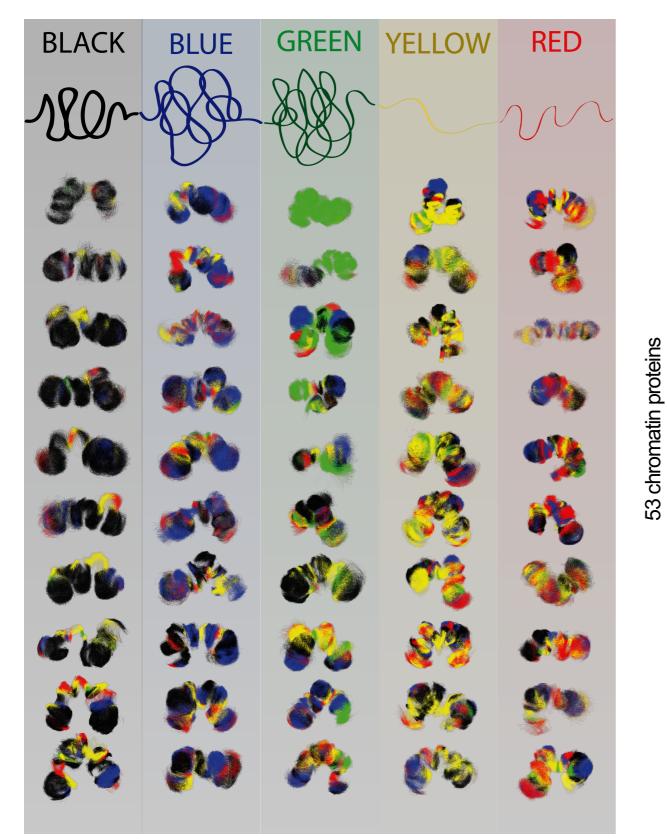
Structural **COLORs**



Structural **COLORs**

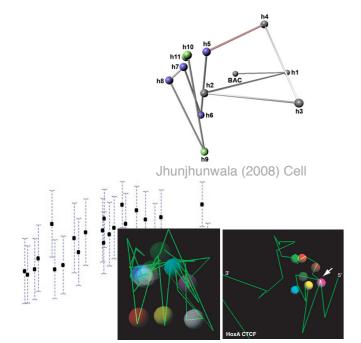


Structural **COLORs**



	Position on chr2L (kb)					
16	000	16200	16400	16600	16800	1700
	1					
MDO45				and the second second second	4 1000	
MRG15 SU(VAR)3-7	al production of the second seco	- sheller are shell and	to onthe later and a second design of the second		a har on a shall be a state of the second	Barg-up characters
SU(VAR)3-9				and an and and and the state of the second	and a second south of a law of the second se	
HP6	supportant and a second second second	- farmers			and a survey of a survey line of the line	and also represent
HP1 LHR		A desta d	and the second	and and the second s		
CAF1	Restary and a complete set of				and the second state of the second second	desired gardelar
ASF1	another and a second a	a station and a state Make		المرجو الجريد كالط ورديد الحا الرهو	and the second dia water and the second	and the second second
MUS209	Antonia and a second second	a glean the second blesses	antid an las on harry allowed	<u>لل بر معلم (ر قمارتم ، بر مرد 88 دو قرر</u> ۵. ممطور خطرین ، روی قرو میش	and a state of the second	the second second second
TOP1 RPII18	A low there is a second s	a standardard	al pilled and a second second second second	wheel the survey of the second s	and a standard on the state of the state of the state of the	State of the state of the
SIR2		- statut transmit			and the state of the second state of the secon	status littlean
RPD3		a strange segment.				datas lalkas
CDK7	alleftingertingentelenerge a	a prover and a state of the sta	and the second as a second second second	and the second states a	and the first state and the life search and because	for the starting
DSP1 DF31	dentities and dentities of a		and a second design of the sec	and the second section of the	white and a settle the set of the	provide contractor
MAX	100010 Acres 40000 Acres 4	-	and an		and a state of the	
PCAF	and a state of the second s	· statestatest	and the second sec	and the second states and an of the	an see well also be a strength of the second s	where wears
ASH2 HP1c	damilian hard a start	· plantage standard and and	which has a been a support of the support of the support	and the second states of the s		interim and day
CtBP	Contraction of the second seco	- Andrewson	and the second second diversion		man a second the set of the	Particle contractor
JRA	Administration of a		and the second s	Bard Man, & provide has des prove	warden and all and and and	And and an other state
BRM	Antona and a state of the	- statement -		A distance and the second	and the second second second second	Induction and second second
ECR BCD	deserves and the second	a photosphot		and the second states of a second		Industry and and
MED31	distribution of the second sec	- management	and a second second second second	م في الم الم الم الم الم الم		when the second
SU(VAR)2-10	to an exterior and the second	A. Susan and and a	and an an a gran and a second and	and her man and and an and		
LOLAL	and the second s	a all an	Abdases it and successive about	A LAL A LAL A		
GAF CG31367	and a second and a second seco	- glasses and a second se	and have some some share	and a second part of the second part of the second s		and the state of the
ACT5C	Section and stands	-		and the second sector and the second		- AND - CHANNE
TIP60		- sterestered	Aller A	want between sur the dame of the		
MNT SIN3A	alantikan seperahahan k	a philosophic and an	A Balance de la companya de la compa	and being and also arrest		desired excertain
TBP	Antiparticipation of the second se		48.44 ···· 4····	atur barren and a harring the		And and a state of the second
DWG		· paragenter de Marine	and a service of the second second			former and state
PHOL	Alexandration and a second sec	a galanan panganapalah ang		and and the second s	a local and the second s	presents developed
PROD BEAF32b	coplean and a group a	in all an approximation of the second	and a superior of the superior	An a second second second second		Antonio alternation
SU(HW)	and and a set of a second like it	a sundariation		and and the second states of the second		and a state of the
LAM		-	and the second	and the second second second second	and a second state of the second s	
D1	antikal. Bala de la companya de la	a de antes de la desta de la property	and the second sec	and a subdivision by dealer . Barra proving		
H1 SUUR	and the second s	the second second				
EFF	A water water and the second					
IAL	and the state of the second	N Sternsteinungen und sterner und sterner				
GRO PHO	increase and a second second	a source and and a second	and some many parts & and	and the second second second second		for the summary set
CTCF	and the second second of the second s	a state of the state of the	- del	the state of the state of the state of the state	a balan sound balance	and a second
PC						
E(Z)	abbreview and a second second		a de la companya de l	Any sectar be addressed able does		
PCL SCE	distant developed and a second	-	and a state of the second	the of the second states and second		and a second
SUE	whether a stranger of a					and the second second

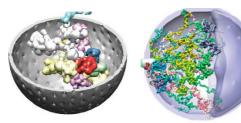
Are the models correct?



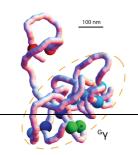


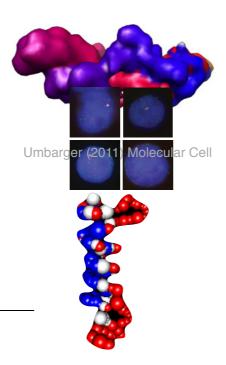


Fraser (2009) Genome Biology Baù (2011) Nature Structural & Molecular Biology Ferraiuolo (2010) Nucleic Acids Research



Kalhor (2011) Nature Biotechnology Tjong (2012) Genome Research





Nucleic Acids Research Advance Access published March 23, 2015

Nucleic Acids Research, 2015 1 doi: 10.1093/nar/gkv221

Assessing the limits of restraint-based 3D modeling of genomes and genomic domains

Marie Trussart^{1,2}, François Serra^{3,4}, Davide Baù^{3,4}, Ivan Junier^{2,3}, Luís Serrano^{1,2,5} and Marc A. Marti-Renom^{3,4,5,*}

¹EMBL/CRG Systems Biology Research Unit, Centre for Genomic Regulation (CRG), Barcelona, Spain, ²Universitat Pompeu Fabra (UPF), Barcelona, Spain, ³Gene Regulation, Stem Cells and Cancer Program, Centre for Genomic Regulation (CRG), Barcelona, Spain, ⁴Genome Biology Group, Centre Nacional d'Anàlisi Genòmica (CNAG), Barcelona, Spain and ⁵Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

Received January 16, 2015; Revised February 16, 2015; Accepted February 22, 2015

ABSTRACT

Restraint-based modeling of genomes has been recently explored with the advent of Chromosome Conformation Capture (3C-based) experiments. We previously developed a reconstruction method to resolve the 3D architecture of both prokaryotic and eukaryotic genomes using 3C-based data. These models were congruent with fluorescent imaging validation. However, the limits of such methods have not systematically been assessed. Here we propose the first evaluation of a mean-field restraint-based reconstruction of genomes by considering diverse chromosome architectures and different levels of data noise and structural variability. The results show that: first, current scoring functions for 3D reconstruction correlate with the accuracy of the models: second, reconstructed models are robust to noise but sensitive to structural variability; third, the local structure organization of genomes, such as Topologically Associating Domains, results in more accurate models; fourth, to a certain extent, the models capture the intrinsic structural variability in the input matrices and fifth, the accuracy of the models can be a priori predicted by analyzing the properties of the interaction matrices. In summary, our work provides a systematic analysis of the limitations of a meanfield restrain-based method, which could be taken into consideration in further development of methods as well as their applications.

INTRODUCTION

expression regulation and replication (1-6). The advent of the so-called Chromosome Conformation Capture (3C) assays (7), which allowed identifying chromatin-looping interactions between pairs of loci, helped deciphering some of the key elements organizing the genomes. High-throughput derivations of genome-wide 3C-based assays were established with Hi-C technologies (8) for an unbiased identification of chromatin interactions. The resulting genome interaction matrices from Hi-C experiments have been extensively used for computationally analyzing the organization of genomes and genomic domains (5). In particular, a significant number of new approaches for modeling the 3D organization of genomes have recently flourished (9–14). The main goal of such approaches is to provide an accurate 3D representation of the bi-dimensional interaction matrices, which can then be more easily explored to extract biological insights. One type of methods for building 3D models from interaction matrices relies on the existence of a limited number of conformational states in the cell. Such methods are regarded as mean-field approaches and are able to capture, to a certain degree, the structural variability around these mean structures (15).

We recently developed a mean-field method for modeling 3D structures of genomes and genomic domains based on 3C interaction data (9). Our approach, called TADbit, was developed around the Integrative Modeling Platform (IMP, http://integrativemodeing.org), a general framework for restraint-based modeling of 3D bio-molecular structures (16). Briefly, our method uses chromatin interaction frequencies derived from experiments as a proxy of spatial proximity between the ligation products of the 3C libraries. Two fragments of DNA that interact with high frequency are dynamically placed close in space in our models while two fragments that do not interact as often will be kept apart. Our method has been successfully applied to model the structures of genomes and genomic domains in eukaryote and prokaryote organisms (17-19) In all of our studies the final models were partially validated by assessing their

m correspondence should be addressed. Tel: +34 934 020 542; Fax: +34 934 037 279; Email: mmarti@pcb.ub.ca

Recent studies of the three-dimensional (3D) conforma-

tion of genomes are revealing insights into the organiza-

tion and the regulation of biological processes, such as gene

© The Author(s) 2015. Published by Oxford University Press on behalf of Nucleic Acids Research. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creati permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited. vecommons.org/licenses/by/4.0/), which

Trussart, et al. (2015), Nucleic Acids Research.

Junier (2012) Nucleic Acids Research

Hu (2013) PLoS Computational Biology

IMGR: a truly integrative method for chromatin path detection