

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

Marc A. Marti-Renom CNAG-CRG · ICREA

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











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



<image/>			$\begin{array}{c} & & & \\ & & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & &$	
			DNA length	
	10 ⁶		10 ⁹	nt
		0	Volume	
10 ⁻³	10 ⁰		10 ³	μm³
			Time	
10 ⁻⁴ 10 ⁻²	10 ⁰	10 ²	10 ³	S
			Resolution	
10 ⁻²		10 ⁻¹		μ

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

Experiments





Computation

Chromosome Conformation Capture

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.



















Lieberman-Aiden, E., et al. (2009). Science, 320(3930), 289–293. Rao, S. S. P., et al. (2014). Cell, 1–29.







TADs Chromosome 14

Lupiáñez, et al. (2015). Cell, 1–15.



TADs are functional units

TADs are functional units

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





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







Biomolecular structure determination 2D-NOESY data

Chromosome structure determination 3C-based data



Sequence @FORTUSP82AJWD1 CCGTCANTTCATTTAAGTTTTAACCTTGCGGCCGTACTCCCCAGGCGGT AMAMAMAAA: : 99@::::??@@::FFAAMACCAA::::BB@@?A? Queres (us ASCII chars) Base-T, Q=':'=25 _____ _ _ _ _ _ _ _ _ _ _ _

FastQ files to Maps

Map analysis

Model building

Model analysis

Serra, F., Baù, D. et al. PLOS CB (2017)



Model representation and scoring



Harmonic

$$H_{i,j} = k \Big(d_{i,j} - d_{i,j}^0 \Big)^2$$



Harmonic Upper Bound

$$\begin{cases} if \ d_{i,j} \ge d_{i,j}^{0}; & ubH_{i,j} = k \left(d_{i,j} - d_{i,j}^{0} \right)^{2} \\ if \ d_{i,j} < d_{i,j}^{0}; & ubH_{i,j} = 0 \end{cases}$$



Harmonic Lower Bound

$$\begin{cases} if \ d_{i,j} \le d_{i,j}^{0}; & lbH_{i,j} = k \left(d_{i,j} - d_{i,j}^{0} \right)^{2} \\ if \ d_{i,j} > d_{i,j}^{0}; & lbH_{i,j} = 0 \end{cases}$$







Parameter optimization

Optimization of the scoring function



Model analysis: clustering and structural features



Accessibility (%)

Density (bp/nm)





Interactions

Angle







Serra, Baù, et al. (2017). PLOS CompBio

ing	
1	
ation	
9	
/sis	

- Baù, D. et al. Nat Struct Mol Biol (2011)
- Umbarger, M. A. et al. Mol Cell (2011)
- Le Dily, F. et al. Genes & Dev (2014)
- Belton, J.M. et al. Cell Reports (2015)
- Trussart M. et al. Nature Communication (2017)
- Cattoni, D. et al. Nature Communication (2017)
- Stadhouders R. et al. Nature Genetics (2018)
- Kojic, A., Cuadrado, A. et al. Nat Struct Mol Biol (2018)
- Beekman R. et al. Nature Medicine (2018)
- Mas, G. et al. Nature Genetics (2018) in press



Automatic analysis and 3Dmodelling of Hi-C data using TADbit reveals structural features of the fly chromatin colors

Marc A. Marti-Renom CNAG-CRG · ICREA

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





Fly Chromatin **COLORs** Filion et al. (2010). Cell, 143(2), 212–224.



Position on chr2L (kb)





Fly Chromatin **COLORs** Hou et al. (2012). Molecular Cell, 48(3), 471–484.







All fly genome @2Kb resolution

Structural properties

RED dense region 3R:18920000-19920000 22% 17% 0% 11% 45% 6%

YELLOW dense region X:15590000-16600000

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





Accessibility (%)

Density (bp/nm)





GREEN dense region 2R:510000-1530000

0% 0% 100% 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%



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%



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







GREEN dense region 2R:510000-1530000

BLUE dense region 3L:210000-1230000

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



BLACK dense region 2L:17500000-18530000

1% 0% 0% 98% 1%





Structural **COLORs**



Structural **COLORs**



		F	Position on ch	r2L (kb)		
16	000	16200	16400	16600	16800	1700
		1	1		1	
MRG15	alanda ana ana ana a	· spinetics are spinetic and the second statements	mental and the second	and the state of the	Contraction of the second state of the second	ob -consulta
	manufacture and an and a	a shared and a state of the sta	number for the first state of the second	and the second sec	and the state of t	the section of the
50(VAR)3-9	here and here and the second	- Annual Contraction of the Annual Contraction of the International Contractional Contr	Contraction of the second sector of the sect			
	Andread and a second	- Handlinger and - Handlinger	and a first one of the second second			the carefulate
IHR		a descent of the second states of the second states and the second		the second se	and the state of t	
CAF1	Restatuted documents of the second	a contact of the state state of the state of	destroyed the second of the ball	and a state as the surger where	and a state of the second seco	
ASF1	another advantation of a	a statement of the Mandal Anna	an and the second second second	and a state of the second second	and a come the allance and the same sur	
MUS209	Antonia traditiona antonia antonia antonia	. desauresessiblesauresestillesaure	an a	man a block of the second states	- she ash del	and the second
TOP1	and the substantial sector of the	a province province the party of the party	and the second state of the second states	and the second states of the s	Marting Maria and a marting of	the street and
RPII18	and any property and a second second second	and the second statement of th	And the state of t	and the second	and a second	
SIR2		· · ·	and the state of t		and the second state of th	
RPD3	manda and an open service of the ser	a strange water of the state of the state of the	And a state of the state of the second	and a state of the state of the second	Ballan age able able as a state of the second	es papers
CDK7	allalangertersenationer a	a strengt and the state of the second state of the second	an and a second second second second	ana a the state of the state	and the second	
DSP1	demostrations of a	a share any restantial and related and and	Anna and a second second second second	and the state of t		dit reprinte
DF31	desires	A ARCHINGTON AND A CONTRACTOR OF A	water the second distance and and have	And the second s	1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1	da amquine
MAX	damps and all a sub-	a strangerungen ander an	to an an a state of the second s	a second at an an after some fitte aster tel	and the state of the second	dip entrutingen
PCAF	and a state of the	· sand and a standard of the s	and the first of the second se	starting of the second s	and the second se	on sponsiture
	dismalify gardenergy applied and	- plantage any production of the second states of the	and the state of t			the spin day
	entrefficients for a state of the	A Address of the Addr	And the state of t			an and have
	and a second	a state of the second sec	the second state of the se			and static days
BRM	dimmentania alla a		and the state of t	All the second second second	and the second sec	and the state of t
FCR	and the state of the second	a stand of the state of the second state of th	manage and a descendent little	and an and and and and and and and and a	and and a star and a star and a star and	
BCD	mercenter and a state of the		anna and a state of the second states	and and the Barris and a stand of the	and see and the second se	and an and
MED31	distant and the second se	- manana marked and and	- And have a strange of the second	and and the second second second	and an and the state of the second state of th	ay distants to
U(VAR)2-10	and the second second	a garante the the second and a second and a second se	and have a second description of the second	In such as the set of some designed and	and a second de la seconda de	
` LÓLAL	the sector of th			and a state of the second state of the second	and a state of the second s	
GAF		- glassester, and the Market stranger and	and the second	الليود منافر حجر مردو فردت أله أليا أحدر جرمنا		
CG31367	Adaption and a second should be	a democratic states and a second states	when had a state of the state of the	proper a la santa da da la despetita con color	and a set of a particular second and a set of a	an spectrum
ACT5C	helicialities is a second restriction of the	· State and a second state of the second state	reading the second second second	and a second starting and a second	and and the second second second second second	to and days
TIP60	freezensperieste state and the second state of	· · · · · · · · · · · · · · · · · · ·	to-organization-official and believe	and the second sec	**************************************	and service and
MNI	abantation and a second advanced a	a philippine philippine and a state of the s	Construction of the second	and the state of the second se	and the state of t	of exception
SIN3A	Admittages share and a	- standard -	and the state of the second second second	and the second se	and the second	the sparse files
IBP	Admistration of the second sec	a harden and harden and a second second	and an and the second second second second			di angeler
	modulingerstrendsenthaltered to		and the second			the openanties
	Algorithe and algorithm of	and the second sec	and the second second second		and the state that we do not set	the surgering
BEAE32b		and the second s	dense ber	In succession of the state		
SU(HW)	and and a safe out a summer shifted to	A sector and a sector of the	All advertising the second second			
LAM		and the state of t	a can the contract of the second of the	and and the second state of the second		
D1						
H1						
SUUR	al difficient south the of the	a show and a second	and the state of the second second second		and a state of the second s	
EFF	to setting the state of the set	a construction of the second second second	a set to be a state of the set of	م م م م م م م م م م الم م م م م م م م م	And the second state of th	
IAL	ى يىنىقۇرىلىردىلېقىدىناتېمرىدىرى	a demandary and second the second from the second	mental production of the second second	and and tree bolloged adapted and a second	and the state of t	in the state
GRO	there we are a second and a second	a dissinguisting in the second se		and a second sec	and a state of the second s	in a superior
PHO	discription of the second seco	· substant and a state of the s	an and the state of the state of the second states	and a subscription of the second data of the second data	and the state of the second state of the secon	ob and date
	and water and a second state of the second s	· Passes and all a colling of the second	andressed and a second second second second		A REAL PROPERTY AND A REAL PROPERTY A REAL PROPERTY AND A REAL PROPERTY AND A REAL PRO	and strangers
	an an a guild a finge of a sector of party of the special sector of the	· ····································	and the state of the state	A DESCRIPTION OF THE OWNER OF THE		the pulper-open
	and an	and the second se	the second state and should be a life of			an alternative
SCE	and an and a state of the state		and the second sec	and the second se		
JOL	a training and a second of					

. . .



3D structural dynamics of the SOX2 locus activation

Marc A. Marti-Renom CNAG-CRG · ICREA

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





Marco di Stefano CNAG-CRG



Ralph St CRG



Ralph Stadhouders, Enrique Vidal & Thomas Graf

Transcription factors dictate cell fate

Graf & Enver (2009) Nature



Transcription factors (TFs) determine cell identity through gene regulation Normal 'forward' differentiation

Cell fates can be converted by enforced TF expression

Transdifferentiation or reprogramming



Interplay: topology, gene expression & chromatin

Stadhouders, R., Vidal, E. et al. (2018) Nature Genetics







Reprogramming from B to PSC Stadhouders, R., Vidal, E. et al. (2018) Nature Genetics

Hi-C maps of reprogramming from B to PSC The SOX2 locus

B cell

50 100 150 200 250



Hi-C maps of reprogramming from B to PSC The SOX2 locus



How does these structural rearrangements interplay with the transcription activity?

What are the main drivers of structural transitions?



Optimal IMP parameters lowfreq=0, upfreq=1, maxdist=200nm, dcutoff=125nm, particle size=50nm (5kb)

TADbit modeling of SOX2 from B cells Hi-C

Models of reprogramming from B to PSC The SOX2 locus



TADdyn. Dynamics of chromatin



Chain-connectivity interaction Bending **Lennard-Jones Potential**



Marco Di Stefano



$$+ U_{br}(i, i + 1, i + 2) + \sum_{j=i+1}^{N} U_{LJ}(i, j)$$

TADdyn: from time-series Hi-C maps to dynamic restraints The SOX2 locus



TADdyn: from time-series Hi-C maps to dynamic restraints The SOX2 locus



TADdyn: from time-series Hi-C maps to dynamic restraints The SOX2 locus



Energy penalty

Transition	Stable	Vanishing	Raising
Β -> Β α	18,612	6,984	7,290
Β α -> D2	18,512	7,390	6,687
D2 -> D4	18,369	6,830	6,893
D4 -> D6	18,971	6,291	7,289
D6 -> D8	20,167	6,093	6,250
D8 -> ES	20,679	5,738	6,173

SOX2 locus structural changes from B to PSC Contacts





















SOX2 locus structural changes from B to PSC Contacts





















SOX2 locus structural changes from B to PSC TAD borders



SOX2 locus structural changes from B to PSC TAD borders

SOX2 locus structural changes from B to PSC Distance to regulatory elements

SOX2 locus structural changes from B to PSC Distance to regulatory elements

SOX2 locus structural changes from B to PSC Structural exposure

SOX2 locus structural changes from B to PSC Structural exposure

SOX2 locus dynamics changes from B to PSC SOX2 displacement

SOX2 locus dynamics changes from B to PSC SOX2 displacement

SOX2 locus dynamics changes from B to PSC SOX2 displacement

Two dimensional trajectories and area explored over 50s of the CCND1 locus recored before -E2 and after +E2 activation.

Germier ,T., et al, Blophys J. 113, 1383–1394 (2017).

A "cage" model for transcriptional activation

Chromosome walking with super-resolution imaging and modeling

Marc A. Marti-Renom CNAG-CRG · ICREA

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

Irene Farabella CNAG-CRG

Guy Nir Harvard Med School

Ting Wu Harvard Med School

Can we walk the chromatin path in the nucleus?

by

Integrating imaging and Hi-C maps with modeling.

by developing a method for

Oligopaint-based modeling of genomes

High-resolution imaging Tracing chromosomes with OligoSTROM & fluidics cycles in PGP1 cells

Beliveau et al. Nat. Comm. 2015

chr19:7,335,095-15,449,189 ~8Mb

homologous 32-42bp

High-resolution imaging Tracing chromosomes with OligoSTROM & fluidics cycles in PGP1 cells

Carl Ebeling Bruker

High-resolution imaging Tracing chr19:7,335,095-15,449,189 ~8Mb 3 9

1,280Kb

1,240Kb

1,800Kb

1,040Kb

520Kb 520Kb 840Kb

٩.

520Kb 360Kb

.

Cell-02

High-resolution imaging XYZ points convolution into a density map

Cell-02 · Segment 1

$$\frac{Z_{N}}{\sqrt{2\pi}^{3}}e^{-\frac{(x-x_{n})^{2}+(y-y_{n})^{2}+(z-z_{n})^{2}}{2\sigma^{2}}}$$

Farabella et al, J Appl Crystallogr. 2015

Density maps Cell-02 · Density map @ 50nm

Area (nm^2) Volume (nm³) Sphericity Overlap (%) Distance (nm)

Farabella et al, J Appl Crystallogr. 2015

Structural features Area, Volume and Sphericity of 19 cells each with 2 homologous resolved

Area

Spatial arrangement Distance and overlap of 19 cells each with 2 homologous resolved

Diff. distance

Diff. overlap

Structural clustering 19 cells each with 2 homologous and 9 segments each (342)

cluster 2

15

10

PGP1 ChIP-seq and Hi-C data from ENCODE and Lieberman-Aiden Lab, respectively

89

Cluster properties A/B compartment properties

Can we walk the chromatin path in the nucleus?

Can we increase the resolution of our data?

by fitting 3D models based on Hi-C interaction maps

YES!

Increasing resolution Rigid body fitting 3D structures based on Hi-C data

Farabella et al, J Appl Crystallogr. 2015 Roseman, 2000; Wriggers & Chacon, Structure 2001

Increasing resolution Flexible fitting 3D structures based on Hi-C data

Increasing resolution Flexible fitting 3D structures based on Hi-C data

Chromosome walking path @10Kb resolution

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

> David Castillo Yasmina Cuartero Marco Di Stefano Irene Farabella Silvia Galan Mike Goodstadt Francesca Mugianesi Julen Mendieta Juan Rodriguez François Serra Paula Soler Aleksandra Sparavier

Multiscale Complex Genomics

MINISTERIO DE ECONOMÍA, INDUSTRIA Y COMPETITIVIDAD

.: Our current sponsors :.

