Inter-chromosomal interaction analysis based on population 3D genome modeling

Chen Fengling

20170218

Genome architectures revealed by tethered chromosome conformation capture and population-based modeling

Reza Kalhor[1,2], Harianto Tjong[1], Nimanthi Jayathilaka[1,2], Frank Alber[1] & Lin Chen[1,3,4]

Received 22 February; accepted 2 November; published online 25 December 2011; doi:10.1038/nbt.2057

Population-based 3D genome structure analysis reveals driving forces in spatial genome organization

Harianto Tjong[a,1], Wenyuan Li[a,1], Reza Kalhor[a], Chao Dai[a], Shengli Hao[a], Ke Gong[a], Yonggang Zhou[a], Haochen Li[a], Xianghong Jasmine Zhou[a], Mark A. Le Gros[b,c,d], Carolyn A. Larabell[b,c,d], Lin Chen[a,e], and Frank Alber[a,2]

[a]Molecular and Computational Biology, Department of Biological Sciences, University of Southern California, Los Angeles, CA 90089; [b]Department of Anatomy, University of California, San Francisco, CA 94148; [c]Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94702; [d]National Center for X-Ray Tomography, Advanced Light Source, Lawrence Berkeley National Laboratory, Berkeley, CA 94702; and [e]Department of Chemistry and Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, Los Angeles, CA 90089

Edited by José N. Onuchic, Rice University, Houston, TX, and approved January 29, 2016 (received for review June 26, 2015)

ARTICLE

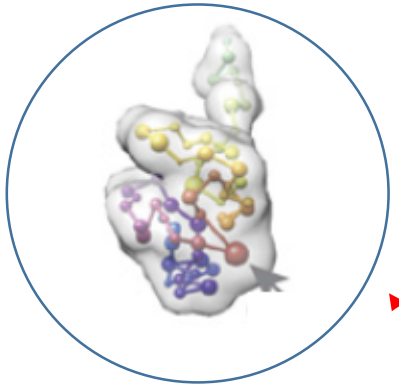
Received 18 Dec 2015 | Accepted 8 Apr 2016 | Published 31 May 2016

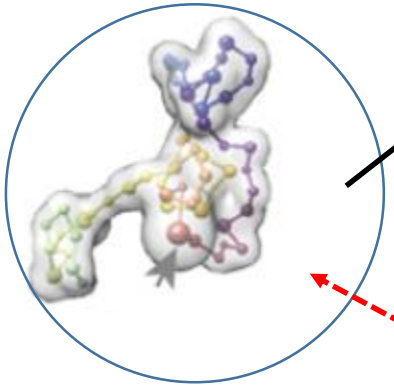
DOI: 10.1038/ncomms11549 OPEN

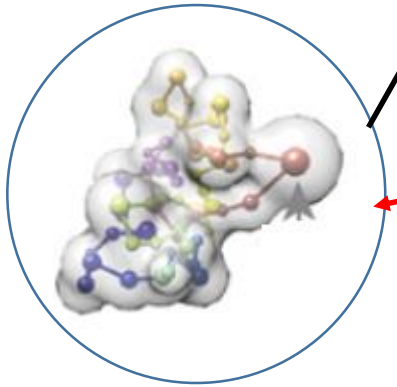
Mining 3D genome structure populations identifies major factors governing the stability of regulatory communities

Chao Dai[1,*], Wenyuan Li[1,*], Harianto Tjong[1,*], Shengli Hao[1], Yonggang Zhou[1], Qingjiao Li[1], Lin Chen[1], Bing Zhu[2], Frank Alber[1] & Xianghong Jasmine Zhou[1]

cell populations





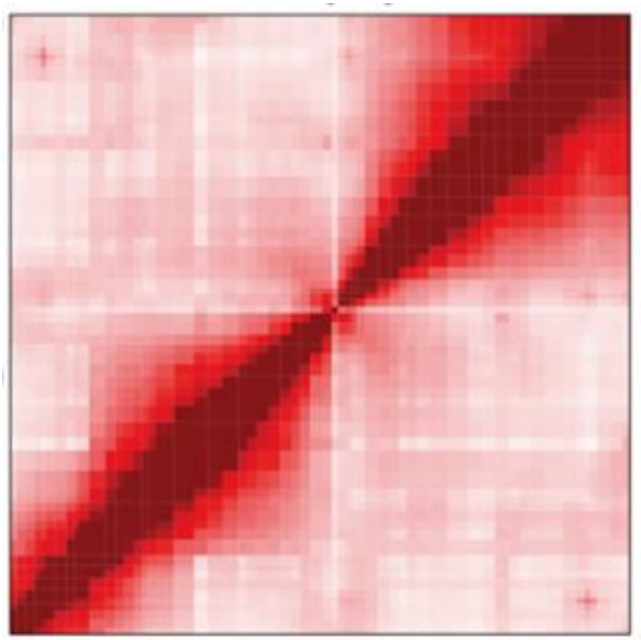


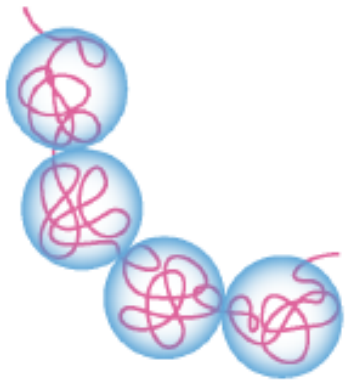
TCC

Conformation capture technologies

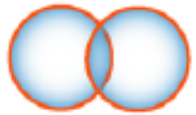
How to compute specified structure from average structure ?

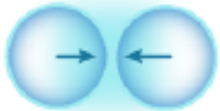
Average structure





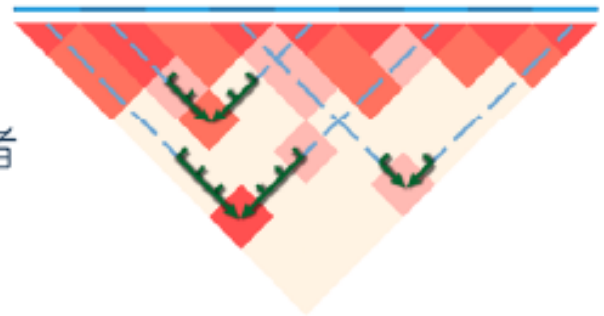
势能





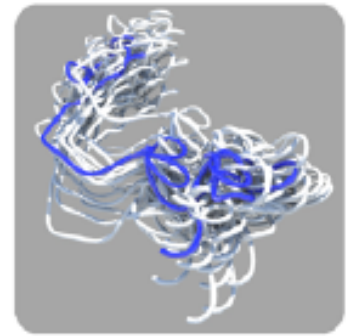
等

或者

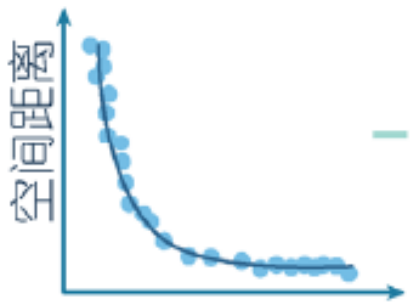


交互频率

计算机模拟



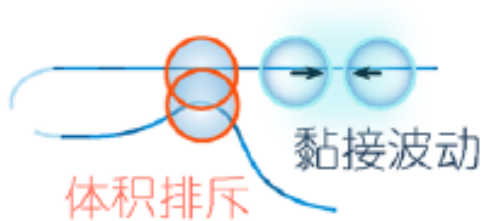
距离转换



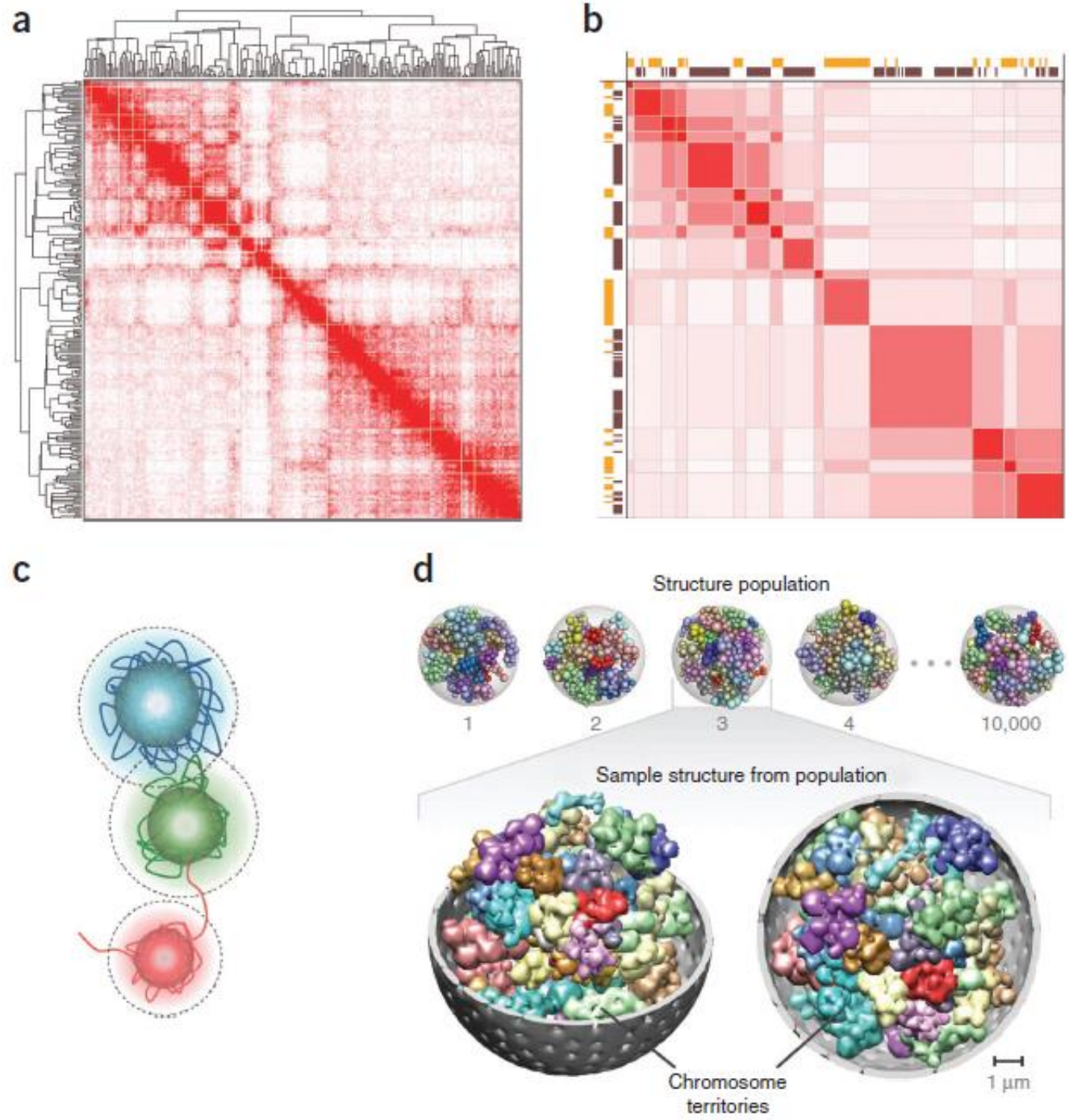
目标函数

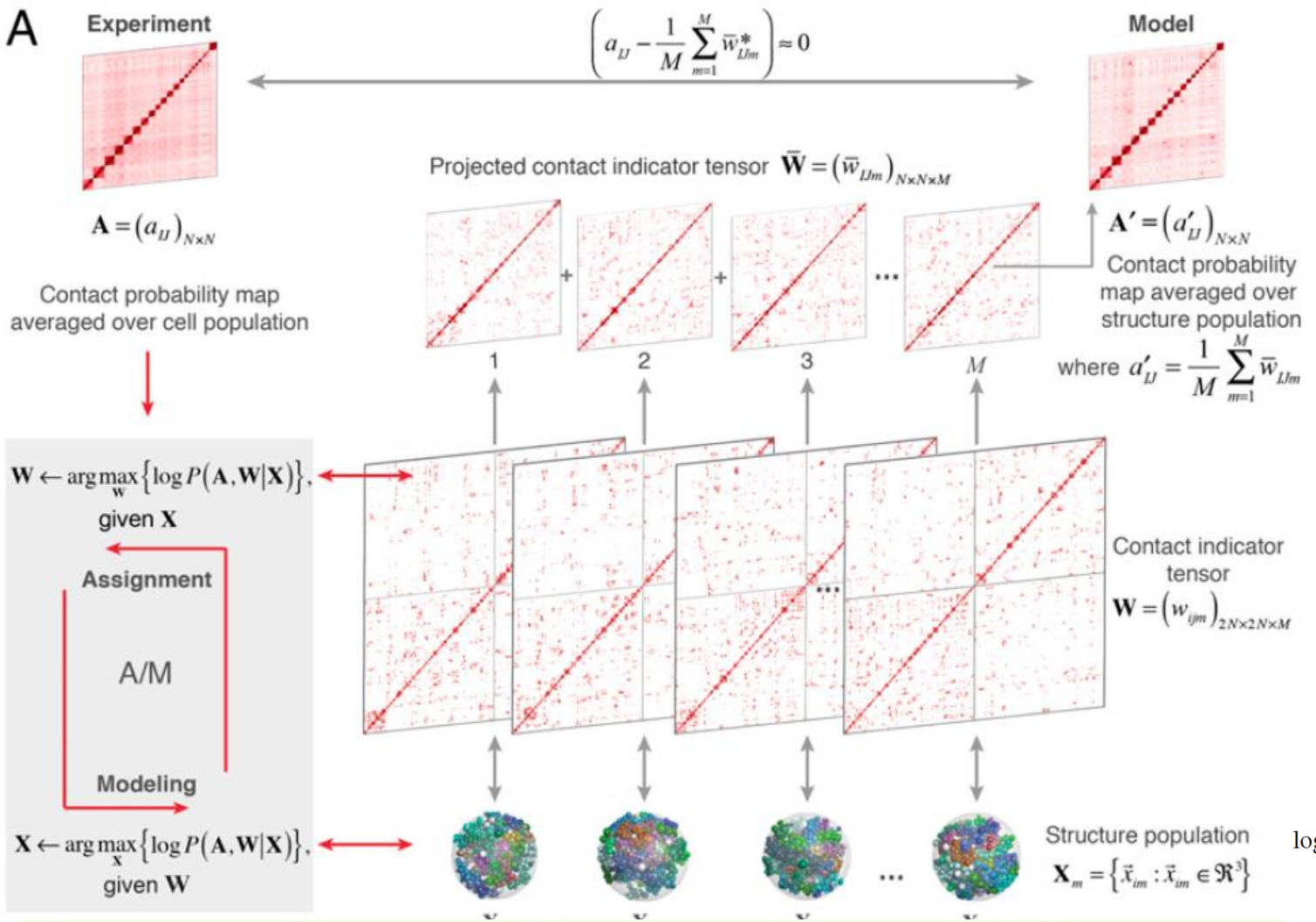
$$(P_1,\cdots,P_n) = arg\ min\ f(||P_i-P_j||,d_{ij})$$

约束条件



有约束优化


a
b
c
d
Structure population
1
2
3
4
10,000
Sample structure from population
Chromosome territories
1 µm



$\mathbf{A}=(a_{IJ})_{N\times N}$ — N domains in the genome, where $0 \le aIJ \le 1$ is the contact probability of two chromosome domains I and J

$\mathbf{X}=\{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_M\}$ — a set of M diploid genome structures

$\mathbf{X}_m=\{\vec{x}_{im}:\vec{x}_{im}\in\Re^3,\quad i=1,\ 2\ \ldots,\ 2N\}$ — a set of 3-dimensional vectors representing the center coordinates of $2N$ spheres

$\mathbf{W}=(w_{ijm})_{2N\times 2N\times M}$ — $wijm = 1$ indicates that the contact between domains i and j in structure m; $wijm = 0$ otherwise

$$\hat{\mathbf{X}}=\arg\max_{\mathbf{X}}\max_{\mathbf{W}}\{\log \mathrm{P}(\mathbf{A},\mathbf{W}\mid\mathbf{X})\}$$

subject to
- spatial constraint I: nuclear volume
- spatial constraint II: excluded volume
- spatial constraint III: 4qtel-NE proximity

$$\log P(\mathbf{A},\mathbf{W}|\mathbf{X})=\log P(\mathbf{A}|\mathbf{W})+\log P(\mathbf{W}|\mathbf{X})=\sum_{\substack{I,J=1\\I\neq J}}^{N}\log P(a_{IJ}\mid a'_{IJ})+\sum_{m=1}^{M}\sum_{\substack{i,j=1\\i\neq j}}^{2N}\log P\left(w_{ijm}\middle|\vec{x}_{im},\vec{x}_{jm}\right)$$

$$\sum_{m=1}^{M}\sum_{\substack{i,j=1\\i\neq j}}^{2N}\left[w_{ijm}\log P(w_{ijm}=1|d_{ijm})+(1-w_{ijm})\log P(w_{ijm}=0|d_{ijm})\right]$$

when d_{ijm} is larger than a distance threshold value (termed as d_{IJ}^{act}), let $w_{ijm}=0$

d_{ijm} is smaller than a distance threshold d_{IJ}^{act}, let $w_{ijm}=1$

ARTICLE

Received 18 Dec 2015 | Accepted 8 Apr 2016 | Published 31 May 2016

DOI: 10.1038/ncomms11549 OPEN

Mining 3D genome structure populations identifies major factors governing the stability of regulatory communities

Chao Dai[1,*], Wenyuan Li[1,*], Harianto Tjong[1,*], Shengli Hao[1], Yonggang Zhou[1], Qingjiao Li[1], Lin Chen[1], Bing Zhu[2], Frank Alber[1] & Xianghong Jasmine Zhou[1]

outline

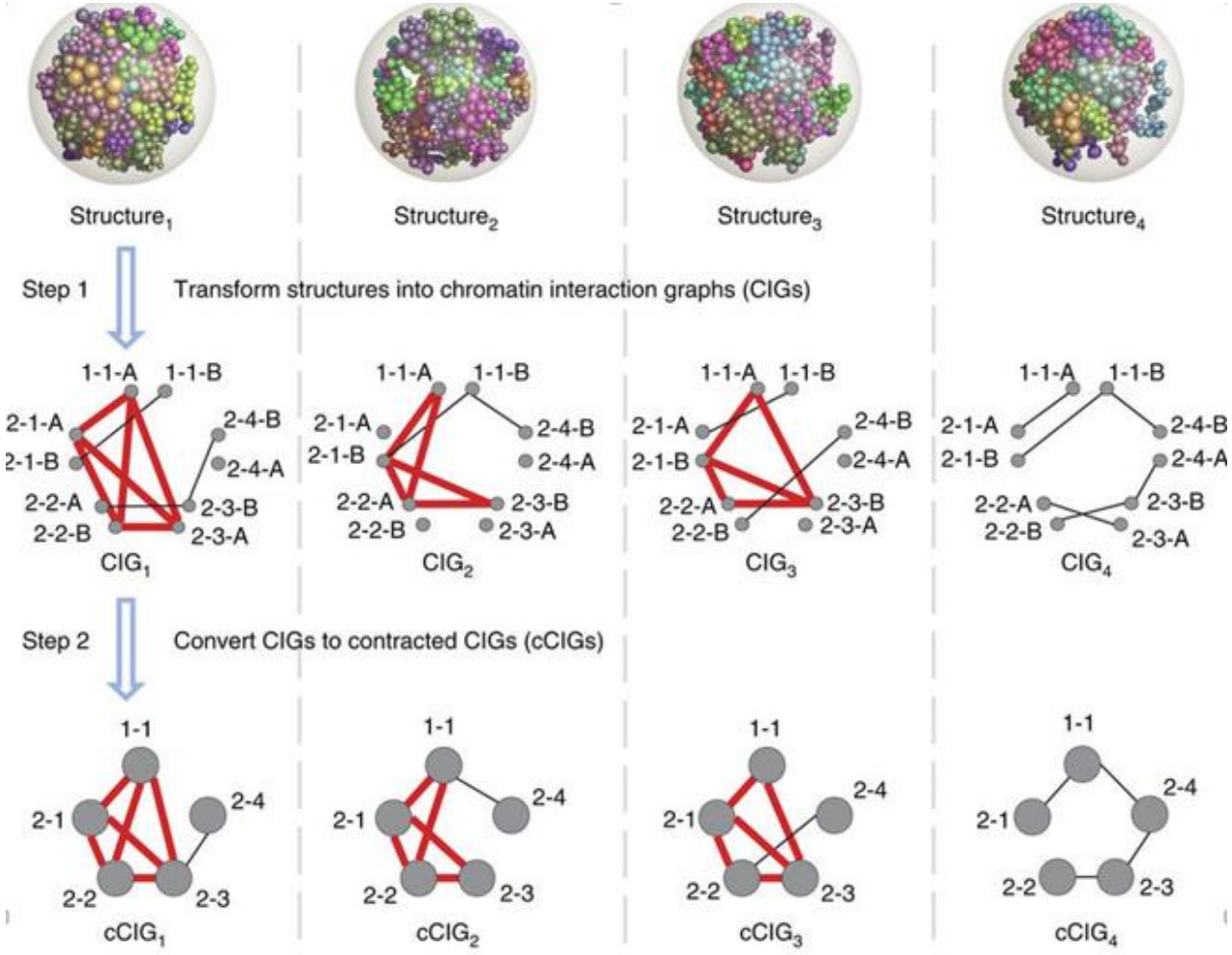
1.A graph-based computational framework for the analysis of 3D genome structure populations

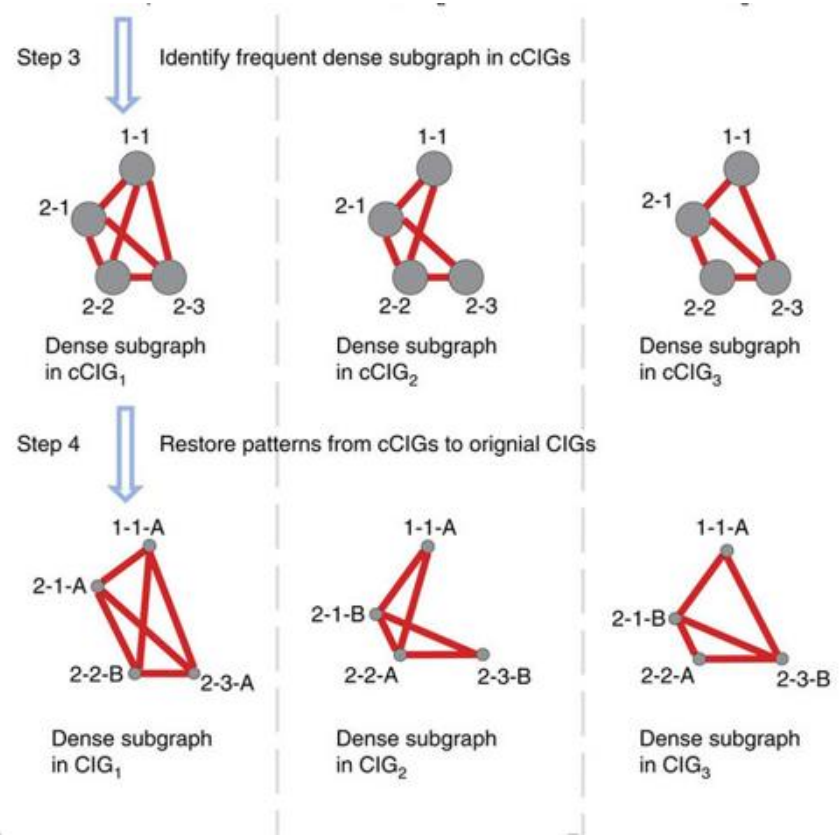
2. frequently occurring chromatin clusters are enriched in binding of specific regulatory factors

3.Two major factors, centromere clustering and transcription factor binding, significantly stabilize such regulatory communities

4.The regulatory communities differ substantially from cell to cell

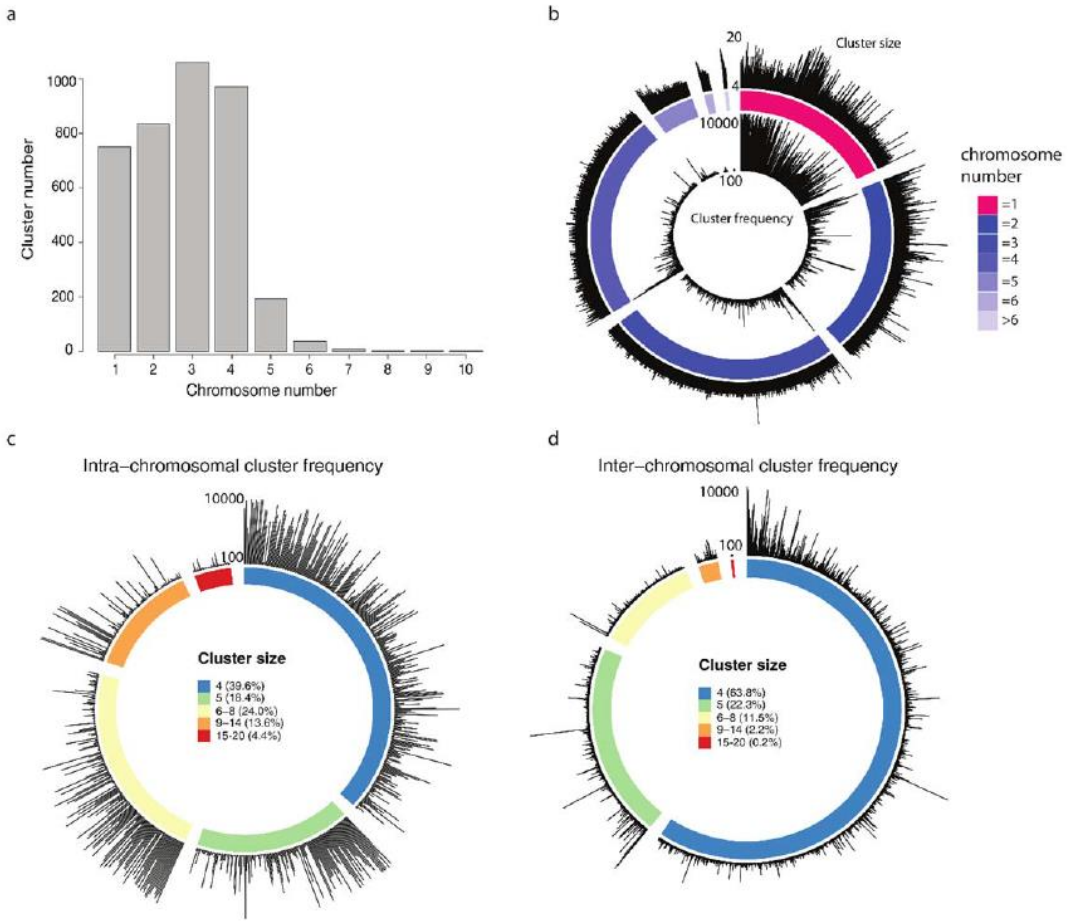
1.Discover frequent spatial clusters in a 3D genome population

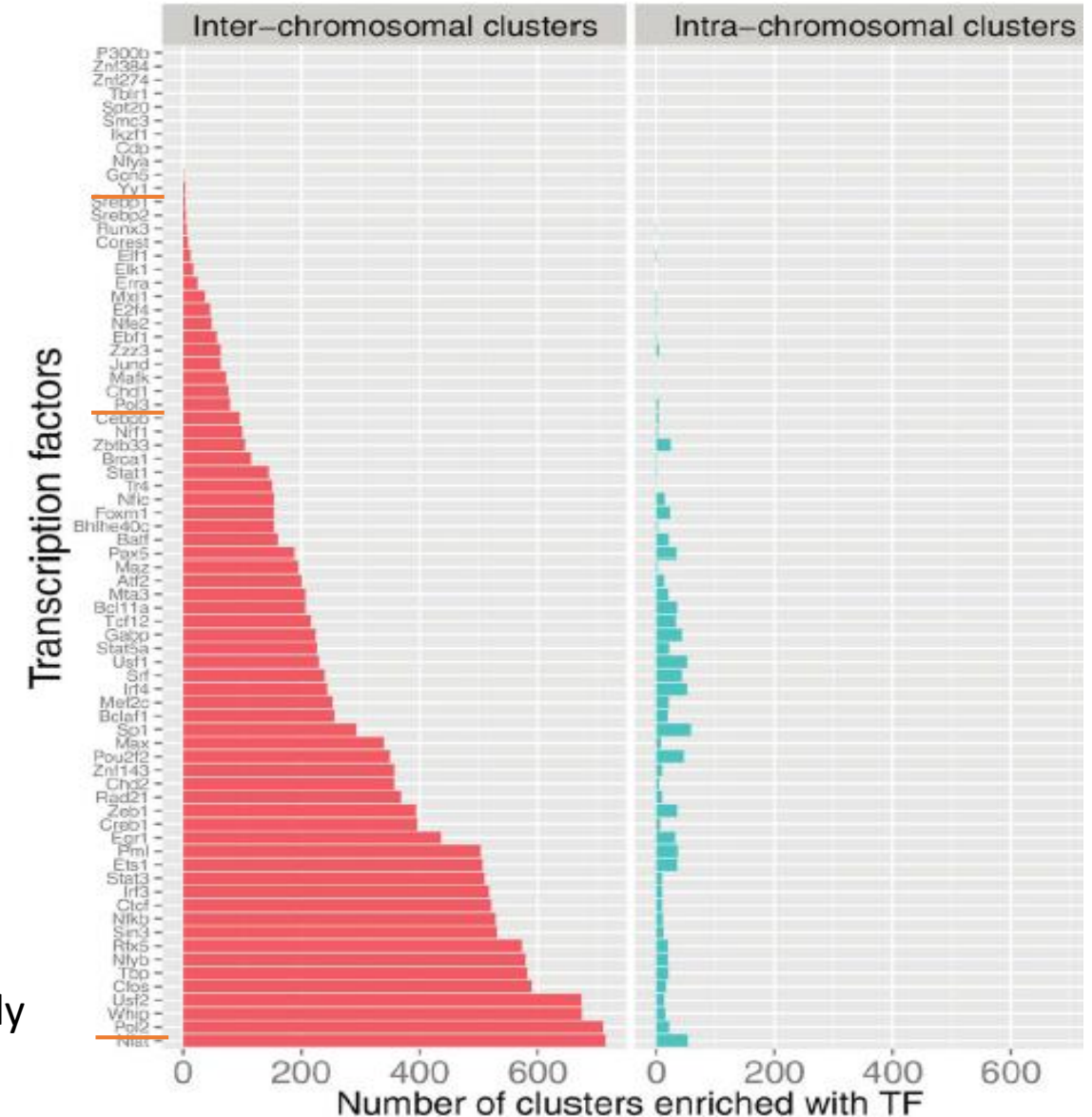




3856 frequent dense subgraphs :
at least 4 nodes, edge density at least 0.6,
occurs in at least 100 genome structures.

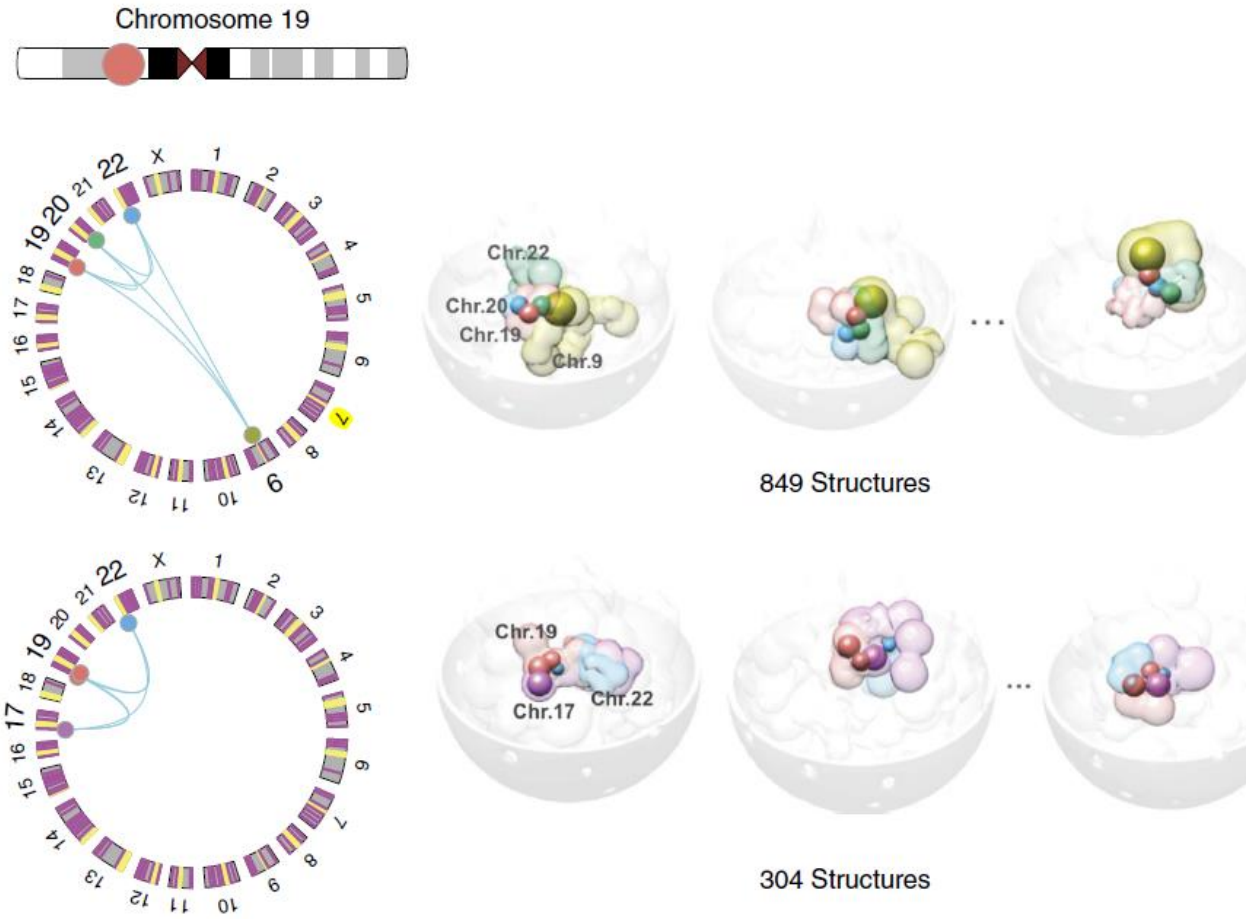
2.Spatial clusters constitute various regulatory communities.





regulatory community :
a frequent spatial cluster whose member domains are significantly co-enriched in binding to the same regulatory factor(s)

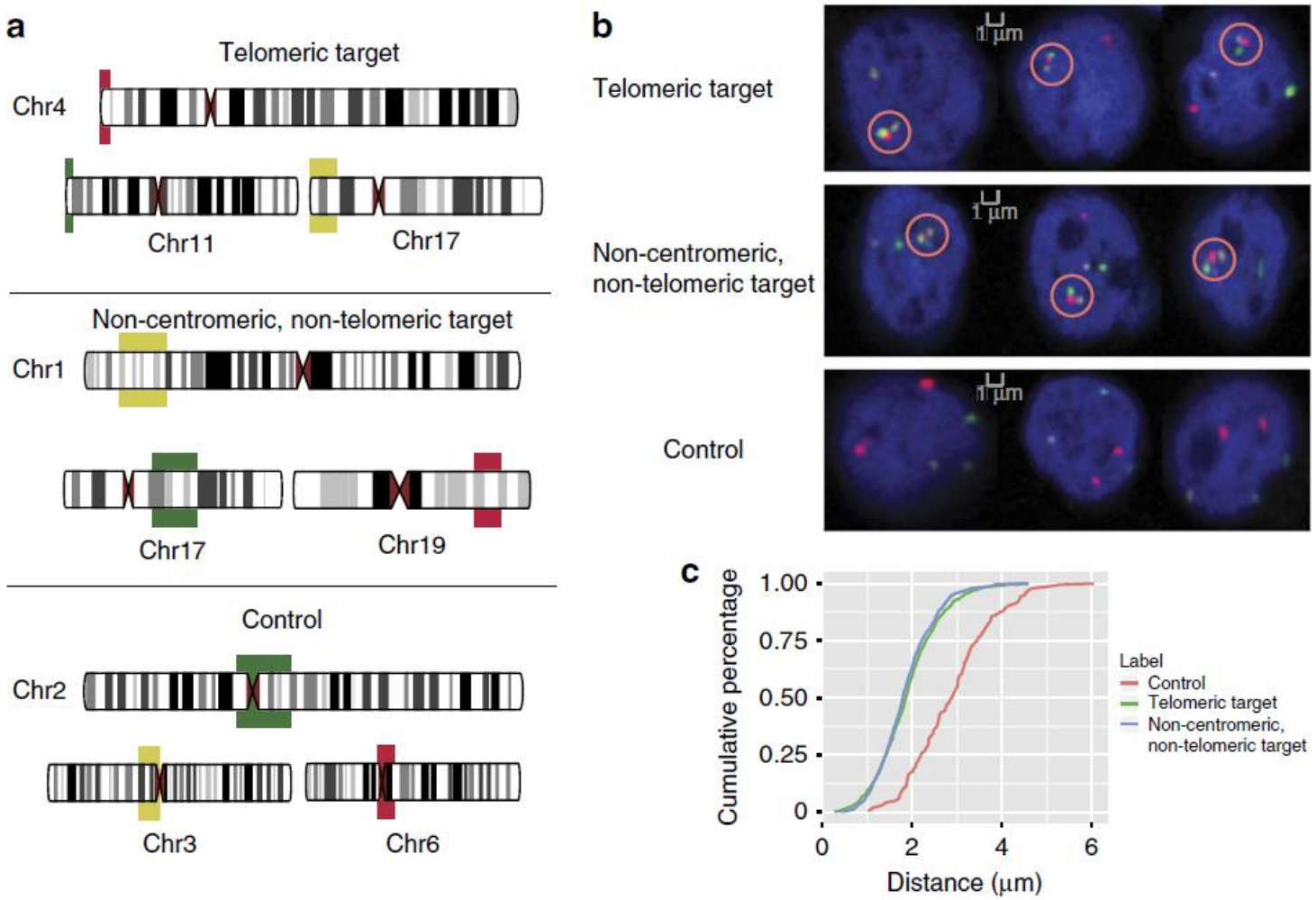
2.Spatial clusters constitute various regulatory communities.



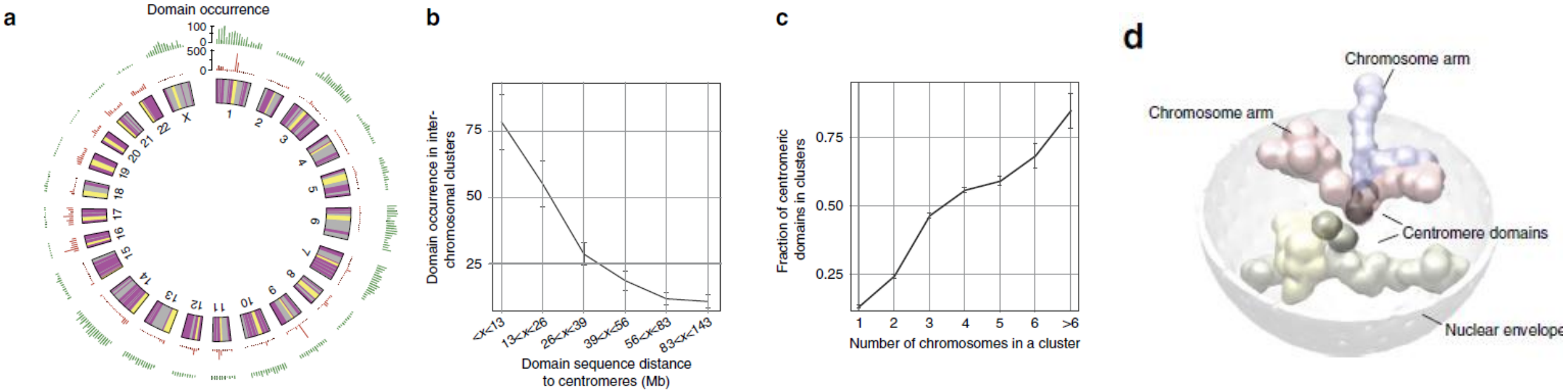
Functional plasticity of chromatin domain.

An active domain in chromosome 19 can participate in two different clusters that are enriched with binding of the same transcription factors, including RNAPII, CTCF, NFYB and CREB1.

Validation of co-localization with 3D FISH

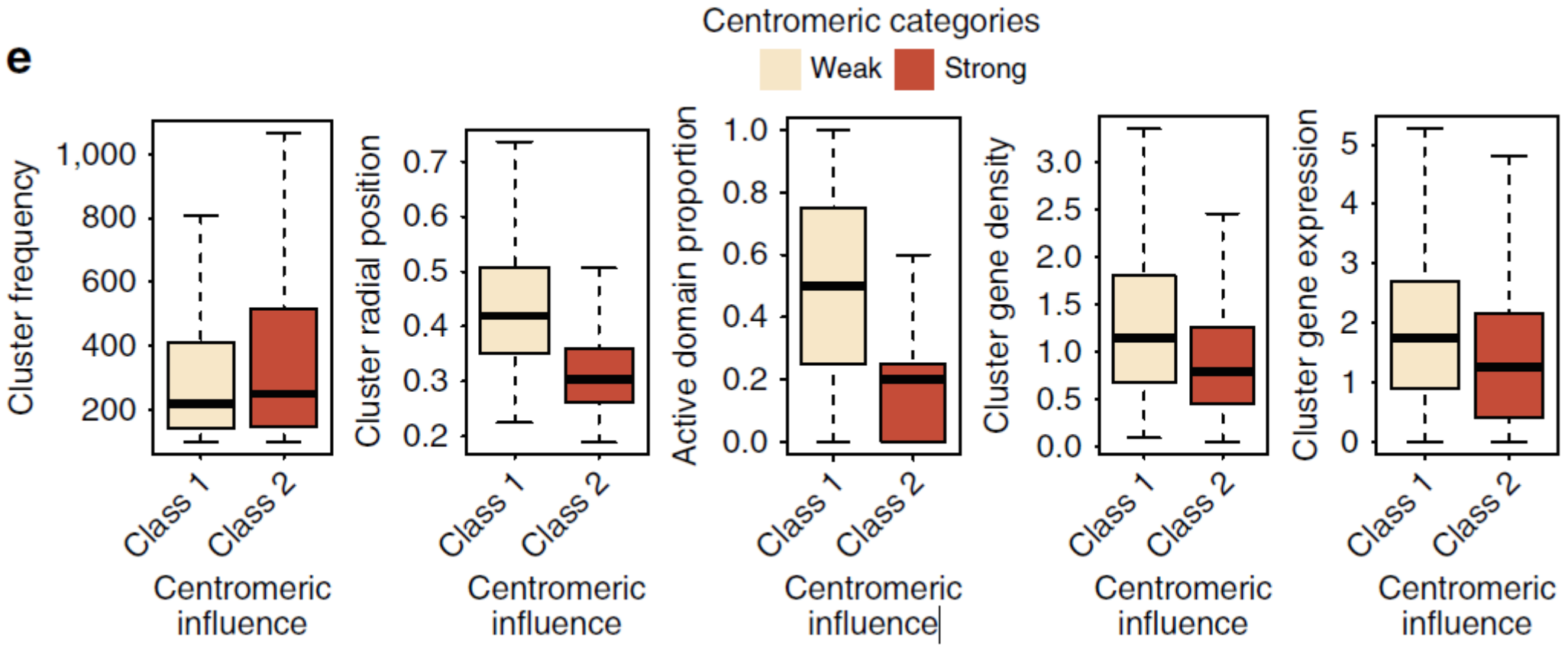


3.Centromeric domains are hubs for inter-chromosomal clusters.

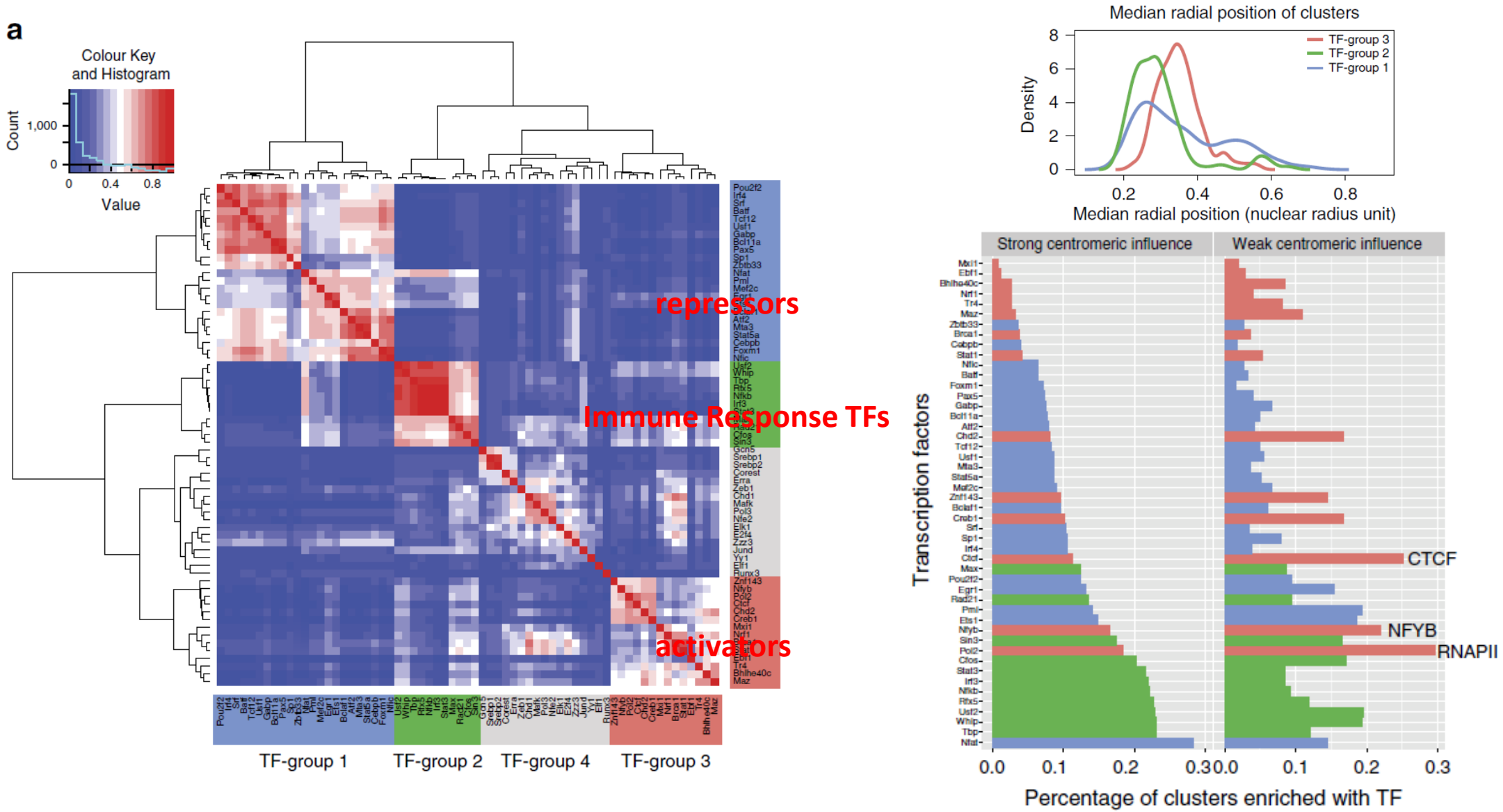


a.the vast majority (87%) of 3,107 inter-chromosomal clusters contain at least 1 centromeric domain.
b.the closer a domain is to the centromere of its chromosome, the more frequently it participates in stable inter-chromosomal clusters
c.clusters involving more chromosomes generally have a higher proportion of centromeric domains

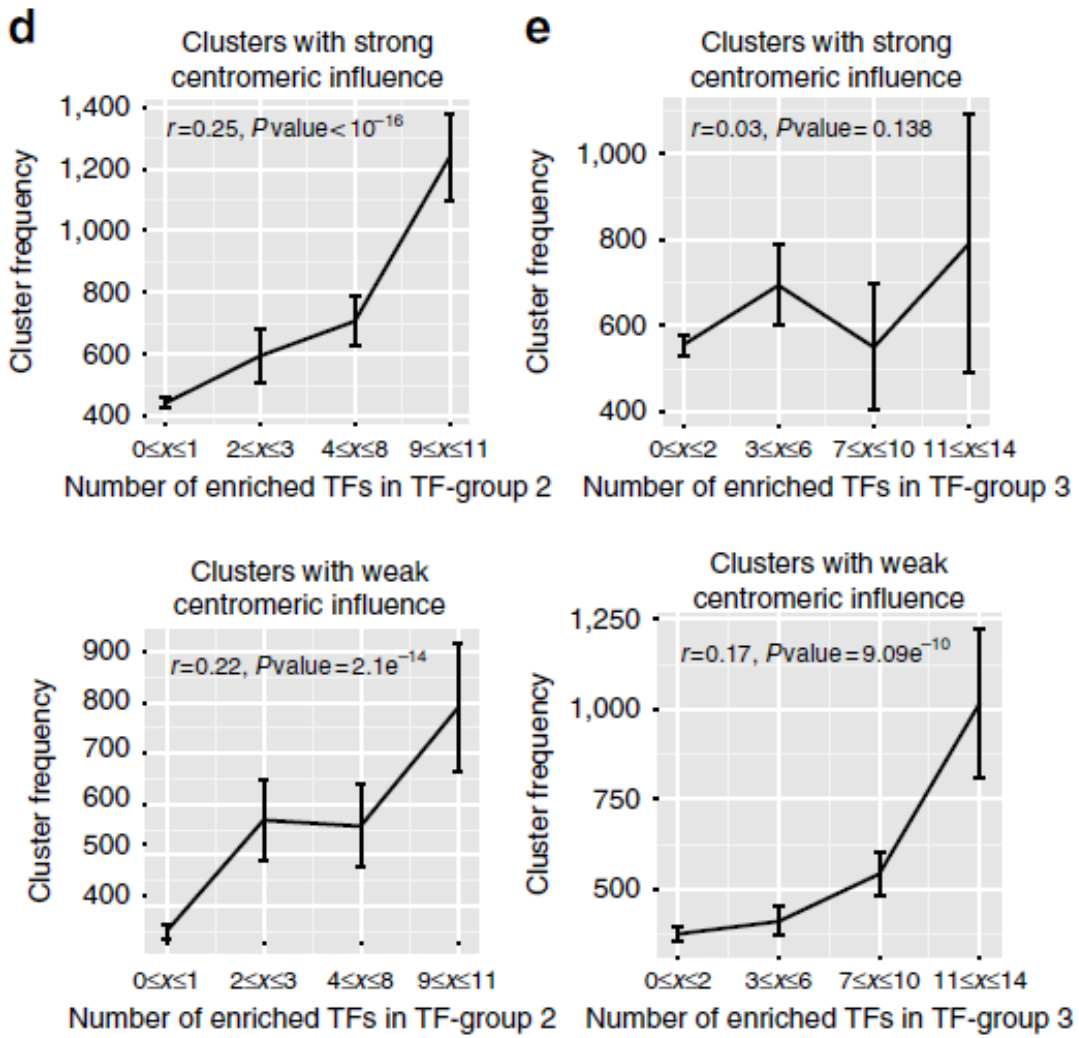
3.Centromeric domains are hubs for inter-chromosomal clusters.

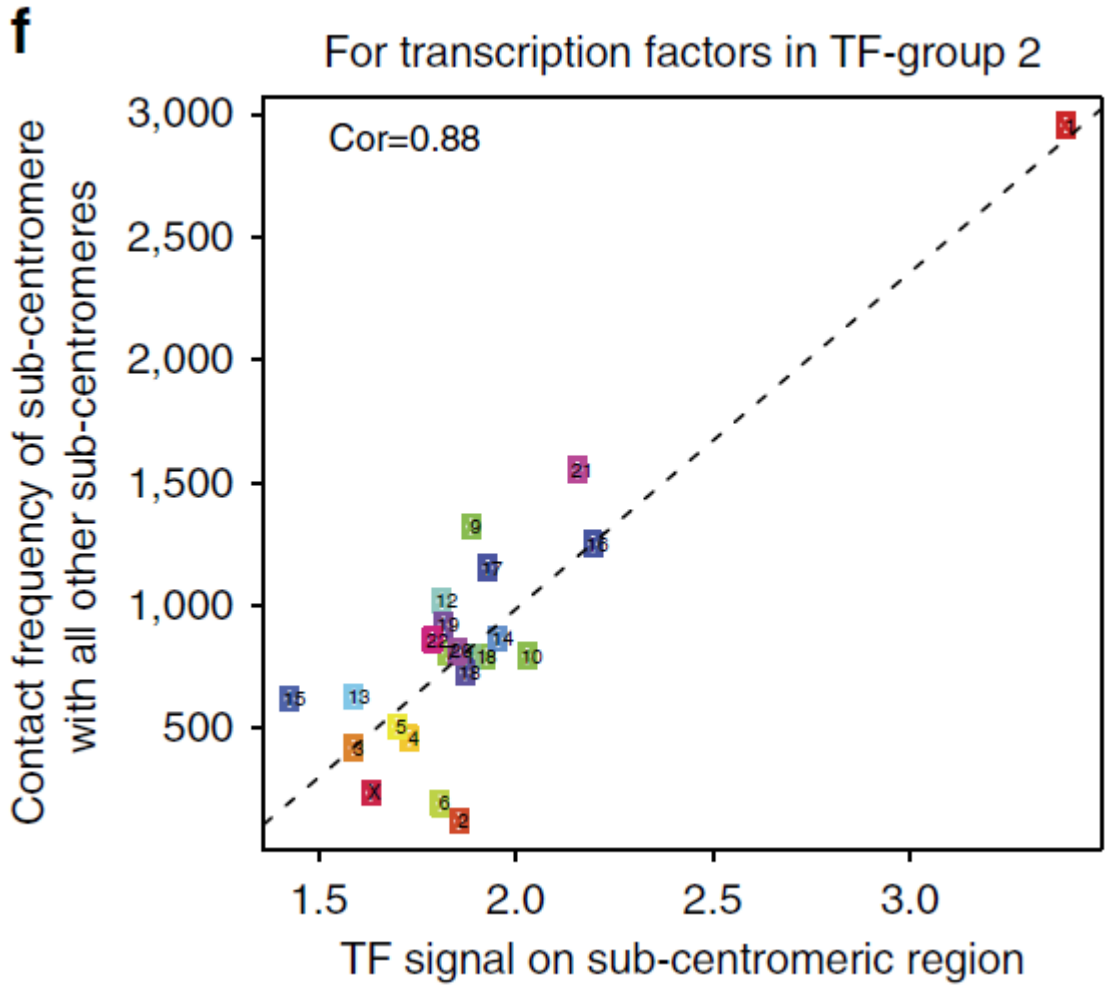


4.Transcription factors may stabilize regulatory communities.

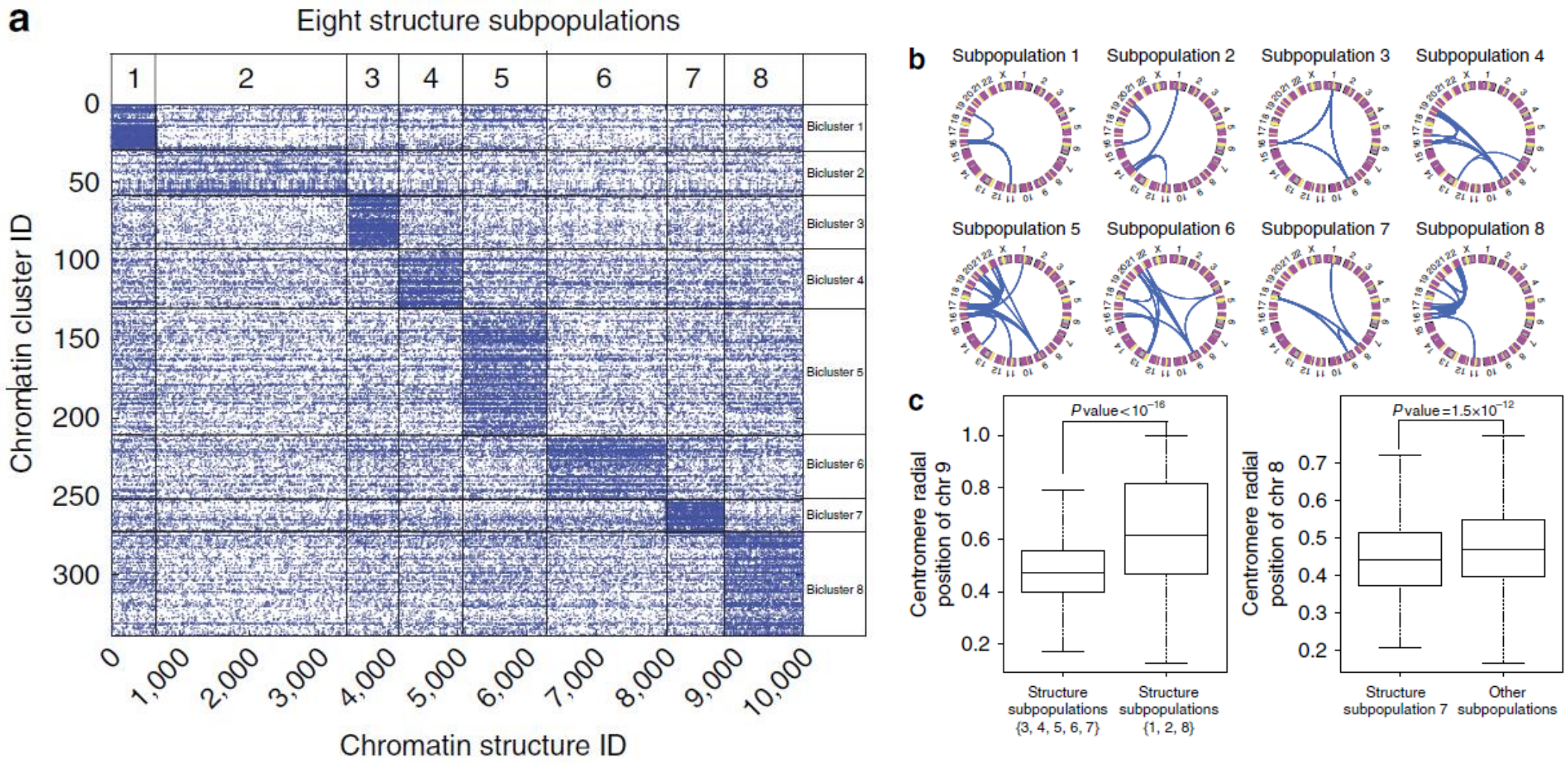


4.Transcription factors may stabilize regulatory communities.





5.The genome structure population contains multiple substates



5.The genome structure population contains multiple substates

Summary

1.A graph-based computational framework for the analysis of 3D genome structure populations

2. frequently occurring chromatin clusters are enriched in binding of specific regulatory factors

3.Two major factors, centromere clustering and transcription factor binding, significantly stabilize such regulatory communities

4.The regulatory communities differ substantially from cell to cell

3D genome modeling opportunity and challenge

a. single cell level (limited capture efficiency)
b. diploid genome
c. Integration: ChIA-PET, Lamin,FISH

at which scale ?(inter chromosome; inter domain; intra domain)
what biological question can we solve?
How much we approach the real structure?
How to verify?
Is it necessary?

Single-cell Hi-C reveals cell-to-cell variability in chromosome structure

Takashi Nagano[1]*, Yaniv Lubling[2]*, Tim J. Stevens[3]*, Stefan Schoenfelder[1], Eitan Yaffe[2], Wendy Dean[4], Ernest D. Laue[3], Amos Tanay[2] & Peter Fraser[1]

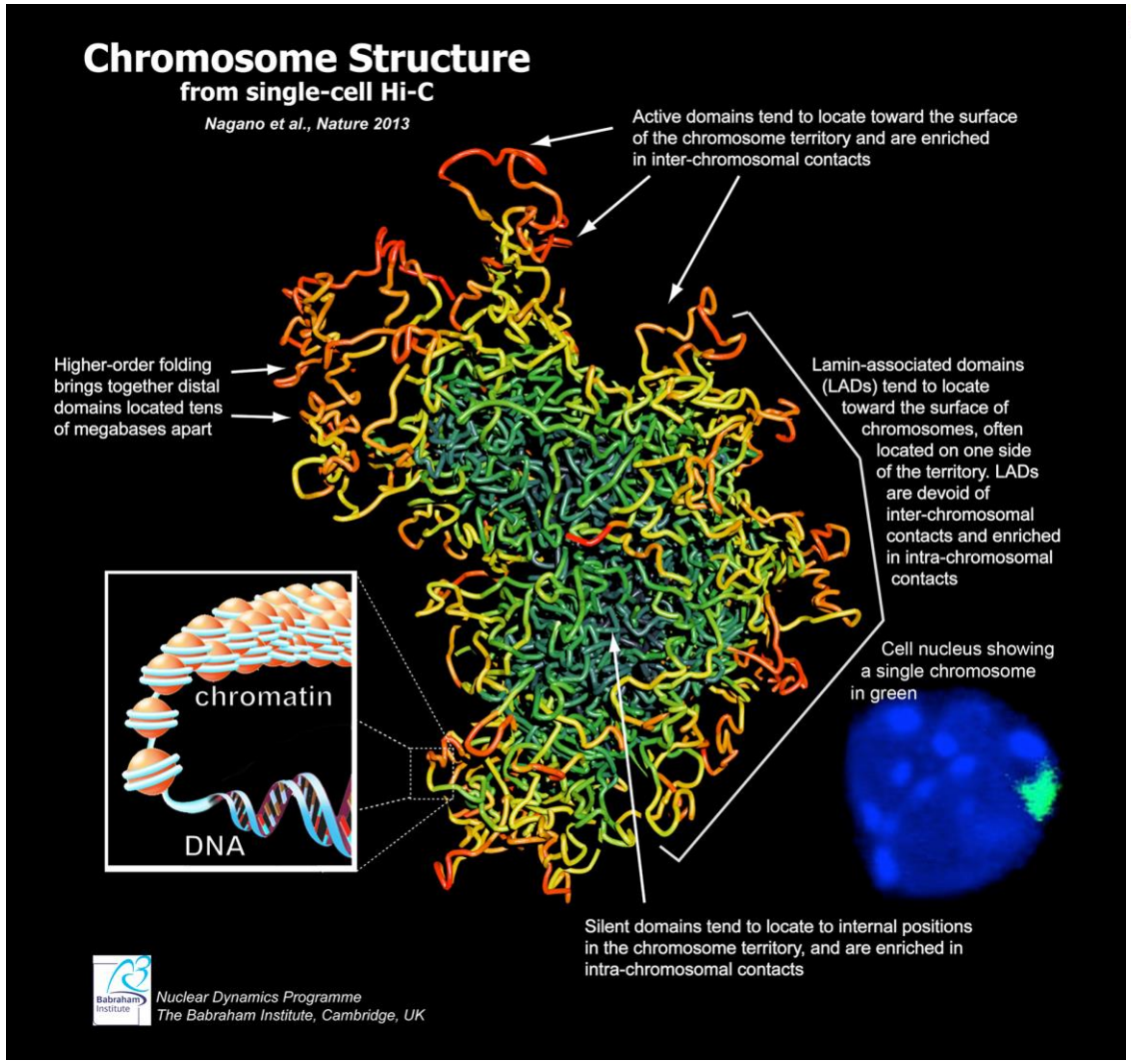
Cell cycle dynamics of chromosomal organisation at single-cell resolution

Massively multiplex single-cell Hi-C

Vijay Ramani[1], Xinxian Deng[2], Ruolan Qiu[1], Kevin L Gunderson[3], Frank J Steemers[3], Christine M Disteche[2,4], William S Noble[1], Zhijun Duan[5,6] & Jay Shendure[1,7]

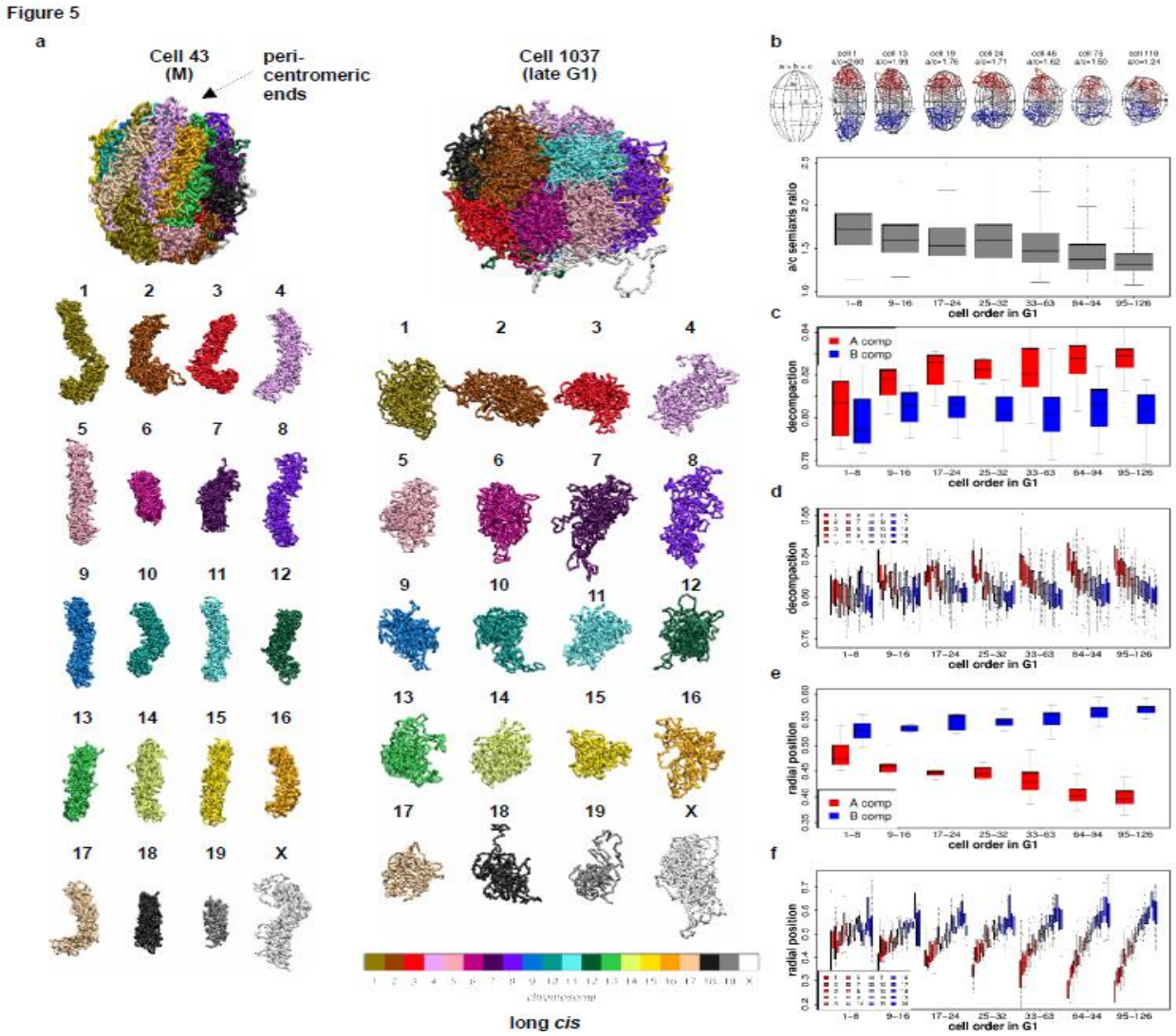
Single-cell Hi-C reveals cell-to-cell variability in chromosome structure

Takashi Nagano[1]*, Yaniv Lubling[2]*, Tim J. Stevens[3]*, Stefan Schoenfelder[1], Eitan Yaffe[2], Wendy Dean[4], Ernest D. Laue[3], Amos Tanay[2] & Peter Fraser[1]



10 Mb resolution

Cell cycle dynamics of chromosomal organisation at single-cell resolution



100 kb resolution

References:

[1].Cell cycle dynamics of chromosomal organisation at single-cell resolution. 2017.

[2].Paulsen, J., et al., Chrom3D: three-dimensional genome modeling from Hi-C and nuclear lamin-genome contacts. Genome Biol, 2017. 18(1): p. 21.

[3].Wang, S., et al., Spatial organization of chromatin domains and compartments in single chromosomes. Science, 2016. 353(6299): p. 598-602.

[4].Tjong, H., et al., Population-based 3D genome structure analysis reveals driving forces in spatial genome organization. Proceedings of the National Academy of Sciences, 2016. 113(12): p. E1663-E1672.

[5].Zou, C., Y. Zhang and Z. Ouyang, HSA: integrating multi-track Hi-C data for genome-scale reconstruction of 3D chromatin structure. Genome Biol, 2016. 17: p. 40.

[6].Dai, C., et al., Mining 3D genome structure populations identifies major factors governing the stability of regulatory communities. Nature Communications, 2016. 7: p. 11549.

[7].Wang, S., J. Xu and J. Zeng, Inferential modeling of 3D chromatin structure. Nucleic Acids Res, 2015. 43(8): p. e54.

[8].Kind, J., et al., Genome-wide maps of nuclear lamina interactions in single human cells. Cell, 2015. 163(1): p. 134-47.

[9].Trieu, T. and J. Cheng, Large-scale reconstruction of 3D structures of human chromosomes from chromosomal contact data. Nucleic Acids Res, 2014. 42(7): p. e52.

[10].彭城等, 染色质三维结构重建及其生物学意义. 中国科学:生命科学, 2014(08): 第794-802页.

[11].Lesne, A., et al., 3D genome reconstruction from chromosomal contacts. Nat Methods, 2014. 11(11): p. 1141-3.

[12].McCord, R.P., et al., Correlated alterations in genome organization, histone methylation, and DNA-lamin A/C interactions in Hutchinson-Gilford progeria syndrome. Genome Res, 2013. 23(2): p. 260-9.

[13].Zhang, Z., et al., 3D chromosome modeling with semi-definite programming and Hi-C data. J Comput Biol, 2013. 20(11): p. 831-46.

[14].Hu, M., et al., Bayesian inference of spatial organizations of chromosomes. PLoS Comput Biol, 2013. 9(1): p. e1002893.

[15].Dekker, J., M.A. Marti-Renom and L.A. Mirny, Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. Nat Rev Genet, 2013. 14(6): p. 390-403.

[16].Nagano, T., et al., Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. Nature, 2013. 502(7469): p. 59-64.

[17].Marti-Renom, M.A. and L.A. Mirny, Bridging the resolution gap in structural modeling of 3D genome organization. PLoS Comput Biol, 2011. 7(7): p. e1002125.

[18].Kalhor, R., et al., Genome architectures revealed by tethered chromosome conformation capture and population-based modeling. Nat Biotechnol, 2011. 30(1): p. 90-8.

[19].Chromosome Territories.

THANK YOU